

Notatki z kursu Sieci Komputerowe

Małgorzata Dymek

2018/19, semestr letni

1 Podstawowe pojęcia

Sieć komputerowa - zbiór połączonych komputerów i urządzeń z możliwością komunikacji. Obejmują również wszystko co umożliwia komunikację i współdzielenie zasobów, w szczególności oprogramowanie, protokoły komunikacyjne, media transmisyjne (kable).

Dwie główne kategorie sieci:

- **LAN** - Local Area Networks
- **WAN** - Wide Area Networks

Przybliżone kryterium rozróżnienia: odległości między łączonymi komputerami. Są technologie charakterystyczne dla sieci lokalnych (Gigabit Ethernet), inne dla sieci rozległych (T1, E1 czy Frame Relay). Inny podział wg kryterium odległości:

- **Nanoscale** – obecnie przedmiot badań, zastosowanie w nanomedycynie (IEEE).
- **NFC** (Near Field Communication) – odległości rzędu centymetrów.
- **BAN** (Body Area Network), **WBAN** (Wireless BAN), inaczej to BSN (Body Sensor Network) – łączenie czujników, „wearable devices”.
- **PAN** (Personal Area Network) – od centymetrów do kilku metrów (beziprowodowe: IrDA, Bluetooth, Wireless USB, oraz przewodowe jak USB, FireWire, Thunderbolt).
- **NAN** (Near-me Area Network) – komunikacja między urządzeniami bezprzewodowymi, które są blisko siebie. Ścieżka komunikacyjna między urządzeniami w jednej sieci NAN może jednak być długa i może nawet obejmować technologie sieci WAN, np. jeśli do NAN należą telefony zarejestrowane w różnych firmach telekomunikacyjnych.
- **SAN** (Storage Area Network).
- **CAN** – Campus Area Network, sieci kampusowe.
- **MAN** (Metropolitan Area Network) – obejmuje obszar miasta, dużego kampusu.

Internet jest globalną siecią komputerową złożoną z wielu sieci komputerowych wykorzystujących **protokół IP** (zestaw protokołów TCP/IP). Od strony logicznej Internet można identyfikować z **przestrzenią adresową protokołu IP** (są w użyciu dwie wersje: IPv4 i nowsza IPv6).

Na każdą sieć komputerową składają się sprzętowe oraz programowe elementy składowe.

1.1 Sprzętowe elementy składowe sieci komputerowych

Podstawowymi elementami sprzętowymi są:

- **Nośniki transmisji** (media transmisyjne)
Nośniki transportu sygnałów przesyłanych przez sieć. Są to na przykład kable koncentryczne, tzw. skrętki, kable miedziane, kable światłowodowe, ale też przestrzeń (przesyłanie fal radiowych, mikrofal, światła).

- **Urządzenia dostępu do nośnika**

Są odpowiedzialne za formatowanie danych tak, by nadawały się do przesyłania poprzez nośnik transmisji, umieszczanie tych danych w nośniku transmisji oraz odbieranie odpowiednio zaadresowanych danych (np. karty sieciowe w sieci LAN).

- **Urządzenia wzmacniające, filtrujące i kierujące przesyłane sygnały**, np. przełączniki warstwy drugiej, routery.

Sygnały umieszczane w nośniku transmisji ulegają zakłóceniom.

- **Tłumienie** (osłabienie siły sygnału)

Sposoby unikania tłumienia: ograniczenie długości połączeń (kabli), zainstalowanie urządzenia, które odczytuje przesyłane sygnały, wzmacnia je i wysyła z powrotem do sieci.

- **Zniekształcenie** (niepożądana zmiana kształtu przebiegu czasowego)

Przeciwdziałanie zniekształceniom polega na przestrzeganiu zaleceń dotyczących nośnika (odpowiedni typ nośnika, poprawna instalacja, odpowiednie długości przewodów) oraz korzystaniu z protokołów obsługujących korektę błędów transmisji.

Zadania filtrujące i kierujące sygnały spełniają takie urządzenia jak mosty (pomosty, bridges już raczej nie używane), koncentratory (hubs) przełączniki (switches), punkty dostępowe (access points), routery (routers), bramy (gateways).

Karty sieciowe pakują dane w tzw. ramki. Ramki są podstawowymi porcjami danych przesyłanymi w sieciach komputerowych.

1.2 Programowe elementy składowe sieci komputerowych.

Elementami programowymi sieci są:

- **Protokoły komunikacyjne** (sieciowe)

Zestawy standardów i zasad obowiązujących przy przesyłaniu danych przez sieć. Określają sposoby komunikowania się urządzeń i programów.

- **Oprogramowanie komunikacyjne**

Implementuje protokoły sieciowe. Są to programy umożliwiające użytkownikom korzystanie z sieci komputerowych np. program telnet, przeglądarki WWW, klienci pocztowi, oprogramowanie umożliwiające mapowanie dysków sieciowych itd.

- **Programy poziomu sprzętowego**

Sterują pracą elementów sprzętowych. Sterowniki, programy obsługi, mikroprogramy umożliwiające działanie takich urządzeń, jak karty sieciowe.

1.3 Ramki

Dane przesyłane są w **porcjach zwanych ramkami**. Urządzenie zapewniające dostęp do nośnika przesyła pewne sygnały, które są **interpretowane jako bity**. Od strony logicznej wysłany ciąg bitów zawiera pewne informacje i może być podzielony na porcje zwane **polami**.

Typowa ramka zawiera następujące pola:

- ogranicznik początku ramki (jest to ustalony wzór bitów)
- tzw. adres fizyczny nadawcy (źródła danych)
- adres fizyczny odbiorcy (miejsca docelowego)
- dane
- ogranicznik końca ramki (sekwencja kontrolna ramki).

Ogranicznik początku ramki być poprzedzony lub może zawierać tzw. **preambułę**, która w pewnych technologiach sieciowych jest stosowana do synchronizacji nadajnika i odbiornika. Wielkość pól określana jest w oktetach (8 bitów, uniknięcia niejednoznaczności "bajtu" mogącego mieć więcej bitów).

Kapsułkowanie - wstawienie danych do struktury ramki. Istnieją różne formaty ramek, różne sposoby kapsułkowania i różne sposoby fizycznego adresowania komputerów.

1.4 Topologia sieci lokalnych

Dwa rodzaje topologii:

- Topologie fizyczne
- Topologie logiczne

Jeżeli przy fizycznej topologii gwiazdy komputer przesyła dane bezpośrednio do komputera docelowego (przełącznik), to mamy logiczną topologię gwiazdy. Jeżeli ramka jest wysyłana do wszystkich dostępnych komputerów (koncentrator), to logicznie jest to topologia magistrali.

1.4.1 Komunikacja między komputerami

Założenia:

- Komputer źródłowy - Komputer 1: IP1, MAC1
- Komputer docelowy - Komputer 2: IP2, MAC2

Połączone switchem. Na komputerze docelowym jest serwer strony WWW2.

Jeżeli na komputerze 1 ktoś spróbuje otworzyć WWW2, to:

- Zadziała system DNS: komputer 1 skontaktuje się ze swoim serwerem DNS i zapyta jaki jest adres IP komputera związanego z nazwą domenową WW2. Serwer DNS znajdzie odpowiedni adres w swoich zasobach i odeśle informację do komputera 1.
- Przeglądarka utworzy komunikat (wg protokołu HTTP). Do komunikatu zostanie dodany nagłówek (wg protokołu TCP), który zawiera m.in. port docelowy (standardowo serwery WWW wykorzystują port o numerze 80) oraz port źródłowy (przeglądarka wykorzystuje porty dynamicznie przydzielane, zwykle o „wysokich” numerach). Komunikat razem z dołączonym nagłówkiem TCP nazywa się segmentem TCP.
- Do segmentu TCP zostanie dodany nagłówek I – w ten sposób powstanie pakiet (datagram) IP. Nagłówek IP zawiera m.in. adres IP źródłowy (IP1) i adres P docelowy (IP2).
- Pakiet musi być przesłany w ramce. Do pakietu musi zostać dodany nagłówek ramki, zawierający źródłowy adres MAC (MAC1 – komputer 1 tworzący ramkę zna swój adres MAC) oraz docelowy adres MAC (powinien to być MAC2). **Komputer 1 nie zna adresu MAC komputera 2.** Zna tylko jego adres IP. W IPv4 do poznania adresu docelowego MAC wykorzystywany jest **protokół ARP** – Address Resolution Protocol.
 - Komputer 1 wysyła specjalną ramkę **ARP Request** (zapytanie ARP, ramka ta NIE zawiera w pakiecie IP), która ma adres docelowy składający się z samych jedynek (48 jedynek: ffff-ff-ff-ff). Adres ten nazywa się **adresem rozgłoszeniowym**. Ramka ARP Request jest przesyłana przez przełącznik do wszystkich przyłączonych komputerów. Ramka ta zawiera zapytanie o adres MAC komputera, którego adres IP jest przesyłany w ramce.
 - Każdy komputer przyłączony do przełącznika ma obowiązek odebrać ramkę wysłaną na adres rozgłoszeniowy MAC. Jednak tylko komputer o zadanym IP odpowie na ARP Request.
 - Odpowiedź to specjalna ramka **ARP Reply** (odpowiedź ARP). Odpowiedź ARP jest wysyłana na adres MAC komputera 1.
- Po tym, jak komputer 1 pozna adres MAC komputera 2, może już zbudować ramkę przeznaczoną do komputera 2. Ramka ta zawiera wcześniej zbudowany pakiet IP (który z kolei zawiera segment TCP, który z kolei zawiera komunikat HTTP). Ramka jest wysyłana do przełącznika, a przełącznik dostarcza ją tylko do komputera 2.

- Przełącznik uczy się adresów MAC przyłączonych komputerów i routerów i zapamiętuje w tablicy przypisanie adresu MAC do konkretnego swojego portu. Jeśli przełącznik dostanie ramkę ze znanym mu adresem MAC, to kieruje tę ramkę tylko do odpowiedniego portu, w przeciwnym wypadku wysyła kopię ramki do wszystkich swoich portów (z wyjątkiem tego, na którym dostał ramkę).
- Komputer 2 (jego karta sieciowa) odbiera ramkę, sprawdza adres MAC docelowy i sumę kontrolną, po czym „wyjmuje” z ramki pakiet IP. Sprawdza adres docelowy IP i „wyjmuje” z pakietu segment TCP. Sprawdza do którego portu należy przekazać zawartość (komunikat HTTP) i ostatecznie „wyjmuje” komunikat http z segmentu i przekazuje do portu 80, na którym nasłuchuje serwer WWW.
- Serwer WWW skonstruuje odpowiedź – stronę WWW zapisaną z wykorzystaniem języka HTML). Strona ta zostanie umieszczona w komunikacie http, który następnie musi być przesłany do komputera 1. Mechanizm jest analogiczny jak poprzednio.

W rzeczywistości zanim może zostać przesłany segment TCP, komputery wykorzystujące ten protokół do komunikacji, muszą zbudować tzw. połączenie TCP.

Nagłówek ramki (numery MAC)	Nagłówek IP (numery IP) 20 bajtów	Nagłówek TCP (numery portów) 20 bajtów	Komunikat HTTP	Suma kontrolna 4 bajty
--------------------------------	---	--	----------------	-------------------------------

W przypadku komunikacji między komputerami rozdzielonymi przynajmniej jednym routerem

- Wszystko do skonstruowania pakietu IP włącznie działa tak samo. Komputer tworzący ramkę musi więc wykorzystując ARP Request poznać MAC adres routera, czyli swojej **bramy domyślnej** (default gateway). Stąd elementem konfiguracji komputera jest podanie mu nie tylko jego adresu IP, ale również adresu IP bramy.
- Ramka jest wysyłana do routera.
- Router (brama) po otrzymaniu ramki „wyjmuje” z niej pakiet IP, zagląda do nagłówka i sprawdza jaki jest adres docelowy IP. Na podstawie tego adresu i tablicy routowania wyznacza router następnego skoku i konstruuje i wysyła do niego nową ramkę, w której umieszcza przesyłany pakiet IP. Analogicznie aż pakiet dotrze w kolejnych ramach do docelowej sieci i do docelowego komputera.

1.5 Model ISO/OSI

OSI (Open Systems Interconnection) utworzony przez Międzynarodową Organizację Normalizacyjną (ISO International Organization for Standardization w Genewie) stanowi **model referencyjny** (wzorcowy).

- ułatwienie zrozumienia procesów zachodzących podczas komunikowania się komputerów
- ułatwienie projektowania protokołów komunikacyjnych

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

Każdą z warstw można rozpatrywać w aspekcie dwóch zasadniczych funkcji: odbierania i nadawania.

- **Warstwa fizyczna** - standard połączenia fizycznego, charakterystyki wydajnościowe nośników. Same media transmisyjne pozostają poza dziedziną jej zainteresowania (czasem określane są terminem warstwa zerowa).

- **Warstwa łączy danych** – grupowanie danych wejściowych (z warstwy fizycznej) w bloki zwane **ramkami** danych („jednostki danych usług warstwy fizycznej”), mechanizmy kontroli poprawności transmisji (FCS).
- **Warstwa sieci** - określenie trasy przesyłania danych między komputerami poza lokalnym segmentem sieci LAN, protokoły trasowane takie jak IP (ze stosu protokołów TCP/IP), IPX (Novell IPX/SPX), DDP (AppleTalk) (adresowanie logiczne).
- **Warstwa transportu** - kontrola błędów i przepływu danych poza lokalnymi segmentami LAN, protokoły zapewniające komunikację procesów uruchomionych na odległych komputerach (np. oprogramowanie klient/serwer), w tym komunikację z zapewnieniem niezawodności dostarczania danych. Protokoły tej warstwy to np. TCP oraz UDP (z TCP/IP), SPX (Novell IPX/SPX), ATP, NBP, AEP (AppleTalk).
- **Warstwa sesji** - zarządzanie przebiegiem komunikacji podczas połączenia między komputerami (sesji).
- **Warstwa prezentacji** - kompresja, kodowanie i translacja między niezgodnymi schematami kodowania oraz szyfrowanie.
- **Warstwa aplikacji** - interfejs między aplikacjami a usługami sieci.

Określona sesja komunikacyjna nie musi wykorzystywać protokołów ze wszystkich warstw modelu. Przy nadawaniu dane kierowane są od warstwy 7 do 1, przy odbiorze od 1 do 7.

Używane są też inne modele warstwowe, np. jeden z takich prostszych modeli – model TCP/IP – jest często wykorzystywany przy opisie zestawu (stosu) protokołów TCP/IP (Internet). Nie ma warstwy prezentacji ani warstwy sesji, ponieważ zestaw TCP/IP nie zawiera żadnych protokołów, które można byłoby przypisać do tych warstw. Nieco inny model „dydaktyczny” został zaproponowany przez Tannenbauma.

Nr warstwy OSI	Nazwa warstwy OSI	Nazwa warstwy TCP/IP	Nazwa warstwy Tannenbaum
7	Aplikacji	Aplikacji	Aplikacji
6	Prezentacji		
5	Sesji		
4	Transportu	Transportu	Transportu
3	Sieci	Intersieci	Sieci
2	Łączy danych	Interfejsu sieciowego	Łączy danych
1	Fizyczna		Fizyczna

1.6 Zestaw (stos) protokołów TCP/IP

Składa się z protokołów, które stanowią obecnie podstawę działania Internetu. TCP/IP jest standardem w komunikacji sieciowej.

Warstwa TCP/IP	Warstwa modelu OSI
Aplikacji	Aplikacji Prezentacji Sesji
Transportu	Transportu
Internetowa (sieci, intersieci)	Sieci
Dostępu do sieci (interfejsu sieciowego)	Łączy danych Fizyczna

Protokoły z zestawu TCP/IP:

- TCP - warstwa transportu,
- UDP - warstwa transportu,
- IP - IPv4 i IPv6, warstwa internetowa,
- ARP - tłumaczy adresy między warstwą internetową a warstwą interfejsu sieciowego, czasami zaliczany do tej ostatniej warstwy,

- ICMP - m.in. komunikaty o problemach,
- IGMP - komunikacja grupowa.

Na warstwę aplikacji składają się komponenty programowe sieci, wysyłające i odbierające informacje przez tzw. porty TCP lub UDP (z warstwy transportu). Protokoły warstwy aplikacji to między innymi:

- FTP (File Transfer Protocol),
- TELNET,
- DNS (Domain Name System) związany z usługą DNS (Domain Name Service).

Dane przechodząc w dół stosu protokołów TCP/IP są opakowywane i otrzymują odpowiedni nagłówek. Porcje danych przesyłane w dół stosu mają różne nazwy:

- **Komunikat** - porcja danych utworzona w warstwie aplikacji i przesłana do warstwy transportu.
- **Segment** - porcja danych utworzona przez oprogramowanie implementujące protokół TCP w warstwie transportu. Zawiera w sobie komunikat.
- **Datagram UDP** - porcja danych utworzona przez oprogramowanie implementujące protokół UDP w warstwie transportu.
- **Datagram** - również porcja danych utworzona w warstwie internetowej przez oprogramowanie implementujące protokół IP. Datagram IP zawiera w sobie segment, bywa nazywany pakietem.
- **Ramka** - porcja danych utworzona na poziomie dostępu do sieci.

Sekwencja zdarzeń przy wysłaniu danych:

- Aplikacja przesyła dane do warstwy transportu.
- Dalszy dostęp do sieci realizowany jest przez TCP albo UDP.
 - TCP realizuje tzw. niezawodne połączenia i kontroluje przepływ danych zapewniając niezawodne dostarczenie danych.
 - UDP nie zapewnia niezawodności, ale jest szybszy.
- Segment lub datagram UDP przesyłany jest do warstwy IP, gdzie protokół IP dołącza między innymi informacje o adresach IP źródła i celu tworząc datagram IP (pakiet).
- Datagram z IP przechodzi do warstwy interfejsu sieciowego, gdzie tworzone są ramki. W sieci LAN ramki zawierają adres fizyczny (przypisany do karty sieciowej) otrzymany z protokołu ARP.
- Ramka przekształcana jest w ciąg sygnałów, który zostaje przesłany przez sieć.

RFC (Request for Comments) - miejsce publikowania oficjalnych standardów internetowych. Niektóre dokumenty RFC nie są standardami, ale pełnią funkcję informacyjną. Każdy dokument RFC posiada swój numer – większe numery oznaczają nowsze dokumenty.

1.6.1 ARP (Address Resolution Protocol)

ARP stosowany jest w sieciach Ethernet (jeśli w warstwie sieci wykorzystywany jest protokół IPv4), był też używany w sieciach Token Ring. W wersji IPv6 protokół ARP nie jest w ogóle wykorzystywany, zastępują go inne mechanizmy.

- Zadaniem ARP jest **odnalezienie adresu fizycznego MAC** na podstawie znanego wprost adresu IP (został wpisany przez użytkownika, lub uzyskany automatycznie na podstawie nazwy domenowej (www) dzięki DNS).
- Ze względu na możliwość wymiany karty sieciowej w komputerze o określonym adresie IP ARP musi być **dynamiczny**.
- ARP jest oparty na **metodzie rozgłoszeniowej** i zasadzie **żądania i odpowiedzi**.

- ARP najpierw sprawdza w swojej pamięci podręcznej (**cache**) czy posiada wpis dla danego IP. Jeśli nie, to zostaje wysłana ramka rozgłoszeniowa **ARP Request** Message w fizycznym segmencie sieci, do którego przyłączony jest nadawca (ARP requestor).
 - Jeśli węzeł docelowy znajduje się w tym samym segmencie sieci lokalnej, ARP requestor pyta wprost o to kto ma docelowy adres IP. ARP responder odpowiada wysyłając ramkę ARP Reply pod adres MAC, z którego przyszło żądanie. Po wymianie ramek zarówno nadawca jak i odbiorca mają uaktualnione tablice w pamięci podręcznej (cache).
 - Jeśli węzeł docelowy znajduje się w innym segmencie sieci datagram jest kierowany do domyślnego routera (IP wpisane w konfiguracji TCP/IP lub wykorzystanie Proxy ARP. ARP requestor pyta o adres IP routera domyślnego, router odpowiada wysyłając ramkę ARP respond i podaje swój adres MAC.
- Po otrzymaniu ARP Request uaktualniane są również pamięci podręczne cache ARP w komputerach, które miały w pamięci podręcznej IP ARP requestor. Otrzymanie ramki ARP request jest zatem metodą aktualizacji wpisów w pamięci podręcznej ARP.
- Wpisy w pamięci podręcznej ARP są usuwane po okresie nieużywania rzędu kilku minut. W przypadku użycia wpisu czas ten może wzrosnąć, z pewnym limitem górnym.
- TCP/IP w Microsoft Windows pozwala na użycie statycznych wpisów w pamięci podręcznej ARP. Jednak są one przechowywane w RAM, więc wyłączeniu komputera przepadają.
- Struktura ramki ARP:
 - Typ sprzętu (2 oktety)
 - Typ protokołu (2 oktety)
 - Długość adresu sprzętu (1 oktet)
 - Długość adresu protokołu (1 oktet)
 - Kod operacji (2 oktety)
 - Adres sprzętu nadawcy (dla Ethernet 6 oktetów)
 - Adres protokołu nadawcy (dla IPv4 4 oktety)

1.6.2 Wykrywanie zduplikowanych adresów IP

Tzw "zbędny ARP"

- Węzeł wysyła ARP Request z zapytaniem o swój własny adres.
 - Jeśli ARP Reply nie nadejdzie to znaczy, że w lokalnym segmencie nie ma konfliktu adresów.
 - Jeśli odpowiedź nadejdzie, oznacza to konflikt.
- Węzeł już skonfigurowany traktowany jest jako węzeł z poprawnym adresem (**węzeł zgodny**, defending node), węzeł wysyłający „zbędny ARP” jest **węzłem konfliktowym** (offending node).
- **Węzeł konfliktowy wprowadza błąd** w pamięci podręcznej ARP komputerów w **całym segmencie** sieci. ARP Reply z węzła zgodnego nie naprawia sytuacji (ramka ARP Reply nie jest ramką rozgłoszeniową), więc zgodny wysyła ARP Request ze swoim adresem po wykryciu konfliktu.

Datagramy IP wysłane na w ramach z niepoprawnym adresem MAC odbiorcy przepadają. Protokół IP nie zapewnia niezawodnej dostawy datagramów i nie spowoduje powtórznego przesłania datagramu w nowej ramce. Za niezawodność odpowiedzialne są protokoły warstwy transportu.

1.6.3 Proxy ARP

Router ze skonfigurowanym mechanizmem Proxy ARP odpowiada na ramki ARP Request w imieniu wszystkich węzłów – komputerów spoza segmentu sieci lokalnej. Może być używany jest np. w sytuacji, gdy komputery w sieci nie mają ustawionego domyślnego routera (domyślna brama, default gateway).

Routery mogą mieć włączoną standardowo opcję Proxy ARP, wówczas jeśli jakiś komputer wyśle ARP Request z adresem spoza danej sieci lokalnej (zwykle to nie następuje), to router odpowie „w imieniu” komputera zewnętrznego.

2 Adresacja IPv4

Adres IP jest przypisywany do karty sieciowej, nie do komputera.

Są trzy typy adresów IPv4:

- **Adresy jednostkowe** (unicast) – pojedynczy interfejs sieciowy (komunikacja one-to-one).
- **Adresy rozgłoszeniowe** (broadcast) – wszystkie węzły w tym samym segmencie sieci (one-to-everyone).
- **Adresy grupowe** (multicast) – jeden lub wiele komputerów w jednej lub w różnych segmentach sieci (one-to-many).

W **adresie IP** zapisanym binarnie można wyróżnić **dwie części**:

- **Identyfikator sieci** (Network ID) - pewna liczba bitów z lewej strony adresu
- **Identyfikator hosta** (Host ID) - pozostałe bity.

Terminem **host** określa się komputer, który jest końcowym konsumentem usług sieciowych.

Granica między identyfikatorem sieci a identyfikatorem hosta może być wyznaczona przez tzw. **maskę sieci**.

Identyfikator sieci

- **nie** może się składać z **samych jedynek**.
- **nie** może się składać z **samych zer**.
- **nie może się powtarzać** w złożonej sieci.
- W **pierwszym oktecie** adresu **nie** może się znaleźć wartość **127** (jest ona zarezerwowana dla adresu tzw. pętli zwrotnej).

Identyfikator hosta

- **nie** może się składać z **samych jedynek**.
- **nie** może się składać z **samych zer**.
- musi być **unikalny** w segmencie sieci lokalnej.

Adres IP, który zawiera **same zera** w części hosta jest traktowany jako **adres sieci**.

Adres ograniczonego rozgłoszenia - **255.255.255.255 = 11111111 11111111 11111111 11111111**

- adres rozgłoszenia w danym segmencie sieci ograniczonym routerami.

Adresy rozgłoszenia do sieci lub podsieci mają jedynki tylko w części hosta.

Adresy nieunikalne, powtarzalne - przykłady:

- adresy rozpoczynające się od liczby 127, które oznaczają zawsze komputer lokalny (zwykle 127.0.0.1).
- adresy tzw. transmisji grupowej.
- grupy tzw. adresów prywatnych.

2.1 Adresowanie oparte na klasach

Pierwszy bajt adresu determinuje do jakiej klasy należy sieć.

Klasa	Adres sieci	Adresy	Zakres 1-go bajtu	Najstarsze bity
A	w.0.0.0	1.0.0.0 - 126.0.0.0	1 – 126	0
B	w.x.0.0	128.0.0.0 - 191.255.0.0	128 – 191	10
C	w.x.y.0	192.0.0.0 - 223.255.255.0	192 – 223	110
D	nie dotyczy	nie dotyczy	224 – 239	1110
E	nie dotyczy	nie dotyczy	240 – 255	11110

Klasa	Ilość sieci	Komp. w sieci	ID sieci	ID hosta	"pierwszy"	"ostatni"
A	126	$2^{24} - 2$	1 bajt	3 bajty	w.0.0.1	w.255.255.254
B	$(191 - 128 + 1) * 256$	$2^{16} - 2 = 65534$	2 bajty	2 bajty	w.x.0.1	w.x.255.254
C	$(192 - 223 + 1) * 256 * 256$	$2^8 - 2 = 254$	3 bajty	1 bajt	w.x.z.1	w.x.z.254

- **Adresy klasy D** - przeznaczone są do transmisji grupowych.
- **Adresy klasy E** - zarezerwowane (nie wykorzystywane normalnie do transmisji pakietów).
- **Adresy pętli zwrotnej** (loopback) - postaci 127.x.y.z (na ogół 127.0.0.1). Cały ruch przesyłany na ten adres nie wychodzi z komputera.

Identyfikator sieci można określić na podstawie adresu IP oraz tzw. **maski sieci**. Jest to **liczba binarna 32 bitowa**, zapisywana podobnie jak adres IP, jednak **maska zawsze z lewej strony ma jedynek, natomiast z prawej ma zera**.

Przykłady masek:

255.0.0.0 = 11111111.00000000.00000000.00000000 = /8,

255.255.0.0 = 11111111.11111111.00000000.00000000 = /16.

Adres sieci = AdresIP & Maska

Dzielenie sieci na podsieci

Dzielenie sieci na fragmenty nazywane **podsieciami** w celu zwiększenia efektywności działania.

- zmniejszenie ruchu w segmentach,
- zmniejszenie tzw. dziedziny rozgłaszania (obszary przekazywania ramek rozgłoszeniowych, tj. z adresem MAC ff-ff-ff-ff-ff-ff)
- zwiększenie bezpieczeństwa. Routery mogą działać jako filtry pakietów, które przepuszczają między sobą tylko pakiety spełniające określone kryteria.

Adres podsieci można określić przez użycie niestandardowych masek sieci (**masek podsieci**).

Dzielenie klasy A - maski 255.255.0.0, 255.255.255.0, dla B - 255.255.255.0.

Routery przechowują w tablicach routowania informacje o znanych im trasach do pewnych sieci. Administrator w trakcie konfiguracji routera wprowadza informacje o sieciach bezpośrednio przyłączonych do routera. Potem, po włączeniu opcji dynamicznego routowania routery wymieniają się informacjami o sieciach wg protokołów routowania i na tej podstawie budują sobie pewien obraz sieci i potrafią wyznaczać trasy datagramów.

Przykład: komputer A: 162.168.1.100, komputer B: 162.168.2.101

Maska 255.255.0.0 - komputery są względem siebie lokalne.

Maska 255.255.255.0 - komputery są względem siebie odległe (przedzielone routerem).

2.2 Adresowanie bezklasowe

Pytanie: czy nie można podzielić sieci na podsieci z **użyciem dowolnej liczby jedynek**? Na przykład, jeśli w sieci klasy B o adresie 149.159.0.0 /16 zastosowalibyśmy maskę podsieci nie 24 bitową (co daje 254 lub 256 podsieci z możliwością adresowania do 254 komputerów) tylko 22 bitową. Otrzymalibyśmy 62

lub 64 sieci, z których każda mogłaby mieć 1022 komputery. Do określenia sieci należy podać adres sieci oraz maskę. Obecnie w Internecie powszechnie jest wykorzystywane adresowanie bezklasowe. Przykład adresowania bezklasowego: 145.217.123.7 /20 (maska: 255.255.240.0)

3 Routing

Routowanie - proces przesyłania pakietów (datagramów IP) od hosta nadawczego do odbiorcy na ogół z wykorzystaniem routerów pośredniczących. Każdy host oraz router podejmuje decyzję jak przesłać datagram, podejmowaną na podstawie tzw. tabel routowania oraz pewnych reguł. **Routing statyczny** - gdy w routerze tabela routowania jest wypełniona wpisami statycznymi i nie zmienia się.

Routing dynamiczny - tabela routowania zmienia się dynamicznie na podstawie protokołów routowania.

Protokoły routowalne: **IPv4**, IPX firmy Novell (należący do stosu IPX/SPX), AppleTalk, IPv6. Routery mogą realizować trasowanie dla pakietów z protokołów innych niż IPv4.

Dla routerów tablica routowania na ogół jest modyfikowana dynamicznie na podstawie **protokołów routowania**. Określają one w jaki sposób routery mają wymieniać między sobą informacje na temat połączeń między routerami w sieci i jak na podstawie tych informacji mają aktualizować swoje tablice routowania.

Protokoły routowania: RIP, RIP2, OSPF, IGRP, EIGRP, BGP. **Brama domyślna** to router, do którego kierowany jest datagram, jeśli nie została znaleziona dla niego lepsza trasa w tablicy routowania.

Typy bezpośrednich połączeń:

- rozgłoszenia - np. Ethernet; działa protokół ARP,
- punkt-punkt - np. analogowa linia telefoniczna; technologie sieci rozległych WAN, ARP nie jest wykorzystywany;
- wielodostęp nierozgłoszeniowy (NBMA) - technologie przełączania pakietów WAN Frame Relay, ATM; ARP nie jest wykorzystywany.

3.0.1 Tablica routowania IP

Zawiera wpisy dotyczące tras do hostów, routerów i sieci. Zwykle w hostach liczba wpisów jest dużo mniejsza niż w routerach i zawiera informacje o bramie domyślnej.

W każdym hostcie i routerze dla każdego przesyłanego pakietu IP na podstawie tablicy routowania wyznaczane są dwie wartości:

- interfejs – reprezentacja fizycznego urządzenia, przez które ma być wysłany pakiet,
- adres następnego skoku - adres IP następnego routera, do którego ma być skierowany datagram lub adres docelowy hosta, jeśli jest bezpośrednio dołączony do nadawcy

Tablica routowania zawiera następujące pola:

- **Przeznaczenie (Destination)**
W koniunkcji (logiczne AND) z polem Maska zawiera informację o zakresach adresów IP, które są dostępne przy użyciu tej trasy. Pole to może zawierać ID sieci lub adres IP konkretnego hosta lub routera. W prawidłowym wpisie nie może zawierać jedynek w miejscu, gdzie w Masce sieci są zera.
- **Maska sieci**
Zawiera maskę sieci, może też zawierać same jedynek (255.255.255.255).
- **Adres następnego skoku**
Adres IP, do którego datagram będzie przesłany w następnym kroku, jeśli zostanie wybrana ta trasa. Bez znaczenia dla połączeń punkt-punkt.

- **Interfejs**

Oznaczenie interfejsu, przez który datagram będzie przesłany do miejsca określonego przez adres następnego skoku.

- **Metryka**

Liczba wskazująca na koszt trasy. Im wyższa wartość, tym „gorsza” trasa.

- **Odległość administracyjna**

Liczba przypisywana na podstawie tego, w jaki sposób router poznał trasę do danego miejsca docelowego. „Wygrywa” trasa z najmniejszą liczbą. Trasa domyślna może być oznaczana przez 0.0.0.0 w polu Przeznaczenie i 0.0.0.0 w polu Maska sieci.

Proces określania trasy na podstawie tablicy routowania:

- Dla każdej trasy w tablicy routowania określa się, czy jest ona **zgodna z adresem IP** w przesyłanym pakiecie. Trasa domyślna jest traktowana zawsze jako zgodna.
- Spośród tras zgodnych wybierana jest ta (lub kilka), dla której w polu **Maska sieci** jest **największa liczba jedynek**. Może się zdarzyć, że jedyną trasą zgodną jest trasa domyślna.
- Spośród tras, które zostały wybrane w punkcie 2 wybierane są trasy o **najmniejszej metryce**.
- Spośród tras wybranych w punkcie 3 wybierana jest dowolna trasa.

Routery potrafią również wykonywać równoważenie obciążeń.

3.0.2 Komunikaty ICMP o przekierowaniu

Komunikaty ICMP o przekierowaniu pozwalają hostom TCP/IP na konfigurację tylko jednego routera – bramy domyślnej nawet w sytuacji, gdy w sieci lokalnej są dwa lub więcej routerów, które są odpowiedzialne za pewne miejsca docelowe. Hosty mogą zacząć pracę z jedną domyślną trasą i uczyć się topologii sieci (w szczególności informacji o routowaniu) poprzez otrzymywanie komunikatów ICMP.

Komunikaty ICMP o przekierowaniu powinny być generowane przez routery, ale korzystać z nich mogą tylko hosty. Jeśli datagram IP zostanie celowo usuwany przez router, to może być wysłany odpowiedni komunikat protokołu ICMP do nadawcy.

4 Protokół IPv4

- Protokół warstwy trzeciej modelu ISO OSI.
- Oprogramowanie implementujące protokół IP jest odpowiedzialne za:
 - adresowanie IP,
 - tworzenie datagramów IP (pakietów)
 - uczestniczenie w kierowaniu ich w sieci z punktu początkowego do punktu docelowego.
- realizuje usługę zawodną. Jeśli komunikacja powinna zawierać mechanizmy niezawodności, to muszą one być dostarczone przez protokoły warstwy wyższej.
- Datagram IP składa się z nagłówka (header) i bloku danych (payload).
 - **Nagłówek** dzięki informacjom w nim zawartym umożliwia obsługę routowania, identyfikację bloku danych, określenie rozmiaru nagłówka i datagramu oraz obsługę fragmentacji. W nagłówku mogą się znaleźć również tzw. opcje rozszerzające. Ma zmienną długość (20 do 60 bajtów, co 4 bajty).
 - **Blok danych** może mieć długość do 65515 bajtów.

4.1 Nagłówek IPv4

- **Wersja** (4 bity) (=0100)
- **Długość nagłówka IP** (IHL – Internet Header Length) (4 bity)
Najczęściej nagłówek ma 20 bajtów, a więc 5 bloków (0101 binarnie). Maksymalnie długość nagłówka może wynosić 60 bajtów.
- **Typ usługi:** TOS (Type of Service) lub DS. (Differentiated Services) (8 bitów)
Zawiera dodatkowe informacje, które mogą być użyte w routingu. Pierwotnie pole TOS było zdefiniowane następująco:
 - Bits 0-2: Precedence.
 - Bit 3: 0 = Normal Delay, 1 = Low Delay.
 - Bits 4: 0 = Normal Throughput, 1 = High Throughput.
 - Bits 5: 0 = Normal Reliability, 1 = High Reliability.
 - Bit 6-7: Reserved for Future Use (0).

TOS było ustawiane przez hosta nadającego i nie było modyfikowane przez routery, miały być używane do obsługi QoS (Quality of Service). W rzeczywistości jego wykorzystanie było problematyczne.

Zmieniono nazwę pola na DS (Differentiated Services) i sześć najstarszych bitów nazwano DSCP. Następnie pozostałe dwa bity przeznaczono na ECN.

- Bits 0-5: DSCP
- Bits 6-7: ECN

ECN jest rozszerzeniem protokołów IP oraz TCP. Umożliwia powiadamianie punktów końcowych IP/TCP o nadchodzącym zatorze bez usuwania pakietów, poprzez ustawienie wartości 11 na bitach ECN. Jest opcjonalny.

Standardowo (bez obsługi ECN) zator w sieci TCP/IP przejawia się usuwaniem pakietów. Dopuszczalne wartości na bitach ECN:

- 00 – Non ECN-Capable Transport, Non-ECT
- 10 – ECN Capable Transport, ECT(0)
- 01 – ECN Capable Transport, ECT(1)
- 11 – Congestion Encountered, CE.

Odbiorca pakietu przesyła informację do źródła, wykorzystując odpowiednie flagi nagłówka TCP (ze względu na to, że zatorowi może przeciwdziałać TCP, nie IP). Pierwotne źródło danych redukuje prędkość transmisji zmniejszając rozmiar okna przeciążeniowego. Protokół TCP wspiera ECN przez wykorzystanie specjalnych trzech flag w nagłówku: TCP: Nonce Sum (NS), ECN-Echo (ECE) oraz Congestion Windows Reduced (CWR).

- **Długość całkowita** (16 bitów)
Na podstawie tego pola oraz pola Długość nagłówka można określić wielkość bloku danych oraz początek tego bloku. Całkowita długość podawana jest w bajtach, maksymalna możliwa długość może wynosić 65535.
- **Identyfikator** (16 bitów)
Identyfikator kolejnych datagramów. Wartość jest wpisywana przez host nadający i dla kolejnych datagramów jest zwiększana.
- **Flagi** (3 bity)
3 bity tworzące dwie flagi używane przy fragmentacji datagramów.
- **Przesunięcie fragmentu** (13 bitów)
Używane przy fragmentacji datagramów.

- **Czas życia (TTL) (8 bitów)**
Określa przez ile łączy może przejść (skoków) datagram zanim zostanie odrzucony przez router. Host docelowy nie sprawdza TTL. Jeśli pakiet jest odrzucany to wysyłany jest komunikat ICMP „Time Expired – TTL Expired”.
- **Protokół (8 bitów)**
Określa do jakiego protokołu warstwy wyższej należy przekazać datagram. Przykładowe wartości to 1 – ICMP, 6 – TCP, 17 – UDP.
- **Suma kontrolna nagłówka (16 bitów)**
Liczona jest tylko dla nagłówka. Jest on dzielony na słowa 16-to bitowe. Są one dodawane a wynik negowany. Wynik umieszczany jest w polu sumy kontrolnej. W miejscu docelowym suma kontrolna jest ponownie obliczana. Ponieważ nagłówek w miejscu docelowym zawiera sumę kontrolną, to ponownie wyliczona suma powinna składać się z samych jedynek. Jeśli jest inna to oprogramowanie IP odrzuca odebrany pakiet (brak komunikatu o błędzie). Po przejściu przez router jest modyfikowane pole TTL, zatem suma kontrolna powinna ulec zmianie.
- **Adres IP źródła (32 bity)**
- **Adres IP docelowy (32 bity)**
- **Dodatkowe opcje i wypełnienie (32 bity + ew. więcej).**
Opcje mogą zająć maksymalnie 40 bajtów i mogą zawierać m.in.:
 - zapis trasy (RR - Record Route)
 - * kod - typ opcji; 1 bajt, RR=7
 - * length - liczba bajtów opcji; 1 bajt, max=39
 - * ptr – numer bajta wolnego miejsca na wpisanie kolejnego adresu IP; 1 bajt, na początku = 4
 - * adres IP 1 (4 bajty)
 - * ...
 - * adres IP 9 (4 bajty)
 - zapis czasu – timestamp
 - * kod - typ opcji; 1 bajt, timestamp = 0x44
 - * length - liczba bajtów opcji; 1 bajt, zwykle 36 lub 40
 - * ptr - numer bajtu wolnego miejsca na kolejny wpis; 1 bajt, na początku = 5
 - * OF - flaga przepełnienia. Jeśli router nie może dopisać swojego czasu, bo nie ma już miejsca, to powiększa OF o jeden; 4 bity, na początku = 0
 - * FL - znacznik: 0 – zapisuj tylko czasy, 1 – zapisuj adres IP i czas, 2 – wysyłający wpisuje adresy IP, router o danym IP wpisuje czas
 - * timestamp 1 (4 bajty)
 - * ...
 - * timestamp 9 (4 bajty)
 - routowanie źródłowe
Normalnie to routery wybierają dynamicznie trasę datagramów. Można jednak określić trasę datagramu w opcjach nagłówka IP.
 - * dokładne – wysyłający komputer określa dokładną trasę, jaką musi przejść datagram. Jeśli kolejne routery na tej trasie są przedzielone jakimś innym routerem, to wysyła komunikat ICMP „source route failed” i datagram jest odrzucany.

- * swobodne – wysyłający określa listę adresów IP, przez jakie musi przejść datagram, ale datagram może przechodzić również przez inne routery.

Pole opcji zawsze zajmuje wielokrotność 4 bajtów, stąd czasem jest uzupełniane zerami.

Za nagłówkiem IP w datagramie znajdują się dane (segment TCP, datagram UDP, komunikat ICMP).

4.2 Fragmentacja datagramów IPv4

MTU (Maximum Transmission Unit) to największa porcja danych, jaka może być przesłana w ramce przez pewną sieć (sieci) przy wykorzystaniu konkretnej technologii. Jeśli datagram IP jest większy niż wynika to z MTU dla warstwy łącza, to IP dokonuje fragmentacji. Najmniejsze MTU po drodze przejścia datagramu nazywa się **ścieżką MTU**. Jeśli nastąpiła fragmentacja, to w miejscu docelowym oprogramowanie warstwy IP składa fragmenty z powrotem w pakiety oryginalnej wielkości. Fragmenty też mogą być dalej dzielone, stają się samodzielnymi pakietami.

Pole identyfikator w nagłówku IP zawiera numer wysłanego pakietu. Pole powinno być inicjowane przez protokół warstwy wyższej. Warstwa IP zwiększa identyfikator o 1 dla kolejnych pakietów.

Pole flagi (3 bity) :

- Bit 0: zarezerwowane, musi być zero
- Bit 1: (DF) 0 = May Fragment, 1 = Don't Fragment.
- Bit 2: (MF) 0 = Last Fragment, 1 = More Fragments.

Jeśli bit 1 jest ustawiony, to znaczy, że pakiet nie może być dzielony. W przypadku konieczności dzielenia jest odrzucany i do nadawcy wysyłany jest komunikat ICMP (typ 3 z polem kod = 4). Pole przesunięcia fragmentu zawiera informację o przesunięciu fragmentu względem początku oryginalnego pakietu. Wyrażane jest w blokach ośmiobajtowych. Jeśli zgubiony zostanie chociaż jeden fragment, wówczas cały wyjściowy pakiet jest odrzucony, więc fragmentacja jest niekorzystna. Do tego może ona bardzo obciążać routery.

4.3 ICMP (Internet Control Message Protocol)

- raportowanie routingu,
- dostarczanie informacji o błędach podczas przesyłania ze źródła do komputera docelowego,
- dostarczanie funkcji sprawdzających możliwość komunikacji komputerów wykorzystaniem protokołu IP,
- pomoc w automatycznej konfiguracji hostów.

Komunikaty ICMP wysyłane są w pakietach IP. W efekcie w ramce znajduje się nagłówek IP, nagłówek ICMP oraz dane komunikatu ICMP.

Struktura komunikatu ICMP

- Typ (1 oktet)
- Kod (1 oktet)
- Suma kontrolna (2 oktety)
- Dane charakterystyczne dla typu (różna długość)

Typy komunikatów ICMP	0	Odpowiedź echa (echo reply)
	3	Miejsce docelowe nieosiągalne (destination unreachable)
	4	Tłumienie źródła (source quench)
	5	Przekierowanie (redirect)
	8	Żądanie echa (echo request)
	9	Ogłoszenie routera (router advertisement)
	10	Wybór routera (router selection)
	11	Przekroczenie czasu (time exceeded)
	12	Problem parametru (parameter problem)

Żądanie i odpowiedź echa Cel – wysłanie prostego komunikatu do węzła IP i odebranie echa tego komunikatu. Bardzo użyteczne przy usuwaniu problemów i naprawianiu sieci. Narzędzia takie jak ping oraz tracert i traceroute używają tych komunikatów ICMP do uzyskania informacji o dostępności węzła docelowego.

Żądanie echa:

- Typ = 8
- Kod = 0
- Suma kontrolna (2 oktety)
- Identyfikator (2 oktety)
- Numer sekwencji (2 oktety)
- Opcjonalne dane (różna długość)

Odpowiedź echa:

- Typ = 0
- Kod = 0
- Suma kontrolna (2 oktety)
- Identyfikator, Numer sekwencji, Opcjonalne dane przepisane z Echo request.

5 Routing dynamiczny

Sposób obsługi routowania przez warstwę IP to **mechanizm routowania** - przeglądanie przez tablicy routowania, podejmowanie decyzji co do przesyłania datagramów IP. Przez pojęcie **polityka routowania** określa się działania procesu routowania podejmowane w celu ustanowienia i bieżącej modyfikacji tablicy routowania, jest ona realizowana z wykorzystaniem protokołów routowania.

Pożądane cechy protokołów routowania to:

- Wyznaczenie **najlepszej trasy** do punktu docelowego, wymaga określenia kryterium porównywania tras.
- **Odporność** (robustness) - protokoły muszą zawsze działać poprawnie.
- **Szybkie osiągnięcie zbieżności** (rapid convergence), czyli stanu, w którym wszystkie routery „widzą” jednakowo topologię sieci. Szybkość określa czas rozpowszechnienia informacji o zmianach.
- **Dopasowanie do zmian** (flexibility), wyznaczanie nowych optymalnych tras.

Internet jest zorganizowany jako grupa tzw. systemów autonomicznych (Autonomous System – AS), z których każdy jest osobno administrowany. W każdym są wewnętrzne protokoły routowania (ang. IGP – interdomain gateway protocol lub IRP – interdomain routing protocol).

Ze względu na sposób działania protokoły routowania wewnętrznego dzielimy na:

- protokoły **wektora odległości** (DV, distance-vector): RIP, RIP2, IGRP, EIGRP
- protokoły **stanu łącza** (link-state): OSPF, OSPF2, IS-IS, NLSP

Uwaga: w starszych opracowaniach firmy Cisco protokół EIGRP był określany jako protokół hybrydowy.

Zewnętrzne protokoły routowania (EGP - exterior gateway protocols) - międzydomenowe protokoły routowania używane między routerami działającymi w różnych systemach autonomicznych. Najważniejszym protokołem zewnętrznym jest BGP. Oprócz protokołów routowania rozważa się protokoły routowalne, takie jak IP, IPX, Apple Talk. Określają sposoby adresowania, umożliwiające dostarczanie pakietów w złożonej sieci komputerowej. Routery mogą wyznaczać trasy dla różnych protokołów routowalnych, nie tylko IP. Każdy z protokołów routowania i każdy z protokołów routowalnych musi być w routerze skonfigurowany.

5.1 Protokoły routowania wektora odległości

Distance vector – wektor odległości. Protokoły wektora odległości są oparte na **algorytmie Bellmana Forda** obliczania najkrótszych ścieżek w grafie.

Węzły grafu oznaczają routery, krawędzie odpowiadają połączeniom między routerami. Połączenia te mają różne koszty, co odpowiada różnym wagom krawędzi grafu. W protokołach routowania wagi nie mogą być ujemne, co oznacza, że można też wykorzystać np. algorytm Dijkstry (jest wykorzystywany w protokołach stanu łącza).

Nie zawsze router jest w stanie wykryć uszkodzenie łącza. Uszkodzeniu może ulec też np. sąsiedni router albo ramka zawierająca pakiet z wektorem odległości i wektor odległości nie dotrze do docelowego routera. Dlatego w protokołach typu wektor odległości trasa jest zaznaczana jako niedostępna, gdy router nie dostanie o niej informacji od sąsiada przez kilka kolejnych rozgłoszeń (np. w RIP 180 sekund, czyli sześć rozgłoszeń). Trasa niedostępna nie jest jeszcze usuwana z tablicy routowania przez kilka rozgłoszeń, np. w RIP przez 90 sekund. Powoduje to większe opóźnienie w czasie uzyskania zbieżności.

5.2 Niekorzystne zjawiska związane z routowaniem wg protokołów wektora odległości

- Pętle routowania.
- Efekt odbijania.
- Zliczanie do nieskończoności.

Taktyki rozwiązania:

- **Dzielony horyzont (split horizon)**
 - Do łącza nie zostanie przekazana informacja o trasach wiodących przez to łącze.
 - Dzielony horyzont nie zawsze, ale zazwyczaj likwiduje **pętle routowania**.
- **Natychmiastowe/wymuszane aktualizacje (triggered updates)**
 - W przypadku zmiany metryki trasy **musi nastąpić rozgłoszenie bez względu na okres rozgłoszeń** charakterystyczny dla danego protokołu.
 - Powoduje to **szybszą zbieżność** i częściowo zapobiega **pętlom routowania**.
- **Zegary hold-down (hold-down timers)**
 - Router po otrzymaniu od sąsiada informacji o dezaktualizacji trasy włącza **specjalny zegar** (hold-down timer).
 - Jeśli przed upływem czasu progowego nastąpi:

- * **aktualizacja od tego samego sąsiada** na trasę aktywną, to **trasa** jest zaznaczana jako **aktywna**.
- * router dostanie **od innego routera informację** o trasie do rozważanego miejsca docelowego z **metryką mniejszą bądź równą** tej zdezaktualizowanej, wówczas następuje **wpis zgłoszonej trasy**.
- * router dostanie **od innego routera informację** o trasie do rozważanego miejsca docelowego z **metryką większą** od tej zdezaktualizowanej, taka trasa **nie jest brana pod uwagę**.

5.3 Protokół RIP

- Protokół typu wektor odległości.
- Metryką w RIP jest liczba skoków (hops) do celu. Można skonfigurować inne wartości dla połączeń, więc np. preferencję tras szybszych.
- Metryka 16 oznacza umownie nieskończoność (miejsce niedostępne), zatem RIP nie jest dobrym protokołem w przypadku dużych sieci.
- Można ustawić trasę domyślną (adres 0.0.0.0).
- W wersji RIP na routerach Cisco można przechowywać więcej niż jedną trasę o takiej samej metryce. Można włączyć równoważenie obciążeń (load balancing) na dwa sposoby:
 - **Process switching** (packet-by-packet load balancing), kosztowny i dlatego nie polecany, każdy pakiet jest kierowany osobno, dla każdego jest przeglądana tablica routowania.
 - **Fast switching** (destination-by-destination), tylko dla pierwszego pakietu z pewnego miejsca źródłowego do pewnego miejsca docelowego przeszukiwana jest tablica routowania. Wyznaczona trasa jest zapisywana w pamięci podręcznej (cache) i kolejne pakiety wędrują tą samą ścieżką.
- Wykorzystywane są następujące **mechanizmy** typowe dla protokołów wektor odległości:
 - Split horizon (+ with poison reverse)
 - Holddown counters (timers)
 - Triggered updates
- **Cztery zegary, liczniki** (timers, counters):
 - **Update timer** (standardowo 30 sekund) – po przesłaniu wektora odległości (routing update) zegar jest zerowany. Po osiągnięciu 30 s. wysyłany jest następny wektor.
 - **Invalid timer** (standardowo 180 sekund) – za każdym razem jak router dostaje uaktualnienie pewnej trasy zegar ten dla trasy jest zerowany. Po osiągnięciu wartości progowej trasa jest zaznaczana jako niepoprawna, ale pakiety jeszcze są kierowane tą trasą.
 - **Hold-down timer** (standardowo 180s.) – po przekroczeniu wartości progowej przez invalid timer trasa jest ustawiana w stan hold-down. Trasa jest ustawiana w stan holddown również gdy router dostanie informację o tym, że sieć jest nieosiągalna (i nie ma innej, osiągalnej trasy).
 - **Flush-timer** (standardowo 240s.) – zegar dla trasy jest zerowany po otrzymaniu informacji o trasie. Po osiągnięciu czasu progowego trasa jest usuwana nawet, jeśli trasa jest jeszcze w stanie hold-down.
- Przewidziano możliwość **odpytywania routera o cały wektor odległości** lub o **trasy** do pewnych miejsc docelowych – takie możliwości są wykorzystywane przez starcie oraz na ogół w celach diagnostycznych.

- **Zalety RIP**

- prostota - procesor nie jest nadmiernie obciążony aktualizacją tablicy routowania i innymi działaniami,
- łatwość konfiguracji.

- **Wady RIP**

- wolne rozprzestrzenianie się informacji o zmianach w topologii sieci (wolna zbieżność),
- stosunkowo częste (co 30s.) przesyłanie dużych porcji informacji w komunikatach RIP, co obciąża sieć.
- Wadą RIPv1 jest to, że nie daje możliwości przesyłania masek. RIP w wersji 1 jest protokołem routowania klasowego. Przyjmowana jest maska taka, jaka jest ustawiona na interfejsie, do którego dotarł wektor odległości z tą informacją. Na granicach klas jednak routery RIP wykonują automatyczną sumaryzację, co przy niepoprawnych konfiguracjach może prowadzić do błędów w routowaniu.

- RIPv2 przekazuje maski podsieci, można stosować sieci bezklasowe i podsieci o zmiennym rozmiarze. Umożliwia prostą autentykację (przez hasła). Przekazuje adresy następnego skoku w komunikatach. Część wad RIP pozostała: 16 jako metryka oznaczająca nieskończoność, brak alternatywnych tras.

5.4 Wiele protokołów routowania. Odległości administracyjne.

W środowisku współdziałania wielu protokołów routowania dla tras, w zależności od tego jakie jest ich źródło, ustawiana jest tzw. **odległość administracyjna**. Odległość jest używana tylko **wewnątrz** przez router, wybierane są trasy o **najmniejszej odległości**.

Standardowe odległości dla tras:

Źródło trasy	Odległość administracyjna
Connected interface	0
Static route	1
Summary EIGRP	5
External BGP	20
Internal EIGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
EGP	140
Internal BGP	200
Unknown	255

Na routerach można również włączyć redystrybucję tras między protokołami routowania według różnych zasad, z odpowiednim przeliczaniem metryk.

6 Protokoły UDP i TCP (warstwa transportowa)

- Informacje o **docelowym i źródłowym procesie** zawarte są w nagłówkach protokołów warstwy transportu.
- Protokoły UDP i TCP umożliwiają dostarczenie danych do procesu działającego na komputerze, przy czym wykorzystywane są tzw. **porty**. Porty są ponumerowane, numer portu jest liczbą dwubajtową.

- Aplikacja rezerwuje na swoje potrzeby pewne zasoby (komunikacyjne bufor w pamięci operacyjnej), ma do dyspozycji tzw. **gniazda**.
- Każde gniazdo jest identyfikowane przez numer IP oraz numer portu. **Interfejs komunikacyjny** umożliwiający komunikowanie się programów z wykorzystaniem TCP/UDP nazywa się **interfejsem gniazd**.

6.1 UDP – User Datagram Protocol

- Prosty protokół bezpołączeniowy warstwy transportu.
- Umożliwia przesyłanie danych między procesami dzięki określeniu adresów IP komputerów oraz 16 bitowych numerów portów.
- Porcja danych zgodna z protokołem UDP nazywana jest datagramem/pakiem UDP.
- Nie zapewnia niezawodności. Ewentualne zapewnienie niezawodności musi być realizowane przez protokoły warstwy aplikacji.
- Niewielki nagłówek (8 bajtów), nie zawiera mechanizmów ustanawiania połączenia ani sterowania przepływem datagramów, zatem jest szybszy od TCP.
- Datagramy UDP mogą być przesyłane w pakietach IP z adresem docelowym przesyłania grupowego.
- Przykłady zastosowań: strumieniowanie audio/wideo, wideokonferencje, transmisje głosu; RIP (port 520).

Aplikacja jest odpowiedzialna za rozmiar wysyłanego datagramu. Jeśli wielkość przekroczy MTU sieci, wówczas datagram IP (zawierający w sobie datagram UDP) jest dzielony (następuje fragmentacja IP).

6.1.1 Porty

Aktualne przyporządkowanie numerów portów można znaleźć na stronie IANA.

- Poprawne numery portów: 1...65535.
- Numer 0 może być używany jako numer portu źródłowego, jeśli nadawca nie oczekuje odpowiedzi.
- IANA dzieli porty na trzy grupy:
 - 1...1023 – porty ogólnie znane, mogą być przypisywane przez procesy systemowe lub działające z uprawnieniami użytkownika uprzywilejowanego.
 - 1024...49151 – zarejestrowane, do użytku przez różne aplikacje.
 - 49152...65535 – dynamiczne lub prywatne, np. dla testowanych aplikacji klient-serwer. Mogą być przydzielane dynamicznie np. dla przeglądarki WWW.

6.1.2 Enkapsulacja datagramu UDP

nagłówek IP 20 bajtów	nagłówek UDP 8 bajtów	dane UDP ...
--------------------------	--------------------------	-----------------

Nagłówek UDP

- Numer portu źródłowego (16 bitów)
- Numer portu docelowego (16 bitów)
- Długość UDP (nagłówek + dane) – wypełniana opcjonalnie (16 bitów)
- Suma kontrolna UDP (16 bitów)
- Dane, jeśli są

Suma kontrolna

- Jedyne mechanizm sprawdzenia poprawności datagramu w UDP.
- Opcjonalna, jeśli datagram UDP jest przesyłany w pakiecie IPv4.
- Obowiązkowa, jeśli datagram UDP jest przesyłany w pakiecie IPv6.
- Dodawane są liczby 16 bitowe, stąd możliwa konieczność dodania do obliczeń dodatkowego bajtu z samymi zerami (jeśli długość datagramu liczona w bajtach jest liczbą nieparzystą).
- Do obliczenia sumy kontrolnej używany jest tzw. pseudonagłówek. Pseudonagłówek nie jest przesyłany.

Pseudonagłówek (12 bajtów) w pakiecie IPv4:

- Adres IP źródła (32 bity)
- Adres IP celu (32 bity)
- 8 zer
- Typ protokołu (8 bitów, UDP = 17)
- Długość UDP (16 bitów)

Algorytm jest taki sam jak dla sumy kontrolnej nagłówka IPv4 (jednak w IPv4 suma kontrolna obejmowała sam nagłówek, w UDP obejmuje również przesyłane za nagłówkiem dane).

6.2 TCP – Transmission Control Protocol

Oprogramowanie TCP tworzy połączenia między dwoma procesami z jednoczesną dwukierunkową transmisją. Punkty końcowe są identyfikowane przez parę: numer IP, numer portu. Połączenie identyfikowane jest przez cztery liczby: dwa numery IP oraz dwa numery portów.

Pomiędzy procesami przesyłane są dwa strumienie 8-bitowych oktetów, po jednym w każdą stronę (**strumień oktetów**). Bajty wysyłane są w segmentach, ale proces docelowy nie jest w stanie z góry określić wielkości nadchodzących porcji. Ilość bajtów danych przesyłanych w jednym segmencie nie powinna być większa niż ustalony MSS (**Maximum Segment Size**).

Cechy TCP

- Partnerzy (procesy) tworzą połączenie z wykorzystaniem mechanizmu (trójfazowego) uzgodnienia.
- Zamknięcie połączenia odbywa się z wykorzystaniem mechanizmu uzgodnienia (zgoda na zamknięcie).
- TCP zapewnia sterowanie przepływem. Informuje partnera o tym ile bajtów danych ze strumienia danych może od niego przyjąć (okno oferowane). Rozmiar okna zmienia się dynamicznie i jest równy rozmiarowi wolnego miejsca w buforze odbiorcy. Zero oznacza, że nadawca musi poczekać, aż program użytkowy odbierze dane z bufora.
- Dane ze strumienia danych dzielone są na fragmenty, które według TCP mają najlepszy do przesłania rozmiar. Jednostka przesyłania danych nazywa się **segmentem**.
- TCP zapewnia **niezawodność** połączenia.

Mechanizmy niezawodności

- **Potwierdzanie otrzymania segmentów z mechanizmem zegara.**
Odebrany segment musi być potwierdzony przez odbiorcę przez wysłanie segmentu potwierdzającego. Jeśli potwierdzenie nie nadejdzie w odpowiednim czasie, segment zostanie przesłany ponownie.

- **Sumy kontrolne.**
Jeśli segment zostanie nadesłany z niepoprawną sumą kontrolną, to jest odrzucany. Nadawca po odczekaniu odpowiedniego czasu prześle segment jeszcze raz.
- **Przywracanie kolejności nadchodzących segmentów.**
Segmenty mogą nadchodzić w kolejności innej niż zostały wysłane, oprogramowanie TCP przywraca prawidłową kolejność przed przekazaniem do aplikacji.
- **Odrzucanie zdublowanych danych.**

6.2.1 Nagłówek TCP

- **Numer sekwencji.**
Dla segmentu z ustawionym tylko znacznikiem SYN w to pole wpisywany jest (losowy) numer sekwencji początkowej (Initial Sequence Number). Taki segment jest wysyłany w celu rozpoczęcia nawiązywania połączenia. Pierwszy oktet przesyłanych danych ze strumienia w rzeczywistości będzie miał numer ISN+1. W kolejnych segmentach w połączeniu TCP w polu numer sekwencji jest numer pierwszego oktetu danych w segmencie. Numer potwierdzenia jest ważny tylko przy włączonym znaczniku ACK. Znacznik ACK jest włączany tylko wówczas, jeśli segment zawiera potwierdzenie odebrania jakiegoś segmentu. Numer potwierdzenia jest to kolejny numer bajta (w strumieniu danych), którego spodziewa się wysyłający potwierdzenie.
- **Długość nagłówka** (przesunięcie danych).
Wielkość nagłówka wyrażona w liczbie bloków 4-bajtowych.
- **Jednobitowe znaczniki (flagi):**
 - NS, CWR, ECE – związane z przeciwdziałaniem przeciążeniom na routerach.
 - NS – jednobitowa suma kontrolna wartości flag związanych z mechanizmem ECN weryfikująca ich integralność;
 - ECE – Jeśli flaga SYN=0, to flaga ECE jest ustawiana na 1 przez odbiorcę segmentu, jeśli segment ten dotarł w pakiecie IP z ustawionym kodem CE w bitach ECN nagłówka IP; Jeśli SYN=1, wówczas ECE=1 oznacza, że mechanizm ECN będzie stosowany (czyli jest to element nawiązania połączenia, w którym druga strona połączenia jest informowana o tym, że można stosować ECN),
 - CWR – flaga potwierdzająca odebranie ECE i zredukowanie wielkości okna przeciążeniowego.
 - URG – wskaźnik ponaglenia
 - ACK – segment potwierdzenia,
 - PSH – segment „push” – wypchnięcie danych,
 - RST – zresetowanie połączenia,
 - SYN – synchronizacja,
 - FIN – nadawca zakończył przesyłanie danych.
- **Rozmiar okna.**
Oznacza liczbę bajtów, które odbiorca może zaakceptować.
- **Suma kontrolna.**
Liczona dla nagłówka i danych, z użyciem pseudonagłówka (algorytm analogiczny jak w UDP oraz IP).
- **Wskaźnik ważności.**
Dodatnie przesunięcie, które musi być dodane do numeru przesunięcia sekwencyjnego pierwszego oktetu danych aby uzyskać numer ostatniego bajta szczególnie ważnych danych w segmencie (dane te są na początku segmentu). Pole to jest brane pod uwagę tylko, jeśli bit URG jest ustawiony na 1.

- **Opcje.**

Rodzaj opcji (bajt), długość opcji (bajt), opcja. Najważniejsza opcja to **MSS**. Może być uzyskana jako MTU (Maximum Transmission Unit) minus rozmiar nagłówka IP oraz TCP.

Specyfika stanu TIME WAIT

Spóźnione segmenty są w czasie 2 MSL odrzucane. Para punktów końcowych definiujących połączenie nie może być powtórnie użyta przed upływem 2MSL. Eliminuje to ewentualne kłopoty związane z odbieraniem z sieci segmentów jeszcze ze starego połączenia.

Półzamknięcie TCP

Strona, która zakończyła połączenie i nie nadaje danych, może dane odbierać od partnera TCP. Takie połączenie nazywane jest połączeniem półzamkniętym (half-closed).

Segmenty RST

Segment RST wysyłany jest przez oprogramowanie implementujące TCP, kiedy nadchodzi segment niepoprawny z punktu widzenia dowolnego połączenia. Segment RST nie jest potwierdzany. W protokole UDP generowany jest komunikat ICMP o tym, że port jest nieosiągalny. Segment RST jest wysyłany również wtedy, gdy przekroczona jest maksymalna dopuszczalna liczba połączeń TCP.

Połączenia półotwarte (połowicznie otwarte)

Jest to połączenie nie poprawnie nawiązane. Występuje, jeśli jedna ze stron przerwała połączenie bez informowania drugiej. Segment z ustawioną na 1 flagą SYN został przesłany od klienta do serwera, serwer odpowiedział segmentem z ustawionymi na 1 flagami SYN i ACK, ale klient nie odpowiedział segmentem z ustawioną na 1 flagą ACK. Jeden ze sposobów atakowania serwisów (np. WWW) polegał na otwieraniu bardzo dużej liczby połączeń półotwartych. Obecnie implementacje TCP są odporne na tego typu ataki. Dopuszczalne jest, by oprogramowanie realizujące TCP mogło sprawdzać stan połączenia przez okresowe przysyłanie segmentów sprawdzających aktywność. Segment taki to zawiera ustawioną na 1 flagę ACK i nie zawiera żadnych danych. Dodatkowo ma on ustawiony numer sekwencyjny na o 1 mniejszy od tego, którego normalnie spodziewa się strona wysyłająca ACK. Partner odpowiada też segmentem z ustawioną na 1 flagą ACK ze standardowo ustawionymi prawidłowymi wartościami numerów sekwencyjnych.

6.2.2 Opcje TCP

- **Koniec listy opcji:** 1 bajt - 0.
- **Brak operacji:** 1 bajt - 1. Opcja ta jest stosowana do dopełnienia pola do wielokrotności 4 bajtów.
- **MSS:** typ = 2, długość = 4, dane – 2 bajty ustawione na mniejszą wartość z wartości MSS podanych przez strony połączenia. MSS podawany jest w segmentach SYN.
- **Skala okna:** typ = 3, długość = 3, dane – 1 bajt = współczynnik skali (licznik przesunięć), określa liczbę mnożeń przez 2 wielkości rozmiar okna, max. 14.
- **Pozwolenie na selektywne potwierdzanie:** typ opcji = 4, długość = 2.
- **Selektywne potwierdzanie:** typ = 5, długość = 10, 18, 26 lub 34 (potwierdzenie jednego, dwóch, trzech lub czterech nieciągłych bloków danych).
- **Sygnatury czasowe:** typ = 8, długość = 10, timestamp – 4 bajty, echo timestamp – 4 bajty. Stosowana jest do określania RTO (retransmission time-out).

6.2.3 Przepływ danych w TCP

Potwierdzenia

- **Skumulowane potwierdzanie**

Segment TCP z ustawionym znacznikiem ACK jest segmentem potwierdzającym. Numer potwierdzenia to numer następnego bajta, który nadawca ACK spodziewa się otrzymać od nadawcy potwierdzanych danych. Taki sposób potwierdzania nie jest korzystny w środowisku z dużą liczbą gubionych segmentów.

- **Opóźnione potwierdzenia**

Potwierdzenie otrzymania segmentu nie jest wysyłane natychmiast. Zalety takiego podejścia są następujące:

- Podczas trwania opóźnienia mogą zostać odebrane następne segmenty i potwierdzenie może je obejmować.
- Można dołączyć ACK do segmentu przesyłanych danych (piggybacking).
- Oprogramowanie TCP realizuje utrzymanie połączenia.

- **Selektywne potwierdzenia**

Realizowane przez jedną z opcji TCP. Przydaje się w łączach które są szerokopasmowe i mają duże opóźnienie. Rozmiar okna może tu być duży, nadawca może na raz transmitować dużą liczbę segmentów. Jeśli pierwszy segment zaginie, to nadawca niepotrzebnie retransmituje wszystkie. Selektywne potwierdzanie temu zapobiega.

Ruchome okna TCP (sliding windows)

Połączenie TCP obejmuje dwa strumienie danych. W każdym strumieniu określony jest nadawca i odbiorca. Kontrolę przesyłania oktetów w strumieniu umożliwiają mechanizmy tzw. przesuwanych (ruchomych) okien, które można sobie wyobrazić jako nałożone na strumień. Dla strumienia określone jest okno nadawcy oraz okno odbiorcy. Nadawca może wysyłać tylko te dane, które są w tej chwili w jego oknie nadawczym, przy czym może to zrobić tylko za zgodą odbiorcy. Okno nadawcze jest przesuwane nad wyjściowym strumieniem bajtów, okno odbiorcze nad strumieniem wejściowym.

Okno nadawcze Nadawca określa, które oktetety w jego strumieniu wyjściowym zostały:

- wysłane i potwierdzone,
- wysłane, ale jeszcze nie potwierdzone,
- jeszcze nie wysłane, ale już znajdujące się w oknie nadawczym
- jeszcze nie wysłane i znajdujące się poza oknem nadawczym

Dane w oknie nadawczym mogą być wysłane przez nadawcę, gdyż odbiorca na to zezwolił, lub jeszcze raz wysłane wskutek realizacji strategii powolnego startu lub zapobiegania zatorom. Okno nadawcze rozciąga się od oktetu, którego spodziewa się partner w ostatnim potwierdzeniu ACK.

Okno nadawcze ma lewą i prawą krawędź. W trakcie potwierdzania kolejnych segmentów lewa krawędź okna nadawczego przesuwa się w prawo powodując tzw. zamykanie okna. Jeśli numer potwierdzenia + rozmiar okna wskazują na konieczność przesunięcia prawej krawędzi, to jest ona przesuwana powodując tzw. otwieranie okna.

Jeśli bufor odbiorczy się wypełni to odbiorca przesyła z potwierdzeniem ostatniego segmentu proponowany rozmiar okna równy zero. Kiedy bufor odbiorczy zostanie całkowicie opróżniony odbiorca wysyła „zbędny” segment ACK bez danych i z numerem potwierdzenia takim jak poprzedni ACK. Ten segment nie jest potwierdzany ani retransmitowany.

Okno odbiorcze Odbiorca określa, które oktetety w jego strumieniu wejściowym zostały:

- odebrane, potwierdzone i pobrane do warstwy aplikacji
- odebrane, potwierdzone, ale jeszcze nie pobrane do warstwy aplikacji
- odebrane, ale jeszcze nie potwierdzone
- nie odebrane, ale znajdujące się w tzw. bieżącym oknie odbiorczym (wolna przestrzeń buforu odbiorczego)
- nie odebrane i znajdujące się poza oknem odbiorczym (nie zostaną odebrane)

Maksymalne okno odbiorcze obejmuje wszystkie oktetety odebrane. Ma ono stałą wielkość równą buforowi do odbierania danych w połączeniu TCP. Bieżące okno odbiorcze obejmuje odebrane niepotwierdzone i będące w bieżącym oknie odbiorczym.

Ze względu na to, że oktety są ponumerowane łatwo można określić, w którym miejscu strumienia powinny się znaleźć. W każdym segmencie jest przekazywany numer pierwszego oktetu danych. Bieżące okno odbiorcze ma lewą i prawą krawędź, analogicznie jak dla okna nadawczego. Prawa krawędź odbiorczego okna bieżącego pokrywa się z prawą krawędzią maksymalnego okna odbiorczego. Jeśli aplikacja odbiera dane z bufora, to krawędź ta przesuwa się w prawo (podobnie jak lewa krawędź maksymalnego okna odbiorczego). Potwierdzenie ACK przesuwa lewą krawędź okna bieżącego w prawo.

6.2.4 Przesyłanie małych segmentów

Tak określa się segmenty o rozmiarze mniejszym od MSS. Przesyłanie małych segmentów zmniejsza efektywność sieci. Aby temu zapobiec stosowane są takie mechanizmy, jak algorytm Nagle'a i unikanie „syndromu głupiego okna”.

- **Algorytm Nagle'a**

„Dopasowuje się” do sieci, w której przesyłane są segmenty.

- TCP powinno przysłać tylko pojedyncze małe niepotwierdzone segmenty.
- W czasie oczekiwania na potwierdzenie, dane są gromadzone w buforze.
- W sieci o szerokim paśmie i małym opóźnieniu akumulacja jest mała, efektywność mniejsza, ale kompensowana szybkością sieci. W sieciach o małym paśmie i dużych opóźnieniach akumulacja zwiększa efektywność.

Algorytm Nagle'a może być wyłączany przez oprogramowanie TCP.

- **Syndrom głupiego okna (SWS)**

- Jeśli odbiorca ma zerowy rozmiar okna (i nadawca też) oraz warstwa aplikacji pobierze 1 bajt, to okno odbiorcze otwiera się o jeden bajt.
- Odbiorca może wysłać ACK z proponowaną wielkością okna równą 1. Gdyby tak było, to druga strona połączenia też powiększy okno nadawcze o jeden bajt i prześle segment zawierający 1 bajt danych itd.
- Aby uniknąć SWS odbiorca ogłasza nowy rozmiar okna dopiero, gdy rozmiar ten wynosi co najmniej MSS lub połowę maksymalnego rozmiaru okna odbiorczego.
- Nadawca unika SWS wstrzymując się z wysyłaniem danych dopóki rozmiar okna proponowanego przez odbiorcę nie jest równy co najmniej MSS. „Interaktywne” dane mogą być wysyłane z flagą PSH wg algorytmu Nagle'a.

Dodatkowa kontrola przepływu po stronie nadawcy

Nadawca może przesłać wszystkie segmenty, które znajdują się w oknie nadawczym. Takie działanie może jednak doprowadzić do zatorów. Poniższe algorytmy bazują na pojęciu okna przeciążeniowego (okna zatoru, congestion window) i zapobiegają zatorom oraz powodują unikanie powtórnej zapaści.

- **Algorytm powolnego startu**

Po otwarciu połączenia lub dłuższym czasie nie przesyłania danych wielkość okna przeciążeniowego ustawiana jest na $2 \cdot MSS$. Każde przychodzące potwierdzenie (ACK) powoduje zwiększenie okna przeciążeniowego o jeden MSS. Może to prowadzić do wykładniczego wzrostu wielkości tego okna. Rozmiar okna nadawczego ustalany jest jako minimum z wielkości okna przeciążeniowego oraz ogłoszonego przez odbiorcę bieżącego okna przeciążeniowego.

- **Algorytm unikania zatoru**

Tu stosuje się wolniejszy wzrost wielkości okna przeciążeniowego, np. o jeden segment na kilka przychodzących ACK. Algorytm ten działa zwykle od pewnego progu (najpierw działa powolny start).

6.2.5 Retransmisje segmentów w TCP

W każdym połączeniu definiowana jest zmienna RTO (Retransmission Time-out). Jeśli TCP nie odbierze ACK w czasie RTO dla pewnego nadanego segmentu, to segment musi być retransmitowany. RTO powinien być większy od stale obliczanego RTT (Round Trip Time). Segmenty ACK (bez danych) nie są potwierdzane. Wyznaczenie prawidłowej wartości RTO jest ważne dla uniknięcia zapaści spowodowanej przeciążeniem (np. RTT rośnie szybciej niż wyliczany RTO, segmenty będą retransmitowane co jeszcze zwiększy ruch w sieci).

Zasady retransmisji Jeśli nie określono innych: dla początkowego segmentu przyjmuje się bieżące znane RTO dla połączenia. Po upływie RTO wartość RTO jest podwajana, a segment wysyłany ponownie.

Potwierdzenie dla retransmitowanego segmentu jest niejednoznaczne, nie wiadomo czy jest to opóźnione potwierdzenie pierwszej kopii, czy potwierdzenie drugiej kopii. Rodzi to problemy z szacowanym RTT.

Niejednoznaczność potwierdzenia – Algorytm Karni Pomiary RTT dla retransmitowanych segmentów są pomijane. Może to rodzić problemy, jeśli RTT gwałtownie wzrośnie. Dlatego tymczasowo stosowane jest dublowanie RTO dla kolejnych retransmitowanych segmentów. Dopiero przyjęcie ACK segmentu nie retransmitowanego powoduje obliczenie RTT i RTO dla tego nowego czasu RTT.

7 Informacje uzupełniające

7.0.1 Szerokość pasma

W sieciach komputerowych termin ten oznacza na ogół liczbę bitów, które mogłyby być przesłane w danej technologii w ciągu sekundy. Może oznaczać również częstotliwość zegara taktującego, wykorzystywanego w danej technologii sieciowej.

7.1 Przepustowość

Przepustowość jest miarą ilości użytecznej informacji dostarczonej z sukcesem przez ścieżkę komunikacyjną. Przepustowość odnosi się do mierzonej efektywności systemu, np. łącze o szerokości pasma 100 Mb/s może osiągnąć przepustowość np. 70 Mb/s, ze względu na implementację, wykorzystane protokoły, szyfrowanie itd.

W zależności od tego czy uwzględnia się narzut związany z technologią, narzut związany z protokołami komunikacyjnymi definiuje się różne odmiany przepustowości.

- **Przepustowość maksymalna**

- maximum theoretical throughput - największa możliwa do przesłania liczba bitów danych w jednostce czasu; do tych bitów NIE wlicza się narzutu warstwy pierwszej i drugiej,
- peak measured throughput - wartość mierzona w rzeczywistym systemie lub na symulatorze w krótkim odcinku czasu,
- maximum sustained throughput - średnia wartość dla długotrwałych obciążeń.

- **Efektywna przepustowość** (Goodput, przepustowość warstwy aplikacji)

- Mierzy liczbę bitów danych efektywnie przesłanych w warstwie aplikacji w jednostce czasu.
- Iloraz liczby bitów dostarczonych w warstwie aplikacji do czasu mierzonego od wysłania pierwszego do dostarczenia ostatniego przesłanego bitu.
- Nie zalicza się narzutu związanego z technologią warstwy drugiej i pierwszej, protokołami z warstw wyższych, niezawodnością.

7.2 Pozostałe

- **Opóźnienie** - w sieciach z przełączaniem pakietów oznacza czas, jaki mija od wysłania pakietu do odebrania go przez adresata. Może być mierzony w jedną stronę (one way) lub w dwie (round-trip delay time).
- **Zmienność opóźnienia** - miara krótkotrwałych zmian w opóźnieniu. Duża zmienność ma bardzo niekorzystny wpływ np. na jakość transmitowanego dźwięku.
- **Czas propagacji** - czas w jakim sygnał pokona pewną odległość w medium transmisyjnym.
- **Czas transmisji** - czas jaki zajmuje umieszczenie pakietu w medium transmisyjnym.
- **Opóźnienie przekazania** - czas jaki mija od odebrania pakietu przez urządzenie (przekazujące pakiet) do momentu, gdy pakiet może być wysłany.
- **Czas przetwarzania** - czas w jakim pakiet jest przetwarzany w urządzeniu, obejmuje np. wykonanie zmian w nagłówkach, wyznaczenie routera następnego skoku itd.
- **Niezawodność sieci**
- **Współczynnik gubienia pakietów**, Packet Loss Ratio (PLR) - iloraz liczby straconych pakietów do całkowitej liczby pakietów przesłanych w pewnym czasie.

8 DNS

Oprócz adresu IP komputer ma przyporządkowaną **nazwę**. Konwencje nazywania komputerów:

- **nazwy hosta**
W czasach gdy istniała sieć ARPAnet każdy komputer miał przydzieloną własną unikalną nazwę. Nazwy wszystkich komputerów były przechowywane w pliku HOSTS.TXT. Plik ten musiał być ręcznie aktualizowany i centralnie zarządzany oraz rozsyłany z centrum do lokalnych komputerów.
- **nazwy DNS**
Domain Name System – system nazw domenowych, który jest rozproszonym systemem przechowującym informacje o nazwach komputerów i ich numerach IP. System ten zawiera mechanizmy tłumaczenia nazw. Dane DNS przechowywane są na serwerach nazw. Serwer DNS jest odpowiedzialny za swój fragment sieci i udostępnia swoje dane innym serwerom. DNS jest systemem hierarchicznym i jego strukturę można przedstawić w postaci drzewa. Termin domena DNS jest często utożsamiany z poddrzewem drzewa DNS. Dostęp do serwera DNS jest realizowany przez mechanizm określany czasem jako resolver.
- **nazwy NetBIOS.**

Istnieją mechanizmy tłumaczące nazwy na numery IP i odwrotnie. Ciałem odpowiedzialnym za koordynację nazw domen górnego poziomu a także odpowiedzialnym za przypisywanie adresów IP jest IANA Internet Assigned Numbers Authority. Ciałem nadzorującym od strony technicznej różne działania związane z uzyskiwaniem (rejestracją) nazw domen, numerów IP, numerów portów jest ICANN - Internet Corporation for Assigned Names and Numbers.

Domeny górnego poziomu

- arpa - specjalna, wykorzystywana do odwzorowania adresów IP w nazwy.
- Domeny podstawowe (generic, gTLD), np: com, edu, gov.
- Domeny geograficzne (krajowe, country-code ccTLD), np: pl, uk, de.

Domeny drugiego poziomu - w wielu krajach domeny drugiego poziomu odzwierciedlają domeny organizacyjne pierwszego poziomu, ale ujmowane na swoim terytorium. Przykłady: edu.pl, com.pl, co.uk, ac.uk.

W DNS (np. w plikach konfiguracyjnych) występują często tzw. **absolutne nazwy domeny**, inaczej w pełni określone nazwy domeny (FQDN). FQDN jest to nazwa domeny zakończona kropką (np. ii.uj.edu.pl.). Jeśli nazwa nie jest zakończona kropką, to może być jakoś uzupełniana.

Obszar, inaczej **strefa** (zone) jest częścią systemu DNS, która jest oddzielnie administrowana. Domeny drugiego poziomu dzielone są na mniejsze strefy. Z kolei te strefy mogą być dalej dzielone. Występuje tu delegowanie zarządzania w dół struktury drzewa. Jednostka odpowiedzialna za zarządzanie daną strefą decyduje ile będzie serwerów DNS w strefie, rejestruje i udostępnia nazwy i numery IP nowych komputerów zainstalowanych w strefie. W tej chwili jest na świecie 13 (typów) serwerów głównych (najwyższego poziomu) zwanych po angielsku root-servers, posiadającymi nazwy od a.root-servers.net do m.root-servers.net.

Poszukiwania w DNS

- Proste, „do przodu” – klient zna nazwę domenową, a chce uzyskać numer IP.
- Odwrotne (reverse) – klient zna adres IP i chce uzyskać nazwę domenową. Przeszukiwanie odwrotne wykorzystuje domenę arpa.in-addr. Jeśli chcemy poznać nazwę domenową komputera, to w systemie DNS adres ten jest reprezentowany jako specyficzna nazwa w domenie arpa.in-addr.

8.1 Typy serwerów DNS

W każdej strefie musi być uruchomiony podstawowy serwer DNS oraz pewna liczba serwerów drugoplanowych, zapewniających usługi w razie awarii serwera podstawowego. Serwer podstawowy pobiera dane z pliku konfiguracyjnego, natomiast serwery drugoplanowe uzyskują dane od serwera podstawowego na drodze tzw. transferu strefy (zone transfer). Serwery drugoplanowe odpytują serwer podstawowy o dane w sposób regularny, zwykle co kilka godzin. Oprócz dwóch wymienionych rodzajów serwerów są jeszcze serwery podręczne (lokalne), których zadaniem jest zapamiętanie na pewien czas w pamięci podręcznej danych uzyskanych od innych serwerów tak, aby kolejne zapytania klientów mogły być obsłużone lokalnie. Serwery DNS działają na portach 53 UDP oraz 53 TCP. Na ogół w warstwie transportu używany jest UDP. Wyjątkiem jest m.in. transmisja danych z serwera podstawowego do drugoplanowego (większe porcje danych) oraz komunikaty w sieciach WAN. Również kiedy w odpowiedzi od serwera (przez UDP) ustawiony jest bit TC (patrz niżej) ponawiane jest zapytanie z wykorzystaniem TCP.

Podział ze względu na sposób uzyskania odpowiedzi poszukiwania

- Przeszukiwanie rekurencyjne – klient oczekuje od serwera żądanej informacji. W przypadku, gdy serwer nie przechowuje żądanej informacji, sam znajduje ją na drodze wymiany komunikatów z innymi serwerami.
- Przeszukiwanie iteracyjne – występuje między lokalnym serwerem DNS a innymi serwerami DNS. Jeśli odpytywany serwer nie zna szukanego adresu IP, odsyła pytającego do innych serwerów (odpowiedzialnych za daną domenę).

Komunikacja klienta z serwerem DNS

Przy odwołaniu do nazwy domenowej system zwykle najpierw sprawdza, czy nie jest to nazwa hosta lokalnego, następnie sprawdza plik hosts - o ile istnieje. Jeśli nie znajdzie odpowiedniego wpisu, to wysyłane jest zapytanie do pierwszego serwera DNS (adres w pliku konfiguracyjnym).

Standardowy sposób poszukiwania

Klient pyta swój domyślny serwer DNS wysyłając zapytanie rekurencyjne. Odpytany serwer realizuje zapytania iteracyjne, zaczynając od serwerów głównych, które odsyłają do serwerów niższego poziomu. Mechanizm ten może się wydawać nieefektywny, ale w rzeczywistości dzięki temu, że serwery DNS zapamiętują na pewien czas informacje uzyskane z innych serwerów DNS (cache), często odpowiedź na zapytanie programu-klienta zostaje znaleziona bardzo szybko.

Dynamiczny DNS (DDNS)

Chyba najważniejsze wpisy w DNS dotyczą serwisów, np. www. Standardowo DNS obsługuje odwzorowanie nazw do statycznych adresów IP. Można jednak skonfigurować odwzorowanie dla adresów zmieniających się dynamicznie. W tym celu należy skorzystać z odpowiednich usługodawców w Internecie, którzy przypisują nazwę do swojego IP i pewnego numeru portu, następnie zapytanie przekierowują do

komputera ze zmiennym IP z ewentualną zmianą portu. Na komputerze ze zmiennym IP należy zainstalować odpowiedni program (klient DDNS), który będzie powiadamiał serwer DDNS o zmianach adresu IP. Oddzielnym problemem, który należy rozwiązać, jest wykorzystanie serwera NAT i przypisywanie adresów prywatnych do serwisu w sieci. Na ogół wystarczy odpowiednie działanie klienta DDNS oraz odpowiednie skonfigurowanie serwera NAT.

8.2 Przykładowe konfiguracje serwerów DNS

Przykładowy plik `/etc/resolv.conf` (Unix) `nameserver 149.156.78.3 nameserver 149.156.78.95 domain xx.yy.edu.pl` Podstawowe wpisy to – po słowie `nameserver` – numery IP kolejnych serwerów DNS, które mają być odpytywane w przypadku braku odpowiedzi od poprzedniego. Linia ze słowem `domain` oznacza domenę domyślną (tzn. co ma być dołączone do nazwy hosta w przypadku, gdy nie określono domeny, np. `ftp gandalf`). Może się pojawić linia podobna do `domain`, ze słowem `search`, po którym jest (maksymalnie sześć) nazw domen. Taka linia oznacza, że poszukiwania mają być prowadzone dla nazw z dołączonymi kolejno nazwami podanych domen.

Plik konfiguracyjny serwera `named` (`bind`): `/etc/named.boot` lub `/etc/named.conf`. Główną częścią BIND jest proces o nazwie „`named`”. Skonfigurowanie `bind` może polegać na utworzeniu m.in. pliku `/etc/named.conf`. Pliki te zawierają opcje działania programu i informacje o plikach stref. W `zone` lub `options` może się pojawić dyrektywa `allow-query`, która ogranicza adresy IP, z których mogą pochodzić zapytania. W katalogu `/var/named` należy utworzyć pliki stref, np. `xx.yy.edu.pl` Po każdorazowej zmianie w plikach konfiguracyjnych należy serwis restartować. Proces `named` powinien być startowany nie z prawami `roota`, tzn. należy utworzyć konto `named`, zamiast domyślnego `shella`.

8.3 Rekordy zasobów

Każdy serwer DNS przechowuje informacje o tej części obszaru nazw DNS, dla której jest autorytatywny (administratorzy są odpowiedzialni za poprawność informacji). Informacje zapisywane są w postaci tzw. rekordów zasobów. Dla zwiększenia wydajności serwer DNS może przechowywać również rekordy zasobów domen z innej części drzewa domen. **Istnieje szereg typów rekordów zasobów:**

SOA (Start of Authority) Rekord uwierzytelnienia – pierwszy rekord w pliku strefy, określa podmiot odpowiedzialny od tego punktu hierarchii „w dół”.

- Serial – pole zawierające numer wersji pliku strefy, zwykle w polu tym odzwierciedlona jest data oraz numer wersji pliku w danym dniu.
- Refresh – określa jak często serwer pomocniczy ma sprawdzać na serwerze podstawowym, czy nie zachodzi potrzeba uaktualnienia plików.
- Retry – czas, po którym serwer pomocniczy będzie ponownie próbował odtworzyć dane po nieudanej próbie odświeżenia.
- Expire – maksymalny limit czasu, przez który serwer pomocniczy może utrzymywać dane w pamięci `cache` bez ich uaktualnienia. Minimum (Default TTL) – domyślny czas, jaki ma być użyty dla rekordów, które nie mają określonego TTL.

8.4 DHCP - Dynamic Host Configuration Protocol

Wadą BOOTP jest statyczny sposób przydzielania numerów IP. Przydział dynamiczny umożliwia pracę (ale nie jednocześnie) wielu komputerów z przydzielonym jednym numerem IP. Serwer DHCP przydziela adresy IP dynamicznie. Obecnie w bardzo wielu sieciach lokalnych komputery nie mają na stałe wpisanych IP, ale pobierają IP od serwera DHCP w momencie startu systemu. Serwer DHCP może wykorzystywać różne sposoby przypisywania adresów:

- przydział statyczny IP do danego komputera (ustawienie „ręczne”, danemu adresowi MAC jest przypisywany stale jeden na stałe wybrany IP),
- automatyczny przydział statyczny przy pierwszym starcie komputera i kontakcie z serwerem,

- przydział dynamiczny, w którym serwer wynajmuje adres IP na określony czas. DHCP umożliwia budowanie systemów konfigurujących się automatycznie. Oprócz przydzielenia adresu IP serwer DHCP przesyła do komputera klienta również