chabińska

Cztery liczby, które zawładnęły światem

Jedną z niewidocznych, ale niezwykle ważnych, usług internetowych jest system DNS (Domain Name System). Na całym świecie mechanizm ten wykorzystuje się wiele milionów razy na sekundę. Udostępnia on adresy komputerów podłączonych do Internetu, dzięki czemu – za pośrednictwem protokołu TCP/IP (Transmision Control Protocol/Internet Protocol) – po Sieci mogą wędrować miliardy pakietów danych.

DNS pełni w Internecie funkcję systemu informacji o adresach oraz tłumaczy adresy symboliczne (np. harry@hotmail.com) na adresy IP jednoznacznie określające adresata. Na najniższym poziomie, w którym przemieszczają się pakiety danych przesyłane przez przeglądarki WWW i pozostałe programy, identyfikatory URL (takie jak www.chip.pl) są zupełnie nieznane. W tym miejscu niepodzielnie królują czteroczłonowe, liczbowe adresy IP, identyfikujące każdy komputer podłączony do Internetu, przyporządkowując mu określony fragment Sieci i odpowiedni numer w ramach tego obszaru.

Zanim przeglądarka WWW uzyska dostęp do wybranej witryny internetowej, musi najpierw poznać przyporządkowany jej adres IP. Taka sama sytuacja występuje w przypadku aplikacji pocztowych, programów do „pogawędek" (chat) czy narzędzi FTP. Dopiero po uzyskaniu adresu IP docelowego komputera możliwe jest nawiązanie z nim łączności i rozpoczęcie komunikacji z użyciem odpowiedniego protokołu. Z tego właśnie względu w sieci Internet tak ważną rolę odgrywa usługa DNS. Gdyby jej nie było, musielibyśmy za każdym razem podawać adresy stron WWW w postaci *http://195.116.104.13/index.htm*.

Początkowo Internet funkcjonował bez usługi DNS. Do przełomu roku 1986/87 podłączone do sieci komputery zapisywały adresy innych maszyn w zwykłym pliku ASCII umieszczonym na lokalnym dysku twardym. Aby odnaleźć adres IP innego komputera, aplikacje internetowe musiały tylko przeszukać taką lokalną listę. To proste i szybkie rozwiązanie okazało się jednak niewystarczające w obliczu gwałtownego rozwoju Sieci. Listy adresowe stawały się bowiem coraz dłuższe i musiały być coraz szybciej dystrybuowane przez serwer prowadzący centralny rejestr. Taka sytuacja systematycznie zwiększała ruch w Internecie, chociaż wielu odbiorców tych informacji nigdy nie korzystało z większości umieszczonych na liście adresów. Z tego właśnie względu opracowany został system DNS, zapewniający wyłącznie doraźną dystrybucję informacji adresowych. Jego zadaniem jest przekształcanie adresów symbolicznych na adresy IP i vice versa na żądanie poszczególnych aplikacji internetowych w momencie, gdy potrzebują one adresu IP, a dysponują jedynie jego zapisem symbolicznym.

Zadania serwera DNS

Zasadniczo mechanizm DNS nie różni się niczym od innych dostępnych usług internetowych. Również w tym systemie istnieją serwery (DNS Servers), które oferują określoną usługę (konwersję adresu) i są wywoływane przez stacje robocze (Clients). W tym przypadku funkcjonujące na poszczególnych komputerach aplikacje internetowe łączą się z serwerami DNS w celu przekształcenia posiadanych nazw poszczególnych komputerów i domen na adresy IP. Aby sprostać temu zadaniu, serwery DNS prowadzą bazy danych, zawierające nazwy i adresy znanych domen internetowych oraz należących do nich hostów, i za pośrednictwem odpowiedniego protokołu udostępniają te informacje. Komputer pyta więc np. „Czy znasz adres www.chip.pl?", na co serwer odpowiada: „Tak, to numer 195.116.104.13" lub udziela odpowiedzi negatywnej, gdyż

nazwa została podana nieprawidłowo bądź dany host i/lub jego domena nie istnieją.

Funkcji serwera DNS nie pełni jeden centralny komputer. Zadanie to zostało rozdzielone pomiędzy wiele maszyn. Żaden z tych serwerów nie musi dysponować informacjami adresowymi dotyczącymi całego Internetu. Poszczególne serwery odpowiadają więc zawsze tylko za pewną część – tzw. strefę (Zone) – przestrzeni adresowej systemu DNS. Jeśli serwer otrzyma zapytanie dotyczące adresu, który nie należy do obsługiwanej przez niego strefy, przekazuje to zlecenie do stojącego wyżej w hierarchii serwera DNS, ten zaś skieruje je do maszyny obsługującej właściwą strefę.

Informacje adresowe dotyczące danej strefy mogą być również udostępniane przez kilka serwerów DNS. Takie rozwiązanie pozwala na odpowiedni podział zadań związanych z ustalaniem adresów sieciowych. Samym operatorom sieciowym zależy na tym, aby przepływ danych związanych z zapytaniami DNS był jak najmniejszy i nie odbywał się pomiędzy różnymi krajami i kontynentami. Znacznie wygodniejszym rozwiązaniem jest udostępnianie użytkownikom własnych serwerów DNS. Wszyscy więksi operatorzy prowadzą dla potrzeb swoich podsieci odrębne serwery DNS, na których przechowywane są kopie danych strefowych, pochodzących z serwera nadrzędnego.

Ważną rolę odgrywa w tym przypadku funkcja cache. Gdy jakieś zapytanie dotyczy adresu spoza danej strefy, musi być ono wprawdzie przekazane do nadrzędnego serwera DNS, ale otrzymana odpowiedź – przed odesłaniem do komputera „pytającego" – może być zarejestrowana w pamięci buforowej. Przy kolejnym zapytaniu dotyczącym tego adresu serwer potrafi już na nie odpowiedzieć na podstawie danych przechowywanych w pamięci cache i nie musi korzystać z usług nadrzędnego serwera DNS.

Tworzenie przestrzeni nazw

Podział wyodrębnionych stref pomiędzy poszczególne serwery oraz przekazywanie zapytań do nadrzędnych serwerów DNS może prawidłowo funkcjonować tylko dlatego, że struktura nazw domen ma układ ściśle hierarchiczny. Na szczycie systemu znajdują się tzw. Top-Level-Domains (TLD), z których pewna część jest już zaliczana do powszechnych dóbr kulturalnych (np. domeny .com i .pl). Te ostatnie reprezentują dwie różne grupy zdefiniowanych już domen TLD. Pierwsza z nich kojarzy ze sobą domeny według ich znaczenia (treści), a druga – według ich pochodzenia geograficznego.

Do pierwszej kategorii należą m.in. domeny .cpm, .org, .net, .mil, natomiast w skład drugiej wchodzi ponad sto dwuliterowych symboli państw (od Afganistanu (af), który nie ma jeszcze w ogóle dostępu do Internetu aż po Zimbabwe (zw)). W globalnej sieci swoją reprezentację ma także Państwo Watykańskie (vk). Wymienione domeny TLD są zdefiniowane na stałe i nie można do nich dodawać kolejnych elementów. Jeśli więc ktoś chce utworzyć nową domenę, musi umieścić ją w ramach jednej z tych grup TLD. Symbol takiej domeny jest automatycznie dodawany do całej nazwy. Domena znanego czasopisma komputerowego nie nazywa się zatem po prostu „chip", lecz „chip.pl".

▶ 176

podstawy

Wszystko o nazwach domen

Reguły syntaktyczne

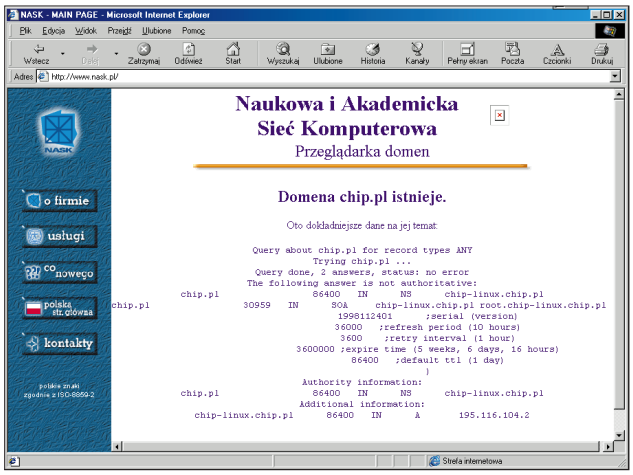
- Top-Level-Domain znajduje się na samym końcu nazwy domeny. Na lewo od niej umieszczona jest Second-Level-Domain, dalej Third-Level-Domain itd. Kolejne nazwy poszczególnych subdomen są oddzielone od siebie za pomocą kropki.
- Na samym początku znajduje się nazwa hosta (jeśli musi być ona dołączona do nazwy domeny), za którą umieszczona jest kropka oddzielająca kolejne subdomeny.
- Cały adres tekstowy złożony z nazwy hosta i wszystkich subdomen nie może być dłuższy niż 255 znaków.
- Żaden pojedynczy element nazwy (nazwa hosta lub subdomeny) nie może przekraczać maksymalnej długości 63 znaków.
- Pierwszy znak w nazwie hosta lub domeny musi być literą (z przedziału a–z lub A–Z).
- Oprócz małych i dużych liter w nazwie hosta lub domeny mogą występować tylko cyfry od 0 do 9 oraz znak minusa (-). Niedozwolone są jakiekolwiek znaki alfabetów narodowych.
- W systemie DNS nie są rozróżniane duże i małe litery. Nazwa *APPLE.COM* oznacza więc tę samą domenę co *apple.com*.

Nazwy hostów

Użytkownicy Internetu mają do czynienia z nazwami domen głównie przy wpisywaniu adresów URL w przeglądarkach webowych lub programach FTP. Znając hierarchiczną strukturę systemu DNS, można dojść do wniosku, że w adresie *www.chip.pl* nazwa *www* oznacza Third-Level-Domain w ramach Second-Level-Domain *chip.pl*. W rzeczywistości jednak *www* jest nazwą komputera (hosta) funkcjonującego w ramach domeny *chip.pl*. Na początku adresu URL – przed nazwą domeny – umieszcza się bowiem nazwę hosta. Zamiast *www* mógłby on zresztą nazywać się zupełnie inaczej, gdyż nigdzie nie zostało określone, że komputer internetowy z zainstalowanym serwerem http musi nosić nazwę *www*. Nazwa ta jest zatem stosowana zwyczajowo. O tym, że dany adres URL należy do sieci WWW, przeglądarka nie dowiaduje się za pośrednictwem umieszczonego w nim słowa *www*, lecz poprzez oznaczenie protokołu transmisji (http://). Z tego samego względu również serwer FTP nie musi koniecznie nosić nazwy *ftp*.

Adresy e-mailowe

System DNS odgrywa także ważną rolę w przypadku internetowej poczty elektronicznej, zwłaszcza przy przekazywaniu wiadomości pomiędzy różnymi serwerami mailowymi. Adres pocztowy typu *martin@chip.pl* nie zawiera bowiem żadnej informacji o nazwie lub adresie IP serwera pocztowego domeny *chip.pl*, do którego musi być przesłana dana wiadomość. Przy transmitowaniu wiadomości e-mailowych system DNS ma więc dodatkowe zadanie polegające na ustaleniu nazwy serwera pocztowego obsługującego daną domenę. Informacja ta może być wprowadzana przez administratorów DNS w odpowiednich miejscach w zbiorach Zone-Files. W tym przypadku stacja robocza użytkownika (klient), zamiast pytać o adres IP określonego hosta, musi dowiedzieć się, jaka jest nazwa serwera pocztowego obsługującego daną domenę, oraz uzyskać jego adres IP. Ten właśnie mechanizm sprawia, że przedstawiony wcześniej adres e-mailowy nie musi posiadać bardziej skomplikowanej postaci *martin@poczta.chip.pl*.



***Naukowa i Akademicka Sieć Komputerowa* zajmuje się administracją domeny .pl. W udostępnionej przez tę instytucję bazie danych możemy sprawdzić, czy wybrana domena .pl została już przez kogoś zarejestrowana. Na przedstawionym zdjęciu widać wynik wyszukiwania hasła „chip.pl"**

Taka domena jest określana mianem Second-Level-Domain. W praktyce domena ta mogłaby nosić również nazwę *chip.com*, gdyby redakcja pisma zarejestrowała ją nie w kategorii geograficznej .PL, ale w komercyjnej. Oprócz tej alternatywy nie ma już właściwie żadnych innych możliwości wyboru. Nie można użyć innej domeny geograficznej, gdyż nie jest dozwolone rejestrowanie się w ramach domeny innego kraju (chyba że znajduje się w nim dany serwer). Zastrzeżone są również „znaczeniowe" domeny TLD: .mil – dla serwerów armii Stanów Zjednoczonych, .edu – dla amerykańskich instytucji edukacyjnych, oraz .gov – dla jednostek administracji USA.

W tym miejscu wyraźnie widać amerykańskie pochodzenie systemu DNS. Chociaż struktura ta jest obecnie dostępna dla wszystkich krajów świata, to w momencie powstawania Internetu Amerykanie byli jedynymi, którzy intensywnie korzystali z jego zasobów, i pierwszymi, którzy dokonywali rejestracji domen TLD.

Przestrzeń nazw systemu DNS nie kończy się jednak na drugim poziomie. W strukturze tej może bowiem funkcjonować niemal dowolna liczba kolejnych poziomów. Opisywana hierarchia ma ścisły związek z podziałem stref pomiędzy poszczególne serwery DNS. Ten, kto zarezerwował dla siebie Second-Level-Domain, otrzymuje bowiem jednocześnie prawo do tworzenia w jej ramach kolejnych subdomen i administrowania nimi. Z faktem tym wiąże się oczywiście również instalacja serwera DNS lub zlecenie tego zadania swemu operatorowi internetowemu. Takie rozwiązanie ma sens głównie w przypadku większych instytucji dysponujących rozproszoną strukturą informatyczną, które chcą wprawdzie połączyć w jedną Second-Level-Domain kilka oddziałów lub budynków, ale każda z tych jednostek ma mieć możliwość administrowania własną Third-Level-Domain. Na tej właśnie zasadzie tworzone są takie nazwy domen, jak *cri.reston.va.us*, którą obsługuje Computer Research Institute (cri) na Uniwersytecie Reston (reston) w amerykańskim stanie Virginia (va.us).

Zalety struktury drzewiastej

Główną zaletą systemu hierarchicznego jest to, że stwarza mniej konfliktów nazw niż system „płaski", jednowymiarowy. Nazwy domen *chip.com* i *chip.pl* różnią się między sobą tak samo jak *chip.pl* i *chip.vogel.pl*, gdyż należą do różnych domen nadrzędnych. Inna zaleta tego systemu wiąże się ze wspomnianym już wcześniej podziałem stref. Pod pojęciem strefy należy rozumieć jedną lub kilka domen wraz z subdomenami, które są administrowane przez jeden serwer DNS.

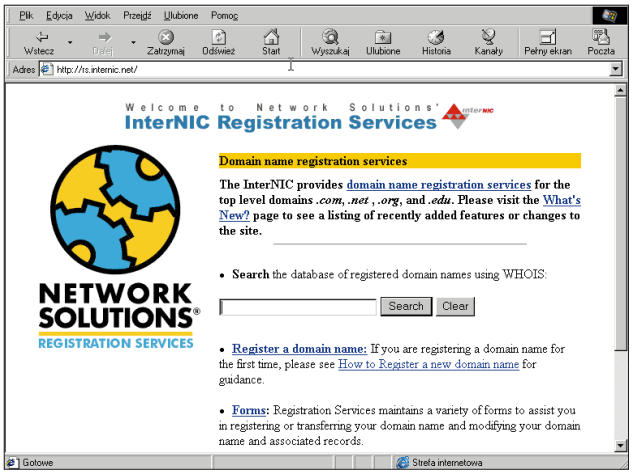
Ściśle rzecz biorąc, zawsze istnieją przynajmniej dwa komputery, które mogą udzielać informacji dotyczących określonej strefy. Koncepcja systemu DNS zakłada bowiem, że w celu zapewnienia niezawodności pracy, oprócz podstawowego DNS-Servera (Primary), musi być zawsze określony także serwer zapasowy (Secondary DNS-Server). W przypadku awarii lub zbyt dużego obciążenia pierwszego serwera zapytania pochodzące ze stacji roboczych (klientów) są przekazywane do drugiego z nich.

Same strefy reprezentowane są na serwerze przez tzw. Zone-Files, czyli zwykłe pliki ASCII, w których znajdują się wykazy nazw wszystkich hostów występujących w ramach danej domeny wraz z odpowiadającymi im adresami IP. Ponadto w tym miejscu mogą być zdefiniowane poszczególne subdomeny należące do bieżącej domeny oraz inne serwery DNS, jeżeli dana subdomena tworzy samodzielną strefę na innym serwerze DNS. Korzystając ze zbioru Zone-File, serwer taki może więc udzielać informacji dla wszystkich zawartych w nim domen. Serwery DNS mogą ponadto wykorzystywać kilka różnych Zone-Files.

Zadania Root-Serverów

Na szczycie całej struktury systemu DNS znajdują się tzw. Root-Level-Servers, które znajdują się w Stanach Zjednoczonych i podłączone są do dużych magistral (Backbones), stanowiących „kręgosłup" całego Internetu. W tym miejscu utrzymywane są niegeograficzne Top-Level-Domains, ale rzadko definiuje się tu konkretne komputery hostów. Zamiast tej informacji dla poszczególnych subdomen typu ibm.com czy white-house.gov wskazywane są tylko odpowiednie serwery DNS, które administrują daną strefą i znajdującymi się w niej hostami i subdomenami. Podobnie dzieje się również w przypadku domen geograficznych, których Second-Level-Domains w poszczególnych krajach są obsługiwane przez centralnego administratora, troszczącego się też o odpowiednie zbiory Zone-Files.

▶ 179



W przeciwieństwie do domen narodowych (np. polskiej .pl) rezerwacji domen z identyfikatorem .com dokonuje się za pośrednictwem serwera webowego serwisu InterNIC, dostępnego pod adresem *http://rs.internic.net/*

Ze względu na fakt, że każdy serwer DNS zna nazwę swojego serwera nadrzędnego, zapytania DNS mogą czasem trafić również do Root-Level-Servera. Dzieje się tak wówczas, gdy żaden z napotkanych wcześniej serwerów DNS nie ma w swojej pamięci podręcznej poszukiwanej nazwy hosta lub nie dysponuje zbiorem Zone-File dla odpowiedniej domeny. W sytuacji gdy także Root-Level-Servers nie potrafią zidentyfikować adresata konkretnego zapytania, korzystają z pomocy znanych im serwerów DNS odpowiedzialnych za daną domenę Top-Level lub Second-Level. Dzięki takiemu rozwiązaniu otrzymane zapytanie może być zawsze przekazane do właściwego serwera.

W najgorszym przypadku zapytanie DNS wędruje więc poprzez kilka serwerów do poziomu Root-Level-DNS-Server, aby potem za pośrednictwem jednego lub kilku następnych serwerów dotrzeć do właściwej domeny. Tutaj trafia wreszcie do serwera DNS, który obsługuje hosta należącego do poszukiwanej subdomeny (lub nawet sub-sub-sub-subdomeny). Operatorzy sieciowi wymieniają jednak między sobą swoje Zone-Files, dzięki czemu droga do kompetentnego serwera DNS bardzo rzadko jest aż tak długa.

Administrowanie nazwami

Zarządzanie poszczególnymi domenami jest zadaniem bardzo pracochłonnym, gdyż na świecie jest już zarejestrowanych ponad pół miliona Second-Level-Domains. Dla różnych domen Top-Level istnieją ponadto różne instytucje rejestrujące, które przyjmują zamówienia na domeny, przydzielają wybrane nazwy, wprowadzają je do odpowiednich Zone-Files i obsługują serwery DNS, które na podstawie tych plików udzielają właściwych informacji. Dla niegeograficznych Top-Level-Domains – włącznie z najważniejszą domeną .com – funkcję tę pełni instytucja InterNIC, której witryna webowa jest dostępna pod adresem rs.internic.net.

Do marca 1998 roku rejestracja domeny .com i jej obsługa przez pierwsze dwa lata kosztowała 100 dolarów, zaś za każdy kolejny rok trzeba było płacić po 50 dolarów. Od kwietnia 1998 roku obowiązuje już obniżona taryfa: 70 dolarów za pierwsze dwa lata eksploatacji i 35 dolarów za każdy następny rok.

Domeną .pl zarządza organizacja Naukowa i Akademicka Sieć Komputerowa. Dokładne informacje dotyczące rejestracji domen w hierarchii .pl można znaleźć na stronie *http://www.nask.pl/*. Jeśli nie zamierzamy sami uruchamiać serwera DNS, wszystkie formalności związane z rejestracją domeny powinien za nas załatwić dostawca usług internetowych, któremu zlecimy prowadzenie dla nas takiego serwera.

Najpierw sprawdźmy, potem rezerwujmy

Zanim złożymy wniosek o rejestrację domeny, powinniśmy upewnić się, czy domena ta nie została już przydzielona komuś innemu. Każda instytucja rejestrująca ma obowiązek udostępniać w tym celu stronę WWW „WHOIS", na której możemy wpisać planowaną nazwę własnej domeny. Jako odpowiedź otrzymamy wówczas dane właściciela tej domeny (nazwa firmy, adres kontaktowy) lub informację o tym, że nie jest ona jeszcze zarejestrowana. Nazwa takiej strony informacyjnej pochodzi od unixowego programu WHOIS, umożliwiającego podłączenie się do systemu DNS z poziomu linii poleceń.

Jeśli chcemy obecnie zarezerwować sobie domenę z zakresu np. pl, to musimy za pośrednictwem swojego operatora przedstawić adres IP, pod którym ma być dostępny serwer tej domeny. Takie rozwiązanie ma zapobiegać masowemu rezerwowaniu atrakcyjnych nazw przez osoby, które nie mają wcale zamiaru wykorzystać ich do prezentacji informacji w Internecie (WWW czy FTP). Przez pewien czas sprytni „biznesmeni" mogli bowiem nieźle zarabiać, rezerwując sobie jako nazwy domen znane znaki firmowe ▶ 180

podstawy

Więcej informacji na temat serwerów DNS

Bastiony Internetu

Takim mianem można określić internetowe Root-Level-Servery, do których trafiają te zapytania dotyczące domen, z którymi wcześniej nie mógł poradzić sobie żaden z niższych rangą serwerów DNS. Każdy taki serwer musi więc znać co najmniej adresy IP odpowiednich Root-Level-Serverów. W celu lepszego rozkładu obciążenia do dyspozycji użytkowników pozostaje dziewięć serwerów tego typu (oznaczonych literami od A do I). Jeśli wszystkie te komputery ulegną awarii lub zostaną unieruchomione na skutek sabotażu, komunikacja w Internecie może stać się prawdziwym problemem.

Nazwa Root-Servera	Adres IP
A.ROOT-SERVERS-NET	198.41.0.4
B.ROOT-SERVERS-NET	128.9.0.107
C.ROOT-SERVERS-NET	192.33.4.12
D.ROOT-SERVERS-NET	128.8.10.90
E.ROOT-SERVERS-NET	192.203.230.10
F.ROOT-SERVERS-NET	39.13.229.241
G.ROOT-SERVERS-NET	192.112.36.4
H.ROOT-SERVERS-NET	128.63.2.53
I.ROOT-SERVERS-NET	192.36.148.17

Inverse Lookup

Uzyskiwanie właściwego adresu IP na podstawie nazwy hosta jest głównym zadaniem systemu DNS. Czasami chcemy jednak wykonać odwrotne zadanie – znając adres IP, ustalić nazwę hosta i jego domeny. Taka sytuacja może się zdarzyć np. przy analizowaniu tworzonych automatycznie przez serwery WWW statystyk dotyczących obsługiwanych klientów. Z uwagi na fakt, że serwery te znają tylko ich adresy IP, to aby uzyskać dodatkowe informacje o pochodzeniu poszczególnych wywołań, należy dokonać konwersji numerów na nazwy.

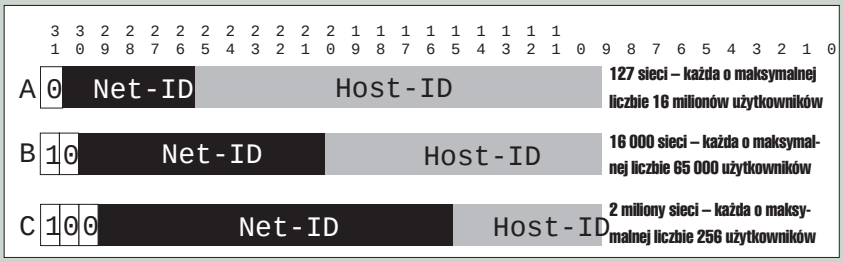
System DNS oferuje do tego celu funkcję Inverse Lookup, która po wcześniejszym przekształceniu do specjalnej postaci danego adresu IP korzysta ze standardowego mechanizmu DNS. Przekształcenie to polega na odwróceniu kolejności poszczególnych liczb w adresie IP i potraktowaniu go jako subdomeny w ramach specjalnej domeny „in-addr.arpa". Aby np. znaleźć nazwę hosta o adresie IP 213.168.65.3, do systemu DNS musi być skierowane zapytanie o adres 3.65.168.213.in-addr.arpa. Przy odrobinie szczęścia otrzymamy wówczas jako odpowiedź nazwę danego hosta oraz jego domeny; nie każdy bowiem serwer DNS ujawnia takie informacje. Taką formę zapytania można przecież wykorzystać również do śledzenia poszczególnych domen i uzyskiwania informacji o nazwach funkcjonujących w ich ramach hostów. Administrator serwera DNS decyduje więc na poziomie strefy, czy funkcja Inverse Lookup jest w niej dozwolona, czy też nie.

podstawy

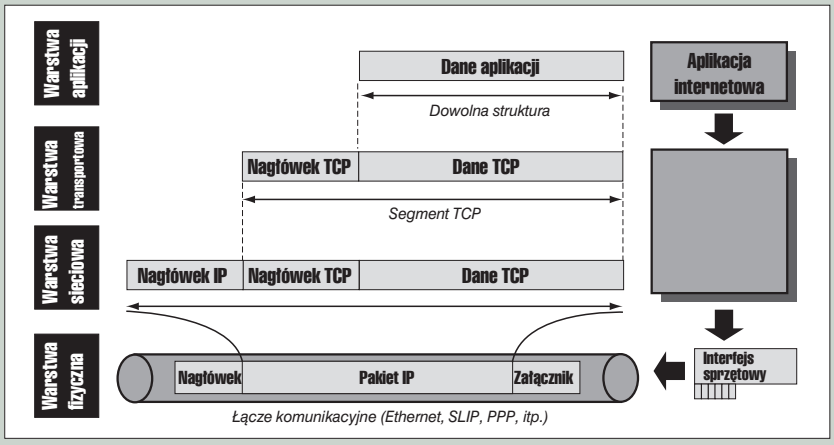
Budowa adresów IP

Z uwagi na fakt, że na 32 bitach adresu IP musi znaleźć się zarówno identyfikator sieci (Net-ID), jak i identyfikator hosta (Host-ID), pojawia się pytanie, jak wiele bitów należy przeznaczyć na każdą z tych informacji. Im więcej bitów zarezerwujemy na Net-ID, tym więcej podsieci będzie mogło funkcjonować w przestrzeni adresowej IP. W takiej sytuacji pozostanie jednak mniej bitów na numer Host-ID, co oznacza, że do danej podsieci będzie można podłączyć mniej hostów. Ze względu na to, że Internet powinien obejmować zarówno duże, jak i małe sieci, kwestia ta nie została rozstrzygnięta jednoznacznie: utworzono bowiem trzy klasy adresów IP (A, B i C).
We wszystkich tych klasach długość numeru Net-ID jest większa niż Host-ID. Adresy klasy A są przeznaczone dla największych sieci, klasy B – dla sieci o średniej wielkości, a C – dla sieci mniejszych. Na podstawie najbardziej znaczących bitów adresu IP oprogramowanie internetowe rozpoznaje klasę, do której należy dany numer.

Regułom tym nie podlegają dwa specjalne adresy IP, z którymi mieliśmy już może kiedyś do czynienia: 127.0.0.1 oraz 127.1.1.1. Oba są zarezerwowane dla funkcji „Loopback" i oznaczają tego samego hosta. Jeśli np. przeglądarka WWW ma skontaktować się z serwerem http umieszczonym na tym samym komputerze, nie musimy znać ani nazwy domeny, ani konkretnego adresu IP – wystarczy jako odpowiedni adres URL wpisać po prostu http://127.0.0.1/.



Tak wygląda struktura adresu IP w poszczególnych klasach A, B i C



Dane aplikacji są umieszczane w pakietach TCP, które z kolei wchodzą w skład pakietów IP

lub inne atrakcyjne hasła. Gdy jakaś firma chciała później wykorzystać dla siebie określoną nazwę, musiała odkupić daną domenę od takiego „handlowca" lub próbować odzyskać prawo do niej na drodze sądowej. W kilku przypadkach to drugie rozwiązanie okazało się nawet skuteczne.

Rozwiązania na przyszłość

Chociaż system DNS działa niezawodnie i codziennie obsługuje miliony zapytań, to za kulisami trwają obecnie gorące dyskusje nad nowym sposobem administrowania systemem oraz nad dodaniem kolejnych domen Top-Level. Na początku 1997 roku organizacja ISOC (Internet Society) opublikowała oficjalne pismo, w którym zakłada utworzenie siedmiu nowych Top-Level-Domains oznaczonych symbolami .firm, .store, .web, .arts, .rec, .info i .nom. Znalazła się tam również propozycja utworzenia możliwie dużej liczby prywatnych instytucji rejestrujących domeny DNS, co ma spowodować powstanie w tym sektorze konkurencji i zapewnić utrzymanie niskich cen za rejestrację i późniejszą administrację domen.
Na razie jednak rząd USA przedłużył o dwa lata wygasającą 30 września umowę z NSI (InterNIC) na zarządzanie domenami com., net. i org. Począwszy od kwietnia 1999 r. na rynek wejdzie pięć kolejnych firm, rejestrujących nazwy w wymienionych domenach. Jednym z warunków przedłużenia umowy z NSI jest przedłożenie do dyspozycji rządu USA wszystkich zgromadzonych dokumentacji, danych i programów związanych z obsługą spornych domen. Materiały te mają zostać przekazane nowej organizacji, która nadzorować będzie funkcjonowanie firm rejestrujących.

Wojenny rodowód

Zapewne niewiele osób zdaje sobie dziś sprawę z tego, że Internet zaczął funkcjonować w połowie lat siedemdziesiątych jako projekt badawczy armii Stanów Zjednoczonych. Projekt ten stracił swój militarny charakter dopiero pod koniec lat osiemdziesiątych, gdy wojsko wycofało się z jego pilotowania, a sieć zaczęła zwiększać swój zasięg – najpierw w Ameryce, a potem w Azji i Europie. Dopiero wówczas pierwotna sieć ARPA-Net przekształciła się w Internet, który funkcjonuje do dzisiaj.

Chociaż techniczne podstawy współczesnego Internetu zostały ustalone jeszcze w epoce „militarnej", to nadal – mimo powstania nowoczesnych programów usługowych i dodatkowych struktur – są one bardzo wyraźnie widoczne. Podstawowym wymogiem, jaki wojsko postawiło przed nową siecią, była bowiem decentralizacja. Takie rozwiązanie miało sprawić, że żaden pojedynczy atak jądrowy przeciwnika nie mógłby unieruchomić całej instalacji.

W konsekwencji współczesny Internet nie jest fizycznie żadną samodzielną siecią, lecz związkiem wielu mniejszych, już istniejących. Do tej ogólnoświatowej struktury mogą więc należeć zarówno małe sieci lokalne (LAN) z kilkoma pecetami, jak i duże sieci dysponujące wieloma tysiącami stacji roboczych. Ideą takiego rozwiązania było to, że każdy komputer pracujący w dowolnej części Internetu może nawiązać łączność z dowolnym innym użytkownikiem Sieci.

Routery zapewniają łączność

Kluczową rolę w takiej strukturze odgrywają tzw. routery, czyli mechanizmy zajmujące się przesyłaniem pakietów danych z jednej części sieci do innej. W każdej części Internetu zainstalowany jest co najmniej jeden router zbierający wszystkie pakiety danych, które nie mogą być dostarczone w ramach danej sieci lokalnej. Routery mogą mieć postać sprzętową lub funkcjonować jako oprogramowanie w środowisku DOS, Windows lub Linux. Część routerów ma stałe połączenie z Internetem, natomiast inne nawiązują łączność w sposób dynamiczny – tylko w razie potrzeby.

Poszczególne pakiety danych w Internecie są tak długo przekazywane z routera do routera i z jednej części sieci do drugiej, aż trafią do tego obszaru, w którym znajduje się adresat danego pakietu. Obsługujący ten obszar router dostarcza następnie określony pakiet do odbiorcy – w taki sposób jakby pochodził on z sieci lokalnej.

Jeśli jedna z dołączonych podsieci ulegnie awarii lub będzie niedostępna z powodu przeciążenia, to pakiety mogą zostać skierowane do adresata (i dotrzeć do niego) inną drogą. W ten sposób Internet spełnia najważniejszy warunek, jaki pierwotnie narzuciła mu armia Stanów Zjednoczonych: może funkcjonować dalej mimo uszkodzenia poszczególnych podsieci.

Koncepcja ogólnoświatowej sieci zawdzięcza jednak swoją uniwersalność temu, że pozwala na łączenie ze sobą dowolnych podsieci o zupełnie różnych strukturach i typach okablowania. W Internecie funkcjonują więc wspólnie instalacje Ethernet, Token-Ring, ATM, ISDN, sieci lokalne i sieci rozległe, zapewniając łączność pomiędzy różnymi systemami operacyjnymi i osprzętem sieciowym. Warunki współpracy tych wszystkich składników Internetu określają odpowiednie protokoły i standardy.

Protokoły sieciowe dbają o to, aby odmienne właściwości poszczególnych podsieci były niewidoczne dla aplikacji internetowych. Do tych różnic zalicza się różne sposoby kodowania fizycznych adresów sieciowych, inne maksymalne wielkości pakietów danych oraz odmienne metody dołączania sum kontrolnych, stosowanych do korekcji błędów. Wszystkimi tymi kwestiami zajmuje się tzw. pakietowa i transportowa warstwa Internetu, która obsługuje transmisję sieciową. Jeśli więc mówimy o sieci TCP/IP, to pod tym pojęciem należy rozumieć dwa podstawowe protokoły internetowe: protokół pakietowy IP oraz protokół transportowy TCP.

Protokół IP

Najniższą warstwę Internetu tworzy tzw. Internet Protocol (IP), który pełni funkcję podstawowego mechanizmu obsługi danych i definiuje dwie bardzo istotne kwestie: budowy pakietów danych oraz tworzenia adresów hostów internetowych.

Pierwsze zagadnienie dotyczy sposobu „opakowania" danych, które mają być przesłane od jednego hosta internetowego do drugiego. Protokół IP definiuje więc 20-bajtowy nagłówek, który – umieszczony przed właściwymi danymi – pełni funkcję etykiety adresowej. Oprócz adresu nadawcy i odbiorcy zawiera on informacje dodatkowe, jak długość pakietu, suma kontrolna czy tzw. licznik stacji, mający zapobiegać ciągłemu krążeniu w Sieci pakietów, których doręczenie do adresata nie jest możliwe.

Druga kwestia dotyczy adresowania hostów internetowych i wiąże się ściśle ze znanymi nam czteroliczbowymi numerami identyfikacyjnymi (np. 201.93.43.2). Pod taką sekwencją kryją się zawsze cztery bajty (32 bity), na których zapisane są dwie informacje: Net-ID i Host-ID. Identyfikator Net-ID opisuje podsieć, do której należy określony host. Poszczególne numery są przydzielane przez centralną instytucję administrującą (IANA – Internet Assigned Numbers Authority), co gwarantuje ich unikatowość. Host-ID określa natomiast numer danego hosta w ramach danej podsieci i może być nadawany przez jej administratora.

Podział adresów IP na identyfikatory Net-ID i Host-ID jest podstawowym warunkiem sprawnego kierowania ruchem pakietów IP (czyli tzw. routingu). Gdy dany host internetowy na polecenie określonej aplikacji zamierza wysłać pakiety danych IP, już na początku – dzięki porównaniu własnego Net-ID z identyfikatorem adresata – rozpoznaje on, czy pakiety te mają pozostać w danej sieci lokalnej. Jeśli tak, to jest on w stanie dostarczyć je do adresata bez pomocy routera. Jeśli natomiast nie, host przesyła je do routera obsługującego daną podsieć w celu przekazania ich do podsieci adresata. Router ten odczytuje Net-ID sieci docelowej z umieszczonego w nagłówku IP adresu odbiorcy, a następnie sprawdza w tzw. tabeli routingu, jaka droga transmisji pakietu byłaby najlepsza. W tabeli tej znajduje się wykaz routerów obsługujących wszystkie znane podsieci oraz adres domyślnego routera, do którego powinny być kierowane te pakiety IP, których Net-ID nie został uwzględniony w tabeli. Tabele takie muszą być stale aktualizowane, w związku z czym ich zawartość jest automatycznie wymieniana pomiędzy poszczególnymi routerami za pomocą specjalnych protokołów routingu.

Z punktu widzenia routera przesłanie pakietów IP z jednej podsieci do drugiej oznacza umieszczenie ich w odpowiedniej dla danej sieci „ramce", gdyż dopiero w takiej postaci są gotowe do transmisji. Ani Ethernet, ani Apple-Talk lub jakakolwiek inna sieć nie potrafi bowiem na poziomie warstwy pakietowej obsługiwać pakietów IP. Przesyłki takie muszą być więc umieszczone w ramkach specyficznych dla danego protokołu, dzięki czemu nie będą się niczym odróżniały od standardowych pakietów określonej sieci. Z uwagi na fakt, że w różnych technologiach sieciowych stosowane są odmienne długości pakietów, routery napotykają poważny problem: otrzymane pakiety IP są często zbyt duże, aby można je było przekazać w całości do dołączonej sieci.

Istotnym elementem protokołu IP jest zatem mechanizm fragmentacji danych. Jeśli sytuacja tego wymaga, routery mogą dzielić wychodzące pakiety IP na kilka specyficznych dla danej sieci przesyłek i wysyłać je kolejno po sobie. Już w nagłówku IP zarezerwowane są odpowiednie pola, które pozwalają rozpoznać dany pakiet jako część większej całości i ustawić go we właściwej kolejności. Na tej podstawie ▶ 182

podstawy

Najpopularniejsze usługi internetowe i ich numery portów

Port	Protokół	Zadanie
20+21	FTP	Transmisja danych
23	TELNET	Wiersz poleceń zdalnego hosta
25	SMTP	Nadawanie przesyłek e-mailowych
53	DNS	Odwzorowywanie nazw domen na postać adresów IP
70	GOPHER	Usługa wyszukiwawcza dla Archive
80	HTTP	World Wide Web
110	POP3	Pobieranie przesyłek e-mailowych
119	NNTP	Usenet
161	SNMP	Zdalne administrowanie urządzeniami sieciowymi
194	IRC	Internet Chat
666	DOOM	Ulubiona gra wraz z odpowiednim numerem portu
27000	QuakeWorld	Internetowa wersja Quake'a

podstawy

IP w środowisku Windows

W Windows oprogramowanie obsługujące protokoły IP, ICMP, ARP, TCP i UDP wchodzi w skład biblioteki WINSOCK.DLL. W przypadku 16-bitowej wersji systemu plik ten nosi nazwę WINSOCK.DLL, a w wersji 32-bitowej – WSOCK32.DLL. Biblioteka ta obsługuje tzw. mechanizm „TCP/IP Stack", który umożliwia aplikacjom nadrzędnym komunikowanie się z innymi aplikacjami internetowymi przy użyciu usług TCP i UDP. W tym celu WINSOCK.DLL udostępnia standardowy zestaw funkcji, za pomocą których mogą być odbierane i nadawane pakiety UDP, otwierane i zamykane kanały TCP oraz transmitowane dane. Zestaw taki nosi nazwę interfejsu WINSOCK. Języki C, C++ i Pascal pozwalają na bezpośrednie odwoływanie się do tego interfejsu, natomiast w przypadku Visual Basica jego funkcję pełni dodatkowy moduł Control.

adresat może więc odtworzyć pierwotną postać wysłanego pakietu IP.

ARP obsługuje transmisję fizyczną

Dostarczenie danego pakietu IP do właściwego hosta w lokalnej sieci lub kolejnego routera wymaga znajomości fizycznego adresu sieciowego komputera docelowego. W tym przypadku nie chodzi o adres IP, który jest tylko adresem logicznym w Internecie, lecz o konkretny adres karty sieciowej określonego komputera. Adres ten zajmuje centralne miejsce w nagłówku danego protokołu sieciowego, który w ramach sieci lokalnej musi poprzedzać właściwy pakiet IP.

Skąd jednak ma pochodzić fizyczny adres sieciowy, gdy znany jest tylko adres IP adresata? Rozwiązanie tego problemu umożliwia protokół ARP (Address Resolution Protocol), który jest obsługiwany przez wszystkie hosty internetowe i ściśle współpracuje z protokołem IP. ARP definiuje prosty mechanizm, za pomocą którego router lub host internetowy może na podstawie podanego adresu IP uzyskać od lokalnej sieci informację o odpowiednim adresie fizycznym. W tym celu wysyłany jest specjalny pakiet ARP Broadcast, który nie jest skierowany do określonego użytkownika, lecz do wszystkich stacji należących do danej sieci lokalnej.

Każda stacja odbiera taki pakiet i sprawdza, czy zawarty w nim adres IP odpowiada jej własnemu numerowi. Jeśli tak, to wysyła do nadawcy pakietu odpowiedź, w której podaje swój fizyczny adres sieciowy. Jeśli nie, to nie udziela po prostu żadnej odpowiedzi. Gdy żadna ze stacji nie odpowie na otrzymany pakiet, oznacza to, że dany adres IP nie jest znany w tej sieci lokalnej. W takim wypadku pakiet IP nie może zostać doręczony do adresata i jego transmisja zostaje przerwana.

ICMP sprawuje kontrolę nad ruchem w sieci

Aby router nie był ciągle zasypywany pakietami IP, których nie może przekazać dalej, Internet został wyposażony w specjalny protokół ICMP, umożliwiający kontrolowanie ruchu w sieci. Przesyłki nadawane za pomocą tego protokołu mają wyższy priorytet, a więc są transmitowane szybciej niż zwykłe pakiety. Protokół ICMP (Internet Control Message Protocol) określa sposób tworzenia i wymiany informacji sterujących pomiędzy routerami lub routerami i hostami. Za pomocą wiadomości ICMP router może więc np. zgłosić, że dana podsieć jest przeciążona i pakiety IP powinny być do niej kierowane przez inny router. Dany router może też – na przesłane poprzez ICMP zapytanie hosta – przekazać mu informację o numerze Net-ID aktualnej podsieci lub o jej masce. Z protokołu ICMP korzysta również popularna funkcja PING, za pomocą której można sprawdzić obecność zdalnego hosta i zmierzyć czas transmisji nadawanych do niego pakietów.

Porty pomagają rozdzielić poszczególne przesyłki

Protokoły IP i ARP wchodzą w skład internetowej warstwy wymiany pakietów. Na wyższym poziomie znajduje się natomiast warstwa transportowa, którą obsługują protokoły UDP i TCP. Gdy przez Internet komunikują się ze sobą takie aplikacje, jak serwery i przeglądarki WWW, muszą one korzystać z pomocy jednego z tych dwóch ostatnich protokołów. Nie mają natomiast w ogóle dostępu do pakietów IP, gdyż pakiety te są wysyłane do hostów internetowych, a nie aplikacji. Jeśli na danym hoście funkcjonuje kilka takich aplikacji, nie istnieje tu możliwość jednoznacznej identyfikacji, czy dany pakiet IP jest kierowany do serwera WWW czy FTP.

Aby dokonać takiej identyfikacji, protokoły UDP i TCP korzystają z pomocy tzw. portów, których funkcję pełni specjalna liczba 16-bitowa (WORD). Podobnie jak pakiety IP, w których właściwe dane poprzedza odpowiedni nagłówek, również pakiety UDP i TCP wykorzystują taki element. Najważniejszą informacją zawartą w nagłówku jest identyfikator portu adresata, do którego ma trafić dany pakiet, oraz identyfikator portu nadawcy. Gdy oprogramowanie sieciowe otrzyma pakiet UDP lub TCP, może za pomocą numeru portu rozpoznać, do jakiej aplikacji należy przesyłka, i skierować nadesłane dane do właściwego programu.

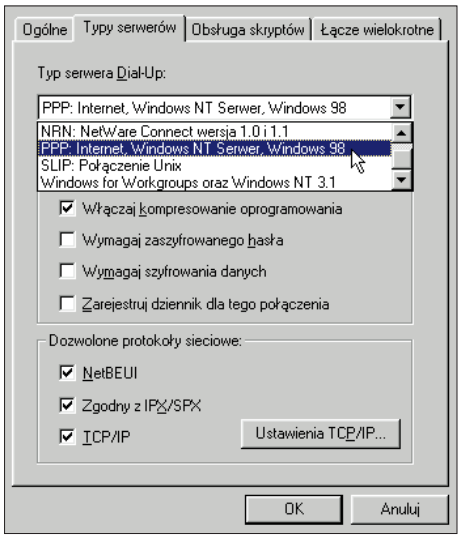
W celu nawiązania komunikacji z serwerem internetowym potencjalny klient potrzebuje nie tylko jego numer IP, ale również numeru portu. Numer ten musi być więc z góry znany. Z tego też względu poszczególne numery portów z przedziału 0-1024 są na stałe przyporządkowane różnym usługom internetowym (takim jak FTP czy HTTP). Gdy zatem w przeglądarce WWW wpiszemy adres URL (np. *http://www.chip.pl*), program ten będzie od razu wiedział, że ma odwołać się do portu numer 80 w komputerze, który kryje się pod nazwą *www.chip.pl*.

Takie rozwiązanie sprawia jednak, że hakerzy mogą stosunkowo łatwo docierać

podstawy

Protokoły komunikacji internetowej

IP	Internet Protocol	Określa strukturę adresów internetowych i pakietów danych
ICMP	Internet Controll Message Protocol	Umożliwia dostrojenie parametrów transmisji i sterowanie przepływem danych pomiędzy routerami
UDP	User Datagram Protocol	Zapewnia transport pakietów danych pomiędzy aplikacjami internetowymi, ale bez gwarancji ich dostarczenia
TCP	Transmission Control Protocol	Określa strukturę zabezpieczonego kanału komunikacyjnego pomiędzy aplikacjami internetowymi
ARP	Address Resolution Protocol	Zapewnia odwzorowanie adresów hostów na postać fizycznych adresów sieciowych
RARP	Reverse Address Resolution Protocol	Umożliwia ustalenie adresu IP dla danego fizycznego adresu sieciowego
SLIP	Serial Line IP Protocol	Zapewnia transmisję pakietów IP za pomocą łączy szeregowych (modem)
PPP	Point-to-Point Protocol	Umożliwia transmisję pakietów IP za pomocą dowolnych łączy Point-to-Point



W przypadku modemów i sieci ISDN najchętniej stosowanym protokołem internetowym jest PPP, gdyż oferuje lepsze możliwości pracy niż jego poprzednicy: SLIP i CSLIP

do całych grup adresów IP i przez nawiązanie kontaktu za pomocą znanych numerów portów sprawdzać, czy pod danym adresem funkcjonuje określony serwer. Jeśli tak, to próbują się do niego włamać, wykorzystując w tym celu słabe punkty danej usługi. Z drugiej strony istnieje dość prosty sposób, pozwalający na ukrycie określonych usług internetowych przed zwykłymi użytkownikami. W tym celu wystarczy po prostu przypisać dany serwer do nietypowego numeru portu. I tak zamiast korzystać z portu numer 80, można uruchomić serwer WWW na porcie 7436. Kontakt z tym serwerem będą mogły nawiązać tylko osoby, które znają ten numer i w swojej przeglądarce wpiszą adres URL w postaci *http://www.chip.pl:7436/*. Podanie oznaczenia protokołu jest w tym wypadku konieczne, gdyż inaczej przeglądarka nie będzie wiedziała, jakiej usługi się pod wyspecyfikowanym portem spodziewać.

Bezpieczna transmisja danych

Pakiety TCP i UDP są transportowane wewnątrz pakietów IP, dzięki czemu kolejne routery mogą obsługiwać je tak samo jak standardowe pakiety IP. Oznacza to również, że w przypadku przeciążenia routera, zablokowania podłączonej podsieci lub nieaktywności odbiorcy wszystkie przesyłki takie czeka podobny los: znikną one z sieci bez jakiegokolwiek śladu. W ten sposób nadawca nigdy nie uzyska informacji o tym, czy dany pakiet rzeczywiście dotarł do adresata.

Większość aplikacji sieciowych musi mieć jednak pewność, że wysłane dane dotarły do odbiorcy bez przekłamań i we właściwej kolejności. Protokoły IP oraz UDP (User Datagram Protocol) nie oferują wprawdzie takich możliwości, ale TCP (Transmission Control Protocol) już tak. Z tego też względu przeważająca część aplikacji internetowych wykorzystuje do transmisji danych właśnie protokół TCP.

W przeciwieństwie do UDP, TCP jest protokołem dość skomplikowanym. W celu uzyskania pewności, że wszystkie wysłane dane zostały prawidłowo przetransmitowane, korzysta on z pomocy sum kontrolnych, numerów bloków oraz potwierdzeń od odbiorcy. Wszystkie przesyłki, które w ciągu określonego czasu nie dotarły do adresata lub uległy jakimkolwiek przekłamaniom, są automatycznie nadawane ponownie. Dopiero wówczas, gdy kilka takich prób nie odniesie skutku i adresat nie nadeśle żadnego potwierdzenia, transmisja danych jest przerywana, a nadawca informowany o jej niepowodzeniu.

Połączenie za pomocą modemu lub sieci ISDN

Większość użytkowników Internetu nie ma stałego połączenia za pośrednictwem sieci lokalnej, lecz łączy się z odpowiednim operatorem systemu za pomocą modemu lub ISDN. Transmisja pakietów IP przy użyciu takiego łącza Point-to-Point (od naszego peceta do komputera węzłowego operatora) wymaga zastosowania dodatkowych protokołów. Są one potrzebne choćby do odseparowania transmitowanych pakietów IP, gdyż poprzez łącza modemowe dane przesyłane są powoli (bajt po bajcie), w związku z czym trzeba w pewien sposób zaznaczyć, gdzie kończy się jeden pakiet, a gdzie zaczyna następny. Najprostszym i najstarszym tego typu protokołem jest SLIP – „Serial Line IP Protocol". Na końcu każdego pakietu IP SLIP dołącza więc odpowiedni bajt rozpoznawczy, dzięki czemu odbiorca danych może rozdzielić poszczególne pakiety.

Nieco bardziej zaawansowanym protokołem jest CSLIP (*Compressed SLIP*), który zmniejsza nieco strumień danych przesyłany łączem szeregowym. W tym celu wykorzystuje on fakt, że nagłówki występujących po sobie pakietów IP i TCP są do siebie podobne, w związku z czym wystarczy przesyłać tylko różnice pomiędzy nimi. Protokół ten nie dokonuje jednak kompresji właściwych danych umieszczonych w pakietach.

Protokoły SLIP i CSLIP przestały być powszechnie stosowane, gdyż nie miały wielu możliwości, które oferował inny protokół tego typu, PPP (Point-to-Point Protocol). Z tego też względu na nowy protokół przeszła już większość operatorów internetowych obsługujących transmisje modemowe i ISDN. PPP wykorzystuje do przesyłania danych nie tylko specjalną ramkę, ale otwiera również oddzielny kanał komunikacyjny dla obu końców łącza Point-to-Point. Przed nawiązaniem połączenia i podczas jego trwania oba urządzenia końcowe mogą za jego pośrednictwem wykorzystywać następujące możliwości (oraz wymieniać między sobą odpowiednie informacje):

- sprawdzenie haseł przed rozpoczęciem właściwej transmisji danych,
- uzgodnienie opcji szyfrowania,
- podłączenie kilku łączy w celu przyspieszenia transmisji,
- przesyłanie pakietów danych pochodzących z innych sieci niż IP (tunneling),
- przypisywanie dynamicznego numeru IP komputerowi inicjującemu połączenie.

Wprawdzie nie każda aplikacja internetowa potrafi obsłużyć wszystkie dostępne funkcje, ale dzięki tak dużym możliwościom protokół PPP pozwala na rozwiązanie praktycznie każdego problemu, jaki może wystąpić po obu stronach łącza Point-to-Point.

oprac. Marcin Pawlak (mt, bj)

info

Grupa dyskusyjna

Pytania, uwagi i komentarze do artykułu można umieścić na liście dyskusyjnej *news://news.vogel.pl/chip.artykuly*

Internet

Rejestracja domeny w hierarchii pl:
http://www.nask.pl/
Rejestracja domeny w hierarchii COM:
http://rs.internic.net/
IANA (Internet Assigned Numbers Authority):
http://www.iana.org/iana/
DNS Resources Directory:
http://www.dns.net/dnsrd/

CHIP CD 2/99 Na dołączonej do tego numeru CHIP-a płycie CD-ROM w dziale *CHIP-offline* | *Internet* | *Usługa DNS* znajdują się dodatkowe materiały do tekstu