Sieci TCP/IP - cz. 1
Wprowadzenie

Ostatnia modyfikacja: 25.03.2020

Wprowadzenie

Wybrane charakterystyki sieci:

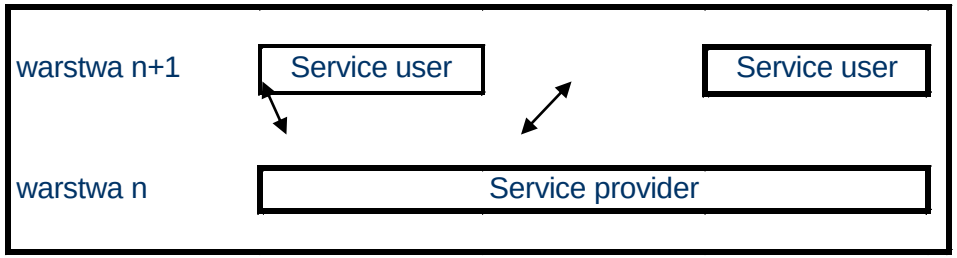
- medium transmisyjne: kabel współosiowy, skrętka, światłowód, fale radiowe ;
- topologia połączeń: szyna, pierścień, gwiazda, drzewo, krata;
- metoda transmisji: simplex, half-duplex, full-duplex;
- dostęp do kanału: centralne odpytywanie (polling), przekazywanie żetonu, dostęp jednoczesny z wykrywaniem kolizji (CSMA/CD), dostęp jednoczesny z unikaniem kolizji (CSMA/CA);
- rozległość: WAN, MAN, LAN.

- Informacje szczegółowe dotyczące sieci TCP/IP – na przedmiotach dotyczących komunikacji sieciowej (np. „Wprowadzenie do sieci TCP/IP”), czy transmisji danych.

Komunikacja sieciowa – model OSI

- Model OSI (ISO) reprezentuje architekturę sieci komunikacyjnej w postaci 7 warstw.
- Warstwa realizuje usługi dla warstwy bezpośrednio wyższej, zamawiając usługi u warstwy bezpośrednio niższej.

7	w. aplikacji
6	w. prezentacji
5	w. sesji
4	w. transportowa
3	w. sieciowa
2	w. łącza danych
1	w. fizyczna



- Wirtualna komunikacja pomiędzy warstwami OSI:

System A		System B
w. aplikacji	<--- prot. w. 7 ------>	w. aplikacji
w. prezentacji	<--- prot. w. 6 ------>	w. prezentacji
w. sesji	<--- prot. w. 5 ------>	w. sesji
w. transportu	<--- prot. w. 4 ------>	w. transportu
w. sieci	<--- prot. w. 3 ------>	w. sieci
w. łącza	<--- prot. w. 2 ------>	w. łącza
w. fizyczna	<== prot. w. 1 ===>	w. fizyczna
	poł. fizyczne	

Warstwa oprogramowania N na komputerze docelowym musi otrzymywać dokładnie ten sam komunikat, który został wysłany przez warstwę N na komputerze nadawczym.

Model odniesienia OSI 7 – c.d.

- **Warstwa 1 -** odpowiada za transmisję ciągu danych (bitów) po fizycznym nośniku informacji (np. kablu).
- **Warstwa 2** - odpowiada za transmisję ramek, wykrywanie błędów, kontrolę przepływu (dla usług połączeniowych). Dwie podwarstwy: kontroli dostępu do nośnika (MAC - Media Access Control) i kontroli łącza logicznego (LLC - Logical Link Control) (Ethernet II, 80.2, 802.3).
- **Warstwa 3** - odpowiada za przesyłanie komunikatów pomiędzy węzłami w sieci oraz realizowanie adresowania i kierowania pakietami (łącznie z obsługą ich podziału czy scalania); musi też zapewniać kontrolę przeciążeń (*congestion control*). (X.25, IP).
- **Warstwa 4** - odpowiada za przesyłanie wiadomości pomiędzy komputerami w sieci, niezależne od realizacji podsieci (relay and route), adresowanie węzłów. Dostępne tryby pracy: datagramowy (*connection-less*) np. UDP lub strumieniowy (*connection - oriented*), np. TCP, TP0,...,TP4. Dla trybu strumieniowego może być dodatkowo: kontrola przepływu (*flow control*), przesyłanie danych pilnych, potwierdzenie odbioru, reset kolejki danych. Usługi: ISO 8072, protokoły: ISO 8073.

Model odniesienia OSI 7 – c.d.

- **Warstwa 5** - odpowiada za organizowanie sesji: otwarcie, zamknięcie, przesyłanie danych (zwykłych i ekspresowych), zarządzanie interakcją (simplex, half-duplex, full-duplex), realizację synchronizacji. Usługi: ISO 8326, protokoły ISO 8327.

- **Warstwa 6** - odpowiada za konwersję danych (*abstract syntax-> transfer syntax->abstract syntax*).

- **Warstwa 7** - umożliwia użytkownikom aplikacji dostęp do środowiska OSI. Usługi są podzielone na grupy (*Common Application Service Elements, Reliable Transfer Service Element, Virtual Terminal Service, File Transfer Access and Management, Electronic Mail Services, Directory Services, Network Management service*). E-Mail, NFS, RPC.

Internet

- **Intersieć** składa się ze zbioru sieci połączonych za pomocą ruterów. Celem budowy intersieci jest zapewnienie jednolitych usług w sieciach heterogenicznych.
- **Internet** – intersieć o zasięgu ogólnoświatowym, oparta na *rodzinie protokołów TCP/IP*.
- Za rozwój i promocję standardów Internetu odpowiada m.in. otwarta organizacja standaryzacyjna Internet Engineering Task Force (**IETF**), współpracująca z World Wide Web Consortium (W3C) oraz ISO/IEC.
- **Standardy** IETF (Proposed, Draft czy Internet Standards (RFC1610)) są upubliczniane w postaci dokumentów Requests For Comments (**RFC**), patrz: https://www.rfc-editor.org/retrieve/, http://tools.ietf.org/html/
- **Architektura** rodziny protokołów TCP/IP nie jest zgodna z modelem referencyjnym OSI (RFC1122).

Model warstwowy TCP/IP

wg RFC 1122

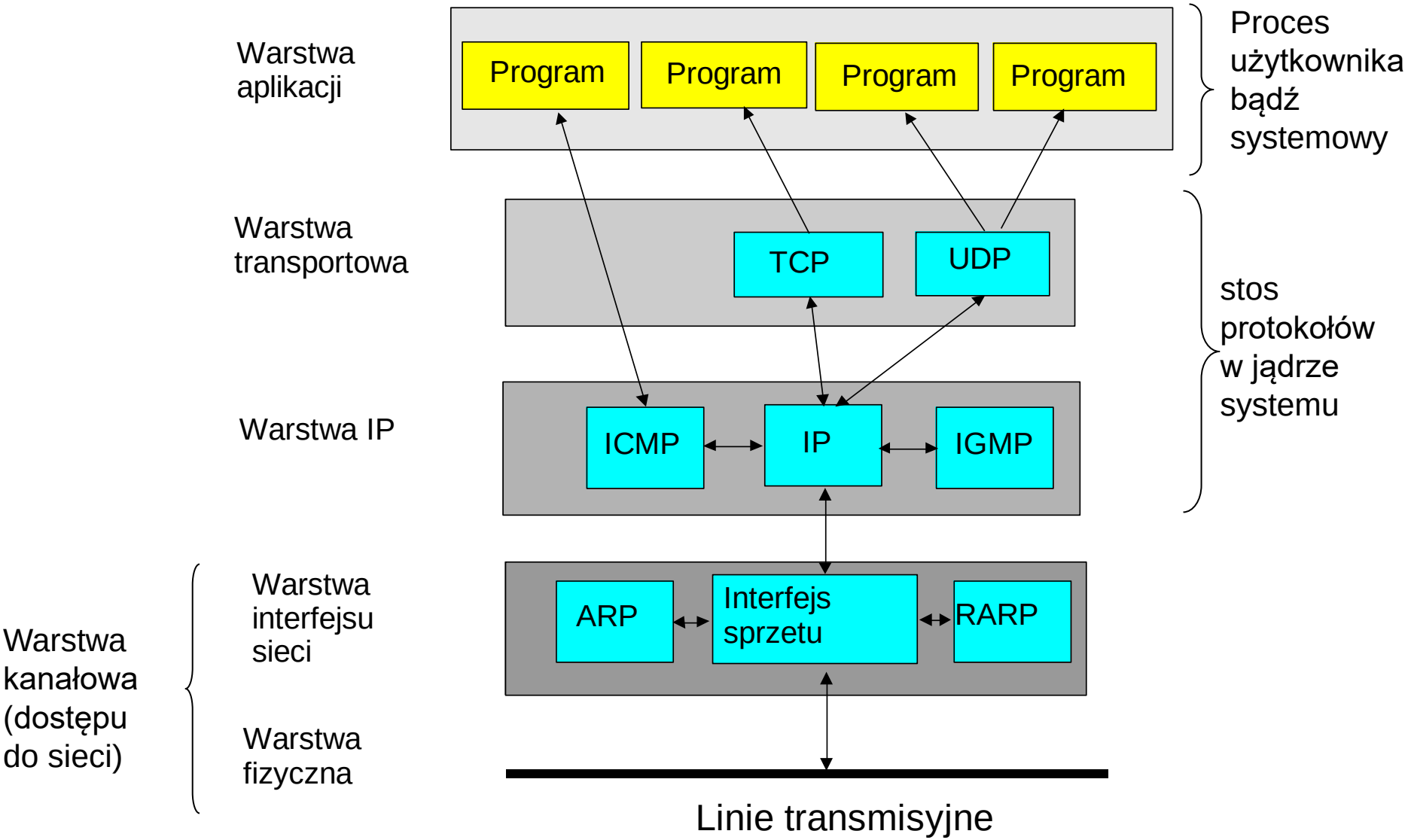
4	w. aplikacji
3	w. transportu
2	w. Internetowa (IP)
1	w. kanałowa (dostępu, *Link layer*)

Model warstwowy TCP/IP

wg Comera

5	w. aplikacji
4	w. transportu
3	w. intersieci (IP)
2	w. interfejsu sieciowego
1	w. sprzętowa (fizyczna)

Model sieci TCP/IP

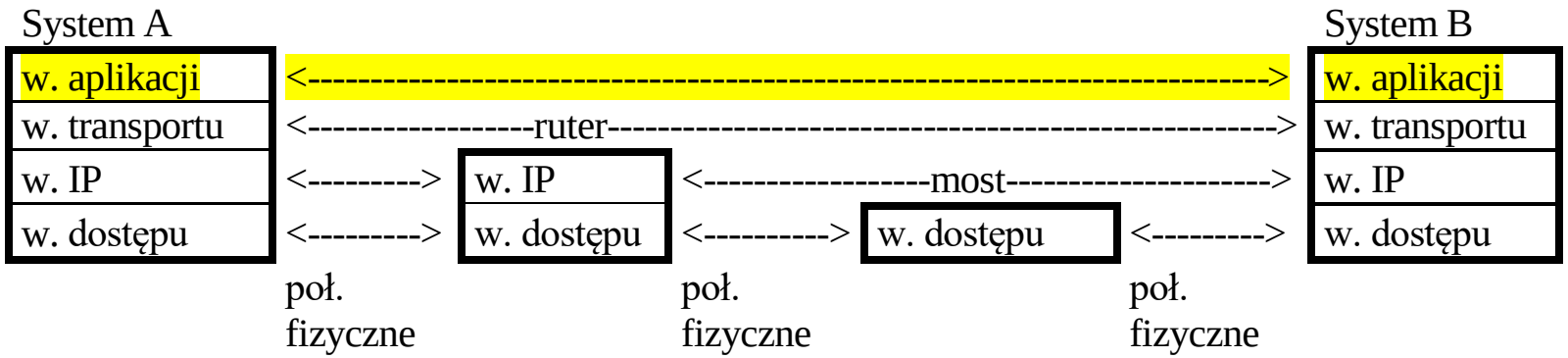


TCP/IP - przykład

Jak zachodzi przekazywanie pakietu sieciowego pomiędzy dwoma komputerami (hostami) podłączonymi do sieci Ethernet.

- Każdy host ma przydzielony unikalny adres IP skojarzony z adresem Ethernet (**Media Access Control** (**MAC**) address)
- Do komunikacji potrzebna jest znajomość obydwóch rodzajów adresów (nadawcy i odbiorcy)
- Usługa sieciowa **Domain Name Service** (**DNS**) może być użyta do pozyskania adresu IP
- Usługa sieciowa **Address Resolution Protocol** (**ARP**) umożliwia znalezienie odwzorowania adresu MAC na adres IP
  - W tym celu host wysyła zapytanie do wszystkich hostów sieci lokalnej (w trybie **Broadcast**). Odpowiada host, który zna odpowiedź.
- Jeżeli host odbiorcy znajduje się w tej samej sieci co nadawcy – można użyć zapytania ARP dla uzyskania jego adresu MAC .i wysłać pakiet bezpośrednio pod ten adres
- Jeżeli host odbiorcy znajduje się w innej sieci pakiet wysyłany jest do rutera, a ten kieruje pakiet dalej.

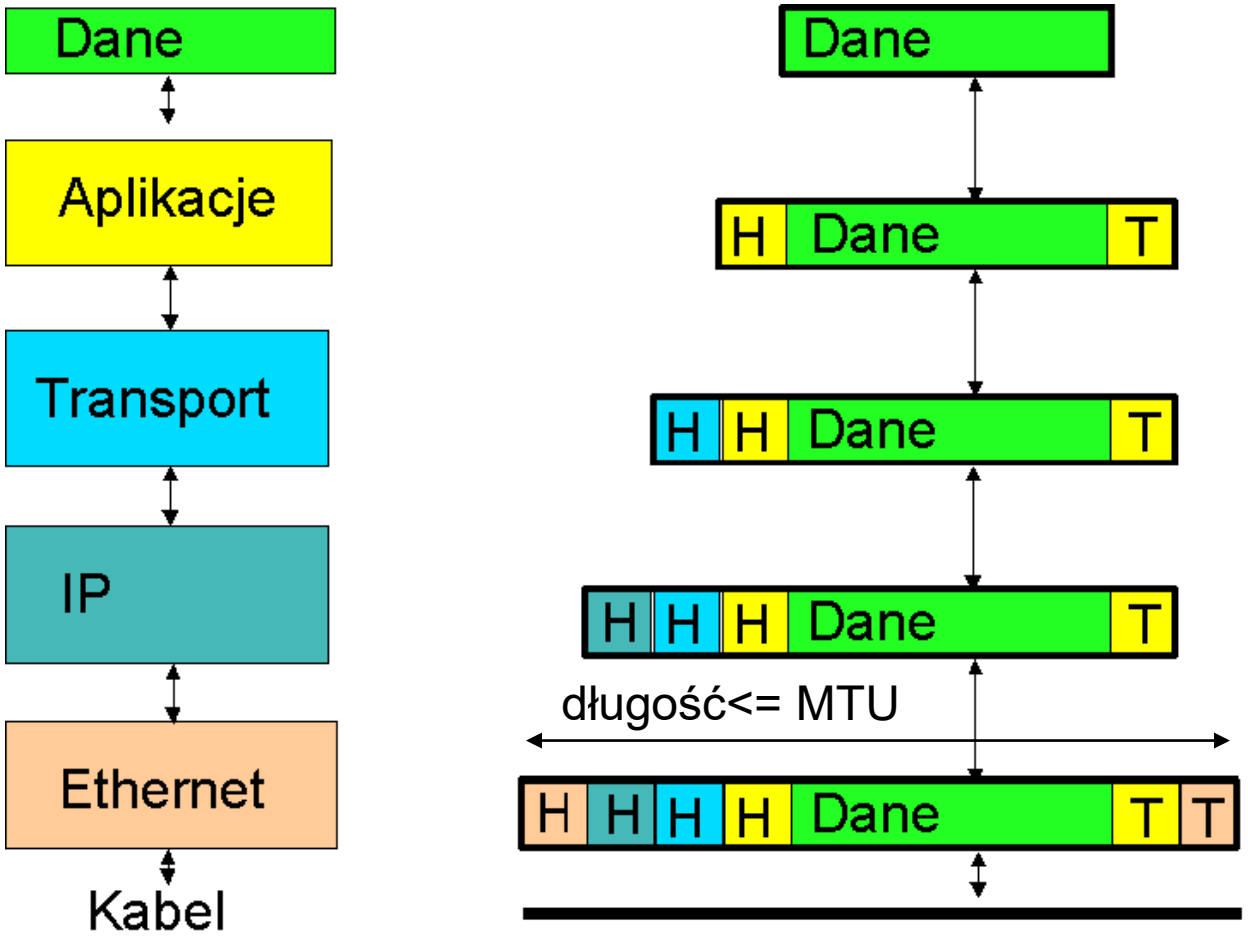
Komunikacja w sieci TCP/IP



- Przykład stosu protokołów

Nr warstwy	Warstwa	Protokół
4	aplikacji	HTTP
3	transportu	TCP
2	Internetowa (IP)	IP
1	dostępu do sieci	Ethernet 802.3u

Przesyłanie danych w sieci TCP/IP (+Ethernet)



Kapsułkowanie: H : header T : trailer

MTU dla Ethernetu: 1500, dla Ethernetu 802.3: 1492 oktetów (wg RFC 1122)

Przykład – protokół TFTP

- Przykład realizacji kapsułkowania – Trivial File Transfer Protocol (RFC1350)

Ethernet	Ethernet	IP	UDP	TFTP	data	Ethernet
8	14	20	8	4		4

<----------Ethernet frame (72 - 1526 octets) ---------->

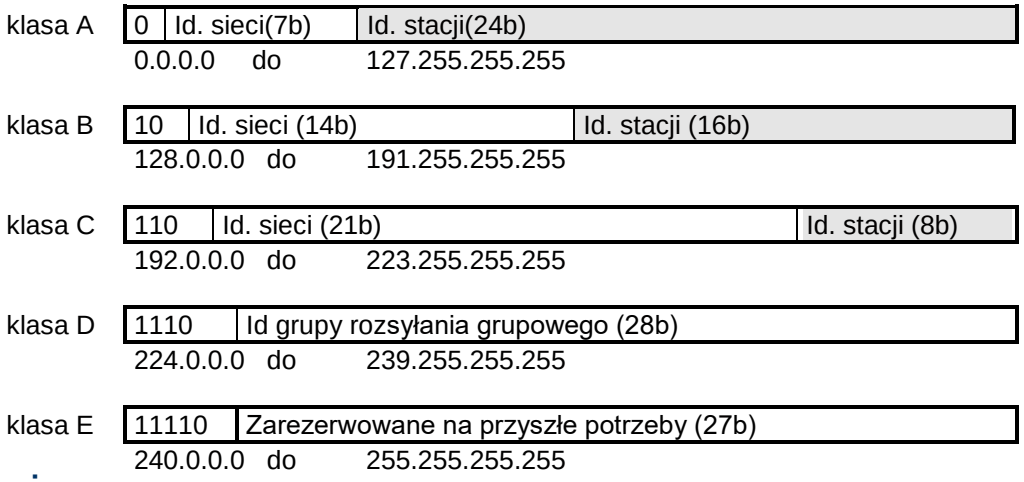
Typy transmisji danych

- **Unicast** – oddzielna kopia danych jest przesyłana ze źródła do każdego odbiorcy, który ich zażąda
- **Multicast** – pojedyncza kopia danych jest wysyłana do grupy odbiorców
- **Broadcast** – pojedyncza kopia danych jest przesyłana do wszystkich odbiorców znajdujących się w tym samym segmencie sieci co nadawca
- **Anycast** – pojedyncza kopia danych jest przesyłana do jednego z odbiorców z grupy odbiorców o tym samym adresie („najbliższego”)
- **Geocast** – doręczanie danych do grup odbiorców w sieci określonej przez położenie geograficzne. Jest to specjalizowana forma transmisji multicast, wykorzystywana do mobilnych sieci ad hoc.

Adresowanie klasowe w sieci TCP/IP

Budowa klas adresów IPv4 →
Klasa D: adresy multicast
Klasa E: adresy zastrzeżone

Adresy IP (a także domeny i protokoły Internetu) od lat 1970tych były zarządzane przez **Internet Assigned Numbers Authority** (IANA), obecnie w ramach the **Internet Corporation for Assigned Names and Numbers** (ICANN)



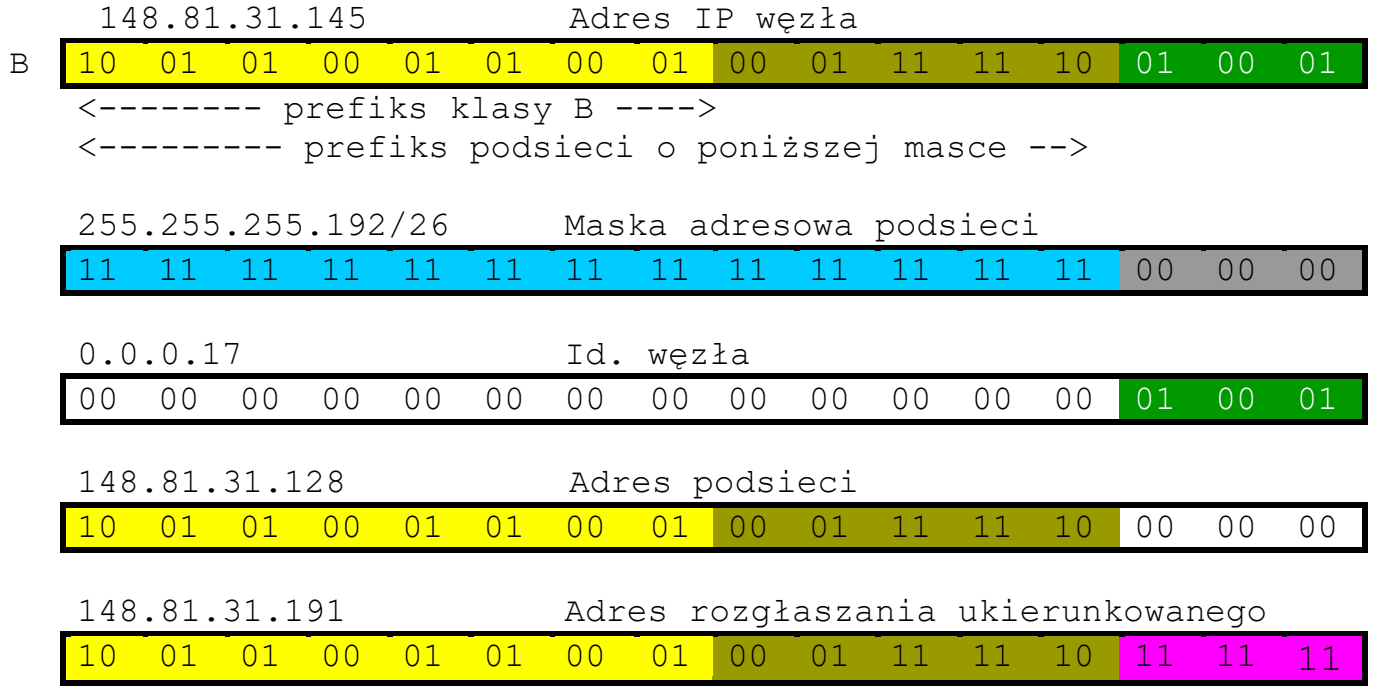
Adresy IP specjalnego przeznaczenia

- Prefix (adres sieci) >0 i suffix (id. stacji)=0 => adres (pod)sieci
- Prefix>0 i id. stacji=same jedynki (binarnie) => adres **rozgłaszania ukierunkowanego** (adresatem są wszystkie hosty wskazanej lokalnej sieci fizycznej)
- Adres 255.255.255.255 => adres **rozgłaszania ograniczonego** (w lokalnej sieci fizycznej)
- Adres 0.0.0.0 => komputer nie zna swojego IP (albo adres rozgłąszania typu UCB)
- Prefix 127/8 oznacza adresy **pętli zwrotnej** (pakiety nie opuszczają komputera); 127.0.0.1 ←→ localhost
- 224.0.0.0-224.0.0.255 – zakres lokalny dla łącza (*link local*), używany przy wykrywaniu niskopoziomowej topologii sieci lub protokołów zarządzania siecią. Pakiety z takim adresami nie są transowane.
- Bloki adresów dla sieci prywatnych (RFC1918):
  - 10.0.0.0 - 10.255.255.255 (1 klasa A; prefix 10/8)
  - 172.16.0.0 - 172.31.255.255 (16 klas A; prefix 172.16/12)
  - 192.168.0.0 - 192.168.255.255 (256 klas C; prefix 192.168/16)

Adresowanie bezklasowe w sieci TCP/IP

Zamiast korzystać z trzech klas adresowych można określić podział adresu (na prefiks i sufiks) w dowolnym miejscu – za pomocą maski adresowej.

Przykład.



CIDR: Classless InterDomain Routing (RFC1519)

Maska CIDR / 26 z przykładu odpowiada w notacji dziesiętnej z kropkami wartości: 255.255.255.192, tzn. prefiks adresu węzła jest 26 bitowy, a sufiks: 6-bitowy

Numery portów

- Adresy w sieci TCP/IP mają postać **nr_IP:nr_portu**, gdzie nr_IP jest numerem IP interfejsu sieciowego (IPv4: 32b), a nr_portu jest liczbą całkowitą (IPv4: 16b), która wyróżnia kanał komunikacji w zbiorze kanałów korzystających z tego samego numeru IP.
- Przy transmisji pojedynczego oktetu (bajtu) reprezentującego liczbę (zapis pozycyjny), pierwszy bit ma największą wagę, np. liczba 170 jest reprezentowana następująco --->

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+
|1 0 1 0 1 0 1 0|
+-+-+-+-+-+-+-+-+
```

- Przy transmisji liczby wielobajtowej pierwszy bit pierwszego bajtu ma największą wagę. Jest to tzw. **porządek sieciowy** (*Big Endian*).
- **Internet Assigned Numbers Authority** (IANA) jest centralnym koordynatorem przydziału liczb występujących w standardach protokołów Internetowych, a reprezentujących m.in: adresy IP, numery protokołów i numery portów (RFC1700).
- Wg IANA **porty ogólnie znane** (*well-known ports, system ports*) mają numery z przedziału 1-1023. ---->
- **Porty zarejestrowane** przez IANA, ale nie nadzorowane przez IANA *(user ports)* mają zakres: 1024-49151
- **Porty dynamiczne** (prywatne, efemeryczne) są z zakresu: 49152-65535.

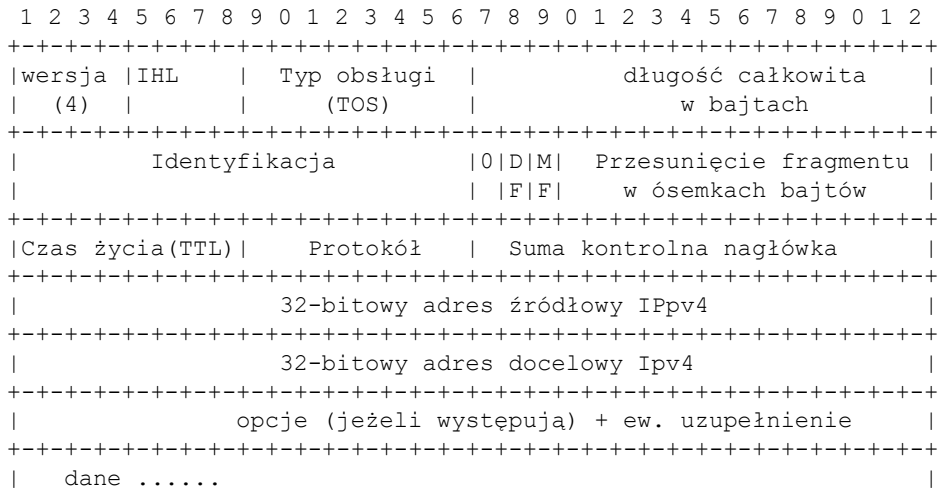
protokół	port(y)
ECHO	7
FTP	20,21
SSH	22
TELNET	23
DNS	53
HTTP	80
POP3	110
IMAP	143
HTTPS	443
X11	6000-6007

Protokoły podstawowe TCP/IP : IP

- **Internet Protocol** (IPv4) realizuje dwie podstawowe funkcje (RFC791):
  - Adresowanie pakietów (**datagramów**)
  - Fragmentacja/defragmentacja
- Protokół definiuje definiuje format pakietów (datagramów), określa operacje przekazywania pakietów pomiędzy sieciami i zbiór reguł, które służą do realizacji idei zawodnego przenoszenia pakietów bez użycia połączenia. Reguły opisują, w jaki sposób węzły i rutery powinny przetwarzać pakiety, jak i kiedy powinny być generowane komunikaty o błędach oraz kiedy pakiety mogą być porzucane.
- Protokół nie gwarantuje obsługi problemów związanych z
  - Doręczaniem pakietów z opóźnieniem, bądź w kolejności innej niż u nadawcy
  - Uszkodzeniem danych
  - Duplikowaniem bądź utratą pakietów
- Długość pakietu IP, które musi obsługiwać stos TCP/IP: 576 do **65535 oktetów**
- Dla sieci fizycznej o małym dozwolonym rozmiarze pakietu (MTU) konieczna jest **fragmentacja** pakietu (datagramu) IP (**scalanie** u odbiorcy)
- Pole ``czas życia'' (TTL) powoduje **porzucanie** niedoręczonego datagramu IP po TTL przejściach przez ruter oraz wysłanie nadawcy komunikatu ICMP o błędzie
- Można zażyczyć sobie trasowanie (rygorystyczne, swobodne), a także zapamiętywanie trasy datagramu IP (pole opcji w datagramie) – o ile rutery to wspierają

Protokoły podstawowe TCP/IP : IPv4 (*)

Format pakietu IP (RFC791)



- IHL*4=długość nagłówka, maks.15*4=60B
- TOS:
  - 3b priorytet (pomijany)
  - 4b rodzaj usługi
  - 1b zawierający 0
- Pole identyfikacja jest używane przy fragmentacji i łączeniu fragmentów
- Flagi:
  - DF (don't fragment) - zakaz fragmentacji
  - MF (more fragments), =0 gdy brak fragmentacji, albo ostatni fragment
- TTL – licznik czasu życia (0-255); przejście każdego rutera (*hop*) dekrementuje licznik.
- Nr protokułu (RFC1700):
  - 1: ICMPv4
  - 2: IGMPv4
  - 6: TCP
  - 17 UDP

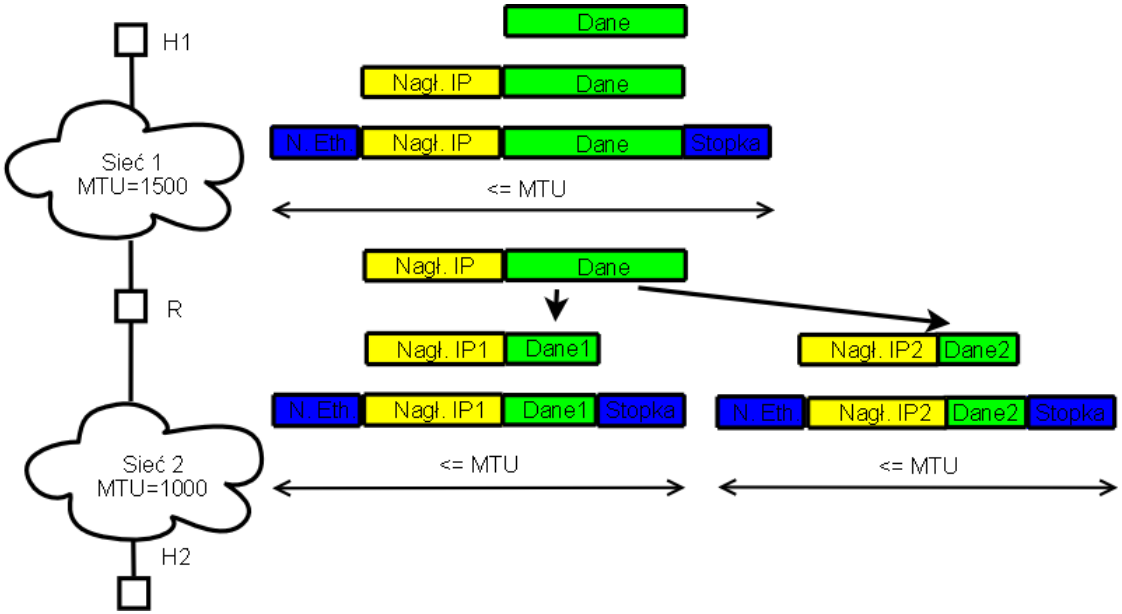
Opcje protokołu IPv4

1. Nic nie rób (wypełnienie)
2. EOL: koniec listy opcji
3. LSRR: wyznaczanie trasy swobodnej przez nadawcę oraz zapisywanie trasy(maks. 9 adresów pośrednich)
4. SSRR: wyznaczanie trasy rygorystycznej ...
5. Zapisywanie znaczników czasowych
6. Zapisywanie trasy
7. Podstawowe zabezpieczenia
8. Rozszerzone zabezpieczenia (RFC2113)
9. Identyfikator strumienia (przestarzałe)
10. Powiadomienie rutera (RFC2113)

Uwaga: opcje mogą być pobierane i ustawiane za pomocą funkcji interfejsu gniazd (**getsockopt, setsockopt**).

Protokół IP: fragmentacja

- Każda technika sieciowa określa maksymalny rozmiar jednostki transmisyjnej (MTU).
- Ruter, który otrzyma pakiet większy od MTU sieci wyjściowej dokonuje **fragmentacji** pakietu IP na kilka mniejszych (RFC791).
- Każdy z fragmentów zawiera część danych i zmodyfikowany nagłówek (pola MF, offset).



- Protokół IP określa, że składanie fragmentów datagramu przez oprogramowanie obsługujące protokół (np. u końcowego odbiorcy) wykorzystuje cztery pola do selekcji fragmentów: identification, source, destination, and protocol. W składaniu wykorzystywana jest informacja pola offset każdego fragmentu..
- Fragmenty pakietów IP mogą podlegać kolejnej fragmentacji. Minimalne MTU=68 octetów
- Każdy odbiorca sieciowy musi być przygotowany na odbiór co najmniej 576 oktetów.danych (w całości bądź we fragmentach)

Protokoły podstawowe TCP/IP : ICMP

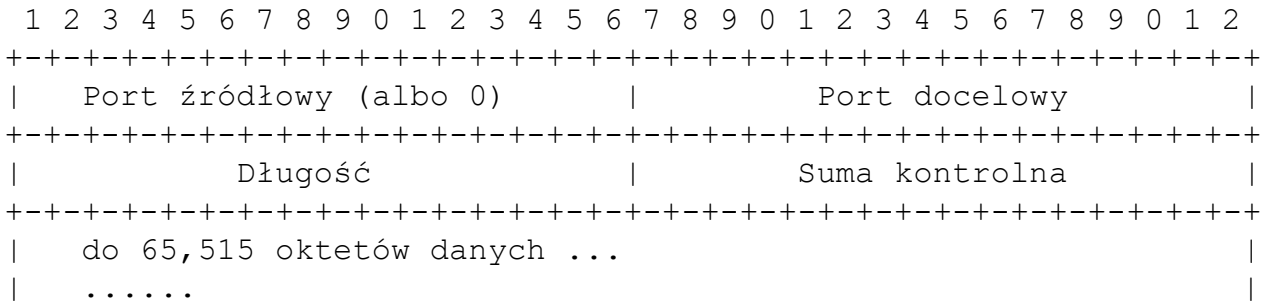
Protokół **ICMP** (RFC792) umożliwia ruterom wysyłanie komunikatów o błędach i komunikatów kontrolnych do nadawców komunikatów IP (które przyszły do rutera). Komunikaty ICMP podróżują w części ``dane'' pakietu IP, a więc nie są doręczane niezawodnie.

Najbardziej użyteczne komunikaty:

- powiadomienie o nieosiągalnym odbiorcy
- przekroczenie terminu pakietu (pole TTL=0)
- tłumienie nadawcy (element kontroli przepływu pakietów IP)
- prośba o echo i odpowiedź z echem (pakiet z odpowiedzią zawiera takie same dane jak pakiet z prośbą)
- zmień trasowanie
- prośba o czas i odpowiedź z czasem (synchronizacja zegarów, szacowanie czasu przesłania)
- problem z parametrami przysłanego pakietu IP

Protokoły podstawowe TCP/IP : UDP

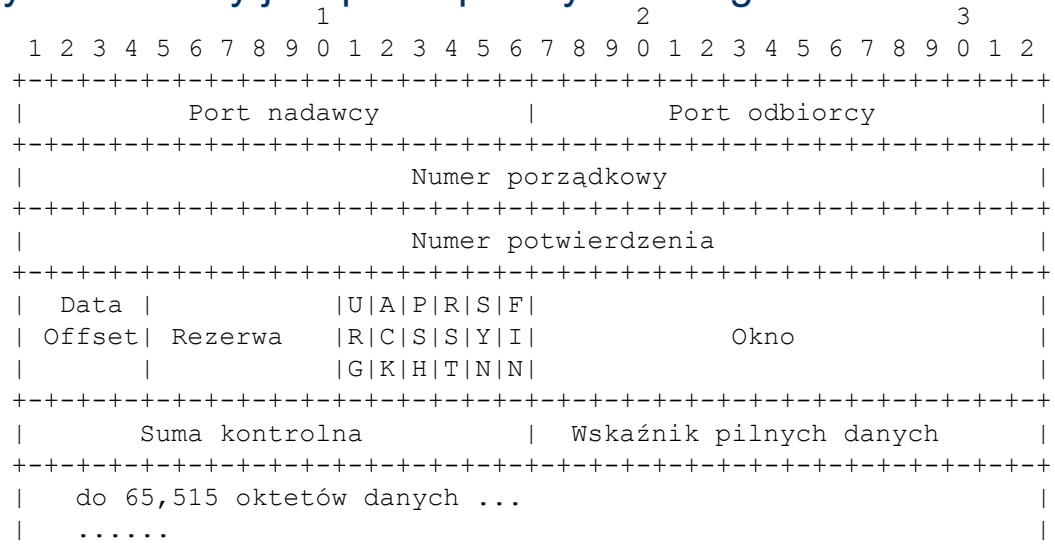
- Protokół **UDP** (RFC768) udostępnia usługę bezpołączeniową, ponieważ nie wymaga istnienia żadnego długotrwałego związku pomiędzy węzłem nadawczym i odbiorczym.
- Struktura datagramu UDP:



- Suma kontrolna jest obliczana dla bloku danych składających się z całego datagramu UDP, IP nadawcy i odbiorcy, pola typu protokołu oraz długości datagramu (patrz RFC768).

Protokoły podstawowe TCP/IP : TCP

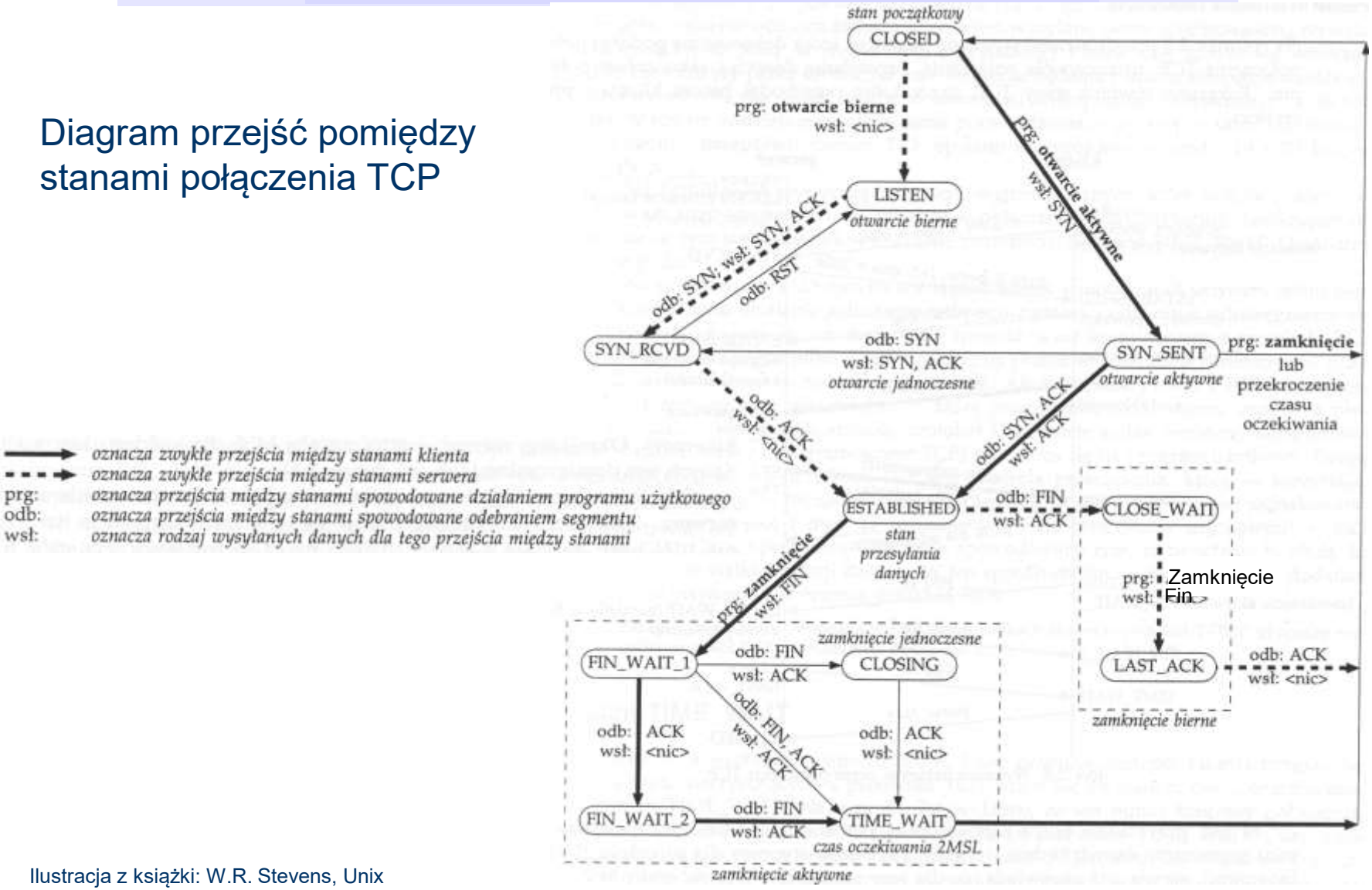
- Protokół **TCP** (RFC793, RFC1122) zapewnia:
  - Dwukierunkową, strumieniową, połączeniową usługę transportową
  - Kontrolę przepływu (oprogramowanie TCP węzła informuje partnera o tym ile bajtów chce przyjąć (okno oferowane))
  - Niezawodność: dostarczanie wszystkich wysłanych danych odbiorcy, we właściwej kolejności
- Strumień danych tworzony jest przez przesyłanie segmentów TCP:



- Dla zapewnienia niezawodnej obsługi TCP używa:
  - Sumy kontrolnej dla każdego segmentu TCP
  - Retransmisji segmentów utraconych
  - Adaptacji szybkości wysyłania do aktualnego opóźnienia sieci i chłonności odbiorcy

Protokół TCP – diagram stanów (*)

Diagram przejść pomiędzy stanami połączenia TCP



Ilustracja z książki: W.R. Stevens, Unix programowanie usług sieciowych, t.1

TCP: bufory, kontrola przepływu, okna

- Oprogramowanie TCP rezerwuje bufor na odbierane dane, a następnie wysyła drugiej stronie informację o jego rozmiarze.
- Gdy przychodzą dane, odbiorca wysyła potwierdzenie z informacją o ilości wolnego miejsca w buforze odbiorczym (tzw. rozmiar okna).
- Nadawca otrzymujący ofertę zerowego okna musi zaprzestać nadawania – aż odbiorca zaoferuje dodatnią wartość rozmiaru okna.
- Jeżeli protokół TCP wykryje zgubiony segment – zamiast próbować zapełnić bufor odbiorcy kolejnymi segmentami, wysyła pojedynczy segment z danymi. Jeżeli potwierdzenie przybędzie bez utraty segmentu – wysyła 2 segmenty z danymi, itd. Wykładniczy wzrost liczby wysyłanych segmentów trwa do osiągnięcia połowy oferowanego przez odbiorcę okna – wtedy zwalnia.
- Normalnie oprogramowanie TCP składa dane w buforze wysyłkowym, by zmniejszyć liczbę wysyłanych segmentów (algorytm Nagle'a RFC896). Przy ustawieniu sygnalizatora PSH segmentu TCP zawartość bufora jest natychmiast ekspediowana.
- Ustawienie przez aplikację sygnalizatora URG pozwala dołączyć 1bajt danych „pilnych" aplikacji do bufora wysyłkowego; w utworzonym segmencie TCP URG=1, a wskaźnik danych pilnych pokazuje pozycję następną po wstawionym bajcie.
- Opis rozszerzeń TCP dla szybkich sieci Internetowych (TCP Window Scale Option, Round-Trip Time Measurement, Protect Against Wrapped Sequence Numbers) można znaleźć w RFC1323.

Protokoły pomocnicze TCP/IP (*)

- Protokół odwzorowywania adresów **ARP** (RFC826) definiuje 2 podstawowe komunikaty:
  - Pytanie o adres IP (wysyłany na adres rozgłoszeniowy sieci)
  - Odpowiedź na pytanie (adres sprzętowy hosta o wskazanym IP) jest odsyłany na adres pytającego

- Protokół odwrotnego odwzorowywania adresów **RARP** (RFC903) służy komputerowi do uzyskiwania swojego adresu IP odpowiadającego adresowi fizycznemu (np.. 48b adres Ethernet)

- Protokół **BOOTP** (RFC951), wywoływany w czasie startu systemu operacyjnego, umożliwia pozyskiwanie od serwera adresu IP, nazwę i adres serwera, IP domyślnego rutera oraz inne dane konfigurujące inicjalizację systemu (np. informację o pliku startowym, który komputer może pobrać za pomocą protokołu TFTP). Pakiet BOOTP jest przenoszony przez datagram UDP (adres nadawcy: same 0, adres odbiorcy: same 1).

- Protokół **DHCP** (RFC2131) rozszerza możliwości BOOTP o automatyczne wynajmowanie puli adresów IP na określony czas.

IPv6

- W 1998 r. organizacja IETF zdefiniowała standard IPv6 (RFC2460), który ma zastąpić powszechnie używany IPv4.
- Podstawowe rozszerzenia
  - Adres 128-bitowy (ponad 7.9e28 razy więcej adresów niż dla IPv4)

  Zapis: 8 grup (rozdzielonych dwukropkiem) po 4 cyfry heksadecymalnie, np.:

  1234:5678:0000:0000:0000:0000:9ABC:DEF0. Short form:

  1234:5678::9ABC:DEF0
  - Maksymalny rozmiar pakietu został powiększony z (2^{16}-1) do (2^{32}-1)
  - Uproszczenie nagłówka → zmniejszenie kosztów przetwarzania nagłówków przez rutery
  - Elastyczność (nagłówki rozszerzające)
  - Możliwa rezerwacja zasobów (gwarancje przepustowości i opóźnienia) i uwierzytelnianie; IPSEC
  - Uproszczenie komunikacji multicast, zdefiniowanie zakresów dla pakietów multicast