1 Część 1 Podstawy sieci

1.1 Rozdział 1 ABC sieci

Mark A. Sportack

W miarę jak przetwarzanie danych na odległość staje się coraz to powszechniejsze, również coraz częstszym elementem praktycznie wszystkich środowisk obliczeniowych stają się sieci komputerowe. Sieć komputerowa jest mechanizmem umożliwiającym komunikowanie się komputerów znajdujących się w różnych miejscach; integralnym elementem owej komunikacji jest wzajemne udostępnianie sobie zasobów. Pomimo wielorakich zastosowań, sieci komputerowe należą jednak do słabiej rozpoznanych obszarów technologii informatycznych - czemu nierzadko towarzyszy swoista aura tajemniczości.

W niniejszym rozdziale przedstawione są różne typy sieci, zasady ich działania, a także omówiony jest sposób, w jaki ich ewolucja wpłynęła na zmiany standardów informatycznych.

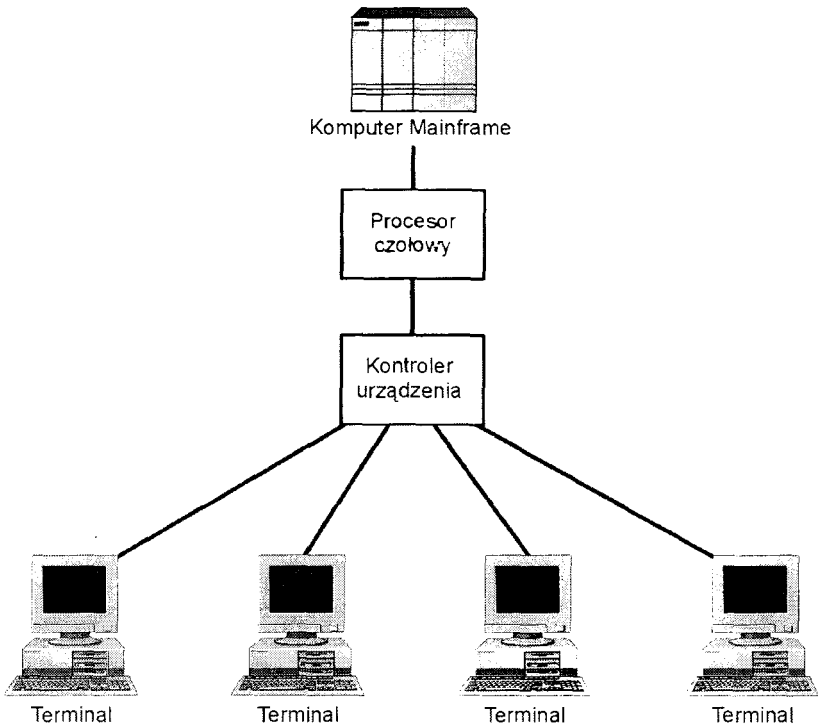
Standardy współdziałania sieciowego ustanawiane są przez różne organizacje. W rozdziale przedstawione są więc zarówno organizacje, ich standardy, jak i związki pomiędzy nimi. .Jednym z ważniejszych standardów jest model *połączonych systemów otwartych OSI).* W modelu tym każda warstwa obsługuje określony zbiór funkcji. Podział sieci na składniki funkcjonalne ułatwia zrozumienie zasad jej działania jako całości. Grupy składników gromadzone są następnie *w warstwy.* A warstwowy model OSI jest dokładnie omówiony na dalszych stronach niniejszego rozdziału.

1.1.1 Ewolucja sieci

U początków swego istnienia sieci komputerowe były zindywidualizowanymi formami połączeń, stanowiącymi integralną część równie zindywidualizowanych rozwiązań obliczeniowych. Przedsiębiorstwa, które w owych przedpecetowych czasach zdecydowały się zautomatyzować funkcje księgowości lub przetwarzania danych, wykonanie całego systemu musiały powierzyć jednemu wykonawcy.

Standardowe konfiguracje składały się z terminali połączonych sprzętowo z kontrolerami urządzeń. Kontrolery te umożliwiały dostęp multipleksowany (czyli wielodostęp) do urządzeń komunikacyjnych pozwalających na przyłączanie urządzeń do sieci głównej *(mainframe).* Urządzenia komunikacyjne skupione były w procesorze czołowym sieci *mainframe.* Procesor czołowy umożliwiał wielu urządzeniom komunikacyjnym współdzielenie pojedynczego kanału dostępu do sieci. Ze względu na różnicę w szybkości przetwarzania danych między portami wejścia/wyjścia a procesorami sieci *mainframe,* rozwiązanie to było najbardziej efektywne ekonomicznie. Przedstawione jest ono na rysunku 1.1.

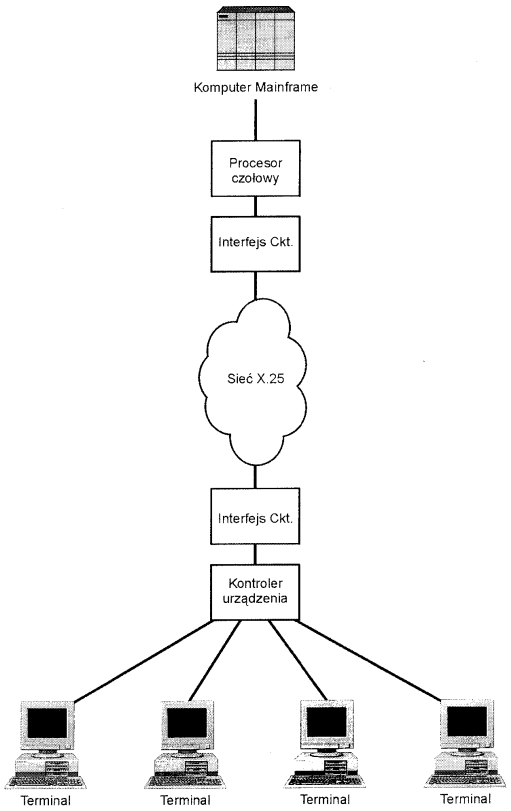
Rysunek 1.1. Sprzętowy dostęp do sieci mainframe.



Do komunikacji z siecią mainframe na dużą odległość wykorzystywane były linie dzierżawione pasm o niskiej częstotliwości. Linia dzierżawiona w takim przypadku łączona była z kanałem wejścia/wyjścia sieci mainframe Ilustruje to rysunek 1.2.

Programy działały jedynie w środowisku obsługiwanym przez pojedynczy system operacyjny. A system operacyjny mógł działać jedynie na podstawie urządzenia jednego, określonego producenta. Nawet terminale użytkowników i urządzenia, za pomocą których były one przyłączane do sieci, musiały być częścią zintegrowanego rozwiązania jednego producenta.
W czasach takich zintegrowanych rozwiązań systemowych nastąpiły dwie ewolucyjne zmiany technologiczne, które zmieniły kierunek rozwoju informatyki. Pierwsza - to pojawienie się prymitywnych poprzedników dzisiejszych komputerów osobistych (czyli komputerów PC lub inaczej pecetów). Innowacyjność tych urządzeń polegała na przeniesieniu miejsca wykonywania obliczeń - z osobnego pokoju, mieszczącego (z trudnością) komputer główny, na biurka każdego z uczestników sieci.

Rysunek 1.2. Dostęp do sieci mainframe za pomocą linii dzierżawionych.



Druga zmiana wynikła z potrzeby poprawy wydajności pracy jaką odczuwali pracownicy centrum badawczego firmy Xerox w Palo Alto (PARC). Dokładnie rzecz biorąc, próbowali oni usprawnić sposób współdzielenia plików i danych pomiędzy ich inteligentne stacje robocze - praktykowane udostępnianie danych przy użyciu dyskietek było bowiem czasochłonne i nieporęczne, niezawodność zaś całego procesu - wielce problematyczna.

Rozwiązanie opracowane w PARC polegało na utworzeniu pierwszej tzw.. sieci lokalnej LAN (ang. Local Area Network) - sieć ta nazwana została Ethernet. Korzystała ona z protokołów współdziałania międzysieciowego wyższych warstw. Jej możliwości rynkowe zostały szybko wykorzystane: pierwotny Ethernet, dziś znany jako Ethernet PARC lub Ethernet I, został zastąpiony przez nieco udoskonaloną wersję - DIX Ethernet, zwaną również Ethernet II. Jej autorzy - firmy Xerox, Digital oraz Intel ustaliły wspólnie „standardy" sieciowe, do przestrzegania których zobowiązały się przy produkcji jej elementów składowych.
Inteligentne urządzenia końcowe w połączeniu z sieciami LAN przyczyniły się do rozpowszechnienia nowego paradygmatu: otwartego, rozproszonego przetwarzania danych.

1.1.2 Organizacje ustanawiające standardy

Sukces sieci Ethernet I oraz Ethernet II uwidacznia, że dawno już zmęczyliśmy się zindywidualizowanym podejściem do sieciowego przetwarzania danych. Jako klienci wymagamy teraz bardziej otwartych środowisk. które umożliwiłyby nam tworzenie własnych rozwiązań z różnych produktów wielu producentów. Współdziałanie międzysieciowe, jak widać na przykładzie Ethernetu, przyczynia się do zwiększenia konkurencyjności i tempa wprowadzania innowacji technologicznych. Wzajemnie zależnymi celami otwartości są:
• niższe koszty,
• większe możliwości,
• współdziałanie produktów różnych producentów.
Współdziałanie produktów różnych producentów wymaga, aby różne platformy rozpoznawały się wzajemnie oraz wiedziały, w jaki sposób mogą się ze sobą komunikować i współdzielić dane. Wymogi te przyczyniły się do rozwoju uniwersalnych standardów dotyczących każdego aspektu sieciowego przetwarzania danych.

Potrzeba standaryzacji zwiększyła wysiłki organizacji zajmujących się normalizowaniem. Dziś za określanie krajowych i międzynarodowych standardów regulujących różne aspekty technologii informatycznych odpowiedzialnych jest wiele różnych organizacji. Najczęściej współpracują one ze sobą w celu ustanowienia jak najbardziej uniwersalnego zbioru standardów. Może to, co prawda, powodować pewne zamieszanie, które jednak -jest nieporównywalne z korzyściami płynącymi z owej standaryzacji.
Poniżej przedstawione są organizacje tworzące standardy i powiązania między nimi.

1.1.2.1 ANSI

Amerykański Narodowy Instytut Normalizacji lub - bardziej dosłownie: Amerykański Instytut Standardów Narodowych (ang. ANSI - The American National Standards Institute) - jest prywatną organizacją niekomercyjną. Jej misją jest ułatwianie rozwoju,
koordynacja oraz publikowanie nieobligatoryjnych standardów. „Nieobligatoryjność" standardów ANSI polega na tym, że organizacja ta nie wdraża aktywnie ani nie narzuca nikomu swoich standardów. Uczestniczy natomiast w pracach organizacji ustanawiających standardy globalne, takich jak IOS, IEC itp., w związku z czym niezgodność z jej standardami powoduje niezgodność ze standardami globalnymi.

1.1.2.2 IEEE

Instytut Elektryków i Elektroników (ang. lEEE- The Institute of Electrical and Electronic Engineers) jest odpowiedzialny za definiowanie i publikowanie standardów telekomunikacyjnych oraz przesyłania danych. Jego największym - jak dotąd - osiągnięciem jest zdefiniowanie standardów sieci LAN oraz MAN. Standardy te tworzą wielki i skomplikowany zbiór norm technicznych, ogólnie określany jako „Project 802" lub jako seria standardów 802.
Celem IEEE jest tworzenie norm, które byłyby akceptowane przez instytut ANSI. Akceptacja taka zwiększyłaby ich forum dzięki uczestnictwie ANSI w globalnych organizacjach określających standardy.

1.1.2.3 ISO

Międzynarodowa Organizacja Normalizacyjna (ang. ISO - lnternational Organization for Standardization) została utworzona w 1946 roku w Szwajcarii, w Genewie-tam też znajduje się dziś jej główna siedziba. Niektóre źródła organizację tę identyfikują za pomocą akronimu IOS. Mimo iż to właśnie ten skrót jest formalnie poprawny, organizacja woli określać się za pomocą bardziej mnemonicznego (łatwiejszego do zapamiętania) skrótu: ISO. Skrót ten pochodzi od greckiego słowa i.sos, które jest odpowiednikiem polskiego „równy" lub „standardowy". Dlatego ten właśnie skrót jest uznawany za skrót Międzynarodowej Organizacji Normalizacyjnej, która - przy okazji - jest niezależnym podmiotem wynajętym przez Organizację Narodów Zjednoczonych (ONZ) do określania standardów międzynarodowych. Zakres jej działania obejmuje praktycznie wszystkie dziedziny wiedzy ludzkiej, poza elektryką i elektroniką. Aktualnie ISO składa się z ponad 90 różnych organizacji standardo-dawczych z siedzibami na całym świecie. Najważniejszym prawdopodobnie standardem ustanowionym przez ISO jest Model Referencyjny Połączonych Systemów Otwartych, czyli model OSI (ang. Open Systems Interconnection Reference Model). Model ten jest szczegółowo omówiony w dalszej części niniejszego rozdziału- w podrozdziale „Model referencyjny OSI".

1.1.2.4 IEC

Międzynarodowa Komisja Elektrotechniczna(ang. IEC-lnternational Electrotechnical C'ornnai.ssion), z siedzibą również w Genewie, została założona w roku 1909. Komisja IEC ustanawia międzynarodowe standardy dotyczące wszelkich zagadnień elektrycznych i elektronicznych. Aktualnie w jej skład wchodzą komitety z ponad 40 państw. W Stanach Zjednoczonych Instytut ANSI reprezentuje zarówno IEC, jak i ISO.
IEC oraz ISO dostrzegły, że technologie informatyczne stanowią potencjalny obszar zazębiania się ich kompetencji; w celu określenia standardów dla technologii informatycznych utworzyły więc Połączony Komitet Techniczny (ang. JTC - Joint Technical Committee).

1.1.2.5 IAB

Komisja Architektury Internetu (ang. IAB - The Internet Architecture Board, uprzednio znana jako Komisja Działań Internetu (Internet Activities Board), zarządza techniczną stroną rozwoju sieci Internet. Składa się z dwóch komisji roboczych: Grupy Roboczej ds. Technicznych Internetu (ang. IETF- Internet Engineering Task Force) oraz Grupy Roboczej ds. Naukowych Internetu (ang. IRTF- Internet Research Task Force). Każda z tych grup, na co wskazują ich nazwy, pracuje indywidualnie. Grupa ds. Naukowych (IRTF) bada nowe technologie, które mogą okazać się wartościowe lub mieć wpływ na rozwój Internetu. Grupa ds. Technicznych (IETF) jest odbiorcą badań Grupy Naukowej. Jest więc odpowiedzialna za ustanawianie standardów technicznych dla Internetu, jak również za określanie nowych standardów dla technologii internetowych, takich jak protokół Internetu (IP).
IP jest protokołem warstwy 3, czyli warstwy sieci. Jako taki jest on z natury protokołem bezpołączeniowym. Nie identyfikuje więc pakietów, które mają być przesłane ponownie (retransmitowane). Nie potrafi też wykonywać wielu procesów związanych z odtwarzaniem prawidłowej sekwencji pakietów, które mogły nadchodzić w w kolejności innej niż zostały nadane. Wygodne korzystanie z protokołu IP, jak z każdego innego protokołu warstwy 3, umożliwia dopiero protokół warstwy 4 (warstwy transportu). Przykładami protokołów warstwy 4, które korzystają z protokołu IP, są TCP, UDP, a także eksperymentalny TTCP. W przypadku używania połączonych protokołów warstwy 3 i 4 do rozdzielania ich nazw używa się ukośnika „/":, na przykład TCP/IP czy też UDP/IP. Jest to, niestety, przyczyną zamieszania dotyczącego warstw 3 i 4, w wyniku czego często nawet osoby z biegłą znajomością zagadnień technicznych używają nazwy TCP/IP, podczas gdy faktycznie chodzi im o IP.
Więcej informacji na temat protokołów działających w warstwach 3 i 4 modelu OSI zawiera rozdział 12 pt. „Protokoły sieciowe".1

1.1.3 Model referencyjny OSI

Organizacja ISO opracowała Model Referencyjny Połączonych Systemów Otwartych (czyli model OSI) w celu ułatwienia realizacji otwartych połączeń systemów komputerowych. Połączenia otwarte to takie, które .mogą być obsługiwane w środowiskach
' Obszerny opis wymienionych protokołów znajduje się również w książce T. Parkera .,1'CP/IP" , wyd.. polskie HELION 1997 (przyp. red.)
wielosystemowych. Omawiany model jest globalnym standardem określania warstw funkcjonalnych wymaganych do obsługi tego typu połączeń.
Opracowany model OSI jawił się wówczas jako na wskroś radykalny. Producenci, skazujący dotąd swych klientów na indywidualnie tworzone architektury, wymagające korzystania z elementów jednego tylko producenta, stanęli nagle wobec wyzwania, jakie stwarzało pojawienie się modelu otwartego - kładącego de .facto kres podobnym praktykom monopolistycznym i postrzeganego tym samym (czemuż się tu dziwić) jako swoisty zamach na ich partykularne interesy. Model OSI zyskał jednak olbrzymie poparcie, a wcześniejsze, zindywidualizowane podejście odeszło w zapomnienie.
Obecnie otwartość komunikacji jest warunkiem koniecznym większości systemów - zastanawiający może wydawać się więc fakt, iż tak naprawdę niewiele jest produktów w pełni zgodnych z modelem OSI. Producenci dostosowują raczej warstwową strukturę tego modelu do nowych standardów - co nie zmienia jego roli jak żywotnego mechanizmu przedstawiającego we właściwy sposób mechanizmy funkcjonowania sieci.
Mimo popularności modelu OSI, istnieje na jego temat wiele nierzadko mylnych przeświadczeń; spróbujmy przyjrzeć się niektórym z nich, analizując - po prostu - fakty, o których traktują.
Tak więc, zgodnie z powszechnym mniemaniem, model OSI został opracowany w paryskiej siedzibie Organizacji Standardów Międzynarodowych (International Standards Organization, czyli lS0). Nie jest to prawdą - został on bowiem opracowany przez Międzynarodową Organizację Normalizacji (lnternational Organization for Standardization, czyli również „ISO", lecz w nieco alegorycznym sensie greckiego isos, o czym wspominaliśmy już przed chwilą.
Opracowanie modelu OSI często, acz niesłusznie, przypisuje się Organizacji Standardów Międzynarodowych. Model ten został bowiem opracowany przez Międzynarodową Organizację Normalizacji.
Jednak w razie ewentualnego testu czy egzaminu, który w przyszłości może przydarzyć się każdemu z nas, pamiętajmy, że w razie braku poprawnej odpowiedzi, odpowiedzią „poprawną" jest Organizacja Standardów Międzynarodowych (International Standards Organization). Jest to odpowiedź niepoprawna, ale często oczekiwana przez (najwidoczniej niedouczone) komisje egzaminacyjne.
Model referencyjny OSI dzieli procesy zachodzące podczas sesji komunikacyjnej na siedem warstw funkcjonalnych, które zorganizowane są według naturalnej sekwencji zdarzeń zachodzących podczas sesji komunikacyjnej.
Model OSI przedstawiony jest na rysunku 1.3. Warstwy od 1 do 3 umożliwiają dostęp do sieci, a warstwy od 4 do 7 obsługują logistycznie komunikację końcową.
Rysunek 1.3. Model referencyjny OSI.

Nazwa warstwy modelu OSI	Numer warstwy
Aplikacji	7
Prezentacji	6
Sesji	5
Transportu	4
Sieci	3
Łącza danych	2
Fizyczna	1

1.1.3.1 Warstwa 1: warstwa fizyczna

Warstwa najniższa nazywana jest warstwa fizyczną.Jest ona odpowiedzialna za przesyłanie strumieni bitów. Odbiera ramki danych z warstwy 2, czyli warstwy łącza danych, i przesyła szeregowo, bit po bicie, całą ich strukturę oraz zawartość.
Jest również odpowiedzialna za odbiór kolejnych bitów przychodzących strumieni danych. Strumienie te są następnie przesyłane do warstwy łącza danych w celu ich ponownego ukształtowania.
Warstwa fizyczna w istocie widzi tylko jedynki i zera. Nie ma wbudowanego mechanizmu określania wagi ani znaczenia otrzymywanych i wysyłanych bitów. Jest zajęta wyłącznie fizycznymi właściwościami elektrycznych i/lub optycznych technik sygnalizowania. Dotyczy to napięcia prądu elektrycznego używanego do przenoszenia sygnałów. rodzaju nośnika i właściwości impedancji, a nawet fizycznego kształtu złącza terminującego nośnik.
Często błędnie zakłada się, że warstwa 1 modelu OSI obejmuje wszystkie jego elementy tworzące lub przenoszące sygnały komunikacji danych. Nie jest to prawdą. Model OSI jest bowiem jedynie modelem funkcjonalnym.
Warstwa 1, czyli warstwa fizyczna, obejmuje więc jedynie procesy i mechanizmy dotyczące przenoszenia sygnałów na nośnik i odbierania z niego sygnałów. Jej dolną granicę stanowi fizyczne złącze nośnika. Warstwa 1 nie obejmuje medium transmisyjnego (czyli nośnika)!
Nośnikami są wszelkie urządzenia przenoszące sygnały generowane przez mechanizmy warstwy 1 modelu OSI. Przykładami nośników są kable koncentryczne, kable światłowodowe i kabel skrętki dwużyłowej.
Niejasności dotyczące nośników wynikać mogą z faktu, że warstwa fizyczna określa wymagane charakterystyki wydajnościowe nośników, na których oparte są procesy i mechanizmy tej warstwy. Zakłada się niejako, że wymagania te są spełnione.

W związku z tym media transmisyjne pozostają poza obszarem zainteresowania warstwy fizycznej i czasem określane są mianem warstwy 0 (zerowej).

1.1.3.2 Warstwa 2: warstwa łącza danych

Druga warstwa modelu OSI nazywana jest warstwą łącza danych. Jak każda z warstw, również i ta pełni dwie zasadnicze funkcje: odbierania i nadawania. Jest ona odpowiedzialna za końcową zgodność przesyłanych danych.
W zakresie zadań związanych z przesyłaniem, warstwa łącza danych jest odpowiedzialna za upakowywanie instrukcji, danych itp. w tzw.. ramki. Ramka jest strukturą rodzimączyli właściwą dla - warstwy łącza danych, która zawiera ilość informacji wystarczającą do pomyślnego przesłania danych przez sieć lokalną do ich miejsca docelowego.
Pomyślna transmisja danych zachodzi wtedy, gdy dane osiągają miejsce docelowe w postaci niezmienionej w stosunku do postaci, w której zostały wysłane. Ramka musi więc również zawierać mechanizm umożliwiający weryfikowanie integralności jej zawartości podczas transmisji.
Wysoka jakość transmisji wymaga spełnienia następujących dwóch warunków:
• Węzeł początkowy musi odebrać od węzła końcowego potwierdzenie otrzymania każdej ramki w postaci niezmienionej.
• Węzeł docelowy przed wysłaniem potwierdzenia otrzymania ramki musi zweryfikować
integralność jej zawartości.
W wielu sytuacjach wysyłane ramki mogą nie osiągnąć miejsca docelowego lub ulec uszkodzeniu podczas transmisji. Warstwa łącza danych jest odpowiedzialna za rozpoznawanie i naprawę każdego takiego błędu.
Warstwa łącza danych jest również odpowiedzialna za ponowne składanie otrzymanych z warstwy fizycznej strumieni binarnych i umieszczanie ich w ramkach. Ze względu na fakt przesyłania zarówno struktury, jak i zawartości ramki, warstwa łącza danych nie tworzy ramek od nowa. Buforuje raczej przychodzące bity dopóki nie uzbiera w ten sposób całej ramki.
Warstwy 1 i 2 są niezbędne do komunikacji każdego rodzaju, niezależnie od tego czy sieć, w której się ona odbywa jest siecią lokalną (LAN), czy też rozległą (WAN).

1.1.3.3 Warstwa 3: warstwa sieci

Warstwa sieci jest odpowiedzialna za określanie trasy transmisji między komputerem-nadawcą, a komputerem-odbiorcą. Warstwa ta nie ma żadnych wbudowanych mechanizmów kontroli korekcji błędów i w związku z tym musi polegać na wiarygodnej transmisji końcowej warstwy łącza danych.
Warstwa sieci używana jest do komunikowania się z komputerami znajdującymi się poza lokalnym segmentem sieci LAN. Umożliwia im to własna architektura trasowania, niezależna od adresowania fizycznego warstwy 2.
Protokołami trasowanymi są: • IP,
• IPX,
• AppleTalk.
Korzystanie z warstwy sieci nie jest obowiązkowe. Wymagane jest jedynie wtedy, gdy komputery komunikujące się znajdują się w różnych segmentach sieci przedzielonych routerem.

1.1.3.4 Warstwa 4: warstwa transportu

Warstwa transportu pełni funkcję podobną do funkcji warstwy łącza w tym sensie, że jest odpowiedzialna za końcową integralność transmisji. Jednak w odróżnieniu od warstwy łączy danych - warstwa transportu umożliwia tę usługę również poza lokalnymi segmentami sieci LAN. Potrafi bowiem wykrywać pakiety. które zostały przez routery odrzucone i automatycznie generować żądanie ich ponownej transmisji.
Inną ważną funkcją warstwy transportu jest resekwencjonowanie pakietów, które mogły zostać przysłane w nieodpowiedniej kolejności. Sytuacja taka może mieć kilka przyczyn. Na przykład, pakiety mogły podążać przez sieć różnymi ścieżkami lub zostać uszkodzone podczas transmisji. Warstwa transportu identyfikuje więc oryginalną sekwencję pakietów i ustawia je w oryginalnej kolejności przed wysłaniem ich zawartości do warstwy sesji.

1.1.3.5 Warstwa 5: warstwa sesji

Piątą warstwą modelu OSI jest warstwa sesji. Jest ona rzadko używana; wiele protokołów funkcje tej warstwy dołącza do swoich warstw transportowych.
Zadaniem warstwy sesji modelu OSI jest zarządzanie przebiegiem komunikacji podczas połączenia między dwoma komputerami. Ów przepływ komunikacji nazywany jest również *sesją.* Warstwa 5 określa, czy komunikacja może zachodzić w jednym, czy w obu kierunkach. Gwarantuje również zakończenie wykonywania bieżącego żądania przed przyjęciem kolejnego.

1.1.3.6 Warstwa 6: warstwa prezentacji

Warstwa prezentacji jest odpowiedzialna za zarządzanie sposobem kodowania wszelkich danych. Nie każdy komputer korzysta z tych samych schematów kodowania danych, więc warstwa prezentacji odpowiedzialna jest za translację między niezgodnymi schematami kodowania danych,`takimi jak na przykład American Standard Code for Information Interchange (ASCII) a Extended Binary Coded Decimal Interchange Code (EBCDIC).
Warstwa prezentacji może być wykorzystywana do niwelowania różnic między formatami zmiennopozycyjnymi, jak również do szyfrowania i rozszyfrowywania wiadomości.

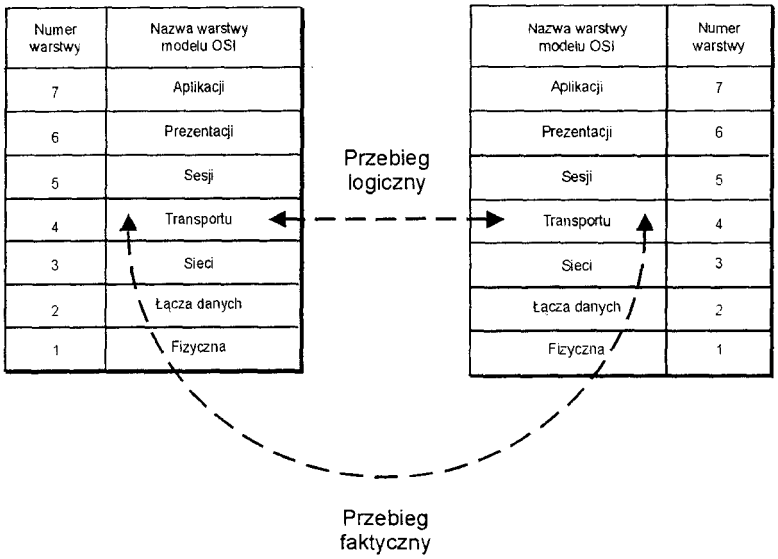
1.1.3.7 Warstwa 7: warstwa aplikacji

Najwyższą warstwą modelu OSI jest warstwa aplikacji. Pomimo sugestywnej nazwy warstwa ta nie obejmuje aplikacji użytkownika, pełniąc raczej rolę interfejsu pomiędzy tą aplikacją a usługami sieci.
Warstwę tę można uważać za inicjującą sesje komunikacyjne. Na przykład, klient poczty elektronicznej mógłby generować żądanie pobrania nowych wiadomości od jej nadawcy. Taka aplikacja kliencka generuje automatycznie żądanie do odpowiedniego protokołu (lub protokołów) warstwy 7 i uruchamia sesję komunikacji w celu otrzymania odpowiednich plików.

1.1.3.8 Zastosowania modelu

Pionowe zorientowanie stosu odzwierciedla funkcjonalny przebieg procesów oraz danych. Każda warstwa wyposażona jest w interfejsy warstw sąsiednich. Komunikacja jest możliwa, gdy komputery przesyłają dane, instrukcje, adresy itd. między odpowiednimi warstwami. Różnice między logicznym przebiegiem komunikacji, a rzeczywistym przebiegiem sesji przedstawione są na rysunku 1.4.
Mimo że model składa się z siedmiu warstw, to określona sesja komunikacyjna nie musi wykorzystywać wszystkich siedmiu, lecz tylko niektóre z nich. Na przykład, komunikacja w ramach jednego segmentu LAN może być przeprowadzana wyłącznie w warstwach 1 i 2 modelu OSI, bez potrzeby korzystania z dwóch pozostałych (3 i 4) warstw komunikacyjnych.
Choć komunikacja w stosie odbywa się w płaszczyźnie pionowej, każdej warstwie wydaje się, że może się komunikować bezpośrednio z odpowiadającymi jej warstwami w komputerach zdalnych. Logiczne rozgraniczenie warstw możliwe jest dzięki temu, że do każdej warstwy stosu protokołów komputera nadającego dodawany jest nagłówek. Nagłówek ten może być rozpoznany i użyty jedynie przez daną warstwę lub jej odpowiedniki w innych komputerach. Stos protokołów komputera odbierającego usuwa kolejne nasłówki, warstwa po warstwie, w miarę jak dane przesyłane są do jego warstwy aplikacji. Proces ten jest przedstawiony na rysunku 1.5.

Rysunek 1.4. *Logiczny i faktyczny przebieg komunikacji warstwowej.*



Na przykład, segmenty danych upakowane przez warstwę 4 komputera nadającego przesyłane są do jego warstwy 3. Warstwa 3 łączy segmenty danych otrzymane z warstwy 4 w pakiety (czyli pakietuje je segmenty), adresuje je i wysyła do protokołu warstwy 3 komputera docelowego za pośrednictwem własnej warstwy 2. Warstwa 2 oprawia dane w ramki, opatrując je adresem rozpoznawanym przez sieć LAN. Ramki te są wysyłane do warstwy 1 w celu zamiany ich na strumień cyfr binarnych (bitów), które następnie są przesyłane do warstwy 1 komputera docelowego. Komputer docelowy odwraca opisany przebieg przekazów, przy czym każda warstwa pobiera i zatrzymuje nagłówki dodane przez jej odpowiednik z komputera nadającego. Zanim przesyłane dane dotrą do warstwy 4 komputera docelowego, przyjmą one na powrót formę nadaną im przez warstwę 4 komputera nadającego. W ten sposób protokoły dwóch warstw 4 wydają się fizycznie graniczyć ze sobą i komunikować bezpośrednio.
Proszę zauważyć, że większość obecnie używanych protokołów używa własnych modeli warstwowych. Mogą one w różnym stopniu odpowiadać podziałowi funkcji określonemu przez model OSI. Modele te bardzo często dzielą funkcje nie między 7, lecz między 5 lub mniej warstw. Często też warstwy wyższe różnią się znacznie od ich odpowiedników modelu OSI.
Każda warstwa 3 (sieci) przesyła dane do warstwy 2 (łącza danych), która z kolei przekształca otrzymane ramki na ciągi bitów wysyłane przez warstwę 1 (fizyczną) komputera nadającego. Warstwa 1 komputera odbierającego ciąg bitów przesyła je do swojej warstwy 2 w celu przetworzenia ich ponownie do postaci ramki. Po pomyślnym zestawieniu otrzymanych danych w ramki, z ramek zdejmowane jest obramowanie, a uzyskany w ten sposób pakiet przesyłany jest do warstwy 3 komputera adresata. Do miejsca przeznaczenia pakiet dochodzi zatem w dokładnie takiej samej postaci, jaka nadana mu została przed wysłaniem. Do poziomu warstw trzecich komunikacja pomiędzy warstwami jest praktycznie bezpośrednia.

1.1.4 Podstawy sieci

Siecią nazwać można wszystko, co umożliwia dwóm lub większej liczbie komputerów komunikowanie się ze sobą i/lub z innymi urządzeniami. Dzięki sieciom można wykorzystywać komputery do współdzielenia informacji, do współpracy przy realizacji zadań, do drukowania, a nawet do bezpośredniego komunikowania się za pomocą indywidualnie adresowanych wiadomości.

Sieci składają się z wielu elementów, takich jak sprzęt i oprogramowanie; niektóre z ich składników są niematerialne.
Przed dalszym zagłębieniem się w materię podstawowych składników sieci zauważyć należy, że sieci jako systemy rozwinęły się w dwóch kierunkach określanych obecnie przez dwie odrębne kategorie sieci: sieci lokalne (LAN - Local Area Networks) oraz sieci rozległe (WAN- Wide Area Networks). Rozróżnienie między nimi jest dość proste. Sieci LAN używane są do łączenia urządzeń, które znajdują się w niewielkiej odległości. Sieci WAN służą do łączenia sieci LAN na znaczne odległości.
Trzecią kategorią są sieci miejskie, czyli sieci MAN (ang.. Metropolitan Area Networks). Mimo że pierwotnie zostały one zdefiniowane przez Projekt 802 instytutu IEEE - ten sam, który określa standard sieci LAN, to sieci MAN bliższe są sieciom WAN, aniżeli LAN. Są one nadal rzadko używane i - w związku z tym - mało poznane.
Istnieje oczywiście wiele wyjątków od obu tych prostych definicji sieci, które podaję wyłącznie jako punkt wyjścia; kolejne, coraz dokładniejsze definicje podawane będą na dalszych stronach tej książki.

1.1.4.1 Sprzętowe elementy składowe

Podstawowymi sprzętowymi składnikami sieci są urządzenia trojakiego rodzaju:
• urządzenia transmisji, • urządzenia dostępu,
• urządzenia wzmacniania przesyłanych sygnałów.
Składniki te są niezbędne do funkcjonowania wszelkiego rodzaju sieci. Pozostała część niniejszego podrozdziału opisuje te składniki oraz sposób, w jaki obsługują one sieci LAN i WAN. Bardziej szczegółowe informacje na ich temat znajdują się w dalszych częściach tej książki.
Urządzenia transmisji
Urządzenia transmisji to nośniki używane do transportu sygnałów biegnących przez sieć do ich miejsc docelowych. Nośnikami są kable koncentryczne, skrętka dwużyłowa,
a także kable światłowodowe. Najczęściej stosowane nośniki sieci LAN przedstawione są szczegółowo w rozdziale 3 pt.. „Warstwa fizyczna".
Nośniki LAN mogą również być niematerialne. Nośnikiem tego rodzaju jest na przykład powietrze, przez które przesyłane są światło, fale radiowe, a także mikrofale. Mimo że atmosfera służy za nośnik wszystkim wspomnianym formom transmisji, przedstawianie jej jako rodzaju nośnika jest raczej niepotrzebne. Więcej sensu ma zatem opisywanie nie powietrza przenoszącego transmisje, lecz mechanizmów je generujących. Mechanizmy te opisane są w rozdziale 4 pt.. „Niezupełnie-fizyczna warstwa fizyczna".Z
Sieci WAN również mają swoje urządzenia transmisji. Urządzenia takie są często określane ze względu na częstotliwości ich zegarów i strukturę ramek (na przykład 1,544 Mbps, linia dzierżawiona Frame Relay). Rodzaj nośnika fizycznego nie determinuje ich rzeczywistej wydajności. Owe urządzenia transmisji szczegółowo omówione są w rozdziałach 14 pt.. „Linie dzierżawione"; l5 pt.. „Urządzenia transmisji w sieciach z komutacją obwodów" oraz 16 pt.. „Urządzenia transmisji w sieciach z komutacją pakietów".

1.1.4.1.1 Urządzenia dostępu

Urządzenia dostępu są odpowiedzialne za:
• formatowanie danych w taki sposób, aby nadawały się one do przesyłania w sieci, • umieszczanie w sieci tak sformatowanych danych,
• odbieranie danych do nich zaadresowanych.
W sieci lokalnej (w sieci LAN) urządzeniami dostępu są karty sieciowe (karty interfejsów sieciowych). Karta sieciowa jest płytką drukowaną, którą instaluje się w jednym z gniazd rozszerzeń („slotów") płyty głównej. Karta taka pełni funkcję portu, za pomocą którego komputer przyłączony jest do sieci. Karty sieciowe oprawiają w ramki dane, których wysłania domagają się aplikacje komputera, a następnie umieszczają te dane, mające postać binarną, w sieci, a także odbierają ramki zaadresowane do obsługiwanych przez nie komputerów. Proces oprawiania danych w ramki i umieszczania ich w sieci opisany jest w rozdziale 5 pt.. „Warstwa łącza danych".
W sieciach rozległych (WAN) urządzeniami dostępu są routery. Routery działają na poziomie warstwy 3 modelu OSI i składają się z protokołów dwojakiego rodzaju: protokołów trasowania i protokołów trasowalnych. Protokoły trasowalne, takie jak protokół IP, używane są do transportowania danych poza granice domen warstwy 2. Protokoły te są szczegółowo opisane w rozdziale 12.
Protokoły trasowania natomiast udostępniają wszystkie funkcje niezbędne do:
' Gwoli ścisłości - trudno przypisywać atmosferze rolę ośrodka. w którym rozchodzą się fale elektromagnetyczne: lala elektromagnetyczna jest zjawiskiem samoistnym. a koncepcja hipotetycznego eteru, mającego spełniać rolę wspomnianego ośrodka. została obalona na początku XX w. głównie dzięki wiekopomnemu doświadczeniu Michelsona i Morleya. Być może właśnie ze względu na ów niezwykły charakter fal elektromagnetycznych połączenie radiowe może być postrzegane jako mniej ..materialne' od galwanicznego połączenia za pomocą przewodów - choć w porównaniu z tymi ostatnimi wymaga zwykle większych nakładów-właśnie- materialnych (przyp. red.)
• określania w sieci WAN ścieżek opt.ymalnych dla każdego adresu docelowego, • odbierania pakietów i przesyłanie ich dalej, do miejsca docelowego z wykorzystaniem owych ścieżek.
Sieci WAN używane są do łączenia kilku sieci LAN. Więcej informacji na ten temat znaleźć można w rozdziale 13 pt.. „Korzystanie z sieci WAN".

1.1.4.1.2 Wzmacniaki

Wzmacniak jest urządzeniem, które odbiera przesyłane sygnały, wzmacnia je i wysyła z powrotem do sieci. W sieciach LAN wzmacniak - częściej zwany koncentratorem umożliwia przyłączanie do sieci wielu urządzeń. Funkcja ta jest dla dzisiejszych sieci LAN o tyle istotna, że często zapomina się o pierwotnym zadaniu koncentratorów - regenerowaniu sygnałów.
A zdolności koncentratorów do regenerowania sygnałów decydują o pomyślnym działaniu sieci LAN w równym stopniu, co ich funkcje tworzenia punktów dostępu do sieci. Okrutna rzeczywistość nieubłaganie dostarcza nam dowodów na wpływ nośników na przesyłane

sygnały. Sygnały elektroniczne umieszczone w sieci ulegają bowiem zakłóceniom, które mogą przyjąć jedną z dwóch form: tłumienia lub zniekształcenia.
Tłumienie sygnału to osłabienie jego siły. *Zniekształcenie* natomiast to niepożądana zmiana jego kształtu. Każda ze wspomnianych form zakłóceń musi być traktowana z osobna i z osobna rozwiązywana.
Tłumienie można eliminować zmniejszając długość kabli na tyle, by moc sygnału umożliwiała mu dotarcie do wszystkich części okablowania. Jeśli jednak kabel musi być długi, to aby uniknąć tłumienia, można na kablu zamontować wzmacniak.
Zniekształcenie stanowi poważniejszy problem związany z przesyłaniem sygnałów. Zniekształcenie sygnałów powoduje uszkodzenie wszelkich danych, które są przy ich użyciu przesyłane. Wzmacniaki nie potrafią rozróżniać sygnałów prawidłowych od zniekształconych, wzmacniają więc wszystkie sygnały. Istnieje na szczęście kilka sposobów eliminowania zniekształceń.
Przede wszystkim należy stosować się do wszelkich zaleceń dotyczących nośnika - szczegółowe informacje na ten temat znajdują się w rozdziale 3 pt.. „Warstwa fizyczna".
W razie wystąpienia zniekształcenia należy określić jego źródło, a następnie przeprowadzić okablowanie jak najdalej od niego. Często zniekształceń uniknąć można, stosując nowoczesne technologie transmisji, które są odporne na zakłócenia, takie jak na przykład kable światłowodowe. Wszelkie dostępne technologie omówione są na dalszych stronach tej książki.
Korzystać można z protokołów sieciowych umożliwiających rozpoznawanie i automatyczną korektę wszelkich ewentualnych błędów transmisji. Protokoły te omówione są w rozdziałach 5 i 12.

1.1.4.2 Programowe elementy składowe

Składnikami programowymi niezbędnymi do utworzenia sieci są:
protokoły - określające sposoby komunikowania się urządzeń i regulujące je,
programy poziomu sprzętowego, nazywane mikroprogramami, sterownikami lub programami obsługi - umożliwiające działanie urządzeniom, takim jak na przykład karty sieciowe,
oprogramowanie komunikacyjne.
Protokoły
Przyłączalność fizyczna sieci jest łatwa do zabezpieczenia. Prawdziwe wyzwanie stanowi natomiast zorganizowanie standardowych sposobów komunikowania się zarówno dla komputerów, jak i dla innych urządzeń przyłączanych do sieci. Sposoby, o których mowa, nazywane są protokoiami. Protokołem posługujemy się na przykład podczas komunikowania się za pomocą telefonu. Zazwyczaj pierwszym wyrażeniem wypowiadanym po podniesieniu słuchawki jest „Halo" (lub jego odpowiednik). Zwykłe pozdrowienie informuje o pomyślnym ustanowieniu połączenia. Następną czynnością jest zwykle odpowiedź potwierdzająca, że łącze komunikacyjne działa w obu kierunkach. Jeśli strony potrafią się rozpoznać w wyniku tej prostej wymiany dwóch słów, wystarcza ona do nawiązania najbardziej nawet intymnych rozmów. Jeśli jednak osoby po przeciwległych stronach kabla telefonicznego nie znają się nawzajem, to do rozpoznania potrzebne są dodatkowe protokoły. Po ustanowieniu połączenia i rozpoznaniu się rozmowa może potoczyć się w kierunku, w jakim została zainicjowana.
Przykładowy protokół zakorzenił się w zwyczajach na tyle, że jego naruszenie interpretowane jest jako brak wychowania lub nawet umyślna nieuprzejmość.
W tej kwestii komputery nie różnią od nas ani trochę. Połączenie ich za pomocą sieci jest zaledwie jednym z wymogów, które muszą być spełnione w celu pomyślnej realizacji rzeczowej komunikacji i udostępniania zasobów. Bezpośrednia komunikacja między dwoma komputerami umożliwia im, a zatem również ich użytkownikom, współdzielenie zasobów. Zakładając, że jedynie garstka osób nie współpracuje podczas pracy z innymi, umożliwienie komputerom współdzielenia informacji oraz innych zasobów stanowiło „skok przez płot" w dziedzinie infrastruktury technologii informacyjnych zbliżający sposób współpracy między komputerami do naturalnego sposobu współpracy grup ludzi.
Protokoły dla sieci LAN nazywane są często architekturami LAN, jako że zawierają one również karty sieciowe. Determinują one w znacznym stopniu kształt, rozmiar oraz mechanikę sieci LAN, które opisane są w części 2 niniejszej książki pt.. „Tworzenie sieci lokalnych" (rozdziały 7 do 12).
Protokoły dla sieci WAN zwykle dostarczane są grupowo i to właśnie dzięki nim korzystać możemy z takiej różnorodności usług sieci rozległych. Protokoły te są dokładnie opisane w rozdziale 12.

1.1.4.2.1 Sterowniki urządzeń

Sterownik urządzenia jest programem poziomu sprzętowego umożliwiającym sterowanie określonym urządzeniem. Sterownik urządzenia można porównać do miniaturowego systemu operacyjnego obsługującego jedno tylko urządzenie. Każdy sterownik zawiera całą logikę oraz wszystkie dane, które są niezbędne do odpowiedniego funkcjonowania obsługiwanego urządzenia. W przypadku karty sieciowej sterownik dostarcza interfejsu dla systemu operacyjnego hosta.
Sterownik zwykle umieszczony jest w oprogramowaniu sprzętowym obsługiwanego urządzenia.

1.1.4.2.2 Oprogramowanie komunikacyjne

Wszystkie uprzednio wspomniane sprzętowe i programowe składniki sieci nie wystarczą do korzystania z niej. Tworzą one jedynie (lub „aż") infrastrukturę oraz mechanizmy pozwalające na korzystanie z sieci. Samo korzystanie z sieci odbywa się pod kontrolą specjalnego oprogramowania sterującego komunikacją.
Zalety „informatyki myszkowej" (przeciągnij i upuść - ang.. *drag and drop)* sprawiły, że oprogramowanie komunikacyjne stało się obecnie tak proste, iż bardzo często użytkownik korzysta z programu komunikacyjnego wcale o tym nie wiedząc. Przykładami takich właśnie prostych programów komunikacyjnych są programy „mapowania" dysków lub udostępniania obszarów (np.. dysków, folderów czy plików)

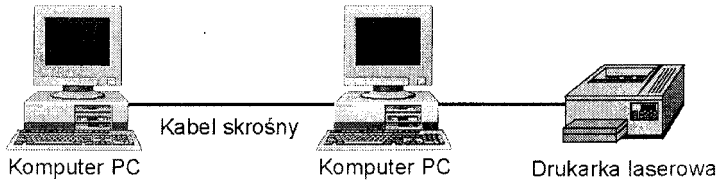
w Windows NT. Innymi, nieco bardziej oczywistymi przykładami, są sieć WWW, protokół HTTP, telnet, tn3270, programy przesyłania plików, a nawet poczta elektroniczna.
Niezależnie od typu aplikacji komunikacyjnej oraz stopnia jej złożoności, to ona jest mechanizmem, który sprawia, że można korzystać z pasma przesyłania utworzonego i udostępnionego przez wcześniej wspomniane składniki sieci.

1.1.4.2.3 Składanie elementów w sieć

Sprzętowe i programowe elementy składowe sieci nawet razem z ich funkcjami nadal nie stanowią sieci. Jest ona tworzona dopiero w wyniku ich zintegrowania. Poniższe przykłady ilustrują niektóre ze sposobów, w jakie składniki sprzętowe i programowe mogą być łączone w celu utworzenia z nich prostych sieci. Ukazują one przy tym dowolność, jaka może towarzyszyć tworzeniu sieci.

1.1.4.2.4 Sieci LAN bez wzmacniaków

Dwa komputery - wyposażone w zgodne karty sieciowe - mogą komunikować się ze sobą bezpośrednio bez użycia wzmacniaka (nazywanego również koncentratorem) przy założeniu, że komputery te znajdują się w niewielkiej odległości. Zależność tę przedstawia rysunek 1.6.
Rysunek 1.6. Dwa komputery komunikujące się bezpośrednio be. u:ycia koncentratora.

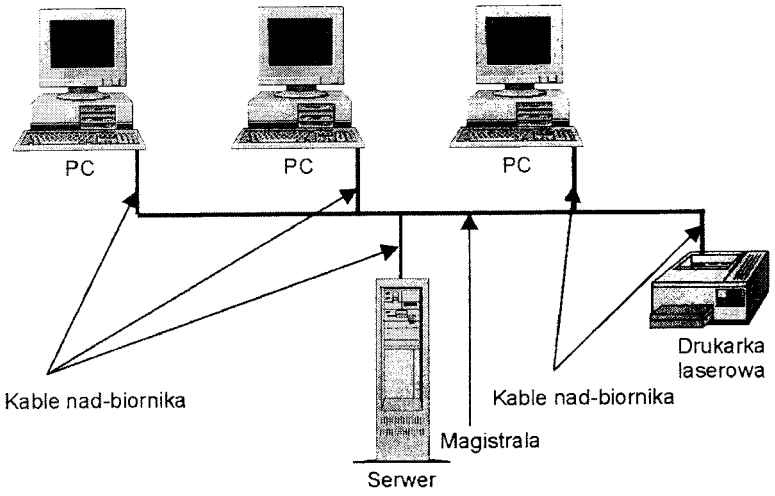


Przykład przedstawiony na rysunku 1.6 podany jest przy założeniu, że kable umożliwiają skrzyżowanie ścieżek wysyłania i odbierania. Ma to miejsce jedynie w przypadku okablowania skrętką dwużyłową. Kabel koncentryczny używa tych samych ścieżek fizycznych zarówno do wysyłania, jak i odbierania. Skrętka dwużyłowa używa dwóch lub więcej par kabli. Kabel podwójny używałby jednej pary przewodów do wysyłania i jednej do odbierania danych. W każdej parze jeden przewód służy do wysyłania sygnałów o napięciu dodatnim, a drugi - ujemnym.
Zakładając, że interfejsy takie są standardowe, 4-parowa skrętka dwużyłowa przeciągnięta między dwoma komputerami powodowałaby, że oba próbowałyby przesyłać dane tą samą parą kabli. Na jednej i tej samej parze przewodów oczekiwałyby również nadejścia ramek. Nie byłyby więc zdolne do komunikowania się. Kabel skrośny krzyżuje pary przewodów służące do wysyłania i odbierania sygnałów, tak że para, której jedno urządzenie używa do wysyłania - jest również parą, na której drugie urządzenie oczekuje transmisji przychodzącej.
Magistralowe sieci LAN
Najprostszy typ sieci LAN oparty jest na magistrali. Nazywa się go magistralową siecią lokalną. Magistrala jest siecią, która do komunikacji w sieci używa karty interfejsu sieciowego (karty sieciowej). Niektóre technologie LAN korzystają z topologii magistrali jako rodzimej części ich protokołu. Innym sposobem utworzenia sieci magistralowej jest usunięcie koncentratora z sieci koncentratorowej (opartej na koncentratorze).
Magistralowa sieć lokalna składa się z następujących elementów składowych:
• medium transmisyjnego, czyli nośnika (jest nim magistrala),
• interfejsu fizycznego lub nad-biornika dla każdego urządzenia przyłączanego do sieci,
• protokołów transmisji i komunikacji,
• oprogramowania umożliwiającego użytkownikom komunikowanie się i udostępnianie zasobów.
Powyższa lista (jak również rysunek 1.7) wskazuje, że sieci magistralowe LAN nie używają wzmacniaka sygnałów. Automatycznie ogranicza to zarówno maksymalne odległości między urządzeniami, jak również ich liczbę.
Rysunek 1.7. Magistralowa sieć lokalna (Bus Lan.
Ograniczenia narzucane przez magistralową sieć LAN. w połączeniu ze względnie niskimi cenami koncentratorów, sprawiły, że dziś jest ona zapomnianym typem sieci. Jednak uznać ją należy jako prawowitą topologię sieci, która dobrze ilustruje funkcjonalność podstawowych składników sieci.

1.1.4.2.5 Sieci oparte na koncentratorze (koncentratorowe)

Koncentrator jest urządzeniem, które, jak sama nazwa wskazuje, znajduje się w centrum sieci. Przykład prostej koncentratorowej sieci lokalnej przedstawiony jest na rysunku 1.8. Sieć LAN oparta na koncentratorach posiada kilka zalet nad siecią tego samego typu, lecz opartą na magistrali. Pierwsza wynika z faktu, że koncentrator wzmacnia odbierany sygnał, umożliwiając tworzenie rozleglejszych sieci LAN, aniżeli jest to możliwe na bazie architektury magistrali. Druga polega na tym, że koncentrator można łączyć z innymi koncentratorami, zwiększając w ten sposób znacznie liczbę urządzeń, które można przyłączyć do sieci lokalnej. Łańcuchowe łączenie koncentratorów pozwala na zwiększanie odległości, na jakie rozciągać można tego rodzaju sieci.
Nie wszystkie koncentratory są jednocześnie wzmacniakami. Niektóre koncentratory to nie-wzmacniające agregatory stacji komputerowych. Są one stosowane tylko do nadawania sieci LAN topologii gwiazdy oraz do łączenia innych koncentratorów. Umożliwiaj przyłączanie do sieci większej liczby użytkowników, ale nie wzmacniają sygnału, czyli nie przyczyniają się do rozszerzenia sieci w przestrzeni.
Koncentratory (czyli wzmacniaki) zostały pomyślnie przystosowane do architektur sieci lokalnych opartych na magistrali, dzięki czemu możliwe stało się opracowanie nowej topologii: topologii magistrali gwiaździstej.

Do utworzenia koncentratorowej sieci LAN prócz koncentratora (lub wzmacniaka) potrzebne są wszystkie podstawowe składniki niezbędne do utworzenia magistralowej sieci LAN. Są one przedstawione w podpunkcie poprzednim - „Magistralowe sieci LAN".



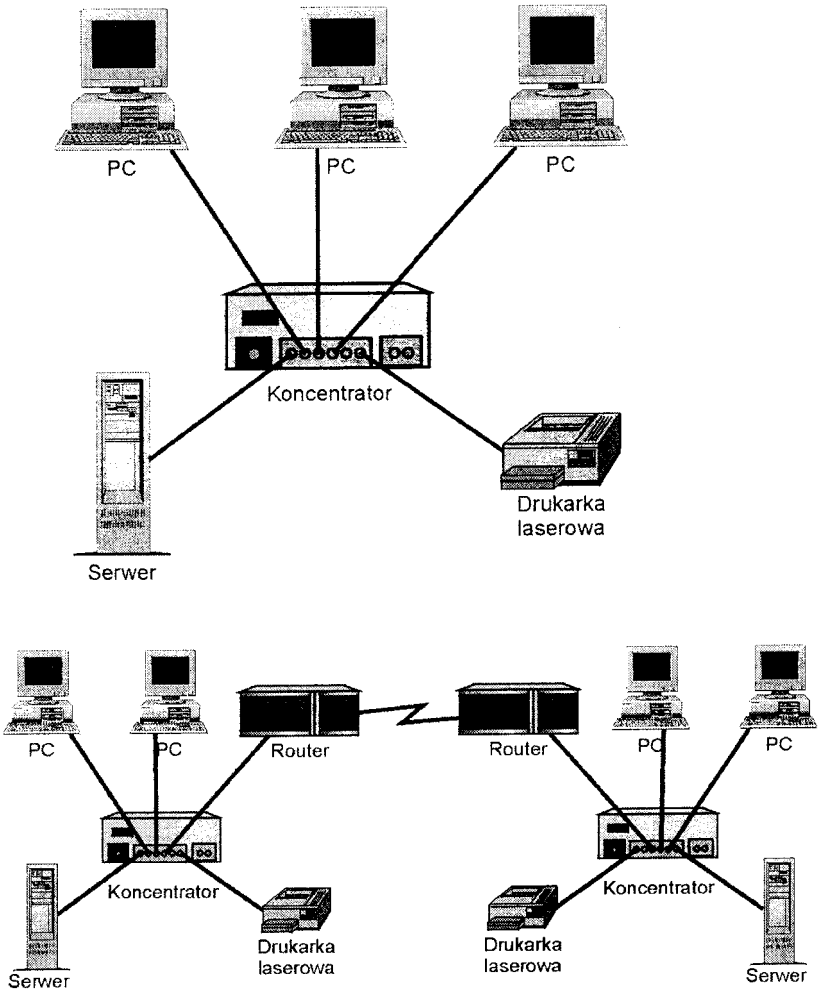
Rysunek 1.8. Koncentratorowa sieć lokalna.

Sieci WAN

Sieci rozległe, czyli sieci WAN, łączą sieci LAN za pomocą urządzeń znanych jako routery. Routery umożliwiają użytkownikom jednej sieci LAN dostęp do zasobów innych, połączonych z nią sieci LAN bez potrzeby utraty własnej odrębności. Innymi słowy, router łączy sieci LAN bez scalania ich w jedną, większą sieć lokalną. Dwie sieci LAN połą czone za pomocą routera tworzą sieć WAN przedstawioną na rysunku 1.9.

Rysunek 1.9. Sieć rozległa (sieć WAN).

Przykładowa sieć WAN mogłaby rozciągać się na dowolną odległość. Sieci LAN mogłyby znajdować się na różnych piętrach tego samego budynku lub na różnych kontynentach. Do rozciągnięcia sieci WAN na rozległym geograficznie obszarze niezbędne jest zastosowanie wielu wzmacniaków.



Inne są wzmacniaki służące do obsługi sieci LAN i inne dla sieci WAN. Wzmacniaki sieci LAN najczęściej nazywane są koncentratorami. Wzmacniaki sieci WAN najczęściej są niewidoczne zarówno dla użytkowników, jak i dla administratorów tych sieci. Stanowią bowiem całkowicie zintegrowaną część infrastruktury komercyjnej, która nadaje sieciom WAN właściwość przyłączalności.

1.1.5 Podsumowanie

Podstawowe składniki sieci przedstawione w niniejszym rozdziale, ich funkcje i zastosowania stanowią dopiero przyczółek, z którego zdobywać będziemy dalsze zagony wiedzy. Choć dokładniej rzecz biorąc, tworzą one fundament, na którym budować będziemy dalsze piętra naszej struktury wiedzy. Składniki te zostały przedstawione w celu ukazania historii rozwoju sieci oraz wyjaśnienia pewnych podstawowych pojęć i definicji właściwych dla sieci.
Podstawy te omówione są w kontekście różnych standardów przemysłowych obecnie obowiązujących w dziedzinach związanych z sieciami. Na stronach tej książki. w celu ułatwienia zrozumienia podstawowych zagadnień sieci, szczególnie często sięgać będziemy do modelu referencyjnego OSI, który posłuży nam za płaszczyznę, w odniesieniu do której porównywane będą wszystkie inne standardy.

1.2 Rozdział 2 Typy i topologie sieci LAN

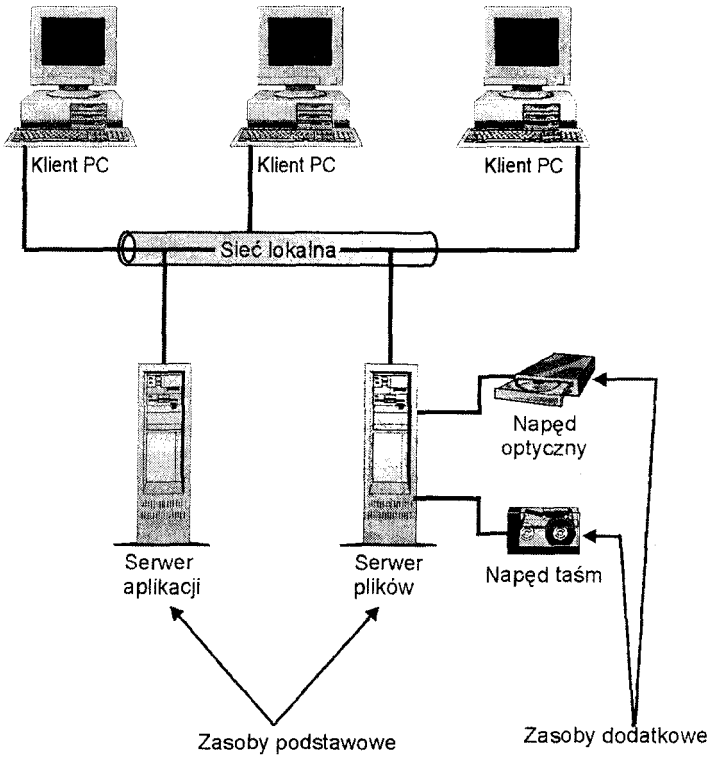
Mark A. Sportack
Sieci lokalne (sieci LAN) rozpowszechniły się do dziś w bardzo wielu - zwłaszcza komercjalnych - środowiskach. Mimo że większość z nas miała już większą lub mniejszą styczność z sieciami, to niewiele osób wie, czym one są i w jaki sposób działają. Łatwo wskazać koncentratory czy przełączniki i powiedzieć, że to one są siecią. Ale one są jedynie częściami pewnego typu sieci.
Jak zostało to wyjaśnione w rozdziale 1 pt.. „ABC sieci", sieci lokalne najłatwiej zrozumieć po rozłożeniu ich na czynniki pierwsze. Często składniki sieci dzielone są na *warstwy* w sposób określony przez model referencyjny OSI, który szczegółowo przedstawiony został w rozdziale 1. Każda warstwa tego modelu obsługuje inny zbiór funkcji.
Niezbędnym warunkiem wstępnym podziału sieci lokalnej na warstwy jest poznanie dwóch jej atrybutów: metodologii dostępu do sieci oraz topologii sieci. Metodologia dostępu do zasobów sieci LAN opisuje sposób udostępniania zasobów przyłączanych do sieci. Ten aspekt sieci często decyduje o jej typie. Dwoma najczęściej spotykanymi typami są: „każdy-z-każdym" oraz „klient-serwer".
Natomiast topologia sieci LAN odnosi się do sposobu organizacji koncentratorów i okablowania. Topologiami podstawowymi sieci są: topologia magistrali, gwiazdy, pierścienia oraz przełączana (czyli komutowana). Wspomniane atrybuty tworzą zarys ułatwiający rozróżnianie warstw funkcjonalnych sieci LAN. Rozdział niniejszy bada wszystkie możliwe kombinacje typów i topologii sieci LAN. Przedstawione są również ich ograniczenia, korzyści z nich płynące oraz potencjalne zastosowania.

1.2.1 Urządzenia przyłączane do sieci LAN

Przed zanurzeniem się w morzu typów i topologii sieci LAN, warto zapoznać się z niektórymi podstawowymi zasobami dostępnymi w sieci LAN. Trzema najbardziej powszechnymi urządzeniami podstawowymi są klienci, serwery oraz drukarki. Urządzeniem podstawowym jest takie urządzenie, które może uzyskać bezpośredni dostęp do innych urządzeń lub umożliwić innym urządzeniom dostęp do siebie.
Serwer to dowolny komputer przyłączony do sieci LAN, który zawiera zasoby udostępniane innym urządzeniom przyłączonym do tej sieci. *Klient* to natomiast dowolny komputer, który za pomocą sieci uzyskuje dostęp do zasobów umieszczonych na serwerze. Drukarki są oczywiście urządzeniami wyjścia tworzącymi wydruki zawartości plików. Do wielu innych urządzeń, takich jak napędy CD-ROM oraz napędy taśm, również można uzyskać dostęp za pomocą sieci, ale mimo to są one urządzeniami dodatkowymi. Wynika to z faktu, że bezpośrednio przyłączone są one do urządzeń podstawowych. Urządzenie dodatkowe (na przykład CD-ROM) znajduje się więc w stosunku *podporządkowania* wobec urządzenia podstawowego (na przykład serwem). Drukarki mogą być podporządkowanymi urządzeniom podstawowym urządzeniami dodatkowymi lub - w razie bezpośredniego podłączenia ich do sieci - urządzeniami podstawowymi. Rysunek 2.1 przedstawia zasoby podstawowe, które można znaleźć w sieci lokalnej, a także zależności między zasobami podstawowymi i dodatkowymi.

1.2.2 Typy serwerów

„Serwer" to słowo ogólnie określające komputer wielodostępny (z którego jednocześnie korzystać może wielu użytkowników). Warto zauważyć. że serwery nie są identyczne, lecz stanowią grupę różnorodnych komputerów. Zwykle wyspecjalizowane są w wykonywaniu określonych funkcji, na które wskazuje przymiotnik dołączany do ich nazwy. Wyróżnić więc można serwery plików, serwery wydruków, serwery aplikacji i inne.
Rysunek 2.1. Podstawowe i dodatkowe =asobv sieci Lan.



Serwery plików

Jednym z podstawowych i dobrze znanych rodzajów serwerów jest serwer plików. Serwer plików jest scentralizowanym mechanizmem składowania plików, z których korzystać mogą grupy użytkowników. Składowanie plików w jednym miejscu zamiast zapisywania ich w wielu różnych urządzeniach klienckich daje wiele korzyści, takich jak:

• Centralna lokalizacja -wszyscy użytkownicy korzystają z jednego, ustalonego magazynu współdzielonych plików. To z kolei daje korzyści dwojakiego rodzaju. Użytkownicy nie muszą przeszukiwać wielu miejsc, w których potencjalnie zapisany mógł zostać potrzebny im plik; pliki są bowiem składowane w jednym miejscu. Zwalnia to również użytkowników z obowiązku logowania się osobno do każdego z tych wielu miejsc i, tym samym, z konieczności pamiętania wielu różnych haseł dostępu. Dzięki temu dostęp do wszystkich potrzebnych plików umożliwia jedno wyłącznie logowanie (jedna rejestracja).

• Zabezpieczenie źródła zasilania - składowanie plików w centralnym serwerze umożliwia również wprowadzanie wielu technik pozwalających na ochronę danych przed zakłóceniami w dostawach prądu elektrycznego. Zmiany częstotliwości lub nawet całkowita przerwa w zasilaniu mogą uszkodzić zarówno dane, jak i sprzęt. Filtracja prądu oraz bateryjne zasilanie dodatkowe, możliwe dzięki zastosowaniu urządzeń „UPS" nieprzerywalnego podtrzymywania napięcia (ang.. Uninterruptible. Power Supply.y), jest dużo efektywniejsze ekonomicznie w przypadku stosowania ich względem jednego serwera. Podobny sposób ochrony w sieci równorzędnej spowodowałby nadmierne (w stopniu nie do zaakcept.owania) podniesienie kosztów ze względu na większą liczbę komputerów wymagających ochrony.

• Zorganizowane archiwizowanie danych - składowanie udostępnianych plików w jednym, wspólnym miejscu znacznie ułatwia tworzenie ich kopii zapasowych; wystarcza bowiem do tego celu jedno tylko urządzenie i jedna procedura działania. Zdecentralizowane przechowywanie danych (na przykład w każdym komputerze osobno) oznacza, że kopie zapasowe danych znajdujących się w każdym komputerze musiałyby być tworzone indywidualnie. Kopie zapasowe są podstawowym sposobem ochrony przed utratą lub uszkodzeniem plików. Urządzeniami służącymi do tworzenia kopii zapasowych są napędy taśm, napędy dysków opt.ycznych, a nawet napędy dysków twardych. Do zapisywania kopii zapasowych używać można również wielu napędów dysków, która to technika znana jest jako *porcjowanie.* Porcjowanie polega na wielokrotnych, jednoczesnych zapisach dokonywanych na różnych dyskach twardych. Mimo że technika ta używana jest głównie w celu uzyskania szybszego odczytu danych, to może ona być również stosowana do tworzenia nowych kopii zapasowych podczas każdej operacji zapisu.

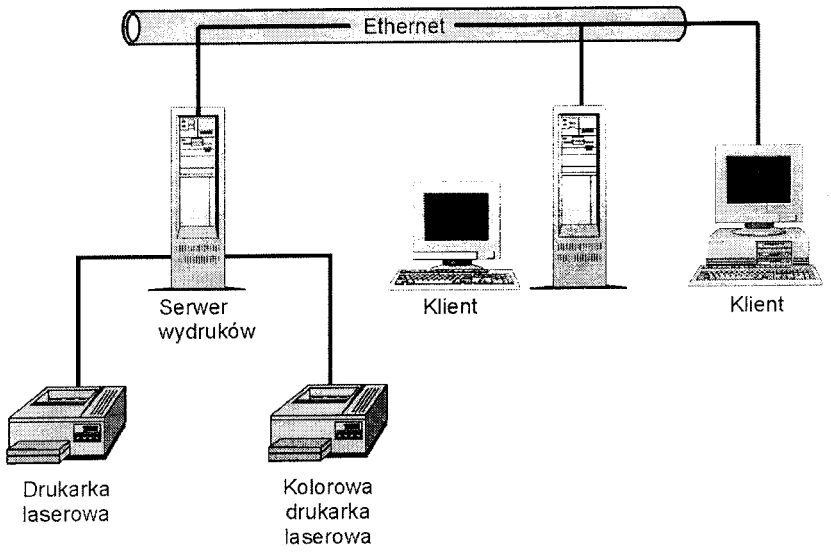
• Szybkość - standardowy serwer stanowi dużo bardziej niż typowy komputerklient niezawodną i w pełni konfigurowalną platformę. Przekłada się to bezpośrednio na znaczną, w stosunku do sieci równorzędnej, poprawę wydajności odczytywania plików.

Zastosowanie serwera plików nie zawsze powoduje zwiększenie szybkości obsługi plików. Dostęp do danych zapisanych lokalnie (czyli w tym samym komputerze, który je odczytuje) uzyskać można dużo szybciej niż do plików zapisanych w komputerze zdalnym i pobieranych za pośrednictwem sieci lokalnej. Wspomniane zwiększenie szybkości zależy od szybkości, z jaką pliki mogą być uzyskiwane z innych urządzeń przyłączonych do sieci równorzędnej, a nie od szybkości, z jaką pliki mogą być pobierane z lokalnego dysku twardego.

Nie wszystkie pliki nadają się do przechowywania w serwerze plików. Pliki prywatne, zastrzeżone i nie nadające się do użycia przez osoby korzystające z sieci najlepiej zostawić na lokalnym dysku twardym. Przechowywanie ich w serwerze plików daje wszystkie uprzednio opisane korzyści, ale może też być powodem niepotrzebnych zagrożeń.

1.2.2.1 Serwery wydruków

Serwery mogą być również używane do współdzielenia drukarek przez użytkowników sieci lokalnej. Mimo że ceny drukarek, zwłaszcza laserowych, od czasu wprowadzenia ich na rynek uległy znacznemu obniżeniu, to w mało której organizacji potrzebna jest drukarka na każdym biurku. Lepiej więc korzystać, za pośrednictwem serwera wydruków, z kilku lub nawet jednej tylko drukarki, udostępniając je (ją) w ten sposób każdemu użytkownikowi sieci.
Jedyną funkcją serwerów wydruków jest przyjmowanie żądań wydruków ze wszystkich urządzeń sieci, ustawianie ich w kolejki i „spoolowanie" (czyli wysyłanie) ich do odpowiedniej drukarki. Proces ten jest przedstawiony na rysunku 2.2.
Rysunek 2.2. Prosty serwer wydruków



Słowo „spool" jest akronimem wyrażenia „Simultaneous Peripheral Operations On Line", oznaczającego po polsku „współbieżne bezpośrednie operacje peryferyjne". Operacje te polegają na tymczasowym buforowaniu (magazynowaniu) na nośniku magnetycznym programów i danych w postaci strumieni wyjściowych w celu późniejszego ich wyprowadzenia lub wykonania.
Każda drukarka przyłączona do serwera wydruków ma swoja własną listę kolejności, czyli kolejkę, która informuje o porządku, w jakim wszystkie żądania są tymczasowo zapisywane i czekają na wydrukowanie. Żądania zwykle przetwarzane są w kolejności, w jakiej zostały otrzymane. Systemy operacyjne klientów, takie jak Windows 95 oraz Windows NT dla stacji roboczej (NT Workstation) firmy Microsoft umożliwiają udostępnianie (współdzielenie) drukarek.
Drukarkę można do sieci LAN przyłączyć również bezpośrednio, czyli bez pośrednictwa serwera wydruków. Umożliwiają to karty sieciowe, za pomocą których drukarki mogą być konfigurowane tak, aby były jednocześnie serwerami kolejek wydruków. W takiej sytuacji serwer wydruków nie jest potrzebny, a wystarczy jedynie przyłączyć odpowiednio skonfigurowane drukarki bezpośrednio do sieci LAN - będzie to wystarczające, o ile nie zamierzamy zbytnio obciążać ich czynnościami drukowania.

1.2.2.2 Serwery aplikacji

Równie często serwery służą jako centralne składy oprogramowania użytkowego. Serwery aplikacji, mimo że na pierwszy rzut oka podobne do serwerów plików, różnią się jednak od nich znacznie. Serwer aplikacji jest miejscem, w którym znajdują się wykonywalne programy użytkowe. Aby móc uruchomić określony program, klient musi nawiązać w sieci połączenie z takim serwerem. Aplikacja jest następnie uruchamiana, ale nie na komputerze-kliencie, lecz na rzeczonym serwerze. Serwery umożliwiające klientom pobieranie kopii programów do uruchomienia na komputerach lokalnych to serwery plików. Pliki w ten sposób wysyłane są co prawda plikami aplikacji, ale serwery spełniają w takim przypadku funkcje serwerów plików.
Serwery aplikacji umożliwiają organizacji zmniejszenie kosztów zakupu oprogramowania użytkowego. Koszty nabycia i konserwacji jednej, wielodostępnej kopii programu są zwykle dużo niższe od kosztów nabycia i konserwacji kopii instalowanych na pojedynczych komputerach.
Instalowanie zakupionego pakietu oprogramowania użytkowego przeznaczonego dla pojedynczej stacji w serwerze plików może być niezgodne z warunkami umowy jego zakupu. W podobny bowiem sposób, w jaki użytkownicy mogą przekazywać sobie nośnik z oryginalną wersją oprogramowania, również serwer udostępnia pojedynczą kopię programu wszystkim użytkownikom sieci, do której jest przyłączony. Uznawane jest to za piractwo komputerowe. Warto zatem upewnić się, że każdy pakiet oprogramowania instalowany na serwerze zakupiony został na podstawie umowy umożliwiającej korzystanie z niego wielu użytkownikom.
Mimo że w zasadzie oprogramowanie użytkowe warto składować na innych serwerach niż pliki danych (na przykład aplikacje na serwerze aplikacji, a pliki na serwerze plików s. od zasady tej odnotować należy jeden, istotny wyjątek.
Niektóre aplikacje tworzą i obsługują duże relacyjne bazy danych. Aplikacje te i ich bazy danych powinny znajdować się obok siebie w serwerze aplikacji.
Przyczyna tego jest dość prosta: zasady pobierania danych z bazy danych różnią się znacznie od sposobu korzystania z plików Worda czy Excela. Aplikacja relacyjnej bazy danych udostępnia żądane dane, zatrzymując wszystkie pozostałe w bazie danych. Aplikacje automatyzujące pracę w biurze takie jak Word czy Excel, wszystkie informacje zapisują w odrębnych plikach, które zwykle nie są wzajemnie zależne, a na pewno nie w złożony sposób. Inaczej w przypadku aplikacji relacyjnych baz danych, które są bezpośrednio

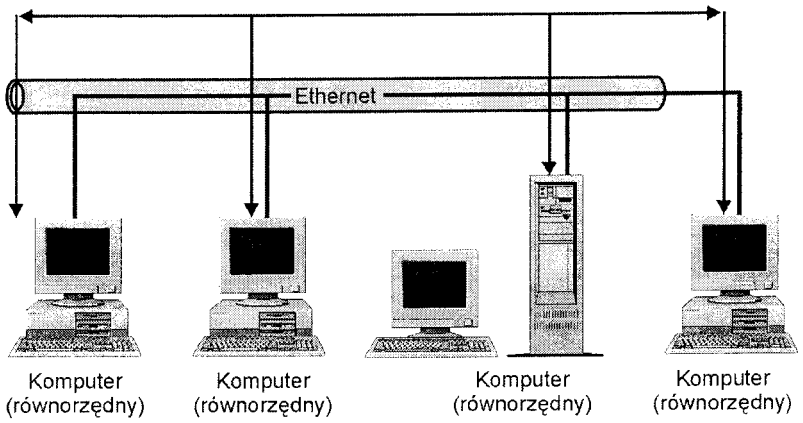
odpowiedzialne za integralność zarówno samych baz danych, jak i ich indeksów. A zarządzanie bazami danych w sieci zwiększa ryzyko uszkodzenia ich indeksów i całej aplikacji.

1.2.3 Typy sieci

Typ sieci opisuje sposób, w jaki przyłączone do sieci zasoby są udostępniane. Zasobami mogą być klienci, serwery lub inne urządzenia, pliki itd.., które do klienta lub serwera są przyłączone. Zasoby te udostępniane są na jeden z dwóch sposobów: równorzędny i serwerowy.

1.2.3.1 Sieci równorzędne (każdy-z-każdym)

Sieć typu każdy-z-każdym obsługuje nieustrukturalizowany dostęp do zasobów sieci. Każde urządzenie w tego typu sieci może być jednocześnie zarówno klientem, jak i serwerem. Wszystkie urządzenia takiej sieci są zdolne do bezpośredniego pobierania danych, programów i innych zasobów. Innymi słowy, każdy komputer pracujący w takiej sieci jest równorzędny w stosunku do każdego innego - w sieciach tego typu nie ma hierarchii.
Rysunek 2.3 przedstawia sieć typu każdy-z-każdym.
Korzyści
Korzystanie z sieci równorzędnej daje cztery główne korzyści
• Sieci typu każdy-z-każdym są w miarę łatwe do wdrożenia i w obsłudze. Są bowiem niczym więcej jak tylko zbiorem komputerów-klientów, obsługiwanych przez sieciowy system operacyjny umożliwiający udostępnianie równorzędne. Tworzenie sieci każdy-z-każdym wymaga jedynie dostarczenia i zainstalowania koncentratora (lub koncentratorów) sieci LAN, komputerów, okablowania oraz systemu operacyjnego pozwalającego na korzystanie z tej metody dostępu do zasobów.
Rysunek 2.3. Śieć typu każdy-z-każdvm.
1 Sieci typu każdy-z-każdym są bardzo tanie w eksploatacji**.** Nie wymagają one drogich i skomplikowanych serwerów dedykowanych, nad którymi należy roztaczać administracyjną opiekę i które trzeba klimatyzować. Brak dedykowanych serwerów eliminuje również towarzyszące im wydatki związane z zatrudnianiem i szkoleniem pracowników, jak również z dodatkowymi kosztami tworzenia pomieszczeń klimatyzowanych wyłącznie dla serwerów. Każdy komputer znajduje się przy biurku lub na nim, a pod opieką korzystającego zeń użytkownika.



• Sieci typu każdy-z-każdym mogą być ustanawiane przy wykorzystaniu prostych systemów operacyjnych, takich jak Windows for Workgroups, Windows95 czy Windows NT.
• Brak hierarchicznej zależności sprawia, że sieci każdy-z-każdym są dużo odporniejsze na błędy aniżeli sieci oparte na serwerach. Teoretycznie w sieci typu klient-serwer serwer jest pojedynczym punktem defektu. Pojedyncze punkty defektu są miejscami, których niesprawność spowodować może awarię całej sieci. W sieciach typu każdy-z-każdym uszkodzenie jednego komputera powoduje niedostępność jedynie przyłączonej do niego części zasobów sieci.
Ograniczenia
Sieci każdy-z-każdym niosą ze sobą również ryzyko i ograniczenia. Niektóre z nich dotyczą sfer bezpieczeństwa, wydajności i administracji.
Sieć każdy-z-każdym charakteryzuje się następującymi słabościami z zakresu bezpieczeństwa.
• Użytkownicy muszą pamiętać wiele haseł, zwykle po jednym dla każdego komputera wchodzącego w skład sieci.
• Brak centralnego składu udostępnianych zasobów zmusza użytkownika do samodzielnego wyszukiwania informacji. Niedogodność ta może być ominięta za pomocą metod i procedur składowania, przy założeniu jednak, że każdy członek grupy roboczej będzie się do nich stosować. Użytkownicy obmyślają
bardzo twórcze sposoby radzenia sobie z nadmiarem haseł. Większość tych sposobów bezpośrednio obniża bezpieczeństwo każdego komputera znajdującego się w sieci równorzędnej.
• Jak każdy zasób sieciowy, również bezpieczeństwo jest w sieci równorzędnej rozdysponowane równomiernie. Na środki bezpieczeństwa charakterystyczne dla tego typu sieci zwykle składają się: identyfikacja użytkownika za pomocą identyfikatora ID i hasła oraz szczegółowe zezwolenia dostępu do określonych zasobów. Struktura zezwoleń dla wszystkich pozostałych użytkowników sieci zależy od „administratora" komputera, dla jakiego są one ustalane.
Mimo że użytkownik każdego komputera w sieci równorzędnej uważany może być za jego administratora, rzadko kiedy posiada on wiedzę i umiejętności potrzebne do sprawnego wykonywania czynności administracyjnych. Jeszcze rzadziej zdarza się, by poziom tych umiejętności

był równy dla małej nawet grupy roboczej. Owa nierówność często staje się przyczyną wielu problemów występujących podczas korzystania z sieci typu każdy-z-każdym.

• Niestety, umiejętności techniczne uczestników grupy roboczej nie są zwykle jednakowe. W związku z tym bezpieczeństwo całej sieci jest wprost proporcjonalne do wiedzy i umiejętności jej technicznie najmniej biegłego uczestnika. Jedną z lepszych metafor, za pomocą której opisać można taką sytuację jest porównanie sieci równorzędnej do łańcucha. Łańcuch mianowicie jest tak mocny, jak jego najsłabsze ogniwo. Również sieć typu każdy-z-każdym jest tak bezpieczna, jak jej najsłabiej zabezpieczony komputer.

Mimo że obciążenie czynnościami administracyjnymi w sieci każdy-z-każdym jest mniejsze niż w sieci klient-serwer, to jest ono rozłożone na wszystkich członków grupy. Jest to przyczyną powstawania niektórych problemów logistycznych. Najpoważniejszymi z nich są:

• Nieskoordynowane i niekonsekwentne tworzenie kopii zapasowych danych oraz oprogramowania. Każdy użytkownik jest odpowiedzialny za swój komputer, w związku z czym jest bardzo możliwe, że każdy będzie wykonywać kopie zapasowe plików w czasie wolnym. wtedy gdy sobie o tym przypomni i gdy będzie mu się chciało (ci, co to już robili, wiedzą, że nie zawsze się chce...).

• Zdecentralizowana odpowiedzialność za trzymanie się ustalonych konwencji nazywania i składowania plików. Zakładając, że nie ma jednego miejsca, w którym informacja byłaby centralnie składowana, ani innego systemu logicznego, za pomocą którego zorganizowane byłyby zasoby przyłączone do sieci LAN, orientowanie się w tym, co jest zapisywane w jakim miejscu może stanowić nie lada wyzwanie.

Jak wszystko inne w sieciach każdy-z-każdym, również efektywność całej sieci zależna jest bezpośrednio od stopnia, do jakiego ustalone metody i procedury są przestrzegane przez wszystkich jej uczestników.

Mniejsza jest również wydajność tego typu sieci, czego przyczyną jest wielodostępność każdego z komputerów tworzących sieć równorzędną. Komputery standardowe, z jakich

zwykle składa się sieć każdy-z-każdym, przeznaczone są bowiem do użytku jako klienci przez pojedynczych użytkowników, w związku z czym nie są najlepiej dostosowane do obsługi wielodostępu. Ze względu na to, wydajność każdego komputera postrzegana przez jego użytkownika zmniejsza się zauważalnie zawsze, gdy użytkownicy zdalni współdzielą jego zasoby.

Pliki i inne zasoby danego hosta są dostępne tylko na tyle, na ile dostępny jest ów host. Innymi słowy, jeśli użytkownik wyłączył swój komputer, jego zasoby są niedostępne dla reszty komputerów znajdujących się w sieci. Problem ten może być rozwiązany przez nie wyłączanie komputerów, co z kolei rodzi wątpliwości dotyczące innych zagadnień, takich jak bezpieczeństwo.

Innym, bardziej subtelnym aspektem wydajności jest skalowalność. Sieć typu każdy-zkażdym jest z natury nieskalowalna. Im większa liczba komputerów przyłączona jest do sieci równorzędnej, tym bardziej staje się ona „krnąbrna" i nieposłuszna

Zastosowania

Sieci typu każdy-z-każdym mają dwa główne zastosowania. Pierwsze - są one idealne dla małych instytucji z ograniczonym budżetem technologii informacyjnych i ograniczonymi potrzebami współdzielenia informacji. Drugie - to zastosowanie tego rodzaju sieci do ściślejszego współdzielenia informacji w ramach grup roboczych wchodzących w skład większych organizacji.

1.2.3.2 Sieci oparte na serwerach (klient-serwer)

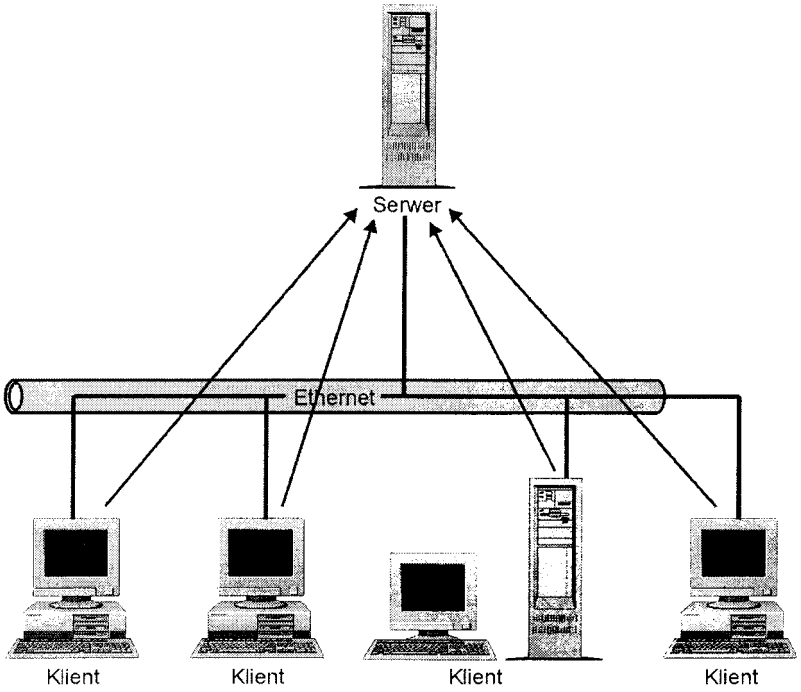
Sieci oparte na serwerach wprowadzają hierarchię, która ma na celu zwiększenie sterowalności różnych funkcji obsługiwanych przez sieć w miarę, jak zwiększa się jej skala. Często sieci oparte na serwerach nazywa się sieciami typu klient-serwer. Rysunek 2.4 ilustruje ową hierarchię klientów i serwerów.

W sieciach klient-serwer zasoby często udostępniane gromadzone są w komputerach odrębnej warstwy zwanych serwerami. Serwery zwykle nie mają użytkowników bezpośrednich. Są one raczej komputerami wielodostępnymi, które regulują udostępnianie swoich zasobów szerokiej rzeszy klientów. W sieciach tego typu z klientów zdjęty jest ciężar funkcjonowania jako serwery wobec innych klientów.

KORZYŚCI

Wiele jest korzyści płynących z opartego na serwerach podejścia do współdzielenia zasobów sieci. Korzyści te bezpośrednio odpowiadają ograniczeniom sieci każdy-z-każdym. Obszarami, w których zastosowanie sieci serwer-klient przynosi korzyści, są więc bezpieczeństwo, wydajność oraz administracja.

Rysunek 2.4. .Sieć typu klient-serwer.



• Sieci oparte na serwerach są dużo bezpieczniejsze niż sieci równorzędne. Przyczynia się do tego wiele czynników. Po pierwsze, bezpieczeństwem zarządza się centralnie. Zasoby przyłączone do sieci nie podlegają już zasadzie „najsłabszego ogniwa w łańcuchu", która stanowi integralną część sieci każdy-z-każdym.

Zamiast tego wszystkie konta użytkowników i ich hasła zarządzane są centralnie i tak są weryfikowane przed udostępnieniem zasobu użytkownikowi. Ułatwia to przy okazji życie użytkownikom, zmniejszając znacznie liczbę haseł, które muszą oni pamiętać (najczęściej do jednego).

Inną korzyścią wynikającą z owej centralizacji zasobów jest fakt, że zadania administracyjne, takie jak tworzenie kopii zapasowych, mogą być przeprowadzane stale i w sposób wiarygodny.

• Sieci oparte na serwerach charakteryzują się większą wydajnością wchodzących w jej skład komputerów ze względu na kilka czynników. Po pierwsze - z każdego klienta zdjęty jest ciężar przetwarzania żądań innych klientów. W sieciach opartych na serwerach każdy klient musi przetwarzać jedynie żądania pochodzące wyłącznie od jego głównego użytkownika.

Co więcej, przetwarzanie to jest wykonywane przez serwer, który jest skonfigurowany specjalnie do wykonywania tej usługi. Zwykle serwer cechuje się większą mocą przetwarzania, większą ilością pamięci i większym, szybszym dyskiem twardym niż komputer-klient. Dzięki temu żądania komputerów-klientów mogą być obsłużone lepiej i szybciej.

Taki sposób organizacji sieci oszczędza również syzyfowego trudu zapamiętywania miejsc w sieci, w których przechowywane są różne zasoby. W sieciach opartych na serwerach liczba „kryjówek", w których dane mogą się przed nami schować, jest ograniczona do liczby serwerów. W środowisku Windows NT zasoby serwerów mogą być łączone z każdym komputerem jako jego osobny dysk logiczny (taki sposób ich łączenia nazywa się mapowaniem). Po zmapowaniu dysku sieciowego można korzystać z jego (zdalnych z perspektywy mapującego) zasobów w taki sam sposób, w jaki korzysta się z zasobów znajdujących się na dysku lokalnym.

• Łatwo również zmieniać rozmiary sieci serwerowych, czyli je skalować. Niezależnie od liczby przyłączonych do sieci klientów, jej zasoby znajdują się bowiem zawsze w jednym, centralnie położonym miejscu. Zasoby te są również centralnie zarządzane i zabezpieczane. W związku z tym wydajność sieci jako całości nie zmniejsza się wraz ze zwiększaniem jej rozmiaru.

Ograniczenia

• Sieć serwerowa ma jedno tylko ograniczenie: zainstalowanie i obsługa tego rodzaju sieci kosztuje dużo więcej niż sieci typu każdy-z-każdym. Owa różnica w cenie ma kilka powodów.

Przede wszystkim, koszty sprzętu i oprogramowania są dużo wyższe ze względu na potrzebę zainstalowania dodatkowego komputera, którego jedynym zadaniem będzie obsługa klientów. A serwery najczęściej są dosyć skomplikowanymi - czyli drogimi urządzeniami.

Również koszty obsługi sieci opartych na serwerach są dużo wyższe. Wynika to z potrzeby zatrudnienia wyszkolonego pracownika specjalnie do administrowania i obsługi sieci. W sieciach każdy-z-każdym każdy użytkownik odpowiedzialny jest za obsługę własnego komputera, w związku z czym nie trzeba zatrudniać dodatkowej osoby specjalnie do realizacji tej funkcji.

Ostatnią przyczyną wyższych kosztów sieci serwerowej jest większy koszt ewentualnego czasu przestoju. W sieci każdy-z-każdym wyłączenie lub uszkodzenie jednego komputera powoduje niewielkie jedynie zmniejszenie dostępnych zasobów sieci lokalnej. Natomiast w sieci lokalnej opartej na serwerze, uszkodzenie serwera może mieć znaczny i bezpośredni wpływ na praktycznie każdego uczestnika sieci. Powoduje to zwiększenie potencjalnego ryzyka użytkowego sieci serwerowej. W celu jego zmniejszenia stosowane są więc różne podejścia, z grupowaniem serwerów w celu uzyskania nadmierności włącznie. Każde z tych rozwiązań - niestety - dalej zwiększa koszty sieci serwerowej.

Zastosowania

Sieci oparte na serwerach są bardzo przydatne, zwłaszcza w organizacjach dużych oraz wymagających zwiększonego bezpieczeństwa i bardziej konsekwentnego zarządzania zasobami przyłączonymi do sieci. Koszty dodane sieci opartych na serwerach mogą jednak przesunąć je poza zasięg możliwości finansowych małych organizacji.

1.2.3.3 Sieci mieszane

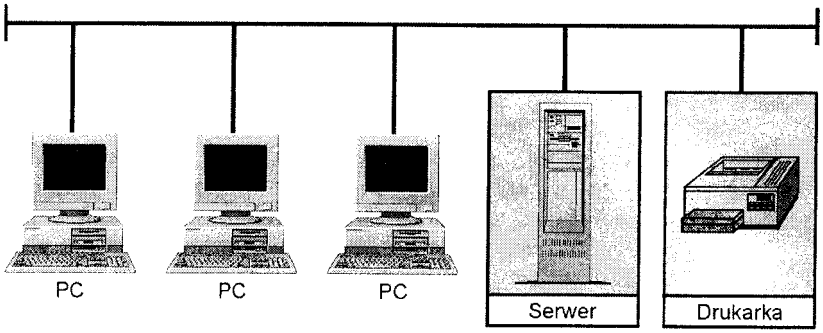
Różnice między sieciami każdy-z-każdym a sieciami opartymi na serwerach nie są tak oczywiste jak sugerują to poprzednie podpunkty. Przedstawione one w nich zostały specjalnie jako odmienne typy sieci, dla tzw. potrzeb akademickich, czyli dla lepszego ich opisania i zrozumienia. W rzeczywistości różnice między typami sieci uległy rozmyciu ze względu na wielość możliwości udostępnianych przez różne systemy operacyjne.
Obecnie standardowo zakładane są sieci będące mieszanką sieci równorzędnych (każdy-z -każdym) i serwerowych (opartych na serwerze). Przykładem tego rodzaju sieci jest sieć o architekturze serwerowej grupującej centralnie zasoby, które powinny być ogólnodostępne. W ramach takiej organizacji sieci, udostępnianie zasobów wewnątrz lokalnych grup roboczych może nadal odbywać się na zasadzie dostępu równorzędnego.

1.2.4 Topologie sieci lokalnych

Topologie sieci LAN mogą być opisane zarówno na płaszczyźnie fizycznej, jak i logicznej. Topologia fizyczna określa geometryczną organizację sieci lokalnych. Nie jest ona jednak mapą sieci. Jest natomiast teoretyczną strukturą graficznie przedstawiającą kształt i strukturę sieci LAN.
Topologia logiczna opisuje wszelkie możliwe połączenia między parami mogących się komunikować punktów końcowych sieci. Za jej pomocą opisywać można, które punkty końcowe mogą się komunikować z którymi, a także ilustrować, które z takich par mają wzajemne, bezpośrednie połączenie fizyczne. Rozdział niniejszy koncentruje się tylko na topologii fizycznej.
Do niedawna istniały trzy podstawowe topologie fizyczne: • magistrali,
• pierścienia, • gwiazdy.
Rodzaj topologii fizycznej wynika z rodzaju zastosowanych technologii sieci LAN. Na przykład, dla sieci Token Ring, z definicji, stosowane były topologie pierścienia. Jednak koncentratory, znane również jako jednostki dostępu do stacji wieloterminalowych (ang.. MSAU- Multi-Station Access Units) sieci Token Ring, rozmyły różnice między topologią pierścienia a topologią gwiazdy dla sieci Token Ring. W wyniku wprowadzenia tychże koncentratorów powstały sieci o topologii pierścieni gwiaździstych. Podobnie wprowadzenie przełączania sieci LAN zmieniło sposób klasyfikowania topologii. Lokalne sieci przełączane, niezależnie od rodzaju ramki i metody dostępu, są topologicznie podobne. Pierścień jednostki dostępu do stacji wieloterminalowej, który do niedawna używany był do przyłączania - na poziomie elektroniki - wszelkich urządzeń do sieci Token Ring, nie pełni już tej funkcji. Zamiast niego każde z przyłączanych urządzeń ma własny minipierścień, do którego przyłączone są tylko dwa urządzenia: ono samo oraz port
przełączania. W związku z tym słowo „przełączane" powinno być dodane do triady nazw podstawowych fizycznych topologii sieci LAN, na określenie kolejnego, czwartego już ich typu.
Przełączniki wprowadzają topologię gwiazdy, bez względu na rodzaj protokołu warstwy łącza danych dla którego są zaprojektowane. Ze względu na przyjęcie terminu „przełącznik" (dzięki nieustannym kampaniom reklamowym producentów przełączników), sieci częściej określa się mianem „przełączanych" (lub „komutowanych") niż „gwiaździstej magistrali" (ang.. star bus) czy „gwiaździstego pierścienia"(ang.. star ring. W związku z tym przełączanie może być uważane za rodzaj topologii. Mimo to na ewentualnym egzaminie MCSE lepiej jednak zaznaczyć odpowiedź „gwiaździsta magistrala (star bus)" i „gwiaździsty pierścień (star ring)".
Przełączanie rozdzieliło dotychczasowe nierozłączki: topologię i technologię sieci LAN. Obecnie bowiem dosłownie wszystkie technologie LAN mogą być zakupione w wersji przełączanej. Ma to niebagatelny wpływ na sposób uzyskiwania dostępu do sieci i, tym samym, na jej ogólną sprawność. Wpływ ten przedstawiony jest bardziej szczegółowo w dalszym punkcie tego rozdziału zatytułowanym „Topologia przełączana".
Mimo że kupić można przełączniki dla sieci lokalnych o dowolnej technologii, na przykład Ethernet, Token Ring, FDDI itd.., to przełączniki te nie pełnią funkcji mostków translacyjnych. Oznacza to, że nie są one zdolne do przełączania ramek pomiędzy sieciami lokalnymi o różnych architekturach.

1.2.4.1 Topologia magistrali

Topologię magistrali wyróżnia to, że wszystkie węzły sieci połączone są ze sobą za pomocą pojedynczego, otwartego (czyli umożliwiającego przyłączanie kolejnych urządzeń) kabla. Kabel taki obsługuje tylko jeden kanał i nosi nazwę magistrali. Niektóre technologie oparte na magistrali korzystają z więcej niż jednego kabla, dzięki czemu obsługiwać mogą więcej niż jeden kanał, mimo że każdy z kabli obsługuje niezmiennie tylko jeden kanał transmisyjny.
Oba końce magistrali muszą być zakończone opornikami ograniczającymi, zwanymi również często terminatorami. Oporniki te chronią przed odbiciami sygnału. Zawsze gdy komputer wysyła sygnał, rozchodzi się on w przewodzie automatycznie w obu kierunkach. Jeśli sygnał nie napotka na swojej drodze terminatora, to dochodzi do końca magistrali, gdzie zmienia kierunek biegu. W takiej sytuacji pojedyncza transmisja może całkowicie zapełnić wszystkie dostępne szerokości pasma i uniemożliwić wysyłanie sygnałów wszystkim pozostałym komputerom przyłączonym do sieci. Przykładowa topologia magistrali przedstawiona jest na rysunku 2.5.
Rysunek 2.5. Standardowa topologia magistrali.

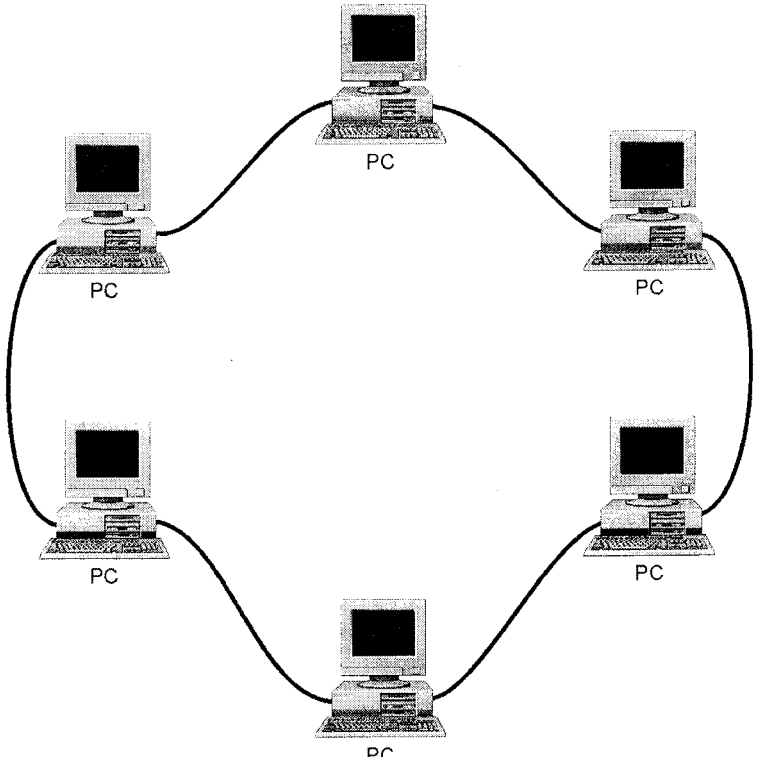


Typowa magistrala składa się z pojedynczego kabla łączącego wszystkie węzły w sposób charakterystyczny dla sieci równorzędnej. Kabel ten nie jest obsługiwany przez żadne urządzenia zewnętrzne. Zatem wszystkie urządzenia przyłączone do sieci słuchają transmisji i przesyłanych magistralą i odbierają pakiety do nich zaadresowane. Brak jakichkolwiek urządzeń zewnętrznych, w tym wzmacniaków, sprawia, że magistrale sieci lokalnych są proste i niedrogie. Jest to również przyczyną dotkliwych ograniczeń dotyczących odległości, funkcjonalności i skalowalności sieci.

Topologia ta jest więc praktyczna jedynie dla najmniejszych sieci LAN. Wobec tego obecnie dostępne sieci lokalne o topologii magistrali są tanimi sieciami równorzędnymi udostępniającymi podstawowe funkcje współdziałania sieciowego. Produkty te są przeznaczone do użytku w domach i małych biurach.
Jedynym od tego wyjątkiem jest specyfikacja IEEE 802.4 magistrali Token Bus sieci LAN. Technologia ta charakteryzuje się względnie dużą odpornością na błędy, dość wysokim stopniem determinizmu i ogólnym podobieństwem do sieci lokalnej Token Ring. Deterministyczne sieci LAN oferują administratorom wysoki stopień kontroli przez określanie maksymalnej ilości czasu, podczas którego ramka danych może znajdować się w transmisji. Podstawową różnicę stanowi oczywiście fakt, że magistrala Token Bus została wdrożona na podstawie topologii magistrali.
Magistrala Token Bus spotkała się z wyjątkowo chłodnym przyjęciem rynku. Jej zastosowanie było ograniczone do linii produkcyjnych zakładów pracy. Topologie magistralowe rozwinęły się jednak w tysiące innych form. Na przykład dwoma wczesnymi formami Ethernetu były lOBase2 oraz lOBase5, oparte na topologii magistrali oraz kablu koncentrycznym. Magistrale stały się również technologią ważną stosowaną do łączenia składników poziomu systemowego i urządzeń peryferyjnych w ramach wewnętrznych architektur komputerów.

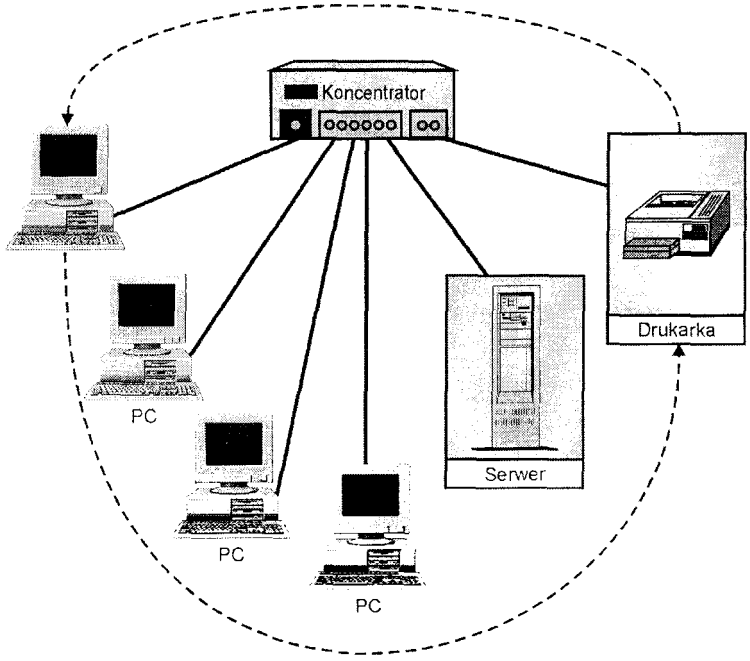
1.2.4.2 Topologia pierścienia

Pierwszą topologią pierścieniową była topologia prostej sieci równorzędnej. Każda przyłączona do sieci stacja robocza ma w ramach takiej topologii dwa połączenia: po jednym do każdego ze swoich najbliższych sąsiadów (patrz rysunek 2.6). Połączenie takie musiało tworzyć fizyczną pętlę, czyli pierścień. Dane przesyłane były wokół pierścienia
Rysunek 2.6. Topologia równorzędna kady-z-każdynz.



w jednym kierunku. Każda stacja robocza działała podobnie jak wzmacniak, pobierając i odpowiadając na pakiety do niej zaadresowane, a także przesyłając dalej pozostałe pakiety do następnej stacji roboczej wchodzącej w skład sieci.

Pierwotna, pierścieniowa topologia sieci LAN umożliwiała tworzenie połączeń równorzędnych między stacjami roboczymi. Połączenia te musiały być zamknięte; czyli musiały tworzyć pierścień. Korzyść płynąca z takich sieci LAN polegała na tym, że czas odpowiedzi był możliwy do ustalenia. Im więcej urządzeń przyłączonych było do pierścienia, tym dłuższy był ów czas. Ujemna strona tego rozwiązania polegała na tym, że uszkodzenie jednej stacji roboczej najczęściej unieruchamiało całą sieć pierścieniową.
Owe prymitywne pierścienie zostały wyparte przez sieci Token Ring firmy IBM, które z czasem znormalizowała specyfikacja IEEE 802.5. Sieci Token Ring odeszły od połączeń międzysieciowych każdy-z-każdym na rzecz koncentratorów wzmacniających. Wyeliminowało to podatność sieci pierścieniowych na zawieszanie się przez wyeliminowanie konstrukcji każdy-z-każdym pierścienia. Sieci Token Ring, mimo pierwotnego kształtu pierścienia (stąd nazwa ang.. *ring - pierścień),* tworzone są przy zastosowaniu topologii gwiazdy i metody dostępu cyklicznego, co przedstawia rysunek 2.7.
Rysunek 2.7. Topologia pierścieniowa o kształcie gwiazdy.

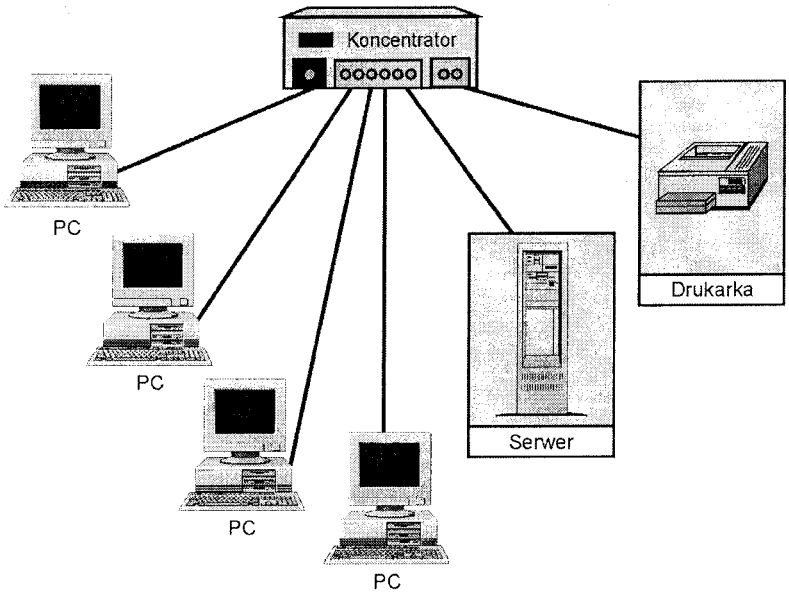


Sieci LAN mogą być wdrażane w topologii gwiazdy, przy zachowaniu - mimo to metody dostępu cyklicznego. Sieć Token Ring zilustrowana na rysunku 2.7 przedstawia pierścień wirtualny tworzony za pomocą metody dostępu cyklicznego zwanej roundrobin. Linie pełne reprezentują połączenia fizyczne, podczas gdy linie przerywane - logiczny przebieg sterowanego dostępu do nośnika.

Token w takiej sieci przesyłany jest do kolejnych punktów końcowych, mimo że wszystkie one są przyłączone do wspólnego koncentratora. I to pewnie dlatego wielu z nas nie potrafi oprzeć się pokusie określenia sieci Token Ring jako mających „logiczną" topologię pierścienia, nawet mimo tego, że fizycznie ujęte są one w kształcie gwiazdy. Pokusie tej ulegli na przykład twórcy kursu i egzaminu Microsoft Networking Essentials. Uważają oni sieć Token Ring za mającą topologię pierścienia, a nie - jak jest w istocie - topologię gwiazdy. Co prawda pierścień występuje, ale na poziomie elektroniki, wewnątrz koncentratora Token Ring, czyli jednostki dostępu do stacji wieloterminalowej.

1.2.4.3 Topologia gwiazdy

Połączenia sieci LAN o topologii gwiazdy z przyłączonymi do niej urządzeniami rozchodzą się z jednego, wspólnego punktu, którym jest koncentrator, co przedstawia rysunek 2.8. Odmiennie niż w topologiach pierścienia- tak fizycznej, jak i wirtualnej - każde urządzenie przyłączone do sieci w topologii gwiazdy może uzyskiwać bezpośredni i niezależny od innych urządzeń dostęp do nośnika. W tym celu urządzenia te muszą współdzielić dostępne szerokości pasma koncentratora. Przykładem sieci LAN o topologii gwiazdy jest I 1OBaseT Ethernet.
Rysunek 2.8. Topologia gwiazdy.



Połączenia w sieci LAN o małych rozmiarach i topologii gwiazdy rozchodzą się z jednego wspólnego punktu. Każde urządzenie przyłączone do takiej sieci może inicjować dostęp do nośnika niezależnie od innych przyłączonych urządzeń. Topologie gwiazdy stały się dominującym we współczesnych sieciach LAN rodzajem topologii. Są one elastyczne, skalowalne i stosunkowo tanie w porównaniu z bardziej skomplikowanymi sieciami LAN o ściśle regulowanych metodach dostępu. Gwiazdy przyczyniły się do dezaktualizacji magistral i pierścieni, formując tym samym podstawy pod ostateczną (obecnie przynajmniej) topologię sieci LAN -topologię przełączaną.

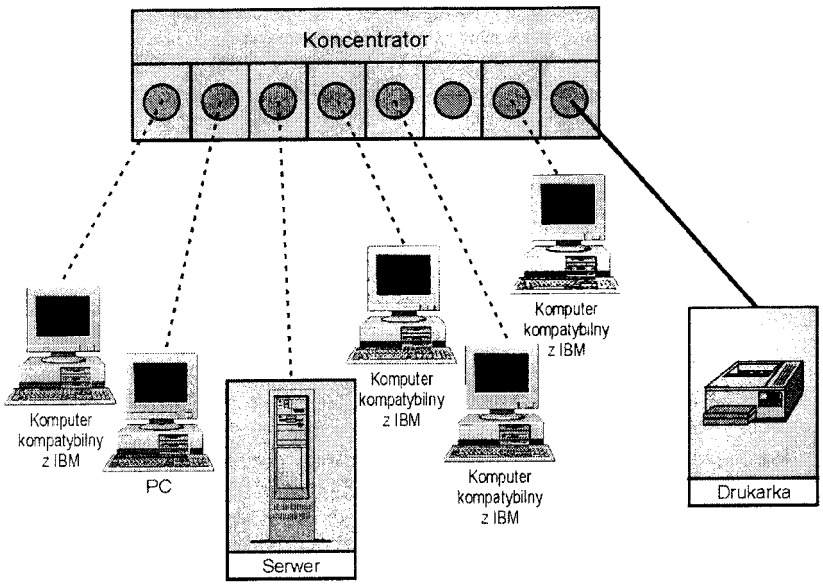
1.2.4.4 Topologia przełączana

Przełącznik jest urządzeniem warstwy łącza danych (warstwy 2 modelu referencyjnego OSI), zwanym również wieloportem. Przełącznik „uczy się" adresów sterowania dostępem do nośnika i składuje je w wewnętrznej tablicy przeglądowej (w tablicy wyszukiwania). Tymczasowo, między nadawcą ramki i jej zamierzonym odbiorcą, tworzone są ścieżki przełączane (czyli inaczej komutowane), a następnie ramki te są przesyłane dalej wzdłuż owych tymczasowych ścieżek.

Typowa sieć LAN o topologii przełączanej pokazana jest na rysunku 2.9, na którym widać wiele połączeń urządzeń z portami koncentratora przełączającego. Każdy port oraz urządzenie, które jest doń przyłączone, ma własną dedykowaną szerokość pasma. Choć pierwotnie przełączniki przesyłały dalej ramki na podstawie ich adresów fizycznych, to postęp technologiczny szybko zmienia ten stan rzeczy. Obecnie dostępne są przełączniki, które potrafią przetwarzać komórki, ramki, a nawet pakiety używające adresów warstwy 3, takie jak protokół IP.

Ramka jest zmiennej długości strukturą, która zawiera przesyłane dane, adresy nadawcy i adresata oraz inne pola danych niezbędne do przesyłania dalej ramek w warstwie 2 modelu referencyjnego. Komórki są podobne do ramek, z jednym wyjątkiem - są one strukturami nie o zmiennej, lecz o niezmiennej długości. Pakiety natomiast są układami protokołów działających na poziomie warstwy 3 modelu referencyjnego OSI. IP oraz IPX to dwa przykłady protokołów warstwy 3, które używają pakietów do opakowywania danych przesyłanych do domen zdalnych.

Rysunek 2.9. Topologia przełączna.:.



Przełączniki poprawiają sprawność sieci LAN na dwa sposoby. Pierwszy polega na zwiększaniu szerokości pasma dostępnego w sieci. Na przykład, przełączany koncentrator Ethernetu o 8 portach zawiera 8 odrębnych domen kolizji, z których każda przesyła dane z prędkością 10 Mbps, co daje łączną szerokość pasma rzędu 80 Mbps.

Drugi sposób zwiększania sprawności przełączanych sieci LAN polega na zmniejszaniu liczby urządzeń, wymuszających udostępnianie wszystkich segmentów pasma szerokości. Każda przełączana domena kolizji składa się jedynie z dwóch urządzeń: urządzenia sieciowego oraz portu koncentratora przełączanego, z którym urządzenie to jest połączone. Wyłącznie te dwa urządzenia mogą rywalizować o szerokość pasma 10 Mbps w seg mencie, w którym się znajdują. W sieciach, które nie korzystają z metody dostępu do nośnika na zasadzie rywalizacji o szerokość pasma-takich jak Token Ring lub FDDI tokeny krążą między dużo mniejszą liczbą urządzeń sieciowych niż ma to zwykle miejsce w sieciach o dostępie opartym na zasadzie rywalizacji.

Jedyny problem dużych sieci przełączanych (komutowanych) polega na tym, że przełączniki nie rozróżniają rozgłoszeniowych transmisji danych. Zwiększenie sprawności

sieci jest wynikiem segmentacji wyłącznie domeny kolizji, a nie domeny rozgłaszania. Nadmierne natężenie rozgłaszania może więc znacznie i niekorzystnie wpłynąć na wydajność sieci LAN.

1.2.4.5 Topologie złożone

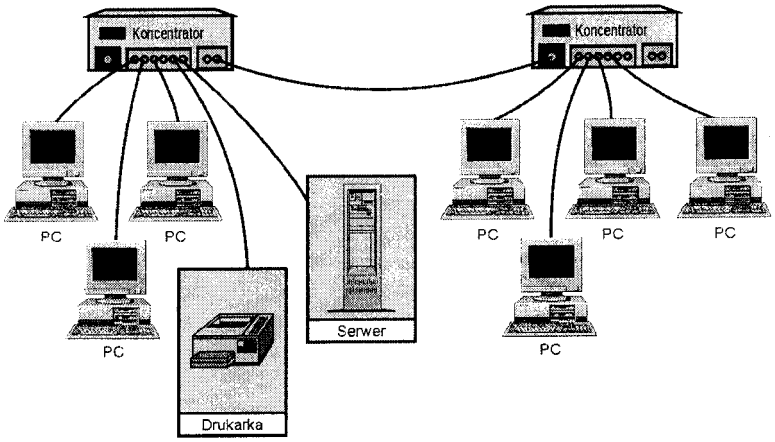
Topologie złożone są rozszerzeniami i/lub połączeniami podstawowych topologii fizycznych. Topologie podstawowe są odpowiednie jedynie do bardzo małych sieci LAN. Skalowalność topologii podstawowych jest bardzo ograniczona. Topologie złożone tworzone są z elementów składowych umożliwiających uzyskanie topologii skalowalnych odpowiadających zastosowaniom.

1.2.4.5.1 Łańcuchy

Najprostszą z topologii złożonych otrzymać można w wyniku połączenia szeregowego wszystkich koncentratorów sieci tak, jak przedstawia to rysunek 2.10. Taki sposób łączenia znany jest jako *łańcuchowanie.* Wykorzystuje ono porty już istniejących koncentratorów do łączenia ich z kolejnymi koncentratorami. Dzięki temu uniknąć można ponoszenia kosztów dodatkowych związanych z tworzeniem odpowiedniego szkieletu.

Rysunek 2.10. Koncentratory łańcuchujące.

Małe sieci LAN mogą być zwiększane (skalowane dodatnio) przez łączenie koncentratorów w łańcuchy (łańcuchowania ich). Łańcuchy dają się łatwo tworzyć i nie wymagają żadnych specjalnych umiejętności administracyjnych. Łańcuchy stanowiły alternatywną, wobec sieci LAN pierwszej generacji, metodę przyłączania urządzeń.



Ograniczenia nakładane przez łańcuchowanie określić można na wiele sposobów. Specyfikacje technologii LAN, takie jak 802.3 Ethernet, ograniczały maksymalny rozmiar sieci LAN ze względu na maksymalną liczbę koncentratorów i/lub wzmacniaków, które
można łączyć ze sobą szeregowo. Maksymalna długość kabla określona przez wymogi warstwy fizycznej pomnożonej przez maksymalną liczbę urządzeń dających się łączyć szeregowo określała maksymalny rozmiar sieci LAN. Rozmiar ten nazywany jest maksymalną średnicą sieci. Zwiększanie rozmiaru sieci ponad tę wartość wpływa negatywnie na działanie sieci LAN. Nowoczesne, wysokowydajne sieci lokalne, takie jak Fast Ethernet, nakładają ścisłe ograniczenia dotyczące średnicy sieci i liczby połączonych w jej ramach wzmacniaków.

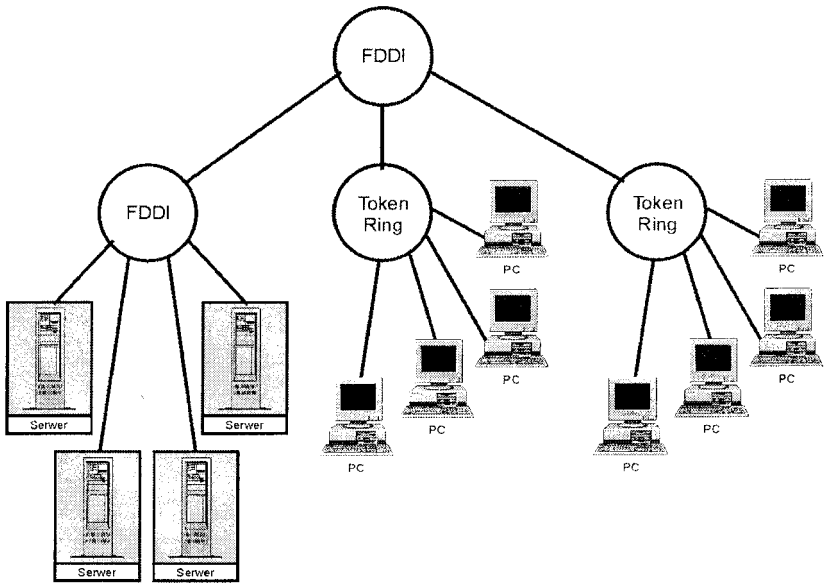
Wzmacniak jest urządzeniem, które odbiera przychodzące sygnały, wzmacnia je do poziomu ich pierwotnej siły i umieszcza z powrotem w sieci. Zwykle funkcje wzmocnienia i powtórzenia sygnału są dołączane również do koncentratorów. Terminy „wzmacniak" oraz „koncentrator" są więc synonimami i jako takie mogą być używane zamiennie.
Sieci łańcuchowane korzystające z dostępu do nośnika na zasadzie rywalizacji mogą powodować problemy na długo przed osiągnięciem maksymalnej średnicy sieci. Łańcuchowanie zwiększa liczbę połączeń i tym samym również liczbę urządzeń możliwych do przyłączenia do sieci LAN. Nie powoduje to zwiększenia całkowitej szerokości pasma ani domen kolizji. Łańcuchowanie zwiększa po prostu liczbę urządzeń współdzielących dostępne w sieci pasmo szerokości. Zbyt wiele urządzeń konkurujących o ten sam zakres pasma może powodować kolizje i szybko uniemożliwić poprawne działanie sieci LAN.
Topologia ta sprawdza się najlepiej w sieciach lokalnych, w skład których wchodzi garstka jedynie koncentratorów i - co najwyżej - niewielkie współdziałanie w ramach sieci rozległych.

1.2.4.5.2 Hierarchie

Topologie hierarchiczne składają się z kilku (więcej niż jednej) warstw koncentratorów. Każda z tych warstw realizuje inną funkcję sieci. Warstwa podstawowa jest w tego rodzaju topologii zarezerwowana dla komunikacji między stacją roboczą a serwerem. Poziomy wyższe umożliwiają grupowanie wielu poziomów użytkownika. Innymi słowy, wiele koncentratorów poziomu użytkownika połączonych jest za pomocą mniejszej liczby koncentratorów wyższego poziomu. Wszystkie koncentratory, niezależnie od poziomu, na którym się znajdują, najczęściej są urządzeniami identycznymi. Różni je tylko warstwa, na której się znajdują, a tym samym ich zastosowanie. Organizacja hierarchiczna jest najodpowiedniejsza dla sieci LAN o rozmiarach średnich do dużych, w których rozwiązuje ona problemy skalowalności i agregacji ruchu w sieci.
Hierarchiczne pierścienie
Rozmiary sieci pierścieniowych mogą być zwiększane przez łączenie wielu pierścieni w sposób hierarchiczny, tak jak przedstawia to rysunek 2.11. Łączność między stacją roboczą a serwerem może być realizowana za pomocą tylu pierścieni o ograniczonych rozmiarach, ile potrzeba do uzyskania odpowiedniego poziomu sprawności. Pierścień poziomu drugiego, zarówno w sieciach Token Ring, jak i FDDI, może być używany do wzajemnego łączenia wszystkich pierścieni poziomu użytkownika oraz do umożliwienia zagregowanego dostępu do sieci rozległych (sieci WAN).

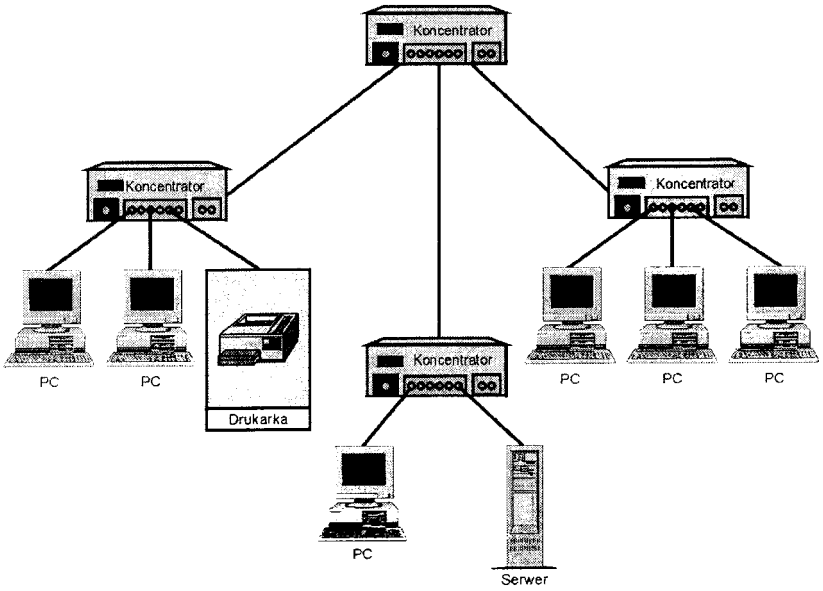
Rysunek 2.11. Topologia hierarchiczna pierścienia..
Sieci lokalne o małych pierścieniach można skalować, dodając hierarchicznie kolejne pierścienie. Rysunek 2.11 ilustruje dwa odrębne pierścienie Token Ring o szybkości przesyłania danych rzędu 16 Mbps każdy (przedstawione logicznie jako pętle), które używane są do łączenia komputerów użytkowników oraz odrębną pętlę FDDI używaną do łączenia serwerów i tworzenia szkieletu.
Hierarchiczne gwiazdy
Topologie gwiazdy również mogą być organizowane hierarchicznie w wiele gwiazd, co przedstawia rysunek 2.12. Hierarchiczne gwiazdy mogą być realizowane jako pojedyncze domeny kolizji lub dzielone przy użyciu przełączników, routerów i mostków na segmenty, z których każdy jest domeną kolizji.
Domena kolizji składa się ze wszystkich urządzeń konkurujących o prawo do transmisji przy użyciu współdzielonego nośnika. Przełączniki, mostki oraz routery dzielą domeny kolizji tworząc w ten sposób wiele mniejszych domen kolizji.
Topologia hierarchiczna gwiazdy używa jednego poziomu do łączenia użytkownika z serwerem, a drugiego - jako szkielet.
Hierarchiczne melanże

Ogólna sprawność sieci może być zwiększona przez nie wypełnianie wszystkich wymagań funkcjonalnych sieci lokalnej na siłę w ramach jednego rozwiązania. Dzisiejsze
nowoczesne koncentratory przełączające umożliwiają łączenie wielu różnych technologii

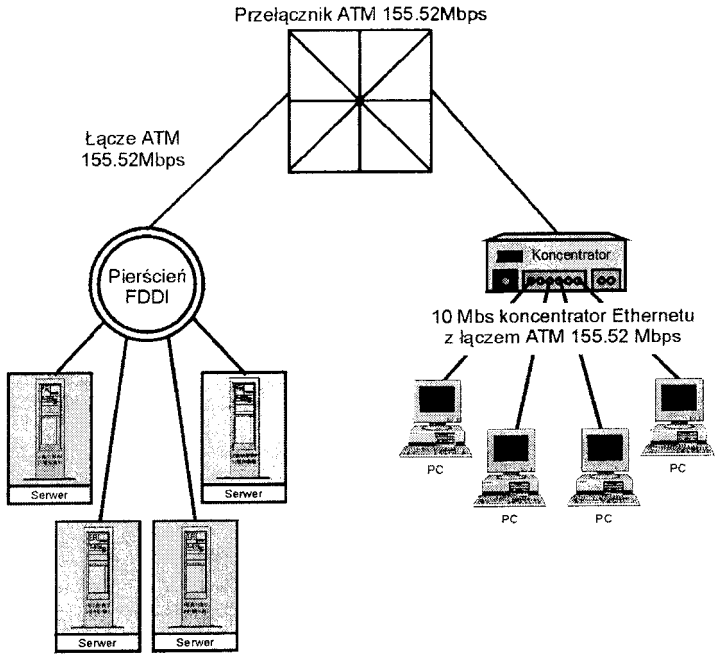


Rysunek 2.12. Topologia hierarchiczna gwiazdy.

Nowe topologie mogą być wprowadzane przez wstawianie odpowiednich płytek logicznych w obudowę koncentratora przełączającego o wielu gniazdach. Topologię hierarchiczną wykorzystywać można do tworzenia topologii mieszanych. co przedstawia n'sunek 2.13.



Rysunek 2.13. Topologia hier archiczna mieszana.



W przedstawionym przykładzie topologii hierarchicznej mieszanej do łączenia koncentratorów poziomu użytkownika używany jest szkielet ATM asynchronicznego trybu przesyłania (ang.. - *Asynchronous Transfer Mode),* do łączenia serwerów - sieć FDDI, a do łączenia stacji roboczych - sieć Ethernet. W podejściu tym sieć LAN dzielona jest na jej składniki funkcjonalne (przyłączania stacji roboczej, przyłączania serwera oraz szkieletu), co umożliwia zastosowanie odpowiedniej technologii dla każdej z realizowanych funkcji. Wspomniane obszary funkcjonalne opisane są szerzej w podrozdziale następnym.

1.2.5 Obwary funkcjonalne sieci LAN

Topologiczne wariacje mogą stanowić ważny sposób opt.ymalizowania wydajności każdego z obszarów funkcjonalnych sieci LAN. Sieci lokalne składają się z czterech odrębnych obszarów funkcjonalnych: przyłączenia stacji (komputera), przyłączenia serwera, przyłączenia sieci WAN oraz ze szkieletu. Każdy obszar funkcjonalny jest najlepiej obsługiwany przez odpowiednie topologie podstawowe i złożone.

1.2.5.1 Przyłączanie stacji

Podstawową funkcją większości sieci LAN jest przyłączanie stacji. Jak nazwa wskazuje, jest to ta część sieci LAN, która używana jest do przyłączania stacji roboczych użytkowników do sieci. Ten obszar funkcjonalny ma zwykle najmniejsze ze wszystkich obszarów wymagania odnośnie właściwości sieci LAN. Istnieją oczywiście wyjątki od tej tendencji, takie jak stacje robocze CAD/CAM, konferencje wideo itp. Ogólnie jednak rzecz biorąc, oszczędności dotyczące kosztów i wydajności poczynione w tym obszarze funkcjonalnym mają mniejsze prawdopodobieństwo negatywnego wpływu na sprawność sieci.

Przyłączanie urządzeń mających różne wymagania względem wydajności sieci może wymagać korzystania z wielu technologii LAN, co przedstawia rysunek 2.14. Na szczęście, obecnie produkowane koncentratory mimo nie dającej się modyfikować obudowy obsługują wiele różnych technologii.

Sieci LAN są podstawowym sposobem łączenia stacji roboczych i ich urządzeń peryferyjnych. Różne wymagania stacji roboczych odnośnie właściwości sieci mogą wymagać stosowania mieszanych rozwiązań topologiczno-technologicznych.

1.2.5.2 Przyłączanie serwera

Serwery są zwykle solidniejsze od stacji roboczych użytkowników. Na ogół są one miejscami dużego natężenia ruchu i muszą obsługiwać wielu klientów i/lub innych serwerów. W przypadku serwerów o dużej przepustowości agregowanie ruchu musi być ujęte w projekcie topologii sieci LAN; inaczej wydajność sieci ulegnie obniżeniu. Również przyłączalność serwerów powinna być bardziej niezawodna niż przyłączalność stacji roboczych, zwłaszcza w zakresie dostępnych szerokości pasma oraz metody dostępu.

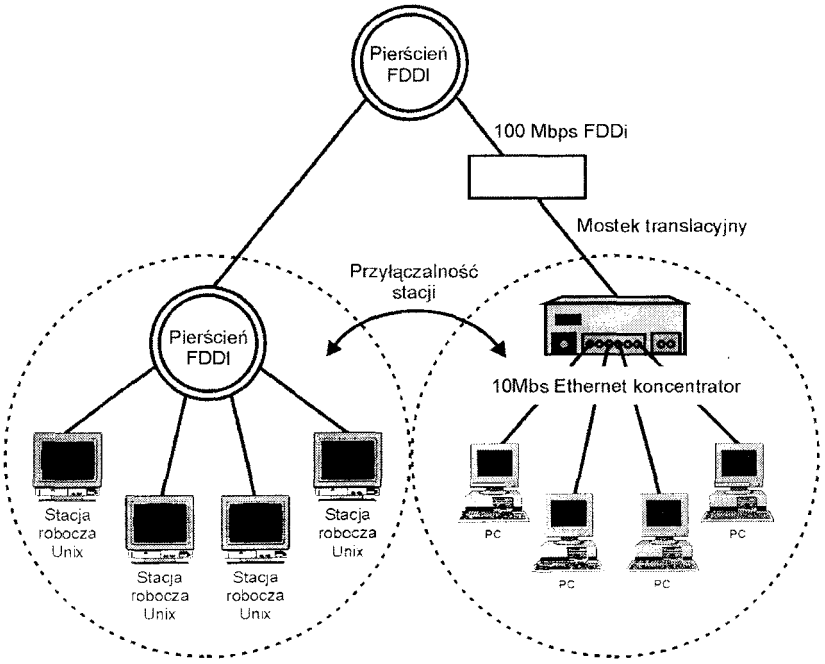
Rysunek 2.14. Przyłączalność stacji w sieci LAN.

Topologie sieci LAN mogą również być zmieniane w celu dostosowania ich do zwiększonych wymagań serwerów i ich grup. Na rysunku 2.15 przedstawiona jest taka przykładowa łączona topologia hierarchiczna. Grupa serwerów połączona jest za pomocą małej pętli FDDI; mniej niezawodne stacje robocze są tu połączone za pomocą Ethernetu.

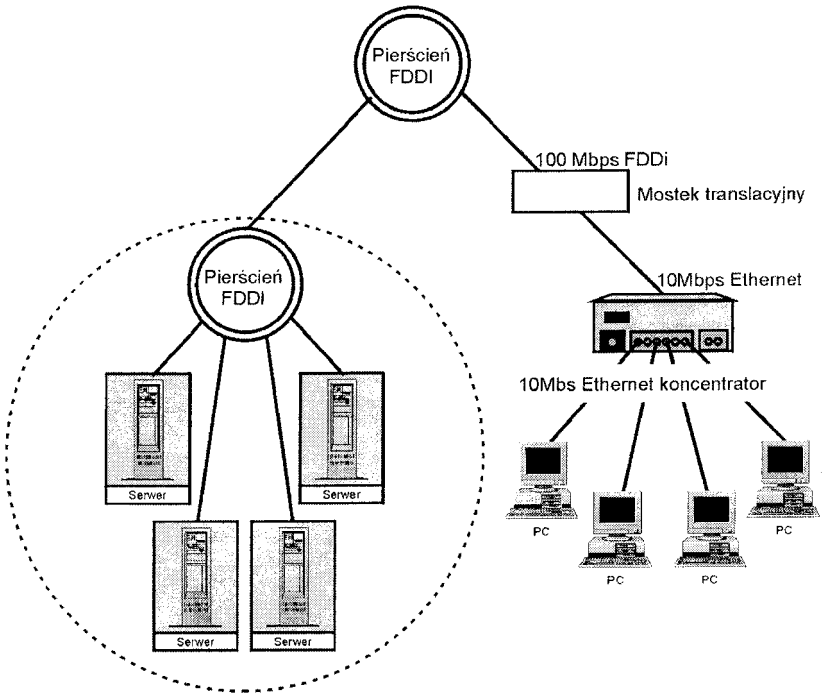
*1.2.6 **Przyłączanie do sieci** WAN*

Często pomijanym aspektem topologii sieci lokalnych (sieci LAN) jest możliwość łą czenia ich z sieciami rozległymi (z sieciami WAN). W wielu przypadkach sieć WAN jest przyłączana za pomocą pojedynczego łącza szkieletu z routerem, tak jak przedstawia to rysunek 2.16.

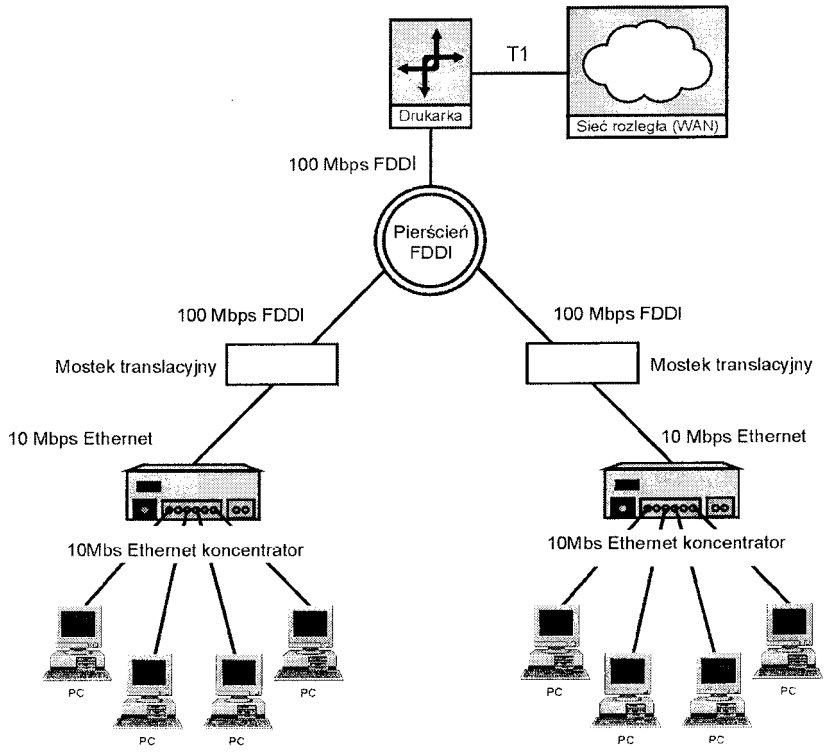
Połączenie sieci LAN z routerem umożliwiające dostęp do sieci WAN stanowi podstawowe ogniwo całej konstrukcji sieci LAN. Wybór nieodpowiedniej technologii w tak krytycznym miejscu może spowodować znaczne obniżenie poziomu wydajności w obsłudze ruchu przychodzącego do sieci LAN i wychodzącego z niej. Technologie LAN używające dostępu do nośnika na zasadzie rywalizacji są zdecydowanie nieodpowiednie do wykonywania tej funkcji.
Sieci obsługujące znaczne natężenia ruchu z sieci WAN do LAN i z LAN do WAN wymagają najbardziej niezawodnych z dostępnych połączeń. Wybrana technologia powinna być niezawodna pod względem jej nominalnej szybkości transmisji oraz stosowanej



Rysunek 2.15. Przyłączalność serwera sieci LAN



Rysunek 2.16. *?y łączanie sieci* WAN \.



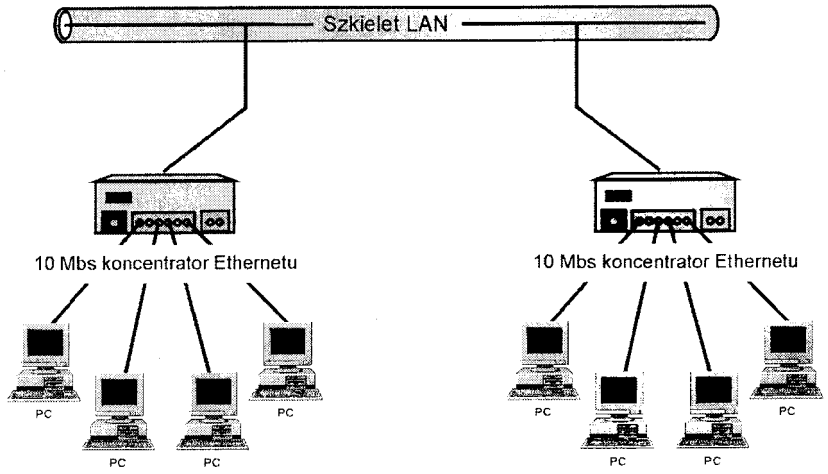
metody dostępu. Za wszelką cenę powinno się tu unikać technologii o dostępie opartym na zasadzie rywalizacji. Stosowanie rywalizacji, nawet w połączeniu z dedykowanym portem przełączanym może powodować problemy w sieciach o dużym natężeniu ruchu. Rozwiązanie takie będzie powodowało zatykanie się sieci LAN.

1.2.7 Przyłączanie do szkieletu

Szkielet sieci LAN tworzą urządzenia używane do łączenia koncentratorów. Szkielet może być tworzony w różnych topologiach za pomocą różnorodnych składników sieciowych, co przedstawia rysunek 2.17.

Rysunek 2.17. Szkielet LAN.

Szkielet sieci LAN realizuje funkcję krytyczną. Łączy on wszystkie zasoby sieci lokalnej oraz ewentualnie sieć lokalną z siecią rozległą. Logiczny opis szkieletu przedstawiony jest na rysunku 2.17. Może on być zrealizowany na jeden z wielu sposobów.

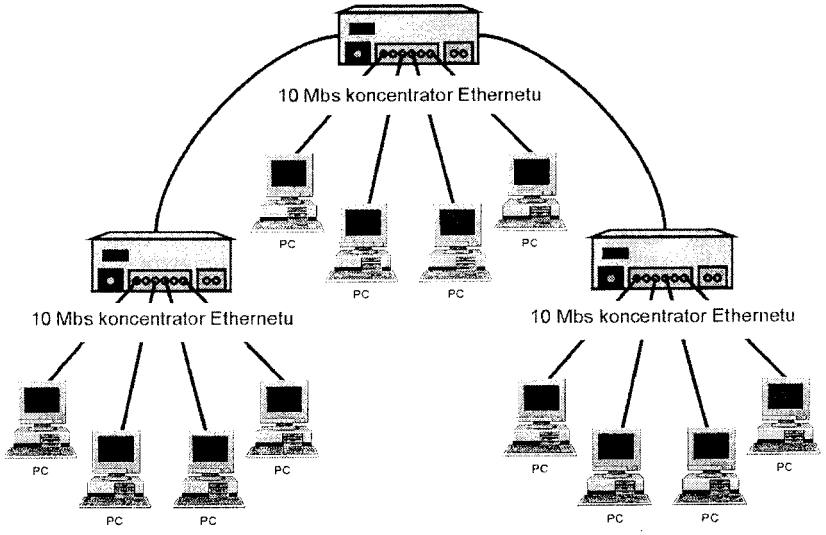


Wybór najodpowiedniejszej topologii szkieletu dla zadanej sieci LAN nie jest łatwy. Niektóre opcje są łatwiejsze do opracowania, dostępniejsze cenowo i łatwiejsze w obsłudze. Inne mogą być z. kolei kosztowne: zarówno do nabycia, jak i w obsłudze. Kolejna ważna różnica polega na skalowalności różnych topologii szkieletu. Rozmiary niektórych można łatwo zwiększać aż do pewnej granicy, po przekroczeniu której wymagają one ponownej inwestycji w celu utrzymania akcept.owalnych poziomów wydajności. Każda opcja musi być sprawdzona indywidualnie stosownie do istniejącej sytuacji oraz wymagań technicznych.

Szkielet szeregowy

Szkielet szeregowy, przedstawiony na rysunku 2:18, jest szeregiem koncentratorów połączonych ze sobą łańcuchowo. Jak zostało to opisane w poprzednich punktach, topologia ta jest odpowiednia wyłącznie do zastosowań w małych sieciach.

Rysunek 2.18. Szkielet szeregowy.



Szkielet rozproszony

Szkielet rozproszony jest rodzajem topologii hierarchicznej, która może być utworzona przez zamontowanie koncentratora szkieletowego w centralnym miejscu sieci. Zwykle centrum topologii okablowania znajduje się w pokoju rozdzielni telefonicznej. Miejsce to jest idealne w celu umieszczenia tam również koncentratora szkieletu rozproszonego. Połączenia wychodzące z tego koncentratora biegną do innych koncentratorów znajdujących się w budynku (w którym sieć jest montowana), co przedstawia rysunek 2.19.

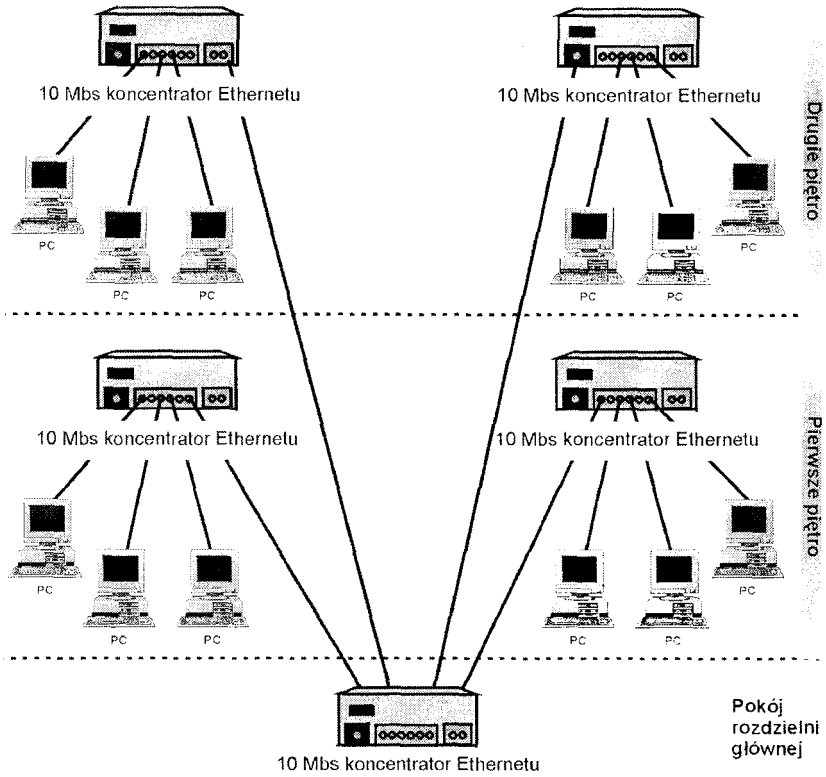
Szkielet rozproszony może być utworzony przez centralne umieszczenie koncentratora szkieletowego. Połączenia rozchodzą się z tego koncentratora do innych koncentratorów znajdujących się w budynku. Szkielet o topologii rozproszonej, w odróżnieniu od szkieletu szeregowego, umożliwia pokrycie siecią LAN dużych budynków bez obaw o przekroczenie maksymalnych średnic sieci.

Jeśli rozważamy utworzenie szkieletu rozproszonego, upewnijmy się najpierw, że znamy topologię okablowania budynku i ograniczenia długości poszczególnych nośników sieci I.A1V. Dla sieci o rozmiarach średnich do dużych i szkielecie rozproszonym jedynym nośnikiem wartym wzięcia pod uwagę jest kabel światłowodowy.

1.2.7.1 Szkielet segmentowy

Topologia szkieletu segmentowego oparta jest na centralnie umieszczonym routerze łączącym wszystkie segmenty sieci LAN w danym budynku. Router skutecznie tworzy wiele domen kolizji i rozgłaszania zwiększając w ten sposób wydajność każdego z segmentów sieci LAN. Routery działają na poziomie warstwy 3 modelu referencyjnego OSI. Funkcjonują one wolniej od koncentratorów. W związku z tym mogą ograniczać efektywną wydajność każdej transmisji rozpoczynającej się w jednym i kończącej się w drugim segmencie sieci LAN.

Rysunek 2.19. Szkielet rozproszony.



Szkielety segmentowe, jak widać na rysunku 2.20 również wprowadzają pojedynczy punkt defektu sieci LAN. Nie jest to poważna wada. Wiele innych topologii również wprowadza pojedynczy punkt defektu do sieci LAN. Jest to natomiast niedomaganie, które trzeba brać pod uwagę podczas planowania topologii sieci.

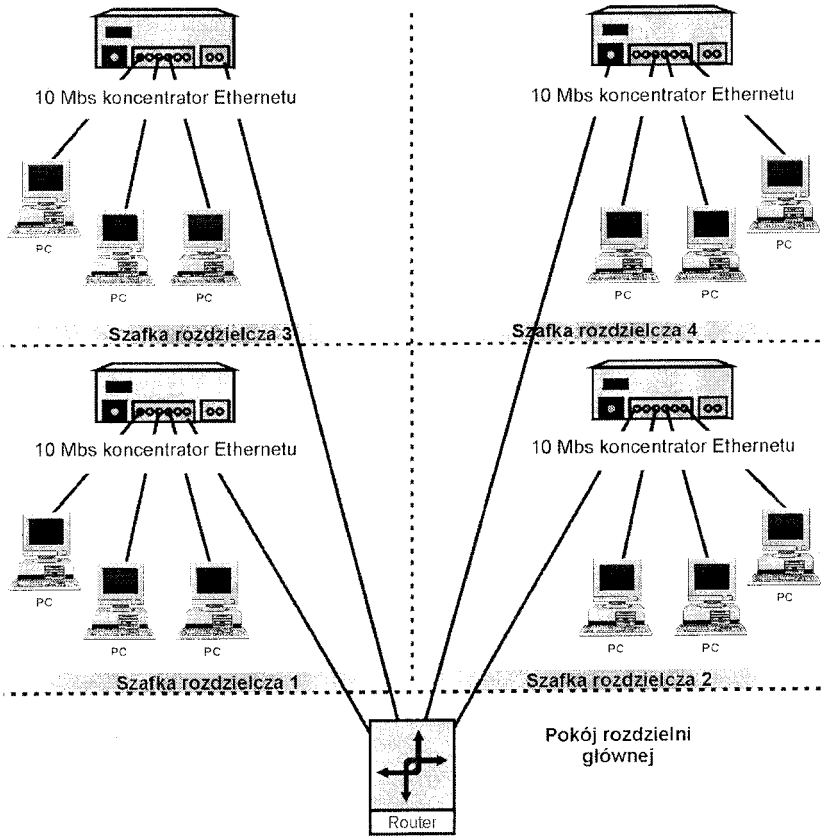
Ważnym aspektem topologii szkieletów segmentowych jest fakt, że użytkownicy rzadko kiedy zgrupowani są w określonych miejscach budynku. Raczej porozrzucani są w różnych, zwykle odległych jego miejscach. Oznacza to, że często znajdują się również po przeciwnych stronach routera szkieletu segmentowego sieci. W takiej sytuacji nawet proste zadania sieciowe wykonywane między uczestnikami jednej grupy roboczej będą często przechodzić przez jej router. W związku z tym podczas projektowania segmentowych szkieletów sieci LAN należy zwrócić szczególną uwagę na minimalizowanie natężenia ruchu przechodzącego przez takie routery i używać je jako agregatory ruchu dla zasobów poziomu LAN, takich jak urządzenia sieci WAN, lecz nie jako mostek.

1.2.7.2 Szkielet równoległy

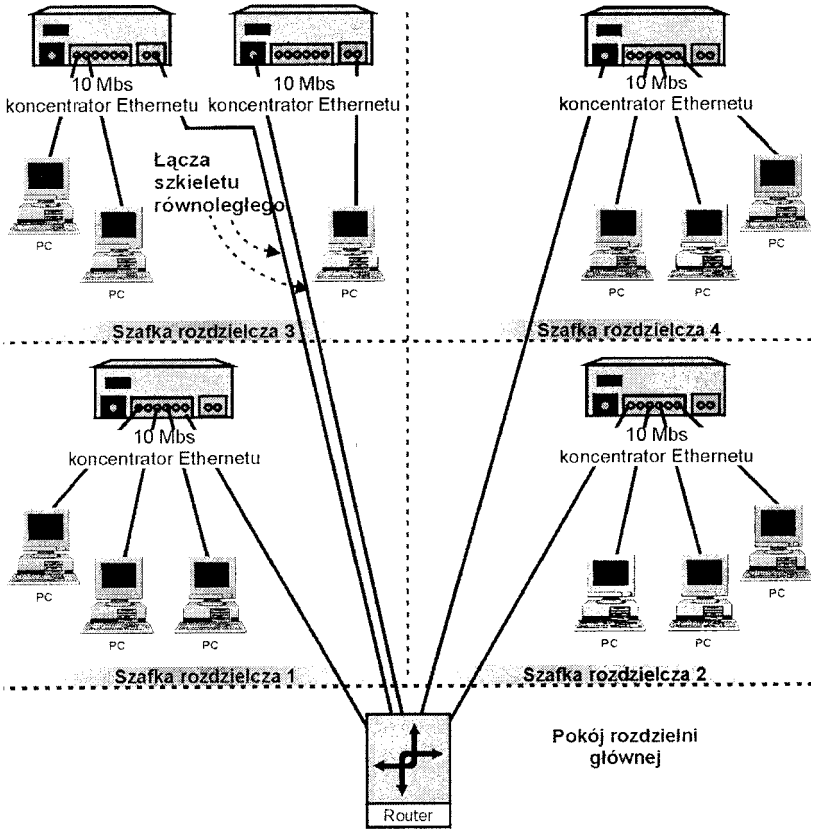
W niektórych sytuacjach, w których szkielety segmentowe stanowią rozwiązanie niepożądane, ich zmodyfikowana wersja może okazać się idealna. Wersja ta nazywana jest szkieletem równoległym. Powodów, dla których warto utworzyć szkielet równoległy, jest wiele. Grupy użytkowników mogą być rozrzucone po całym budynku. Niektóre z nich oraz niektóre aplikacje mogą mieć wysokie wymagania odnośnie bezpieczeństwa sieciowego. Jednocześnie również potrzebny może być wysoki stopień dostępu do sieci. Niezależnie od przyczyny, poprowadzenie połączeń równoległych od routera szkieletu segmentowego do szafki rozdzielczej umożliwia obsługę z tej szafki wielu segmentów, co przedstawione jest na rysunku 2.21.

Topologia szkieletu równoległego jest modyfikacją szkieletu segmentowego. Zarówno sam pokój rozdzielni jak i pojedyncza szafka rozdzielcza mogą obsługiwać wiele seg mentów. Powoduje to nieznaczne podniesienie kosztu utworzenia sieci, ale umożliwia zwiększenie wydajności każdego segmentu i zaspokaja zwiększone wymagania wobec sieci, takie jak wymogi bezpieczeństwa.

Dokładne zrozumienie potrzeb i wymagań klienta odnośnie każdego z obszarów funkcjonalnych sieci LAN jest kluczowe do utworzenia idealnej topologii. Potencjalne sposoby łączenia różnych składników sieci ogranicza jedynie wyobraźnia. Ciągłe innowacje technologiczne będą stale zwiększać różnorodność topologiczną rozwiązań dostępnych dla projektantów sieci.



Rysunek 2.20. Szkielet segmentowy



Rysunek 2.21. Topologia szkieletu równoległego.

Wiele spośród technologii złożonych przedstawionych jest w niniejszym rozdziale wyłącznie w celu pogłębienia wiedzy praktycznej.. Wielu producentów używa i tak innych nazw na określenie tych technologii, więc zamiast koncentrować się na nazwach, lepiej skupić się na ich właściwościach (mocnych i słabych stronach) oraz na zasadach ich działania.

1.2.8 Podsumowanie

Ważnym aspektem sieci LAN jest sposób obsługi dostępu do nośnika. I choć jest on bardziej funkcją sieciowych systemów operacyjnych aniżeli sprzętu sieciowego, to wpływa bezpośrednio na przebieg ruchu i sprawność sieci LAN.
Również topologia LAN jest zależna bezpośrednio od skuteczności sieci LAN. Istnieją cztery podstawowe topologie, które można mieszać, dopasowywać, układać i łączyć na bardzo wiele różnych sposobów. Zrozumienie korzyści i ograniczeń topologii podstawowych, a także potrzeb i wymagań użytkowników, jest istotne przy wyborze tej, która najlepiej je zaspokaja.
Tematy zaprezentowane w niniejszym rozdziale powinny udowodnić, że sieć składa się z więcej niż tylko sprzętu i okablowania. Równie ważnymi jak składniki fizyczne elementami składowymi sieci są sposób zorganizowania owych składników fizycznych oraz sposób dostępu do urządzeń przyłączonych do sieci.

1.3 Rozdział 3 Warstwa fizyczna

Marle A. Sportack
W rozdziale poprzednim pt.. „Typy i topologie sieci LAN" przedstawiono różne typy (określające metodę dostępu do zasobów) oraz topologie (określające sposób organizacji koncentratorów i okablowania), które mogą być stosowane w sieciach LAN. Zagadnienia te przedstawione były w sposób skrótowy, bez zagłębiania się w istotne szczegóły warstwy fizycznej.
Warstwa fizyczna, w postaci określonej przez Model Referencyjny OSI, składa się ze wszystkich procesów, mechanizmów, elektroniki oraz protokołów, które potrzebne są urządzeniu obliczającemu w celu wysyłania i odbierania binarnych strumieni danych. W specyfikacji warstwy fizycznej technologii LAN zamieszczone są oczekiwania odnośnie wydajności nośnika łączącego komunikujące się ze sobą urządzenia. Model jednak nie określa rodzaju samego nośnika.
W niniejszym rozdziale określimy rolę i znaczenie warstwy fizycznej sieci LAN, a następnie poszerzymy zakres rozważań i analizy warstwy fizycznej, wykraczając poza ramy określone przez model. Prócz opisu funkcjonalności warstwy fizycznej, dokładnie zbadamy okablowanie światłowodowe, koncentryczne i wiele rodzajów skrętek dwużyłowych. Opis każdego rodzaju nośnika obejmuje jego właściwości i ograniczenia.

1.3.1 Warstwa 1: warstwa fizyczna

Fundamentem, na którym zbudowany jest model referencyjny OSI, jest jego warstwa fizyczna. Określa ona wszystkie składniki sieci niezbędne do obsługi elektrycznego i/lub opt.ycznego wysyłania i odbierania sygnałów. Warstwa fizyczna składa się z czterech obszarów funkcjonalnych:
• mechanicznego, • elektrycznego, • funkcjonalnego, • proceduralnego.
Wspólnie obejmują one wszystkie mechanizmy potrzebne do obsługi elektrycznej i/lub opt.ycznej transmisji danych, takie jak techniki sygnalizacyjne, napięcie prądu elektrycznego przenoszącego sygnał, rodzaje nośników i odpowiadające im właściwości impedancji, elektroniczne składniki kart sieciowych, a nawet fizyczny kształt złącza używanego do terminacji nośnika.
Bardzo częste wśród użytkowników jest błędne przekonanie, wedle którego warstwa fizyczna modelu referencyjnego OSI dotyczy wszystkich składników sieci, które wytwarzają i/lub przenoszą sygnały. Jest to nieprawda. Pamiętać należy, że warstwa fizyczna jest najniższą warstwą stosu protokołów. Stos ułożyć można jedynie na odpowiednim urządzeniu. W związku z tym warstwa podstawowa jest więc ograniczona do opisu i/lub specyfikacji procesów i mechanizmów niezbędnych temu urządzeniu do komunikowania się z innymi kompatybilnymi urządzeniami. Warstwa fizyczna nie obejmuje wszystkiego, co nie składa się na komputer-hosta. Specyficznymi przykładami mechanizmów, które potrzebne są do obsługi przesyłania danych, lecz które znajdują się poza zakresem warstwy fizycznej, są:
• nośniki fizyczne, • koncentratory, • routery,
• przełączniki.
Wszystkie te składniki niezbędne są do efektywnego transportowania sygnałów pomiędzy komunikującymi się urządzeniami, lecz znajdują się poza zakresem warstwy fizycznej. Niższą granicą tej warstwy jest fizyczny port przyłączający urządzenie do nośnika. Zadaniem modelu nie jest definiowanie specyfikacji dla całej sieci i jej elementów składowych, lecz określenie jedynie niektórych właściwości sieci związanych z jej wydajnością. Model nie obejmuje więc niczego, co znajduje się pomiędzy fizycznymi złączami dwóch komunikujących się urządzeń, w tym i nośników, które w związku z tym nazywane są warstwą zerową. Niejasności dotyczące zakresu warstwy fizycznej modelu wynikają stąd, że specyfikuje ona wymagania odnośnie wydajności nośników nie określając samych nośników.

1.3.1.1 Funkcje warstwy fizycznej

Warstwa fizyczna przesyła i odbiera sygnały zaadresowane dla wszystkich protokołów jej stosu oraz aplikacji, które je wykorzystują. Musi ona więc wykonywać kilka istotnych funkcji-w szczególności:
Aby móc nadawać wiadomości, musi ona:
• zamieniać dane znajdujące się w ramkach na strumienie binarne
• wykonywać taką metodę dostępu do nośnika, jakiej żąda warstwa łącza danych • przesyłać ramki danych szeregowo (czyli bit po bicie) w postaci strumieni binarnych
W celu odbierania wiadomości konieczne jest natomiast:
• oczekiwanie na transmisje przychodzące do urządzenia hosta i do niego zaadresowane,
1 odbiór odpowiednio zaadresowanych strumieni
• przesyłanie binarnych strumieni do warstwy danych w celu złożenia ich z powrotem w ramki.
Lista ta, jak widać, nie obejmuje żadnych sposobów weryfikowania integralności danych. Warstwa, której dotyczy nie posiada bowiem mechanizmu służącego rozpoznawaniu znaczenia wysyłanych ani otrzymywanych danych. Służy wyłącznie przesyłaniu jedynek i zer. Nie

potrafi w związku z tym samodzielnie określać poprawności jakichkolwiek strumieni bitów. Ciężar wykonywania tych czynności przekazany jest protokołom warstw wyższych.

Sygnały kodowania

Zadaniem warstwy fizycznej jest kodowanie danych w formie, w której mogą być one następnie przesłane przez medium transmisyjne (nośnik). Formy te muszą być różne dla różnych nośników ponieważ każdy nośnik ma tylko sobie właściwe charakterystyki. Istnieje wiele różnych technik fizycznego kodowania danych, ale wszystkie kodowane i przenoszone są za pomocą fal elektromagnetycznych.

Fala elektromagnetyczna jest fizyczną formą energii, którą opisuje spektrum elektromagnetyczne. Spektrum to rozciąga się od poziomu zera oscylacji poprzez zakresy częstotliwości słyszalnych dla ludzkiego ucha, widzialnych dla ludzkiego oka, aż po fale szkodliwe dla obu tych i wszystkich pozostałych zmysłów, takie jak promienie X i promienie gamma.

Spektrum zdefiniowane i szczegółowo opisane jest w rozdziale 4 pt.. „Niezupełnie-fizyczna warstwa fizyczna".

Oscylacje fal elektromagnetycznych dają się przedstawić w postaci symetrycznego wzoru tworzonego przez kolejne cykle zmian wartości sygnału między pozytywnymi a negatywnymi. Miarą szybkości, z jaką te zmiany zachodzą jest częstotliwość. Częstotliwość mierzona jest w hercach (Hz). Jeden cyk(, czyli herc, odpowiada jednej 360-stopniowej zmianie w ciągu jednej sekundy.

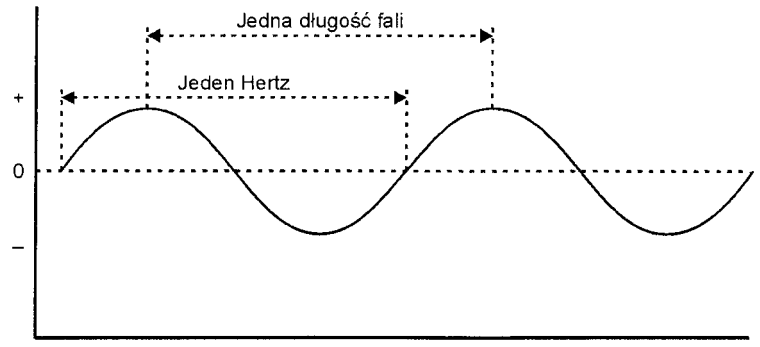
Skrótem, za pomocą którego przedstawiane są tysiące herców na sekundę, jest kHz, miliony - MHz, a miliardy - GHz.

Innym systemem mierzenia intensywności energii elektromagnetycznej jest długość fali. Miara ta określa odległość pomiędzy szczytami kolejnych fal. Prócz sposobu pomiaru obu jednostek fizycznych, różnica między nimi dotyczy również zakresu ich zastosowań. Hz używany jest do określania częstotliwości fal niższych zakresów, podczas gdy długość fali zarezerwowana jest dla fal wyższych zakresów, takich jak fale świetlne i promienie X oraz gamma. Fale mogą mieć różne długości: od niewielu milionowych metra do wielu metrów. Wymienione różnice są jedynymi charakterystykami różniącymi omawiane miary - poza tym są one do siebie podobne - stanowią wszak różne sposoby mierzenia tych samych własności.

Rysunek 3.1 przedstawia poglądowo drganie o częstotliwości jednego herca z zaznaczoną długością fali.

Rysunek 3.1. Jeden herc a jedna długość fali.

Niezależnie od rodzaju użytej miary częstotliwości, właściwości fizyczne fal oraz sposób ich propagacji (czyli rozprzestrzeniania) zmieniają się w miarę przechodzenia do kolejnych obszarów spektrum. Ogólnie rzecz biorąc, ze wzrostem częstotliwości, wzrasta również możliwość kodowania danych. Dzieje się tak dlatego, że liczba zmian stanu w ciągu sekundy dla wyższych częstotliwości jest wyższa niż dla częstotliwości niższych. A to właśnie zmiany stanu używane są do kodowania danych.



Kolejną zmienną właściwością fizyczną sygnału jest jego odporność na zakłócenia. Częstotliwości niższe są zwykle odporniejsze na zakłócenia i potrafią przenikać prawie każdy rodzaj materii nie zanikając przy tym do końca. W miarę zwiększania częstotliwości zmniejsza się gęstość materii, która wymagana jest do całkowitego wytłumienia sygnału. Na poziomie właściwym falom świetlnym sygnały mogą być bezpośrednio i całkowicie wytłumione przez najrzadsze nawet materiały. Oczywiście, po przejściu do zakresu fal o częstotliwości wyższej niż świetlna - np. do promieni X - zależność ta przestaje obowiązywać. Jednak dla celów kodowania i przesyłania danych w sieciach LAN wspomniane uogólnienie uznamy za obowiązujące.

Ostatnią charakterystyką, której zmiany będą nas interesować, jest zależność odwrotna między długością fali a jej częstotliwością, co oznacza, ze im większa jest długość fali, tym niższa jej częstotliwość. Fala o długości na przykład I metra oscyluje między dodatnim a ujemnym ekstremum powoli. Natomiast fala o długości 850 nm (czyli nanometrów, lub inaczej milionowych części metra) oscyluje szybko, czyli jej częstotliwość, a tym samym zdolności do przenoszenia danych, są dość wysokie.

Niezależnie od natężenia fali, nośnik realizuje dwie funkcje: • umożliwia transmisję fali,

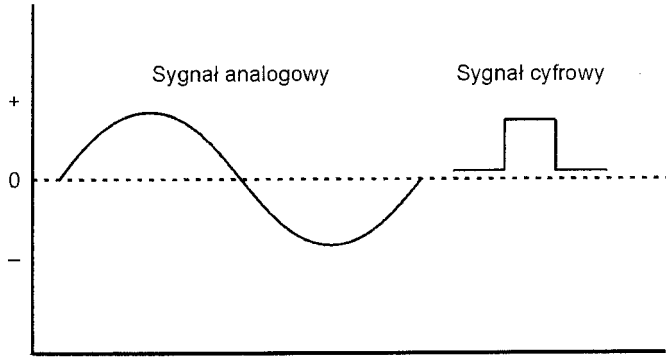
• w miarę możliwości chroni falę przed zakłóceniami.

Zawsze, gdy dane - zarówno analogowe, jak i cyfrowe - przenoszone są za pomocą wibracji elektrycznych, serię wibracji nazwać możemy sygnałem. Sygnały mogą pojawiać się we wszystkich częstotliwościach, choć ogólnie przyjęto, że fale o częstotliwości poniżej 300 kHz nie oferują zakresu pasma, które wystarczałoby większości urządzeń elektrycznych.

Krzywizna fali przedstawiona na rysunku 3.2 jest analogowa w swym kształcie. Szczyty i doliny sygnałów cyfrowych nie są bowiem spłaszczone, z ostrymi przejściami między fazami, lecz zmieniają się stopniowo w sposób naturalny, „organiczny". Porównanie sygnału cyfrowego a sygnału analogowego przedstawiono na rysunku 3.2.

Rysunek 3.2. Sygnał analogowy a sygnał cyfrowy

Warstwa fizyczna obsługiwana przez kartę sieciową wymaga nośnika określonego typu. Mając dany typ nośnika, warstwa ta potrafi określić częstotliwość, której ma używać, sposób (kodowania danych oraz szerokość pasma, które może być obsługiwane przy zastosowaniu kabla o określonej długości maksymalnej.



Szerokość pasma
Jest to szerokość kanału komunikacji mierzona w hercach. To ona decyduje o ilości danych, które mogą być kodowane przy użyciu elektroniki warstwy fizycznej, dla dowolnego zakresu częstotliwości. Szerokość pasma wynika z różnicy między wyższą i niższą granicą częstotliwości kanału komunikacyjnego. Na przykład szerokość pasma od 902 do 928 Megaherców wynosi 26 MHz. Szerokość ta jest podstawowym czynnikiem ograniczającym szybkości sygnalizowania i przesyłania danych wszelkich technologii komunikacyjnych, które używałyby tego pasma.
Proszę nie mylić szerokości pasma z częstotliwością! W podanym przykładzie
pasmo ma szerokość 26 MHz (928-902). Czymś innym jest jednak efektywna częstotliwość spektralna rzędu 26 MHz. Uzyskanie takiej wydajności, wymaga pasma o częstotliwości nie mniejszej niż 26 MHz. Gdyby jednak pasmo znajdowało się w zakresie częstotliwości niższych, na przykład od 26 MHz do 52 MHz, to uzyskanie z niego wspomnianej wydajności częstotliwościowej byłoby praktycznie niemożliwe. Przykładowe pasmo 902-928 MHz nie zostało wybrane przypadkowo. Jest to pasmo, w którym działa wiele pagerów oraz innego rodzaju urządzeń o niskim paśmie i szerokim zakresie transmisji. Nieefektywność tego zakresu łatwo dostrzec, analizując działanie urządzeń wchodzących w skład bezprzewodowych sieci lokalnych funkcjonujących w tym zakresie. Z pasma o szerokości 26 MHz potrafią one „wyżyłować" transmisje rzędu ledwie 2 Mbps.

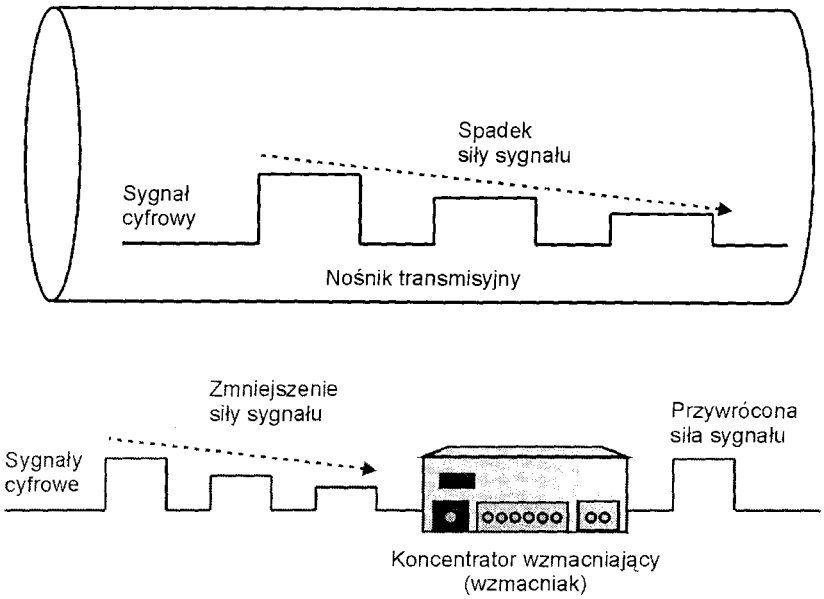
Nadużywanie oraz zwiększanie zakresu znaczeniowego terminu „pasmo częstotliwości" może być przyczyną mylenia go z innymi pojęciami, takimi jak „szybkość przesyłania danych" czy „częstotliwość transmisji". Terminy starsze, takie jak „szybkość bodowa", przyczyniają się do dalszej komplikacji opisu przesyłania danych. Bod (ang.. Baud), to liczba dyskretnych (odrębnych i dających się policzyć) sygnałów przesyłanych w ciągu sekundy. Jest to artefakt pochodzący z wczesnych dni komunikacji, kiedy modemy 300 bodowe zastępowały I 10 bodowe złącza akustyczne. Modemy owe potrafiły przesyłać 300 bitów na sekundę (bps). Przy wyższych szybkościach sygnalizowania w jednym bodzie zakodowane może być wiele bitów. W związku z tym modemy 2400 bps w rzeczywistości przesyłają 1200 bodów, umieszczając po dwa bity w jednym bodzie.
W środowiskach sieci lokalnych rzeczywista częstotliwość transmisji jest najczęściej ignorowana. Potencjalna szerokość pasma dla częstotliwości przesyłania danych oraz częstotliwości obsługiwanych przez nośnik jest dokładniej mierzona za pomocą liczby faktycznie przesyłanych bitów na sekundę (bps) niż za pomocą herców (Hz). Herc, będący miarą liczby cykli na sekundę jest zwykle używany do opisywania bardzo niskich częstotliwości, takich na przykład, które są używane do komunikacji głosowej. Przy tak niskich częstotliwościach liczba Hz zwykle odpowiada liczbie bitów na sekundę.
Dla wyższych częstotliwości zależność między Hz i bps staje się niejasna. Na przykład, potencjalna szerokość pasma kabla światłowodowego określana jest za pomocą długości fali. Kabel taki o średnicy 62,5 mikrona może być opisany jako kabel 850 nanometrów (nm), gdzie liczba opisująca odpowiada maksymalnej długości fali, jaką kabel może obsługiwać na określonej długości (zwykle równej 100 metrom).
Urządzenia komunikacyjne stosujące wyższe częstotliwości są zwykle bardziej skomplikowane i używają bardziej zaawansowanych technik kodowania. Techniki te pozwalają na kodowanie więcej niż jednego bitu w jednym hercu oraz mogą być wykorzystane do wywoływania wrażenia niezawodności określonej technologii. Dobrym przykładem wywoływania dobrego wrażenia za pomocą zmieniania miar szybkości komunikacji są sieci Gigabit Ethernet. Sieci te (opisywane bardziej szczegółowo w rozdziale 8 pt.. „Szybsze sieci Ethernet") utworzone zostały przez połączenie warstwy łącza danych specyfikacji IEEE 802.3 z warstwą fizyczną kanału światłowodowego. Mimo nazwy jego maksymalna szybkość transmisji wynosi nie 1 Gigabit, lecz I Gigabod. Częstotliwości używane w tego typu sieciach różnią się w zależności od rodzaju wybranego nośnika.
Gigabit Ethernet używa schematu kodowania znanego jako 4B/SB, które to oznaczenie informuje, iż warstwa fizyczna tej architektury tworzy 5-bitowy wzór dla każdego 4bitowego łańcucha znaków, który ma być przesyłany. Owa 20% nadmiarowość w warstwie fizycznej automatycznie zmniejsza maksymalne pasmo przesyłania do 800 Mbps. Po wyłączeniu z transmisji bitów organizujących ramki i protokoły Gigabit Ethernetu, bitów organizujących strukturę pakietu i protokoły IP oraz bitów wszelkich innych protokołów obsługujących warstwy od 4 do 7, okazuje się, że szybkość efektywna („netto" - można powiedzieć) jest jeszcze niższa. Nawet przyjmując, że „brutto" wynosi ona 800 Mbps, nadal daleko jej do 1 Gigabitu. Tylko jak tu uczciwie nazwać sieć „800 Mbps Ethernet", jeśli nazwie tej zdecydowanie brakuje tak potrzebnego w marketingu czaru, jaki niewątpliwie posiada nazwa „Gigabit Ethernet".
Reasumując, pasmem przesyłania jest maksymalna ilość danych, które mogą być przesłane za pomocą określonego nośnika.

1.3.1.2 Znaczenie odległości

Potencjalna szerokość pasma każdego typu nośnika ograniczona jest zarówno przez częstotliwość, jak i przez odległość, które nośnik ten potrafi obsługiwać. Odległość jest czynnikiem krytycznym z kilku powodów. Po pierwsze, im większa odległość, tym więcej czasu upływa zanim sygnał dotrze do swego miejsca przeznaczenia. Po drugie, większa odległość zwiększa również rozpraszanie sygnału, czego efektem jest powolne, acz nieuchronne zmniejszanie siły sygnału. W końcu osiąga ona poziom minimalny, poniżej którego sygnał staje się niezrozumiały dla odbiorcy. Ta forma pogarszania się jakości sygnału znana jest jako tłumienie.
Kable znacznej długości bezpośrednio zwiększają ekspozycję kabli i sygnałów, które są nimi przesyłane na zakłócenia spowodowane szumami elektromagnetycznymi powodowanymi przez lampy fluoroscencyjne, przewody napięcia zmiennego itp. Ten rodzaj pogarszania się jakości sygnału to zniekształcanie.

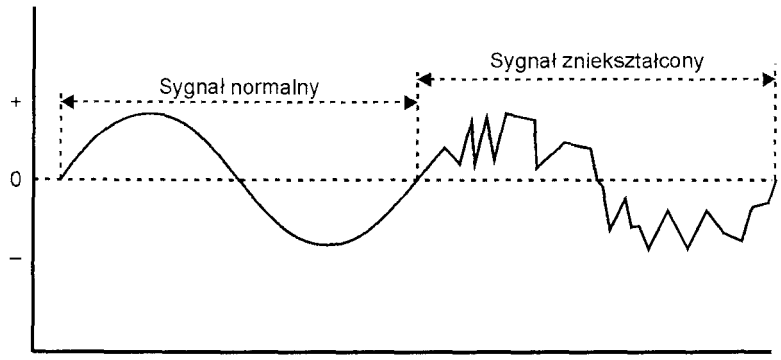
1.3.1.3 Tłumienie

Jednym z ubocznych skutków przesyłania prądu elektrycznego, w której to postaci sygnały są przesyłane, jest powolne, lecz ciągłe zmniejszanie siły sygnału. Prócz stałego oddawania energii w postaci promieniowania, energia sygnału zużywana jest również na przemieszczanie go w nośniku. Pamiętać należy, że sygnał jest falą elektromagnetyczną, która -jak każdy inny rodzaj prądu elektrycznego - w miarę poruszania się w nośniku zużywa własną energię do pokonywania jego oporów. Wynikiem tego jest nieustanne osłabianie siły sygnału, bez zmiany jego kształtu.
Wszystkie sygnały, w tym również opt.yczne, są sygnałami elektromagnetycznymi. Mogą one przemieszczać się nie tylko przez metale będące przewodnikami, takie jak miedź, lecz również przez inne przewodniki, takie jak kable światłowodowe. Tyle że fale płynące przez światłowód mają dużo wyższą częstotliwość niż fale poruszające się w przewodnikach metalicznych. Z tego powodu również fale świetlne są podatne na tłumienie. Szklany nośnik, w którym podróżują fale świetlne, ma bardzo niewielką oporność elektryczną. Tłumienie więc powodowane jest przez jedną z dwóch przyczyn:
• rozpraszaniem się sygnału promieniście w kierunku na zewnątrz kabla,
• kolizjami z nieczystościami znajdującymi się światłowodzie. Pamiętać należy, że światło jest niezwykle wrażliwym rodzajem promieniowania. Ulega całkowitemu wytłumieniu zawsze, gdy natrafia na nieprzeźroczyste ciało stałe.
Na rysunku 3.3 przedstawiony jest tłumiony sygnał.
Rysunek 3.3. Tłumienie sygnału.
Im dłuższy kabel, tym więcej oporów sygnał musi pokonać na swej drodze. Opory te wytłumiają (osłabiają) stopniowo sygnał, tak że po przebyciu pewnej drogi dane niesione przez ów sygnał przestają być czytelne dla ewentualnego odbiorcy. W związku z tym warstwa fizyczna określa szereg specyfikacji, których przestrzeganie gwarantuje, że sytuacje takie nie będą się zdarzać.
Tłumienie nie stanowi problemu w sieciach, w których kable są na tyle krótkie, że siła sygnału jest wystarczająca do tego, by dotrzeć do wszystkich przyłączonych do niej urządzeń. Jeśli wymagane są dłuższe kable, można na nich zamontować wzmacniaki. Wzmacniakierm jest każde urządzenie, które odbiera sygnał przychodzący. wzmacnia jego siłę, a następnie wysyła go dalej. Funkcjonalny opis wzmacniaka zawiera rysunek 3.4.
Rysunek 3.4. Regeneracja sttumionego sygnału a pomocą wzmacniaka.
Często zapomina się o tej podstawowej funkcji wzmacniaka. Dla tych, którzy o niej zapomnieli, wzmacniak jest po prostu koncentratorem, a koncentrator służy wyłącznie do udostępniania portów służących przyłączaniu do sieci wielu urządzeń. Nie jest to, jak już wcześniej wspomniałem, zgodne z prawdą. Poza tym pamiętać należy, że wzmacniaki mają również swoje słabe strony. Nie potrafią mianowicie rozpoznawać, które z przychodzących danych i struktur ramek są uszkodzone, a tym samym nie potrafią przywrócić im postaci pierwotnej. Wszelkie błędy i zniekształcenia otrzymane przez wzmacniak są przezeń wzmacniane.
Zniekształcenie



Zniekształcenie jest dotkliwym problemem dotyczącym przesyłania sygnałów, który polega na niepożądanej zmianie kształtu sygnału, zachodzącej podczas jego transmisji. Jeśli zniekształceniu uległy dane lub zawierające je ramki, dane te stają się bezużyteczne. Mimo zniekształcenia, Komputer-adresat nadal potrafi rozpoznać, że to on jest adresatem

i przesyła strumień danych do protokołów warstwy łącza danych. Protokoły te rozpoznają zniekształcenie i informują nadawcę uszkodzonej ramki o potrzebie ponownego jej przesłania. Taki sposób usuwania zniekształceń nazywany jest sposobem biernym. Przykład zniekształconego sygnału przedstawiony-jest na rysunku 3.5.
Rysunek 3.5. Sygnał
-nieks_talconv.
Istnieje kilka sposobów aktywnego zapobiegania zakłóceniom:
• Ścisłe przestrzeganie wszelkich zasad instalacji, które zostały załączone razem z nośnikiem. Używać należy odpowiedniej jakości przewodów i upewnić się, że zostały one zainstalowane i zakończone („terminowane'°) zgodnie z zaleceniami producenta. Należy też zwrócić uwagę na to, aby długość kabli nie przekraczała ustalonych limitów.
• Rozpoznawanie potencjalnych przyczyn zakłóceń oraz prowadzenie kabli z dala od ich źródeł.
• Używanie protokołów, które zdolne są do rozpoznawania i automatycznego korygowania wszelkich ewentualnych błędów transmisji. Protokoły te omówione zostały w rozdziale 5 pt.. „Warstwa łącza danych" oraz w rozdziale 12 pt.. „Protokoły sieciowe".
Jednym z rodzajów zniekształceń sygnałów jest przesłuch (ang.. crosstalk). Przesłuch jest zjawiskiem charakterystycznym dla okablowania podwójną skrętką. W kablach tego typu w osłonce znajduje się nie jeden przewód, lecz ich para, przy czym jeden wspomaga wykonywanie innej funkcji niż drugi: na przykład jeden służy do wysyłania danych, a drugi do ich odbierania. Przesłuch polega na niepożądanym przedostawaniu się sygnałów z jednego przewodu do drugiego, umieszczonego w tej samej osłonce. Ta forma zakłóceń określana jest czasem jako przesłuch zbliżony, czyli NEXT (skrót od ang.. Near-End-End Cross-Talk).
• Mniejsze zakłócenia sygnału mogą zostać usunięte za pomocą urządzenia, które zamiast sygnał powtarzać, regeneruje go. Przykładem urządzenia powtarzającego uszkodzone sygnały jest prosty wzmacniak. Wzmacnia i powtarza otrzymany sygnał, niezależnie od tego, czy jest on poprawny, czy zniekształcony.



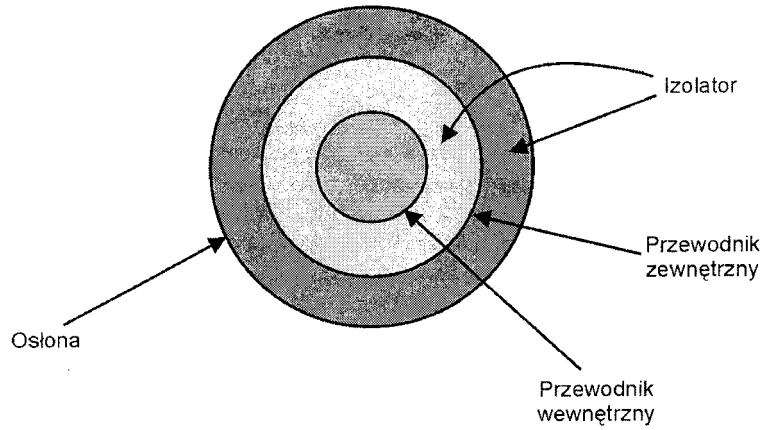
• Inaczej działa router, który jest urządzeniem inteligentnym potrafiącym wykorzystać zasady logiki do przetwarzania każdego pakietu. Jeśli struktura nagłówka pakietu została uszkodzona podczas przesyłania, nie będzie mógł pakietu tego przesłać dalej. Z powodu braku trasy dla pakietów uszkodzonych są one przez router usuwane, przy czym pozostawiany jest protokół kontroli hosta-nadawcy w celu rozpoznania uszkodzenia i przeprowadzenia ponownej transmisji pakietu.
• Jeśli natomiast pakiet jest otrzymany w stanie nieuszkodzonym i opatrzony jest adresem, który router potrafi rozpoznać, trasowany (kierowany) jest on dalej. Trasowanie polega na wyszukaniu w tablicy tras adresu docelowego w celu określenia zarówno następnego hopu (przejście pakietu przez jeden router to jeden hop), jak i co nawet ważniejsze interfejsu, z którego pakiet powinien zostać wysłany. Następnie pakiet jest wysyłany. Routery regenerują przy tym pakiety, nie poprzestając wyłącznie na ich wzmacnianiu.
Routery działają w warstwie 3 modelu referencyjnego OSI. W związku z tym otrzymują i wysyłają pakiety, a nie ramki. Ramki są strukturami właściwymi dla warstwy 2, warstwy łacza danych.

1.3.2 Nośniki transmisji fizycznej

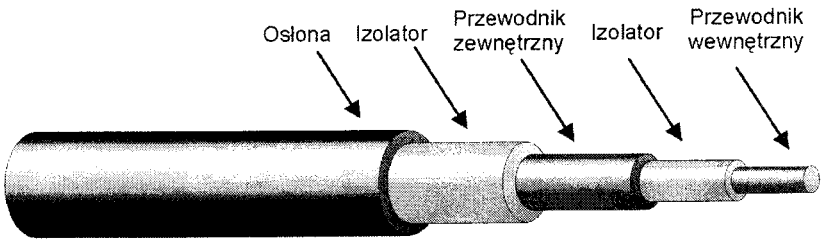
Aby warstwa fizyczna mogła w niezawodny sposób wykonywać swoje funkcje odbierania i wysyłania, przyjęte muszą zostać pewne założenia dotyczące przestrzeni znajdującej się pomiędzy interfejsami dwóch komunikujących się urządzeń. Protokoły warstwy fizycznej muszą na przykład przyjąć określony poziom wydajności właściwy dla obsługiwanego przez nie typu nośnika. Warstwa fizyczna spodziewa się, że faktyczna wydajność nośnika będzie zgodna z wybranymi założeniami, niezależnie od tego, co trzeba będzie uczynić, aby zgodność tę utrzymać.
Na nośniki transmisji składają się wszelkie sposoby przesyłania sygnałów generowanych przez mechanizmy warstwy I modelu OSI. Ze względu na tą definicję nośniki można podzielić na materialne i niematerialne. Nośniki niematerialne omówione są rozdziale 4 pt.. „Niezupełnie-fizyczna warstwa fizyczna".
Materialnymi nośnikami transmisji są: • kabel koncentryczny,
• skrętka dwużyłowa,
• kabel światłowodowy.
• Wymienione nośniki, ich mocne i słabe strony oraz zastosowania w sieciach LAN przedstawione są w pozostałej części niniejszego rozdziału.

1.3.2.1 Kabel koncentryczny

Kabel koncentryczny, często nazywany „koncentrykiem", składa się z dwóch koncentrycznych (czyli współosiowych) przewodów. Kabel ten jest dosłownie współosiowy, gdyż przewody dzielą wspólną oś. Najczęściej spotykany rodzaj kabla koncentrycznego składa się z pojedynczego przewodu miedzianego biegnącego w materiale izolacyjnym. Izolator (lub inaczej dielektryk) jest okolony innym cylindrycznie biegnącym przewodnikiem, którym może być przewód lity lub pleciony, otoczony z kolei następną warstwą izolacyjną. Całość osłonięta jest koszulką ochronną z polichlorku winylu (PCW) lub teflonu. Przekrój poprzeczny przedstawiony jest na rysunku 3.6.
Rysunek 3.6. Przekrój poprzeczny typowego kabla koncentrycznego.
Perspektywa boczna kabla koncentrycznego, ukazująca jego anatomię, przedstawiona jest na rysunku 3.7.
Mimo że kable koncentryczne wyglądają podobnie, mogą charakteryzować się różnymi stopniami impedancji. Oporność ta mierzona jest za pomocą skali RG (ang.. Radio Grade). Na przykład, specyfikacja (1OBase-2-2 Ethernet używa kabla RG-58, którego oporność wynosi 50 omów dla kabla o średnicy I centymetra. W specyfikacji warstwy fizycznej l 1OBase-2-2 przekłada się to na szybkość przesyłania rzędu 10 Mbps dla odległości do 185 metrów.



Rysunek 3.7. Perspektywa boczna kabla koncentrycznego



Istnieją również inne rodzaje kabli koncentrycznych o zastosowaniach specjalnych, takie jak kable koncentryczne z krążkami zabezpieczającymi, ale nie są one stosowane w sieciach lokalnych, więc nie będziemy ich tu omawiać.
Zaletą kabli koncentrycznych jest to, że potrafią obsługiwać komunikację w pasmach o dużej szerokości bez potrzeby instalowania wzmacniaków. Kabel koncentryczny był pierwotnym nośnikiem sieci Ethernet.
Od tego czasu został on zupełnie wyparty przez specyfikacje warstwy fizycznej Ethernetu oparte na skrętce dwużyłowej.
Przyczyny tej dezaktualizacji są proste. Kabel koncentryczny jest dość wrażliwą strukturą. Nie znosi ostrych zakrętów ani nawet łagodnie przykładanej siły gniotącej. Jego struktura łatwo bowiem ulega uszkodzeniu, co powoduje bezpośrednie pogorszenie transmisji sygnału.
Dodatkowymi czynnikami zniechęcającymi do stosowania kabli koncentrycznych są ich koszt i rozmiar. Okablowanie koncentryczne jest droższe aniżeli skrętka dwużyłowa ze względu na jego bardziej złożoną budowę. Każdy koncentryk ma też co najmniej 1 cm średnicy. W związku z tym zużywa on olbrzymią ilość miejsca w kanałach i torowiskach kablowych, którymi prowadzone są przewody. Niewielka nawet koncentracja urządzeń przyłączonych za pomocą kabli koncentrycznych zużywa całe miejsce, którym przewody mogą być prowadzone.
Dziś zastosowanie koncentryków ogranicza się do przesyłania sygnałów szerokopasmowej telewizji kablowej jej abonentom.

1.3.2.2 Skrętka dwużyłowa

Okablowanie skrętką dwużyłową, od dawna używane do obsługi połączeń głosowych, stało się standardową technologią używaną w sieciach LAN. Skrętka dwużyłowa składa się z dwóch dość cienkich przewodów o średnicy od 4 do 9 mm każdy (czasem kable oznaczane są wg miary AWG, czyli American Wire Gauge - wg tej skali przewody takie mają rozmiary od 18 do 24 AWG). Przewody pokryte są cienką warstwą polichlorku winylu (PCW) i splecione razem. Skręcenie przewodów razem równoważy promieniowanie, na jakie wystawiony jest każdy z dwóch przewodów znosząc w ten sposób zakłócenia elektromagnetyczne (nazywane EMI), które inaczej przedostawałyby się do przewodnika miedzianego.
Grubość (czyli średnica) przewodu wpływa bezpośrednio na jego sprawność. Większa średnica przewodu oznacza szersze potencjalne pasmo komunikacji i większą długość maksymalną kabla. Niestety, w miarę wzrostu szerokości pasma ze wzrostem średnicy przewodu wzrastają również jego właściwości tłumienia. Nie można więc stosować kabli o dowolnej średnicy i długości, lecz takie, które umożliwiają

zachowanie równowagi między szerokością pasma i długością przewodu. Określanie tej równowagi jest jednym z ważniejszych punktów specyfikacji warstwy fizycznej. Specyfikacja nie precyzuje, w jaki sposób kabel ma być prowadzony, lecz określa jego grubość, rodzaje terminatorów (oporników ograniczających), maksymalne długości kabla i szerokość jego pasma.

Nie skręcone, lecz proste przewody miedziane podobnej średnicy znane są jako anteny - sprawniej wychodzi im chwytanie otaczającego je promieniowania elektromagnetycznego niż zachowywanie go. Jak już wspomniałem i powiem dokładniej w dalszej części niniejszego rozdziału, skrętka dwużyłowa jest skręcona nie bez przyczyny; a jest nią fakt, że skręcenie likwiduje dużą część zakłóceń sygnału: „na zakłócenie dobre skręcenie".

Dostępne są różne rodzaje, rozmiary i kształty skrętki dwużyłowej, począwszy od ‚jednoparowego" (czyli dwużyłowego) kabla telefonicznego aż do 600-parowych (1200 żyłowych) kabli dalekosiężnych. Niektóre z tych rozmaitości, takie na przykład jak powiązanie par przewodów razem, służą zwiększaniu pojemności kabla. Inne z kolei mają na celu zwiększenie jego przepustowości (wydajności). Niektórymi z wariantów zwiększających wydajność są:

• zwiększanie średnicy przewodnika,
• zwiększanie stopnia skręcenia (liczby skręceń w jednostce odległości),
• stosowanie kilku różnych stopni skręcenia na poziomie skręcania w wielożyłowe wiązki,
• ochrona par przewodów za pomocą metalowych osłonek.

W sieciach LAN najczęściej stosowane są cztery pary przewodów połączone razem w wiązki, które osłonięte są wspólną koszulką z PCW lub teflonu. Teflon jest dużo droższy i sztywniejszy (czyli np. nieporęczny w układaniu), ale nie wydziela trujących oparów podczas spalania (czyt. w razie ewentualnego pożaru). Ze względu na to kable kładzione we wszelkich kanałach dostarczających powietrze do pomieszczeń, w których znajdują się ludzie, muszą mieć osłonę z teflonu.

Dwoma najczęściej stosowanymi rodzajami skrętek ośmiożyłowych są ekranowana i nieekranowana.

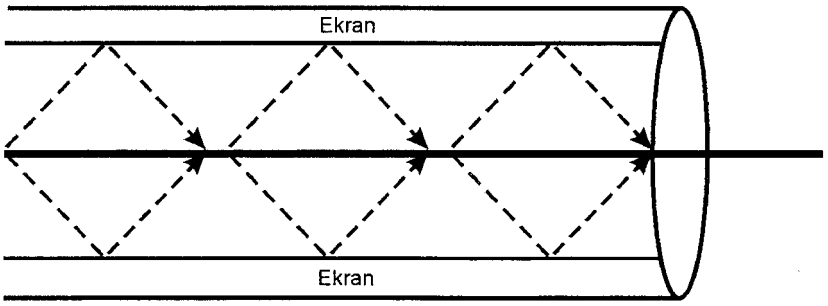
Ekranowana skrętka dwużyłowa

Ekranowana skrętka dwużyłowa (ang.. STP - shielded twisted pair) ma dodatkową warstwę folii lub metalowego przewodu oplatającego przewody. Taki ekran osłonowy znajduje się bezpośrednio pod powierzchnią koszulki kabla. Powodem wprowadzenia ekranowania była potrzeba użycia skrętek dwużyłowych w środowiskach podatnych na zakłócenia elektromagnetyczne i zakłócenia częstotliwościami radiowymi. W praktyce przeszkadza to jednak skrętce w poprawnym funkcjonowaniu. Podczas przesyłania sygnału przewodem miedzianym wytwarzane jest bowiem promieniowanie elektromagnetyczne. Omawiana sytuacja przedstawiona jest na rysunku 3.8, który będzie punktem odniesienia podczas dalszych rozważań dotyczących konsekwencji stosowania ekranowania dla skrętek dwużyłowych.

Rysunek 3.8. Normalne promieniowanie elektromagnetyczne.

Ekranowanie przewodu za pomocą metalowej osłony chroni przed promieniowaniem zewnętrznym. Zatrzymuje ono, niestety, również promieniowanie indukowane, czyli wytwarzane przez ten przewód podczas przesyłania nim sygnału. Zamiast więc rozchodzić się normalnie, jak ilustruje to rysunek 3.8, promieniowanie to zostaje odbite przez ekran i skierowane z powrotem do przewodu miedzianego, co może, z dużym prawdopodobieństwem, powodować uszkodzenie sygnału. Sytuację ta przedstawia rysunek 3.9.

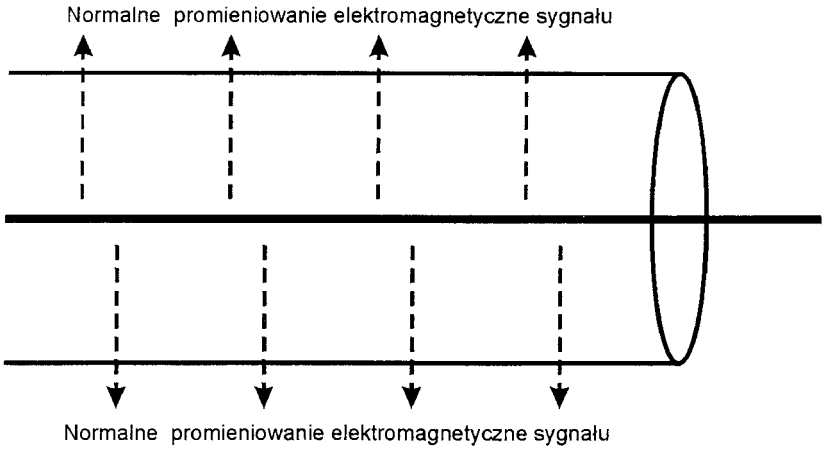
Rysunek 3.9. Ekranowane promieniowanie elektromagnetyczne.



Ekranowanie jest niezbędne do pomyślnego przesyłania sygnałów kablami biegnącymi na wprost, w których sygnały przewodzone są jednym przewodnikiem (takimi jak kabel koncentryczny), ale dla skrętek dwużyłowych jest ono częściej szkodliwe niż korzystne. Tym bardziej, że skrętka dwużyłowa korzysta z mechanizmu korekcji błędów polegającego na wzajemnym znoszeniu się wszelkiego rodzaju indukowanych zakłóceń.

Nieekranowana skrętka dwużyłowa

Również skrętka dwużyłowa nazywana UTP (ang.. Unshielded Twisted Pair) dostępna jest w wielu wersjach różniących się formą, rozmiarem oraz jakością. Jak wcześniej wspominałem, rozmiar dotyczy liczby par połączonych razem w jedną wiązkę. Rozmiarem standardowym okablowania sieci LAN jest kabel czteroparowy, czyli ośmiożyłowy.



Przewody ośmiożyłowej skrętki nieekranowanej podzielone są na cztery grupy po dwa przewody. Każda para składa się z przewodu dodatniego i ujemnego. Przewody nazywane są też wyprowadzeniami. Wyprowadzenia wykorzystuje się parami. Na przykład, jedna para wyprowadzeń obsługuje tylko wysyłanie, a inna tylko odbieranie sygnałów. Pozostałe wyprowadzenia w większości sieci lokalnych nie są używane.

Sieci lokalne, które przez ośmiożyłową skrętkę nieekranowaną przesyłają sygnały z prędkościami 100 Mbps i większymi, wykorzystują wszystkie jej osiem przewodów.

Kategorie wydajności

Okablowanie skrętką dwużyłową jest towarem handlowym. Oczekiwać można od niego w miarę niezmiennych właściwości - niezależnie od tego, kto jest jego producentem. Jest to możliwe dzięki pewnej normalizacji, która zaszła i stale zachodzi w przemyśle telekomunikacyjnym. Co ciekawe, standardów dotyczących skrętki dwużyłowej nie wprowadziła żadna konkretna organizacja, lecz powstały one w wyniku luźnej współpracy ANSI, FCC, EIA oraz wielu innych organizacji standardo-dawczych - dziś standardy te dotyczą nie tylko okablowania jako całości, lecz nawet jego elementów, takich jak terminatory.

Przyznać trzeba, że jest to odejście od standardowego sposobu ustanawiania standardów -tym bardziej, że istniejące specyfikacje nie podają nawet definicji skrętki dwużyłowej. Istnieją, jak wspomniałem, pewne wytyczne dotyczące okablowania i jego składników, ale skrętka dwużyłowa najpełniej określana jest za pomocą kategorii wydajności. Kategorie te definiowane są nie przez standardy fizyczne, lecz przez wydajność, z jaką działają. Innymi słowy, aby producent skrętki dwużyłowej mógł uczciwie sprzedawać ją jako zgodną z określoną kategorią, wystarczy, że udowodni odpowiedni poziom wydajności, niezależnie od budowy kabla, jego grubości i wszystkich innych szczegółów.

Pierwotnie istniała seria pięciu testów, które decydowały o zaszeregowaniu skrętki do odpowiedniej kategorii. Ponumerowane były od 1 do 5, a kabel, który spełnił ich wymagania, mógł być oznaczony jako kabel kategorii x, (ang.. Category x lub Cat-x-x), gdzie x jest numerem najwyższego pomyślnie złożonego testu. Dwie z owych pięciu kategorii okazały się najbardziej popularne wśród użytkowników - były to kategoria 3 i 5. Kategorie 1 i 2 zostały oficjalnie uznane za przestarzałe w roku 1995. Kategoria 4 oferuje pośredni (między 3 a 5) poziom wydajności, ale rzadko kiedy jest stosowana.

Kategoria 3 UTP (skrętki dwużyłowej nieekranowanej) oferuje pasmo o szerokości 16 MHz, które umożliwia przesyłanie sygnałów z prędkością do 10 Mbps na odległość maksymalną 100 metrów. Kategoria 4 obsługuje pasmo o szerokości 20 MHz, a kategoria 5 o szerokości 100 MHz. Szerokości pasm informują, dlaczego kategoria 4 nie zyskała dużej popularności wśród użytkowników. Była mianowicie postrzegana jako oferująca zbyt mały wzrost wydajności w stosunku do kategorii 3. Jeśli bowiem komuś nie wystarczała szerokość pasma oferowana przez kategorię 3, nie wystarczała mu też z pewnością kategoria 4, lecz dopiero 5. Przy założeniu, że wymagania dotyczące maksymalnej odległości są spełnione, kable kategorii 5 umożliwiają przesyłanie danych z prędkością 100 Mbps, 155 Mbps, a nawet 256 Mbps. Oczywiście zwiększanie

maksymalnej szybkości przesyłania danych zmniejsza jednocześnie odległość maksymalną. Jak widać, rozwiązania o wydajności wyższej niż 3 lecz niższej niż 5 okazały się nieefektywne ekonomicznie.

Zagadnienia specjalne

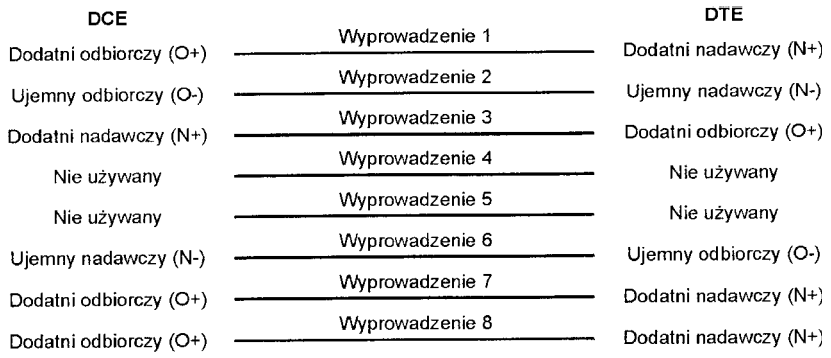
Skrętka dwużyłowa, niezależnie od jej kategorii i typu, używa odrębnych przewodów dla wyprowadzeń dodatnich i ujemnych, osobno dla funkcji wysyłania i funkcji odbioru. Aby dwa urządzenia mogły się komunikować, muszą one uzgodnić, które z nich będzie wysyłać, a które odbierać za pomocą jakich wyprowadzeń. Istnieje zatem możliwość nieodpowiedniego połączenia urządzeń ze sobą.

W normalnych warunkach urządzenia, które mają być połączone bezpośrednio (na przykład komputer i koncentrator), mają komplementarne (odpowiadające sobie) interfejsy, które udostępniają funkcję krzyżowania.

Przykład takiego złącza, zilustrowany na rysunku 3.10, przedstawia przypisania wyprowadzeń dla interfejsu RJ-45-45 sieci Ethernet 10 Mbps łączącego urządzenie komunikacyjne z końcowym. Urządzeniami komunikacyjnymi, dla przypomnienia. są na przykład porty koncentratora, a urządzeniami końcowymi na przykład karty sieciowe stacji roboczych oraz serwerów.

Owe interfejsy komplementarne umożliwiają bezpośrednie łączenie urządzeń bez powodowania konfliktów między nadawaniem i odbieraniem sygnałów. W normalnych warunkach urządzenie komunikacyjne można zawsze połączyć z urządzeniem końcowym i na odwrót, przy użyciu kabla ośmiożyłowego, którego przewody nie krzyżują się. Kabel taki nazywa się kablem prostym.

Rysunek 3.10. *Połączenie urządzeń komunikacyjnego i końcowego za pomocą kabla prostego.*

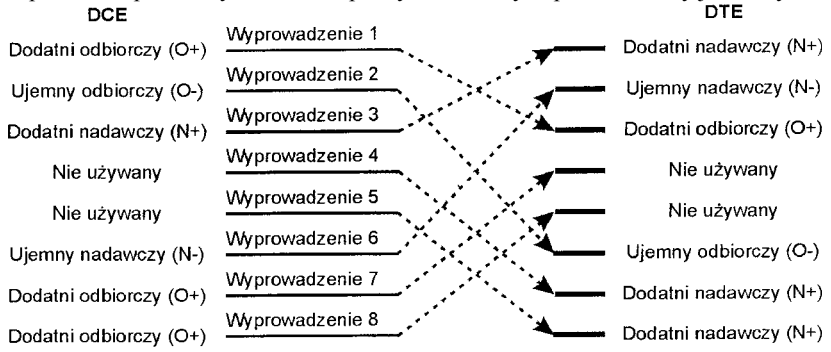


Terminy „urządzenie komunikacyjne" i „urządzenie końcowe" dotyczą każdego portu koncentratora wieloportowego z osobna. W związku z tym każde urządzenie wieloportowe może być jednocześnie urządzeniem komunikacyjnym i końcowym. Na przykład, niektóre z portów koncentratora wieloportowego mogą być urządzeniami komunikacyjnymi, podczas gdy pozostałe będą urządzeniami końcowymi. Porty są portami wyjścia, a tym samym urządzeniami komunikacyjnymi, gdy służą do przyłączania innych koncentratorów skonfigurowanych jako urządzenia komunikacyjne. Te z portów, które służą do

przyłączania urządzeń obliczeniowych, funkcjonują jako porty wejścia, czyli jako urządzenia końcowe. Starsze koncentratory miały zarezerwowany specjalny port służący do obsługi połączeń wyjściowych (port wyjścia). Niektóre inne porty takich koncentratorów można było ustawiać jako wejściowe lub wyjściowe w zależności od potrzeb.

Obecna generacja przełączanych koncentratorów o mieszanej architekturze odprowadziła port wyjścia prosto do lamusa. Dziś do tworzenia rozszerzonych sieci LAN używane mogą być szerokopasmowe magistrale przełączane. Do łączenia dwóch koncentratorów wystarczy więc jedynie kabel skrośny. Odpowiedni sposób używania kabli prostych i skrośnych przedstawiony jest na rysunku 3.11.

Rysunek 3.11. *Złącze komunikacyjne dwóch urządzeń końcowych za pomocą kabla skrośnego (RJ-45 10 Base-T).*

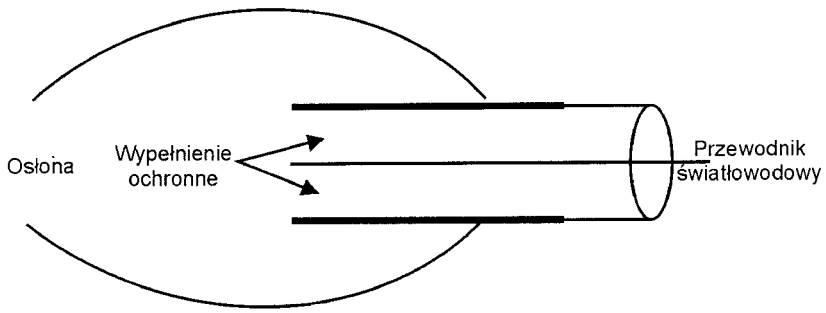


Kable skrośne muszą utrzymywać polaryzację przewodu fizycznego. Napięcia dodatnie i ujemne muszą być rozdzielone. Krzyżowanymi przewodami są więc: przewód dodatniego bieguna danych nadawanych (N+) z dodatnim przewodem odbioru (O+) oraz przewód ujemnego bieguna danych nadawanych (N-) z ujemnym przewodem odbioru (O-).

1.3.2.3 Kabel światłowodowy

Kable światłowodowe potrafią przenosić wyższe częstotliwości spektrum elektromagnetycznego - a mianowicie światło. Dostępne są one w bardzo wielu kształtach, rozmiarach i kategoriach długości fal. Na rysunku 3.12. przedstawiony jest ogólny schemat przekroju wzdłużnego kabla światłowodu.

Rysunek 3.12. Schemat przekroju wzdłużnego kabla światłowodu.

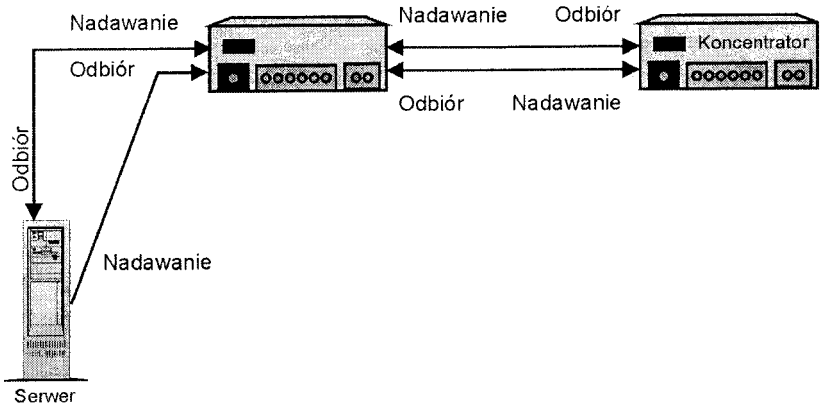


Kabel światłowodowy ma tylko trzy ewidentnie stałe atrybuty:

• W osi centralnej kabla biegnie dużej czystości nośnik opt.yczny zdolny do niezawodnego przenoszenia wzorów świetlnych na duże odległości

• Nieobecność sygnału elektrycznego oraz przewodnika miedzianego oznacza, że transmisje światłowodowe są względnie bezpieczne. W odróżnieniu od przewodu miedzianego, kabla światłowodowego nie można szpuntować ani podłączyć się do niego w żaden inny sposób. Pamiętać warto, że odpowiednio mocne ściśnięcie go powoduje rozbicie jego szklanych struktur.

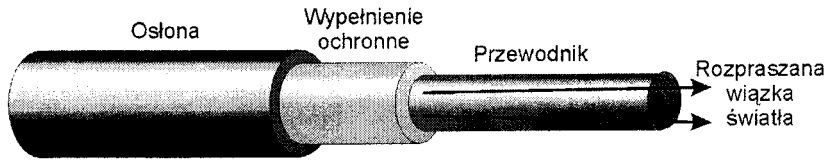
• Nośnik opt.yczny, jakim jest włókno szklane, pokryty jest koncentryczną ochronną warstwą plastiku.
Poza wymienionymi właściwościami charakteryzującymi wszystkie rodzaje kabli światłowodowych, mogą się one różnić prawie wszystkim. Zacznijmy od tego, że nośnikiem najczęściej jest szkło, ale równie dobrze może nim być opt.ycznej jakości plastik. Średnica światłowodu waha się od 5 mikronów do rozmiarów, które cokolwiek łatwiej dostrzec gołym okiem. Wszystkie rodzaje światłowodów zwykle biegną wiązkami po dwie i więcej par. Również włókna światłowodów dostępne są w wielu różnych rodzajach - z czego nie wszystkie nadają się do wykorzystania w sieciach LAN! Typowy kabel światłowodowy do zastosowania w sieciach lokalnych ma zwykle średnicę 62,5 mikronów. Obsługuje on wielomodową komunikację sterowaną za pomocą diody świetlnej.
Włókna zwykle opisywane są za pomocą pary liczb. Na przykład, oznaczeniem najczęściej w sieciach LAN stosowanego włókna szklanego jest 62,5/125. Pierwsza liczba podaje średnicę w mikronach, podczas gdy druga wyraża, również w mikronach, średnicę warstwy plastiku ochronnego.
Kable światłowodowe wykorzystywane są parami: jeden służy do wysyłania sygnałów, a drugi do ich odbioru. Ich zastosowanie w sieciach LAN zwykle ogranicza się do łączenia serwerów i koncentratorów. Oba te zastosowania oraz kierunki przesyłania sygnałów przedstawione są na rysunku 3.13.
Rysunek 3.13. Zastosowanie oraz kierunki przesyłania sygnałów w sieci LAN.



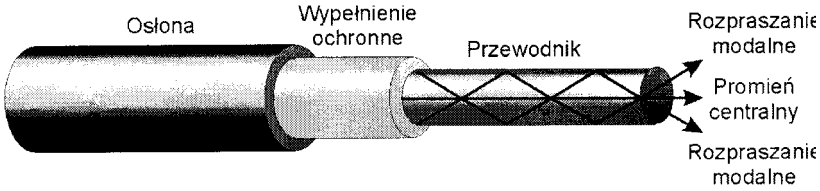
Światło może promieniować na zewnątrz, tak jak w przypadku latarni, lub może być skupione kierunkowo - jak w przypadku latarki. Nawet jednak światło ukierunkowane poddawane jest do pewnego stopnia rozpraszaniu. Każdy, kto kiedykolwiek używał latarki, wie, że rozmiar koła oświetlanego promieniami latarki zwiększa się wraz ze wzrostem odległości. Zjawisko rozpraszania światła niesie ze sobą poważne implikacje dotyczące przesyłania sygnałów opt.ycznych.
Jedną z nich jest podział transmisji światłowodowych na dwa rodzaje: • wielomodowe,
• jednomodowe.
Każdy z nich do komunikacji używa innych technik transmisyjnych. Mod, dla wyjaśnienia, jest rodzajem fali elektromagnetycznej.

1.3.3 Transmisja wielomodowa

Transmisja wielomodowa jest sterowana za pomocą diody świetlnej lub inaczej diody LED (ang.. *Light Emitting Diode* - dosł. „dioda emitująca światło"). A popsuta dioda LED, jak podaje stary żart, staje się diodą DED (ang.. *Dark Emitting Diode),* czyli „diodą emitująca ciemność". Skrót „DED" brzmi przy tym podobnie do przymiotnika „dead" (ang.. *martwa, zdechła),* co nadaje żartowi dodatkowy smaczek.
Diody świetlne są źródłem światła niespecjalnie skoncentrowanego; w związku z tym wymagają dość szerokiej ścieżki transmisji. Mają one też dosyć niską (jak dla światła) częstotliwość, więc szerokość ich pasma przesyłania również nie jest największa. Kluczową właściwością diod świetlnych jest ich niezdolność do wysyłania skoncentrowanej wiązki światła. Wysyłane światło ulega zatem rozpraszaniu. Stopień rozpraszania nakłada praktyczne ograniczenia na długość okablowania światłowodowego sterowanego za pomocą diody świetlnej.
Rysunek 3.14 przedstawia normalne rozpraszanie w kablu światłowodowym wiązki wysłanej przez diodę świetlną.
Rysunek 3.14. Rozpraszanie światła w wielomodowym
1s włóknie światłowodu.



Rozpraszanie przesyłanej wiązki świetlnej powoduje, że niektóre z jej promieni odbijają się od szklanej ściany nośnika. Kąt odbicia jest niewielki, w związku z czym światło nie ucieka do warstwy ochronnej, lecz odbijane jest pod kątem padania. Odbity promień porusza się pod tym samym kątem w kierunku środka przewodnika, napotykając po drodze promienie centralnej części wiązki światła, od których znów się odbija. Sytuacja ta jest przedstawiona na rysunku 3.15.



Rysunek 3.15. Odbicie jest powodem wielomodowości.

Warto pamiętać, że odbijana część promienia niesie ten sam sygnał, który niesiony jest przez jego centralną część, tyle że - ze względu na częste odbicia - promienie odbijane pokonać muszą drogę dłuższą niż promienie centralnej części wiązki. A że prędkość światła jest stała i wynosi ok. 300 000 km/s, to promienie centralnej części wiązki przybywają do celu na długo przed promieniami, które uległy wielokrotnemu odbiciu, czyli modami (stąd nazwa).

Ważniejszą nawet implikacją rozpraszania wielomodowego jest fakt zderzania się poszczególnych fotonów ze sobą. Ciągłe odbijanie się promieni niesie ze sobą możliwość przekroczenia w końcu centralnej osi przewodnika i wejście w konflikt z innymi sygnałami transmisji. Oznacza to, że przesyłanie wielomodowe jest podatne na tłumienie.

Wielomodowość transmisji może być również spowodowana przez nieodpowiednią terminację kabla światłowodu i/lub w wyniku nieodpowiedniego umocowania złączy w gniazdach interfejsów sprzętowych. Terminacja, która nie jest koncentryczna, zawsze zwiększa modowe rozpraszanie wiązki światła. Zwiększa ją również nieodpowiednie zamocowanie terminatorów w interfejsie sprzętu. Wielomodowości powodowanej tymi czynnikami nie da się przewidzieć. Zatem mimo iż wielomodowe rozpraszanie wiązki światła jest czymś normalnym, co należy brać pod uwagę i na czym nawet można polegać, to wspomniane dwa czynniki mogą owo rozpraszanie niespodziewanie zwiększyć. Powodem tego jest wejście promienia pod kątem w stosunku do osi centralnej przewodnika, w wyniku czego:

• kąt odbicia jest dużo większy niż w przypadku normalnego rozproszenia wiązki, • jedynie niewielka część wiązki przesyłana jest równolegle do osi kabla.

A to z kolei powoduje, że niewielka ilość danych osiąga swoje miejsce docelowe. Jeśli kąty odbicia są duże, to ich droga może ulec wydłużeniu do tego stopnia, że promienie centralne sygnałów wysyłanych później osiągną miejsce przeznaczenia wcześniej. Może to wprowadzić w błąd urządzenie odbierające sygnały, a nawet jeśli tak się nie stanie, to ich rozpoznawanie na pewno zajmie sporą część ograniczonych wszakże zasobów komputera.

Krótko mówiąc, ze sztucznego zwiększenia modowości (będącego - dla przypomnienia - efektem zainstalowania nieodpowiednich terminatorów lub nieodpowiedniego ich zainstalowania) nic dobrego wyniknąć nie może.

Ograniczenia kabli światłowodowych są równoważone przez ich zalety; takie jak na przykład cena, która w porównaniu z kosztem systemów jednomodowych jest bardzo

niska. Kable systemu wielomodowego są przy tym łatwiejsze do terminacji, gdyż są kilkakrotnie od kabli systemów jednomodowych dłuższe.

1.3.4 Transmisja jednomodowa

Włókna jednomodowe używają iniekcyjnej diody laserowej (ILD). Słowo LASER jest w zasadzie akronimem nieco dłuższej ang.ielskiej nazwy „Light Amplification by Stimulated Emission of Radiation", co po polsku oznacza dosłownie „wzmacnianie światka przez wymuszoną emisję promieniowania" - czyli po prostu laser. Lasery znane są ze znacznej koncentracji wiązki promieni. Wiązka ta ulega rozproszeniu, ale w stopniu niezauważalnym dla odległości właściwych sieciom lokalnym.

W systemach jednomodowych do wysyłania sygnału przez szklany nośnik stosowany jest laser. Dzięki temu, że sygnał ten prawie wcale nie ulega rozpraszaniu, nawet najbardziej zewnętrzne części jego wiązki nie zaczynają nawet dotykać wewnętrznych ścianek włókna przewodzącego, nie mówiąc o jakichkolwiek odbiciach. Strumień danych przesyłany jest więc równolegle do osi przewodnika na całej jego długości i dociera do miejsca przeznaczenia w jednym modzie, czyli w całości w jednym punkcie czasu.

Włókna jednomodowych kabli światłowodowych mają zwykle od 5 do 10 mikronów średnicy i otoczone są ochronnym wypełnieniem o średnicy 125 mikronów. Wysokie koszty kabli i sprzętu laserowego w połączeniu w dużą szerokością udostępnianego pasma sprawiają, że technologia ta bardziej nadaje się do wykorzystania przy tworzeniu wysokiej jakości infrastruktur informacyjnych niż do sieci lokalnych. Największe zastosowanie jak dotychczas znalazła w komercyjnych sieciach telefonicznych.

1.3.5 Podsumowanie

Warstwa fizyczna modelu referencyjnego OSI dzieli ramki danych na strumienie jedynek i zer (czyli włączeń i wyłączeń), które są następnie przesyłane w sieci. Mimo że nośnik nie jest częścią warstwy fizycznej, warstwa ta musi określić wymagania co do wydajności przewodnika znajdującego się między komunikującymi się urządzeniami. Najczęściej jest nim kabel - jako taki jest więc istotnym dodatkiem do warstwy fizycznej. Decyzja co do tego, który kabel jest najbardziej odpowiedni, zależy wyłącznie od sytuacji. Rozważyć należy wiele czynników, w tym takie jak:

• fizyczna struktura kabla,
• możliwość dołączania nowych kabli,
• wymagania odnośnie wydajności poszczególnych części sieci, • osadzenie podstawy kabla,
• przepisy przeciwpożarowe i inne regulacje,
• specyfikacje warstwy fizycznej obsługiwane przez architekturę sieci.

Najprawdopodobniej najlepszym z rozwiązań okaże się połączenie dwóch, może trzech rodzajów kabli. Obecnie standardem *de facto* na rynku są: dla przyłączania stacji nieekranowana skrętka dwużyłowa kategorii 5, a dla łączenia koncentratorów i serwerów - kabel światłowodowy 62,5/125 mikronów (oznaczony jako 850 nm).
Całkiem możliwe, że żadne z przedstawionych rozwiązań nie będzie odpowiadało odczuwanym potrzebom. Jeśli tak, to rozwiązanie znajduje się prawdopodobnie w rozdziale następnym, zatytułowanym „Niezupełnie-fizyczna warstwa fizyczna".

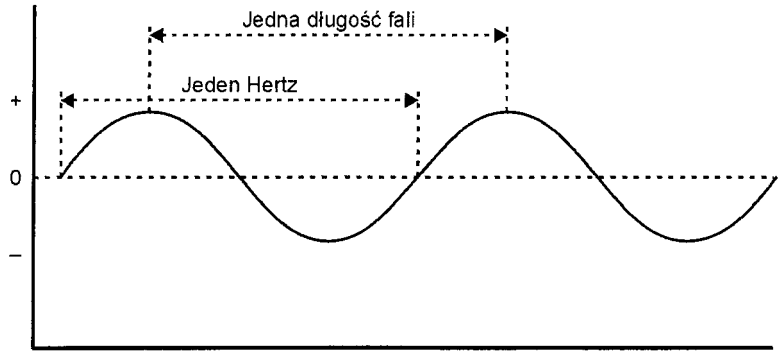
1.4 Rozdział 4 Niezupełnie-fizyczna warstwa fizyczna

Mark A.Sportack
Sieci LAN mogą być tworzone również na podstawie niematerialnej warstwy fizycznej. I są. Od kilku już lat sieci tego typu znane są jako bezprzewodowe sieci LAN. Jednak dopiero niedawno - wraz z przyjęciem specyfikacji IEEE 802.11 - ustanowione zostały pierwsze standardy dotyczące tego sposobu przesyłania danych. Uprzednio, ze względu na brak standardów, każdy producent mógł tworzyć sobie własne, nie współpracujące z innymi „standardy" systemów - niczym pułapki zastawiane na niebacznych klientów. Różnice funkcjonalne między tymi „standardami" skupiały się wokół technologii oraz technik transmisji. Nowy standard IEEE 802.11 zintegrował obsługę wielu z tych uprzednio nietypowych rozwiązań.
W niniejszym rozdziale zapoznamy się z wieloma niematerialnymi implementacjami warstwy fizycznej, ich technikami transmisji oraz właściwościami. Przyjrzymy się również niektórym z bardziej widocznych punktów specyfikacji 802.11.

1.4.1 Spektrum elektromagnetyczne

Sieci lokalne, które przesyłają dane i protokoły za pomocą niekablowych przewodników nazywane są bezprzewodowymi sieciami LAN. Aby dokładniej zgłębić ich naturę, zapoznamy się najpierw ze zjawiskiem spektrum elektromagnetycznego. Spektrum elektromagnetyczne określa fizyczne właściwości przesyłania w zależności od częstotliwości fali nośnej.
„Częstotliwość" jest jednym z tych zagadkowych słów, których znaczenie każdy rozumie, ale które mało kto potrafi zdefiniować. Częstotliwość to coś więcej niż miejsce na skali radia. To częstość, z jaką prąd elektryczny zmienia stan (z dodatniego na ujemny lub na odwrót). Częstość ta mierzona jest zwykle w hercach (Hz). Jeden cykl, czyli jeden
herc, odpowiada zmianie na sinusoidzie fali o 360%. Innymi słowy jeden cykl rozpoczyna się napięciem zerowym, które rośnie do osiągnięcia wartości maksymalnej, a następnie zmniejsza swą wartość i mijając pierwotny poziom zerowy osiąga minimum, aby następnie wzrosnąć z powrotem do zera. To jest właśnie jeden cykl.
Do określania częstotliwości wyższych, takich jak światło i promienie X oraz gamma, używana jest długość fali, czyli odległość między szczytem fali a jej dołem.
Herc oraz długość fali przestawione są na rysunku 4.1.
Rysunek 4.1. .Jeden herc a jedna długość fali.

Herc, uniwersalna miara częstotliwości wibracji elektrycznych, czyli cykli na sekundę, wzięła nazwę od Heinricka Hertza. Pan Hertz odkrył istnienie fal radiowych w roku 1883.



Skala spektrum elektromagnetycznego rozciąga się między 0 Hz a 10'° Hz. Dla porównania - ludzkie ucho rozpoznaje wibracje o częstotliwości od 20 do ok. 16 000 - 20 000 Hz. Granice zakresu słyszalnego są różne dla różnych osób i zmieniając się w czasie. Ludzkie ucho najlepiej dostosowane jest do odbierania wibracji głosowych, których energia przenoszona jest z częstotliwością od 3 000 do 4 000 Hz.
Spektrum rozciąga się daleko poza zakres słyszalny dla ucha ludzkiego. Opuściwszy go, przechodzi do zakresu fal świetlnych, czyli kolejno do zakresu fal podczerwonych, widzialnych, ultrafioletowych oraz promieni X i promieni gamma. W miarę wzrostu częstotliwości fal świetlnych, liczba zer wartości służących do ich przedstawiania zwiększa się lawinowo. A że zależność między długością fali i jej częstotliwością jest odwrotna, to wartościom używanym do określania długości fali również przybywa zer, tyle że po przecinku.
' Autor oryginału czyni tu (i w kilku innych miejscach książki) pewne uproszczenie myślowe, utożsamiając „jeden cykl" drgania z „jednym hertzem". Należy o tym pamiętać, czytając np. o „liczbie zmian poziomu sygnału w jednym hertzu" czy też o „liczbie hertzów w czasie 1 milisekundy%'. Pomimo wątpliwego uzasadnienia poprawności takich uproszczeń zdecydowaliśmy się pozostawić je w tekście, gdyż nie mają one większego znaczenia dla zasadniczej treści niniejszej książki (przyp. red.).
Całe spektrum elektromagnetyczne oraz obie miary przedstawione są na rysunku 4.2.
Rysunek 4.2. Spektrum elektromagnetyczne.

1.4.1.1 Charakterystyki spektrum

Charakterystyka propagacji (rozchodzenia się) fal zmienia się w spektrum wraz ze wzrostem częstotliwości. Fale charakteryzowane są przez trzy główne właściwości, przynajmniej w zakresie bezprzewodowych sieci LAN, a są nimi:

• łamliwość,

• kierunkowość, • szerokość pasma.

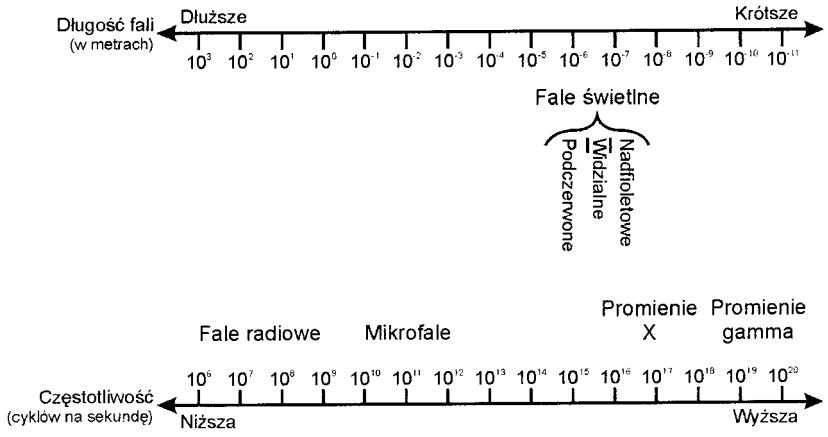
Niższe częstotliwości odznaczają się wybitną trwałością. Potrafią przenikać - w różnym stopniu, ale potrafią- przez nieprzezroczyste ciała stałe. Na przykład fale radiowe potrafią przenikać przez wszystkie materiały - poza najgęstszymi. Czyli nie są łamliwe.

Fale radiowe rozchodzą się we wszystkich kierunkach. Przedstawia to rysunek 4.3.

Fale radiowe o wyższych częstotliwościach zachowują się podobnie jak fale świetlne. Widzialna część spektrum poprzedzona promieniowaniem podczerwonym jest bardzo wąska, a zaraz za nią na skali widzimy nadfioletową część spektrum.

Wyższe częstotliwości sygnałów są bardzo łamliwe. Im są wyższe, tym gorzej przenikają przez nieprzeźroczyste ciała stałe, niezależnie od gęstości tych ciał. Fale takie nie muszą również rozchodzić się w sposób przedstawiony na rysunku 4.3, lecz bardziej ukierunkowany - taki na przykład, jaki przedstawiony jest na rysunku 4.4.

Możliwość kierunkowego przesyłania wibracji elektromagnetycznych wzrasta wraz ze wzrostem częstotliwości. Im więcej wibracja ma herców, tym bardziej można ją skoncentrować. Na przykład, lampa wysyła światło we wszystkich kierunkach. Żarówki latarek również wysyłają światło we wszystkich kierunkach, a do ich koncentrowania



Rysunek 4.3. Wielokierunkowe promieniowanie sygnału.

Rysunek 4.4. Kierunkowo zorientowane przesyłanie fal świetlnych.

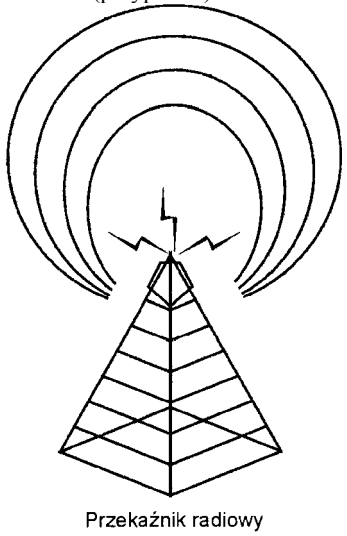
i ukierunkowywania najczęściej stosuje się zwierciadło wklęsłe. Mimo to światło żarówki nie daje się skoncentrować do tego stopnia, co na przykład światło lasera. Fale o niskich częstotliwościach są bowiem trudniejsze do ukierunkowania niż fale mające wysokie częstotliwości.

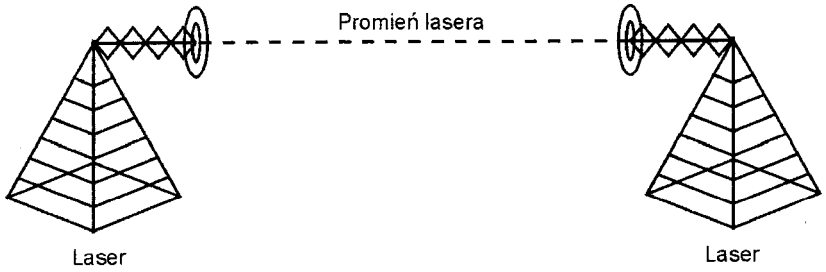
1.4.1.2 Spektrum a szerokość pasma

Szerokość pasma jest kolejnym szeroko stosowanym terminem, który jest równie niezrozumiany. Jest to dosłownie szerokość kanału komunikacyjnego mierzona w hercach. Pasmo jest bowiem słowem, za pomocą którego pierwotnie nazywane było to, co obecnie nazywa się kanałem. Kanał natomiast to zakres częstotliwości służący określonym celom komunikacyjnym.

Szerokość pasma jest więc różnicą między najwyższą i najniższą częstotliwością kanału komunikacyjnego. Na przykład, pasmo od 902 do 928 megaherców (MHz - milion herców) zarezerwowane jest2 do nieregulowanej komunikacji radiowej obsługującej potrzeby przemysłowe, naukowe i medyczne. Szerokość pasma tej części spektrum wynosi 26 MHz.

W LISA (przyp. red.)





Miarą częstotliwości w komunikacji głosowej lub analogowej jest herc. Dla komunikacji cyfrowej lepszym jednak wskaźnikiem jest liczba bitów na sekundę (bps). Natomiast szybkości przesyłania danych w systemach światłowodowych wyrażane są za pomocą długości fali, chyba że opisywana jest konkretna technologia, taka jak FDDI, kiedy miarą częstotliwości jest liczba przesyłanych bitów na sekundę.
Jednostki „Hz" i „bps" są ze sobą silnie skorelowane dodatnio, jako że w ciągu jednego cyklu przesłany może być jeden bit. Niestety, ze względu na różnice między technikami modulacji wymaganymi do przetwarzania sygnałów z postaci analogowej na cyfrową i odwrotnie, liczba bitów jest zawsze mniejsza od liczby cykli. Okazuje się więc, że szerokość pasma jest tylko miarą częstotliwości, nie informującą wiarygodnie o szybkości, z jaką przesyłane są dane. Teoretycznie jednak możliwe jest przesyłanie danych z szybkością równą liczbie cykli (czyli szerokości pasma), w związku z tym szerokość pasma została uznana za miarę potencjalnej szybkości przesyłania bitów danych.
Zawsze, gdy przesyłane są drgania elektryczne, niezależnie od tego, czy w postaci analogowej, czy cyfrowej, ich serię nazywamy sygnałem. Sygnały mogą być emitowane w różnych pasmach częstotliwości, choć zostało ogólnie przyjęte, że sygnały w częstotliwościach niższych niż 300 000 Hz nie są zdatne do wykorzystania przez żadne urządzenia elektryczne.

1.4.2 Co to oznacza?

Celem tego badania fizycznych właściwości różnych fragmentów spektrum elektromagnetycznego jest stworzenie odpowiedniego kontekstu, który umożliwi następnie przeprowadzenie wartościowej analizy różnych rozwiązań bezprzewodowego przesyłania sygnałów w sieciach lokalnych.
Każda technologia transmisji działa na podstawie konkretnej części spektrum, które w dużej mierze decyduje o jej właściwościach fizycznych. Podstawowe zależności między właściwościami fizycznymi przedstawione są w tabeli 4.1.
Tabela 4.1.
Zależne od spektrum właściwości sygnału.
Właściwości fizyczne technologii transmisji z kolei narzucają określoną specyfikę działania bezprzewodowych sieci lokalnych. Właściwości bezprzewodowych sieci LAN dotyczą:
• maksymalnego efektywnego zakresu,
• możliwości przenikania materiałów o różnych strukturach i formach, • maksymalnej szybkości transmisji.

Niska częstotliwość	Wysoka częstotliwość
Wielokierunkowe promieniowanie przesyłanego sygnału	Możliwość skoncentrowania przesyłanego sygnału
Trwały sygnał	Łamliwy sygnał
Mała szerokość pasma	Duża szerokość pasma

Wybór spektrum wpływa na maksymalne rozmiary efektywnego działania bezprzewodowych sieci LAN na wiele sposobów. Jednym z nich jest Komisja Łączności3, która określa maksymalne długości efektywnej transmisji, zwłaszcza dla trwalszych niskich częstotliwości. Jej zarządzenia ograniczają moc z jaką przesyłane mogą być sygnały w poszczególnych pasmach. Mimo że nie są one natury technicznej, są równie rzeczywiste i równie wiążące.
Stosowana technika transmisji jest decydująca dla określenia maksymalnej obsługiwanej szybkości przesyłania danych. Szybkość przesyłania danych jest zawsze niższa od szybkości transmisji, gdyż nie uwzględnia ona informacji dodatkowych, takich jak protokoły sieci LAN, a poza tym sposoby wprowadzania danych do sieci są często nieefektywne.
Wspomniane powyżej zagadnienia tworzą podstawy funkcjonalnych różnic między poszczególnymi technologiami transmisji bezprzewodowych. Ważne jest, by umieć dostrzec ich wpływ na wydajność sieci oraz zrozumieć konsekwencje technik transmisyjnych przed wybraniem odpowiedniej wersji bezprzewodowej sieci LAN.

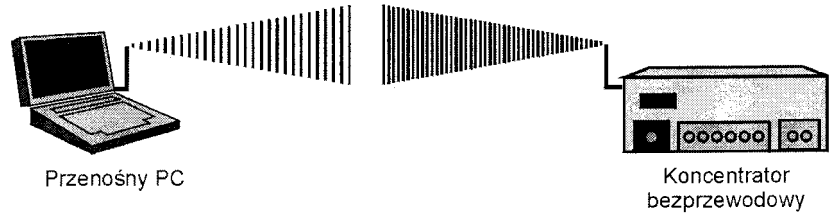
1.4.3 Bezprzewodowe sieci LAN

Stosowanie fal radiowych lub świetlnych w celu transmisji ramek i protokołów w sieci lokalnej bez użycia kabli (przewodów) nazywane jest Bezprzewodowa komunikacja w sieci LAN. Nazwa ta jest jednak myląca. Wiele rzekomo bezprzewodowych sieci LAN nadal korzysta z kabli - tyle, że z mniejszej ich ilości.
Różnorodność technologii transmisji bezprzewodowych w sieci LAN widać co najmniej na czterech przykładach przedstawionych poniżej. Można bowiem zrealizować:
• bezprzewodowe przyłączanie stacji,

• bezprzewodowe połączenia między komputerami sieci każdy-z-każdym, • bezprzewodowe połączenia międzywęzłowe,
• bezprzewodowe mostkowanie.
Każde z tych rozwiązań korzysta z technologii transmisji bezprzewodowych w inny sposób.

1.4.3.1 Bezprzewodowe łączenie stacji

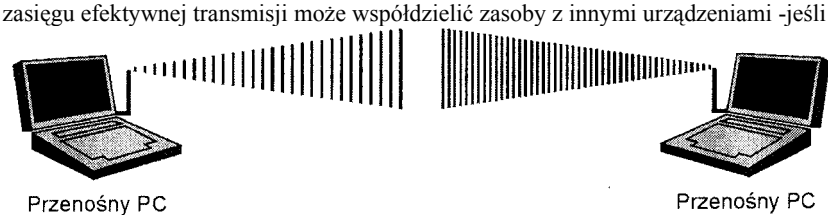
Bezprzewodowe przyłączanie stacji umożliwia użytkownikom komputerów przenośnych
korzystanie z nich w celu ustanowienia połączenia z siecią LAN bez potrzeby wykorzystywania dedykowanego, kablowego połączenia z koncentratorem. Przewody nadal są potrzebne do przyłączania poszczególnych stacji do nad-biornika. Omawiane rozwiązanie przedstawione jest na rysunku 4.5 przedstawiającym lapt.opa połączonego z przenośną anteną za pośrednictwem gniazda (i karty) PCMCIA.
W USA (przyp. red.)
PCMCIA jest obecnie szeroko znaną nazwą urządzenia wielkości karty kredytowej, które wkłada się do odpowiedniego portu komputera przenośnego, co umożliwia wykonywanie kilku dodatkowych funkcji.
Rysunek 4.5. Bezprzewodowe przyłączanie stacji.
Antena komputera przenośnego przedstawionego na rysunku 4.5 przesyła dane upakowane w ramkach do anteny urządzenia zwanego koncentratorem bezprzewodowym, jak również odbiera dane od niej przychodzące. Na rysunku anteny komputera i koncentratora są przedstawione jako do nich przymocowane, ale w rzeczywistości najczęściej anteny są oddzielnymi modułami, które mogą - a w przypadku koncentratora zwykle powinny - znajdować się daleko od urządzenia, które obsługują. Pozwala to na umieszczenie ich w miejscu umożliwiającym lepszą transmisję (czyli zwykle wyżej), co zwykle poprawia wydatnie jakość połączeń. Bezprzewodowy koncentrator wyposażony jest w port, który łączy go fizycznie z bardziej już konwencjonalnym, bo opartym na kablach, szkieletem sieci LAN. Jedyną bezprzewodową częścią tego rozwiązania jest połączenie między pecetem, a jego koncentratorem.



Rozwiązanie powyższe przydaje się osobom znajdującym się często poza biurem.

1.4.3.2 Bezprzewodowe łączenie komputerów w sieci każdy-z-każdym

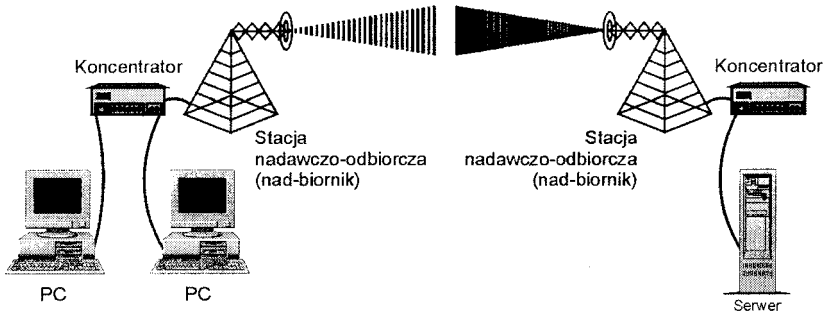
Proste (lecz o małej przepustowości) bezprzewodowe sieci LAN łączyć mogą komputery równorzędne. Rysunek 4.6 przedstawia bezprzewodową sieć każdy-z-każdym, która może być złożona w dość prosty sposób i dzięki której każde urządzenie znajdujące się w zasięgu efektywnej transmisji może współdzielić zasoby z innymi urządzeniami -jeśli oczywiście posiada odpowiednie uprawnienia.



Rysunek 4.6. Bezprzewodowe połączenia między komputerami w sieci
każdy z każdym
Dość nieformalna natura sieci każdy-z-każdym jest bezpośrednim powodem bardzo niskiej wydajności bezprzewodowych sieci równorzędnych.

1.4.3.3 Bezprzewodowe łączenie koncentratorów

Kolejny sposób wykorzystania technologii bezprzewodowych przedstawiony jest na rysunku 4.7. Wykorzystuje on jedno urządzenie nadawania i jedno urządzenie odbioru, a do każdego z nich przyłączona jest grupa stacji roboczych. Stacje te przyłączone są do konwencjonalnego koncentratora. Port wyjścia takiego koncentratora jest przyłączony do nad-biornika radiowego używanego do komunikacji z podobnym nad-biornikiem również przyłączonym do szkieletu sieci LAN za pomocą kabla. Jedyną bezprzewodową częścią tego rozwiązania jest komunikacja koncentratora z koncentratorem. Zaletą takiego podejścia jest to, że likwiduje ono potrzebę prowadzenia kabli od szaf rozdzielczych do stacji użytkowników. Same stacje mogą natomiast być w prosty sposób połączone ze sobą przy użyciu kabli „terminowanych" (zakończonych opornikami ograniczającymi „terminatoram i").
Rysunek 4.7. Bezprzewodowe łączenie
koncentratorów.
Rozwiązanie zilustrowane na rysunku 4.7 niesłychanie ułatwia tworzenie sieci w budynkach, które są trudne do okablowania, czyli na przykład zabytkowych lub zbudowanych tak, że wprowadzenie dodatkowych instalacji łączyłoby się z nadmiernymi kosztami, czyli na przykład mających stropy z betonu bez kanałów do prowadzenia kabli.



Mimo że w przykładzie przedstawionym na rysunku 4.7 do komunikacji używane są fale radiowe, można je zastąpić dowolnymi falami innego rodzaju, choćby takimi, jak podczerwone czy laserowe.

1.4.3.4 Bezprzewodowe mostkowanie

Technologie bezprzewodowe mogą również znaleźć zastosowanie w innych częściach sieci LAN. Na przykład dwie okablowane sieci LAN mogą być ze sobą połączone za pomocą bezprzewodowego mostu. Sytuację tę przedstawia rysunek 4.8.

Mostkowanie bezprzewodowe umożliwia łączenie sieci LAN znajdujących się względnie blisko siebie, ale mających również średnice bliskie maksymalnym. Eliminuje to koszty zakupu dwóch routerów, jak również miesięczne koszty dzierżawienia linii, która w innym razie wymagana byłaby do połączenia tych sieci.

Rysunek 4.8. Bezprzewodowe mostkowanie bezprzewodowych sieci l.4N.

Mostki bezprzewodowe obsługują pasma szerokości 2 Mbps na odległość kilku kilometrów, czyli więcej niż zaoferować może linia dzierżawiona T1 (o szerokości pasma 1,544 Mbps) bez okresowych kosztów dzierżawienia. Mostkowanie bezprzewodowe jest dość ekonomiczne: za porównywalne koszta początkowe (2 urządzenia nadawczoodbiorcze lub 2 routery) z połączenia międzysieciowego korzystać można za darmo tak długo, jak pozostaje ono sprawne.

Przedstawione przykłady ilustrują niektóre z najbardziej typowych implementacji technologii bezprzewodowych w sieciach LAN. Istnieją jeszcze inne sposoby ich wdrażania, ale związane są one z technologią transmisji pojedynczej, więc przedstawione są w jej kontekście.

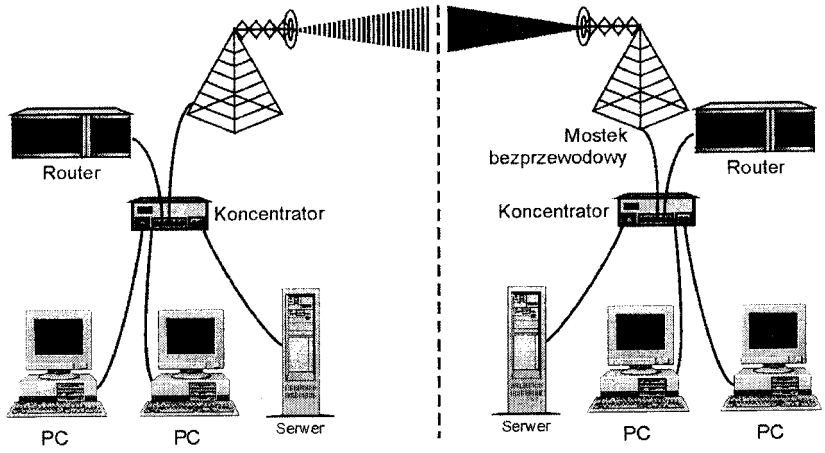
1.4.3.5 Technologie transmisji

Istnieją cztery różne technologie transmisji, z których każda używa innej części spektrum elektromagnetycznego. Są nimi:

• radio szerokopasmowe
• radio wąskopasmowe, radio pojedynczego pasma
• podczerwień • laser

Efektywne zakresy transmisji pierwszych dwóch metod ograniczone są przez zarządzenia komisji FCC4 dotyczącymi maksymalnych dopuszczalnych mocy transmisji (w watach). Zasięgi dwóch ostatnich technologii również limitowane są przez komisję FCC, choć bardziej nawet ogranicza je ich niezdolność do przenikania wszelkich ciał stałych o jakiejkolwiek gęstości. Fale radiowe na przykład mogą przenikać przez większość ścian budynków - poza najbardziej gęstymi. Do osłabienia promieniowania o częstotliwościach

° W USA (przyp. red.)



radiowych w stopniu dającym się zauważyć „gołym uchem" potrzebne są zwykle ściany żelbetonowe wielokrotnie wzmacniane (o gęstej konstrukcji stalowej). Natomiast promieniowanie świetlne, takie jak podczerwone czy laserowe, jest dużo bardziej łamliwe: kartka papieru lub nawet kłęby dymu czy mgła potrafią je całkowicie wytłumić.

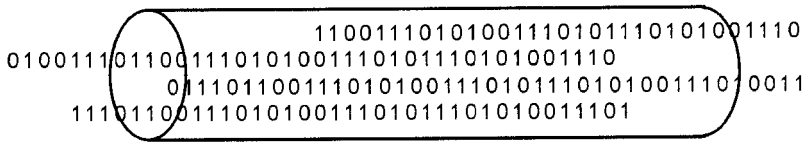
1.4.4 Częstotliwość radiowa szerokiego spektrum

W Stanach Zjednoczonych pasma 902-928 MHz oraz 2.4-2.4835 GHz spektrum elektromagnetycznego przypisane zostały do użytku przemysłowego, naukowego i medycznego. Ostatnio również pasmo 5.725-5.850 GHz zostało w USA udostępnione jako pasmo transmisji radiowych szerokiego spektrum.
Korzystanie z tych zakresów nie podlega licencjonowaniu, co oznacza, że komisja FCC - poza ustaleniem wytycznych dotyczących urządzeń, za pomocą których z pasm tych można korzystać - nie ingeruje w to, kto ich używa. Inaczej jest z pozostałą częścią spektrum. Zwykle osoba lub jednostka organizacyjna (na przykład stacja radiowa) w zamian za zobowiązanie się do przestrzegania wytycznych ustalonych przez Komisję otrzymuje pozwolenie na korzystanie z określonego zakresu pasma w określonym regionie geograficznym.
W połączeniu z technologiami transmisji szerokopasmowej używać można głównie dwóch technik transmisji. Określają one sposób korzystania z „niezupełnie-fizycznych" nośników transmisji. Technikami tymi są:
• niebezpośrednia sekwencja częstotliwości,
• bezpośrednia sekwencja częstotliwości.

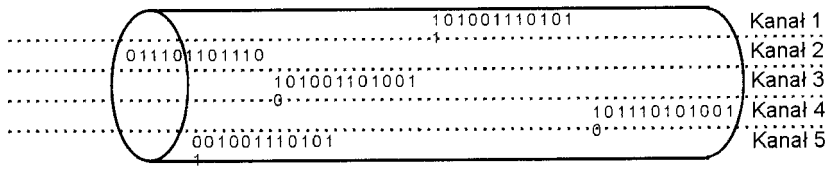
1.4.4.1 Niebezpośrednia sekwencja częstotliwości

Jest to technika używana tylko w połączeniu z systemami transmisji radiowych szerokiego spektrum. Rozproszone spektrum udostępnia bowiem zakres nieregulowanych częstotliwości radiowych. Niebezpośrednią sekwencję częstotliwości przedstawić można jako złoty środek między transmisjami wąskopasmowymi (w paśmie podstawowym) i szerokopasmowymi (w paśmie szerokim). Porównanie pasm podstawowego i szerokiego umożliwi nam lepsze dostrzeżenie różnic między nimi. Transmisje pasmem podstawowym korzystają z całej dostępnej szerokości pasma jako jednego kanału transmisyjnego: jeden sygnał przesyłany jest całym pasmem. Na rysunku 4.9 przedstawiony jest strumień binarny całkowicie wypełniający kanał, którym jest przesyłany. Dobrym przykładem mechanizmu przesyłania pasmem podstawowym jest Ethernet. Nośnikiem transmisji jest tu zawsze cała dostępna szerokość pasma, niezależnie od tego, czy ma ona 10, 100 czy więcej MHz.
Rysunek 4.9. Transmisja pasmem podstawowym.

Natomiast transmisja pasmem szerokim dzieli dostępne pasmo na wiele mniejszych kanałów. Każdy taki kanał obsługuje przesyłanie innego sygnału. Przedstawia to rysunek 4.10. Transmisja szerokopasmowa używana jest na przykład do przesyłania sygnałów telewizji kablowej. Pojedynczy kabel koncentryczny dostarcza pasmo, które zostało skanalizowane. W każdym kanale niesione są inne sygnały, mimo że przesyłane są za pomocą jednego nośnika.
Niebezpośrednia sekwencja częstotliwości, tak jak przesyłanie szerokopasmowe, dzieli pasmo na wiele kanałów. Każdy z tych kanałów może jednocześnie przesyłać jeden sygnał. Różnica w stosunku do przesyłania sygnałów pasmem szerokim polega na tym, że sygnał przeskakuje z kanału na kanał w zadanym tempie i zgodnie z zadaną sekwencją. Przedstawia to rysunek 4.11.
Rysunek 4.10. Transmisja szerokopasmowa.



Rysunek 4.11. Transmisja wykorzystująca niebezpośrenią sekwencję c=ęstotliwości.



Niebezpośrednia sekwencja częstotliwości ma swoje zalety. Po pierwsze, minimalizuje wpływ zakłóceń na jakość transmisji. Zakłócenia, zwłaszcza radiowe i elektromagnetyczne, mogą zniekształcać przesyłane sygnały. Zwykle zakłócenie dotyczy pewnej określonej i w miarę niezmiennej częstotliwości. Niebezpośrednia sekwencja częstotliwości zmniejsza możliwość szkodliwego wpływu tego rodzaju zakłóceń na transmisję.
Ważniejszą jeszcze korzyścią, którą oferuje transmisja tego rodzaju, jest możliwość umieszczania jednostek wielodostępnych w obszarze działania innych tego rodzaju urządzeń. Zachodzenie się obszarów działania takich urządzeń (np. bezprzewodowych koncentratorów) przedstawione jest na rysunku 4.12. Gdyby w przedstawionej sytuacji dane przysyłane były za pomocą jednej częstotliwości, przekaźniki wchodziłyby we wzajemny konflikt i przepustowość każdego z nich uległaby znacznemu obniżeniu. Korzystanie z szerokopasmowego systemu sekwencji niebezpośredniej częstotliwości zmniejsza możliwość kolizji. Dzięki temu znacznie można zwiększyć gęstość rozmieszczenia użytkowników bez zmniejszania wydajności sieci LAN.
Największą jednak korzyścią płynąca z transmisji danych za pomocą niebezpośredniej sekwencji częstotliwości jest zapewniane przez nie bezpieczeństwo. Każdy, kto chciałby zapoznać się z treścią transmisji, musi wykonać trzy czynności:
• obejść zabezpieczenia fizyczne chroniące przed dostępem do ograniczonego faktycznego pasma częstotliwości,

Rysunek 4.12. Zachodzenie się obszar-ów transmisji.

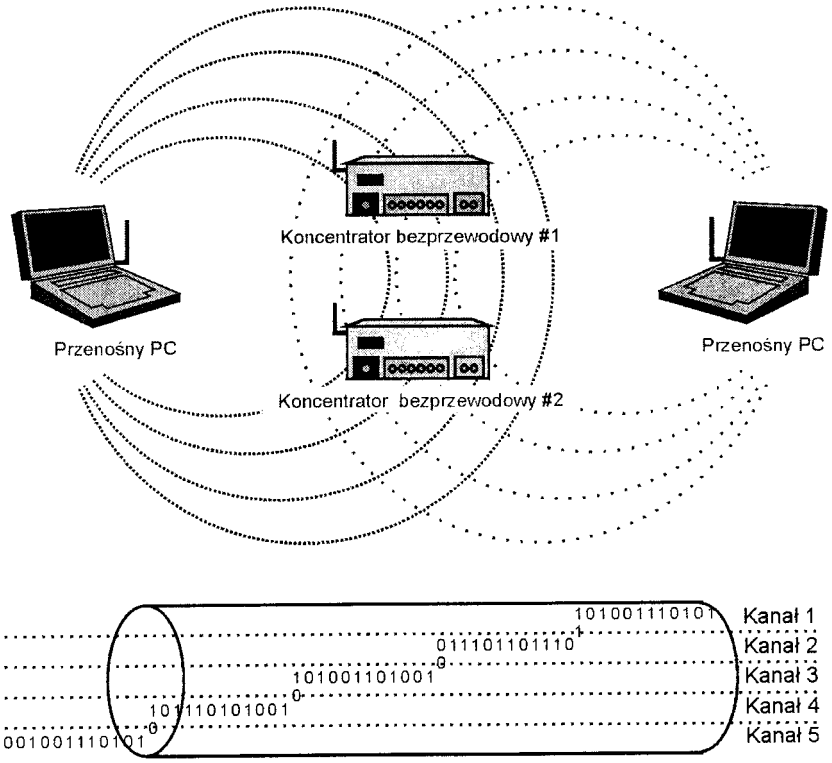
• przechwycić sygnały przesyłane różnymi kanałami,
• zrozumieć przypadkową z pozoru sekwencję transmisji.

Radiowe transmisje szerokopasmowe są pomysłem aktorki Hedy Lamarr. W 1940 roku wymyśliła ona taki sposób zabezpieczania transmisji radiowych. Dwa lata później otrzymała prawa patentowe do tego pomysłu. Niestety, minąć musiało kilka dekad, zanim pierwsi mężczyzni dojrzeli do wykorzystania jej pomysłu.

1.4.4.2 Sekwencja bezpośrednia

Również sekwencję bezpośrednią stosować można wyłącznie razem z systemami transmisji częstotliwości radiowych szerokiego spektrum. W odróżnieniu jednak od niebezpośredniej sekwencji częstotliwości, w ramach której kanały transmisji zmieniane są w pseudolosowy sposób, metoda sekwencji bezpośredniej do przesyłania sygnałów wykorzystuje dostępne kanały po kolei, czyli sekwencyjnie (stąd nazwa - „sekwencja" bezpośrednia). Zasada ta przedstawiona jest na rysunku 4.13. Większa prostota jej algorytmu powoduje, że łatwiej jest obejść zabezpieczenia systemu wykorzystującego do przesyłania danych sekwencję bezpośrednią niż niebezpośrednią.

Rysunek 4.13. Transmisja przy użyciu sekwencji bezpośredniej.



Obchodzenie zabezpieczeń systemu transmisji o sekwencji bezpośredniej szerokiego spektrum nadal wymaga od obchodzącego pokonania zabezpieczeń fizycznych chroniących pasmo transmisyjne. I tak jak w przypadku techniki niebezpośredniej sekwencji częstotliwości, również tu podsłuchiwać trzeba wszystkie kanały jednocześnie, tyle że sekwencja odczytywanych dzięki temu sygnałów nie będzie quasi-przypadkowa, lecz bezpośrednia (stąd nazwa - sekwencja bezpośrednia).

Instytut IEEE ostatnio zaproponował normalizację bezprzewodowych sieci LAN, wprowadzając specyfikację 802.11, która dotyczy obsługi obu systemów transmisji w widmie rozproszonym: zarówno bezpośredniej, jak i niebezpośredniej sekwencji częstotliwości. Szczegółowo specyfikacja ta omówiona jest w podrozdziale „Standard IEEE 802.11", w dalszej części niniejszego rozdziału.

Zalety

Sygnały przesyłane przy użyciu szerokiego spektrum są trudniejsze do podsłuchania, jako że rozszerzony jest zakres, w którym wysyłane są podsłuchiwane dane, a poza tym każdy bit transmisji przesyłany jest innym pasmem - według ustalonego algorytmu. Złamanie tego rodzaju zabezpieczeń jest prawie niemożliwe. Wymaga bowiem uzyskania fizycznego dostępu do efektywnego pasma transmisji i rozpoznania wzoru, za pomocą którego wybierane są kanały, którymi sygnały są następnie wysyłane.

Udostępnienie ostatnio przez komisję FCC częstotliwości 2.4-2.4835 GHz oraz 5.7255.850 GHz oznacza, że technologie szerokiego spektrum nie muszą być ograniczone do niskich szerokości pasma. Zakresy wyższych częstotliwości umożliwiają osiąganie wydajności zbliżonych do wydajności przewodowych sieci LAN.

Systemy szerokiego spektrum są przy tym względnie tanie z tego względu, że pasma przesyłania nie muszą być licencjonowane. W związku z tym producenci urządzeń do takich systemów mogą je dostarczyć po dużo niższej cenie niż ich odpowiedniki przeznaczone tylko do obsługi dedykowanego pasma. Poza tym nie trzeba składać podań do komisji FCC o przyznanie licencji na korzystanie z określonej częstotliwości na określonym obszarze. Dzięki temu taka sieć lokalna może działać zarówno szybciej, jak i taniej niż podobna sieć utworzona na podstawie technologii częstotliwości dedykowanych.

Wady

Jednym z większych problemów korzystania z częstotliwości radiowych jest fakt, że nie pozwalają one na prowadzenie komunikacji w pełni dupleksowanej (czyli komunikacji dwukierunkowej) za pomocą jednej tylko częstotliwości. Doświadczył tego każdy, kto zapoznał się z półdupleksową (jednokierunkową) naturą komunikacji radiowej podczas na przykład korzystania z krótkofalówki. Krótkofalówki wykorzystują metodę „naciśnijżeby-mówić" transmisji. Tylko jedna strona połączenia może mówić za jednym razem. Wszyscy pozostali uczestnicy połączenia mogą słuchać. Radio, jak widać, umożliwia nadawanie lub odbieranie, ale nigdy nie umożliwia wykonywania obu czynności jednocześnie.
W sieciach Ethernet jednokierunkowa natura fal radiowych wymusza odejście od standardowego protokołu CSMA/CD (wielodostępu do łącza sieci z badaniem stanu kanału i (/) wykrywaniem kolizji) na rzecz protokołu CSMA/CA (wielodostępu do łącza sieci z badaniem stanu kanału i (n unikaniem kolizji}.

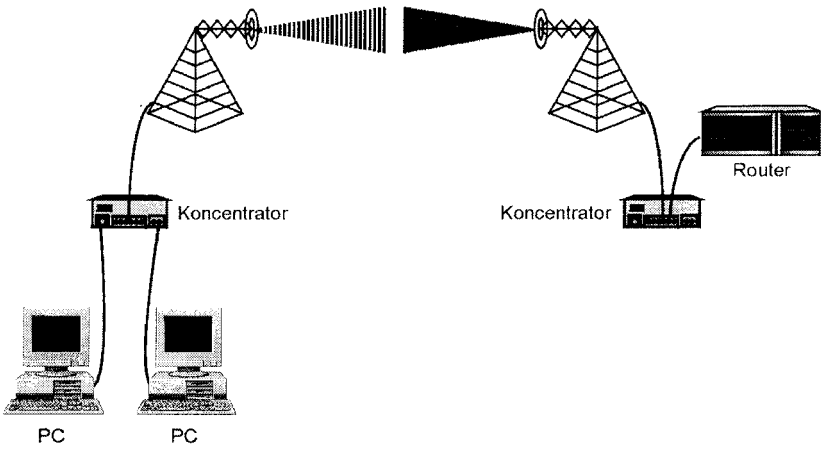
Ze względu na to oraz na narzuty wynikające z rozszerzenia pasma transmisji, rzeczywista przepustowość takiej sieci zmniejsza się do około 2 Mbps. Maksymalna efektywna przepustowość kablowego Ethernetu po odjęciu nagłówków CSMA/CD wynosi około 5-5,5 Mbps. Mało. W nowoczesnych sieciach powoduje to powstawanie wąskich gardeł i umożliwia korzystanie jedynie z najmniej wymagających aplikacji.
Jednym z rozwiązań problemu półdupleksowej natury fal radiowych jest wykorzystywanie do komunikacji dwóch kanałów: jednego do nadawania sygnałów, a drugiego do ich odbierania. Dzięki temu przepustowość sieci może być utrzymana na poziomie przewodowej sieci LAN.
Drugim problemem dotyczącym transmisji w częstotliwościach radiowych szerokiego spektrum jest brak licencji komisji FCC, która jasno określałaby prawa do częstotliwości na danym terenie. W związku z tym w niektórych miejscach trzeba będzie zmniejszyć moc nadawania sygnałów, aby uniknąć kolizji między falami emitowanymi przez sieć naszą i naszych sąsiadów. Jeśli mimo to sieci będą sobie wzajemnie przeszkadzać nie będzie innego wyjścia, jak zaakceptować tę obniżoną wydajność. Dla sieci LAN typową strefą transmisji jest 200-300 metrów. Im bardziej jest ona skoncentrowana, tym mniejsze jest prawdopodobieństwo konkurowania z innymi radiowymi nośnikami tej samej częstotliwości.
Bezprzewodowe mosty umożliwiają transmisję sygnałów na odległość 5 do 8 km. Istnieje więc prawdopodobieństwo, ze będą musiały konkurować z innymi mostami jak pagery i inne podobne im urządzenia konkurują między sobą o pasmo 902-928 MHz. Możliwość wystąpienia konfliktu wzrasta wraz ze wzrostem gęstości urządzeń.

1.4.5 Jednopasmowa częstotliwość radiowa

Techniką odmienną od transmisji szerokiego spektrum są transmisje przy użyciu pojedynczego pasma. Oba sposoby dotyczą radiowej części Spektrum Elektromagnetycznego, lecz pasmo pojedyncze używa (jak nazwa wskazuje) tylko jednego kanału, którym zwykle jest kanał częstotliwości mikrofalowych. W najniższych zakresach mikrofale zachowują się podobnie jak fale radiowe, podczas gdy ich najwyższe pasma przejawiają pewne właściwości charakterystyczne dla światła.
Korzystanie z częstotliwości dedykowanych oznacza potrzebę uzyskania licencji od komisji FCC przed rozpoczęciem legalnego nadawania na tych falach. Technologia komunikowania się za pomocą częstotliwości dedykowanych zapoczątkowana była przez firmę Motorola, która reklamowała ją pod nazwą Altair oraz Altair II. Motorola uzyskała wyłączny dostęp do częstotliwości 18-19 GHz w większości miejskich okręgów USA.
Motorola działa wobec komisji FCC jako przedstawiciel każdego klienta, który zdecyduje się używać ich technologii, oszczędzając im przy okazji wiele czasu i nerwów.
Wdrażanie tego typu technologii ułatwione jest o tyle, że do korzystania z niej wystarczą przewodowo łączone szkielety LAN, sterowniki urządzeń oraz odpowiednie oprogramowanie. Przykładowa implementacja przedstawiona jest na rysunku 4.14.
Rysunek 4.14. powa
=pr=emodowa sieć AN \ o dedykowanym stale.
Moc transmisji wynosi około 25 miliwatów, czyli jest zbyt niska aby powodować jakiekolwiek problemy zdrowotne. Tak mała moc w połączeniu z dużą łamliwością mikrofal ogranicza ich efektywny zasięg do około 30 metrów dla wolnej przestrzeni i około 10 metrów w przypadku obecności przeszkody. Za przeszkodę uznano trzy przenośne biurowe ścianki działowe.
Całkowita szerokość pasma wynosi około 15 Mbps. Po odjęciu standardowych narzutów Ethernetu oraz uwzględnieniu konwersji przewód-powietrze-przewód zmniejsza się ona do poziomu 5.5 Mbps - porównywalnego z poziomem przepustowości współdzielonego przewodowego Ethernetu.

1.4.6 Podczerwień

Kolejna technologia transmisji bezprzewodowej oparta jest na podczerwieni, czyli zakresie częstotliwości znajdującym się między widoczną częścią spektrum elektromagnetycznego a najkrótszymi mikrofalami. Mimo że podczerwień jest rodzajem światła, to obejmuje ono również częstotliwości niewidoczne dla ludzkiego oka - podczerwień jest bowiem światłem niewidzialnym, które nie przenika ciał stałych, lecz odbija się od nich.
Istnieją dwa rodzaje światła podczerwonego: rozproszone i kierunkowe, z których pierwsze przypomina rodzaj promieniowania wysyłanego przez latarnię, a drugi, przez latarkę. Mniejsza koncentracja promieni rozproszonego światła podczerwonego umożliwia również mniejszą potencjalną szybkość przesyłania danych.
Większość urządzeń gospodarstwa domowego (o ile nie wszystkie), które używają promieniowania podczerwonego, wykorzystuje podczerwień kierunkową. Urządzenie



wysyłające tego rodzaju promieniowanie (np. pilot) musi być wymierzone w kierunku odbiornika (np. telewizora, magnetowidu, wieży itp.), aby polecenia (np. wprowadzane za pomocą przycisków pilota) mogły być przez niego rozpoznane. Ta sama zasada dotyczy sieci LAN wykorzystujących kierunkowe promienie podczerwone.

Sieci LAN wykorzystujące rozproszone promieniowanie podczerwone wysyłają je we wszystkich kierunkach. Promienie odbijają się od ścian, sufitów i wszystkich ciał stałych obecnych w pomieszczeniu, umożliwiając w ten sposób ustanowienie połączenia z urzą dzeniem, które niekoniecznie znajduje się w zasięgu wzroku.

Zalety

Komunikacja podczerwona korzysta z sygnału o niespecjalnej trwałości - ze światła. Podczerwień nie przenika ciał stałych o najmniejszej nawet gęstości. Korzystanie z niej nie wymaga więc uzyskiwania zezwoleń komisji FCC, która ustanowiła jednak wytyczne i ustaliła parametry funkcjonalne urządzeń fizycznych wykorzystujących Spektrum Elektromagnetyczne. Przestrzeganie tych wytycznych przez różnych producentów bezprzewodowych sieci LAN gwarantuje wzajemną zgodność ich produktów. Ewentualni użytkownicy takich sieci oszczędzają dzięki temu sporo czasu, którego nie muszą poświęcać na żmudne ustalenia szczegółowych specyfikacji technicznych i nużące papierowe formalności.

wady

Komunikacja `w linii wzroku' nakłada duże ograniczenia na wiele form transmisji opartych na falach świetlnych. Wiele pomieszczeń biurowych jest zupełnie nieprzystosowanych, jako, że nawet pojedyncza przenośna ścianka działowa powoduje całkowite wytłumienie sygnału.

Rozproszone światło podczerwone częściowo rozwiązuje problem ograniczenia transmisji do linii wzroku i wykorzystuje właściwości odbijania do przenoszenia promieni w miejsca niedostępne dla transmisji „w linii wzroku". Niestety, przy każdym odbiciu promienia sygnał jest osłabiany. W związku z tym mały jest zarówno efektywny obszar działania systemu opartego na podczerwieni - o średnicy poniżej 30 metrów, jak i obsługiwana przepustowość.

1.4.7 Laser

Transmisje oparte na promieniowaniu laserowym w sieciach LAN są rodzajem systemu światłowodu bez okablowania światłowodowego. Nie jest to idealne porównanie, ponieważ większość sieci LAN używa systemów optycznych sterowanych nie za pomocą lasera, ale dobrze oddaje zasadę działania i obraz rzeczy.

Koszt technologii laserowej wyklucza wykorzystywanie jej do przyłączania do sieci poszczególnych stacji. Technologię tę lepiej więc stosować do łączenia grup stacji z jednostką dostępu, która wysyła i odbiera sygnały laserowe w imieniu całego zespołu. W ten sposób wysokie koszty lasera zostają rozłożone na wiele stacji roboczych, zwiększając ekonomiczną efektywność instalacji.

Decydując się na taki sposób korzystania z lasera, najlepiej zamontować go pod samym sufitem, jak najdalej od ludzi, a to ze względu na dwie przyczyny: po pierwsze zmniejsza to prawdopodobieństwo przypadkowego uszkodzenia wzroku, a po drugie zmniejsza prawdopodobieństwo zakłócenia sygnału przez ludzkie ruchy.

Urządzenia laserowe mogą również być używane do łączenia sieci LAN znajdujących się na przykład na przeciwnych stronach parkingu. Jak pamiętamy z podrozdziału „Bezprzewodowe sieci LAN", rozwiązanie takie umożliwia przesyłanie większej liczby bitów na sekundę po niższym-niż przy wykorzystaniu linii dzierżawionych i routerów-koszcie.

Zalety

Światło laserowe jest bardzo mocno skoncentrowane i nie ulega rozpraszaniu. W związku z tym może być wykorzystywane do łączenia na odległości dłuższe niż światło podczerwone. Umożliwia to korzystanie z lasera jako bezprzewodowej technologii mostkowania. Oparte na świetle laserowym bezprzewodowe sieci LAN działają podobnie jak tego samego rodzaju sieci oparte na falach radiowych. Sposób ten przedstawiony został na rysunku 4.8.

Wady

Zarówno laser, jak i podczerwień są formami promieniowania świetlnego i jako takie obarczone są pewnymi niedostatkami, w szczególności cierpią one na łamliwość fal. Przesyłane niekablowo światło laserowe oraz kierunkowe promienie podczerwone rozróżniają dwie podstawowe właściwości:

• światło laserowe używa innych części spektrum niż światło podczerwone, • światło lasera jest światłem koncentrowanym sztucznie.
Z nich wynikają kolejne różnice. Po pierwsze, systemy laserowe są dużo droższe niż porównywalne systemy wykorzystujące podczerwień. Wymagają większej mocy do wygenerowania i skoncentrowania sygnału. Oddają one również więcej ciepła.
Urządzenia laserowe używają częstotliwości z zakresu światła widzialnego. Ze względu na to nadają się tylko do łączenia urządzeń znajdujących się „w linii wzroku". Ich sy
gnał może być łatwo wytłumiony przez dym, mgłę, a nawet krople deszczu na instalacji zewnętrznej.
Tłumienie jest zjawiskiem charakterystycznym dla sygnałów elektrycznych. Polega ono na zmniejszaniu się siły sygnału w miarę jego poruszania się przez nośnik. Proszę pamiętać, że sygnał jest drganiem elektrycznym. Taka sama energia elektryczna zużywana jest do napędzania sygnałów biegnących w nośniku. Sygnał jest więc źródłem własnej energii i stale zmniejsza się w miarę przebywania swej drogi. Sygnały optyczne często uważane są za „inne" od sygnałów elektrycznych. Nie są one bowiem oscylacjami prądu elektrycznego, lecz pulsami światła i ciemności . Niemniej jednak są one drganiami bardzo wysokiej częstotliwości tego samego spektrum elektromagnetycznego. W związku z tym sygnały optyczne również podlegają tłumieniu.
Jednak tłumienie optyczne jest raczej przejawem zanikania sygnału ze względu na zderzenia z nieczystościami medium transmisyjnego.
Największym problemem związanym z używaniem lasera jest możliwość uszkodzenia siatkówki oka i jego nerwów. Z tego względu, rozmieszczenie urządzeń laserowych powinno być dokładnie zaplanowane.

1.4.8 Standard IEEE 802.11

Instytut IEEE zakończył pracę nad standardem dla bezprzewodowych sieci LAN (czyli sieci WLAN - ang.. *wireless LAN*. Proces powstawania tego standardu był długotrwały. Dodatkowo komplikowała go potrzeba uwzględnienia technologii już znajdujących się na rynku.
Ostateczna wersja standardu określała metodę (MAC) sterowania dostępem do nośnika oraz wiele warstw fizycznych. Jak w każdym podejściu warstwowym, funkcje każdej z warstw oddzielone są od funkcji warstw sąsiadujących. Innymi słowy, funkcje warstwy MAC nie zależą od szybkości przesyłania danych ani od żadnej innej właściwości specyfikacji warstwy fizycznej.
Standard ten obejmuje kilka podfunkcji, w tym mechanizmy sterowania dostępem (zarówno oparte na rywalizacji, jak i na niej nie oparte) dla kilku „niezupełnie fizycznych" nośników. Każdy nośnik ma własną, oddzielną specyfikację.

1.4.8.1 Dostęp do nośnika

Metodą dostępu do nośnika specyfikacji 802.11 jest CSMA/CA (wielodostęp do łącza sieci z badaniem stanu kanału i unikaniem kolizji). Ta metoda dostępu używana jest również w Ethernecie II, czyli Ethernecie DIX.
CSMA/CA dba o to, by nie wchodzić w konflikt, „słuchając" najpierw, zanim zacznie „mówić". W przypadku sieci WLAN, stacja, która chce nadawać, musi najpierw posłuchać kanału. Sygnał pusty oznacza, że urządzenie może zacząć transmisję. Jeśli natomiast sygnał zawiera ramkę z danymi, urządzenie musi czekać dalej na odpowiedni moment.
Schemat CSMA/CA próbuje zabezpieczyć odpowiedni dostęp do pasma, jednocześnie starając się unikać konfliktów. Aby umożliwić odpowiedni stopień dostępu do pasma przesyłania, wykorzystuje on dwie różne techniki taktowania:
• minimalne opóźnienie propagacji,
• losowy odstęp antykolizyjny.
Minimalne opóźnienie propagacji używane jest do zabezpieczenia, że żadna ze stacji nie zajmie całego dostępnego pasma częstotliwości. Po przesłaniu każdej ramki, a przed wysłaniem kolejnej, urządzenie nadające odczekać musi co najmniej przedział czasu określony jako minimalny.
Metoda losowego odstępu antykolizyjnego wykorzystywana jest w dwóch sytuacjach. Pierwsza polega na odczekiwaniu przez stację nadającą odpowiedniego przedziału czasu po pomyślnym wysłaniu ramki, a przed rozpoczęciem nasłuchu kanału. Jeśli następnie w wyniku nasłuchu okaże się, że kanał jest wolny, przykładowa stacja może wysłać kolejną ramkę. Jeśli jednak kanał jest zajęty, musi ona odczekać kolejny odstęp czasu przed ponownym rozpoczęciem nasłuchiwania i jeszcze raz, dopóki nie okaże się, że kanał jest wolny i że można nim przesłać wysyłaną ramkę. Może się wydawać, że taki sposób dostępu faworyzuje stację nadającą. Nie jest tak ze względu na to, że czas trwania odstępu antykolizyjnego ustalany jest losowo. Taki sposób dostępu wykorzystywany jest przez wszystkie stacje przyłączone do takiej sieci.

1.4.8.2 Warstwy fizyczne

Specyfikacja 802.11 I obsługuje 3 warstwy fizyczne przy użyciu tego sposobu dostępu do nośnika, a są to:
• podczerwień rozproszona,
• spektrum szerokie o bezpośredniej sekwencji częstotliwości,
• spektrum szerokie o niebezpośredniej sekwencji częstotliwości.

1.4.8.2.1Spektrum szerokie o bezpośredniej sekwencji częstotliwości

Specyfikacja 802.11 obsługuje przesyłanie danych w sekwencji bezpośredniej szerokiego spektrum z przepustowością wynoszącą od 1 do 2 Mbps. Różne szybkości przesyłania danych korzystają z różnych technik modulowania sygnałów. Standardowo wykorzystuje się zakres 2.4-2.4835 GHz, w ramach którego wyznaczyć można 4 podzakresy:
• 2.400 - 2.425 GHz • 2.414 - 2.440 GHz • 2.429 - 2.455 GHz • 2.443 - 2.470 GHz
Każdy podzakres ma szerokość 26 MHz i zachodzi się z zakresem sąsiadującym na szerokość 11 MHz.

1.4.8.2.2 Spektrum ***szerokie o niebezpośredniej sekwencji częstotliwości***

Zgodne ze standardem bezprzewodowe sieci lokalne wykorzystujące technikę szerokiego spektrum o niebezpośredniej sekwencji częstotliwości nie umożliwiają przesyłanie danych z szybkością od 1 do 2 Mbps, w zależności od wybranej techniki modulacji. Wzory skoków mają podobną strukturę i nie wpływają na szybkość, z jaką przesyłanie danych może być obsługiwane.
Obie techniki stosują ten sam schemat obsługi jednego zakresu. A polega on na tym, że dostępna szerokość pasma dzielona jest na 79 podzakres o szerokości 1 MHz każdy. Jeden podzakres obsługuje przynajmniej 2,5 hopa na sekundę (hps) przy wykorzystaniu
dowolnego wzorca sekwencji. Minimalna szybkość przesyłania danych (wyrażana w hopach na sekundę) ma na celu zabezpieczenie przed podziałem danych niezgodnym ze stosowanym schematem i przesłaniem ich przy użyciu innego podzakresu.

1.4.9 Podsumowanie

Jeśli specyfikację IEEE 802.11 uda się pomyślnie wprowadzić, producenci szybko zmodyfikują swoje produkty tak, aby były z nią zgodne. Przyczyni się to do wystąpienia dwóch zjawisk:
• bezprzewodowe sieci LAN staną się współoperacyjne (będą mogły współdziałać na wielu platformach),
• koszty zakładania bezprzewodowych sieci LAN zmniejszą się ze względu na zjawisko ekonomii skali, które jest następstwem zmiany technologii zindywidualizowanych na standardowe.
Bezprzewodowe sieci LAN były, jak dotąd, zbiorem nie współdziałających produktów i technologii. Nie zdobyły więc szerokiej akceptacji, a ich występowanie ograniczone było do nisz rynkowych. Standaryzacja wspólnego sposobu dostępu do nośnika oraz warstwy fizycznej przez instytut IEEE obiecuje poprawę tego stanu rzeczy w przyszłości. Kluczowym wskaźnikiem stopnia współoperacyjności jest cena takiej sieci za jedną stację roboczą, która będzie się zmniejszać w miarę zwiększania się możliwości współdziałania tego rodzaju systemów. W takiej sytuacji bezprzewodowe sieci LAN mają szansę stania się instalacjami typowymi.
Trudno się spodziewać, ażeby bezprzewodowe sieci LAN stały się technologią zamienną w stosunku do przewodowych sieci LAN. W czasie, w którym wydajność sieci bezprzewodowych osiągnęła poziom sprawności typowych sieci przewodowych, ten ostatni stał się już nieaktualny, a dokładnie rzecz biorąc, zwiększył się kilkukrotnie. I jak na razie nie wygląda na to, by różnica między wydajnościami sieci tych typów miała ulec zmniejszeniu.
Jednakże, sieci bezprzewodowe charakteryzują się pewnymi właściwościami, których sieci przewodowe dostarczyć nie potrafią. Odpowiednio wykorzystywane, mogą być bardzo efektywne.

1.5 Rozdział 5 Warstwa łącza danych

Mark A. Sportack
Drugą warstwą modelu referencyjnego OSI, jak pamiętamy z rozdziału 1 pt.. „ABC Sieci", jest warstwa łącza danych. Warstwa ta jest interfejsem między sprzętem a oprogramowaniem.
W niniejszym rozdziale poznamy dokładnie naturę, strukturę i funkcje ramek danych w sieci LAN i przyjrzymy się kilku najważniejszym architekturom sieci LAN - będą to: Ethernet, Token Ring oraz FDDI. Zapoznamy się z różnicami w sposobach uzyskiwania dostępu do nośnika fizycznego. Rozdział ten zamkniemy przeglądem technologii warstwy łącza danych oraz zagadnień związanych z jej zainstalowaniem.

1.5.1 Warstwa 2 modelu OSI

Drugą z określanych przez model OSI warstw jest warstwa łącza danych. Pełni ona zadania związane z wysyłaniem i odbiorem danych.
Ze względu na wysyłanie danych, jest ona odpowiedzialna za spakowanie instrukcji, danych itp., do postaci ramek. Ramka jest strukturą zawierającą taką ilość informacji, która wystarcza do przesłania danych za pomocą sieci (LAN lub WAN) do ich miejsca przeznaczenia. Pomyślna transmisja danych następuje wtedy, gdy dane osiągną miejsce docelowe w stanie niezmienionym (bez uszkodzeń). W związku z tym ramki zawierają mechanizmy sprawdzania integralności własnej zawartości.
Węzeł nadawania, czyli nadajnik, musi otrzymać potwierdzenie, że dane dotarły do miejsca przeznaczenia w stanie nieuszkodzonym. Ramki, które weszły w konflikt lub zostały uszkodzone podczas przesyłania danych, muszą być transmitowane ponownie, dopóki nie dotrą one w całości do miejsca przeznaczenia. Wszelkie ramki, których otrzymanie nie zostało potwierdzone przez ich wskazanego odbiorcę, również muszą być przesyłane ponownie.
Warstwa łącza danych jest odpowiedzialna również za ponowne przetwarzanie wszelkich - przybywających do niej z warstwy fizycznej - strumieni binarnych z powrotem
do postaci ramek. Przed przesłaniem tak utworzonej ramki do warstwy następnej wykonuje ona matematyczny sprawdzian zgodności nowo utworzonych ramek z danymi oryginalnymi. Tego rodzaju sprawdzian nazywany jest Cykliczną Kontrolą Nadmiarową CRC (ang.. *Cyclic Redundancy Check).*
Wykrywanie błędów przy użyciu sum kontrolnych jest podobne do cyklicznej kontroli nadmiarowej, choć mniej od niej skomplikowane. Sumy kontrolne tworzone są przez dodanie binarnej wartości każdego znaku alfanumerycznego wchodzącego w skład bloku danych. Wartość ta przesyłana jest do miejsca przeznaczenia za pomocą odrębnego pola. Odbiorca ponownie przelicza sumy kontrolne i porównuje ich wartość z wartością zapisaną w polu sumy kontrolnej. Jeśli porównywane liczby są różne, oznacza to wystąpienie błędu. Wystąpienie kilku błędów spowodować może uzyskanie tych samych sum kontrolnych - błędy takie nie są przy użyciu tej techniki wykrywane.
Cykliczna Kontrola Nadmiarowa (czyli CRC) jest bardziej niezawodnym sposobem wykrywania błędów. Wartość CRC uzyskiwana jest bowiem w nieco inny - niż suma kontrolna - sposób. Dane odczytywane są jako ciąg binarny; uzyskana w ten sposób liczba jest dzielona przez wcześniej określoną wartość, dając w ten sposób wartość CRC. Zmniejsza to prawdopodobieństwo uzyskania tej samej wartości CRC w wyniku wystąpienia kilku błędów transmisji.

Ramki, które nie przejdą pomyślnie kontroli CRC, są usuwane, a urządzenie wskazane w polu „Adres nadawcy" tej ramki proszone jest o ponowienie transmisji. Dzięki wykorzystaniu tego mechanizmu rozpoznawania błędów, protokoły warstw 2 i 3 nie muszą już zajmować się sprawdzaniem, czy ramka została dostarczona i czy została dostarczona w całości. Protokoły warstwy 4, takie jak TCP oraz SPX polegają na własnych mechanizmach rozpoznawania i korekcji błędów, niezależnie od wyników kontroli uzyskanych przez warstwę 2.

1.5.2 Ramki

Ramka jest strukturą wykorzystywaną do przesyłania bloków danych w sieci. Rozmiar i struktura ramki zależy od rodzaju używanego przez sieć protokołu warstwy sprzętowej (np. protokołu Ethernet, Token Ring itp.). Ramka jest podobna do koperty listowej. Wiadomo, że każdy rodzaj koperty ma określone i stałe rozmiary. Zawartość kopert tego samego rodzaju może jednak różnić się rozmiarem, wagą, zawartością, pilnością itp.
Rozmiar koperty nie mówi nic o sposobie dostarczenia jej do miejsca przeznaczenia. W środowiskach sieciowych procesy służące przesyłaniu ramek nazywane są protokołami. Protokoły istnieją również dla warstwy 3 modelu referencyjnego OSI. Protokoły warstwy 3 modelu OSI umieszczają więc protokoły w pakietach i umożliwiają przesyłanie ich poza obręb sieci lokalnych. Protokoły te opisane są w rozdziale 12 pt. „Protokoły sieciowe".

1.5.2.1 Składniki typowej ramki

Typowa ramka składa się jedynie z tylu pól - czyli podstruktur - ile niezbędnych jest do zagwarantowania bezpiecznego dostarczenia ramki wskazanemu odbiorcy. Najczęściej spotykanymi polami są:
• Ogranicznik początku ramki, • Adres źródłowy (nadawcy), • Adres docelowy (adresata),
• Dane,
• Sekwencja kontrolna ramki.

1.5.2.1.1 Definicja ramki

Gdy już ustalimy rozmiar kopert - kontynuując kopertową analogię - możemy zacząć wykorzystywać infrastrukturę służącą do masowego ich przetwarzania. Standaryzacja rozmiaru kopert jest więc niezbędnym etapem budowy infrastruktury umożliwiającej ich przesyłanie - niezależnie od tego, kto jest ich producentem.
Struktury warstwy 2 0 określonym, stałym rozmiarze nazywane są komórkami. Przykładem komórkowego protokołu warstwy 2 jest protokół ATM. Używa on 53-oktetowych komórek: 5 oktetów zawiera całą informację nagłówka, a pozostałe 48 oktetów przenosi informacje. Komórki przenoszą zawsze 48 oktetów danych, podczas gdy ramki mogą przenosić różną ich ilość.
Protokoły warstwy 2 sieci LAN nie określają jednego rozmiaru ramek, lecz ich rozmiary maksymalne i minimalne. Ramki są więc strukturami o zmiennej długości. Pozwala to protokołom na maksymalizację efektywności każdej transmisji przez zoptymalizowanie stosunku ilości narzutu informacyjnego do ilości przesyłanej informacji. Stosunek ten jest miarą skuteczności transmisyjnej protokołu.
Początek każdej ramki jest identyfikowany przez ustalony wzór bitów. Wzór ten jest ogranicznikiem początku ramki. Koniec ramki identyfikowany jest za pomocą ogranicznika końca ramki lub sekwencji kontroli ramki.

1.5.2.1.2 Adres źródłowy i docelowy

Podobnie jak koperty opatrywane są adresami nadawcy i adresata, tak również ramka posiada swój adres źródłowy i docelowy. Adresem źródłowym jest kod adresu maszynowego nadawcy. Adresem odbiorcy jest adres maszynowy adresata.
Umieszczanie tej informacji na początku ramki oszczędza urządzeniom będącym potencjalnymi adresatami problemu związanego z otwieraniem każdej z nich, sprawdzaniem jej zawartości i określaniem, czy faktycznie jest ono jej adresatem. Proces taki ang.ażowałby zarówno wiele zasobów, jak i czasu, czego efektem byłoby obniżenie wydajności działającej w ten sposób sieci.
Informacje adresowe umieszczone są jak najbliżej początku ramki. Przyspiesza to znacznie proces przekierowywania ramek: sprawdzanie pierwszych 18 oktetów ramki trwa dużo krócej niż sprawdzanie pierwszych 1500 oktetów. Informacje znajdujące się w omawianych polach używane są w celu kierowania ramki do wskazanego adresata oraz do informowania nadawcy, że ramki nie da się dostarczyć do jej miejsca docelowego.
Zadaniem różnych elementów nagłówka ramki jest dostarczanie podstawowej informacji potrzebnej do zidentyfikowania odbiorcy oraz nadawcy, a także do określenia, czy ramka powinna być przesłana ponownie. Jedynym polem wspólnym dla wszystkich ramek jest pole Dane. Pole to ma różną długość; wszystkie pozostałe pola ramki mają natomiast ściśle określoną długość - określoną przez specyfikacje protokołu. Jest ono przyczyną, dla której ramka w ogóle istnieje. Jeśli znajdujące się w tym polu dane nie zostaną dostarczone lub ulegną uszkodzeniu przed osiągnięciem miejsca docelowego, wyrzucona zostaje cała ramka.

1.5.2.1.3 Ramki - podsumowanie ramowe

Sieci LAN są w zasadzie mechanizmami przesyłania ramek. Aby móc swoje zadanie wykonywać efektywnie, sieci LAN wymagają, aby ramki miały określony kształt i strukturę. Standaryzacja zapewnia, że różne składniki sieci wyprodukowane przez różnych producentów mogą ze sobą współdziałać.
Te same standardy tworzą również wspólną podstawę, dzięki której możliwa jest konwersja ramek sieci różnych typów, jak na przykład sieci Ethernet i Token Ring.

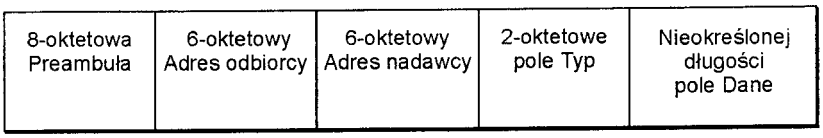
1.5.2.2 Ewolucja struktur ramek firmowych

Pierwszą na świecie siecią lokalną był Ethernet PARC firmy Xerox. Technologia ta pojawiła się jako wewnątrzbiurowa technologia transmisji pasmem podstawowym służąca do łączenia stacji roboczych. Została utworzona przez pracowników sławnego centrum badawczego w Palo Alto (ang.. PARC - Palo Alto Research Center) do użytku własnego jako alternatywa dla współdzielenia danych przy użyciu dyskietek. Był to więc prosty mechanizm, którego prostotę odzwierciedlała również struktura ramek. Wszystkie rozwiązania sieciowe dostępne obecnie na rynku są rozwinięciami tej technologii.

1.5.2.2.1 Ramka sieci PARC Ethernet firmy Xerox

Naukowcy z PARC postanowili, że protokół warstwy drugiej będzie przenosić protokoły warstw wyższych, takie jak protokół IP, protokół Xerox XNS i inne. Protokoły przenoszone same ograniczałyby ilość danych przenoszonych jednorazowo. Więc zamiast poświęcać czas na tworzenie protokołu warstwy sprzętowej, wystarczało jedynie udostępnienie 2-oktetowego pola Typ, które wskazywało rodzaj protokołu wyższej warstwy znajdującego się wewnątrz ramki (inaczej mówiąc, będącego jej klientem). Pozwalało to bardziej wyrafinowanym protokołom na określanie rozmiarów ramek.

Oktet składa się z ośmiu cyfr bitowych (bitów) niezależnie od rodzaju zapisanej przy ich użyciu informacji. Ale strzeżcie się, powiadam, ekspertów, którzy na określenie tych struktur używają słów „wyrażenie bitowe" lub „bajt", albowiem spędzają oni zbyt wiele czasu wśród informatyków. Tylko patrzeć, jak zaczną używać słów „rekordy" lub „pliki" na określenie ilości przesyłanych danych.
„Zmajstrowana" domowym sposobem sieć Ethernet firmy Xerox była strukturą nieskomplikowaną polegającą na protokołach klienckich w celu określenia długości pola Dane. Jak przedstawia to rysunek 5.1, ramka Ethernet PARC składała się z:
• 8-oktetowej Preambuły,
• 6-oktetowego Adresu fizycznego (adresu MAC) odbiorcy, • 6-oktetowego Adresu fizycznego (adresu MAC) nadawcy,
• 2-oktetowego pola Typ, które identyfikuje protokół kliencki osadzony w polu Dane • pola Dane o nieokreślonej i zmiennej długości.
Rysunek 5.1. Ramka sieci PARC Ethernet.



Protokół ten rozprowadza pakiety do wszystkich urządzeń przyłączonych do sieci LAN. Urządzenia te muszą więc konkurować o dostępne pakiety. W celu ułatwienia tej rywalizacji stosowana jest technika CSMA, czyli wielodostępu do łącza sieci z badaniem stanu kanału. Każde urządzenie, które chce przesłać dane, musi najpierw nasłuchiwać, czy sieć jest dostępna, czy też należy poczekać, gdyż zajęta jest przesyłaniem ramek nadawanych przez inne urządzenie. Jeśli urządzenie wykryje sygnał nośny, może rozpocząć transmisję. Obniża to znacznie liczbę kolizji ramek wysyłanych przez różne urządzenia, choć nie eliminuje ich zupełnie. Usuwanie skutków kolizji i innych zdarzeń powodujących niedostarczenie ramek znajduje się w gestii urządzeń końcowych i nie jest wykonywane przez protokół sieciowy.

1.5.2.2.2 Ramka sieci DIX Ethernet

Potencjał handlowy sieci Ethernet Xerox PARC został szybko dostrzeżony, a jej ramka i protokół zostały poprawione w celu dostosowania sieci do szerszego rynku. Powstała w ten sposób druga generacja sieci lokalnych znana jako Ethernet II lub DIX Ethernet, gdzie DIX jest skrótem od nazw firm sponsorujących jej powstanie: Digital, Intel oraz Xerox.
Xerox, właściciel technologii oraz strażnik swoich firmowych „standardów", przypisał 2-oktetowy kod typu do określania protokołów klienckich, których przykładami są XNS
firmy Xerox, IPX firmy Novell, a także IP oraz DECNet. Protokoły te są protokołami warstw wyższych niż protokoły (pierwotnej) sieci Ethernet i mają one inne wymagania co do rozmiarów przesyłanych wiadomości. Sieć Ethernet nie mogła pozwolić sobie na nie kontrolowanie rozmiarów ramek, jeśli miała przy tym obsługiwać bardziej skomplikowane metody dostępu zdolne do wykrywania kolizji. W związku z tym ustanowione zostały ograniczenia rozmiarów ramek.
Ramka sieci DIX Ethernet (Ethernet II) składała się z:
• 8-oktetowej Preambuły,
• 6-oktetowego Adresu odbiorcy,
• 6-oktetowego Adresu nadawcy,
• 2-oktetowego pola Typ, które identyfikowało opakowany przez ramkę protokół transportu warstw wyższych
• pola Dane o rozmiarze co najmniej 50 oktetów, lecz nie większym niż 1486 oktetów.
Warto zapamiętać, że mimo określenia minimalnej długości ramki, standard DIX nadal polegał na polu Typ wykorzystywanym przez pole Typ sieci PARC Ethernet. Pole to wciąż służyło do identyfikacji protokołu, który z kolei był wykorzystywany do określania długości ramki. Klienckie (czyli te przenoszone) protokoły transportu (protokoły warstwy 3) nadal miały swoje własne wymagania co do rozmiarów pakietów, ale sieć DIX Ethernet korzystała z dużo bardziej skomplikowanej metody dostępu do nośnika niż jej poprzedniczka - z metody wielodostępu do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji, czyli znana CSMA/CD, która nakładała dość precyzyjne ograniczenia czasowe.
Rysunek 5.2. Ramka DIX Ethernet.

8-oktetowa Preambuła	6-oktetowy Adres odbiorcy	6-oktetowy Adres nadawcy	2-oktetowe pole Typ	Pole Dane <1486 oktetów	Wypełnienie

Ta metoda dostępu wymaga od stacji sprawdzania, przed wysłaniem danych, czy kablami nie są już wysyłane jakieś sygnały. Jeśli sieć wygląda na pustą, stacja może rozpocząć nadawanie. Niestety, transmitowane w przewodzie miedzianym sygnały potrzebują czasu na dotarcie do miejsca docelowego. Zatem zdarza się, że stacja rozpocznie wysyłanie swoich impulsów w sieci, która wygląda na niewykorzystywaną, po to tylko, aby kilka mikrosekund później zderzyły się z nimi sygnały wysłane przez inną stację. Zderzenie takie nazywane jest kolizją.

Dodanie funkcji wykrywania kolizji (czyli CD) do metody wielodostępu do łącza sieci z badaniem stanu kanału (czyli CSMA) umożliwiło nowym sieciom LAN usuwanie skutków kolizji, dzięki czemu nie muszą one w tym zakresie polegać wyłącznie na urządzeniach końcowych. Dzięki metodzie CSMA/CD stacje mogą wykryć kolizję, wstrzymać nadawanie i - po upływie odpowiednio długiej przerwy - rozpocząć je na nowo. Czas trwania tej przerwy jest określany za pomocą algorytmu binarnego wykładniczego algorytmu postępowania w przypadku kolizji (czyli algorytmu BEB - ang.. Binary Exponential Back-off- algorithm).

Wykrywanie kolizji w sieci Ethernet DIX polega na kontrolowaniu najdłuższego czasu potrzebnego do przesłania sygnału w sieci LAN tam i z powrotem. Dla sieci Ethernet 10 MHz czas ten wynosi nie więcej niż 50 mikrosekund. Oznacza to, że stacja musi nadawać przez czas wystarczający do wysłania sygnału na drugi koniec i z powrotem, czyli przez dłużej niż 50 mikrosekund. Wystarcza to na przesłanie 500 bitów. A, że w oktecie znajduje się 8 bitów, oznacza to, że minimalny rozmiar pakietów umożliwiający działanie wykrywania kolizji wynosi 62,5 oktetów. Firma Xerox zaokrągliła minimalny rozmiar ramki dla sieci Ethernet II do 64 oktetów.

Każda ramka, której ładunek (a pamiętajmy, że jego rozmiar określany jest w tego typu sieciach przez protokoły transportu warstw wyższych) po dodaniu wszystkich narzutów da nam ramkę mającą mniej niż 64 oktety, zostanie wypełniona zerami do osiągnięcia przez nią wielkości minimalnej. Rozwiązuje to problem minimalnego czasu przy wykrywaniu kolizji, ale zmusza każdy protokół do rozpoznawania, które dane są informacją, a które wypełnieniem. Ramka sieci DIX Ethernet nadal polega na zawartości pola Typ w celu identyfikowania protokołów warstw wyższych, a tym samym długości pola Dane.

Mimo że do Ethernetu DIX dodano funkcje mające umożliwić poszerzenie jego rynku zbytu, to jedyna większa zmiana polegała na ustanowieniu minimalnych i maksymalnych ograniczeń rozmiarów ramki. Xerox, twórczyni Ethernetu, zachowała więc prawa do technologii i w związku z tym sama ustanawiała i publikowała standardy. Taki sposób standaryzacji spełnił swój cel - Ethernet stał się produktem handlowym. Niestety, taki sposób ustanawiania i utrzymywania standardów nie wytrzymuje próby czasu. Przedsiębiorstwo działające w środowisku konkurencyjnym nie jest najlepszą organizacją do tego, by utrzymywać standardy dotyczące towarów handlowych. Będzie ona pod ciągłą presją działania we własnym imieniu. Ze względu na to, aby sieć Ethernet mogła stać się prawdziwie wziętą technologią handlową, odpowiedzialność za standaryzowanie jej musiała być scedowana na bezstronną ekonomicznie jednostkę organizacyjną.

1.5.3 Projekt IEEE 802

Takim ciałem ustanawiającym standardy, odpowiedzialnym za wiele standardów obsługujących dzisiejsze wysoko wydajne sieci, jest instytut IEEE. Prace standaryzacyjne dotyczące sieci rozpoczęto w lutym 1980 roku (stąd mnemoniczny numer 802 na określenie standardu), kiedy to instytut IEEE utworzył Komitet Standardów Sieci Lokalnych i Miejskich, nazywany również Projektem 802.

Celem ustanawiania standardów jest stworzenie zasad umożliwiających wszystkim typom sieci LAN swobodne przesyłanie danych między sobą oraz odseparowanie nośników fizycznych od protokołów sieci LAN. Umożliwia to na używanie tego samego rodzaju kabli bez jakiejkolwiek różnicy dla współoperacyjności.

Komitet nakreślił zadania wymagane do realizacji całego celu i do ich realizacji powołał grupy zadaniowe. Wynikami działań tych grup są następujące standardy:

Sieci MAN to inaczej sieci miejskie (ang.. Metropolitan Area Network). Standardy sieci MAN są również określone w ramach Projektu 802 IEEE.

• 802.1 określa przegląd i architekturę niezbędną dla współdziałania między sieciami LAN i WAN. Stanowi podstawę wszystkich pozostałych inicjatyw Projektu 802 i określa standardy sterowania sieciami oraz LAN/WAN mostkowania sieci zgodnych z Projektem 802.

• 802.2 określa standard warstwy łącza danych (warstwy 2) dla telekomunikacji i wymiany informacji między systemami LAN i WAN. Specyfikacja ta dostarcza również kompatybilność wstecz niezbędną do łączenia niestandardowych wersji Ethernetu z jego wersją 802.3.

• 802.3 ustanawia nowy standard sieci LAN umożliwiający wielodostęp do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji (czyli CSMA/CD). Właściwą nazwą tego nowego rodzaju sieci LAN jest CSMA/CD, lecz wszyscy i tak mówią o niej Ethernet.

• 802.4 określa standard warstwy fizycznej dla topologii magistrali sieci LAN o metodzie dostępu do sieci na zasadzie przesyłania tokenu. Taki rodzaj sieci LAN nazywany jest Token Bus. Obsługuje on przesyłanie danych z szybkością l , 2, 5, i 10 Mbps.

• 802.5 ustanawia standardy dla metody dostępu Token Ring oraz fizycznych technik sygnalizowania.

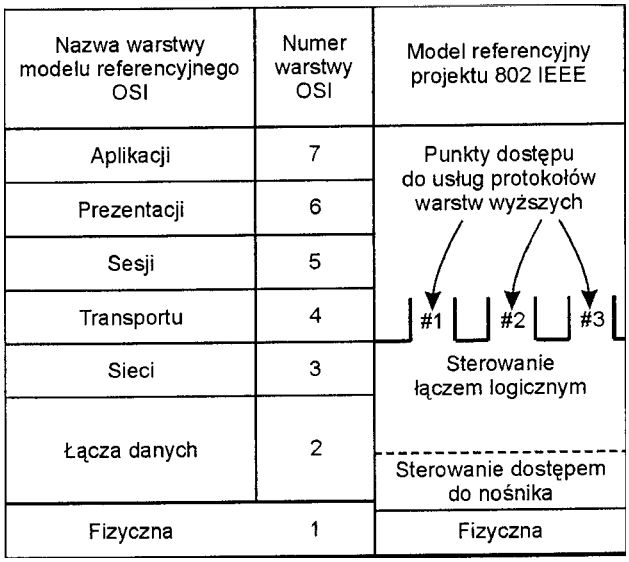
Mimo że nie jest to pełna lista standardów Projektu 802, wymienione 5 pozycji przedstawia kierunek, w jakim standaryzowane są sieci LAN i WAN. Każdy z tych standardów może współdziałać z pozostałymi, wykorzystując nieskomplikowane metody konwersji ramek, ze względu na ich wspólną podstawę. Standardy te wyszły daleko poza zakres standardów Xeroksa dezaktualizując sieci typu DIX Ethernet.

1.5.3.1 **Sterowanie łączem logicznym w standardzie** IEEE 802.2

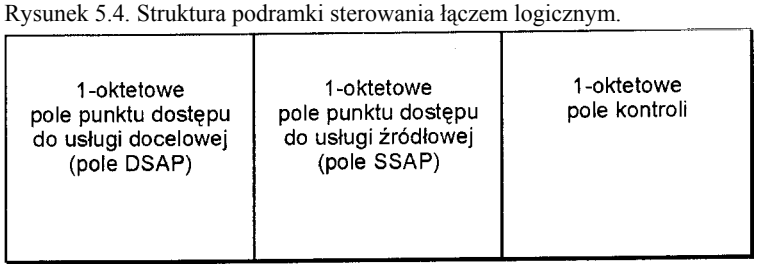
Projekt 802 IEEE zorganizował swoje standardy wokół trójpoziomowej hierarchii protokołów, które odpowiadają dwóm najniższym warstwom: fizycznej oraz łącza danych. Owymi trzema poziomami są: warstwa firyczna, warstwa sterowania dostępem do nośnika (warstwa MAC) oraz warstwa sterowania łączem logicznym (warstwa LCC). Stan taki jest właściwy dla wszystkich sieci LAN zgodnych z projektem

802. Specyfikacja adresowania warstwy MAC pozwala na stosowanie adresów 2- lub 6-oktetowych, przy czym standardem są adresy 6-oktetowe. Adres 2-oktetowy używany jest prawie wyłącznie na wykładach.
Model referencyjny IEEE 802 różni się od Modelu referencyjnego OSI pod dwoma zasadniczymi względami. Po pierwsze, warstwa fizyczna modelu 802 jest podzbiorem swojego odpowiednika z modelu OSI (czyli obejmuje mniejszą część sieci). A po drugie, warstwa łącza danych modelu IEEE podzielona jest na dwa odrębne poziomy: sterowania dostępem do nośnika (MAC) oraz sterowania łączem logicznym (LLC). Omawiana korelacja przedstawiona jest na rysunku 5.3.

Rysunek 5.3. Korelacja między modelem OSI a modelem IEEE 802.



Jednym z mechanizmów umożliwiających współdziałanie sieciom LAN o różnych topologiach jest podramka LCC. Jest to trójpolowa struktura dołączona do pola Dane ramki modelu 802. Używana jest do identyfikowania protokołu docelowego w urządzeniach wieloprotokołowych. Pole to jest niezbędne ze względu na ograniczone możliwości protokołu 802.3 do identyfikowania protokołów zewnętrznych. Podramka LCC składa się z następujących elementów:
• 1-oktetowego pola Punktu dostępu do usługi docelowej (pola DSAP), • 1-oktetowego pola Punktu dostępu do usługi źródłowej (pola SSAP),
• 1-oktetowego pola Kontroli.
Podramka zilustrowana jest na rysunku 5.4.
Rysunek 5.4. Struktura podramki sterowania łączem logicznym.



Punkty dostępu do usługi wskazują, dla którego z protokołów warstw wyższych pakiet jest przeznaczony. Protokołom przypisywane są wartości szesnastkowe, które umieszczane są w polach DSAP oraz SSAP pakietu.
Sterowanie łączem logicznym umożliwia adresowanie i kontrolowanie łącza danych. Określa ono, które mechanizmy są używane do adresowania przesyłanych danych, a które
do kontrolowania danych wymienianych między urządzeniem nadawczym i odbiorczym. Odbywa się to za pomocą trzech usług sterowania LLC:
• nie potwierdzana usługa bezpołączeniowa, • potwierdzana usługa bezpołączeniowa,
• usługa połączeniowa.
Usługi te udostępniane są przy użyciu punktów dostępu do usługi, które znajdują się między warstwami sieci oraz łącza danych.
Pierwsza usługa LLC, nie potwierdzana usługa bezpołączeniowa, jest propozycją skromną, ale użyteczną. Często protokołom warstwy 4, takim jak TCP, SPX itp., przypisywane są zadania sterowania przepływem danych i inne funkcje związane z niezawodnością. W związku z tym nie ma sensu duplikować tych samych zadań, tyle że na poziomie warstwy 2. Omawiana usługa zmniejsza więc ilość narzutów wykorzystywanych do przesyłania danych. Podczas korzystania z niektórych aplikacji, zwłaszcza tych, dla których czas ma decydujące znaczenie, takich jak konferencje audio lub wideo, zauważyć możemy znaczne obniżenie wydajności.
Następną usługą LLC jest potwierdzana usługa bezpołączeniowa, która dostarcza potwierdzenie otrzymania danych, bez żadnych narzutów związanych z zarządzaniem połączeniem. Mniejsze narzuty oznaczają szybsze dostarczanie danych. Dodanie gwarancji dostarczenia pakietów stwarza bardzo przydatną usługę, której zastosowania są praktycznie nieograniczone.
Ostatnia usługa LLC, usługa zorientowana połączeniowo, udostępnia mechanizmy warstwy łącza danych służące ustanawianiu i utrzymywaniu połączeń. Przydaje się ona zwłaszcza urządzeniom nieinteligentnym, które nie mają protokołów warstwy 4 czy 3, i które w

związku z tym nie mogą dostarczać funkcji przez te protokoły wykonywanych na poziomie 2. Omawiana funkcja kontroli wymaga, aby warstwa łącza danych utrzymywała tablicę śledzącą aktywne połączenia.
Wybór usługi zależy od rodzaju aplikacji i zwykle nie jest widoczny dla jej użytkownika. Jako że omawiane usługi są częścią podstawowej specyfikacji 802.2, dostępne są we wszystkich sieciach zgodnych z projektem 802 i mogą być wykorzystywane podczas współdziałania różnych sieci zgodnych z serią 802.

1.5.3.2 Protokół dostępu do podsieci (protokół SNAP) standardu IEEE 802.2

Jednym ze sposobów zapewniania zgodności wstecz we wczesnych, niestandardowych wersjach sieci LAN, takich jak PARC Ethernet czy DIX Ethernet, była struktura podramki, która dostarczała mechanizmu służącego do identyfikowania przenoszonych protokołów warstw wyższych. Zarówno ramki Ethernetu PARC, jak i Ethernetu DIX zawierały pole Typ. Miejsce to zostało na nowo wykorzystane przez IEEE w standardzie 802.3 sieci Ethernet.
Sieć tego typu jest dużo bardziej skomplikowana niż jej poprzedniczki, w związku z czym znika większość przyczyn, dla których to pole zostało wyodrębnione. Jednak niektóre protokoły warstwy 4 polegają na tym polu w celu wykonywania owych funkcji na poziomie warstwy 2. Aby więc ustanowić zgodność z istniejącymi protokołami, należało opracować strukturę podramki, która potrafiłaby identyfikować protokoły warstw wyższych. W efekcie powstała podramka protokołu dostępu do podsieci, czyli protokołu SNAP (ang.. Sub-Network Access Protocol, która przedstawiona jest na rysunku 5.5.
Rysunek 5.5. Struktura podramki S.\ AP 802.2.

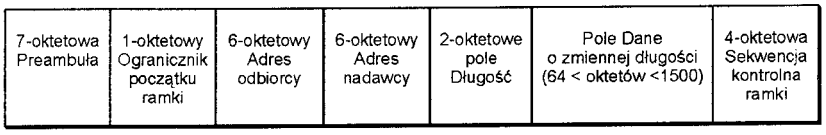
1-oktetowe pole punktu dostępu do usługi docelowej (pole DSAP)	1-oktetowe pole punktu dostępu do usługi źródłowej (pole SSAP)	1-oktetowe pole kontroli	3-oktetowe pole Identyfikatora organizacyjnie unikatowego	2-oktetowe pole Identyfikatora protokołu

Specyfikacja 802.2 określa strukturę również ramki SNAP. Ramka do standardowej
ramki 802.2 dodaje 5-oktetowe pole zawierające 3-oktetowe pole identyfikatora organizacyjnie unikatowego oraz 2-oktetowe pole Typ protokołu. Podramka SNAP jest rozszerzeniem podramki LLC i musi być używana wspólnie z nią. Może być używana w każdej sieci zgodnej z 802.

1.5.3.3 Ramka sieci Ethernet standardu IEEE 802.3

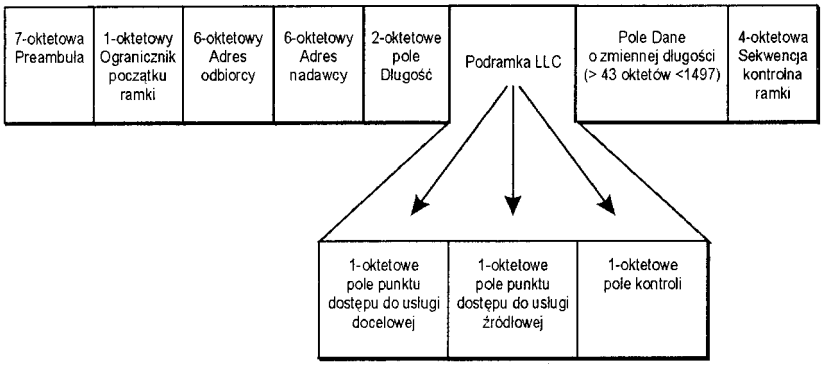
Projekt 802 określił standard, na którym opierają się wszystkie rodzaje ramek ethernetowych. Ramki te mogą mieć co najmniej 64 i co najwyżej 1500 oktetów (liczone w całości, włącznie z nagłówkami i samym ładunkiem danych). Nagłówki używane są do rozpoznawania nadawcy i odbiorcy każdego z tych pakietów. Jedynym ograniczeniem tego rodzaju identyfikacji jest to, że każdy adres musi być unikatowy i mieć 6 oktetów długości.
Pierwsze 12 oktetów każdej ramki zawiera 6-oktetowy Adres docelowy (adresata) i 6oktetowy adres (nadawcy). Są to maszynowe (fizyczne) kody adresowe poziomu sprzętowego znane też jako adresy MAC. Adresami tymi często są unikatowe „adresy uniwersalnie administrowane", które każdej karcie sieciowej nadawane są automatycznie podczas jej produkcji. Ten typ adresów MAC przedstawiany jest za pomocą sześciu par cyfr szesnastkowych rozdzielanych dwukropkami. Pierwsze dwie liczby są identyfikatorem producenta, o przyznanie którego, jak i zakresu adresów MAC, każdy producent kart sieciowych musi zwrócić się do IEEE.
Adresy te kartom mogą być też nadawane podczas ich instalowania. Tak przypisane adresy nazywane są „adresami administrowanymi lokalnie". Adresy te mogą być użyte do identyfikacji numeru pokoju, działu itp. Korzystanie z adresów administrowanych lokalnie dostarcza administratorom sieci ważnych informacji ułatwiających rozwiązywanie problemów. Ich obsługa bywa czasochłonna i nieraz bardzo uciążliwa.
Ramki zgodne z projektem 802 mogą zawierać adres określonego urządzenia lub odnosić się do grup stacji roboczych przy użyciu wspólnych identyfikatorów. Przesyłanie danych do grup urządzeń nazywane jest „rozgłaszaniem" lub „multicastingiem".
W normalnych warunkach, karty sieciowe Ethernetu otrzymują tylko te ramki, których adresy docelowe odpowiadają ich unikatowym adresom MAC lub spełnią kryteria rozgłaszania. Większość kart sieciowych można ustawić na działanie w „trybie podsłuchu". Otrzymywać one wtedy będą wszystkie ramki przesyłane w sieci, niezależnie od ich adresu docelowego. Stwarza to możliwość wystąpienia zagrożeń bezpieczeństwa wszystkich pozostałych użytkowników magistrali nadawania takiej sieci, a dla użytkownika korzystającego z tak ustawionej karty - możliwość wystąpienia problemów z wydajnością. Choć większość zmian wprowadzonych w standardzie 802.3 dotyczy samego protokołu, to jedna ważna zmiana dotyczyła ramki 802.3. Komitet 802 uznał bowiem, że ramka nie powinna polegać na poprawnym działaniu innych protokołów, w związku z czym wymienił 2-oktetowe pole Typ na 2-oktetowe pole Długość.
Po ustanowieniu minimalnej i maksymalnej długości pola, na podstawie przedziału czasu niezbędnego do transmisji tam i z powrotem (co, jak podaję w punkcie „Ramka sieci DIX Ethernet" wcześniej w tym rozdziale, jest niezbędne dla celów wykrywania kolizji), rozmiar ramki przestał być zależny od protokołów klienckich. Grupa robocza 802.3 przedefiniowała 2-oktetowe pole Typ tak, aby bezpośrednio wskazywało ono długość ramki, i przesunęła identyfikację protokołów do warstwy LCC.
W ramce podstawowej sieci Ethernet 802.3 pole Typ zostało więc wymienione na pole Długość. W celu identyfikacji typu protokołu, jeśli jest to niezbędne, stosowana jest podramka 802.2. Kolejną właściwością ramki 802.3 nieobecną w poprzednich jej wersjach jest wymóg, aby jej całkowity (od początku pola adresu docelowego do końca pola sekwencji kontrolnej ramki) rozmiar mieścił się w granicach od (minimum) 64 do (maksimum) 1500 oktetów.
Preambuła jest 7-oktetowym ciągiem znaków poprzedzającym każdą ramkę służącym do synchronizacji transmisji. Za nim znajduje się Ogranicznik początku ramki, czyli ogranicznik SFD (ang.. *Start of Frame Delimiter),* którego nazwa w zasadzie wyjaśnia jego zadanie:

wskazuje on początek ramki wszystkim znajdującym się w sieci LAN urządzeniom. Ogranicznikiem początku ramki jest bitowa cyfra „11", po której następuje powtarzalna sekwencja „10", czyli „1010101010 itd.".
Czasem Ogranicznik początku ramki uważany jest za integralną część Preambuły, a nie część ramki, co zwiększa taką Preambułę do rozmiaru 8 oktetów. Stanowi to kolejną subtelną różnicę między sieciami Ethernet PARC czy DIX a standardem 802.3. Te pierwsze powtarzają bowiem sekwencję „10101010" przez całe 8 oktetów Preambuły. Wzór taki używany był zarówno do synchronizacji, jak i do określania początku ramki. Następnym mechanizmem jest Sekwencja kontrolna ramki (ang.. FCS - *Frame Check Sequence).* W tym polu komputer wysyłający ramkę zapisuje matematycznie wyprowadzaną wartość. Komputer docelowy również wie, w jaki sposób wartość taką obliczyć. Porównanie tych wartości umożliwia sprawdzenie integralności ramki. Ramki podczas przesyłania mogą zostać uszkodzone przez wiele różnych czynników. Zakłócenia elektromagnetyczne, przesłuch itp. mogą uszkodzić zawartość ramek bez zatrzymywania samej transmisji.
Po otrzymaniu ramki komputer docelowy sprawdza pole Sekwencji kontrolnej ramki w procesie cyklicznej kontroli nadmiarowej (ang.. CRC - *Cyclical Redundancy Check).* Komputer-adresat przeprowadza te same obliczenia co komputer-nadawca i porównuje otrzymaną wartość do tej, która zapisana jest w polu sekwencji kontrolnej. Jeśli wartości są te same, komputer docelowy wie, że dane dotarły doń w stanie niezmienionym. Jeśli nie, komputer docelowy żąda ponownej transmisji ramek.
Ramka podstawowa sieci Ethernet przedstawiona jest na rysunku 5.6. Ramki tego rodzaju używane są rzadko - jeśli w ogóle. Częściej stanowią one podstawę ramek rozszerzonych, wykorzystujących podramki 802.2 LLC i/lub podramki SNAP.
Rysunek 5.6. Ramka podstawowa Ethernet IEEE
802. 3.



1.5.3.3.1Struktura ramki LCC Ethernet

Ethernetowa ramka LLC jest połączeniem ramki 802.3 i podramki 802.2 LCC. W tej wersji do ramki podstawowej Ethernetu dołączone są 3 pola: pole punktu dostępu do usługi docelowej (pole DSAP), pole punktu dostępu do usługi źródłowej (pole SSAP) i pole kontroli.
Ramka LLC Ethernet ma następującą strukturę:
• 7-oktetową Preambułę sygnalizującą początek ramki,
• 1-oktetowy Ogranicznik początku ramki,
• 6-oktetowy Adres odbiorcy,
• 6-oktetowy Adres nadawcy,
• 2-oktetowe pole Długość określające całkowitą (włącznie z nagłówkami LLC oraz SNAP) długość pola Dane,
• 1-oktetowe pole punktu dostępu do usługi docelowej (pole DSAP) identyfikujące przewidywalny punkt dostępu do usługi LLC urządzenia docelowego,
• 1-oktetowe pole punktu dostępu do usługi źródłowej (pole SSAP) identyfikujące punkt dostępu do usługi LLC urządzenia źródłowego,
• 1- lub 2-oktetowe pole kontroli, wskazujące typ przenoszonej ramki LLC,
• pole Dane składające się z oktetów od 42 do 1496 lub od 43 do 1497, w zależności od długości poprzedzającego je pola Kontroli,
• 4-oktetowa Sekwencja kontrolna ramki używana do sprawdzania integralności ramki.
Struktura ramki Ethernet LLC przedstawiona jest na rysunku 5.7.
Rysunek 5.7. Ramka Ethernet IEEE 802. 3
podramką 802.2 LLC.

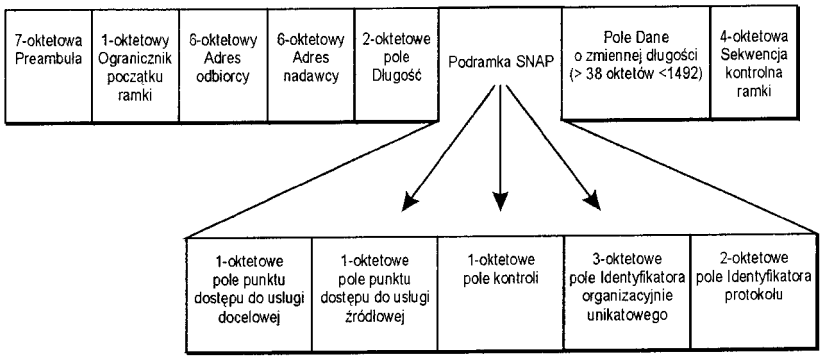


Ramka Ethernet LLC integruje struktury podramki 802.2, czyli nagłówki, i pozwala na identyfikację protokołu wyższego rzędu, który jest docelowym odbiorcą zawartości ramki. Zachowanie mechanizmu identyfikowania protokołów warstw wyższych umożliwia kompatybilność wstecz z wcześniejszymi wersjami Ethernetu, którego ramki zawierają osobne mechanizmy identyfikowania protokołów.
Całkowita długość ramki Ethernet LLC nie może być mniejsza niż 64 oktety (nie licząc Preambuły i Ogranicznika początku ramki), gdyż inaczej mechanizm CSMA/CD wielodostępu do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji nie działa poprawnie. Jeśli więc

pole Dane nie ma wystarczającej długości, do jego końca dołączanych jest tyle zer, ile potrzeba do osiągnięcia przez ramkę rozmiarów minimalnych. Maksymalnym rozmiarem ramki jest 1518 oktetów liczone włącznie z polami Preambuły i Ogranicznika początku ramki.

1.5.3.3.2 Struktura ramek Ethernet SNAP

Ramka Ethernet SNAP jest połączeniem ramki 802.3 i podramki 802.2 protokołu dostępu do podsieci (czyli podramki podsieci SNAP - ang.. Sub-Network Access Protocol. W tej implementacji SNAP dodaje 5-oktetowe pole identyfikacji protokołu. Pole to dołączane jest do ramki za nagłówkiem LLC. Składa się z 3-oktetowego identyfikatora organizacyjnie unikatowego i 2-oktetowego pola Typ. Pola te wskazują, dla których z protokołów warstw wyższych ramka jest przeznaczona.
Ramka SNAP składa się z następujących pól:
• 7-oktetowej Preambuły wskazującej początek ramki,
• 1-oktetowego Ogranicznika początku ramki, • 6-oktetowego Adresu odbiorcy,
• 6-oktetowego Adresu nadawcy,
• 2-oktetowego pola Długość określającego całkowitą (włącznie z nagłówkami LLC oraz SNAP) długość pola Dane,
• 1-oktetowego pola Punktu dostępu do usługi docelowej (pola DSAP) identyfikującego przewidywalny punkt dostępu do usługi LLC urządzenia docelowego,
• I-oktetowego pola Punktu dostępu do usługi źródłowej (pola SSAP) identyfikującego punkt dostępu do usługi LLC urządzenia źródłowego,
• 1- lub 2-oktetowego pola Kontroli, wskazującego typ przenoszonej ramki LLC, • 5-oktetowej podramki SNAP, w skład której wchodzi 3-oktetowe pole Identyfikatora organizacyjnie unikatowego oraz 2-oktetowe pole Typu protokołu identyfikujące przesyłany protokół wyższego poziomu,
• pola Dane składającego się z oktetów od 38 do 1492 lub od 37 do 1491, w zależności od długości poprzedzającego je pola Kontroli,
• 4-oktetowej Sekwencji kontrolnej ramki używanej do sprawdzania integralności ramki.
Podramka SNAP Ethernet przedstawiona jest na rysunku 5.8.
Rysunek 5.8. Ramka Ethernet !IEEE 802. 3
- z podramką SNAP 802.2 _ ł P 802. 2.
Podramka SNAP Ethernet została zaprojektowana w celu zwiększenia możliwości podramek LLC w zakresie udostępniania kompatybilności wstecz, z poprzednimi wersjami Ethernetu.



Całkowita długość ramki SNAP Ethernet musi mieć więcej niż 64 oktety. W przeciwnym razie mechanizm CSMA/CD wielodostępu do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji nie będzie funkcjonował poprawnie. Górną granicą rozmiaru ramki Ethernet SNAP jest 1518 oktetów liczonych z uwzględnieniem Preambuły i Ogranicznika początku ramki.

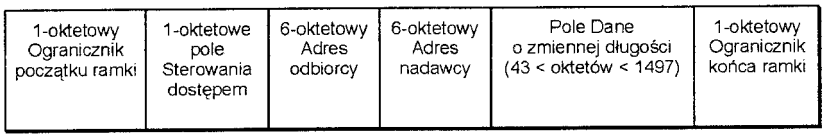
1.5.3.4 Sieci Token Ring standardu IEEE 802.5

Instytut IEEE znormalizował również format oraz protokół wiadomości, dzięki czemu uzyskana została bardziej przewidywalna wersja sieci Token Ring, znana jako standard 802.5. Sieci Token Ring istniały od połowy lat 70. głównie jako w technologia własna IBM-u. Specyfikacja Token Ring 802.5 jest więc prawie identyczna z siecią Token
Ring firmy IBM. Dlatego też nazwa „Token Ring" używana jest na określenie zarówno produktów firmy IBM opracowanych przed wprowadzeniem standardów, jak i tych zgodnych ze standardem IEEE 802.5.
We wdrażaniu standardów aktywny udział brały duże firmy, takie jak IBM (Token Ring 802.5) i General Motors (Token Bus 802.4). Technologia Token Ring oferuje szybszy i pewniejszy dostęp do sieci niż protokół 802.3, jednak za cenę wyższego kosztu sieci liczonego na jedną stację. Organizacje wymagające szybkiego dostępu do informacji uznały Token Ring za jedyne odpowiednie rozwiązanie. Protokół 802.3 zapewnia, co prawda, że dostarczony zostanie każdy pakiet, ale w tym celu może wymagać kilku prób jego przesłania. W związku z tym nie może zapewnić dostarczenia każdej ramki na czas. Topologia Token Ring natomiast, ze względu na jej uwarunkowania i pierścieniowy kształt, umożliwia uporządkowany sposób komunikacji.

1.5.3.4.1 Struktura ramki IEEE 802.5

Struktura ramki 802.5 Token Ring składa się z dwóch części: tokenu i ramki danych. Jak pamiętamy, ramka danych składa się z trzech 1-oktetowych pól.
Ramki tokenów oraz danych przedstawione są na rysunku 5.9. Rysunek ten przedstawia „surową" strukturę Token Ring, bez podramek LLC i SNAP. Podobnie do ramki specyfikacji 802.3, również Token Ring rzadko używany jest w swej podstawowej formie.

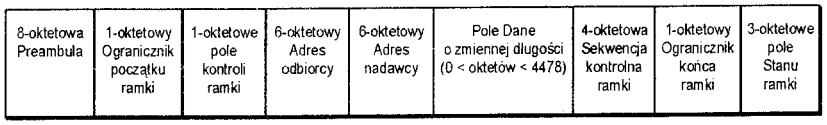
Rysunek 5.9. Ramka danych Token Ring sieci IEEE. 80Z. 5.



Ramki tokenów i ramki danych maja trzy takie same 1-oktetowe pola: Ogranicznika początku, Sterowania dostępem oraz Ogranicznika końca. Pole Sterowania dostępem jest kluczowe dla działania Token Ringu. Zawiera ono osiem bitów, z których jeden musi zostać odwrócony w celu dezaktywacji tokenu i zamiany go na sekwencję Początku ramki.
Po zamianie tokenu do postaci ramki danych, składa się on (w zasadzie już „ona") z dziewięciu różnych pól i podpól. Pole pierwsze to Ogranicznik początku ramki, który wskazuje, gdzie rozpoczyna się ramka. Następnym polem jest pole Sterowania dostępem. Pole to informuje urządzenia zgodne ze specyfikacją 802.5 o tym, czy mogą one przesyłać dane, czy też nie. Zawiera ono również bity dla systemów rezerwacji oraz określania priorytetów Token Ring. Polem kolejnym jest pole kontroli ramki. W nim umieszczone są bity identyfikujące protokół transportu. To pole używane jest do rozróżniania między ramkami danych a ramkami kontroli.
Następne dwa pola to adresy fizyczne nadawcy oraz odbiorcy. Każdy z nich składa się z 6-oktetowego pola. Adresy te są zgodne z wcześniej omówioną specyfikacją Projektu 802 i są identyczne z tymi, które stosowane są w sieciach Ethernet. Długość pola Dane dla sieci Token Ring jest zmienna i wynosi od 0 do 4099. Polem ostatnim jest wskazujący koniec ramki 1-oktetowy Ogranicznik końca ramki.
Co prawda, nie jest to przedstawione na rysunku, ale Token Ring, tak samo jak wcześniej omówiona ramka Ethernet korzysta z podramek LLC oraz SNAP. Pozwala to sieciom Token Ring na łączenie się - za pomocą mostków translacyjnych - z sieciami Ethernet bez obniżania swej zdolności przesyłania otrzymanych ramek otrzymanych do protokołów warstw wyższych. Podramki LLC oraz SNAP przydają się zwłaszcza w środowiskach łączących sieci Token Ring oraz Ethernet.

1.5.4 Architektura FDDI

Oficjalna nazwa tej standardowej architektury sieci LAN opracowanej przez amerykański instytut ANSI brzmi „Fiber Distributed Data Interface" (co po polsku znaczy „Złącze Danych Przenoszonych Światłowodem"). Nazwa jest dosyć długa, więc architektura ta lepiej znana jest jako FDDI. najczęściej wymawiane - dla dalszego uproszczenia - jako „fiddy". Mimo że przez wielu uważana jedynie za szybszą wersję sieci Token Ring, różni się od niej fundamentalnie - w zakresie topologii, zarządzania, a nawet struktur tokenów i ramek. Najprostsza wersja ramki przenoszącej dane w sieci FDDI przedstawiona jest na rysunku 5.10. Charakteryzuje się ona następującą strukturą:
• 8-oktetową Preambułą wskazująca początek ramki
• 1-óktetowym polem Ogranicznika początku ramki, który wskazuje początek zawartości ramki,
• 1-oktetowym polem Kontroli ramki, które określa typ ramki, czyli token, adres fizyczny (MAC) lub logiczny (LLC), ramkę priorytetu, itp.
• 6-oktetowym Adresem fizycznym (MAC) odbiorcy, • 6-oktetowym Adresem fizycznym (MAC) nadawcy,
• polem Dane o zmiennej długości - nie przekraczającej jednak 4478 oktetów,
• 4-oktetową Sekwencją kontrolną ramki stosowaną do sprawdzania integralności ramki,
• półoktetowej długości (czterobitowym) Ogranicznikiem końca ramki,
• 3-oktetowym polem Stanu ramki zawierającym 3 jednooktetowe podpola: Błąd (ang.. Error), Zgodność adresu (ang.. Adress-match) oraz Skopiowane (ang.. Copied).
Rysynek 5.10. Ramka FDDI

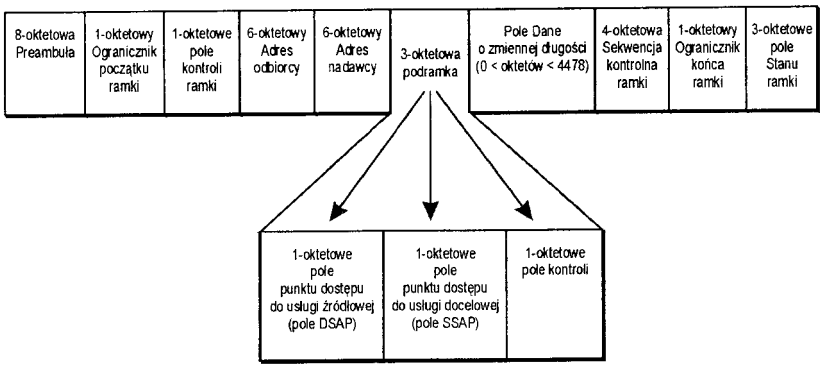


Maksymalna długość ramki FDDI nie może wynosić więcej niż 4500 oktetów, włącznie ze wszystkimi danymi i składnikami ramki. Zwykle z ramki FDDI korzysta się w połączeniu z jednym z dwóch podformatów: LLC lub SNAP. Ramka o tak utworzonym formacie również nie może mieć więcej niż 4500 oktetów (nie licząc Preambuły).

1.5.4.1 Struktura ramki FDDI LLC

Architektura FDDI, bazując na warstwie LLC (czyli warstwie sterowania łączem logicznym) standardu IEEE 802.2 umożliwia obsługę logicznego sterowania łączem.
Standardy z grupy 802 zostały przyjęte przez instytut ANSI. Dzięki temu - mimo że standard FDDI nie został opracowany przez instytut IEEE może on korzystać z mechanizmów podramek przez IEEE opracowanych. Zwiększa to możliwości sieci FDDI w zakresie funkcjonowania w roli szerokopasmowego szkieletu łączącego sieci LAN.
Przedstawioną na rysunku 5.11 ramkę LLC cechuje następująca struktura:
• 8-oktetowa Preambuła wskazująca początek ramki
• 1-oktetowe pole Ogranicznika początku ramki, który wskazuje początek zawartości ramki,
• I-oktetowe pole kontroli ramki, które określa t<-p ramki, czyli token, adres fizyczny (MAC) lub logiczny (LLC), ramkę priorytetu itp.
• 6-oktetowy Adres fizyczny (MAC) odbiorcy, • 6-oktetowy Adres fizyczny (MAC) nadawcy,
• pole Dane o zmiennej długości - nie przekraczającej jednak 4478 oktetów,
• 4-oktetowa Sekwencja kontrolna ramki stosowana do sprawdzania integralności ramki,

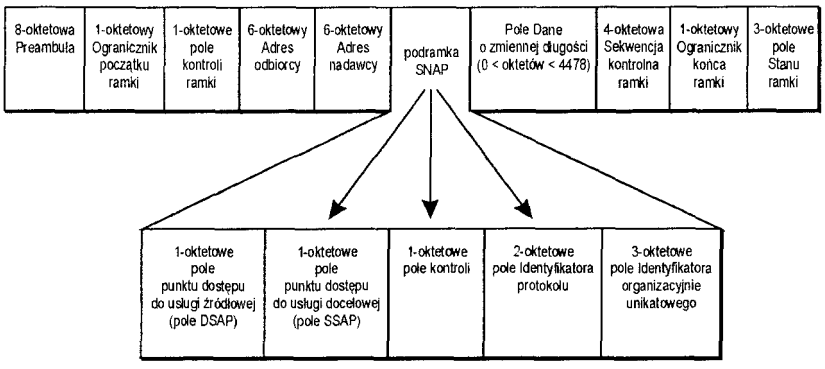
• półoktetowej długości (czterobitowy) Ogranicznik końca ramki,
• 3-oktetowe pole Stanu ramki zawierające 3 jednooktetowe podpola: Błąd (ang.. Error), Zgodność adresu (ang.. Adress-match) oraz Skopiowane (ang.. Copied); każde z tych pól ma wartość S (od ang.. Set - ustawione) lub R (od ang.. Reset wyzerowane).
Rysunek 5.11. Ramka FDDI
i podramką 802.2 LLC.



Ramka FDDI może zawierać struktury podramki 802.2 LLC, na które składają się pola DSAP, SSAP oraz Pole kontroli.

1.5.4.2 Struktura ramki FDDI SNAP

Architektura FDDI - na podstawie warstwy LLC sterowania łączem logicznym standardu 802.2 - umożliwia również obsługę SNAP. Ramka SNAP dodaje do struktur podramki 3-oktetowe pole Identyfikacji protokołu oraz 2-oktetowe pole Typu. Ramka taka przedstawiona jest na rysunku 5.12.
Ramka FDDI SNAP składa się z następujących pól:
• 8-oktetowej Preambuły, która wskazuje początek ramki
• 1-oktetowego pola Ogranicznika początku ramki, który wskazuje początek zawartości ramki,
• 1-oktetowego pola Kontroli ramki, które określa typ ramki, czyli token, adres fizyczny (MAC) lub logiczny (LLC), ramkę priorytetu itp.
• 6-oktetowego Adresu fizycznego (MAC) odbiorcy, • 6-oktetowego Adresu fizycznego (MAC) nadawcy,
• 5-oktetowej podramki SNAP, składającej się z 3-oktetowego pola Identyfikatora organizacyjnie unikatowego i 2-oktetowego pola Typu protokołu określających przenoszony protokół wyższego poziomu
• pola Dane o zmiennej długości - nie przekraczającej jednak 4473 oktetów,
• 4-oktetowej Sekwencji kontrolnej ramki stosowanej do sprawdzania integralności ramki,
• półoktetowej długości (czterobitowego) Ogranicznika końca ramki,
• 3-oktetowego pola Stanu ramki zawierającego 3 jednooktetowe podpola: Błąd (ang.. Error), Zgodność adresu (ang.. Adress-match) oraz Skopiowane (ang.. Copied); każde z tych pól ma wartość S (od ang.. Set - ustawione) lub R (od ang.. Reset-wyzerowane).
Rysunek 5.12. .?::Ramka FDDI
- z podramką ?.S-_-ł P 802. 2.
Ramka FDDI SNAP dodaje do ramki FDDI LLC 3-oktetowy Identyfikator organizacyjnie unikatowy oraz 2-oktetowe pole Typu bezpośrednio za nagłówkiem LLC, a przed polem Danych.



1.5.5 Zasady sterowania dostępem do nośnika

Choć przesyłanie jedynek i zer za pomocą nośnika jest zadaniem pierwszej warstwy fizycznej modelu referencyjnego OSI, to kontrolowanie dostępu do nośnika jest funkcją warstwy Łącza danych. Regulacja ta znana jest jako MAC, czyli „Media Access Control" (Sterowanie dostępem do nośnika).
Specyfikacje IEEE wyróżniają 3 różne sposoby uzyskiwania dostępu do nośnika. Są to:
• na zasadzie rywalizacji,

• na zasadzie priorytetu żądań,
• na zasadzie przesyłania tokenu.
W sieciach o komunikacji współdzielonej, takich jak na przykład Ethernet CSMA/CD (czyli sieć Ethernet umożliwiająca wielodostęp do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji), wszystkie przyłączone urządzenia współdzielą jedną warstwę MAC. Warstwa ta stanowi domenę nadawania warstwy 2.

1.5.5.1 Dostęp do nośnika na zasadzie rywalizacji

Dostęp do nośnika na zasadzie rywalizacji jest częścią opisanego wcześniej schematu CSMA/CD (wielodostępu do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji) standardu 802.3. Jak sama nazwa tej metody wskazuje, wymaga ona od stacji sprawdzania, czy medium transmisyjne jest wolne, czy też inne urządzenie przesyła już za jego pośrednictwem jakieś dane. Stacja może przesyłać dane za pomocą nośnika, jeśli w trakcie badania jego stanu wykrywa ona sygnał nośny (sygnał wolny).
Jedynie sieci stosujące metodę transmisji szerokopasmowej mają sygnał nośny. W systemie opartym na paśmie podstawowym, takim jak na przykład każda z sieci 10Base, zamiast sygnału nośnego przesyłana jest seria zmian napięcia.
Wykrycie sygnału nośnego w trakcie badania stanu kanału nie zawsze zapewnia bezkolizyjną transmisję danych. Jak bowiem pamiętamy z punktu „Ramka sieci DIX Ethernet", przesyłanie danych odbywa się nie w punkcie, lecz w przedziale czasu. Nawet jeśli transmisja zachodzi przy prędkości równej prędkości przewodzenia, przesłanie ramki 802.3 w sieci LAN może trwać do 50 mikrosekund. W związku z tym stacja może, rozpoznawszy sygnał wolny, rozpocząć nadawanie sygnału, w który kilka mikrosekund później uderzy inny, wcześniej wysłany sygnał. Zajście kolizji wymaga ponownego przesłania obu biorących w niej udział sygnałów, a prawdopodobieństwo jej zaistnienia rośnie znacznie wraz ze wzrostem natężenia ruchu w sieci.
Wersja 802.3 CSMA/CD potrafi wykrywać kolizje, co ujmuje warstwom wyższym zadań związanych z identyfikowaniem kolizji i ponownym przesyłaniem danych. Obie stacje (których sygnały uległy kolizji) rozpoznają zajście kolizji na poziomie warstwy 2, wstrzymują przesyłanie danych na krótką chwilę, po upływie której transmisja jest kontynuowana.

1.5.5.2 Dostęp do nośnika na zasadzie priorytetu żądań

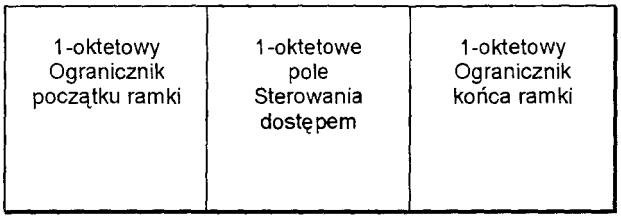
Standard IEEE 802.12 określa sieć 100 Mbps, która używa metody dostępu do nośnika na zasadzie priorytetu żądań, korzysta z ramki formatu Token Ring lub Ethernet (lecz nigdy obu) oraz topologii gwiazdy. Ze względu na elastyczną obsługę ramek, sieć LAN 802.12 określona została mianem „VG-AnyLAN". VG-AnyLAN jako nośnik wykorzystywać może cztery pary nieekranowanej miedziowej skrętki dwużyłowej (UTP) kategorii 3, ekranowanej skrętki dwużyłowej kategorii 5, jak również okablowania światłowodowego. Sieć tego typu obsługuje do trzech warstw kaskadowo łączonych wzmacniaków 0 odległości między wzmacniakiem i stacją roboczą nie większej niż 100 metrów. Średnica sieci może mieć do 1300 metrów średnicy.
Specyfikacja zapewnia zgodność z SNMP przy użyciu zmiennych MIB, czyli zmiennych bazy informacji zarządzania (która to baza znajduje się w każdym węźle i zawiera informacje o jego zasobach). Zmienne te podobne są do tych, które używane są w sieciach Ethernet oraz Token Ring. Muszą być, bo na zbiór zmiennych MIB 802.12 składają się zbiory informacji MIB zarówno sieci Ethernet, jak i Token Ring.
Metoda dostępu do nośnika oparta jest na priorytecie żądań - co oznacza, że wykorzystuje cykliczną metodę arbitrażu, czyli przyznawania dostępu, w której wzmacniak centralny regularnie sprawdza stan przyłączonych do niego portów. Sprawdzanie to odbywa się według kolejności portów, a ma na celu określenie portów, które zgłaszają żądania transmisji. Po sprawdzeniu, które z portów żądają dostępu do nośnika, wzmacniak ustala, czy żądania mają normalny, czy wysoki priorytet. Przypisywanie zgłoszeniom priorytetu wysokiego ma na celu umożliwienie natychmiastowej obsługi urządzeń, które tego wymagają. Porty nie przesyłające danych lub mające transmisje zawieszone wysyłają sygnał wolny.
Sygnał wolny jest przez wzmacniak usuwany, gdy stacja wysyłająca go została wyznaczona jako następna w kolejności przesyłania danych. Po ustaniu sygnału wolnego stacja zaczyna wysyłać dane.
Wtedy wzmacniak informuje inne stacje, że mogą one otrzymać wiadomość przychodzącą. Następnie wzmacniak odczytuje adres odbiorcy otrzymanej ramki, sprawdza go w swojej tabeli konfiguracyjnej i przesyła ramkę dalej, tak jak uczyniłby to każdy port pracujący w trybie podsłuchu (czyli, jak pamiętamy, pobierający wszystkie, a nie tylko do siebie zaadresowane ramki).
Wzmacniak centralny, lub inaczej główny, kontroluje działanie domeny priorytetów. Może składać się z nie więcej niż trzech warstw wzmacniaków połączonych ze sobą kaskadowo. Kaskadowo połączone wzmacniaki mogą bowiem działać jak jeden wielki wzmacniak. Wtedy wzmacniak centralny przesyła cały ruch do poszczególnych wzmacniaków niższej warstwy, a te z kolei sprawdzają swoje porty aktywne na obecność żądań transmisji pakietów.
Żadna stacja nie może wykonywać dwóch transmisji z rzędu, jeśli zawieszone żądania transmisji innych stacji mają taki sam priorytet. We wzmacniaku centralnym żądanie o wysokim priorytecie nie przerwie rozpoczętej już realizacji żądania. We wzmacniaku niższej warstwy żądanie o priorytecie normalnym zostaje zawłaszczone w celu
umożliwienia jak najszybszej realizacji żądania o priorytecie wyższym. Aby zapewnić, że żadne z żądań nie będzie wiecznie ignorowane, żądania o priorytecie normalnym, które oczekują dłużej niż 250 ms (2,5s), automatycznie uzyskują priorytet wysoki.

1.5.5.3 Dostęp do nośnika na zasadzie pierścienia

Specyfikacje projektu 802 zawierają 3 różne protokoły oparte na technologii tokenu, z których każdy ma swoją własną metodę dostępu. Protokołami tymi są 802.4 Token Bus, 802.5 Token Ring oraz FDDI. Pierwszy z nich, protokół 802.4 Token Bus, odznaczał się wąskim zakresem zastosowań: nadawał się do obsługi linii produkcyjnych. Dawno już odszedł w zapomnienie. W związku z tym w dalszej części rozdziału przedstawione są jedynie metody FDDI oraz Token Ring dostępu do nośnika.

1.5.5.3.1 Dostęp do nośnika w sieci Token Ring 802.5

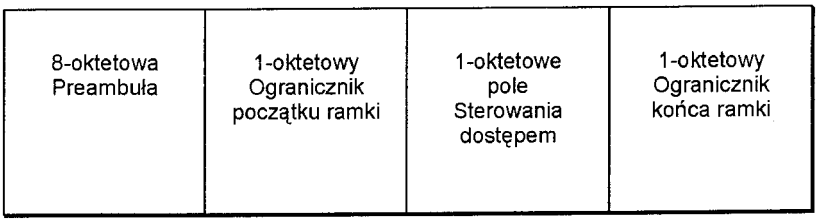
W sieciach zgodnych ze specyfikacją Token Ring specjalna ramka, znana również jako „token", przesyłana jest do kolejnych urządzeń wchodzących w skład pierścienia. Token może być przesyłany wyłącznie wtedy, gdy pierścień jest wolny. Ramka ta, przedstawiona na rysunku 5.13, składa się z trzech jednooktetowych pól oraz specjalnej sekwencji bitów. Urządzenie, które otrzymało token, ale nie ma potrzeby przesyłania danych, może token zatrzymać na 10 ms lub na dłużej, jeśli wartość domyślna została zmieniona. Jeśli czas upłynął, a urządzenie nie musiało nic przesyłać, oddaje ono kontrolę nad tokenem, który przekazywany jest do następnego urządzenia w sieci.
Rysunek 5.13. Ramka tokenu sieci Token Ring
lEEE 802.5.



Token Ring 802.5 używa tego tokenu do sterowania dostępem do nośnika. Token składa się z Ogranicznika początku, pola Sterowania dostępem oraz Ogranicznika końca. Pole Sterowania dostępem zawiera osiem bitów, z których jeden musi zostać odwrócony w celu dezaktywacji tokenu i zamiany go tym samym na sekwencję Początku ramki. Przesyłane dane oraz inne ważne pola dołączane są do tak zmodyfikowanej ramki tokenu i przesyłane w sieci. Dopóki ramka z dołączonymi do niej danymi przesyłana jest w sieci, dopóty token nie jest przekazywany następnemu uczestnikowi sieci Token Ring.
Łatwo można policzyć maksymalny czas przez jaki urządzenie działające w takiej sieci może oczekiwać na możliwość wysłania danych. Czas ten można zwiększać bądź zmniejszać, odpowiednio dodając bądź odejmując węzły wchodzące w skład sieci. Sieci Token Ring są więc idealne do zastosowań wymagających dokładnego określenia długości czasu opóźnienia (czasu propagacji).

1.5.5.3.2 Dostęp do nośnika w sieci FDDI

Sieci FDDI w celu regulacji dostępu do nośnika korzystają nie - jak sieci Token Ring z metody priorytetu/rezerwacji, lecz używają do tego celu protokołu synchronizowanego. Efektem tego jest bardziej przewidywalna sieć LAN, która eliminuje potrzebę umieszczania pola Kontroli w ramce.
W sieciach FDDI dostęp do nośnika kontrolowany jest za pomocą tokenu przedstawionego na rysunku 5.14, który (jak widać) składa się z pól Preambuły, Ogranicznika początku ramki, Kontroli ramki oraz Ogranicznika końca ramki.
Rysunek 5.14. Ramka tokenu FDDI.



Sieci FDDI, w odróżnieniu od sieci Token Ring, mogą korzystać jednocześnie z wielu tokenów i/lub ramek danych. Każda stacja, która otrzymuje token, zmienia zawartość pola Kontroli ramki, a następnie dodaje do niej swoje dane i wysyła token do sieci, już jako ramkę danych. Jednocześnie tworzy następny token i wysyła go w pierścień zaraz za ramką danych. Token może być przechwycony przez każdą następną (w pierścieniu) stację, która zgłasza żądanie przesyłania danych.
Ramka danych nie jest wyciągana z sieci przez jej adresata. Odbiorcy po prostu kopiują te ramki, które są dla nich przeznaczone w miarę jak przesyłane są one w pierścieniu. W końcu dane wysłane docierają do nadawcy, który następnie je niszczy.
Obsługa synchronicznej transmisji wielu ramek danych pozwala sieciom FDDI na uzyskiwanie efektywności większej od efektywności sieci Token Ring, zwłaszcza w sieciach dużych lub często używanych.

1.5.6 Wybór technologii LAN

Każdy ze standardów sieci LAN grupy 802 określa własną strukturę szkieletu, metodę dostępu oraz fizyczne nośniki transmisji. W celu wybrania odpowiedniej technologii LAN należy poznać i zrozumieć korzyści oraz ograniczenia, jakie niesie ze sobą każdy typ ramki.
Zebrane są one w niniejszym podrozdziale.

1.5.6.1 Sieć Ethernet 802.8

Sieci Ethernet, Fast Ethernet i Gigabit Ethernet umożliwiają przesyłanie danych z prędkościami odpowiednio 10 Mbps, 100 Mbps i 1Gbps. Specyfikacje tej grupy sieci wykorzystują różne warstwy fizyczne w połączeniu z nieomal jednakową dla wszystkich specyfikacji Warstwą dostępu do nośnika. Drobne różnice między mechanizmami dostępu poszczególnych specyfikacji wynikają z różnic między ich warstwami fizycznymi.

Dostęp do nośnika oparty jest tu na zasadzie chaotycznej rywalizacji, która to metoda nie pozwala na efektywne zwiększanie rozmiarów (skalowanie) sieci. Sprawność sieci może zostać znacznie zwiększona przez wykorzystanie przełączników do obsługi tych protokołów. Zastosowanie przełączników zmniejsza rozmiar domeny kolizji nie zmniejszając rozmiaru jej domeny nadawania. W sieciach LAN 802.3 o portach przełączanych domena kolizji ograniczona jest do dwóch tylko urządzeń: urządzenia końcowego i portu koncentratora do którego jest ono przyłączone. Zmniejsza to w znacznym stopniu problem skalowalności będący poważnym ograniczeniem tego typu sieci.

1.5.6.2 Sieć Token Ring 802.5

Sieć Token Ring oferuje przesyłanie danych z prędkością 4 i 16 Mbps z możliwością prognozowania opóźnień ze względu na przewidywalną metodę dostępu do nośnika. Prócz tego token takiej sieci wyposażony jest w bity priorytetu służące do nadawania ramce wysokiego priorytetu, co ułatwia realizację bardziej krytycznych zadań.

1.5.6.3 **Sieć** FDDI

Sieć FDDI umożliwia przesyłanie danych z prędkością 100 Mbps oraz topologię podwójnych, przeciwbieżnych pierścieni, potrafiących samodzielnie usuwać skutki uszkodzeń. Sieć FDDI może być (i przez niektórych jest) uważana za szybszą wersję sieci Token Ring, jako że również w tego rodzaju sieci token przesyłany jest w ściśle określony sposób. Ta właściwość oraz cechy wspólne specyfikacji 802.1 i 802.2 są jednak wszystkim, co sieci te mają ze sobą wspólnego.
Taktowanie zegara oraz wynikająca stąd metodologia taktowanego dostępu do nośnika odróżnia sieci FDDI od sieci Token Ring i sprawia, że nadają się one do zupełnie różnych zastosowań w środowiskach sieci LAN. Ich podwójne i przeciwbieżne pierścienie potrafią automatycznie i logicznie złączyć się ze sobą w celu ominięcia uszkodzonej części kabla. Dzięki temu sieci tego typu odznaczają się wrodzoną tolerancją błędów. Wadą takiego rozwiązania jest znaczne obniżenie sprawności sieci w przypadku uszkodzenia (np. przerwania kabla).
Inną ważną różnicą między sieciami FDDI oraz Token Ring jest przepustowość sieci FDDI. W odróżnieniu od sieci Token Ring, sieci FDDI potrafią obsługiwać przesyłanie wielu ramek danych jednocześnie. Pozwala to na wykorzystanie większego zakresu szerokości pasma. Bardziej szczegółowe informacje na temat FDDI oraz ich właściwości znaleźć można w rozdziale 10 pt. „FDDI".

1.5.6.4 Sieć VG-AnyLAN 802.12

Sieć VG-AnyLAN umożliwia łączenie ramek o formatach FDDI oraz Token Ring. Tego rodzaju sieć jest w dużym stopniu niezależna od rodzaju zastosowanego nośnika, jako że umożliwia przesyłanie danych za pomocą czterech par nieekranowanej skrętki dwużyłowej kategorii 3, nieekranowanej skrętki kategorii 5, ekranowanej skrętki kategorii 5 oraz kabla światłowodu o średnicy 62,5 mikrona.
Zastosowanie dostępu do nośnika na zasadzie priorytetu żądań lokuje sieć tego typu pomiędzy technologiami transmisji wąsko- i szerokopasmowych. Dokładnie rzecz biorąc, technologia tego typu ustanawia architekturę priorytetów pozwalającą pakietom krytycznym na uzyskanie, w razie potrzeby, odpowiedniej szerokości pasma. Architektura ta nie posiada mechanizmów umożliwiających rezerwowanie szerokości pasma.
Sieci VG-AnyLAN charakteryzują również dwa inne potencjalnie ważne ograniczenia. Po pierwsze, wymaga ona aż czterech par skrętki dwużyłowej, co może zmusić użytkowników okablowania lOBase-T do ponownego okablowania swoich stacji. W związku z tym - mimo że technologia ta została zaprojektowana specjalnie do wykorzystania za skrętką dwużyłową kategorii 3 - ci użytkownicy, którzy nie mają położonych czterech par skrętki kategorii 3, nie będą mogli jej używać. Drugim ograniczeniem jest brak współpracy z „prawdziwym Ethernetem" ze względu na inną metodę dostępu do nośnika.

1.5.7 Podsumowanie

Sieci lokalne określane są przez warstwę łącza danych. Jej dwoma głównymi składnikami są: struktura ramek oraz metoda dostępu do nośnika. Każda z architektur sieci LAN składa się z innej kombinacji obsługiwanych nośników transmisji, konwencji tworzenia ramek i metod sterowania dostępem do nośnika. Podczas wybierania architektury sieci (takiej jak Ethernet, Token Ring, FDDI itp.) szczegółowo rozważyć należy każde z wymienionych kryteriów dla każdego z dostępnych rozwiązań w celu określenia, czy spełniają one wymagania odnośnie wydajności tworzonej sieci LAN.

1.6 Rozdział 6 Mechanizmy dostępu do nośnika

Mark A. Sportack
W rozdziale 5 pt. „Warstwa łącza danych" określone i omówione zostały różne struktury przenoszenia danych warstwy 2. W niniejszym rozdziale opisane są różne sposoby umieszczania tych ramek na nośniku. Omówionymi sposobami są: przesyłanie tokenu, rywalizacja, priorytet żądań i przełączanie. Każdy z tych sposobów jest unikatowym zestawieniem właściwości i składników funkcjonalnych również omówionych w niniejszym rozdziale. W rozdziale podane są również specyficzne przykłady omawianych architektur sieci LAN.

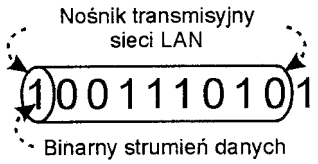
1.6.1 Dostęp do nośnika

Każda sieć musi w jakiś sposób regulować dostęp do nośnika. Mechanizm regulacji dostępu do nośnika realizowany jest przez warstwę 2 modelu referencyjnego OSI, czyli warstwę danych. Mimo że potrzeba sterowania dostępem jest jedna i ta sama, to sposoby jej zaspokajania mogą być bardzo różne. I tak w sieciach LAN dostęp do nośnika regulowany może być na jeden z czterech różnych sposobów:
• rywalizacji,

• przesyłania tokenu, • priorytetu żądań,
• przełączania.
Każda z tych metod dostępu do nośnika opisana jest szczegółowo w dalszej części niniejszego rozdziału.

1.6.2 Dostęp do nośnika na zasadzie rywalizacji

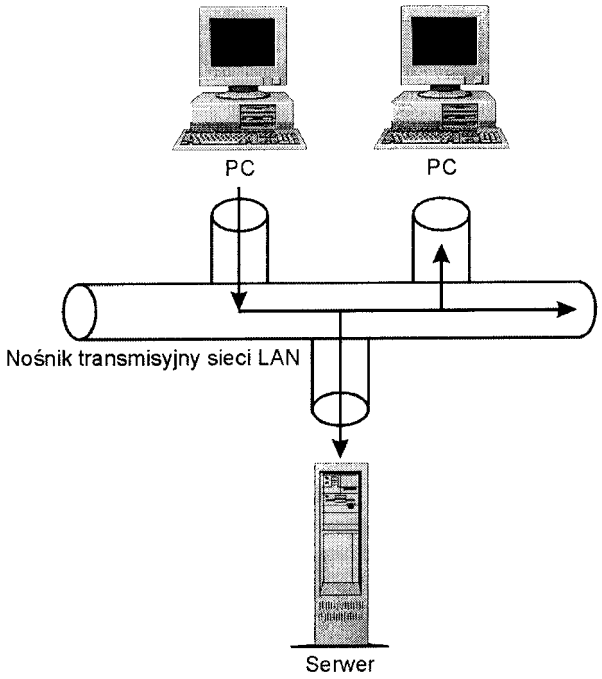
Sieć LAN, która używa rywalizacji jako podstawy do przyznawania prawa do transmisji, określana jest jako wykorzystująca metodę dostępu do nośnika na zasadzie rywalizacji. Wszystkie urządzenia konkurujące ze sobą o dostępne pasmo szerokości tworzą domenę kolizji. Omawiana metoda przyjęta została w setkach różnych wersji sieci Ethernet. Niektórymi z nich są:
• Ethernet 11, czyli DIX Ethernet,
• IEEE 802.3 - 10 Mbps Ethernet (CSMA/CD), • IEEE 802.3 - 100 Mbps Fast Ethernet,
• IEEE 802.3z- 1 Gbps Gigabit Ethernet.
Dostęp na zasadzie rywalizacji jest dość prostym sposobem regulowania dostępu, gdyż nie posiada on żadnych scentralizowanych mechanizmów regulacyjnych. Zamiast tego każde urządzenie przyłączone do sieci przyjmuje na siebie ciężar samodzielnego przeprowadzenia transmisji. Za każdym razem, kiedy urządzenie chce przesyłać dane, musi sprawdzić, czy kanał transmisyjny jest wolny, czy też nie. Jeśli nie, to urządzenie, które właśnie o mały włos wysłałoby dane, musi swój „zamysł" porzucić i odczekać określony przedział czasu przed podjęciem ponownej próby wysłania. Wszystkie urządzenia konkurują więc o dostęp do nośnika na zasadach i według logiki ustanowionej przez warstwę fizyczną. Wszystkie urządzenia konkurujące ze sobą o dostęp do nośnika tworzą domenę rywalizacji.
Domena rywalizacji jest czasem nazywana domeną kolizji. Nie ma w tym żadnego błędu, jako że kolizje mogą zachodzić tylko między urządzeniami LAN, które współzawodniczą o ten sam zakres dostępnej szerokości.
W definicji dostępu do nośnika na zasadzie rywalizacji domyślnie założono, że wszystkie urządzenia przyłączone do sieci mogą dane odbierać i wysyłać w tym samym zakresie częstotliwości. Nośniki transmisji mogą jednocześnie obsługiwać jeden tylko sygnał, który zajmuje całą dostępną szerokość pasma transmisyjnego. Innymi słowy, nośniki transmisyjne obsługują transmisje pasem podstawowym. Fakt ten został zilustrowany na rysunku 6.1.
Rysunek 6.1. Transmisja pasmem podstawowym.



Technologia transmisji pasmem podstawowym całą komunikację prowadzi z wykorzystaniem tylko jednego kanału. Wynikają stąd dwojakiego rodzaju konsekwencje:
• tylko jedno urządzenie może przesyłać dane w danej chwili,
• urządzenie może informacje albo odbierać, albo wysyłać, ale nigdy obie te czynności nie występują jednocześnie; taki sposób działania nazywany jest półdupleksem.

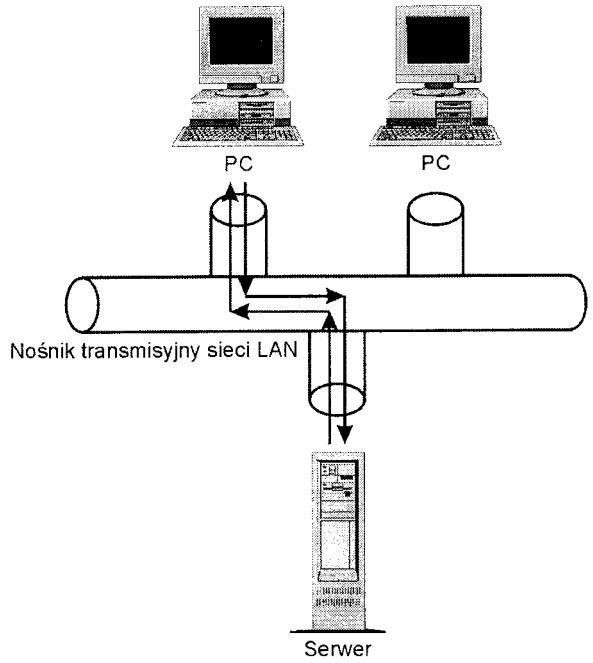
1.6.2.1 Półdupleks a pełny dupleks

W sieci wykorzystującej półdupleks tylko jedno urządzenie może przesyłać dane w określonej chwili; pozostałe muszą czekać, wsłuchując się co pewien czas w stan kanału transmisyjnego. Sytuacja taka przedstawiona została na rysunku 6.2.
Rysunek 6.2. Półdupleksowa transmisja pasmem podstawowym.
Termin „sieć pełnego dupleksu" implikuje, że dostępna szerokość pasma jest w pewien sposób podzielona na odrębne kanały. Do wydzielenia odrębnego kanału użyć można poszczególnych przewodów (żył) kabla wielożyłowego będącego nośnikiem. Typowa pełnodupleksowa sieć LAN korzysta z technologii przełączania. Niezależnie od sposobu uzyskania pełnodupleksowości, charakteryzuje się ona tym, że urządzenia mogą jednocześnie wysyłać dane i je odbierać. Pełnodupleksowe łącze komutowane (przełączane) przedstawione jest na rysunku 6.3.
Warto zauważyć, że w sieciach pełnodupleksowych o dostępie do nośnika na zasadzie rywalizacji tylko jedno urządzenie w obrębie określonej domeny kolizji może transmitować sygnały w określonym czasie. Na korzystanie jednocześnie z wielu urządzeń pozwala dopiero wprowadzenie portu przełączanego dla każdej pary urządzeń. W takiej sytuacji liczba urządzeń w domenie kolizji ograniczona zostaje do dwóch: urządzenia i portu, do którego jest ono przyłączone.



Rysunek 6.3. Pełnodupłeksowa transmisja przełączana (komutowana).

1.6.2.1.1 *Podstawa to* timing



Termin „timing" po ang.ielsku oznacza zarówno odpowiednie umiejscowienie czegoś w wymiarze czasu (na przykład opowiedzenie odpowiedniego w danym momencie żartu czy dostarczenie przesyłki na czas), jak i „taktowanie" - w sensie wyznaczania liczby cykli w jednostce czasu (na przykład: komputer taktowany procesorem Alpha 666 MHz); w niniejszym rozdziale interesuje nas jedynie druga z wymienionych ewentualności.

Niezależnie od liczby urządzeń sieci LAN umieszczonych w obszarze jednej domeny kolizji, a nawet od tego, czy sieć obsługuje transmisję pół-czy pełno dupleksową, muszą istnieć pewne mechanizmy regulujące dostęp urządzeń do nośnika. Dostęp do nośnika na zasadzie rywalizacji jest właśnie jednym z takich mechanizmów. Ta prosta metoda dostępu do nośnika została spopularyzowana razem z siecią Ethernet II. W specyfikacji Ethernet IEEE 802.3 została ona określona dużo bardziej rygorystycznie. Ogólnie rzecz biorąc, dostęp do nośnika na zasadzie rywalizacji jest mieszaniną timingu (taktowania) i matematyki.

Specyfikacja IEEE 802.3 opisuje architekturę LAN, która korzysta z metodologii dostępu CSMA/CD, czyli z metody wielodostępu do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji. Architektura ta nosi również nazwę CSMA/CA.

Aby taktowanie mogło być skoordynowane dla wszystkich urządzeń, przesyłane ramki warstwy 2 powinny mieć podobną długość. Długość ramek jest jednak z natury dynamiczna. W związku z tym unikanie kolizji może stanowić całkiem skomplikowane zadanie.

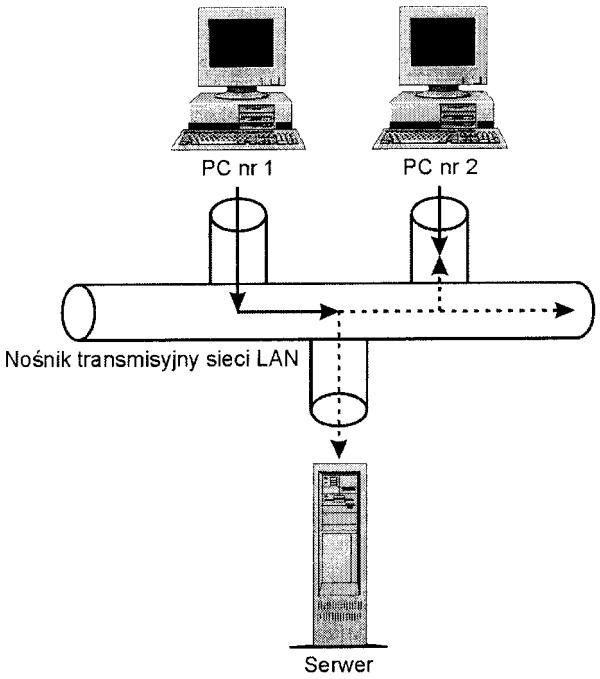
Sposób, w jaki 802.3 CSMA/CD - czyli obecnie popularna wersja Ethernetu - usiłuje uniknąć kolizji, polega na ustalaniu górnych i dolnych granic rozmiaru ramki. Ramka musi zatem mieć więcej niż 64 oktety, lecz nie więcej niż 1524 oktety, włącznie z przenoszonymi danymi oraz nagłówkami. Dla transmisji o różnych prędkościach te górne i dolne granice długości ramki wyznaczają bezpośrednio maksymalne i minimalne czasy transmisji dla ramek zgodnych ze standardem.
Ramki CSMA/CD, które zamiast 6-oktetowych używają 2-oktetowych adresów nadawcy i odbiorcy, mogą mieć długość całkowitą nie przekraczającą 1516 oktetów. Więcej informacji na temat ramek sieci Ethernet znaleźć można w rozdziale 5 pt. „Warstwa łącza danych". Szczegółowe informacje na temat różnych implementacji Ethernetu znajdują się w rozdziale 7 pt. „Ethernet" oraz w rozdziale 8 pt. „Szybsze sieci Ethernet".
W sieciach sterujących dostępem na zasadzie rywalizacji ilość czasu potrzebna do przesłania ramki przez sieć może być użyta do rozpoznania kolizji. Warto pamiętać, że dostęp na zasadzie rywalizacji zakłada, iż dane przesyłane są w sieci za pomocą pasma podstawowego, tak więc przesyłane ramki muszą zostać pomyślnie dostarczone do wszystkich końców sieci, tak aby upewnić się przed rozpoczęciem transmisji kolejnych ramek, że wszyscy uczestnicy sieci otrzymali ramki poprzednie.
Ramka może zostać zniszczona w wyniku wejścia w kolizję w dowolnym miejscu sieci LAN. Prawdopodobieństwa zajścia kolizji zwiększają dwa czynniki:
• liczba urządzeń przyłączonych do sieci,
• fizyczny rozmiar sieci.
Im więcej urządzeń przyłączonych jest do sieci, tym większa zachodzi między nimi rywalizacja o dostępny zakres pasma przesyłania. A im dłuższa sieć, tym więcej czasu zajmuje przesłanie ramki do końca sieci. Oba czynniki należy określić tak, aby pozwalały na osiągnięcie odpowiednich poziomów wydajności sieci.
Liczba urządzeń przyłączonych do sieci jest stale (patrząc z perspektywy czasu) zmniejszana, a to przez zastosowanie takich mechanizmów segmentacji, jak mostki, routery i przełączniki. Mimo że wszystkie te urządzenia działają w różny sposób, wszystkie one zmniejszają efektywny rozmiar domeny rywalizacji sieci LAN. A sam fizyczny rozmiar sieci określany jest przez specyfikacje warstw fizycznych architektur różnych sieci LAN.
Ustalenie dokładnej wartości najdłuższego czasu transmisji pierwszy raz wykonaliśmy dla sieci DIX Ethernet - wyszło nam 50 ms. Aby zapewnić pomyślne dostarczenie ramki, urządzenie wysyłające sygnał musiało kontynuować przesyłanie danych przynajmniej przez ten czas. Gwarantowało to bezkolizyjne przenoszenie ramek do najdalszych zakątków sieci (przy założeniu, że sieć LAN została utworzona zgodnie ze specyfikacjami określającymi maksymalną długość ścieżki sygnału i średnicę sieci). Dzięki temu bowiem operatorzy sieci mogli być pewni, że każde z urządzeń przyłączonych do sieci mogło wykryć przesyłaną kablem ramkę i wstrzymać się z transmisją.
Niestety, do przesłania sygnału przewodem potrzeba czasu. Zdarza się więc, że urządzenie rozpocznie nadawanie sygnału, z którym po kilku nanosekundach w kolizję wejdzie inny sygnał. Oba urządzenia rozpoznają zajście kolizji, zatrzymują transmisję i rozpoczynają ją od nowa, odczekawszy odpowiednio długi przedział czasu. Na tym w zasadzie polega wykrywanie kolizji. Przykładową sytuację kolizyjną przedstawia rysunek 6.4. PC numer 1 rozpoczął transmisję, ale nie dotarła ona do PC numer 2, który sprawdziwszy, że kanał transmisyjny jest wolny, również rozpoczął nadawanie. Oba wysłane sygnały „zderzą się" ze sobą, w wyniku czego pecety przerwą wysyłanie sygnałów i po odczekaniu okresu o pseudolosowej długości podejmą kolejną próbę transmisji.
50 mikrosekund wystarcza do wysłania 500 bitów z prędkością 10 Mbps. Podzieliwszy 500 (bitów) na 8 (bitów w oktecie) daje 62,5. Oznacza to, że pakiety musza mieć przynajmniej 62,5 oktetów długości, aby wykrywanie kolizji mogło działać. Xerox zaokrąglił minimalny rozmiar ramki sieci Ethernet DIX do 64 oktetów. Konwencja ta została utrzymana również dla specyfikacji 802.3.
Ze względu na to, że rozmiar ładunku danych (czyli rozmiar danych przenoszonych przez ramkę) określany jest przez protokół transportu warstwy wyższej, zdarza się od czasu do czasu, że ramka ma mniej niż 64 oktety długości. Takie ramki wypełniane są zerami do osiągnięcia rozmiaru minimalnego. Wypełnianie zlikwidowało problem taktowania (timingu) wykrywania kolizji, lecz zmusiło każdy protokół do rozpoznawania, co jest wypełnieniem ramki, a co niesioną przez nią informacją. Ramki Ethernetu DIX do rozpoznawania przenoszonego protokołu warstwy wyższej, a zatem również do ustalania długości własnego pola Dane, używają pola Typ. Mimo że sieć 802.3 nie rozpoznaje pola Typ w taki sposób, jej przykrótkie ramki nadal wypełniane są zerami aż do osiągnięcia minimalnego rozmiaru 64 oktetów.
Do mechanizmu wykrywania kolizji CSMA/CD podobny jest mechanizm standardowej wielkości przerw między ramkami. Wersja 802.3 sieci Ethernet korzysta z 96-oktetowej przerwy między ramkami. Wszystkie urządzenia o tym wiedzą i wszystkie się tego spodziewają. W przeciwieństwie do różnych pokrętnych domysłów mających na celu wyjaśnienie obecności tej struktury w sieciach, przerwa międzyramkowa spełnia funkcję podobną do sygnału nośnego w sieciach telefonicznych. Wskazuje urządzenie nadające i umożliwia ilość czasu wystarczającą do kontynuowania i umieszczania danych w ramkach i przesyłania ich. Po to właśnie w dzisiejszych sieciach Ethernet urządzenie wysyła ramkę, a po niej 96-bitową przerwę, która przez pozostałe urządzenia przyłączone do sieci interpretowana jest jako część strumienia bitów wysyłanego przez urządzenie nadające. Nie mogą więc one rozpocząć nadawania, dopóki przesyłane są bity przerwy, kiedy to dopiero urządzenie nadające oddaje kontrolę nad pasmem przesyłania.

1.6.2.1.2 Kolizje

Kolizje wśród administratorów sieci LAN cieszą się złą sławą. Już samo słowo „kolizja" ma negatywny wydźwięk. Nic dziwnego, wszak kolizje są sytuacjami niepoprawnymi. Jednak są one też naturalną częścią metodologii dostępu do nośnika opartej na zasadzie rywalizacji. We wszystkich tego rodzaju sieciach kolizje się zdarzają. Istnieją też jednak mechanizmy wykrywania i usuwania skutków kolizji. W związku z tym nie warto poświęcać zbyt wiele uwagi wskaźnikom liczb kolizji i nie panikować za każdym razem, gdy na koncentratorze zaświeci się żółte światełko. Jeśli sieć została zbudowana
poprawnie, czyli m.in. zgodnie z ograniczeniami warstwy fizycznej, takimi jak ograniczenia maksymalnych odległości dla nośników każdego z zastosowanych typów, sieć automatycznie powróci do stanu normalnego.
Jeśli wydajność sieci znacznie obniży się, najlepszym sposobem obniżenia liczby kolizji jest zmniejszenie liczby urządzeń dla domeny kolizji. Do tego celu wykorzystać można - jak pamiętamy - mostki, routery, przełączniki.

Warto pamiętać, że ze względu na chaotyczną naturę dostępu do nośnika opartego na zasadzie rywalizacji, technologie tego typu są nieodpowiednie do zastosowań wymagających określonego czasu reakcji. Dla zastosowań takich niezbędne są bardziej przewidywalne sposoby uzyskiwania dostępu do pasma przesyłania sieci LAN. Sieci kierujące się w dostępie do nośnika zasadą rywalizacji są za to idealne do bardziej tradycyjnych form wykonywania obliczeń sieciowych, takich jak emulowanie terminala, udostępnianie plików i drukarek itp.
Rysunek 6.4. Kolizja.

1.6.3 Dostęp **do nośnika na zasadzie pierścienia**



Najpopularniejszym sposobem dostępu do nośnika jest przesyłanie tokenu. Przesyłanie tokenu jest zjawiskiem charakterystycznym dla sieci LAN opartych na topologii pierścienia. Specyficznymi przykładami tego typu sieci są różne wersje sieci FDDI oraz Token Ring.

Token to specjalna ramka, która jest przesyłana w jednym kierunku do kolejnych urządzeń wchodzących w skład pierścienia. Token może być przesyłany tylko wtedy, gdy sieć jest wolna. Ramka tokenu ma najczęściej długość kilku oktetów i zawiera specjalny wzór bitów. Wzór ten jest zmieniany w celu zmiany tokena w sekwencję początku ramki informującą urządzenia znajdujące się w dalszej części pierścienia o tym, że otrzymana właśnie ramka jest ramką danych. Zaraz po sekwencji początku ramki umieszczone są w niej pary adresów odbiorcy i nadawcy.

Token uznawany jest przez wszystkie urządzenia za element decydujący o dostępie do nośnika. Jeśli token przesyłany jest do urządzenia, które nie ma akurat potrzeby wysyłania czegokolwiek, urządzenie to może przytrzymać token przez 10 ms lub dłużej - jeśli zmieniona została wartość domyślna. Czas ten ma pozwolić urządzeniu, które właśnie weszło w posiadanie tokena, na zakończenie umieszczania w ramkach danych otrzymanych od protokołów warstw wyższych. Aby umieścić jakiekolwiek dane w sieci, urządzenie musi znajdować się w posiadaniu tokena. Jeśli go nie ma, musi poczekać, aż otrzyma go od sąsiada poprzedzającego go w pierścieniu.

Jeśli czas upłynął, a urządzenie nie musiało nic przesyłać, oddaje ono kontrolę nad tokenem, który przekazywany jest do następnego urządzenia w sieci. Ogranicznik początku ramki może być przekonwertowany z powrotem do postaci tokenu tylko przez to urządzenie, które go umieściło w sieci. W końcu token dociera do urządzenia, które go utworzyło. Urządzenie to zmienia token do postaci pola Początku ramki. Zwykle wykonywane jest to po skopiowaniu przez urządzenie odbierające niesionych przez tę ramkę danych i zmodyfikowaniu jej wzoru bitowego w celu poinformowania urządzenie wysyłające ramkę o pomyślnym jej otrzymaniu. Tak zmodyfikowana ramka danych kontynuuje swą podróż dokoła pierścienia, aż do powrotu do swego nadawcy, który otrzymawszy potwierdzenie pomyślnego dostarczenia zawartości, albo trzyma token przez pewien określony czas, albo używa go do przenoszenia kolejnych danych. Omawiany schemat przesyłania tokenu przedstawiony jest na rysunku 6.5.

Maksymalny czas wymagany, zanim urządzenie rozpocznie nadawanie, daje się policzyć. W tym celu należy pomnożyć maksymalną ilość czasu, jaką każdy węzeł może trzymać token, przez liczbę węzłów przyłączonych do sieci. Do tak uzyskanej wartości dodać należy czas potrzebny tokenowi do przejścia przez całą sieć. Mimo że tak uzyskana wartość nie uwzględnia czasu potrzebnego do przeprowadzenia operacji we/wy, czasu przetwarzania, czasu pozycjonowania głowic dysku twardego ani opóźnień związanych z działaniem procesora, to umożliwia w miarę dokładne oszacowanie maksymalnej wartości odczuwanego opóźnienia.

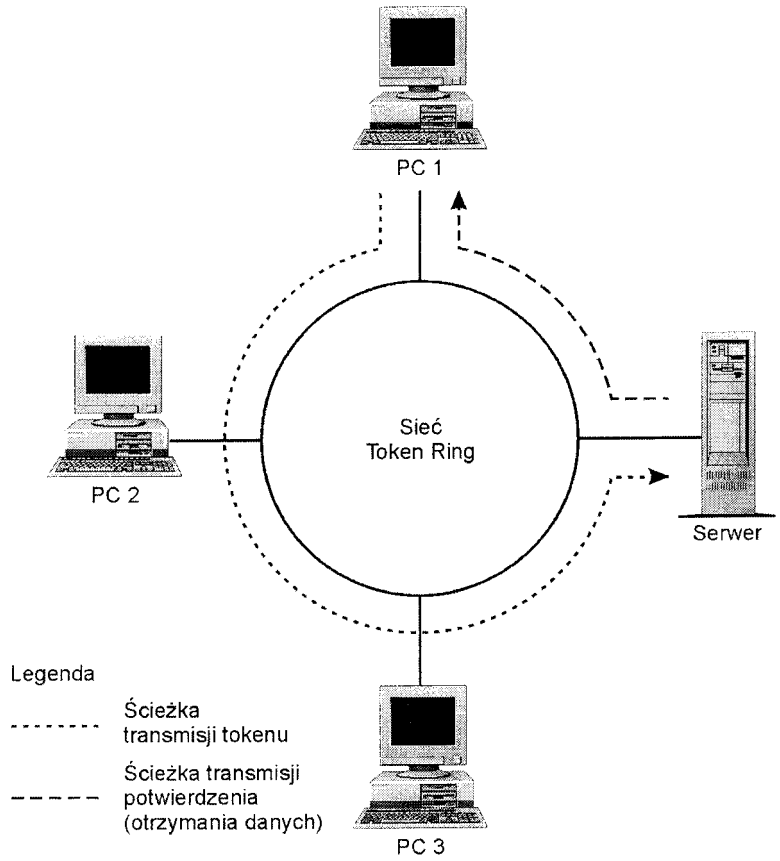
Czas ten można wydłużać lub skracać przez zwiększanie lub zmniejszanie liczby węzłów w sieci. W związku z tym sieci oparte na przesyłaniu tokenu nadają się idealnie do zastosowań wymagających przewidywalnej wartości opóźnień.

Dostęp do nośnika na zasadzie pierścienia w sieciach FDDI

Sieci FDDI korzystają ze schematu przesyłania tokenu opisanego w punkcie poprzednim z subtelną, lecz znaczącą różnicą. Stacje nie muszą już wstrzymywać się z dalszą pracą do czasu otrzymania przez nadawcę potwierdzenia pomyślnego dostarczenia

Rysunek 6.5. Przesyłanie tokenu.

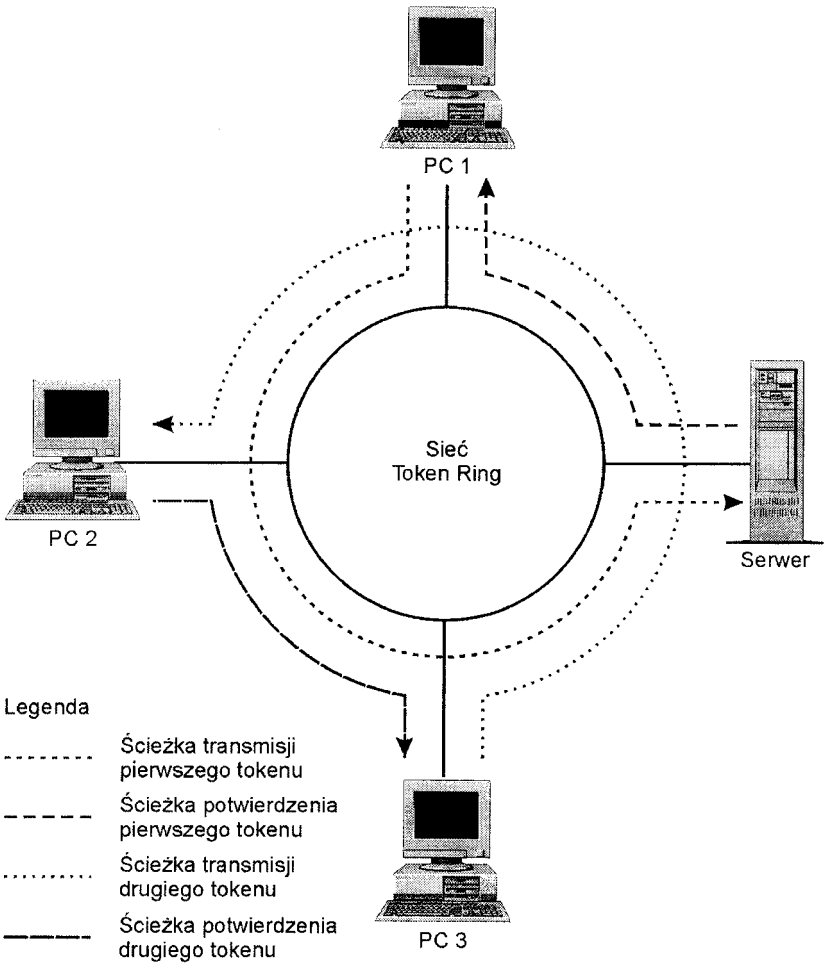
przesyłanej ramki. Zamiast tego sieci FDDI korzystają z mechanizmu szybkiego uwalniania tokenu, który pozwala innym urządzeniom przesyłać dane, mimo że uprzednio wysłana ramka nadal znajduje się w drodze (do miejsca docelowego). Mechanizm szybkiego uwalniania jest mechanizmem dość prostym. Bezpośrednio po wysłaniu ramki (zawierającej token zamieniony na pole początku ramki) urządzenie przesyłające dane wysyła drugi token. Dzięki temu kolejne stacje pierścienia nie muszą już wstrzymywać się z przesyłaniem danych do czasu, aż ramka z danymi powróci do jej nadawcy. Kontrola nad nośnikiem transmisyjnym w wersji FDDI dostępu do nośnika jest więc przekazywana dużo szybciej. Omawiany proces przedstawiony jest na rysunku 6.6.



Korzyści płynące z zastosowania szybkiego uwolnienia tokenu są dość oczywiste; dzięki temu, następne urządzenie uzyskuje możliwość przesłania danych dużo wcześniej, co oznacza, że może ono zdjąć nowo utworzony token z sieci i zamienić go na ogranicznik początku ramki, nawet zanim jeszcze wcześniejsza ramka dotrze do swego adresata.
Drugą korzyścią stosowania tego schematu jest zwiększenie wydajności działania sieci. Maksymalna obsługiwana przepustowość sieci wyposażonej w mechanizm szybkiego uwalniania zbliża się więc do teoretycznej przepustowości maksymalnej.
Rysunek 6.6. Przesyłanie tokenu przy użyciu szybkiego uwolnienia.

1.6.4 Dostęp do nośnika na zasadzie priorytetu żądań

Metoda dostępu na zasadzie priorytetu żądań wykorzystywana jest w sieciach odpowiadających specyfikacji IEEE 802.12 100 Mbps VG-AnyLAN. DPAM jest metodą cyklicznego przyznawania prawa dostępu, w której centralny wzmacniak (koncentrator) regularnie sprawdza stan portów do niego przyłączonych. Sprawdzanie to wykonywane jest w kolejności portów i ma na celu określenie, które z nich zgłaszają żądania transmisji. Po rozpoznaniu zgłoszenia koncentrator określa jego priorytet, który może być normalny lub wysoki. Powodem wprowadzania priorytetów jest potrzeba umożliwienia uprzywilejowanego dostępu do nośnika procesom, które obsłużone muszą być w określonym czasie. Każdy port nie przeprowadzający transmisji przesyła sygnał wolny (nośny). Do portów takich należą również wszystkie urządzenia nie wysyłające w danym momencie danych oraz urządzenia, których zgłoszenia transmisji są chwilowo zawieszone.



Ów sygnał wolny jest przez wzmacniak usuwany w momencie wybrania urządzenia jako kolejnego do rozpoczęcia transmisji. Innymi słowy, wzmacniak identyfikuje stację następną w kolejce do przeprowadzenia transmisji, a następnie nakazuje jej zaprzestanie wysyłania sygnału wolnego. Port może rozpocząć transmisję dopiero po zaprzestaniu wysyłania sygnału wolnego.

Również wówczas wzmacniak informuje pozostałe stacje, że mogą one otrzymać wiadomość przychodzącą. Następnie odczytuje on adres odbiorcy otrzymanego pakietu, sprawdza go w swojej tabeli konfiguracyjnej i przesyła ramkę dalej tak, jak uczyniłby to każdy port pracujący w trybie podsłuchu.

Wzmacniak centralny, lub inaczej główny, kontroluje działanie domeny priorytetów. Może składać się z nie więcej niż trzech warstw wzmacniaków połączonych ze sobą kaskadowo. Kaskadowo połączone wzmacniaki mogą bowiem działać jak jeden wielki wzmacniak. Wtedy wzmacniak centralny przesyła cały ruch do poszczególnych wzmacniaków warstwy niższej, a te z kolei sprawdzają swoje porty aktywne na obecność żądań transmisji pakietów.

Żadna stacja nie może wykonywać dwóch transmisji pod rząd, jeśli zawieszone żądania transmisji innych stacji mają taki sam priorytet. We wzmacniaku centralnym żądanie o wysokim priorytecie nie przerwie rozpoczętej już realizacji żądania. We wzmacniaku niższej warstwy żądanie o priorytecie normalnym zostaje zawłaszczone w celu umożliwienia jak najszybszej realizacji żądania o priorytecie wyższym. Aby zapewnić, że żadne z żądań nie będzie wiecznie ignorowane, żądania o priorytecie normalnym, które oczekują dłużej niż 250 ms (2,5s), automatycznie uzyskują priorytet wysoki.

Ta metoda dostępu do nośnika wykorzystywana jest przez specyfikację IEEE 802.12 dla sieci 100 Mbps, o ramkach formatu Token Ring lub Ethernet (ale nigdy obu jednocześnie) oraz topologii gwiazdy. Sieci tego rodzaju znane są jako sieci VG-AnyLAN (skrót od ang.. voice grade wiring, any LAN architecture, czyli okablowanie jakości telefonicznej, sieć LAN o dowolnej architekturze). Jako nośnik transmisji wykorzystywać mogą one cztery pary zarówno ekranowanej, jak i nieekranowanej skrętki dwużyłowej (UTP) kategorii 3, skrętki dwużyłowej kategorii 5 oraz kabla światłowodu. Sieć tego typu obsługuje do trzech warstw kaskadowo łączonych wzmacniaków o odległości między wzmacniakiem i stacją roboczą nie większej niż 100 metrów. Średnica sieci może mieć do 1300 metrów średnicy.

Sieci typu VG-AnyLAN okazały się niemalże zupełnym fiaskiem. Mimo że dostęp do nośnika na zasadzie priorytetu żądań jest technicznie sprawniejszą metodą dostępu do nośnika sieci LAN niż CSMA/CD, czyli wielodostęp do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji, to szybsze sieci LAN, takie jak Fast Ethernet czy Gigabit Ethernet, oferują dużo prostsze niż VG-AnyLAN sposoby rozwoju technologii sieci Ethernet 10 Mbps CSMA/CD.

1.6.5 Dostęp do nośnika w komutowanych sieciach LAN

Prócz przedstawionych trzech podstawowych sposobów dostępu do nośnika istnieje też metoda czwarta; choć ściśle rzecz biorąc, nie jest ona metodą dostępu. Jest natomiast

metodą coraz częściej stosowaną zamiast przedstawionych typów w celu zwiększenia sprawności i wydajności sieci LAN. Przełączanie zmienia dotychczasowe zasady dotyczące topologii i metodologii dostępu sieci LAN.

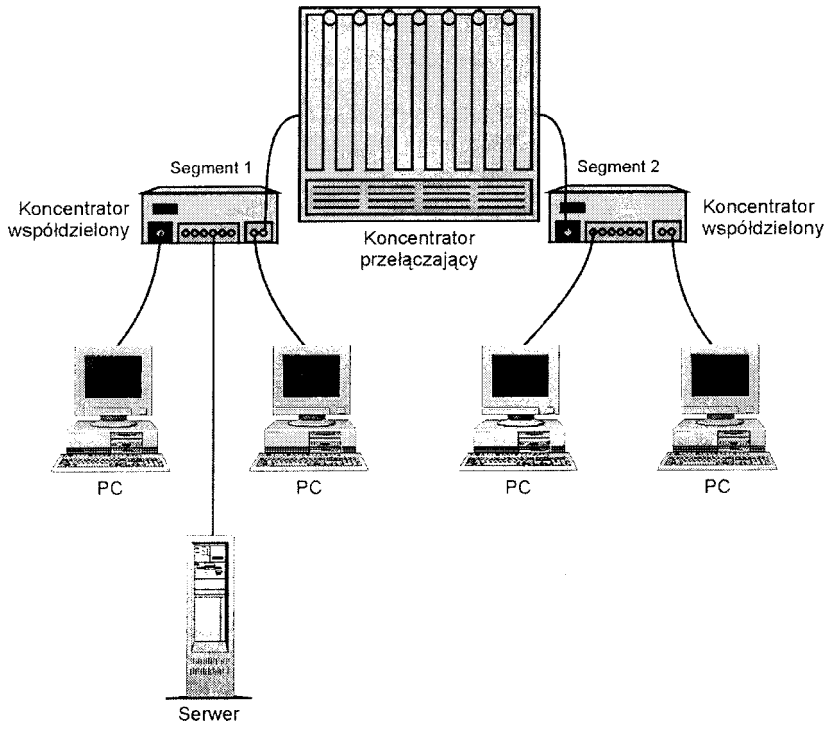
Jak pamiętamy z rozdziału 2 pt. „Typy i topologie sieci LAN", przełącznik jest wieloportowym urządzeniem warstwy łącza danych (warstwy 2 modelu referencyjnego OSI). Przełącznik „uczy się" adresów i „zapamiętuje" je w wewnętrznej tabeli . Tworzy między nadawcą i odbiorcą ścieżki przełączane, którymi następnie przesyłane są dane.

Przełączanie może być stosowane zarówno do wzajemnego łączenia współdzielonych koncentratorów, jak i poszczególnych urządzeń. Segmentowanie koncentratorów współdzielonych za pomocą koncentratora przełączającego znane jest jako przełączanie segmentów. Przedstawione jest to na rysunku 6.7.

Rysunek 6.7. Przełączanie segmentów.

W sieciach LAN o przełączanych portach każdy z portów koncentratora przełączającego połączony jest z jednym urządzeniem, którym może być serwer, stacja robocza czy drukarka. Sytuacja ta przedstawiona jest na rysunku 6.8.

Każde urządzenie przyłączone do portu komutowanego (czyli przełączanego) ma własną domenę warstwy 2, którą współdzieli jedynie z tym portem. Przełączanie może być używane do zwiększania sprawności sieci LAN o architekturach stosujących dostęp do nośnika zarówno na zasadzie rywalizacji, jak i na zasadzie przesyłania tokenu.



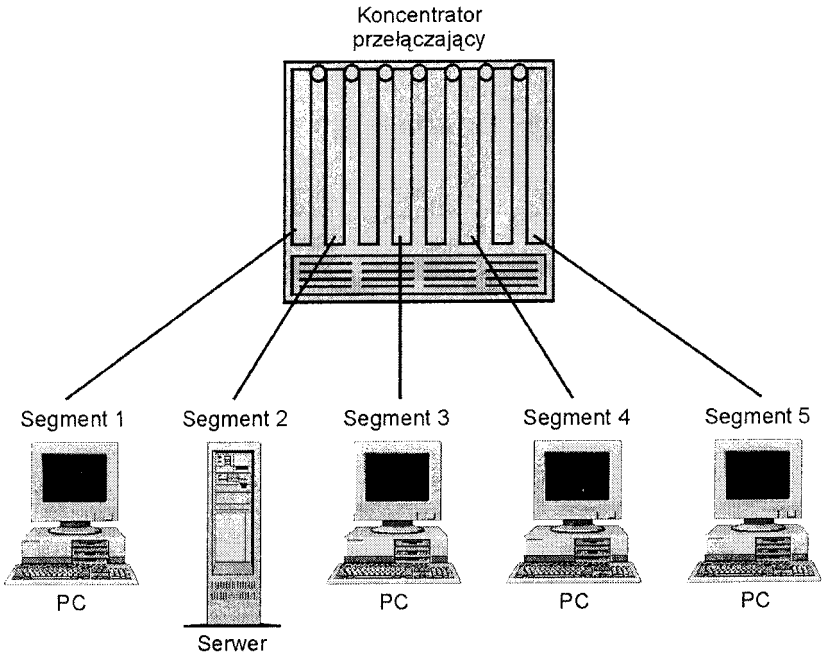
Sieci przełączane o dostępie do nośnika opartym na zasadzie rywalizacji

W protokołach wykorzystujących mechanizm rywalizacji przełączanie portu zmniejsza rozmiar domeny kolizji do dwóch tylko urządzeń: przełączanego portu oraz urządzenia, które za jego pośrednictwem jest przyłączone do sieci.

W zależności od rodzaju sprzętu znajdującego się na obu końcach sieci CSMA/CD, mogą one obsługiwać połączenia pełno- lub tylko półdupleksowe.

Rysunek 6.8. Port przełączany (komutowany).

Komutowane połączenia pełnodupleksowe pozwalają na uzyskanie maksymalnej sprawności możliwej do osiągnięcia przy określonej szybkości transmisji. Każde z urządzeń wchodzących w skład dwuurządzeniowej domeny (czyli zarówno port, jak i przyłączone doń urządzenie) ma swoje własne kanały sieciowe i własną odrębną szerokość pasma służącą do odbierania i wysyłania sygnałów. Na przykład, w sieciach Ethernet 10 Mbps wykorzystujących okablowanie skrętką dwużyłową jedna para przewodów doprowadzona do każdego urządzenia przełączanego służy do nadawania, a druga do odbierania sygnałów.



Rozdzielenie funkcji nadawania i odbioru eliminuje wszelkiego rodzaju konkurencję. W implementacjach pełnodupleksowych para przewodów używana przez przełącznik do nadawania jest jednocześnie parą, za pośrednictwem której urządzenie przyłączone do koncentratora odbiera dane, i na odwrót - para przewodów służąca przełącznikowi do odbioru jest parą, której urządzenie do niego przyłączone używa do nadawania. W rozwiązaniu tym nie ma miejsca na kolizje. Jest ono przedstawione na rysunku 6.3, we wcześniejszej części niniejszego rozdziału.

Niektórzy producenci wprowadzają klientów w błąd, twierdząc, iż pełnodupleksowe łącza sieci komutowanych (przełączanych) oferują pasmo przesyłania o dwukrotnie większej szerokości. Technika ta zwiększa oczywiście wydajność, ale niekoniecznie przez podwojenie szerokości pasma transmisyjnego. Zarówno bowiem funkcja nadawania, jak i odbioru sygnałów obsługiwane są przez tę samą kartę sieciową i muszą współzawodniczyć o cykle procesora, do którego dostęp uzyskują przez wspólny port magistrali znajdujący się na płycie głównej. W związku z tym w wyniku zastosowania techniki pełnego dupleksu nie należy oczekiwać podwojenia sprawności systemu, lecz raczej przesunięcia wąskiego gardła w inne miejsce sieci.

Przełączanie sieci wykorzystujących metodę przesyłania tokenu

Przełączanie portu może usprawnić działanie sieci LAN opartych na przesyłaniu tokenu w taki sam sposób, w jaki usprawnia ono działanie sieci korzystających z metody dostępu opartego na zasadzie rywalizacji. Liczba urządzeń, które przesyłają tokeny, jest ograniczona do absolutnego minimum: portu oraz urządzenia do niego przyłączonego. Jedyna różnica polega na tym, że dostęp do nośnika jest regulowany nie przez konkurencję, lecz za pomocą przesyłania tokenu.

W odróżnieniu od sieci LAN wykorzystujących konkurencję, sieci LAN oparte na przesyłaniu tokenu uniemożliwiają innym urządzeniom komunikowanie się do czasu powrotu ramki do nadawcy. Zasada ta nie przestaje obowiązywać również w środowisku przełączanym. Ze względu na to, poprawa sprawności wynikająca z wprowadzenia pełnodupleksowości w sieciach wykorzystujących przesyłanie tokenu nie jest tak znaczna, jak w przypadku sieci wykorzystujących konkurencję jako sposób sterowania dostępem do nośnika.

1.6.6 Podsumowanie

W wąskopasmowych (czyli wykorzystujących pasmo podstawowe) sieciach LAN regulowanie dostępu do nośnika staje się funkcją niezwykle istotną. Dostęp ten próbowano dotychczas regulować za pomocą trzech różnych sposobów - przez rywalizację, za pomocą przesyłania tokenu oraz w oparciu o priorytety żądań. W praktyce używa się jedynie dwóch ze wspomnianych sposobów, a mianowicie rywalizacji oraz przesyłania tokenu. Metoda dostępu do nośnika oparta na priorytecie żądań została prawie całkowicie zapomniana. Co ciekawe, również pozostałe dwie metody stają się coraz to bardziej przestarzałe w miarę wprowadzania przełączanych sieci LAN.

Każdy z przedstawionych sposobów dostępu do nośnika charakteryzuje się określonym zestawem zalet i wad, które powinny zostać dobrze poznane i zrozumiane, zanim zostaną one użyte jako fundament dla dalszej wiedzy w zakresie technologii informacyjnych.

1.7 Rozdział 7 Ethernet

Mark A. Sportack

Ethernet, stworzony jako prowizoryczny mechanizm pomagający naukowcom odkrywać nowe technologie, okazał się jedną z najbardziej wartościowych i trwałych technologii informatycznych. Wszedł w trzecią dekadę istnienia, podlegając w międzyczasie znacznej ewolucji. Niektóre z tych zmian sprawiły, że praktycznie niemożliwe jest podanie zwięzłej definicji Ethernetu.

W niniejszym rozdziale opisane są różne potencjalne znaczenia terminu „Ethernet" oraz implementacje warstwy fizycznej sieci Ethernet 10 Mbps. Opisane zostały także niektóre ograniczenia tych sieci.

1.7.1 Różne rodzaje sieci Ethernet

Pierwotnie Ethernet oznaczał sieć lokalną utworzoną przez naukowców w centrum badawczym firmy Xerox w Palo Alto. Ethernet (jak nazwali go naukowcy) nie był śmiałą, nową technologią, stworzoną z myślą o jej ogromnym potencjale rynkowym, lecz raczej prostym narzędziem, ułatwiającym naukowcom wymianę danych podczas odkrywania i wdrażania nowych technologii.
Ów pierwotny Ethernet był siecią niezbyt wyszukaną. W dużym stopniu jej działanie, w tym również wielkość ramek, opierało się na lepiej zdefiniowanych protokołach warstwy sieci i transportu. Była to sieć półdupleksowa, w której urządzenia łączone były za pomocą grubego kabla koncentrycznego. Prędkość przesyłania sygnału wynosiła 10 Mbps. Obecnie ten typ sieci znany jest jako PARC Ethernet lub Ethernet I. Nazwy te zostały wprowadzone dopiero po utworzeniu innych, nowych form Ethernetu w celu umożliwienia rozróżniania ich. Dziś Ethernet w swej oryginalnej postaci jest przestarzały i wspomina się o nim tylko dla celów historycznych i porównawczych.
Odkrywszy potencjalną wartość tej technologii, firma Xerox pozyskała partnerów w celu wprowadzenia jej na rynek. Były nimi firmy Intel oraz DEC (Digital Equipment Corporation). Firmy te wspólnie dokonały szeregu ulepszeń sieci PARC Ethernet i uczyniły ją czymś w rodzaju standardu otwartego. Tak zmieniony Ethernet nazwano Ethernet II.

Ethernet II, znany także jako DIX Ethernet, od pierwszych liter nazw jego twórców (Digital, Intel i Xerox), nie mógł być uważany za prawdziwy standard otwarty, ponieważ był kontrolowany przez trzech głównych wytwórców jego komponentów.
Jednym z pierwszych kroków było zatwierdzenie Ethernetu jako samodzielnego protokołu sieciowego, który do określenia rozmiarów ramki nie musiałby już korzystać z protokołów warstwy sieci i transportu.
Osiągnięcie to było ważne. Istotniejsze jednak było udoskonalenie metodologii dostępu do nośnika. Oryginalny Ethernet używał bardzo prymitywnej metody „słuchaj zanim zaczniesz mówić", znanej też jako wielodostęp do łącza sieci z badaniem stanu kanału lub metoda CSMA (ang.. Carrier Sense, Multiple Access). Jej istota polegała na tym, że stacja, która chciała przesyłać dane, musiała najpierw upewnić się, że jest to możliwe, „nasłuchując", czy linie przesyłowe (kanały) są wolne. Usprawnienie polegało na dodaniu możliwości wykrywania kolizji. Nowa metodologia dostępu do nośnika zastosowana w Ethernecie II nazwana została więc wielodostępem do łącza sieci z badaniem stanu kanału i wykrywaniem kolizji (ang.. CSMA/CD - Carrier Sense, Multiple Access with Collision Detection). Więcej informacji o metodzie CSMA/CD można znaleźć w rozdziale 6 pt. „Mechanizmy dostępu do nośnika".
W lutym 1980 roku instytut IEEE wziął na siebie odpowiedzialność za przekształcenie rozwijającego się Ethernetu w prawdziwy standard otwarty. Prawdę mówiąc, jego celem nie było wyłącznie znormalizowanie Ethernetu, lecz dalej idące i ambitniejsze znormalizowanie technologii sieciowych i - co za tym idzie - umożliwienie im współdziałania. W celu utworzenia standardów dla sieci MAN i LAN uruchomiono projekt 802, z wieloma podkomisjami i technicznymi grupami roboczymi.
Projekt 802 rozbił typową sieć na składniki funkcjonalne i pogrupował je w logicznie po sobie następujące warstwy. Wprowadził standardy otwarte dla adresowania na poziomie sprzętowym, zarządzania siecią i monitorowania jej. Standardy te stały się podstawą wielu architektur sieci lokalnych, takich jak:
• Ethernet 10 Mbps o dostępie CSMA/CD,
• Token Ring,
• Token Bus,
• inne.
Fakt, że architektury te posiadały wspólną podstawę dla adresowania na poziomie sprzętowym, zarządzania siecią i monitorowania, oznaczał, że można było tworzyć sieci o topologiach mieszanych, nie uzgadniając współoperacyjności między różnymi platformami sieciowymi. Co ważniejsze, można było zarządzać różnymi platformami i monitorować je za pomocą wspólnego zestawu narzędzi odpowiadających każdemu z owych standardów.
Stworzona przez IEEE wersja Ethernetu została formalnie nazwana 802.3 CSMA/CD. Nazwa ta jest jednak niefunkcjonalna i niemal zawsze mówi się po prostu Ethernet. Metodologia dostępu do nośnika CSMA/CD została zachowana, tak samo jak oryginalnie
stosowany kabel koncentryczny oraz półdupleksowy tryb transmisji. Następnie dodano kolejną specyfikację dla kabla koncentrycznego, a po pewnym czasie również inne specyfikacje warstwy fizycznej - dla skrętki dwużyłowej i dla światłowodu.

Obecnie bardziej stosowne staje się określanie Ethernetu za pomocą przymiotników innych niż CSMA/CD. W ciągu ostatnich dwóch lat oryginalną specyfikację 802.3 rozszerzono tak, aby obejmowała również wersję Ethernetu 100 Mbps. Spowodowało to konieczność wprowadzenia nowych specyfikacji dla warstwy fizycznej, a także pomniejszych modyfikacji mechanizmów dostępu do nośnika. Co ważniejsze, dodano również obsługę transmisji pełnodupleksowej (czyli w obu kierunkach jednocześnie). Pełnodupleksowy Ethernet pozwala urządzeniu wykorzystywać do nadawania jedną fizyczną ścieżkę, którą tworzy jedna z dwóch par skrętek dwużyłowych znajdujących się w kablu ośmiożyłowym przy jednoczesnym odbieraniu danych przychodzących drugą parą przewodów. Pełnodupleksowy Ethernet o konfiguracji przełączanej za pomocą portu skutecznie zapobiega konfliktom związanym z dostępem do nośnika na zasadzie rywalizacji. Urządzenie nadawcze może dzięki temu umieszczać ramki w sieci z prędkością przewodzenia. Specyfikacja Ethernet IEEE 802.3 jak więc widać uległa znacznym zmianom, w związku z czym określanie sieci tego typu mianem CSMA/CD jest błędne.
Prędkość przewodzenia to prędkość, z jaką elektrony przepływają przez przewód miedziany. Jest to maksymalna prędkość teoretyczna, z jaką dane (zamienione na strumień elektronów) mogą być przesyłane przez medium transmisyjne.
Specyfikacje utworzone zostały nie tylko dla Ethernetu o podstawowym paśmie transmisji 10 Mbps, lecz również dla Ethernetu szerokopasmowego, którego jednym z przykładów jest specyfikacja lOBroad36. Na szczęście dla jasności wywodów, szerokopasmowe sieci Ethernet leżą poza głównym nurtem zastosowań, więc nie musimy niepotrzebnie komplikować sobie nimi życia. Kursy dotyczące podstaw

sieci w ogóle nie obejmują tego tematu. Sieci tego rodzaju tylko powierzchownie bowiem przypominają zwykły Ethernet. W niniejszym rozdziale skupimy się zatem wyłącznie na wariantach pracujących w paśmie podstawowym.
Aby uniknąć nieporozumień, w książce na określenie serii standardów 802.3, jako ogólnego, łącznego zestawu technologii używamy terminu „rodzina standardów Ethernetu". Terminu „Ethernet" używamy na określenie standardów 802.3 działających z prędkością 10 Mbps. Natomiast termin „Fast Ethernet" dotyczy standardów 802.3, działających z prędkością 100 Mbps. Podobnie termin „Gigabit Ethernet" opisuje wprowadzane dopiero standardy pracujące z prędkością 1024 Mbps. Każda z technologii ethernetowych omawianych na stronach tej książki omawiana jest odrębnie - jak choćby opis specyfikacji interfejsu międzynośnikowego (ang.. MDI - Medium Dependent Interface) warstwy fizycznej (na przykład 1OBaseT). Selektywność taka jest konieczna, ponieważ obecnie Ethernet jest licznym zbiorem zróżnicowanych standardów zarówno już istniejących, jak i dopiero powstających.

1.7.2 Obsługiwany sprzęt

Zanim zagłębimy się na dobre w zawiłości Ethernetu, rozpatrzmy różne składniki sprzętowe używane do budowy sieci Ethernet. Komponenty te są niezależne od rodzaju nośnika, zatem wspominać o nich będziemy w różnych kontekstach w dalszej części niniejszego rozdziału. Omówmy je więc, nakreślając tło odpowiednie do przedstawienia następnie ich roli w sieci.
Na sprzęt, który może być używany do obsługi sieci Ethernet, składają się: • karty sieciowe,
• koncentratory wzmacniające,
• koncentratory nie wzmacniające, • mosty,
• routery.

1.7.2.1 Karty sieciowe

Karta sieciowa to płytka drukowana instalowana w wolnym gnieździe magistrali (ang.. *l/O bus)* komputera. Z tyłu karty znajduje się fizyczny interfejs dla określonego rodzaju złącza. Każdy rodzaj złącza zaprojektowany jest dla konkretnego nośnika. Karta zapewnia połączenie między wewnętrznymi zasobami systemu komputerowego a zasobami zewnętrznymi, przyłączonymi do sieci. Zawiera układy logiczne warstwy łącza sieciowego oraz warstwy fizycznej.

1.7.2.2 Wzmacniaki

Wzmacniak (repeater) jest względnie prostym urządzeniem, które wzmacnia sygnał wejściowy, nie zmieniając jego kształtu. Działa ono wyłącznie na poziomie warstwy 1 (fizycznej) modelu referencyjnego OSI.
Funkcje wzmacniania i powtarzania sygnału zostały zebrane w urządzeniach wieloportowych, używanych do łączenia urządzeń w sieci LAN. Popularnie są one nazywane koncentratorami, choć tak naprawdę są to właśnie wzmacniaki wieloportowe.

1.7.2.3 Koncentratory nie wzmacniające

Koncentrator nie wzmacniający jest bardzo podobny do wzmacniającego. Jedyna w zasadzie między nimi różnica polega na tym, że koncentrator nie wzmacniający nie wzmacnia ani nie powtarza sygnałów. Koncentratory tego rodzaju służą w zasadzie wyłącznie do łączenia wielu stacji roboczych, umożliwiając tworzenie sieci o topologii gwiazdy.

1.7.2.4 Mosty

Most jest mechanizmem warstwy 2 (warstwy łącza danych) umożliwiającym łączenie dwóch segmentów sieci lokalnej. Ponieważ mosty działają w warstwie 2, nie rozpoznają one protokołów wyższych warstw, osadzonych w ramkach, które są przez mosty przekazywane. Do przesyłania ramek mosty wykorzystują adresy fizyczne (adresy MAC). Przykładowo, most „uczy się", które adresy fizyczne są przyłączone do sieci przy użyciu jego portów. Dzięki temu w razie odebrania przez most ramki z adresem fizycznym nie występującym w segmencie sieci, dla którego została ona wygenerowana, szuka on tego adresu w swojej tablicy mostkującej (która jest zestawieniem adresów fizycznych oraz numerów portów) i wysyła ramkę do odpowiedniego segmentu sieci.
Zestawianie razem kilku mostów - czy nawet używanie ich w środowisku wykorzystującym adresy MAC - jest dobrym sposobem poznania ograniczeń tych urządzeń. W większości przypadków nie są one w stanie sprostać wzrastającym wymaganiom stawianym urządzeniom sieciowym i w takich warunkach stają się czymś w rodzaju muzealnej ciekawostki.

1.7.2.5 Routery

Router nie jest, jak się niektórym wydaje, urządzeniem warstwy 2 sieci Ethernet. Jest bowiem mechanizmem przesyłania pakietów funkcjonującym na poziomie warstwy 3. Routery obsługują interfejsy wszystkich standardowych technologii LAN i mają liczne zastosowania. Głównym zastosowaniem routerów jest łączenie sieci lokalnej z sieciami spoza jej domeny. Mato trzy ważne implikacje dla konstrukcji sieci.
Po pierwsze, sieć rozległa (WAN) wyraźnie przekracza zakres domeny sieci LAN. Do łączenia różnych sieci lokalnych rozproszonych po tak dużych regionach geograficznych, na jakie pozwalają technologie transmisji dalekiego zasięgu, niezbędne są routery. Technologiami takimi są dedykowane linie dzierżawione oraz obwody przełączane.
Po drugie, wiele domen sieci lokalnych może koegzystować we (względnej) wzajemnej bliskości. Pojedynczy budynek biurowy może na przykład zawierać wiele sieci LAN przeznaczonych dla wielu różnych grup roboczych. Jeśli względy bezpieczeństwa uzasadniają pewien stopień separacji, ale nie w stopniu uniemożliwiającym wzajemne łączenie sieci, to sieci te najlepiej łączyć nie za pomocą mostów, lecz przy

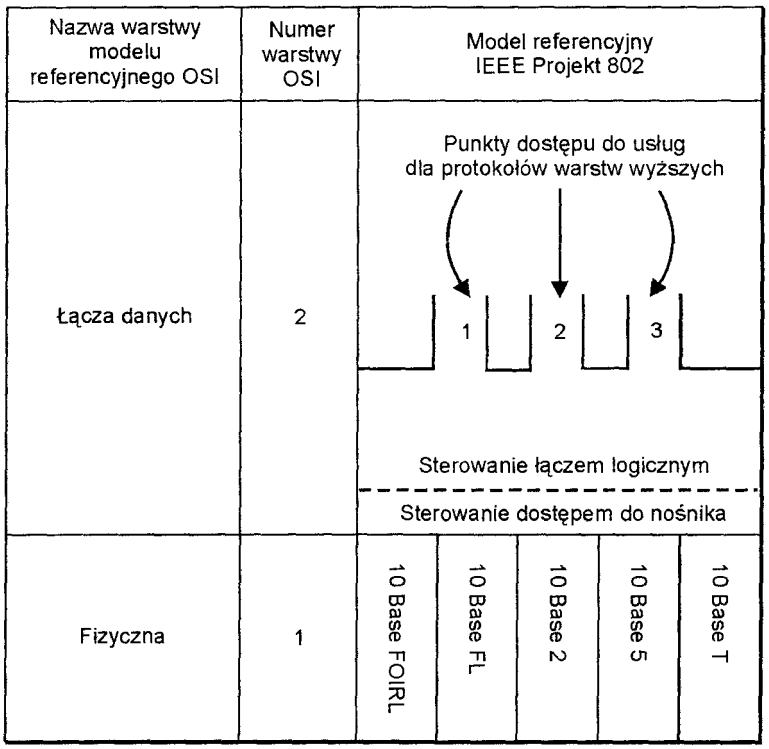
użyciu routerów. Routery zapewniają bowiem nieco większe bezpieczeństwo, dzięki mechanizmom, takim jak listy kontroli dostępu, i pozwalają na efektywne łączenie sieci z zachowaniem integralności ich domen warstwy 2: tak kolizji, jak i rozgłaszania.
Trzecia konsekwencja wynika z wymagań dotyczących wydajności. Sieci Ethernet mogą osiągnąć tak wielkie rozmiary, że zaczną ulegać samosegmentacji. Na przykład, w budynku może istnieć sieć Ethernet, w której działają 1024 urządzenia. Jeśli obsługiwane przez tę sieć aplikacje używają transmisji rozgłoszeniowych warstwy 2, to całkiem możliwe, że łączenie segmentów sieci LAN mostami lub przełącznikami wcale nie zwiększy jej wydajności. Umożliwi jedynie segmentację domeny kolizji sieci LAN, lecz nie jej domeny nadawania. W takiej sytuacji router stanowić może jedyne praktyczne rozwiązanie problemu.

1.7.3 Funkcje warstwowe

Członkowie IEEE rozpoczęli swe wysiłki standaryzacyjne od zgrupowania niezbędnych funkcji sieci lokalnych i metropolitalnych w moduły czy też warstwy, bazując na kolejności zdarzeń następujących podczas normalnej sesji komunikacyjnej. Jak wiele innych organizacji normalizacyjnych, również oni utworzyli swój własny stos protokołów, nie przystający ściśle do modelu referencyjnego OSI. Rysunek 7.1 ilustruje kompletny model IEEE, z podziałem funkcji na warstwy, podwarstwy, a nawet oddzielne moduły w ramach różnych warstw.

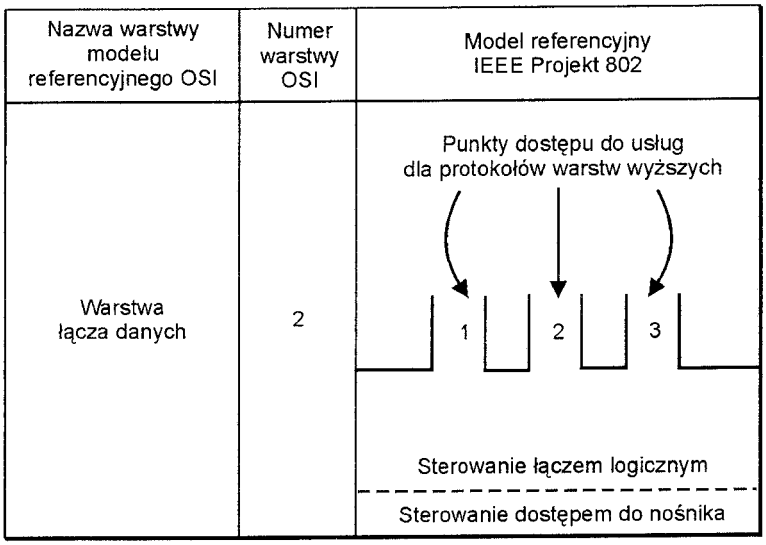
1.7.3.1 Funkcje warstwy łącza danych

Specyfikacje serii IEEE 802 dzielą warstwę łącza danych modelu OSI (czyli jego drugą warstwę) na dwie odrębne części. Ich nazwy pochodzą od nazw kontrolowanych przez nie funkcji, a są to:
• sterowanie łączem logicznym (ang.. LLC- Logical Link Control,
• sterowanie dostępem do nośnika (ang.. MAC- Media Access Control.
Związki między nimi przedstawia rysunek 7.2.
Rysunek 7.1. Schemat blokowy porównujący specyfikacje IEEE l0 Mbps Ethernet oraz modelu referencyjnego OSI.



Rysunek 7.2. Składniki warstwy łącza-a danych wg IEEE.

Wspólnie warstwy LLC oraz MAC tworzą serce Ethernetu. Umożliwiają one umieszczanie danych w ramkach (czyli ich opakowywanie) oraz adresowanie danych, co pozwala na przesyłanie ich do miejsca przeznaczenia. Warstwy te posiadają również mechanizmy wykrywania błędów i są odpowiedzialne za inicjowanie retransmisji uszkodzonych lub utraconych ramek. Krótko mówiąc, sterują nadawaniem i odbieraniem ramek danych, aczkolwiek nie przeprowadzają rzeczywistej transmisji, gdyż za nią odpowiada warstwa fizyczna.



1.7.3.1.1 Sterowanie łączem logicznym

Warstwa LCC jest wyższym z dwóch składników warstwy łącza danych. Izoluje ona protokoły wyższej warstwy od właściwej metody dostępu oraz nośnika. Należy pamiętać, że specyfikacje serii 802 zapewniają współoperacyjność różnym architekturom sieci lokalnych. Sterowanie łączem danych jest mechanizmem uniezależniającym protokoły warstw sieci i transportu od różnych odmian architektury sieci LAN. Dzięki niemu protokoły wyższych warstw nie muszą wiedzieć, czy będą przesyłane poprzez Ethernet, Token Ring, czy też Token Bus. Nie muszą również wiedzieć, jakiej specyfikacji warstwy fizycznej będą używać. Sterowanie LLC udostępnia wspólny interfejs dla wszystkich architektur i odmian sieci LAN zgodnych ze specyfikacją 802.

1.7.3.1.2 Sterowanie dostępem do nośnika

Warstwa MAC jest niższym składnikiem warstwy łącza danych w architekturze IEEE. Odpowiada za połączenie z warstwą fizyczną oraz zapewnia udany przebieg nadawania i odbioru. Składają się na nią dwie funkcje: nadawania i odbioru.

Warstwa sterowania dostępem do nośnika odpowiada za opakowywanie wszystkich danych otrzymywanych z warstwy LLC w ramki. Prócz danych, ramka zawiera strukturę oraz wszystkie adresy potrzebne do przesłania jej do miejsca przeznaczenia. Warstwa MAC jest także odpowiedzialna za przeprowadzanie testu integralności danych, używanego do sprawdzania, czy zawartość ramki nie została uszkodzona lub zmieniona podczas transmisji.

Warstwa sterowania dostępem do nośnika zawiera również mechanizmy potrafiące określać - na podstawie mechanizmów warstwy fizycznej - czy pasmo komunikacyjne jest dostępne, czy też nie. Jeśli jest dostępne, ramki danych są przekazywane warstwie fizycznej do przesłania. Jeśli nie, warstwa MAC uruchamia swój binarny wykładniczy algorytm zwrotny, który generuje pseudolosowy czas oczekiwania, po upływie którego dopiero może nastąpić kolejna próba transmisji.

Ostatnią ważną funkcją warstwy sterowania dostępem do nośnika jest monitorowanie statusu transmitowanych ramek polegające na wykrywaniu wszelkich znaków sygnalizujących zajście konfliktu. Gdy warstwa MAC wykryje konflikt jednej ze swoich ramek, określa, które dane muszą być ponownie wysłane, uruchamia algorytm zwrotny i ponownie próbuje wysłać ramkę. Algorytm zwrotny jest powtarzany, dopóki próba wysłania ramki nie zakończy się powodzeniem. Jest to jednocześnie siłą, jak i słabością Ethernetu. Gwarantuje protokołom wyższej warstwy, że ich dane zostaną dostarczone. Niestety, chaotyczna natura mechanizmu dostępu do nośnika opartego na zasadzie rywalizacji może spowodować, że to zadanie to będzie trudne i czasochłonne.

1.7.3.2 Funkcje warstwy fizycznej

Podobnie jak warstwa łącza danych, również warstwa fizyczna modelu OSI została przez instytut IEEE podzielona na odrębne składniki. Uzyskana w ten sposób modularność zapewnia elastyczność w adaptowaniu nowych technologii. Gdyby tak nie było, po każdej zmianie medium transmisyjnego (nośnika) należałoby dostosowywać całą warstwę pierwszą. Dzięki podejściu modularnemu, modyfikacji w takim przypadku wymaga jedynie mechanizm odpowiedzialny za połączenie z nowym medium transmisyjnym. Pozostałe funkcje warstwy fizycznej mogą być używane bez wprowadzania żadnych zmian. Rysunek 7.3 ilustruje składniki warstwy fizycznej w specyfikacji IEEE. Ich omówienie znajduje się w dalszej części rozdziału.

Wyróżniamy cztery następujące składniki warstwy fizycznej:

• Fizyczna podwarstwa sygnałowa (ang.. PLS-Physical Signaling Sublayer)

• Interfejs jednostki przyłączeniowej (ang.. AUI-Attachment Unit lnterface)

• Fizyczne przyłącze nośnika (ang.. PMA -Physical Medium Attachment)

• Interfejs międzynośnikowy (ang.. *MDI - Medium Dependent Interface)*

Razem komponenty te w pełni definiują przebieg transmisji między dwoma urządzeniami przyłączonymi do sieci. Definicja obejmuje rodzaje kabli (w tym minimalne oczekiwane poziomy wydajności), złączy kablowych, przypisania wyprowadzeń kabla (tylko dla skrętek

dwużyłowych), poziomy napięć (dla transmisji sygnału elektrycznego) lub długości fali świetlnej (dla transmisji światłowodowej), taktowanie, a nawet fizyczny interfejs sieciowy - tj.. nad-biornik (ang.. transceiver) - przeprowadzający zarówno nadawanie, jak i odbiór.
Rysunek 7.3. Skladniki warstwy /i.yc=nej w specyfikacji lEEE.

Nazwa warstwy modelu referencyjnego OSI	Numer warstwy OSI	Model referencyjny IEEE Projekt 802
Fizyczna	1	Fizyczna podwarstwa sygnałowa
		Interfejs jednostki przyłączeniowej
		Fizyczne przyłączenie nośnika
		Interfejs międzynośnikowy

Fizyczna podwarstwa sygnałowa (PLS) - jest mechanizmem lokalnym terminali (DTE) wykorzystujących okablowanie typu 1 OBaseT określającym schemat sygnalizowania oraz złącze kabla nad-biornika.
Interfejs jednostki przyłączeniowej (AUI) - określa specyfikacje nośnika.
Fizyczne przyłącze nośnika (PMA) - definiuje procesy operacyjne i specyfikacje nadbiornika.
Interfejs międzynośnikowy (MDI) - jest najbardziej zauważalną częścią warstwy fizycznej 802.3. Istnieje wiele interfejsów MDI, z których każdy opisuje mechanizmy niezbędne do obsługi transmisji przez różne nośniki.
Elementy AUI, PMA oraz MDI są często wbudowane w jedno urządzenie, określane w specyfikacji IEEE jako jednostka przyłączania nośnika lub jako jednostka MAU (ang.. *media attachment unit),* która to jednostka jest niczym innym jak kartą sieciową.
Interfejs międzynośnikowy nie określa samego nośnika! Definiuje raczej wszystkie mechanizmy i procesy obsługujące transmisję poprzez określony typ nośnika. Definicje te przyjmują pewien określony, minimalny poziom wydajności, który musi być przez nośnik zapewniony, a także określają rodzaje fizycznych połączeń z nośnikiem. Nie definiują jednak samego medium transmisyjnego.

1.7.3.3 Interfejsy międzynośnikowe warstwy fizycznej

IEEE definiuje pięć różnych interfejsów międzynośnikowych MDI dla sieci Ethernet działającej w paśmie podstawowym 10 Mbps. Interfejsy te pogrupowane są w moduły określające wszystkie aspekty warstwy fizycznej w stosunku do różnych nośników. Z pięciu interfejsów MDI dwa oparte są na kablu koncentrycznym, dwa na światłowodzie i jeden na miedzianej skrętce dwużyłowej. Kilka podstawowych ograniczeń specyfikacji dotyczy wszystkich interfejsów.
Na przykład, nośniki wszystkich typów mają określoną maksymalną ścieżkę sygnału. Maksymalna dozwolona ścieżka sygnału obejmuje cztery wzmacniaki. Pozwala to na połączenie pięciu segmentów kabla, przy założeniu, że co najmniej dwa z tych seg mentów służą wyłącznie jako połączenia między wzmacniakami! Ograniczenie to często nazywane jest regułą 5-4-3, ponieważ pozwala na połączenie pięciu segmentów, z wykorzystaniem czterech wzmacniaków, przy czym urządzenia przyłączać można do najwyżej trzech z owych segmentów.
Powyższe ograniczenie można łatwo przełożyć na maksymalne średnice sieci dla każdego z nośników fizycznych, mnożąc liczbę połączeń przez maksymalny zasięg transmisji dla każdego z tych nośników, a następnie sumując wszystkie wyniki.
Specyfikacja 802.3 nakłada też ograniczenie liczby mostów jednego segmentu LAN do siedmiu. Ograniczenie to dotyczy wszystkich rodzajów nośników.
Kolejne ograniczenie dotyczy liczby urządzeń w segmencie. Do każdego segmentu, niezależnie od rodzaju nośnika, przyłączonych może zostać do 1024 urządzeń. W sieciach, które mają obsługiwać więcej urządzeń, stosować trzeba mosty, przełączniki lub routery, w celu zwielokrotnienia liczby segmentów.
Dwa kable koncentryczne pozwalają na uzyskanie pewnej różnorodności topologicznej, niedostępnej innym interfejsom międzynośnikowym, dzięki temu, że kabel koncentryczny może być szpuntowany. „Szpuntowanie" kabla koncentrycznego polega na przebiciu (lecz nie przerywaniu) dwóch przewodów kabla i zamontowaniu w tym miejscu rozgałęźnika, który zapewnia ciągłość oryginalnego kabla, umożliwiając jednocześnie przyłączenie nowego kabla. Ów nowy kabel może z kolei być używany do przyłączania innych urządzeń sieciowych: wzmacniaków, serwerów, drukarek i komputerówklientów. Sieci Ethernet zgodne ze specyfikacją 802.3, mające okablowanie koncentryczne, mogą obsługiwać do 64 urządzeń na każdym rozgałęźniku. Odległość między rozgałęźnikami musi wynosić przynajmniej 2,5 metra. Skrętka dwużyłowa oraz światłowód są nośnikami transmitującymi „z punktu do. punktu", w związku z czym nie mogą one być szpuntowane. Rozgałęzianie sieci o tego typu nośnikach jest możliwe przy użyciu koncentratorów.
Okablowanie sieci Ethernet składa się z dwóch kategorii funkcjonalnych: nad-biornika oraz magistrali. Ich nazwy pochodzą z okresu początków Ethernetu (kiedy nie istniały jeszcze wzmacniaki). Wszystkie sieci Ethernet miały wtedy topologię magistrali. Dziś nad-biorniki

oraz magistrale wychodzą z użycia, a w związku z nowymi technologiami dawne kategorie funkcjonalne zyskują nowe nazwy, odpowiednio: „okablowanie stacji" (ang.. station wiring oraz „szkielet" (ang.. backbone).

1.7.3.3.1 1OBase2

lOBase2, jak i większość interfejsów międzynośnikowych Ethernetu, wywodzi swoją nazwę z następującej konwencji: szybkość sygnału (w Mbps) + metoda transmisji (w tym wypadku transmisja pasmem podstawowym) + maksymalna długość kabla w metrach, zaokrąglona do ł00, a następnie podzielona przez 100.
Rozpatrzmy przykładowo specyfikację o nazwie „lOBase2". Opisuje ona protokół sieciowy dla pasma podstawowego i prędkości lOMbps. Stosowany jest 50-omowy kabel koncentryczny o maksymalnej długości 185 metrów. 185 po zaokrągleniu daje 200. Dzieląc 200 na 100 otrzymujemy 2 - ostatnią cyfrę nazwy interfejsu.
Sieci 1 OBase2 mogą być rozszerzane poza granicę 185 metrów za pomocą wzmacniaków, mostów lub routerów. Używając routerów do segmentacji Ethernetu, tworzy się segmenty lOBase2, które mogą być rozgałęziane do 30 razy, przy czym każde z rozgałęzień może obsłużyć do G4 urządzeń.

1.7.3.3.2 10Base5

Jak wskazuje nazwa, maksymalna długość kabla koncentrycznego IOBase5 wynosi 500 metrów. Ten interfejs MDI wykorzystuje dużo grubszy koncentryk niż lOBase2. Skuteczność transmisji w przewodzie miedzianym jest bowiem funkcją grubości przewodnika. Im większa jest jego średnica, tym większą osiąga się szerokość pasma. W rezultacie, kabel lOBase5 może być rozgałęziany do 100 razy, przy zachowaniu maksymalnej liczby 64 urządzeń dla każdego rozgałęzienia.
Matematyka podpowiada, że możliwe jest zbudowanie segmentu lOBase5 zawierającego 6400 urządzeń. Niestety, ograniczenie liczby urządzeń w segmencie do 1024 odnosi się do wszystkich specyfikacji 802.3.

1.7.3.3.3 1OBaseT

Gdyby 1 OBaseT odnieść do konwencji nazewniczych dotyczących kabli koncentrycznych, nazwa ta powinna brzmieć 1OBasel, gdyż długość segmentu jest ograniczona do 100 metrów. Z jakichś powodów IEEE wyłamała się z konwencji i oznaczyła ten interfejs literą „T", symbolizującą jego nośnik fizyczny: skrętkę dwużyłową (twisted pair).
Specyfikacja 1OBaseT, wbrew powszechnemu przekonaniu, nie określa rodzaju kabla. Dotyczy ona natomiast specjalnej techniki sygnalizowania dla nieekranowanej skrętki dwużyłowej wykorzystującej cztery przewody spełniające wymogi trzeciej kategorii wydajności. Nazwy przewodów wskazują na ich funkcję oraz biegunowość. Jedna para przewodów obsługuje dodatnie i ujemne bieguny obwodu nadawania. Druga para obsługuje dodatnie i ujemne bieguny obwodu odbioru. Stąd nazwy czterech przewodów są następujące:
• N+ dla dodatniego przewodu nadawania, • N- dla ujemnego przewodu nadawania, • O+ dla dodatniego przewodu odbioru,
• O- dla ujemnego przewodu odbioru.
Specyfikacja 1OBaseT przypisuje tym przewodom odpowiedni układ wyprowadzeń kabla (czyli styków lub pinów złącza) skrętki dwużyłowej. Wzmacniaki/koncentratory 1OBaseT używają przyporządkowań wyprowadzeń, które umożliwiają tworzenie łączy z portami kart sieciowych. Funkcje pinów ethernetowych wzmacniaków i innego rodzaju urządzeń przesyłania danych, czyli komunikacyjnych (ang.. *DCE - Data Communications Equipment)* pokazuje tabela 7.1.
Tabela 7.1.
Numery i funkcje pinów urządzeń komunikacyjnych.
Funkcje pinów ethernetowych kart sieciowych i terminali, czyli urządzeń końcowych (ang.. *DTE-Dala terminal Equipment)* pokazuje tabela 7.2.

Numer pinu urządzenia komunikacyjnego	Funkcja
1	O+
2	O-
3	N+
4	nie używany
5	nie używany
6	N-
7	nie używany
8	nie używany

W normalnych warunkach urządzenie końcowe zawsze jest połączone z urządzeniem komunikacyjnym. Komplementarność interfejsów tych urządzeń pozwala łączyć je bezpośrednio za pomocą kabla, bez obaw o konflikty między nadawaniem i odbiorem. Kabel łączący składa się z czterech par przewodów biegnących prosto od jednego urządzenia do drugiego. Nie dziwi więc, że nazywa się go „kablem prostym". Istnieją jednak pewne wyjątki od tego prostego schematu łączenia urządzeń komunikacyjnych i końcowymi.
Tabela 7.2.
Numery i funkcje pinów urządzeń końcowych.

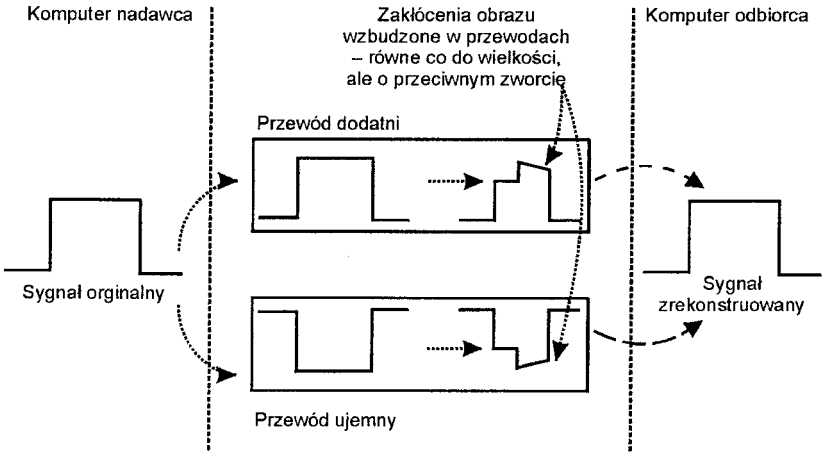
Pierwszy wyjątek stanowi na przykład łączenie ze sobą nowoczesnych koncentratorów, czyli urządzeń komunikacyjnych z urządzeniami komunikacyjnymi. Starsze koncentratory wyposażone były w oddzielne porty wyjściowe, skonfigurowane jako urządzenia końcowe. Ich wyłącznym zastosowaniem było umożliwianie tworzenia łączy z portami wejściowymi innych koncentratorów, mostów itd.. Jednak w większości dzisiejszych koncentratorów zrezygnowano z tego rozwiązania na korzyść szerokopasmowych, w pełni oczkowanych, tablic połączeniowych (ang.. backplanes) umieszczonych w solidnych obudowach mających wiele gniazd rozszerzeń. Producenci takich tablic uważają, że wystarczają one do połączenia ze sobą wszystkich przyłączonych do nich koncentratorów. Dlatego wszystkie porty są skonfigurowane jako wyjściowe, tzn. ich wyprowadzenia mają konfigurację urządzeń komunikacyjnych.

Numer pinu DTE	Funkcja
1	N+
2	N-
3	O+
4	nie używany
5	nie używany
6	O-
7	nie używany
8	nie używany

Łączenie ze sobą dwóch koncentratorów posiadających wyłącznie porty komunikacyjne stanowi jednak problem. Obydwa używają bowiem tej samej pary przewodów do nadawania i oba oczekują na sygnał na pozostałej parze. W efekcie każdy pakiet wysłany przez takie łącze albo ulega kolizji, albo jego odbiór nie zostaje potwierdzony przez adresata. Jest to oczywiście sytuacja niepożądana. Problem ten rozwiązuje połączenie omawianych urządzeń komunikacyjnych za pomocą kabla skrośnego zmieniającego przewody, którymi sygnały są przesyłane.
Drugim wyjątkiem jest bezpośrednie łączenie dwóch urządzeń wyposażonych w karty sieciowe, bez pośrednictwa koncentratorów/wzmacniaków. W przeciwieństwie do sytuacji poprzedniej, tu mamy do czynienia z połączeniem dwóch urządzeń końcowych. Efekt - którym jest brak komunikacji - pozostaje jednak ten sam.
Sytuacja taka jest dość niezwykła. W zasadzie, za wyjątkiem testowania, nie ma wielu powodów, aby w ten sposób łączyć ze sobą dwa urządzenia. Jednak może to ulec zmianie. Opracowywanie zaawansowanych metod grupowania technologicznie niezaawansowanych architektur obliczeniowych tworzy potrzebę niezawodnego łączenia w sieć dwóch urządzeń znajdujących się blisko siebie. Przyjmując, że początkowo większość takich grup będzie ograniczona do dwóch komputerów, instalowanie między nimi koncentratora wzmacniającego będzie niecelowe. A zastosowanie go zwiększyłoby jedynie opóźnienie
propagacji. Bez koncentratora grupa tego rodzaju działa wydajniej! Przykład takiej grupy stanowią komputery łączone bezpośrednio za pomocą kart sieciowych i kabla skrośnego.
Kable skrośne muszą zachowywać biegunowość fizycznego przewodu. Napięcia dodatnie i ujemne muszą być oddzielone, ponieważ są one podawane bezpośrednio na układy sygnałowe wykorzystywane przez Ethernet. Krzyżuje się tylko następujące przewody: dodatni nadawania (N+) z dodatnim odbioru (O+) i ujemny nadawania (N-) z ujemnym odbioru (O-).
Przesyłanie sygnałów różnicowych
Stosowanie dla każdej funkcji doprowadzeń czy też przewodów, zarówno dodatnich, jak i ujemnych umożliwia korzystanie z techniki sygnalizacyjnej znanej jako przesyłanie sygnałów różnicowych. Ma ona na celu wyeliminowanie wpływu zakłóceń elektromagnetycznych na dane. Zakłócenia tego typu względnie łatwo wzbudzają się w przewodach i stanowią jeden z najpoważniejszych problemów związanych z transmisją przy użyciu nieekranowanej skrętki dwużyłowej.
Urządzenie nadawcze wykorzystujące sygnał różnicowy, przesyła odwrotne postaci tego samego sygnału dwoma różnymi przewodami. Przewodem N+ przesyłany jest sygnał, a przewodem N- jego inwersja. Na drugim końcu kabla sygnał N-jest ponownie odwracany. Sygnały N+ i N-teoretycznie powinny więc być wtedy równe. Każde zakłócenie, przynajmniej teoretycznie, jest indukowane jednocześnie i w tej samej wielkości w obydwu przewodach. Tak więc jeśli zakłócenie spowoduje zwiększenie „szczytu" w przewodzie N+ kabla, to zwiększy ono również, o tę samą wartość, „dolinę" w przewodzie N-, i na odwrót. Dodanie obydwu sygnałów różnicowych na końcu kabla umożliwia uzyskanie na powrót pierwotnego sygnału. „Szczyt" i „dolina" znoszą się bowiem nawzajem całkowicie.
Omawiany proces przedstawiony jest na rysunku 7.4. Oryginalny sygnał płynie przewodem N+, a jego inwersja przewodem N-. Zakłócenie przedstawia się jako niewielkie wzniesienie na płaszczyźnie jednego z kodowanych (w kodzie Manchester) bitów przesyłanych przewodem N+, W przewodzie N- zakłócenie wytwarza tej samej wielkości dolinę. Na końcu kabla dodatnie i ujemne „szczyty" i „doliny" znoszą się, przywracając oryginalny kształt sygnału.
Skrętka dwużyłowa jest dość szczególnym typem okablowania: jej standardy dotyczą nie atrybutów fizycznych, lecz raczej poziomów sprawności. Przed zakupem okablowania warto zatem zwrócić uwagę na jej kategorię wydajności. Kategorie są numerowane od 1 do 5 i poprzedzone skrótem Cat#. Obecnie jedynie kable Cat 3 i Cat 5 są odpowiednie dla sieci LAN.
Więcej informacji na temat kategorii wydajności można znaleźć w rozdziale 3 pt. „Warstwa fizyczna".
Rysunek 7.4. Likwidacja zakłóceń za pomocą sygnału różnicowego.

1.7.3.3.4 10BaseFL

Specyfikacja lOBaseFL umożliwia transmisję w paśmie podstawowym z prędkością 10 Mbps przez wielofunkcyjny kabel światłowodowy o średnicy 62,5/125 mikrona. Maksymalna długość kabla wynosi 2 km. Podobnie jak skrętka dwużyłowa, również światłowód nie może być rozgałęziany. Jest on bowiem nośnikiem łączącym „z punktu do punktu".



lOBaseFL może służyć do łączenia wzmacniaków ze sobą, a nawet do łączenia serwerów ze wzmacniakiem. Połączenie tego typu jest nieco droższe niż porównywalne z nim połączenie 1OBaseT, ale może być stosowane w sieciach o większych rozmiarach.

1.7.3.3.5 10BaseFOIRL

Najnowszym dodatkiem do specyfikacji serii 802.3 jest lOBaseFOIRL. Ten dość spory skrót pamięciowy oznacza transmisję w paśmie podstawowym z prędkością 10 Mbps, z wykorzystaniem łączy światłowodowych pomiędzy wzmacniakami. Z tej definicji wynika, że ta technologia jest ograniczona wyłącznie do połączeń między wzmacniakami. Innymi słowy, jest to połączenie „koncentrator z koncentratorem" za pomocą okablowania światłowodowego. Nie można do niego przyłączać żadnych innych urządzeń.
lOBaseFOIRL wykorzystuje kabel światłowodowy o średnicy 8,3 mikrona, który musi być sterowany przez iniekcyjną diodę laserową, czyli diodę ILD (ang.. injection laser diode). To połączenie sprzętu i nośnika zapewnia efektywną transmisję sygnałów w paśmie podstawowym z prędkością 10 Mbps na odległość do 5 km. Niestety, zarówno iniekcyjne diody laserowe, jak i jednofunkcyjny kabel światłowodowy są dość drogie, co ogranicza zainteresowanie nimi na rynku.
Dość powszechnie mówi się, że kable światłowodowe mają średnicę 9 mikronów. W rzeczywistości, średnica ta wynosi 8,3 mikrona i jest zaokrąglana do kolejnej liczby całkowitej. Co dziwne - nie dotyczy to światłowodu wielofunkcyjnego o średnicy 62,5 (którego średnicę podaje się in extenso).

1.7.3.4 Mieszanie typów nośników

Różne specyfikacje warstwy fizycznej i ich nośniki nie powinny być traktowane jako wzajemnie się wykluczające. Jednym z głównych celów wprowadzenia rodziny standardów 802 dla sieci LAN/MAN było właśnie ułatwienie współpracy między sieciami LAN o różnych architekturach i pochodzącymi od różnych producentów. Tak więc byłoby dość dziwne, gdyby specyfikacje nie uwzględniały topologii zbudowanych z wykorzystaniem nośników mieszanych.
Dość powszechnie dopasowuje się poszczególne typy nośników do wymagań różnych obszarów funkcjonalnych sieci LAN. Historycznie rzecz biorąc, wyróżnia się trzy obszary funkcjonalne:
• przyłączalność stacji,
• przyłączalność serwera,
• połączenia między wzmacniakami.
Jak wiemy, rozwój technologii grupowania niezaawansowanych technologicznie architektur obliczeniowych przyczynił się do utworzenia czwartego obszaru funkcjonalnego sieci LAN - obszaru połączeń między grupami. W odróżnieniu od tradycyjnego połączenia serwer-wzmacniak, w tym obszarze funkcjonalnym kluczowym warunkiem sukcesu jest zmniejszenie opóźnień propagacji przy jednoczesnym zwiększeniu szerokości pasma. Dlatego do połączeń między grupami powinno się stosować nośniki o jak najmniejszym czasie propagacji i jak największej szerokości pasma.
Także pozostałe trzy obszary funkcjonalne mają własne wymagania dotyczące wydajności. Na przykład, połączenia stacji - przy założeniu, że miejsce pracy jest okablowane strukturalnie - są zwykle bardzo krótkie i wykorzystują kable miedziane. Dobrym wyborem może być tu zarówno kabel lOBase2, jak i 1OBaseT, choć kabel 1OBaseT jest zdecydowanie lepszy. Nawet względnie cienkie okablowanie lOBase2 może szybko zapełnić kanały, którymi biegną kable.
Przyłączalność serwera w dużym stopniu przypomina przyłączalność stacji, ale serwer funkcjonuje jako punkt skupienia ruchu sieciowego. Tak więc w tym obszarze funkcjonalnym szerokość pasma nigdy nie jest za duża. W tym przypadku dobrym rozwiązaniem jest kabel 1 OBaseT, a najlepszym - lOBaseFL.
Połączenia między wzmacniakami stanowią coś w rodzaju szkieletu sieci lokalnej. Połączenia takie mogą być bardzo długie. Dodatkowo szkielety sieci LAN są również naturalnymi punktami skupienia ruchu sieciowego. (Taka jest ich podstawowa funkcja!).
Dlatego konieczne jest stosowanie nośnika, który może łączyć na dużą odległość, a jednocześnie zapewnia odpowiednią szerokość pasma. Jako logiczny wybór nasuwają się tu specyfikacje światłowodowe, choć kabel 1OBaseT również może się sprawdzić (oczywiście przy

uwzględnieniu ograniczeń dotyczących maksymalnej długości tego kabla). Jeśli koniecznie chcesz wykorzystać instalacje z kablem koncentrycznym, to I OBase5 dobrze nadaje się na szkielet sieci LAN.

1.7.4 Ramka Ethernetu IEEE 802.3

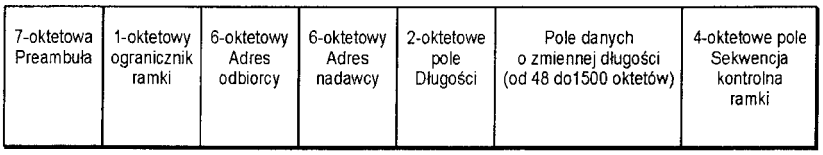
Projekt 802 zdefiniował podstawę normalizacyjną dla wszystkich rodzajów ramek ethernetowych. Minimalna długość ramki może wynosić 64 oktety, a maksymalna - 1518 oktetów, przy czym do długości wlicza się część użyteczną (dane) i wszystkie nagłówki, z wyjątkiem Preambuły i ogranicznika początku ramki. Nagłówki służą do zidentyfikowania nadawcy i odbiorcy każdego z pakietów. Jedynym ograniczeniem tej identyfikacji jest to, że adres musi być unikatowy i 6-oktetowy.

W pierwszych 12 oktetach każdej ramki zawarty jest 6-oktetowy adres docelowy (adres odbiorcy) i 6-oktetowy adres źródłowy (adres nadawcy). Adresy te są fizycznymi kodami adresowymi urządzeń, znanymi jako adresy MAC. Adres taki może być albo unikatowym adresem administrowanym globalnie, automatycznie przypisanym każdej karcie sieciowej przez jej producenta, albo adresem ustalonym podczas instalacji. Ten drugi adres znany jest także jako adres administrowany lokalnie. Adresy takie, choć potencjalnie użyteczne, były jednak wyjątkowo trudne do utrzymania. Z tego powodu już się ich nie używa.

Techniczna grupa robocza 802.3 szukała standardu, który sam w sobie byłby kompletny i nie zależałby od właściwego działania innych protokołów. Dlatego 2-oktetowe pole Typ, występujące w starszych wersjach Ethernetu, zastąpiono 2-oktetowym polem Długość. Pole to określa długość pola danych w ramce. Podstawową ramkę Ethernetu IEEE 802.3 przedstawia rysunek 7.5.

Rysunek 7.5. Podstawowa ramka Ethernetu

IEEE 802. 3.



Podstawowa ramka IEEE 802.3 często nazywana jest surową ramką ethernetową, jako że właściwie rzadko jest używana w tej postaci. Ramka ta służy jako podstawa dla dodatkowych podnagłówków, ułatwiających identyfikację protokołu protokołom wyższych warstw w komputerze odbiorcy.

W podstawowej ramce Ethernetu IEEE 802.3 tradycyjne pole Typ zastąpiono polem Długość. Jeśli to konieczne, do identyfikacji typu protokołu stosuje się, zamiast pola

Typ, podramkę 802.2. Inną zmianą, odróżniającą ramkę 802.3 od jej poprzedniczek, było wymaganie, aby długość ramki mieściła się w przedziale od 64 do 1518 oktetów, licząc od początku pola Adres Odbiorcy do końca pola Sekwencja Kontrolna Ramki.

Pewne wątpliwości dotyczą dokładnej długości ramki ethernetowej. Ich źródłem są trzy czynniki:

• różnorodność typów ramek ethernetowych, • zmienna długość pola danych,

• niezgodność co do tego, jak traktować Preambułę i Ogranicznik Początku Ramki.

Jak pamiętamy z rozdziału 5, istnieje pięć różnych typów ramek ethernetowych: PARC Ethernet, DIX Ethernet, podstawowa ramka 802.3, 802.3 z podramką LLC i 802.3 z podramką SNAP. W ramach tych pięciu odmian reprezentowane są trzy różne zestawy minimalnych i maksymalnych rozmiarów ramki.

Uwagę zwraca celowe unikanie podawania określonych rozmiarów ramki. Przy zmiennej długości pola danych najlepszym rozwiązaniem jest określanie minimalnej i maksymalnej dozwolonej wielkości ramki. Faktyczny rozmiar ramki zawarty jest w tym przedziale wartości.

Ostatnia wątpliwość dotyczy Preambuły i Ogranicznika początku ramki, czyli ogranicznika SFD (ang.. *Start of Frame Delimiter).* Preambuła jest 7-oktetowym ciągiem znaków poprzedzającym każdą ramkę i umożliwiającym synchronizację transmisji. Jest ona nieodłącznym dodatkiem do ramki, ale rzadko uważa się ją za składnik tej ostatniej.

Bezpośrednio po Preambule następuje 1-oktetowy Ogranicznik początku ramki (SFD). Czasem uważa się go za integralną część Preambuły, a nie za odrębny element ramki. Nie trzeba długo szukać, by znaleźć źródła, określające ogranicznik SFD jako oficjalny początek ramki.

By uniknąć nieporozumień, wszystkie rozmiary ramek podawane w tym rozdziale są opatrzone komentarzem opisującym początkowe i końcowe pola ramek.

1.7.5 Struktura ramki Ethernet LLC

Ramka Ethernet LLC jest kombinacją ramki 802.3 i podramki 802.2 LLC. W tej implementacji ramka LLC dodaje trzy pola do podstawowej ramki Ethernetu: Punkt dostępu usługi docelowej, Punkt dostępu usługi źródłowej i pole kontroli.

Ponieważ w ramce 802.3 zrezygnowano z pola Typ, a dodano pole Długość, nie można było określić, który protokół został osadzony w części użytecznej. Dlatego ramka odebrana przez komputer obsługujący wiele protokołów komunikacyjnych mogłaby zostać skierowana do niewłaściwego protokołu! Innymi słowy, ramka 802.3 mogła dotrzeć do miejsca przeznaczenia, ale nie było gwarancji, że zostanie dostarczona do protokołów wyższej warstwy w komputerze docelowym.

Aby wyjaśnić tę sprawę, IEEE opracowało strukturę i standard podramki 802.2. Ta nowa ramka bywa określana jako podramka 802.2 lub ramka LLC. Rysunek 7.6 pokazuje, jak wygląda ta nowa ramka osadzona w ramce Ethernet 802.3. Ma ona następującą strukturę:

• 7-oktetowa Preambuła, sygnalizująca początek ramki

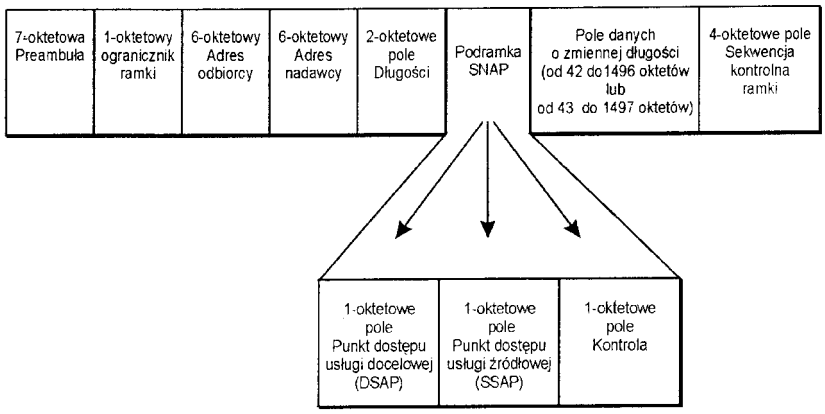
• 1-oktetowy Ogranicznik początku ramki, wskazujący że zaczyna się właściwa zawartość ramki

• 6-oktetowy adres MAC odbiorcy

• 6-oktetowy adres MAC nadawcy

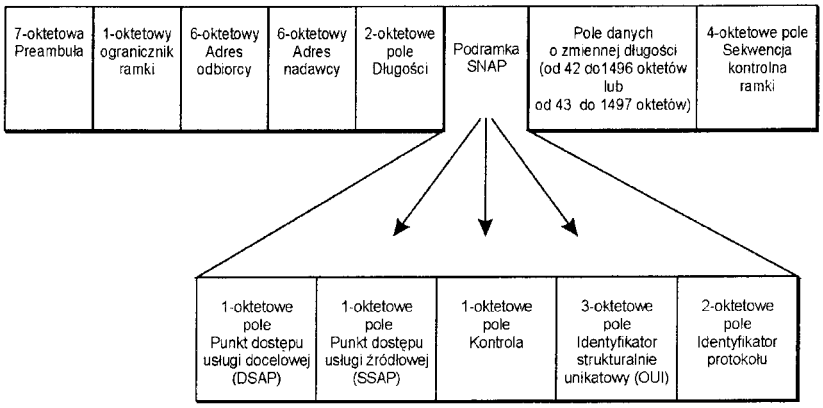
• 2-oktetowe pole Długość, określające całkowitą długość pola danych, wliczając także nagłówki LLC i SNAP

• 1-oktetowe pole Punkt dostępu usługi docelowej (ang.. *DSAP - Destination Service Access Point),* określające przewidziany punkt dostępu do usług LLC w komputerze odbiorcy
• 1-oktetowe pole Punkt dostępu usługi źródłowej (ang.. SSAP - Source Service Access Point), określające punkt dostępu do usług LLC w komputerze nadawcy • 1- lub 2-oktetowe pole Kontrola, wskazujące typ przenoszonej ramki LLC
• Pole danych zawierające albo od 42 do 1496, albo od 43 do 1497 oktetów danych, w zależności od długości pola Kontrola
• 4-oktetowe pole Sekwencja kontrolna ramki, używane do sprawdzania integralności ramki
Aby prawidłowo funkcjonował mechanizm CSMA/CD, całkowita długość ramki Ethernet LLC musi wynosić przynajmniej 64 oktety (pomijając Preambułę i SFD). By zapewnić to minimum, w razie potrzeby na końcu pola danych dodaje się zera. Górna granica długości ramki wynosi 1518 oktetów (nie licząc Preambuły i SFD).
Rysunek 7.6. Ramka Ethernet LLC'.



1.7.6 Struktura ramki Ethernet SNAP

Po wprowadzeniu podramki LLC zaczęto zastanawiać się nad adekwatnością tej struktury. Chodziło o to, że nie mogła ona identyfikować wszystkich protokołów wyższej warstwy, jakie ewentualnie mogłyby być potrzebne.
IEEE wróciło do pracy i wprowadziło podramkę, tzw.. protokół dostępu podsieci (ang.. SNAP - Sub-Network Access Protocol). SNAP wprowadza dodatkowe, S-oktetowe pole identyfikacji protokołu. W ramce pole to zajmuje miejsce za nagłówkiem LLC. Składa się z 3-oktetowego pola identyfikatora strukturalnie unikatowego (ang.. OUI - Organizationally Unique Identifier) i 2-oktetowego pola Typ. Pola te określają, dla którego protokołu wyższej warstwy w komputerze odbiorcy dana ramka jest przeznaczona.
Ramka Ethernet SNAP jest przedstawiona na rysunku 7.7. Zawiera ona następujące pola: • 7-oktetowa Preambuła, sygnalizująca początek ramki
• 1-oktetowy Ogranicznik początku ramki, wskazujący, że zaczyna się właściwa zawartość ramki
• 6-oktetowy adres MAC odbiorcy • 6-oktetowy adres MAC nadawcy
• 2-oktetowe pole Długość, określające całkowitą długość pola danych, z nagłówkami LLC i SNAP
• 1-oktetowe pole Punkt dostępu usługi docelowej (DSAP), określające przewidziany punkt dostępu do usług LLC w komputerze odbiorcy
• 1-oktetowe pole Punkt dostępu usługi źródłowej (SSAP), określające punkt dostępu do usług LLC w komputerze nadawcy
• 1- lub 2-oktetowe pole Kontrola, wskazujące typ transmitowanej ramki LLC
• 5-oktetową podramkę SNAP, zawierającą 3-oktetowe pole Identyfikator strukturalnie unikatowy i 2-oktetowe pole Typ protokołu, identyfikujące przenoszony protokół wyższej warstwy
• Pole danych zawierające albo od 37 do 1491, albo od 38 do 1492 oktetów danych, w zależności od długości pola Kontrola
• 4-oktetowe pole Sekwencja kontrolna ramki, używane do sprawdzania integralności ramki
Ramka Ethernet SNAP integruje struktury lub nagłówki podramki 802.2 i umożliwia identyfikację protokołów wyższego poziomu, dla których przeznaczona jest zawartość ramki. Zapewnia to wsteczną kompatybilność z wcześniejszymi wersjami Ethernetu, których ramki zawierały oddzielne mechanizmy identyfikacji protokołu.
Aby mechanizm CSMA/CD mógł działać prawidłowo, całkowita długość ramki Ethernet SNAP musi wynosić przynajmniej 64 oktety. Górną granicą długości ramki jest 1518 oktetów (wliczając Preambułę i SFD).
Rysunek 7.7. .Ramka Ethernet SNAP P.
Więcej informacji o różnych strukturach ramek dla Ethernetu możesz znaleźć w rozdziale 5.



1.7.7 Prognozowanie opóźnień

Ethernetowy protokół CSMA/CD wymaga, by całkowity czas transmisji i potwierdzenie przyjęcia wynosił co najwyżej 51,2 mikrosekundy, przy prędkości przesyłania sygnału wynoszącej 10 Mbps. Każdy element składowy sieci, wliczając w to media transmisyjne i urządzenia fizyczne, ma własną charakterystykę opóźnienia propagacji. Dlatego dobrze jest obliczyć całkowite opóźnienie dla sieci Ethernet, zanim się ją zbuduje.

1.7.7.1 Szacowanie opóźnień propagacji

Poniższe tabele zawierają szacunkowe wartości opóźnień propagacji dla wymienionych urządzeń i nośników. Tabela 73 przedstawia szacunkowe opóźnienia dla urządzeń różnych typów.
Tabela 7.3.
Szacunkowe opóź'nienie w zależności od rodzaju urządzenia.

Urządzenie	Szacunkowe opóźnienie (w mikrosekundach)
Wzmacniak światłowodowy	1,55
Wzmacniak wieloportowy	0,1
Nad-biornik światłowodowy	0,20
Nad-biornik dla skrętki dwużyłowej	0,27

Wymienione wartości opóźnień propagacji dla różnych rodzajów urządzeń są jedynie wartościami szacunkowymi! Na wydajność może wpływać wiele czynników, jak np. temperatura, wilgotność, a nawet wiek danego urządzenia elektronicznego i/lub to, od jakiego producenta ono pochodzi. Dlatego nie można precyzyjnie przewidzieć wartości opóźnienia dla danego rodzaju urządzeń. Nie można również ustalić jednej wartości, która obowiązywałaby przez cały cykl życia urządzenia.
Tabela 7.4 przedstawia szacunkowe opóźnienia przypadające na każdy metr długości powszechnie stosowanych nośników.
Tabela 7.4.
S=achnkowe opóźnienia w zależ ności od rodzaju nośnika.

1.7.7.2 Prognozowanie opóźnień Ethernetu

Rodzaj nośnika	Opóźnienie (w mikrosekundach na metr)
10Base2	0,00514
10Base5	0,00433
Nieekranowana skrętka dwużyłowa (UTP)	0,0057
Ekranowana skrętka dwużyłowa (STP)	0,0057
Światłowód	0,005

Prognozowanie opóźnień dla sieci LAN jest procesem bardzo prostym. Po pierwsze, należy zaplanować rozmieszczenie urządzeń sieciowych, uwzględniając długość połączeń między urządzeniami i rodzaj nośnika użytego dla każdego z połączeń. Następnie trzeba policzyć, ile urządzeń danego rodzaju ma działać w sieci i pomnożyć liczbę urządzeń przez opóźnienie przewidziane dla urządzeń tego rodzaju. Proces ten należy powtórzyć dla wszystkich rodzajów urządzeń, a wyniki zsumować. W ten sposób szacuje się prognozowaną wielkość opóźnień sprzętowych w sieci.

Liczbę otrzymaną powyżej trzeba zwiększyć o prognozę opóźnień wprowadzanych przez kable. Obliczenie jej jest równie łatwe. Dla każdego połączenia należy pomnożyć jego długość (w metrach) przez opóźnienie (w mikrosekundach) przewidywane dla danego nośnika, a następnie zsumować wyniki uzyskane dla wszystkich połączeń. W ten sposób otrzymuje się pełną prognozę opóźnień wprowadzanych przez kable. Suma opóźnień dla kabli i urządzeń daje całkowite opóźnienie dla sieci. Aby protokół CSMA/CD działał prawidłowo, wartość tego opóźnienia musi być mniejsza niż 51,2 mikrosekundy. Jeśli jest większa lub równa tej wartości, prawdopodobnie wystąpią problemy w działaniu sieci. Jeśli całkowite opóźnienie jest bliskie, ale wciąż mniejsze niż 51,2 mikrosekundy, można się spodziewać trudności, które będą się pojawiać w miarę starzenia się infrastruktury sieci.

Należy unikać pokusy szukania granic możliwości operacyjnych sieci. Nie jest to gra warta świeczki, zwłaszcza w przypadku sieci wspomagających procesy gospodarcze, a nie zabawy heurystyczne.

1.7.8 Podsumowanie

Ethernet jest bogatym i różnorodnym zbiorem technologii. Sieci Ethernet mogą pracować w paśmie podstawowym lub mogą być szerokopasmowe, pełnodupleksowe lub półdupleksowe. Mogą wykorzystywać jeden z pięciu różnych nośników (lub więcej niż jeden, jeśli odważysz się wyjść poza uświęcone standardy!) i pracować z prędkościami z zakresu od 10 Mbps do I Gbps.

W rozdziale tym dokonaliśmy przeglądu popularniejszych specyfikacji Ethernetu. Przegląd ten obejmuje różne konwencje ramek, typowe komponenty sprzętowe, typy nośników, ograniczenia protokołów, sposób prognozowania opóźnień, a nawet schemat różnicowego przesyłania sygnałów, wykorzystywany przy transmisji przez nieekranowaną skrętkę dwużyłową.

1.8 Rozdział 8 Szybsze sieci Ethernet

Mark A. Sportack

Wszechobecny Ethernet na tyle zadowalająco spełniał wymagania stawiane sieciom, że IEEE, jego „opiekun", regularnie go udoskonalał i aktualizował. Dwie najważniejsze aktualizacje dotyczyły zwiększenia prędkości przesyłania sygnałów.

W pierwszej z tych szybszych sieci Ethernet prędkość wynoszącą pierwotnie l0 Mbps podniesiono do wartości 100 Mbps. Wymagało to opracowania całkowicie nowej warstwy fizycznej i wprowadzenia niewielkich zmian w warstwie łącza danych, która musiała zostać dopasowana do nowej warstwy fizycznej. Wydaje się, że ten nowy standard, nazwany Fast Ethernet, pobił sieci ATM (ang.. Asynchronous Transfer Mocie - Tryb Transferu Asynchronicznego) na rynku szybkich sieci LAN.

Druga z tych aktualizacji jest nieco bardziej radykalna. Większość specyfikacji warstwy 2 Ethernetu przeszczepiono na kompletnie inną warstwę fizyczną, zapożyczoną z technologii kanału światłowodowego 1 Gbps. Ta modyfikacja, ciągle opracowywana przez komisję roboczą IEEE 802.3z, odmieni Ethernet CSMA/CD w wielu aspektach. Jedna z najbardziej interesujących propozycji zmian zakłada rezygnację ze znanej metody dostępu do nośnika CSMA/CD.

W tym rozdziale zostaną przedstawione różnice pomiędzy Fast Ethernetem, Gigabit Ethernetem i tradycyjnym, półdupleksowym Ethernetem CSMA/CD, 10 Mbps, który zrodził powyższe warianty.

1.8.1 Fast Ethernet

We wczesnych latach dziewięćdziesiątych publikacje branżowe pełne były entuzjastycznych recenzji nowego, rewolucyjnego protokołu LAN: ATM (trybu transferu asynchronicznego). Pierwotnie miał to być protokół dla sieci rozległych (WAN), przeznaczony do obsługi ruchu między centralami telefonicznymi. Wyobrażano sobie, że ATM będzie wielkim unifikatorem. Mógł zintegrować sieci LAN i WAN tak, że różnice między nimi stałyby się akademickie. Wszystkie inne technologie LAN, łącznie z Ethernetem, były skazane na odejście do lamusa.

Centrale telefoniczne są budynkami, w których zbiegają się i są podłączone do wielkich rozdzielni telefonicznych miedziane linie telefoniczne idące z mieszkań i biur. Każda centrala jest połączona z inną lub wieloma innymi centralami. Razem cała ta sieć tworzy przemysłową strukturę telefonii.

Ta wielka i wspaniała wizja została zniweczona przez liczne przyziemne czynniki, z których najbardziej istotnym było powolne tempo rozwoju standardów (i produktów) ATM LAN oraz fakt, że Ethernet nie chciał się poddać. Ethernet, zmęczony, stary, oparty na zasadzie rywalizacji protokół, czekał cicho w ukryciu, aż przemysłowe konsorcjum konkurentów zapadło się pod swoją własną masą i pogrążyło w morzu konfliktów. Równie groźnym wyzwaniem stojącym przed ATM Forum były trudności z zapewnieniem solidnej kompatybilności wstecznej z istniejącymi, bardzo niepodobnymi infrastrukturami sieci LAN. Z początku producenci LAN zaczęli po cichu mówić o Ethernecie 100 Mbps jako o pośredniej alternatywie dla sieci lokalnych ATM.

Rozwiązanie pośrednie było niezbędne, gdyż istniejące sieci lokalne wykazywały oznaki starzenia w stosunku obsługiwanych przez nie procesorów i aplikacji. ATM nadal uważany był za rozwiązanie docelowe, ale polityka wewnątrz ATM Forum, jak też praktyczne ograniczenia, spowalniały jego rozwój. W międzyczasie klienci głośno domagali się wydajniejszej technologii LAN.

Podczas gdy świat czekał na ATM, Ethernet, co zrozumiałe, mógł zostać szybko odświeżony. Protokoły warstwy łącza danych mogły zostać zachowane, a prędkość przesyłania sygnału zwiększyłaby się o rząd wielkości. Należało tylko zmodyfikować warstwę fizyczną - konieczne było wprowadzenie szeregu nowych interfejsów fizycznych, dostosowanych do zwiększonej częstotliwości taktowania. Wydawało się to dość łatwe. Nowy Ethernet byłby dla administratorów sieci lokalnych rozwiązaniem alternatywnym wobec migracji do ATM-u.

Zgłoszono wiele propozycji, ogólnie znanych pod nazwą Fast Ethernet. Gdy koncept został wspólnie przemyślany, rozmaite propozycje zebrano w dwóch rywalizujących grupach. Obydwie zostały przez IEEE przyjęte do rodziny standardów 802. Jedna z nich, znana dziś jako Fast Ethernet, jest po prostu tradycyjnym protokołem CSMA/CD 802.3 ze zwiększoną o rząd wielkości prędkością sygnału. Fast Ethernet został znormalizowany jako rozszerzenie istniejącego standardu 802.3.

Druga, konkurencyjna propozycja otrzymała nazwę VG-AnyLAN i stała się standardem 802.12. Technologia ta, choć bardziej zaawansowana technicznie (ma możliwość dostosowania do transmisji izochronicznej), pozostała nieco mniej eleganckim odejściem od

Ethernetu 10 Mbps. Z tego i wielu innych względów, w tym z powodu braku aplikacji koniecznie wymagających transmisji izochronicznej, sieć VG-AnyLAN nie zdobyła znaczącej pozycji na rynku.

1.8.1.1 Nośniki Fast Ethernetu

Rozszerzenia standardu 802.3 (do 100 Mbps) obejmują trzy różne interfejsy międzynośnikowe (MDI):
• 100BaseTX - określa oryginalną specyfikację IOOBaseX dla Kategorii 5 nieekranowanej skrętki dwużyłowej (UTP) i dla ekranowanej skrętki dwużyłowej (STP) Typu I.
• 100BaseFX - określa Ethernet 100 Mbps z okablowaniem światłowodowym. • 100BaseT4 - opisuje Ethernet 100 Mbps z okablowaniem UTP Kategorii 3,4 i 5.
Jak dowodzą konwencje nazewnicze zastosowane w przypadku tych trzech interfejsów, Fast Ethernet zaadoptował dla tych interfejsów własną formę skróconej notacji. Jeśli nie znasz kodu, skróty te mogą wydawać się bardzo zagmatwane. Na szczęście kod jest względnie prosty i w pewien sposób mnemoniczny. Pierwsza liczba, w tym wypadku 100, określa szybkość przepływu danych dla danej specyfikacji. Potem następuje słowo: „Base" lub „Broad". „Base" dotyczy transmisji w paśmie podstawowym, natomiast „Broad" opisuje transmisję w technologii szerokopasmowej. Ostatnie znaki mogą być znakami alfabetycznymi, numerycznymi i/lub specjalnymi. Są one najmniej mnemoniczną częścią konwencji nazewniczej i określają fizyczne medium transmisyjne wykorzystywane w danej specyfikacji.
Termin „IOOBaseX" jest stosowany zarówno w .odniesieniu do 100BaseTX, jak i 1OOBaseFX
. Podobnie jest w przypadku interfejsów dla skrętki dwużyłowej, 100BaseTX i 1OOBaseT4, nazywanych czasami „100BaseT". Należy jednak podkreślić, że 100BaseX i 100BaseT nie są fizycznymi interfejsami! Są one ogólnymi nazwami grupy podobnych interfe j sów.
Dalsze nieporozumienia pociąga za sobą niefortunna konwencja nazewnicza, przyjęta przez komitet roboczy 802.3 dla schematów sygnalizacyjnych wykorzystywanych przez interfejsy Fast Ethernetu. Schematami tymi są 100Base4T+ i 100BaseX. Tak, właśnie 100BaseX. Tego samego terminu użyto dla opisania schematu sygnalizacyjnego oraz dwóch interfejsów fizycznych.
Aby uniknąć nieporozumień, w dalszej części tego rozdziału nazwy poszczególnych interfejsów są przytaczane w pełnym brzmieniu. Terminy „100BaseT" i „100BaseX" są zarezerwowane dla przypadków, kiedy używa się ich w odniesieniu do obydwu (opisywanych przez te nazwy) fizycznych wariantów jednocześnie. Oprócz tego, gdy termin „100BaseX" pojawia się w kontekście schematu sygnalizacyjnego, jest z nim właśnie utożsam iany.

1.8.1.1.1 100BaseTX

Pierwsza klasyfikacja nośnika dla Fast Ethernetu nosi nazwę 100BaseTX. Obejmuje ona kable ekranowanej skrętki dwużyłowej (STP) Kategorii I i nieekranowanej skrętki dwużyłowej (UTP) Kategorii 5. Druga klasyfikacja, 100BaseFX, dotyczy światłowodu, a trzecia, 100BaseT4, Ethernetu 100 Mbps z kablami UTP Kategorii 3, 4 i 5. Wprowadzono również zmiany w warstwie sterowania dostępem do nośnika (MAC), aby mechanizmy tej warstwy, pierwotnie przeznaczone dla sieci 10 Mbps, mogły pracować przy prędkości 100 Mbps.
Ponieważ standard ten jest rozszerzeniem specyfikacji Ethernetu IEEE 802.3, włożono wiele wysiłku, aby produkt ten w bardzo dużym stopniu przypominał 1 OBaseT. Przykładowo, 1OBaseT może wykorzystywać dwie pary przewodów UTP Kategorii 3. Ważne więc jest, aby Fast Ethernet również mógł korzystać z dwóch par przewodów, tak aby przejście na ten standard nie pociągało za sobą konieczności wymiany okablowania stanowisk. Istotnie, Fast Ethernet może wykorzystywać dwie pary przewodów, ale muszą to być przewody Kategorii 5, a nie Kategorii 3. Wydaje się, że redukuje to zalety tego rozwiązania, gdyż nie wszystkie stanowiska okablowane przewodami UTP Kategorii 3 posiadają pełną instalację z czterema parami przewodów. Dlatego wielu administratorów sieci lokalnych staje przed problemem wymiany okablowania, pomimo że dostępne są technologie sieci LAN 100 Mbps, obsługujące transmisję przewodami Kategorii 3.
Warto zauważyć, że choć specyfikacja IEEE 802.3 definiuje liczne interfejsy fizyczne dla Ethernetu 10 Mbps, to Fast Ethernet najbardziej przypomina 1OBaseT.

1.8.1.1.2 100BaseFX

100BaseFX jest światłowodowym odpowiednikiem 100BaseTX. Mają one wspólny schemat sygnalizacyjny i technikę kodowania danych, ale wykorzystują różne nośniki fizyczne. 100BaseFX może obsługiwać transmisję danych z szybkością 100 Mbps na odległość do 400 metrów, wykorzystując dwie żyły kabla światłowodowego o średnicy 62,5/125 mikronów. Ogromnie rozszerza to zasięg sieci Fast Ethernet, a najlepszym zastosowaniem 100BaseFX jest łączenie ze sobą wzmacniaków.

1.8.1.1.3 100BaseT4

100BaseT4 umożliwia transmisję danych z szybkością 100 Mbps przez cztery pary przewodów telefonicznych na odległość do 100 metrów. Przewody telefoniczne muszą odpowiadać co najmniej Kategorii 3 UTP. Możliwe jest także przeprowadzanie transmisji z wykorzystaniem UTP Kategorii 4 i 5.
4T+ nie obsługuje wiązek okablowania poziomego Kategorii 3, zawierających 25 par przewodów. Ten rodzaj kabla był dość powszechnie stosowany w systemach okablowania wielu starszych budynków biurowych.
Jedną z najważniejszych różnic funkcjonalnych między IOOBaseT4 a jego „rodzeństwem" jest to, że specyfikacja ta nie obsługuje sygnału ciągłego występującego między ramkami. Ten sygnał ciągły znany jest jako odstęp między ramkami. Zwykle urządzenie nadawcze wykorzystuje sygnał ciągły, aby utrzymać uprawnienie do transmisji. Inne urządzenia interpretują ten nic nie znaczący łańcuch 96 bitów jako sygnał zajętości linii.
Jeśli chcą nadawać, ale „widzą" ten łańcuch, uruchamiają swoje binarne wykładnicze algorytmy zwrotne i czekają. Dlatego specyfikacja, która nie obsługuje sygnału ciągłego, ma mniejsze wymagania co do poboru mocy. Jest to szczególnie ważne w przypadku laptopów, a także w innych warunkach, w których zużycie mocy jest czynnikiem istotnym.

1.8.1.2 Schematy sygnalizacyjne

Fast Ethernet używa dla swoich interfejsów skrętki dwużyłowej dwóch różnych schematów sygnalizacyjnych. Obydwa schematy, dość nieszczęśliwie nazwane „100BaseX" i „100Base4T+", obsługują transmisję danych z szybkością 100 Mbps. Szybkość taką można uzyskać przy maksymalnej odległości między koncentratorem a stacją roboczą wynoszącej 100 metrów i odległości między koncentratorami wynoszącej 10 metrów. Zastosowanie światłowodu pozwala zwiększyć odległość między koncentratorami.
Termin „szybkość transmisji danych" nie powinien być utożsamiany z prędkością przesyłania sygnału! Fast Ethernet, o którym powszechnie mówi się, że obsługuje prędkość 100 Mbps, w rzeczywistości przesyła sygnał z prędkością 125 Mbps. Odejmując narzuty warstwy fizycznej, w tym technikę kodowania, otrzymujemy szybkość transmisji danych wynoszącą 100 Mbps. Tak samo jest w przypadku Ethernetu 10 Mbps: przesyła on sygnały z prędkością 12,5 Mbps, ale sieciowa szybkość transmisji danych wynosi 10 Mbps.

1.8.1.2.1 100Base4T+

Schemat sygnalizacyjny 100Base4T+ pozwala fizycznym interfejsom 100BaseT korzystać z większości istniejącego okablowania Kategorii 3. Zastrzeżenie „większość" jest konieczne, gdyż ten schemat sygnalizacyjny wymaga czterech par przewodów. Trzy pary służą do przenoszenia danych w obydwu kierunkach, w trybie półdupleksowym, podczas gdy czwarta para pracuje wyłącznie w trybie odbioru i służy do wykrywania kolizji.
Jeśli czwarta para wykryje sygnał na drugim końcu kabla, informuje protokoły warstwy fizycznej, że w tej chwili nie powinny wysyłać danych. Pozostałe trzy pary służą do nadawania i odbioru danych, lecz nigdy nie robią tych dwóch rzeczy jednocześnie.
Każda z tych par może obsługiwać transmisję danych z szybkością 33,33 Mbps. Razem daje to użyteczną szerokość pasma równą 100 Mbps. (Niech matematycy nie narzekają: 3*33,33 jest wystarczająco bliskie 100).
Trójprzewodowe ścieżki są konieczne ze względu na schemat dekodujący 8B6T wykorzystywany przez 100Base4T+. Istota jego działania polega na tym, że podwarstwa MAC przekazuje oktety danych binarnych warstwie fizycznej. Warstwa fizyczna konwertuje każdy oktet czy też 8 bitów (to jest właśnie część 8B nazwy 8B6T) na 6 znaków trójkowych. Każda grupa 6 znaków trójkowych jest nazywana grupa 6T. Grupy 6T są następnie rozdzielane na trzy dostępne kanały szeregowe - czyli trzy pary przewodów służące do nadawania i odbioru danych.
Dla czytelników „zakochanych" w szczegółach technicznych: każdy znak trójkowy (z każdego oktetu powstaje ich 6) jest przesyłany w ciągu około 40 nanosekund, jeśli prędkość sygnału wynosi 125 Mbps.

1.8.1.2.2 100BaseX

100BaseX został zaadaptowany do pracy z Ethernetem CSMA/CD z pełnodupleksowego mechanizmu sygnalizacyjnego FDDI. Mechanizm FDDI obsługuje dwa interfejsy fizyczne: kabel UTP Kategorii 5 i wielofunkcyjny światłowód 62.5/125 mikrona. Dlatego IOOBaseX obsługuje te same rodzaje kabli i ma takie same ograniczenia dotyczące odległości jak FDDI.
Interfejs FDDI nazywany jest także interfejsem skrętki dwużyłowej zależnym od nośnika warstwy fizycznej (ang.. TP-PMD - Twisted Pair Physical Medium Dependent Interface). Wykorzystuje on dwie pary skrętki UTP Kategorii 5 do obsługi pełnodupleksowej komunikacji z szybkością 100 Mbps na odległość do 100 metrów. Jeśli wykorzystywany jest kabel z czterema parami przewodów, pozostałe dwie pary mogą teoretycznie służyć do komunikacji telefonicznej, ale nie mogą być używane przez inną szybką sieć LAN.
IOOBaseX obsługuje także interfejs FDDI F-PMD (ang.. Fiber Medium Dependent), czyli wielofunkcyjny kabel światłowodowy 62,5/125 mikrona. Ten interfejs zapewnia pełnodupleksową transmisję danych z szybkością 100 Mbps na odległość do 400 metrów.
Schemat sygnalizacyjny 100BaseX wykorzystuje schemat kodowania 4B/SB. Przyjmuje on półbajty (ang.. *nibbles),* czyli 4 bity lub pół oktetu danych z podwarstwy MAC i koduje je w 5-bitowe znaki przeznaczone do transmisji. Znaki te są znane jako grupy kodowe. Grupy kodowe są właściwe wyłącznie mechanizmom IOOBaseX i nie mają znaczenia poza tym kontekstem.
5-bitowe pole binarne ma 32 możliwe kody. Kody te reprezentują 16 grup kodowych szesnastkowego zestawu znaków (od 0 do F). Dodatkowo 4 grupy kodowe są wykorzystywane jako mechanizmy kontrolne transmisji. Pozostałych 12 możliwych kodów wykorzystuje się jako wypełniacz między strumieniami danych, wypełnienie strumienia danych lub pozostają one niezdefiniowane.

1.8.1.3 Maksymalna średnica sieci

Współoperacyjność koncentratora osiąga się dzięki wzmacniakowi. Specyfikacje dla 100BaseT obejmują uniwersalny wzmacniak obsługujący obydwie warstwy fizyczne. W sieci IOOBaseT można używać maksymalnie dwóch uniwersalnych wzmacniaków, oddalonych od siebie o nie więcej niż 10 metrów.
Segmentacja z wykorzystaniem routerów może również zwiększyć potencjalną średnicę sieci, pozwalając korzystać z więcej niż dwóch wzmacniaków.

1.8.1.4 Podsumowanie sieci Fast Ethernet

Fast Ethernet jest rozszerzeniem specyfikacji IEEE 802.3 do 100 Mbps. Właściwie jest bardzo podobny do Ethernetu 1 OBaseT, ale działa o rząd wielkości szybciej. Zwiększona szybkość transmisji danych wymusiła znaczne zmiany w warstwie dostępu do nośnika.
Fast Ethernet szybko zadomowił się w środowisku sieci lokalnych. Wielu producentów wspomogło ten proces, oferując karty sieciowe (NIC) obsługujące dwie szybkości transmisji: 10 i 100 Mbps. Takie karty albo mogą automatycznie wybierać optymalną prędkość, uwzględniając typ okablowania i odległość od koncentratora, lub też prędkość może być wybierana ręcznie.

1.8.2 Gigabit Ethernet

Nowa propozycja zwiększenia prędkości przesyłania sygnałów w sieci Ethernet znana jest jako Gigabit Ethernet. Standard Gigabit Ethernet został opracowany przez IEEE pod auspicjami podgrupy zadaniowej 802.3z. Produkty, które pojawiły się wcześniej, opierają się na szkicach roboczych proponowanej specyfikacji. W jakim stopniu te produkty będą odpowiadać przyjętemu standardowi - to okaże się w przyszłości.
Ponieważ standard ten znajduje się „pod parasolem" specyfikacji 802.3, czynione są wszelkie starania, aby byk jak najbardziej kompatybilny ze swoim wolniejszym „rodzeństwem". Dla przykładu: Gigabit Ethernet dalej używa tego samego protokołu CSMA/CD oraz formatu i rozmiaru ramki co inne sieci Ethernet. Takie rozwiązania umożliwiają kompatybilność w tył z Ethernetem 10 i w przód z Ethernetem 100 Mbps.
Gigabit Ethernet ma początkowo służyć jako szkielet łączący ze sobą przełączniki 10/1OOBaseT. Nieco dalej idą propozycje, aby łączył wysokowydajne serwery z siecią LAN. Przewiduje się nawet, że Gigabit Ethernet mógłby ewentualnie służyć do łączenia stacji roboczych za pomocą kabli UTP Kategorii 5 o długości do 100 m.

1.8.2.1 Interfejsy fizyczne

W celu przyspieszenia prac nad standardem zespół 802.3z postanowił wykorzystać wiele elementów specyfikacji ANSI warstwy fizycznej kanału światłowodowego. Kanał światłowodowy, oryginalnie opracowany jako technologia nowej generacji kanałów mainframe, został znormalizowany i przekształcony w technologię LAN. Jak dotąd rynek słabo zareagował na światłowodowe sieci lokalne.
Pomimo braku sukcesu rynkowego, kanał światłowodowy stanowił sprawdzoną i perspektywiczną warstwę fizyczną, mogącą być podstawą dalszych prac zespołu 802.3z. Ponieważ intencją IEEE jest pozyskanie jak najszerszego poparcia dla swoich standardów, każdy standard IEEE jest przedstawiany instytutowi ANSI, aby ten rozważył możliwość uczynienia go standardem narodowym (czyli standardem o zakresie szerszym od przemysłowego). Tak więc sensowne jest objęcie standardem elementów warstwy fizycznej już wcześniej ratyfikowanych przez ANSI.
Rynkowe niepowodzenia kanału światłowodowego są raczej wynikiem słabości jego warstwy łącza danych oraz marketingu w stosunku do istniejących technologii LAN. Specyfikacja warstwy fizycznej tej technologii dała inżynierom IEEE okazję do nadania rozpędu ich pracy nad rozwojem Gigabit Ethernetu.
Jednym z pomniejszych problemów jest prędkość przesyłania sygnału. W kanale światłowodowym, gdzie sygnał osiąga prędkość 1,063 Gbps, wykorzystuje się schemat kodowania, który generuje dwa dodatkowe bity na każdy przesyłany oktet. Stosowanie tego schematu (znanego jako „8B/1 0B."), podobnie jak zbliżonego do niego schematu 4B/5B, zmniejsza użyteczną szerokość pasma do wartości poniżej 1 Gbps. W rzeczywistości wartość ta wynosi około 800 Mbps. Wciąż jest to znaczący wzrost w porównaniu z większością istniejących technologii LAN, ale nie ma tu psychologicznego wpływu, jaki miałoby złamanie bariery 1 Gbps. Oczekuje się, że za jakiś czas technologie transmisji sygnałów umożliwią transmisję z prędkością 1,25 Gbps. Przyjmując, że wciąż używany będzie schemat kodowania 8B/IOB, szerokość pasma sieci wyniesie I Gbps.
Gigabit Ethernet pozwoli wybierać między czterema nośnikami, z których każdy ma własną specyfikację interfejsu fizycznego. Są to:
• miedziany kabel koncentryczny,
• wielofunkcyjny kabel światłowodowy,
• jednomodowy kabel światłowodowy o średnicy 8,3/125 mikrona, • nieekranowana skrętka dwużyłowa (UTP) Kategorii 5.
Interfejs wielofunkcyjnego kabla światłowodowego w rzeczywistości obsługuje dwa różne rodzaje kabla: kabel zgodny z konwencją północnoamerykańską o średnicy 62,5 mikrona i kabel zgodny z konwencją ogólnoświatową o średnicy 50 mikronów.
Przyjęcie technologii sygnalizacyjnych warstwy fizycznej dla standardu kanału światłowodowego ANSI ułatwiło prace normalizacyjne, dotyczące wymienionych nośników. Kanał światłowodowy był podstawą wszystkich interfejsów warstwy fizycznej, z wyjątkiem UTP Kategorii 5. Interfejsy te są opisane w następnych czterech podpunktach. Określenie maksymalnej długości połączeń dla każdego interfejsu międzynośnikowego miało na celu zapewnienie odpowiedniej wydajności większości istniejących instalacji kablowych przy założeniu najbardziej niekorzystnej realizacji.

1.8.2.1.1 1000BaseSX

1000BaseSX to proponowana przez IEEE 802.3z specyfikacja wielofunkcyjnej transmisji, wykorzystującej lasery krótkofalowe. Lasery krótkofalowe to takie lasery, które wytwarzają światło o długości fali 850 nanometrów.
1000BaseSX może wykorzystywać dwa różne nośniki: kable światłowodowe o średnicy 50 mikronów i 62,5 mikrona. Kabel o średnicy 50 mikronów może przesyłać sygnał z prędkością 1 Gbps na odległość do 550 metrów. Długość jednego segmentu kabla o średnicy 62,5 mikrona nie może przekroczyć 260 metrów.

1.8.2.1.2 1000BaseLX

1000BaseLX jest proponowaną specyfikacją transmisji wykorzystującej lasery długofalowe. Przesyłane fale świetlne o długości 1300 nanometrów są uważane za fale długie.
1000BaseLX może wykorzystywać trzy różne media transmisyjne: • wielofunkcyjny kabel światłowodowy o średnicy 62,5 mikrona, • wielofunkcyjny kabel światłowodowy o średnicy 50 mikronów, • jednofunkcyjny kabel światłowodowy o średnicy 8,3 mikrona.
W swojej aktualnej postaci wielofunkcyjny kabel o średnicy 62,5 mikrona zapewnia połączenie na odległość do 440 metrów. Kabel o średnicy 50 mikronów może mieć długość nie większą niż 550 metrów. Specyfikacja jednofunkcyjnego kabla 8,3 mikrona, która wydaje się być najdroższa w produkcji i instalacji, obsługuje transmisję sygnału z prędkością 1 Gbps na odległość do 3 kilometrów.

1.8.2.1.3 1000BaseCX

1000BaseCX określa proponowaną przez grupę 802.3z specyfikację dla transmisji wykorzystującej wysokiej jakości ekranowaną skrętkę dwużyłową lub kabel koncentryczny. Niezależnie od nośnika, maksymalna odległość dla takiej transmisji wynosi 25 metrów.
Tak mała odległość transmisji znacząco ogranicza ten interfejs fizyczny. Jednym z proponowanych zastosowań 1000BaseCX jest łączenie ze sobą gigabitowych przełączników za pomocą tanich przewodów miedzianych, co stanowiłoby alternatywę dla połączeń światłowodowych.
Współtwórcy Gigabit Ethernetu nie wspominają jednak, że przesyłanie sygnałów z tak dużą prędkością przez przewody miedziane jest dość problematyczne. Miedziane instalacje przesyłowe mogą zdziałać cuda, jeśli dostarczy się im wystarczająco dużo energii elektrycznej: im więcej energii „wpompuje" się do sieci, tym większa będzie szerokość pasma. Niestety, miedź jest wspaniałym radiatorem elektromagnetycznym. Wyższa moc i częstotliwość transmisji skutkuje wysokim poziomem promieniowania elektromagnetycznego (ang.. EMI - Electromagnetic Radiation). Emisja promieniowania EMI jest ściśle regulowana przez Federalną Komisję Komunikacyjną (FCC - Federal Communications Commission), gdyż dotyczy pasm częstotliwości podlegających kontroli tej komisji.
Dla sieci skutek jest taki, że przy wykorzystywaniu dzisiejszych schematów sygnalizacyjnych gigabitowe sygnały mogą być przesyłane miedzianymi przewodami tylko na bardzo małe odległości.

1.8.2.1.4 1OOOBaseT

Przedstawione wcześniej opisy fizycznych interfejsów Gigabit Ethernetu są tak bliskie rzeczywistości, jak tylko to możliwe. Uzyskały `one końcową postać projektową, co oznacza, że są już zamknięte dla nowych rozwiązań.
1000BaseT to jednak zupełnie inna historia. Jak dotąd nie powstała technologia, która podczas transmisji danych przewodami UTP Kategorii 5 na odległość do 100 metrów pozwalałaby przekroczyć granicę szybkości 100 Mbps. Tak więc nie istniała specyfikacja interfejsu fizycznego, która mogłaby zachęcać do prac nad tym standardem.
Nad standardem 1000BaseT pracuje oddzielny zespół zadaniowy. Zarówno zespół, jak i potencjalny standard nazywane są „802.3ab". Celem zespołu jest uzyskanie wydajności Fast Ethernetu wykorzystującego cztery pary przewodów UTP Kategorii 5, ale przy szybkości transmisji sygnałów wynoszącej 1024 Mbps. Zadanie to wyodrębniono z prac zespołu 802.3z nad Gigabit Ethernetem, po prostu ze względu na ilość pracy, jaka musiała zostać wykonana. Należy pamiętać, że większą część Gigabit Ethernetu utworzono, mieszając i dopasowując istniejące specyfikacje warstwy łącza danych i warstwy fizycznej różnych technologii sieciowych. Dlatego wysiłki ograniczyły się przede wszystkim do pogodzenia pomniejszych różnic.
Jak dotąd zespół opracował trzy propozycje. Dwie są bardzo podobne i dotyczą transmisji w paśmie podstawowym. Prawdopodobnie zostaną połączone w jedną całość. Trzecia propozycja jest całkowicie odmienna. Oparta jest na metodzie transmisji w paśmie przepustowym. Jest to metoda dużo bardziej złożona i bardziej elegancka technicznie niż transmisja w paśmie podstawowym. Niestety, wymaga również większego poboru mocy i większej liczby bramek logicznych na karcie sieciowej. Nie ma także takiej odporności na zakłócenia, jaką oferuje technika przesyłania sygnałów różnicowych w przypadku transmisji w paśmie podstawowym z wykorzystaniem UTP. Te czynniki bezpośrednio przekładają się na mniejsze poparcie, jakim ta propozycja cieszy się wśród producentów, w porównaniu z propozycjami transmisji w paśmie podstawowym.
Pasmo przepustowe (ang.. passband) to metoda multipleksowania przedziałów częstotliwości. W odróżnieniu od pasma szerokiego (ang.. broadband), liczba kanałów i przypisane im częstotliwości są ograniczone tylko do takich, które mogą przejść przez dany filtr bez całkowitego wytłumienia.
Praca zespołu 802.3ab wydaje się być czynnikiem istotnym dla przyszłości Gigabit Ethernetu jako ekonomicznie efektywnej technologii łączenia stacji roboczych. Ogromna większość istniejących instalacji kablowych wykorzystuje do łączenia stacji skrętkę dwużyłową. Większość kabli to kable Kategorii 3, ale wraz z przemieszczaniem się grup użytkowników, zmianą środowiska ich stanowisk i zaang.ażowaniem w wiele innych czynności, zainstalowane okablowanie UTP Kategorii 3 coraz bardziej się zużywa. Przypuszczalnie, ze względu na wspomnianą wielość czynności, nowym okablowaniem, które będzie musiało je obsługiwać, będzie UTP Kategorii 5. Jeśli kabel UTP Kategorii 5 się rozpowszechni, koszty przejścia na Gigabit Ethernet (korzystający z interfejsu 1000BaseT) automatycznie spadną.
Zespół 802.3ab wie, że stoi przed zastraszająco trudnym wyzwaniem. Musi on opracować efektywną ekonomicznie i nadającą się do produkcji technologię sygnalizacyjną, która umożliwi czterem parom przewodów UTP Kategorii 5 transmisję sygnałów z gi
gabitowymi prędkościami, daleko przewyższającymi 100 Mbps, nie wywołując przy tym oburzenia komisji FCC! Jak dotąd, każda para może przesyłać sygnały z prędkością
do 100 Mbps, wykorzystując konwencjonalne techniki sygnalizacyjne. Zespół 802.3ab musi więc załatać całkiem sporą dziurę.
Przewidując nieunikniony sukces, zespół 802.3z opracował interfejs logiczny, który będzie istniał między warstwą MAC a warstwą fizyczną. Interfejs ten odsprzęga od podwarstwy MAC stary schemat kodowania, pochodzący z kanału światłowodowego. Ma to umożliwić włączenie do standardów 802.3z innych schematów kodowania, np. takich jak ten, który opracuje zespół 802.3ab.

1.8.2.2 Co jeszcze nowego?

Słuszne jest pytanie: „Co jeszcze nowego, oprócz potrzeby wprowadzenia całkowicie nowych interfejsów warstwy fizycznej?" Odpowiedź, jak dotąd, brzmi: „Niewiele". Wiele trudu kosztowało zachowanie możliwie największej części różnych istniejących specyfikacji 802.3. Zrobiono to, aby zapewnić ciągłą współoperacyjność w ramach rodzin specyfikacji 802.2 i 802.3. Zauważalne zmiany zaszły jeszcze w dwóch obszarach (innych niż prędkość sygnału, szybkość transmisji danych i interfejsy fizyczne!): chodzi tu 0 odstęp między ramkami i metodę dostępu do nośnika opartą na zasadzie rywalizacji.

1.8.2.2.1 Odstęp między ramkami

Zmiana rozmiaru odstępu między ramkami jest jedyną znaczącą zmianą w protokole 802.3, wymuszoną przez Gigabit Ethernet. Jak wyjaśniono w rozdziale 7 pt. „Ethernet", protokół CSMA/CD 802.3 zapewniał odstęp między ramkami o rozmiarze 96 bitów. W rzeczywistości odstęp ten jest strumieniem losowych znaków binarnych. Jego jedynym przeznaczeniem (w przeciwieństwie do tego, o czym

głoszą plotki, które można znaleźć w grupach dyskusyjnych), jest podtrzymywanie zajętości linii. Strumień ten został wbudowany w protokół po to, by umożliwić nadającej stacji utrzymanie linii w czasie, gdy stacja ta przygotowuje do wysłania kolejną ramkę danych.
W sieci Gigabit Ethernet 96 bitów przepływa tak szybko, że elementy elektroniczne karty sieciowej gubią się, stosując przyjętą taktykę zatrzymywania. Co więcej, przy gigabitowej prędkości sygnału minimalny rozmiar ramki wynoszący 64 oktety jest zbyt mały, żeby mógł zapewnić wykrycie kolizji w sieci. Pamiętajmy, że minimalny rozmiar ramki ustalono na podstawie czasu, jaki ramce zajmie dotarcie do wszystkich stacji w sieci LAN działającej z prędkością 10 Mbps. Dopasowanie do gigabitowej transmisji oznacza znaczne zwiększenie minimalnego rozmiaru ramki. Uniemożliwiłoby to jednak współoperacyjność z innymi, wolniejszymi sieciami Ethernet 802.3. Uznano to za niemożliwe do zaakceptowania.
Rozwiązaniem obydwu dylematów było zwiększenie rozmiaru odstępu między ramkami. W Gigabit Ethernecie odstęp ma długość 512 oktetów. Można podnieść argument wątpliwej wydajności stosowania 512-oktetowej przerwy między ramkami, skoro ich rozmiary mogą być tak małe, jak 64 oktety. Jednak w tym wypadku nie chodziło o wydajność. Celem była współoperacyjność. Aby ułagodzić narzekających malkontentów, nowe rozwiązanie zostało włączone do protokołu 802.3z. Znane jest ono jako sekwencjonowanie ramek (ang.. Frame Bursting) i pozwala stacji nadającej wysłać maksymalnie do 8192 oktetów za każdym razem, gdy uda jej się uzyskać dostęp do pasma.

1.8.2.2.2 Dostęp do nośnika na zasadzie rywalizacji

Choć nie całkiem nowa - w tym sensie, że poprzedzała inicjatywę 802.3z - ethernetowa metoda dostępu do nośnika na zasadzie rywalizacji stała się dyskusyjna. Gdy IEEE opracowywało opisane wcześniej wersje Fast Ethernetu, wprowadzono zastrzeżenie, że transmisja w tej sieci ma być pełnodupleksowa.
Transmisja pełnodupleksowa polega na tym, że urządzenie może nadawać i odbierać równocześnie, choć oddzielnymi kanałami. Nadajnik jednego urządzenia (na przykład przełącznika) jest połączony dedykowanym kanałem z odbiornikiem innego urządzenia (na przykład karty sieciowej komputera) i odwrotnie. Eliminuje to rywalizację o dostęp do tego samego pasma pomiędzy funkcjami nadawania i odbioru, co było wcześniej bolączką urządzeń pracujących w sieci Ethernet.
W środowisku komutowanym kolizje dotyczą tylko dwóch urządzeń: przełącznika i urządzenia, z którym jest on połączony. Prowadzenie transmisji pełnodupleksowej łączem komutowanym sprawia, że rywalizacja o dostęp do pasma staje się niepotrzebna. Każde urządzenie może uzyskać pozwolenie na nadawanie i umieścić ramki w nośniku niemal bez opóźnień. Protokół CSMA/CD staje się zbyteczną pozostałością z dawnych czasów Ethernetu.
Gigabit Ethernet bazuje na tym właśnie sukcesie transmisji pełnodupleksowej. W praktyce wielu producentów projektuje swoje wyroby wyłącznie jako pełnodupleksowe. Jest to testamentem sukcesu pełnodupleksowego, komutowanego środowiska sieciowego.

1.8.2.3 Zbyt dobre, aby mogło być prawdziwe?

Choć technicznie interesujący, a nawet ekscytujący, Gigabit Ethernet musi być rozpatrywany w kontekście przewidywanego środowiska jego pracy. Podmioty zajmujące się jego tworzeniem widzą w nim technologię łączenia ze sobą koncentratorów, serwerów, a być może nawet stacji roboczych. Jego użyteczność musi być więc oceniana w zakresie wydajności i wymagań stawianych przez każdy z tych obszarów funkcjonalnych sieci LAN.
Dzisiejsze stacje robocze i serwery są w większości niedopasowane do pasma oferowanego przez Gigabit Ethernet. Nawet uwzględniając spowalniający wpływ schematu kodowania 8B/ I 0B., Gigabit Ethernet wciąż oferuje przepustowość 800 Mbps. To sprawia, że sieć LAN przestaje być czynnikiem ograniczającym wydajność. Ograniczenie stanowią dwa czynniki wewnątrz komputera. Pierwszym jest fakt, że nawet SCSI-3 może działać tylko z prędkością do 40 MB/s (megabajtów na sekundę). Przekłada się to na prędkość sygnału 320 Mbps (megabitów na sekundę). Tak więc każda czynność w sieci Gigabit Ethernet, wymagająca odwołań do urządzeń peryferyjnych przyłączonych do magistrali SCSI, będzie ograniczona wydajnością tej magistrali.
Funkcje, które mogą być obsługiwane bez udziału urządzeń peryferyjnych SCSI, będą limitowane ograniczeniami wydajności magistrali PCI (ang.. Peripheral Component Interconnect). PCI stała się dominującą architekturą magistrali we/wy w tanich komputerach. Jest wykorzystywana w maszynach jedno- i wielostanowiskowych. W praktyce większość komputerów z magistralą PCI nie doświadczy dużego wzrostu wydajności (lub wręcz nie doświadczy go wcale) po przejściu z Ethernetu 100 Mbps na Gigabit Ethernet. Nie
pozwolą na to ich elementy składowe. Tak więc Gigabit Ethernet jest rozwiązaniem nieodpowiednim dla dzisiejszych architektur tanich komputerów.
Gigabit Ethernet nadaje się najlepiej do łączenia solidniejszych komputerów średnich (ang.. mid-rang.e) i komputerów mainframe. Te architektury są znacznie lepiej dopasowane do przepustowości sieci Gigabit Ethernet. Koncentratory przełączające, które mogą obsługiwać łącze Gigabit Ethernetu, są w stanie wykorzystywać go jako szkielet sieci LAN.

1.8.3 Podsumowanie

Fast Ethernet sprawdził się jako praktyczny i ekonomiczny środek podniesienia jakości infrastruktur sieci LAN. Bardzo ważne jest dotrzymanie kroku wciąż rosnącym wymaganiom dotyczącym przyłączalności i przepustowości. Największym zagrożeniem dla specyfikacji Fast Ethernet jest to, że korzysta ona z całkowicie odmiennych schematów sygnalizacyjnych dla różnych rodzajów kabli. Tak więc, płytka krzemowa wykorzystywana przez złącze kabla UTP Kategorii 5 jest zupełnie różna od płytki dla UTP Kategorii 3. Każda grupa użytkowników wdrażająca Fast Ethernet z kablami Kategorii 3 inwestuje w technologie telefoniczne i w rzeczywistości utrudnia migrację do przyszłościowych technologii sieci LAN.
Gigabit Ethernet, choć nieukończony, wydaje się być zdolny do dalszego przedłużenia życia tego starego protokołu. Kluczowa różnica między dwiema wymienionymi wyżej technologiami jest taka, że Gigabit Ethernet zajmuje niszę technologiczną. Jego przepustowość o kilka rzędów wielkości przekracza możliwości dzisiejszych komputerów. Układy we/wy i magistrale systemowe wykorzystywane w tych komputerach stają się czynnikami ograniczającymi, nie pozwalającymi w pełni wykorzystać przepustowości oferowanej przez Gigabit

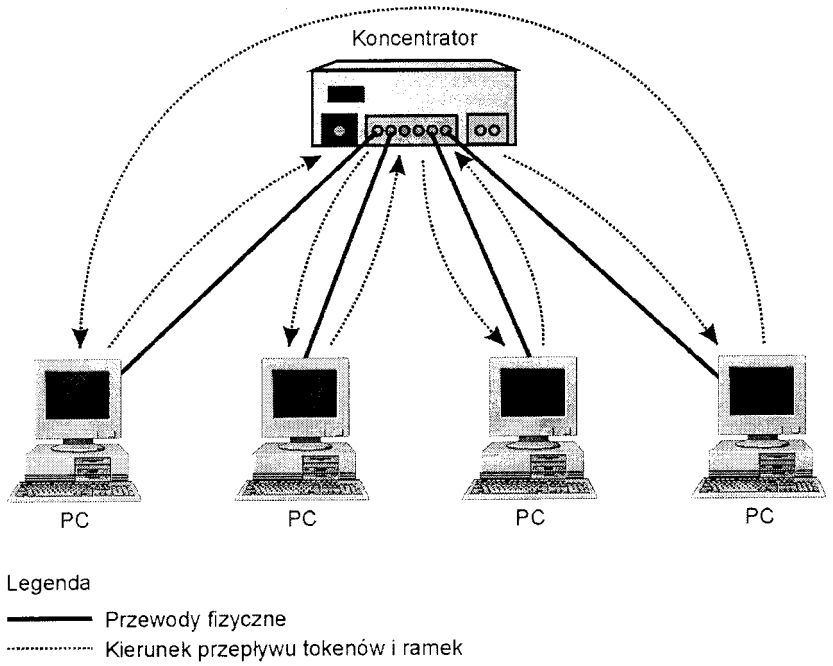
Ethernet. W przyszłości może on stać się technologią LAN nadającą się do bardziej ogólnego zastosowania, ale do tego czasu jego użyteczność wydaje się być ograniczona.
Więcej aktualnych informacji o rozwoju Gigabit Ethernetu można znaleźć na stronie: http://www.Gigabit-Ethernet.org

1.9 Rozdział 9 Token Ring

Mark A. Sportack
Token Ring jest kolejną architekturą sieci LAN znormalizowaną przez IEEE. Ma wiele cech wspólnych z Ethernetem i innymi architekturami sieci LAN należącymi do rodziny standardów sieciowych IEEE 802. W rezultacie może z nimi współpracować, korzystając z mostu tłumaczącego.
Token Ring również przeszedł sporo aktualizacji od momentu powstania. Początkowo był technologią dostosowaną do pasma 4 Mbps, później przepustowość podniesiono do 16 Mbps. Dziś istnieją propozycje rozwiązań, które zwiększą prędkość sygnału w sieci Token Ring do 100 lub nawet 128 Mbps, a być może do 1 Gbps.
W rozdziale opisywane są mechanizmy Token Ringu. Przedstawiono wiele struktur ramek, specyfikacje nośników fizycznych, mechanikę działania, a także krótkie spojrzenie w przyszłość tej architektury.

1.9.1 Przegląd

W swej znormalizowanej formie Token Ring jest solidną i wysoce deterministyczną architekturą LAN. Nazwę zawdzięcza swojemu okrężnemu schematowi dostępu do nośnika. W odróżnieniu od Ethernetu, z jego chaotyczną i nieuregulowaną metodą wielodostępu, Token Ring pozwala w danym czasie nadawać tylko jednemu urządzeniu. Nie występują więc kolizje.
Dostęp do nośnika jest przyznawany poprzez przekazywanie tokenu w ustalony sposób. Token może być tylko jeden i jest on modyfikowany przez urządzenie transmitujące w celu utworzenia nagłówka ramki danych. Gdyby nie było tokenu, nie dałoby się utworzyć nagłówka ramki danych i transmisja byłaby niemożliwa. Urządzenie odbierające kopiuje dane przesyłane w ramce, zmieniając przy tym (negując) niektóre bity nagłówka ramki i w ten sposób potwierdzając odbiór. Sama ramka dalej krąży w pierścieniu, aż powróci do swojego nadawcy. Urządzenie, które wysłało ramkę, pobiera ją teraz z sieci i usuwa z niej dane oraz adresy. Jeśli urządzenie chce przesłać więcej danych, może to zrobić. Jeśli nie, nagłówek ramki jest przekształcany z powrotem w token i umieszczany w medium transmisyjnym, przez które podróżuje do następnego urządzenia.
Przekazywanie tokenu odbywa się w przedstawiony na rysunku 9.1 sposób okrężny, co daje każdemu urządzeniu okazję do nadawania. Aby zapewnić, że żadna stacja nie zmonopolizuje łącza, stosuje się mechanizm znany jako zegar przetrzymywania tokenu, śledzący i regulujący maksymalną ilość czasu, przez który dowolna stacja może mieć prawo do nadawania. Ten mechanizm czasowy jest przydatny także podczas przywracania normalnego działania sieci w wypadku, *gdy* stacja posiadająca token przestanie działać. Innym korzystnym efektem ubocznym związanym z techniką dostępu do nośnika za pomocą przekazywania uprawnienia jest to, że sieci Token Ring mogą być skalowane co do rozmiaru i natężenia transmisji w o wiele bardziej elegancki sposób niż sieci wykorzystujące zasadę rywalizacji.
Rysunek 9.1. Tokeny
przeka-zywane sekwencyjnie e po drodze okrężnej.



1.9.2 Standaryzacja sieci Token Ring

Token Ring, jak na dzisiejsze standardy informatyczne, jawi się jako technologia wręcz „starożytna". Został stworzony przez firmę IBM jako technologia centrum danych dla pracujących w sieci komputerów *mainframe.* Po raz pierwszy przedstawiono go instytutowi IEEE do

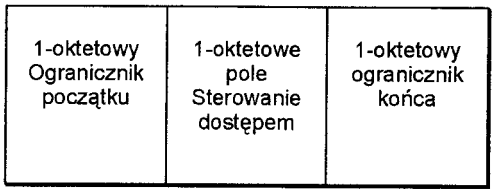
standaryzacji w roku 1969. Gdy pojawiły się komputery osobiste (PC), zauważono, że Token Ring może posłużyć do łączenia ich ze sobą. Przyspieszyło to włączenie Token Ringu do projektu IEEE 802.
Standaryzacja w ramach projektu 802 wymusiła dokonanie pewnych zmian w warstwie łącza danych, tak aby mogło obsługiwać adresowanie sprzętowe i połączenia mostowe z innymi architekturami LAN 802.
IEEE nazwało Token Ring specyfikacją 802.5. Jest ona niemal identyczna ze specyfikacją Token Ringu firmy IBM. Oprócz wspomnianych wcześniej zmian sprzętowych, IEEE znormalizowała format wiadomości oraz protokoły warstwy 2. Nawiasem mówiąc, IBM był głównym orędownikiem wysiłków standaryzacyjnych IEEE.
Token Ring zaoferował solidniejsze, właściwsze (czasowo) i bardziej deterministyczne podejście do działania sieci niż protokół Ethernet 802.3, choć przy wyższych kosztach przypadających na pojedyncze stanowisko. Przedsiębiorstwa, których aplikacje wymagały dostarczania danych w odpowiednim czasie, uznały Token Ring za jedyne rozwiązanie stosowne do potrzeb. Choć protokół 802.3 zapewnia, że pakiet zostanie prawidłowo przesłany, może to jednak wymagać kilku prób transmisji. Nie można więc zagwarantować granic czasowych dostarczenia przesyłki. Token Ring może to zagwarantować dzięki swojej deterministycznej topologii o kształcie pierścienia i uporządkowanej metodzie dostępu.

1.9.3 Struktura ramki Token Ring

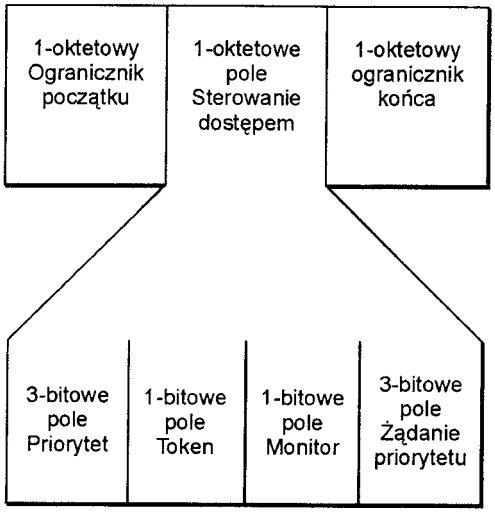
Urządzenia Token Ringu nie mogą nadawać niczego bez tokenu. Podstawowy token służy dwóm celom:
• Jest używany do przyznawania przywilejów dostępu.
• Podstawowa ramka tokenu jest przekształcana w nagłówki rozmaitych, specjalizowanych ramek.
W rzeczywistości każda funkcja (w tym także przesyłanie danych) wymaga ramki 0 określonej strukturze.
Token Ring obsługuje następujące'rodzaje ramek: • Ramkę Token
• Ramkę danych
• Ramkę danych LLC
• Ramki zarządzania MAC • Ramkę przerwania

1.9.3.1 Ramka Token

Token Ring IEEE 802.5 wykorzystuje do sterowania dostępem do nośnika specjalną sekwencję bitów, znaną jako token. Token zawiera następujące pola: Ogranicznik Początku, Sterowanie Dostępem i Ogranicznik Końca. Każde pole ma długość 1 oktetu (8 bitów). Ramka Token jest przedstawiona na rysunku 9.2.
Rysunek 9.2. Ramka Token IEEE 802.5.



Ramka ta jest przekazywana od urządzenia do urządzenia i przydziela prawa transmisji urządzeniom w pierścieniu. Ponieważ istnieje tylko jedna ramka Token, w danym momencie tylko jedna stacja może nadawać czy też podejmować próbę nadawania.
Pole sterowania dostępem
Pole sterowania dostępem jest kluczowym polem tokenu. Każdy z jego ośmiu bitów jest znaczący. Jak widać na rysunku 9.3, pole sterowania dostępem zawiera:
• 3-bitowe pole Priorytet • I-bitowe pole Token • 1-bitowe pole Monitor
• 3-bitowe pole Żądanie priorytetu
Rysunek 9.3. Pole sterowania dostępem.
Pole bitowe Priorytet wskazuje priorytet samego tokenu. Może przyjmować wartość z zakresu od 000 do 111 i jest ustawiane przez nadającą stację. Nie może być zmieniane przez inne stacje. Tylko stacje o priorytecie równym lub wyższym niż wartość tego pola mogą je modyfikować. Bit Token jest bitem, którym należy manipulować, aby zmienić token w sekwencję początku ramki (o czym już wcześniej wspominam w tym rozdziale). Faktycznie, bit ten ustawiony jako 1, mówi innym stacjom, że token jest teraz częścią ramki. Oznacza to, że jest w tej chwili używany. Tokeny, jak te pokazane na rysunku 9.2, kiedy krążą po pierścieniu, wywołując stacje, mają ten bit ustawiony na 0.



Pole Żądanie Priorytetu pozwala stacjom żądać usługi o wyższym priorytecie. Dotyczy to stacji posiadających dane o wysokim priorytecie, które muszą być przesłane tak szybko, jak to możliwe. Stacje mogą oznajmiać o swoich potrzebach dotyczących priorytetu, ustawiając bity żądania priorytetu odpowiednio do ważności ich danych (dopuszczalny jest zakres wartości od 000 do 111). W ten sposób informują nadawcę tokenu o swoich potrzebach. Nadawca ustawia następnie pole Priorytet, stosownie do otrzymanej właśnie wartości pola Żądanie priorytetu. W ten sposób stacja z najwyższym priorytetem może zarezerwować sobie następny token, niezależnie od tego, ile stacji znajduje się między nią a nadawcą tokenu.
Bity priorytetu są uporządkowane wg znaczenia od lewej strony do prawej. Bit położony najbardziej na lewo jest bitem najbardziej znaczącym. Oznacza to, że niesie ze sobą wyższy priorytet niż bity środkowy i prawy. Dlatego 010 jest wyższym priorytetem niż 001.

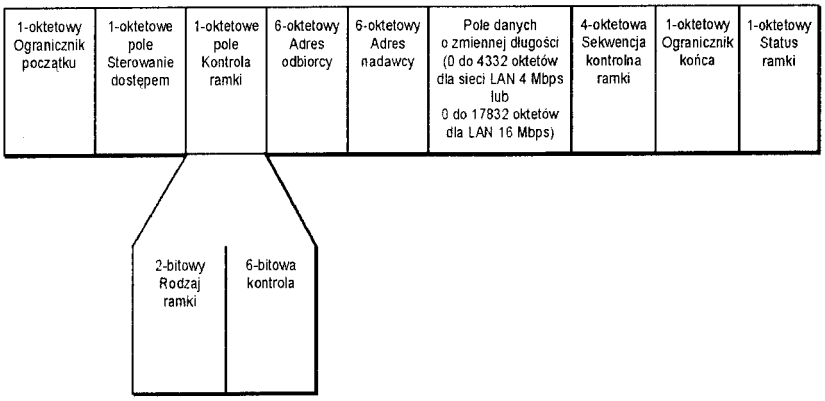
1.9.3.2 Ramka danych

Minimalna długość ramki danych w sieci Token Ring wynosi 21 oktetów. Rozmiar maksymalny zależy od prędkości sygnału w pierścieniu. Czas potrzebny na przesłanie ramki musi być mniejszy niż ustalony czas przetrzymywania tokenu. Czas ten jest domyślnie ustawiany na 10 milisekund. W Token Ringu pracującym z szybkością 4 Mbps daje to maksymalną długość ramki danych równą 4500 oktetów. Przy szybkości 16 Mbps ramki danych mogą mieć długość do 18000 oktetów.
Struktura ramki danych Token Ringu 802.5 składa się z dwóch części: ramki Token i ramki danych. Kiedy urządzenie przechwytuje token i zmienia wartość bitu Token, czyni pierwszy krok w kierunku utworzenia ramki danych. Kolejnym krokiem jest wstawienie innych pól, wymaganych przez osadzoną w protokole strukturę ramki danych, i nadanie im wartości. Kompletna ramka danych jest przedstawiona na rysunku 9.4.
Rysunek 9.4. Ramka danych IEEE 802.. s.

1-oktetowy Ogranicznik początku	1-oktetowe pole Sterowanie dostępem	1-oktetowe pole Kontrola ramki	6-oktetowy Adres odbiorcy	6-oktetowy Adres nadawcy	Pole danych o zmiennej długości (0 do 4332 oktetów dla sieci LAN 4 Mbps lub 0 do 17832 oktetów dla LAN 16 Mbps)	4-oktetowa Sekwencja kontrolna ramki	1-oktetowy Ogranicznik końca	1-oktetowy Status ramki

Jak dowodzi rysunek 9.4, trzy 1-oktetowe pola ramki Token pozostają w ramce danych. Do tej podstawowej struktury dodaje się sześć innych pól oraz podpól.
Pierwszym polem jest Ogranicznik początku, określający początek ramki. Po nim następuje pole Sterowanie dostępem (opisane w podpunkcie „Pole sterowania dostępem") i 8-bitowe pole Kontrola ramki. Pole to przechowuje bity „rodzaju" identyfikujące protokół transportu. Pole to służy także do rozróżniania ramek danych i ramek sterowania. Pierwsze dwa bity określają typ ramki: ramka danych lub ramka zarządzania MAC. Następnych 6 bitów informuje odbiorcę o priorytecie jednostki danych protokołu (ang..
PDU- Protocol Data Unit) i/lub jednostki MAC PDU. Jeśli jest to ramka MAC, pole to określa także dokładnie rodzaj ramki zarządzania MAC. Pole Kontrola Ramki przedstawione jest na rysunku 9.5.
Następne dwa pola to adresy fizyczne MAC odbiorcy i nadawcy. Każdy z nich ma 6oktetów. Te adresy MAC („sterowanie dostępem do nośnika") odpowiadają opisanej poprzednio specyfikacji projektu 802 i są identyczne z używanymi w sieciach Ethernet.
Pole danych w sieci opartej na tokenie ma zmienną długość, zależną od prędkości sygnału w pierścieniu. Sieci Token Ring 4 Mbps mogą obsługiwać pola danych o długości od 0 do 4332 oktetów. Sieci Token Ring 16 Mbps mogą obsługiwać pola danych o długości od 0 do 17832 oktetów. Wartości te reprezentują maksymalny rozmiar ramki (przy czasie przetrzymywania tokenu ustalonym na 10 milisekund), pomniejszony o 168 oktetów podstawowej struktury ramki.
Rysunek 9.5.
Pole Kontrola ramki.
Ostatnie trzy pola ramki danych to: 32-bitowa Sekwencja kontrolna ramki (FCS), 8bitowy Ogranicznik końca i 8-bitowy Status ramki. Sekwencja kontrolna ramki zawiera matematycznie wyprowadzoną wartość - sumę kontrolną- obliczoną na podstawie długości i zawartości ramki. Odbiorca i nadawca stosują wobec ramki ten sam algorytm. Jeśli odbiorca uzyska taką samą wartość sumy kontrolnej jak wartość

przechowywana w polu FCS (która została obliczona przez nadawcę), to może przyjąć, że zawartość ramki nie uległa zmianie podczas transmisji.



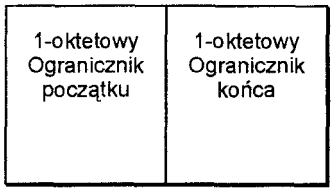
Ostatnie dwa oktety, obejmujące Ogranicznik końca i Status ramki, są nazywane sekwencją końca ramki (ang.. End of Frame Sequence).
Tak wygląda podstawowa, czy też „surowa", postać ramki danych Token Ring. W praktyce wykorzystuje się ją razem z mechanizmami sterowania łączem logicznym specyfikacji IEEE 802.2. Specyfikacja ta odwołuje się do dodatkowej struktury podramki, dodanej do ramki danych w celu identyfikacji protokołu wyższej warstwy, dla którego przeznaczona jest zawartość ramki. Jest to istotne w dzisiejszym środowisku wieloprotokołowej komunikacji i obliczeń. Ramka danych Token Ring z dodaną podramką 802.2 znana jest jako ramka danych Token Ring LLC. Więcej informacji o strukturze podramki 802.2 można znaleźć w rozdziale 7 pt. „Ethernet".

1.9.3.3 Ramki zarządzania MAC

Protokół Token Ring IEEE 802.5 ustanawia czterech agentów zarządzania siecią (ang.. NMA - Network Management Agents). Agenci przebywają w każdej stacj i Token Ringu i są wykorzystywani w zwykłych czynnościach zarządzania pierścieniem. Agentami tymi są:
• monitory: aktywny (ang.. AM - Active Monitor) lub oczekujący (ang.. SM Standby Monitor)
• monitor błędów pierścienia (ang.. *REM-Ring Error Monitor)*
• serwer raportu konfiguracji (ang.. CRS- Configuration Report Server) • serwer parametrów pierścienia (ang.. RPS - Ring Parameter Server)
Skoro transmisja w pierścieniu jest możliwa tylko przy użyciu ramki, której podstawą jest token, nie powinno być wielką niespodzianką, że każdy z tych agentów może generować kilka różnych, wysoce wyspecjalizowanych rodzajów ramek zarządzania MAC. Faktycznie, w sieci Token Ring IEEE ta czwórka agentów może generować i używać 25 różnych ramek MAC! Jeśli potrzebna jest jeszcze większa kontrola, to opracowane przez IBM rozszerzenie struktur ramek MAC IEEE dodaje kolejne 17 ramek MAC, co w sumie daje 42 różne ramki MAC.
Każda ramka MAC wykonuje określoną funkcję zarządzania siecią. Oto niektóre z tych funkcji:
• lobe test (test podłączenia stacji końcowej), • inicjalizacja pierścienia,
• czyszczenie pierścienia, • token zgłoszenia,
• różne funkcje monitora aktywnego.
Względnie duża liczba ramek sprawia, że bezcelowe jest szczegółowe ich omawianie. Wystarczy powiedzieć, że ramki MAC służą do zbierania miar wydajności sieci, które mogą być dostarczane do zgodnych ze standardami produktów zarządzania siecią. Wiele z tych miar opisanych jest w kontekście eksploatacyjnym w dalszej części tego rozdziału.

1.9.3.4 Ramka przerwania

Ramka przerwania zawiera wyłącznie pola ograniczników początku i końca ramki. Choć z powodu braku danych i adresów taka struktura może wydawać się bezużyteczna, to ramka przerwania znajduje zastosowanie - jest wykorzystywana do natychmiastowego zakończenia transmisji.
Strukturę ramki przerwania przedstawia rysunek 9.6.
Rysunek 9.6. Ramka przerwania IEEE 802. 5.



1.9.3.5 Sekwencja wypełniania

Jedynym nie bazującym na ramce nośnikiem transmisyjnym w sieci Token Ring jest sekwencja wypełniania. Nie posiada ona ograniczników początku i końca. Jest po prostu dowolnym ciągiem zer i jedynek.

Sekwencja wypełniania jest generowana przez stację nadającą. Pamiętajmy, że jest tylko jeden token, więc w danym momencie tylko jedna stacja może nadawać.
Sekwencja wypełniania jest wykorzystywana w połączeniu z różnymi, wymienionymi uprzednio typami ramek, aby zapobiec wystąpieniu w pierścieniu czasu ciszy (ang.. *quiet time).* Czas ciszy następuje, gdy ani jedna ramka lub token nie przemieszcza się przez pierścień. Czas ciszy jest interpretowany przez monitor aktywny i wszystkie inne stację jako przerwanie pierścienia. W wyniku tego inicjowane są mechanizmy samoczynnej naprawy izolujące obszar awarii i identyfikujące nieosiągalne stacje. Może to szkodliwie wpływać na wydajności sieci, jeśli pierścień nie został uszkodzony. Na przykład, jeśli stacja, która odebrała token, przetrzymuje go, umieszczając jednocześnie w ramkach dane otrzymane z wyższych protokołów, w pierścieniu nastaje cisza. Inne stacje mogą interpretować ten brak aktywności jako stan awaryjny.
Aby zapobiec niewłaściwej reakcji na czas ciszy, stacja transmitująca podczas przygotowywania rzeczywistej ramki lub tokenu generuje losową sekwencję zer i jedynek. Sekwencja wypełniania jest nadawana przed i po wysłaniu ramki lub tokenu. Sekwencja ta może również być wykorzystywana przez nadającą stację do zatrzymywania czasu. Wysyłając losowy ciąg bitów zamiast rzeczywistej ramki lub tokenu, nadająca stacja może zatrzymać swój zegar przetrzymywania tokenu. Pozwala to przetrzymywać token dłużej niż byłoby to możliwe winnym przypadku.

1.9.4 Funkcjonowanie sieci Token Ring

Przegląd różnych struktur ramek Token Ringu powinien pokazać, że jest to dość złożona i bardzo solidna architektura sieci LAN. Szybki przegląd mechaniki jego działania powinien stworzyć odpowiednie tło dla szczegółowej analizy fizycznych i logicznych komponentów Token Ringu.
Token Ring wykorzystuje token do przydzielania dostępu do nośnika. Tokeny są rozpoznawane i obsługiwane przez wszystkie stacje pracujące w sieci. Token może być tylko jeden i tylko jego posiadacz może nadawać.
Token jest przekazywany od stacji do stacji w określonej kolejności i tylko w jednym kierunku. Ponieważ pierścień nie ma jasno zdefiniowanego początku i końca, token po prostu ciągle po nim krąży. Mechanizm ten znany jest jako wywoływanie metodą okrężną lub inaczej metodą *round-robin.* Każda stacja, która otrzyma token i chce nadawać, może przekształcić jego strukturę bitową w sekwencję początku ramki (ang. *SOF - Start of Frame).* Token służy więc do utworzenia ramki danych. Nadająca stacja zmienia sekwencję SOF, dodaje potrzebne dane, adresuje je i umieszcza z powrotem w sieci. Jeśli stacja nie chce nadawać, może po prostu z powrotem umieścić token w sieci - wtedy otrzyma go kolejna stacja. Gdy ramka dotrze do miejsca przeznaczenia, urządzenie odbierające nie wyciąga ramki z sieci, lecz po prostu kopiuje jej zawartość do bufora w celu dalszego wewnętrznego przetwarzania. W oryginalnej ramce zmieniany jest bit pola sterowania dostępem, co informuje nadawcę, że ramka została odebrana. Potem ramka kontynuuje swoją podróż przez pierścień, dopóki nie powróci do urządzenia, które ją wysłało. Gdy urządzenie ją odbierze, uznaje się, że transmisja zakończyła się sukcesem; zawartość ramki jest kasowana, a sama ramka jest z powrotem przekształcana w token.
Taka jest istota działania sieci Token Ring. Oczywiście jest to wersja mocno uproszczona i nie opisuje szczegółów różnych kroków i procesów. Regułami podstawowego działania urządzeń w sieci Token Ring rządzi jedna ze stacji znajdujących się w pierścieniu. Jest to tzw.. monitor aktywny (AM). Jest on szczegółowo opisany w dalszej części rozdziału, w punkcie „Monitor aktywny".
Znając powyższy opis mechanizmu przekazywania tokenu, można się oczywiście domyślić, że urządzenie zgodne ze standardem 802.5 jest półdupleksowe. Oznacza to, że może działać tylko w jednym z dwóch trybów: nadawania lub odbioru. Urządzenie nasłuchujące po prostu przekazuje token do następnego urządzenia w pierścieniu. Jeśli token został przekształcony w sekwencję początku ramki, urządzenie nasłuchujące sprawdza, czy ramka jest przeznaczona dla niego. Jeśli tak, buforuje dane i przesyła już zmodyfikowany token z powrotem do nadawcy ramki. Nadawca musi wtedy potwierdzić, że transmisja ramki zakończyła się sukcesem, zamienić sekwencję SOF z powrotem w token i umieścić go w sieci.
W trybie nadawania, jak już to wcześniej opisano, urządzenie zmienia strukturę bitów tokenu, aby utworzyć sekwencję początku ramki. Następnie urządzenie dołącza do niej niezbędne dane i nagłówki. Metodologia ta, w przeciwieństwie do stosowanej w Ethernecie, działa wydajniej przy dużym ruchu w sieci. Dzieje się tak dlatego, że zezwolenia na transmisję nie są przydzielane chaotycznie (jak w sieciach Ethernet),a maksymalna liczba oktetów w ramce nie jest ograniczona.

1.9.4.1 Sprzęt

Token Ring używa podstawowego zestawu komponentów sprzętowych, z których można zbudować wiele topologii obsługujących dostęp do nośnika za pomocą przekazywania tokenu. Oprócz niezbędnych kart sieciowych (NIC), do komponentów sprzętowych zalicza się:
• kabel dalekosiężny,
• kabel stacji końcowej,
• jednostki dostępu do stacji wieloterminalowej,
• jednostkę sprzęgania dalekosiężnego.
Wszystkie składniki opisane są w następnych podpunktach.

1.9.4.1.1 Kabel dalekosiężny

Kabel dalekosiężny stanowi szkielet sieci Token Ring. Jest to kabel łączący ze sobą wszystkie koncentratory (czyli „jednostki dostępu do stacji wieloterminalowej" - w języku Token Ringu).
Może to być kabel światłowodowy albo skrętka dwużyłowa, ekranowana lub nieekranowana. Skrętka dwużyłowa oferuje dodatkową korzyść: zapewnia rezerwową ścieżkę transmisji. Połączenie osiąga się, wykorzystując jedną parę; pozostałe pary w kablu UTP nie są używane. Jeśli nastąpi awaria (zakładając, że uszkodzeniu ulegnie tylko jedna para przewodów), można użyć drugiej pary do wykonania obejścia wokół części uszkodzonej.

1.9.4.1.2 Kabel stacji końcowej

Kable stacji końcowych używane są do przyłączania pojedynczych stacji do portu w koncentratorze. Podobnie jak w przypadku kabli dalekosiężnych, mogą to być światłowody lub skrętki dwużyłowe (ekranowane lub nie).
Warto zauważyć, że w większości systemów okablowania skrętką dwużyłową zainstalowanych w budynkach przemysłowych kabel stacji końcowej nie jest kablem pojedynczym. Jest to raczej szereg kabli połączonych razem tak, aby tworzyły one ciągłą ścieżkę.

1.9.4.1.3 Jednostki dostępu do stacji wieloterminalowej

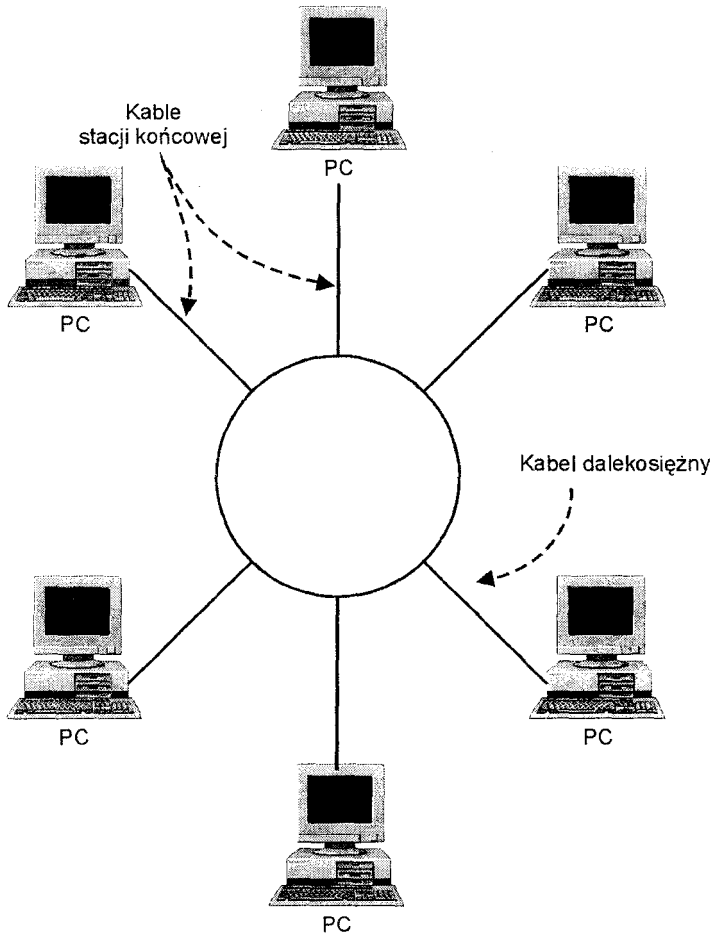
Urządzenie służące zarówno jako wzmacniak, jak i punkt dostępu dla wielu stacji (innymi słowy koncentrator), znane jest jako jednostka dostępu do stacji wieloterminalowej (ang. MSAU- Multi-Station Access Unit). Urządzenia te, jak większość koncentratorów, mogą być łączone ze sobą, aby utworzyć większą sieć.
Tak jak w przypadku koncentratorów w sieci Ethernet, należy uważać, żeby nie połączyć ze sobą dwóch portów urządzeń komunikacyjnych DCE. Jednostki MSAU posiadają porty oznaczone jako Ring In (RI) i Ring Out (RO). Oczywistym powinno być, że porty RI obsługują połączenia przychodzące: są to porty DCE. Porty RO są łączone z portami RI innych koncentratorów: są to porty terminali DTE. Próba połączenia ze sobą dwóch jednostek MSAU poprzez łączenie ich portów RI lub RO (RI z RI lub RO z RO) nie powiedzie się, chyba że zastosuje się kabel skrośny.
Typowy koncentrator (MSAU) ma od 8 do 24 portów RI i/lub RO. Bardziej formalnie porty te są nazywane jednostkami sprzęgania dalekosiężnego lub inaczej jednostkami TCU (ang. Trunk Coupling Units).
Jednostki sprzęgania dalekosiężnego
Jednostki sprzęgania dalekosiężnego (TCU) to porty fizyczne oraz układy elektroniczne i logiczne pomagające tym portom obsługiwać połączenia z innymi stacjami i koncentratorami. Porty TCU posiadają inteligentne układy elektroniczne, pozwalające na przyłączanie i odłączanie stacji do i od pierścienia. Umożliwia to dynamiczne i automatyczne zarządzanie elementami pierścienia.
Stacje, które nie są aktywne z jakiegokolwiek powodu, nie są faktycznie odłączane od sieci Token Ring. Zamiast tego TCU rozpoznaje stan nieaktywny stacji i omija ją (elektrycznie), kiedy przekazuje tokeny i ramki przez pierścień.

1.9.4.2 Topologia

Opisane właśnie komponenty fizyczne są modułami konstrukcyjnymi sieci Token Ring. Istnieją różne topologie, czyli sposoby rozmieszczania tych komponentów. Podstawową topologią jest pierścień: jednokierunkowa droga transmisji, bez wyraźnie określonego początku lub końca. W sieci Token Ring sam pierścień może być albo fizyczny, albo logiczny.
Wczesne implementacje Token Ringu oparte były na wiązkowych kablach stacji końcowych odgałęziających się od kabla dalekosiężnego. Cały pierścień składał się wyłącznie z takiego okablowania przyłączonego do równorzędnych stacji. W topologii przedstawionej na rysunku 9.7 występuje pierścień fizyczny.
Wykorzystanie w pierścieniu wzmacniaków, znanych także jako koncentratory lub jednostki MSAU, zaowocowało topologią fizycznej gwiazdy, będącą podstawą logicznego pierścienia. Topologia ta została przedstawiona wcześniej w tym rozdziale, na rysunku 9.1.
W sieci Token Ring o podstawowej topologii gwiazdy jednostka MSAU pełni rolę szkieletu; nie ma tu kabla dalekosiężnego. Stacje są przyłączane do portów TCU koncentratora za pomocą kabli stacji końcowej.
Topologię taką można rozszerzać, dodając koncentratory i okablowanie dalekosiężne. Rozszerzanie sieci Token Ring może być skomplikowanym przedsięwzięciem. Parametry wydajności takiej sieci są o wiele bardziej rygorystyczne niż w przypadku sieci Ethernet. Liczba stacji w domenie nie może być większa niż 260. To ograniczenie oraz czas, który może być potrzebny na obsłużenie takiej liczby urządzeń za pomocą jednego tokenu, mogą uczynić pozornie łatwą ekspansję przedsięwzięciem o wiele bardziej złożonym.

1.9.4.3 Dynamiczna **przynależność do pierścienia**

Ważnym aspektem normalnego działania sieci Token Ring jest możliwość obsługi dynamicznych zmian przynależności do sieci. Zmiany przynależności mogą mieć miejsce w dwóch przypadkach:
Rysunek 9.7. Topologia pierścienia fizycznego.
• zwykłego przyłączania i odłączania stacji,



• awarii sieci.

Aby zrozumieć mechanizm tego ważnego aspektu sieci Token Ring, należy przyjrzeć się procesom zmiany przynależności.

1.9.4.4 Przyłączanie stacji

Prosta czynność włączenia zasilania stacji nie powoduje automatycznego dołączenia jej do sieci. Zanim stacja zostanie zaakceptowana, musi przejść szereg testów połączeniowych. Testy te wykonywane są automatycznie, przy każdej inicjalizacji karty sieciowej Token Ring (czyli karty NIC).

Pierwszy z tych testów znany jest jako Lobe test, czyli test podłączenia stacji końcowej. Testowana stacja musi wysłać szereg testowych ramek sterowania dostępem. W ten sposób sprawdzana jest ciągłość fizyczna nośnika na odcinku między stacją a koncentratorem.

Jeśli ten test zakończy się powodzeniem, stacja musi następnie fizycznie włączyć się do pierścienia. W tym celu musi wysłać do koncentratora niskonapięciowy sygnał DC (prądu stałego), tzw. prąd pozorny (ang. *phantom current).* Sygnał ten ma wystarczająco niskie napięcie, żeby nie wpływał w żaden sposób na sygnały danych, które mogą przepływać przez kabel stacji końcowej. Stąd bierze się jego nazwa „prąd pozorny".

MSAU posiada obwody przekaźnikowe, zapobiegające wywoływaniu nieużywanych portów. Niestety, warstwa fizyczna nie pozwala na odróżnienie portu nieużywanego od portu używanego, ale przyłączonego do nieaktywnego urządzenia. Dlatego Token Ring wykorzystuje prąd pozorny do powiadomienia koncentratora, że urządzenie nieaktywne staje się aktywne. Koncentrator odpowiada, pozwalając stacji włączyć się fizycznie w elektroniczne obwody pierścienia poprzez wcześniej nieczynny port. Fizyczne włączenie musi nastąpić w ciągu 5 milisekund. Jest to maksymalny czas, przez jaki może występować przerwa w sieci, nie powodując uaktywnienia procesów naprawczych.

Proces fizycznego włączenia pozwala przyłączanej stacji określić prędkość sygnału w pierścieniu i sprawdzić, czy w sieci jest już obecny monitor aktywny (AM). Jeśli go nie ma, stacja staje się monitorem aktywnym po zakończeniu procesu przyłączania. Stacja uzyskuje tę wiedzę poprzez badanie sieci na obecność którejś z ramek zarządzania MAC, a dokładnie:

• Ramki monitor aktywny obecny (ang. AMP - Active Monitor Present)

• Ramki czyszczenie pierścienia (ang. PRG - Purge Ring)

Jeśli stacja wykryje obecność jednej z tych ramek, wie, że w pierścieniu działa monitor aktywny. Jeśli nie wykryje żadnej z nich, wysyła token zgłoszenia i rozpoczyna arbitraż, pełniąc obowiązki monitora aktywnego.

Następnym krokiem w procesie przyłączania stacji do pierścienia jest sprawdzenie, czy żadna inna stacja nie używa adresu stacji przyłączanej. Czynność tę wykonuje inna ramka zarządzania MAC, znana jako ramka Test Podwójnego Adresu (ang. DAT Duplicate Address Test). Jest to prosta ramka, zaadresowana sama do siebie. Jeśli powraca do swego nadawcy z potwierdzeniem otrzymanym od innej stacji, stacja nadawcza odłącza się od pierścienia, dezaktywując swoje połączenia i protokoły sieciowe.

Jeśli jednak ramka przejdzie przez cały pierścień i niezmieniona powróci do nadawcy, można być pewnym, że żadna inna stacja w pierścieniu nie wykorzystuje adresu stacji przyłączanej. Ale stacja ta wciąż nie przynależy do pierścienia! W kolejnym kroku musi ona

zidentyfikować swojego poprzedniego aktywnego sąsiada (ang. NAUN - Nearest Active Upstream Neighbor) i następnego aktywnego sąsiada (ang. NADN- Nearest Active Downstream Neighbor).
Te procesy identyfikacyjne pozwalają stacji określić jej miejsce w pierścieniu, dając jej względne punkty odniesienia. Przy czym mechanizm umożliwiający identyfikację sąsiadów nie jest właściwością specyficzną wyłącznie dla procesu przyłączania stacji do pierścienia. Jest to jeden ze stałych obowiązków monitora aktywnego. Regularnie wysyła on ramki zarządzania MAC „Monitor Aktywny Obecny" do następnego sąsiada w pierścieniu. Stacja ta akceptuje ramkę, ustawiając odpowiednie bity jej nagłówka: „Adres Rozpoznany" i „Ramka Skopiowana". Stacja zapamiętuje również adres monitora
aktywnego, a następnie wysyła ramkę „Monitor Oczekujący" obecny (ang. SMP Standby Monitor Present) do następnego sąsiada. Sąsiad powtarza ten proces. Efekt jest taki, że wszystkie stacje są na bieżąco informowane o każdej zmianie przynależności do sieci.
Ostatnim etapem uaktywniania stacji w pierścieniu jest proces żądania inicjalizacji. Przyłączana stacja żąda podania różnych parametrów operacyjnych pierścienia, wysyłając
ramkę MAC Żądanie Parametrów. Jej adresatem jest specjalny serwer, nazywany Serwerem Parametrów. Jeśli taki serwer nie istnieje lub jest niedostępny, stacja przyjmuje parametry domyślne. Po zakończeniu tego procesu stacja zostaje włączona do pierścienia.
Odłączanie stacji
Jeśli udało Ci się przebrnąć przez tę skomplikowaną sekwencję zadań związanych z przyłączaniem stacji, możesz teraz spokojnie odpocząć. Proces odłączania stacji jest o wiele prostszy! Jeśli z jakiegoś powodu stacja jest odłączana od sieci, port jednostki MSAU wykrywa brak prądu pozornego i jednostka MSAU automatycznie otwiera swoje przekaźniki, wyłączając port. Port ten wraz ze wszystkim, co może być do niego przyłączone, zostaje odizolowany elektrycznie od pierścienia. Ten prosty proces nie zakłóca działania pierścienia.

1.9.4.5 Awarie

Jeśli stacja wykryje w pierścieniu awarię, odpowiada, wysyłając znaki kierunkowe. Znak kierunkowy to specjalna ramka wysyłana przez stację, aby poinformować inne stacje o awarii. Stacja wysyłająca ramkę zna swoje względne położenie w pierścieniu i wykorzystuje potwierdzone znaki kierunkowe do określenia rozmiarów uszkodzenia. Przyjmuje się, że stacje, które nie potwierdzą ramki znaku kierunkowego, są odcięte przez awarię. W ten sposób protokół Token Ring umożliwia w pewnym stopniu automatyczną detekcję i naprawę awarii występujących na poziomie sieci.

1.9.4.6 Monitor aktywny

Wiele czynności zarządczych w sieci Token Ring wykonywanych jest przez jedną stację, znaną jako monitor aktywny (AM). Może nim być dowolna stacja w pierścieniu. Zwykle jest to pierwsza stacja, która została uaktywniona, ale po uaktywnieniu większej liczby stacji rola ta może zostać przekazana innej stacji. AM monitoruje cały ruch w sieci, zapewniając przestrzeganie reguł protokołu pierścienia. Jest także odpowiedzialny za inicjowanie wszelkich działań, które mogą być konieczne do przezwyciężenia problemów wynikłych z awarii lub naruszenia protokołu.
Wszystkie stacje sieci Token Ring mogą pełnić rolę monitora aktywnego, ale tylko jedna może być nim w danym momencie. Pozostałe stacje nazywa się monitorami oczekującymi (SM).
Do obowiązków monitora aktywnego należą:
• inicjalizacja pierścienia poprzez wysłanie ramki MAC czyszczenia pierścienia podczas uruchamiania
• tworzenie tokenów • taktowanie sieci
• zapewnianie, że ramki ani tokeny nie okrążą pierścienia więcej niż raz. Uzyskuje się je za pomocą odwrócenia bitu Monitor pola Sterowanie Dostępem Ramki/ Tokenu.
Monitor aktywny obsługuje także wiele innych funkcji zarządzania pierścieniem, w tym buforowanie opóźnień. Protokół Token Ring wymaga, żeby sieć LAN była wystarczająco duża, aby mogła pomieścić cały token. Niektóre z mniejszych sieci przy swoich prędkościach transmisji mogą być zbyt małe. Buforowanie opóźnień polega na wykorzystywaniu buforu pamięci do symulowania większej sieci.
Bufor opóźnienia musi mieć wielkość odpowiadającą minimum 24 czasobitom. Czaso bit to ilość czasu potrzebna na transmisję jednego bitu informacji. Bufor oferowany przez monitor aktywny jest znany jako założony bufor minimalnego czasu oczekiwania (ang. Assured Minimum Latency Buffer). Buforu o takiej wielkości wymaga protokół 802.5, by zagwarantować, że pierścień LAN jest wystarczająco duży, aby mógł pomieścić jeden token. Skoro token ma długość 24 bitów, a prędkość przesyłania sygnału może wynosić 4 lub 16 Mbps, pierścień musi być wystarczająco duży, żeby czas jednego przebiegu (okrążenia) wynosił odpowiednio 250 nanosekund lub 62,5 nanosekundy. Nie powinno to być problemem, ale by zapobiec nieprzewidzianym zachowaniom ekstremalnie
małych pierścieni, monitor aktywny odpowiada za zapewnienie odpowiedniego buforowania.

1.9.4.7 Wybór nowego monitora aktywnego

Proces wyboru monitora wykorzystuje ramkę MAC zgłoszenia tokenu, aby przyznać którejś ze stacji rolę monitora aktywnego. Proces ten jest inicjowany przez dowolną stację monitora oczekującego, gdy wykryje ona prawdopodobną awarię monitora aktywnego. Proces ten mogą wyzwolić liczne symptomy. Mogą nimi być:
• niepowodzenia monitora aktywnego przy próbie transmisji dobrej ramki Token, występujące przynajmniej raz na 2,6 sekundy
• niepowodzenia monitora aktywnego przy próbie transmisji ramki MAC Monitor aktywny obecny, występujące przynajmniej raz na 15 sekund
• nieudane próby oczyszczenia pierścienia • niezgodność z zegarem
• nieobecność ramek MAC Monitor aktywny obecny lub Monitor oczekujący obecny po przyłączeniu stacji do sieci. Normalnie takie ramki są wysyłane w ciągu 18 sekund od udanego przyłączenia stacji.

Również inne czynniki mogą uruchomić proces wyboru nowego monitora aktywnego, ale powyższe przykłady powinny odpowiednio zilustrować, w jaki sposób stacje SM
stale monitorują pracę monitora AM. Niewłaściwe wykonywanie obowiązków przez monitor AM może być spowodowane wieloma czynnikami, takimi jak zanik zasilania w monitorze AM, uszkodzenia sprzętu lub oprogramowania czy też po prostu fizyczne odłączenie monitora od sieci.
Niezależnie od przyczyny, w wypadku, gdy monitor oczekujący wykryje niewłaściwe działanie monitora aktywnego, natychmiast generuje ramkę MAC zgłoszenia tokenu. Następnie wszystkie stacje przystępują do wyboru nowego monitora aktywnego.
Proces wyboru monitora aktywnego wymaga, by każda stacja wysłała ramkę MAC token zgłoszenia do wszystkich innych stacji w pierścieniu. Adresy nadawców tych ramek są porównywane co do wartości liczbowej z adresami MAC stacji odbierających w celu znalezienia stacji mającej adres o najmniejszej wartości liczbowej. Jeśli adres stacji jest mniejszy od adresu otrzymanej ramki zgłoszenia tokenu, stacja wysyła dalej tę ramkę i przestaje wysyłać własną. Jeśli jednak adres stacji jest większy niż adres otrzymanej ramki, stacja odrzuca otrzymaną ramkę i wysyła własną. Tym sposobem nowy AM jest wybierany na podstawie uzgodnienia numerycznych adresów stacji.

1.9.5 Co dalej z Token Ringiem?

Token Ring - ostatnimi laty przyćmiony przez nowsze i/lub szybsze, cieszące się większym rozgłosem architektury LAN - zaczyna wychodzić z cienia. Takie odrodzenie stało się możliwe dzięki wykorzystaniu technologii komutujących do przesyłania ramek, a także dzięki licznym próbom dalszego zwiększania prędkości sygnału w sieci Token Ring.

1.9.5.1 Przełączanie a dedykowane sieci Token Ring

Najważniejszym ulepszeniem jest przełączanie na poziomie portu. Zamiast koncentratora wzmacniającego, do łączenia urządzeń w pierścieniu wykorzystuje się przełącznik. Każdy port ma własne dedykowane pasmo. Rywalizacja o tokeny ogranicza się do dwóch urządzeń: stacji i przełączanego portu w koncentratorze, do którego stacja jest przyłączona. Wykorzystanie odpowiedniego nośnika fizycznego umożliwia temu dedykowanemu połączeniu obsługę pełnodupleksowej transmisji. Innymi słowy, przełączany port i połączona z nim stacja mogą jednocześnie wysyłać i odbierać dane oddzielnymi ścieżkami przewodów.
Pełnodupleksowe technologie przełączania portów stanowią podstawę dedykowanego Token Ringu (ang. *DTR - Dedicated Token Ring)*. W sieci DTR ciężar spoczywa na przełączniku, który musi tworzyć tablice śledzące adresy MAC i korelujące je z numerami przełączanych portów. Każda ramka odebrana przez przełącznik wyzwala mechanizm przeszukiwania tablicy w celu określenia odpowiedniej ścieżki, na którą należy przełączyć ramkę. Jest to ogromne ulepszenie w stosunku do tradycyjnej, półdupleksowej transmisji we wspólnym Token Ringu.
W tym kontekście DTR oznacza „dedykowany Token Ring". Każdy jednak, kto kiedykolwiek używał modemu i zadał sobie trud przyjrzenia się ładnym światełkom na jego płycie czołowej, wie że o wiele bardziej popularnym rozwinięciem skrótu DTR jest Data Terminal Ready Terminal danych gotów.

1.9.5.2 Zwiększanie szybkości transmisji

Podjęto kilka prób przyspieszenia transmisji w sieci Token Ring z 16 Mbps do 100 Mbps lub nawet powyżej tej wartości. Według jednej z propozycji prędkość sygnału ma wynosić 128 Mbps. Inne mówią o 100 Mbps, przy czym warstwa fizyczna mogłaby obsłużyć nawet do 1 Gbps!
Odpowiedzialność za stworzenie standardu High Speed Token Ring spoczywa na komitecie IEEE 802.5. Podzieli on całe zadanie na trzy części składowe, które będą przedstawiane publicznie, w miarę kończenia poszczególnych etapów.

1.9.5.2.1 100 Mbps przy wykorzystaniu nośników miedzianych

Pierwsza część składowa zadania będzie miała na celu zdefiniowanie standardu sieci Token Ring 100 Mbps z okablowaniem miedzianym. Plan działania zakłada przeniesienie istniejących rozwiązań, tak aby jak najszybciej można było skorzystać z pierwszych owoców pracy nad standardem High Speed Token Ring. Tak więc podstawą warstwy fizycznej sieci Token Ring 100 Mbps będzie 1 OOBaseTX. Warstwa łącza danych 802.5 zostanie przeszczepiona na interfejs międzynośnikowy warstwy fizycznej 802.3, aby utworzyć pierwszy High Speed Token Ring.
Na pierwszy rzut oka taka hybrydyzacja może wydać się herezją, zwłaszcza ludziom od dawna przywiązanym do Ethernetu bądź Token Ringu. Pamiętajmy, że modularność standardów IEEE była zamierzona; nie ma wzajemnych zależności między nimi. Chociaż 100BaseTX został zaprojektowany, aby obsługiwać warstwę łącza danych sieci Ethernet, nie został z nią tak silnie sprzężony, aby nie mógł współpracować z innymi specyfikacjami warstwy łącza danych.
Jak wyjaśniono w przeglądzie zamieszczonym w rozdziale 8 pt. „Szybsze sieci Ethernet", 100BaseTX umożliwia transmisję poprzez kabel UTP Kategorii 5 z prędkością 100 Mbps na odległość do 100 metrów. Ma wbudowany mechanizm automatycznego uzgadniania prędkości, który pozwala zmniejszyć prędkość, jeśli z jakiegoś powodu dane połączenie nie może podtrzymać prędkości maksymalnej.
Specyfikacje 100BaseTX i 100BaseFX zostały stworzone podczas opracowywania rozszerzenia standardu 802.3 CSMA/CD do prędkości 100 Mbps. Rozszerzenie to jest lepiej znane jako Fast Ethernet. Więcej informacji o Fast Ethernecie, 100BaseTX czy 100BaseFX można znaleźć w rozdziale 8 pt. „Szybsze sieci Ethernet".

1.9.5.2.2 100 Mbps przy wykorzystaniu światłowodu

Drugą specyfikacją będzie Token Ring 100 Mbps, wykorzystujący jako medium transmisyjne kabel światłowodowy. Podobnie jak jej oparte na przewodach miedzianych „rodzeństwo", ta specyfikacja Token Ringu będzie oparta na 100BaseFX.

1 Gbps

Na koniec komitet 802.5 skoncentruje się na wersji Token Ringu 1 Gbps. Prawdopodobnie także w tym przypadku będzie wiele zapożyczeń z wykonanych lub właśnie wykonywanych prac nad przyspieszeniem prędkości sygnału w sieci Ethernet. Oczekuje się, że Token Ring 1 Gbps będzie wykorzystywał światłowód, ale wersja z przewodami miedzianymi nie będzie mu znacznie ustępować.

Token Ring 1 Gbps, tak jak Gigabit Ethernet, wydaje się być bardziej atrakcyjną technologią niż przejście z 16 na 100 Mbps. Na razie nie jest jasne, jak duży będzie zakres niezbędnej nowelizacji warstwy łącza danych. Wydaje się nieprawdopodobne, żeby w bliskiej przyszłości pojawiły się karty sieciowe 4/16/100/1 Gbps, oferujące automatyczne dopasowywanie prędkości. Różnice w warstwie fizycznej, a także w części sterowania dostępem do nośnika warstwy łącza danych, konieczne dla obsługi prędkości 1 Gbps, wydają się być zbyt wielkie, żeby jedna płytka drukowana (ang. PCB printed circuit board mogła obsługiwać także niższe prędkości.

1.9.5.2.3 Będzie działać?

Choć perspektywa zwiększenia prędkości sygnału i zastosowania technologii komutujących wygląda niezwykle obiecująco, pozostaje uporczywe pytanie: „Czy to będzie działać?". Przełączanie portów, choć jest wykorzystywane przy obsłudze dedykowanego Token Ringu, może w rzeczywistości uczynić High Speed Token Ring rozwiązaniem dyskusyjnym. Oryginalną przewagą sieci Token Ring nad siecią Ethernet było wykorzystywanie uporządkowanej, deterministycznej metodologii dostępu do nośnika. Token Ring mógł zawsze wykorzystywać większą część dostępnego pasma niż Ethernet. Jak wyjaśniono w rozdziale 7, wprowadzenie pełnodupleksowych sieci Ethernet z przełączanym portem zmieniło tę sytuację. W dupleksowych przełączanych łączach Ethernetu nie ma rywalizacji o dostęp do pasma. Urządzenie nadawcze może wprowadzać ramki do sieci niemal bez opóźnień. Dziś sieć Ethernet może wykorzystywać ponad 98% możliwej prędkości sygnału. Dlatego różnice między Ethernetem 100 Mbps a Token Ringiem 100 Mbps stają się znacznie bardziej subtelne, a nawet subiektywne.

Podstawowa strategia wprowadzania produktów obsługujących kilka prędkości również została zastosowana dla produktów 802.3. Zamiarem wytwórców jest produkowanie dla bazie Token Ringu urządzeń 4/16/100 Mbps, mogących automatycznie ustalać najwyższą możliwą prędkość sygnału, jaką jest w stanie obsłużyć dany światłowód lub kabel miedziany. Sztuką będzie rozszerzenie tego zakresu do prędkości 1 Gbps.

Wielu wytwórców LAN ma nadzieję, że wysiłki te powstrzymają odpływ klientów, którzy wybierają Ethernet, uważając, że Token Ring nie ma przed sobą przyszłości. Dostarczając klientom perspektyw, można zdziałać nieco więcej, niż tylko zmniejszyć ich odpływ do Ethernetu.

Również inne problemy zagrażają rynkowej akceptacji standardu High Speed Token Ring. Najważniejszymi z nich są technologie FDDI i przełączania portów. FDDI jest deterministyczną, bazującą na pierścieniu architekturą LAN, działającą z prędkością 100 Mbps. FDDI zawsze będzie wydajniejsze od Token Ringu I00 Mbps ze względu na stosowane w nim rozwiązanie tzw. szybkiego wyzwalania. Szybkie wyzwalanie, jak jest to wyjaśnione w rozdziale 10 pt. „FDDI", pozwala nadającej stacji zrezygnować z kontroli nad medium transmisyjnym dzięki generowaniu nowego tokenu natychmiast po wysłaniu ramki danych. Dlatego następne urządzenie w pierścieniu może otrzymać token i rozpocząć transmisję własnych danych, zanim jeszcze usunięta zostanie z sieci pierwsza ramka danych.

Podsumowując - z teńhnicznego punktu widzenia istnieją tylko drobne wątpliwości co do tego, czy High Speed Token Ring będzie działał. Nie wiadomo jednak, czy będzie to sieć praktyczna, efektywna ekonomicznie i co najważniejsze, czy zostanie zaakceptowana przez rynek.

1.9.6 Podsumowanie

Token Ring długo był uważany za solidniejszą i doskonalszą technicznie architekturę LAN niż Ethernet. Jednak ostatnio nie został w takim stopniu znowelizowany jak standard 802.3. W rezultacie ucierpiała na tym jego pozycja na rynku. Ostatnie próby odnowienia tej starzejącej się architektury wyglądają obiecująco, ale tylko czas rozstrzygnie, czy nie są to próby zbyt słabe i spóźnione, aby uratować Token Ring.

Niektóre rozwiązania Token Ringu nadal są wartościowe, choć są nieco tłumione przez postęp technologiczny. Token Ring wciąż oferuje lepszy mechanizm nadawania priorytetów dostępu niż mechanizm stosowany w sieciach Ethernet. Lepsze jest również wykrywanie i korygowanie uszkodzeń. Korzystniejsza jest także proporcja części użytecznej danych zawartych w ramce do pozostałej części ramki. Pytanie, czy mechanizmy te wystarczają, aby uzasadnić dalszy rozwój i istnienie tego standardu, może być przedmiotem bardzo emocjonalnej debaty.

Pomimo niepewności co do perspektyw, Token Ring w swojej obecnej formie posiada wiele zalet, ale również wiele ograniczeń.

1.9.7 Zalety Token Ringu

Współdzielony Token Ring posiada wiele zalet w porównaniu z innymi architekturami LAN. Na przykład, współdzielony Token Ring może zaoferować wysoce deterministyczną wydajność dzięki temu, że nie wykorzystuje rywalizacji jako metody dostępu. Można obliczyć maksymalny czas, jaki mija od momentu, kiedy stacja chce nadawać, do momentu, gdy otrzymuje token umożliwiający transmisję. Ten czas dostępu można zmniejszyć w przewidywalnym stopniu poprzez zmniejszenie liczby urządzeń w pierścieniu.

Uporządkowana metodologia dostępu daje także inne korzyści. Inaczej niż we współdzielonym Ethernecie, który może być tak obciążony kolizjami, że będzie wykorzystywany zaledwie w 20%, współdzielony Token Ring działa wydajniej przy większym obciążeniu sieci. Jeśli obciążenie zbliża się do maksymalnej, obsługiwanej wielkości, wydajność spada, ale w sposób przewidywalny i kontrolowany, co jest bardzo miłe dla użytkowników, zwłaszcza w porównaniu z analogiczną sytuacją w sieci Ethernet.

Token Ring wyróżnia się także jeśli chodzi o monitorowanie działania sieci. Specyfikacja jego warstwy fizycznej dostarcza kilku ważnych mechanizmów. Są to m.in. agenci zarządzania stacją (SMT), zajmujący się zbieraniem danych i raportowaniem. Istnieją również

mechanizmy automatycznego wykrywania awarii sprzętu i informowania o nich innych stacji w pierścieniu. Warstwa fizyczna dostarcza także kilku mechanizmów dostrajania działania pierścienia (trzeba tylko wiedzieć, jak z nich korzystać!).
Wreszcie Token Ring może obsługiwać ramki o rozmiarach do 18 kB. W sieciach, których aplikacje charakteryzują się dużymi transferami plików, oznacza to wyjątkowo mały stosunek narzutu (część ramki nie zawierająca danych) do części użytecznej (przesyłanych danych).

1.9.8 Ograniczenia Token Ringu

Token Ring posiada również wady. Jak wspomniano w poprzednim punkcie, dostrajanie działania Token Ringu wymaga dogłębnego zrozumienia protokołu. Ponieważ jednak możliwości dostrajania są większe niż w innych protokołach, ograniczenie to jest nieco mniej dokuczliwe.
Nieco bardziej znaczącym ograniczeniem jest mała liczba urządzeń obsługiwanych przez Token Ring. Podczas gdy Ethernet może obsłużyć do 1024 urządzeń, Token Ring ogranicza ich liczbę do 260.

1.10 Rozdział 10 FDDI

Mark A. Sportack
Jedną ze starszych i solidniejszych technologii LAN jest interfejs danych przesyłanych światłowodowo, czyli interfejs FDDI (ang. Fiber Distributed Data Interface). Standard FDDI został znormalizowany w połowie lat 80. jako specyfikacja ANSI X3T9.5. W tym czasie zaczynały się pojawiać wysokowydajne UNIX-owe stacje robocze. Potrzebowały one sieci o większej wydajności niż ta, którą oferowały sieci będące ówcześnie na rynku. Zmobilizowało to instytut ANSI do opracowania specyfikacji odpowiedniej sieci lokalnej.
W miarę dojrzewania środowiska sieci lokalnych, z jednorodnej sieci zaczęły wyłaniać się różne obszary funkcjonalne. Każdy z nich obsługiwał określone zadanie: przyłączalność serwera, przyłączalność stacji roboczych, łączenie ze sobą koncentratorów itd. We wszystkich tych obszarach wzrastały wymagania dotyczące przepustowości. FDDI, ze swoją dużą szybkością transmisji danych i potencjalną niezawodnością działania, stało się naturalnym wyborem dla łączenia serwerów, a także dla łączenia ze sobą koncentratorów w szkielecie sieci LAN.
W tym rozdziale omawiany jest interfejs FDDI, jego nośniki fizyczne, ograniczenia dotyczące odległości, struktury ramek, mechanika i korzyści wynikające z jego stosowania. Stanowi to tło dla rozważań na temat roli FDDI w obecnych i przyszłych środowiskach sieciowych.

1.10.1 FDDI

FDDI jest akronimem nazwy Fiber Distributed Data Interface, ale tej długiej nazwy nikt nie używa. W zasadzie większość ludzi nawet nie wymawia poszczególnych liter F-D-D-I; łączą głoski i wymawiają je jako „fidi". FDDI jest solidną i niezawodną technologią sieci LAN, której powstanie datuje się na połowę lat 80. Cechuje się ona szybkością transmisji danych 100 Mbps i dwoma przeciwbieżnymi pierścieniami. Pierścienie mogą mieć rozpiętość do 200 kilometrów i wykorzystują kable światłowodowe. Dostęp do nośnika jest regulowany przez przekazywanie tokenu, podobnie jak w sieci Token Ring. Token może się poruszać tylko w jednym kierunku.
W wypadku awarii sieci, wzmacniaki i/lub stacje są wstanie wykryć uszkodzenie, określić obszar sieci, z którym utracono łączność, i automatycznie (ale tylko logicznie; nie fizycznie) połączyć razem obydwa pierścienie. Jest to tzw. zawijanie (ang. wrapping lub wrap-around), przywracające łączność w możliwie największej części sieci.
Zdolność autonaprawy i duża szybkość transmisji danych czynią FDDI jedyną technologią LAN odpowiednią dla aplikacji wymagających dużej przepustowości i/lub wysokiej niezawodności. Stan taki utrzymuje się od ponad 10 lat. Każda sieć lokalna, której zadaniem jest przesyłanie danych z szybkością ponad 16 Mbps, musi korzystać z FDDI. Dla każdej sieci, która nie może pozwolić sobie na przestoje, jedyną rozsądną opcją jest FDDI. Niestety, ponieważ medium transmisyjnym FDDI jest światłowód, jest to także opcja najdroższa. Ogranicza to implementację FDDI do najbardziej wyspecjalizowanych środowisk, wymagających dużej przepustowości lub niezawodności.
Opracowano już inne technologie LAN, mogące osiągać prędkość 100 Mbps lub większą. Konkurenci - ATM i Fast Ethernet - byli w stanie dorównać lub przewyższyć FDDI pod względem szybkości transmisji. Wymusiło to znaczną obniżkę cen FDDI. Dziś FDDI nie jest już elitarną technologią, jaką było dawniej. Wciąż jest to technologia dość wyspecjalizowana, ale dość powszechnie pojawiająca się w mieszanych topologiach sieci lokalnych. Jest używana przede wszystkim do łączenia serwerów z wieloprotokołowymi przełączanymi koncentratorami, a także do łączenia przełączanych koncentratorów w szkielecie sieci LAN.

1.10.1.1 Składniki funkcjonalne

FDDI obejmuje cztery odrębne składniki funkcjonalne. Każdy z nich jest określany przez własną serię specyfikacji. Składnikami tymi są:
• Sterowanie dostępem do nośnika (ang. MAC- Media Access Control) • Protokół warstwy fizycznej (ang. PHY- Physical Layer Protocol)
• Nośnik warstwy fizycznej (ang. PMD - Physical Layer Medium) • Zarządzanie stacją (ang. SMT- Station Management)
Rysunek 10.1 przedstawia porównanie powyższych składników z modelem referencyjnym OSI.
Sterowanie dostępem do nośnika
Jak widać na rysunku 10.1, najwyższą warstwą FDDI jest sterowanie dostępem do nośnika (MAC). Jest ona równoważnikiem warstwy łącza danych w modelu referencyjnym OSI. Podwarstwa MAC jest odpowiedzialna za określanie metodologii dostępu do nośnika oraz definiowanie wielu formatów ramek. Dodatkowo odpowiada również za generowanie tokenu i ramki, zarządzanie nimi, adresowanie fizyczne MAC, a nawet za przeprowadzanie detekcji i korekcji błędów przy odbiorze ramek danych.
Rysunek 10.1. Zbiór protokołów FDDI kontra model referencyjny OSI.

1.10.1.1.1 Protokół warstwy fizycznej

<table>
<tr><th>Nazwa warstwy modelu referencyjnego OSI</th><th>Numer warstwy OSI</th><th colspan="2">FDDI</th></tr>
<tr><td>Aplikacji</td><td>7</td><td colspan="2" rowspan="5">Brak określonych specyfikacji</td></tr>
<tr><td>Prezentacji</td><td>6</td></tr>
<tr><td>Sesji</td><td>5</td></tr>
<tr><td>Transportu</td><td>4</td></tr>
<tr><td>Sieci</td><td>3</td></tr>
<tr><td>Łącza danych</td><td>2</td><td>MAC</td><td rowspan="2">SMT</td></tr>
<tr><td>Fizyczna</td><td>1</td><td>PHY
PMD</td></tr>
</table>

Protokół warstwy fizycznej (PHY) FDDI odpowiada górnej podwarstwie warstwy fizycznej modelu referencyjnego OSI. Odpowiada za przyjmowanie bitowego strumienia danych i przekształcanie go na format bardziej odpowiedni do transmisji. Proces ten nosi nazwę „kodowania" (ang. encoding). Wykorzystywany jest schemat kodowania 4 bity/5 bitów. Schemat ten przyjmuje 4-bitowe półbajty z warstwy MAC i każdy z nich koduje jako 5-bitowy znak. Ten właśnie znak jest transmitowany. Należy zauważyć, że skoro warstwa MAC jest odpowiedzialna za generowanie ramek i umieszczanie w nich danych, to każda cząstka ramki jest kodowana w 5-bitowe znaki.
Warstwa PHY odpowiada także za taktowanie sieci LAN. FDDI jest taktowane częstotliwością 125 MHz. Warstwa PHY generuje sygnał taktujący transmisję i synchronizuje go we wszystkich stacjach przyłączonych do sieci.

1.10.1.1.2 Medium transmisyjne warstwy fizycznej

Medium transmisyjne warstwy fizycznej (PMD) określa wszystkie atrybuty nośnika, czyli: • Rodzaj nośnika
• Poziom sygnału transmisyjnego • Dopuszczalny poziom błędów • Rodzaje złączy fizycznych
Pierwotnie FDDI wykorzystywało tylko jeden nośnik warstwy fizycznej (PMD): wielofunkcyjny kabel światłowodowy o średnicy 62,5/125 mikrona. Do początku lat 90. FDDI opierało się wyłącznie na technologii światłowodowej. Wtedy wysoki koszt kabla światłowodowego zaczął niekorzystnie wpływać na udział FDDI w rynku. Odpowiedź wydawała się oczywista: należało opracować wykorzystujący przewody miedziane nośnik PMD, który mógłby obsługiwać protokoły FDDI.
W czerwcu 1990 r. ANSI sformowało komitet roboczy, który miał opracować specyfikację skrętki dwużyłowej PMD (ang. *TP-PMD).* Oryginalnie specyfikacja TP-PMD była zastrzeżonym produktem, który przenosił warstwę 2 FDDI na warstwę fizyczną nieekranowanej skrętki dwużyłowej (UTP) Kategorii 5. Produkt końcowy otrzymał nazwę interfejsu przesyłania danych przewodem miedzianym, czyli interfejsu CDDI (ang. *Copper Distributed Data Interface).* Specyfikacja ta stała się standardem ANSI w roku 1994.
Opracowano również jednofunkcyjną wersję światłowodu (ang. *SMF-PMD).* Jest ona znacznie droższa niż jej wielofunkcyjny odpowiednik, gdyż wykorzystuje kabel o średnicy 8,3 mikrona oraz laser zamiast diody świecącej (LED). Na jej korzyść przemawia jednak fakt, że może zapewnić integralność sygnału na dużo większych odległościach do 60 kilometrów, w porównaniu z 2 kilometrami dla wersji wielofunkcyjnej.
Chociaż termin „laser" wszedł do powszechnego użytku jako rzeczownik, właściwie jest to akronim. Opisuje on fizyczny proces, w którym powstaje związana z laserami skoncentrowana energia. Akronim pochodzi od nazwy „Light Amplification through Stimulated Emission of Radiation" - Wzmocnienie wiązki światła poprzez wymuszoną emisję promieniowania. „Laser" jest jednak określeniem o wiele prostszym.

1.10.1.1.3 Zarządzanie stacją (SMT)

Zarządzanie stacja (SMT) jest oddzielnym modułem, obejmującym pełny zestaw protokołów FDDI. Komunikuje się bezpośrednio z warstwami MAC, PHY i PMD, aby monitorować i zarządzać działaniami stacji i pierścienia. Specyfikacja ANSI X3T9.5 definiuje trzy obszary funkcjonalne SMT:
• Obsługa ramek SMT,
• Sterowanie połączeniem,
• Sterowanie pierścieniem.
Razem obszary te obejmują wiele różnych usług, istotnych dla normalnego działania stacji i pierścienia FDDI; najważniejszymi z nich są:
• Przyłączanie stacji,
• Odłączanie stacji, • Zbieranie statystyk,
• Identyfikacja uszkodzeń, • Naprawa uszkodzeń.
Choć dana stacja może mieć wiele wystąpień warstw MAC, PHY i PMD (co jest zwykłą sytuacją w przypadku podwójnie przyłączanych stacji), moduł SMT może mieć tylko jedno.

1.10.2 Tworzenie sieci FDDI

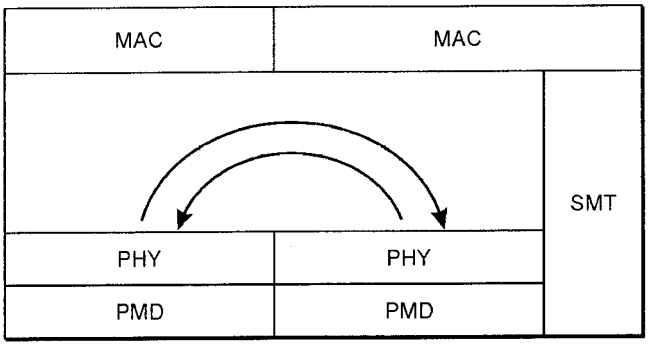
Przyjęło się uważać, że sieć FDDI ma topologię podwójnego, przeciwbieżnego pierścienia. Prawda jest taka, że istnieje kilka różnych sposobów konstruowania sieci FDDI. Podwójny pierścień jest tylko jedną z wielu form. Aby budować bardziej efektywne sieci FDDI, trzeba poznać różne rodzaje portów i sposoby przyłączania stacji do sieci.

1.10.2.1 Typy portów i metody przyłączania

FDDI rozpoznaje cztery różne typy portów:
• Port A: podstawowe wejście, dodatkowe wyjście, • Port B: podstawowe wyjście, dodatkowe wejście, • Port M: główny (master) port koncentratora,
• Port S: podporządkowany (slave) port dla pojedynczo przyłączanych urządzeń. Wymienione rodzaje portów mogą być łączone ze sobą na różne sposoby. Zanim tego spróbujesz, powinieneś poznać różne rodzaje obsługiwanych połączeń. Dwie podstawowe metody używane do przyłączania urządzeń FDDI do sieci to:
• Podwójne przyłączenie, • Pojedyncze przyłączenie.
W metodach tych można używać wzmacniaków. Przyłączenia mogą być dokonywane przy różnych konfiguracjach portów. Zwiększa to różnorodność i funkcjonalność metod budowy i wykorzystywania sieci lokalnych FDDI.

1.10.2.1.1 Stacje podwójnie przyłączane

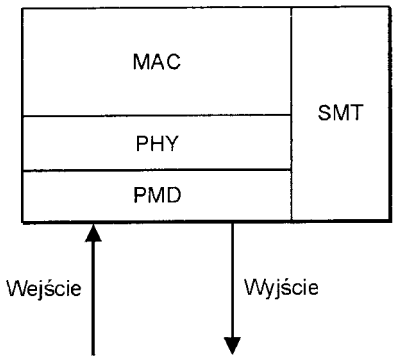
Stacje podwójnie przyłączane (ang. DAS- Double-attached Stations) mają dwa zestawy interfejsów nośnika. Pozwala to fizycznie przyłączyć urządzenie DAS do każdego z dwóch pierścieni FDDI. Rysunek 10.2 przedstawia sposób przyłączenia stacji DAS do sieci. Każde urządzenie DAS ma dwa zestawy portów interfejsu nośnika, z których każdy zawiera porty A i B. Każdy port posiada fizyczne złącza dla dwóch nośników fizycznych. Tak więc do urządzenia DAS przyłącza się cztery kable światłowodowe.
Koncentrator (ang. concentrator) to urządzenie, które grupuje wiele połączeń sieci LAN na wspólnej płytce elektrycznej. Najpowszechniejszym typem koncentratora LAN jest tzw. hub (w języku polskim nazywany po prostu koncentratorem). Koncentratory także mogą być podwójnie przyłączane. Dlatego poprawne jest stosowanie zwrotu „podwójnie przyłączane" (DA) zarówno w odniesieniu do koncentratorów, jak i do stacji, bez wyszczególniania urządzeń.
Rysunek 10.2. Stacja podwójnie przyłączana.



Jak pokazano na rysunku 10.2, fizyczne urządzenie staje się integralną częścią dwóch pierścieni, ponieważ karta sieciowa (NIC) zapewnia fizyczną ciągłość dwóch pierścieni między portami A i B. Łącząc stacje DAS, można stworzyć równoprawną (ang. peer-to-peer
) sieć lokalną, nie wykorzystując wzmacniaków. W tym celu należy port A jednego urządzenia połączyć z portem B drugiego urządzenia i na odwrót. Wadą takiego rozwiązania jest fakt, że każde urządzenie DAS musi mieć włączone zasilanie, a także musi funkcjonować, aby pierścienie były kompletne. FDDI może zawijać uszkodzony fragment pierścienia, ale wpływa to bezpośrednio na wydajność całego pierścienia. Co więcej, jeśli wiele stacji zostanie jednocześnie wyłączonych (z zasilania) lub w inny sposób przestaną działać, sieć może zmienić się w dwie lub więcej mniejszych par pierścieni.

1.10.2.1.2 Stacje pojedynczo przyłączane

Stacje pojedynczo przyłączane (ang. SAS - Single-attached Stations) unikają potencjalnych problemów działania związanych ze stacjami DAS dzięki temu, że nie stosują zawijania. Każde urządzenie SAS ma tylko jeden interfejs komunikacyjny, „S"; z dwoma portami nośników. Do nadawania i odbioru używa się dwóch oddzielnych światłowodów. Końce obydwu są przyłączone do koncentratora, który zapewnia połączenie z obydwoma pierścieniami. Rysunek 103 przedstawia pojedynczo przyłączoną stację i jej koncentrator.
Rysunek 10.3. Stacja pojedyncz-o przyłączana.



1.10.2.1.3 Prawidłowe połączenia portów

Aby zdobyć kompletną wiedzę o różnych rodzajach połączeń w sieciach FDDI, należy uzupełnić opis rodzajów portów i metod przyłączania krótkim przeglądem prawidłowych połączeń portów. Wszystkie prawidłowe kombinacje połączeń portów są przedstawione w tabeli 10.1.
Tabela 10.1.
Prawidłowe kombinacje połączeń portów.
Odwrócenie kombinacji połączeń portów przedstawianych w tabelach 10.1 i 10.2 niczego nie zmienia - przykładowo B i A również jest dozwoloną kombinacją połączeń portów, nie różniącą się funkcjonalnie od opisanej w tabeli kombinacji A i B.

Kombinacja połączeń portów	Zastosowanie
A i B	Połączenie urządzeń DAS w podwójnym pierścieniu bez wzmacniaka.
A i M.	Połączenie urządzenia DAS z koncentratorem w konfiguracji podwójnego przyłączenia.
B i M	Połączenie urządzenia DAS z koncentratorem w konfiguracji podwójnego przyłączenia.
M i S	Połączenie urządzenia SAS z koncentratorem.
S i S	Równoprawne połączenie urządzeń SAS.

W tabeli 10.2 przedstawione są kombinacje połączeń portów, które uważa się za niepożądane. Kombinacje takie co prawda działają, lecz nie są optymalne.
Tabela 10.2.
N Iepożądane kombinacje połączeń portów.

Kombinacja połączeń portów	Konsekwencje
A i A	Można ustanowić połączenie równoprawne, ale skutkuje to poplątanymi pierścieniami.
B i B	Można ustanowić połączenie równoprawne, ale skutkuje to poplątanymi pierścieniami.
A i S	Jest to tzw. pierścień zawijany. Można go utworzyć „ręcznie", ale lepiej pozostawić to protokołom FDDI, które stworzą logiczne połączenie.
B i S	Jest to tzw. pierścień zawijany. Można go utworzyć „ręcznie", ale lepiej pozostawić to protokołom FDDI, które stworzą logiczne połączenie.

Więcej informacji o pierścieniach zawijanych można znaleźć w następnym punkcie pt. „Topologie i implementacje".
Jedyną kombinacją połączeń portów, która jest nieprawidłowa i niedozwolona, jest połączenie M i M. Tworzy ona tzw. „pierścień drzew", który nie jest zbyt użyteczny. Termin ten zostanie wyjaśniony w następnym punkcie pt. „Topologie i implementacje".

1.10.2.2 Topologie i implementacje

Opisane wcześniej rodzaje portów i metody przyłączania występują w różnych odmianach na poziomie topologii i implementacji. Wbrew uparcie pokutującemu mitowi, FDDI to nie tylko podwójne przeciwbieżne pierścienie. Jest to, być może, najważniejsza topologia, ale istnieje także wiele innych, użytecznych topologii i implementacji. Niektórymi spośród najpowszechniej spotykanych odmian sieci FDDI są:'
• Podwójny pierścień
• Podwójny pierścień z drzewami • Pojedyncze drzewo

• Podwójne kierowanie docelowe • Cykliczne zawijanie

Pierwsze cztery topologie cechują się różnymi zaletami i ograniczeniami działania. Piąta topologia - cykliczne zawijanie - występuje tylko w wypadku awarii sieci.

Podwójny pierścień

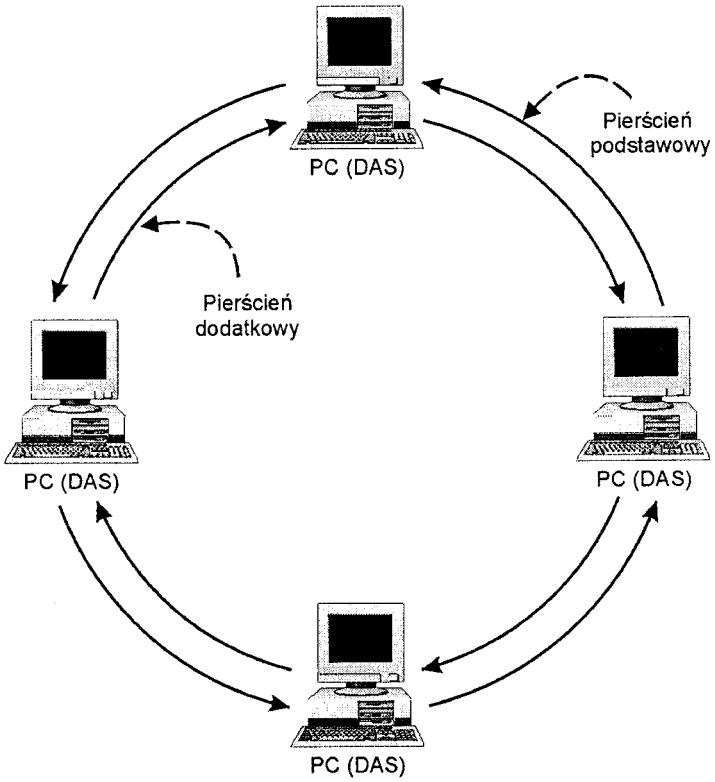
Podstawową topologię podwójnego pierścienia, czasem nazywaną „podwójnym pierścieniem bez drzew", tworzą podwójnie przyłączane stacje, bezpośrednio połączone jedna z drugą. Tworzy to parę równoprawnych pierścieni, zilustrowaną na rysunku 10.4.

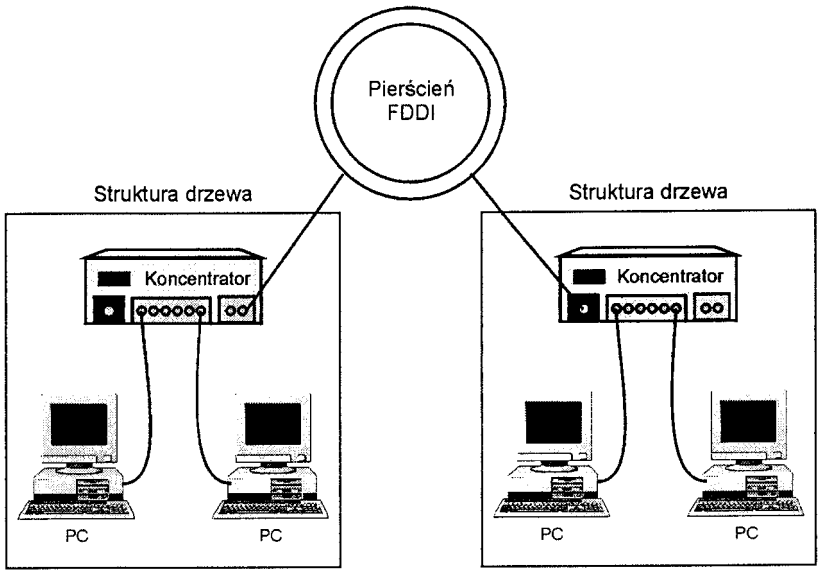
Wady rozwiązania polegającego na uzależnieniu funkcjonowania całego pierścienia od wszystkich stacji go tworzących powinny być oczywiste. Pierścienie w równym stopniu zależą od każdego z urządzeń. Jeśli któreś z nich zostanie wyłączone lub z innych przyczyn przestanie działać, pierścienie fizyczne są zagrożone. FDDI wykrywa i automatycznie ogranicza rozmiar uszkodzenia, ale faktem jest, że ryzyko jest nieodłącznie związane z tą topologią. Jej stosowanie powinno ograniczać się do małych, wysoko wyspecjalizowanych środowisk.

Podwójny pierścień z drzewami

Topologia podwójnego pierścienia z drzewami jest rozwinięciem topologii podwójnego pierścienia. Cechują ją drzewopodobne „wyrostki" odchodzące od podwójnych pierścieni FDDI. Topologia ta wymaga stosowania koncentratorów podwójnie przyłączanych i pojedynczo przyłączanych oraz stacji pojedynczo przyłączanych. Rysunek 10.5 przedstawia topologię podwójnego pierścienia z drzewami.

Rysunek 10.4. Podwójny pierścień.





Rysunek 10.5. Podwójny pierścień drzewami.

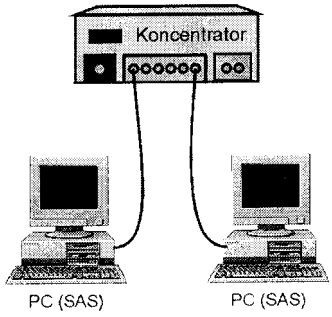
Kluczowa różnica między tą topologią a podstawową topologią podwójnego pierścienia jest taka, że urządzenia nie muszą być przyłączone bezpośrednio do pierścieni. Urządzenia SAS mogą być przyłączone do pojedynczo przyłączanych koncentratorów. Z kolei te koncentratory są przyłączone do koncentratorów DAS, które stanowią szkielet tandemu pierścieni.
Topologia ta gwarantuje niezawodność identyczną jak w topologii podwójnego pierścienia (automatyczne zawijanie w razie awarii), ale jej koszt jest niższy. Elementy SAS, czyli koncentratory i karty sieciowe, są znacznie tańsze od swoich odpowiedników DAS (wykorzystywanych w topologii podwójnego pierścienia).

1.10.2.2.1 Pojedyncze drzewo

Topologia pojedynczego drzewa, jak sugeruje jej nazwa, składa się wyłącznie z jednej, przypominającej drzewo, grupy urządzeń. Nie występuje tu podwójny pierścień, nie ma również żadnych elementów DAS. Drzewo można jednak uważać za logiczny pierścień, gdyż FDDI wykorzystuje okrężną- opartą na przekazywaniu tokenu - metodę dostępu do nośnika. Tokeny wciąż krążą po sieci, ale topologia bazuje na koncentratorze, więc ma kształt gwiazdy.
Oczywistą wadą jest brak ścieżki zapasowej. Rzutuje to bezpośrednio na niezawodność sieci. Topologia ta ma jednak wiele zalet. Po pierwsze, koszt budowy sieci FDDI w topologii pojedynczego drzewa jest dużo niższy niż w przypadku innych topologii. Wpływ na to mają dwa czynniki:
• Wszystkie urządzenia (koncentratory i stacje) występują w relatywnie tanim wariancie SAS.
• Koszt okablowania szkieletu sieci LAN jest mniejszy o połowę, gdyż używa się dwóch światłowodów zamiast czterech.
Inną znaczącą zaletą jest niezawodność. Choć może się to wydawać sprzeczne z wcześniejszą uwagą o braku drugiego pierścienia, stosowanie wyłącznie urządzeń pojedynczo przyłączanych ma istotny wpływ na niezawodność całej sieci lokalnej. Wpływ awarii jednego z urządzeń SAS na całą sieć jest dużo mniejszy niż w przypadku awarii urządzenia podwójnie przyłączanego. Jeśli stacja SAS ulegnie awarii, nie ma to wpływu na resztę sieci. Jeśli awarii ulegnie koncentrator SAS, to w najgorszym wypadku urządzenia do niego przyłączone zostaną odizolowane od reszty sieci. Awaria nie wyzwala mechanizmu zawijania. Mechanizm ten, choć ceniony jako rozwiązanie zwiększające niezawodność, obniża wydajność, podwajając niemal długość kabla w sieci. W pewnych sytuacjach mogłoby się zdarzyć, że automatyczna naprawa przy pomocy zawijania byłaby mniej pożądana niż zwykłe odizolowanie kilku stacji w przypadku awarii koncentratora.
Rysunek 10.6 przedstawia topologię pojedynczego drzewa.
Rysunek 10.6. Pojedyncze drzewo.

1.10.2.2.2 Podwójne kierowanie docelowe

Podwójne kierowanie docelowe to specjalny sposób wykorzystania podwójnego przyłączenia, zapewniający rezerwowe ścieżki fizyczne poprowadzone do istotnych zasobów sieciowych. Zasobami tymi mogą być serwery plików i/lub aplikacji, mosty czy nawet stacja robocza szefa! Proszę zauważyć, że podwójne kierowanie docelowe niekoniecznie musi obejmować każde urządzenie w sieci, więc tak naprawdę nie jest topologią. Jest opcjonalnym środkiem wdrażania połączeń sieci LAN. Może on być stosowany raczej w wąskim zakresie, dla pojedynczych urządzeń, a nie dla wszystkich urządzeń w sieci.



Ta implementacja może być wykorzystywana tylko w topologii podwójnego pierścienia z drzewami. Każde urządzenie, które ma być kierowane podwójnie, z definicji musi być podwójnie przyłączane. Musi także być połączone z siecią poprzez koncentrator DAS. Podwójne kierowanie docelowe umożliwia kluczowym urządzeniom posiadanie głównego oraz rezerwowego (mniej pożądanego z perspektywy protokołu FDDI) połączenia z siecią.
Protokoły zarządzania stacją dla urządzenia kierowanego podwójnie uaktywniają połączenie główne, natomiast połączenie rezerwowe pozostawiają w trybie pracy jałowej (ang. Standby Mode). Każde polaczenie kończy się w innym koncentratorze DAS. Protokoły zarządzania stacją mogą wykryć tę różnicę między dwoma połączeniami, wykorzystując mechanizmy odkrywania sąsiada. Protokół zarządzania stacją uaktywnia wtedy połączenie poprzez port A, jako ścieżkę główną, a połączenie poprzez port B pozostawia w stanie spoczynku. Jeśli z jakiegokolwiek powodu nastąpi utrata połączenia poprzez port A, protokół zarządzania stacją uaktywni połączenie rezerwowe.
Rysunek 10.7 przedstawia serwer skonfigurowany dla podwójnego kierowania docelowego w topologii podwójnego pierścienia z drzewami.

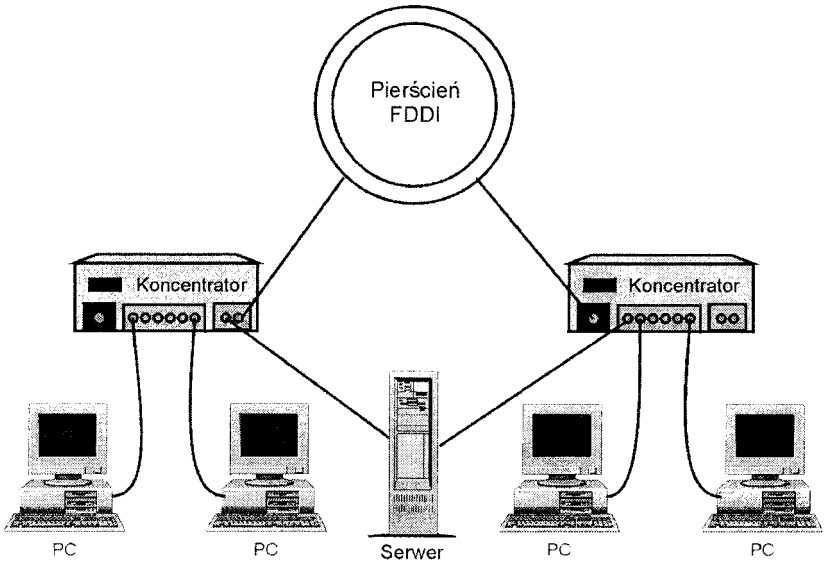
1.10.2.2.3 Cykliczne zawijanie

Cykliczne zawijanie nie jest właściwie oddzielną topologią, którą mógłbyś utworzyć. Jest raczej automatycznie konstruowane przez mechanizmy zarządzania stacją FDDI w wypadku, gdy awarii ulegnie stacja lub przewody łączące. Obszar awarii jest izolowany dzięki

natychmiastowemu logicznemu połączeniu pierścienia głównego z rezerwowym przed i za miejscem uszkodzenia. Z definicji tej wynika, że zawijanie mogą stosować tylko topologie bazujące na podwójnym pierścieniu.

Rysunek 10.7. Podwójne kierowanie docelowe.

Choć mechanizm naprawy jest podobny w obydwu przypadkach, istnieje między nimi jedna, zasadnicza różnica. W wypadku uszkodzenia przewodu wszystkie stacje w sieci mogą pozostać aktywne. Za to awaria stacji zmniejsza o jeden liczbę aktywnych urządzeń w sieci.



Na rysunku 10.8 awaria kabla dotknęła stację 2. Jej sąsiedzi, stacje 1 i 3, omijają uszkodzenie

, przenosząc swoją transmisję do pierścienia dodatkowego w celu zachowania integralności pętli. W nowym pierścieniu całkowita długość nośnika jest niemal dwukrotnie większa niż w pierścieniu oryginalnym. Dlatego w praktyce maksymalna długość nośnika w topologii podwójnego pierścienia powinna być zawsze co najwyżej połową maksymalnej długości dopuszczalnej dla nośnika danego rodzaju.

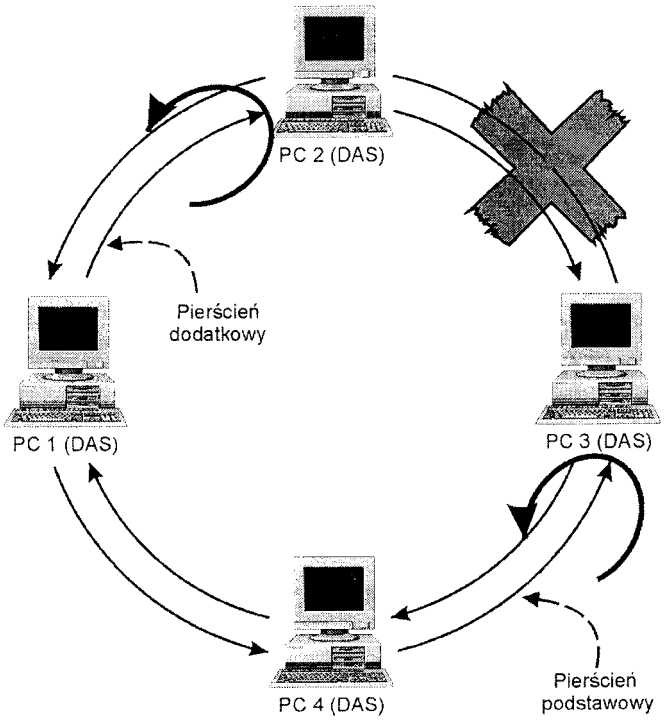
Aczkolwiek rysunek 10.8 przedstawia awarię kabla w podstawowej ścieżce stacji, podobne awarie mogą wystąpić w szkielecie sieci lokalnej. Podwójnie przyłączane koncentratory również mogą używać rezerwowego pierścienia, aby ominąć punkt uszkodzenia kabla.

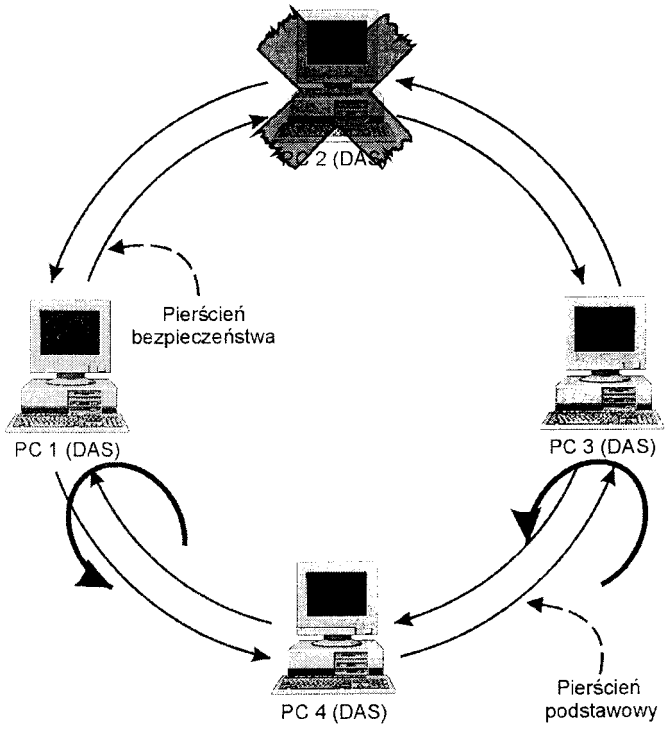
Gdyby stacja 2 na rysunku 10.8 była całkowicie niezdolna do pracy, zawijanie wyglądałoby nieco inaczej. Pierścień nie dochodziłby już do tej stacji, ale byłby zawijany w stacjach 1 i 3. Pokazuje to rysunek 10.9.

Aby zapobiec wymuszaniu zawijania w wypadku awarii stacji, można wykorzystywać urządzenia znane jako optyczne przełączniki obejściowe. Urządzenia te są instalowane pomiędzy stacją a koncentratorem. W razie awarii stacji przełączniki zapewniają ciągłość ścieżki transmisyjnej z pominięciem stacji.

Rysunek 10.8. Zawijanie pierścienia (awaria kabla).

Rysunek 10.9. Zawijanie pierścienia awaria stacji).





1.10.3 Rozmiar sieci

FDDI zostało zaprojektowane jako solidna sieć, mogąca obsługiwać stacje robocze o wysokiej wydajności. Aby utrzymać wysoką wydajność wszystkich przyłączonych urządzeń, FDDI musi narzucać ścisłe limity dotyczące wielkości sieci. Rozmiar sieci mogą określać następujące czynniki:

• Liczba przyłączonych urządzeń • Fizyczny rozmiar pierścienia

• Odległości fizyczne między urządzeniami

Wszystkie czynniki są tak samo ważne przy tworzeniu sieci lokalnej, jeśli ma ona wykorzystać potencjalną wydajność FDDI.

1.10.3.1 Maksymalna liczba urządzeń

Pierścień FDDI może obsłużyć maksymalnie 500 urządzeń. Powyższe ograniczenie wynika z maksymalnego dopuszczalnego czasu propagacji, przy którym protokoły FDDI mogą prawidłowo funkcjonować. Każde połączenie wprowadza własne, mierzalne opóźnienie propagacji. Skumulowane czasy ponad tysiąca fizycznych połączeń przekraczają prognozę opóźnień FDDI.

Policzenie 500 urządzeń może wydawać się dość prostą czynnością, jednak trudność leży w tym, żeby precyzyjnie zidentyfikować „urządzenie". W konfiguracji z podwójnym przyłączaniem każde urządzenie wymaga dwóch połączeń fizycznych i w efekcie liczy się to jako dwa połączenia. Każde urządzenie, w którym te połączenia są terminowane, liczy się jako dodatkowe połączenie. Tak więc port koncentratora i urządzenie do niego przyłączone w rzeczywistości stanowią dwa urządzenia.

Podwójnie przyłączany koncentrator szkieletowy, do którego nie są przyłączone żadne stacje, liczy się jako dwa połączenia. Jego porty liczą się jako urządzenie tylko wtedy, gdy są wykorzystywane. Podwójnie przyłączana stacja, niezależnie od tego, czy obydwa jej połączenia idą do tego samego, czy do dwóch różnych koncentratorów, liczy się jako dwa urządzenia. Stacja pojedynczo przyłączana jest traktowana jako jedno urządzenie.

Długość pierścienia

Standard ANSI X3T9.5 nie precyzuje jawnie maksymalnej długości pierścienia. Nie powinno to być niespodzianką, jeśli wiemy, że warstwa fizyczna (wbrew popularnej opinii) nie obejmuje samego nośnika. Trzymając się definicji warstwy fizycznej, określonej przez model referencyjny OSI, standard ANSI ustala parametry działania tak, by dla danego nośnika narzucały odległości maksymalne.

W pierścieniu zbudowanym na wielofunkcyjnym kablu światłowodowym całkowita długość ścieżki transmisyjnej musi być mniejsza niż 200 kilometrów. Nie powinno to być znaczącym ograniczeniem, chyba że sieć FDDI ma obejmować wielki obszar geograficzny, np. ma być siecią miejską (MAN). Warto jednak zwrócić uwagę na termin „całkowita długość ścieżki transmisyjnej".

W skowach tych kryją się dwie ważne implikacje. Po pierwsze, duży pierścień, mierzący 190 kilometrów, działa poprawnie, dopóki awaria nie wymusi zawijania. Wtedy długość pierścienia zwiększa się do około 380 kilometrów i cała sieć przestaje działać. Dlatego projektując sieć LAN, należy zawsze podzielić maksymalną dopuszczalną długość ścieżki transmisyjnej na pół..

Po drugie, termin „całkowita długość ścieżki transmisyjnej" należy rozumieć dosłownie. Żeby określić całkowitą długość pierścienia, trzeba dodać długości wszystkich odcinków światłowodu. Czyli nie tylko głównego pierścienia, ale też wszystkich odgałęzień łączących stacje.

Odległość między napędami

Odległość między napędami to maksymalna odległość między dowolnymi dwoma urządzeniami. Zjawisko tłumienia sygnału występuje zawsze, niezależnie od użytego nośnika fizycznego. Tak więc odległość między urządzeniami musi być wystarczająco małca, by zagwarantować integralność sygnału.
Dla wielofunkcyjnego kabla światłowodowego maksymalna odległość między napędami wynosi 2 kilometry. Dla światłowodu jednofunkcyjnego wzrasta do 60 kilometrów. Jednak nośniki miedziane są znacznie bardziej ograniczone. Ekranowana skrętka dwużykowa (STP) i nieekranowana skrętka dwużyłowa (UTP) Kategorii 5 mogą łączyć na odległość co najwyżej 100 metrów.

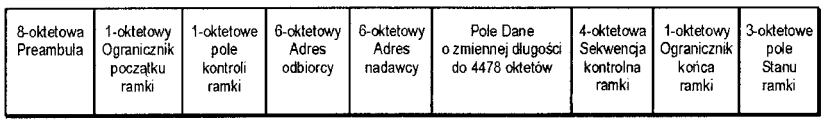
1.10.4 Ramki FDDI

FDDI w znacznym stopniu przypomina Token Ring: wszystkie funkcje związane z medium transmisyjnym muszą być umieszczone w ramce. FDDI ma wiele typów ramek używanych podczas zwykłej pracy i konserwacji. Są to takie ramki jak:
• podstawowa ramka danych, • ramka danych LLC,
• ramka danych LLS SNAP, • ramka Token,
• zestaw ramek zarządzania stacją.

1.10.4.1 Ramka danych

Najbardziej znaną spośród tych ramek jest „surowa" ramka danych. Jest ona przedstawiona na rysunku 10.10. Ramka FDDI ma długość maksymalną 9000 znaków, wliczając w to dane i pozostałe elementy składowe ramki. Jest to podstawowa ramka FDDI. Zwykle występuje w jednym z dwóch podformatów: LLC i SNAP. W żadnym z tych formatów nie może być dłuższa niż 4500 oktetów- nie wliczając w to Preambuły.

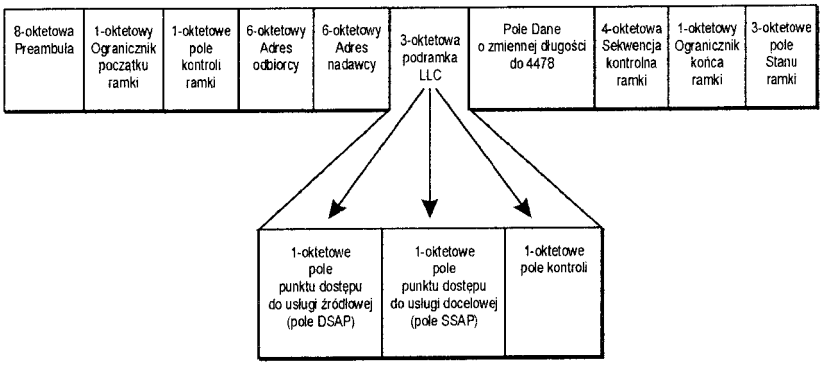
Rysunek 10.10.
Ramka danych
FDDI



Długości ramek i pól FDDI są często podawane w znakach, oktetach lub bajtach. Terminy te nie są całkowicie wymienne. Zrozumienie mechanizmów schematu kodującego FDDI może rozjaśnić różnice między nimi. Niestety, mało kto bada FDDI na takim poziomie szczegółowości. Protokoły warstwy fizycznej FDDI przed wysłaniem kodują każdy półoktet danych w 5-bitowy znak czy też symbol. Tak więc każdy 8-bitowy bajt (używając terminu programistów) danych, przekazany z warstwy aplikacji, staje się 10 bitami lub 1,25 oktetu. Dlatego terminy bajt i oktet nie są wymienne!
Ramka i jej elementy składowe mogą być mierzone w oktetach albo znakach. Nie oznacza to, że oktet i znak są synonimami. Przykładowo, warstwa MAC generuje ramki o maksymalnej długości 4500 oktetów. Oktety te, zarówno dane jak i reszta ramki, są transmitowane jako 5-bitowe znaki. Na poziomie fizycznym oktety są dzielone na połowy i każda połowa jest tłumaczona na 5-bitowy znak binarny. Podczas transmisji ramka ma więc maksymalną długość 9000 znaków po 5-bitów każdy. Tak więc podczas transmisji każda ramka MAC o maksymalnej wielkości ma długość 5625 oktetów.
Podstawowa ramka zawiera następujące pola:
• 8-oktetową Preambułę, sygnalizującą początek ramki
• 1-oktetowy Ogranicznik początku ramki, sygnalizujący początek zawartości ramki
• 1-oktetowe pole Kontrola ramki, sygnalizujące typ ramki, taki jak token, MAC, LLC, ramka priorytetu itd.
• 6-oktetowy adres MAC odbiorcy
• 6-oktetowy adres MAC nadawcy
• Pole danych o zmiennej długości, mogące zawierać maksymalnie do 4478 oktetów
• 4-oktetową Sekwencję kontrolną ramki, używaną do sprawdzania integralności ramki
• Półoktetowy (4-bity) Ogranicznik końca
• 3-oktetowe pole Status ramki, zawierające trzy jednooktetowe podpola: Błąd (ang. Error), Zgodność adresu (ang. Address-match), Skopiowana (ang. Address-match
). Każde z tych pól może mieć wartość „S" (od ang. „Set" - ustawione) lub „R" (od ang. „Reset" - wyzerowane).

1.10.4.2 Ramka danych LLC

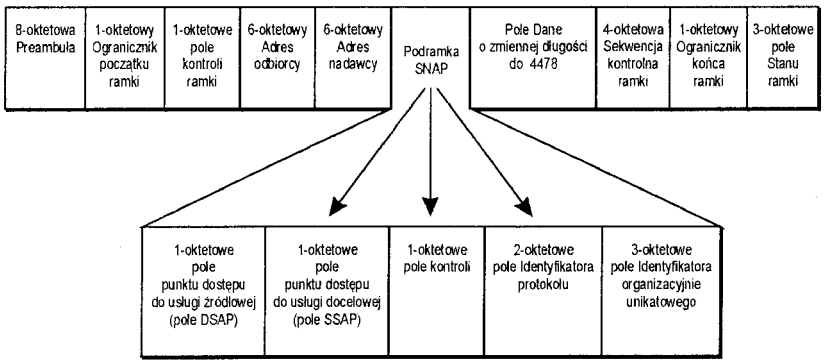
Podstawowa ramka danych FDDI może być również wykorzystywana do obsługi funkcji, określonych w specyfikacji IEEE 802.2 jako sterowanie łączem logicznym (ang. LLC- Logical Link Control. Ramkę LLC, przedstawioną na rysunku 10.11, tworzy się, dodając składającą się z trzech pól podramkę LLC do ramki FDDI. Dodatkowe trzy pola to: Punkt dostępu usługi docelowej, czyli pole DSAP (ang. Destiantion Service Access Point), Punkt dostępu usługi źródłowej, czyli pole SSAP (ang. Source Service Access Point) oraz pole Kontrola. Pola te poprzedzają bezpośrednio pole danych i wlicza się je do ładunku użytecznego ramki.
Rysunek 10.11. Ramka FDDI
L podramką b'02.2 LLC.



Ramka danych FDDI LLC ma następującą strukturę:
• 8-oktetowa Preambuła, sygnalizująca początek ramki
• 1-oktetowy Ogranicznik początku ramki, sygnalizujący początek zawartości ramki • 1-oktetowe pole Kontrola ramki, sygnalizujące typ ramki, taki jak token, MAC, LLC, ramka priorytetu itd.
• 6-oktetowy adres MAC odbiorcy
• 6-oktetowy adres MAC nadawcy
• 3-oktetowa podramka LLC, zawierająca 1-oktetowe pola DSAP, SSAP i Kontrola • Pole danych o zmiennej długości, mogące zawierać maksymalnie do 4475 oktetów
• 4-oktetowa Sekwencja kontrolna ramki, używana do sprawdzania integralności ramki
• Półoktetowy (4-bity) Ogranicznik końca
• 3-oktetowe pole Status ramki, zawierające trzy jednooktetowe podpola: Błąd (ang. Error), Zgodność adresu (ang. Address-match), Skopiowana (ang. Copied), przy czym każde z nich może mieć wartość „S" (od ang. „Set" - ustawione), albo „R" (od ang. „Reset" - wyzeromane).
Pierwotnym przeznaczeniem struktury LLC było zwiększenie możliwości Ethernetu w zakresie kierowania odebranych ramek do odpowiedniego protokołu/aplikacji. Była to kluczowa czynność w wieloprotokołowych urządzeniach, gdyż oryginalna specyfikacja Ethernetu powstała w czasach, gdy istniało tylko kilka protokołów komunikacyjnych. Oczywiście, FDDI nie ma ograniczeń wczesnego Ethernetu: obsługując ramkę LLC, może współpracować z ethernetowymi klientami za pomocą mostu tłumaczącego warstwy MAC, co nie zmniejsza sprawności sieci.

1.10.4.3 **Ramka danych LLC** SNAP

FDDI obsługuje również podramkę LLC SNAP. Ramkę FDDI z podramką SNAP standardu IEEE tworzy się dodając do ramki FDDI LLC 3-oktetowy identyfikator strukturalnie unikatowy i 2-oktetowe pole Typ. Te dodatkowe pola są umieszczone pomiędzy nagłówkiem LLC a polem danych. Wlicza się je do całkowitej długości pola danych. Ramka danych LLC SNAP jest przedstawiona na rysunku 10.12.
Rysunek 10.12. Ramka FDDI
_- podramką 802.2 SNAP.
Jak widać na rysunku 10.12, ramka FDDI SNAP ma następującą strukturę: • 8-oktetowa Preambuła, sygnalizująca początek ramki

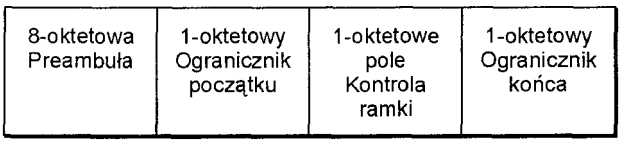


• l-oktetowy Ogranicznik początku ramki, sygnalizujący początek zawartości ramki
• 1-oktetowe pole Kontrola ramki, sygnalizujące typ ramki, taki jak token, MAC, LLC, ramka priorytetu itd.
• 6-oktetowy adres MAC odbiorcy
• 6-oktetowy adres MAC nadawcy
• 3-oktetowa podramka LLC, zawierająca 1-oktetowe pola DSAP, SSAP i Kontrola
• 5-oktetowa podramka SNAP, zawierająca 3-oktetowe pole Identyfikator strukturalnie unikatowy i 2-oktetowe pole Typ Protokołu, identyfikujące protokół wyższego poziomu
• Pole danych o zmiennej długości, mogące zawierać maksymalnie do 4470 oktetów
• 4-oktetowa Sekwencja kontrolna ramki, używana do sprawdzania integralności ramki
• Półoktetowy (4-bity) Ogranicznik końca

• 3-oktetowe pole Status ramki, zawierające trzy jednooktetowe podpola: Błąd, Zgodność adresu oraz Skopiowana, a każde z nich może mieć wartość „S" lub „R".
FDDI obsługuje struktury podramek LLC i SNAP jedynie po to, by zapewnić możliwość łączenia się z Ethernetem za pomocą mostu. Jest to potrzebne tylko przy łączeniu Ethernetu z FDDI poprzez most, przy wykorzystaniu protokołów Novell IPX/SPX lub AppleTalk. Niemniej jednak zwiększa to możliwości wykorzystania FDDI jako szkieletu w sieci LAN z wieloma topologiami.

1.10.4.4 Ramka Token

Ramka Token FDDI zawiera następujące pola:
• 8-oktetową Preambułę, sygnalizującą początek ramki
• 1-oktetowy Ogranicznik początku ramki, sygnalizujący początek zawartości ramki
• l-oktetowe pole Kontrola ramki, sygnalizujące typ ramki, taki jak token, MAC, LLC, ramka priorytetu itd.
• Półoktetowy (4-bity) Ogranicznik końca
Ramka Token jest przedstawiona na rysunku 10.13.
Rysunek 10.13. Ramka token FDDI.



Token jest przekazywany wzdłuż pierścienia tylko w jednym kierunku. Stacja posiadająca token może wykorzystać go, aby uzyskać dostęp do nośnika. W tym celu zmienia sekwencję bitów w polu Kontrola Ramki, co powoduje przekształcenie tokenu w ramkę danych.
W odróżnieniu od Token Ringu, FDDI wykorzystuje mechanizm szybkiego wyzwalania (ang. *quick release).* Natychmiast po przekształceniu tokenu w ramkę danych i wysłaniu jej, urządzenie wysyłające generuje nowy token. Jest on wysyłany do następnego urządzenia w pierścieniu. Może ono wykorzystać token do wysłania danych lub po prostu przekazać go dalej. Tak czy inaczej, urządzenie, które pierwsze nadawało, szybko odstępuje kontrolę nad medium transmisyjnym. Kolejna stacja nie musi czekać z nadawaniem, aż ramka danych powróci do swojego nadawcy.
Mechanizm szybkiego wyzwalania jest jednym z lepszych sposobów zwiększania wydajności, charakterystycznych dla FDDI. Dzięki niemu FDDI zawsze będzie mieć przewagę nad Token Ringiem, nawet gdy w tym drugim protokole zwiększy się szybkość sygnału, a co za tym idzie - szybkość transmisji danych.

1.10.4.5 Ramki SMT

FDDI, podobnie jak Token Ring, wykorzystuje ramki do realizacji funkcji zarządzania stacją (SMT). Podstawowe funkcje SMT to:
• ramka zgłoszenia (ang. CF- Claim Frame), służąca zgłaszaniu żądań,
• ramki echa (ang. ECF - Echo Frames), mające na celu testowanie echa (czyli potwierdzania odbioru),
• ramki informacji o sąsiadach (ang. NIF- Neighbor Information Frames), służące do uzyskiwania informacji o sąsiadach
• ramki informacji o statusie (ang. SIF-Status Information Frame.s), służące do uzyskiwania informacji o statusie
• ramki raportujące o statusie (ang. SRF-Status Reporting Frames), dostarczające (rozpowszechniające) informacje o statusie
• ramki odmowy dostępu (ang. RDF - Request Denied Frames), służące do przekazywania odmowy dostępu niewłaściwym żądaniom
• ramki zarządzania parametrami (ang. *RMF- Parameter Management Frames),* umożliwiające zdalny dostęp do innych stacji.
Razem ramki te umożliwiają wielu protokołom zarządzania stacją obsługę normalnych działań w sieci FDDI. Właściwe mechanizmy działania pierścienia są wyjaśnione w następnym podrozdziale tego rozdziału - „Mechanika sieci FDDI".

1.10.5 Mechanika sieci FDDI

Jak dotąd, mechanizmy przekazywania ramek w FDDI powinny być zrozumiałe. Podobnie jak w Token Ringu, wszystkie czynności są wykonywane za pomocą ramek.
Podstawowe mechanizmy FDDI najłatwiej zrozumieć, badając protokoły, dzięki którym stacja staje się aktywna i jest włączana do pierścienia. Inne protokoły, m.in. inicjalizacja pierścienia i odłączanie stacji, są ich uzupełnieniem. Procesy te pokazują, jaką rolę w działaniu sieci odgrywają różne elementy FDDI, takie jak SMT, MAC, PMD i PHY.

1.10.5.1 Inicjalizacja stacji

Logicznym punktem wyjścia jest proces, dzięki któremu stacja jest inicjalizowana przed włączeniem jej do pierścienia. Na inicjalizację składa się szereg testów, weryfikujących
fizyczną integralność połączenia z pierścieniem (lub drzewem) poprzez medium transmisyjne, a także gotowość pierścienia i przyłączonych do niego stacji. Za przebieg inicjalizacji odpowiada protokół zarządzania połączeniem fizycznym lub - inaczej - protokół PCM (ang. Physical Connection Management). Protokół ten jest jednym z elementów składowych funkcji zarządzania stacją (ang. SMT - Station Management).
Protokół PCM inicjuje serię testów, które kończą się po udanym dołączeniu stacji do pierścienia. Testy te zaczynają się po wykryciu odpowiedniego połączenia z siecią. Inicjacja stacji rozpoczyna się od wysłania znaków ciszy do protokołu PHY najbliższej (zgodnie z

kierunkiem ruchu w pierścieniu) stacji. Stacja odbierająca zatrzymuje swoją transmisję i wchodzi w stan wstrzymania (ang. Break State). Stan ten został tak nazwany, gdyż musi przerwać wszystkie transmisje i przyjąć przychodzące znaki ciszy.
Symbole są 5-bitowymi strukturami binarnymi, używanymi do kodowania danych. Ponieważ możliwych niepowtarzalnych struktur jest więcej (32) niż znaków heksadecymalnych (16), łatwo zauważyć, że niektóre znaki nie mogą być wykorzystywane do kodowania danych. Symbole te są zarezerwowane dla funkcji sieciowych. Przykładami mogą być: symbol uciszenia i symbol stopu.
Stan wstrzymania jest po prostu stanem przejściowym, który trwa tylko tak długo, aby odbiorca mógł wstrzymać swoje transmisje. Gdy to nastąpi, stacja wchodzi w Stan uciszenia linii (ang. Quite Line State). Stan ten charakteryzuje się tym, że zarówno nowo aktywowana stacja, jak i jej najbliższy sąsiad równocześnie wysyłają do siebie symbole uciszenia (ang. Quiet Symbols) . W ten sposób mogą się przekonać, że są zsynchronizowane.
Po zsynchronizowaniu obydwie stacje zaczynają w powtarzalny sposób wysyłać inny symbol - symbol stopu (ang. Halt Symbol. Symbole te służą do synchronizowania zegarów transmisyjnych obydwu stacji. Gdy zegary zostaną zsynchronizowane, stacje wychodzą ze stanu stopu i wchodzą w tzw.. stan przygotowania (ang. .Next State).
Podczas stanu przygotowania stacje wymieniają informacje o swoich portach. Jak dotąd, każda ze stacji wiedziała tylko, że druga istnieje i że ich zegary transmisyjne są zsynchronizowane. Poza tym nie miały żadnej innej informacji odnośnie drugiej stacji. Wymieniając informacje o portach (jak na przykład porty A, B, M lub S), mogą się lepiej poznać. Informacje te służą również do określenia, jaka kombinacja połączeń portów między nimi istnieje. Dane te są niezbędne do późniejszych testów inicjalizacji stacji.
Po udostępnieniu sobie tych danych obie stacje wchodzą w stan rozpoczęcia komunikacji (ang. Signal State). Między stanem przygotowania a stanem rozpoczęcia komunikacji występuje krótki stan przejściowy, znany jako stan bezczynności (ang. Idle Line State). Podczas tego stanu przejściowego obie stacje wysyłają szereg symboli pustych (ang. Idle Symbols), które mówią obydwu stacjom, że nadawcy są gotowi do odbioru sygnałów.
Potem następuje test poprawności połączenia (ang. Link Con'idence Te.st). Ten protokół wymaga, aby każda stacja sprawdziła, czy druga stacja posiada warstwę MAC. Jeśli tak, obydwie stacje wchodzą w fazę testowania transmisji ramek i przekazywania tokenu. Jeśli nie, wymieniają one kolejną serię symboli pustych. Jeśli te testy zakończą się sukcesem, stacje mogą przejść do stanu połączenia. Stan ten ma na celu wyłącznie zapewnienie tego, by obydwie stacje zaczęły działać jednocześnie.
Po udanym zakończeniu tej wymiany protokoły zarządzania połączeniem fizycznym obydwu stacji wchodzą w stan aktywny. Ta zmiana stanu kończy proces inicjalizacji stacji i sygnalizuje funkcji zarządzania stacją gotowość włączenia stacji do pierścienia.

1.10.5.2 Inicjalizacja pierścienia

Po zakończeniu procesu inicjalizacji stacji należy przeprowadzić inicjalizację pierścienia. Wymaga to określenia, która stacja ma wygenerować pierwszy token, i ustalenia operacyjnego czasu rotacji tokenu (T POR). Stacje muszą poprosić o prawo wypuszczenia pierwszego tokenu. O przyznaniu tego prawa decyduje porównanie domyślnych wymagań czasowych każdej stacji. Wymagania te przechowuje wartość, zwana zegarem rotacji tokenu (ang. TRT - Token Rotation Timer), określająca, jak często token musi docierać do danej stacji. Zgłaszanie ofert rozpoczyna się, gdy pierwsza aktywna stacja generuje ramkę zgłoszenia.
Ramka zgłoszenia zawiera adres stacji, która ją wysłała, oraz wartość TRT dla tej stacji. Ramka jest wysłana do kolejnego urządzenia w pierścieniu. Urządzenie odbiera ją i porównuje swoją wartość TRT z wartością TRT zapisaną w ramce. Jeśli TRT odbiorcy jest mniejsze, odrzuca on otrzymaną ramkę zgłoszenia i wysyła swoją własną. W przeciwnym wypadku odbiorca wysyła otrzymaną ramkę do następnej stacji. Proces trwa, dopóki jedna ze stacji nie otrzyma z powrotem własnej ramki zgłoszenia. Oznacza to, że jej wartość TRT jest najmniejsza i stacja ma prawo wysłać pierwszy token. Jej TRT staje się operacyjnym zegarem rotacji tokenu (OTRT - Operational Token Rotation Timer) dla całego pierścienia.
Taka jest istota deterministycznej natury FDDI: każda stacja ma możliwość określenia maksymalnego czasu, jaki może upłynąć między wizytami tokenu. Wydajność całego pierścienia można zwiększyć, zmniejszając jego rozmiar (mierzony jako liczba przyłączonych stacji) lub wykorzystując TRT do polepszenia czasów obiegu tokenu.

1.10.6 Podsumowanie

Przez lata FDDI było jedyną stabilną, dojrzałą, dobrze działającą i szybką technologią LAN. Niestety, wysoka cena ograniczała jej zastosowania do małych nisz rynkowych, gdzie wymagana była wysoka wydajność. Ostatnimi laty pojawiła się na rynku konkurencja dla FDDI. Początkowo zwolennicy FDDI mogli z przekonaniem twierdzić, że jest to jedyna stabilna, szybka sieć lokalna. Po pewnym czasie jej rywale również osiągnęli stabilność i standaryzację produktów, potwierdzając w ten sposób ich konkurencyjność.
Dziś FDDI wciąż wyróżnia się niezawodnością i wysoce deterministyczną naturą, dzięki czemu dobrze nadaje się do zastosowania w szkielecie sieci LAN i do łączenia serwerów z siecią.

1.11 Rozdział 11.ATM

Mark A. Sportack

Tryb transferu asynchronicznego, znany powszechnie pod akronimem ATM (ang. *Asynchronous Transfer Mode),* został pierwotnie utworzony przez Międzynarodowy Komitet Konsultacyjny ds.. Telefonii i Telegrafii, czyli przez komitet CCITT (franc. *Comitee Consulatif Internationale de Telegraphique et Telephonique)* jako mechanizm (asynchronicznego) transferu dla szerokopasmowej sieci cyfrowej usług zintegrowanych, czyli sieci B-ISDN (ang. *Broadband Integrated Services Digital Network).* Zastosowanie tego trybu miało być ograniczone do przeprowadzania transmisji między centralami telefonicznymi, ale we wczesnych latach 90-tych uznano, że jego wysoka przepustowość i małe opóźnienia czynią go idealnym mechanizmem dla sieci LAN nowej generacji. Teoretycznie, mógł on obsługiwać biurowe

wideokonferencje (wymagające dużej szybkości i szerokiego pasma) z taką samą łatwością, z jaką mógł obsłużyć tradycyjne aplikacje interaktywne. Co więcej, za pomocą ATM można by było idealnie zintegrować sieci LAN i WAN.
Skrót CCITT oznacza Międzynarodowy Komitet Konsultacyjny ds. Telefonii i Telegrafii.. Nazwa komitetu została niedawno zmieniona na Międzynarodową Unię Telekomunikacyjną (ITU).
Ta wielka wizja unifikacji sieci wyzwoliła niespotykaną aktywność zarówno w na rynku, jak i w sferze badań i rozwoju. Powstało przemysłowe konsorcjum, tzw. ATM Forum, dostarczające propozycje standardów komitetowi CCITT. Nowa technologia byka tak obiecująca, że wkrótce liczbę członków ATM Forum zaczęto liczyć w setkach. Polityczne machinacje, nieuniknione w tak ogromnym konsorcjum, szybko jednak sparaliżowały prace nad standardami. Zaczęło się wydawać, że ATM pozostanie tym, czym pierwotnie miał być - technologią łączenia centrali telefonicznych.
Dziś, po tym, jak wiele stracono w stosunku do przyspieszonych i przełączanych wersji Ethernetu, FDDI i Token Ringu, ATM Forum wydaje się odzyskiwać utracony impet. Być może znowu stanie się siecią LAN nowej generacji. W tym rozdziale opisana jest struktura komórek ATM, podstawowe usługi i mechanizmy działania, w tym emulacja sieci lokalnej oraz komutowanie pakietów IP w sieci ATM.

1.11.1 Podstawy sieci ATM

ATM odwraca tradycyjny paradygmat sieci. W sieciach tradycyjnych, bezpołączeniowe pakiety wysyłane ze stacji nadawczych niosą ze sobą dodatkową informację, która pozwala tylko zidentyfikować ich nadawcę i miejsce przeznaczenia. Tak więc, stacje mogą być względnie proste. Sama sieć została natomiast obarczona uciążliwym zadaniem rozwiązania problemu dostarczenia pakietu do odbiorcy.
ATM jest tego przeciwieństwem. Ciężar spoczywa na stacjach końcowych, które ustanawiają między sobą wirtualną ścieżkę. Przełączniki znajdujące się na tej ścieżce mają względnie proste zadanie - przekazują komórki wirtualnym kanałem poprzez przełączaną sieć, wykorzystując do tego informacje zawarte w nagłówkach tych komórek. Tak brzmi najbardziej uproszczone wyjaśnienie sposobu działania sieci ATM. Aby lepiej opanować ATM, trzeba poznać naturę jego połączeń logicznych, a także niektóre spośród najbardziej podstawowych aspektów tej sieci, jak szybkość transmisji, obsługiwane media transmisyjne, topologie i interfejsy. Powyższe tematy są omówione w niniejszym podrozdziale i stanowią odpowiednie tło dla bardziej dogłębnego zbadania mechaniki działania ATM, które kończy ten rozdział.

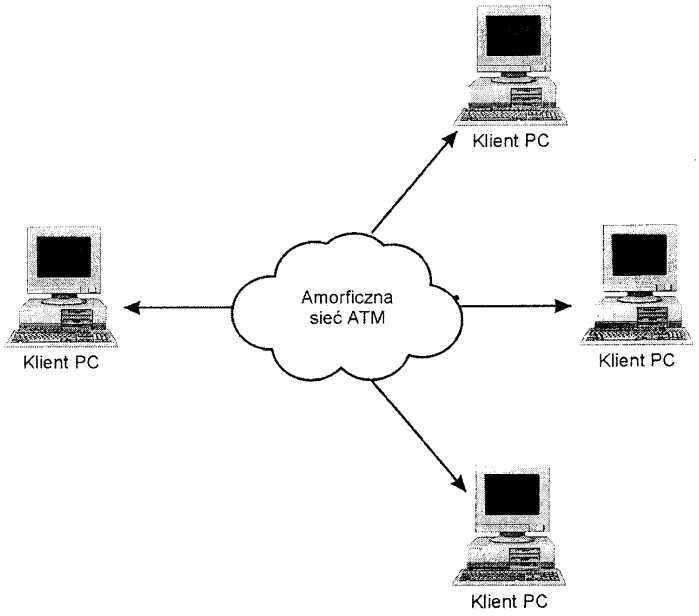
1.11.2 Połączenia wirtualne

W sieci ATM można ustanawiać dwa rodzaje połączeń wirtualnych:
• Obwód wirtualny • Ścieżkę wirtualną
Obwód wirtualny jest połączeniem logicznym pomiędzy dwoma urządzeniami końcowymi poprzez sieć przełączaną. Urządzenia te komunikują się, przesyłając komórki danych tam i z powrotem, poprzez obwód logiczny. Ścieżka wirtualna to zgrupowanie logiczne tych obwodów. Zdolność rozpoznawania takich zgrupowań umożliwia przełącznikom ATM przeprowadzanie operacji na całej grupie - nie muszą zarządzać każdym obwodem logicznym z osobna.
Każda komórka ATM zawiera zarówno informacje ścieżki wirtualnej (ang. VPI - Virtual Path Information), jak też informacje obwodu wirtualnego (ang. VCI - Virtual C.'ircuit Information). Przełącznik ATM używa tych informacji do przekazywania otrzymanych komórek do odpowiedniego następnego urządzenia. By to wykonać, przełącznik ATM musi tworzyć i zachowywać tablice przełączające. Podobnie jak stare tablice mostkujące, również nowe tablice przełączające nie są niczym więcej jak listą kojarzącą informacje VPI i VCI z fizycznymi interfejsami przełącznika.
Gdyby wskazać najważniejszą cechę sieci opartych na trybie ATM -to jest nią niewątpliwie fakt, iż sieć ta jest siecią połączeniową. Może ona być wykorzystywana do obsługi protokołów bezpołączeniowych, takich jak TCP/IP czy IPX/SPX, ale robi to w kontekście połączenia logicznego.

1.11.3 Typy połączeń

ATM jest protokołem połączeniowym, mogącym obsługiwać następujące rodzaje połączeń:
• połączenie dwupunktowe,
• połączenie jednej stacji z wieloma.
Połączenia dwupunktowe wiążą razem dwa urządzenia wirtualnym połączeniem poprzez sieć przełączników ATM. Transfer danych może być jedno- lub dwukierunkowy. Ten rodzaj połączenia przedstawia rysunek ł1.1 .
Rysunek 11.1. Połączenie dwupunktowe w sieci ATM.
Drugi rodzaj połączenia - jednej stacji z wieloma- jest nieco bardziej skomplikowany. Obsługuje on tylko połączenia jednokierunkowe, z jednego punktu nadawania do wielu odbiorców. Ilustruje to rysunek 11.2.

Rysunek 11.2. Połączenie, jednej stacji z wieloma w sieci ATM.



Niestety, protokoły ATM nie są dobrze przystosowane do obsługi dwukierunkowej komunikacji z jednej stacji do wielu. Nie mogą też w ogóle obsługiwać komunikacji wielu stacji z wieloma. Obie formy komunikacji można by było względnie łatwo wdrożyć za pomocą protokołów warstwy 2 dla wspólnych nośników, gdyż takie protokoły miałyby złożone mechanizmy sterowania dostępem do nośnika, które mogłyby służyć równocześnie wielu klientom.

Połączenia dwóch stacji i jednej stacji z wieloma można ustanawiać na dwa sposoby: za pomocą komutowanego lub stałego obwodu wirtualnego. Początkowo ATM Forum koncentrowało się na stałych obwodach wirtualnych, czyli inaczej obwodach PVC (ang. *Permanent Virtual Circuits).* W niemal ten sam sposób, jak w przypadku sieci Frame Relay, połączenia tworzone przy użyciu „chmary" przełączników, mogą być definiowane programowo. Podczas normalnej pracy połączenia te zawsze pozostają aktywne, niezależnie od natężenia ruchu.

Nieco bardziej kłopotliwe pozostaje opracowanie protokołów dla komutowanych obwodów wirtualnych (ang. *SVC - Switched Virtual Circuits).* Są one tworzone na żądanie pomiędzy dwoma lub więcej punktami końcowymi. Komórki są przesyłane powstałym w ten sposób kanałem logicznym. Po zakończeniu przesyłania, kanał jest demontowany i sieć ponownie może korzystać z całego pasma.

1.11.4 Szybkości przesyłania danych

ATM może działać przy wielu różnych szybkościach transmisji. Pierwotnie, szybkości te miały bazować na specyfikacjach linii nośnika optycznego (ang. *OC - Optical Carrier).* Było to zgodne z pierwotnym przeznaczeniem ATM jako protokołu transportu asynchronicznego dla sieci B-ISDN. Tak więc podstawowa szybkość transmisji ATM została ustalona, zgodnie ze standardem OC-3, na 155,52 Mbps. Przewidziano również obsługę szybkości 51,84 Mbps, zgodnej ze standardem OC-1, a także rozszerzenie skali w górę, do 2,488 Gbps, zgodnie ze specyfikacją OC-48. Zapewniało to możliwość dostosowania ATM do przyjętej architektury publicznych sieci komutowanych.

Trwają prace nad zwiększeniem szybkości ATM do 10 Gbps, choć nie przewiduje się wykorzystywania takiej szybkości w sieciach lokalnych.

Adaptacja ATM do środowiska sieci lokalnych spowodowała konieczność wprowadzenia pewnych zmian do oryginalnej specyfikacji. Między innymi należało umożliwić obsługę nośników miedzianych, które byłyby zgodne ze specyfikacjami nośników optycznych. Dodatkowo trzeba było dostosować ATM do pracy z szybkością mniejszą niż 51,84 Mbps (standard OC-I). Opracowano dwie propozycje: 25,6 i 25,9 Mbps. Szybkość 25,9 Mbps była bardziej logiczną propozycją, gdyż w razie potrzeby mogła być przeskalowana wzwyż, do architektury OC.

Specyfikacja 25,6 Mbps została utworzona z chipsetu IBM dla sieci Token Ring. Wierzono, że taka podstawa uczyni standard ATM 25,6 Mbps bardziej wiarygodnym od innych propozycji, nie mających podobnej, dobrze opracowanej technicznie spuścizny. Jednak zaowocowało to również niestandardową (choć względnie funkcjonalną) szybkością transmisji danych. Specyfikacja ta zapewnia łączność za pomocą kabla UTP Kategorii 3 na odległość do 100 metrów. Ostatecznie propozycja 25,6 Mbps została przyjęta przez ATM Forum.

Bazujący na przewodach miedzianych wariant OC-1 był przeznaczony do transmisji z szybkością 51,84 Mbps, przy wykorzystaniu kabla UTP Kategorii 5 o maksymalnej długości 100 metrów. Niestety, w specyfikacji tej zastosowano nową technologię modulacji, znaną jako modulacja amplitudowo-fazowa bez fali nośnej, czyli modulacja CAP lub inaczej CAP-M (ang. *Carrierless Amplitude Phase Modulation).* Mimo że sprawdzała się ona w warunkach laboratoryjnych, okazało się, że wytwarzanie jej na dużą skalę jest praktycznie niemożliwe.

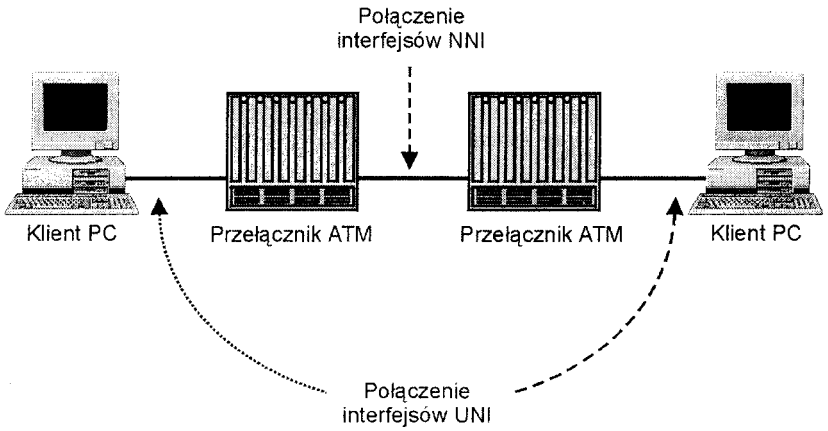
Pełną prędkość ATM, czyli 155,52 Mbps, również uzgodniono z dwoma nowymi interfejsami nośników fizycznych dla środowiska sieci LAN. Pierwszym z nich była skrętka UTP Kategorii 5, o maksymalnej długości 100 metrów, zaś drugim wielofunkcyjny światłowód o średnicy 62,5 mikrona. Maksymalna długość takiego kabla wynosi 2 kilometry.

1.11.5 Topologia

W odróżnieniu od wielu dzisiejszych protokołów sieciowych warstwy 2, które są implementowane w przełącznikach, ATM został od razu zaprojektowany jako sieć komutowana. Dlatego pracuje w topologii gwiazdy. Każde urządzenie jest bezpośrednio przyłączone do przełącznika ATM (nie jest to koncentrator ani wzmacniak) i ma własne dedykowane połączenie z tym przełącznikiem.
Tu powstaje interesująca kwestia: jeśli każda stacja ma własne, dedykowane połączenie z przełącznikiem, to czy te dedykowane połączenie kończy się w tym przełączniku? Jest oczywiste, że gdzieś w sieci musi być wykorzystywane multipleksowanie, zwłaszcza dla połączeń opuszczających LAN i wchodzących do sieci WAN. Mało która organizacja może pozwolić sobie na zapewnienie dedykowanego połączenia 155,52 Mbps wszystkim swoim członkom.
Przełączniki mogące obsługiwać jednoczesne transmisje z pełną szybkością dla wszystkich przyłączonych urządzeń nazywane są nieblokującymi. Przełączniki nieblokujące wymagają płyty, której przepustowość musi być co najmniej tak duża, jak zagregowane pasmo obsługiwanych stacji. Oznacza to, że płyta musi działać z szybkością co najmniej 155,52 razy większą od liczby portów w przełączniku, aby można go było zakwalifikować jako nieblokujący. Na przykład, przełącznik ATM z ośmioma portami musi mieć płytę pracującą z szybkością 8*155,52, czyli 1,244 Gbps, by można było określić go mianem nieblokującego. Wszystkie 8-portowe przełączniki 155,52 Mbps, których płyta ma pasmo mniejsze niż 1,244 Gbps, musiałyby blokować część transmisji, gdyby wszystkie 8 stacji próbowało nadawać równocześnie.

1.11.6 Interfejsy ATM

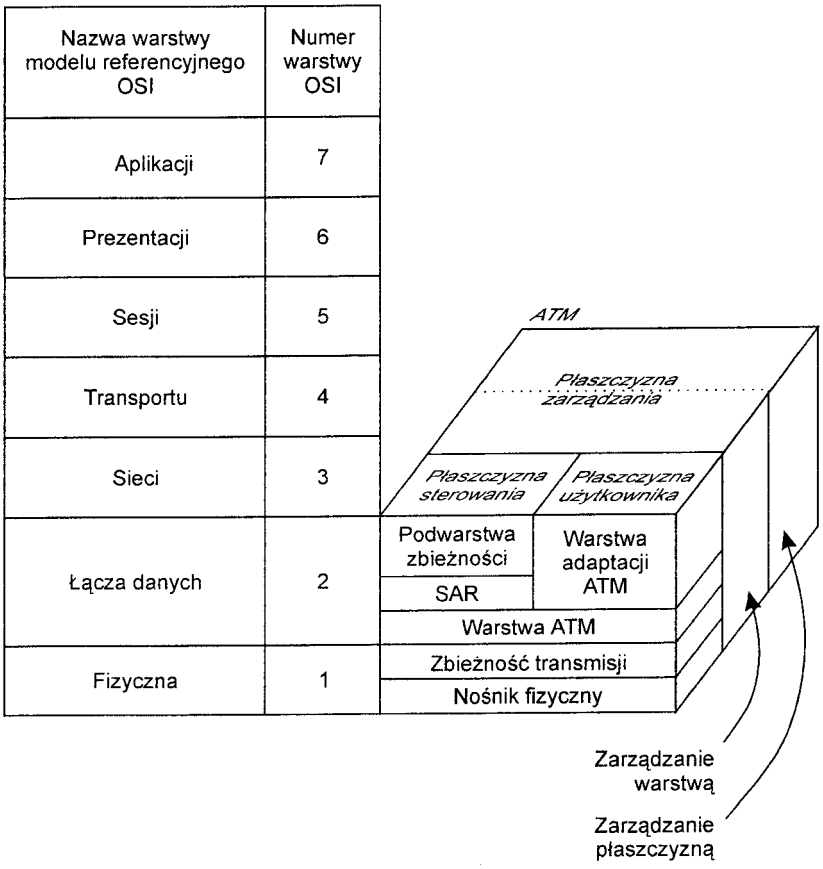
Istnieje wiele typów wysoko wyspecjalizowanych interfejsów ATM. Wiele z nich jest blisko ze sobą powiązanych. Dzielą je tylko drobne, subtelne różnice, wymuszone względami prawnymi, nota bene wciąż nękającymi przemysł telekomunikacyjny w USA. Bliższe przyjrzenie się dwóm spośród interfejsów ATM powinno zapewnić kontekst niezbędny do zrozumienia architektury ATM. Dwa najbardziej znane interfejsy to:
• Interfejs użytkownik-sieć (ang. UNI - User-to-Network Interface)
• Interfejs międzysieciowy (ang. NNI-Network-to-Network Interface)
Nazwy te mówią same za siebie. Interfejsy UNI służą do łączenia sprzętu użytkownika z siecią ATM, podczas gdy interfejsy NNI są potrzebne do łączenia ze sobą przełączników ATM. Rysunek 11.3 przedstawia obydwa interfejsy.
Rysunek 11.3. Połączenia interfejsów ATM: użytkownik-sieć i międz ysieciowego.



Ponieważ interfejs ATM UNI wywodzi się z szerokopasmowej technologii ISDN, został opracowany tak, aby opisywał podstawowe wyposażenie stacji. Tak więc interfejs UNI może być wykorzystywany przez prywatny przełącznik ATM, który łączy się z publiczną siecią telekomunikacyjną ATM.
ATM Forum jest także bliskie ukończenia prac nad prywatnym interfejsem międzysieciowym (ang. P-NNI - Private Network-to-Network Interface). Protokoły tego interfejsu umożliwią dynamiczne trasowanie w prywatnych sieciach ATM z wieloma przełącznikami.

1.11.7 Model ATM

Jakiż protokół byłby kompletny bez własnego modelu referencyjnego? ATM nie jest więc wyjątkiem. Model referencyjny protokołu ATM wprowadza jednakże innowacyjne podejście do znanego modelu warstwowego. Wykorzystuje bowiem koncepcję płaszczyzn. Płaszczyzny, w odróżnieniu od warstw, są trójwymiarowe. Nie wdając się w szczegóły geometryczne, wystarczy powiedzieć, że każda płaszczyzna reprezentuje oddzielny pakiet protokołów. Razem obejmują komplet protokołów ATM. Model referencyjny protokołu ATM jest przedstawiony na rysunku 11.4, gdzie porównuje się go z modelem referencyjnym OSI.
Rysunek 11.4. Model referencyjny ATM a model ręferencyjny OSI.



Chociaż rysunek 11.4 przedstawia podwarstwę zbieżności jako stałą część modelu ATM, jest ona wymagana tylko dla pewnego podzbioru warstw adaptacyjnych. Wskutek tego pewne klasy usług mogą obejść się bez korzystania z podwarstwy zbieżności. Problem ten jest rozwinięty w dalszej części tego podrozdziału.

Jak widać na rysunku 1 1.4, ATM ogranicza się do warstw 1 i 2 modelu referencyjnego OSI. Powyżej znajdują się tradycyjne aplikacje i protokoły transportu. Mechanizmy wyższej warstwy zawierają normy przesyłania danych i komunikacji między urządzeniami końcowymi. Odnoszą się one również do płaszczyzn użytkownika i sterowania.

1.11.7.1 Warstwa fizyczna

Na wersję warstwy fizycznej OSI dla standardu ATM składają się dwie podwarstwy: • Zbieżności transmisji (ang. TC-Transmission Convergence)

• Nośnika fizycznego (ang. *PM-Physical Medium)*

Warstwa fizyczna ATM odpowiada za wysyłanie i odbieranie danych. Podczas wysyłania danych mechanizm zbieżności transmisji przyjmuje komórki z warstwy ATM, generuje sumę kontrolną, którą można wykorzystać do wczesnego wykrywania błędów (w ograniczonym stopniu), a następnie bit po bicie wysyła nagłówek komórki i dane użyteczne poprzez medium transmisyjne.

Przy odbiorze danych warstwa fizyczna przyjmuje nadchodzące komórki, sprawdza sumę kontrolną, aby ustalić, czy nagłówek nie został uszkodzony podczas przesyłania, a następnie przekazuje bity do warstwy ATM, która z powrotem tworzy z nich komórki.

Podwarstwa zbieżności transmisji

Podwarstwa zbieżności transmisji (TC) jest najwyższym ze składników warstwy fizycznej ATM. Jest odpowiedzialna za wiele funkcji, w szczególności:

• określanie komórki,

• generowanie sekwencji kontroli błędów sprzętowych, • rozdzielenie szybkości transmisji komórek,

• funkcje dostosowywania, generowania i odzyskiwania ramki transmisyjnej. „Określanie komórki" jest nazwą funkcji zachowywania integralności granic odbieranej komórki. Jest to funkcja kluczowa dla udanego wyodrębniania komórek z otrzymywanego strumienia bitów.

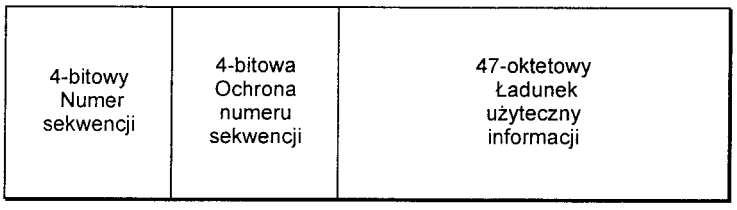
Kontrola błędów sprzętowych (ang. HEC - Hardware Errors Control to po prostu wartość sumy kontrolnej. Stacje transmitujące muszą generować i sprawdzać tę wartość na podstawie zawartości nagłówka komórki. Odbiorcy muszą ponownie obliczyć tę wartość, by określić prawdopodobieństwo uszkodzenia zawartości podczas przesyłania. Niestety, suma kontrolna jest generowana tylko na podstawie 5-oktetowego nagłówka, a nie całej, 53-oktetowej komórki. Stanowi to potwierdzenie starego argumentu, że ATM w rzeczywistości nie zapewnia wczesnego wykrywania błędów, choć posiada odpowiedni mechanizm. Jak okaże się w dalszej części rozdziału, wczesna korekcja błędów jest integralną częścią wielu składników warstwy łącza danych.

Rozdzielenie szybkości transmisji komórek jest mechanizmem stosowanym do dopasowania szybkości transmisji komórek w warstwie ATM do szybkości transmisji w interfejsie nośnika. Na przykład, jeśli warstwa ATM działa z pełną szybkością 155,52 Mbps, ale interfejs nośnika może działać tylko z szybkością 25,6 Mbps, podwarstwa zbieżności transmisji odpowiada za jej spowolnienie. Przeważnie dokonuje się tego

przez usunięcie nie przyporządkowanych komórek. Podwarstwa zbieżności może również wstawić puste komórki, aby przyspieszyć strumień komórek w warstwie ATM, jeśli jest on wolniejszy niż szybkość transmisji komórek przez medium transmisyjne.
Ostatni zestaw funkcji, za które odpowiada podwarstwa zbieżności transmisji, może wydawać się nieco intuicyjny. Funkcje te dotyczą ujmowania transmisji w ramki. Dostosowywanie ramki transmisyjnej to proces, w którym podwarstwa TC upakowuje komórki w ramkach akceptowanych przez nośnik fizyczny, który będzie przesyłać dane. Podobnie generowanie i odzyskiwanie ramek są funkcjami pomocniczymi, tworzącymi i utrzymującymi odpowiednią strukturę ramek warstwy fizycznej.
Jednostki danych protokołu (ang. PDU - Protocol Data Units), tworzone i wykorzystywane przez różne mechanizmy warstwy łącza danych ATM, także są nazywane ramkami. Znacznie bardziej szczegółowy opis jednostek PDU znajduje się w punkcie następnym - „Warstwa adaptacji ATM".
Podwarstwa nośnika fizycznego
Podwarstwa nośnika fizycznego zawiera wszystkie funkcje związane z nośnikiem. Obejmuje to synchronizowanie taktowania transmisji w obwodzie wirtualnym, a jeśli trzeba, także wysyłanie i odbieranie bitów. Warto zauważyć, że ponieważ podwarstwa nośnika fizycznego jest zależna od medium transmisyjnego, istnieją osobne jej specyfikacje dla każdego z obsługiwanych typów nośnika.

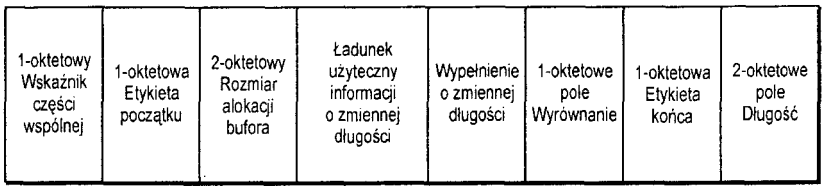
1.11.7.2 **Warstwa adaptacji** ATM

Jak widać na rysunku 11.4, warstwa łącza danych OSI dla standardu ATM obejmuje dwie podwarstwy: warstwę ATM i warstwę adaptacji ATM (ang. AAL - ATM Adaptation Layer). Warstwa AAL jest pakietem protokołów położonych na wyższym poziomie warstwy łącza danych ATM. Choć jest nazywana „warstwą AAL", właściwie obejmuje trzy mechanizmy. Oprócz najbardziej oczywistego mechanizmu AAL są to również mechanizmy podwarstwy zbieżności (CS) oraz segmentacji i ponownego składania (ang. SAR - Segmentation and Reassembly).
Podwarstwa zbieżności pełni rolę swoistego „lejka", przez który dane przeznaczone do transmisji dostarczane są z protokołów warstwy 3, a także - przez który odebrane dane są przekazywane tym protokołom. Podwarstwa zbieżności odpowiada za prawidłowe przekształcanie pochodzących z wyższych protokołów żądań usług AAL, SAR i ATM, dotyczących zarówno danych wchodzących, jak i wychodzących. Takie przekształcanie potrzebne jest w przypadku protokołów TCP/IP, IPX/SPX, a nawet innych protokołów warstwy 2, takich jak Frame Relay.
Podwarstwa SAR jest mechanizmem, który faktycznie zmienia struktury danych, otrzymane z protokołów wyższych warstw, w 48-oktetowe struktury dokładnie wypełniające pola danych użytecznych ATM. Te struktury danych są przekazywane do warstwy ATM, gdzie stają się częścią użyteczną komórek.
Zaplanowano pięć różnych warstw AAL, oznaczonych numerami od 1 do 5. Każda ma obsługiwać jedną z czterech różnych klas usług ATM, znanych jako klasy A,B,C i D. Każda klasa winny sposób obsługuje transmisję i dlatego wymaga osobnego zestawu protokołów w stacjach końcowych. Każda warstwa adaptacji ATM umieszcza dane otrzymane z warstwy SAR w strukturze (jeszcze nie komórce) zwanej jednostka danych protokołu segmentacji i ponownego złożenia (ang. SAO-PDU). Jednostki SAO-PDU są przekazywane do warstwy ATM, gdzie opatruje się je 5-oktetowymi nagłówkami, formując w ten sposób znajome, 53-oktetowe komórki ATM.
Choć zaplanowano pięć warstw AAL, obecnie pozostały tylko trzy. Warstwa AAL 2 nie została nigdy ukończona, zaś warstwy 3 i 4 połączono w jedną i nazwano AAL 3/4. Pozostały więc warstwy AAL 1, 3/4 i 5.
Usługa klasy A
Usługa klasy A korzysta z warstwy AAL 1. Jest to komunikacja połączeniowa, synchroniczna, ze stałą szybkością transmisji bitów (ang. CBR - Constant Bit Rate). Takie charakterystyki transmisji są niezbędne do obsługi komunikacji izochronicznych, takich jak przesyłanie głosu czy nawet wysokiej jakości transmisje wideo. Dodatkowo warstwa AAL 1 może być wykorzystywana do emulacji łączy DS-1 w obwodach N-carrier. Warto zauważyć, że dane są dostarczane do warstwy AAL I przy stałej szybkości transmisji bitów; AAL nie może sama uporządkowywać chaosu - nad stałą szybkością przesyłania bitów muszą czuwać aplikacje. Jeśli aplikacja nie jest w stanie utrzymać od początku do końca stałej szybkości transmisji danych, protokoły ATM mogą to skorygować tylko w nieznacznym stopniu.
Więcej informacji o łączach DS-1 i systemie N-carrier można znaleźć w rozdziale 14 pt. „ Linie dzierżawione".
Jednostki SAO-PDU tworzone przez warstwę AAL I, oprócz części użytecznej, mają jeszcze dwa 4-bitowe pola. Polami tymi są: Numer Sekwencji (ang. SN - Sequence Number) i Ochrona Numeru Sekwencji (ang. SNP - Sequence Number Protection). Numer sekwencji jest 4-bitowym licznikiem zerowanym dla każdej jednostki danych segmentowanej przez mechanizm SAR. Pole SNP to 4-bitowa liczba cyklicznej kontroli nadmiarowej (ang. CRC - Cyclical Redundancy Check), obliczana wyłącznie na podstawie 4 bitów pola Numer Sekwencji. Pola te zmniejszają rozmiar części użytecznej komórki AAL 1 do 47 oktetów. Rysunek I 1.5 przedstawia przykład jednostki AAL 1 SAO-PDU.
Rysunek 11.5. Jednostka AAL I SAO-PDU
Usługa klasy B



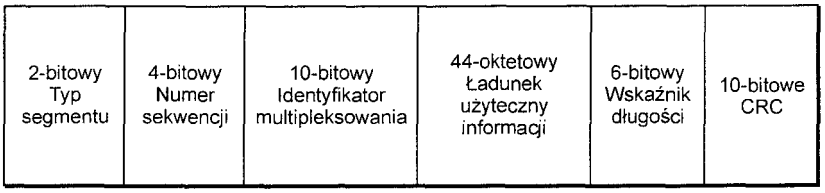
Komunikacja klasy B jest podobna do komunikacji klasy A. Obie są połączeniowe i synchroniczne. Jedyna różnica jest taka, że klasa B nie wymaga stałej szybkości przesyłania bitów. Zamiast niej wykorzystuje zmienną szybkość transmisji bitów (ang. UBR - Variable Bit Rate). Inaczej mówiąc, aplikacja ma wysoką tolerancję czasową, a jej

transmisje, chociaż muszą być dokładnie synchronizowane, przychodzą jednak w nieregularnych odstępach. Tak subtelna różnica wystarcza, aby wyodrębnić kolejną warstwę adaptacji - AAL 2.
Niektóre aplikacje wideo, zwłaszcza te, które wykorzystują algorytmy odświeżające wyłącznie zmienione piksele, są idealne dla klasy B. Niestety, ATM Forum nigdy nie ukończyło specyfikacji warstwy AAL 2. Z tego powodu wiele źródeł wspomina tylko o warstwach AAL 1,3,4 i 5, choć 3 i 4 zostały połączone i oznaczone jako „ 3/4". Bez specyfikacji określającej jej strukturę każda próba zilustrowania warstwy AAL 2 SAO-PDU jest czystą spekulacją- dlatego też próby takiej nie podjęto.
Usługa klasy C
Komunikacja klasy C jest komunikacją połączeniową i wykorzystuje zmienną szybkość transmisji bitów. Co ważniejsze, nie ma wymagań co do taktowania, które trzeba by uzgadniać między stacjami końcowymi; jest to komunikacja asynchroniczna. Pierwotnie komunikacja klasy C miała być zaspokajana przez warstwę AAL 3. Podczas prac rozwojowych zdecydowano jednak połączyć AAL 3 i AAL 4 w pojedynczą warstwę AAL, która mogłaby obsługiwać zarówno klasę C, jak i D.
Warstwa powstała z połączenia została w niezbyt wyszukany sposób nazwana 3/4 (wymawiane jako „trzy-cztery", a nie jako „trzy czwarte"). Spoglądając wstecz i wiedząc, że podzbiór klasy C zasilił jeszcze jedną warstwę AAL (AAL 5), wydaje się jednak, że być może wymowa „trzy czwarte" byłaby bardziej trafna!
Wiele protokołów, w tym protokoły warstwy 2, takie jak Frame Relay i X.25 oraz zestawy protokołów wyższej warstwy, jak TCP/IP i IPX/SPX, reprezentuje komunikację klasy C. Choć protokoły te są z zasady bezpołączeniowe, ATM wymaga, żeby połączenie zostało ustanowione, zanim komórki będą mogły być przesyłane między dwiema stacjami. Po nawiązaniu wirtualnego połączenia nie ma żadnych ograniczeń co do zawartości części użytecznej tych komórek. Protokół połączeniowy może dokonywać opakowywania danych otrzymanych z protokołów bezpołączeniowych i przesyłać je w postaci pakietów.
Stwarza to konieczność ścisłej zbieżności funkcji protokołów wyższych warstw w ramach mechaniki warstwy ATM. Tu przydatna okazuje się podwarstwa zbieżności (CS). Podwarstwa ta tworzy jednostkę CS-PDU (jednostkę danych protokołu podwarstwy zbieżności) dla warstwy AAL 3/4. Ta jednostka CS-PDU jest pokazana na rysunku 1 1.6.
Rysunek 11.6. Jednostka AAL 3/4 CS-PDG'.

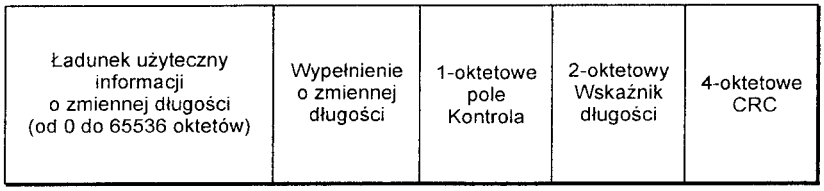


Jak widać na rysunku 11.6, jednostka AAL 3/4 CS-PDU ma następującą strukturę: • 1-oktetowy wskaźnik części wspólnej (ang. CPI-Common Part Indicator) • ł-oktetowa etykieta początku (ang. BTag- Beginning Tag)
• 2-oktetowy wskaźnik rozmiaru alokacji bufora (ang. BASize - Buffer Allocation Size Indicator)
• zmiennej długości ładunek użyteczny informacji (ktoś mądrze uniknął stosowania wobec tego pola szeroko rozpowszechnionego skrótu „IP"!)
• wypełnienie o zmiennej długości (ang. PAD-padding) • 1-oktetowe pole wyrównanie (ang. AL -Alignment)
• 1-oktetowa etykieta końca (ang. ETag- End Tag) • 2-oktetowe pole długość
Każde z tych pól jest ważne dla funkcjonowania warstwy adaptacji ATM. Pole CPI wskazuje użycie pozostałych pól jednostki CS-PDU, w zależności od wykonywanej operacji. Pole BTag zawiera wartość numeryczną. Taka sama wartość jest umieszczona w polu ETag. Razem określają one początek i koniec każdej jednostki CS-PDU. Jest to konieczne, ponieważ operacje przesyłania do następnych stacji mogą skutkować dalszym podziałem jednostki CS-PDU na nieokreśloną jak dotąd liczbę jednostek SAO-PDU.
Wskaźnik rozmiaru alokacji bufora służy do informowania odbierającej warstwy AAL o tym, jakiej wielkości bufor jest niezbędny do ponownego złożenia odbieranej jednostki CS-PDU. Długość pola Ładunek użyteczny informacji nie może być większa niż maksymalna wartość, jaką można zapisać w poprzednim polu. Ponieważ pole BASize jest strukturą binarną o długości 2 oktetów, maksymalna wartość, jaką można w tym polu zapisać, wynosi 2 do potęgi 16, czyli 65536. Taka jest maksymalna wielkość części użytecznej informacji. Oczywiście jest to dużo więcej niż wynosi maksymalna wielkość części użytecznej komórki ATM, więc niezbędna jest dalsza segmentacja.
Czasem wskaźnik rozmiaru alokacji bufora nie opisuje rzeczywistej wielkości części użytecznej, tylko jest automatycznie ustawiany na dużą (jeśli nie maksymalną) wartość. Jest to przydatne w aplikacjach strumieniowych, gdzie warstwa AAL może nie znać dokładnego rozmiaru jednostki PDU.
Pole PAD jest wykorzystywane, jeśli ładunek użyteczny informacji nie jest równy wielokrotności liczby 32. Służy jako wypełnienie, mające zapewnić stałą, konsekwentną, logiczną długość i zawiera wyłącznie zera.
Kolejnym polem jest Wyrównanie. Jest to 1-oktetowe pole, wykorzystywane tylko do powiększenia jednostki CS-PDU do 4 oktetów. Pole zawiera same zera i funkcjonuje jako wypełnienie strukturalne.
Dwa ostatnie pola to Etykieta Końca i Długość. Pole ETag, jak zostało wcześniej wyjaśnione, zawiera taką samą wartość co pole BTag. Pole Długość przechowuje rzeczywistą długość pola Ładunek Użyteczny Informacji.
Następnie jednostka CS-PDU zostaje przekazana do podwarstwy SAR, gdzie jest dzielona na odpowiednią liczbę jednostek SAO-PDU. Jednostka SAO-PDU ma następującą strukturę:
• 2-bitowy Typ segmentu (ang. ST - Segment Type)
• 4-bitowy Numer sekwencji (SN)
• 10-bitowa Identyfikacja multipleksowania (ang. *MID - Multiplexing Identification)*
• 44-oktetowy Ładunek użyteczny informacji
• 6-bitowy Wskaźnik długości (ang. *LI - Length Indicator)*

• 10-bitowe pole cyklicznej kontroli nadmiarowej (CRC)
Jak można obliczyć, jednostka SAO-PDU ma długość 48-oktetów - dokładnie tyle samo, ile wynosi rozmiar części użytecznej komórki ATM. Aby stworzyć komórkę ATM, wystarczy już tylko dodać odpowiedni, 5-oktetowy nagłówek. Jednostka SAO-PDU jest przedstawiona na rysunku 11.7.
Rysunek 11.7. Jednostka AAL 3/4 SAO-PDU



Ponieważ większość z tych pól jest już znana, wystarczy, jeśli opisane zostaną tylko pola dotychczas nie prezentowane. Jedynym takim polem jest Identyfikacja Multipleksowania (MID). Pole MID służy do identyfikowania wszystkich komórek utworzonych z danej jednostki CS-PDU. Struktury te mogą mieć długość do 65536 oktetów, więc można podzielić je na 1365 komórek! Pole MID zapewnia, że komórki są jednoznacznie identyfikowalne, jako przynależące do tej samej jednostki CS-PDU. Teoretycznie odbiorca mógłby otrzymywać jednocześnie wiele strumieni komórek pochodzących z różnych źródeł - stąd niezbędne jest istnienie pola MID.
Warstwa AAL 3/4 jest dość złożona i niewygodna w użyciu. Dlatego opracowano warstwę AAL 5, obsługującą tylko pewną część komunikacji klasy C. AAL 5 jest tą warstwą adaptacji, która będzie używana najczęściej. Taki stan przewiduje się, przyjmując dwa dość ryzykowne założenia. Po pierwsze, że ATM będzie rzeczywiście wykorzystywany, a po drugie, że większość obsługiwanych aplikacji będzie polegać na tradycyjnych, bezpołączeniowych protokołach komunikacyjnych.
Warstwa AAL 5 została zaprojektowana jako prosta i wydajna warstwa adaptacji, ukierunkowana na najczęściej wykorzystywane obszary warstwy AAL 3/4. Również tworzy jednostki CS-PDU, choć dużo bardziej sprawniejsze niż te opisywane wcześniej.
Pierwszym polem jednostki AAL 5 CS-PDU jest Ładunek użyteczny informacji. Pole to może mieć długość od 0 do 65536 oktetów, choć 0-oktetowy ładunek użyteczny wydaje 'się być bezcelowy. Potem następuje pole wypełnienia PAD. Wypełnienie stosuje się, by zapewnić, że wielkość części użytecznej będzie wielokrotnością liczby 48, co ułatwia tworzenie komórek w warstwie ATM. Następnym polem jest 1-oktetowa Kontrola, która jak dotąd nie doczekała się zastosowania, więc domyślnie jest ustawiana na zero. Dwa ostatnie pola to Długość i CRC. I znów, pole Długość ma 2 oktety i określa maksymalny rozmiar pola części użytecznej informacji. Pole CRC ma długość 4 oktetów, a jego wartość jest obliczana na podstawie całej zawartości jednostki CS-PDU (od początku części użytecznej informacji do końca pola Długość).
Tę jednostkę CS-PDU przedstawia rysunek 11.8.
Rysunek 11.8. Jednostka AAL
5 CS'-PD(i.



Jednostka AAL 5 CS-PDU jest przekazywana do podwarstwy SAR, gdzie jest przekształcana w jednostkę SAO-PDU. Jednostka AAL 5 SAO-PDU jest niezbyt spektakularną strukturą, z częścią użyteczną o długości zaledwie 48 oktetów, bez specyficznych nagłówków czy stopek. Prostota tej warstwy AAL czyni ją względnie niedrogą i łatwą w implementacji. Fakt, że warstwa AAL 5 jest przystosowana do tego, by obsługiwać większą część ruchu w sieci ATM, również stawia ją w roli „wołu roboczego" warstw adaptacji.
Usługa klasy D
Usługa klasy D to bezpołączeniowy, asynchroniczny transfer danych. Jest użyteczna przy przesyłaniu komunikacji sieci LAN lub SMDS przez szkielet ATM. AAL 3/4 jest kombinowaną warstwą adaptacji, obsługującą komunikację klasy C i D. Właściwie sensowne byłoby stwierdzenie. że jedyną różnicą między klasami C i D jest to, że jedna jest połączeniowa, podczas gdy druga jest bezpołączeniowa.

1.11.8 Warstwa ATM

Warstwa ATM i warstwy adaptacji ATM odpowiadają warstwie łącza danych modelu referencyjnego OSI. Warstwa ATM jest położona niżej niż warstwy adaptacji i odpowiada za ustanawianie połączeń wirtualnych oraz za przekazywanie za ich pomocą komórek otrzymanych z warstwy AAE - Funkcjonowanie warstwy ATM w dużym stopniu zależy od tego, czy jest umiejscowiona w stacji końcowej, czy w urządzeniu przełączającym. Dlatego sensowne jest rozpatrzenie tych alternatyw niezależnie od siebie.
Stacja końcowa
Warstwa ATM w stacji końcowej powinna mieć możliwość sygnalizowania innym stacjom, że ma do przesłania dane dla nich przeznaczone, oraz uzgodnienia ze stacją (bądź stacjami) konstrukcji komutowanego obwodu wirtualnego, czyli obwodu SVC (ang. Switched Virtual Circuit). Wydawać się to może nieskomplikowane, ale materiału dotyczącego samych technik sygnalizowania wspomagających konstruowanie obwodu starczyłoby na kilka osobnych książek.

Zanim warstwa ATM będzie mogła przyjmować dane z warstwy AAL, musi zostać utworzony obwód logiczny. Dane z AAL przychodzą w postaci jednostki danych protokołu, która musi zostać przekształcona w komórkę. Dokonuje się tego łatwo, poprzez dodanie pól nagłówka i wypełnienie ich zawartością.

Przełącznik

Funkcjonowanie warstwy ATM w przełączniku jest dużo prostsze niż jej odpowiednika w stacji końcowej. Po odebraniu komórki z któregoś ze swoich portów musi sprawdzać wartości identyfikatora ścieżki wirtualnej (VPI) oraz identyfikatora obwodu wirtualnego (VCI), zawarte w nagłówku komórki i porównywać je ze swoją tablicą przyporządkowań VPI/VCI. Przeglądanie tablicy służy dwóm ważnym celom. Po pierwsze, określa, jakie powinny być nowe przyporządkowania VPI/VCI. Po drugie, identyfikuje port, do którego komórka musi być przesłana.

Takie są podstawowe funkcje warstwy ATM przełącznika. Dodatkowe funkcje mogą obejmować ustawianie pola Wskaźnik Typu Ładunku Użytecznego (ang. PT!- Payload Type Indicator)-jeśli w sieci są zatory- i implementowanie wszelkich taktyk kierowania ruchem, które w efekcie tego można zastosować. Warstwa ATM musi także zapewnić możliwość buforowania i porządkowania komórek, co staje się koniecznością w przypadku, gdy wiele portów wejściowych rywalizuje o dostęp do jednego portu wyjściowego.

„Przekazywanie komórki do portu" oznacza, że warstwa ATM przekazuje komórkę elementom warstwy fizycznej tego portu. Warstwa fizyczna oblicza wtedy ponownie wartość pola HEC, a następnie tworzy i wysyła strumień bitów. W efekcie postać przesyłanej komórki różni się znacznie od postaci, jaką komórka ta miała w momencie jej otrzymania jej nagłówek zostaje znacząco modyfikowany.

1.11.9 Komórka

Po zapoznaniu się ze sposobem, w jaki warstwy AAL tworzą podstawowe jednostki PDU, zbadanie struktury właściwych komórek ATM jest proste. Niemniej jednak komórka jest podstawową strukturą transportu danych w ATM. W odróżnieniu od większości sieci LAN, ta struktura warstwy 2 ma ustaloną, niezmienną długość. Komórka ATM ma zawsze 53 oktety. Może się to wydawać liczbą dość niezwykłą, zwłaszcza dla tych, którzy przyzwyczaili się do systemów liczbowych o podstawie 2 i/lub 16. Nie bez powodu! Długość komórki ATM wywodzi się z wypróbowanej zębem czasu zasady kompromisu,

a nie z matematyki. Północnoamerykańscy współtwórcy standardu CCITT ATM optowali za 64-oktetową częścią użyteczną, reprezentanci Europy i Azji uważali, że 32-oktetowy ładunek użyteczny będzie bardziej odpowiedni. Komitet CCITT wyciągnął średnią arytmetyczną i przyjął jako standard 48-oktetową część użyteczną z 5-oktetowym nagłówkiem - stąd ta niezbyt intuicyjna długość. W rzeczywistości istnieją dwie różne struktury komórek, jedna dla interfejsu UNI, druga dla NNI. Różnice występują w formacie 5-oktetowego nagłówka.

1.11.9.1 **Struktura komórki** UNI

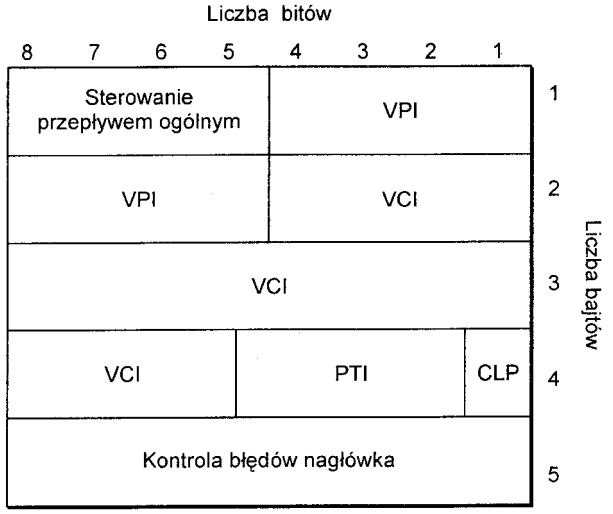
Komórka UNI ma następującą strukturę:

• 4-bitowe sterowanie przepływem ogólnym (ang. GFC- Generic Flow Control), • 8-bitowy identyfikator ścieżki wirtualnej (VPI),
• 16-bitowy identyfikator kanału wirtualnego (VCI),
• 3-bitowy wskaźnik typu części użytecznej (PTI),
• 1-bitowy priorytet straty komórki (ang. CLP - Cell Loss Priority),
• 8-bitowa kontrola błędów nagłówka (ang. HEC - Header Error Check),
• 48-oktetowy ładunek użyteczny.

Komórka UNI jest przedstawiona na rysunku 1 ł.9.

Rysunek 11.9. Struktura komórki UNI.

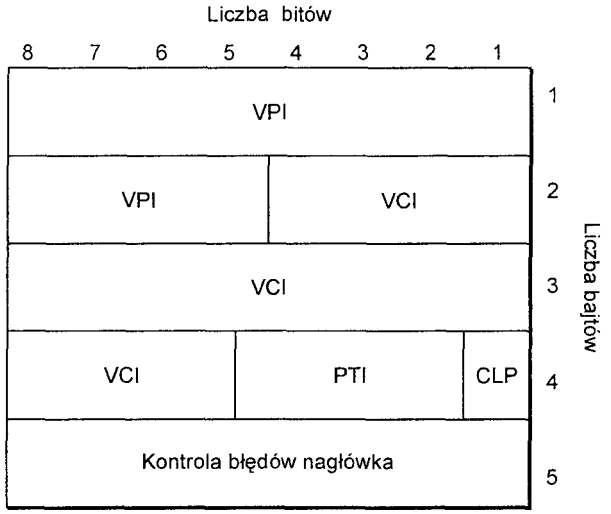
Pole sterowania przepływem ogólnym zostało przewidziane jako mechanizm regulowania przepływu danych (komórek) wyłącznie w interfejsie UNI. Ustalenie potrzeby wprowadzenia tego mechanizmu oraz długości jego pola było łatwym zadaniem. Określenie właściwej metody i procesów, za pomocą których mógł być regulowany przepływ danych między dwoma stacjami końcowymi, okazało się znacznie bardziej zniechęcające. Ze wszystkich usług wyższego poziomu, które należało zdefiniować, aby dostosować ATM



do środowiska LAN, sterowanie przepływem okazało się być jedną z najbardziej kontrowersyjnych. Dziś uważa się, że pole GFC ma jedynie wartość lokalną, tzn. jego zawartość nie jest przenoszona przez cały obwód wirtualny.
Pole VPI komórki UNI ma długość 8 bitów. Służy do identyfikowania ścieżek wirtualnych, choć jest to w większym stopniu funkcja NNI. Dlatego pole UNI VPI jest domyślnie wypełniane zerami i może otrzymać inne wartości w celu usprawnienia funkcji zarządzania siecią. Gdy komórka dotrze do obszarów NNI sieci, pole VPI może być wykorzystywane przez przełączniki NNI. Należy pamiętać, że przełączniki nie są prostymi wzmacniakami: przetwarzają komórkę i przepisują jej nagłówek, zanim prześlą ją dalej.
Identyfikator kanału wirtualnego jest 16-bitowym polem, jednoznacznie identyfikującym wirtualne połączenie, ustanowione podczas procesu sygnalizacji. Przełącznik ATM wykorzystuje interfejs VPI w połączeniu z VCI, do skierowania odebranych komórek do odpowiedniego portu fizycznego.
Wskaźnik typu części użytecznej jest 3-bitowym polem, w którym każdy bit jest znaczący. • Pierwszy bit wskazuje, czy ładunek użyteczny komórki pochodził z płaszczyzny użytkownika, czy z płaszczyzny sterowania.
• Drugi bit wskazuje, czy komórka napotkała zator w sieci.
• Trzeci bit służy do sygnalizowania ostatniej komórki z serii pochodzących z jednej jednostki AAL 5 PDU.
Pole Priorytet Straty Komórki ma długość I bitu. Wskazuje, czy komórka powinna, czy nie powinna być odrzucana, jeśli przechodząc przez sieć, napotka duże zatory. Tradycyjne aplikacje wymagają gwarancji integralności podczas dostarczania i zlecają protokołom warstwy 4 porządkowanie pakietów dostarczonych w niewłaściwej kolejności. Ważniejsze jest, aby pakiety przybyły nienaruszone, niż żeby przybyły na czas i w odpowiedniej kolejności.
Zupełnie inaczej jest w przypadku czułych na opóźnienia aplikacji nowej generacji, takich jak wideokonferencje. Takie aplikacje wymagają, by dane przybywały na czas. Jeśli pakiety są spóźnione albo uszkodzone podczas transmisji, są po prostu odrzucane przez odbierającą je aplikację. Czynnik czasu ma tutaj najwyższy priorytet. Tak więc bit priorytetu straty komórki dostarcza sieci mechanizmu rozróżniania komórek wysoko i nisko uwarunkowanych czasowo. Komórki „wysoko ceniące" czas, które zostaną odrzucone przez aplikację (z powodu opóźnienia), nie muszą przebywać całej drogi. Sieć może po prostu wcześniej je odrzucić.
Ostatnim polem komórki UNI jest Kontrola Błędów Nagłówka. To 8-bitowe pole zawiera sumę kontrolną, obliczoną na podstawie zawartości pól nagłówka komórki (z wyjątkiem samego pola HEC).

1.11.9.2 Struktura komórki NNI

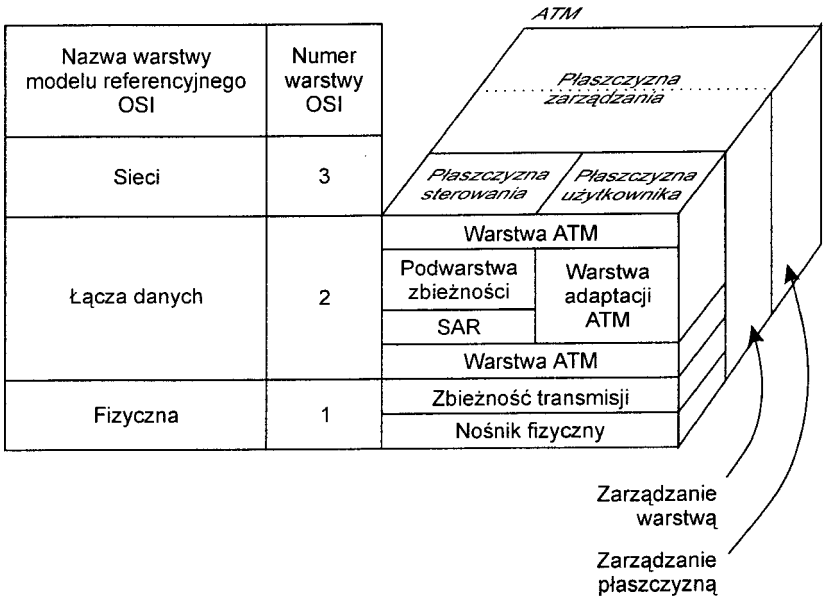
Komórka NNI została zaprojektowana dla potrzeb standardu ATM, który pierwotnie miał pełnić funkcję transportu pomiędzy centralami telefonicznymi w szerokopasmowych sieciach ISDN. Funkcja ta odzwierciedla się w jej strukturze. Komórka NNI zawiera następujące pola:
• 12-bitowy Identyfikator ścieżki wirtualnej (VPI) • 16-bitowy Identyfikator kanału wirtualnego (VCI) • 3-bitowy Wskaźnik typu części użytecznej (PTI) • 1-bitowy Priorytet straty komórki (CLP)
• 8-bitową Kontrolę błędów nagłówka (HEC) • 48-oktetową część użyteczną
Komórka ta jest przedstawiona na rysunku 1 1.10.
Rysunek 11.10. Struktura komórki NNI
Wszystkie pola komórki NNl występują również w komórce UNI, więc nie warto ponownie ich omawiać. Główne różnice między polami komórek UNI i NNI to: brak 4-bitowego pola sterowania przepływem i rozszerzenie pola VPI do 12-bitów. Rozszerzenie pola VPI ma zasadnicze znaczenie w szkieletach sieci. Główną rolą szkieletu jest skupianie ruchu sieciowego. Tak więc może on posiadać wiele ścieżek wirtualnych, złożonych z niezliczonych obwodów wirtualnych, występujących w sieci.



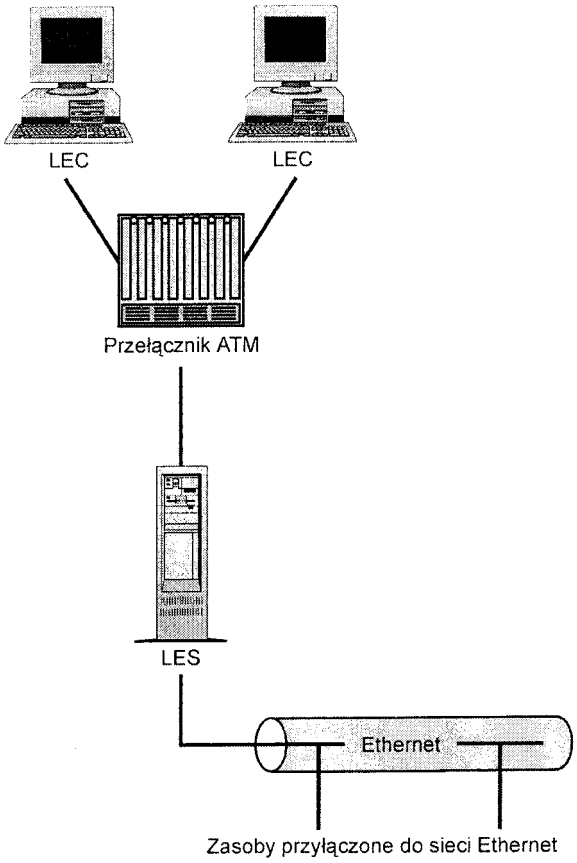
1.11.10 Emulacja sieci LAN

Przez długi czas uważano, że jedną z najpoważniejszych przeszkód dla zaakceptowania ATM jako technologii LAN była ogromna liczba zainstalowanych aplikacji, przeznaczonych dla istniejących technologii LAN. Było to prawdą, zwłaszcza w odniesieniu do specyfikacji IEEE 802.3 (Ethernet) i IEEE 802.5 (Token Ring). Sieci lokalne ATM różnią się od tych sieci lokalnych IEEE w trzech aspektach:

• Sieci LAN ATM są sieciami połączeniowymi, podczas gdy inne sieci zachęcają aplikacje do wykorzystywania komunikacji bezpołączeniowej.
• Ethernet i Token Ring mogą z łatwością przeprowadzać nadawanie i multicasting przez wspólne nośniki.
• Adresy MAC IEEE opierają się na seryjnych numerach wytwórców, a nie na topologii LAN.
ATM Forum stawiło czoła tym wyzwaniom, opracowując nową usługę ATM, emulującą środowisko sieci LAN lEEE. Usługa ta jest nazywana emulacją sieci lokalnej (ang. LANE - Local Area Network Emulation).
Sieci zbudowane na podstawie zestawu protokołów LANE są znane jako emulowane sieci lokalne (ELAN). Sieci ELAN mogą emulować Ethernet 802.3 lub Token Ring 802.5, ale nie jednocześnie, gdyż nie został zdefiniowany mechanizm mostkowania tłumaczącego.
Emulacja LANE umożliwia systemom końcowym - którymi mogą być dowolne urządzenia przyłączone do obrzeży sieci - bezpośrednie połączenie z siecią LAN ATM. Ich aplikacje programowe działają, jak gdyby były połączone bezpośrednio z siecią Ethernet lub Token Ring. Taka możliwość ma kluczowe znaczenie dla organizacji, które chcą łagodnie i stopniowo przejść z innych sieci na sieć LAN ATM.
LANE umożliwia klientom ATM LAN wykorzystywanie istniejących aplikacji sieci LAN za pomocą warstwy programowej zainstalowanej w urządzeniach emulujących. Warstwa programowa staje się integralną częścią stosu protokołów ATM takiego urządzenia. Powstały w ten sposób stos jest przedstawiony na rysunku 11.11.
Rysunek 11.11. Stos protokołów LANE



Urządzenie emulujące wykorzystuje nowy interfejs - interfejs LANE użytkownika z siecią (ang. LUNI - *LANE User to Network Interface).* Ten termin to tylko wierzchołek góry lodowej akronimów. Na przykład, emulowany LAN jest nazywany „ELAN". Każdy klient w ELAN jest klientem emulacji LAN, w skrócie LEC (ang. *LAN Emulation Client).* Poszczególni klienci LEC są identyfikowani poprzez ich adresy MAC.
Każdy ELAN musi mieć tzw, usługę emulacji LAN, czyli LES (ang. LAN Emulation Service). Usługa LES może być zlokalizowana w przełączniku lub stacji końcowej. Niezależnie od tego, LES składa się z serwera konfiguracji emulacji LAN (ang. LECS GAN Emulation Configuration Server) oraz usługi „nadawanie a nieznany serwer" (ang. BUS - Broadcast and Unknown Server).
Jakby tego było mało, istnieje nawet nowy akronim dla typu połączenia logicznego, ustanawianego pomiędzy klientami LEC a usługą LES w sieci ELAN. Akronimem tym jest VCC, co oznacza: Połączenie kanału wirtualnego (ang. Virtual Channel Connection). Połączeń takich wymagają płaszczyzny sterowania i użytkownika.
Rysunek 11.12 z pewnością ułatwi zrozumienie wzajemnych zależności pomiędzy powyższymi akronimami.
Rysunek 11.12.
Typowa konfiguracja LANE.



Na rysunku tym przedstawiono dwóch klientów LEC i jedną usługę LES (zlokalizowaną na serwerze). Serwer LES pełni rolę wrót do sieci LAN IEEE. Klienci mogą dzięki usłudze LES uzyskiwać dostęp do zasobów tej sieci LAN. Usługa ta przeprowadza emulację MAC.

1.11.11 Podsumowanie

Oryginalna koncepcja ATM pozostaje jedyną technologią dla sieci LAN i WAN, mogącą obsługiwać zarówno izochroniczną komunikację o dużej szerokości pasma i małych opóźnieniach, jak też bardziej tradycyjną, interaktywną i masową transmisję danych. Wyzwaniami stojącymi przed ATM i ATM Forum pozostają niezmiennie:

• Ukończenie specyfikacji w ramach zdominowanego przez politykę korporacyjną procesu ustalania standardów otwartych.

• Opracowanie efektywnego ekonomicznie zbioru rozwiązań ATM, który pozwoli na pozyskanie klientów z innych platform o ustalonej pozycji na rynku.

Przez wiele lat ciężko było określić, które z tych wyzwań ma większy ciężar gatunkowy; dzisiaj wydaje się, że ATM Forum odzyskało utracony rozpęd i szybko zmierza do ukończenia pakietu otwartych specyfikacji ATM. To dobra wiadomość.

Złą wiadomością jest natomiast ta, że środowiska konkurencyjne dla ATM nie uległy bynajmniej stagnacji. W ciągu kilku minionych lat wiele technologii komunikacyjnych zostalo poddanych przedłużającym życie modyfikacjom i nowelizacjom. Pod względem możliwości technicznych i stosunku ceny do wydajności są one obecnie dla ATM jeszcze groźniejszą konkurencją niż kiedykolwiek. Tak więc, choć wydaje się, iż ATM sprostał pierwszemu wyzwaniu, sprostanie drugiemu z nich jawi się jako zadanie jeszcze trudniejsze. I tylko czas pokaże, czy ATM Forum uda się zrealizować swoją wizję ATM.

1.12 Rozdział 12 Protokoły sieciowe

Mark A. Sportack

Termin „protokoły sieciowe" odnosi się przede wszystkim do protokołów warstwy 3 modelu OSI. Protokoły zapewniają adresowanie, dzięki któremu dane mogą być dostarczane na nieokreślone odległości, poza domenę sieci lokalnej nadawcy. Przeważnie protokoły warstwy 3 wykorzystują do transportu danych strukturę znaną jako pakiet.

Choć protokoły warstwy 3 dostarczają mechanizmów niezbędnych do wysyłania pakietów, nie są na tyle wyszukane, aby mieć pewność, że pakiety zostały rzeczywiście odebrane i to we właściwym porządku. Zadania te pozostawiono protokołom transportowym warstwy 4. Protokoły te przyjmują dane z wyższych warstw i osadzają je w segmentach, które przekazują warstwie 3.

W tym rozdziale opisane są funkcje i wzajemne oddziaływania między stosami protokołów warstwy 3 i 4, a następnie badane są zawiłości niektórych najpopularniejszych protokołów sieciowych.

1.12.1 Stosy protokołów

Stos protokołów to komplet powiązanych protokołów komunikacyjnych, oferujących użytkownikowi mechanizmy i usługi potrzebne do komunikacji z innymi maszynami włączonymi do sieci. Z perspektywy użytkownika stos protokołów jest tym, co czyni sieć zdatną do użycia.

W poprzednich rozdziałach omówiono pierwszą i drugą warstwę stosu protokołów (tj. warstwę fizyczną i warstwę łącza danych). Są one mocno zintegrowane i powiązane ze sobą. Warstwę fizyczną narzuca wybrana architektura warstwy łącza danych, jak Ethernet, Token Ring itd.

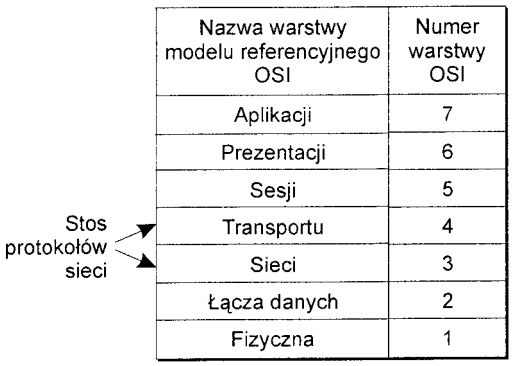
W obecnej epoce sieci i systemów otwartych wybór określonej architektury LAN nie ogranicza możliwości wyboru protokołów wyższego poziomu. Stos protokołów powinien oferować mechanizmy sprzęgające z istniejącymi, znormalizowanymi środkami dostępu do sieci dla protokołów warstwy łącza danych.

Podobnie jak to było w przypadku warstw pierwszej i drugiej, warstwa 3 modelu referencyjnego OSI jest ściśle powiązana z warstwą 4. Warstwa 3 to warstwa sieci, zaś warstwa 4 jest warstwą transportu. Są one przedstawione na rysunku 12.1. Razem zapewniają one mechanizmy umożliwiające przesyłanie informacji między urządzeniami nadawcy i odbiorcy, z wykorzystaniem sieci komunikacyjnej sięgającej poza domenę warstwy 2. Zapewniają też inne funkcje, takie jak zmienianie porządku pakietów otrzymanych w niewłaściwej kolejności lub ponowną transmisję pakietów, które nie dotarły do odbiorcy lub dotarły uszkodzone.

Protokoły te, aby faktycznie przesłać dane, muszą wykorzystywać architekturę LAN warstwy 1 i 2, muszą też mieć pewne środki sprzęgające je z tymi warstwami. Używają do tego techniki opakowywania. Protokół warstwy 4 umieszcza w segmentach dane otrzymane od protokołów wyższej warstwy. Segmenty te są przekazywane do odpowiedniego protokołu warstwy 3. Protokół warstwy 3 bezzwłocznie opakowuje segment strukturą pakietu z adresami nadawcy i odbiorcy, po czym przekazuje pakiet protokołowi warstwy 2. Warstwa 2 umieszcza ten pakiet danych warstwy 3 w ramce, opatrując go przy tym adresowaniem, przeznaczonym dla urządzeń warstwy 3, takich jak routery i przełączniki IP. Umieszczenie pakietu warstwy 3 (IP) w ramce warstwy 2 (Ethernet) przedstawia rysunek 12.2. Strukturę pakietu zalicza się do części pola danych ramki. choć w rzeczywistości pakiet to nie dane, tylko struktura innej warstwy.

Rysunek 12.1. Warstwy sieci

1 transportu w modelu referencyjnym OSI.



Nazwa warstwy modelu referencyjnego OSI	Numer warstwy OSI
Aplikacji	7
Prezentacji	6
Sesji	5
Transportu	4
Sieci	3
Łącza danych	2
Fizyczna	1

Rysunek 12.2. Umieszczenie pakietu IP w ramce Ethernet.

7-oktetowa Preambuła	1-oktetowy Ogranicznik początku ramki	6-oktetowy Adres odbiorcy	6-oktetowy Adres nadawcy	2-oktetowe pole Długość	Min. 20-oktetowy nagłówek IPX	Min. 20-oktetowy nagłówek TCP	Pole Dane o zmiennej długości od 46 do 1482 oktetów	4-oktetowa Sekwencja kontrolna ramki

Ramki służą do transportowania i adresowania danych dla międzysieciowego urządzenia warstwy 3, znajdującego się na krawędzi warstwy 2 (dziedziny ramek). Urządzenie - zwykle jest to router - przyjmuje obramowany pakiet, usuwa ramkę i czyta informację adresową przeznaczoną dla warstwy 3. Informacja ta służy do ustalenia następnego „skoku" na drodze do miejsca przeznaczenia pakietu. Pakiet jest następnie kierowam do kolejnego punktu. Ostatni router przed miejscem przeznaczenia pakietu musi ponownie umieścić pakiet w strukturze ramki warstwy 2, zgodnej z architekturą sieci LAN w miejscu przeznaczenia.

Warstwa 3 zapewnia tylko międzysieciowy transport danych. Warstwa 4 (transportu) wzbogaca mechanizmy sieciowe warstwy 3 o gwarancje niezawodności i integralności końcowej (czyli na całej długości połączenia). Warstwa transportu może zagwarantować wolne od błędów dostarczenie pakietów i odpowiednie ich uszeregowanie, jak też zapewnić odpowiednią jakość usług. Przykładem mechanizmu warstwy 4 jest protokół TCP (ang. Transmission Control Protocol). TCP niemal zawsze występuje razem ze swoim odpowiednikiem z warstwy 3, protokołem internetowym IP, jako „TCP/IP". Wykorzystywanie przez aplikacje warstw 3 i 4 do przesyłania danych do innych komputerów/aplikacji sugeruje, że komputery nadawcy i odbiorcy nie są przyłączone do tej samej sieci lokalnej, niezależnie od tego, jaka odległość je dzieli. Dwie różne sieci muszą być ze sobą połączone, by obsłużyć żądaną transmisję. Dlatego mechanizmy komunikacyjne warstwy 2 są niewystarczające i muszą być rozszerzone o adresowanie warstwy 3. Choć warstwy 3 i 4 istnieją właśnie w tym celu (tzn. w celu łączenia ze sobą różne sieci) , to aplikacje mogą przesyłać do siebie dane, wykorzystując protokoły tych warstw, nawet jeśli są przyłączone do tej samej sieci i podsieci LAN. Na przykład, jeśli komputery nadawcy i odbiorcy są przyłączone do tej samej sieci lokalnej, mogą się ze sobą komunikować, wykorzystując tylko ramki i protokoły warstwy 2. Niektóre aplikacje mogą jednak wymagać wspomagania swojej komunikacji pewnymi właściwościami protokołów wyższej warstwy.

Istnieją dwa rodzaje protokołów sieciowych działających w warstwie 3: protokoły trasowane i protokoły trasujące. Protokoły trasowane to te, które umieszczają dane oraz informacje użytkownika w pakietach i są odpowiedzialne za przesłanie pakietów do odbiorcy. Protokoły trasujące stosowane są pomiędzy routerami i określają dostępne trasy, komunikują o nich i przeprowadzają nimi pakiety protokołów

trasowanych. Protokoły trasujące są dokładniej omówione w rozdziale I3 pt. „Sieci WAN". W niniejszym rozdziale koncentrujemy się na najpopularniejszych protokołach trasowanych.

1.12.2 Protokół Internetu, wersja 4 (Ipv4)

Protokół Internetu (IP)' został opracowany około 20 lat temu dla Departamentu Obrony USA (ang. Department of Defense). Departament Obrony szukał sposobu połączenia różnych rodzajów posiadanych komputerów i sieci je obsługujących w jedną, wspólną sieć. Osiągnięto to za pomocą warstwowego protokołu, który odizolował aplikacje od sprzętu sieciowego. Protokół ten używa modelu nieco różniącego się od modelu referencyjnego OSI. Jest on znany jako model TCP/IP.

Należy w tym miejscu przypomnieć, iż rozpowszechnione obecnie tłumaczenie angielskiego terminu Internet Protocol (którego akronimem ‚jest właśnie IP) jako ‚.protokół internetowy" ‚jest konsekwencją rozpowszechnienia się zastosowań Internetu - wierniejszym wydaje się bowiem używany pierwotnie odpowiednik „protokół międzysieciowy'. Sam mianowicie wyraz „Internet", nim stał się (bodą] najpopularniejszą) nazwą własną, oznaczał po prostu współpracę pomiędzy sieciami ujętą, trzeba przyznać, w charakterystyczną dla anglosasów lakoniczną formę słowną(przyp. red.)

W odróżnieniu od modelu OSI, model TCP/IP bardziej koncentruje się na zapewnianiu przyłączalności niż na sztywnym przywiązaniu do warstw funkcjonalnych. Uznając znaczenie hierarchicznego uporządkowania funkcji, jednocześnie zostawia projektantom protokołu sporą elastyczność odnośnie implementacji. W konsekwencji model OSI znacznie lepiej wyjaśnia mechanikę komunikacji między komputerami, ale TCP/IP stał się protokołem współdziałania międzysieciowego preferowanym przez rynek.

Elastyczność modelu referencyjnego TCP/IP, w porównaniu z modelem OSI, przedstawia rysunek 12.3.

Rysunek 12.3. Porównanie modeli referencyjnych OSI i TCP/IP.

Nazwa warstwy modelu referencyjnego OSI	Numer warstwy OSI	Nazwa równoważnej warstwy TCP/IP
Aplikacji	7	Procesu/ aplikacji
Prezentacji	6	
Sesji	5	Hosta z hostem
Transportu	4	
Sieci	3	Internetu
Łącza danych	2	Dostępu do sieci
Fizyczna	1	

Model referencyjny TCP/IP, opracowany długo po tym, jak powstały protokoły, które opisuje, oferuje dużo większą elastyczność niż model OSI, gdyż wyraża raczej hierarchiczne uporządkowanie funkcji, a nie ich ścisłą strukturę warstwową.

1.12.2.1 **Analiza** TCP/IP

Stos protokołów TCP/IP zawiera cztery warstwy funkcjonalne: dostępu do sieci, Internetu, warstwę host-z-hostem i warstwę procesu/aplikacji. Te cztery warstwy luźno nawiązują do siedmiu warstw modelu referencyjnego OSI, nie tracąc na funkcjonalności.

1.12.2.2 Warstwa procesu/aplikacji

Warstwa aplikacji dostarcza protokoły zdalnego dostępu i współdzielenia zasobów. Znane aplikacje, jak Telnet, FTP, SMTP, HTTP i wiele innych, znajdują się i działają w tej warstwie i są uzależnione od funkcjonalności niższych warstw.

Warstwa „host-z-hostem"

Warstwa „host-z-hostem" protokołu IP luźno nawiązuje do warstw sesji i transportu modelu OSI. Obejmuje dwa protokoły: protokół sterowania transmisją (ang. TCP Transmission Control Protocol) i protokół datagramów użytkownika (ang. UDP - User Datagram Protocol). Obecnie, w celu dostosowania do coraz bardziej zorientowanego

na transakcje charakteru Internetu, definiowany jest trzeci protokół. Protokół ten nosi próbną nazwę protokołu sterowania transmisją i transakcją (ang. T/TCP - Transaction Transmission Control Protocol).

Protokół TCP zapewnia połączeniową transmisję danych pomiędzy dwoma lub więcej hostami, może obsługiwać wiele strumieni danych, umożliwia sterowanie strumieniem danych, kontrolę błędów, a nawet ponowne porządkowanie pakietów, otrzymanych w niewłaściwej kolejności.

Nagłówek protokołu TCP ma długość co najmniej 20 oktetów i zawiera następujące pola: • Port Źródłowy TCP: 16-bitowe pole portu źródłowego przechowuje numer

portu, który inicjuje sesje komunikacyjne. Port źródłowy i adres źródłowy IP funkcjonują jako adres zwrotny pakietu.

• Port Docelowy TCP: 16-bitowe pole portu docelowego jest adresem portu, dla którego przeznaczona jest transmisja. Port ten zawiera adres interfejsu aplikacji w komputerze odbiorcy, do której przesyłany jest pakiet danych.

• Numer Sekwencji TCP: 32-bitowy numer sekwencji jest wykorzystywany przez komputer odbierający do zrekonstruowania rozproszonych, rozbitych, podzielonych danych i przywrócenia im pierwotnej postaci. W sieci dynamicznie trasowanej może się zdarzyć, że

niektóre pakiety pójdą innymi trasami i dotrą w niewłaściwej kolejności. Pole numeru sekwencji kompensuje tę niekonsekwencję w dostarczaniu.
• Numer Potwierdzenia TCP: TCP używa 32-bitowego potwierdzenia (ACK) pierwszego oktetu danych zawartego w następnym oczekiwanym segmencie. TCP może obliczyć ten numer, zwiększając numer ostatniego otrzymanego oktetu o liczbę oktetów w każdym segmencie TCP. Numer używany do identyfikowania każdego ACK jest numerem sekwencji potwierdzanego pakietu.
• Wyrównanie Danych: 4-bitowe pole przechowujące rozmiar nagłówka TCP, którego miarą jest 32-bitowa struktura danych, znana jako „skowo".
• Zarezerwowane: 6-bitowe pole, zawsze ustawione na zero. Pole to jest zarezerwowane dla jeszcze nie wyspecyfikowanego, przyszłego zastosowania.
• Flagi: 6-bitowe pole flagi zawiera sześć 1-bitowych flag, które umożliwiają realizację funkcji sterowania, takich jak pole pilność, potwierdzenie pola znaczącego, pchanie, zerowanie połączenia, synchronizacja numerów sekwencyjnych i zakończenie wysyłania danych.
• Rozmiar Okna: 16-bitowe pole używane przez komputer docelowy w celu poinformowania komputera źródłowego o tym, ile danych jest gotów przyjąć w jednym segmencie TCP.
• Suma Kontrolna (16-bitów): Nagłówek TCP zawiera również pole kontroli błędów, znane jako „suma kontrolna". Komputer źródłowy oblicza wartość tego pola na podstawie zawartości segmentu. Komputer docelowy przeprowadza identyczne obliczenie. Jeśli zawartość pozostała nienaruszona, wynik obydwu obliczeń będzie identyczny, co świadczy o prawidłowości danych.
• Wypełnienie: do tego pola dodawane są zera, tak aby długość nagłówka TCP była zawsze wielokrotnością 32 bitów.
Protokół datagramów użytkownika (UDP) jest innym protokołem IP warstwy host-z-hostem (odpowiadającej warstwie transportu modelu OSI).Protokół UDP zapewnia proste i mające niewielki narzut transmisje danych, tzw. „datagramy". Prostota datagramów czyni UDP protokołem nieodpowiednim dla niektórych aplikacji, za to doskonałym dla aplikacji bardziej wyszukanych, które same mogą zapewnić funkcjonalność połączeniową.
Protokół UDP może być wykorzystywany przy wymienianiu takich danych, jak nadane nazwy NetBIOS, komunikaty systemowe itd., gdyż wymiany te nie wymagają sterowania strumieniem danych, potwierdzeń, ponownego uporządkowywania ani innych funkcji dostarczanych przez protokół TCP.
Nagłówek protokołu UDP ma następującą strukturę:
• Numer portu źródłowego UDP: Port źródłowy jest numerem połączenia w komputerze źródłowym. Port źródłowy i adres źródłowy IP funkcjonują jako adres zwrotny pakietu.
• Numer portu docelowego UDP: Port docelowy jest numerem połączenia w komputerze docelowym. Port docelowy UDP jest wykorzystywany do przekazywania pakietu odpowiedniej aplikacji, po tym jak pakiet dotrze do komputera docelowego.
• Suma kontrolna UDP: Suma kontrolna jest polem kontroli błędów, którego wartość jest obliczana na podstawie zawartości segmentu. Komputer docelowy wykonuje taką samą funkcję matematyczną jak komputer źródłowy. Niezgodność dwóch obliczonych wartości wskazuje na wystąpienie błędu podczas transmisji pakietu.
• Długość komunikatu UDP: Pole długości komunikatu informuje komputer docelowy o jego rozmiarze. Daje to komputerowi docelowemu kolejny mechanizm, wykorzystywany do sprawdzania poprawności wiadomości.
Główną różnicą funkcjonalną pomiędzy TCP a UDP jest niezawodność. Protokół TCP charakteryzuje się wysoką niezawodnością, natomiast UDP jest prostym mechanizmem dostarczania datagramów. Ta fundamentalna różnica skutkuje ogromnie zróżnicowaniem zastosowań tych dwóch protokołów warstwy host-z-hostem.
Warstwa Internetu
Warstwa Internetu protokołu IPv4 obejmuje wszystkie protokoły i procedury potrzebne do przesyłania danych pomiędzy hostami w wielu sieciach. Pakiety przenoszące dane muszą być trasowalne. Odpowiada za to protokół Internetu (IP).
Nagłówek protokołu IP ma następujący rozmiar i strukturę:
• Wersja: Pierwsze cztery bity nagłówka IP identyfikują wersję operacyjną protokołu IP, np. wersję 4.
• Długość Nagłówka Internetu: Następne cztery bity nagłówka zawierają jego długość, wyrażoną w wielokrotnościach liczby 32.
• Rodzaj Usługi: Następne 8 bitów to 1-bitowe flagi, które mogą być używane do określania parametrów pierwszeństwa, opóźnienia, przepustowości i niezawodności tego pakietu danych.
• Długość Całkowita: 16-bitowe pole przechowujące całkowitą długość datagramu IP, mierzoną w oktetach. Prawidłowe wartości mogą mieścić się w przedziale od 576 do 65536 oktetów.
• Identyfikator: Każdemu pakietowi IP nadaje się unikatowy, 16-bitowy identyfikator.
• Flagi: Następne pole zawiera trzy 1-bitowe flagi, wskazujące, czy dozwolona jest fragmentacja pakietów i czy jest ona stosowana.
• Przesunięcie Fragmentu: 8-bitowe pole mierzące przesunięcie fragmentowanej zawartości względem początku całego datagramu. Wartość ta jest mierzona za pomocą 64-bitowych przyrostów.
• Czas Życia (ang. TTL - Time to Live): Pakiet IP nie może „włóczyć się" w nieskończoność po sieci WAN. Musi mieć ograniczoną liczbę skoków, które może wykonać (patrz niżej). Wartość 8-bitowego pola TTL jest zwiększana o jeden przy każdym skoku, jaki pakiet wykonuje. Gdy osiągnie wartość maksymalną, pakiet jest niszczony.
Pakiety IP są trasowane przez różne sieci za pomocą urządzeń znanych jako routery. Każdy router, przez który przechodzi pakiet, jest liczony jako jeden skok. Ustalenie maksymalnej liczby skoków zapewnia, że pakiety nie będą stale wykonywać pętli w dynamicznie trasowanej sieci.
• Protokół: 8-bitowe pole identyfikujące protokół, następujący po nagłówku IP, taki jak VINES, TCP, UDP itd.
• Suma Kontrolna: 16-bitowe pole kontroli błędów. Komputer docelowy lub jakikolwiek inny węzeł bramy w sieci, może powtórzyć działania matematyczne na zawartości pakietu, przeprowadzone wcześniej przez komputer źródłowy. Jeśli dane po drodze nie uległy zmianie, wyniki obydwu obliczeń są identyczne. Pole sumy kontrolnej informuje również komputer docelowy o ilości przychodzących danych.
• Adres Źródłowy IP: jest adresem IP komputera źródłowego. • Adres Docelowy IP: jest adresem IP komputera docelowego.

• Wypełnienie: do tego pola dodawane są zera. tak aby długość nagłówka TCP byka zawsze wielokrotnością 32 bitów.
Te pola nagłówka świadczą o tym, że protokół IPv4 warstwy Internetu jest protokołem bezpołączeniowym - urządzenia kierujące pakietem w sieci mogą samodzielnie ustalać idealną ścieżkę przejścia przez sieć dla każdego pakietu. Nie występują również żadne potwierdzenia, sterowanie strumieniem danych czy też funkcje porządkowania kolejności, właściwe protokołom wyższych warstw, takim jak TCP. Protokół IPv4 pozostawia te funkcje protokołom wyższego poziomu.
Warstwa Internetu musi także obsługiwać inne funkcje zarządzania trasą oprócz formatowania pakietów IP. Musi zapewnić mechanizmy tłumaczące adresy warstwy 2 na adresy warstwy 3 i odwrotnie. Te funkcje zarządzania trasą są dostarczane przez protokoły równorzędne z IP; protokoły trasujące opisane w rozdziale 1 pt. „ABC sieci". Są to: wewnętrzny protokół bramowy (ang. *IGP - Interior Gateway Protocols),* zewnętrzne protokoły bramowe (ang. *EGP - Exterior Gateway Protocols),* protokół rozróżniania adresów (ang. *ARP - Address Resolution Protocol),* odwrócony protokół rozróżniania adresów (ang. *RARP - Reverse Address Resolution Protocol)* i protokół komunikacyjny sterowania Internetem (ang. *ICMP - Internet Control Message Protocol).*

1.12.2.3 Typowe działanie protokołu IPv4

Warstwa aplikacji opatruje pakiet danych nagłówkiem, identyfikując docelowy host i port. Protokół warstwy host-z-hostem (TCP lub UDP, w zależności od aplikacji) dzieli ten blok danych na mniejsze, łatwiej dające sobą kierować kawałki. Do każdego kawałka dołączony jest nagłówek. Taką strukturę nazywa się „segmentem TCP".
Pola nagłówka segmentu są odpowiednio wypełniane, a segment jest przekazywany do warstwy Internetu. Warstwa Internetu dodaje informacje dotyczące adresowania, rodzaju protokołu (TCP lub UDP) i sumy kontrolnej. Jeśli segment był fragmentowany, warstwa Internetu wypełnia również to pole.
Komputer docelowy odwraca właśnie opisane działania. Odbiera pakiety i przekazuje je swojemu protokołowi warstwy host-z-hostem do ponownego złożenia. Jeśli to konieczne, pakiety są ponownie grupowane w segmenty danych, przekazywane odpowiedniej aplikacji.

1.12.2.4 Schemat adresowania protokołu IP

Protokół IPv4 wykorzystuje 32-bitowy, binarny schemat adresowania, w celu identyfikowania sieci, urządzeń sieciowych i komputerów przyłączonych do sieci. Adresy te, znane jako adresy IP, są ściśle regulowane przez internetowe centrum informacji sieciowej (ang. InterNIC - Internet Network Information Center). Choć administrator sieci ma możliwość dowolnego wybierania nie zarejestrowanych adresów IP, taka praktyka jest niewybaczalna. Komputery mające takie „podrobione" adresy IP mogą działać prawidłowo tylko w obrębie swej własnej domeny. Próby dostępu do Internetu z pewnością wykażą ograniczenia takiego krótkowzrocznego działania. Skutki mogą być bardzo różne w zależności od wielu rozmaitych czynników, ale na pewno będą to skutki niepożądane.
Każda z pięciu klas adresów IP jest oznaczona literą alfabetu: klasa A, B, C, D i E. Każdy adres składa się z dwóch części: adresu sieci i adresu hosta. Klasy prezentują odmienne uzgodnienia dotyczące liczby obsługiwanych sieci i hostów. Choć są to adresy binarne, zwykle przedstawia się je w tzw. formacie dziesiętnym kropkowym (np. 135.65.121.6), aby ułatwić człowiekowi ich używanie. Kropki rozdzielają cztery oktety adresu.
Notacja dziesiętna kropkowa odnosi się do konwersji adresu binarnego na dziesiętny system liczbowy. Kropka („.") służy do oddzielania numerów węzła i sieci. Na przykład, 100.99 odnosi się do urządzenia 99 w sieci 100.
• Adres IP klasy A: Pierwszy bit adresu klasy A jest zawsze ustawiony na „0". Następne siedem bitów identyfikuje numer sieci. Ostatnie 24 bity (np. trzy liczby dziesiętne oddzielone kropkami) adresu klasy A reprezentują możliwe adresy hostów. Adresy klasy A mogą mieścić się w zakresie od 1.0.0.0 do 126.0.0.0. Każdy adres klasy A może obsłużyć 16777214 (=22`24-2) unikatowych adresów hostów.
• Adres IP klasy B: Pierwsze dwa bity adresu klasy B to „10". Następne 16 bitów identyfikuje numer sieci, zaś ostatnie 16 bitów identyfikuje adresy potencjalnych hostów. Adresy klasy B mogą mieścić się w zakresie od 128.1.0.0 do 191.254.0.0. Każdy adres klasy B może obsłużyć 65534 (=2'16-2) unikatowych adresów hostów.
• Adres IP klasy C: Pierwsze trzy bity adresu klasy C to „110". Następne 21 bitów identyfikuje numer sieci. Ostatni oktet służy do adresowania hostów. Adresy klasy C mogą mieścić się w zakresie od 192.0.1.0 do 223.255.254.0. Każdy adres klasy C może obsłużyć 254 (=28-2) unikatowe adresy hostów.
• Adres IP klasy D: Pierwsze cztery bity adresu klasy D to „I 110". Adresy te są wykorzystywane do multicastingu , ale ich zastosowanie jest ograniczone. Adres multicast jest unikatowym adresem sieci, kierującym pakiety do predefiniowanych grup adresów IP. Adresy klasy D mogą pochodzić z zakresu 224.0.0.0 do 239.255.255.254.
Pewna niejasność definicji klasy D adresu IP przyczynia się do potencjalnej rozbieżności pomiędzy jej rozumieniem a stanem faktycznym. Choć IETF zdefiniowało klasy C i D jako oddzielne, różniące się pod względem zakresów liczbowych i zamierzonej funkcjonalności, to wcale nie tak rzadko zdarza się, że zakres adresu klasy D jest utożsamiany z zakresem adresu klasy C. Jest to podejście nieprawidłowe - ale najwidoczniej narzucane przez pewne kursy certyfikacyjne.
• Adres IP klasy E: Faktycznie - zdefiniowano klasę E adresu IP, ale InterNIC zarezerwował go dla własnych badań. Tak więc żadne adresy klasy E nie zostały dopuszczone do zastosowania w Internecie.
Duże odstępy między tymi klasami adresów marnowały znaczną liczbę potencjalnych adresów. Rozważmy dla przykładu średnich rozmiarów przedsiębiorstwo, które potrzebuje 300 adresów IP. Adres klasy C (254 adresy) jest niewystarczający. Wykorzystanie dwóch adresów klasy C dostarczy więcej adresów niż potrzeba, ale w wyniku tego w ramach przedsiębiorstwa powstaną dwie odrębne domeny. Z kolei zastosowanie adresu klasy B zapewni potrzebne adresy w ramach jednej domeny, ale zmarnuje się w ten sposób 65534 - 300 = 65234 adresy.
Na szczęście nie będzie to już dłużej stanowić problemu. Został opracowany nowy, międzydomenowy protokół trasujący, znany jako bezklasowe trasowanie międzydomenowe (ang. CIDR - Classless Interdomain Routing), umożliwiający wielu mniejszym klasom adresowym działanie w ramach jednej domeny trasowania.

Adresowanie IP wymaga, by każdy komputer miał własny, unikalny adres. Maski podsieci mogą kompensować ogromne odstępy między klasami adresowymi, dostosowując długość adresów hosta i/lub sieci. Za pomocą tych dwóch adresów można trasować dowolny datagram IP do miejsca przeznaczenia.
Ponieważ TCP/IP jest w stanie obsługiwać wiele sesji z pojedynczego hosta, musi on zapewnić możliwość adresowania specyficznych programów komunikacyjnych, które mogą działać na każdym z hostów. TCP/IP wykorzystuje do tego numery portów. IETF przypisało kilku najbardziej powszechnym aplikacjom ich własne, dobrze znane numery portów. Numery te są stałe dla każdej aplikacji na określonym hoście. Innym aplikacjom przypisuje się po prostu dostępny numer portu.

1.12.2.5 Wnioski dotyczące IPv4

Protokół IPv4 ma już prawie dwadzieścia lat. Od jego początków Internet przeszedł kilka znaczących zmian, które zmniejszyły efektywność IP jako protokołu uniwersalnej przyłączalności. Być może najbardziej znaczącą z tych zmian była komercjalizacja Internetu. Przyniosła ona bezprecedensowy wzrost populacji użytkowników Internetu. To z kolei stworzyło zapotrzebowanie na większą liczbę adresów, a także potrzebę obsługi przez warstwę Internetu nowych rodzajów usług. Ograniczenia IPv4 stały się bodźcem dla opracowania zupełnie nowej wersji protokołu. Jest ona nazywana IP, wersja 6 (IPv6), ale powszechnie używa się również nazwy Następna generacja protokole Internetu (ang. IPng- next generation of Internet Protocol).

1.12.3 Protokół Internetu, wersja 6 (IPv6)

Protokół IPv6 ma być prostą, kompatybilną „w przód" nowelizacją istniejącej wersji protokołu IP. Intencją przyświecającą tej nowelizacji jest wyeliminowanie wszystkich słabości ujawniających się obecnie w protokole IPv4, w tym zbyt małej liczby dostępnych adresów IP, niemożności obsługiwania ruchu o wysokich wymaganiach czasowych i braku bezpieczeństwa w warstwie sieci.
Protokół IPv6 był pierwotnie określany jako „IP: następna generacja" lub „Ipng" - co przydawało mu nieco tajemniczości z pogranicza science fiction. Podczas opracowywania specyfikacji protokół ten otrzymał oficjalną nazwę „IP wersja 6" (IPv6).
Dodatkowym bodźcem dla opracowania i rozwoju nowego protokołu IP stało się trasowanie, które w ramach protokołu IPv4 jest skrępowane jego 32-bitową architekturą adresowądwupoziomową hierarchią adresowania i klasami adresowymi. Dwupoziomowa hierarchia
adresowania „host.domena" po prostu nie pozwala konstruować wydajnych hierarchii adresowych, które mogłyby być agregowane w routerach na skalę odpowiadającą dzisiejszym wymaganiom globalnego Internetu.
Następna generacja protokołu IP - IPv6 - rozwiązuje wszystkie wymienione problemy. Będzie oferować znacznie rozszerzony schemat adresowania, aby nadążyć za stałą ekspansją Internetu, a także zwiększoną zdolność agregowania tras na wielką skalę.
IPv6 będzie także obsługiwać wiele innych właściwości, takich jak: transmisje audio i/lub wideo w czasie rzeczywistym, mobilność hostów, bezpieczeństwo końcowe (czyli na całej długości połączenia) dzięki mechanizmom warstwy Internetu - kodowaniu i identyfikacji, a także autokonfiguracja i autorekonfiguracja. Oczekuje się, że usługi te będą odpowiednią zachętą dla migracji, gdy tylko staną się dostępne produkty zgodne z IPv6. Wiele z tych rozwiązań wciąż wymaga dodatkowej standaryzacji, dlatego też przedwczesne byłoby ich obszerne omawianie.
Jedynym jednakże aspektem protokołu IPv6, który wymaga szerszego omówienia, jest adresowanie. 32-bitowa długość adresu w protokole IPv4 teoretycznie umożliwiała zaadresowanie około 4 miliardów (2"-1) urządzeń. Niewydajne podsieciowe techniki maskowania i inne rozrzutne praktyki roztrwoniły niestety ów zasób.
Protokół IPv6 wykorzystuje adresy 128-bitowe i teoretycznie jest w stanie zwiększyć przestrzeń adresową protokołu o czynnik 2'G - co daje astronomiczną liczbę
340.282.366.920.938.463.463.374.607.431.768.21 1.456
potencjalnych adresów. Obecnie zajęte jest około 15% tej przestrzeni adresowej -reszta jest zarezerwowana dla - bliżej nie określonych - przyszłych zastosowań.
W rzeczywistości przypisanie i trasowanie adresów wymaga utworzenia ich hierarchii. Hierarchie mogą zmniejszyć liczbę potencjalnych adresów, ale za to zwiększają wydajność protokołów trasujących zgodnych z IPv6. Jedną z praktycznych implikacji długości adresu IPv6 jest to, że usługa nazwy domeny (ang. DNS- Domain Name Service), stanowiąca w wersji IPv4 jedynie wygodny luksus, tutaj staje się absolutną koniecznością.
Usługa nazwy domeny jest narzędziem sieciowym, odpowiedzialnym za tłumaczenie (wygodnych dla użytkowników) mnemonicznych nazw hostów na numeryczne adresy IP.
Równie znacząca, jak zwiększona potencjalna przestrzeń adresowa, jest jeszcze większa elastyczność, na jaką pozwalają nowe struktury adresowe IPv6. Protokół ten uwalnia się od adresowania bazującego na klasach. Zamiast tego rozpoznaje on trzy rodzaje adresów typu unicast, adres klasy D zastępuje nowym formatem adresu multicast oraz wprowadza nowy rodzaj adresu; przed przejściem do dalszej części wykładu nieodzowne staje się wyjaśnienie szczegółów tych koncepcji.

1.12.3.1 Struktury adresów *unicast* IPv6

Adresowanie *unicast* zapewnia przyłączalność od jednego urządzenia końcowego do drugiego. Protokół IPv6 obsługuje kilka odmian adresów unicast.

1.12.3.1.1 Adres dostawcy usług internetowych (ISP)

Podczas gdy protokół IPv4 z góry przyjął grupy użytkowników wymagających przyłączalności. IPv6 dostarcza format adresu unicast, specjalnie przeznaczony dla dostawców usług internetowych, w celu przyłączania indywidualnych użytkowników do Internetu. Te oparte na dostawcach adresy unicast oferują unikatowe adresy dla indywidualnych użytkowników lub małych grup, uzyskujących dostęp do Internetu

za pośrednictwem dostawcy usług internetowych. Architektura adresu zapewnia wydajną agregację tras w środowisku użytkowników indywidualnych.
Format adresu unicast ISP jest następujący:
• 3-bitowa flaga adresu unicast ISP, zawsze ustawiana na „010" • Pole ID rejestru, o długości „n" bitów
• Pole ID dostawcy, o długości „m" bitów • Pole ID abonenta, o długości „o" bitów • Pole ID podsieci, o długości „p" bitów
• Pole ID interfejsu, o długości 128-3- (n+m+o+p) bitów
Litery n,m,o,p oznaczają zmienne długości pól. Długość pola ID interfejsu stanowi różnicę długości adresu (128) i łącznej długości pól poprzedzających, wraz z trójbitową flagą. Przykładem adresu tego typu może być 010:0:0:0:0:x, gdzie „x" może być dowolną liczbą. Ponieważ większość nowej przestrzeni adresowej dopiero musi zostać przypisana, adresy te będą zawierać mnóstwo zer. Dlatego grupy zer mogą być zapisywane skrótem w postaci podwójnego dwukropka (::) - skróconą formą adresu O1O:O:O:O:O:x jest więc 010::x.
Inne rodzaje adresów unicast są przeznaczone do użytku lokalnego. Adresy użytku lokalnego mogą być przypisane do urządzeń sieciowych w samodzielnym Intranecie lub do urządzeń w Intranecie, którym potrzebny jest dostęp do Internetu.
Adres użytku lokalnego dla łącza
Adres użytku lokalnego dla łącza jest przeznaczony dla pojedynczego łącza, do celów takich jak konfiguracja auto-adresu, wykrywanie sąsiadów, a także w przypadku braku routerów. Adresy lokalne dla łącza mają następujący format:
• 10-bitowa flaga adresu lokalnego, zawsze ustawiana na „1 I 111 I 1011"
• Zarezerwowane, nienazwane pole, mające długość „n" bitów, ale ustawiane domyślnie na wartość „0"
• Pole ID interfejsu o długości 1 18 - n bitów
ID interfejsu może być adresem MAC karty sieciowej Ethernetu. Adresy MAC, będące teoretycznie adresami unikalnymi, mogą być skojarzone z przedrostkami standardowego adresu IP w celu utworzenia unikalnych adresów dla mobilnych lub zastępczych użytkowników. Przykładem adresu użytku lokalnego dla łącza z adresem MAC mógłby być: 1111111011:O:adres mac.

1.12.3.1.2 Adres użytku lokalnego dla miejsca

Adresy lokalne dla miejsca są przeznaczone do stosowania w pojedynczym miejscu. Mogą być używane w miejscach lub organizacjach, które nie są przyłączone do globalnego Internetu. Nie muszą żądać czy też „kraść" przedrostka adresu z przestrzeni adresowej globalnego Internetu. Zamiast tego mogą używać adresów protokółu IPv6 lokalnych dla miejsca. Gdy organizacja łączy się z globalnym Internetem, może utworzyć unikatowe adresy globalne, zastępując przedrostek lokalny dla miejsca przedrostkiem abonenta, zawierającym identyfikatory rejestru, dostawcy i abonenta.
Adresy lokalne dla miejsca mają następujący format:
• 10-bitowa flaga użytku lokalnego, zawsze ustawiana na „1111111011"
• Zarezerwowane, nienazwane pole, mające długość „n" bitów, ale ustawiane domyślnie na wartość „0"
• Pole ID podsieci o długości „m" bitów
• Pole ID interfejsu o długości 1 18 - (n+m) bitów
Przykładem adresu lokalnego dla miejsca jest: 1 I 11 l 1101 l:podsieć:interfejs.

1.12.3.2 Struktury zastępczych adresów *unicast* IPv6

Dwa specjalne adresy unicast protokołu IPv6 zostały określone jako mechanizmy przejściowe, umożliwiające hostom i routerom dynamiczne trasowanie pakietów IPv6 przez infrastrukturę sieci protokołu IPv4 i na odwrót.

1.12.3.2.1 Adres unicast *IPv6 zgodny z IPv4*

Pierwszy typ adresu unicast nosi nazwę „adres IPv6 zgodny z IPv4". Ten zastępczy adres unicast może być przypisywany węzłom IPv6, a jego ostatnie 32 bity zawierają adres IPv4. Adresy takie mają następujący format:

1.12.3.2.2 Adres unicast *IPv6 wzorowany na IPv4*

80 bitów	16 bitów	32 bity
000...0000	00...00	adres IPv4

Drugi, podobny typ adresu IPv6, również zawierający adres IPv4 w ostatnich 32 bitach, jest znany jako „adres IPv6 wzorowany na IPv4". Adres ten jest tworzony przez router
o podwójnym protokole i umożliwia węzłom pracującym wyłącznie z protokołem IPv4 tunelowanie przez infrastrukturę sieci z protokołem IPv6. Jedyna różnica między adresami unicast IPv6 wzorowanymi na IPv4 a adresami unicast IPv6 zgodnymi z IPv4 jest taka, że adresy wzorowane na IPv4 to adresy tymczasowe. Są one automatycznie tworzone przez routery o podwójnym protokole i nie mogą być przypisane do żadnego węzła. Format takiego adresu wygląda następująco:
Obydwa adresy unicast, zarówno wzorowany na IPv4, jak i zgodny z IPv4, mają zasadnicze znaczenie dla tunelowania. Tunelowanie umożliwia przesyłanie pakietów przez niedostępny w inny sposób rejon sieci dzięki umieszczaniu pakietów w obramowaniu akceptowalnym na zewnątrz.

80 bitów	16 bitów	32 bity
000...0000	FF...FF	adres IPv4

1.12.3.3 Struktury adresów *anycast* IPv6

Adres anycast, wprowadzony w protokole IPv6, jest pojedynczą wartością przypisaną do więcej niż jednego interfejsu. Zwykle interfejsy te należą do różnych urządzeń. Pakiet wysłany pod adres anycast jest trasowany tylko do jednego urządzenia. Jest on wysyłany do najbliższego - według zdefiniowanej przez protokoły trasujące miary odległości interfejsu o tym adresie. Na przykład, strona WWW (World Wide Web) może być powielona na kilku serwerach. Dzięki przypisaniu tym serwerom adresu anycast żądania połączenia z tą stroną WWW są automatycznie trasowane do tylko jednego serwera najbliższego względem użytkownika.
W środowisku trasowanym „najbliższy" interfejs może nie być tym, który jest najbliżej w sensie fizycznego położenia. Routery wykorzystują przy obliczaniu tras zaskakująco szeroki zestaw metryk. Określanie najkrótszej trasy jest uzależnione od aktualnie używanego protokołu trasującego oraz od jego metryk.
Adresy anycast są tworzone (pobierane) z przestrzeni adresów unicast i mogą przybrać formę dowolnego typu adresu unicast. Tworzy się je, przypisując po prostu ten sam adres unicast więcej niż jednemu interfejsowi.

1.12.3.4 Struktury adresów *multicast* IPv6

Protokół IPv4 obsługiwał multicasting, ale wymagało to stosowania niejasnego adresowania klasy D. Protokół IPv6 rezygnuje z adresów klasy D na korzyść nowego formatu adresu, udostępniającego tryliony możliwych kodów grup multicast. Każdy kod grup identyfikuje dwóch lub więcej odbiorców pakietu. Zakres pojedynczego adresu multicast jest elastyczny. Każdy adres może być ograniczony do pojedynczego systemu. do określonego miejsca, powiązany z danym łączem sieciowym lub rozpowszechniany globalnie.
Należy zauważyć, że nadawanie adresów IP również zostało wyeliminowane i zastąpione nowym multicastingowym formatem adresu.

1.12.3.5 Wnioski dotyczące IPv6

Pomimo potencjalnych korzyści związanych z protokołem IPv6, migracja z IPv4 nie jest wolna od ryzyka. Rozszerzenie długości adresu z 32 do 128 bitów automatycznie ogranicza współoperacyjność protokołów IPv4 i IPv6. Węzły „tylko-IPv4" nie mogą współdziałać z węzłami „tylko-IPv6", ponieważ architektury adresowe nie są kompatybilne w przód. To ryzyko biznesowe, w połączeniu z nieustanną ewolucją protokołu IPv4, może stanowić przeszkodę dla rynkowej akceptacji protokołu IPv6.

1.12.4 Wymiana IPX/SPX Novell

Zestaw protokołów firmy Novell bierze nazwę od swoich dwóch głównych protokołów: międzysieciowej wymiany pakietów (ang. IPX-Internet Packet Exchange) i sekwencyjnej wymiany pakietów (ang. SPX-Sequenced Packet Exchange). Ten firmowy stos protokołów został oparty na protokole systemów sieciowych firmy Xerox (ang. XNS-Xerox's Network System), wykorzystywanym w pierwszej generacji sieci Ethernet. Wymiana IPX/SPX zyskała na znaczeniu we wczesnych latach 80. jako integralna część systemu Novell Netware. Netware stał się faktycznym standardem sieciowego systemu operacyjnego (ang. NOS - Network Operating System) dla sieci lokalnych pierwszej generacji. Novell uzupełnił swój system zestawem aplikacji biznesowych i klienckich narzędzi łączności.
Protokół IPX w dużym stopniu przypomina IP. Jest bezpołączeniowym protokołem datagramowym, który nie wymaga ani nie zapewnia potwierdzenia każdego transmitowanego pakietu. Protokół IPX polega na SPX w taki sam sposób, w jaki protokół IP polega na TCP w zakresie porządkowania kolejności i innych usług połączeniowych warstwy 4. Rysunek 12.4 przedstawia stos protokołów IPX/SPX w porównaniu z modelem referencyjnym OSI.
Protokoły IPX i SPX Novella są funkcjonalnym ekwiwalentem warstw modelu OSI, odpowiednio warstw 3 i 4. Pełny zestaw protokołów IPX/SPX, składający się z czterech warstw, funkcjonalnie odpowiada innym warstwom modelu OSI.

1.12.4.1 Analiza IPX/SPX

Stos protokołów IPX/SPX obejmuje cztery warstwy funkcjonalne: dostępu do nośnika, łącza danych, Internetu i aplikacji. Te cztery warstwy luźno nawiązują do siedmiu warstw modelu referencyjnego OSI, nie tracąc nic na funkcjonalności.
Rysunek 12.4. Ponównanie modelu OSI z modelem IP,Y'SPX
Warstwa aplikacji

<table>
<tr><th>Nazwa warstwy modelu referencyjnego OSI</th><th>Numer warstwy OSI</th><th colspan="5">Opis równoważnej warstwy IPX/SPX</th></tr>
<tr><td>Aplikacji</td><td>7</td><td rowspan="3">RIP</td><td rowspan="3">SAP</td><td rowspan="3">NCP</td><td rowspan="3">NLSP</td><td rowspan="5">Rozmaite protokoły</td></tr>
<tr><td>Prezentacji</td><td>6</td></tr>
<tr><td>Sesji</td><td>5</td></tr>
<tr><td>Transportu</td><td>4</td><td colspan="4">SPX</td></tr>
<tr><td>Sieci</td><td>3</td><td colspan="4">Międzysieciowa wymiana pakietów</td></tr>
<tr><td>Łącza danych</td><td>2</td><td colspan="5">Interfejs otwartego łącza danych</td></tr>
<tr><td>Fizyczna</td><td>1</td><td colspan="5">Dostęp do nośnika</td></tr>
</table>

Warstwa aplikacji Novella obejmuje trzy warstwy - aplikacji, prezentacji i sesji - modelu OSI, choć niektóre z jej protokołów aplikacyjnych rozciągają ten stos w dół, aż do warstwy sieci. Głównym protokołem warstwy aplikacji w tym stosie jest protokół rdzenia NetWare (ang. NCP - NetWare Core Protocol. Protokół NCP można bezpośrednio sprzęgać zarówno z protokołem SPX, jak i IPX. Jest wykorzystywany do drukowania, współdzielenia plików, poczty elektronicznej i dostępu do katalogów.

Innymi protokołami warstwy aplikacji są między innymi: protokół informacyjny trasowania (ang. *RIP - Routing Information Protocol,* firmowy protokół ogłoszeniowy usługi (ang. *SAP - Service Advertisement Protocol* i protokół obsługi łącza systemu Netware (ang. *NLSP - Netware Link Services Protocol.*

Protokół RIP jest domyślnym protokołem trasującym systemu NetWare. Jest to protokół trasowania wektora odległości wykorzystujący tylko dwie metryki: kwanty (ang. ticks) i skoki (ang. hops). Kwant jest miarą czasu, zaś liczba skoków, jak już wyjaśniono wcześniej w tym rozdziale, jest licznikiem routerów, które manipulowały trasowanym pakietem. Na tych dwóch metrykach opiera się wybór ścieżki trasowania protokołu IPX. Podstawową metryką są kwanty - skoki rozstrzygają tylko w przypadku, gdy dwie ścieżki (lub więcej) mają taką samą wartość znaków kontrolnych.

RIP jest bardzo prostym protokołem trasującym. Oprócz ograniczonej liczby metryk wektora odległości, cechuje się też wysokim poziomem narzutu sieciowego. Narzut ten powstaje, ponieważ aktualizacje tabeli trasującej RIP są nadawane co 60 sekund. W wielkich lub mocno obciążonych sieciach taka szybkość aktualizacji może mieć szkodliwe działanie.

SAP jest unikatowym protokołem firmowym, który Novell udanie zastosował do polepszenia związku klienta z serwerem. Serwery wykorzystują protokół SAP do automatycznego wysyłania w sieć informacji o udostępnianych przez nie usługach natychmiast po tym. jak uaktywnią się w sieci. Oprócz tego okresowo nadają informacje SAP, aby dostarczać klientom i innym serwerom informacje o swoim statusie i usługach.

Transmisje SAP generowane przez serwer informują o statusie i usługach tego serwera. Transmisje zawierają nazwę i typ serwera, jego status operacyjny, a także numery sieci, węzła i gniazda. Routery mogą przechowywać informacje z transmisji SAP i rozprowadzać je do innych segmentów sieci. Klienci także mogą inicjować zgłoszenie SAP, gdy potrzebują określonej usługi. Ich żądanie jest rozsyłane po całym segmencie sieci. Hosty mogą wtedy odpowiedzieć i dostarczyć klientowi informacje SAP wystarczające do określenia, czy usługa jest dostępna w rozsądnej odległości.

Niestety, SAP jest dojrzałym protokołem, który coraz gorzej funkcjonuje we współczesnych sieciach. Tak jak w przypadku protokołu RIP, ogłoszenia o usługach są nadawane co 60 sekund. Przy dzisiejszych ogromnych, jednorodnych, komutowanych sieciach LAN, taka częstość nadawania może być problematyczna.

Najnowszym protokołem warstwy aplikacji jest protokół obsługi łącza systemu Netware (NLSP). Jest to protokół trasowania w zależności od stanu łącza, którym Novell zamierza zastąpić starzejące się protokoły RIP i SAP. Protokół NLSP aktualizuje trasy tylko wtedy, gdy zaszły jakieś zmiany.

1.12.4.2 Protokoły warstwy Internetu

Warstwa Internetu wymiany IPX/SPX luźno nawiązuje do warstw sieci i transportu modelu referencyjnego OSI. IPX jest w przeważającej części protokołem warstwy 3 (sieci), choć może też być bezpośrednio sprzęgany z warstwą aplikacji. SPX jest wyłącznie protokołem warstwy 4 (transportu) i nie może być bezpośrednio sprzęgnięty z interfejsem ODI warstwy łącza danych. Musi przekazywać dane poprzez protokół IPX sprzęgnięty z ODI. IPX i SPX funkcjonują jako protokoły podwarstw we wspólnej warstwie Internetu.

SPX jest protokołem połączeniowym i może być wykorzystywany do przesyłania danych między klientem serwerem, dwoma serwerami czy nawet dwoma klientami. Tak jak w przypadku protokołu TCP, protokół SPX zapewnia niezawodność transmisjom IPX, zarządzając (administrując) połączeniem i udostępniając sterowanie strumieniem danych, kontrolę błędów i porządkowanie kolejności pakietów.

Nagłówek SPX ma następujący rozmiar i strukturę:

• Sterowanie połączeniem: Pierwszy oktet (8 bitów) nagłówka SPX zawiera cztery 2-bitowe flagi, sterujące dwukierunkowym przepływem danych przez połączenie SPX.

• Typ strumienia danych: Następnych osiem bitów nagłówka definiuje typ strumienia danych.

• Identyfikacja połączenia źródłowego: 16-bitowe pole identyfikacji połączenia źródłowego identyfikuje proces odpowiedzialny za inicjowanie połączenia.

• Identyfikacja połączenia docelowego: 16-bitowe pole identyfikacji połączenia docelowego służy do identyfikowania procesu, który zaakceptował żądanie (zgłoszenie) połączenia SPX.
• Numer sekwencji: 16-bitowe pole numeru sekwencji dostarcza protokołowi SPX hosta docelowego informację o liczbie wysłanych pakietów. To sekwencyjne numerowanie może być wykorzystywane do zmiany kolejności odebranych pakietów, gdyby przybyły w niewłaściwej kolejności.
• Numer potwierdzenia: 16-bitowe pole numeru potwierdzenia wskazuje następny oczekiwany segment.
• Liczba alokacji: 16-bitowe pole liczby alokacji jest wykorzystywane do śledzenia liczby pakietów wysłanych, ale nie potwierdzonych przez odbiorcę.
• Dane: Ostatnie pole nagłówka SPX zawiera dane. W jednym pakiecie SPX można przesłać do 534 oktetów danych.
Protokołem warstwy sieci dla sieci Novell jest IPX. Protokół ten zapewnia bezpołączeniowe usługi dostarczania datagramów. Przygotowuje pakiety protokołu SPX (lub pakiety innych protokołów) do dostarczenia przez wiele sieci, dołączając do nich nagłówek IPX. W ten sposób powstaje struktura zwana datagramem IPX. Nagłówek tego datagramu zawiera wszystkie informacje niezbędne do skierowania pakietów do miejsca przeznaczenia, niezależnie od tego, gdzie mogłoby się ono znajdować.
Długość nagłówka IPX wynosi 1 1 oktetów. Jego struktura obejmuje następujące pola: • Suma Kontrolna: Nagłówek IPX zaczyna się od 16-bitowego pola dziedziczenia, które istnieje tylko po to, aby zapewnić kompatybilność wsteczną z protokołem XNS. Protokół XNS wykorzystywał to pole do kontrolowania błędów, ale IPX domyślnie ustawia to pole na „FFFFH", a wykrywanie (i korekcję) błędów transmisji pozostawia protokołom wyższego poziomu.
• Długość Pakietu: 16-bitowe pole określające długość datagramu IPX, wliczając nagłówek i dane. Pole to jest sprawdzane w celu weryfikacji integralności pakietu. • Sterowanie Transportem: 8-bitowe pole wykorzystywane przez routery pod
czas przesyłania datagramu. Przed wysłaniem IPX ustawia to pole na „0". Każdy router, który odbiera i przesyła dalej datagram, zwiększa wartość pola o jeden. • Typ Pakietu: 8-bitowe pole identyfikujące typ pakietu zawartego w datagramie
IPX. Pole to umożliwia hostowi docelowemu przekazanie zawartości do następnej, odpowiedniej warstwy protokołów. Typy mogą obejmować RIP, NCP, SPX, błąd itd.
• Numer Sieci Docelowej: 32-bitowe pole określające numer sieci, w której znajduje się węzeł docelowy.
• Węzeł Docelowy: 48-bitowe pole zawierające numer węzła, w którym znajduje się komputer docelowy.
• Numer Gniazda Docelowego: Ponieważ IPX umożliwia wiele jednoczesnych połączeń z jednym systemem, istotne jest określenie numeru gniazda procesu lub programu odbierającego pakiety. Informacji takiej dostarcza to 16-bitowe pole.
• Numer Sieci Źródłowej: 32-bitowe pole określające numer sieci, w której znajduje się węzeł źródłowy.
• Adres Węzła Źródłowego: 48-bitowe pole zawierające numer węzła, w którym znajduje się komputer źródłowy.
• Numer Gniazda Źródłowego: 16-bitowe pole, określające numer gniazda procesu lub programu wysyłającego pakiety.

1.12.4.3 Typowe działanie protokołów IPX/SPX

Protokół SPX tworzy i utrzymuje połączeniowy strumień bitów między dwoma przyłączonymi do sieci urządzeniami. Protokół przyjmuje duże bloki danych z protokołów wyższych warstw i dzieli je na łatwiejsze w kierowaniu kawałki, nie przekraczające długości 534 oktetów. Do danych dołączany jest nagłówek SPX i w ten sposób powstają segmenty danych SPX. Segmenty przekazywane są protokołowi warstwy Internetu, czyli protokołowi IPX. IPX umieszcza segmenty w polu danych swoich pakietów i wypełnia wszystkie pola nagłówka IPX.
Pola nagłówka IPX obejmują adresowanie sieci, długość, sumę kontrolną i inne informacje nagłówkowe. Następnie pakiet przekazywany jest warstwie łącza danych.
Rysunek 12.5 pokazuje umiejscowienie nagłówków IPX i SPX w ramce Ethernet 802.3. Jest to struktura używana do przekazywania danych pomiędzy dwiema podwarstwami warstwy Internetu sieci Novell.
Rysunek 12.5. Struktura ramki Ethernet 82.3. =cnrierającej nagłówki IPX/SPX.

7-oktetowa Preambuła	1-oktetowy Ogranicznik początku ramki	6-oktetowy Adres odbiorcy	6-oktetowy Adres nadawcy	2-oktetowe pole Długość	30-oktetowy nagłówek IPX	Nagłówek SPX o zmiennej długości	Pole Dane o zmiennej długości od 46 do 1482 oktetów	4-oktetowa Sekwencja kontrolna ramki

Komputer docelowy odwraca opisane wyżej działania. Odbiera pakiety i przekazuje je własnemu protokołowi SPX do ponownego złożenia. Jeśli to konieczne, pakiety są ponownie grupowane w segmenty danych, przekazywane odpowiedniej aplikacji.

1.12.4.4 Warstwy łącza danych i dostępu do nośnika

W systemie Netware odpowiednikami warstw fizycznej i łącza danych OSI są warstwy dostępu do nośnika i łącza danych. Warstwa łącza danych jest bezpośrednio kompatybilna ze standardem interfejsu otwartego łącza danych (ODI). Podobnie warstwa dostępu do nośnika jest bezpośrednio kompatybilna ze wszystkimi popularnymi, znormalizowanymi protokołami dostępu do nośnika.
Ta niskopoziomowa zgodność z przemysłowymi standardami otwartymi sprawia, że system Netware ze stosem protokołów IPX/SPX może być implementowany niemal wszędzie.

1.12.4.5 **Adresowanie** IPX

Adresy IPX mają długość 10 oktetów (80 bitów). Jest to znacznie więcej niż 32 bity adresu IPv4, ale mniej niż 128 bitów adresu IPv6. Każdy adres składa się z dwóch części składowych: numeru sieci o maksymalnej długości 32 bitów oraz 48-bitowego numeru węzła. Numery te są

wyrażane w notacji kropkowo-szesnastkowej. Na przykład, la2b.0000.3c4d.5e6d mogłoby być prawidłowym adresem IPX, w którym „la2b" reprezentuje numer sieci, a „0000.3c4d.5e6d" jest numerem węzła.
Adresy IPX mogą być tworzone przez administratora sieci. Jednakże tak utworzone numery po znalezieniu się w sieci mogą spowodować występowanie konfliktów adresów. Wymyślanie numerów sieci obciąża administratora obowiązkiem ich utrzymywania i administrowania nimi. Lepszym rozwiązaniem jest więc pozyskanie zarejestrowanych numerów sieci IPX od firmy Novell.
Jako numer hosta IPX wykorzystuje się zwykle powszechnie przypisywany adres (adres MAC) karty sieciowej (NIC). Ponieważ adresy te są unikatowe, przynajmniej w teorii i w stopniu zależnym od zapewnienia jakości przez producenta, oferują wygodną i unikatową numerację hostów.
Podobnie jak IP, protokół IPX może obsługiwać wiele jednoczesnych sesji. Stwarza to potrzebę identyfikowania określonego procesu lub programu, który bierze udział w danej sesji. Identyfikację osiąga się dzięki stosowaniu 16-bitowego numeru „gniazda" w nagłówku IPX. Numer gniazda jest analogiczny do numeru portu w protokole TCP/IP.

1.12.4.6 **Wnioski dotyczące** IPX/SPX

Firma Novell Inc. zaobserwowała, jak pozycja rynkowa będącego jej własnością stosu protokołów IPX/SPX słabnie pod naporem konkurencji. Gdy dostępne stały się stosy protokołów otwartych, takich jak OSI, IP i inne, pozycja IPX/SPX bardzo na tym ucierpiała. Dostępne w handlu pakiety oprogramowania wspomagającego prace biurowe również wpłynęły na sprzedaż produktów firmy Novell. Będące jej własnością, ściśle połączone ze sobą serie produktów zapewniły początkowy sukces, ale stały się ciężarem w warunkach rynku ceniącego otwartość i współoperacyjność.
Novell zademonstrował swoje zaangażowanie w staraniach o odzyskanie utraconej pozycji, czyniąc IPv6 domyślnym protokołem przyszłych wersji systemu Netware. Aby pomyślnie wprowadzić tę zmianę strategii, Novell musi zapewnić kompatybilność między protokołami IPv6 i IPX/SPX. By osiągnąć ten cel, Novell blisko współpracował z Grupa Roboczą ds. Technicznych Internetu podczas projektowania IPv6. Dzięki temu wiele usług IPX stało się integralną częścią Ipv6.
Przygotowawszy grunt pod przyszłość, Novell musi teraz umożliwić bezbolesną migrację obecnego stosu protokołów i zestawu aplikacji do nowego środowiska. Co więcej. powinien także dostarczyć produkty i usługi podnoszące wartość wykorzystywania platformy sieci otwartej. Dla firny Novell wizją na przyszłość jest dostarczenie usługi katalogów sieciowych (ang. NDS - Network Directory Service) i powiązanych produktów dla dwóch grup użytkowników: środowiska Internetu i korporacyjnych intranetów.
Usługa NDS oferuje jeden, globalny, logiczny widok na wszystkie usługi i zasoby sieciowe. Umożliwia to użytkownikom dostęp do tych usług i zasobów po wykonaniu pojedynczego logowania, niezależnie od lokalizacji użytkownika czy zasobów.

1.12.5 Pakiet protokołów AppleTalk firmy Apple

Gdy komputery Apple zyskały większą popularność, a ich użytkownicy zaczęli z nich korzystać w sposób coraz bardziej wyszukany, nieunikniona stała się konieczność połączenia ich w sieć. Nie jest niespodzianką, że sieć opracowana przez Apple jest tak przyjazna użytkownikowi jak komputery tej firmy. Z każdym komputerem Apple sprzedawany jest AppleTalk, czyli stos protokołów pracy sieciowej firmy Apple, a także niezbędny sprzęt.
Przyłączenie do sieci jest równie proste jak wetknięcie wtyczki do złącza sieciowego i włączenie zasilania komputera. AppleTalk jest siecią równoprawną dostarczającą proste funkcje jak wspólne korzystanie z plików i drukarek. Inaczej niż w sieciach klient/serwer, funkcjonalności sieci równoprawnej nie ograniczają żadne sztywne definicje. Każdy komputer może działać jednocześnie jako serwer i klient.
AppleTalk został także przyjęty przez wielu innych producentów systemów operacyjnych. Nierzadko spotyka się możliwość obsługi stosu protokołów AppleTalk na komputerach innych niż Apple. Pozwala to klientom wykorzystywać AppleTalk i komputery Apple do tworzenia lub przyłączania się do istniejących sieci klient/serwer, innych niż sieci Apple.

1.12.5.1 Analiza AppleTalk

Stos protokołów AppleTalk obejmuje pięć warstw funkcjonalnych: dostępu do sieci, datagramową, sieci, informacji o strefach i aplikacji. Stos protokołów AppleTalk dość wiernie naśladuje funkcjonalność warstw transportu i sesji modelu referencyjnego OSI. Warstwy fizyczna i łącza danych zostały rozbite na wiele odrębnych warstw, specyficznych ze względu na ramki. AppleTalk integruje warstwy aplikacji i prezentacji, tworząc pojedynczą warstwę aplikacji. Rysunek 12.6 przedstawia to powiązanie funkcjonalne.
Stos protokołów AppleTalk odwzorowuje funkcjonalność warstw sieci, transportu i sesji modelu referencyjnego OSI, ale pozostałe cztery warstwy zawiera w dwóch.

1.12.5.1.1 Warstwa aplikacji sieci AppleTalk

AppleTalk łączy w pojedynczej warstwie aplikacji funkcjonalność warstw aplikacji i prezentacji modelu OSI. Ponieważ AppleTalk jest dosyć prostym stosem protokołów, warstwę tę zajmuje tylko jeden protokół. Jest to protokół dostępu do plików sieci AppleTalk (ang. AFP - AppleTalk Filing Protocol. Protokół AFP dostarcza usługi plików sieciowych aplikacjom istniejącym oddzielnie od stosu protokołów, takim jak poczta elektroniczna, kolejkowanie wydruków itd. Każda aplikacja uruchamiana na komputerze Apple musi przejść przez protokół AFP, jeśli chce wysłać informacje do sieci lub je z niej odebrać.
Rysunek 12.6. Porównanie modelu referencyjnego OSI i AppleTalk.

Nazwa warstwy modelu referencyjnego OSI	Numer warstwy OSI	Opis równoważnej warstwy AppleTalk
Aplikacji	7	Aplikacji
Prezentacji	6	
Sesji	5	Sesji
Transportu	4	Transportu
Sieci	3	Datagramowa
Łącza danych	2	Dostępu do sieci
Fizyczna	1	

1.12.5.1.2 Warstwa sesji sieci AppleTalk

Warstwa sesji w sieci AppleTalk obejmuje pięć podstawowych protokołów, dostarczających takie usługi, jak pełnodupleksowa transmisja, logiczne rozróżnianie nazw i adresów, dostęp do drukarki. ustalanie kolejności pakietów i inne.

Pierwszym protokołem warstwy sesji jest protokół strumienia danych sieci AppleTalk (ang. ADSP - AppleTalk Data Stream Protocol). Protokół ten dostarcza pełnodupleksowe usługi połączeniowe w wysoce niezawodny sposób, poprzez ustanawianie logicznego połączenia (sesji) pomiędzy dwoma komunikującymi się procesami na komputerach klientów. Protokół ADSP również zarządza tym połączeniem, dostarczając usługi sterowania strumieniem danych, zarządzania kolejnością i potwierdzania transmitowanych pakietów. Protokół ADSP wykorzystuje adresy gniazd do ustanowienia logicznego połączenia procesów. Po ustanowieniu tego połączenia dwa systemy mogą wymieniać dane.

Innym protokołem warstwy sesji sieci AppleTalk jest protokół sesji sieci AppleTalk (ang. ASP-AppleTalkSession Protocol). Protokół ten zapewnia niezawodne dostarczanie danych, wykorzystując sekwencyjne zarządzanie sesją, a także usługi transportowe protokołu transportu sieci AppleTalk (ang. ATP-Apple Talk Transport Protocol), który jest protokołem warstwy transportu.

Protokół trasowania AppleTalk (ang. AURP - AppleTalk Update-Based Routing Protocol) jest wykorzystywany w większych sieciach AppleTalk. Protokół ten służy przede wszystkim do zarządzania trasą i wymianą informacji pomiędzy urządzeniami trasującymi, zwłaszcza routerami bramek zewnętrznych.

Warstwa sesji sieci AppleTalk zawiera także protokół dostępu do drukarki (ang. PAP Priivter Access Protocol). Choć protokół ten został pierwotnie opracowany dla administrowania dostępem do drukarek sieciowych, może być wykorzystywany w rozmaitych wymianach danych. Zapewnia dwukierunkową sesję między dwoma urządzeniami, uzupełnioną o sterowanie strumieniem danych i zarządzanie kolejnością.

Ostatnim z protokołów warstwy sesji sieci AppleTalk jest protokół informacji o strefach (ang. ZIP - Zone Information Protocol). Zapewnia on mechanizm logicznego grupowania

indywidualnych urządzeń sieciowych z wykorzystaniem nazw przyjaznych dla użytkownika. Te grupy logiczne są nazywane strefami. W rozszerzonej sieci komputery mogą być rozrzucone po wielu sieciach, ciągle będąc zgrupowanymi w strefie. Jednak w małych, nie rozszerzonych sieciach, może być zdefiniowana tylko jedna strefa.

Protokół ZIP korzysta z protokołu wiązania nazw (ang. NBP -Name Binding Protocol), który jest protokołem warstwy transportu, do tłumaczenia nazw na numery sieci i węzła, a także z protokołu transportu ATP do aktualizowania informacji o strefach.

Pięć wymienionych protokołów warstwy sesji zapewnia klientom AppleTalk logiczne połączenia i transfery danych między komputerami, niezależnie od tego, jak bardzo są od siebie oddalone.

1.12.5.1.3 Warstwa transportu sieci AppleTalk

Warstwa transportu sieci AppleTalk oferuje usługi transportowe wszystkim warstwom wyższych poziomów. W warstwie tej istnieją cztery odrębne protokoły. Najczęściej używanym spośród nich jest protokół transportu AppleTalk (ATP).

Protokół ATP zapewnia niezawodny mechanizm dostarczania pakietów między dwoma komputerami. ATP korzysta z pól sekwencji i potwierdzenia, znajdujących się w nagłówku pakietu, aby zapewnić, że pakiety nie zaginą na drodze do miejsca przeznaczenia.

Kolejnym ważnym protokołem warstwy transportu AppleTalk jest protokół wiązania nazw (NBP). Jak wspominam wcześniej, NBP wykorzystuje protokół ZIP do tłumaczenia nazw przyjaznych dla użytkownika na rzeczywiste adresy. Protokół NBP przeprowadza faktyczną translację nazw stref na adresy sieci i węzłów. Protokół ten obejmuje cztery podstawowe funkcje:

• Rejestracja nazwy: Funkcja ta rejestruje unikalną nazwę logiczną w bazie rejestrów NBP.

• Przeglądanie nazw: Funkcja ta jest udostępniana komputerowi, który prosi o adres innego komputera. Prośba jest zgłaszana i zaspokajana w sposób jawny. Jeśli w prośbie podawana jest nazwa obiektu, protokół NBP zmienia tę nazwę w adres numeryczny. NBP zawsze przystępuje do zaspokajania takich próśb, przeglądając numery węzłów lokalnych. Jeśli żaden z nich nie pasuje, protokół NBP rozsyła prośbę do innych, połączonych ze sobą sieci AppleTalk. Jeśli wciąż nie można znaleźć pasującego adresu, czas prośby mija i proszące urządzenie otrzymuje komunikat o błędzie.

• Potwierdzenie nazwy: Żądania potwierdzenia są używane do weryfikacji związku obiektu z adresem.

• Usunięcie nazwy: W każdej sieci urządzenia są czasowo wyłączane lub odłączane. Gdy wystąpi taka sytuacja, wysyłane jest żądanie usunięcia nazwy, a tablice „obiekt-nazwa-adresowanie" są uaktualniane automatycznie.

Kolejnym protokołem warstwy transportu jest protokół echa sieci AppleTalk (ang. AEP - AppleTalk Echo Protocol. Służy on do określania dostępności systemu i obliczania czasu transmisji i potwierdzenia przyjęcia (ang. RTT- Round Trip Transmit Time).

Ostatnim protokołem warstwy transportu jest protokół utrzymania wyboru trasy (ang. *RTMP - Routing Table Maintenance Protocol).* Ponieważ AppleTalk stosuje w swojej warstwie sieci protokoły trasowane, musi zapewnić zarządzanie (administrowanie) tablicami trasowania. Protokół RTMP dostarcza routerom zawartość dla ich tablic trasowania.

1.12.5.1.4 Warstwa datagramowa sieci AppleTalk

Warstwa datagramowa sieci AppleTalk, analogiczna do warstwy 3 (sieci) modelu OSI, zapewnia bezpołączeniowe dostarczanie pakietowanych datagramów. Jest podstawą dla ustanawiania komunikacji i dostarczania danych przez sieć AppleTalk. Warstwa datagramowa jest również odpowiedzialna za zapewnianie dynamicznego adresowania węzłów sieciowych, jak też za rozróżnianie adresów MAC dla sieci IEEE 802.

Podstawowym protokołem tej warstwy jest protokół dostaw datagramów (ang. DDP Datagram Delivery Protocol). Zapewnia on transmisję danych przez wiele sieci w trybie bezpołączeniowym. Dostosowuje swoje nagłówki w zależności od miejsca przeznaczenia przesyłki. Podstawowe elementy pozostają stałe; dodatkowe pola są dodawane w razie potrzeby.

Datagramy, które mają być dostarczone lokalnie (innymi słowy w obrębie tej samej podsieci), wykorzystują tzw.. „krótki nagłówek". Datagramy, które wymagają trasowania do innych podsieci, wykorzystują format „rozszerzonego nagłówka". Format rozszerzony zawiera adresy sieci i pole licznika skoków.

Nagłówek DDP składa się z następujących pól:

• Liczba Skoków: Pole to zawiera licznik, zwiększany o jeden po każdym przejściu pakietu przez router. Pole liczby skoków jest wykorzystywane tylko w rozszerzonym nagłówku.

• Długość Datagramu: Pole zawiera długość datagramu i może służyć do sprawdzenia, czy nie został on uszkodzony podczas transmisji.

• Suma Kontrolna DDP: Jest to pole opcjonalne. Kiedy jest używane, zapewnia pewniejszą metodę wykrywania błędów niż proste sprawdzanie długości datagramu. Weryfikacja sumy kontrolnej wykrywa nawet niewielkie zmiany zawartości, niezależnie od tego, czy długość datagramu uległa zmianie.

• Numer Gniazda Źródłowego: To pole identyfikuje proces komunikujący w komputerze, który zainicjował połączenie.

• Numer Gniazda Docelowego: To pole identyfikuje proces komunikujący w komputerze, który odpowiedział na żądanie (prośbę) połączenia.

• Adres Źródłowy: Pole zawierające numery sieci i węzła komputera źródłowego. Jest używane tylko w rozszerzonym formacie nagłówka i umożliwia routerom przesyłanie datagramów przez wiele podsieci.

• Adres Docelowy: Pole zawierające numery sieci i węzła komputera docelowego. Jest używane tylko w rozszerzonym formacie nagłówka i umożliwia routerom przesyłanie datagramów przez wiele podsieci.

• Typ DDP: Pole identyfikujące zawarty w datagramie protokół wyższej warstwy. Jest wykorzystywane przez warstwę transportu komputera docelowego do identyfikowania odpowiedniego protokołu, do którego powinna być przesłana zawartość.

• Dane: Pole to zawiera przesyłane dane. Jego rozmiar może wynosić od 0 do 586 oktetów.

Warstwa datagramowa zawiera także protokół używany do przekształcania adresów węzłów w adresy MAC dla komputerów przyłączonych do sieci IEEE 802. Jest to protokół rozróżniania adresów sieci AppleTalk (ang. AARP - AppleTalk Address Resolution Protocol). Może być także używany do określania adresu węzła danej stacji. Protokół AARP przechowuje swoje informacje w tablicy odwzorowywania adresów (AMT). Stosownie do dynamicznego przypisywania numerów węzłów, tablica ta jest stale i automatycznie aktualizowana.

1.12.5.1.5 Warstwa łącza danych sieci AppleTalk

Warstwa łącza danych sieci AppleTalk odwzorowuje funkcjonalność warstw fizycznej i łącza danych modelu OSI. Funkcjonalność ta jest zintegrowana w podwarstwach specyficznych dla ramek. Na przykład, „EtherTalk" jest protokołem warstwy łącza danych, zapewniającym całkowitą funkcjonalność warstw fizycznej i łącza danych modelu OSI w ramach jednej podwarstwy. Podwarstwa ta umożliwia opakowywanie AppleTalk w strukturze ramki Ethernetu zgodnej z 802.3.

Istnieją podobne podwarstwy AppleTalk dla Token Ringu (znane jako „TokenTalk") i dla FDDI („FDDITaIk"). Protokoły te są nazywane „protokołami dostępu" ze względu na oferowane przez nie usługi dostępu do sieci fizycznej.

EtherTalk używa protokołu dostępu szeregowego, znanego jako „protokół dostępu do łącza EtherTalk" (ang. *ELAP - Ether Talk Link Access Protocol)* do pakowania danych i umieszczania ramek zgodnych z 802.3 w nośniku fizycznym. Taka konwencja nazewnicza i funkcjonalność protokołu dostępu szeregowego dotyczy również pozostałych protokołów dostępu. Na przykład, TokenTalk korzysta z „protokołu dostępu do łącza TokenTalk" (ang. *TLAP - Token Talk Link Access Protocol).*

Oprócz protokołów dostępu pasujących do standardów przemysłowych, firma Apple oferuje własny protokół sieci lokalnych, należący do warstwy łącza danych. Jest on znany pod nazwą „LocalTalk". LocalTalk działa z szybkością 230 Kbps, korzystając ze skrętki dwużyłowej. Wykorzystuje, jak można się spodziewać, protokół dostępu do łącza LocaITalk (ang. LLAP - Local Talk Link Access Protocol) do składania ramek i umieszczania ich w sieci. Protokół LLAP zawiera również mechanizmy zarządzania dostępem do nośnika, adresowania na poziomie łącza danych, opakowywania danych oraz reprezentacji bitowej dla transmisji ramki.

1.12.5.2 Schemat adresowania sieci AppleTalk

Schemat adresowania sieci AppleTalk składa się z dwóch części: numeru sieci i numeru węzła.

Numery sieci mają zwykle długość 16 bitów, choć w przypadku sieci nie rozszerzonych lub rozszerzonych w małym stopniu może być stosowane numerowanie jednoskładnikowe (8 bitów). Numery te muszą być zdefiniowane przez administratora sieci i używane przez

AppleTalk do trasowania pakietów między różnymi sieciami. Numer sieci „0" jest zarezerwowany przez protokół do wykorzystania przy pierwszym przyłączaniu nowych węzłów sieci. Numer sieci musi mieć wartość z zakresu od 00000001 do FFFFFFFF.
Numery węzłów są liczbami 8-bitowymi - dopuszczalny zakres adresów dla hostów. drukarek, routerów i innych urządzeń wynosi od l do 253; numery 0, 254 i 255 są zarezerwowane przez AppleTalk do wykorzystania w rozszerzonych sieciach. Węzły są numerowane dynamicznie przez warstwę łącza danych sieci AppleTalk.
Adresy AppleTalk są wyrażane w notacji kropkowo-dziesiętnej. Jak już wyjaśniono wcześniej w tym rozdziale. adres binarny jest zamieniany na dziesiętny system liczbowy. a kropka (.) służy do oddzielania numerów węzła i sieci. Na przykład, 100.99 odnosi się do urządzenia 99 w sieci I00. Początkowe zera zostały pominięte.
Wnioski dotyczące AppleTalk
AppleTalk jest firnowym stosem protokołów, przeznaczonym specjalnie dla pracujących w sieci komputerów osobistych firmy Apple. Jego przyszłość jest bezpośrednio związana z losami firmy Apple Corporation i kierunkami rozwoju jej technologii. Tak jak w przypadku firmowego stosu protokołów Novella, warstwy fizyczna i łącza danych służą do zapewnienia zgodności z technologiami sieciowymi opartymi na ustanowionych standardach. Jedynym wyjątkiem jest warstwa fizyczna LocalTalk, która może połączyć ze sobą komputery Apple, używając skrętki dwużyłowej przy szybkości do 230 Kbps.

1.12.6 NetBEUI

Ostatnim, zasługującym na omówienie protokołem jest NetBEUI. Ta niewygodna nazwa jest częściowo skrótem, a częściowo akronimem. Oznacza rozszerzony interfejs użytkownika NetBIOS (co z kolei jest skrótem od Podstawowego sieciowego systemu wejścia-wyjścia). Interfejs NetBEUI został opracowany przez IBM i wprowadzony na rynek w 1985 roku. Jest stosunkowo małym, ale wydajnym protokołem komunikacyjnym LAN.
Wizja (inny IBM dotycząca obliczeń rozproszonych zakładała w tamtych czasach segmentację sieci LAN, opartą na potrzebie wspólnej pracy. Poszczególne segmenty obsługiwałyby środowisko powiązane procesami pracy. Dane, do których potrzebny był dostęp, ale znajdujące się poza segmentem, mogły być odnalezione za pomocą pewnego rodzaju bramy aplikacji. Ze względu na takie pochodzenie nie powinno dziwić, że NetBEUI najlepiej nadaje się do małych sieci LAN. Wizja ta wyjaśnia również, dlaczego protokół NetBEUI nie jest trasowalny.
Protokół ten obejmuje warstwy 3 i 4 modelu referencyjnego OSI. Rysunek 12.7 przedstawia odpowiednie porównanie.
Rysunek 12.7. Porównanie modelu referencyjnego OSI
NetBEUI

Nazwa warstwy modelu referencyjnego OSI	Numer warstwy OSI	
Aplikacji	7	
Prezentacji	6	
Sesji	5	
Transportu	4	
Sieci	3	NetBEUI
Łącza danych	2	
Fizyczna	1	

Jak dowodzi rysunek 12.7, NetBEUI ustanawia komunikację pomiędzy dwoma komputerami i dostarcza mechanizmy zapewniające niezawodne dostarczenie i odpowiednią kolejność danych.
Ostatnio firma Microsoft wypuściła protokół NetBEUI 3.0. Jest to ważne z kilku powodów. Po pierwsze, wersja 3.0 jest bardziej tolerancyjna dla wolniejszych środków transmisji niż wersje wcześniejsze. Posiada też możliwość w pełni automatycznego dostrajania się. Najbardziej znaczącą zmianą w NetBEUI 3.0 jest wyeliminowanie samego protokołu NetBEUI. W sieciowych systemach operacyjnych firmy Microsoft został on zastąpiony protokołem ramki NetBIOS (ang. NBF- NetBIOS Frame). Zarówno NetBEUI, jak i NBF są ściśle związane z NetBIOS. Dlatego NetBEUI 3.0 (NBF) jest całkowicie kompatybilny i może współpracować z wcześniejszymi wersjami Microsoft NetBEUI.
NetBEUI, niezależnie od wersji, jest integralną częścią sieciowych systemów operacyjnych firmy Microsoft. Jeśli podejmiesz próbę uruchomienia systemu Windows NT 3.x (lub wyższego), Windows for Workgroups 3.11 czy nawet LAN Manager 2.x bez zainstalowanego protokołu NetBEUI, komputer nie będzie mógł się komunikować.

1.12.6.1 Wnioski dotyczące NetBEUI

NetBEUI jest wyłącznie protokołem transportu sieci LAN dla systemów operacyjnych firmy Microsoft. Nie jest trasowalny. Dlatego jego implementacje ograniczają się do domen warstwy 2, w których działają wyłącznie komputery wykorzystujące systemy operacyjne firmy Microsoft. Aczkolwiek staje się to coraz mniejszą przeszkodą, to jednak skutecznie ogranicza dostępne architektury obliczeniowe i aplikacje technologiczne.
Zalety korzystania z protokołu NetBEUI są następujące:
• Komputery korzystające z systemów operacyjnych lub oprogramowania sieciowego firmy Microsoft mogą się komunikować
• NetBEUI jest w pełni samodostrajającym się protokołem i najlepiej działa w małych segmentach LAN

• NetBEUI ma minimalne wymagania odnośnie pamięci
• NetBEUI zapewnia doskonałą ochronę przed błędami transmisji, a także powrót do normalnego stanu w razie ich wystąpienia
Wadą protokołu NetBEUI jest fakt, że nie może być trasowany i niezbyt dobrze działa w sieciach WAN.

1.12.7 Podsumowanie

Protokoły sieciowe są umiejscowione powyżej warstwy łącza danych. Prawidłowo zaprojektowane i skonstruowane są niezależne od architektur sieci LAN (opisanych w części II pt. „Tworzenie sieci LAN") i zapewniają całościowe zarządzanie transmisjami w domenach sieci LAN.

1.13 Rozdział 13 Sieci WAN

Mark A. Sportack
W sieciach rozległych (WAN) wykorzystywane są routery, protokoły routingu i urzą dzenia transmisji. Odpowiednio skonstruowane sieci WAN umożliwiają połączenie sieci lokalnych, bez względu na dzielące je odległości. Kluczowym zagadnieniem jest tu „odpowiednie skonstruowanie". Projektowanie, budowanie i administrowanie sieciami WAN wymaga opanowania zupełnie innych umiejętności niż w przypadku administrowania aplikacjami typu klient-serwer i sieciami lokalnymi. W niniejszym rozdziale opisane są różne składniki sieci rozległych. względy decydujące o kosztach, a także korzyści płynące ze stosowania każdego z tych składników.

1.13.1 Funkcjonowanie technologii WAN

Technologie sieci rozległych (WAN) oraz ich składników nieustannie zyskują na ważności. Zaledwie kilka lat temu jedynym wymaganiem stawianym sieci WAN było połączenie sieciowe dwóch lub więcej lokalizacji. Choć obecnie nadal jest to ważna funkcja sieci WAN, szybko pojawiają się nowe możliwości zastosowań. Na przykład - firma, w której praca odbywa się tylko w jednej lokalizacji, może potrzebować niezawodnego połączenia z siecią Internet, wykorzystywanego do marketingu, obsługi klienta i wielu innych funkcji. Innym przykładem może być rozproszenie pewnych operacji lub funkcji bądź współpraca pomiędzy firmami, powodująca konieczność połączenia ze sobą prywatnych sieci lokalnych.
Niestety, sieci rozległe znacząco różnią się od sieci lokalnych. Większość technologii sieci LAN jest ściśle powiązanych ze standardami przemysłowymi. Sieci WAN są natomiast strukturami wieloskładnikowymi, zbudowanymi przy wykorzystaniu różnorodnych technologii - zarówno standardowych, jak i bardzo specyficznych. Ponadto wiele konkurencyjnych technologii różni się znacznie funkcjami, wydajnością i kosztami. Najtrudniejszym etapem budowania sieci WAN jest dopasowanie odpowiednich technologii w sposób umożliwiający spełnienie zasadniczych wymagań użytkownika. Wymaga to głębokiego zrozumienia każdego aspektu zastosowania poszczególnych składników sieci WAN.

Do bazy technologicznej sieci rozległych należą:
• urządzenia transmisji,
• sprzęt komunikacyjny, w tym jednostki CSU i DSU, • adresowanie międzysieciowe,
• protokoły routingu.
Dla każdej z powyższych kategorii można wybierać z zaskakująco szerokiej gamy dostępnych technologii. Ponadto, każda technologia istnieje w kilku odmianach zależnych od producentów, modeli i konfiguracji. Przed wybraniem producentów i określonych produktów każda technologia powinna zostać sprawdzona pod kątem możliwych do osiągnięcia wydajności względem stawianych wymagań i spodziewanego obciążenia sieci WAN. Choć szczegółowe badanie ofert wszystkich producentów wykracza poza zakres niniejszej książki, opis każdej technologii stanowić będzie punkt odniesienia, umożliwiający ocenę istniejących produktów.

1.13.1.1 Korzystanie z urządzeń transmisji

Najbogatszą gamę rozwiązań dostępnych dla projektanta sieci WAN stanowią urządzenia transmisji. Istniejące urządzenia mają różne przepustowości, występują w wielu odmianach, a także różnią się kosztami. Na przykład przepustowość (szerokość pasma) urządzeń transmisji może wahać się od 9,6 kilobita na sekundę (Kbps) do ponad 44,736 megabitów na sekundę (Mbps). Owe urządzenia transmisji obsługują strumień cyfrowych informacji, płynący ze stałą i z góry określoną szybkością transmisji. Urządzenia te mogą korzystać z różnorodnych nośników fizycznych. takich jak skrętka dwużyłowa czy kable światłowodowe, a także obsługują wiele formatów ramek.
Specyfikacja DS-3 odnosi się do szybkości transmisji 44,736 Mbps, którą to wartość w niniejszej książce zaokrąglamy (dla wygody) do 45 Mbps.
Również sposób realizowania połączeń jest różny, zależnie od danego urządzenia. Istnieją dwa podstawowe typy urządzeń: urządzenia komutowania obwodów oraz komutowania pakietów. Wymienione typy obejmują wszystkie wersje urządzeń, choć innowacje technologiczne mogą w pewien sposób zacierać granicę podziału. Technologie te są pokrótce opisane w niniejszym rozdziale, co może ułatwić wybranie odpowiedniego typu sieci WAN. Bardziej szczegółowe informacje na temat urządzeń transmisji stosowanych w łączach dzierżawionych można znaleźć w rozdziale 14 pt. „Linie dzierżawione".

1.13.1.2 Urządzenia komutowania obwodów

Komutowanie obwodów jest metodą komunikacji, w której tworzone jest przełączane, dedykowane połączenie między dwiema stacjami końcowymi. Dobrym przykładem sieci

z komutowaniem obwodów jest system telefoniczny. Aparat telefoniczny jest na stałe połączony z centralą telefoniczną, należącą do lokalnego operatora usług telekomunikacyjnych. Istnieje wielu operatorów i jeszcze więcej central telefonicznych, więc połączenie między dwoma dowolnymi aparatami telefonicznymi tworzone jest z serii pośrednich połączeń między centralami telefonicznymi. Połączenie to jest fizycznym obwodem, dedykowanym danemu połączeniu na czas trwania sesji komunikacyjnej. Po zakończeniu sesji fizyczne połączenie między centralami przestaje istnieć, a zasoby sieci są zwalniane dla następnej rozmowy telefonicznej.
Zestawianie dedykowanych obwodów fizycznych między centralami jest istotą komutowania obwodów. Każda jednostka transmisji, niezależnie od tego, czy jest to komórka, ramka, czy dowolna inna konstrukcja, przebywa w infrastrukturze sieci tę samą fizyczną drogę. Opisywana koncepcja może być realizowana na kilka różnych sposobów. W kolejnych podrozdziałach przedstawione są trzy przykłady urządzeń komutowania obwodów: linie dzierżawione, ISDN i Switched 56.
Linie dzierżawione
Linie dzierżawione należą do najbardziej niezawodnych i elastycznych urządzeń komutowania obwodów. Obwody te noszą nazwę „linii dzierżawionych", ponieważ są one wynajmowane od operatora telekomunikacyjnego za miesięczną opłatę.
W Ameryce Północnej podstawową usługą cyfrowych linii dzierżawionych jest system T-Carrier. System ten udostępnia pasmo o przepustowości 1,544 Mbps, które można podzielić na 24 niezależne kanały, przesyłane przez dwie pary przewodów. Każdy kanał ma przepustowość 64 Kbps i może być dzielony na jeszcze mniejsze, np. o szybkości transmisji 9,6 Kbps. Linia o przepustowości 1,544 Mbps jest nazywana linią T-1. W systemie T-Carrier dostępne są również szybsze połączenia. Na przykład linia T-3 ma przepustowość 44,736 Mbps.
Łącza dzierżawione są często nazywane łączami stałymi lub „prywatnymi", ponieważ ich całe pasmo przenoszenia jest zarezerwowane dla podmiotu wynajmującego linię.
Dodatkowe informacje na temat linii dzierżawionych i systemu T-Carrier można znaleźć w rozdziale 14.

1.13.1.3 Cyfrowa sieć usług zintegrowanych (ISDN)

ISDN jest formą cyfrowej technologii komutacji obwodów, która umożliwia jednoczesne przesyłanie głosu i danych przez jedno fizyczne łącze, w której połączenie jest nawiązywane zależnie od potrzeb. Potrzeby te można realizować przy użyciu złączy ISDN stopnia podstawowego (BRI) lub głównego (PRI).
Złącze BRI pracuje z przepustowością 144 Kbps, w formacie znanym jako „2B+D". „ 2B" odnosi się do dwóch kanałów B o przepustowości 64 Kbps, które można wykorzystać jako jedno połączenie logiczne o szybkości 128 Kbps. Kanał D ma przepustowość 16 Kbps i pełni funkcje kontrolne, wykorzystywane np, przy nawiązywaniu i przerywaniu połączenia.
Złącze PRI jest zwykle udostępniane poprzez linie T-I przy szybkości transmisji 1,544 Mbps. Przepustowość ta jest najczęściej dzielona na 23 kanały B o szerokości 64 Kbps i 1 kanał D o szerokości 16 Kbps. Zamiast kanałów B i D (albo łącznie z nimi) można stosować szybsze kanały H o szerokości 384, 1 536 lub I 920 Kbps.
Kanał H3 o przepustowości 1 920 Kbps można stosować jedynie w Europie, gdzie standardową szybkością transmisji jest 2,048 Mbps (zamiast stosowanej w Stanach Zjednoczonych, Kanadzie i Japonii szybkości 1,544 Mbps). Próby wykorzystania kanału H3 przy dostępie do łącza o przepustowości 1,544 Mbps prowadzą do powstania nieużytecznych kanałów.
Choć z technicznego punktu widzenia ISDN jest systemem komutowania obwodów, może obsługiwać ponadto komutowanie pakietów, a częściowo nawet i łącza stałe.
Switched 56
Kolejną odmianą systemu komutowania obwodów, tworzącego połączenie zależnie od potrzeb, jest Switched 56. System ten udostępnia szybkość transmisji 56 Kbps między dwoma dowolnymi punktami korzystającymi z tej usługi. Podobnie jak w przypadku pozostałych systemów telefonicznych, przed nawiązaniem połączenia nie istnieje żaden obwód łączący owe punkty. Obwód taki jest zestawiany w chwili nawiązania połączenia między punktem źródłowym i docelowym. Użytkownicy nie znają rzeczywistych ścieżek połączenia w infrastrukturze telekomunikacyjnej, a informacje na ten temat nie mają dla nich znaczenia. Opisywany obwód przestaje istnieć po zakończeniu połączenia.
Ponieważ system Switched 56 nie ma charakteru łącz dedykowanych, jest on przystępną alternatywą dla linii dzierżawionych. Użytkownik płaci proporcjonalnie do korzystania z usługi, a nie za luksus posiadania całego pasma zarezerwowanego dla własnych potrzeb, niezależnie od stopnia jego wykorzystania. Wadą tego systemu jest jego mała wydajność. Obwody w systemie Switched 56 muszą być zestawiane w chwili żądania połączenia, co zajmuje określony czas. Dlatego połączenie przez dzierżawioną linię 56 Kbps może być nawiązane o wiele szybciej niż przy wykorzystaniu systemu Switched 56. Po nawiązaniu połączenia wydajność obu typów łączy powinna być zbliżona.
Switched 56 jest już technologią dojrzałą, powoli wychodzącą z użycia. Początkowo oferowała ona wysoko wydajne połączenia, przewyższające jakością możliwości modemów i tradycyjnych linii telefonicznych - przy stosunkowo małych kosztach, niższych w porównaniu z liniami dzierżawionymi. Obecnie postępy w technologiach sygnałowych umożliwiły modemom zmniejszenie tych różnic; co prawda Switched 56 wciąż ma większą wydajność od tak zwanych modemów 56 Kbps (pomimo swojej nazwy nie potrafią one osiągnąć i utrzymać takiej szybkości transmisji), ale nie jest to znacząca różnica. Obecnie system Switched 56 jest chyba najlepszym rozwiązaniem stosowanym awaryjnie zamiast linii dzierżawionych.

1.13.1.4 Urządzenia komutowania pakietów

W urządzeniach komutowania pakietów jest stosowany wewnętrzny format pakietów, wykorzystywany do opakowywania transportowanych danych. W odróżnieniu od urządzeń komutowania obwodów, urządzenia komutowania pakietów nie zestawiają dedykowanego połączenia pomiędzy dwiema lokalizacjami. Zamiast tego urządzenia dostępu klienta zapewniają połączenie z infrastrukturą operatora telekomunikacyjnego. Pakiety są przesyłane niezależnie od rodzaju połączenia przy wykorzystaniu istniejącej komercyjnej sieci

komutowania pakietów (PSN). W następnych podrozdziałach omówione są dwa przykłady sieci komutowania pakietów: stary i dobrze znany standard X.25 i jego młodszy krewny, Frame Relay.

1.13.1.4.1 X.25

X.25 jest bardzo starym protokołem komunikacyjnym dla sieci WAN, opracowanym przez organizację CCIT (znaną obecnie jako ITU - Międzynarodowa Unia Telekomunikacyjna). Operatorzy telekomunikacyjni po raz pierwszy udostępnili go jako komercyjną usługę we wczesnych latach 70.

Specyfikacje ITU oznaczane są niekiedy prefiksem ITU-T. Przyrostek T określa specyfikację jako należącą do standardów telekomunikacyjnych organizacji ITU.

Protokół X.25 można stosować zarówno w komutowanych, jak i w stałych obwodach wirtualnych. Komutowane obwody wirtualne (SVC) są zestawiane zależnie od potrzeb, a ich dekompozycja następuje natychmiast po zakończeniu sesji komunikacyjnej. Stałe kanały wirtualne (PVC) są z góry określonymi połączeniami logicznymi, łączącymi dwa punkty za pomocą sieci komutowanej. Zaletą obwodów SVC jest ich elastyczność i możliwość połączenia na żądanie dwóch dowolnych punktów sieci X.25. Ich ograniczeniem jest czas nawiązywania połączenia, jaki trzeba odczekać przed rozpoczęciem wymiany danych z innym urządzeniem w sieci.

Kanały PVC nie są tak elastyczne i konieczne jest ich uprzednie zdefiniowanie. Ich podstawową zaletą jest brak okresu nawiązywania połączenia. Dlatego też kanały PVC są zwykle wykorzystywane do obsługi komunikacji między urządzeniami, które wymieniają dane regularnie lub w sposób ciągły; obwody SVC wykorzystuje się do sporadycznej komunikacji.

Protokół X.25 zaopatrzono w skuteczne mechanizmy wykrywania i korekcji błędów, zapewniające wysoką niezawodność przy przesyłaniu za pośrednictwem elektromechanicznych urządzeń komutacyjnych infrastruktury telekomunikacyjnej. W protokole X.25 wydajność została poświęcona na rzecz niezawodności. Obecnie, w epoce komunikacji cyfrowej i optycznej, mechanizmy wykrywania i korekcji błędów protokołu X.25 nie mają już tak wielkiego znaczenia, stanowiąc raczej pewien zbędny narzut. Funkcje te aktualnie przejęły urządzenia komunikacyjne, więc nie ma konieczności ich pełnienia przez każde urządzenie w sieci. Aplikacje, które wciąż wymagają korzystania z protokołu X.25, mogą osiągnąć lepszą wydajność przy emulacji protokołu przez różne urządzenia transmisyjne.

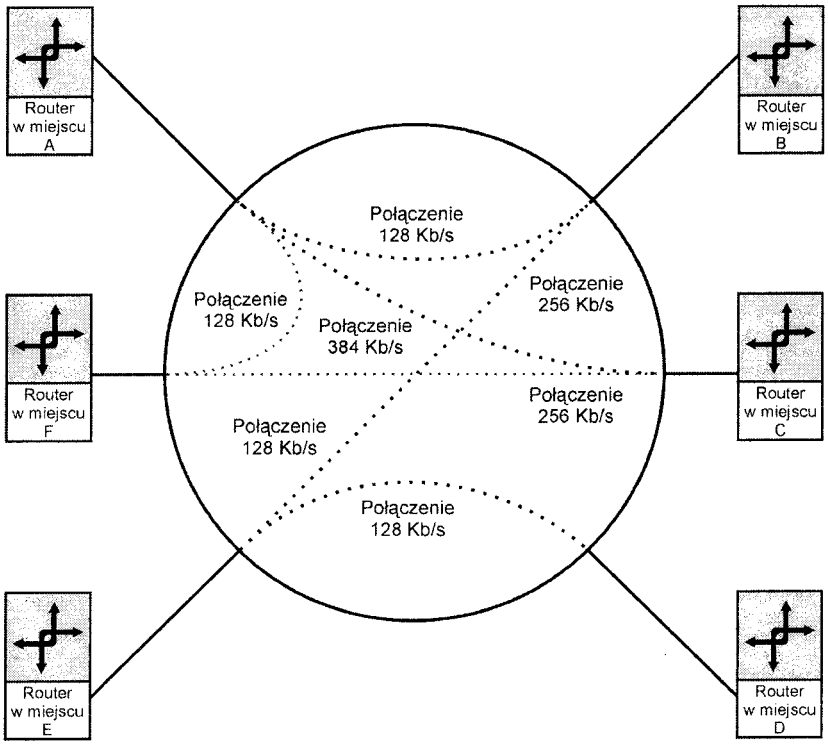
1.13.1.4.2 Frame Relay

Frame Relay jest szybszą odmianą komutowania pakietów X.25, obsługującą krótsze pakiety i mniej mechanizmów sprawdzania błędów. Obecnie Frame Relay obsługuje przesyłanie pakietów wyłącznie przez stałe kanały wirtualne (PVC) pomiędzy końcowymi routerami sieci. Planowana jest również obsługa przez ten protokół obwodów SVC, choć żaden dostawca usług nie określił jeszcze czasu realizacji tej zapowiedzi.

Punkty końcowe kanałów PVC są określane przez identyfikatory DLCI *(Data Link Connection Identifiers)* i mają przypisany umowny wskaźnik szybkości przesyłania informacji (CIR) przez sieć Frame Relay. Pary DLCI mają również przypisaną minimalną dostępną szerokość pasma, z możliwością czasowego przekroczenia tej granicy po spełnieniu określonych warunków. Korzystanie z identyfikatorów DLCI w sieciach Frame Relay ilustruje rysunek I3.1.

Rysunek 13.1. Pary logiczne połączeń Frame Relay.

Sieci rozległe Frame Relay są budowane przez zapewnienie stałego połączenia między punktem roboczym a najbliższą centralą oferującą tę usługę. W centrali dzierżawiona linia kończy się na przełączniku Frame Relay, który połączony jest w częściowe lub pełne oczka sieci z pozostałymi przełącznikami tego typu, tworzącymi komercyjną infrastrukturę Frame Relay danego operatora. Podobnie jak przełączniki głosowe centrali telefonicznej tworzące publiczną sieć telefoniczną (PSTN), przełączniki Frame Relay są niewidoczne dla użytkowników oraz wykorzystywanych aplikacji.



Podstawową zaletą protokołu Frame Relay jest redukcja kosztów połączenia sieciowego lokacji rozproszonych geograficznie przez zminimalizowanie długości własnych połączeń, wymaganych do uzyskania dostępu. Dostępne komercyjnie łącza mają przepustowość 1,544 Mbps, ze wskaźnikami CIR wykorzystywanymi do tworzenia logicznych połączeń z wieloma lokalizacjami, mających mniejszą szybkość transmisji.

Ceną za minimalizację kosztów urządzeń dostępu do linii dzierżawionych jest spadek wydajności. Protokół Frame Relay charakteryzuje się znacznym narzutem informacji „administracyjnych" (dotyczących ramek i protokołu), sumującym się z narzutami związanymi z liniami dzierżawionymi. Standardowym założeniem przy ustalaniu parametrów DLCI i CIR dla połączenia Frame Relay jest subskrypcja maksymalnie 1,024 Mbps z 1,544 Mbps dostępnego pasma. Gwarantuje to, że każdy identyfikator DLCI będzie miał przydzieloną stosowną szybkość przesyłania danych oraz że dostępna będzie rezerwa pasma na chwilowe przekroczenie tej szybkości.

Subskrypcja pasma polega na przydzieleniu pasma kanałom wydzielonym z większego pasma transmisyjnego. W przypadku protokołu Frame Relay każdy identyfikator DLCI ma przypisaną subskrypcję pasma. Subskrypcja ta nosi nazwę wskaźnika CIR.

Możliwe jest zdefiniowanie szeregu identyfikatorów DLCI z sumarycznym wskaźnikiem CIR, większym od dostępnej szerokości pasma transmisyjnego. Kontynuując przykład protokołu Frame Relay przesyłanego przez linię T-1, można by skonfigurować wskaźniki CIR na sumaryczną przepustowość 2,048 Mbps przy dostępnej szerokości pasma 1,544 Mbps. Praktyka taka nosi nazwę nadmiernej subskrypcji, a jej stosowanie nie jest wskazane. U jej podstaw leży założenie, że w dowolnym momencie nie wszystkie identyfikatory DLCI są aktywne, dzięki czemu nie jest wykorzystywana cała przepustowość określona wskaźnikami CIR. Założenie to nie jest zupełnie bezpodstawne, lecz stosowanie nadmiernej subskrypcji może sporadycznie wywoływać obniżenie wydajności usług w okresach szczytowego obciążenia sieci. 0 ile jest to możliwe, należy unikać operatorów telekomunikacyjnych, którzy nagminnie stosują nadmierną subskrypcję obwodów.

1.13.1.5 Urządzenia komutowania komórek

Technologią blisko spokrewnioną z komutowaniem pakietów jest komutowanie komórek. Komórka różni się od pakietu długością struktury. Pakiet jest strukturą danych o zmiennej długości, podczas gdy komórka jest strukturą danych o stałej długości. Najbardziej znaną technologią komutowania komórek jest tryb transferu asynchronicznego (ATM). Choć technicznie ATM jest obecnie technologią komutowania obwodów, najlepiej jest umieścić ją w oddzielnej kategorii.

Technologię ATM zaprojektowano z myślą o wykorzystaniu szybszych urządzeń transmisyjnych, takich jak architektury T-3 lub SONET.

1.13.1.6 Tryb transferu asynchronicznego (ATM)

Pierwotnie technologia ATM była projektowana jako mechanizm transportu asynchronicznego dla szerokopasmowego ISDN. Projektanci kierowali się założeniem, że krótkie czasy oczekiwania i duża szybkość transmisji spowodują, iż technologia ta równie dobrze sprawdzi się w sieciach lokalnych. Późniejsze trendy rynkowe niemal całkowicie ugruntowały jej reputację jako technologii sieci LAN, doprowadzając nawet do zanegowania możliwości zastosowań tej technologii w sieciach WAN.

Jako technologia komutowania komórek sieci rozległych, ATM jest dostępna komercyjnie po postacią łącz o szybkości 1,544 Mbps (DS-1) lub 44,736 Mbps (DS-3), choć dostęp do tych łącz nie jest taki sam we wszystkich obszarach geograficznych. Początkowo technologia ATM sieci rozległych była dostępna wyłącznie przez stałe obwody wirtualne, podobnie jak DLCI lub Frame Relay. Ostatecznie jednak stanie się ona technologią komutowania, umożliwiającą przesyłanie pojedynczych komórek bez narzutu wymaganego do zestawienia stałego obwodu wirtualnego lub rezerwowania szerokości pasma.

1.13.1.7 Wybór sprzętu komunikacyjnego

Sprzęt komunikacyjny potrzebny do zbudowania sieci WAN można podzielić na trzy podstawowe kategorie: sprzęt dostarczony przez klienta (CPE), urządzenia pośredniczące (ang. premises edge oehicles) oraz urządzenia przesyłania danych (DCE). W podanym kontekście DCE odnosi się do sprzętu operatora telekomunikacyjnego. W takiej sytuacji użytkownik nie ma zbyt dużego wpływu na wybór sprzętu DCE, dlatego też nie jest on opisany w tym podrozdziale.

CPE odnosi się do fizycznych mechanizmów komunikacyjnych łączących sprzęt: routery, sieci LAN, komutatory i inne urządzenia z komercyjną siecią telekomunikacyjną operatora.

Urządzenia pośredniczące są mechanizmami łączącymi sieci LAN z CPE. Pracują one na warstwach Layer 2 i 3 modelu referencyjnego OSI i są odpowiedzialne za przesyłanie i odbieranie pakietów, bazując na adresach międzysieciowych. Pełnią one w telekomunikacji rolę mechanizmów oddzielających sieci LAN od sieci WAN. Zarówno CPE, jak i urządzenia pośrednie dostarczane są przez klienta.

Operatorzy telekomunikacyjni dostarczają, rzecz jasna, znaczną ilość sprzętu obsługującego komunikację z użytkownikami. Sprzęt ten pozostaje niewidoczny dla użytkowników i administratorów sieci lokalnych i jako taki nie jest omawiany w niniejszej książce.

Akronim CPE może być odczytany jako „Customer-Provided Equipment" (Sprzęt dostarczony przez klienta) lub „Customer Premises Equipment" (Sprzęt w siedzibie klienta). Obie wersje można przyjąć za poprawne, gdyż mają takie samo znaczenie.

1.13.1.8 **Sprzęt własny klienta** (CPE)

CPE jest sprzętem pracującym na warstwie fizycznej, kodującym sygnały i przesyłającym je do urządzeń transmisyjnych. Sprzęt ten najczęściej jest dostarczany przez użytkowników i instalowany w należących do nich pomieszczeniach, po ich stronie linii demarkacyjnej. Linia ta. nazywana w skrócie „demarc", jest oficjalną granicą między instalacją operatora telekomunikacyjnego a instalacją użytkownika przyłączonego do infrastruktury operatora.

„Demarc" jest zwykle modułową skrzynką połączeń, oznaczoną numerami identyfikacyjnymi obwodów. Właścicielem skrzynki i wszystkich znajdujących się w niej elementów jest operator telekomunikacyjny. Użytkownik jest odpowiedzialny za całe wyposażenie podłączone do modułowego gniazda skrzynki - właśnie ów dostarczony przez użytkownika sprzęt określany jest akronimem CPE.

Typ sprzętu CPE zależy od technologii transmisji. Dwie najczęściej spotykane formy CPE to jednostki CSU/DSU oraz interfejs PAD.

1.13.1.8.1 Jednostka obsługi kanału / jednostka obsługi danych (CSU/DSU)

Typowa sieć WAN zbudowana jest na bazie linii dzierżawionych, czyli transmisyjnych urządzeń komutowania obwodów. Dlatego też typowy sprzęt dostarczony przez klienta znany jest jako CSU; DSU (jednostka obsługi kanału / jednostka obsługi danych). W odniesieniu do sprzętu tego typu przyjęto założenie, że urządzenie transmisyjne jest linią dzierżawioną - nie jest możliwe uzyskanie połączenia poprzez wybranie numeru telefonu.

Urządzenia CSU/DSU to sprzęt komunikacyjny znajdujący się na końcu kanałowych i cyfrowych urządzeń transmisyjnych. Zakończenie linii najczęściej ma postać modułowego gniazda. Sprzęt CSU.'DSU umożliwia również połączenie szeregowe z routerem, znajdującym się w siedzibie użytkownika, co pokazane jest na rysunku 13.2.

Nierzadko zdarza się, że niektóre osoby zaliczają routery do kategorii sprzętu własnego klienta. Przyczyną owej pomyłki jest założenie, że router jest zasadniczym elementem umożliwiającym komunikację z zewnętrznymi ośrodkami i że musi on być dostarczony przez użytkownika. Założenia te są z reguły poprawne, ale router nie jest zaliczany do urządzeń telekomunikacyjnych. Dlatego też jest on zaliczany nie do sprzętu CPE, lecz do wyposażenia pośredniczącego (premises edge vehicle).

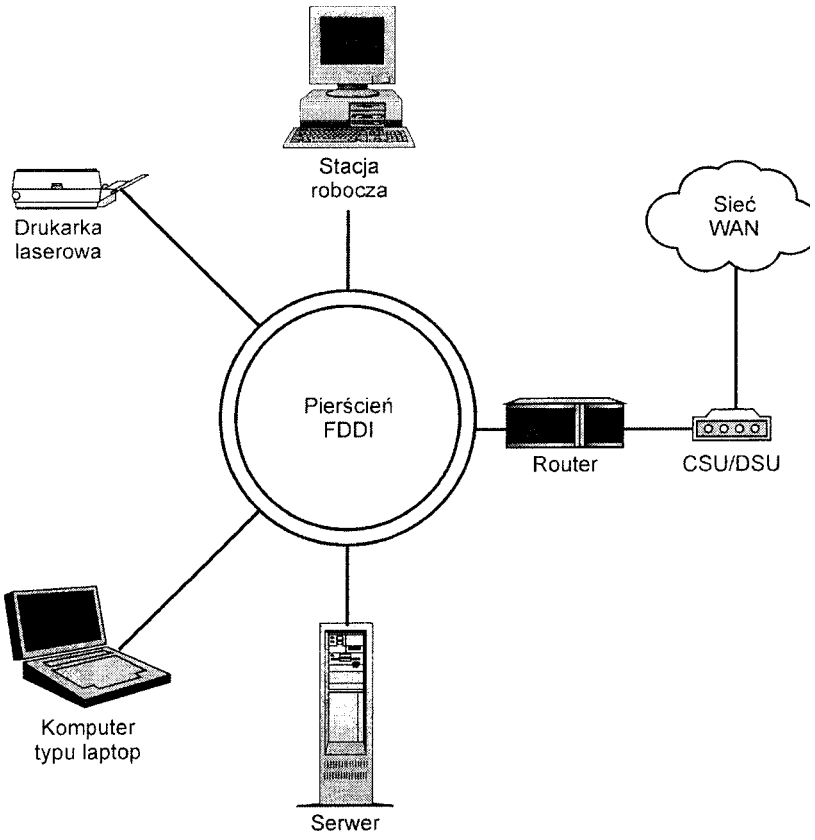
Jednostki CSU/DSU pełnią więcej funkcji niż tylko wysyłanie i odbieranie fizycznych sygnałów. Zależnie od marki i modelu, jednostki CSU/DSU mogą również wykonywać regulowanie łączy i odpowiadać na sygnały diagnostyczne z centrali. Jednostki te mają zasadnicze znaczenie we wszelkich obwodach dzierżawionych, obsługujących transmisję z szybkością 56 Kbps i większą.

Rysunek 13.2. Jednostka CSU/DSU łącząca router
urządzeniem transmisyjnym.

1.13.1.8.2 Interfejs zestawiania i dekompozycji pakietów (PAD)

Sprzęt transmisyjny wykorzystujący komutowanie pakietów, do ich tworzenia i dekompozycji może wymagać dodatkowych urządzeń. Urządzenia te znane są jako PAD (akronim od angielskiego „Packet Assembler/Disassembler"). Dobrym przykładem technologii sieciowej, w której wykorzystywane są interfejsy PAD, jest sieć X.25. W sieci takiej do połączenia ośrodków użytkowników z infrastrukturą komutowanej sieci operatora telekomunikacyjnego najczęściej wykorzystuje się urządzenia transmitujące z szybkością 9,6 Kbps. Urządzeniami końcowymi w tych stosunkowo powolnych urządzeniach były interfejsy PAD.

Obecne technologie komutowania pakietów z reguły korzystają z urządzeń transmisyjnych komutowania obwodów. Na przykład Frame Relay, choć jest bezpośrednim następcą protokołu X.25, nie wykorzystuje interfejsów PAD. Zamiast tego sieci LAN można połączyć poprzez Frame Relay, korzystając z logicznych podkanałów wydzielonych z linii T-I. Jeśli założyć, że linia T-I zapewnia szerokość pasma 1,544 Mbps, należy ją zakończyć u klienta jednostką CSU/DSU, niezależnie od obsługiwanej technologii transmisji. Dlatego też sieci WAN zbudowane na podstawie Frame Relay mają routery i jednostki CSU/DSU w każdej lokacji. Jednostki te współpracują z urządzeniami T-1, łączącymi je za pośrednictwem sieci Frame Relay.



1.13.1.9 Urządzenia pośredniczące *(Premises Edge Vehicles)*

Wyposażenie *premises edge vehicle* służy do łączenia sieci lokalnej klienta z urządzeniami CPE. W środowisku typowej sieci LAN urządzeniem tym jest router. Routery pełnią funkcję granicy między siecią LAN i WAN. Dlatego ich podstawowym zadaniem jest komunikacja z innymi routerami o znanych adresach międzysieciowych. Adresy te są przechowywane w tablicach routingu, umożliwiających powiązanie adresów z fizycznym interfejsem routera, który należy wykorzystać w celu uzyskania połączenia ze wskazanym adresem.

1.13.2 Adresowanie międzysieciowe

W sieciach WAN istnieje stałe zapotrzebowanie na urządzenie adresujące, znajdujące się poza strukturą sieci LAN. Adresy międzysieciowe są elementami warstwy Layer 3, warstwy fizycznej modelu referencyjnego OSI. Adresy te są wykorzystywane w celu uzyskania dostępu i wymiany danych z hostami w innych podsieciach sieci WAN.
Architektura adresów jest określona przez trasowalny protokół wykorzystywany w sieci WAN. Spośród dostępnych możliwości można wymienić protokoły IPv4, IPv6, IPX czy AppleTalk. Każdy z nich ma unikatowy schemat adresowania. Dlatego też wybór protokołu decyduje o możliwej do zastosowania hierarchii adresów.

1.13.2.1 Zapewnianie adresowania unikatowego

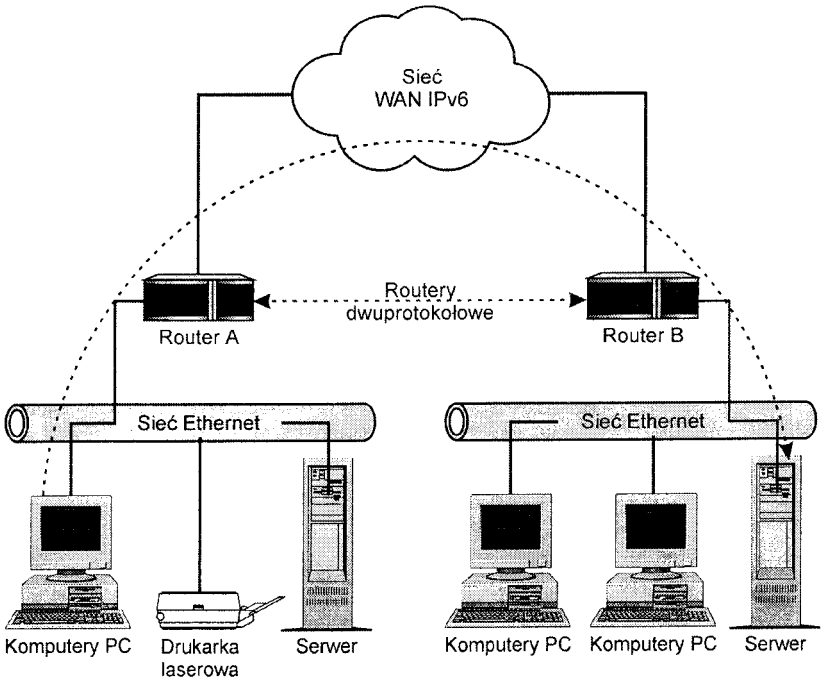
Najważniejszym zagadnieniem adresowania międzysieciowego jest jego unikatowość. Poza jedynym wyjątkiem w postaci protokołu IPv6, każdy protokół sieciowy wymaga, aby w dowolnym momencie istniał tylko jeden punkt końcowy o danym adresie. Powtarzające się adresy międzysieciowe są przyczyną pojawiania się błędów routingu i naruszają spójność operacji sieciowych klienta.
Protokół IPv6 ma nową architekturę adresowania, znaną jako anycast. Adresy anycast mogą być łatwo utworzone (czasem nawet w niezamierzony sposób), jeśli ten sam adres jest przypisany do wielu urządzeń. Gdy do sieci dotrze pakiet z określonym adresem anycast, jest on po prostu przesyłany do najbliższego urządzenia o takim adresie. Dlatego też urządzenia o adresach anycast muszą być całkowicie wymienne, zarówno pod względem ich obsługi, jak i działania.
Teoretycznie, jeśli sieć WAN nie będzie połączona z Internetem lub innymi sieciami, to adresy międzysieciowe mogą być wybierane w dowolny sposób. Ogólnie rzecz ujmując, dowolność wybierania adresów międzysieciowych stanowi przejaw krótkowzroczności i poważnego zaniedbania obowiązków. Na poparcie tego stwierdzenia w maju 1993 roku został opublikowany dokument Request for Comment (RFC) numer 1597, w którym
przedstawiony jest plan mający zapobiec takim praktykom. Zostały w nim określone i zarezerwowane wyłącznie dla wewnętrznych potrzeb sieci trzy obszary adresów. W obszarach tych znajdują się adresy klasy A, B i C protokołu IPv4. Omawiane obszary to:
• 10.0.0.0 - 10.255.255.255,
• 172.16.0.0 - 172.31.255.255,
• 192.168.0.0 - 192.168.255.255.

Wymienione obszary zostały zarezerwowane przez organizację IANA (Internet Assigned Numbers Authority) do wykorzystania w sieciach prywatnych. W dokumencie RFC numer 1597 znajduje się zastrzeżenie, że wymienione adresy nie mogą być wykorzystywane podczas bezpośredniego dostępu do Internetu. Firmy wykorzystujące owe adresy. które chcą uzyskać dostęp do Internetu, mogą wykorzystać jako serwer pośredniczący serwer proxy o unikatowym i zarejestrowanym adresie IP. Innym rozwiązaniem jest wykorzystanie konwersji adresu sieciowego (NAT).
Podczas stosowania adresów zarezerwowanych w dokumencie RFC nr 1597 w dalszym ciągu konieczne jest zapewnienie każdemu urządzeniu unikatowego adresu z domeny prywatnej sieci. Adresy te nie muszą być unikatowe w skali globalnej, a jedynie lokalnie.

1.13.2.2 Współdziałanie międzysieciowe z wykorzystaniem różnych protokołów

1.13.2.2.1 Tunele

Nie w każdej sieci WAN istnieje luksus korzystania z pojedynczego trasowanego protokołu. Sieci wieloprotokołowe stwarzają pewne zasadnicze trudności, zwłaszcza w przypadku połączeń poprzez protokoły o niwielkim podobieństwie; do rozwiązania owych trudności posłużyć się można tunelami i bramami.
Tunele są stosunkowo prostymi konstrukcjami, które można wykorzystać do przesyłania danych między normalnie niekompatybilnymi obszarami sieci. Pakiety danych są opakowywane za pomocą ramek rozpoznawanych w sieci, przez którą są transportowane. Pierwotne ramki i formatowanie pozostaje bez zmian, lecz traktowane jest jako „dane".
Po dotarciu pakietu do celu host odbiorcy rozpakowuje go, ignorując „opakowanie". W ten sposób zostaje odtworzona pierwotna postać pakietu, wraz z oryginalnym adresowaniem międzysieciowym. Na rysunku 13.3 przedstawione jest tunelowanie pakietów protokołu IPv4 przesyłanych przez obszar sieci IPv6. Z powodu istniejących różnic w długości adresów protokoły te nie są bezpośrednio kompatybilne. Aby przezwyciężyć ten problem, pakiety protokołu IPv4 są opakowywane w protokół IPv6 przez router A. co umożliwia ich przesłanie przez sieć WAN opartą na protokole IPv6. Router B usuwa opakowanie IPv6 i przesyła odtworzony pakiet IPv4 do hosta docelowego w rozpoznawalnej przez niego formie.
Rysunek 13.3. Tunelowanie pakietów lPv4 przez obszar sieci IPv6.
Dodatkowe informacje na temat adresowania internetowego trasowanych protokołów można znaleźć w rozdziale 12, zatytułowanym „Protokoły sieciowe".



1.13.2.2.2 Bramy

Jeśli sieć WAN wymaga połączenia podsieci o różnych trasowanych protokołach, to na granicach jej poszczególnych obszarów sieci należy umieścić bramy. Brama (ang. gateway)' jest urządzeniem zdolnym do tłumaczenia struktur adresowania między dwoma różnymi protokołami. Rolę bramy mogą pełnić routery lub hosty. Jedynym kryterium wyboru jest możliwość tłumaczenia przez urządzenie architektur adresów między dwoma protokołami.
Routery mogą wykonywać omawianą konwersję na dwa sposoby. Pierwszym z nich jest korzystanie z dwóch różnych protokołów trasowania. Wymaga to, aby router obliczał trasy i przesyłał informacje o nich, a następnie przesyłał pakiety obu protokołów. Routery są projektowane do pracy w środowiskach wieloprotokołowych, więc wykonywanie opisywanego zadania nie powinno sprawiać żadnych kłopotów.

Inny sposób to obsługa przez router zintegrowanego protokołu, zdolnego do jednoczesnego trasowania dwóch różnych protokołów i adresów. Przykładem tej postaci protokołu routingu są pojawiające się serie protokołów typu „ng", zaprojektowane do obsługi przenoszenia danych między protokołem IPv4 i IPv6, np. OSPFng i RIPng.

W literaturze polskiej spotyka się również termin „śluza' jako odpowiednik gateway- co w pewnym stopniu odzwierciedla spełnianą przez nie funkcję pośrednictwa pomiędzy dwoma peotokołami (PrzYP red.).

1.13.3 Korzystanie z protokołów trasowania

Protokoły trasowania dynamicznego są wykorzystywane przez routery do pełnienia trzech podstawowych funkcji:

• wyszukiwanie nowych tras,

• przekazywanie do innych routerów informacji o znalezionych trasach, • przesyłanie pakietów za pomocą owych routerów.

Protokoły trasowania dynamicznego podzielone są na trzy obszerne kategorie: protokoły wektora odległości, protokoły zależne od stanu złącza oraz protokoły hybrydowe. Każda z tych kategorii jest omówiona w następnych podrozdziałach. Podstawową różnicą między nimi jest sposób pełnienia dwóch pierwszych spośród trzech wspomnianych funkcji. Jedyną alternatywą trasowania dynamicznego jest trasowanie statyczne, opisane w jednym z następnych podrozdziałów.

1.13.3.1 Trasowanie na podstawie wektora odległości

Trasowanie może być oparte na algorytmach wektora odległości (nazywanych również algorytmami Belhmana-Forda), wymagających okresowego przesyłania przez routery kopii tablic trasowania do najbliższych sąsiadów w sieci. Każdy odbiorca tablicy dodaje do niej wektor odległości (własną ,.wartość" odległości) i przesyła ją do najbliższych sąsiadów. Proces ten przebiega we wszystkich kierunkach jednocześnie między bezpośrednio sąsiadującymi routerami.

Ten wieloetapowy proces umożliwia każdemu routerowi poznanie innych routerów oraz stworzenie sumarycznego obrazu „odległości" w sieci. Na przykład, jednym z pierwszych protokołów opartych na wektorze odległości jest RIP (Routing Information Protocol). Protokół ten do określenia następnej najlepszej ścieżki dla dowolnego pakietu wykorzystuje dwie metryki odległości. Wartości tych metryk zależą od czasu, ponieważ mierzone są znakami kontrolnymi (ang. „ticks") i liczbą skoków (ang. „hop count").

Do określenia optymalnych tras między dowolną parą punktu źródłowego i docelowego routery mogą korzystać z zaskakującej różnorodności metryk. „Odległość" mierzona daną metryką może nie mieć nic wspólnego z odległością w sensie geometrycznym - może na przykład odnosić się do czasu, liczby skoków routera lub podobnych parametrów.

Następnie uzyskana sumaryczna tablica odległości wykorzystywana jest do uaktualnienia tablic trasowania każdego routera. Po zakończeniu opisywanego procesu routery uzyskują informacje na temat odległości do zasobów sieciowych. Informacje te nie zawierają żadnych konkretnych danych na temat pozostałych routerów czy rzeczywistej topologii sieci.

Takie podejście może w określonych warunkach spowodować pojawienie się problemów z protokołami opartymi na wektorach odległości. Przykładowo, po awarii łączy routery potrzebują pewnej ilości czasu na poznanie nowej topologii sieci. W czasie trwania tego procesu sieć może być podatna na niespójne trasowanie, a nawet nieskończone pętle.

Pewne zabezpieczenia mogą ograniczyć owe zagrożenia, lecz nie zmienia to faktu, że w trakcie „dostrajania się" sieci wydajność przesyłania danych jest niestabilna. Dlatego też starsze protokoły, które powoli dostosowują się do zmian w sieci, mogą nie być odpowiednie dla dużych i skomplikowanych sieci WAN.

1.13.3.2 Trasowanie na podstawie stanu łącza

Algorytmy trasowania na podstawie stanu łącza, ogólnie określane jako protokoły „najpierw najkrótsza ścieżka" (ang. SPF - shortest path.first), utrzymują złożoną bazę danych opisującą topologię sieci. W odróżnieniu od protokołów wektora odległości, protokoły stanu łącza zbierają i przechowują pełną informację na temat routerów sieci, a także o sposobie ich połączenia.

Uzyskanie tych informacji jest możliwe dzięki wymianie pakietów LSP (ang. Link-State Packet) z innymi bezpośrednio połączonymi routerami. Każdy router, który wymienił pakiety LSP buduje na ich podstawie topologiczną bazę danych. Następnie wykorzystywany jest algorytm SPF w celu obliczenia dostępności punktów docelowych sieci. Informacja ta jest wykorzystywana do uaktualnienia tablicy trasowania. Opisywany proces umożliwia wykrywanie zmian w topologii sieci, które mogły powstać w wyniku awarii składników sieci lub jej rozbudowy. W rzeczywistości wymiana pakietów LSP nie jest przeprowadzana okresowo, lecz dopiero po wystąpieniu w sieci określonego zdarzenia.

Trasowanie w oparciu o stan łącza ma dwie cechy, które mogą stwarzać zagrożenia. Po pierwsze, w trakcie początkowego procesu poznawania sieci trasowanie to może przeciążyć łącza transmisyjne, znacznie obniżając możliwości sieci w zakresie transportowania danych. Wspomniane obniżenie wydajności ma charakter przejściowy, ale jest niestety mocno odczuwalne.

Inny problem polega na tym, że omawiana metoda trasowania wymaga dużej pamięci i szybkiego procesora. Z tego powodu routery skonfigurowane do obsługi trasowania na postawie stanu łącza są stosunkowo drogie.

1.13.3.3 Trasowanie hybrydowe

Ostatnią formą trasowania dynamicznego jest praca hybrydowa. Choć istnieją „otwarte" zrównoważone protokoły hybrydowe, ta forma trasowania jest niemal całkowicie związana z zastrzeżonym produktem jednej firmy - Cisco Systems, Inc. Protokół o nazwie EIGRP (ang. Enhanced Interior Gateway Routing Protocol) został zaprojektowany z zamiarem połączenia najlepszych cech protokołów opartych na wektorze odległości i stanie łącza, przy jednoczesnym ominięciu ich ograniczeń wydajności i innych wad.

Protokoły hybrydowe korzystają z metryk wektorów odległości, lecz szczególny nacisk jest w nich położony na metryki dokładniejsze niż w konwencjonalnych protokołach opartych na wektorach odległości. Również szybsze jest dostosowywanie się do zmian w sieci, przy

ominięciu narzutów spotykanych przy uaktualnieniu stanów łączy. Zrów noważone protokoły hybrydowe nie pracują okresowo, lecz w przypadku wystąpienia określonych zdarzeń w sieci, co oszczędza szerokość pasma dla użytecznych aplikacji.

1.13.3.4 Trasowanie statyczne

Router zaprogramowany do trasowania statycznego przesyła pakiety przez z góry określone porty. Po skonfigurowaniu routerów statycznych nie jest konieczne poznawanie tras ani przesyłanie jakichkolwiek informacji na ich temat. Rola tych urządzeń została ograniczona wyłącznie do przesyłania pakietów.
Trasowanie statyczne sprawdza się jedynie w przypadku bardzo małych sieci, w których przesyłanie danych do wszelkich punktów docelowych odbywa się po tej samej ścieżce. W takiej sytuacji trasowanie statyczne może być najlepszym rozwiązaniem. ponieważ nie wymaga ono dodatkowej szerokości pasma na poznawanie tras i komunikację z innymi routerami.
W miarę rozrastania się sieci i powstawania dodatkowych połączeń utrzymanie trasowania statycznego staje się coraz bardziej pracochłonne. Po każdej zmianie w dostępności routerów lub urządzeń transmisyjnych sieci WAN konieczne jest ich ręczne sprawdzenie i zaprogramowanie. Sieci WAN o bardziej skomplikowanej topologii, gdzie możliwe jest korzystanie z wielu ścieżek, wymagają stosowania trasowania dynamicznego. Stosowanie trasowania statycznego w takich sieciach przeczyłoby w ogóle sensowi istnienia wielokrotnych ścieżek.

1.13.3.5 Wybór protokołu

Protokół trasowania powinien być wybrany w sposób uważny i z uwzględnieniem długoterminowych konsekwencji dokonanego wyboru. Wybranie protokołu bezpośrednio wpływa na rodzaj stosowanego routera oraz na wydajność działania sieci WAN. Znajdujące się w poprzednich podrozdziałach opisy trasowania statycznego oraz różnych klas protokołów powinny w pełni uświadomić następstwa wybrania każdego z tych rozwiązań. Dzięki temu możliwe jest zawężenie wyboru do jednej kategorii lub klasy protokołów.
Następnym krokiem jest określenie, czy w sieci WAN mają być wykorzystane routery jednego czy też kilku producentów. O ile jest to możliwe, zalecane jest korzystanie ze sprzętu jednego producenta. Przyczyna tego jest całkiem prosta: otwarte protokoły trasowania zostawiają producentom pewien margines na modyfikacje. Z tego powodu wersja protokołu trasowania danego producenta raczej nie będzie w 100% wymienna z protokołem innego producenta. Jednym z lepszych przykładów takiej sytuacji są szczegółowo udokumentowane różnice między protokołami OSPF (ang. Open Shortest Path First) firm Bay Networks i Cisco System.
Jeśli producent routera zostanie wybrany przed protokołem trasowania, należy uświadomić sobie wynikające z tego ograniczenia wyboru protokołów. Niektóre protokoły trasowania są produktami zastrzeżonymi, co oznacza, że można je uzyskać wyłącznie u jednego producenta.

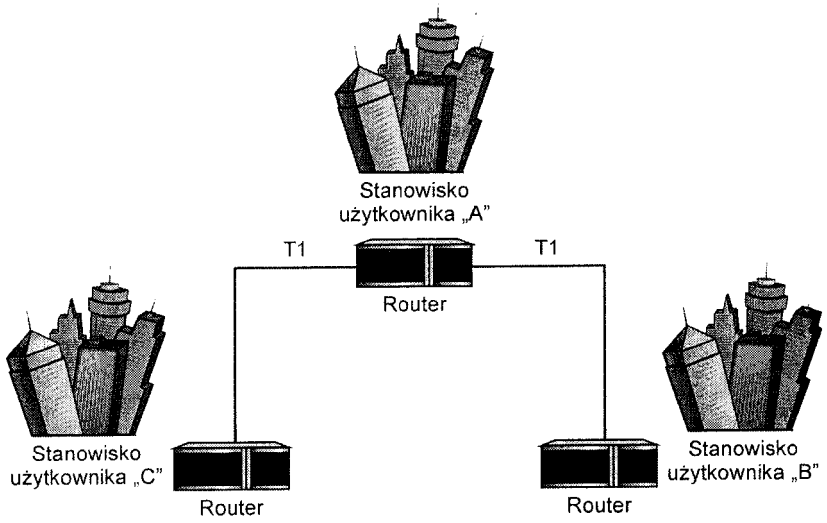
1.13.4 Topologie WAN

Topologia sieci WAN opisuje organizację urządzeń transmisyjnych względem lokalizacji połączonych za ich pomocą. Istnieje wiele różnych topologii, z których każda charakteryzuje się innym wskaźnikiem kosztów, wydajności i możliwości rozbudowy. Ponadto topologie bezpośrednio bazujące na urządzeniach transmisyjnych mogą charakteryzować się dodatkową specjalizacją funkcjonalną. Najbardziej rozpowszechnionymi topologiami stosowanymi w sieciach WAN są:
• każdy-z-każdym,
• pierścienia,
• gwiazdy,
• oczek pełnych,
• oczek częściowych,
• wielowarstwową, w tym dwu- i trójwarstwową,
• hybrydową.
Choć niektóre z tych topologii kojarzone są raczej z sieciami LAN, to równie dobrze sprawdzają się w sieciach WAN. Wszystkie wymienione topologie są opisane i zilustrowane w dalszej części tego podrozdziału. Można tu również znaleźć informacje na temat względnych kosztów, wydajności, możliwości rozbudowy oraz wymagań technologicznych każdej topologii.

1.13.4.1 Topologia każdy-z-każdym

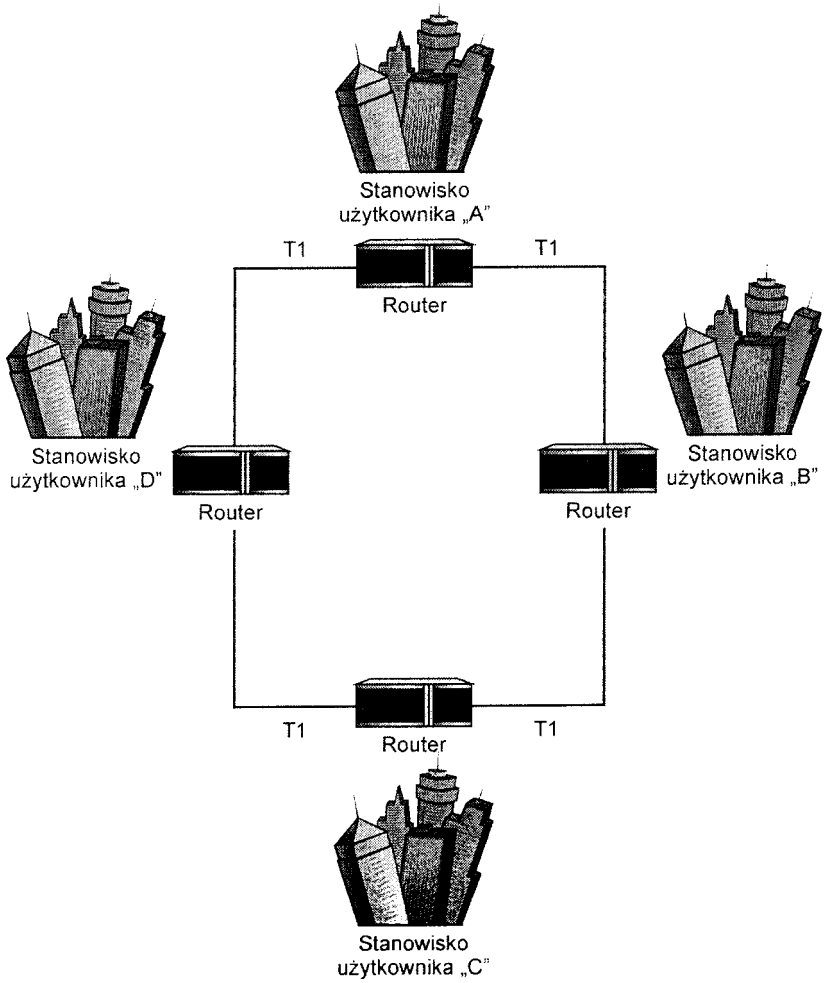
Sieć rozległa o topologii „każdy-z-każdym" może być zbudowana na bazie linii dzierżawionych lub dowolnych innych urządzeń transmisyjnych. Omawiana topologia sieci WAN jest stosunkowo prostym sposobem połączenia niewielkiej liczby punktów. Sieci WAN, składające się tylko z dwóch lokacji, można połączyć wyłącznie w taki sposób. Na rysunku 13.4. jest przedstawiona niewielka sieć rozległa o topologii każdy-z-każdym.
Omawiana topologia jest najtańszym rozwiązaniem dla sieci WAN o niewielkiej liczbie połączonych lokalizacji. Ponieważ każda lokalizacja ma co najwyżej dwa połączenia z resztą sieci, możliwe jest zastosowanie trasowania statycznego. Choć konfiguracja trasowania statycznego jest dosyć pracochłonna, pozwala jednak uniknąć narzutów charakterystycznych dla protokołów trasowania dynamicznego. Jeśli założyć, że w tak prostej topologii nie ma większej liczby dostępnych tras, korzyści płynące z zastosowania trasowania dynamicznego są raczej ograniczone.
Rysunek 13.4. Sieć WAN
o topologii każdy-z-każdym, zbudowana na podstawie linii dzierżawionych.

W trasowanej sieci WAN liczba routerów między danym punktem a lokalizacją docelową jest bezpośrednio związana z czasem, w ciągu którego pakiet znajduje się w drodze do celu. Dlatego możliwe jest opracowanie metryki odpowiadającej liczbie routerów znajdujących się na określonej ścieżce w sieci. Metryka ta nosi nazwę „liczby skoków". Przejście pakietu przez jeden router jest liczone jako jeden skok.
Niestety, sieci rozległe o topologii każdy-z-każdym mają dwa podstawowe ograniczenia. Po pierwsze, nie poddają się one zbyt dobrze rozbudowie. W miarę pojawiania się w sieci nowych lokalizacji liczba skoków między dowolną ich parą staje się bardzo niestała i ma tendencję rosnącą. Skutkiem tego są zmienne poziomy wydajności komunikacji między dowolną daną parą lokacji. Rzeczywisty stopień zmienności wydajności w znacznym stopniu zależy od szeregu czynników, do których należą m.in.:
• rzeczywista odległość między lokacjami, • typ i szybkość urządzenia transmisyjnego,
• stopień wykorzystania urządzenia transmisyjnego.
Drugim ograniczeniem tego rozwiązania jest podatność na awarie składników sieci. Między daną parą lokalizacji istnieje tylko jedna ścieżka przepływu informacji. Wskutek tego awaria sprzętu lub urządzenia transmisyjnego w dowolnym punkcie sieci typu każdy-z-każdym może doprowadzić do podzielenia sieci WAN. W zależności od przepływu informacji i stosowanego typu trasowania, taka awaria może poważnie zakłócić komunikację w całej sieci WAN.
Inną ważną konsekwencją braku dodatkowych tras w topologii każdy-z-każdym jest marnowanie czasu i pracy procesorów przez protokoły trasowania dynamicznego, obliczające trasy i przesyłające pakiety w sieci. Dzieje się tak dlatego, że obliczona trasa między danymi dwoma punktami nigdy nie ulega zmianie. Z tego powodu ręczne zdefiniowanie tras może - ujmując rzecz statystycznie - poprawić wydajność sieci.



1.13.4.2 Topologia pierścienia

Topologię pierścienia można w prosty sposób uzyskać z topologii każdy-z-każdym, dodając jedno urządzenie transmisyjne i po jednym porcie w dwóch routerach. To niewielkie zwiększenie kosztów pozwala uzyskać zwiększenie liczby tras w małych sieciach, dzięki czemu można w nich zastosować protokoły trasowania dynamicznego. Zakładając, że koszt większości urządzeń transmisyjnych zależy od odległości przesyłania danych, rozsądnie jest tak zaprojektować pierścień sieci, aby zminimalizować całkowitą długość łączy. Omawiana topologia sieci WAN zilustrowana jest na rysunku 13.5.
Rysunek 13.5. Sieć WAN połączona u' pierścień.
Sieć WAN o topologii pierścienia, zbudowaną z linii transmisyjnych łączących pary punktów, można wykorzystać do połączenia niewielkiej liczby lokalizacji, zapewniając jednocześnie zwiększenie liczby tras przy minimalnym wzroście kosztów. Istnienie



w sieci wielu potencjalnych tras oznacza, że wykorzystanie protokołów trasowania dynamicznego zapewni elastyczność nieosiągalną przy trasowaniu statycznym. Protokoły trasowania dynamicznego potrafią automatycznie wykryć i dostosować się do niekorzystnych zmian w warunkach pracy sieci WAN, wyszukując trasy omijające uszkodzone połączenia.

Również topologia pierścienia ma pewne podstawowe ograniczenia. Zależnie od geograficznego rozmieszczenia lokacji, dodanie jeszcze jednego urządzenia transmisyjnego zamykającego pierścień może okazać się zbyt kosztowne. W takich sytuacjach alternatywą dedykowanych linii dzierżawionych może być technologia Frame Relay, pod warunkiem. że jej ograniczenia wydajności są możliwe do przyjęcia przy projektowanych obciążeniach sieci.

Drugim ograniczeniem topologii pierścienia jest mała możliwość rozbudowy sieci. Dodanie do sieci WAN nowych lokalizacji bezpośrednio zwiększa liczbę skoków wymaganych do uzyskania dostępu do innych punktów pierścienia. Przeprowadzenie takiego procesu dodawania może również wymagać zamówienia nowych obwodów. Na przykład, jeśli do sieci przedstawionej na rysunku 13.5. zostanie dodana lokalizacja X, znajdująca się w pobliżu lokalizacji C i D, konieczne staje się usunięcie obwodu od lokalizacji C do D. W celu zachowania integralności sieci należy zamówić dwa nowe połączenia: jedno łączące lokacje C i X oraz drugie, między lokalizacjami D i X.

Topologia pierścienia przy jej ograniczeniach lepiej się sprawdza przy łączeniu jedynie bardzo małej liczby lokacji. Jedyną cechą przemawiającą na jej korzyść względem topologii każdy-z-każdym jest zapewnienie dodatkowych tras do każdej lokacji w sieci.

1.13.4.3 Topologia gwiazdy

Odmianą topologii każdy-z-każdym jest topologia gwiazdy, nazwana tak od jej charakterystycznego kształtu. Gwiazda jest budowana przez połączenie wszystkich lokalizacji z jedną lokalizacją docelową. Można by się spierać, że w istocie jest to topologia dwuwarstwowa. Cechą odróżniającą topologię gwiazdy od dwuwarstwowej jest fakt, że centralny router topologii gwiazdy, oprócz obsługi sieci WAN, może być również wykorzystany do wzajemnego połączenia miejscowych sieci LAN.

W przypadku topologii dwuwarstwowej, opisanej w dalszej części niniejszego rozdziału, router drugiej warstwy powinien być wykorzystywany wyłącznie do połączenia urządzeń transmisyjnych z innych lokacji. Co ważniejsze, topologia dwuwarstwowa zapewnia wielość tras przez obsługę rozbudowy sieci z wieloma punktami koncentracji.

Sieć o topologii gwiazdy można zbudować, korzystając z niemal każdego dedykowanego urządzenia transmisyjnego, włączając w to Frame Relay i prywatne linie łączące dwa punkty. Sieć WAN o topologii gwiazdy jest przedstawiona na rysunku 13.6.

Sieci WAN o topologii gwiazdy i z urządzeniami transmisyjnymi łączącymi punkt z punktem są znacznie łatwiejsze w rozbudowie od sieci o topologii pierścienia lub każdy-z

-każdym. Dodanie lokacji do gwiazdy nie wymaga przebudowy istniejących łączy transmisyjnych. Jedyne co trzeba zrobić, to zapewnić nowe połączenie między centralnym routerem sieci a routerem w nowej lokalizacji.

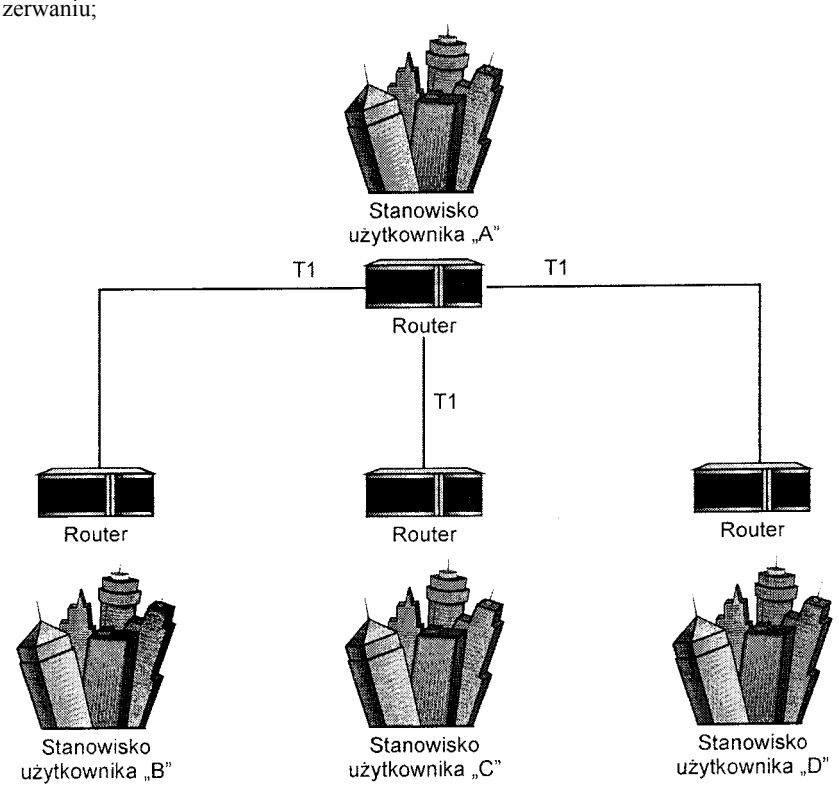
Rysunek 13.6. Sieć IVAN
o topologii gwiazdy.

Topologia gwiazdy pozwala rozwiązać problemy rozbudowy obecne w sieciach każdyz-każdym, wykorzystując router do wzajemnego połączenia, czyli *skoncentrowania* wszystkich routerów sieci. Rozbudowa ta odbywa się przy niewielkim wzroście liczby routerów, ich portów i urządzeń transmisyjnych, w porównaniu z topologią każdy-zkażdym podobnych rozmiarów. Topologię gwiazdy można zbudować przy wykorzystaniu nawet mniejszej liczby urządzeń niż w przypadku topologii pierścienia, co jest zilustrowane rysunkami 13.7. i 13.8.

Możliwość rozbudowy topologii gwiazdy jest ograniczona liczbą portów możliwych do obsłużenia przez router w centralnym punkcie gwiazdy. Przekroczenie tego ograniczenia wymaga albo przebudowania sieci w topologię dwuwarstwową, albo wymiany istniejącego routera na znacznie większy.

Inną zaletą topologii gwiazdy jest lepsza wydajność sieci. Teoretycznie topologia gwiazdy zawsze przewyższa wydajnością topologię pierścienia i każdy-z-każdym. Przyczyną tego jest fakt, iż wszystkie urządzenia w sieci są odległe od siebie tylko o trzy skoki: router w lokacji użytkownika, centralny router sieci i router lokacji docelowej. Ten poziom stałości jest charakterystyczny tylko dla topologii gwiazdy. Omawiana topologia ma dwie wady:

• istnienie pojedynczego punktu awaryjnego: oznacza to, że w przypadku awarii centralnego routera sieci WAN cała komunikacja ulegnie zerwaniu;

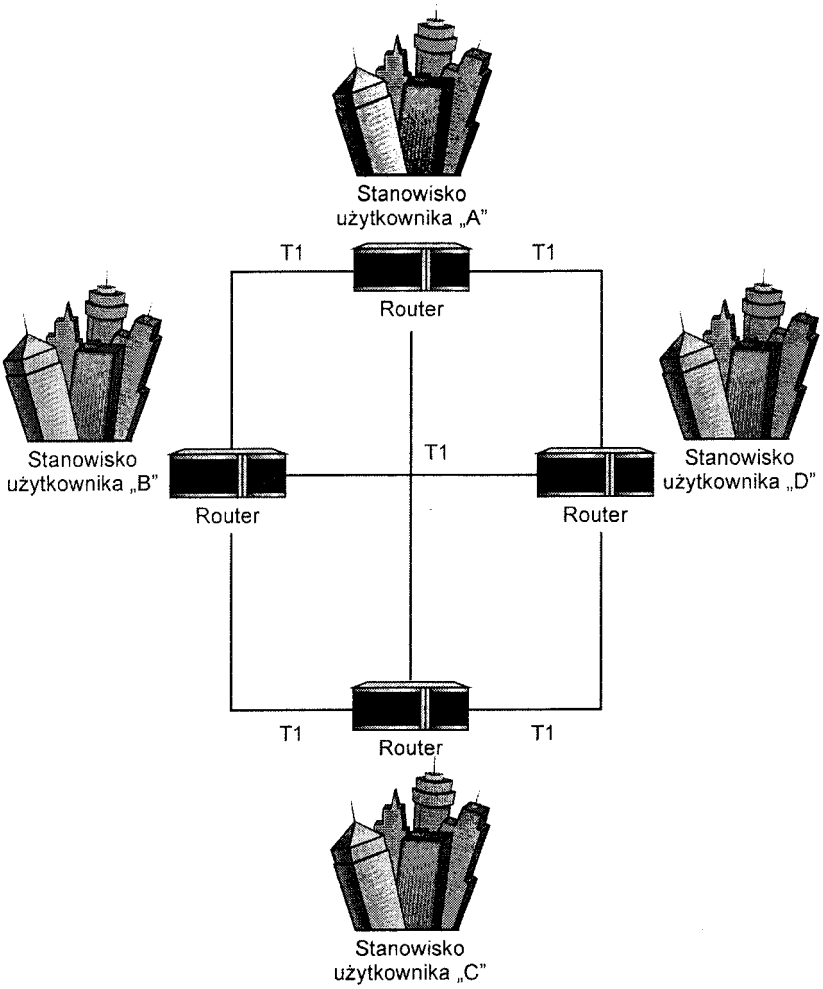


• brak dodatkowych tras: jeśli centralny router ulegnie awarii, komunikacja jest zerwana do chwili usunięcia problemu; protokoły trasowania dynamicznego nie są w stanie obliczyć nowych tras przez sieć, ponieważ trasy takie nie istnieją!

1.13.4.4 Topologia oczek pełnych

Maksymalną niezawodnością charakteryzuje się topologia oczek pełnych. Daje ona największą znaną niezawodność i odporność na uszkodzenia. W sieci takiej każdy węzeł jest bezpośrednio połączony z wszystkimi pozostałymi. Dzięki temu istnieje obfita liczba dodatkowych tras do każdej lokacji. Można się domyślić, że stosowanie w takiej sieci trasowania statycznego jest zupełnie nierealne. W sieci takiej praktycznie jest się zmuszonym do wybrania jednego z protokołów trasowania dynamicznego, umożliwiających obliczanie tras i przesyłania pakietów w sieci. Sieć WAN o topologii oczek pełnych jest przedstawiona na rysunku 13.7.

Rysunek 13.7. Sieć rozległa
o topologii oczek pełnych.



Topologia ta zapewnia zminimalizowanie liczby skoków między dowolnymi dwoma komputerami w sieci. Inną jej zaletą jest możliwość korzystania praktycznie z każdej technologii transmisyjnej.

Jednak nawet topologia oczek pełnych ma pewne praktyczne ograniczenia. Przykładowo, sieci WAN o takiej topologii są dosyć drogie w budowie. Każdy router musi być na tyle duży, aby miał liczbę portów i urządzeń transmisyjnych wystarczającą do połączenia z każdym innym routerem w sieci WAN. Oprócz drogiej budowy, sieć taka charakteryzuje się również wysokimi opłatami miesięcznymi. Ponadto ma ona ograniczone (choć duże) możliwości rozbudowy. Routery mają ograniczoną liczbę portów, które mogą być obsługiwane. Dlatego też sieci o topologii oczek pełnych są rozwiązaniami raczej utopijnymi, o ograniczonej możliwości praktycznego wykorzystania.

Możliwym do zastosowania rozwiązaniem jest połączenie ograniczonej liczby routerów wymagających szybkiego dostępu do sieci. Inne potencjalne rozwiązanie to zastosowanie topologii oczek pełnych jedynie we fragmentach sieci WAN, takich jak centralne części sieci wielowarstwowych lub ściśle powiązane ośrodki robocze. Dokładniejsze informacje na ten temat znajdują się w podrozdziale zatytułowanym „Topologie hybrydowe".

1.13.4.5 Topologia oczek częściowych

Sieci WAN można również zbudować w „częściowej" topologii oczek. Oczka częściowe to bardzo elastyczne topologie, mogące przyjąć różnorodne formy. Topologie oczek częściowych najlepiej opisać jako sieci o routerach powiązanych ze sobą ściślej niż w przypadku jakiejkolwiek topologii podstawowej; w topologii oczek częściowych nie wszystkie punkty sieci są bezpośrednio połączone, jak to było w przypadku oczek pełnych; przykład sieci w takiej topologii jest pokazany na rysunku 13.8.

Sieci WAN o topologii oczek częściowych można łatwo rozpoznać po często stosowanym połączeniu poszczególnych węzłów sieci ze wszystkimi pozostałymi węzłami. Sieci takie pozwalają zminimalizować liczbę skoków między użytkownikami bardzo rozbudowanych sieci WAN. W odróżnieniu od sieci oczek pełnych, oczka częściowe umożliwiają zredukowanie kosztów budowy i eksploatacji przez ograniczenie liczby połączeń z mniej obciążonymi segmentami sieci WAN. Dzięki temu topologia oczek częściowych lepiej nadaje się do rozbudowy i jest tańsza od topologii oczek pełnych.

1.13.4.6 Topologia dwuwarstwowa

Topologia dwuwarstwowa jest odmianą podstawowej topologii gwiazdy: miejsce pojedynczego routera centralnego zajmują tu (co najmniej) dwa routery. Eliminuje to podstawową wadę topologii gwiazdy (tj. zupełną katastrofę w przypadku awarii centralnego routera), zachowując jednocześnie możliwości rozbudowy i nie zmniejszając wydajności.

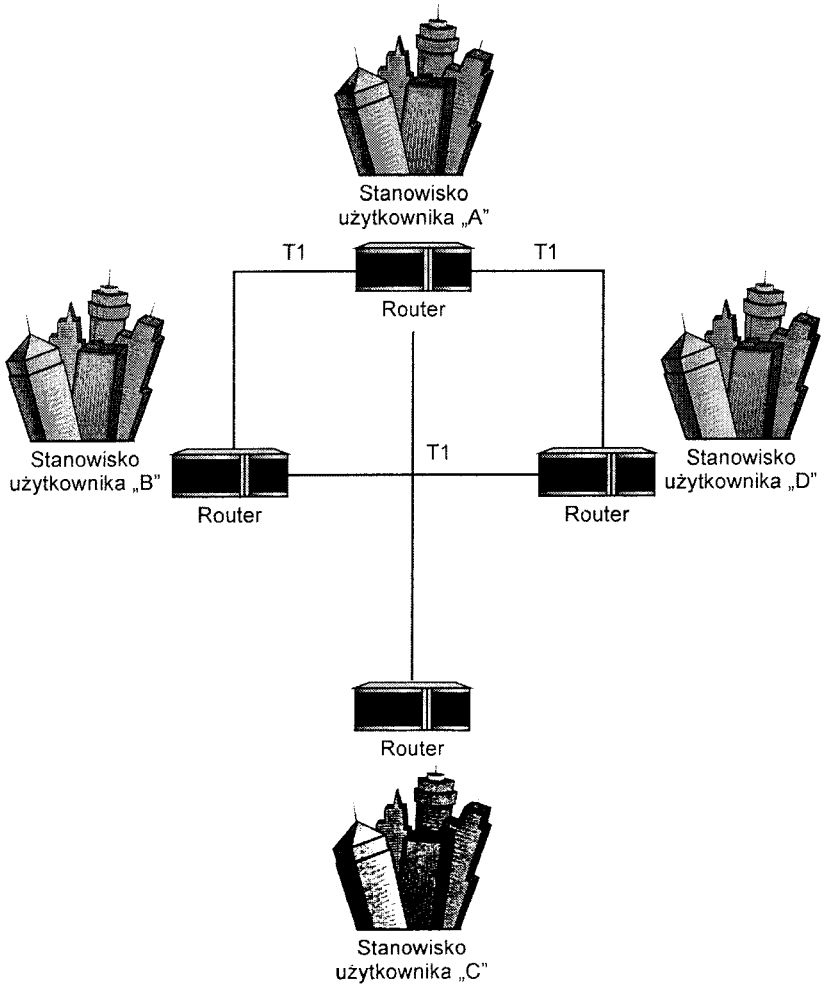
Na rysunku 13.9 jest przedstawiony schemat sieci WAN o typowej topologii dwuwarstwowej. Największa możliwa liczba skoków zwiększa się o jeden, jako efekt umieszczenia dodatkowego routera centralnego. Jednak w odróżnieniu od sieci każdy-z-każdym, przedstawionej na rysunku 13.4, parametr liczby skoków nie ulega pogorszeniu po dodaniu do sieci nowych lokalizacji.

Rysunek 13.8. Topologia oczek c=ęściowych.

Dwuwarstwowa sieć WAN zbudowana na podstawie dedykowanych łącz wykazuje lepszą odporność na uszkodzenia od sieci o topologii gwiazdy- przy równie dużych możliwościach jej rozbudowy. Omawiana topologia może być stosowana w wielu zbliżonych odmianach, różniących się przede wszystkim liczbą centralnych routerów oraz sposobem ich wzajemnego połączenia. Jeśli w sieci znajdują się więcej niż dwa routery centralne, projektant sieci powinien wybrać podtopologię warstwy routerów centralnych. Routery te mogą być połączone w topologii oczek pełnych, oczek częściowych lub każdy-zkażdym.

Niezależnie od wybranej podtopologii, hierarchiczne, wielowarstwowe topologie najlepiej sprawdzają się, jeśli spełnione są wymienione poniżej podstawowe warunki:

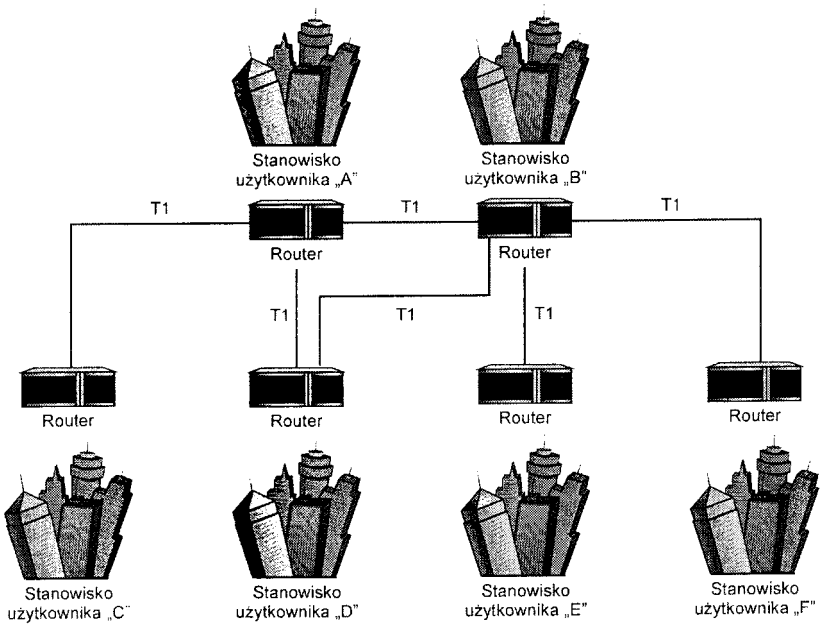
• Warstwa routerów centralnych powinna być przeznaczona wyłącznie na potrzeby tych routerów; oznacza to, że nie może być ona wykorzystana do bezpośredniego łączenia ośrodków użytkowników.



Rysunek 13.9. Dwuwarstwowa sieć 1-VAN.

• Routery w ośrodkach użytkowników powinny być połączone wyłącznie z węzłami centralnymi, bez wzajemnych połączeń w konfiguracji każdy-z-każdym.

• Routery użytkowników nie mogą być łączone z routerami centralnymi w sposób przypadkowy; ich położenie powinno być dobrane w sposób optymalny; zależnie od geograficznego rozmieszczenia użytkowników i wykorzystywanych urządzeń transmisyjnych, bezpieczniejsze może okazać się umieszczenie węzłów centralnych tak, aby zminimalizować odległości od lokalizacji użytkowników.



Ponieważ trasowanie w sieci skupia się na jednym lub więcej routerach, stosowanie tej topologii może być kosztownym przedsięwzięciem. Dlatego rozwiązanie to jest przede wszystkim wykorzystywane w większych firmach.

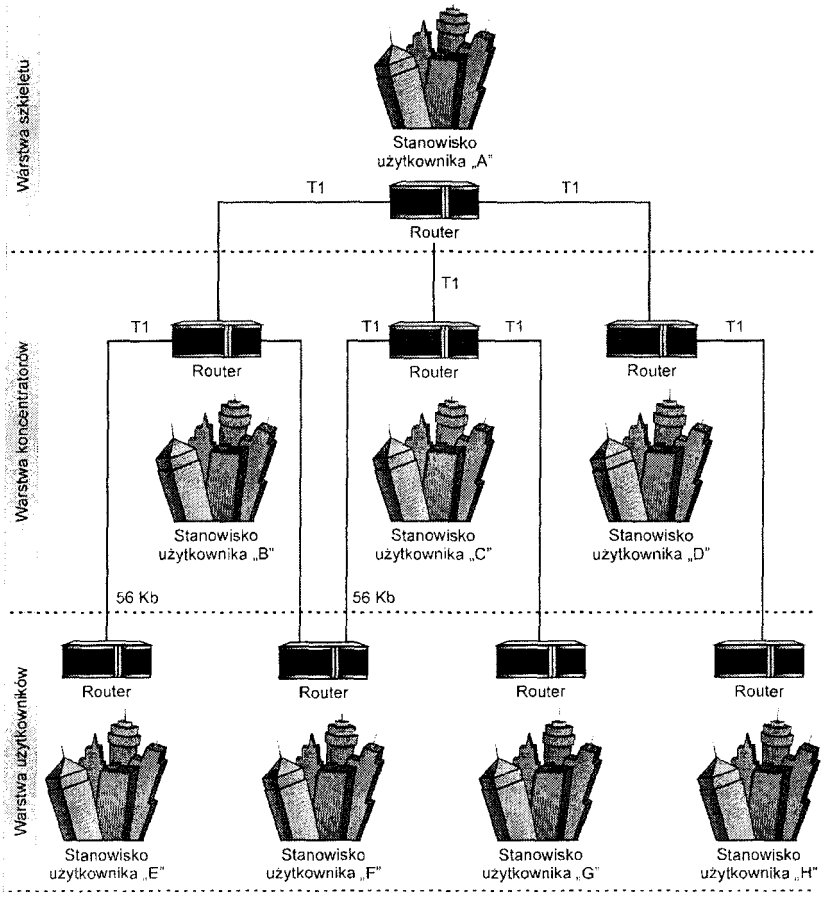
1.13.4.7 Topologia trójwarstwowa

Architektura dwuwarstwowa może okazać się nieodpowiednia dla tych sieci WAN, w których zachodzi potrzeba połączenia bardzo dużej liczby lokalizacji lub które są zbudowane na bazie mniejszych routerów, obsługujących jedynie kilka połączeń szeregowych. Aby zwiększyć możliwości rozbudowy sieci do wymaganego poziomu może więc okazać się konieczne dodanie trzeciej warstwy. Przykład sieci w topologii trójwarstwowej przedstawia rysunek 13.10.

Trójwarstwowe sieci WAN zbudowane na bazie dedykowanych urządzeń transmisyjnych są jeszcze bardziej odporne na awarie i mają większe możliwości rozbudowy niż sieci dwuwarstwowe. Sieci trójwarstwowe są jednak drogie w budowie, eksploatacji

Rysunek 13.10. Sieć WAN
o topologii trójwarstwowej.

i utrzymaniu, powinny być więc wykorzystywane jedynie do łączenia bardzo dużej liczby lokalizacji. W takiej sytuacji nierozsądne wydaje się tworzenie bardzo dużej sieci WAN, w której najwyższa (szkieletowa) warstwa routerów ma topologię inną niż topologia oczek pełnych.

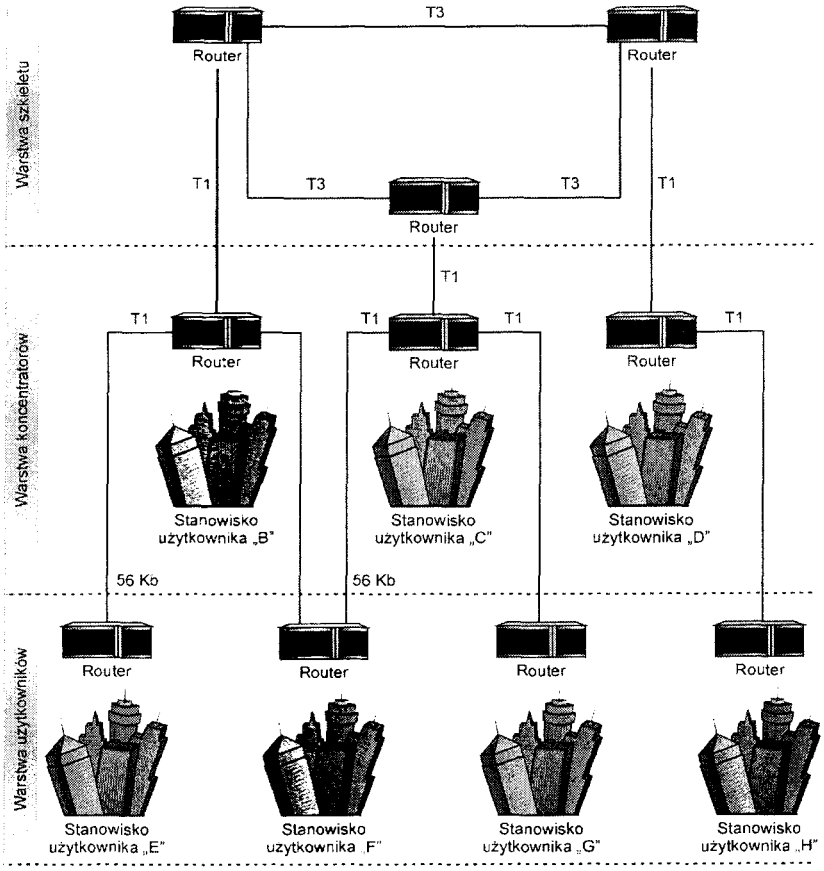


1.13.4.8 Topologie hybrydowe

Łączenie wielu topologii jest szczególnie przydatne w większych, bardziej złożonych sieciach. Pozwala to administratorom dostosować sieci WAN do istniejącego rozkładu obciążeń, zamiast wymuszać dopasowanie komunikacji do sztywnego modelu topologicznego. Innymi słowy, podstawowe topologie przedstawione w niniejszej części rozdziału są czymś więcej niż tylko szkolnymi modelami, mającymi na celu li tylko pobudzenie twórczego myślenia. Nie istnieją ograniczenia różnorodności topologii stosowanych w sieciach WAN. Skuteczność każdej topologii oraz późniejsze łączenie różnych technologii sieci WAN zależy bezpośrednio od danej sytuacji oraz wymagań dotyczących wydajności.

Tendencje do hybrydyzacji występują szczególnie w sieciach wielowarstwowych. Sieci WAN mogą być hybrydyzowane przez zastosowanie topologii oczek pełnych lub częściowych w warstwie routerów szkieletu, co jest opisane we wcześniejszej części rozdziału; w zasadzie trudno jest podać tę jedną „właściwą" lub jakąś „niewłaściwą" metodę konstruowania topologii hybrydowej. Jeden z przykładów sieci w topologii hybrydowej przedstawiony jest na rysunku 13.1 I; z braku miejsca pominięte zostały ikony budynków w warstwie szkieletu. Rysunek 13.11. Topologia hybrydowa.

Wielowarstwowa sieć WAN może posłużyć do utworzenia wydajnej topologii hybrydowej przez zorganizowanie topologii oczek pełnych tylko na warstwie szkieletu. Dzięki temu szkielet sieci staje się odporny na awarie, zapewniając przy okazji częściową minimalizację liczby skoków w całej sieci znaną z sieci o topologii oczek pełnych i jednoczesne uniknięcie kosztów oraz ograniczeń jej rozbudowy.



Połączenie szkieletu wielowarstwowej sieci WAN w topologię oczek pełnych jest tylko jedną z odmian topologii hybrydowej. Również inne hybrydy mogą być wysoce skuteczne. Kluczowym zagadnieniem jest wyszukanie topologii oraz podtopologii, które można łącznie wykorzystać w celu zaspokojenia określonych wymagań dotyczących sieci.

1.13.5 Projektowanie własnych sieci WAN

Projektowanie sieci WAN wymaga pomyślnego połączenia w całość technicznych składników opisanych w niniejszym rozdziale. „Pomyślne" połączenie w całość oznacza, że wydajność gotowej sieci odpowiada założeniom projektowym i oczekiwaniom użytkowników albo je przekracza. Dlatego tak ważne jest określenie i ocena kryteriów wydajności przed rozpoczęciem projektowania sieci.

1.13.5.1 **Kryteria oceny wydajności sieci** WAN

Jakość sieci WAN można ocenić, stosując wiele różnych kryteriów (metryk). Wiele z nich można uznać za obiektywne, a ich uzyskanie polega na bezpośredniej analizie protokołów monitorowania sieci, wbudowanych praktycznie w każde urządzenie sieciowe. Inne metryki należy uznać za mniej obiektywne lub niemal niemożliwe do wcześniejszego oszacowania. Niektóre z bardziej rozpowszechnionych metryk to:

• czas przydatności elementu, • natężenie ruchu,
• opóźnienia i czasy oczekiwania,
• szybkości wykorzystania zasobów.

Każda z tych metryk jest szczegółowo omówiona w dalszej części rozdziału.

1.13.5.1.1 Czas przydatności elementu

Każdy fizyczny składnik sieci WAN może być monitorowany, a jego dostępność mierzona za pomocą metryki zwanej przydatnością elementu i odzwierciedlającą przeciwieństwo czasu przestoju. „Przydatność elementu" określa stosunek czasu, przez jaki urządzenie pracuje i jest sprawne, do czasu dostępności wymaganego przez użytkownika. Częstą praktyką jest podawanie czasu przydatności elementu przy obciążeniu przez 7 dni w tygodniu przez 24 godziny na dobę, nawet jeśli użytkownik wymaga dostępności jedynie przez 5 dni w tygodniu po 12 godzin. Warto o tym pamiętać i odpowiednio dopasowywać metryki do zgłaszanych przez użytkowników wymagań co do wydajności sieci.

Wszystkie urządzenia elektroniczne, nawet te najbardziej niezawodne, ulegają kiedyś awarii. Większość producentów określa stopień niezawodności swego sprzętu za pomocą parametru MTBF (ang. Mean Time Between Failures - średni czas między awariami). Wskaźnik ten osiąga najczęściej wartość dziesiątek tysięcy godzin, co powinno oznaczać lata niezawodnej pracy; w rzeczywistości ów statystyczny optymizm korygowany jest bezlitośnie przez rozmaite czynniki eksploatacyjne, w szczególności:

• zakres średnich temperatur w środowisku pracy,
• „czystość" napięcia zasilającego,

• szeroko pojęty sposób odchodzenia się w urządzeniem - zarówno podczas jego przechowywania, jak w czasie rzeczywistej eksploatacji.
Innymi słowy, rzeczywisty „przebieg" nie jest stały! Monitorowanie i śledzenie czasu przydatności elementu poszczególnych urządzeń pozwala zaprezentować użytkownikom, jak dobrze są spełniane ich wymagania dotyczące dostępu do sieci.
Możliwe jest również śledzenie poza okresami pracy trendów w danych dotyczących czasu przydatności elementu, co pozwala odszukać te składniki infrastruktury sieci, które mogą być przyczyną problemów. Trendy te mogą dostarczyć informacji na temat ogólnej niezawodności określonego typu lub marki sprzętu, co następnie można wykorzystać do wskazania składników o podwyższonym ryzyku uszkodzenia.
Termin dostępność czasami jest wykorzystywany do ogólnego opisania łącznego czasu przydatności sieci. Nie jest to jednak najlepsza metryka. Teoretycznie dostępność sieci powinna być liczbowym odpowiednikiem gotowości sieci. W praktyce „dostępność" jest wskaźnikiem tak mglistym, że jego wartość w zasadzie nie ma większego znaczenia. Przykładowo: jeśli dany router ulegnie awarii, użytkownicy w jego lokacji nie mają dostępu do całej sieci. Sieć pozostaje jednak dostępna dla użytkowników ze wszystkich pozostałych lokalizacji. Nie mogą oni uzyskać dostępu do hostów z uszkodzonego obszaru, ale nic nie przeszkadza im w komunikowaniu się z innymi hostami w sieci. Stopień dostępności sieci zależy więc w znacznym stopniu od lokalizacji i wymagań użytkowania, dlatego też liczbowe wyrażanie dostępności sieci może być bardziej uciążliwe niż przydatne.

1.13.5.1.2 Natężenie ruchu

Jedną z ważniejszych metryk każdej sieci WAN jest spodziewane natężenie obsługiwanego ruchu. Natężenie to zmienia się w czasie, zależnie od cyklu pracy przedsiębiorstw, pór roku i innych czynników - tak więc, jak każdą zmieniającą się wielkość, mierzyć je można wieloma miaram statystycznymi, z których najważniejszymi są wartość średnia i wartość maksymalna:
• Maksymalne natężenie, jakie sieć ma obsługiwać nazywane jest również natęże niem .szczytowym. Jak wskazuje nazwa, jest to największe spodziewane nasilenie ruchu, jakie sieć musi być zdolna obsłużyć.
• Średnie natężenie to - po prostu - natężenie, przy jakim przyjdzie pracować sieci - a konkretnie: jej określonemu składnikowi (lokalizacji) - w typowych warunkach.
Określenie dwóch wspomnianych wartości natężenia ruchu ma podstawowe znaczenie przy dobieraniu szybkości urządzeń transmisyjnych sieci WAN, a także routerów. Jeśli na przykład oczekuje się, że w trakcie dnia roboczego dowolna lokacja będzie obciążała sieć ruchem o natężeniu 100 Kbps, to jest oczywiste, że urządzenie transmisyjne o przepustowości 56 Kbps nie jest dla niej wystarczająco szybkie.
Jeśli w sieci wykorzystana jest jedna ze złożonych topologii opisanych w poprzedniej części rozdziału, konieczne jest określenie sumarycznych schematów i natężeń ruchu obsługiwanego przez szkieletowe routery i urządzenia transmisyjne sieci.
Opóźnienie
Opóźnienie to jedna z częściej stosowanych metryk odzwierciedlających wydajność sieci. Odpowiada ona odcinkowi czasu oddzielającemu dwa zdarzenia. W przypadku wymiany informacji zdarzenia te z reguły oznaczają wysłanie i odebranie danych. Dlatego opóźnienie jest czasem potrzebnym na przesłanie w sieci pakietu z punktu źródłowego do docelowego. Przy takiej definicji opóźnienie jest zjawiskiem sumarycznym, uzależnionym od wielu czynników. Trzy najważniejsze czynniki to:
• Opóźnienia propagacji: termin ten odnosi się do łącznego czasu wymaganego na przesłanie (propagację) danych przez wszystkie urządzenia transmisyjne sieci znajdujące się na ścieżce transportu. Wydajność i ilość tych urządzeń transmisyjnych ma bezpośredni wpływ na sumaryczne opóźnienie każdej transmisji. Dodatkowym czynnikiem wpływającym na opóźnienie propagacji jest natężenie ruchu. Im bardziej obciążone jest dane urządzenie transmisyjne, tym mniejsza szerokość pasma dostępna dla nowej transmisji. Opóźnienia propagacji są zjawiskiem typowym dla instalacji naziemnych, niezależnie od tego, czy nośnikiem jest światłowód, czy przewód miedziany.
• Opóźnienia komunikacji satelitarnej: niektóre urządzenia transmisyjne bazują na łączności satelitarnej. Wymagają one przesłania sygnału do satelity i z powrotem na Ziemię. Z powodu potencjalnie ogromnych odległości między naziemnymi urządzeniami transmisyjnymi i satelitami opóźnienia te mogą być nawet całkiem spore.
• **Opóźnienia w przesyłaniu:** opóźnienie w przesyłaniu przez sieć jest łącznym czasem potrzebnym na odebranie, buforowanie, przetwarzanie i przesłanie da**nych** przez każde fizyczne urządzenie. Rzeczywiste opóźnienie w przesyłaniu każdego urządzenia może zmieniać się w czasie. Urządzenia, które pracują przy niemal maksymalnym obciążeniu, mają zwykle większe opóźnienia w przesyłaniu niż porównywalne urządzenia o małym obciążeniu. Ponadto wartości opóźnień mogą dodatkowo wzrastać z powodu zbyt dużego obciążenia lub błędów pracy sieci. Opóźnienia w przesyłaniu często nazywane są *czasem oczekiwania* poszczególnych składników.
Stopień wykorzystania zasobów
Stopień, w jakim wykorzystywane są różne fizyczne zasoby sieci WAN, jest coraz częściej stosowany jako wiarygodny wskaźnik ilustrujący, jak dobrze lub jak źle pracuje sieć w porównaniu z wymaganiami. Szczególnie uważnie powinny być obserwowane następujące dwie podstawowe kategorie wskaźników wykorzystania zasobów:
• wskaźniki zaangażowania procesora i pamięci routera,
• wskaźniki wykorzystania urządzeń transmisyjnych.

1.13.5.1.3 Zasoby routera

Routery należą do najważniejszych składników sieci WAN. W odróżnieniu od urządzeń transmisyjnych, znajdują się one poza sferą zainteresowań operatora telekomunikacyjnego. Zatem odpowiedzialność za ich działanie ponosi użytkownik. Na szczęście router jest urządzeniem inteligentnym, wyposażonym we własny procesor i pamięć. Zasoby te są niezastąpione przy obliczaniu tras w sieci WAN i przesyłaniu pakietów. Można je również wykorzystać do monitorowania wydajności routera.
Jeśli obciążenie procesora lub pamięci sięga 100%, ujemnie wpływa to na wydajność. Takie zwiększenie obciążenia, prowadzące do spadku wydajności, może być wywołane wieloma przyczynami. Jednym z przykładów może być nagły wzrost transmisji z sieci LAN do WAN. Sieci lókalne mogą pracować z szybkością do 1 Gbps (zwykle jest to szybkość 10, 16 lub 100 Mbps). Każda z tych szybkości jest znacznie większa od przepustowości typowych urządzeń transmisyjnych sieci WAN, zapewniających jedynie 1,544 Mbps szerokości pasma. Taka

różnica szerokości pasma musi być buforowana w pamięci routera. Każdy router może wyczerpać swoje zasoby, jeśli wystarczająco długo będzie obciążony transmisjami z sieci LAN.
Jeśli takie sytuacje zdarzają się rzadko, można je uznać za dopuszczalne odchylenia. Należy je monitorować, lecz nie zmuszają one do wymiany urządzeń. Jednak jeśli przeciążenie zasobów powtarza się lub staje się regułą, konieczne jest podjęcie czynności zaradczych. Zwykle wymagana jest zmiana routera na większy lub wyposażenie istniejącego w większą pamięć. Jeśli pamięć routera jest bez przerwy zajęta niemal w 100%, najwyższy czas, aby kupić dodatkową pamięć.
Rozwiązanie problemu przeciążenia procesora może nie być tak proste jak rozszerzenie pamięci. Praktycznie istnieją tylko dwa wyjścia z opisywanej sytuacji:
• wymiana procesora routera (lub - routera w ogóle) na nowy o większych możliwościach,
• zbadanie schematów ruchu w sieci WAN i sprawdzenie, czy możliwe jest zmniejszenie obciążenia danego routera.
Regulację ruchu w sieci można praktycznie przeprowadzić jedynie w dużych sieciach rozległych o złożonych topologiach zapewniających wielość tras. Jednak nawet w takich sieciach, jeśli problemy dotyczą routera w lokalizacji użytkownika (nie routera szkieletowego), jedynym środkiem zaradczym jest wymiana urządzenia na szybsze.

1.13.5.1.4 Stopień wykorzystania urządzeń transmisyjnych

Możliwe jest również monitorowanie wykorzystania urządzeń transmisyjnych. Wykorzystanie to zwykle jest wyrażane procentowym zużyciem szerokości pasma. Na przykład, jeśli stosowane jest łącze T-1, wykorzystanie jego 30% oznacza, że aktualnie jest wykorzystywana taka właśnie część dostępnego pasma o szerokości 1,544 Mbps.
Wskaźniki te mogą być trudne do przeanalizowania, a czasami mogą być wręcz mylące. Na przykład, często zdarza się, że oprogramowanie zarządzające siecią pobiera informacje na temat wykorzystania zasobów co pewien okres czasu. Może to być godzina, pięć minut lub dowolny inny czas. Jeśli częstotliwość próbkowania jest zbyt mała (w stosunku
do dynamiki zmian), krótkotrwałe wahania wykorzystania pasma mogą być gubione. Z kolei zbyt częste próbkowanie może doprowadzić do powstania ogromnej ilości nieistotnych informacji. Sztuka polega na dobraniu odpowiedniej częstotliwości, pozwalającej na zebranie istotnych danych na temat działania sieci w stosunku do oczekiwań użytkownika.
Poza samym wybraniem szybkości próbkowania pozostaje jeszcze zagadnienie okna próbkowania. Okno próbkowania jest czasem, w którym pobierane są próbki informacji. Ustalenie okna próbkowania polega na określeniu częstotliwości próbkowania oraz czasu trwania próbkowania. Okno próbkowania powinno być ustalone na podstawie wymagań użytkowników dotyczących dostępności sieci WAN. Jeśli próbkowanie wykorzystania sieci odbywa się 7 dni w tygodniu przez 24 godziny na dobę, a użytkownicy pracują jedynie przez 10 godzin dziennie i po 5 dni w tygodniu, zebrane dane statystyczne nie będą odzwierciedlały rzeczywistego stopnia zaspokojenia wymagań użytkowników.
Wskaźniki wykorzystania są doskonałym narzędziem statystycznym do monitorowania i mierzenia stanu urządzeń transmisyjnych. Nie są one jednak jedynymi metrykami wy dajności sieci. Sieć jest sprawna tylko wtedy, gdy spełnia oczekiwania użytkowników. Dlatego też lepszy obraz sprawności sieci dają perspektywiczne, złożone wskaźniki będące połączeniem różnych metryk wydajności.

1.13.5.2 Koszt sieci WAN

Ważnym elementem wszelkich kryteriów oceny wydajności jest koszt. Koszty posiadania i korzystania z sieci rozległej obejmują początkowe koszty budowy oraz miesięczne opłaty za jej eksploatację. Nie jest zatem niespodzianką, że składniki sieci o dużych możliwościach są znacznie droższe od elementów mniejszych i, co tu ukrywać - mniej pewnych. Dlatego też projektowanie sieci WAN jest zadaniem ekonomicznym, w którym osiąga się równowagę między kosztami i wydajnością.
Osiągnięcie tej równowagi może być bardzo trudne. Nikt nie chce projektować sieci WAN, której użytkownicy będą niezadowoleni, ale też nikt nie chce sieci, której koszt przekroczy zakładany budżet! Na szczęście szereg wskazówek może pomóc administratorom wybrać projekt sieci WAN, która zaspokoi bieżące wymagania, umożliwi przyszłą rozbudowę oraz zmieści się w zakładanym budżecie:
• Kapitał jest inwestowany w routery i inny sprzęt, który staje się stałą częścią sieci. Po uruchomieniu tych urządzeń, ich wymiana stwarza spore problemy. Ponadto, zależnie od planu amortyzacji, może się okazać, że korzystanie ze sprzętu powinno potrwać pięć lub więcej lat! Może to doprowadzić do kupienia większego lecz stosunkowo mało popularnego routera. Dodatkowy sprzęt (pamięć, procesory i interfejsy) można dodawać w późniejszym czasie, w miarę potrzeb. Rozwiązanie to umożliwia przyszły rozwój przy niewielkich kosztach dodatkowych oraz krótkim (lub wręcz zerowym) czasie nieaktywności sieci.
• Urządzenia transmisyjne można stosunkowo łatwo zastąpić innymi. Koszt ich użytkowania należy do kosztów stałych, nie do inwestycji, nie ma więc możliwości odpisów amortyzacyjnych. Wymianę urządzeń transmisyjnych można przeprowadzać tak często, jak na to pozwala umowa z operatorem telekomunikacyjnym. Zatem można sprawdzać możliwości zaspokojenia oczekiwań co do wydajności za pomocą różnych dostępnych urządzeń i technologii transmisyjnych.
Korzystanie z powyższych wskazówek może pomóc w zaspokajaniu bieżących i przyszłych wymagań użytkowników bez przekraczania budżetu.

1.13.6 Podsumowanie

Sieci rozległe są skomplikowanymi konstrukcjami, które nie zawsze przystają do opisanych w publikacjach czy otwartych standardów. Projektowanie, budowa i obsługa sieci WAN mogącej stale spełniać oczekiwania użytkowników może okazać się zadaniem niemal niewykonalnym. Sukces leży w zrozumieniu możliwości, ograniczeń i kosztów różnych technologii składowych sieci. Dzięki temu możliwa staje się ich poprawna integracja. Głównym celem jest zbudowanie sieci, w której każdy składnik jest dobrze dopasowany do możliwości pozostałych elementów i mieści się w ograniczeniach budżetowych.
Rozdział 14, zatytułowany „Linie dzierżawione", poświęcony jest szczegółowemu omówieniu linii dzierżawionych.

1.14 Rozdział 14 Linie dzierżawione

Mark A. Sportack

Sieci WAN są zwykle zbudowane z cyfrowych urządzeń transmisyjnych dzierżawionych od operatora telekomunikacyjnego. Urządzenia te udostępniają użytkownikom dedykowaną szerokość pasma dochodzącą do niemal 45 Mbps. Takie urządzenia transmisyjne noszą nazwę *linii d.-ierżawionych.* Są one integralną częścią każdej sieci WAN, choć wiedza na ich temat jest mało rozpowszechniona. Ich modułowe połączenia prowadzą do ogromnej, złożonej infrastruktury telekomunikacyjnej, w której wykorzystywany jest szereg standardów.

Owej infrastrukturze oraz obsługującym ją standardom mało kto - za wyjątkiem oczywiście samego operatora telekomunikacyjnego - poświęca jakąkolwiek uwagę. Do niedawna zresztą wiedza ta nie była też nikomu do niczego potrzebna. Z czasem jednak linie dzierżawione stały się integralną częścią współczesnych sieci WAN. Dlatego też warto je lepiej poznać.

W niniejszym rozdziale można znaleźć informacje na temat standardów linii dzierżawionych przyjętych w Europie i w Ameryce Północnej, a także na temat podstawowych zasad leżących u podstaw ich działania. Rozjaśnią one nieco aurę niejasności i tajemniczości otaczającą linie dzierżawione oraz umożliwią lepsze projektowanie, obsługę i usuwanie problemów związanych z sieciami WAN.

1.14.1 Przegląd linii dzierżawionych

Choć linie dzierżawione są powszechnie uważane za urządzenia transmisji cyfrowej dla sieci przesyłania danych, pierwotnie były one projektowane jako cyfrowe urządzenia do komunikacji głosowej. Przekształcanie sygnałów dźwiękowych na format cyfrowy ma wiele zalet. Należy do nich lepsze odtwarzanie osłabionych sygnałów oraz wydajniejsze sumowanie wielu strumieni danych we wspólnym nośniku transmisyjnym. Sumowanie to jest nazywane multipleksowaniem.

W miarę upływu lat połączenie następujących trzech czynników:

• powiększenie bazy cyfrowego sprzętu transmisyjnego w infrastrukturze operatorów telekomunikacyjnych,

• powstanie mechanizmów obsługi telefonii cyfrowej,

• pojawienie się zapotrzebowania klientów na wysokowydajną komunikację cyfrową,

wymusiło znaczne zmiany w sposobie korzystania z linii dzierżawionych.

Początkowo operatorzy telekomunikacyjni z wielką niechęcią dzierżawili takie linie klientom potrzebującym ich do skoncentrowanej komunikacji głosowej, np. dla central telefonicznych czy dużych biurowców z prywatną centralą rozdzielczą. W miarę rozwoju rozproszonych technologii sieciowych zaczęto jednak doceniać wartość linii dzierżawionych jako urządzeń transmisji danych. Przejście z cyfrowego przesyłania głosu do przesyłania danych cyfrowych było naturalne; liczby dwójkowe pozostają dwójkowymi niezależnie od tego, co reprezentują. Linie dzierżawione odniosły tak duży sukces jako środki wymiany danych, że ich pierwotne przeznaczenie stało się jedynie historyczną ciekawostką.

1.14.2 Techniki multipleksowania

Linie dzierżawione były pierwotnie projektowane z myślą o przesyłaniu wielu kanałów głosowych przez jedno urządzenie transmisyjne o większej pojemności. Wymagało to opracowania mechanizmu mogącego pobrać wiele nadchodzących strumieni komunikacji, sterować ich przesyłaniem przez wspólne urządzenie transmisyjne oraz rozdzielić je na pierwotne składniki w celu dostarczenia do poszczególnych punktów docelowych. Multipleksowanie pozwoliło na wydajniejsze pośrednie przełączanie rozmów. Istnieje kilka sposobów multipleksowania:

• dzielenie dostępnej szerokości pasma w czasie (tzw.. multipleksowanie czasowe); • dzielenie dostępnej szerokości pasma na podczęstotliwości (tzw. multiplekso wanie podziału c.-ęstotlivrości).

Istnieją również inne metody multipleksowania, np. multipleksowanie podziału spektrum, lecz techniki czasowe i częstotliwościowe są najbardziej przydatne w opisywanych liniach dzierżawionych

Obie wymienione metody multipleksowania są dokładniej opisane w następnych częściach rozdziału.

1.14.2.1 Multipleksowanie czasowe

Multipleksowanie czasowe polega na dzieleniu dostępnej szerokości pasma na odcinki czasu. W danym przedziale czasu urządzenie komunikacyjne zajmuje całą szerokość pasma. Następnie odcinki te są przydzielane na podstawie z góry określonego algorytmu. Złożoność systemu multipleksowania czasowego niech zilustruje fakt, iż długość odcinka w przypadku łączy T-1 (linia dzierżawiona o przepustowości 1,544 Mbps) wynosi 125 mikrosekund.

System multipleksowania czasowego ma wydajną konstrukcję, gdyż stacja transmisyjna ma dostęp do całego pasma. Jeśli stacja nie transmituje danych, jej odcinek czasu jest przydzielany innej stacji żądającej dostępu. W ten sposób maksymalizowane jest wykorzystanie pasma przenoszenia urządzenia transmisyjnego.

System taki wymaga narzutów koniecznych do utrzymania synchronizacji. W przypadku utraty synchronizacji transmitowane dane ulegają uszkodzeniu. Istnieją dwie metody synchronizowania transmisji przez łącza multipleksowane czasowo:

• dodawanie bitów,

• dodawanie kanałów.

Obie wymienione metody są skuteczne i wykorzystywane we współczesnych liniach dzierżawionych. W Europie jako standard stosowane jest dodawanie kanałów, a w Ameryce Północnej -- dodawanie bitów.

Multipleksowanie podziału częstotliwości

Systemem alternatywnym do multipleksowania czasowego jest multipleksowanie podziału częstotliwości. Multipleksowanie podziału częstotliwości jest techniką podziału dostępnej szerokości pasma na pasy pod-częstotliwości. Każdy z tych pasów jest dedykowany określonemu urządzeniu lub użytkownikowi. Każdy pas jest przypisany na stałe, przez co pozostaje niedostępny dla innych urządzeń, nawet

kiedy nie jest wykorzystywany! To ograniczenie wydajności jest dostatecznym powodem do wyraźnego preferowania multipleksowania czasowego nad multipleksowani podziału częstotliwości.

1.14.3 Cienie i blaski linii dzierżawionych

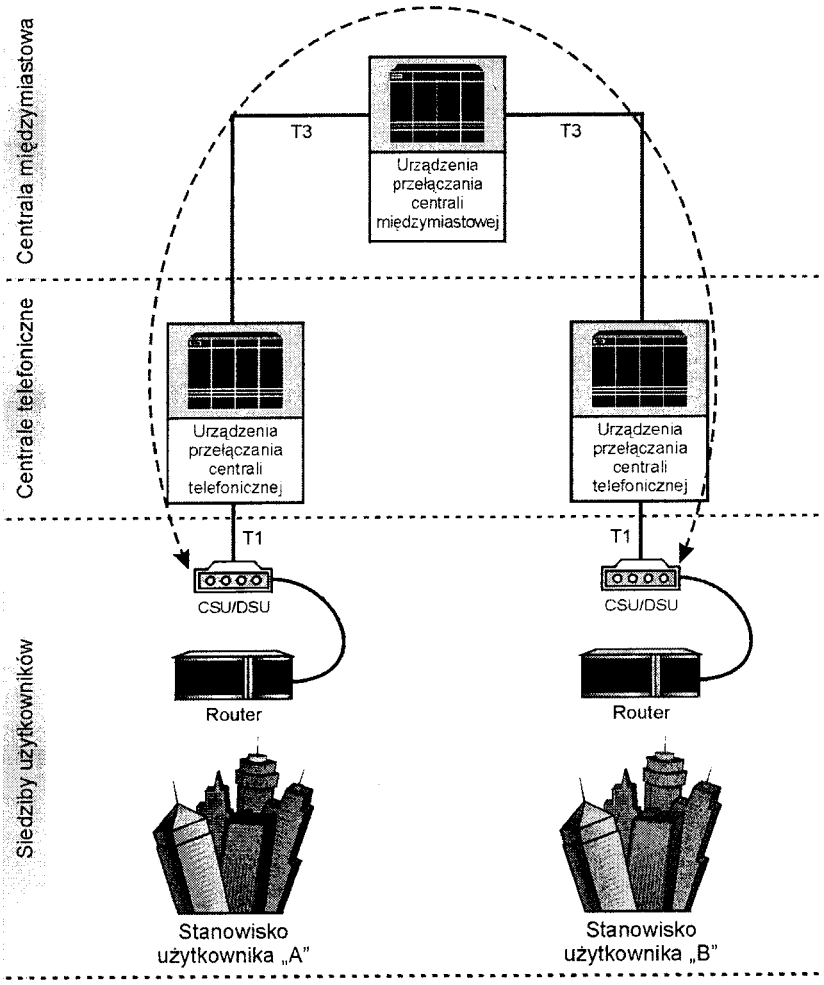
Jedną z bardziej rozpowszechnionych form linii dzierżawionych jest *prywatne połącze nie dwupunktowe.* Z funkcjonalnego punktu widzenia linie dzierżawione umożliwiają uzyskanie dedykowanego połączenia między dwiema stacjami przy minimalnym narzucie komunikacji lub przesyłania ramek. Jednak linie te nie tworzą bezpośredniego połączenia dwóch punktów.

Przynajmniej jeden operator telekomunikacyjny udostępniał rozwiązanie tańsze od linii dzierżawionych, które zapewniało bezpośrednie połączenie dwóch punktów. Usługa ta była droższa w budowie i instalacji, lecz miesięczne opłaty za korzystanie były znacznie mniejsze. Niedogodnością tego rozwiazania był brak możliwości zdalnej analizy problemów. Każde zgłoszenie awarii wymagało wysłania pracownika serwisu.

Rozwiązanie takie wydawać się może marnotrawieniem zasobów. I faktycznie, można przytoczyć argumenty na dowód tego, że jest ono znacznie mniej wydajne od bezpośredniego łączenia par stanowisk użytkowników. Jednak obniżenie wydajności tej metody jest równoważone możliwościami rozbudowy dostarczanych usług. Bezpośrednie łącza transmisyjne stają się bezwartościowe, gdy tylko zanika potrzeba wzajemnej komunikacji między połączonymi stanowiskami.

Centrale przełączania, co przedstawione jest na rysunku 14.1, działają podobnie jak główny szkielet infrastruktury oraz pośredni szkielet infrastruktury w systemie okablowania budynku, tworząc elastyczny system okablowania wielokrotnego użytku, dzięki któremu w dowolnym punkcie budynku istnieją takie same możliwości korzystania ze wspólnego sprzętu komunikacji głosowej i wymiany danych.

Rysunek 14.1. Linie dzierżawione nie sa połączeniami bezpośrednimi.



Połączenie stanowisk wielu użytkowników z najbliższymi centralami operatora telekomunikacyjnego daje w rezultacie infrastrukturę możliwą do ponownego wykorzystania. Niezależnie od zmian wymagań dotyczących komunikacji lub schematów natężenia ruchu, najważniejsze samo połączenie centrali z stanowiskiem użytkownika jest nadal użyteczne.

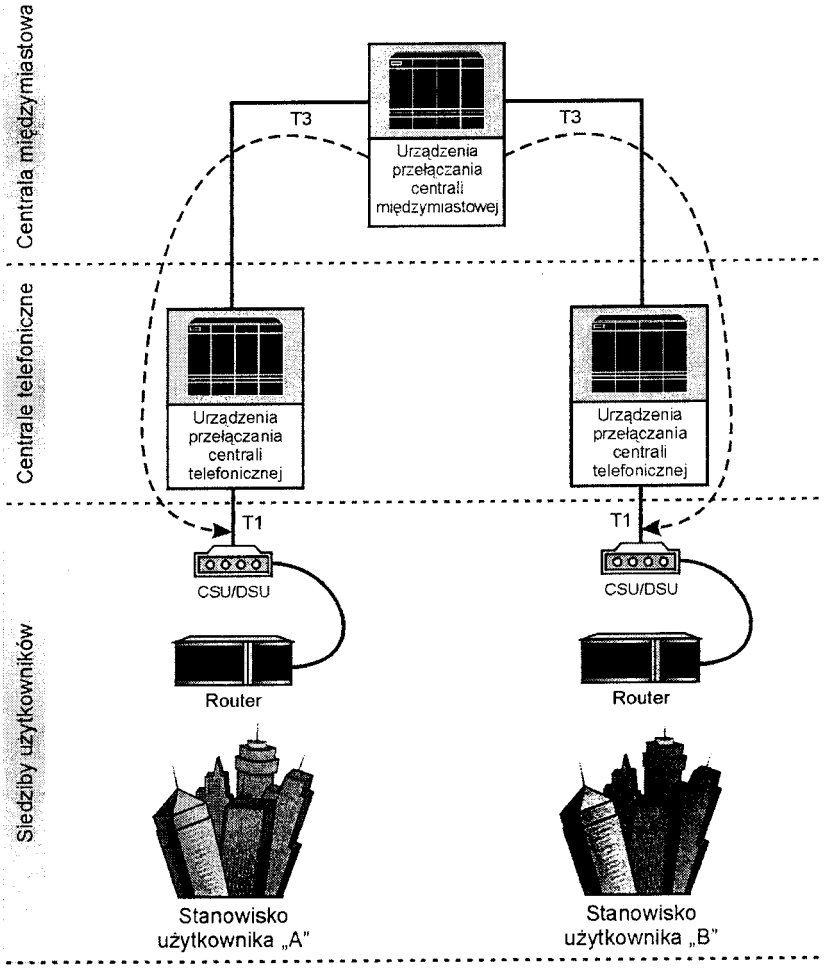
Dodatkową zaletą łączenia stanowiska za pośrednictwem central telekomunikacyjnych jest możliwość przeprowadzania zdalnej diagnostyki. Gdyby operator instalował okablowanie bezpośrednio między dwoma stanowiskami użytkownika, przeprowadzenie

nawet bieżącej analizy w przypadku usterek wymagałoby wysłania technika na miejsce awarii. Mając jedną lub więcej central na trasie linii dzierżawionej, możliwe jest wygodne przeprowadzanie przez operatora różnorodnych diagnoz w całym obwodzie połączenia. Zdalne diagnozowanie pozwala zaoszczędzić czas i koszt ponoszony w innym wypadku w celu odnalezienia usterki sieci.

Aby wykonać zdalną diagnozę, operator telekomunikacyjny próbuje nawiązać łączność z urządzeniami własnymi klientów na obu końcach połączenia, jak jest to pokazane na rysunku 14.2.

Rysunek 14.2. Ginie dzierżawione są tak projektowane aby timożliwiały zdalną diagnostykę.

Jeśli uda mu się połączyć z obydwoma zestawami urządzeń klientów (zwykle są to jednostki obsługi kanałów / obsługi danych), to można w miarę bezpiecznie założyć, że usterka nie dotyczy linii dzierżawionej. Jeśli operator nie może nawiązać łączności z jedną z owych jednostek, obszar awarii jest ograniczony do jednego segmentu linii transmisyjnej.



W każdej z opisanych sytuacji zdalna diagnostyka umożliwia znaczną redukcję czasu i kosztów usuwania problemów.

1.14.4 Topologia linii dzierżawionych

Rzeczywista topologia linii dzierżawionej może być nadspodziewanie zawiła. Schemat` linii dzierżawionych na rysunkach 14.1 i 14.2 zostały celowo uproszczone. W rzeczywistości niemal wszystkie linie dzierżawione - poza najkrótszymi - wymagają budowy obwodów biegnących przez co najmniej dwie centrale. Dokładny przebieg linii dzierżawionej zależy bezpośrednio od położenia geograficznego oraz liczby operatorów uczestniczących w zestawianiu połączenia. Na szczęście, dokładny kształt topologii linii dzierżawionej rzadko ma jakiekolwiek znaczenie.

W następnych częściach rozdziału opisane są dwa zagadnienia. Ich znajomość jest niezbędna do zrozumienia topologii linii dzierżawionych, jakimi są: centrale rozdzielcze oraz zdecentralizowana infrastruktura telefonii.

Centra przełączania

Wśród pojęć i terminów związanych z komunikacją istnieje bałagan będący skutkiem niepełnego ich zrozumienia oraz upływu czasu. Termin „centrala telefoniczna" (ang. CO-Central Office) nie jest tu wyjątkiem. Obecnie odnosi się on do wszystkich urządzeń przełączania obsługiwanych przez operatorów telekomunikacyjnych. W rzeczywistości centrale telefoniczne (znane również jako lokalne centrale telefovriczne) stanowią najniższą warstwę wielowarstwowej topologii przełączania, tworzącej komercyjną infrastrukturę telefoniczną. Urządzenia klienta są fizycznie przyłączone do infrastruktury telekomunikacyjnej operatora za pośrednictwem tej najniższej warstwy.

Dlatego też warstwa ta jest najbardziej widoczna dla użytkowników oraz personelu obsługującego komunikację zatrudnionego poza przemysłem telekomunikacyjnym. Zatem nie jest niespodzianką, że terminem „centrala telefoniczna" określane są wszystkie rodzaje telekomunikacyjnych centrów przełączania.

Pozostałe cztery warstwy zapewniają wielość tras w bardzo dużej i skomplikowanej topologii hybrydowej. Pięć warstw przełączania stanowią zatem:

• regionalne centra przełączania, • przekrojowe centra przełączania, • podstawowe centra przełączania, • centrale rozliczeń połączeń,
• centrale telefoniczne.

Warstwy te są ponumerowane stosownie do ich klasy. W USA przed podziałem firmy Bell System w styczniu 1984 roku komunikacja telefoniczna była oparta na centrach przełączania klasy 5: centra przełączania linii dzierżawionych. W takich warunkach lokalne rozmowy telefoniczne były praktycznie bezpłatne, ponieważ firma Bell System

dopłacała do nich, wykorzystując zyski z połączeń zamiejscowych. Obecnie taki hierarchiczny model jest bardziej celem teoretycznym niż ściśle przystającym do rzeczywistej architektury.
Centrale rozliczeń połączeń były odpowiedzialne za łączenie central znajdujących się w małej odległości od siebie. Dzięki temu lokalne połączenia telefoniczne mogły być nawiązywane pomimo różnych central źródłowych i docelowych. Centrale rozliczeń połączeń umożliwiały również łączenie z płatnymi, zamiejscowymi, obszarami infrastruktury telekomunikacyjnej.
Trzy najwyższe warstwy odpowiadają różnym stopniom podziału geograficznego, wykorzystywanego do naliczania opłat. Regionalne centra przełączania stanowiły szkielet całej sieci; w każdym regionie kraju znajdowało się centrum przełączania połączone innymi tego rodzaju centrami.
Przekrojowe i podstawowe centra przełączania znajdowały się na niższych poziomach agregacji geograficznej. Ta skomplikowana architektura została zaprojektowana w celu udostępnienia możliwie najwydajniejszej ścieżki między dowolną parą punktów końcowych sieci. Rozmiar i złożoność sieci pozwala na wykorzystywanie różnych tras połączeń, co łagodzi skutki ewentualnego przeciążenia lub awarii części sieci.
W Stanach Zjednoczonych znajduje się około 20 000 różnego rodzaju centrów przełączania. Ilustracje oddające tę topologię byłyby mało praktyczne. Równie nierealne byłoby odwzorowanie na ilustracji map topologicznych linii dzierżawionych biegnących przez owe centra przełączania.
Pięć klas centrów przełączania zostało zaprojektowanych tak, aby zapewnić obsługę dużej liczby abonentów, uproszczenie naliczania opłat oraz wielość tras w telekomunikacji Ameryki Północnej z czasów sprzed demonopolizacji. Choć pod tymi względami architektura ta sprawdziła się dobrze, demonopolizacja wymusiła wprowadzenie radykalnych zmian.

1.14.4.1 Infrastruktura telefonii po podziale rynku

Podział demonopolizacyjny jest nazwą nadaną złożonemu procesowi prawnemu, przeprowadzonemu w celu rozbicia firmy Bell System na w pełni niezależne jednostki. Dziś każda organizacja obłożona jest poważnymi ograniczeniami mającymi na celu stymulowanie konkurencji oraz innowacji technologicznych, takimi jak np.:
• ograniczenia obszaru geograficznego, na którym jednostka może działać, • ograniczenia rodzaju usług, które może oferować każda nowa organizacja.
Łącznie ograniczenia te spowodowały znaczne podzielenie infrastruktury telekomunikacyjnej i przyczyniły się do powstania niektórych nowych terminów i pojęć. Najbardziej wyraźne podziały powstały między komunikacją lokalną i zamiejscową, czyli regionalną i ponadregionalną. Operatorom oferującym połączenia lokalne nie wolno prowadzić komunikacji ponadregionalnej. Sfera ich działania ograniczona jest do obszaru regionu.
Operatorzy, tacy jak AT&T, ograniczeni do komunikacji dalekosiężnej, znani są jako *operatory ponadregionalni* (ang. *interexchange carriers - IXC)*. Nazwa ta pochodzi od terminu „region", który przed demonopolizacją stanowił obszar wolny od opłat. W obecnym środowisku cały ruch między obszarami lokalnymi musi odbywać się za pośrednictwem operatorów ponadregionalnych.
Amerykańskie prawo wymaga, aby operatorzy lokalni zapewniali wszystkim operatorom ponadregionalnym taki sam dostęp do swoich zasobów. Idea ta, znana jako „równy dostęp", wymagała powstania *punktów obecności* (POP). POP jest formą obecności operatorów ponadregionalnych w ramach centrów przełączania będących własnością operatorów LEC. Punkty POP zapewniają, że cały ruch między regionami LATA jest obsługiwany przez operatorów ponadregionalnych, czyli wszelki ruch adresowany do punktu docelowego w innym obszarze LATA musi być kierowany do punktów POP operatora ponadregionalnego, nawet jeśli początkowy i docelowy obszar LATA jest obsługiwany przez tego samego operatora LEC. W ten sposób sieci komunikacji lokalnej pozostają całkowicie oddzielone od zamiejscowych, z punktami POP pełniącymi rolę granicy między nimi.
Sam punkt POP może być tak prostym urządzeniem jak panel przyłączeniowy przewodów czy światłowodów, łączący urządzenia przełączające operatora LEC z urządzeniami operatora ponadregionalnego, bądź też może to być komutator telekomunikacyjny.
Tak więc po podziale demonopolizacyjnym w sieci nadal istnieje pięć klas przełączania, z tą różnicą, że obecnie jest to niejednolita struktura prywatnych sieci, obsługiwanych przez różnych lokalnych i ponadregionalnych operatorów. Układ taki w naturalny sposób zmusza do współpracy i konkurencji. Taka infrastruktura telefonii jest wykorzystywana do obsługi wszystkich połączeń telefonicznych na terenie Stanów Zjednoczonych, włącznie z liniami dzierżawionymi.

1.14.5 Standardy sygnałów cyfrowych

Podobnie jak każdy inny standard technologii sieciowej, również linie dzierżawione umożliwiają przesyłanie danych w sposób zgodny ze standardowymi schematami transmisyjnymi. Schematy te określają szybkości transmisji i rodzaje nośników, a także formaty ramek i metody multipleksowania.
Istnieje wiele takich schematów, różniących się technologią i obszarami ich zastosowań. Wśród bardziej rozpowszechnionych standardów można wymienić:
• hierarchię ANSI sygnału cyfrowego, • hierarchię ITU sygnału cyfrowego,
• system optycznych nośników SONET,
• system synchronicznego sygnału transportowego SONET.
Standardy te są omówione w pozostałej części podrozdziału.

1.14.5.1 Hierarchia ANSI sygnału cyfrowego

We wczesnych latach 80. organizacja ANSI (American National Standards In.stitute) ustanowiła standardy transmisji sygnałów cyfrowych. Rodzina powstałych standardów znana jest jako DSH (ang. Digital Signal Hierarchy - hierarchia sygnału cyfrowego). W skład tej hierarchii

wchodzi pięć specyfikacji, oznaczonych jako DS-0 = DS-4. W tabeli i4.1 wymienione są szerokości pasma i liczba obsługiwanych kanałów głosowych każdej z tych specyfikacji.
Tabela 14.1.
Standardy sygnałów cyfrowych ANSI
Standard DS-0 określa minimalną szerokość pasma wymaganą do przesłania głosu w postaci cyfrowej. Zakładając, że komunikacja głosowa odbywa się przy wykorzystaniu obwodów o paśmie 8 kHz, a dla cyfrowej postaci sygnału wymagane jest ośmiokrotnie większe pasmo, szerokość pasma DS-0 została ustalona na 64 Kbps. Nie ten standard jest jednak podstawą linii dzierżawionych, a pełni ją standard DS-1. DS-0 opisuje jedynie parametry kanału głosowego, wydzielonego z innych specyfikacji sygnałów cyfrowych.

Standard sygnału cyfrowego	Szerokość pasma	Liczba kanałów głosowych
DS-0	64 Kbps	1
DS-1	1,544 Mbps	24
DS-1C	3,152 Mbps	48
DS-2	6,312 Mbps	96
DS-3	44,736 Mbps	672
DS-4	274,176 Mbps	4032

Przyjęte przez ANSI standardy DS zostały zastosowane w systemach telefonicznych pod nazwą .systemu T-Carrier. Istnieje bezpośrednie powiązanie między obwodami T-Carrier i ich odpowiednikami standardu DS. Na przykład, standard DS-1 jest reprezentowany przez obwody transmisyjne T-I. T-3 to fizyczna realizacja standardu DS-3. Dodatkowe informacje na temat systemu T-Carrier można znaleźć w podrozdziale zatytułowanym „System T-Carrier", w dalszej części rozdziału.
Nawet po bardzo pobieżnym przejrzeniu tabeli 14.1 łatwo można rozpoznać przynajmniej jeden ze standardów. Linie T-1 stały się tak powszechne, że niemal każda osoba mająca styczność z techniką sieciową rozpozna szerokość pasma 1,544 Mbps i powiąże ją z tym typem urządzeń transmisyjnych. Niektóre osoby mogą rozpoznać nawet szerokość pasma 44,736 Mbps linii T-3. Pozostałe standardy mogą nie być tak łatwo rozpoznawane.
Standard DS-IC, z funkcjonalnego punktu widzenia, jest połączeniem dwóch sygnałów DS-1, przesyłanych przez wspólne urządzenie transmisji. Standardy DS-2 i DS-4 nigdy nie zdobyły większej popularności. Główną tego przyczyną był niekorzystny stosunek
ceny do wydajności. Usługi w standardzie DS-3 również są drogie, lecz cechują się lepszą od innych relacją ceny do pasma przenoszenia.
Obecnie niemal zawsze, gdy wymagane jest pasmo przekraczające możliwości standardu DS-3, stosowana jest technologia SONET (opisana w dalszej części rozdziału). Dlatego też z systemu T-Carrier wykorzystywane są jedynie linie T-1, częściowa T-1 oraz T-3. Pozostałe specyfikacje T-n przestały być używane.

1.14.5.2 Hierarchia ITU sygnału cyfrowego

Hierarchia sygnałów cyfrowych przedstawiona w tabeli 14.1 została przyjęta przez organizację ANSI jako standard. Jako że ANSI jest organizacją amerykańską, jej standardy mogą nie być akceptowane w innych częściach świata. Na przykład w Europie organizacja ITU (dawniej CCITT) utworzyła własną rodzinę standardów sygnałów cyfrowych. Ich nazwy pochodzą od nazwy komitetu, który zalecił je organizacji ITU: Conference of European Posts and Telecommunications Administration (CEPT). Owe standardy CEPT są wyszczególnione w tabeli 14.2.
Tabela 14.2.
Standardy sygnałów cyfrowych ITU .

Standard sygnałów cyfrowych	Szerokość pasma	Liczba kanałów głosowych
CEPT-1	2,048 Mbps	30
CEPT-2	8,448 Mbps	120
CEPT-3	34,368 Mbps	480
CEPT-4	139,264 Mbps	1 920
CEPT-5	565,148 Mbps	7 680

Choć w standardach ITU wykorzystywane jest pasmo podstawowe o takiej samej co DS-0 szerokości 64 Kbps, stosowane są zupełnie inne wielokrotności tego pasma. Dlatego europejska wersja linii T-I znana jest jako E-1 i ma przepustowość 2,048 Mbps zamiast 1,544 Mbps. Połączenia sieciowe między Europą i Ameryką Północną zawsze sprawiały problemy związane z różnymi standardami, zwykle przejawiające się występowaniem kanałów niezdatnych do wykorzystania.
Po wykonaniu prostych obliczeń łatwo zauważyć, że standard CEPT-1 (stosowany w europejskich urządzeniach transmisyjnych E-1) ma w rzeczywistości szerokość pasma wystarczającą na 32 kanały po 64 Kbps każdy. Standard CEPT-1 i obwody E1 obsługują maksymalnie 30 możliwych do wykorzystania kanałów. Pozostałe dwa kanały są zarezerwowane dla celów synchronizacji i sygnalizowania. Jest to podejście odmienne od przyjętego przez ANSI w specyfikacjach DS. ANSI narzuca umieszczenie impulsów czasowych i ramek w każdym kanale, redukując w ten sposób dostępną szerokość pasma w stosunku do podawanej szybkości transmisji.

1.14.6 Systemy nośników SONET

SONET jest akronimem słów *Synchronous Optical NETwork* (Synchroniczna sieć optyczna). W istocie jest to szereg systemów transmisyjnych opartych na technologii optycznej. Systemy te wykorzystują zestaw wysoko specjalizowanych technologii, zaprojektowanych specjalnie do celów telekomunikacyjnych. Powstały one z myślą o zapewnieniu współpracy między systemami przełączania różnych producentów oraz buforowaniu różnic pojemności i szybkości transmisji wielu różnorodnych systemów.

Aby wywiązać się z tego zadania, opracowane zostały fizyczne interfejsy, konwencje przesyłania ramek oraz dwie rodziny standardów sygnalizacji. Standardy te umożliwiają komunikację z szybkością od 51,84 Mbps do 2,48 Gbps.

SONET jest standardem opracowanym przez ANSI, który został przyjęty, lecz nie w całości zaadoptowany, przez organizację ITU. Wersja przyjęta przez ITU różni się szczegółami mającymi istotne znaczenie i nosi nazwę synchronicznej hierarchii cyfrowej (ang. SDH - Synchronous Digital Hierarchy). Podstawowa szybkość transmisji w systemie SDH to 155,52 Mbps przy podstawowej szybkości 51,84 Mbps stosowanej w systemie SONET.

Inną, choć już stosunkowo niewielką różnicą jest fakt, że standardy przesyłania sygnałów przez nośniki miedziane, które w standardzie SONET nazywane są sygnałami transportu synchronicznego, w systemie Synchronicznej hierarchii cyfrowej znane są pod nazwą modułów transportu.

Standard SONET obsługuje dwa systemy transmisji, opisane w następujących częściach rozdziału:

• System nośników optycznych (OC).

• System sygnałów transportu synchronicznego (STS).

1.14.6.1 System nośników optycznych

Ostatnią grupą standardów w niniejszym przeglądzie jest system OC oraz związany z nim, oparty na nośnikach miedzianych, system STS. W tabeli 14.3 przedstawione są standardy OC, ich szerokość pasma oraz liczba mieszczących się w nich kanałów DS-0 i DS-1.

Teoretycznie tabela ta może być powiększana niemal w nieskończoność, przez dalsze mnożenie podstawowego pasma OC-1. W miarę udoskonalania technologii sygnałowej, oczekiwanie włączenia do standardu coraz większych wielokrotności pasma OC-1 wydaje się uzasadnione. W obecnej chwili istnieją już urządzenia umożliwiające transmisję w paśmie OC-192! Również prawdopodobne jest, że wyższe wielokrotności zostaną dopasowane do standardu SONET i SDH.

Tablica 14.3.

Szerokości pasma systemu OC

System sygnałów transportu synchronicznego

Linia nośnika optycznego	Szerokość pasma	Liczba kanałów DS-0	Liczba kanałów DS-1
OC-1	51,84 Mbps	672	28
OC-3	155,52 Mbps	2 016	84
OC-9	466,56 Mbps	6 048	252
OC-12	622,08 Mbps	8 064	336
OC-18	933,12 Mbps	12 096	504
OC-24	1,244 Gbps	16 128	672
OC-36	1,866 Gbps	24 192	1 008
OC-48	2,488 Gbps	32 256	1 344

Szybkości OC systemu SONET można również stosować w systemach sygnałów elektrycznych, korzystających z miedzianego okablowania. Tyle że zamiast wyrażać szybkości przesyłania sygnałów za pomocą jednostek OC nośnika optycznego, tu oznacza się je jednostkami STS synchronicznego przesyłania sygnałów (ang. Synchronous Transport Signa!). Mimo różnicy rodzaju nośnika zachowana jest relacja jeden do jednego między tymi standardami. Innymi słowy. standard STS-1 jest bezpośrednim odpowiednikiem standardu OC-1; STS-3 ma szybkość 155.52 Mbps standardu OC-3, i tak dalej.

Warto pamiętać, że standardy STS-n osiągnęły poziom STS-48, przy którym szerokość pasma wynosi 2,488 Gbps. Osiągnięcie takiej szybkości przy wykorzystaniu miedzianego nośnika sygnałów elektrycznych stwarza poważne problemy techniczne. Już powyżej szybkości 155.52 Mbps przesyłanie sygnałów na odległości większe niż 100 metrów jest mocno utrudnione. Dlatego też jedynie standardy STS-1 i STS-3 można uznać za zdatne do użycia. Pozostałe standardy pozostają raczej w sferze teoretycznych dywagacji.

1.14.7 System T-Carrier

Standardy sygnałów cyfrowych wymienione w tabeli 14.1 są stosowane w fizycznych systemach nośników. Najbardziej rozpowszechniony w Ameryce Północnej jest system T-Carrier. Linie dzierżawione obsługiwane w tym systemie oznaczane są za pomocą litery „T". Na przykład, linia dzierżawiona dostępna w systemie T-Carrier, która spełnia parametry standardu DS-1, nazywana jest linią T-1. Pojęcia te są często mylone, nawet przez osoby najlepiej obeznane z techniką! Skutkiem tego terminy DS-n i T-n (gdzie n jest określoną liczbą) są często używane zamiennie, co jest błędem. Terminy te nie są wymienne. DS-n jest nazwą standardu, a T-n - obwodu (jednego z wielu typów) spełniającego wymagania tego standardu.

System T-Carrier pierwotnie był projektowany do obsługi multipleksowanej komunikacji głosowej przez pojedyncze urządzenia transmisyjne. Urządzenia te były stosowane do transportu rozmów telefonicznych między różnymi centralami. Jeden obwód T-1 może służyć do przesyłania 24 rozmów telefonicznych jednocześnie. Ponieważ głos w tych rozmowach jest przesyłany w postaci cyfrowej, na drodze do punktu docelowego można go wzmocnić i zregenerować. Dlatego głos w postaci cyfrowej ma czystsze brzmienie niż dźwięk analogowy, którego jakości nie można poprawić.
Przekształcenie sygnałów dźwiękowych w postać cyfrową sprawia pewne trudności techniczne. System T-Carrier, w porównaniu z poprzednimi rozwiązaniami, ma szereg udoskonaleń. Zmiany te opracowano z myślą o poprawieniu jakości transmisji cyfrowego dźwięku. W ten sposób system T-Carrier stał się lepszym mechanizmem transportu cyfrowych transmisji, bez preferowania komunikacji głosowej czy wymiany danych. Poprawki te zostały wprowadzone w technikach kodowania sekwencyjnego i w formatach ramek.

1.14.7.1 Usługi T-Carrier

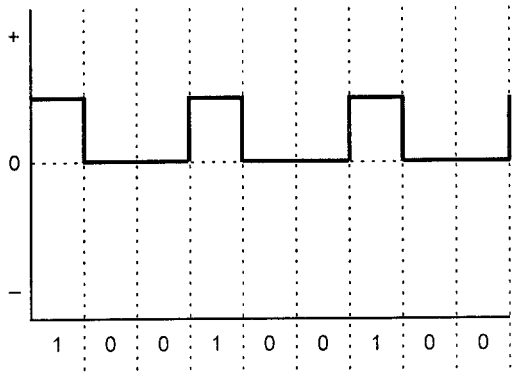
Jak pamiętamy z wcześniejszej części rozdziału, zatytułowanej „Hierarchia sygnału cyfrowego ANSI", jedynie dwa standardy DS dostępne są jako usługi komercyjne. Są to linie T-I i T-3. Dostępna jest również trzecia usługa - Częściowe linie T-I. Jest to odmiana specyfikacji DS-1, nie pochodząca od DS-0.
Urządzenie transmisyjne T-3 ma pasmo o szerokości 44,736 Mbps. Pasmo to można podzielić na 672 oddzielne kanały po 64 Kbps każdy lub 28 kanałów odpowiadających szybkością łączu T-1. Linie T-3 z reguły są bardzo drogie, zwłaszcza w przypadku dużych dystansów. W przypadku aplikacji sieciowych wymagających pasma szerszego niż jest dostępne przez parę linii T-1, linia T-3 może okazać się tańszym rozwiązaniem.
Urządzenia transmisyjne T-1 są fundamentem systemu T-Carrier. Ta podstawowa usługa udostępnia użytkownikom pasmo o szerokości 1,544 Mbps. Pasmo to można podzielić na 24 kanały o szerokości 64 Kbps lub całe pozostawić dla aplikacji sieciowych wymagających dużej przepustowości.
Jeśli w sieci dostęp do całego pasma T-1 jest zbyt drogi lub nie jest potrzebny, bardziej atrakcyjna może okazać się usługa częściowej linii T-I. Usługa ta jest skonstruowana na podstawie obwodu T-1. Pasmo tego obwodu jest dzielone na części za pomocą multipleksera kanałów (ang. channel bank). Obwód podrzędny najczęściej jest obwodem o przepustowości 56 Kbps, lecz można stosować nawet obwody 9,6 Kbps. Takie kanały o zmniejszonej szybkości można następnie udostępniać różnym klientom.
Dzielenie kanału T-1 wymaga zachowania zależności czasowych dla każdego podkanału. Jedna z technik nosi nazwę kradzieży bitów. Kradzież bitów pozwala utrzymać synchronizację kanału przez wykorzystanie najmniej znaczącego bitu w co ósmej ramce strumienia jako impulsu. Zabieranie tych bitów powoduje, że cała ramka z kradzionym bitem nie nadaje się do wykorzystania. Powoduje to zmniejszenie przepustowości kanału DS-0 z dostępnych 64 Kbps do jedynie 56 Kbps. U operatora można zamówić linie z czystym kanałem 64 Kbps. Urządzenia te to po prostu kanały DS-0 wydzielone z linii T-1, nie posiadające własnych mechanizmów synchronizacji.
Ramki, o których tu mowa, znacznie różnią się od ramek wykorzystywanych w zaawansowanych i wielofunkcyjnych sieciach LAN. Ramki w systemie DS są bardzo oszczędne. Ich funkcje są ograniczone do synchronizacji transmisji oraz, w niektórych przypadkach, wykrywania błędów i monitorowania sieci. Dlatego ramka w sygnale cyfrowym to najczęściej pojedynczy bit wstawiany do strumienia danych co ustalony okres czasu.

1.14.7.2 Kodowanie sygnału

Ponieważ T-Carrier był pierwszym systemem cyfrowym, wymagał zupełnie nowej techniki kodowania sygnału. Jednym z problemów w opracowaniu tej techniki było zachowanie synchronizacji linii przy możliwie najmniejszym narzucie.

1.14.7.2.1 Jednobiegunowe kodowanie binarne

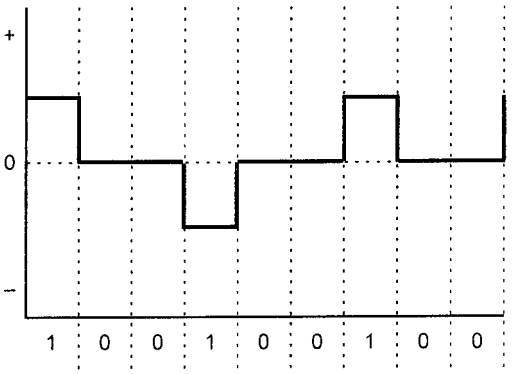
Najprostszą metodą kodowania cyfrowego jest kodowanie jednobiegunowe. Innymi słowy, albo sygnał reprezentuje stan włączony (polaryzacja dodatnia), albo wyłączony (brak polaryzacji). Choć jest to prosta technika, jej poważną wadą jest fakt, iż ciąg kolejnych zer lub jedynek nie powoduje zmian sygnału. Sytuacja ta jest zilustrowana na rysunku 14.3.
Rysunek 14.3. Jednobiegunowe kodowanie danych binarnych.
Przy typowym osłabieniu sygnału przesyłanego na duże odległości można stosunkowo łatwo utracić sygnały synchronizacji lub bity danych. Może to niekorzystnie wpłynąć na zakodowane informacje. Dlatego kodowanie jednobiegunowe nie nadaje się zbytnio do transmisji długodystansowych.
Dwubiegunowe kodowanie binarne
Alternatywą kodowania jednobiegunowego jest kodowanie dwubiegunowe. W kodowaniu dwubiegunowym, jak wskazuje nazwa, wykorzystywane są obie polaryzacje sygnału elektrycznego. Dzięki temu między kolejnymi bitami sygnał musi przejść przez zero



(brak napięcia). Przejście to jest ogólnie nazywane powrotem do zera. W ten sposób zapewnia się rozdzielenie transmitowanych bitów.
W systemie T-Carrier wykorzystywana jest metoda kodowania dwubiegunowego o nazwie inwersja znaku zmiennego (AMI). W metodzie AMI jedynki są reprezentowane za pomocą obu biegunów napięcia, a zera - brakiem napięcia. Ciąg bitów z rysunku 14.3 jest przedstawiony na rysunku 14.4 w postaci kodowania dwubiegunowego.
Rysunek 14.4. Dwubiegunowe kodowanie danych binarnych J
Choć w tej metodzie problem kodowania kolejnych jedynek został rozwiązany, system T-Carrier nadal jest podatny na utratę zgodności czasowej w przypadku ciągu kolejnych zer. Aby temu przeciwdziałać, w systemie T-Carrier zastosowano protokół znany jako *zasada gęstości jedynkowej.* Zasada ta ogranicza liczbę kolejno przesyłanych zer do piętnastu. Aby wymusić przestrzeganie tej zasady, została opracowana technika, powodująca zastąpienie ośmiu kolejnych zer za pomocą z góry określonego wzoru. Wzór ten nosi nazwę *słowa.* Technika została ochrzczona (niezbyt zgrabną) nazwą *bipolarnej substytucji o śmiozerowej,* w skrócie B8ZS.
Oczywiste jest, że zarezerwowanie dowolnego ciągu ośmiu bitów jako zamiennika zer powoduje, że ciąg ten może być przypadkowo odczytany ze strumienia bitów. Spowodowałoby to przekształcenie ośmiu pełnoprawnych bitów w zera. Dlatego w celu odróżnienia umownego słowa reprezentującego osiem zer od normalnych bitów informacji naruszona została zasada zmienności dwubiegunowej. Aby zasygnalizować początek umownego słowa odpowiadającego ośmiu zerom, przesyłane są dwie jedynki o dodatniej polaryzacji napięcia. Sytuacja ta jest zilustrowana na rysunku 14.5.
Technika BBZS jest metodą utrzymania zależności czasowych, wykorzystywaną w liniach T-1. Jest ona stosowana przez jednostki obsługi kanałów / obsługi danych znajdujące się na obu końcach linii dzierżawionej.

1.14.7.3 Formaty ramek

W systemie T-Carrier stosowane są trzy różne postaci ramek:
• format D-4,



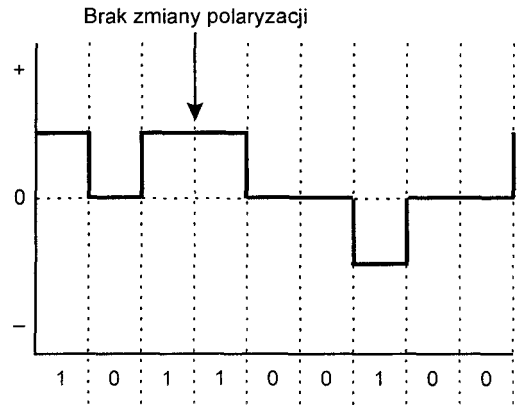
Rysunek 14.5. Brak zmiany polaryzacji sygnalizuje początek amownego słowa reprezentującego osiem zer.
• format ramek rozszerzonych SuperFrame (ESF),
• format M I -3.

1.14.7.3.1 Format D-4

Najbardziej rozpowszechnionymi ramkami w obwodach T-1 są ramki D-4, nazwane tak od kanałów D-4, w których je wykorzystano. W technice tej bit synchronizacji umieszczany jest przed co 24. oktetem danych. Jeśli liczba 24 wydaje się znajoma, to dlatego, że jest to liczba 8-bitowych kanałów przesyłanych przez linie T-1. Dlatego jedna ramka D-4 składa się z:
• 12-bitowego sygnału wyrównania ramek,
• I-bitowego sygnału synchronizacji, • 192-bitowego pola danych.
192-bitowe pole danych składa się oktetów, po jednym z każdego kanału w linii T-I. Po wykonaniu obliczeń można zauważyć, że ramka D-4 ma do 205 bitów. Jednak powszechnie przyjmuje się, że ramka D-4 ma 193 bity długości. Owa rozbieżność wynika z faktu, że sygnał wyrównania ramek, podobnie jak nagłówka w sieci Ethernet, nie jest traktowany jako część ramki.

1.14.7.3.2 Format ESF

Firma AT&T, autor wielu technologii i standardów telefonii stosowanych w Ameryce Północnej, opracowała we wczesnych latach 80. format ramek rozszerzonych ESF. Format ten wprowadzono w kanałach D-5, następcy standardu D-4.
Format ESF został zaprojektowany w celu dalszego poprawienia wydajności urządzeń transmisyjnych T-1. Istotą tej techniki jest zmniejszenie liczby bitów wyznaczających ramki znane z metody D-4. Jeśli w przypadku 24 ramek D-4 wymagane było użycie 24 bitów synchronizacji, ramki ESF wymagają jedynie sześciu. Co czwarta 193-bitowa ramka ma bit synchronizacji. Sześć innych bitów wykorzystuje się do korekcji błędów, a pozostałych 12 umożliwia monitorowanie sieci bez zakłócania jej pracy. Oznacza to, że monitorowanie nie powoduje zwiększenia narzutu w strumieniu bitów.



1.14.7.3.3 Format M1-3

W urządzeniach T-3 wykorzystywana jest zupełnie inna technika ramek, nazywana M 1-3. Nazwa tej techniki pochodzi od jej zasady działania: multipleksowania kanałów DS-1 w formacie DS-3. Teoretycznie 28 kanałów DS-1 wymaga 43,232 Mbps sumarycznej szerokości pasma. Wartość ta wraz z marginesem na wyrównanie ramek; wykrywanie błędów i synchronizację, dobrze pasuje do 44,736 Mbps standardu DS-3.
Format M1-3 należy do rodziny trzech standardów interfejsów multipleksowanych, znanych pod wspólną nazwą MX-3. Dwa pozostałe standardy to MC-3 i M2-3.
Standard MC-3 umożliwia multipleksowanie 14 kanałów DS-1C w jednym kanale DS.-3. Z kolei M2-3, jak wskazuje nazwa, służy do multipleksowania siedmiu kanałów DS-2 w jednym DS-3. Skoro standardy DS-1C i DS-2 nie zdobyły popularności, podobny los spotkał ich multipleksowane interfejsy.
Ramka formatu M 1-3 ma 4 760 bitów, z których 4 704 służą do przesyłania danych. Pozostałych 56 bitów służy do korekcji błędów i synchronizacji.

1.14.8 Podsumowanie

Linie dzierżawione są bardzo ważnym składnikiem współczesnych sieci WAN, lecz wiedza na ich temat jest skomplikowana i mało rozpowszechniona. Podobnie jak w przypadku innych technologii sieciowych, linie dzierżawione mają swoje technologie warstwy fizycznej i protokoły warstwy łącza danych oraz struktury ramek. Znając te aspekty linii dzierżawionych oraz ich równie skomplikowane topologie, można nabrać większej biegłości w projektowaniu, budowie, obsłudze i usuwaniu problemów w sieciach WAN.

1.15 Rozdział 15 Urządzenia transmisji w sieciach z komutacją obwodów

Tony Northrup
Wszyscy niemal spotykamy się codziennie z urządzeniami komutowanej transmisji obwodów, jednak większość z nas nawet o tym nie wie. Z samej natury sieci z komutacją obwodów wynika niezawodność i pewność przeprowadzanych przez nie transmisji danych. Ten wysoki poziom niezawodności sprawia, że z sieci z komutacją obwodów korzystają tak popularne technologie, jak Frame Relay oraz ATM.
Najbardziej rozpowszechniona sieć na świecie, czyli publiczna komutowana sieć telefoniczna, opiera się na zasadzie, że rozmowy w sieci powinny mieć zarezerwowane pasmo i muszą podążać tą samą ścieżką dopóki połączenie nie zostanie przerwane. Choć może nie jest to najlepsza metoda dla wszystkich typów danych, to jest wyjątkowo dobrze przystosowana do multimediów oraz transmisji głosu, obrazu i elektronicznych konferencji w czasie rzeczywistym. Gdy technologie te staną się bardziej popularne w nowoczesnych sieciach, wraz z nimi zyska popularność środowisko sieci z komutacją obwodów.

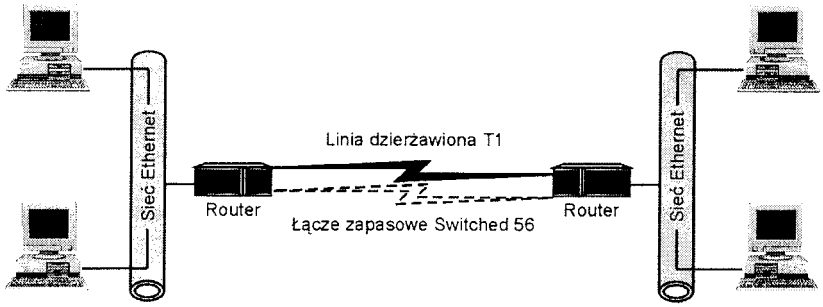
1.15.1 Sieci Switched 56

Sieć Switched 56 jest najtańszą usługą cyfrową w sieci WAN. Działa z szybkością 56 Kbps i wykorzystuje standardowe okablowanie skrętki dwużyłowej. Podobnie jak w publicznej komutowanej sieci telefonicznej (ang. PSTN - Public Switched Telephone Network), połączenia przeprowadza się, wybierając 7-cyfrowy numer telefoniczny i łącząc się z innym obwodem sieci Switched 56 lub linią cyfrowej sieci usług zintegrowanych (ang. ISDN Integrated
Services Digital Network).

Główną zaletą sieci Switched 56 jest fakt, że połączenia są aktywne tylko wtedy, gdy są potrzebne, a podstawą opłaty jest wykorzystanie linii. Powszechną praktyką dostawców usług jest stosowanie takiej samej taryfy dla obwodów sieci Switched 56 i dla standardowych łączy telefonicznych. Stąd, jeśli w Twoim obszarze usługowym lokalne rozmowy telefoniczne są darmowe, to połączenia w sieci Switched 56 również mogą być darmowe.
Słabą stroną sieci Switched 56 są jej małe możliwości rozwojowe - jeśli chcesz zmienić połączenie na szybsze, musisz przejść na inną technologię. Inną wadą technologii Switched 56 jest jej niedostępność - dziś w większej części kraju nie ma dostawców tej usługi.

1.15.1.1 Najczęstsze zastosowania sieci Switched 56

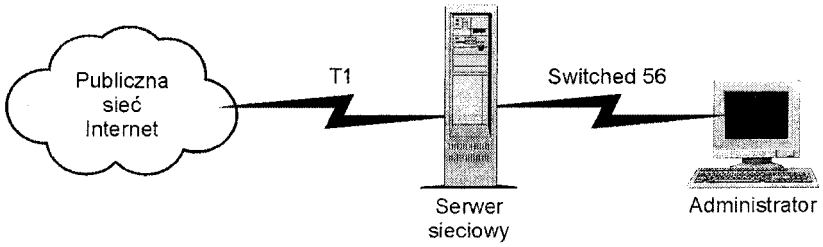
Sieć Switched 56 jest powszechnie używana w połączeniach o małej szerokości pasma, które nie muszą pozostawać aktywne przez cały czas. Sieć ta jest również często używana jako łącze zapasowe linii dzierżawionej, co pokazuje rysunek 1 S.1.
Rysunek 15.1. Sieć Switched 56 jako łącze zapasowe e linii dzierżawionej.



Innym popularnym zastosowaniem sieci Switched 56 jest tzw. autonomiczny dostęp. Termin ten oznacza połączenie między organizacją a siecią lub serwerem, które jest całkowicie oddzielone od podstawowego interfejsu sieciowego. Dla przykładu rozpatrzmy często zachodzącą sytuację dotyczącą korporacyjnego serwera sieciowego. Zwykle serwer taki ma bezpośrednie połączenie z Internetem przez linię dzierżawioną, którą ogół użytkowników wykorzystuje do uzyskiwania dostępu do informacji. Jeśli administratorzy chcą uzyskać dostęp do serwera bez konieczności podawania informacji identyfikacyjnych w publicznej sieci Internet, muszą w tym celu utworzyć autonomiczne połączenie za pomocą łącza Switched 56, co przedstawione jest na rysunku 15.2. Korzystając z tego łącza mogą uzyskać na żądanie dostęp do serwera i są całkowicie oddzieleni od sieci publicznej. Takie połączenie działa również jako łącze zapasowe w przypadku awarii interfejsu podstawowego.

1.15.1.2 Technologie Switched 56

W połączeniach Switched 56 wykorzystywane są różne technologie, określone w specyfikacji EIA/TIA-596. Tylko jedna z nich jest naprawdę popularna.
Rysunek 15.2. Sieć Switched 56 jako niedrogie pouczenie autonomiczne.



Technologia ta, nazwana usługą switched 56, typu III, została opracowana przez firmę Northern Telecom i wykorzystuje standardowy przewód telefoniczny. Jest to wielką zaletą, ponieważ wprowadzający tę technologię mogą wykorzystać istniejącą infrastrukturę okablowania, co znacznie zmniejsza koszty instalacji.
Dane są przesyłane przez kabel z szybkością 64 Kbps. Może się to wydać niespodzianką, ponieważ liczba „56" w nazwie technologii dotyczy właśnie szybkości przesyłania danych. Pozostałe 8 Kbps jest wykorzystywane przez konfigurację wywołań i sygnalizowanie.

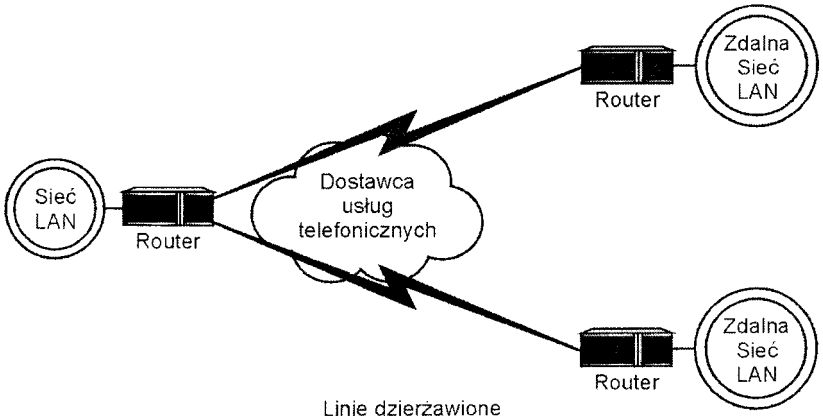
1.15.2 Sieci Frame Relay

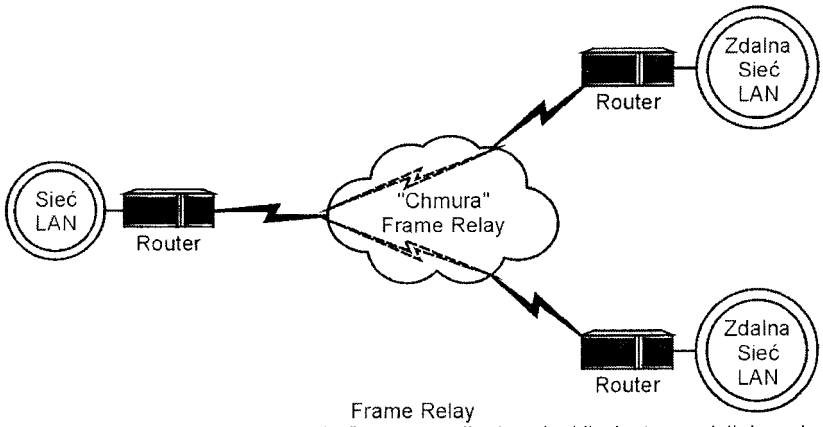
Frame Relay jest siecią z komutacją pakietów, powszechnie używaną jako łącze sieci WAN do przyłączania odległych stanowisk. Ponieważ jest to sieć z komutacją pakietów, emuluje sieć z komutacją obwodów, stosując stałe kanały wirtualne (ang. PVC- Permanent Virtual Circuits), które wyznaczają ścieżkę wśród wielu przełączników.
Sieć Frame Relay istnieje jedynie w dwóch najniższych warstwach modelu OSI. Na każdym końcu łącza znajdują się routery, które przyłączają poszczególne sieci do sieci Frame Relay. Oddalone stanowiska płacą jedynie za linię dzierżawioną łączącą je z dostawcą natomiast dostawca odpowiada za komunikację między stanowiskami. Wewnątrz sieci występuje nadmiarowość połączeń, redukująca koszty i narzut administracyjny w przedsiębiorstwach, które zamawiają publiczne usługi sieci Frame Relay. Przedsiębiorstwa korzystające z usług

operatora sieci Frame Relay znacznie zmniejszają koszty narzutu sieci WAN. Co więcej, niezawodność sieci Frame Relay zmniejsza koszty związane z czasami przestoju.
Sieć Frame Relay jest często wykorzystywana jako udoskonalenie przestarzałej sieci X.25. Standard X.25 powstał w czasach, gdy sieci oparte były na analogowych systemach transmisji i korzystały z okablowania miedzianego - dlatego sieć X.25 była zawodna i podatna na błędy. Aby uporać się z tymi problemami, do standardu X.25 włączono wiele protokołów i mechanizmów kontrolowania błędów. Sieć Frame Relay została zaprojektowana w nowej erze pracy sieciowej, opartej na transmisjach cyfrowych i nośnikach światłowodowych, poświęcającej zbędny narzut mechanizmów kontroli błędów na korzyść szybkości.
Inną zaletą sieci Frame Relay jest kojarzenie wielu połączeń - czy też obwodów wirtualnych - w pojedynczym łączu, takim jak linia dzierżawiona.
Ostatnimi laty sieć Frame Relay jest wykorzystywana coraz częściej jako środek transportu ruchu w sieci łączący klika oddzielnych linii dzierżawionych w pojedynczy obwód Frame Relay. Przenoszone są nawet tradycyjne dane analogowe, np. głos, co stanowi dla odległych stanowisk bardziej ekonomiczną alternatywę wobec opłacania rozmów międzymiastowych.

1.15.2.1 Frame Relay a linie dzierżawione

Powszechnie uważa się, że sieć Frame Relay.stanowi konkurencję dla tradycyjnej linii dzierżawionej. Być może najważniejszym czynnikiem rozwoju sieci Frame Relay jest szybkie zastępowanie nią dzierżawionych linii telefonicznych.
Zwykle sieć Frame Relay jest tańszą alternatywą dla tworzenia sieci WAN. Organizacja, która decyduje się na linie dzierżawione, naraża się na wysoki koszt wszystkich połączeń, z których jedno lub wiele może być bardzo długie. Z kolei organizacja, która chce pracować w sieci WAN, wykorzystując połączenia przez sieć Frame Relay, odpowiada tylko za linię dzierżawioną łączącą organizację z samą „chmurą" Frame Relay. Jest to opcja szczególnie atrakcyjna dla organizacji, które wykonują połączenia na wielkie odległości, ponieważ opłaty w sieci Frame Relay nie zależą od odległości między punktami końcowymi. Zwykle punktem równowagi cenowej dla linii dzierżawionej i Frame Relay jest odległość między 16 a 32 km (10 - 20 mil). Dla połączeń na odległość mniejszą niż 16 km tańsza jest linia dzierżawiona, natomiast przy odległości większej niż 32 km opcją atrakcyjniejszą finansowo jest Frame Relay.
Dzięki połączeniu z siecią Frame Relay organizacja fizyczna może łączyć się z wieloma różnymi partnerami za pomocą pojedynczego połączenia z „chmurą', co pokazuje rysunek 15.3. Dodatkową korzyścią związaną z taką architekturą jest brama do Internetu - usługa oferowana w publicznych „chmurach" Frame Relay, umożliwiająca podłączonym klientom dostęp do Internetu, często oferująca również ochronę za pomocą systemów „firewal l".
Instalowanie nowej, dzierżawionej linii telefonicznej jest procesem kosztownym i czasochłonnym. Za to łączenie się z kimkolwiek za pomocą sieci Frame Relay wymaga tylko skonfigurowania kolejnego stałego lub komutowanego obwodu wirtualnego. Unika się w ten sposób kosztów i komplikacji związanych z doprowadzaniem do budynku kolejnej linii dzierżawionej, konfigurowaniem oddzielnego routera i konserwacją połączenia. Daje to właścicielom połączeń Frame Relay większą elastyczność i przyłączalność w porównaniu z właścicielami linii dzierżawionych. Proces konfiguracji jest znacznie szybszy, co ułatwia przedsiębiorstwom szybsze dopasowywanie się do zmiennych warunków działania.
W samych liniach dzierżawionych nie występuje nadmiarowość łączy -linia dzierżawiona jest po prostu pojedynczym połączeniem między dwoma punktami. Uszkodzenie linii spowodować może całkowitą utratę łączności. Co gorsza, przedsiębiorstwa telefoniczne często popełniają pomyłki i często nękane są awariami sprzętu, które powodują zrywanie połączeń. Sieć Frame Relay nie boryka się z tego rodzaju problemami, ponieważ jest siecią komutowaną- jeśli łącze między dwiema częściami sieci zostanie przerwane,
Rysunek 15.3. Porównanie .sieci Frame Relay i linii dzierżawionych
ruch będzie po prostu ponownie trasowany, a jedynym tego skutkiem będzie niewielkie opóźnienie. Jeśli przełącznik wewnątrz „chmury" przestanie działać prawidłowo, inne przełączniki wykryją problem i zaczną trasować ramki inną ścieżką. Wrodzona nadmiarowość łączy „chmury" Frame Relay czyni z niej usługę znacznie bardziej niezawodną od linii dzierżawionych.





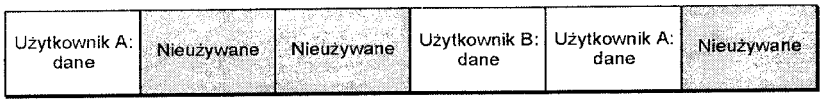
Frame Relay
wymaga tylko jednego połączenia fizycznego dla dowolnej liczby transmisji danych

Linie dzierżawione gwarantują szerokość pasma dzięki procesowi znanemu jako „multipleksowanie z podziałem czasu". Multipleksowanie z podziałem czasu oferuje pojedynczemu klientowi pewien wycinek dużego łącza sieci i jest to zawsze ten sam wycinek. Na przykład, łącze TI zawsze oferuje pasmo przenoszenia 1,54 Mbps. Z jednej strony jest to zaletą, gdyż upraszcza planowanie, ale z drugiej strony jest to wadą, ponieważ pasmo jest zawsze przypisane do połączenia, niezależnie od tego, czy jest używane, czy też nie. Co więcej, szerokości pasma nie można powiększyć wtedy, gdy jest to wymagane, czy też w okresach szczytowego natężenia ruchu. Opłaty nie kształtują się ekonomicznie, gdyż organizacja płaci za całą szerokość pasma w łączu, niezależnie od tego, jaka część tego pasma jest wykorzystywana.

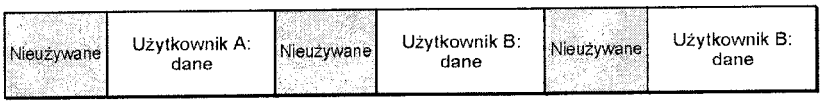
Sieć Frame Relay obchodzi powyższe ograniczenia dzięki wykorzystywaniu statystycznego multipleksowania. Rysunek 15.4 przedstawia porównanie multipleksowania z podziałem czasu i multipleksowania statystycznego. Multipleksowanie statystyczne pozwala sieci Frame Relay nie angażować całego wymaganego pasma na samym początku połączenia. Choć multipleksowanie to oferuje właściwość tzw.. umownego wskaźnika informacji (ang. CIR - Committed Information Rate), gwarantującą przedsiębiorstwom określoną wielkość pasma, to pozwala również rozszerzać pasmo w momentach zwiększonego wykorzystania sieci, przy czym nie płaci się za rozszerzone pasmo, gdy nie jest ono używane. Opłaty zależą od ruchu sieciowego, więc przedsiębiorstwa o mniejszym ruchu płacą mniej, a te o większych potrzebach mogą korzystać z opcji rozszerzenia pasma.

Rysunek 15.4. Multiplelcsowanie z odziałem czasu w porównaniu
z multipleksowaniem statystycznym.

Różnice w sposobie zapewniania pasma mają istotny wpływ na stosowane w obydwu przypadkach metody inżynieryjne. Linie dzierżawione muszą być tak projektowane, aby odpowiadały potrzebom ruchu szczytowego, trudnym do przewidzenia i często gwałtownie się zmieniającym wraz z rozwojem przedsiębiorstwa. Sieci Frame Relay dzięki umownemu wskaźnikowi informacji i możliwości pakietowania, mogą być projektowane tak, aby zaspokajały wymagania przeciętnego natężenia ruchu. Wymagania te można łatwiej przewidzieć, poza tym dzięki możliwości pakietowania są one bardziej elastyczne.



Multipleksowanie z podziałem czasu



Multipleksowanie statystyczne

Wreszcie, „chmura" Frame Relay jest także korzystniejsza dla dostawcy, ponieważ pozwala na wspólne użytkowanie pasma. Dostawca powinien gwarantować, że dysponuje wystarczającym pasmem, by klienci mogli w pełni wykorzystać umowny wskaźnik informacji, ale pozostała część pasma może być wspólnie wykorzystywana przez wielu klientów. Dostawcy linii dzierżawionych (zwykle są to przedsiębiorstwa telefoniczne) muszą zawsze zapewniać każdemu klientowi całą szerokość pasma, jaka może mu być kiedykolwiek potrzebna. Dlatego dostawcy linii dzierżawionych wykorzystują pasmo w znacznie mniejszym stopniu niż dostawcy sieci Frame Relay. Rysunek 15.5 dostarcza szczegółowego opisu różnic między Frame Relay a liniami dzierżawionymi.

1.15.2.2 Rozszerzone Frame Relay

Identyfikatory łącza danych (ang. DLCI - Data Link Connection Identifier) identyfikują wyłącznie połączenie między routerem i siecią Frame Relay - aby zidentyfikować ostateczny punkt docelowy, identyfikatory te muszą być statycznie kojarzone z punktami docelowymi w „chmurze" Frame Relay. Nie przewidziano protokołu rozróżniania adresów, który dynamicznie kojarzyłby identyfikatory łącza danych z punktami docelowymi w sieci Frame Relay.

Rysunek 15.5. Porównanie linii d.ier=awionvclr i Frame Relay.

Linie dzierżawione	Frame Relay
Architektura oczek pełnych wymaga n*(n-1)/2 łączy fizycznych	Architektura oczek pełnych wymaga n*(n-1)/2 stałych lub komutowanych obwodów wirtualnych, ale tylko n łączy fizycznych
Trudne i drogie dodawanie i usuwanie linii	Dodawanie i usuwanie wymaga tylko dodania obwodu wirtualnego
Mniejsza niezawodność, brak połączeń zapasowych	Większa niezawodność, połączenia zapasowe
Małe wykorzystanie pasma	Większe wykorzystanie pasma

Rozszerzenia LMI (ang. Local Management Interface) czynią Frame Relay siecią bardziej solidną i dostarczają wielu zaawansowanych właściwości, z których wcale nie najmniej znaczącą jest adresowanie globalne. Przypisując określony identyfikator łącza danych do określonego routera, adresowanie globalne sprawia, że identyfikatory łącza danych są unikatowe w obrębie „chmury". Dzięki temu „chmura" Frame Relay jest w większym stopniu kompatybilna z popularnymi protokołami wyższego poziomu, jak TCP/1P, które przypisują każdemu węzłowi odrębny identyfikator.
Innym pożytecznym rozszerzeniem LMI jest multicasting. Identyfikatory łącza danych z zakresu od 1019 do 1022 są zarezerwowane jako adresy multicast, które oznaczają w „chmurze" wiele punktów końcowych.

1.15.2.3 Stałe a komutowane kanały wirtualne

W sieci Frame Relay występują dwa rodzaje obwodów (kanałów) wirtualnych: stałe obwody wirtualne (PVC) i komutowane obwody wirtualne (SVC). Stały obwód wirtualny jest tworzony statycznie podczas konfiguracji i zapewnia, że dane przesyłane między dwoma punktami będą zawsze podążać tą samą ścieżką, dzięki czemu właściwości przesyłania danych będą bardziej stabilne. Komutowane obwody wirtualne obliczają ścieżkę za każdym razem, gdy połączenie jest ustanawiane, dzięki czemu obwody mogą omijać miejsca awarii sieci. Ponieważ jednak za każdym razem, gdy ustanawiane jest połączenie, wykorzystywana jest inna ścieżka, mogą się zmieniać charakterystyki wydajności, takie jak fluktuacja i opóźnienie. Na rysunku 15.6 wyszczególnione są różnice między stałymi i komutowanymi kanałami wirtualnymi.
Niektóre implementacje sieci Frame Relay obsługują także multicasting, który umożliwia serwerowi wysyłanie danych jednocześnie do wielu odbiorców. Multicasting jest przeprowadzany w „chmurze" Frame Relay - przełączniki wewnątrz „chmury" gwarantują że wszyscy odbiorcy otrzymają pakiety, które zamówili, a jednocześnie starają się jak najefektywniej wykorzystać dostępne pasmo.
Sieć Frame Relay jest bardzo elastyczna i może być wykorzystywana w łączach o szybkości od 56 Kbps do 'C3 (45 Mbps). Co więcej, w publicznych „chmurach" Frame Relay opłaty zależą od ruchu, dzięki czemu nie płaci się niepotrzebnie za rzadziej wykorzystywane łącza.
Rysunek 15.6. Porównanie stałych i komutowanych kanałów
wirtualnych.

1.15.2.4 Format podstawowej ramki Frame Relay

Stałe kanały wirtualne	Komutowane kanały wirtualne
Połączenie ustanawiane tylko raz	Połączenie ustanawiane na podstawie każdego wywołania
Ramki zawsze podążają tą samą ścieżką	Ramki mogą podążać inną ścieżką za każdym wywołaniem
Stałe i niezawodne	W mniejszym stopniu stałe i niezawodne
Połączenie jest zawsze skonfigurowane, niezależnie od tego, czy jest używane, czy nie	Połączenie jest likwidowane, gdy nie ma już więcej informacji do przesłania
Trwałe i gwarantowane połączenia, ale słabe wykorzystanie	Lepsze wykorzystanie "chmury" Frame Relay
Trudno nimi zarządzać	Łatwiejsze w utrzymaniu
Wymaga sztywnej architektury sieci i większej liczby połączeń dla wysokich poziomów nadmiarowości	Dopuszcza elastyczną architekturę sieci wymagając mniejszej liczby połączeń dla sieci o gęstych oczkach
Opłaty według przewidywalnej, miesięcznej stawki	Opłaty zależą od wykorzystania pasma, czasu trwania transmisji i liczby przesłanych ramek

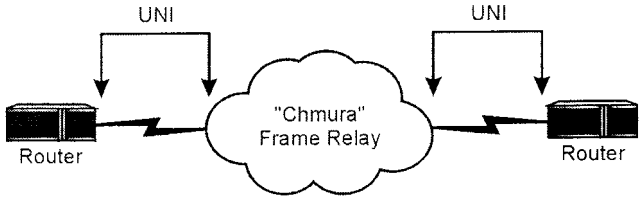
Do grupy danych, które są przełączane wewnątrz sieci, odnosi się termin „ramka", ponieważ sieć Frame Relay istnieje tylko w dwóch najniższych warstwach modelu OSI, a nie w warstwie sieci, gdzie grupy danych są nazywane „pakietami".
Pakiety z poszczególnych sieci LAN są przekazywane do sieci Frame Relay, ta sieć jednak nie czyni różnicy między poszczególnymi protokołami warstwy sieci. Sieć Frame Relay obsługuje te protokoły opakowując je i opatrując nagłówkami Frame Relay z jednej i stopkami Frame Relay z drugiej strony. Ze względu na wydajność całkowity rozmiar takiego pakunku wynosi tylko sześć bajtów.
Dziesięć bitów tego „pakunku" przeznaczonych jest na adres warstwy łącza danych, będący odpowiednikiem adresu MAC w sieci Ethernet. Adres łącza danych nazywany jest identyfikatorem łącza danych, w skrócie DLCI. Routery nie wykorzystują identyfikatora DLCI do prostego określania miejsca przeznaczenia, lecz raczej do identyfikowania kanału na obydwu końcach sieci. Warto zauważyć, że żadna ze stron stałego kanału wirtualnego nie musi używać tego samego identyfikatora łącza danych - identyfikator ten jest znaczący tylko dla danej pary urządzenie końcowe / urządzenie komunikacyjne (DTE/DCE). Tym właśnie identyfikatory łącza danych różnią się od adresów MAC, które są powszechnie unikatowe. Zakres identyfikatorów łącza danych jest zbyt mały, by mogły one być powszechnie unikatowe.
Trzy kolejne bity pakunku wykorzystywane są jako flagi zawiadamiające o zatorach w sieci. Jeden bit umożliwia nadanie ramce niższego priorytetu, zaś pozostałe dwa bity są zarezerwowane dla nieokreślonych jeszcze zastosowań.
W praktyce każdy interfejs powinien mieć od jednego do czterdziestu ośmiu identyfikatorów łącza danych.

1.15.2.5 Projektowanie sieci Frame Relay

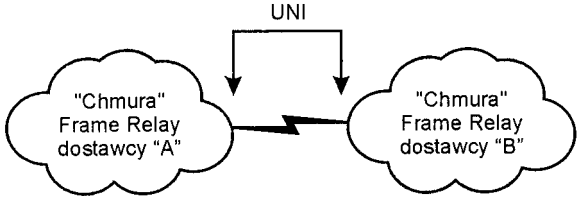
Jedną z zalet sieci Frame Relay jest dostrajalna, skalowalna szerokość pasma. Podczas zamawiania usług Frame Relay dostawca określa umowny wskaźnik informacji (CIR) precyzujący gwarantowaną szerokość pasma między dwiema dowolnymi lokalizacjami.

1.15.2.6 UNI a NNI

Dwie części specyfikacji Frame Relay definiują protokoły dla połączeń interfejsu użytkownik-sieć, czyli interfejsu UNI (ang. User-to-Network Interface) oraz dla połączeń interfejsu międzysieciowego NNI (ang. Network-to-Network Interface). Interfejs użytkownik-sieć został opracowany przez instytut ANSI i sekcję standardów Międzynarodowej Unii Telekomunikacyjnej (ITU-T). Dostarcza on specyfikacji, dzięki którym urządzenie końcowe może porozumiewać się z urządzeniem komunikacyjnym, jak to przedstawia rysunek 15.7. Zakres tego standardu jest funkcjonalnie zbliżony do specyfikacji X25. Specyfikacje interfejsu użytkownik-sieć nie różnią się od standardów poczty Stanów Zjednoczonych, dotyczących adresowania listów - opisują określony format adresu docelowego, ale nie wdają się w szczegóły dotyczące sposobu przesyłania listu do adresata, czyli tego, co dzieje się z nim po wrzuceniu do skrzynki. Poczta daje natomiast pewne gwarancje dotyczące usługi, takie jak czas dostarczenia i niezawodność, a osoba wysyłająca list musi ufać, że poczta wywiąże się ze zobowiązań.
Rysunek 15.7. Interfejsy użytkownik-sieć łączące użądzenie końcowe z ...-ad-eniem komunikacyjnym
Sieć Frame Relay działa w ten sam sposób: sieć i użytkownik końcowy komunikują się korzystając ze standardów UNI. Sieć określa pewne gwarancje jakości usług, takie jak umowny wskaźnik informacji „CIR" (Commited Information Rate), umowny rozmiar pakietu „Bc" (ang. Committed Burst Size) oraz nadmiarowy rozmiar pakietu „Be" (ang. Excess Burst Size). W gestii dostawcy leży sposób. w jaki sieć wywiązuje się ze swoich obietnic.

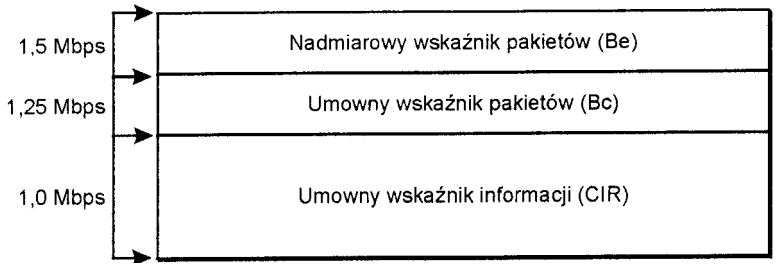


Interfejs międzysieciowy pozwala dwóm sieciom wymieniać dane, przy czym nie jest konieczna znajomość struktury poszczególnych sieci.
Specyfikacja interfejsu
międzysieciowego została opracowana przez forum Frame Relay. Dzięki interfejsowi międzysieciowemu dostawcy Frame Relay mogą zwiększyć użyteczność sieci, udostępniając klientom bramy do sieci innych dostawców. Organizacje mogą współdziałać pomimo. iż nie korzystają z usług tego samego dostawcy. Zakres standardu interfejsu międzysieciowego przedstawia rysunek 15.8.
Rysunek 15.8. Interfejsy międzysieciowe łączące sieci różnych dostawców.



1.15.2.7 Przekraczanie szybkości przesyłania informacji

Moc sieci Frame Relay zależy od dwóch parametrów: umownego wskaźnika pakietów „Bc" (ang. *Commited Burst Rate)* i nadmiarowego wskaźnika pakietów „Be" (ang. *Excess Burst Rate),* ustalanych podczas tworzenia stałego obwodu wirtualnego. Wskaźnik Bc określa dodatkową wielkość ruchu, jaką operator sieci Frame Relay może przesłać powyżej limitu ustalonego w umownym wskaźniku informacji (CIR). Nie oznacza to, że operator sieci Frame Relay zgadza się robić to za darmo. Oznacza to, że przez pewien czas może on zaakceptować ruch większy niż przewidziany w umownym wskaźniku informacji, ale tylko pod warunkiem, że w danym okresie czasu średnia wartość natężenia ruchu nie przekroczy wielkości ustalonej we wskaźniku CIR. Wskaźnik Be określa maksymalną wielkość ruchu, jaką może zaakceptować dostawca, ale ruch o natężeniu z zakresu pomiędzy wskaźnikami Bc i Be nie jest gwarantowany. Rysunek 15.9 pokazuje związki między trzema wskaźnikami: CIR, Bc i Be. Każdy z nich musi zostać określony podczas zamawiania usługi u dostawcy.
Rysunek 15.9. l'orównnnie wskaźników Be, Bc i CIR.



Rozważmy następujący przykład: oddalone biuro jest połączone ze swoim dostawcą Frame Relay łączem T1, a z centralą firmy stałym obwodem wirtualnym o wskaźniku CIR równym 1,0 Mbps, wskaźniku Bc - I,25 Mbps i wskaźniku Be - 1,5 Mbps. Maksymalna przepustowość łącza T1 wynosi około 1,5 Mbps, czyli jest niemal półtora razy większa niż wskaźnik CIR. Użytkownik, dajmy na to Teresa, wysyła do centrali firmy

zestawienie kosztów, korzystając z usług protokołu transferu plików (FTP). FTP wykorzystuje całe pasmo obwodu i przekazuje dane do „chmury" Frame Relay z szybkością 1,25 Mbps.

Ta „eksplozja" danych trwa tylko jedną sekundę. W ciągu następnej sekundy natężenie ruchu spada poniżej wartości 0,5 Mbps. Przeciętna przepustowość w okresie dwóch sekund jest mniejsza niż wartość wskaźnika CIR, więc warunki umowy zostały dotrzymane i gwarantowane jest przesłanie wszystkich danych.

A teraz rozważmy nieco inną sytuację: Teresa wysyła do centrali bazę danych o wielkości 100 MB. Transfer danych jest ograniczony ze względu na maksymalną przepustowość łącza T1, wynoszącą 1,5 Mbps. Przełącznik Frame Relay, który otrzymuje dane, wylicza, że przeciętna szybkość przesłania danych przekracza wartość wskaźnika CIR. Przełącznik może teraz zareagować na dwa sposoby. Najkorzystniej będzie, jeśli prześle dane do centrali, ale ustawi w nagłówku ramki bit Odrzucenie Dozwolone, czyli bit DE (ang. Discard Eligible). Bit ten służy do oznaczania danych o niskim priorytecie - jeśli gdzieś w sieci wystąpi zator, dane o niskim priorytecie mogą zostać porzucone na korzyść ramek o wyższym priorytecie.

Druga możliwość jest taka, że przełącznik może po prostu porzucić pakiety, które przekraczają wskaźnik CIR. Zwykle dzieje się tak, gdy w sieci jest tłok.

Przełączniki mogą akceptować ruch aż do wielkości określonej wskaźnikiem Be, ale nie w dłuższym czasie. Wciąż wymaga się, by przepustowość w danym czasie nie przekraczała wartości wskaźnika CIR.

1.15.2.8 Sterowanie przepływem w sieci Frame Relay

Sterowanie przepływem jest możliwe dzięki dwóm polom nagłówka ramki. Są to pola: „Zawiadomienie węzła odbierającego o napotkanym zatorze" (ang. FECN - Forward Explicit Congestion Notification) i „Zawiadomienie węzła nadającego o napotkanym zatorze" (ang. BECN - Backward Explicit Congestion Notification). Pola te umożliwiają przełącznikom przesłanie prośby do węzłów końcowych, aby te ograniczyły wykorzystanie pasma, zanim zajdzie konieczność porzucenia ramek. W ten sposób sieć może zmniejszyć natężenie ruchu, zapobiegając przeciążeniu. Rozciągnięcie przesyłania plików w czasie jest dla użytkownika korzystniejsze niż konieczność ponownego przesłania ramek porzuconych przez sieć Frame Relay.

Przełącznik Frame Relay ustawia bit FECN, gdy stwierdzi zator w sieci i chce powiadomić o tym węzeł odbierający. Wtedy partner odbierający powinien ograniczyć wykorzystanie sieci. Przełącznik może również ustawić bit BECN w ramkach powracających do węzła, który wysyła zbyt dużo danych. Wykorzystywanie bitu BECN jest efektywniejsze niż korzystanie z bitu FECN, ponieważ host powiadamiany jest bezpośrednio.

Dopiero niedawno urządzenia końcowe zaczęły obsługiwać bity FECN i BECN. Jeśli sprzęt klienta nie reaguje na żądania, bity te nie są użyteczne.

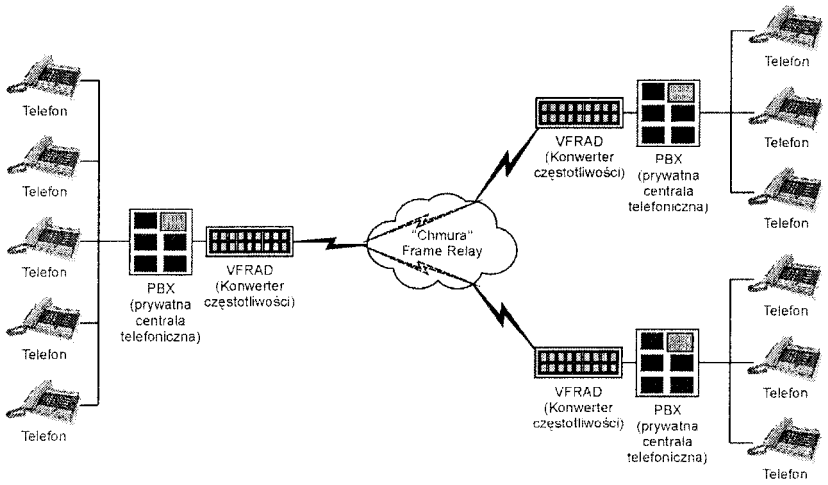
1.15.2.9 Przesyłanie głosu za pomocą Frame Relay

Nową cechą sieci Frame Relay jest zdolność multipleksowania różnych rodzajów danych w jednym obwodzie fizycznym. Czynniki ekonomiczne sprzyjają wprowadzeniu przesyłania głosu za pomocą Frame Relay, gdyż taka możliwość pozwala oddalonym biurom przeprowadzać rozmowy telefoniczne bez wykorzystywania publicznej sieci telefonicznej lub linii dzierżawionych. Organizacje, które zdecydowały się wykorzystywać

Frame Relay do przesyłania danych i zainwestowały w sprzęt sieci Frame Relay. mogą łatwo wdrożyć przesyłanie głosu za pomocą tej sieci.

Tradycyjnie głos przesyłany jest dedykowanymi liniami analogowymi lub liniami cyfrowymi, wykorzystującymi multipleksowanie z podziałem czasu. Dzięki temu każda rozmowa ma zagwarantowane pasmo i jakość połączenia nigdy nie spada. Generalnie. pojedyncze połączenie telefoniczne wymaga pasma około 64 Kbps.

Przesyłanie głosu siecią Frame Relay stanowi ciekawe wyzwanie, gdyż sieci z komutacją pakietów wykorzystują multipleksowanie statystyczne, które sprawdza się w przypadku przesyłania danych, ale może zawodzić w aplikacjach czasu rzeczywistego (takich jak przesyłanie głosu), jeśli w sieci wystąpi zator. Aby zmniejszyć szerokość pasma, jakiej wymaga każdy obwód przekazujący głos i poprawić ogólne wykorzystanie łączy, stosuje się kompresję strumienia głosu, usuwając pauzy i nadmiarowe (powtarzające się) informacje. Zazwyczaj do zapewnienia wysokiej jakości transmisji wystarcza tylko 22% typowej szerokości pasma, dostępnej w obwodach przesyłających głos.
Aby przesłać te dane (głos) siecią Frame Relay, informację należy skompresować (jak opisano powyżej) i umieścić w ramce danych. Następnie multiplekser Frame Relay umieszcza ramki w tym samym połączeniu wyjściowym, którym mogą wychodzić inne dane podróżujące tym samym łączem fizycznym. L.ącząc głos i dane w tych samych obwodach, organizacje znacznie zmniejszają koszty, jakie ponosiłyby w przypadku utrzymywania oddzielnych łączy i urządzeń sieciowych. Proces ten jest przedstawiony na rysunku 15.10.
Rysunek 15.10. Sieć Frame Relay przenoszenia głosu łącząca wiele miejsc.



Kompresja i multipleksowanie przynoszą oczywiście korzyści, ale mają również efekty niepożądane. Kompresja, zwłaszcza w zatłoczonych sieciach, może być stratna. Termin „kompresja stratna' oznacza kompresję, w której jakość poświęca się w zamian za większą wydajność. Choć rozmowa telefoniczna może się odbywać pomimo kompresji stratnej, to jej jakość jest niższa od tej, do jakiej większość ludzi przywykła, korzystając z usług droższych sieci, takich jak publiczna komutowana sieć telefoniczna. Kolejną wadą jest brak standardów naliczania opłat dla sieci Frame Relay przenoszącej głos. Istniejące systemy, które śledzą rozmowy i naliczają opłaty, mogą wymagać zmiany, gdy zdecydujemy się korzystać z sieci Frame Relay. W sieciach z komutacją pakietów, takich jak Frame Relay, występują również opóźnienia i fluktuacje, jeszcze bardziej obniżające jakość. Aby lepiej obsługiwać aplikacje czasu rzeczywistego, takie jak przesyłanie głosu, zaprojektowano sieć ATM. W tym przypadku jest ona lepszą, choć droższą alternatywą.

1.15.3 Sieci prywatne, publiczne i hybrydowe (mieszane)

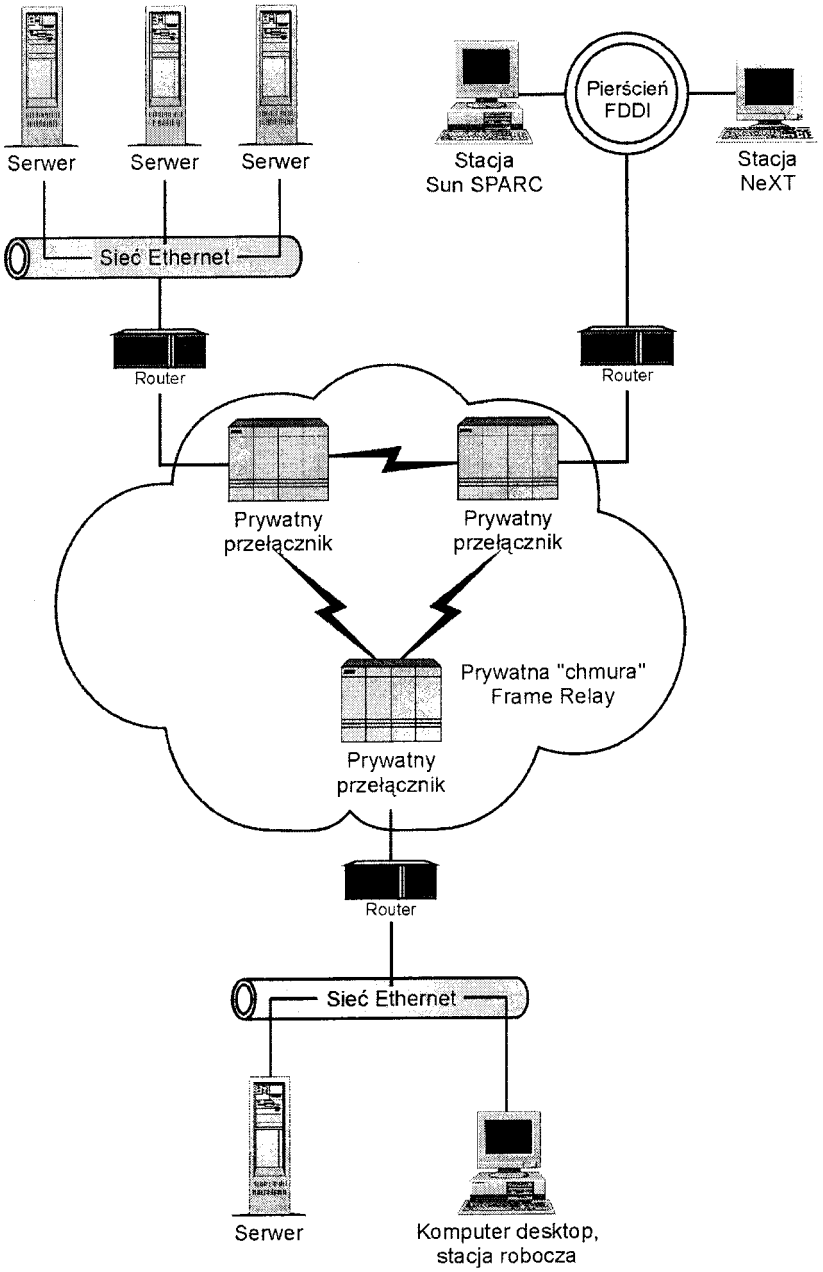
Sieć Frame Relay stała się tak potężną technologią, że zyskała solidną bazę klientów, którą tworzą organizacje pragnące wykorzystać ją zarówno w sieciach publicznych, jak i prywatnych.

1.15.3.1 Prywatne sieci Frame Relay

Prywatne sieci Frame Relay, jak ta pokazana na rysunku 15.1 1, oferują następujące korzyści:
• Sieć Frame Relay może korzystać z istniejącego sprzętu sieciowego, przedłużając jego czas życia i chroniąc inwestycje.
• Dzięki wspólnemu korzystaniu z pasma Frame Relay pozwala organizacji lepiej wykorzystywać istniejące obwody.
• Ponieważ całe wyposażenie skupia się w rękach jednej organizacji, może ona zmniejszyć ryzyko narażenia danych sieciowych na niebezpieczeństwo.
• Frame Relay oferuje większą elastyczność i kontrolę nad siecią.

1.15.3.2 Publiczne sieci Frame Relay

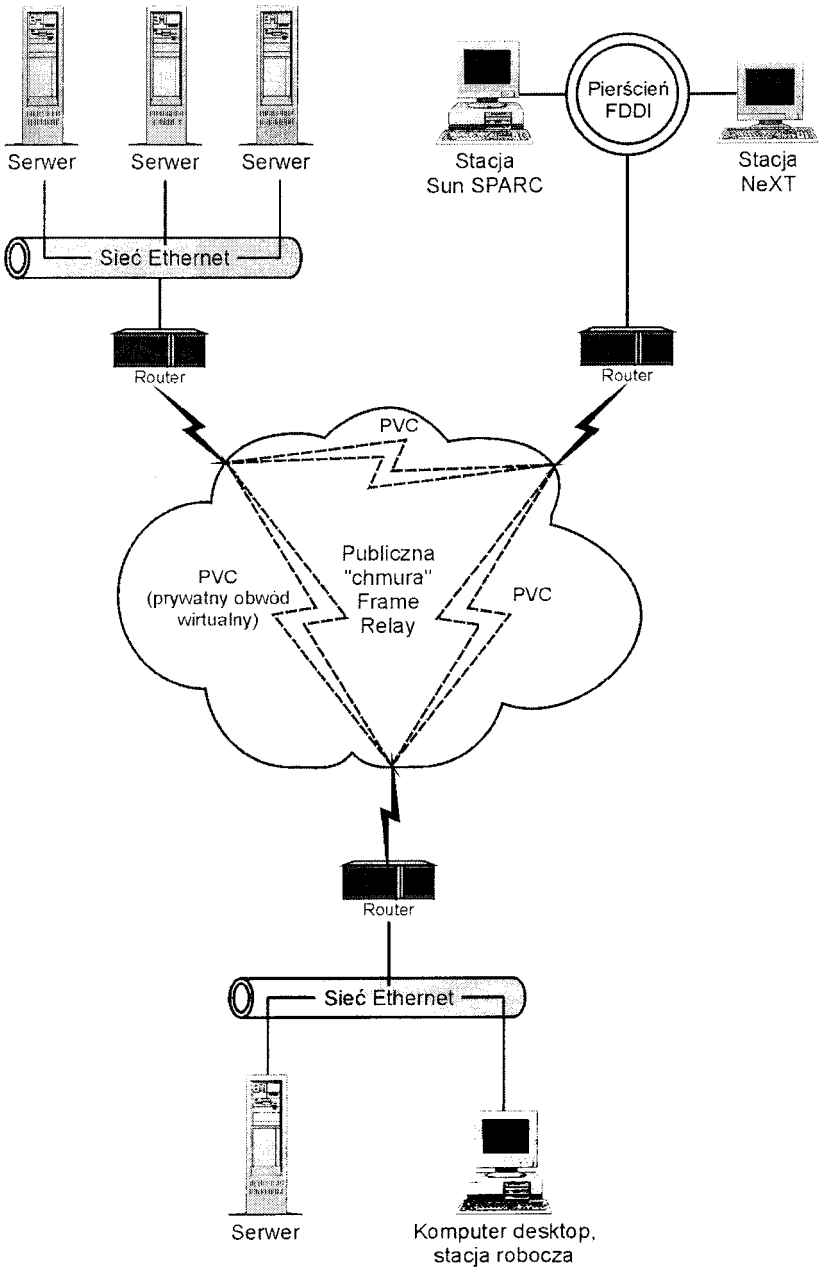
Publiczne sieci Frame Relay.,jak ta pokazana na rysunku 15.12, oferują następujące korzyści:
• Publiczne sieci Frame Relay zmniejszają koszt własności, ponieważ szkieletem sieci zarządza dostawca usługi.
• Publiczne sieci Frame Relay zwykle mają duży zasięg geograficzny, udostępniając tanie połączenia wielu oddalonym miejscom. Umożliwiają nawet dostęp dial-in, czyli przy użyciu modemu.
• Publiczne sieci Frame Relay ułatwiają współdziałanie z niezależnymi organi
zacjami.
Rysunek 15.11. Przykład prywatnej sieci Frame Relay.
Hybrydowe sieci Frame Relay
Hybrydowe sieci Frame Relay, jak ta pokazana na rysunku 15.13, oferują następujące korzyści:



Rysunek 15.12. Publiczna sieć Frame Relay.

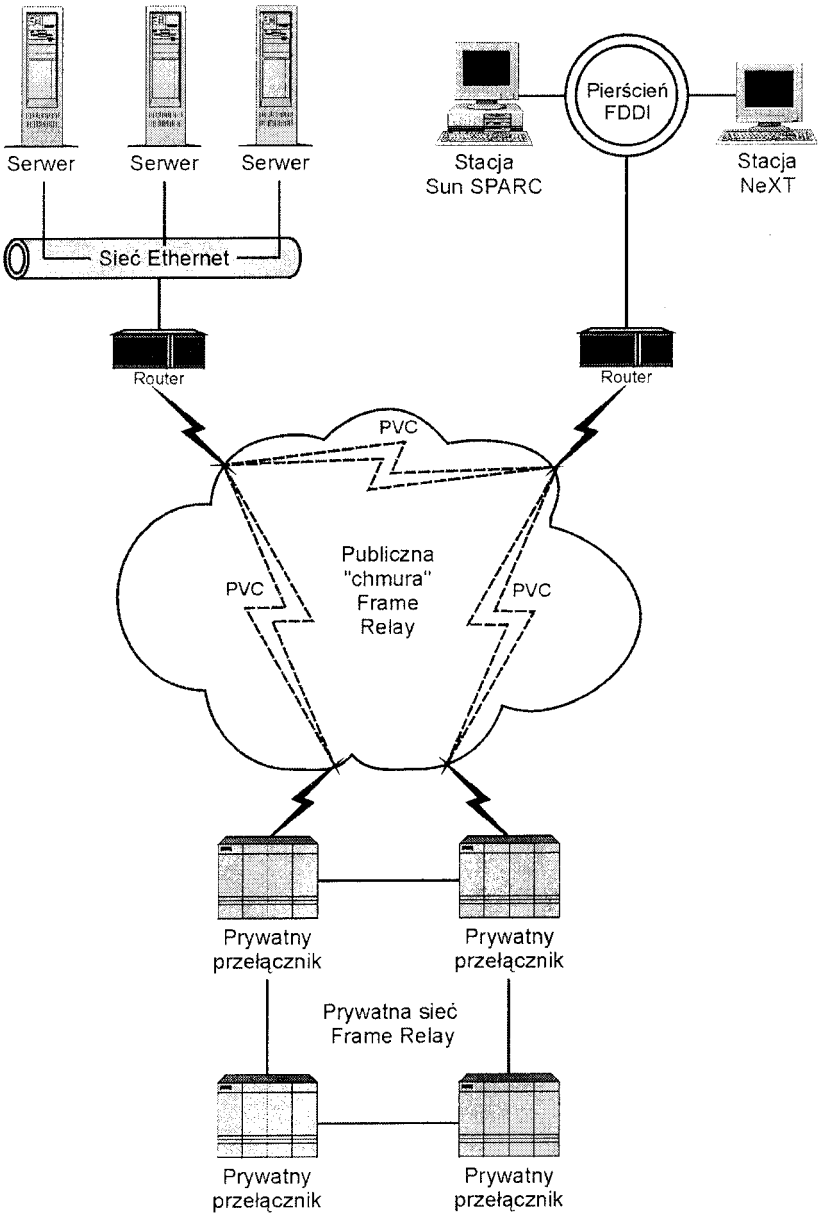
• Hybrydowe sieci Frame Relay zmniejszają koszt własności, ponieważ szkieletem sieci zarządza dostawca usługi.

• Hybrydowe sieci Frame Relay zwiększają kontrolę i bezpieczeństwo poszczególnych sieci.



• Hybrydowe sieci Frame Relay oferują „złoty środek" w proporcji ceny do wydajności.

Rysunek 15.13. Hybrydowa sieć France Relay.



1.15.3.3 Współdziałanie międzysieciowe przy zastosowaniu ATM

Tryb transferu asynchronicznego, znany lepiej jako ATM (ang. Asynchronous Transfer Mode) jest szybko rozwijającą się technologią przeznaczoną dla wielu obecnych zastosowań sieci Frame Relay. Wysoka prędkość, jakość usług i skalowalność oferowane przez technologię ATM są dla wielu przedsiębiorstw zachętą do tworzenia sieci na niej opartych. Aby dostosować się do tego trendu i ułatwić przejście, większość dostawców publicznych sieci Frame Relay oferuje - lub planuje zaoferować - przezroczys współdziałanie
z siecią ATM. Korzystając z bram, takich jak pokazana na rysunku 15.14, organizacja może udoskonalić niektóre części swojej sieci Frame Relay, zachowując jednocześnie pełnię możliwości współdziałania z odległymi biurami ciągle wykorzystującymi istniejące sieci Frame Relay. Pozwala to także przedsiębiorstwu przejść na technologię ATM przy wykorzystaniu istniejącego szkieletu sieci i utrzymać połączenia Frame Relay z innymi partnerami.
Wiele nowych sieci wciąż jest tworzonych na podstawie Frame Relay, a nie ATM. Współdziałanie z siecią ATM przez bramę stopniowo przyzwyczaja przedsiębiorstwo do środowiska bardzo szybkiej pracy sieciowej i pozwala nabrać doświadczenia personelowi obsługującemu sieć. Rozwijając najpierw szkielet na podstawie dobrze rozwiniętej technologii Frame Relay, organizacja unika wysokich kosztów rozruchu, które wiązałyby się z całkowitym przejściem na sieć ATM.

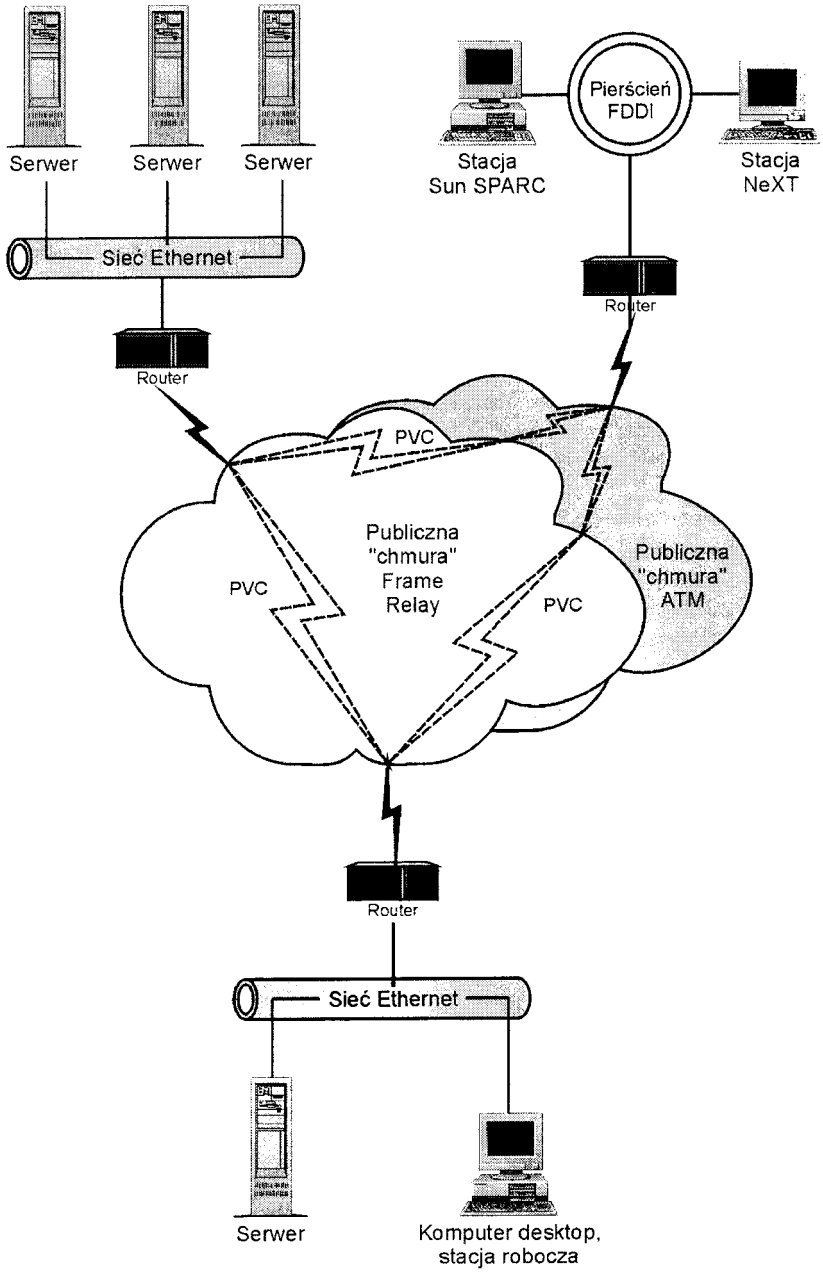
1.15.4 ATM

Tryb transferu asynchronicznego, znany również jako szybkie przełączanie pakietów (ang. Fast Packet Switching), został opracowany jako szerokopasmowa alternatywa dla cyfrowej sieci usług zintegrowanych (ISDN). ATM działa z szybkością z zakresu od 1,54 Mbps do 622 Mbps i jest dobrze przystosowany do transmisji danych, głosu i obrazu. Jedną z głównych zalet ATM-u jest to, że dobrze adaptuje się do nowych technologii, oferując takie właściwości jak ‚jakość usług" (ang. Quality of Service). Inną zaletą jest możliwość uzyskania większej szerokości pasma w czasie porcjownia (pakietowania) danych, dzięki wykorzystywaniu czasu przestoju na łączu.

ATM nie przełącza pakietów, jak czynią to konwencjonalne sieci - ATM przełącza komórki. Komórka różni się od pakietu, ponieważ ma ustaloną długość, co widać na rysunku 15.15. W sieci A"fM długość ta wynosi 53 bajty (48 bajtów danych i 5 bajtów nagłówka). Ustalona długość komórki sprawia. że urządzenia przełączające mogą działać dużo szybciej niż konwencjonalne przełączniki pakietów o zmiennej długości, gdyż przetwarzania wymaga tylko nagłówek komórki. Konwencjonalne przełączniki pakietów o zmiennej długości muszą czytać i przetwarzać każdy bit przychodzący przez przewód, aby mogły określić początek i koniec każdego z pakietów.

Rysunek 15.14. Brama Frame Relay ATM.

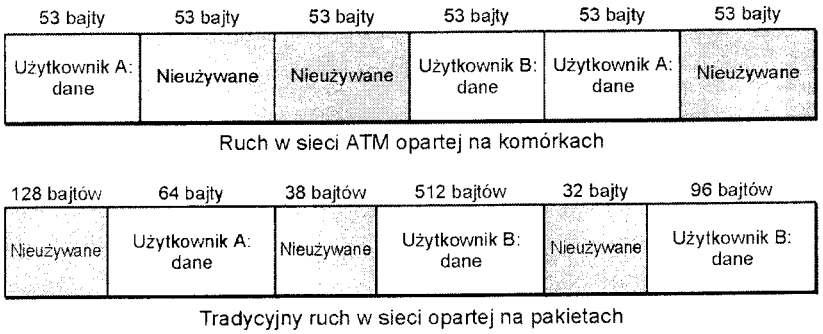
1.15.4.1 Historia ATM



ATM opracowano, aby rozwiązać problemy, z jakimi borykały się konwencjonalne przełączniki pakietów o zmiennej długości. Głównym celem forum IETF ATM było opracowanie technologii, która byłaby elastyczna i skalowalna.

Rysunek 15.15. Porównanie komórek i pakietów.

Odkąd istnieją sieci, ich menedżerowie chcą skonsolidować liczne obwody szkieletowe. Każdy obwód wymaga dodatkowego czasu administracji, wprowadza zwiększoną złożoność, niepewność i koszt. łącze szkieletowe między biurem korporacyjnym a oddalonym stanowiskiem składa się z kanału transmisji głosu, łączącego prywatne centrale telefoniczne, i kanału transmisji danych, łączącego routery. Wraz z rozpowszechnianiem się takich technologii, jak wideokonferencje, powstaje konieczność stosowania dodatkowych kanałów.



Konstruktorzy ATM - czyli XVIII grupa badawcza ITU-T - dostrzegła ten problem i zdecydowała nie odwoływać się do żadnej szczególnej technologii, ale raczej zaprojektować ATM tak, aby mógł pracować z istniejącymi oraz jeszcze nie opracowanymi technologiami. Specyfikacje Międzynarodowej Unii Telekomunikacyjnej są uznawane na całym świecie. W ten właśnie sposób spełnia się marzenie inżynierów i menedżerów sieci - technologia sieciowa wychodząca naprzeciw ich teraźniejszym i przyszłym potrzebom.
Aby obsłużyć tak wiele różnich technologii, ATM musi mieć wyjątkowo małe opóźnienie i fluktuację. Zasadniczo sieć musi być „przezroczysta" i w żaden sposób nie może spowalniać ani przyspieszać ruchu sieciowego. Dlatego ATM musi działać z odpowiednio wysoką szybkością, by nie stał się wąskim gardłem, jak to dzieje się w przypadku wielu sieci WAN. Nie może również niemal wcale tracić pakietów, czy to w wyniku przekłamań danych podczas transmisji. czy też porzucania pakietów z powodu przepełnienia buforów.
ATM musi również obsługiwać zmienne prędkości przesyłania bitów, odpowiadające umownym wskaźnikom informacji (takim jak ustalane w sieci Frame Relay) dla aplikacji czasu rzeczywistego i umożliwiające rozszerzanie pasma w celu lepszego dopasowania do warunków ruchu w sieci. Czyli krótko - ATM ma być wszystkim dla wszystkich.
Choć sieci ATM są obecnie wykorzystywane w praktyce, to standard ten wciąż jest w dużym stopniu niekompletny i ciągle trwają nad nim prace. Obecne implementacje ATM polegają w tym zakresie na producentach, którzy dostarczają niektóre brakujące standardy. Dlatego bardzo ważne jest współdziałanie producentów i należy czynić wszelkie wysiłki, aby wdrażać nowe projekty sieci ATM na jednorodnym sprzęcie.
Zaletą sieci ATM są nieblokujące przełączniki. Przełącznik „nieblokujący" to taki, który jest w stanie utrzymać maksymalną przepustowość w każdym ze swoich interfejsów.
Na przykład, przełącznik, który ma osiem portów, 155 Mbps każdy, wymagałby zagregowanego pasma o szerokości 1,25 Gbps.

1.15.4.2 ATM - sedno sprawy

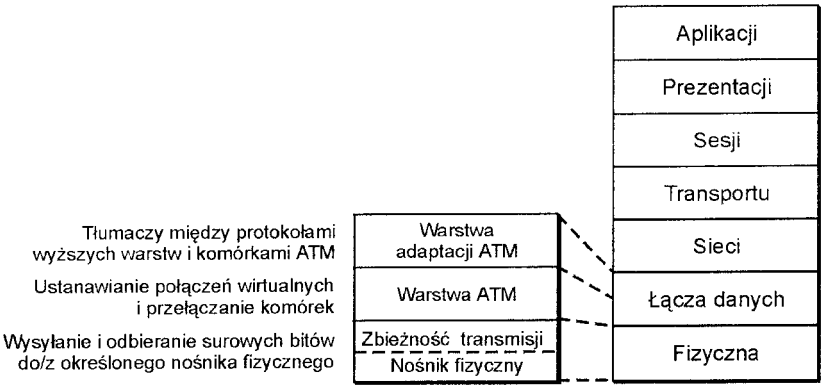
W niniejszym punkcie opisane są szczegóły dotyczące faktycznego działania sieci ATM; nie wszystkie się tu zmieściły, ale wystarczy ich do utworzenia solidnej podstawy dla wszystkich zainteresowanych tą technologią.

1.15.4.3 **Warstwy** ATM

Dla zapewnienia większej elastyczności specyfikacje ATM dzieli się na kilka warstw. Choć nie miały one odpowiadać warstwom modelu OSI, podstawowe warstwy ATM są w przybliżeniu analogiczne do dwóch najniższych warstw - fizycznej i łącza danych. Rysunek 15.16 przedstawia związek między ATM i modelem OSI.
Rysunek 15.16. Porównanie ATM
modelem OSI.

1.15.4.3.1 Warstwa fizyczna

Warstwa fizyczna ATM bezpośrednio odpowiada warstwie fizycznej opisanej w modelu OSI - jest zależna od nośnika i odpowiada za faktyczne transmitowanie bitów. Warstwa ta dzieli się na podwarstwę nośnika fizycznego i podwarstwę zbieżności transmisji. Podwarstwa nośnika fizycznego odpowiada za umieszczanie bitów w nośniku (zwykle światłowodzie) i zapewnia taktowanie.



Podwarstwa zbieżności transmisji odpowiada za rozgraniczanie komórek, kontrolowanie błędów nagłówka i utrzymywanie odpowiedniej struktury ramki w zależności od nośnika.

1.15.4.3.2 Warstwa ATM i warstwa adaptacji ATM (warstwa łącza danych)

Warstwa ATM jest odpowiedzialna za ustanawianie połączeń w sieci ATM. Odwzorowuje adresy warstwy sieci na adresy warstwy łącza danych. Opracowano cztery schematy adresowania - format publiczny jest taki sam jak w wąskopasmowych cyfrowych sieciach usług zintegrowanych.

Warstwa adaptacji ATM (ang. AAL - ATM Adaptation Layer) zapewnia funkcjonalność translacyjną (czyli po prostu - tłumaczy) pomiędzy protokołami wyższych warstw a właściwymi komórkami ATM. Przykładowo, warstwa AAL dzieli pakiety TCP/IP na 48-bajtowe segmenty i opakowuje je nagłówkiem ATM.

Do obsługiwania różnych typów danych przeznaczono różne specyfikacje warstwy adaptacji ATM. Są one przedstawione na rysunku 15.17.

Rysunek 15.17. Zestawienie różnych specyfitcacji warstwy AAL.

Specyfikacja AALI została przeznaczona do przenoszenia ruchu analogowego, który normalnie byłby przenoszony przez obwody dedykowane. Wymaga nośnika niższej warstwy, który obsługuje sterowanie taktowaniem, np. takiego jak SONET. Specyfikacja zapewnia gwarantowane pasmo między urządzeniami końcowymi.

	Wymaga taktowania	Szybkość przesyłania danych	Tryb połączenia	Typy transmisji
AAL1	Tak	stała	połączeniowa	głos, wideo i inne transmisje konwersacyjne
AAL3/4	Nie	zmienna	połączeniowa	dane
AAL4	Nie	zmienna	bezpołączeniowa	dane
AAL5	Nie	zmienna	połączeniowa	dane

Specyfikacja AAL3/4 przesyła przez sieć ATM pakiety usług SMDS (ang. Switched Multimegabit Data Service.

Specyfikacja AALS jest warstwą adaptacji ATM najczęściej używaną w sieciach połączeniowych. Służy do emulowania klasycznego protokołu IP w sieciach ATM i sieciach LAN.

1.15.4.4 Format komórki ATM

Komórka ATM ma długość 53 bajtów, z czego 5 bajtów zajmuje nagłówek.

Wyróżnia się dwa formaty nagłówka. Typ nagłówka interfejsu użytkownik-sieć (ang. UNI - User-Network Interface) jest stosowany dla wszystkich komórek przesyłanych między użytkownikami końcowymi a siecią ATM. Jest bardzo podobny do typu nagłówka interfejsu sieć-węzeł (ang. :NNI - Network-Node Interface), ale istnieją między nimi dwie ważne różnice, co widać na rysunku 15.18.

Rysunek 15.18. .°oróYrnanie

Formatów komórek *-et*

	UNI	NNI
Pole VPI	8 bitów	12 bitów
Sterowanie przepływem ogólnym	4 bity	Nie istnieje

Jak można zobaczyć na rysunku 15.19, nagłówek składa się z sześciu pól: • Zarezerwowane

• Kontrola błędów nagłówka (ang. HEC - Header Error Control) • Sterowanie przepływem ogólnym

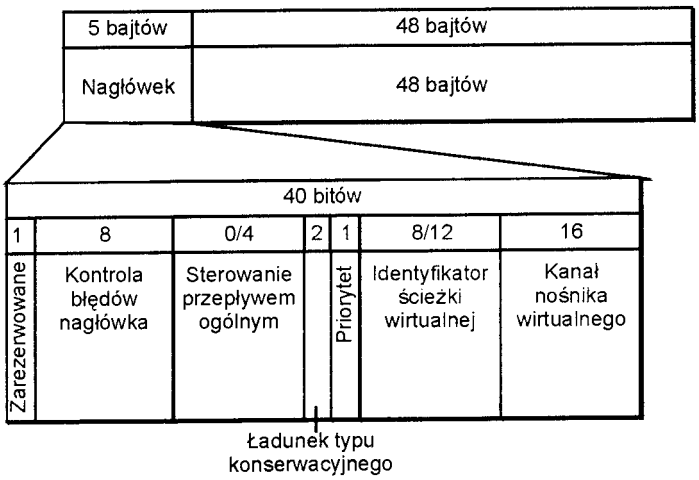
• Ładunek typu konserwacyjnego

• Wskaźnik typu priorytetu (ang. PTI -- Priority Type Identifier)

• Identyfikator ścieżki wirtualnej (ang. *VPI - Virtual Path Identifier)*

Rysunek 15.19. Struktura komórki ATM.

Pola te dostarczają przełącznikom ATM informacji potrzebnych do trasowania komórek między punktem źródłowym a punktem docelowym.



Zarezerwowane (1 bit)
Pierwszy bit komórki ATM pozostawiono dla przyszłych zastosowań. Choć z jednym bitem niewiele można zrobić, może on być w przyszłości wykorzystany do zmiany znaczenia pozostałych pól nagłówka.

1.15.4.4.1 Kontrola błędów nagłówka (HEC) (8 bitów)

W odróżnieniu od wielu innych kanałów WAN, ATM nie oferuje kontroli błędów. Za to jego nośniki transmisyjne są wyjątkowo niezawodne. Dzieje się tak ponieważ tryb transferu asynchronicznego zwykle wykorzystywany jest w bardzo szybkich sieciach światłowodowych, jak SONET czy FDDI, gdzie nośnik fizyczny jest niemal pozbawiony wad. Podczas fazy projektowania ATM zdecydowano, że informacje kontroli błędów umieszczone w nagłówku ramki zajmowałyby zbyt dużo miejsca i w efekcie spowolniłyby przełączanie, gdyż każdy przełącznik musiałby odczytywać całą komórkę, obliczać sumę kontrolną i porównywać ją z odpowiednim polem w nagłówku. Cała korekcja błędów i retransmisja komórek musi być wykonywana przez klienta i serwer, wykorzystujących protokoły takie jak TCP/IP i związane z nim pole cyklicznej kontroli nadmiarowej (CRC).

Choć ATM pozostawia kontrolę błędów protokołom wbudowanym, to nie może na nich polegać, jeśli chodzi o kontrolę własnego, 5-bajtowego nagłówka. Gdyby w nagłówku wystąpił błąd, pakiet mógłby zostać przesłany do niewłaściwego przełącznika, stając się potencjalnym źródłem problemów. Przykładowo, zmiana nawet jednego bitu w nagłówkowym polu Identyfikatora Ścieżki Wirtualnej/ Identyfikatora Kanału Wirtualnego mogłaby spowodować błędne trasowanie komórki.

Aby wykrywać błędy w swoim nagłówku, ATM wykorzystuje pojedynczy bajt parzystości, zwany Kontrolą Błędów Nagłówka (HEC). Dzięki niemu przełącznik może wykryć zmianę nawet jednego bitu i odpowiednio skorygować błąd. W ten sposób błędy w pojedynczych bitach są naprawiane i pakiet jest przesyłany dalej bez przerywania usługi. Jeśli w nagłówku uszkodzonych jest więcej bitów niż jeden, cała komórka jest odrzucana, a jej nadawca nie jest o tym powiadamiany. Odpowiedzialność za wykrycie utraty informacji spoczywa na wbudowanym protokole sieciowym.

1.15.4.4.2 Sterowanie przepływem ogólnym (0 lub 4 bity)

Funkcja ta jest przeznaczona do ustanawiania komutowanych obwodów wirtualnych i nie została jak dotąd wdrożona w standardach ATM. Pozwoli ona użytkownikowi utworzyć połączenie z jednym lub wieloma punktami docelowymi.

1.15.4.4.3 Ładunek typu konserwacyjnego (2 bity)

Bity te informują przełącznik, czy komórka służy do konserwacji sieci, czy jest częścią zwykłego ruchu. Dzięki temu komórki mogą być wykorzystywane w ten sam sposób, w jaki są używane pakiety ICMP i NSMP w sieciach TCP/IP - do testowania funkcjonalności i mierzenia wydajności.

1.15.4.4.4 Wskaźnik typu priorytetu (PTI) (1 bit)

Bit ten pozwala sieci ATM rozróżniać komórki o odmiennych wymaganiach odnośnie jakości usług. Konkretne znaczenie bitu może być określone dla komórki, kanału lub ścieżki.

1.15.4.4.5 Identyfikator ścieżki wirtualnej / Identyfikator kanału wirtualnego (VPI/VCI) (8 lub 12 bitów)

Pole VPI/VCI przechowuje ścieżkę i kanał punktu docelowego komórki. W ramach sieci ATM jest to informacja unikatowa.

1.15.4.4.6 Identyfikatory ścieżki wirtualnej (VPI), a identyfikatory kanału wirtualnego (VCI)

Identyfikator ścieżki wirtualnej opisuje całą trasę między punktami końcowymi. Identyfikator kanału wirtualnego dotyczy bezpośredniego połączenia miedzy dwoma przełącznikami. Za każdym razem, gdy przełącznik odbiera komórkę, może zmienić obie te wartości zanim przekaże ją do następnego przełącznika. Proste przejrzenie identyfikatorów VPI i VCI dostarcza przełącznikowi informacji potrzebnych do określenia portu docelowego i następnego skoku VPI/VCI.

1.15.4.5 Połączenia ATM

ATM zapewnia dwa rodzaje połączeń: stałe połączenia wirtualne (ang. PVC- Permanent Virtual Connections) i komutowane połączenia wirtualne (ang. SVC - Switched Virtual Connections). Stałe połączenia wirtualne są utrzymywane przez dostawcę usług ATM i działają w podobny sposób jak połączenia Frame Relay. Stałe połączenie wirtualne wymaga starannego projektowania sieci i wykorzystuje takie topologie jak linie dzierżawione.

Podczas zamawiania stałego połączenia wirtualnego należy ustalić z dostawcą kilka jego właściwości. Każdemu stałemu połączeniu wirtualnego przypisana jest klasa usług (ang. COS - C'lass of Service). Specyfikacje klasy usług obejmują wybór między zmienną lub stałą szybkością przesyłania bitów. Klasa usług jest określana za każdym razem, gdy tworzone jest komutowane połączenie wirtualne. Pozwala to uzyskać większą elastyczność. gdyż między organizacjami można utworzyć wiele komutowanych połączeń wirtualnych, każde o innej klasie usług.

Inną właściwością, którą należy określić, jest maksymalna liczba przesyłanych komórek (ang. PC'R - Peak Cell Rate). Komórki ATM mają ustalony rozmiar, więc maksymalna liczba przesyłanych komórek określa szerokość pasma przydzielonego danemu obwodowi. Co więcej, maksymalna liczba przesyłanych komórek może być inna dla każdego kierunku ruchu w obwodzie, czyniąc obwód niesymetrycznym. Maksymalna liczba przesyłanych komórek przypomina umowny wskaźnik informacji (CIR) określany dla obwodów Frame Relay.

Asymetria ruchu sieciowego występuje w wielu relacjach klient-serwer. Doskonałym przykładem jest popularny serwer sieciowy: klient wysyła do serwera proste żądanie GET/PUT, a serwer odpowiada przesyłając długą stronę tytułową. Jeśli takie byłoby typowe zastosowanie obwodu, to maksymalna liczba komórek mogłaby być mniejsza dla klienta niż dla serwera.

Komutowane połączenia wirtualne w dużej mierze wciąż są opracowywane. Umożliwią stacjom końcowym dynamiczne tworzenie połączeń i zmniejszą potrzebę planowania sztywnych topologii.

1.15.4.5.1 Jakość usług

Jakość usług (ang. QOS- Quality of Service) jest ważnym pojęciem w sieciach ATM. Termin ten opisuje właściwości ruchu sieciowego, takie jak maksymalna i długookresowa szerokość pasma, gwarantowane dla danego połączenia. Dane, które przekraczają te określone limity, mogą zostać porzucone (lub nie), w zależności od konfiguracji poszczególnych przełączników ATM.

Aby zapobiec utracie danych, przełączniki ATM wykorzystują kształtowanie ruchu dostosowujące dane do uprzednio określonych ograniczeń. Na przykład, jeśli węzeł końcowy wysyła większą ilość danych niż określone połączenie może obsłużyć, odbierający przełącznik ATM kolejkuje tyle danych, ile może, i wysyła je w ilości dopasowanej do ograniczeń pasma.

Alternatywą dla kształtowania danych jest strategia ruchu. Według tej strategii, dane przychodzące, które nie odpowiadają ograniczeniom, są porzucane lub oznaczane jako dane o niższym priorytecie. Komórki o niższym priorytecie mogą później zostać porzucone, jeśli w połączeniu wystąpi zator.

1.15.4.5.2 Sygnalizowanie

Stałe połączenia wirtualne w sieciach ATM są ustanawiane za pomocą sygnalizowania. Dla porównania, połączenia TCP/IP są ustanawiane za pomocą pakietów SYN i ACK, które są inną formą sygnalizowania. W celu ustanowienia połączenia router wysyła do routera docelowego komórkę sygnalizowania, zawierającą specjalne informacje, takie jak parametry jakości usług.

Każdy przełącznik znajdujący się na ścieżce sprawdza komórkę sygnalizowania i określa, czy jest w stanie obsłużyć połączenie. Jeśli tak, to wysyła komórkę w dalszą drogę do miejsca przeznaczenia; jeżeli nie. powiadamia o tym nadawcę, wysyłając odpowiednią wiadomość - i połączenie jest odrzucane.

Jeśli wszystko przebiega prawidłowo; komórka sygnalizowania osiąga miejsce przeznaczenia. Jeśli router docelowy również może sprostać określonym parametrom jakości usług, wysyła z powrotem wiadomość potwierdzającą. W drodze powrotnej każdy przełącznik odbiera potwierdzenie i odnotowuje utworzenie połączenia, przypisując mu wszelkie niezbędne zasoby. Gdy wiadomość dotrze do routera inicjującego połączenie, identyfikatory ścieżki i obwodu wirtualnego są zapamiętywane do późniejszego wykorzystania i połączenie zostaje ustanowione.

1.15.4.5.3 Zamawianie obwodów ATM

Zamawianie dowolnego rodzaju obwodu nigdy nie jest tak proste, jak być powinno. Dostawcy mogą być powolni. niekomunikatywni lub mogą nie dysponować pasmem odpowiednim dla Twoich potrzeb. Wybory, jakich dokonujesz podczas zamawiania obwodu, mają wpływ na użyteczność sieci WAN w przeciągu całego „czasu życia" obwodu. Czas poświęcony na staranne planowanie zawsze się kiedyś zwraca.

1.15.4.5.3.1 Dostęp fizyczny

W tanich połączeniach ATM o małej szerokości pasma routery mogą wykorzystywać łącze Tl. Multiplekser umożliwia przesyłanie tym łączem głosu, obrazu i danych, ale prawdopodobnie transmisje te napotkają ograniczenia pasma.

Dostęp do ATM może być zintegrowany z istniejącym łączem TI, dzięki technice znanej jako „emulacja Tl". Polega ona na podzieleniu łącza 1'3 na 28 pojedynczych kanałów. Organizacje posiadające infrastrukturę łącza T3 mogą zintegrować ATM z siecią WAN, nie korzystając z obwodów dedykowanych, chroniąc w ten sposób poczynione wcześniej

inwestycje.
Jeśli wymagane jest pasmo z zakresu między TI a T3, można łączyć wiele (kanałów Tl, uzyskując w ten sposób pasmo o szerokości 3 Mbps lub większej. Demultiplekser. czyli IMUX (ang. *Inverse multiplexer)* płynnie rozdziela informacje pomiędzy poszczególne obwody.
Powszechnie dostępna, ale droga linia dzierżawiona T3 zapewnia przyłączalność sieci ATM, która może działać z szybkością 45 Mbps. Pełną prędkość 155 Mbps można osiągnąć w obwodach OC3, ale nie są one dostępne wszędzie - jest to uzależnione od dostępności synchronicznej sieci światłowodowej (SONET).
Choć sieć ATM traci wiele ze swych zalet, gdy łączymy ją z siecią Frame Relay, to dostępna jest możliwość konwersji między tymi dwoma sieciami. Powinna ona zainteresować te organizacje, które chcą wykorzystać istniejące połączenia sieciowe i ochronić inwestycje poczynione w sieć Frame Relay. Brama między ATM i siecią Frame Relay tłumaczy adresy i formaty ramek.
Bramy TCP/1P dają użytkownikom sieci ATM wygodny dostęp do Internetu. Pakiety TCP/IP przeznaczone dla bramy są dzielone i umieszczane w komórkach ATM. Następnie są kierowane bezpośrednio do bramy TCP/IP, która ponownie zestawia oryginalne pakiety. kieruje je do Internetu i czeka na odpowiedź. Konwersacje są przechowywane w pamięci bramy, aby umożliwić przesłanie odpowiedzi do odpowiednich miejsc przeznaczenia.

1.15.4.6 Współdziałanie przy użyciu emulacji LAN

W rozdziale 11 pt. „ATM", przedstawiono emulację LAN w sieci ATM, czyli LANE. Emulacja LAN pozwala na szybką pracę w sieciach LAN, zapewniając funkcjonalność, której brak jest w sieci ATM: rozgłaszanie (nadawanie), rozróżnianie nazw itp. Łącząc szkielet ATM z emulacją sieci LAN, można bezpośrednio dostarczać do lokalnego komputera wideokonferencje i aplikacje wykorzystujące szerokie pasmo. Ani szkielet, ani topologia LAN nie stanowią wąskiego gardła, zaś jakość usług jest gwarantowana na całej długości połączenia.

1.15.4.7 Migrowanie do sieci ATM

Choć zastępowanie istniejących połączeń WAN połączeniami ATM oferuje wiele korzyści, to jednak jest skomplikowanym i drogim przedsięwzięciem. Na szczęście istnieje wiele sposobów czynienia procesu migracji łatwiejszym.
Jak w każdej migracji, dobrze jest robić tylko jeden krok naraz. Zacznij od zainstalowania obwodu ATM jako łączącego dwa miejsca kanału zapasowego dla obwodu sieci WAN. Jeśli po kilku tygodniach łącze pozostanie stabilne, zamień te obwody rolami, tak aby łącze ATM stało się podstawową metodą komunikacji, a stary obwód działał jako sprawdzone i wiarygodne łącze zapasowe.
Łącze po łączu, obwód po obwodzie, zastępuj istniejące połączenia WAN połączeniami ATM. W tym okresie koszty rosną, ponieważ płacisz za podwójną łączność, ale i tak są o wiele mniejsze niż koszty związane z przestojem sieci WAN. Gdy cała sieć okaże się stabilna i nadająca się do użytku, usuń stare połączenia, kończąc w ten sposób migrację.

1.15.5 Podsumowanie

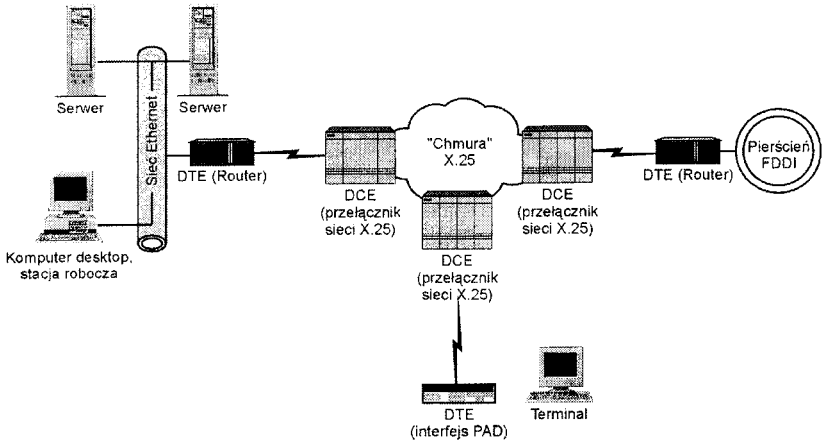
Technologie Frame Relay i ATM oferują dwie najpowszechniejsze metody łączenia oddalonych sieci. Zapewniając współdziałanie między sobą, sieciami X.25 i Internetem, pozwalają na wielką elastyczność w zakresie poszczególnych połączeń fizycznych. Obie sieci obsługują transmisje cyfrowe i analogowe, ale ATM został zaprojektowany specjalnie dla aplikacji czasu rzeczywistego, takich jak transmisje głosu i wizji.

1.16 Rozdział 16 Urządzenia transmisji w sieciach z komutacją pakietów

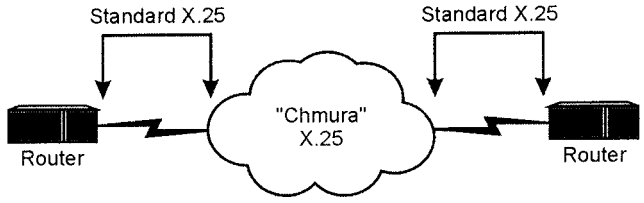
Tony Northrup
Transmisje w sieciach z komutacją pakietów różnią się od transmisji w sieciach z komutacją obwodów, opisanych w rozdziale 15 pt. „Urządzenia transmisji w sieciach z komutacją obwodów". Sieci z komutacją pakietów indywidualnie trasują każdy pakiet przez sieć, a nie przez wstępnie zadaną ścieżkę przełączników, jak to było w przypadku sieci z komutacją obwodów. Takie rozwiązanie oferuje większą elastyczność, ponieważ pakiety mogą być trasowane tak. by omijały miejsca uszkodzeń. Z drugiej jednak strony, trasa musi być obliczana oddzielnie dla każdego pakietu, a więc przełączanie w sieciach z komutacją pakietów jest znacznie wolniejsze niż w sieciach z komutacją obwodów.

1.16.1 Sieci X.25

Sieci X.25 są wykorzystywane przez klientów, którzy zawierają z operatorem kontrakt na wykorzystanie jego sieci z komutacją pakietów (ang. PSN - Packet Switched Network); związane z tym opłaty zależą zazwyczaj opłaty zależą od wykorzystania sieci, dlatego też mogą się zmieniać z upływem czasu.
Sam protokół X.25 określa interfejs pomiędzy węzłem sieci, takim jak router (urządzenie końcowe, inaczej DTE) a siecią z komutacją pakietów (urządzeniami przesyłania danych, inaczej DCE). Elementy te są przedstawione na rysunku 16.1. Sieć X.25 jest odpowiedzialna za transmitowanie ruchu do oddalonego partnera o podobnej konfiguracji, standard X.25 nie określa jednak sposobu, w jaki powinno to być przeprowadzane!
Rysunek 16.1. Elementy sieci X.25.



Każdy dostawca określa najwydajniejszą metodę przenoszenia ruchu między jego klientami. Tak więc bardziej stosowny niż „sieć X.25" byłby termin „połączenie X.25". Ponieważ jednak X.25 jest najbardziej znanym standardem, jego nazwa została przyjęta jako nazwa dla całej sieci. Rysunek 16.2 ilustruje zakres tego standardu.
Rysunek 16.2. Standard X 25.



Komunikacja rozpoczyna się, gdy jeden z klientów sieci X.25 z komutacją pakietów chce połączyć się z innym. Następuje wywołanie w podobny sposób, jak ma to miejsce w standardowej telefonii. Jeśli wywoływany system zaakceptuje połączenie, obydwa systemy zaczynają przesyłać między sobą dane. Po przesłaniu żądanych informacji jedna ze stron kończy połączenie, tak jak w rozmowie telefonicznej jedna ze stron odkłada słuchawkę.

1.16.1.1 Historia X.25

Sieć X.25 została pierwotnie zaprojektowana w roku 1976 przez spółki telefoniczne i innych dostawców sieciowych jako pewna metoda przesyłania danych cyfrowych liniami analogowymi. Ponieważ zaprojektowało ją kilka różnych spółek telefonicznych, specyfikacja X.25 współpracuje poprawnie z różnymi typami sieci przenoszących ruch. Od tamtej pory specyfikacja X.25 jest stale rozwijana, a nowelizacje wprowadzane są co cztery lata.

1.16.1.2 Zalety i wady sieci X.25

X.25 jest starym standardem, co jest zarówno jego największą zaletą, jak i wadą. Maksymalna szybkość obsługiwana przez X.25 wynosi 56 Kbps. Rzadko zdarza się, by taka szybkość odpowiadała dzisiejszym wymaganiom pracy sieciowej, ale wysoki stopień kompatybilności sprawia, że wielu klientów wciąż jest przyłączonych do „chmur" X.25. Jest to również najbardziej globalny standard - w większości miejsc na świecie są oferujący go dostawcy; w istocie administruje tym agenda ONZ, a mianowicie Międzynarodowa Unia Telekomunikacyjna (ang. ITU-T-Ihterrcational Telecommunications Union).
Sieć X.25 jest znacznie tańsza niż sieci najbardziej do niej podobne: Frame Relay i ATM. Syć może w tym przypadku prawdziwy jest stary aksjomat „dostajesz to, za co płacisz".

1.16.1.3 Najczęstsze zastosowania

Sieć X.25, choć wielu uważa ją za zbyt starą i powolną, by mogła być użyteczna, wciąż jest szeroko wykorzystywana w takich aplikacjach, jak systemy kart kredytowych, bankomaty, przetwarzanie wniosków kredytowych, bazy danych z zapisami medycznymi, monitorowanie stanu zapasów. Ostatnio w sieci X.25 udoskonalono przeprowadzanie komunikacji IP, wykorzystując enkapsulację, podobnie jak wykorzystuje się ją w sieciach ATM i Frame Relay.
Początkowo najczęściej dołączanymi do „chmury" X.25 węzłami były proste terminale tekstowe. Terminale były przyłączane do sieci za pośrednictwem usługi tłumaczącej znanej jako interfejs zestawiania i rozkładania pakietów (ang. PAD - Packet Assembler/ Disassembler). Interfejs PAD zapewnia połączenie z siecią i tłumaczy tekst dla/z terminalu na pakiety, które może zaakceptować urządzenie komunikacyjne sieci X.25. Poszczególne zalecenia wykorzystywane przez interfejs PAD opracował Międzynarodowy Komitet Konsultacyjny ds. Telefonii i Telegrafu (CCITT). Są to zalecenia X.28, X.3 i X.29.

1.16.1.4 Porównanie z modelem OSI

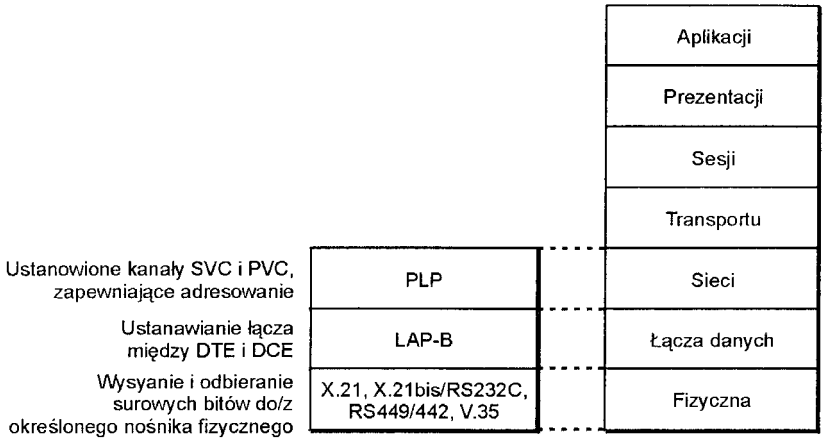
Protokół X.25 odpowiada trzem najniższym warstwom modelu OSI: fizycznej, łącza danych i sieci, co pokazuje rysunek 16.3. Standard nie specyfikuje żadnych wyższych warstw, takich jak aplikacje właściwe dla X.25.

1.16.1.4.1 Warstwa fizyczna

Rozmaite protokoły zapewniają fizyczne połączenia między urządzeniami końcowymi (DTE) a urządzeniami komunikacyjnymi (DCE). Można wśród nich wyróżnić między innymi protokoły X.21, X.21bis/RS232C, RS449/442 i V.35.

Rysunek 16.3. Porównanie standardu X 25
modelen a OSI.

1.16.1.4.2 Warstwa łącza X.25 w warstwie łącza danych modelu OSI



Odpowiednik warstwy łącza danych dla standardu X.25 jest opisywany przez oddzielny protokół - LAPB. LAPB (ang. Link Access Protocol Balanced - protokół symetrycznego dostępu do łącza) jest pewnym, kontrolującym błędy protokołem przesyłania ramek między urządzeniem końcowym i urządzeniem komunikacyjnym. Protokół ten jest odpowiedzialny za inicjalizowanie łącza między tymi urządzeniami, a także za umieszczanie pakietów w ramkach przed przekazaniem ich do warstwy fizycznej.

Protokół symetrycznego dostępu do łącza dość wiernie odpowiada warstwie łącza danych modelu OSI, ponieważ jest on wersją protokołu HDLC, który z kolei jest standardem OSI. Podobieństwo do protokołu HDLC można rozpoznać w strukturze danych, przedstawionej na rysunku 16.4.

Rysunek 16.4. Struktura ramki protokotu X.25.

8	8	8	Zmienne	16	8
Flaga 01111110	Adres	Kontrola	Dane	Suma kontrolna	Flaga 01111110

Standard LAPB obejmuje specjalne pola synchronizujące początek i koniec ramki. Flagi te zawierają specjalną sekwencję bitów, 01 1 I I 10 (dwójkowo) lub 126 (dziesiętnie). Urządzenia komunikacyjne na każdym z końców obwodu wirtualnego wiedzą, że ta szczególna sekwencja bitów nigdy nie pojawia się na linii wejściowej, chyba że oznacza początek lub koniec ramki. Próba przesłania takiego bajtu przez sieć X.25 spowoduje zastąpienie go innym znakiem specjalnym.

Drugim bajtem nagłówka ramki jest adres, który kieruje ramką między urządzeniami DCE i DTE.

Trzeci bajt, kontrolny, określa format ramki. Informuje, czy jest to ramka informacji, nadzoru, czy też ramka nienumerowana. Przenosi także numery sekwencyjne nadawania/ odbioru.

Bajtem następującym po polu danych jest sekwencja kontrolna ramki (ang. FCS- Frame Control Sequence), czyli suma kontrolna CRC. Gdy host odbiera ramkę, pierwszą czynnością jaką wykonuje jest sprawdzenie, czy suma kontrolna odpowiada danym zawartym w ramce. Jeśli ramka nie przejdzie tego testu, host wysyła pakiet odrzucenia (czyli pakiet REJ - ang. Reject) z powrotem do nadawcy, prosząc o ponowne wysłanie ramki. Jeśli suma kontrolna jest prawidłowa, wysyłane są pakiety gotowości odbiornika (czyli pakiety RR - ang. Receiver Ready) lub brak gotowości odbiornika (czyli RNR - ang. Receiver Not Ready), wskazujące, czy urządzenie może przyjąć więcej danych, czy nie.

Suma kontrolna jest kluczem do niezawodności sieci X.25, ale też przyczynia się do jej niskiej wydajności. Inaczej niż w większości protokołów, odpowiedzialność za transmitowanie pakietu spada na sieć. Tym sposobem sieć X.25 gwarantuje 100% niezawodność przesyłania danych między węzłami końcowymi. Odzwierciedla to całkiem inną filozofię niż w przypadku większości nowoczesnych protokołów, które zamiast przekazywać zawiadomienie o odrzuceniu bezpośrednio do hosta, który wysłał dane, czynią host odpowiedzialnym za ponowne wysłanie danych. Jest to istotne uzgodnienie, ponieważ wszystkie dane wysyłane przez sieć muszą być przechowywane w pamięci podręcznej sprzętu sieciowego aż do czasu, gdy odbiór zostanie potwierdzony.

Aby wyjaśnić tę różnicę, rozważmy sieć TCP/IP. Gdy użytkownik wysyła żądanie ściągnięcia strony sieciowej, polecenie to jest przekazywane z przeglądarki pracującej w warstwie aplikacji do protokołu TCP, działającego w warstwie sesji. Podczas dodawania informacji nagłówkowych, TCP przechowuje w pamięci dane otrzymane od użytkownika. Następnie dane są przekazywane ..w dół" modelu OSI,

do warstwy fizycznej, przez którą są przesyłane do węzła odbierającego. Gdy węzeł docelowy odbierze pakiety, sprawdza integralność danych, wwkorzystując pole CRC. Jeśli zostanie znaleziony błąd, węzeł odbierający prosi węzeł nadający o ponowne przesłanie danych, które ten zachował w pamięci podręcznej. Gdy prawidłowo przesłane pakiety zostaną potwierdzone, do węzła nadającego wysyłane jest zawiadomienie o tym fakcie i pakiety są usuwane z pamięci podręcznej.
W sieci X.25 urządzenie końcowe nie musi niczego przechowywać w pamięci, ponieważ uzgodnione jest, że sieć robi to, co jest konieczne, aby wszystko, co zostanie wysłane przez „chmurę", dotarło do odbierającego urządzenia końcowego. Narzut ów obciąża sprzęt sieciowy i ogranicza ruch możliwy do obsłużenia przez ten sprzęt. Nowoczesne sieci, takie jak ATM i Frame Relay, mocno różnią się w tym względzie od X.25; gwarantują one bowiem jedynie, że „zrobią co mogą", by dostarczyć transmisję. Wykrycie ewentualnego błędu i ponowienie transmisji jest już sprawą protokołów wyższych warstw. Protokół symetrycznego dostępu do łącza ma do wykonania trzy różne zadania. Pierwszym z nich jest konfiguracja łącza -inicjowanie połączenia przez obwód wirtualny. Drugim zadaniem jest przesyłanie informacji - faktyczne wysyłanie i odbieranie informacji użytecznej dla użytkowników końcowych. Gdy połączenie nie jest już potrzebne, protokół symetrycznego dostępu do łącza przeprowadza rozłączenie.
Standard LAPB określa trzy różne formaty ramek, o których przeznaczeniu decyduje wartość w polu Kontrola. Każdy typ ramki nadaje się do przenoszenia innego rodzaju danych. Ramki informacji przenoszą dane uporządkowane sekwencyjnie i nadające się do transmisji pełnodupleksowej. W tym celu wykorzystywana jest ramka numeru sekwencyjnego wysłania i numeru sekwencyjnego odbioru. Ramki nadzoru zapewniają kontrolę informacji. Są to ramki sterowania strumieniem danych i potwierdzania informacji. Ramki nienumerowane nie zawierają sekwencji informacji i służą do celów kontrolnych, takich jak rozpoczynanie i zatrzymywanie połączenia.

1.16.1.5 Poziom pakietu w warstwie sieci modelu OSI (X.25)

Poziom pakietu X.25 właściwie istnieje tylko w warstwie sieci, ale nazwa ta jest używana w odniesieniu do całego zestawu protokołów. Warstwa ta nawiązuje połączenie i zapewnia ustanowienie wywołania, transfer danych, sterowanie strumieniem danych, usuwanie błędów oraz kasowanie wywołania. Każdy pakiet może mieć maksymalnie 128 bajtów.
Warstwa sieci odpowiada za zarządzanie każdym obwodem wirtualnym i może jednocześnie utrzymywać 128 połączeń. Pola tworzące nagłówek warstwy sieci są przedstawione na rysunku 16.5.
Rysunek 16.5. Struktura pakietu X.25.

1	1	1	1	4	8	3	1	3	1	Zmienne
Q	D	0	1	Grupa (Group)	Kanał (Channel)	Odb (Rec)	M	Nad (Send)	0	Dane (Data)

(w nawiasie podane są nazwy angielskie)

Pierwszy bit, bit Q, wskazuje, czy dane są danymi kwalifikowanymi (ang. Qualified data). Jest on przeznaczony do wykorzystania przez zawarty w pakiecie protokół w celu oddzielenia pakietów sterowania od pakietów danych. Drugi bit, D, sygnalizuje, czy pakiet ma znaczenie lokalne i podróżuje między dwoma urządzeniami komunikacyjnymi (wartość 0), czy też globalne i podróżuje między dwoma urządzeniami końcowymi (wartość 1). Obydwa te bity, uzupełnione sekwencją `O1' tworzą czterobitowe pole identyfikatora formatu ogólnego, czyli pole GFI (ang. General Format Identifier).
Następne dwa pola to numer grupy kanałów logicznych (numer LGN - ang. Logical Channel Group Number) oraz numer kanału logicznego (numer LCN - ang. Logical Channel Number). Często te 12 kolejnych bitów traktuje się jako pojedyncze pole identyfikatora kanału logicznego (czyli pole LCI - ang. Logical Channel Identifier). Pole LCI wskazuje na określone połączenie z siecią z komutacją pakietów. Ostatni bajt nagłówka pakietu dzieli się na podpola numer odbioru i numer nadania, które są wykorzystywane w sekwencji śledzenia.
Jednym z bitów tego bajtu jest bit M, używany do grupowania bloków powiązanych informacji, które są zbyt duże, aby zmieściły się w pojedynczym pakiecie. M oznacza „More" („Więcej"). Wartość 1 informuje, że dane zostały podzielone między wiele pakietów. Bit przyjmuje wartość 0, gdy pakiet jest ostatnim z grupy. Proces ten może być znany z innych sieci; jest bardzo podobny do fragmentacji pakietów w sieciach IP.
Zalecenie CCITT X.121 definiuje dokładny format adresów warstwy sieci, czyli międzynarodowych numerów danych (numerów IDN - ang. lnternational Data Numbers). Format ten jest przedstawiony na rysunku 16.6. Adresy warstwy sieci identyfikują unikatowe miejsce docelowe w sieci X.25. Pierwsze cztery pola adresu IDN są kodem identyfikacyjnym sieci danych (ang. DNIC - Data Network Identificatio~ Code). Pierwsze pola cyfry kodu DNIC są specyficzne dla danego kraju, natomiast ostatnie pole identyfikuje określoną publiczną sieć komutowaną (czyli sieci PSN - ang. Public Switched
Network). Pozostała część adresu służy do identyfikowania określonego węzła w sieci i nosi nazwę krajowego numeru terminala (numeru NTN - ang. National Terminal Number).
Rysunek 16.6. Struktura adresu ‚Y".121.

4	4	3	1	10
Długość adresu wywoływanego DTE	Długość adresu wywołującego DTE	Kod kraju	PSN	Krajowy numer terminala

Po ustanowieniu połączenia adres X.121 nie jest już dłużej potrzebny. Zamiast niego do identyfikowania połączenia wykorzystuje się identyfikator kanału logicznego (LCI). Jest on podobny do identyfikatora DLCI w sieci Frame Relay, ponieważ ma znaczenie wyłącznie

lokalne. Adresy X.121 nie są w ogóle potrzebne w przypadku obwodów wirtualnych skonfigurowanych ręcznie - do identyfikowania określonego obwodu używa się identyfikatora kanału logicznego.
Protokół warstwy sieci jest odpowiedzialny za konfigurowanie między dwoma oddalonymi urządzeniami końcowymi połączeń sieciowych w „chmurze" X.25. Stanowi to kontrast wobec adresowania warstwy łącza, które łączy urządzenie końcowe z urządzeniem komunikacyjnym.
Warstwa sieci X.25 łączy urządzenia końcowe ze sobą; warstwa łącza - urządzenia końcowe z urządzeniami komunikacyjnymi.
Pojedyncze routery korzystają z usług protokołu X.25, by ustanowić połączenie ze zdalnym urządzeniem końcowym. Terminal nadający inicjuje połączenie, wysyłając do swojego urządzenia komunikacyjnego pakiet żądania wywołania (ang. call request packet), zawierający adres X.121 docelowego urządzenia końcowego. Urządzenie komunikacyjne odpowiada za przesłanie tego pakietu przez „chmurę" X.25. Gdy pakiet osiągnie miejsce przeznaczenia, odbierające urządzenie komunikacyjne kieruje go do odbierającego urządzenia końcowego, które sprawdza, czy pakiet jest przeznaczony właśnie dla niego, i decyduje, czy ma wziąć udział w konwersacji. Jeśli zaakceptuje wywołanie, wysyła pakiet akceptacji wywołania i w ten sposób zostaje ustanowione połączenie.
Po zakończeniu procesu ustanawiania połączenia obie strony są gotowe do wysyłania danych (w obydwu kierunkach). Od tego momentu, adres X.121 nie jest już używany. Połączenie ogranicza się do urządzeń komunikacyjnych i jest identyfikowane przez numer kanału logicznego.

1.16.1.6 Różne typy sieci

W ciągu wielu lat użytkowania sieci X.25, została ona przeniesiona do kilku różnych platform sieciowych. Każda z tych platform wymaga oddzielnego protokołu warstwy sieci.
Różne sieci nie mogą korzystać z tego samego interfejsu fizycznego, choć wielu dostawców obsługuje przyłączalność dwupunktową poprzez publiczną sieć transmisji danych.
Publiczna sieć transmisji danych (PDN)
Usługi X.25 PDN, zdefiniowane przez RFC 1356, zapewniają połączenia transmisyjne między klientami. Połączenia te są zwykle używane do przesyłania ruchu IP i OSI z wykorzystaniem enkapsulacji protokołu. Tym sposobem sieci X.25 stały się użyteczne raczej jako łącza WAN pomiędzy odległymi sieciami, a nie jako prosty środek transportu obsługujący ruch między terminalami.
Kiedy dostawca doprowadza ruch IP do Internetu, staje się dostawcą usług internetowych (ang. ISP - Internet Service Provider). Dostępna szerokość pasma w tego typu sieci jest zbyt mała dla wysokowydajnych transferów plików do i z Internetu, ale umożliwia wygodny dostęp do przedsiębiorstw posiadających łącze X.25.
Sieć transmisji danych Departamentu Obrony (DDN)
Sieć DDN jest częścią składową systemu komunikacji Departamentu Obrony i służy przede wszystkim do komutowania danych rządowych i łączenia odległych baz wojskowych. MILNET, część Internetu, a także sieci rządowe, klasyfikowane jako sieci o ściśle ograniczonym dostępie, tworzą razem sieć transmisji danych Departamentu Obrony. Siecią tą administruje Agencja Departamentu Obrony ds. Systemów Informacyjnych.
Punkt-z-punktem
Połączenie X.25 punkt-z-punktem (czyli połączenie dwupunktowe) jest w istocie prywatną siecią X.25 biegnącą między dwoma systemami. Systemy te mogą być połączone w dowolny sposób, poczynając od kabla szeregowego, a kończąc na dzierżawionej linii 56 Kbps. Może wydawać się dziwne, że ktoś mógłby wybrać protokół z komutacją pakietów do połączenia dwupunktowego, ale jest to prosty i niedrogi sposób rozszerzania istniejących sieci.

1.16.1.7 Specyfikacje X.25 (RFC 1356)

Trudno jest zapoznać się z rzeczywistymi specyfikacjami, na których opiera się standard X.25; w dodatku, jakby nie dość było komplikacji, powszechnie używany termin „X.25" w rzeczywistości określa zestaw standardów. Na szczęście, większość menedżerów sieci rzadko musi odwoływać się do rzeczywistych specyfikacji, aby połączyć się z sieciami X.25. Jedynie producenci sprzętu, projektujący wyposażenie sieciowe, oraz dostawcy oprogramowania, tworzący programy zgodne z X.25, muszą znać dokładne szczegóły opisywane w tych dokumentach.

1.16.1.7.1 ITU-T (dawniej CCITT)

Międzynarodowa Unia Telekomunikacyjna dostarcza specyfikacji (lub, jak je nazywa, „zaleceń") dla zestawu protokołów X.25. Można znaleźć je w Internecie pod adresem http: %/www. itu. int. Zaleceniami o szczególnym znaczeniu są Rec.X.25 Version 10/1996 i Rec.X.l21 Version 10/1996.

1.16.1.7.2 IETF

Grupa IETF (ang. Internet Engineering Task Force), czyli Grupa Robocza ds. Technicznych Internetu opracowała kilka standardów mających znaczenie dla środowiska sieci X.25. Zalicza się do nich zalecenia RFC 1356, 877 i 1236. Pełne teksty tych zaleceń są na bieżąco dostępne w Internecie.

1.16.1.7.3 RFC 877, transmisja datagramów IP w publicznej sieci transmisji danych

Jest to oryginalna specyfikacja dla enkapsulacji pakietów IP w sieciach X.25. Nakłada także ograniczenia na wykorzystywanie standardu X.25 do przesyłania datagramów IP.

1.16.1.7.4 RFC 1236, konwersja adresów IP na X.121 dla sieci DDN

Powyższe zalecenie określa sposób tłumaczenia adresów IP klasy B i C na adresy X.121.

1.16.1.7.5 RFC 1356, wieloprotokołowe połączenie X.25 i ISDN w trybie pakietu

Całkowicie zastępując oryginalne zalecenie RFC 877, RFC 1356 wyszczególnia metody wykorzystywane do przesyłania pakietów IP w sieciach X.25. Najważniejszą zmianą w stosunku do RFC 877 jest zwiększenie maksymalnego rozmiaru pakietu do 1600 bajtów.

1.16.1.8 Migrowanie z sieci X.25

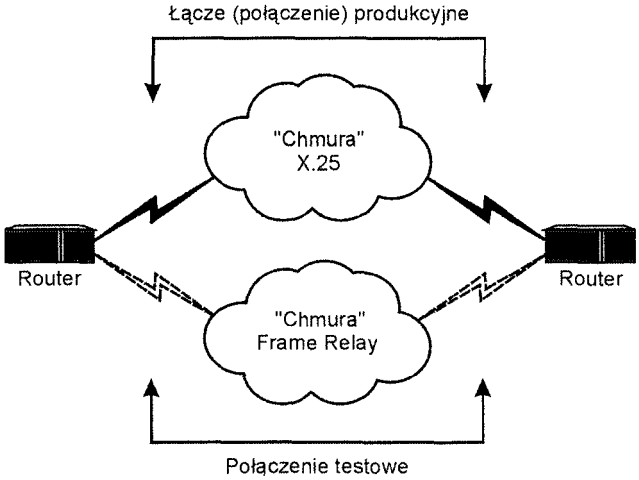
X.25 jest starzejącym się standardem i ma wiele słabych punktów, które sprawiają, że w nowoczesnych konfiguracjach nie jest popularnym protokołem. Faktycznie, wiele organizacji posiadających sieci X.25 uważa, że szerokość pasma jest dla nich ograniczeniem i spogląda w stronę nowych technologii: ATM i Frame Relay. Obie zapewniają usługi podobne jak w przypadku X.25, ale oferują większą szybkość i nie gwarantują niezawodności.

Kardynalną regułą migracji do innych sieci jest znana zasada „dziel i rządź". Wydziel część sieci i utwórz dodatkowe łącze, równoległe do istniejącego połączenia X.25, jak pokazuje rysunek 16.7. Przesyłaj te same dane, z tej samej aplikacji, za pomocą nowego obwodu, w taki sam sposób, jak dotychczas. Technikę tę stosuj przez kilka miesięcy, dopóki całkowicie nie upewnisz się, że nowa platforma sieci działa w sposób efektywny, niezawodny i stabilny.

Po osiągnięciu tego stanu zamień rolami obydwa połączenia, tak by sieć, do której migrujesz, stała się siecią podstawową, a X.25 siecią zapasową. Gdyby coś poszło źle, nie wahaj się powrócić do starego, pewnego połączenia - po to właśnie je pozostawiłeś.

Rysunek 16.7. Równoległe łącza X25 i Frame Relay.

Znów poczekaj kilka miesięcy, obserwując działanie nowego połączenia, zanim będziesz kontynuować migrację z pozostałej części sieci. Jeśli w nowej sieci wystąpi nierozwiązywalny problem, lepiej przekonać się o nim, tracąc jedną małą część sieci niż tracąc całą łączność. Gdy wszystkie jej części wejdą w pełny tryb produkcyjny, obwody X.25 mogą zostać usunięte.



Taka technika zapewnia oczywiste korzyści, umożliwiając przetestowanie nowego wyposażenia pod względem niezawodności. Daje również obsługującemu personelowi szansę przystosowania się do szybkości nowej technologii. Tworzenie mocnej bazy wiedzy jest kluczowym składnikiem każdej migracji; wiele problemów związanych z nowymi technologiami wynika raczej z błędów i niewystarczającego doświadczenia niż z samych technologii.

Wiele organizacji może nigdy nie zerwać całkowicie swoich związków z siecią X.25. Potrzebne może okazać się zachowanie połączeń umożliwiających utrzymywanie współpracy z organizacjami, które dopiero będą migrować, lub ze stanowiskami zdalnymi, którym migracja taka nie odpowiada. Aczkolwiek wrota X.25 do sieci Frame Relay i ATM zapewniają wsteczną kompatybilność, to nie można niestety gwarantować tego dla każdej aplikacji.

1.16.2 Podsumowanie

X.25 jest międzynarodowym standardem komunikacji o niskiej szybkości. Choć wielu ludzi w branży uważa, że jest to standard przestarzały, wciąż jest on szeroko wykorzystywany i bardzo popularny. Sieć X.25 uważa się za sieć z komutacją pakietów, ponieważ każdy pakiet jest trasowany oddzielnie, a nie kierowany na jedną ścieżkę wraz z całym ruchem w określonym obwodzie wirtualnym. Zestaw protokołów X.25 odwzorowuje trzy najniższe warstwy modelu OSI. Zapewnia niezawodną, sekwencyjną transmisję danych do dowolnego miejsca na świecie przy użyciu standardowego okablowania miedzianego.

1.17 Rozdział 17 Modemy i technologie Dial-Up

James M. Spann

W niniejszym rozdziale opisana jest zdolność innych urządzeń, znanych jako modemy, do rozszerzania zasięgu sieci. Kable i topologie przedstawione wcześniej pokazują, jak połączyć razem sieci na ograniczoną odległość. Aby komunikować się na większe odległości, trzeba rozważyć dodatkowe rozwiązania. Niniejszy rozdział przedstawia jedną z możliwych opcji.

1.17.1 Sposób działania modemu

Słowo „modem" jest telekomunikacyjnym terminem określającym urządzenia zmieniające sygnał elektroniczny komputera w sygnał tonowy, który można przesyłać łączem komunikacyjnym; morfologicznie jest ono zbitką słów „MOdulator" i „DEModulator". Modulowanie sygnału oznacza konwertowanie sygnału z jednego stanu do drugiego.

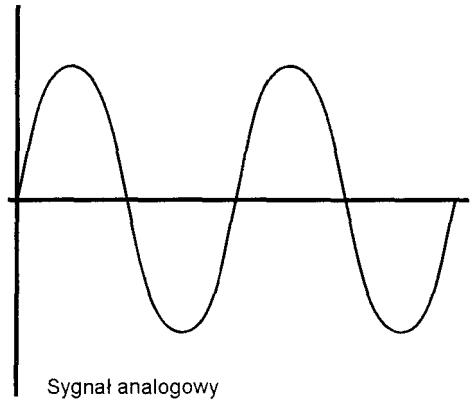
Słowo „modem" właściwie jest względnie nowym terminem w telekomunikacji. Oryginalnie urządzenia takie były znane jako „zestawy transmisji danych". Przed rokiem 1977 przedsiębiorstwa, które chciały podłączyć urządzenie do linii telefonicznej, musiały korzystać ze sprzętu dostarczanego przez spółki telefoniczne, które były wtedy monopolistami w tym względzie. Od roku 1977 na mocy zarządzenia Federalnej Komisji Łączności osoby fizyczne i prawne mogą korzystać z własnego wyposażenia służącego do połączeń z liniami telefonicznymi. Do dziś wszystkie produkowane w USA modemy wciąż muszą odpowiadać temu zarządzeniu FCC, znanemu jako „Part 68".

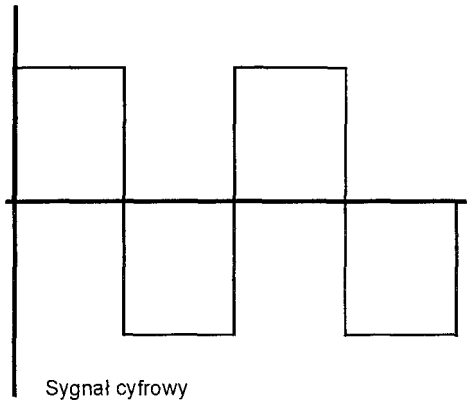
Komputer nie jest zdolny do bezpośredniego przyłączenia do linii telefonicznej, gdyż posługuje się ‚językiem cyfrowym", składającym się z elektronicznych sygnałów dyskretnych, podczas gdy zwykła linia telefoniczna może przesyłać tylko impulsy analogowe lub dźwięk. Sam komputer PC, będąc urządzeniem cyfrowym, porozumiewa się w języku znanym jako binarny, który koduje informacje wykorzystując albo stan „niski", reprezentowany przez zero albo stan „wysoki", reprezentowany przez jedynkę. Transmisję

cyfrową zawsze tworzą stany wysoki lub niski; nie ma tu stanów pośrednich. Z kolei transmisja analogowa ma stale zmieniającą się amplitudę natężenia sygnału oraz częstotliwość (liczbę zmian w ciągu jednego cyklu). Przykłady sygnałów cyfrowych i analogowych przedstawia rysunek 17.1.

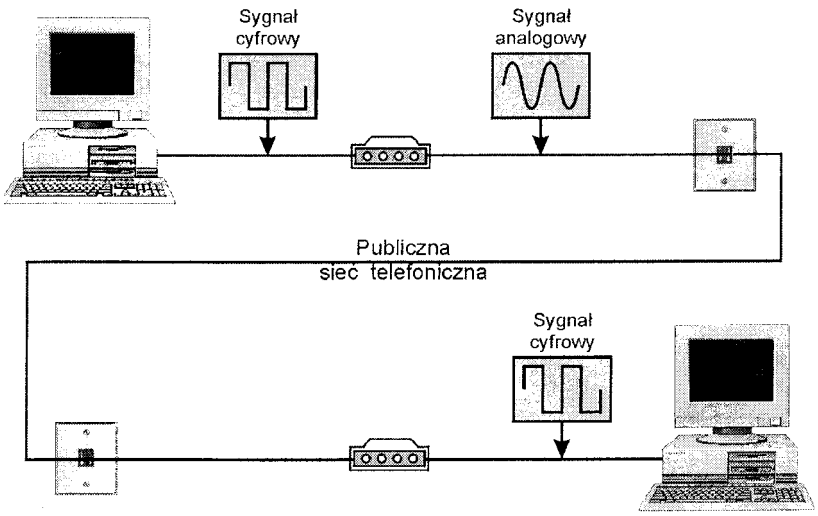
Rysunek 17.1. Sygnały cyfrowe
analogowe.

Zwykłe kable sieciowe są wystarczające przy mniejszych odległościach. Gdy komputery chcą komunikować się ze sobą na większe odległości, konieczne jest jednak wykorzystanie innego nośnika transmisji. Najbardziej logicznym i dostępnym wyborem jest standardowa linia telefoniczna. Niestety, jak już powiedziano, nie można wziąć komputera posługującego się językiem cyfrowym i przeprowadzić komunikacji z innym komputerem, jeśli używany nośnik komunikacyjny obsługuje wyłącznie transmisję analogową. W przypadku dwóch maszyn komunikujących się za pośrednictwem linii telefonicznej modemy muszą znajdować się na obydwu końcach linii. Komputer nadający musi posiadać modem, by zmienić tworzone przez siebie sygnały cyfrowe w sygnały analogowe, wysyłane linią telefoniczną. Również komputer odbierający musi mieć modem, aby odebrać transmisję analogową i przekształcić ją z powrotem na sygnał cyfrowy, który komputer może zrozumieć. Rysunek 17.2 przedstawia transmisję między dwoma komputerami z wykorzystaniem linii telefonicznej.





Rysunek 17.2. komunikacja przy użyciu linii
telefonicznej Digital.



Transmisja danych zachodzi zwykle między dwoma urządzeniami znanymi jako DCE (urządzenie przesyłania danych, urządzenie komunikacyjne) i DTE (urządzenie końcowe, terminal). Modem jest przykładem urządzenia DCE. Sam z kolei łączy się z komputerem osobistym, który jest terminalem DTE. Różnorodność dostępnego wyposażenia i rozmaitość aplikacji po obydwu stronach interfejsu DTE/DCE zwiększa potrzebę znormalizowania jego charakterystyk mechanicznych, elektrycznych i funkcjonalnych. Niektóre ze standardów są omawiane w dalszej części tego rozdziału.

1.17.2 Bity i body

Aby modem mógł modulować sygnał cyfrowy otrzymany z komputera na sygnał składający się z tonów analogowych, który można przesłać linią telefoniczną, musi mieć jakąś metodę zmieniania amplitudy, częstotliwości i fazy przesyłanych tonów. Liczba zmian jednego z wymienionych stanów linii w ciągu jednej sekundy nazywana jest „szybkością modulacji".

Termin „bod" służy do przedstawiania mogących wystąpić unikatowych stanów linii. Każdy unikatowy stan linii w większości nowoczesnych dokumentacji bywa określany mianem „symbolu". Każdy z tych symboli może reprezentować określony wzór bitów informacji. Reprezentowany wzór bitowy często określany jest jako „token". Sama nazwa „bod" pochodzi od nazwiska francuskiego wynalazcy Emila Baudot, który w 1875 roku stworzył 5-bitowy kod służący reprezentowaniu alfabetu. Każdy 5-bitowy wzór był tokenem reprezentującym jedną z liter alfabetu.

Niestety, trzeba dostosować się do pewnych ograniczeń, jakie niesie ze sobą fakt, że głównym przeznaczeniem publicznej sieci telefonicznej jest przesyłanie sygnałów dźwiękowych, a nie danych. Ponieważ większość energii wytwarzanej przez ludzki głos przypada na częstotliwości od 300 do 3300 herców (Hz), obwody komunikacyjne zaprojektowano tak, by przenosiły częstotliwości z zakresu od 0 do 4000 Hz, czyli 4 kHz.

We wczesnej erze transmisji modemowej wielu ludzi uważało, że terminy „bity na sekundę" i „body" są równoważne. Było to prawdą, ale tylko ze względu na sposób wysyłania sygnału. Kiedy każdy bod (lub zmiana stanu linii) w komunikacie reprezentuje jeden bit informacji do wysłania, szybkość modulacji wyrażana w bodach jest równa liczbie bitów na sekundę. Wyobraźmy sobie komunikat transmitowany z szybkością 300 bodów (300 zmian stanu linii na sekundę). Jeśli każda zmiana stanu linii reprezentuje jeden bit, wtedy komunikat jest transmitowany z szybkością 300 bitów na sekundę (bps).

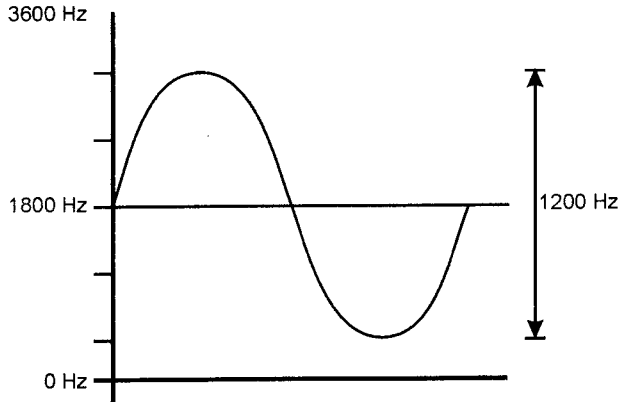
Aby wysłać więcej niż jeden bit informacji na jeden bod, opracowano schematy modulacji pozwalające na osiąganie większych szybkości transmisji w bitach na sekundę. By wysłać dane obwodem telefonicznym, modem zwykle używa częstotliwości nośnej położonej mniej więcej w połowie pasma głosowego i nakłada na tę częstotliwość dane, odpowiednio ją modulując. Na wyjściu procesu modulacji powinno być spektrum przypominające obciętą część całego spektrum akustycznego, mieszczącą się w paśmie telefonicznym.

Wyobraź sobie, że chcesz wysłać 2400 symboli na sekundę. Aby to zrobić, musisz wytworzyć 2400 cykli na sekundę (Hz). By wykorzystać częstotliwość 2400 Hz do wysłania transmisji, musisz wybrać 1800 Hz jako częstotliwość nośną i modulować ją w górę lub w dół 0 1200 Hz, tak by cały zakres wynosił 2400 Hz. Daje to całkowitą szerokość pasma pomiędzy 600 Hz a 3000 Hz, co przypada na dopuszczalny zakres pasma typowej linii telefonicznej (0-4000 Hz). Zakres ten jest przedstawiony na rysunku 17.3. Co ciekawe, taka forma modulacji jest wykorzystywana przez standard V.32, omawiany w dalszej części tego rozdziału.

Rysunek 17.3. Pasmo o szerokości 2-100 Hz
modulowane na częstotliwości nośnej 1200 Hz.

Z czasem jakość sieci telefonicznej uległa poprawie, co dało większą użyteczną szerokość pasma w typowej linii telefonicznej. Przykładem może tu być standard V.34, omawiany w dalszej części rozdziału, który wykorzystuje częstotliwość nośną 1959 Hz i moduluje 3429 symboli na sekundę. Całkowite pasmo zawiera się w przedziale od 244 Hz do 3674 Hz.

Aby obwód głosowy wydajnie przenosił dane, technika modulacji powinna dopasowywać kształt fali do charakterystyki kanału. Proces dopasowywania jest źródłem wielu pomysłowych rozwiązań. Technika modulacji musi być tak obmyślona, aby maksymalizowała



ilość transmitowanych danych i minimalizowała efekty szumów i zakłóceń. Na podstawie wielu międzynarodowo uznanych metod i szybkości transmisji ustalono dzisiejsze standardy transmisji danych.

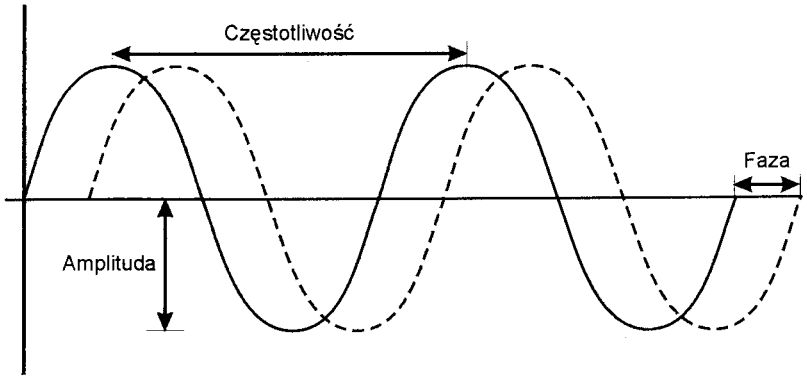
Różne schematy modulacji pozwalają, aby stany linii reprezentowały różne pary cyfr binarnych. Wykorzystywanie schematu modulacji, który pozwala czterem różnym częstotliwościom reprezentować pary binarne 00, Ol, 10 i 11, w istocie umożliwia dostarczanie 1200 bitów na sekundę, nawet jeśli szybkość modulacji wynosi tylko 600 bodów. Większe prędkości transmisji można osiągnąć, pozwalając, by większa liczba stanów linii (lub „znaków") reprezentowała więcej kombinacji bitów.

1.17.3 Typy modulacji modemów

Proces wykorzystywania nośnika - w tym przypadku linii telefonicznej - do przenoszenia informacji między dwoma punktami określa się mianem modulacji. Analogowy dźwięk głosu ludzkiego stale zmienia swoją częstotliwość i amplitudę. Jeśli przyjrzeć się jego obrazowi na oscyloskopie, ludzki głos przypomina szereg sinusoid. Każdy szereg impulsów, dźwięków. fal lub napięć może być przedstawiony za pomocą sinusoidy. Poszczególne sinusoidy charakteryzują się określoną amplitudą, częstotliwością i fazą.

Amplituda sinusoidy jest jej wysokość, której odpowiada wartość na osi rzędnych (osi y). Częstotliwość fali sinusoidalnej to - w przybliżeniu - szybkość powtarzania się przebiegu na osi odciętych (osi x). Faza sinusoidy jest brana pod uwagę tylko wtedy, gdy porównujemy ją z inną sinusoidą o takiej samej częstotliwości i amplitudzie - stanowi ona po prostu wielkość przesunięcia (na osi x) pomiędzy obydwiema sinusoidami. Rysunek 17.4 przedstawia podstawowe wielkości charakterystyczne fali sinusoidalnej.

Rysunek 17.4. Charakterystyki fali sinusoidalnej.



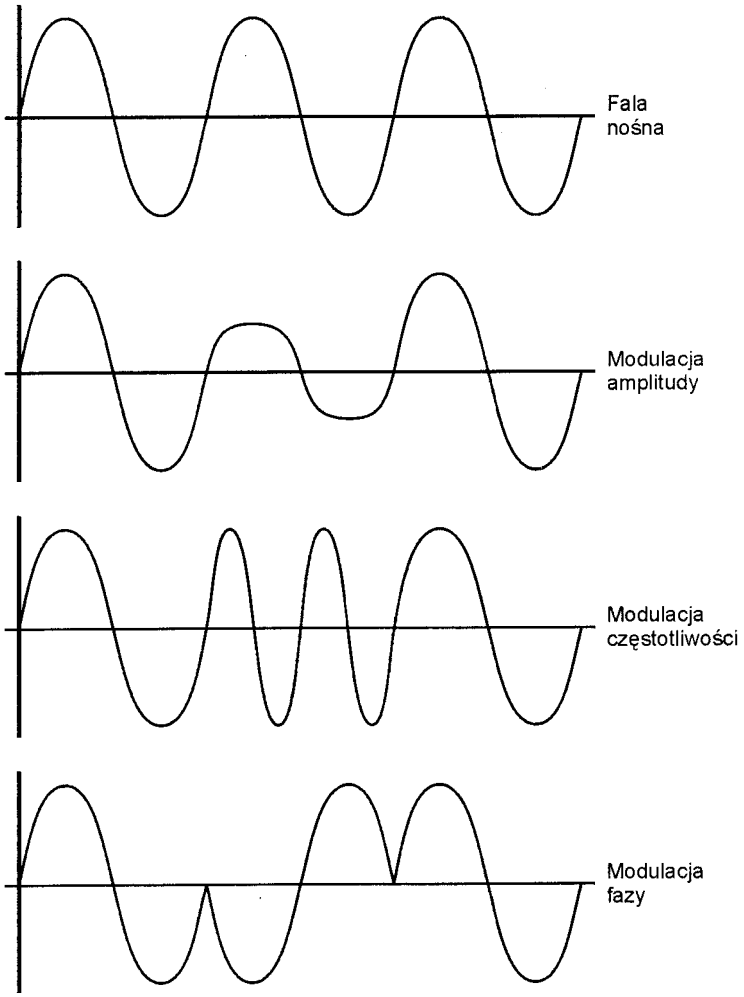
Modemy wysyłają informacje, modulując falę nośną, czyli zmieniając jedną z określonych charakterystyk sinusoidy. Modulacja analogowa może występować w trzech postaciach: modulacji amplitudy, częstotliwości lub fazy. Techniki te przedstawia rysunek 17.5.

Rysunek 17.5. Trzy rodzaje technik modulacji
analogowej.

Jak widać na rysunku 17.5, pierwszym typem modulacji analogowej jest modulacja amplitudy. Modulacja amplitudy różnicuje poziom sygnału, by określić, czy przesyłany bit jest zerem, czy jedynką. Na rysunku 17.5 sygnał o niskiej amplitudzie reprezentuje zero lub „odstęp", zaś sygnał o wysokiej amplitudzie przedstawia jedynkę lub „znak".

Drugim typem modulacji analogowej jest modulacja częstotliwości. Aby wskazać zero lub jedynkę, wykorzystuje ona zmianę liczby powtórzeń fali sinusoidalnej zachodzących w ciągu jednej sekundy. Na rysunku 17.5 sinusoida o określonej częstotliwości reprezentuje zero lub „odstęp", a sygnał o częstotliwości dwa razy większej reprezentuje jedynkę lub „znak".

Ostatnim typem modulacji jest modulacja fazy. Wykorzystuje ona zmianę w fazie fali, by przedstawić zero lub jedynkę. Faza sinusoidy to jej względna pozycja na osi odciętych (osi x). Na rysunku 17.5 sinusoida przedstawia zero, a przesunięcie w fazie wskazuje jedynkę (wielkość przesunięcia w fazie wynosi tu 180° - przyp. red.).



Kombinacje tych trzech form modulacji mogą być wykorzystywane do tworzenia większej liczby możliwych symboli, a co za tym idzie, większej przepustowości. Przykładem tego jest kwadraturowa modulacja amplitudy (ang. QAM- Quadrature Ampditude Modulation). W modulacji QAM, kombinacja przesunięć fazy i zmian amplitudy może reprezentować 16 różnych stanów, co pozwala wysyłać cztery bity na jeden bod. Jeśli szybkość modulacji w bodach wynosi 2400, można efektywnie przesłać 9600 bps (2400 zmian na sekundę x 4 bity na zmianę = 9600 bps).

Inną techniką modulacji jest impulsowa modulacja amplitudy (ang. PAM- Pulse Amplitude Modulation). Modulacja PAM jest metodą stosowaną w połączeniach nowych modemów 56 Kbps - nie wymagają one wstępnej konwersji sygnału analogowego na cyfrowy w swoich strumieniach danych (w przeciwieństwie do starszych modemów, które taką modulację muszą wykonywać). Nowe modemy nadal używają modulacji QAM, by odsyłać dane z powrotem przy szybkościach do 33,6 Kbps. Więcej informacji o technologii 56 Kbps można znaleźć w opisach standardów przedstawionych na końcu tego rozdziału.

1.17.3.1 Asynchronicznie i synchronicznie

Kiedy dwa urządzenia przystępują do komunikacji między sobą, muszą. w pewien sposób sterować przepływem danych, tak aby wiedziały, gdzie rozpoczynają się i kończą wysyłane znaki. Strumień danych wysyłany przez modem na drugi koniec połączenia może wykorzystywać jedną z dwóch form koordynacji. Jednym ze sposobów sterowania taktowaniem sygnałów wysyłanych i odbieranych na obydwu końcach jest asynchroniczne przesyłanie danych.

Komunikacja asynchroniczna jest najbardziej rozpowszechnioną formą stosowaną w konwencjonalnych modemach. W takiej komunikacji informacja (znak, litera, liczba lub symbol) przesyłana z jednego urządzenia do drugiego jest przedstawiana jako strumień bitów. Każdy strumień bitów jest oddzielony od innych bitem startu i bitem stopu. Dzięki stosowaniu bitu startu i bitu stopu dla każdego transmitowanego znaku, urządzenie wie, kiedy wysyła lub odbiera znak, i nie są potrzebne zewnętrzne sygnały taktujące, które sterowałyby przepływem danych.

Jednym z zastrzeżeń wobec komunikacji asynchronicznej jest to, że około 20 do 25 procent przesyłanych danych służy jako informacja sterująca do „synchronizowania" konwersacji między urządzeniami.

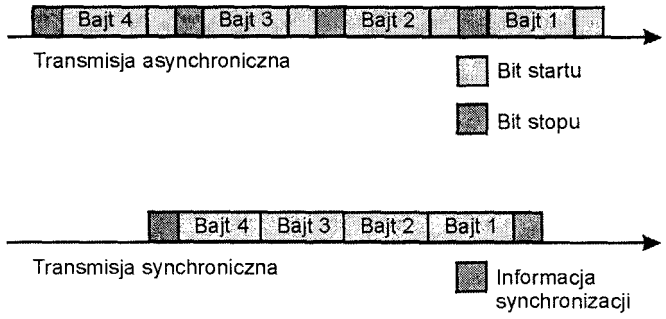
Alternatywą dla komunikacji asynchronicznej jest komunikacja synchroniczna. W komunikacji synchronicznej musi występować sygnał taktujący, sterujący transmisją bloków znaków, zwanych „ramkami". W transmisji nie używa się bitów startu i stopu. Znaki synchronizacji służą do rozpoczęcia transmisji oraz do sprawdzania jej dokładności. Rysunek 17.6 przedstawia porównanie komunikacji asynchronicznej i synchronicznej. Protokoły wykorzystywane w transmisjach synchronicznych spełniają funkcje nieobecne w protokołach asynchronicznych. Przykładami takich funkcji mogą być:

• Kontrolowanie dokładności wysyłania informacji • Formatowanie danych w ramki

• Dodawanie informacji sterujących

Rysunek 17.6. Porównanie komunikacji asynchronicznej i synchronicznej.

Protokoły synchroniczne są używane w środowiskach cyfrowych. Świat analogowy zwykle wykorzystuje komunikację asynchroniczną. Większość komunikacji sieciowej odbywa się w sposób synchroniczny. Do najpopularniejszych protokołów synchronicznych należą: protokół bisynchroniczny (czyli binarny synchroniczny protokół komunikacji ang. Binary Synchronous Communication Protocol, sterowanie synchronicznym łączem transmisji danych (sterowanie SDLC - ang. Synchronous Data Link Control) oraz sterowanie wysokopoziomowym łączem transmisji danych (sterowanie HDLC - ang. High-Level Data Link Control.



1.17.4 Standardowe interfejsy modemów

Choć istnieją liczne „zalecane standardy", bez wątpienia najważniejszym w świecie technologii modemów jest interfejs RS-232. Istnieje kilka wersji tego interfejsu, a każda z nich jest wyróżniona literą występującą po oznaczeniu RS-232. Najpowszechniejszą implementacją standardu interfejsu RS-232 jest wersja RS -232C.

Specyfikacja RS-232 jest standardem stowarzyszenia przemysłu elektronicznego, czyli stowarzyszenia EIA (ang. Electronic Industries Association). Jest to standard bardzo podobny do standardu V.24 opracowanego przez ITU-T (sekcję standardów telekomunikacyjnych Międzynarodowej Unii Telekomunikacyjnej), znaną dawniej jako CCITT. Na standard RS-232 składają się cztery podstawowe obszary informacji:

• mechaniczne charakterystyki interfejsu,

• sygnały elektryczne wykorzystywane w interfejsie, • funkcje wszystkich sygnałów,

• podziały sygnałów dla specyficznych aplikacji.

RS-232 i jego odpowiedniki zapewniają szeregową transmisję danych przez interfejs. W interfejsie szeregowym bity tworzące dane są wysyłane bit po bicie, synchronicznie lub asynchronicznie.

Choć złącze DB-25 jest powszechnie kojarzone z interfejsem RS-232C, w rzeczywistości nie zostało ono zdefiniowane jako część nowelizacji C; określiła je nowelizacja D

specyfikacji RS-232. Każdemu z pinów złącza przypisano pewną funkcję związaną z przesyłaniem różnego rodzaju sygnałów. Oto one:

• Pin l: uziemienie ochronne (gdzie masą jest obudowa) (pin PG - Protective Ground)

• Pin 2: wysyłanie danych (pin TD - Transmit Data) • Pin 3: odbieranie danych (pin RD - Receive Data)

• Pin 4: żądanie nadawania (pin RTS - *Request to Send)*

• Pin 5: gotowość modemu do transmisji (pin CTS - Clear to Send) • Pin 6: przyłączenie modemu do sieci (pin DSR - Data Set Ready) • Pin 7: sygnał masy (pin SG -Signal Ground)

• Pin 8: wykrywanie nośnika danych (pin DCD - *Data Carrier Detect)* • Pin 9: zarezerwowany

• Pin 10: zarezerwowany • Pin 1 1: nie przypisany

• Pin 12: wykrywanie wtórnego nośnika danych (pin SDCD - Secondary Data Carrier Detect)

• Pin I 3: wtórna gotowość do nadawania (pin SCS - Secondary Clear to Send) • Pin 14: wtórne wysyłanie danych (pin STD - Secondary Transmit Data)

• Pin I 5: zegara nadawania (pin TC - Transmit Clock)

• Pin 16: wtórne odbieranie danych (pin SRD - Secondary Receive Data) • Pin 17: zegar odbioru (pin RC - Receive Clock)

• Pin 18: nie przypisany

• Pin 19: wtórne żądanie nadawania (pin SRS - Secondary Request to Send)

• Pin 20: gotowość komputera do transmisji danych (DTR-Data Terminal Ready) • Pin 21: określanie jakości odbioru (pin SQD - ang. Signal Quality Detector) • Pin 22: wywołanie stacji (pin RI - Ring Indicator)

• Pin 23: wybór szybkości transmisji (pin DRS - Data Rate Select) • Pin 24: zegar zewnętrzny (pin EC - External Clock)

• Pin 25: nie przypisany

Dla przeprowadzenia transmisji danych konieczne jest wystąpienie następujących zdarzeń: • Oprogramowanie komunikacyjne komputera PC podaje napięcie na pin 20 (DTR), aby wskazać, że komputer jest gotów do wysłania danych. W tym samym czasie modem podaje napięcie na pin 6 (DSR), by powiadomić komputer, że modem może przyjmować dane lub instrukcje.

• Komputer wysyła do modemu przez pin 2 (TD) polecenie uaktywnienia linii (żeby modem „podniósł słuchawkę") i wybrania określonego numeru. Modem odpowiada potwierdzeniem na pinie 3 (RD).

• Po ustanowieniu połączenia pomiędzy modemami, modem wysyła sygnał do komputera na pinie 8 (CD), powiadamiając go, że istnieje ścieżka komunikacyjna i można rozpocząć przesyłanie danych.

• Gdy komputer jest gotów do wysłania danych, sygnalizuje to modemowi poprzez pin 4 (RTS), nazywany „żądaniem nadawania". Jeśli modem nie jest zajęty, odpowiada komputerowi sygnałem na pinie 5 (CTS), powiadamiając go, że może rozpocząć nadawanie danych pinem 2 (TD).
Główną wadą interfejsu RS-232 jest ograniczenie odległości do 15 metrów (50 stóp). Zwykle nie jest to problemem dla połączenia między komputerem a modemem, ale może się nim stać, jeśli modem musi być umieszczony w pewnej odległości od komputera. Ponieważ zwykłym limitem przepustowości interfejsu RS-232 jest 19200 bps, musisz używać krótszego kabla, jeśli chcesz osiągnąć wyższą przepustowość, możliwą przy dzisiejszych modemach. Większość producentów modemów zaleca stosowanie kabla RS-232 o długości 4 m (12 stóp) lub krótszego, a bardzo często używa się kabla o długości 2 m (6 stóp).
Aczkolwiek RS-232 jest najpowszechniej wybieranym interfejsem dla połączeń modemowych, należy wspomnieć również o kilku innych. Są nimi RS-422, RS-423; RS-449 i RS-530.
Standard RS-422 i jego odpowiednik, X.27 (V.11), obejmują elektryczne charakterystyki obwodów zrównoważonych (różnicowych), czyli takich, w których dodatnie i ujemne linie sygnału są odizolowane od masy. Obwód zrównoważony jest mniej podatny na zakłócenia, oferuje większą szybkość transmisji i większą długość kabli.
RS-422 jest przeznaczony dla aplikacji wykorzystujących skrętkę dwużyłową na odległość do 1200 m (4000 stóp) i przy szybkości transmisji do 100 000 bps. Przy odległości 12 m (40 stóp) lub mniejszej można osiągnąć szybkość 10 000 000 bps. Takie charakterystyki umożliwiają połączenie urządzeń w obrębie zakładu bez potrzeby korzystania z drogich urządzeń do transmisji danych.
Standard RS-423 i jego odpowiednik, X.26 (V.10), określają niezrównoważone charakterystyki elektryczne, podobne do charakterystyk interfejsu RS-232. Nowy standard pozwala jednak na przesyłanie danych z szybkością od 100 000 bps na odległość do 12 m, a na odległość do 60 m - z prędkością do 10 000 bps. Szybkość transmisji danych obecnego standardu niezrównoważonego jest z grubsza ograniczona do 20 000 bps na odległości do 15 metrów.
Standardy RS-422 i RS-423 określają tylko elektryczne charakterystyki interfejsu, natomiast standard towarzyszący, RS-449, określa funkcjonalne i mechaniczne wymagania dla implementacji. Choć w zamierzeniu te nowe standardy miały zastąpić RS-232, jak dotąd tak się nie stało.
RS-449 i standardy towarzyszące znacznie różnią się od starszego standardu RS-232. 10 nowych funkcji na poziomie sterowania interfejsem umożliwia testowanie, wybieranie szybkości i działania rezerwowe. Być może najbardziej znaczącymi nowymi funkcjami są lokalne i zdalne sygnały pętli zwrotnej. Pozwalają one do pewnego stopnia diagnozować błędy sprzętu oraz obwodu przez udostępnienie pętli zwrotnej do urządzenia końcowego, do analogowej części lokalnego urządzenia komunikacyjnego lub do cyfrowej części zdalnego urządzenia końcowego.
RS-530 został wprowadzony jako standard działający przy szybkościach transmisji od 20000 bps do 2000000 bps i wykorzystujący takie same 25-pinowe złącze DB-25, jak standard RS-232. Głównym ulepszeniem w standardzie RS-530 jest to, że nie jest on specyfikacją elektryczną lecz raczej odwołuje się do dwóch innych standardów, RS422 i RS-423. Te nowe standardy wykorzystują zwiększoną wydajność, możliwą teraz dzięki technologii obwodów zintegrowanych.

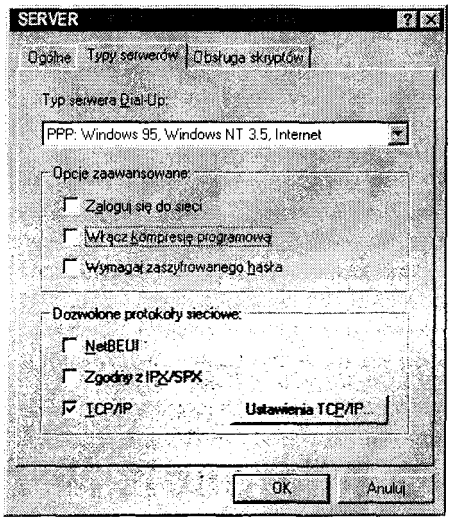
1.17.5 Standardy ITU-T (CCITT) modemów

Jedna z najbardziej podstawowych reguł dotyczących komunikacji brzmi następująco: aby komunikacja była efektywna i wydajna, muszą istnieć reguły, dzięki którym urządzenia komunikujące się rozpoznają wszystkie elementy uczestniczące w procesie komunikacji. Te reguły porządkujące są znane w świecie przesyłania danych jako „protokoły" i „standardy".
W środowisku modemów istnieje wiele standardów, z których większość została przyjęta przez ITU-T. Standardy te są powszechnie znane jako „zalecenia serii V". Poniżej wymieniono kilka spośród najważniejszych zaleceń serii V.
• V.22: Standard V.22 określa dupleksowy modem 1200 bps przeznaczony do wykorzystywania w publicznej komutowanej sieci telefonicznej oraz w obwodach linii dzierżawionych. Strukturalnie jest on podobny do standardu Bell System 212A, ale nie jest kompatybilny wstecz ze standardem 212A o szybkości transmisji rzędu 300 bps. Kolejna różnica polega na tym, że standard 212A przy niższej szybkości 300 bps wykorzystuje modulację kluczem (kluczowanie) z przesuwem częstotliwości (czyli modulację FSK - ang. Frequency Shift Keyed, natomiast standard V.22 dla swojej niskiej szybkości (już nie 300 bps, lecz 600 bps) używa kluczowania z przesuwem nie częstotliwości, lecz fazy (czyli modulację PSK - Phase Shift Keyed).
• V.22 bis: Standard V.22 bis („drugi") opisuje dupleksowy modem działający z szybkością 2400 bps, wykorzystujący technikę podziału częstotliwości przystosowaną do publicznej komutowanej sieci telefonicznej. Może być także używany w połączeniach dwupunktowych w dwuprzewodowych obwodach linii dzierżawionych.
• V.26: Standard V.26 określa modem 2400 bps dla obwodu linii dzierżawionych. Jest to pełnodupleksowy modem o szybkości modulacji 1200 bodów, stosujący modulację dwubitową kluczowaną z przesuwem fazy (czyli modulację DPSK - Dibit Phase Shift Keyed.
• V.26 bis: Standard V.26 bis określa modem pracujący z szybkością 2400 bps lub 1200 bps, wykorzystywany w instalacjach nie dzierżawionych.
• V.26 tertio: Standard V.26 tertio („trzeci") dodał do poprzednich standardów przede wszystkim technikę niwelacji odbić.
• V.27: Modemy zgodne ze standardem V.27 pracują z szybkością 4 800 bps, wykorzystują różnicowe kluczowanie z przesuwem fazy i mogą działać w trybie pełnodupleksowym lub półdupleksowym. Wyposażone są także w ręcznie regulowany korektor. Wykorzystywana technika modulacji rozpoznaje osiem różnych faz i jest przeznaczona dla instalacji dedykowanych lub dzierżawionych. Wersja trzecia standardu V.27 umożliwia połączenie z instalacjami nie dedykowanymi.
• V.29: Modemy standardu V.29 to z kolei pełnodupleksowe, czteroprzewodowe modemy pracujące z szybkością 9 600 bps, przeznaczone dla instalacji dedykowanych. Częstotliwością nośną modemów tego typu jest 1 700 Hz, a ich szybkość modulacji wynosi 2 400 bodów.
• V.32: Standard V.32 opisuje rodzinę dwuprzewodowych, dupleksowych modemów, działających z szybkością do 9 600 bps. Modemy te mogą być używane w publicznej komutowanej sieci telefonicznej, jak również w obwodach linii dzierżawionych. Zgodne z tym standardem

modemy wykorzystują modulację kwadraturowo-amplitudową (modulację QAM - ang. Quadrature Amplitude Modulation), a dzięki stosowaniu techniki niwelacji odbić mogą działać pełnodupleksowo.
• V.32 bis: Standard V.32 bis jest bardzo podobny do opisanego powyżej standardu V.32. Jedyną ważną różnicą jest szybkość. Modemy te mogą komunikować się z szybkością do 14 400 bps, osiągniętą dzięki przesyłaniu 6 bitów z prędkością 2 400 bodów.
• V.32 tertio: Specyfikacja V.32 tertio właściwie nie jest standardem, ale propozycją firmy AT&T. Pozwala na komunikację z szybkością do 19 200 bps i działa między dwoma modemami obsługującymi tę specyfikację.
• V.33: Choć standard V.33 jest bardzo podobny do standardu V32 dla usług komutowanych, to jest przeznaczony dla instalacji dedykowanych. Inną ważną różnicą jest to, że V.33 posiada opcję multipleksowania, pozwalającą łączyć klika źródeł danych w jeden strumień danych o szybkości 14400 bps.
• V.34: Standard V.34 miał pierwotnie wyspecyfikować szybkości transmisji sygnałów do 28800 bps, ale został tak zmodyfikowany, by objął szybkości do 33600 bps. Jedną z nowszych charakterystyk zaimplementowanych w klasie V.34 jest zdolność modemu do ciągłego monitorowania kanału komunikacyjnego i zwiększania lub zmniejszania wynegocjowanej szybkości transmisji, zgodnie ze zmianami warunków na linii. V.34 jest także zgodny ze standardem kompresji V.42, który umożliwia uzyskanie większej przepustowości połączenia. Jest także wstecznie kompatybilny z wcześniejszymi standardami modemów, tak że może negocjować połączenia ze starszymi, wolniejszymi modemami. V.34, tak jak poprzednie standardy, może pracować z publiczną komutowaną siecią telefoniczną, jak też w konfiguracji z linią dzierżawioną.
• V.42: Standard V.42 określa procedury wykrywania i korygowania błędów dla urządzeń komunikacyjnych. Właściwie definiuje on dwa typy wykrywania i korygowania błędów: protokoły LAPB i MNP klasy 4. Protokół symetrycznego dostępu do łącza (LAPB) jest metodą używaną przez protokół warstwy sieci (warstwy 3 modelu OSI), znany jako HDLC (sterowanie wysokopoziomowym łączem danych). LAPB jest metodą wykrywania i korygowania błędów preferowaną przez standard V.42. MNP klasy 4 jest metodą drugorzędną, dostępną jako alternatywa.
• V.42 bis: Standard V.42 bis wzbogaca możliwości wykrywania i korygowania błędów oferowane przez standard V.42. Zapewnia kompresję danych przez urządzenie komunikacyjne wykorzystujące procedury korekcji błędów opisane przez standard V.42.
• 56K: Choć w czasie. gdy powstaje ten tekst, nie istnieje jasna definicja standardu 56 K, to dostępne są dwie rywalizujące technologie, umożliwiające użytkownikowi uzyskanie szybkości 53 000 bps dla transmisji przychodzących (ang. downstream) i szybkości 33 600 bps dla transmisji wychodzących (ang. upstream). Mimo że technologia ta teoretycznie umożliwia uzyskanie szybkości 56 000 bps transmisji wychodzących, to komisja FCC ogranicza szybkość tej transmisji do 53 000 bps. Zwiększona prędkość wynika z możliwości wyeliminowania części szumu związanego z konwersją analogowo-cyfrową (nazywanego „szumem kwantowania") dla połączeń wychodzących w sytuacjach, w których połączenie z modemu docelowego z powrotem do sieci będzie połączeniem całkowicie cyfrowym. Tak więc standard 56 K eliminuje konieczność przeprowadzania konwersji analogowo-cyfrowej dla połączeń wychodzących, usuwając w ten sposób szum kwantowania i zwiększając przepustowość.

1.17.6 Modemy a Microsoft Networking

Modemy mogą być stosowane w środowisku sieci Microsoft w celu rozszerzenia zasięgu sieci. Aby użyć modemu do ustanowienia zdalnego połączenia z siecią, użytkownik musi najpierw skonfigurować modem. Zarówno w Windows 95, jak i w Windows NT 4.0 modem konfiguruje się za pośrednictwem appletu' Modem, dostępnego w Panelu sterowania. Po zainstalowaniu i odpowiednim skonfigurowaniu modemu można ustawić właściwości zdalnego połączenia z siecią.
Użytkownik systemu Windows 95 w celu połączenia się ze zdalną siecią lub urządzeniem musi korzystać z apletu Dial-Up Networking. Każde połączenie utworzone przy użyciu Dial-Up Networking może być następnie modyfikowane przez zmienianie właściwości połączeń. Właściwości karty Typy serwerów przykładowego połączenia Dial-Up Networking przedstawione są na rysunku 17.7. Połączenie Dial-Up Networking charakteryzuje kilka właściwości, za pomocą których można konfigurować sposób wykonywania tego połączenia. Każde połączenie może mieć inne ustawienia dla każdej
Termin "applet" oznacza dosłownie "mała aplikacja"; pomimo zwyczajowego kojarzenia go ze środowiskiem Java oznaczać więc może również dowolny program zbyt skromny na to, by uznać go za typową aplikację (przyp. red.).
z dostępnych kart właściwości. Na rysunku 17.7 właściwości zostały ustawione tak, aby ustanowić połączenie z ISP (dostawcą usług internetowych). Ten konkretny dostawca używa połączenia protokołu PPP a jego usługa działa najlepiej, jeśli nie jest wybrana żadna z zaawansowanych opcji. Inni dostawcy usług internetowych w celu ustanowienia połączenia mogą wymagać innych ustawień różnych zaawansowanych opcji. W dolnej części karty Typy serwerów znajduje się panel Dozwolone protokoły sieciowe. Większość połączeń korzysta wyłącznie z protokołu TCP/IP. Połączenie może być oczywiście skonfigurowane tak, że nie będzie używać żadnego lub będzie używać kilku albo wszystkich dozwolonych protokołów. Warto zaznaczać tylko te protokoły, które są do ustanowienia połączenia potrzebne, upraszcza to bowiem proces ustanawiania połączeń, a także ułatwia rozwiązywanie problemów oraz powoduje zwiększenie wydajności.
Rysunek 17.7. Właściwości połączenia sieciowego Windows Dial-tlp Networking.
Również użytkownicy Windows NT 4.0 muszą skonfigurować applet Dial-Up Networking w celu łączenia się ze zdalnym komputerem. Jeśli system NT ma służyć jako serwer, do którego dołączeni są klienci, wtedy musi być zainstalowana i uruchomiona usługa dostępu zdalnego (ang. RAS - Remote Access Service). Administrowaniem usługą dostępu zdalnego kieruje narzędzie „Remote Access Admin", dostępne jako jedno z narzędzi administracyjnych usługi RAS. W systemie Windows NT (wersja dla stacji roboczej) połączenie uzyskać może jednocześnie tylko jeden zdalny użytkownik. System Windows NT Server obsługuje do 256 jednoczesnych zdalnych połączeń. Po ustanowieniu zdalnego połączenia, za pomocą linii telefonicznej uzyskujemy dostęp do wszystkich zasobów sieci, tak jakbyśmy byli do niej przyłączeni lokalnie.



Usługę dostępu zdalnego w systemie Windows NT 4.0 zainstalować można, klikając przycisk Dodaj znajdujący się na karcie Usługi okna Sieci. Aby dodać usługę dostępu zdalnego do listy zainstalowanych usług, trzeba uprzednio zainstalować i skonfigurować modem, a także wybrać klika opcji konfiguracyjnych, takich jak protokoły, których używać mogą rozmówcy łączący się z zewnątrz, a także użytkownicy wewnętrzni. Jeśli modem nie został jeszcze w systemie Windows NT zainstalowany podczas instalacji usługi RAS, program konfiguracyjny podpowiada, by to zrobić.

Do innych opcji dostępnych wraz z usługą RAS należy możliwość określenia, czy porty (takie jak COM1, COM2, COM3, COM4, LPT I itd.) mogą być wykorzystywane jako porty połączeń wyjściowych, wejściowych, czy też jako oba. Każdy protokół wybrany jako wejściowy ma opcję konfiguracyjną dopuszczającą inne zmiany w konfiguracji. Każdy protokół ma inną listę ustawień, a najszerszy zestaw opcji dodatkowych ma TCP/IP. Inną opcją usługi RAS jest możliwość żądania zaszyfrowanej weryfikacji tożsamości, wybranie której umożliwia korzystanie z szyfrowania danych.

1.17.7 Podsumowanie

Wraz z rozwojem technologii modemów rośnie również możliwość zwiększania zasięgu nowoczesnych sieci oraz szybkości, z jaką one działają. W niedalekiej przyszłości takie technologie jak modemy kablowe, asymetryczne cyfrowe łącza abonenckie oraz połączenia bezprzewodowe otworzą przed nami nowe możliwości rozszerzania sieci i łączenia ich ze sobą. Aby jednak te coraz to nowsze i lepsze technologie mogły odnieść sukces i zostać szeroko zaakceptowane, niezbędne są coraz to nowsze i lepsze standardy i protokoły regulujące proces komunikacji.

1.18 Rozdział 18 Usługi dostępu zdalnego (RAS)

Arthur Cooper

Czym są usługi dostępu zdalnego? Dziś użytkownicy oczekują, że będą w stanie korzystać z tych samych aplikacji i usług niezależnie od tego, gdzie się w danym czasie znajdują. To wymaganie odnośnie przyłączalności postawiło niezwykle ciężkie zadanie przed departamentami technologii informatycznych (ang. IT - Information Technology departments). By unaocznić. dlaczego istnieje aż takie zapotrzebowanie na niezawodne usługi dostępu zdalnego, w rozdziale tym opisano ciąg zdarzeń, które doprowadziły do powstania nowoczesnych usług RAS. Przedstawiono również rozwiązania zdalnego dostępu wykorzystywane we wszystkich erach rozwoju usług RAS. Warto zobaczyć, gdzie byliśmy, by ujrzeć, dokąd zmierzamy.

1.18.1 Historia korzystania z sieci o dostępie zdalnym

W późnych latach 50. dla większości korporacji, organizacji militarnych i innych agencji używających systemów z komputerami mainframe stało się jasne, że możliwości dostępu do danych w tych systemach powinny zostać rozszerzone poza tradycyjny zasięg przyłączonych terminali. Komputerowe centra danych w latach S0. były miejscami mistycznymi, pełnymi ludzi w białych laboratoryjnych fartuchach. Stan taki trwał aż do czasu, gdy pojawił się system mainframe IBM i narodziło się nowoczesne centrum danych.

Te centra danych uczyniły usługi komputerowe dostępnymi dla mas, ale użytkownicy tych wczesnych centrów byli zmuszeni zmagać się z wieloma niemądrymi regułami i wymaganiami nakładanymi przez personel. Wielokrotnie reguły te nakładały na użytkowników poważne ograniczenia. Niewygodne godziny, konieczność przestrzegania wielu ustalonych formatów danych i mnóstwo innych wymagań mogło sprawić, że użytkownicy czuli się niemalże sługami centrów danych i mieli poczucie bagatelizowania ich racji. Stało się więc jasne, że sposób korzystania z komputerów musi ulec zmianie.

W tym samym czasie następowała rewolucja w świecie komputerów. Rząd USA zobligował Agencję ds. Perspektywicznych Badań Obronnych (ang. DARPA - Defence Advanced Research Projects Agency), by rozpoczęła testowanie nowych sposobów łączenia ze sobą niepodobnych systemów komputerowych. Podczas gdy kadra menedżerska zaczęła przenosić terminale z centrów danych do swoich biur, DARPA eksperymentowała z dalekosiężnymi połączeniami między centrami danych. W większych korporacjach, w których komputery

mainframe były odpowiedzialne za śledzenie (monitorowanie) budżetów i zapasów, kierownictwo zaczęło zdawać sobie sprawę z wagi decentralizacji tych informacji.
Gdy członkowie zarządu i kierownicy średniego szczebla byli zmuszeni korzystać z danych w postaci wielkich, niewygodnych, tygodniowych wydruków, dane których używali były tylko tak dobre, jak ostatni wydruk. W wielu przypadkach siedmiodniowe opóźnienie informacji było nie do przyjęcia. Prowadziło to do większej liczby wydruków i wielu kierowników czuło, że korzystanie z tych raportów na co dzień, czy nawet co tydzień, jest stratą czasu. Coś wreszcie musiało się zmienić - dostęp do danych w systemach mainframe musiał osiągnąć poziom użytkownika i musiało to nastąpić szybko.
Nagle stało się coś wielkiego. Dostępne stały się sterowniki łącza i regulatory łącza. Urządzenia te pozwoliły centrum danych zdalnie łączyć terminale z komputerem mainframe. Pierwszymi użytkownikami tych terminali byli sami menedżerowie centrum danych. Następnie terminale zaczęły się pojawiać w całym miejscu pracy. Wkrótce informacje dostępne tylko w jednym miejscu w organizacji były doręczane tylko ludziom naprawdę ich potrzebującym. Wyobraźcie sobie to!

1.18.1.1 Lata siedemdziesiąte

We wczesnym okresie przetwarzania danych bardzo rzadko można było znaleźć organizację rozproszoną na dużym obszarze geograficznym. Z tego powodu nie było specjalnego zainteresowania łączeniem w sieć dużych komputerów mainframe. Korporacje, uniwersytety i agencje rządowe były jedynymi poważnymi użytkownikami komputerów mainframe, zaś spośród tych trzech jednostek tylko rząd był zainteresowany łączeniem komputerów rozproszonych geograficznie. Kiedy jednak agencja DARPA zaczęła doskonalić ideę komputerów komunikujących się z innymi, niepodobnymi do nich komputerami, sektor prywatny zaczął zwracać na to uwagę.
Modem stał się wtedy wyjątkowo ważnym urządzeniem, a pierwsze modemy nie były właściwie niczym więcej jak urządzeniami sprzęgającymi akustycznie, które umożliwiały sprzęgnięcie telefonu z portem systemu komputerowego. Użytkownik przyłączony do centrum danych mógł teraz, korzystając z połączenia sieciowego istniejącego między centrami, uzyskać dostęp do informacji w innych centrach danych. Narodziła się nowoczesna sieć transmisji danych.
Zawodowi informatycy zaczęli podłączać do systemów mainframe coraz bardziej inteligentne terminale. W tym samym czasie zaczęto wyposażać terminale we wbudowaną pamięć. Mogły one dzięki temu przechowywać informacje konfiguracyjne. W ten sposób, gdy terminal był wykorzystywany do ustanowienia połączenia lub sesji z komputerem, wspólne polecenia konfiguracyjne mogły być automatycznie przesyłane między systemem komputerowym i terminalem.
To udogodnienie tylko zaostrzyło apetyty wielu użytkowników mainframe. Na scenie pojawiły się systemy minikomputerowe.
Użytkownicy usług transmisji danych zaczęli eksperymentować z łączeniem - przy użyciu modemu - minikomputerów z dużymi komputerami *mainframe.* Pomysł przyjął się dobrze. Pozwoliło to na przesyłanie danych tam i z powrotem między różnego typu komputerami.

1.18.1.2 Lata osiemdziesiąte

W latach 80. komputery mainframe zaczęły tracić na znaczeniu. W raku 1981 wprowadzono na rynek komputer osobisty IBM PC. Dostępnych stało się kilka systemów operacyjnych dla sieci lokalnych (LAN), zaś jeśli chodzi o sprzedaż, oczywistym liderem stała się firma Novell. Gdy sieci lokalne małych komputerów stały się standardem w miejscach pracy, wzrosła potrzeba zdalnego dostępu do tych sieci z dowolnego miejsca, za pomocą połączeń dial-up. Modemy stały się (stosunkowo) tanie i dostępne, zaś użytkownicy zaczęli wymagać dobrej jakości połączeń dial-up dla swoich pecetów.
Firmy Novell, IBM i Microsoft zaczęły przedstawiać rozmaite pakiety oprogramowania osadzone w ich systemach operacyjnych. Przeznaczeniem tych pakietów było dostarczanie usługi dial-up. Za pomocą programowych emulatorów, takich jak ProComm, użytkownicy komputerów osobistych mogli łączyć się z dużymi systemami komputerowymi lub sieciami, wykorzystującymi pojedyncze modemy lub ich grupy.
Sieć komputerowa Agencji ds. Badań Perspektywicznych (ang. ARPANET - Advenced Research Projects Agency Network) przeobraziła się w roku 1969 w sieć Internet. Wkrótce zaczęły ją eksploatować uniwersytety i korporacje prywatne i w roku 1984 Internet łączył ze sobą 1000 komputerów mainframe. W latach 1987 - 1994 liczba podłączonych do Internetu hostów wzrosła z 10000 do dwóch milionów. Użytkownicy zaczęli domagać się dostępu do Internetu ze swoich domów, a firmy Novell i Microsoft były szczęśliwe, mogąc im go dostarczyć. Zdalny dostęp w latach 90. mógł już wiernie powielać i emulować usługi sieci LAN, na których użytkownicy zaczęli polegać.

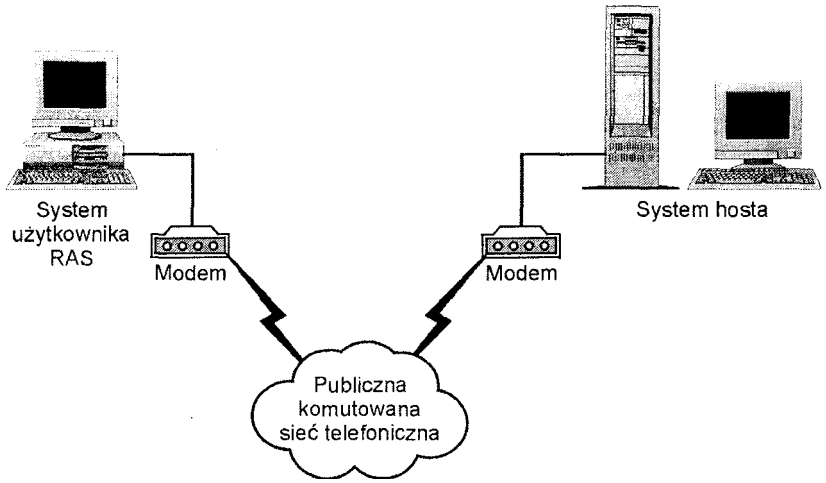
1.18.1.3 Szaleństwo lat dziewięćdziesiątych

Dziś akronim RAS jest używany niemal tak często jak akronim PC. Użytkownicy wymagają, by sieci z usługami RAS były w stanie zapewnić podczas łączenia się z serwerami i systemami sieciowymi takie same możliwości, jakie są dostępne w sieciach LAN. Microsoft wbudował usługę RAS w swój system operacyjny Windows NT, na skutek czego w ciągu ostatnich kilku lat całkiem spora część PKB' jest nie tylko konsumowana, ale również wytwarzana w gospodarstwach domowych pracowników firm bardzo wielu różnych działów gospodarki. Opracowanych zostało dotychczas wiele pakietów producentów niezależnych. Usługa RAS stała się jedną z najważniejszych technologii świata komputerów.
PKB - produkt krajowy brutto (przyp. red)
Z tych właśnie powodów każdy pragnący odnieść sukces informatyk powinien w pełni zrozumieć zasady, na jakich działają usługi RAS. Wiele przedsiębiorstw to przedsiębiorstwa geograficznie rozproszone, a zdalny dostęp jest najlepszym sposobem, w jaki pracownicy mogą utrzymywać stały związek z korporacyjnymi zasobami obliczeniowymi. Dziś zdalny dostęp oferuje możliwość korzystania z poczty elektronicznej, realizowanie wspólnych projektów i wiele innych aplikacji. Usługi zdalnego dostępu zyskały należne im miejsce w środowisku technologii informacyjnych. W następnych punktach opisane jest działanie i właściwości usług RAS.

1.18.2 Ustanawianie połączeń zdalnych

Dla zapewnienia połączenia dial-up konieczne jest zastosowanie protokołów standaryzujących sposób, w jaki każdy koniec połączenia uzgadnia przepływ danych pomiędzy hostem a użytkownikiem. Na rysunku 18.1 pokazany jest przykład takiego połączenia.

Rysunek 18.1. Połączenie Użytkownik-host



Na użytek tego rozdziału termin host odnosi się do serwera, systemu komputerowego lub sieci wykorzystywanej przez użytkownika RAS. Termin użytkownik odnosi się do małego komputera czy też peceta, który inicjuje zdalne połączenie z hostem.

W początkowym okresie połączeń dial-up komputery PC emulowały bezpośrednio przyłączone terminale. Dzięki tego typu połączeniom dial-up, komputer PC mógł wyglądać i działać tak samo, jak terminal bezpośrednio przyłączony do hosta. W efekcie jedynym oprogramowaniem, jakiego potrzebował komputer użytkownika, było oprogramowanie wybierające numer (telefoniczny) oraz program symulujący sygnały znakowe wysyłane przez terminal.

ASCII to skrót od nazwy American Standard Code for Information Interchange. Jest to standardowy 7-bitowy kod znakowy służący do wymiany informacji. Skrót EBCDIC (ang. Extended Binary Coded Decimal Interchange Code - rozszerzony dziesiętny kod wymiany o dwójkowym zapisie) odnosi się natomiast do 8-bitowego kodu znakowego opracowanego przez IBM.

Pierwsze systemy dial-up były efektywne, ale powolne. Generalnie, szybkości połączeń modemowych były mniejsze niż 1200 bps. Niektóre z wcześniejszych połączeń były po prostu akustycznym sprzęganiem urządzeń, gdzie słuchawka telefoniczna była połączona z akustycznym sprzęgiem przypominającym uszy myszki Miki.

1.18.2.1 Ewolucja standardów protokołów

Jedynymi protokołami wchodzącymi w grę w epoce wczesnych, zdalnych połączeń dwupunktowych (użytkownik-host) były protokoły określające elektromechaniczne standardy pomiędzy modemami i liniami telefonicznymi. W Stanach Zjednoczonych firma Bell Telephone Company ustanowiła standardy definiujące połączenia modemu z modemem. Pierwszymi standardami były Bell 103 i Bell 212A. Później zostały one przyjęte na całym świecie i CCITT opracowało standardy V.21 i V.22. Standardy te określają transmisję między dwoma połączonymi modemami.

CCITT jest akronimem francuskiej nazwy Comitee Consulatif Internationale de Telegraphique et Telephonique (Międzynarodowy Komitet Konsultacyjny ds. Telefonii i Telegrafii). Standardy „V" oraz „X" dotyczą łączy transmisji danych. Zwykle standardy „V" dotyczą łączy danych i modemów wykorzystujących obwody telefoniczne lub głosowe. Standardy „X" natomiast zwykle dotyczą danych cyfrowych. Od czasu do czasu standardy te są nowelizowane lub usuwane. Komitet CCITT definiuje standardy zalecane dla całego świata. Ostatnio CCITT uległo reorganizacji. Zmieniono również jego nazwę z francuskiej na angielską, więc komitet przestał być komitetem, a stał się unią, a dokładnie Międzynarodową Unią Telekomunikacyjną (ang International Telecommunication Union).

1.18.2.2 **Zestaw poleceń** AT

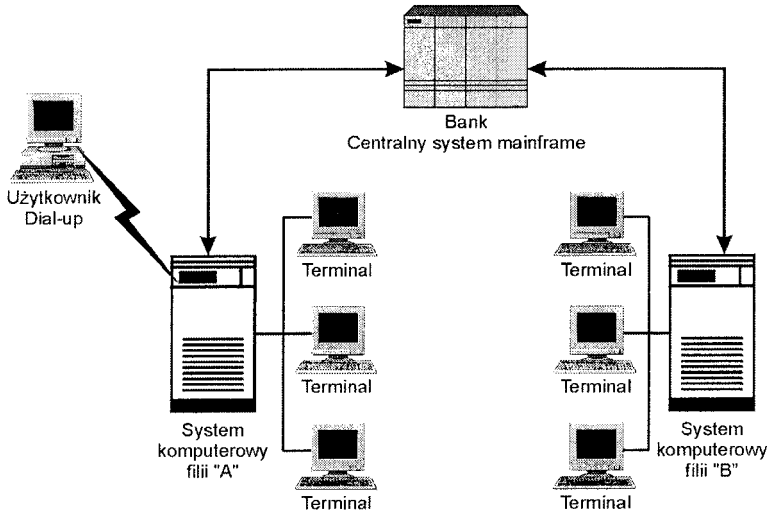
Firma Hayes Inc. wprowadziła na rynek Hayes Smartmodem i znormalizowała zestaw poleceń AT (ang. attention - uwaga) do wykorzystania w komputerach PC i modemach dial-up. Zestaw poleceń AT jest używany po dzień dzisiejszy. Obecnie jest on wbudowany w większość programów emulacji terminali dla komputerów PC. To właśnie zestaw poleceń AT umożliwia komputerowi wydawanie modemowi poleceń podniesienia słuchawki, wybierania numeru itp. Pamiętajmy jednak, że początkowo użytkownicy musieli ręcznie wprowadzać polecenia AT, by zmusić modem do zrobienia czegokolwiek.

Takie połączenie ograniczało użytkowników, gdyż mogli wykonywać tylko te funkcje, które mógł wykonywać system host. Jeśli terminale połączeń bezpośrednich w centrach danych nie miały dostępu do zasobów sieci, zwykle nie mieli go również użytkownicy dial-up.

Może najlepszym przykładem usługi zdalnego dostępu do terminala jest metoda pracy sieciowej stosowana wówczas w bankowych systemach komputerowych. Łączono w sieć mniejsze host-systemy w filii banku z większymi hostami mainframe w banku macierzystym. Mniejsze systemy w filiach były tak skonfigurowane, by aktualizowały dane w większych, centralnych hostach.

Wielokrotnie banki konfigurowały systemy filialne tak, aby przekazywały one dane „do góry", do centralnego hosta, w czasie rzeczywistym. Dokonywano tego za pomocą modemów i dedykowanych linii telefonicznych. Jednakże przepływ danych „w dół", z hosta centralnego do systemów oddziałów czy filii, zachodził tylko raz dziennie. Rysunek 18.2 przedstawia przykład takiego rozwiązania. Filia „A" i filia „B" są

połączone z systemem centralnym. Nie są jednak połączone bezpośrednio ze sobą. Dane stale przepływają z systemów filialnych do systemu centralnego, ale w drugą stronę płyną tylko raz dziennie. Dlatego dane w systemach filialnych pochodzą zawsze z poprzedniego dnia roboczego.
Rysunek 18.2. Po(qc<enia systemu centralnego zfiliami
Zasadniczo cały dostęp dial-up odbywał się za pośrednictwem filii. W tamtych czasach dostęp dial-up próbował naśladować działanie terminali bezpośrednio podłączonych przewodami do systemu mainframe. W następnym punkcie opisane są procesy wyższego poziomu, zachodzące wtedy, gdy użytkownicy inicjują sesję zdalnego dostępu do systemu hosta.



1.18.2.3 Protokoły połączeń zdalnych

Aby zbadać procesy wyższego poziomu zachodzące przy zdalnym połączeniu, trzeba poznać dwa najbardziej rozpowszechnione protokoły połączeń używane w dzisiejszych połączeniach zdalnego dostępu. Wstępem do dyskusji o tych protokołach może być krótkie podsumowanie procesu zachodzącego, gdy ustanowiona jest sesja zdalnego dostępu użytkownik-host.

1.18.2.4 Ustanawianie sesji

Aby zainicjować zdalne połączenie, jedno z urządzeń końcowych (zwykle użytkownik) wybiera numer (czyli wywołuje) drugiego (zwykle hosta). Dziś wykonuje to komputer użytkownika, uruchamiając określony program wywołujący. System Windows 95 na przykład ma wbudowaną funkcję sieciową dial-up. Po tym, jak modemy uzgodnią połączenie poprzez linię telefoniczną i ustali się między nimi sygnał nośny, sygnały cyfrowe na porcie wyjściowym modemu przekazują tę informację z powrotem do komputera użytkownika.
Sygnał nośny to tzw. „pilot" albo główny sygnał analogowy, wysyłany przez modem do innego modemu poprzez łącze komunikacyjne, gdy między dwoma urządzeniami zostanie ustanowione połączenie dla transmisji danych.
Następnie użytkownik wywołuje program używany do komunikowania się z hostem. Może to być pakiet emulacji terminala, taki jak ProComm, lub też może to być wbudowany program wywołujący. 1 znów, system Windows 95 ma wbudowane oprogramowanie uzgadniające w ramach swojej funkcji sieciowej dial-up. Wszystko, czego użytkownik może potrzebować, aby uzgodnić procesy wyższego poziomu w systemie hosta, jest wykonywane właśnie wtedy. Czasem użytkownik może wywołać unikatowy program napisany specjalnie dla określonego przedsiębiorstwa.

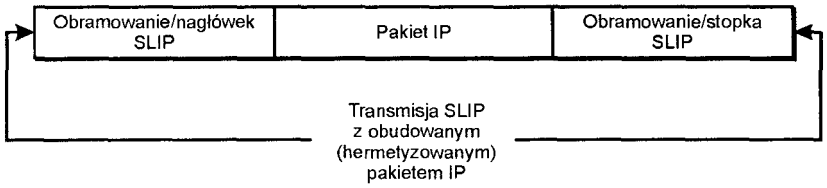
1.18.2.5 Protokoły dostępu sieci TCP/IP

Większość ludzi korzystających z połączeń zdalnego dostępu łączy się z Internetem lub z Intranetem swojego przedsiębiorstwa. Z tego powodu pakietem protokołów używanym do przekazywania danych tam i z powrotem jest TCP/IP. Dziś, kiedy TCP/IP jest wykorzystywany jako główny pakiet protokołów transportowych, jeden z protokołów, SLIP albo PPP, będzie wykorzystywany jako główny protokół dostępu dla sesji dial-up pomiędzy komputerem PC a serwerem RAS.

1.18.2.5.1 SLIP

We wczesnych latach 80. Internet stawał się wyjątkowo ważnym środkiem masowego przekazu. Być może najważniejszy protokół opracowany w tamtym okresie stworzyła korporacja 3COM. Firma ta opracowała technologię, która umożliwiła przesyłanie pakietów TCP/IP poprzez linie telefoniczne. TCP/IP wciąż jest pakietem protokołów, dzięki któremu dane przemieszczają się z jednego miejsca do drugiego.
Głównym mechanizmem przenoszącym dane w świecie TCP/IP jest pakiet IP. Nie wdając się w długie dyskusje na temat protokołu IP i adresowania IP'2, można powiedzieć. że jest to bezpołączeniowy protokół Internetu przemieszczający pakiety TCP i UDP (protokołu datagramów użytkownika) z jednego miejsca do drugiego. Pakiety TCP i UDP niosą dane wysłane przez użytkownika. Innymi słowy, IP można uważać za nośnik, za pakiety TCP i UDP za przedmioty przenoszone.

Każde połączenie dial-up, które mogłoby przenosić pakiety IP, musiało być skonfigurowane tak, by przesyłało wszystkie znaki pakietu IP. Korporacja 3COM wprowadziła nowy protokół, nazwany protokołem komunikacji szeregowej Internetu (ang. SLIP Serial Line Internet Protocol). W rzeczywistości nie był to wcale protokół; jednak działał jak protokół. SLIP nigdy nie stał się standardowym protokołem Internetu. Protokół SLIP jedynie umieszcza pakiet IP w ramkach i przesyła go z jednego punktu sieci do drugiego. Przedstawia to rysunek 18.3. Protokół ten nie zapewnia adresowania identyfikacji typu pakietu, kontroli błędów ani kompresji. Jednak braki te są jego zaletą. gdyż sprawiają, że protokół SLIP jest bardzo łatwy do wdrożenia i obsługi.
Rysunek 18.3. Ramka SLIP pakietów IP.
W 1984 roku Rick Adams zaimplementował protokół SLIP w systemie Unix Berkeley i „upublicznił" go. Zdalny dostęp do usług Internetu stał się powszechny.



Do dnia dzisiejszego wiele programów połączeniowych dla komputerów PC, przeznaczonych do łączności z Internetem, ma opcję SLIP. Brak specyfikacji standardu SLIP oznacza, że nie został dla niego określony maksymalny rozmiaru pakietu. W końcu lat 80. powstało wiele odmian protokołu SLIP, ale wszystkie zapewniały tylko ramkę do przenoszenia pakietów IP liniami szeregowymi. Pamiętajmy, że komputer użytkownika nadal musi komunikować się z modemem za pomocą poleceń AT. Nadal wywołuje i ustanawia połączenie z modemem hosta. Tyle tylko, że dziś utrzymywanie łączności między modemami odbywa się według standardów ITU-T.
Protokół SLIP wchodzi do gry dopiero wtedy, gdy powstanie już stabilne połączenie między modemami, a użytkownik i host ustanowią połączenie. Gdy to nastąpi, protokół SLIP umożliwi przekazanie pakietów IP poprzez to szeregowe połączenie użytkownik-host Zapamiętaj': SLIP nie przekazuje żadnych informacji adresowych. Oznacza to, że każdy komputer (zarówno host, jak i komputer użytkownika) musi znać adres drugiego, by efektywnie przesyłać między sobą pakiety IP. Zawsze jest to funkcją dowolnego typu oprogramowania wywołującego SLIP w komputerze użytkownika i oprogramowania SLIP w systemie hosta.
' Technologii TCP/IP poświęcona jest w całości książka T.Yarkera „Protokół TCP/IP" wyd.. polskie HELION 1998 (przyp. red.)
Protokół SLIP ma wiele ograniczeń, a użytkownicy wymagają wolnego od błędów połączenia - zwłaszcza podczas wywoływania systemu hosta. Fakt, że protokół SLIP nie mógł przeprowadzać kompresji, adresowania czy kontrolowania błędów, sprawił, że użytkownicy zaczęli domagać się lepszego protokołu połączeniowego. Pragnienia użytkowników spełniły się wraz z narodzinami protokołu PPP.

1.18.2.5.2 Protokół PPP

Protokół PPP (ang. *point-to point protocol,* czyli protokół z punktu do punktu lub inaczej - protokół dwupunktowy, był kolejnym projektem powstałym z inicjatywy użytkowników. Podstawowym przeznaczeniem tego protokołu jest zapewnienie połączenia między dwoma równoprawnymi urządzeniami przy wykorzystaniu portów szeregowych. Typowe zastosowanie protokołu PPP ogranicza się do połączenia dial-up użytkownik-host. Jednak niektórzy wykorzystują go w połączeniach host-host oraz host-router. W tym rozdziale opisany jest protokół PPP stosowany w połączeniach dial-up dostępu zdalnego.
Protokół PPP został znormalizowany jako prawdziwy protokół Internetu. SLIP nigdy nie osiągnął takiego statusu - co nie zmienia faktu, iż wielu użytkowników wciąż święcie wierzy w protokół SLIP. Niestety, czasem ilość błędów występujących podczas korzystania z niego może być astronomiczna. Jeśli więc połączenie dostępu zdalnego ma poprawnie przekazywać pakiety IP, do jego obsługi lepiej wybrać protokół PPP.
Protokół PPP także umieszcza pakiety w swojej ramce, tak jak robi to protokół SLIP, ale oprócz prostego ramkowania danych dzieją się jeszcze inne rzeczy. Protokół PPP ustanawia również sesję sterowania łączem pomiędzy użytkownikiem i hostem, a sesja ta jest kontrolowana przez protokół sterowania łączem, czyli protokół LCP (ang. *Link Control Protocol).* Protokół ten konfiguruje łącze między użytkownikiem, a hostem dla sesji PPP.
Jest to pierwsza czynność sesji PPP. Pakiety protokołu sterowania łączem są przesyłane między użytkownikiem a hostem. Testują one i konfigurują ustanowione łącze. Aby upewnić się, że łącze jest stabilne, testowane są takie jego właściwości, jak jakość łącza, echo łącza itp. Istnieją trzy klasy pakietów protokołu LCP:
• pakiety ustanawiania łącza (ang. Link Establishment packets), wykorzystywane do tworzenia i konfigurowania połączenia,
• pakiety zakończenia łącza (ang. Link Termination Packets), wykorzystywane do przerywania połączenia,
• pakiety utrzymania łącza (ang. Link Maintenance Packets), wykorzystywane do zarządzania łączem oraz wykrywania i usuwania usterek łącza.
Jeśli protokół LCP upewni się, że łącze działa i że działa stabilnie, informację tę przekazuje głównemu protokołowi PPP w celu zadeklarowania gotowości łącza. Gdy to nastąpi. protokół PPP wysyła pakiet zwany protokołem sterowania siecią (ang. NCP - Network Control Protocol). Pakiet sterowania siecią jest specyficzny dla typu danych, które mają być przekazane łączem PPP.
Pakiet sterowania siecią jest wysyłany przez inicjatora połączenia PPP (przeważnie użytkownika, który wywołał hosta). Pakiet sterowania siecią informuje hosta o rodzaju ruchu, który ma być przekazywany łączem PPP. Protokół sterowania siecią PPP dla datagramów IP nazywany jest protokołem sterowania IP, czyli protokołem IPCP (ang. IP Control Protocol). Protokół ten odpowiedzialny jest za konfigurowanie, a także za rozpoczynanie i kończenie działań protokołu IP na obydwu końcach łącza. Pakiety IPCP nie mogą być jednak wysłane, zanim protokół sterowania łączem nie zakończy fazy negocjacji. Gdy faza ta zostanie zakończona, protokół IPCP sygnalizuje, że jest gotów rozpocząć przesyłanie pakietów i datagramów IP.

Dopiero wtedy użytkownik i host mogą rozpocząć wysyłanie i odbieranie datagramów i pakietów IP. Jak wiemy, mechanizm przesyłania w postaci protokołu IP jest ostoją dzisiejszego Internetu. Dlatego nie ma sensu omawiać innych protokołów sterowania siecią, które protokół PPP mogą wykorzystywać do przesyłania danych. Warto jednak pamiętać, że wiele jest wersji protokołu PPP, a w niektóre z nich wbudowane są inne protokoły sterowania siecią.

1.18.2.5.3 Trendy bieżące

Protokoły SLIP i PPP bez wątpienia są obecnie „wołami roboczymi" sesji zdalnego dostępu. Rzadko można spotkać użytkowników usług RAS wykorzystujących inne protokoły do przesyłania danych między sobą a hostem. Jest ich mniej więcej 1 na 100. Równie jednak ważne jak zrozumienie zasad działania protokołu PPP jest poznanie usług dostarczanych przez serwery RAS po ustanowieniu połączenia PPP.
Znaczenie terminu usługi używane w niniejszym rozdziale różni się od znaczenia terminu „usługi" odnoszącego się do konfigurowania „usługi" w systemach Windows 95 lub Windows NT Server. Są to zupełnie różne terminy.

1.18.3 Usługi transportu zdalnego

W świecie usług dostępu zdalnego użytkownikom zależy na możliwości połączenia się z systemem hosta, wykorzystania tego systemu do swoich potrzeb i likwidacji połączenia RAS. Ze względu na rozmiar tego rozdziału niemożliwe jest omówienie wszystkich usług, z jakich użytkownik może korzystać podczas zdalnego połączenia. Zamiast tego przedstawiony jest przegląd najczęściej wykorzystywanych obecnie usług.

1.18.3.1 W jaki sposób obecnie łączą się użytkownicy usług dostępu zdalnego

Zwykle większość przedsiębiorstw należy do jednej z trzech kategorii:
• Przedsiębiorstwo duże - ma przynajmniej dwadzieścia pięć odrębnych ośrodków. Każdy z nich zwykle posiada sieć LAN, zaspokajającą potrzeby pracy sieciowej pracowników tej lokacji.
• Przedsiębiorstwo średnie - może mieć więcej niż dwie, trzy lokalizacje geograficzne, ale nie za wiele (mniej niż dwadzieścia pięć). W każdej z siedzib jest albo sieć LAN, albo „centralny" system komputerowy w lokacji (siedzibie) głównej. Pozostałe stanowiska mogą łączyć się z głównym komputerem za pomocą połączeń dial-up.
• Przedsiębiorstwo małe - firma mająca tylko jedną siedzibę, w której może być, choć nie musi, sieć LAN.
Duża organizacja ma zazwyczaj zainstalowaną wewnętrzną sieć Intranet. Użytkownicy korporacyjnego Intranetu zwykle korzystają z komputerów osobistych w swoich miejscach pracy. Część z nich jednak pracuje poza zakładem, w rozjazdach czy w domu, a dzięki dostępowi zdalnemu do sieci (i fizycznemu do komputera) może stale pozostawać w kontakcie z firmą, nawet podczas przenoszenia się z jednego miejsca winne.
Aby ich praca podczas korzystania z połączeń zdalnego dostępu mogła być wydajna, połączenia te powinny zapewniać im usługi identyczne jak użytkownikom komputerów przyłączonych bezpośrednio do sieci LAN. Czy pracownik siedzi przy biurku w swoim biurze, korzystając z bezpośredniego połączenia LAN, czy też łączy się z hotelu, zawsze musi widzieć to samo na ekranie swojego komputera osobistego. Dopuszczając pewne wyjątki, taki właśnie powinien być cel wszystkich rozwiązań dostępu zdalnego implementowanych w korporacyjnym intranecie.
Standard ten nie zawsze dotyczy pracowników przedsiębiorstw średnich. Niektórzy regularnie łączą się z komputerem znajdującym się w siedzibie głównej. W takim wypadku łączenie powinno odbywać się w ten sam sposób, niezależnie od tego, czy odbywa się z domu, czy z biura. Wiele przedsiębiorstw średnich obsługiwanych jest przez jednego centralnego hosta. Natomiast użytkownicy w małych firmach często łączą się z dostawcą usług internetowych (ISP) i do komunikowania się z innymi, wykorzystują pocztę elektroniczną.
Z drugiej strony, małe korporacje mogą mieć lepsze komputery, sieci LAN i przyłączalność niż duże korporacje. Dzieje się tak dlatego, że koszty ulepszania na małą skalę są niższe. W dodatku niektóre duże korporacje, ze względu na bezpieczeństwo i poufność, nie pozwalają na dostęp dial-up do swoich systemów. Niezależnie od szczegółów, można dokonać kilku obserwacji ogólnych dotyczących większości połączeń zdalnych.
• Przeważnie połączenia zdatne są ustanawiane w celu uzyskania dostępu do sieci LAN, sieci Intranet przedsiębiorstwa lub Internetu.
• Po ustanowieniu połączenia przekazywaniem danych przez to połączenie kieruje głównie protokół PPP.
• Najczęstszą usługą przesyłania danych, wykorzystywaną w tego typu połączeniach jest TCP/IP.
Od każdej reguły istnieją wyjątki i powyższe obserwacje również mogą być kwestionowane - w większości przypadków ustalenia powyższe są jednak bardzo bliskie prawdy. Następny punkt omawia usługi, jakie protokół TCP/IP może wykonać dla swoich zdalnych użytkowników.

1.18.3.2 Protokół TCP/IP - „wół roboczy" połączeń zdalnych

Żadna dyskusja na temat usług zdalnego dostępu nie byłaby kompletna bez omówienia protokołu TCP/IP. TCP/IP jest w świecie komputerów i telekomunikacji jednym z najbardziej niewłaściwie używanych terminów. TCP/IP ma dwa znaczenia. Po pierwsze, jest to połączenie protokołów. Powstało ono z połączenia protokołu sterowania transportem (TCP) z protokołem Internetu (IP). Po drugie, TCP/IP to podstawa dla pakietu innych protokołów stosowanych w Internecie i w korporacyjnych Intranetach.
Omawianie tych innych protokołów nie mieści się w zakresie tematycznym tego rozdziału. Jednak te najbardziej rozpowszechnione, takie jak:
• Telnet,
• protokół przesyłania plików, czyli protokół FTP (ang. File Transfer Protocol, • prosty protokół przesyłania poczty, czyli protokół SMTP (ang. Simple Mail Transfer Protocol,

są częścią pakietu TCP/IP. Internet był poligonem badawczym dla tych protokołów i na długo przed utworzeniem modelu referencyjnego połączonych systemów otwartych (modelu OSI) protokół TCP/IP był systemem całkowicie otwartym. Jego „otwartość" oznacza, że dowolny system zgodny z TCP/IP może komunikować się z dowolnym innym systemem uważanym za zgodny z TCP/IP.

Kiedy protokół TCP/IP zostanie w pełni zaimplementowany w połączeniach zdalnych użytkownik-host, wszystkie cudowne „zabawki", do których użytkownicy przywykli w swoich biurowych sieciach LAN, udostępniane są również użytkownikom zdalnych komputerów PC. Przeglądarki, takie jak Netscape, programy poczty elektronicznej, jak Eudora, programy FTP, jak WS-FTP i inne, mogą być wtedy implementowane na komputerze każdego zdalnego użytkownika.

Taki jest obecnie stan, jeśli chodzi o wykorzystywanie usług dostępu zdalnego, i wygląda na to, że przez bardzo długi czas sytuacja ta nie ulegnie zmianie. TCP/IP ma ograniczenia, ale wciąż jest protokołem numer jeden, wykorzystywanym na całym świecie do przesyłania danych. W przyszłości zmieni się zapewne kształt pakietu IP - w zasadzie cały przemysł sieciowy czeka na wdrożenie adresowania IP następnej generacji (IPv6) we wszystkich hostach internetowych. Protokół IPv6 używa 128-bitowego schematu adresowania, a nie - jak dotychczas - 32-bitowego.

Protokół TCP/IP jest obecnie „wołem roboczym" numer 1 implementowanym we wszystkich połączeniach zdalnych. Po ustanowieniu takiego połączenia TCP/IP zostaje uaktywniony przez równorzędne procesy w systemach komputerowych użytkownika i hosta. Dlatego system użytkownika ma przypisany adres IP. Adres ten służy do prawidłowego adresowania i kontroli przepływu danych od użytkownika do hosta i od hosta do użytkownika.