Informatyka 1 (EZ1E2008)

Politechnika Białostocka - Wydział Elektryczny

Elektrotechnika, semestr II, studia niestacjonarne I stopnia

Rok akademicki 2020/2021

Wykład nr 9 (28.05.2021)

dr inż. Jarosław Forenc

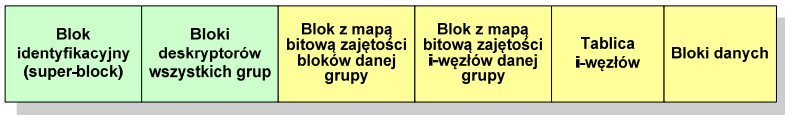
Plan wykładu nr 9

- Zarządzanie dyskowymi operacjami we-wy
  - systemy plików (ext2)
- Zarządzanie pamięcią operacyjną
  - partycjonowanie, stronicowanie, segmentacja, pamięć wirtualna
- Sieci komputerowe
  - definicja, podział, topologie i media transmisyjne
  - model referencyjny ISO/OSI, model protokołu TCP/IP
- Algorytmy komputerowe
  - definicje, podstawowe cechy, sposoby opisu
  - rekurencja, złożoność obliczeniowa
- Algorytmy sortowania
  - proste wstawianie, proste wybieranie, bąbelkowe

ext2

- pierwszy system plików w Linuxie: Minix (14-znakowe nazwy plików i maksymalny rozmiar wynoszący 64 MB)
- system Minix zastąpiono nowym systemem nazwanym rozszerzonym systemem plików - ext (ang. extended file system), a ten, w styczniu 1993 r., systemem ext2 (ang. second extended file system)
- w systemie ext2 podstawowym elementem podziału dysku jest blok
- wielkość bloku jest stała w ramach całego systemu plików, określana na etapie jego tworzenia i może wynosić 1024, 2048 lub 4096 bajtów
- w celu zwiększenia bezpieczeństwa i optymalizacji zapisu na dysku posługujemy się nie pojedynczymi blokami, a grupami bloków

ext2



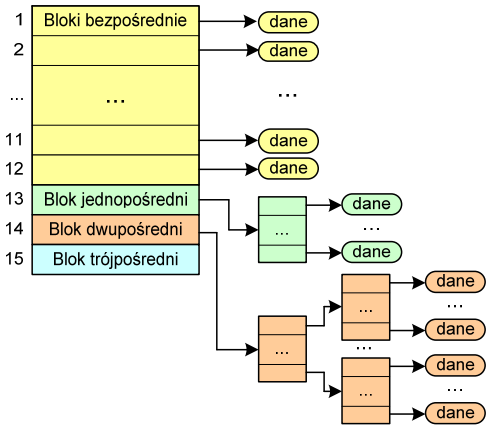
- każda grupa bloków zawiera ten sam blok identyfikacyjny oraz kopie bloków z deskryptorami wszystkich grup
- blok identyfikacyjny zawiera informacje na temat systemu plików (np. rodzaj systemu plików, rozmiar bloku)
- deskryptor grupy opisuje grupę bloków (np. położenie bloków z mapami bitowymi, liczba wolnych bloków, liczba katalogów w grupie)
- blok z mapą bitową zajętości bloków danej grupy - tablica bitów, zajmuje jeden blok (np. dla bloku o rozmiarze 1 kB opisuje 8096 bloków danych)
- blok z mapa bitową zajętości i-węzłów danej grupy - tablica bitów, każdy bit zawiera informację czy dany i-węzeł jest wolny czy zajęty

ext2 - i-węzeł

- pliki na dysku reprezentowane są przez i-węzły (ang. i-node)
- każdemu plikowi odpowiada dokładnie jeden i-węzeł, który jest strukturą zawierającą m.in. następujące pola:
  - numer i-węzła w dyskowej tablicy i-węzłów
  - typ pliku: zwykły, katalog, łącze nazwane, specjalny, znakowy
  - prawa dostępu do pliku: dla wszystkich, grupy, użytkownika
  - liczba dowiązań do pliku
  - identyfikator właściciela pliku
  - identyfikator grupy właściciela pliku
  - rozmiar pliku w bajtach (max. 4 GB)
  - czas utworzenia pliku
  - czas ostatniego dostępu do pliku
  - czas ostatniej modyfikacji pliku
  - liczba bloków dyskowych zajmowanych przez plik

ext2 - i-węzeł

- położenie pliku na dysku określają w i-węźle pola:
  - 12 adresów bloków zawierających dane (w systemie Unix jest ich 10) - bloki bezpośrednie
  - 1 adres bloku zawierającego adresy bloków zawierających dane - blok jednopośredni (ang. single indirect block)
  - 1 adres bloku zawierającego adresy bloków jednopośrednich - blok dwupośredni (ang. double indirect block)
  - 1 adres bloku zawierającego adresy bloków dwupośrednich - blok trójpośredni (ang. triple indirect block)



ext2

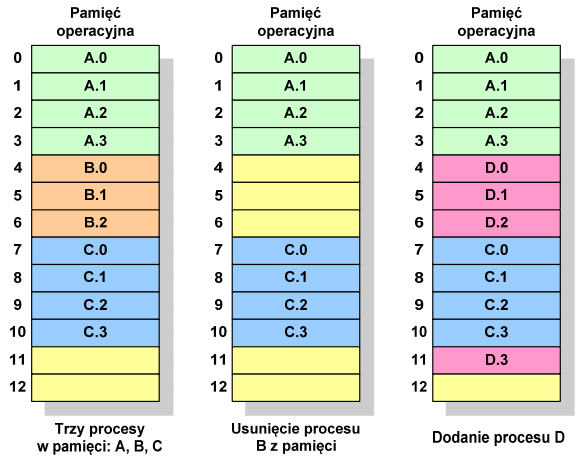
- nazwy plików przechowywane są w katalogach, które w systemie Linux są plikami, ale o specjalnej strukturze
- katalogi składają się z ciągu tzw. pozycji katalogowych o nieustalonej z góry długości
- każda pozycja opisuje dowiązanie do jednego pliku i zawiera:
  - numer i-węzła (4 bajty)
  - rozmiar pozycji katalogowej (2 bajty)
  - długość nazwy (2 bajty)
  - nazwa (od 1 do 255 znaków)

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 pamięci wirtualnej, segmentacja pamięci wirtualnej
  - stronicowanie i segmentacja pamięci wirtualnej

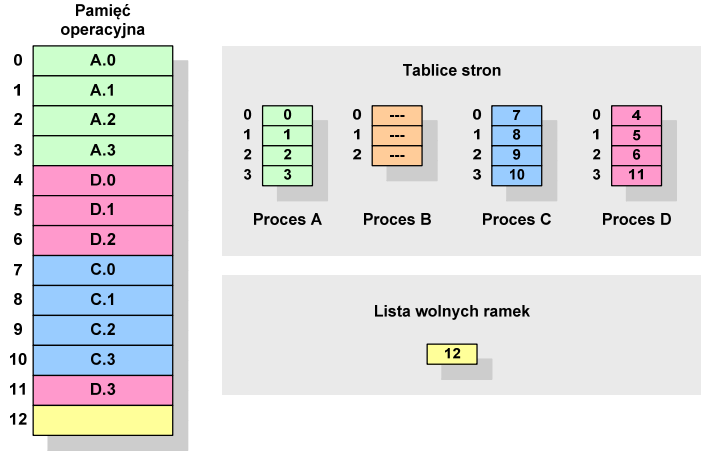
Proste stronicowanie

- 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



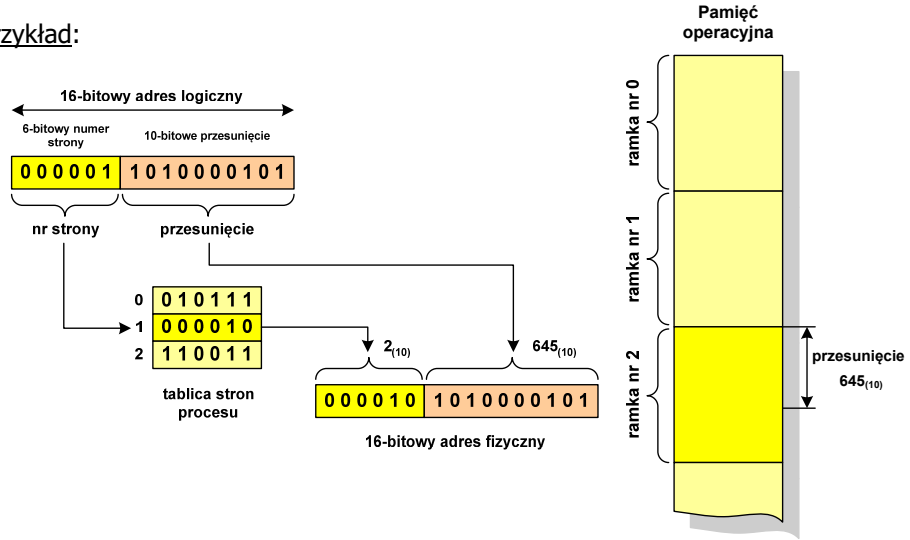
Proste stronicowanie

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



Proste stronicowanie

Przykład:

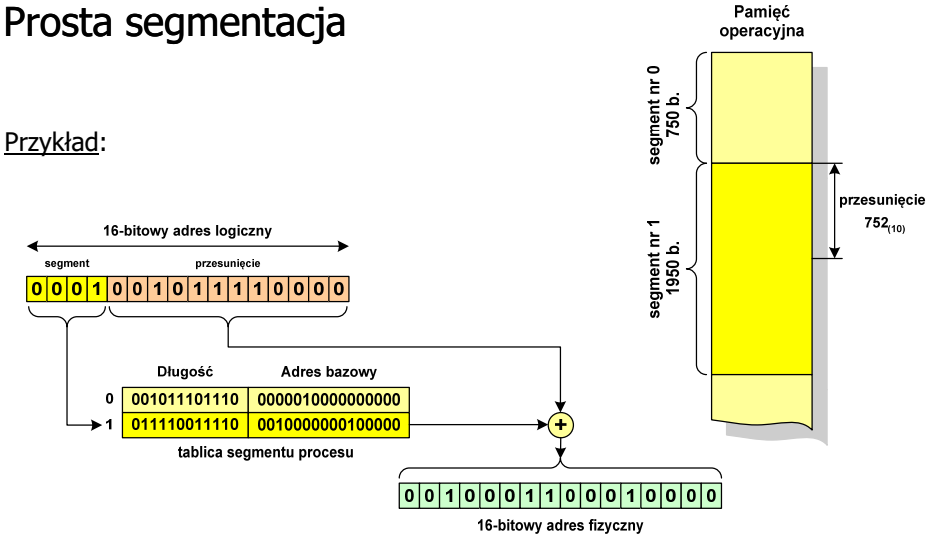


Prosta segmentacja

- 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?

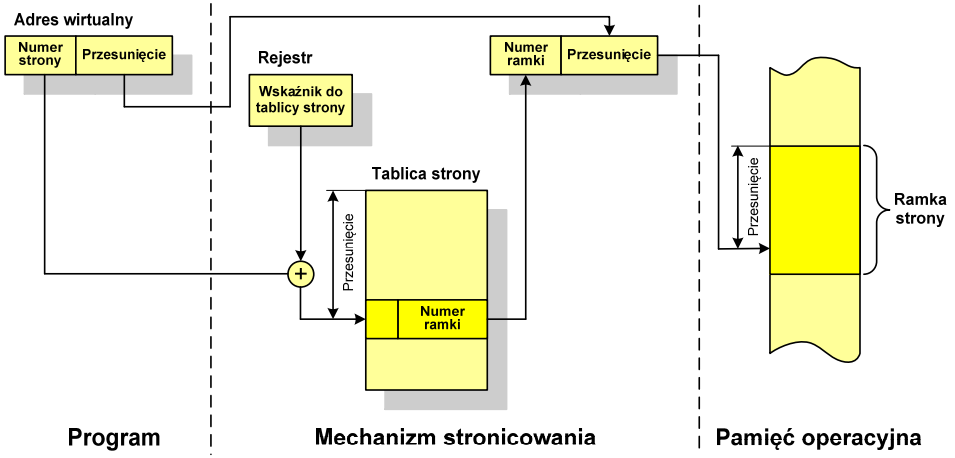
- generowanie przerwania sygnalizującego błąd w dostępie do pamięci
- zmiana stan procesu na zablokowany
- wstawienie do pamięci operacyjnej fragment procesu zawierający adres logiczny, który był przyczyną błędu
- zmiana stanu procesu na uruchomiony

Dzięki zastosowaniu pamięci wirtualnej:

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

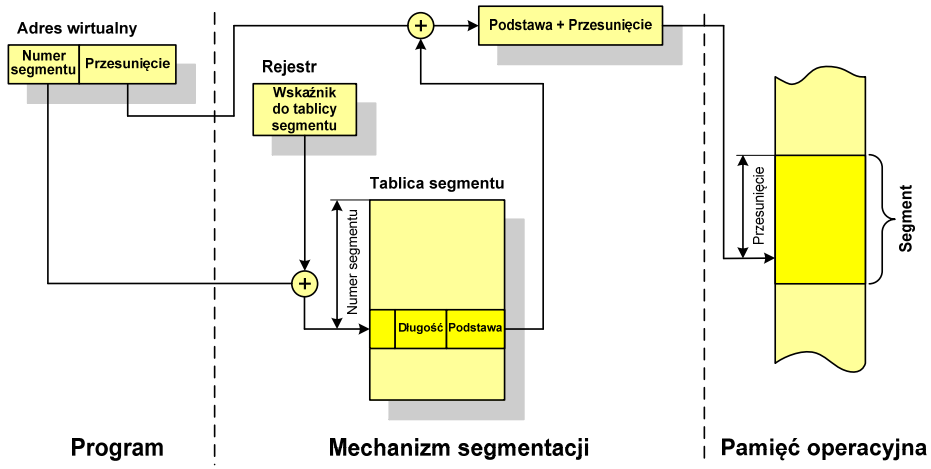
Stronicowanie pamięci wirtualnej

- odczytanie strony wymaga translacji adresu wirtualnego na fizyczny



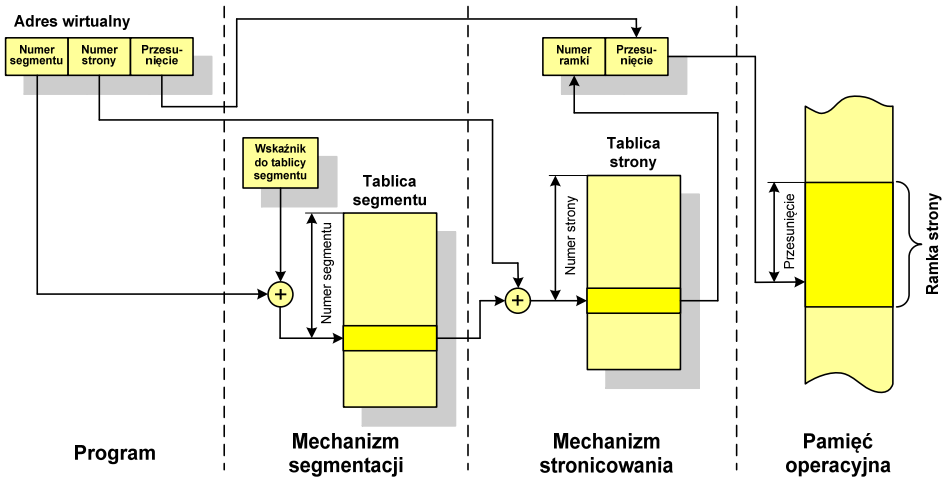
Segmentacja pamięci wirtualnej

- mechanizm odczytania słowa z pamięci obejmuje translację adresu wirtualnego na fizyczny za pomocą tablicy segmentu



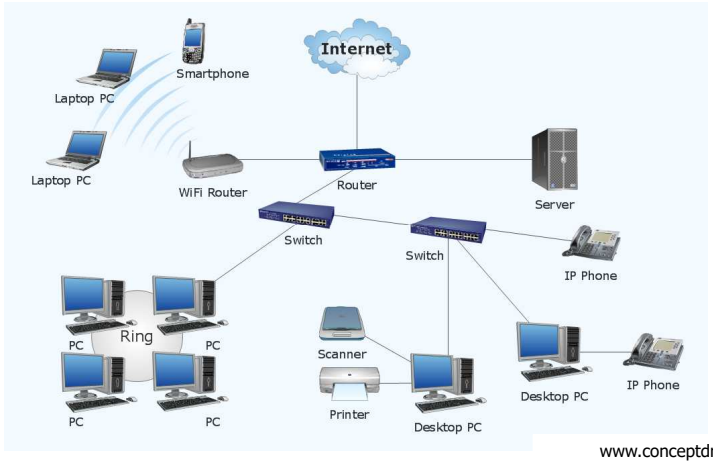
Stronicowanie i segmentacja pamięci wirtualnej

- tłumaczenie adresu wirtualnego na adres fizyczny:



Sieć komputerowa

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



www.conceptdraw.com

Podział sieci w zależności od ich rozmiaru

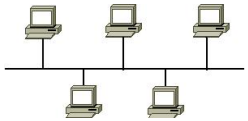
- LAN (Local Area Network) - sieć lokalna, łączy komputery znajdujące się na określonym, niewielkim obszarze (kilka budynków, przedsiębiorstwo), wykonana jest w jednej technologii (np. Ethernet)
- MAN (Metropolitan Area Network) - sieć miejska, obejmuje zasięgiem aglomerację lub miasto łącząc oddzielne 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 - ogólnoświatowa 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

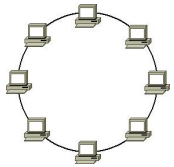
- Topologia sieci - określa strukturę sieci
  - zbiór zasad fizycznego łączenia elementów sieci (topologia fizyczna)
  - zbiór reguł komunikacji poprzez medium transmisyjne (topologia logiczna)
- Topologia fizyczna - opisuje sposoby fizycznego łączenia ze sobą komputerów (układ przewodów, media transmisyjne)
- 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

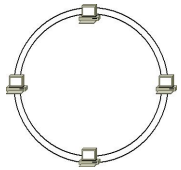
- topologia magistrali (bus) - wszystkie komputery podłączone są do jednego współdzielonego medium transmisyjnego (najczęściej kabla koncentrycznego)



- topologia pierścienia (ring) - komputery połączone są pomiędzy sobą odcinkami kabla tworząc zamknięty pierścień (sieci światłowodowe, sieci LAN)

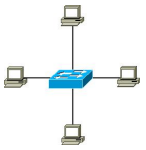


- topologia podwójnego pierścienia (dual-ring) - komputery połączone są pomiędzy sobą odcinkami kabla tworząc dwa zamknięte pierścienie (większa niezawodność, sieci: szkieletowe, MAN, Token Ring, FDDI)

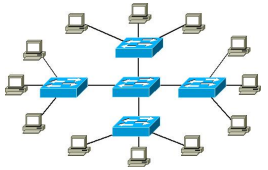


Topologie sieci komputerowych

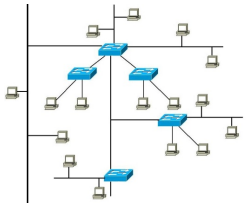
- topologia gwiazdy (star) - komputery podłączone są do jednego punktu centralnego (koncentrator, przełącznik), obecnie jest to najczęściej stosowana topologia sieci LAN



- 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



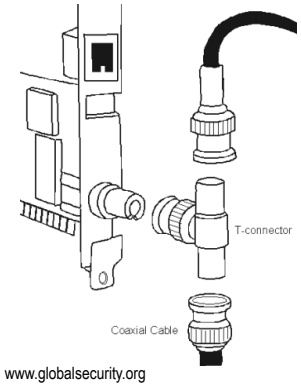
Media transmisyjne - przewód koncentryczny

- Ethernet gruby (Thick Ethernet), 10Base-5, 10 Mb/s
  - kabel RG-8 lub RG-11, impedancja falowa: 50 Ω, grubość: 1/2”
  - max. odległość między stacjami: 500 m
- Ethernet cienki (Thin Ethernet), 10Base-2 , 10 Mb/s
  - kabel RG-58, impedancja falowa: 50 Ω, grubość: 1/4”
  - max. odległość między stacjami: 185 m

Złącze BNC

Trójnik (Łącznik T)

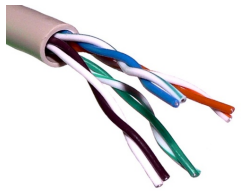
Terminator



www.globalsecurity.org

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



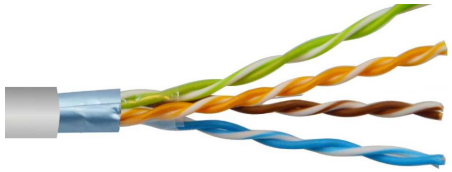
U/UTP - skrętka nieekranowana (UTP)

RJ-45

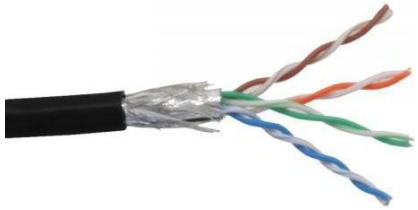
RJ-11

Media transmisyjne - skrętka

- F/UTP (dawniej FTP) - skrętka foliowana



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



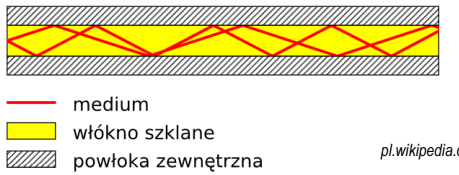
Media transmisyjne - światłowód

- 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

lan-networks.cba.pl

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

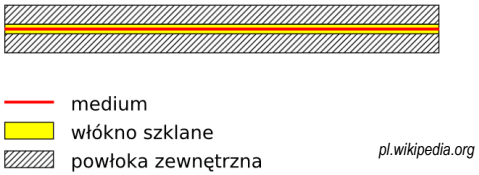


pl.wikipedia.org

- 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



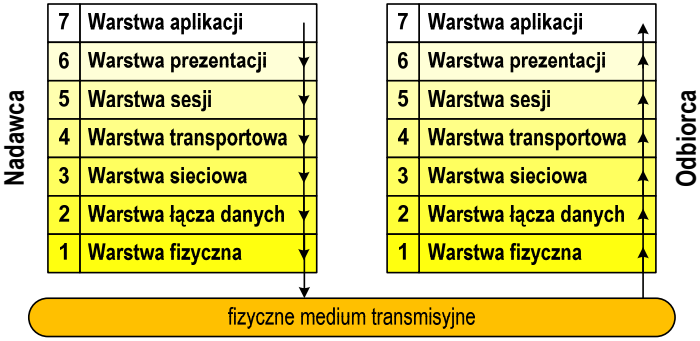
pl.wikipedia.org

- 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

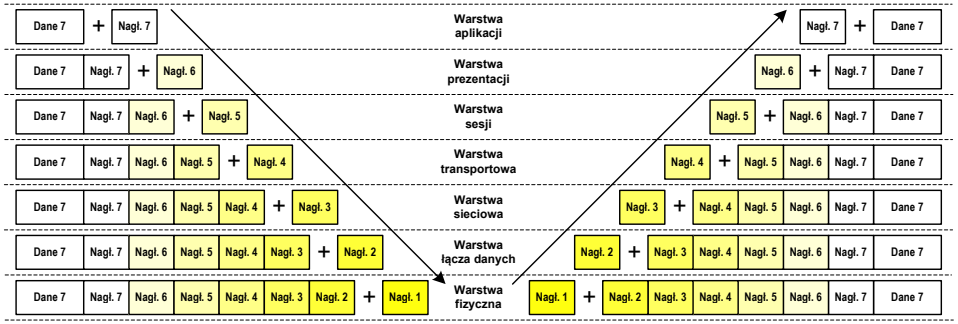
- w latach 70-tych nie istniały ogólne standardy dotyczące sieci komputerowych - każdy producent tworzył własną sieć
- w 1984 roku Międzynarodowa Organizacja Normalizacyjna (ISO) przyjęła model sieciowy, dzięki któremu producenci mogliby opracowywać współpracujące ze sobą rozwiązania sieciowe
- ISO OSI 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

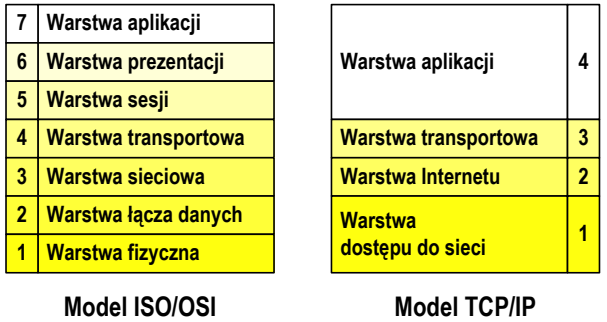
Model ISO/OSI



- 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

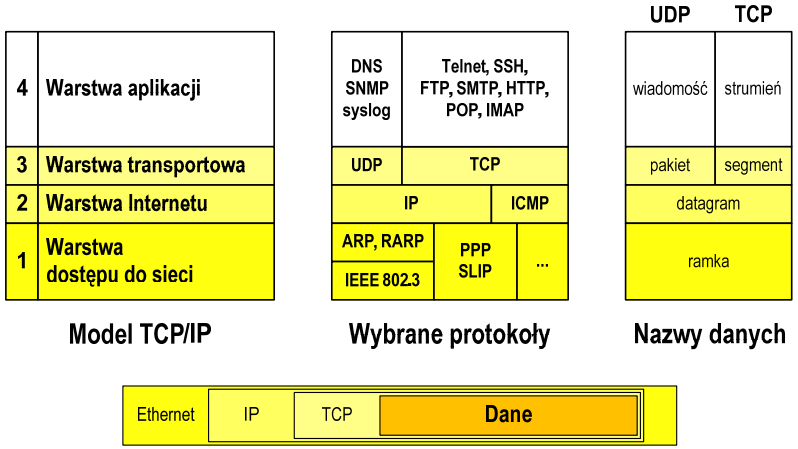
Model ISO/OSI a model TCP/IP

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



Model TCP/IP

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

- Preambuła - naprzemienny ciąg bitów 1 i 0 informujący o ramce
- Adres docelowy / źródłowy - 6-bajtowe liczby będące adresami sprzętowymi komunikujących się interfejsów sieciowych (MAC - Media Access Control)

`00:23:76:09:41:3B` — producent karty / numer egzemplarza

`FF:FF:FF:FF:FF:FF` — adres docelowy rozgłoszeniowy

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 jedynkami (bitowo)
- FCS (Frame Check Sequence) - 4 bajty kontrolne (CRC - Cyclic Redundancy Check) wygenerowane przez interfejs nadający i sprawdzane przez odbierający

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

- wysłanie ramki wymaga znajomości adresu MAC odbiorcy
- do określenia adresu MAC na podstawie numeru IP stosowany jest protokół ARP (Address Resolution Protocol)
- protokół ARP stosowany jest także do zapobiegania zdublowaniu adresów IP
- aktualną tablicę translacji ARP wyświetla polecenie: arp -a

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:

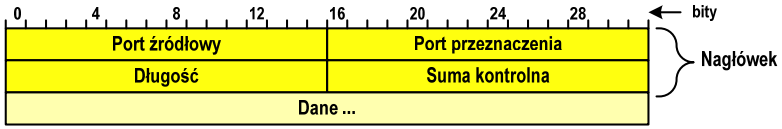
port	protokół
20	FTP (dane)
21	FTP (polecenia)
22	SSH
23	Telnet
25	SMTP (mail)

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 pakiet

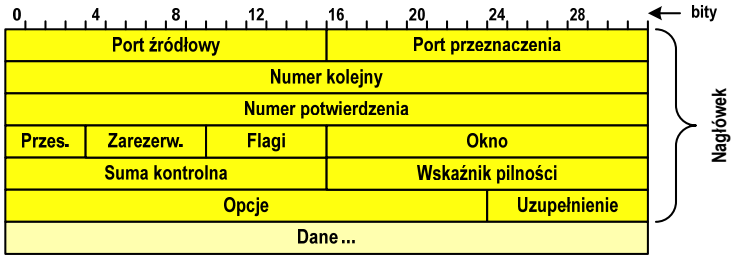


- Port źródłowy (Source port) - numer portu nadawcy
- Port przeznaczenia (Destination port) - numer portu odbiorcy
- Długość (Length) - całkowita długość pakietu w bajtach (nagłówek + dane)
- Suma kontrolna (Checksum) - tworzona na podstawie nagłówka i danych

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 Internetu - 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

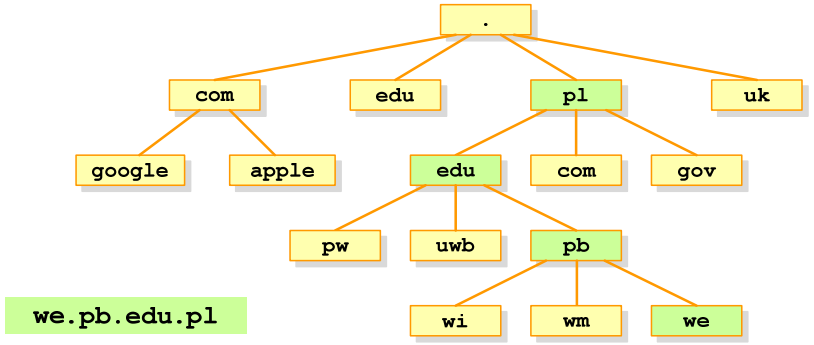
`we.pb.edu.pl` → `213.33.95.2`

- wykorzystuje port o numerze 53

Warstwa aplikacji

DNS (Domain Name System)

- przestrzeń nazw w Internecie oparta jest na modelu odwróconego drzewa



- zarządzaniem przestrzenią nazw domenowych zajmuje się w świecie ICANN, zaś w Polsce - NASK

Algorytm - definicje

Definicja 1

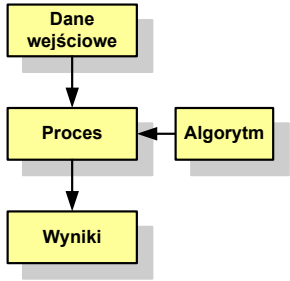
- Skończony, uporządkowany ciąg jasno zdefiniowanych czynności, koniecznych do wykonania pewnego zadania

Definicja 2

- Opis rozwiązania problemu wyrażony za pomocą operacji zrozumiałych i możliwych do zrealizowania przez wykonawcę

Definicja 3

- Ściśle określona procedura obliczeniowa, która dla właściwych danych wejściowych zwraca żądane dane wyjściowe zwane wynikiem działania algorytmu



Definicja 4

- Metoda rozwiązania zadania

Algorytmy

- Słowo „algorytm” pochodzi od nazwiska matematyka perskiego z IX wieku - Muhammada ibn-Musy al-Chuwarizmiego (po łacinie pisanego jako Algorismus)
- Badaniem algorytmów zajmuje się algorytmika
- „Przetłumaczenie” algorytmu na wybrany język programowania:
  - implementacja algorytmu
  - kodowanie algorytmu
- Sposoby opisu algorytmów
  - opis słowny w języku naturalnym lub lista kroków (opis w punktach)
  - schemat blokowy
  - pseudokod (nieformalna odmiana języka programowania)
  - wybrany język programowania

Opis słowny algorytmu

- Podanie kolejnych czynności, które należy wykonać, aby otrzymać oczekiwany efekt końcowy
- Przypomina przepis kulinarny z książki kucharskiej lub instrukcję obsługi urządzenia, np.

Algorytm: Tortilla („Podróże kulinarne” R. Makłowicza)

Dane wejściowe: 0,5 kg ziemniaków, 100 g kiełbasy Chorizo, 8 jajek

Dane wyjściowe: gotowa Tortilla

Opis algorytmu: Ziemniaki obrać i pokroić w plasterki. Kiełbasę pokroić w plasterki. Ziemniaki wrzucić na gorącą oliwę na patelni i przyrumienić z obu stron. Kiełbasę wrzucić na gorącą oliwę na patelni i przyrumienić z obu stron. Ubić jajka i dodać do połączonych ziemniaków i kiełbasy. Dodać sól i pieprz. Usmażyć z obu stron wielki omlet nadziewany chipsami ziemniaczanymi z kiełbaską.

Lista kroków

- Uporządkowany opis wszystkich czynności, jakie należy wykonać podczas realizacji algorytmu
- Krok jest to pojedyncza czynność realizowana w algorytmie
- Kroki w algorytmie są numerowane, operacje wykonywane są zgodnie z rosnącą numeracją kroków
- Jedynym odstępstwem od powyższej reguły są operacje skoku (warunkowe lub bezwarunkowe), w których jawnie określa się numer kolejnego kroku
- Przykład (instrukcja otwierania wózka-specerówki):

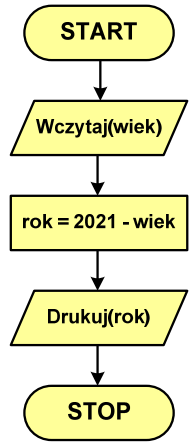
  Krok 1: Zwolnij element blokujący wózek

  Krok 2: Rozkładaj wózek w kierunku kółek

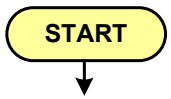
  Krok 3: Naciskając nogą dolny element blokujący aż do zatrzaśnięcia, rozłóż wózek do pozycji przewozowej

Schemat blokowy

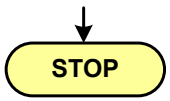
- Zawiera plan algorytmu przedstawiony w postaci graficznej
- Na schemacie umieszczane są bloki oraz linie przepływu (strzałki)
- Blok zawiera informację o wykonywanej operacji
- Linie przepływu (strzałki) określają kolejność wykonywania bloków algorytmu
- Przykład: wyznaczanie roku urodzenia na podstawie wieku (algorytm liniowy)



Schemat blokowy - symbole graficzne

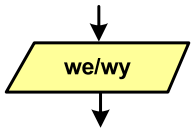


- blok startowy, początek algorytmu
- wskazuje miejsce rozpoczęcia algorytmu
- ma jedno wyjście
- może występować tylko jeden raz

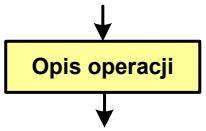


- blok końcowy, koniec algorytmu
- wskazuje miejsce zakończenia algorytmu
- ma jedno wejście
- musi występować przynajmniej jeden raz

Schemat blokowy - symbole graficzne

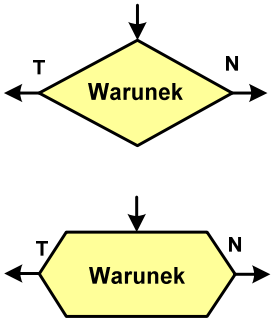


- blok wejścia-wyjścia
- poprzez ten blok wprowadzane są (czytane) dane wejściowe i wyprowadzane (zapisywane) wyniki
- ma jedno wejście i jedno wyjście

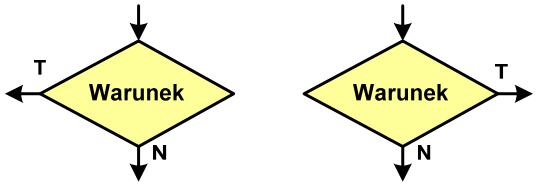


- blok wykonawczy, blok funkcyjny, opis procesu
- zawiera jedno lub kilka poleceń (elementarnych instrukcji) wykonywanych w podanej kolejności
- instrukcją może być np. operacja arytmetyczna, podstawienie
- ma jedno wejście i jedno wyjście

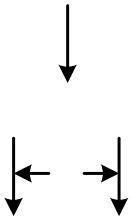
Schemat blokowy - symbole graficzne



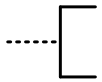
- blok warunkowy (decyzyjny, porównujący)
- wewnątrz bloku umieszcza się warunek logiczny
- na podstawie warunku określana jest tylko jedna droga wyjściowa
- połączenia wychodzące z bloku:
  - T lub TAK - gdy warunek jest prawdziwy
  - N lub NIE - gdy warunek nie jest prawdziwy
- wyjścia mogą być skierowane na boki lub w dół



Schemat blokowy - symbole graficzne



- linia przepływu, połączenie, linia
- występuje w postaci linii zakończonej strzałką
- określa kierunek przemieszczania się po schemacie
- łączy inne bloki występujące na schemacie
- linie pochodzące z różnych części algorytmu mogą zbiegać się w jednym miejscu



- komentarz
- dodanie do schematu dodatkowego opisu

Pseudokod i język programowania

Pseudokod:

- Pseudokod (pseudojęzyk) - uproszczona wersja języka programowania
- Często zawiera zwroty pochodzące z języków programowania
- Zapis w pseudokodzie może być łatwo przetłumaczony na wybrany język programowania

Opis w języku programowania:

- Zapis programu w konkretnym języku programowania
- Stosowane języki: Pascal, C, C++, Matlab, Python
  (kiedyś - Fortran, Basic)

Największy wspólny dzielnik - algorytm Euklidesa

- NWD - największa liczba naturalna dzieląca (bez reszty) dwie (lub więcej) liczby całkowite

NWD(1675,3752) = ?

Algorytm Euklidesa - przykład

a	b	Dzielenie większej liczby przez mniejszą	Zamiana
1675	3752	b/a = 3752/1675 = 2 reszta 402	b = 402
1675	402	a/b = 1675/402 = 4 reszta 67	a = 67
67	402	b/a = 402/67 = 6 reszta 0	b = 0
67	0	KONIEC	

NWD(1675,3752) = 67

Algorytm Euklidesa - lista kroków

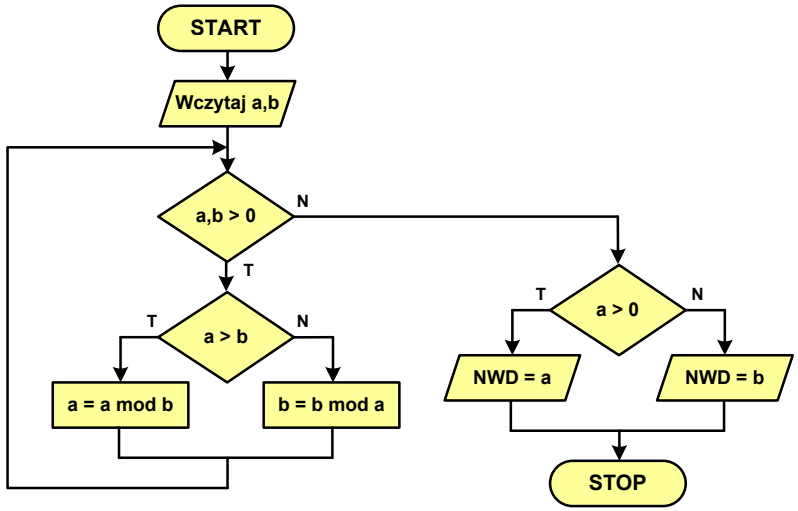
Dane wejściowe: niezerowe liczby naturalne a i b

Dane wyjściowe: NWD(a,b)

Kolejne kroki:

1. Czytaj liczby a i b
2. Dopóki a i b są większe od zera, powtarzaj krok 3, a w przeciwnym przypadku przejdź do kroku 4
3. Jeśli a jest większe od b, to weź za a resztę z dzielenia a przez b, w przeciwnym przypadku weź za b resztę z dzielenia b przez a
4. Przyjmij jako największy wspólny dzielnik tę z liczb a i b, która pozostała większa od zera
5. Drukuj NWD(a,b)

Algorytm Euklidesa - schemat blokowy



Algorytm Euklidesa - pseudokod

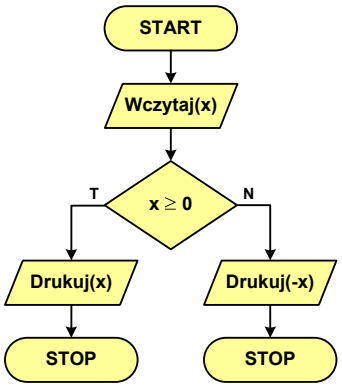
```
NWD(a,b)
while a>0 i b>0
 do if a>b
       then a ← a mod b
       else b ← b mod a
if a>0
   then return a
   else return b
```

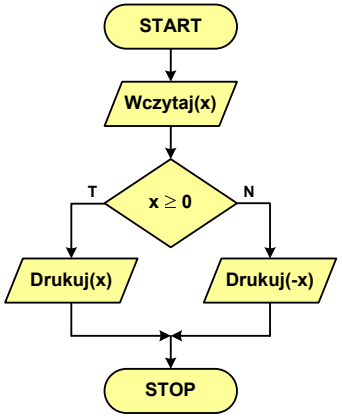

Algorytm Euklidesa - język programowania (C)

```c
#include <stdio.h>

int main(void)
{
    int a = 1675, b = 3752, NWD;

    while (a>0 && b>0)
        if (a>b)
            a = a % b;
        else
            b = b % a;

    if (a>0)
        NWD = a;
    else
        NWD = b;

    printf("NWD = %d\n",NWD);
}
```

Wartość bezwzględna liczby - schemat blokowy

$$|x| = \begin{cases} x & dla \quad x \geq 0 \\ -x & dla \quad x < 0 \end{cases}$$





Równanie kwadratowe - schemat blokowy

$$ax^2 + bx + c = 0$$

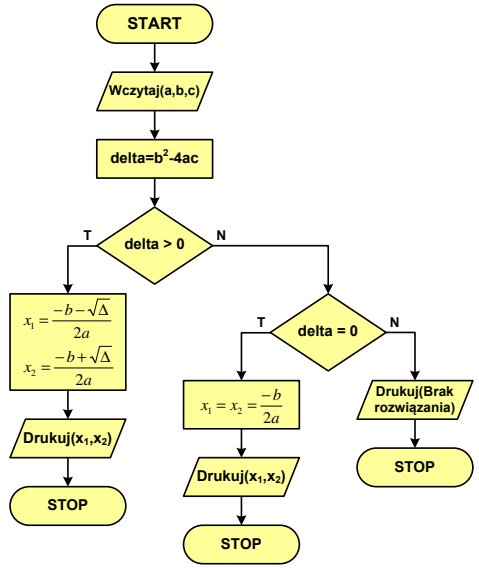
$$\Delta = b^2 - 4ac$$

$$\Delta > 0:$$

$$x_1 = \frac{-b - \sqrt{\Delta}}{2a}, \quad x_2 = \frac{-b + \sqrt{\Delta}}{2a}$$

$$\Delta = 0:$$

$$x_1 = x_2 = \frac{-b}{2a}$$

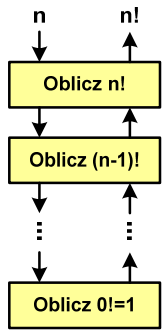


Rekurencja

- Rekurencja lub rekursja - jest to odwoływanie się funkcji lub definicji do samej siebie
- Rozwiązanie danego problemu wyraża się za pomocą rozwiązań tego samego problemu, ale dla danych o mniejszych rozmiarach
- W matematyce mechanizm rekurencji stosowany jest do definiowania lub opisywania algorytmów
- Silnia:

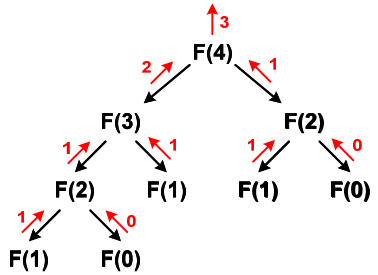
$$n! = \begin{cases} 1 & dla \quad n = 0 \\ n(n-1)! & dla \quad n \geq 1 \end{cases}$$

```
int silnia(int n)
{
    return n==0 ? 1 : n*silnia(n-1);
}
```



Rekurencja - ciąg Fibonacciego

$$F_n = \begin{cases} 0 & dla \quad n = 0 \\ 1 & dla \quad n = 1 \\ F_{n-1} + F_{n-2} & dla \quad n > 1 \end{cases}$$




```
int F(int n)
{
    if (n==0) return 0;
    else
        if (n==1) return 1;
        else
            return F(n-1) + F(n-2);
}
```

Złożoność obliczeniowa

- W celu rozwiązania danego problemu obliczeniowego szukamy algorytmu najbardziej efektywnego czyli:
  - najszybszego (najkrótszy czas otrzymania wyniku)
  - o możliwie małym zapotrzebowaniu na pamięć
- Problem: Jak ocenić, który z dwóch różnych algorytmów rozwiązujących to samo zadanie jest efektywniejszy?
- Do oceny efektywności służy złożoność obliczeniowa algorytmu (koszt algorytmu) czyli ilość zasobów potrzebnych do jego działania (czas, pamięć)
- Miarą złożoności czasowej jest liczba podstawowych (dominujących) operacji (porównanie, podstawienie, operacja arytmetyczna) - pozostałe operacje są pomijane
- Miarą złożoności pamięciowej jest liczba wykorzystanych komórek pamięci (bajty lub liczba zmiennych określonego typu)

Złożoność obliczeniowa

- Złożoność obliczeniowa algorytmu jest funkcją opisującą zależność między liczbą danych a liczbą operacji wykonywanych przez ten algorytm
- W praktyce stosuje się oszacowanie powyższej funkcji - są to tzw. notacje (klasy złożoności):
  - O (duże O) - najbardziej popularna
  - Ω (omega)
  - Θ (theta)

Notacja O („duże O”)

- Wyraża złożoność matematyczną algorytmu
- Do wyznaczenia złożoności bierze się pod uwagę liczbę dominujących operacji wykonywanych w algorytmie
- Przykład zapisu: $O(n^2)$
  - po literze O występuje wyrażenie w nawiasach zawierające literę n, która oznacza liczbę elementów, na których działa algorytm
- W funkcji opisującej złożoność bierze się pod uwagę tylko najistotniejszy składnik, np.

  $f(n) = n^2+2n \rightarrow O(n^2)$ $\quad f(n) = n^2+n-5 \rightarrow O(n^2)$
- W powyższych przykładach dla dużego n wpływ składnika liniowego i stałego na wartość funkcji jest nieistotny w porównaniu ze składnikiem głównym n^2

Notacja O („duże O”)

- Porównanie najczęściej występujących złożoności:

Elementy (n)	O(log n)	O(n)	O(n log n)	$O(n^2)$	$O(n^3)$	$O(2^n)$
10	3	10	33	100	1 000	1024
100	7	100	664	10 000	1 000 000	$1{,}27 \cdot 10^{30}$
1 000	10	1 000	9 966	1 000 000	10^9	$1{,}07 \cdot 10^{301}$
10 000	13	10 000	132 877	10^8	10^{12}	$1{,}99 \cdot 10^{3010}$

O(log n) - logarytmiczna (np. przeszukiwanie binarne)
O(n) - liniowa (np. porównywanie łańcuchów znaków)
O(n log n) - liniowo-logarytmiczna (np. sortowanie szybkie)
$O(n^2)$ - kwadratowa (np. proste algorytmy sortowania)
$O(n^3)$ - sześcienna (np. mnożenie macierzy)
$O(2^n)$ - wykładnicza (np. problem komiwojażera)

Sortowanie

- Sortowanie polega na uporządkowaniu zbioru danych względem pewnych cech charakterystycznych każdego elementu tego zbioru (wartości każdego elementu)
- W przypadku liczb, sortowanie polega na znalezieniu kolejności liczb zgodnej z relacją ≤ lub ≥

Przykład:

- Tablica nieposortowana:

6	4	5	2	3	1

- Tablica posortowana zgodnie z relacją ≤ (od najmniejszej do największej liczby):

1	2	3	4	5	6

- Tablica posortowana zgodnie z relacją ≥ (od największej do najmniejszej liczby):

6	5	4	3	2	1

Sortowanie

- W przypadku słów sortowanie polega na ustawieniu ich w porządku alfabetycznym (leksykograficznym)

Przykład:

- Tablica nieposortowana:

Paweł	Piotr	Adrian	Ela	Ola	Henryk

- Tablice posortowane:

Adrian	Ela	Henryk	Ola	Paweł	Piotr

Piotr	Paweł	Ola	Henryk	Ela	Adrian

Sortowanie

- W praktyce sortowanie sprowadza się do porządkowanie danych na podstawie porównania - porównywany element to klucz

Przykład:

- Tablica nieposortowana (imię, nazwisko, wiek):

Piotr	Ola	Paweł	Jan	Ela	Magda
Kowalski	Nowak	Wójcik	Kamiński	Król	Mazur
25	18	23	20	22	15

- Tablica posortowana (klucz - nazwisko):

Jan	Piotr	Ela	Magda	Ola	Paweł
Kamiński	Kowalski	Król	Mazur	Nowak	Wójcik
20	25	22	15	18	23

- Tablica posortowana (klucz - wiek):

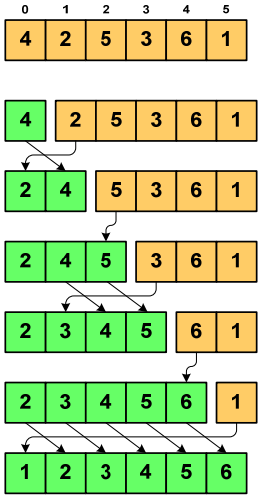
Magda	Ola	Jan	Ela	Paweł	Piotr
Mazur	Nowak	Kamiński	Król	Wójcik	Kowalski
15	18	20	22	23	25

Sortowanie

- Po co stosować sortowanie?
  - posortowane elementy można szybciej zlokalizować
  - posortowane elementy można przedstawić w czytelniejszy sposób
- Przykładowe algorytmy sortowania
  - proste wstawianie (insertion sort)
  - proste wybieranie (selection sort)
  - bąbelkowe (bubble sort)
  - szybkie (quick sort)
  - przez scalanie (merge sort)
  - kubełkowe / przez zliczanie (bucket sort)

Proste wstawianie (insertion sort)

Przykład:

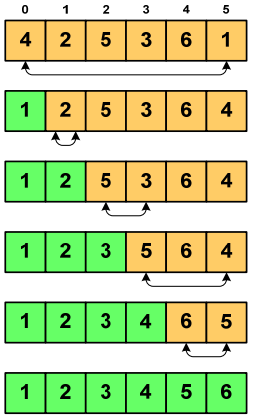


Program w języku C:

```
int main(void)
{
   int tab[N],i,j,tmp;
   // ...
   for (i=1; i<N; i++)
   {
      j=i;
      tmp=tab[i];
      while (tab[j-1]>tmp && j>0)
      {
         tab[j]=tab[j-1];
         j--;
      }
      tab[j]=tmp;
   }
}
```


Proste wybieranie (selection sort)

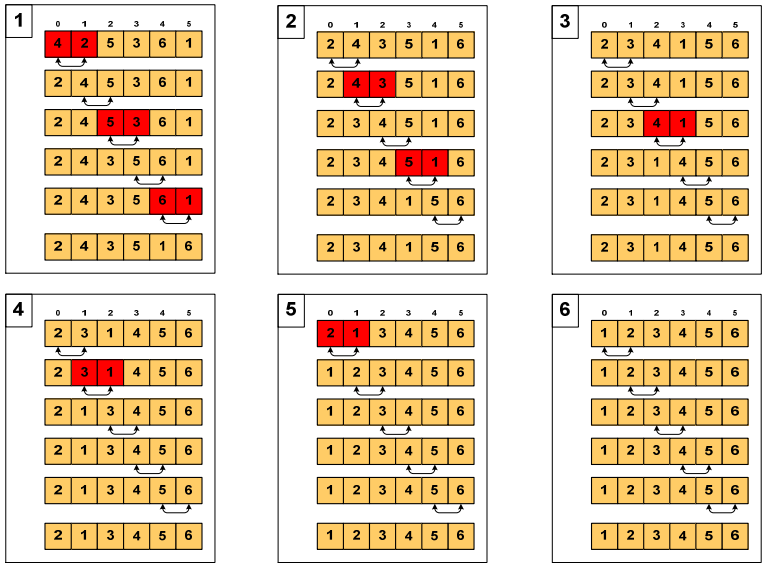
Przykład:



Program w języku C:

```
int main(void)
{
   int tab[N],i,j,k,tmp;
   // ...
   for (i=0;i<N-1;i++)
   {
      k=i;
      for (j=i+1; j<N; j++)
         if (tab[k]>=tab[j])
            k = j;
      tmp = tab[i];
      tab[i] = tab[k];
      tab[k] = tmp;
   }
}
```

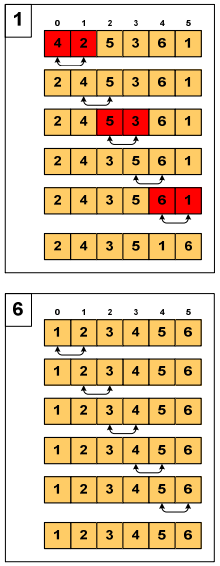

Bąbelkowe (bubble sort)



Bąbelkowe (bubble sort)

Program w języku C:

```
int main(void)
{
   int tab[N],i,j,tmp,koniec;
   // ...
   do {
      koniec=1;
      for (i=0;i<N-1;i++)
         if (tab[i]>tab[i+1])
         {
            tmp=tab[i];
            tab[i]=tab[i+1];
            tab[i+1]=tmp;
            koniec=0;
         }
   } while (!koniec);
}
```



Koniec wykładu nr 9

Dziękuję za uwagę!