

Informatyka 1 (ES1F1002)

Politechnika Białostocka - Wydział Elektryczny
Elektrotechnika, semestr I, studia stacjonarne I stopnia
Rok akademicki 2024/2025

Wykład nr 14 (24.01.2025)

dr inż. Jarosław Forenc

Plan wykładu nr 14

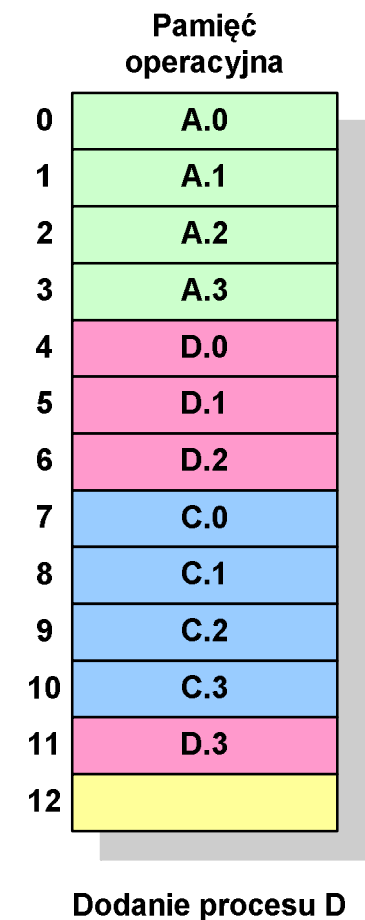
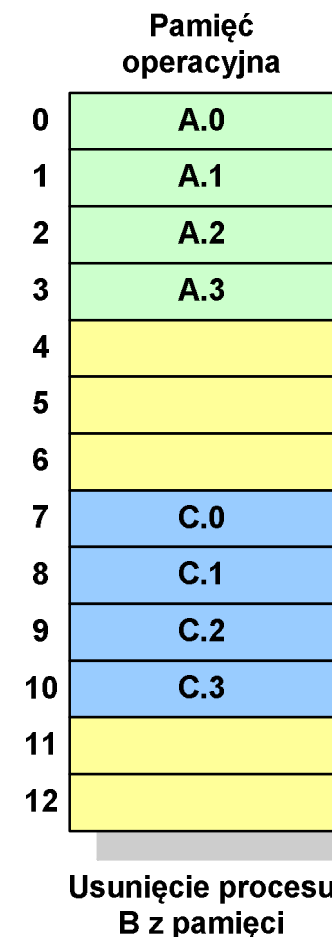
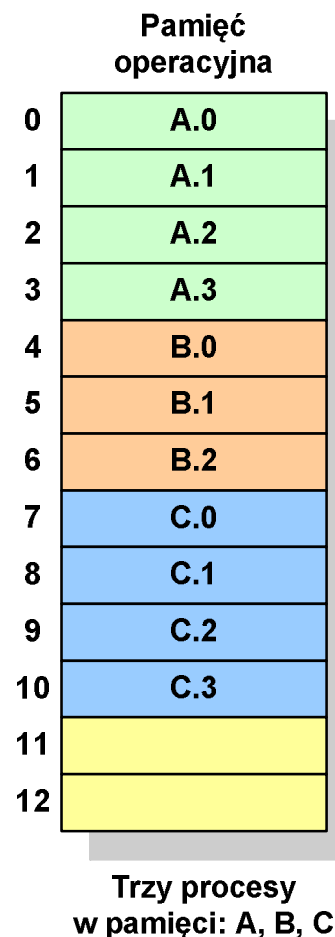
- Zarządzanie pamięcią operacyjną
 - partycjonowanie, stronicowanie, segmentacja, pamięć wirtualna
- Sieci komputerowe
 - definicja, podział, topologie i media transmisyjne
 - model referencyjny ISO/OSI, model protokołu TCP/IP
- Operacje wejścia-wyjścia w języku C, pliki
 - strumienie, typy standardowych operacji wejścia-wyjścia
 - operacje na plikach: otwarcie i zamknięcie pliku
 - format (plik) / tryb otwarcia tekstowy i binarny

Zarządzanie pamięcią

- zarządzanie pamięcią polega na wydajnym przenoszeniu programów i danych do i z pamięci operacyjnej
- w nowoczesnych wieloprogramowych systemach operacyjnych zarządzanie pamięcią opiera się na **pamięci wirtualnej**
- pamięć wirtualna bazuje na wykorzystaniu **segmentacji** i **stronicowania**
- z historycznego punktu widzenia w systemach komputerowych stosowane były/są następujące metody zarządzania pamięcią:
 - proste stronicowanie
 - prosta segmentacja
 - **stronicowanie i segmentacja pamięci wirtualnej**

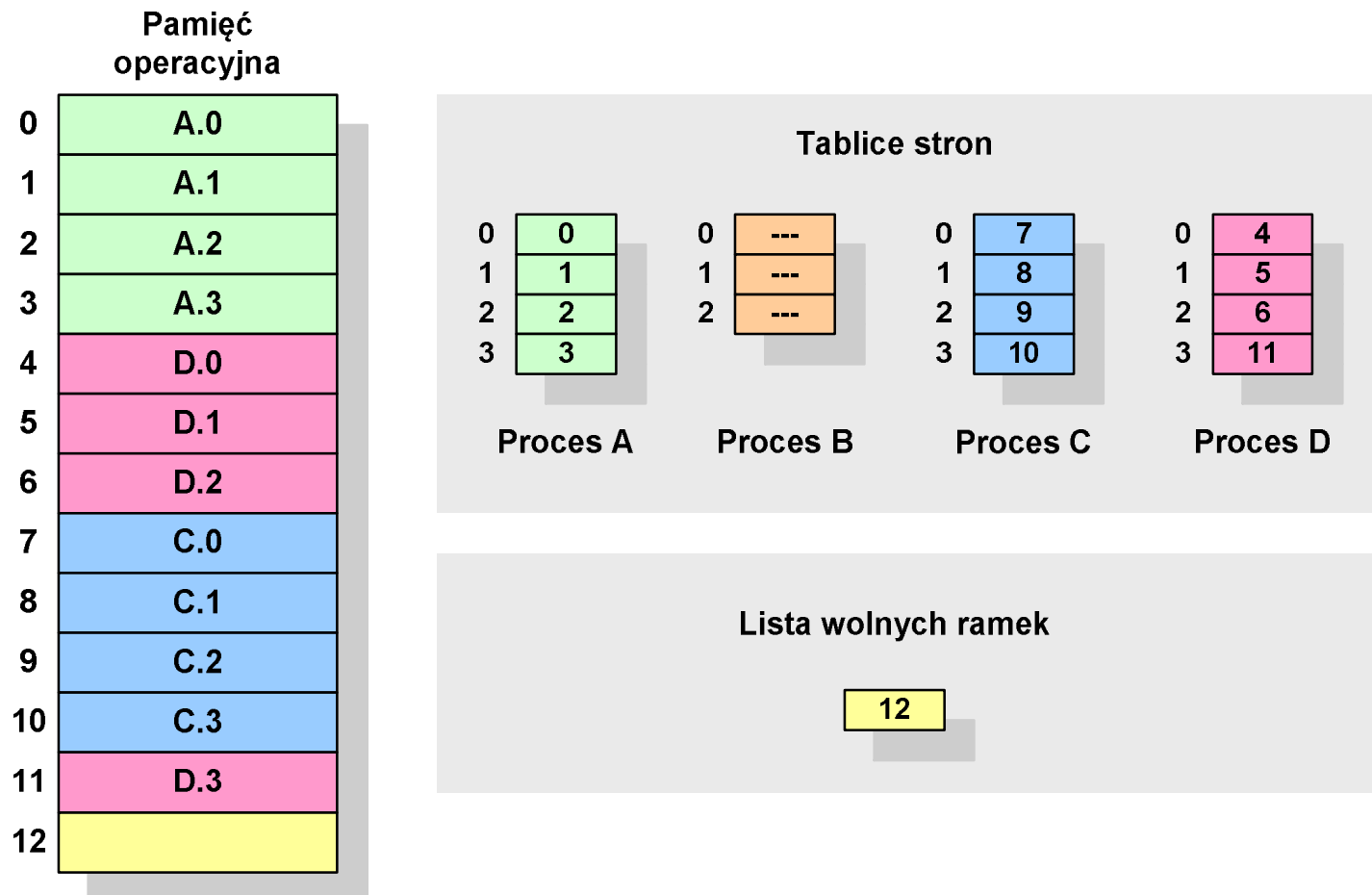
Proste stronicowanie

- ❑ pamięć operacyjna podzielona jest na jednakowe bloki o stałym niewielkim rozmiarze nazywane **ramkami** lub **ramkami stron** (page frames)
- ❑ do tych ramek wstawiane są fragmenty procesu zwane **stronami** (pages)
- ❑ aby proces mógł zostać uruchomiony wszystkie jego strony muszą znajdować się w pamięci operacyjnej



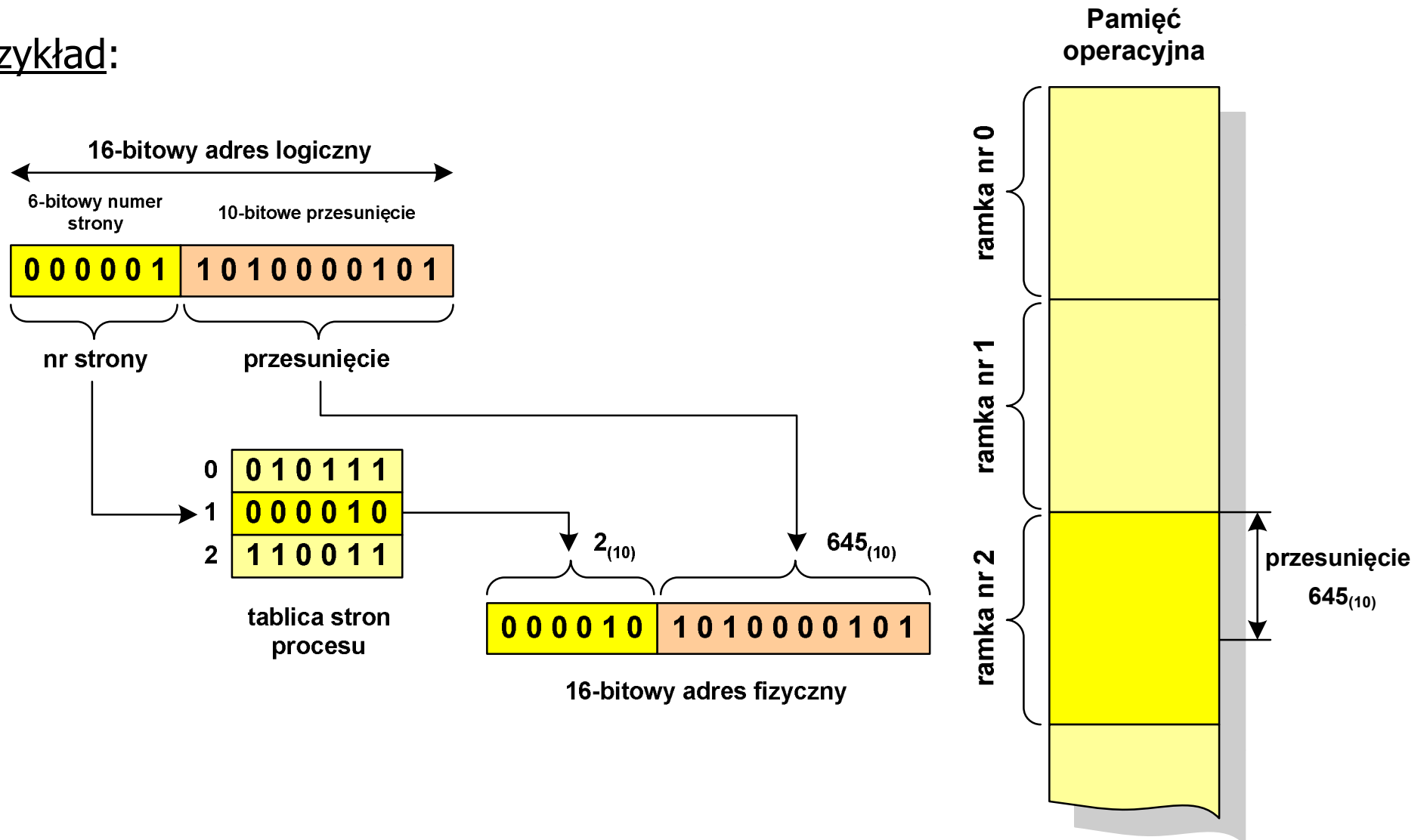
Proste stronicowanie

- dla każdego procesu przechowywana jest **tablica strony** (page table) zawierająca lokalizację ramki dla każdej strony procesu



Proste stronicowanie

Przykład:

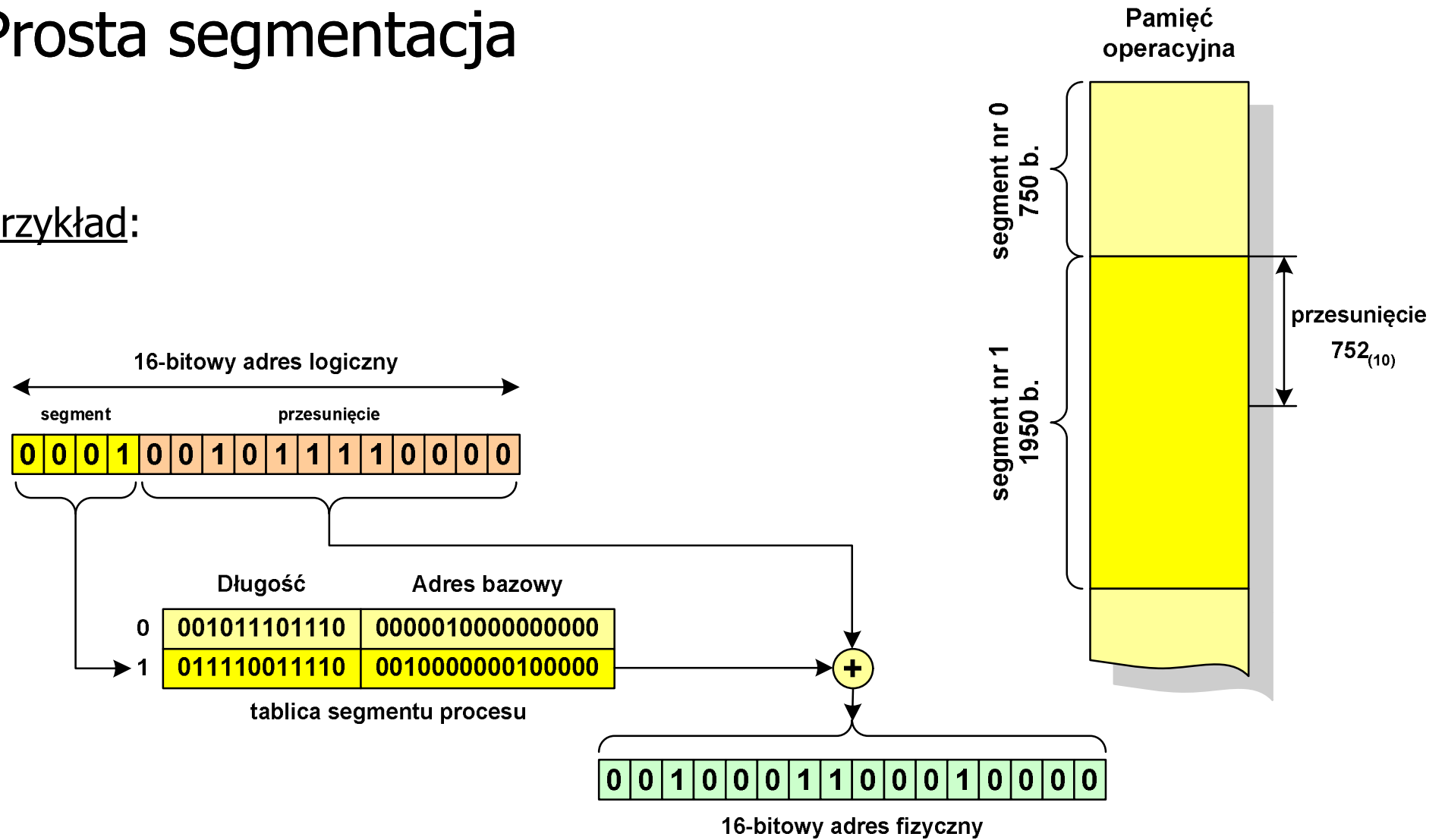


Prosta segmentacja

- polega na podzieleniu programu i skojarzonych z nim danych na odpowiednią liczbę **segmentów** o **różnej długości**
- ładowanie procesu do pamięci polega na wczytaniu wszystkich jego segmentów do partycji dynamicznych (nie muszą być ciągłe)
- segmentacja jest widoczna dla programisty i ma na celu wygodniejszą organizację programów i danych
- **adres logiczny** wykorzystujący segmentację składa się z dwóch części:
 - numeru segmentu
 - przesunięcia
- dla każdego procesu określana jest **tablica segmentu procesu** zawierająca:
 - długość danego segmentu
 - adres początkowy danego segmentu w pamięci operacyjnej

Prosta segmentacja

Przykład:



Pamięć wirtualna

- **pamięć wirtualna** umożliwia przechowywanie stron/segmentów wykonywanego procesu w pamięci dodatkowej (na dysku twardym)

Co się dzieje, gdy procesor chce odczytać stronę z pamięci dodatkowej?

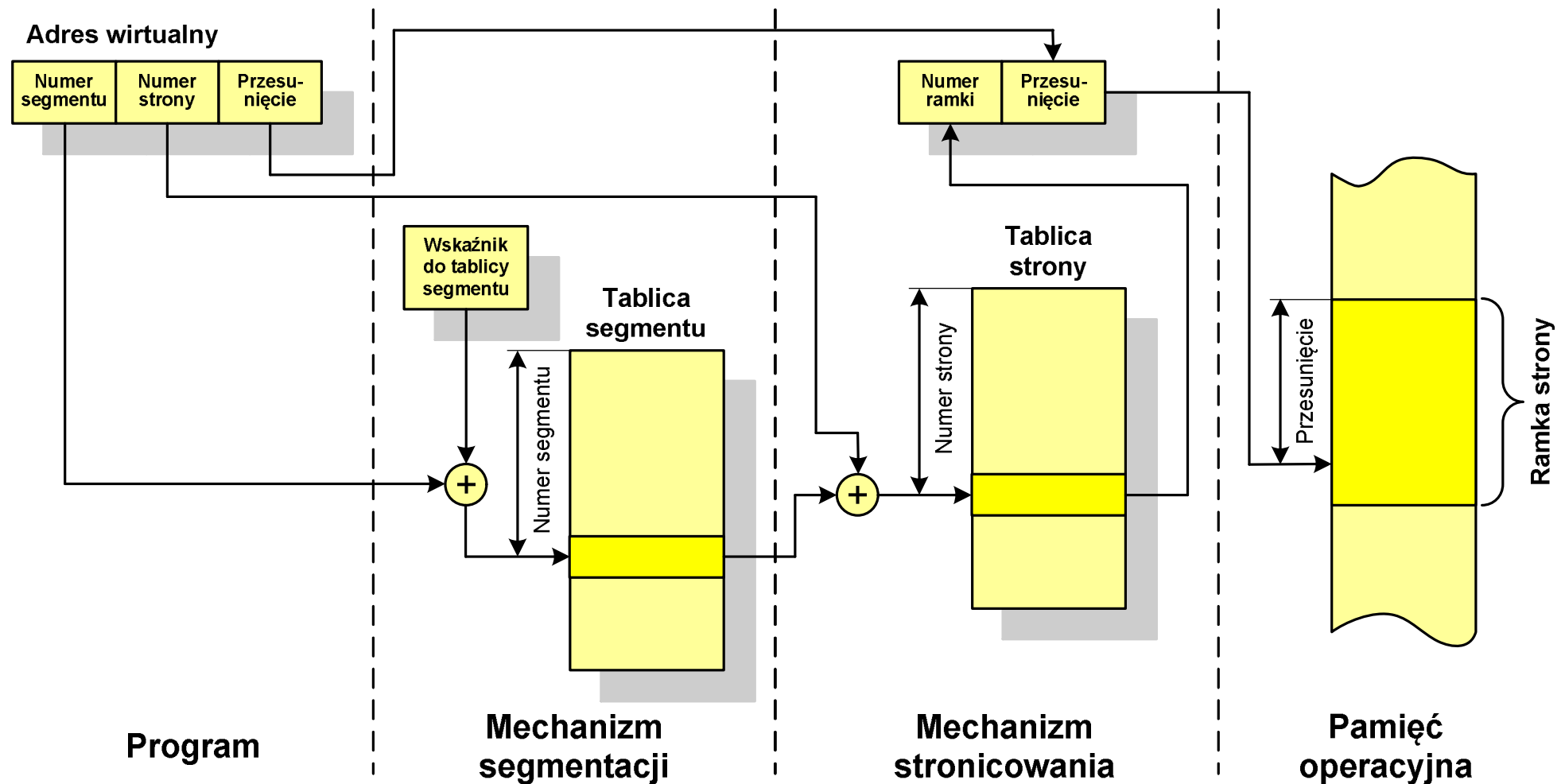
- generowanie przerwania sygnalizującego błąd w dostępie do pamięci
- zmiana stan procesu na zablokowany
- wstawienie do pamięci operacyjnej fragment procesu zawierający adres logiczny, który był przyczyną błędu
- zmiana stanu procesu na uruchomiony

Dzięki zastosowaniu pamięci wirtualnej:

- w pamięci operacyjnej może być przechowywanych więcej procesów
- proces może być większy od całej pamięci operacyjnej

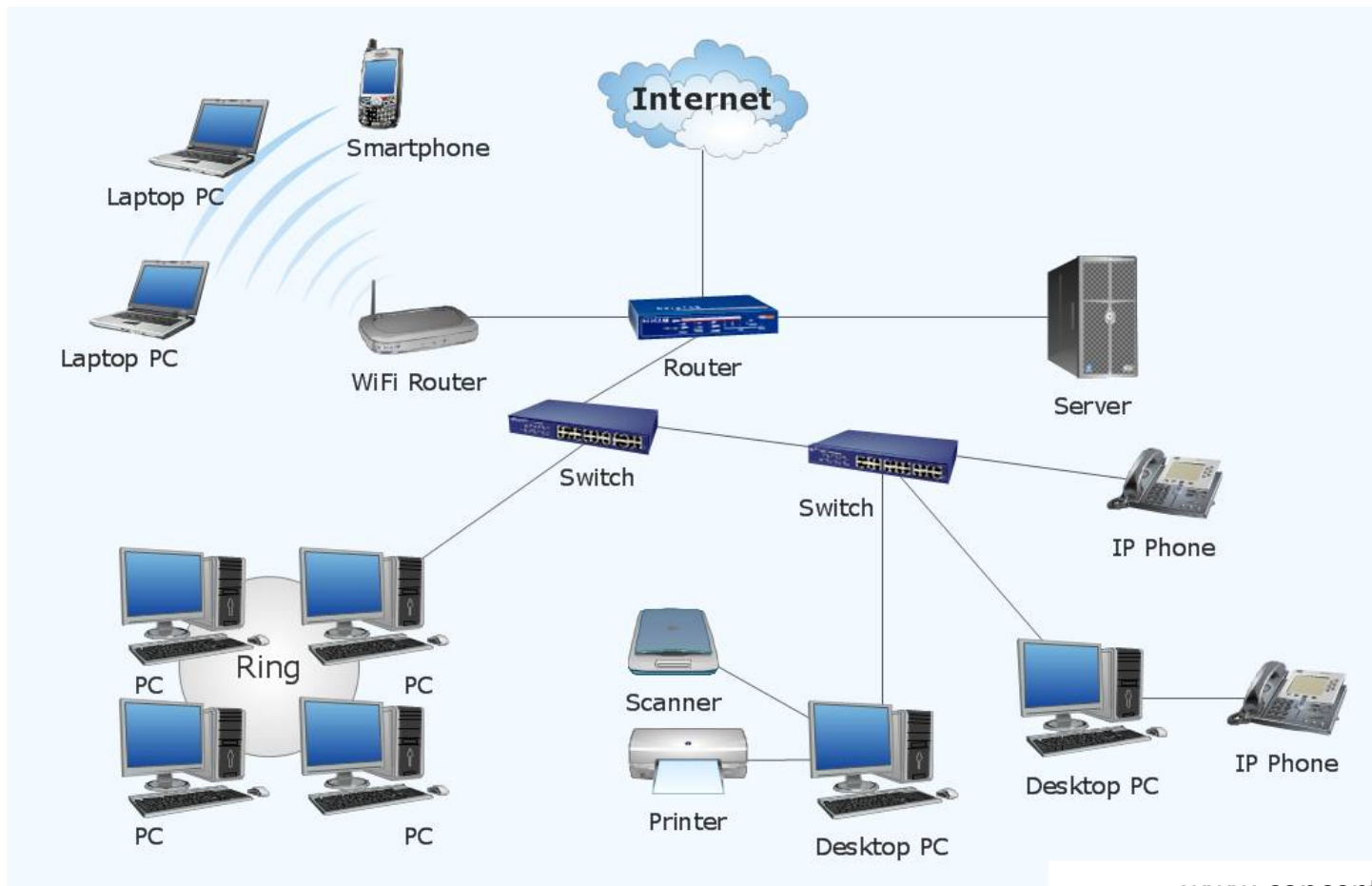
Stronicowanie i segmentacja pamięci wirtualnej

- tłumaczenie adresu wirtualnego na adres fizyczny:



Sieć komputerowa

- **Sieć komputerowa** - zbiór komputerów i innych urządzeń umożliwiających wzajemne przekazywanie informacji oraz udostępnianie zasobów



Podział sieci w zależności od ich rozmiaru

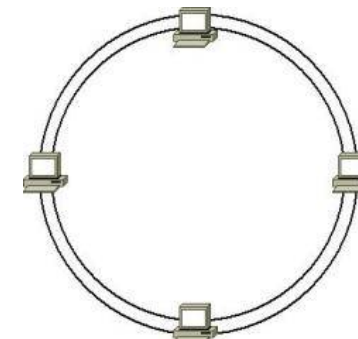
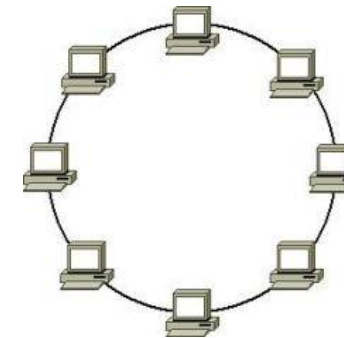
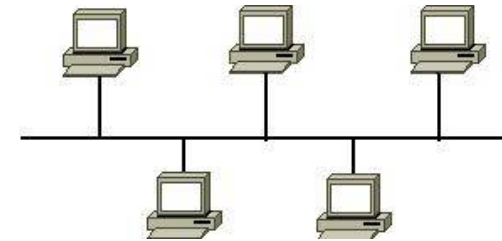
- **LAN (Local Area Network)** - sieć lokalna, łączy komputery znajdujące się na określonym, niewielkim obszarze (kilka budynków, przedsiębiorstwo), wykonana jest w jednej technologii (np. Ethernet)
- **MAN (Metropolitan Area Network)** - sieć miejska, obejmuje zasięgiem aglomerację lub miasto łącząc oddzielne sieci LAN (np. Biaman)
- **WAN (Wide Area Network)** - sieć rozległa, łączy ze sobą sieci MAN i LAN na obszarze wykraczającym poza jedno miasto (POL-34, Pionier)
- **Internet** - ogólnosiwiatowa sieć komputerowa łączące ze sobą wszystkie rodzaje sieci („sieć sieci”)
- **Intranet** - sieć podobna do Internetu, ale ograniczająca się do komputerów w firmie lub organizacji

Topologie sieci komputerowych

- **Topologia sieci** - określa strukturę sieci
 - zbiór zasad fizycznego łączenia elementów sieci (topologia fizyczna)
 - zbiór reguł komunikacji poprzez medium transmisyjne (topologia logiczna)
- **Topologia fizyczna** - opisuje sposoby fizycznego łączenia ze sobą komputerów (układ przewodów, media transmisyjne)
- **Topologia logiczna** - opisuje sposoby komunikowania się hostów za pomocą urządzeń topologii fizycznej; standardy komunikacji definiowane przez IEEE:
 - IEEE 802.3 - 10 Mb Ethernet
 - IEEE 802.3u - 100 Mb Ethernet
 - IEEE 802.3z - 1 Gb Ethernet
 - IEEE 802.5 - Token Ring
 - IEEE 802.11 - Wireless LAN
 - IEEE 802.14 - Cable Modem

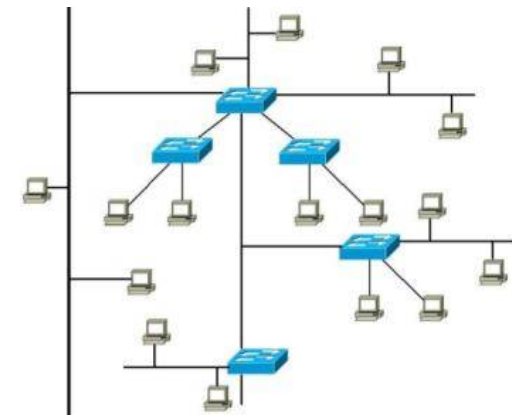
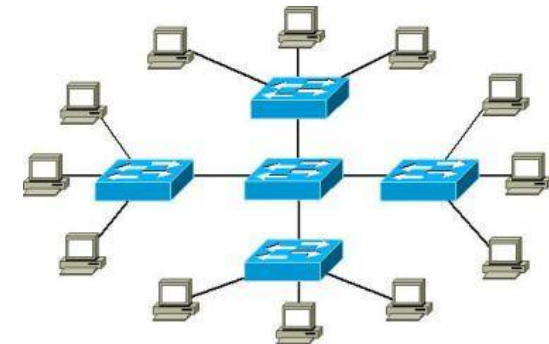
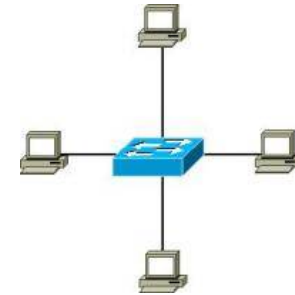
Topologie sieci komputerowych

- **topologia magistrali (bus)** - wszystkie komputery podłączone są do jednego współdzielonego medium transmisyjnego (najczęściej kabla koncentrycznego)
- **topologia pierścienia (ring)** - komputery połączone są pomiędzy sobą odcinkami kabla tworząc zamknięty pierścień (sieci światłowodowe, sieci LAN)
- **topologia podwójnego pierścienia (dual-ring)** - komputery połączone są pomiędzy sobą odcinkami kabla tworząc dwa zamknięte pierścienie (większa niezawodność, sieci: szkieletowe, MAN, Token Ring, FDDI)



Topologie sieci komputerowych

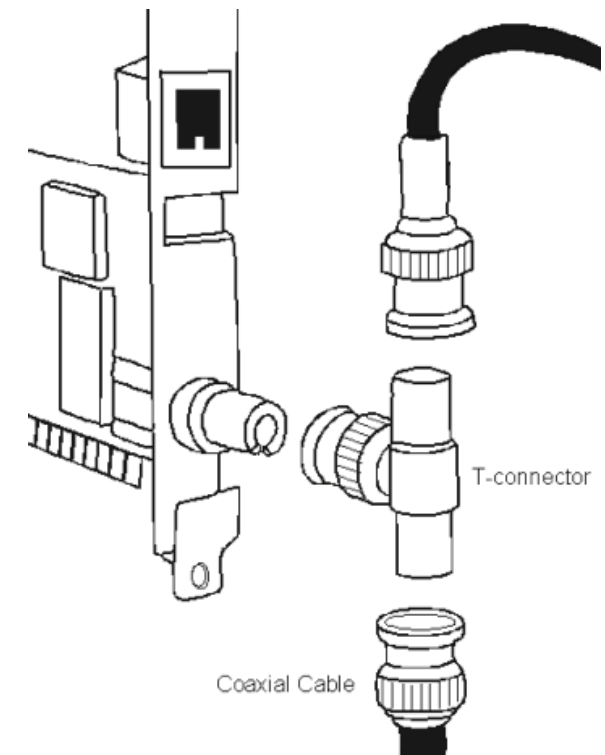
- **topologia gwiazdy (star)** - komputery podłączone są do jednego punktu centralnego (koncentrator, przełącznik), obecnie jest to najczęściej stosowana topologia sieci LAN
- **topologia rozszerzonej gwiazdy (extended star)** - posiada punkt centralny i punkty poboczne (stosowana w rozbudowanych sieciach lokalnych)
- **topologia hierarchiczna (drzewa)** - jest kombinacją topologii gwiazdy i magistrali



Media transmisyjne - przewód koncentryczny

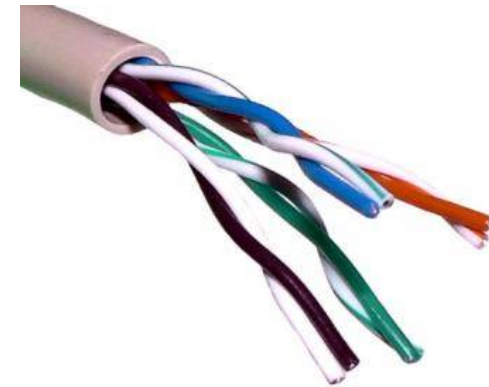
- **Ethernet gruby (Thick Ethernet)**, 10Base-5, 10 Mb/s
 - kabel RG-8 lub RG-11, impedancja falowa: 50Ω , grubość: 1/2"
 - max. odległość między stacjami: 500 m

- **Ethernet cienki (Thin Ethernet)**, 10Base-2, 10 Mb/s
 - kabel RG-58, impedancja falowa: 50Ω , grubość: 1/4"
 - max. odległość między stacjami: 185 m



Media transmisyjne - skrętka

- **Skrętka** - typ przewodu do przesyłania informacji, zbudowany z jednej lub kilku par przewodów skręconych ze sobą i umieszczonych we wspólnej izolacji
- Sposób oznaczania kabli: **xx/yyTP**
 - **xx** - sposób ekranowania całego przewodu
 - **yy** - sposób ekranowania pojedynczej pary
 - TP - Twisted Pair
- Jako **xx** i **yy** może występować:
 - **U** - nieekranowane (ang. unshielded)
 - **F** - ekranowane folią (ang. foiled)
 - **S** - ekranowane siatką (ang. shielded)
 - **SF** - ekranowane folią i siatką



U/UTP - skrętka nieekranowana (UTP)



RJ-45



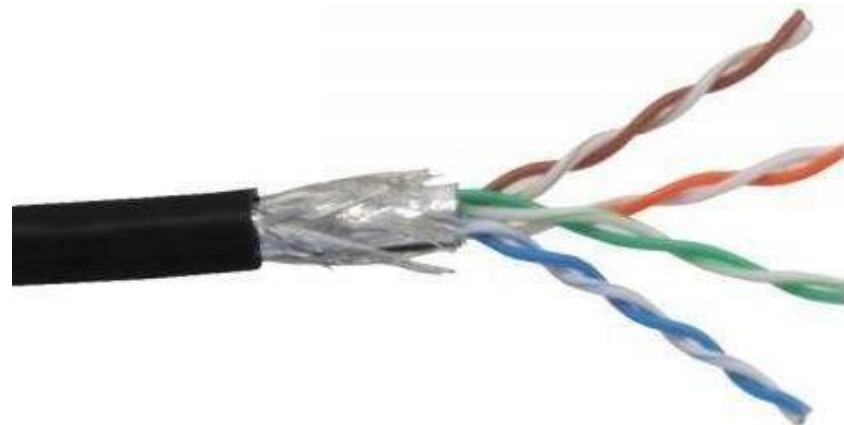
RJ-11

Media transmisyjne - skrętka

- **F/UTP** (dawniej FTP) - skrętka foliowana

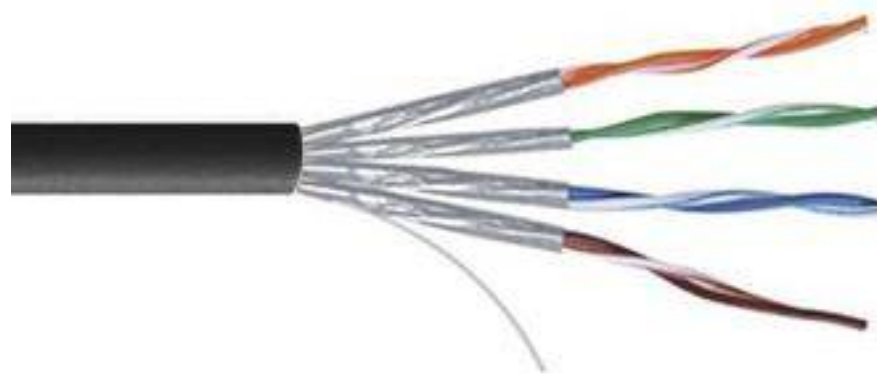


- **SF/UTP** (dawniej STP) - skrętka ekranowana folią i siatką

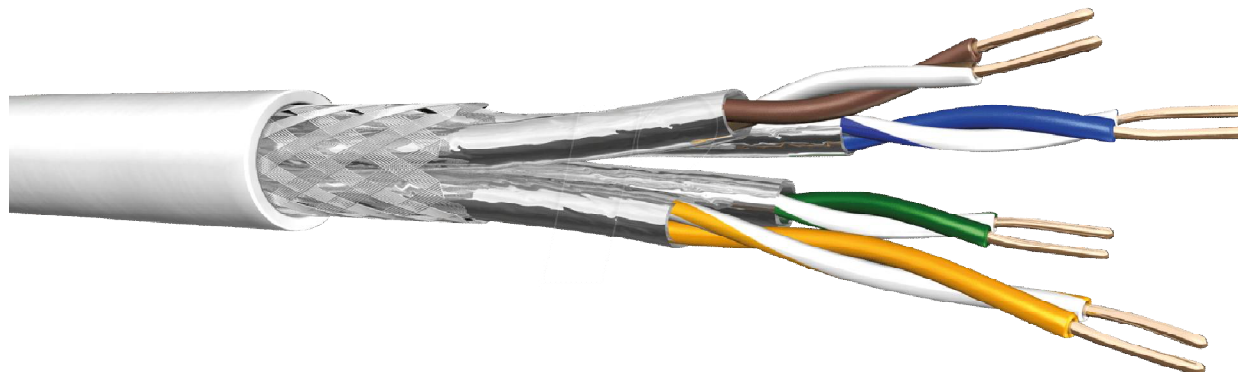


Media transmisyjne - skrętka

- **U/FTP** - skrętka z każdą parą w osobnym ekranie z folii



- **S/FTP** (dawniej SFTP) - skrętka z każdą parą foliowaną dodatkowo w ekranie z siatki



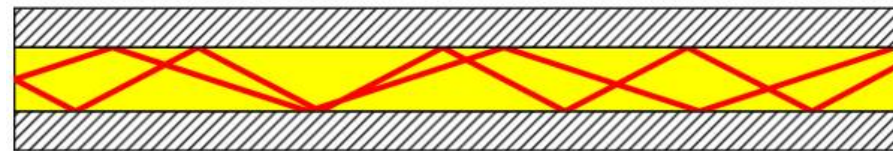
Media transmisyjne - światłowód

- **światłowód (fiber optic cable)** przesyła impulsy świetlne między nadajnikiem i odbiornikiem
- nadajnik przekształca sygnały elektryczne na świetlne, a odbiornik przekształca sygnały świetlne na elektryczne
- impulsy świetlne są przenoszone przez **włókno optyczne** składające się z dwóch rodzajów szkła o różnych współczynnikach załamania światła
- budowa światłowodu:
 - rdzeń (core), średnica: 9 μm lub 50 μm
 - płaszcz zewnętrzny (cladding), średnica: 125 μm
 - pokrycie zewnętrzne
- promień światła wędrując w rdzeniu pada na płaszcz pod pewnym kątem i następuje **zjawisko całkowitego odbicia wewnętrznego światła** - umożliwia to transmisję strumienia światła przez włókno



Media transmisyjne - światłowody wielomodowe

- w światłowodzie **wielomodowym** (**multi mode fiber**) promień światła może zostać wprowadzony pod różnymi kątami - modami
- fala świetlna o takiej samej długości może rozchodzić się wieloma drogami



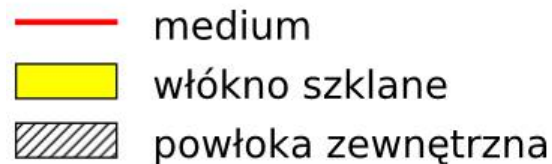
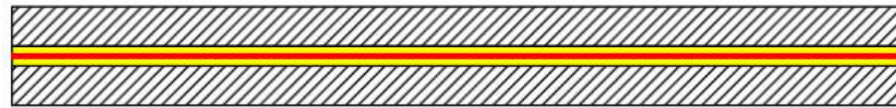
- medium
- włókno szklane
- ▨ powłoka zewnętrzna

pl.wikipedia.org

- źródło światła: diody LED
- długość fali świetlnej (850 nm i 1300 nm)
- ze względu na dyspersję maksymalna długość kabla to 5 km

Media transmisyjne - światłowody jednomodowe

- w światłowodzie **jednomodowym** (**single mode fiber**) propaguje tylko jeden mod



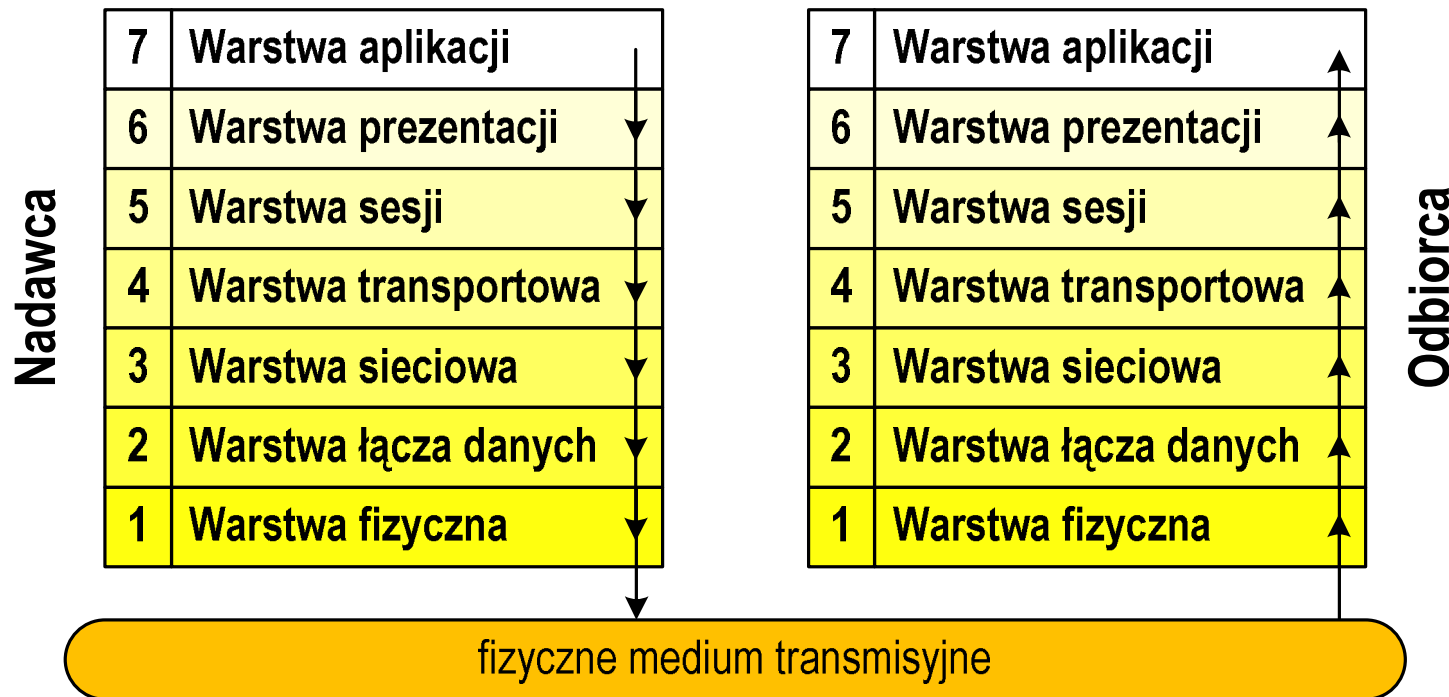
pl.wikipedia.org

- źródło światła: dioda laserowa
- długość fali świetlnej (1300 nm i 1500 nm)
- długość kabla: do 100 km
- wyższy koszt od światłowodów wielomodowych

Model ISO/OSI

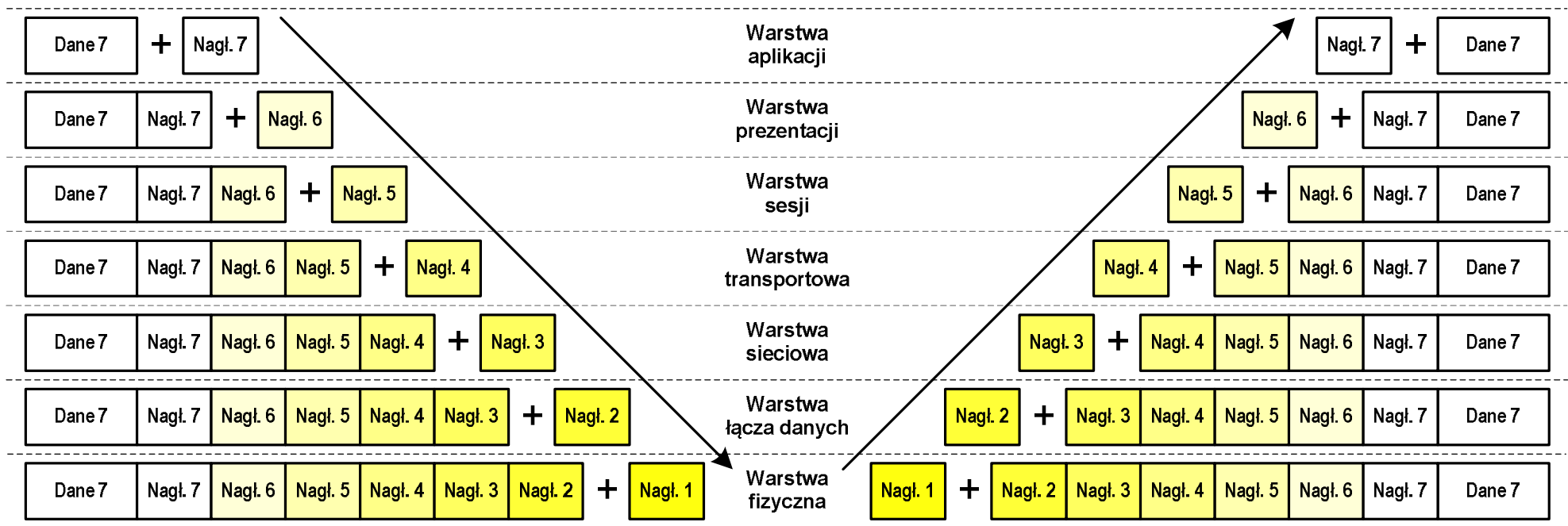
- w latach 70-tych nie istniały ogólne standardy dotyczące sieci komputerowych - każdy producent tworzył własną sieć
- w 1984 roku Międzynarodowa Organizacja Normalizacyjna (ISO) przyjęła model sieciowy, dzięki któremu producenci mogliby opracowywać współpracujące ze sobą rozwiązania sieciowe
- **ISO OSI RM - ISO Open Systems Interconnection Reference Model**
- głównym założeniem modelu jest podział systemów sieciowych na współpracujące ze sobą **7 warstw** (layers)
- struktura tworzona przez warstwy nazywana jest **stosem** protokołu wymiany danych

Model ISO/OSI



- ❑ wierzchołek stosu odpowiada usługom świadczonym bezpośrednio użytkownikowi przez aplikacje sieciowe, zaś dół odpowiada sprzętowi realizującemu transmisję sygnałów
- ❑ dane przekazywane są od wierzchołka stosu nadawcy przez kolejne warstwy, aż do warstwy pierwszej, która przesyła je do odbiorcy

Model ISO/OSI



- przy przechodzeniu do warstwy niższej, warstwa dokleja do otrzymanych przez siebie danych nagłówek z informacjami dla swojego odpowiednika na odległym komputerze (odbiorcy)
- warstwa na odległym komputerze interpretuje nagłówek i jeśli trzeba przekazać dane wyżej - usuwa nagłówek i przekazuje dane dalej

Model ISO/OSI a model TCP/IP

- w przypadku protokołu TCP/IP tworzącego Internet stosuje się uproszczony model czterowarstwowy

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

Model ISO/OSI

Warstwa aplikacji	4
Warstwa transportowa	3
Warstwa Internetu	2
Warstwa dostępu do sieci	1

Model TCP/IP

Model TCP/IP

- z poszczególnymi warstwami związanych jest wiele **protokołów**
- protokół** - zbiór zasad określających format i sposób przesyłania danych

4	Warstwa aplikacji
3	Warstwa transportowa
2	Warstwa Internetu
1	Warstwa dostępu do sieci

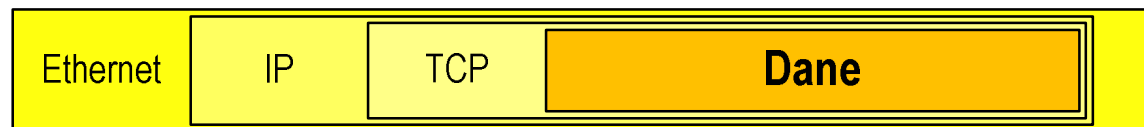
Model TCP/IP

DNS SNMP syslog	Telnet, SSH, FTP, SMTP, HTTP, POP, IMAP	
UDP	TCP	
IP		ICMP
ARP, RARP	PPP SLIP	...
IEEE 802.3		

Wybrane protokoły

UDP	TCP
wiadomość	strumień
pakiet	segment
datagram	
ramka	

Nazwy danych



Warstwa dostępu do sieci

- standard **IEEE 802.3 (Ethernet)** - 1985 r.
- dane przesyłane w postaci ramek Ethernet, format ramki Ethernet II (DIX):

8B	6B	6B	2B	46 - 1500 B	4B
Preambuła	Adres docelowy	Adres źródłowy	Typ	Dane	FCS

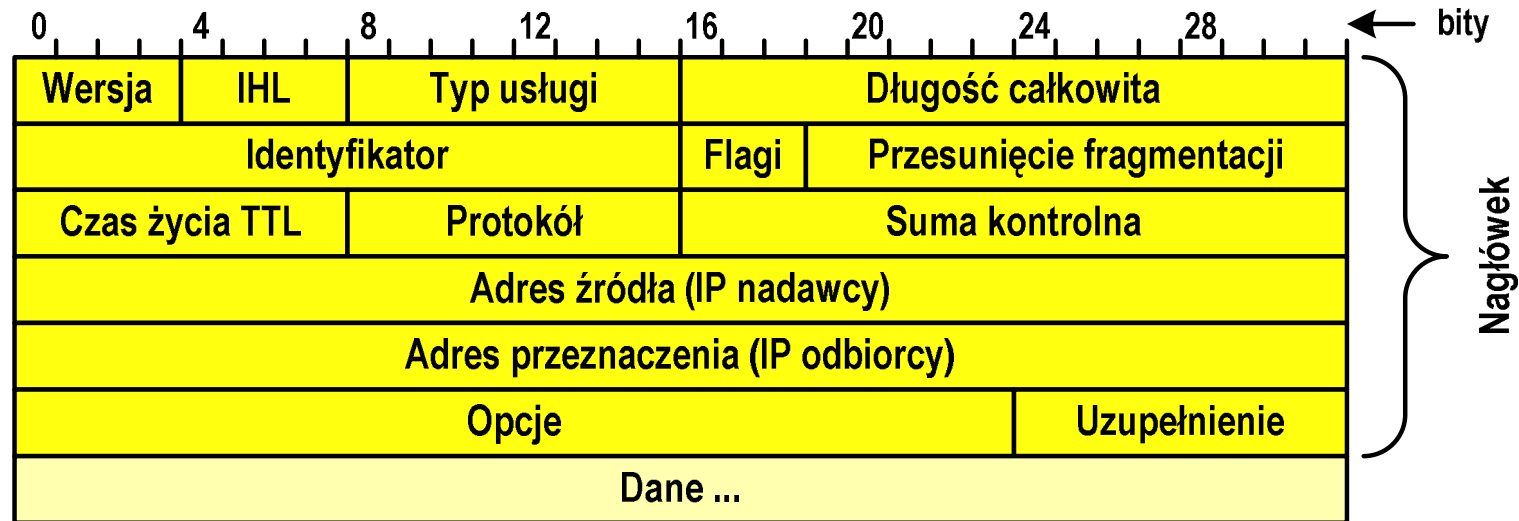
- **Typ** - numer protokołu warstwy wyższej, która odbierze dane po zakończeniu obróbki przez standard Ethernet
- **Dane** - przesyłane dane, jeśli ilość danych jest mniejsza od 46 bajtów, wprowadzane jest uzupełnienie jedynekami (bitowo)
- **FCS (Frame Check Sequence)** - 4 bajty kontrolne (CRC - Cyclic Redundancy Check) wygenerowane przez interfejs nadający i sprawdzane przez odbierający

Warstwa Internetu

- najważniejsza część Internetu to protokół **IP (Internet Protocol)**:
 - definiuje format i znaczenie pól **datagramu** IP
 - określa schemat adresowania stosowany w Internecie
 - zapewnia wybór trasy przesyłania datagramu (routing)
 - zapewnia podział danych na fragmenty i łączenie ich w całość w przypadku sieci nie akceptujących rozmiaru przenoszonych danych

- cechy protokołu:
 - **bezpołączeniowy** - nie ustanawia połączenia i nie sprawdza gotowości odbiorcy danych
 - **niepewny** - nie zapewnia korekcji i wykrywania błędów transmisji

Warstwa Internetu - datagram IP



- **Wersja (Version)** - numer wersji protokołu IP (IPv4, nowsza - IPv6)
- **Identyfikator (Identification), Flagi (Flags), Przesunięcie fragmentacji (Fragment offset)** - pola używane w przypadku podziału datagramu na części (fragmenty)
- **Adres źródła (Source Address)** - adres IP źródła danych
- **Adres przeznaczenia (Destination Address)** - adres IP odbiorcy danych

Warstwa Internetu - adresy IP

- adres IP komputera zajmuje 4 bajty (32-bitowa liczba całkowita)
- najczęściej zapisywany jest w postaci 4 liczb z zakresu od 0 do 255 każda, oddzielonych kropkami, np.

213 . 33 . 95 . 114

11010100 . 00100001 . 01011111 . 01110010

- adres składa się z dwóch części:
 - identyfikującej daną sieć w Internecie
 - identyfikującej konkretny komputer w tej sieci
- do roku 1997 wyróżnienie części określającej sieć i komputer w sieci następowało na podstawie tzw. **klas adresów IP**

Warstwa Internetu - klasy adresów IP

Klasa A	0nnnnnnn . hhhhhhhh . hhhhhhhh . hhhhhhhh sieć (max. 126) komputer (max. 16 777 214)	Zakres IP od: 1.0.0.0 do: 126.255.255.255
Klasa B	10nnnnnn . nnnnnnnn . hhhhhhhh . hhhhhhhh sieć (max. 16 382) komputer (max. 65 534)	Zakres IP od: 128.1.0.0 do: 191.255.255.255
Klasa C	110nnnnn . nnnnnnnn . nnnnnnnn . hhhhhhhh sieć (max. 2 097 150) komputer (max. 254)	Zakres IP od: 192.0.0.0 do: 223.255.255.255
Klasa D	1110xxxx . xxxxxxxx . xxxxxxxx . xxxxxxxx multicast - adresy transmisji grupowej, np. wideokonferencje	Zakres IP od: 224.0.0.0 do: 239.255.255.255
Klasa E	1111xxxx . xxxxxxxx . xxxxxxxx . xxxxxxxx zarezerwowane na potrzeby badawcze	Zakres IP od: 240.0.0.0 do: 255.255.255.255

Warstwa Internetu - maska sieci

- klasy adresów IP zostały zastąpione **bezklasowym routowaniem międzydomenowym** CIDR (Classless Inter-Domain Routing)
- do określenia liczby bitów odpowiadających sieci i liczby bitów odpowiadających hostowi stosowana jest **maska sieci**

IP: 212.33.95.114

11010100.00100001.01011111.01110010

Maska: 255.255.255.192

11111111.11111111.11111111.11000000

Adres sieci: 212.33.95.64

11010100.00100001.01011111.01000000

Broadcast: 212.33.95.127

11010100.00100001.01011111.01111111

Pierwszy host: 212.33.95.65

11010100.00100001.01011111.01000001

Ostatni host: 212.33.95.126

11010100.00100001.01011111.01111110

Warstwa Internetu - adresy IP

□ adresy specjalne

0.0.0.0

- adres sieci dla całego Internetu

255.255.255.255

- adres rozgłoszeniowy dla całego Internetu

127.0.0.1

- adres pętli (loop-back address) - stosowany do komunikacji z lokalnym komputerem (localhost)

□ adresy prywatne (nierutowalne) - nie są przekazywane przez routery

10.0.0.0 – 10.255.255.255

- klasa A

172.16.0.0 – 172.31.255.255

- klasa B

192.168.0.0 – 192.168.255.255

- klasa C

Warstwa transportowa - porty

- protokoły warstwy transportowej zapewniają dostarczenie danych do **konkretnych aplikacji** (procesów) w odpowiedniej kolejności i formie
- identyfikacja przynależności danej transmisji do procesu odbywa się na podstawie **numeru poru** (liczba 16-bitowa, zakres: **0 ÷ 65535**)
- numery portów przydzielane są przez organizację **IANA** (Internet Assigned Numbers Authority):
 - **0 ÷ 1023** - zakres zarezerwowany dla tzw. **dobrze znanych portów** (well-know port number)
 - **1024 ÷ 49151** - porty zarejestrowane (registered)
 - **49152 ÷ 65535** - porty dynamiczne/prywatne (dynamic/private)
- połączenie numeru IP komputera i portu, na którym odbywa się komunikacja, nazywa się **gniazdem** (socket)

Warstwa transportowa - porty

- wybrane dobrze znane porty:

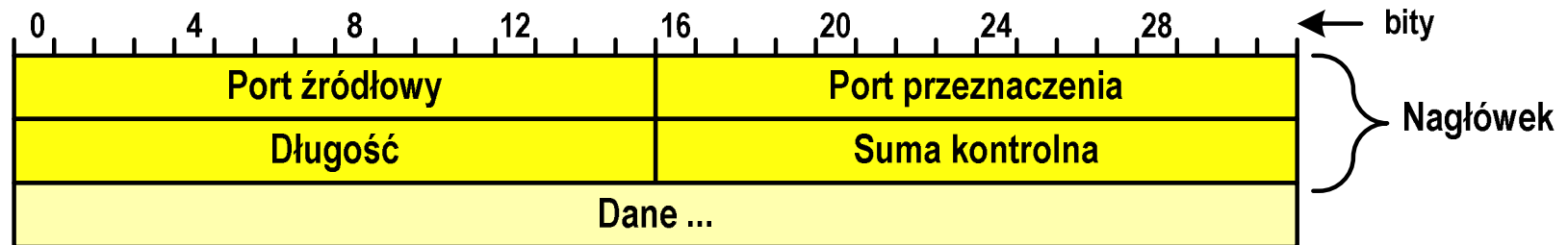
port	protokół
20	FTP (dane)
21	FTP (polecenia)
22	SSH
23	Telnet
25	SMTP (mail)

port	protokół
53	DNS
80	HTTP (www)
110	POP3 (mail)
119	NNTP (news)
143	IMAP (mail)

- w warstwie transportowej funkcjonują dwa podstawowe protokoły:
 - **UDP** (User Datagram Protocol)
 - **TCP** (Transmission Control Protocol)

Warstwa transportowa - protokół UDP

- UDP wykonuje usługę **bezpołączeniowego** dostarczania datagramów:
 - nie ustanawia połączenia
 - nie sprawdza gotowości odbiorcy do odebrania przesyłanych danych
 - nie sprawdza poprawności dostarczenia danych
- jednostką przesyłanych danych jest **paket**

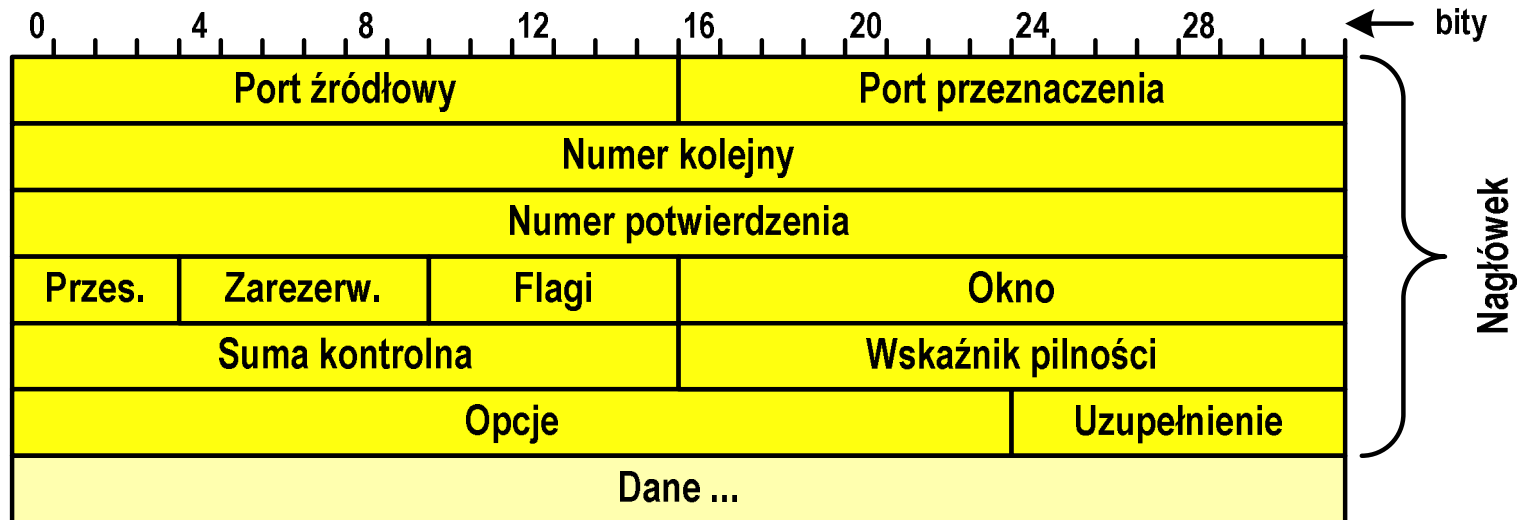


- **Port źródłowy** (Source port) - numer portu nadawcy
- **Port przeznaczenia** (Destination port) - numer portu odbiorcy
- **Długość** (Length) - całkowita długość pakietu w bajtach (nagłówek + dane)
- **Suma kontrolna** (Checksum) - tworzona na podstawie nagłówka i danych

Warstwa transportowa - protokoły UDP i TCP

- **UDP** stosowany jest, gdy ilość przesyłanych danych w pakiecie jest niewielka
- pakiet **UDP** zawiera bardzo mało informacji kontrolnych, zatem opłacalne jest jego stosowanie w powiązaniu z aplikacjami samodzielnie dbającymi o kontrolę poprawności transmisji
- **TCP** (Transmission Control Protocol) jest protokołem **niezawodnym** i **połączeniowym**, działa na strumieniach bajtów
- **TCP** sprawdza czy dane zostały dostarczone poprawnie i w określonej kolejności
- jednostką przesyłanych danych stosowaną przez TCP jest **segment**

Warstwa transportowa - segment TCP



- Port źródłowy (Source port) - numer portu nadawcy
- Port przeznaczenia (Destination port) - numer portu odbiorcy
- Numer kolejny (Sequence number) - identyfikator określający miejsce segmentu przed fragmentacją
- Numer potwierdzenia (Acknowledgment number) - identyfikator będący potwierdzeniem otrzymania danych przez odbiorcę

Warstwa aplikacji

- zawiera szereg procesów (usług, protokołów) wykorzystywanych przez uruchamiane przez użytkownika aplikacje do przesyłania danych
- większość usług działa w architekturze **klient-serwer** (na odległym komputerze musi być uruchomiony serwer danej usługi)

DNS (Domain Name System)

- świadczy usługi zamieniania (rozwiązywania) nazwy komputera na jego adres IP

we.pb.edu.pl



213.33.95.2

- wykorzystuje port o numerze 53
- przekształcone nazwy przechowywane są także na komputerze osobistym:
ipconfig /displaydns

Operacje wejścia-wyjścia w języku C

- Operacje wejścia-wyjścia nie są elementami języka C
- Zostały zrealizowane jako funkcje zewnętrzne, znajdujące się w bibliotekach dostarczanych wraz z kompilatorem
- **Standardowe** wejście-wyjście (strumieniowe)
 - plik nagłówkowy **stdio.h**
 - duża liczba funkcji, proste w użyciu
 - ukrywa przed programistą szczegóły wykonywanych operacji
- **Systemowe** wejście-wyjście (deskryptorowe, niskopoziomowe)
 - plik nagłówkowy **io.h**
 - mniejsza liczba funkcji
 - programista sam obsługuje szczegóły wykonywanych operacji
 - funkcje bardziej zbliżone do systemu operacyjnego - działają szybciej

Strumienie

- Standardowe operacje wejścia-wyjścia opierają się na **strumieniach** (ang. **stream**)
- Strumień jest pojęciem abstrakcyjnym - jego nazwa bierze się z analogii między przepływem danych, a np. wody
- W strumieniu dane płyną od źródła do odbiorcy
- Strumień może być skojarzony ze zbiorem danych znajdujących się na dysku (plik) lub zbiorem danych pochodzących z urządzenia znakowego (klawiatura)
- Strumienie reprezentowane są przez zmienne będące wskaźnikami na struktury typu **FILE** (definicja w pliku **stdio.h**)
- Podczas pisania programów nie ma potrzeby bezpośredniego odwoływania się do pól tej struktury

Strumienie

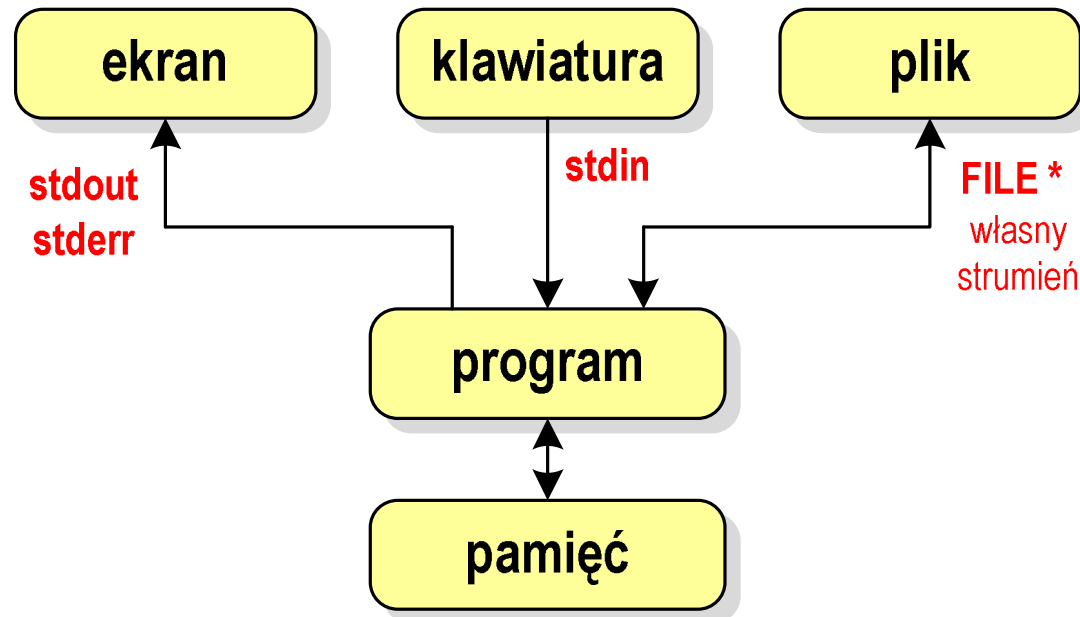
- W każdym programie automatycznie tworzone są i otwierane trzy standardowe strumienie wejścia-wyjścia:
 - **stdin** - standardowe wejście, skojarzone z klawiaturą
 - **stdout** - standardowe wyjście, skojarzone z ekranem monitora
 - **stderr** - standardowe wyjście dla komunikatów o błędach, skojarzone z ekranem monitora

```
_CRTIMP FILE * __cdecl __iob_func(void);  
  
#define stdin (&__iob_func()[0])  
#define stdout (&__iob_func()[1])  
#define stderr (&__iob_func()[2])
```

- Funkcja **printf()** niejawnie używa strumienia **stdout**
- Funkcja **scanf()** niejawnie używa strumienia **stdin**

Strumienie

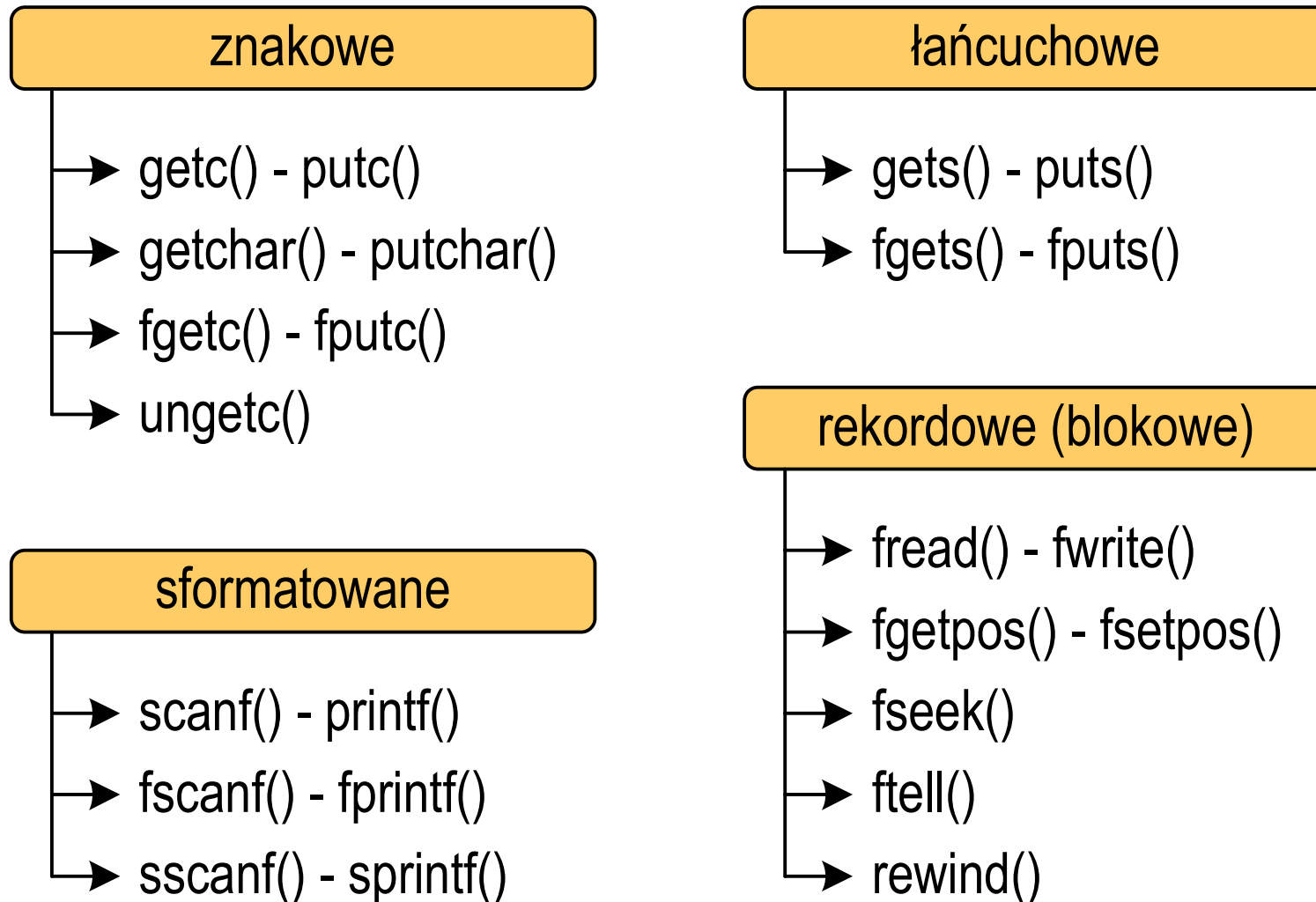
■ Współpraca programu z „otoczeniem”



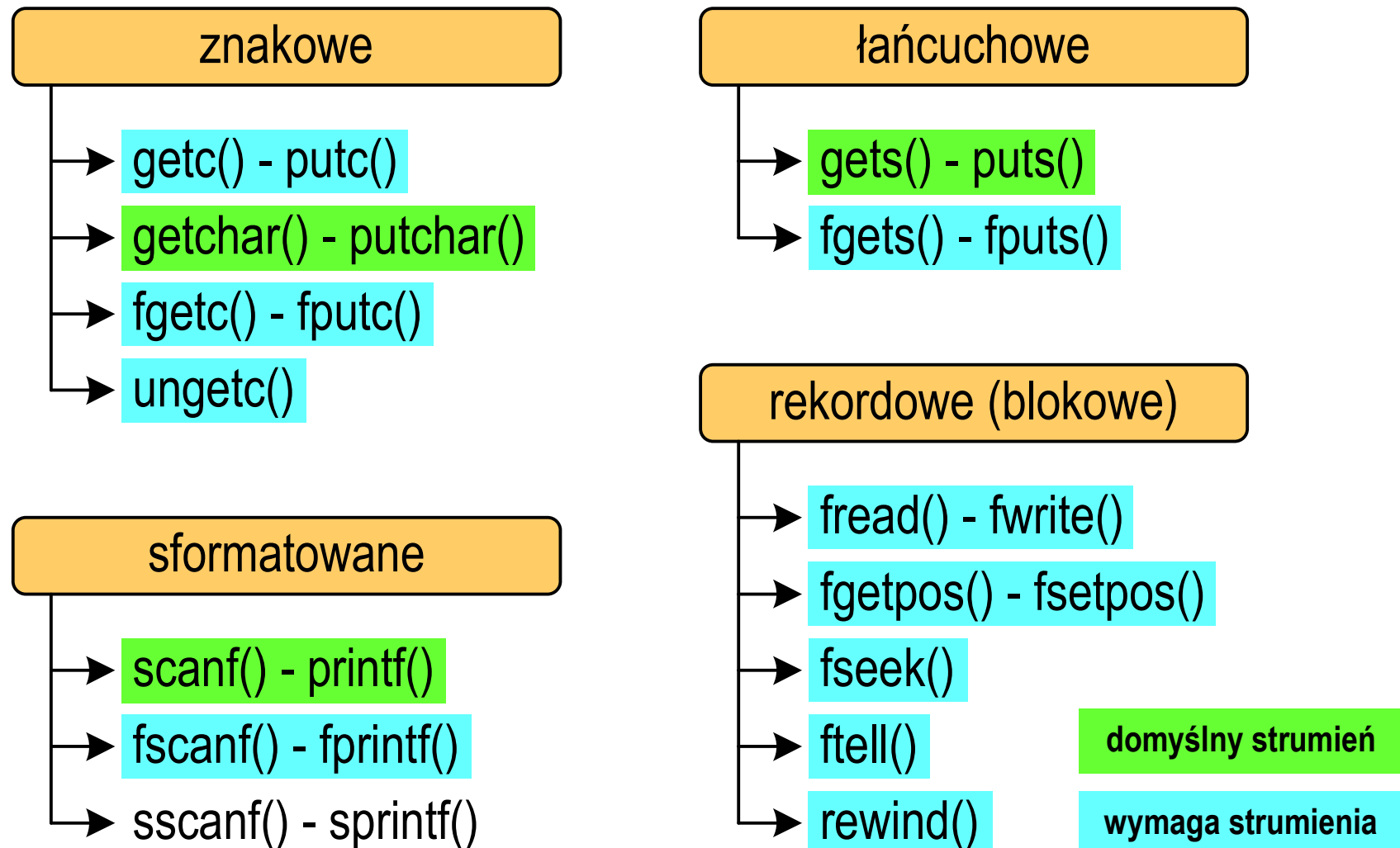
■ Standardowe funkcje wejścia-wyjścia mogą:

- domyślnie korzystać z określonego strumienia (**stdin**, **stdout**, **stderr**)
- wymagać podania strumienia (własnego, **stdin**, **stdout**, **stderr**)

Typy standardowych operacji wejścia-wyjścia



Typy standardowych operacji wejścia-wyjścia



Operacje na plikach

- Strumień wiąże się z plikiem za pomocą **otwarcia**, zaś połączenie to jest przerywane przez **zamknięcie** strumienia
- Operacje związane z przetwarzaniem pliku zazwyczaj składają się z trzech części

1. Otwarcie pliku (strumienia):

- funkcje: **fopen()**

2. Operacje na pliku (strumieniu), np. czytanie, pisanie:

- funkcje dla plików tekstowych: **fprintf(), fscanf(), fgetc(), fputc(), fgets(), fputs()...**

- funkcje dla plików binarnych: **fread(), fwrite(), ...**

3. Zamknięcie pliku (strumienia):

- funkcja: **fclose()**

Otwarcie pliku - fopen()

```
FILE* fopen(const char *fname, const char *mode);
```

- **fname** - nazwa pliku, może zawierać całą ścieżkę dostępu do pliku
- **mode** - tryb otwarcia:
 - **"r"** - odczyt
 - **"w"** - zapis - jeśli pliku nie ma to zostanie on utworzony, jeśli plik istnieje, to jego poprzednia zawartość zostanie usunięta
 - **"a"** - zapis (dopisywanie) - dopisywanie danych na końcu istniejącego pliku, jeśli pliku nie ma to zostanie utworzony
 - **"t"** - otwarcie w trybie tekstowym (domyślnie)
 - **"b"** - otwarcie w trybie binarnym
- **fopen()** zwraca wskaźnik na strukturę **FILE** skojarzoną z otwartym plikiem lub **NULL**, gdy otwarcie nie powiodło się

Otwarcie pliku - fopen()

- Otwarcie pliku w trybie tekstowym, tylko odczyt

```
FILE *fp;  
fp = fopen("dane.txt", "r");
```

- Otwarcie pliku w trybie binarnym, tylko zapis

```
fp = fopen("c:\\baza\\data.bin", "wb");
```

- Otwarcie pliku w trybie tekstowym, tylko zapis

```
fp = fopen("wynik.txt", "wt");
```

Zamknięcie pliku - fclose()

FCLOSE

stdio.h

```
int fclose(FILE *fp);
```

- Zamyka plik wskazywany przez **fp**
- Zwraca **0 (zero)** jeśli zamknięcie pliku było pomyślne
- W przypadku wystąpienia błędu zwraca **EOF**

```
#define EOF (-1)
```

- Po zamknięciu pliku, wskaźnik **fp** może być wykorzystany do otwarcia innego pliku
- W programie może być jednocześnie otwartych wiele plików

Przykład: otwarcie i zamknięcie pliku

```
#include <stdio.h>

int main(void)
{
    FILE *fp;

    fp = fopen("plik.txt", "w");
    if (fp == NULL)
    {
        printf("Bład otwarcia pliku.\n");
        return (-1);
    }

    /* przetwarzanie pliku */

    fclose(fp);

    return 0;
}
```


Format (plik) tekstowy i binarny

- Dane w pliku tekstowym zapisane są w postaci kodów ASCII
- Deklaracja i inicjalizacja zmiennej **x** typu **int**:

```
int x = 123456;
```

- W pamięci komputera zmienna **x** zajmuje 4 bajty:

00000000 00000001 11100010 01000000 (2)

- Po zapisaniu wartości zmiennej **x** do pliku **tekstowego** znajdzie się w nim 6 bajtów zawierających kody ASCII kolejnych cyfr

00110001 00110010 00110011 00110100 00110101 00110110 (2)

'1' '2' '3' '4' '5' '6' znaki

Format (plik) tekstowy i binarny

- Dane w pliku tekstowym zapisane są w postaci kodów ASCII
- Deklaracja i inicjalizacja zmiennej **x** typu **int**:

```
int x = 123456;
```

- W pamięci komputera zmienna **x** zajmuje 4 bajty:

00000000	00000001	11100010	01000000
----------	----------	----------	----------

 (2)

- Po zapisaniu wartości zmiennej **x** do pliku **binarnego** znajdą się w nim 4 bajty o takiej samej zawartości jak w pamięci komputera

00000000	00000001	11100010	01000000
----------	----------	----------	----------

 (2)

Format (plik) tekstowy i binarny

- Elementami pliku tekstowego są **wiersze** o różnej długości
- W systemach DOS/Windows każdy wiersz pliku tekstowego zakończony jest parą znaków:
 - **CR** (carriage return) - powrót karetki, kod ASCII - $13_{(10)} = 0D_{(16)} = '\r'$
 - **LF** (line feed) - przesunięcie o wiersz, kod ASCII - $10_{(10)} = 0A_{(16)} = '\n'$
- Załóżmy, że plik tekstowy ma postać:

```
Pierwszy wiersz pliku
Drugi wiersz pliku
Trzeci wiersz pliku
```

- Rzeczywista zawartość pliku jest następująca:

```
50 69 65 72 77 73 7A 79|20 77 69 65 72 73 7A 20 | Pierwszy wiersz
70 6C 69 6B 75 0D 0A|44|72 75 67 69 20 77 69 65 | pliku██Drugi wie
72 73 7A 20 70 6C 69 6B|75 0D 0A| 54 72 7A 65 63 | rsz pliku██Trzec
69 20 77 69 65 72 73 7A|20 70 6C 69 6B 75 0D 0A| i wiersz pliku██
```


Format (plik) tekstowy i binarny

- W systemie Linux każdy wiersz pliku tekstowego zakończony jest tylko jednym znakiem:
 - **LF** (line feed) - przesunięcie o wiersz, kod ASCII - $10_{(10)} = 0A_{(16)} = '\n'$
- Załóżmy, że plik tekstowy ma postać:

```
Pierwszy wiersz pliku
Drugi wiersz pliku
Trzeci wiersz pliku
```

- Rzeczywista zawartość pliku jest następująca:

```
50 69 65 72 77 73 7A 79|20 77 69 65 72 73 7A 20 | Pierwszy wiersz
70 6C 69 6B 75 0A 44 72|75 67 69 20 77 69 65 72 | pliku■Drugi wier
73 7A 20 70 6C 69 6B 75|0A 54 72 7A 65 63 69 20 | sz pliku■Trzeci
77 69 65 72 73 7A 20 70|6C 69 6B 75 0A | wiersz pliku■
```

- Pliki **binarne** nie mają ściśle określonej struktury

Tryby otwarcia pliku: tekstowy i binarny

```
FILE *fp1, *fp2;  
fp1 = fopen("dane.txt", "r"); // lub "rt"  
fp2 = fopen("dane.dat", "rb")
```

- Różnice pomiędzy trybem tekstowym i binarnym otwarcia pliku dotyczą innego traktowania znaków **CR** i **LF**
- W trybie **tekstowym**:
 - przy odczycie pliku para znaków **CR**, **LF** jest tłumaczona na znak nowej linii (**LF**)
 - przy zapisie pliku znak nowej linii (**LF**) jest zapisywany w postaci dwóch znaków (**CR**, **LF**)
- W trybie **binarnym**:
 - przy odczycie i zapisie para znaków **CR**, **LF** jest traktowana zawsze jako dwa znaki

Koniec wykładu nr 14

Dziękuję za uwagę!