

Systemy operacyjne

Zarządzanie pamięcią operacyjną

Wykład prowadzą:
Jerzy Brzeziński
Dariusz Wawrzyniak



Systemy operacyjne

Plan wykładu

- Pamięć jako zasób systemu komputerowego
 - hierarchia pamięci
 - przestrzeń adresowa
- Wsparcie dla zarządzania pamięcią na poziomie architektury komputera
- Podział i przydział pamięci
- Obraz procesu w pamięci
- Stronicowanie
- Segmentacja

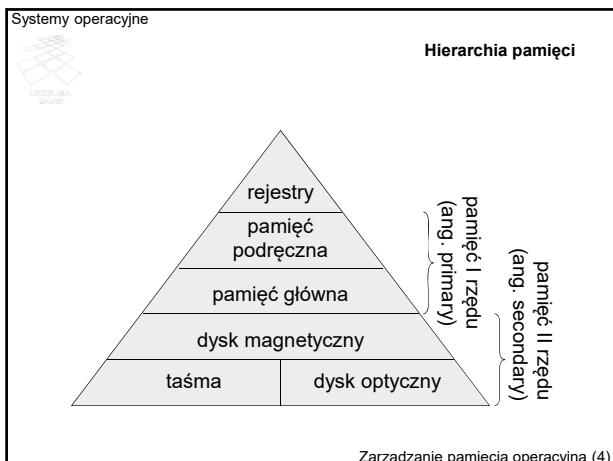
Zarządzanie pamięcią operacyjną (2)

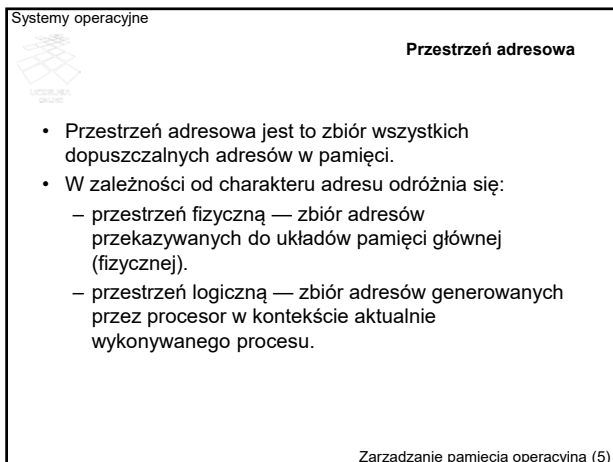
Systemy operacyjne

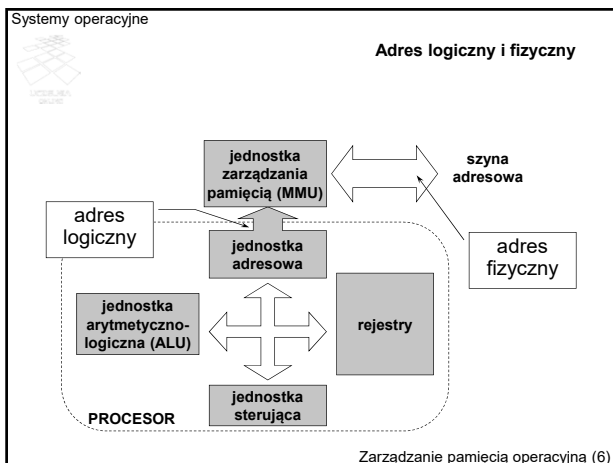
Pamięć jako zasób systemu komputerowego

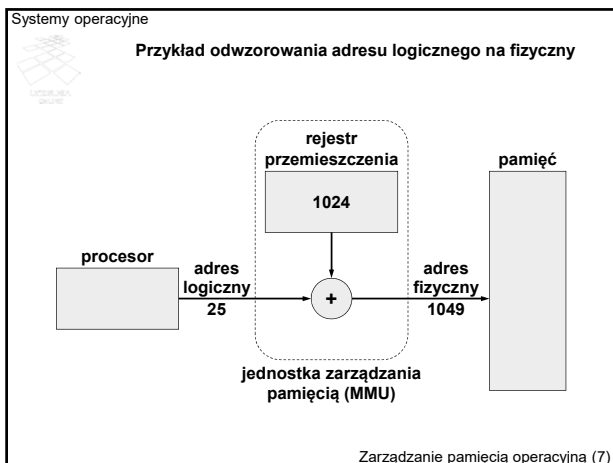
- Pamięć jest zasobem służącym do przechowywania danych i programów.
- Z punktu widzenia systemu pamięć jest zasobem o strukturze hierarchicznej (począwszy od rejestrów procesora, przez pamięć podręczną, pamięć główną, skończywszy na pamięci masowej), w której na wyższym poziomie przechowywane są dane, stanowiące fragment zawartości poziomu niższego.
- Z punktu widzenia procesu (również procesora) pamięć jest zbiorem bajtów identyfikowanych przez adresy, czyli tablicą bajtów, w której adresy są indeksami.

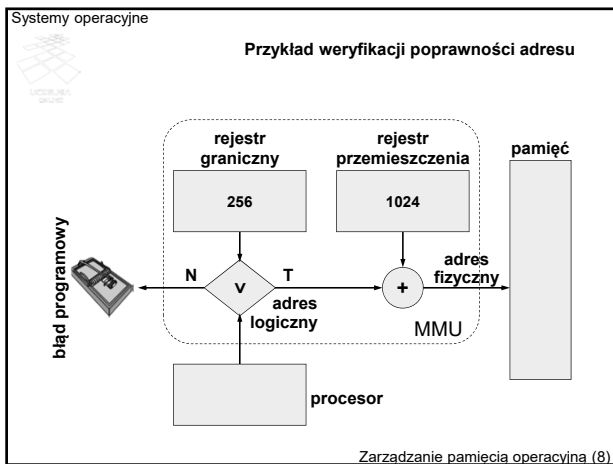
Zarządzanie pamięcią operacyjną (3)

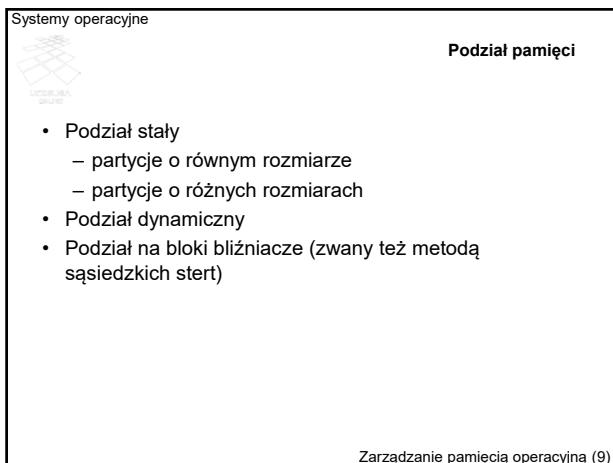


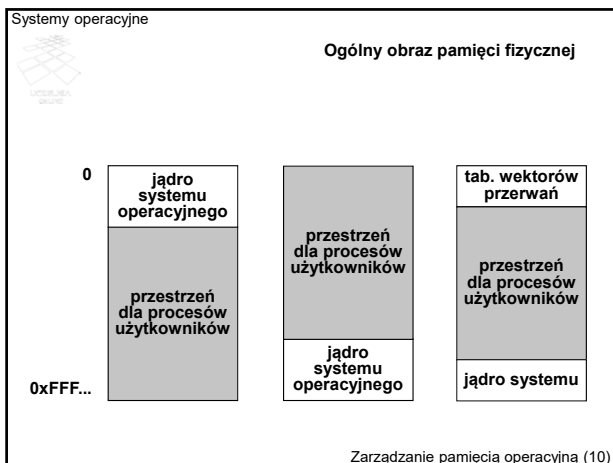


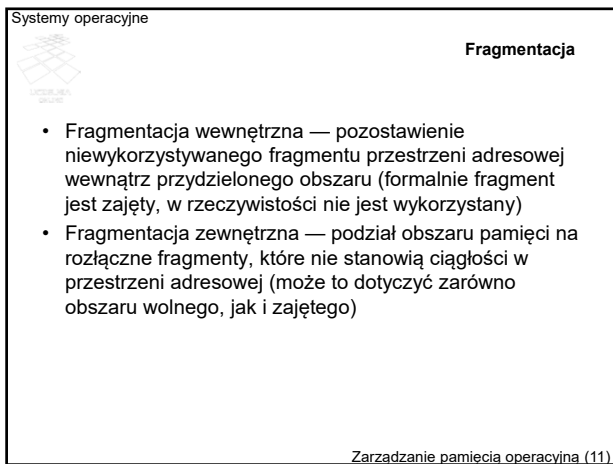


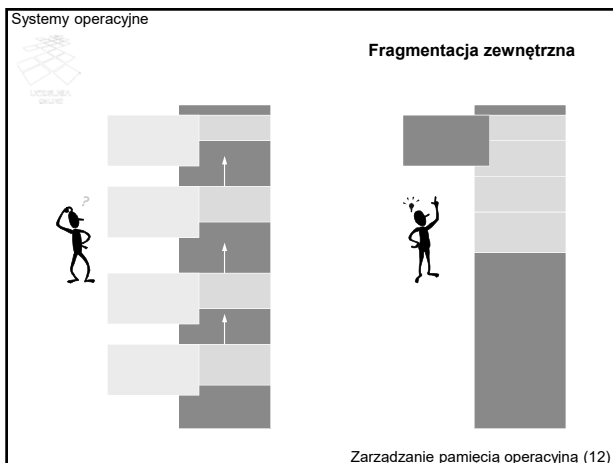


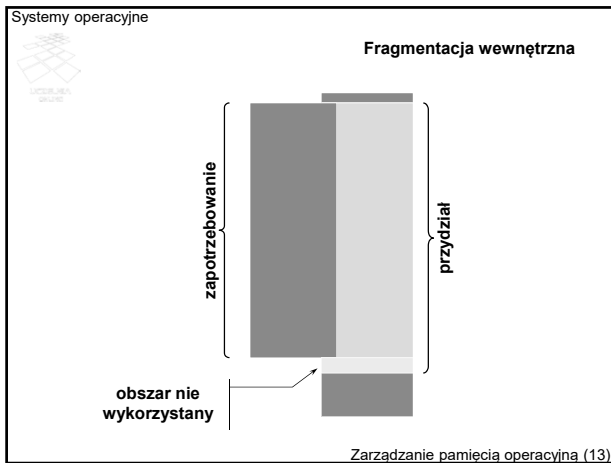


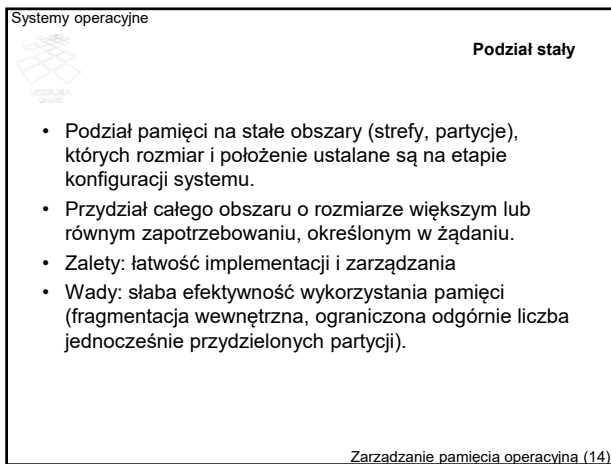


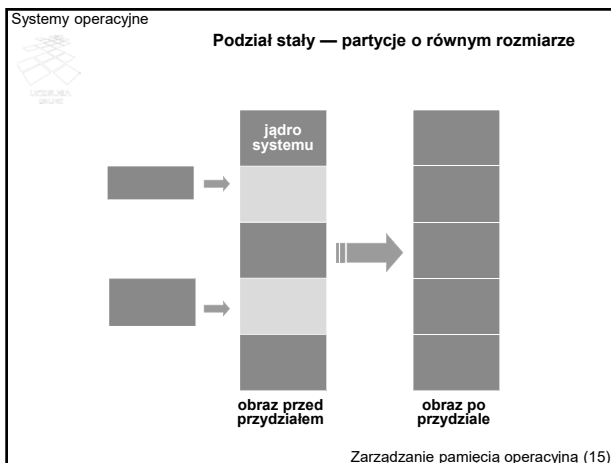


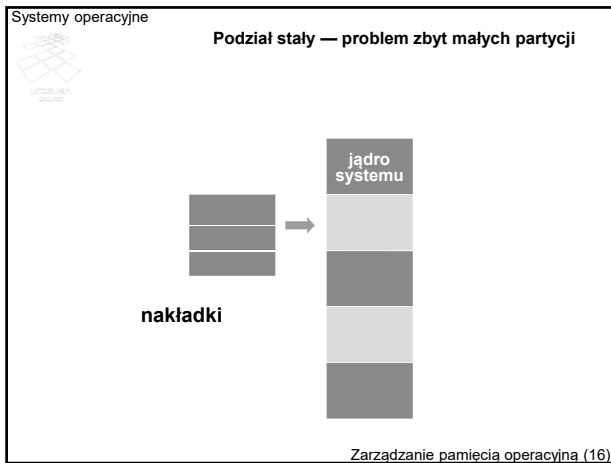


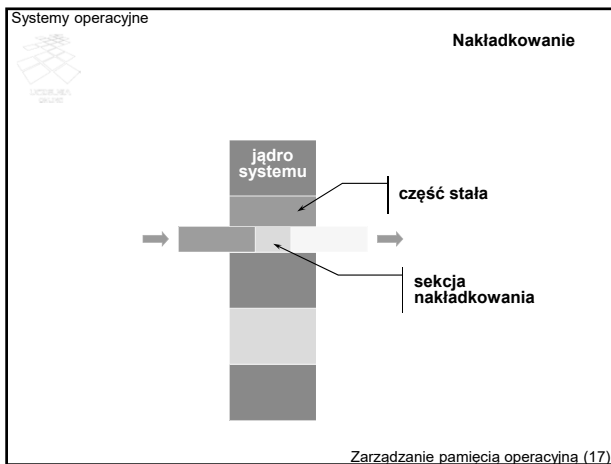


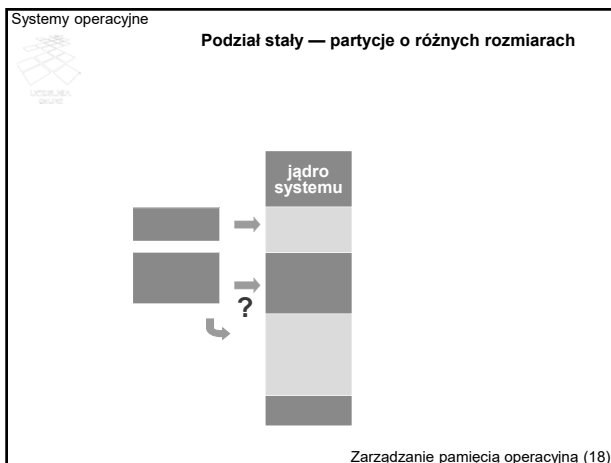












Systemy operacyjne

Podział dynamiczny

- Podział pamięci tworzony jest w czasie pracy systemu stosownie do żądań procesów.
- Proces ładowany jest w obszar o rozmiarze dosyć dokładnie odpowiadającym jego wymaganiom.
- Zalety: lepsze wykorzystanie pamięci (brak fragmentacji wewnętrznej)
- Wady: skomplikowane zarządzanie, wynikające z konieczności utrzymywania odpowiednich struktur danych w celu identyfikacji obszarów zajętych oraz wolnych.

Zarządzanie pamięcią operacyjną (19)

Systemy operacyjne

Obraz pamięci przy podziale dynamicznym

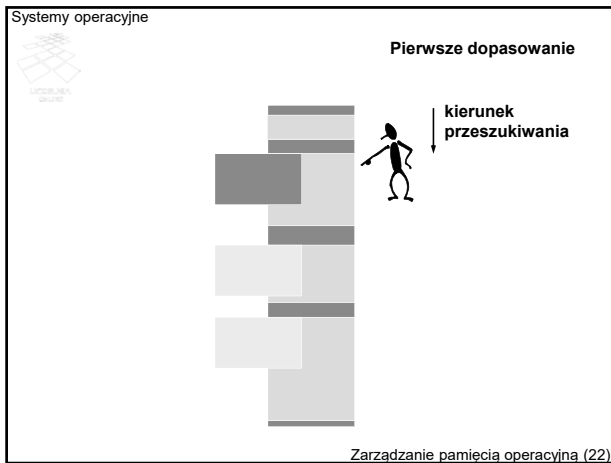
Zarządzanie pamięcią operacyjną (20)

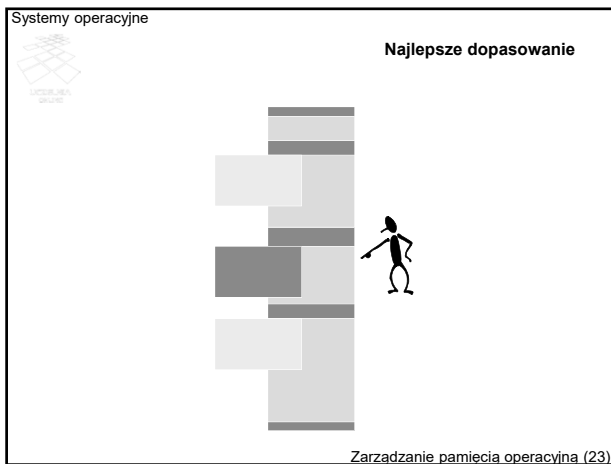
Systemy operacyjne

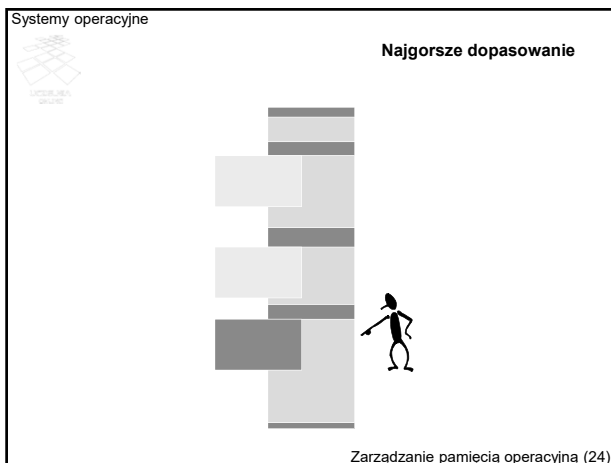
Podział dynamiczny — problem wyboru bloku

- Pierwsze dopasowanie (ang. first fit) — przydziela się **pierwszy** wolny obszar (tzw. dziurę) o wystarczającej wielkości. Poszukiwanie kończy się po znalezieniu takiego obszaru.
- Najlepsze dopasowanie (ang. best fit) — przydziela się **najmniejszy** dostatecznie duży wolny obszar pamięci. Konieczne jest przeszukanie wszystkich dziur.
- Następne dopasowanie — podobnie jak pierwsze dopasowanie, ale poszukiwania rozpoczyna się od miejsca ostatniego przydziału.
- Najgorsze dopasowanie (ang. worst fit) — przydziela się **największy** wolny obszar pamięci. Konieczne jest przeszukanie wszystkich dziur.

Zarządzanie pamięcią operacyjną (21)







Systemy operacyjne

System bloków bliźniaczych

- Pamięć dostępna dla procesów użytkownika ma rozmiar 2^U .
- Przydzielany blok ma rozmiar 2^K , gdzie $L \leq K \leq U$.
- Początkowo dostępny jest jeden blok o rozmiarze 2^U .
- Realizacja przydziału obszaru o rozmiarze s polega na znalezieniu lub utworzeniu (przez połowienie) bloku o rozmiarze 2^i takim, że $2^{i-1} < s \leq 2^i$.

Zarządzanie pamięcią operacyjną (25)

Systemy operacyjne

System bloków bliźniaczych — przykład

kolejne
żądania
przydziału

| | | | | |
|------|--------|--------|--------|--------|
| 1 MB | 512 KB | 512 KB | 512 KB | 256 KB |
| | 256 KB | 256 KB | 256 KB | 256 KB |
| | 128 KB | 128 KB | | 256 KB |
| | 128 KB | 128 KB | | 128 KB |
| | | | 128 KB | |
| | | | | 128 KB |

100 KB 240 KB 64 KB 250 KB

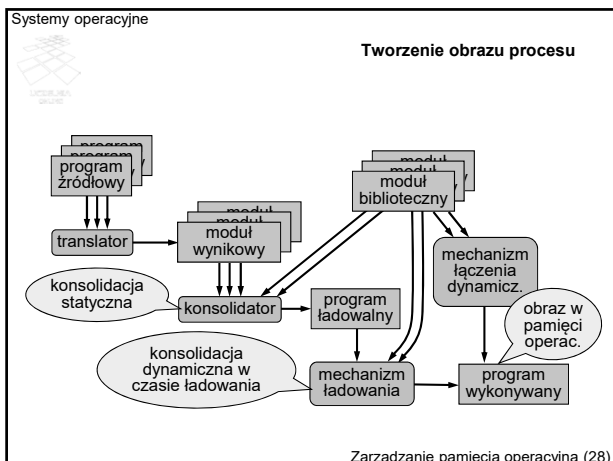
Zarządzanie pamięcią operacyjną (26)

Systemy operacyjne

Obraz procesu w pamięci

- Tworzenie obrazu
 - Kompilacja
 - Konsolidacja
 - Ładowanie kodu
- Relokacja
- Ochrona
- Współdzielenie

Zarządzanie pamięcią operacyjną (27)



Systemy operacyjne

Wiązanie i przekształcanie adresów

- W modelu von Neumana adresy dotyczą rozkazów (instrukcji) oraz zmiennych (danych).
- Jeśli w programie źródłowym występują adresy, to mają one najczęściej postać symboliczną (etykiety w assemblerze) lub abstrakcyjną (wskaźniki w C lub Pascalu).
- Adresy związane z lokalizacją pojawiają się na etapie translacji i są odpowiednio przekształcane aż do uzyskania adresów fizycznych.

Zarządzanie pamięcią operacyjną (29)

Systemy operacyjne

Translacja

- W wyniku translacji (kompilacja, asemlacja) powstaje przemieszczalny moduł wynikowy (relocatable object module), w którym wszystkie adresy liczone są względem adresu początku modułu.
- Gdyby program składał się z jednego modułu, a jego docelowa lokalizacja w pamięci była z góry znana, na etapie translacji mogłyby zostać wygenerowane adresy fizyczne.

Zarządzanie pamięcią operacyjną (30)

Systemy operacyjne

Konsolidacja

- Konsolidacja statyczna — odniesienia do innych modułów zamieniane są na adresy w czasie tworzenia programu wykonywalnego.
- Konsolidacja dynamiczna
 - w czasie ładowania — w czasie ładowania programu następuje załadowanie modułów bibliotecznych i związanie odpowiednich adresów,
 - w czasie wykonania — załadowanie modułów bibliotecznych i związanie adresów następuje dopiero przy odwołaniu się do nich w czasie wykonywania programu.

Zarządzanie pamięcią operacyjną (31)

Systemy operacyjne

Konsolidacja statyczna

- W czasie łączenia modułów przemieszczalnych w jeden program ładowalny zwany też modulem absolutnym (ang. absolute module), do adresów przemieszczalnych dodawane są wartości, wynikające z przesunięcia danego modułu przemieszczalnego względem początku modułu absolutnego.
- Gdyby docelowa lokalizacja programu w pamięci była z góry znana, na etapie tym mogłyby zostać wygenerowane adresy fizyczne.

Zarządzanie pamięcią operacyjną (32)

Systemy operacyjne


Ładowanie kodu (1)

plik z programem ładowalnym

pamięć operacyjna

Zarządzanie pamięcią operacyjną (33)

Systemy operacyjne




Ładowanie kodu ⁽²⁾

- Ładowanie absolutne — program ładowany jest w ustalone miejsce w pamięci, znane w momencie tworzenia programu ładowalnego.
- Ładowanie relokowalne — fizyczna lokalizacja procesu ustalana przy ładowaniu do pamięci.
- Ładowanie dynamiczne w czasie wykonania — fizyczna lokalizacja procesu w pamięci może ulec zmianie w czasie wykonywania.

Zarządzanie pamięcią operacyjną (34)

Systemy operacyjne




Współdzielenie pamięci

- Efektywność wykorzystania pamięci
 - współdzielenie kodu programu
 - współdzielenie kodu funkcji bibliotecznych
- Kooperacja procesów
 - synchronizacja działań procesów
 - komunikacja pomiędzy procesami (współdzielenie danych)

Zarządzanie pamięcią operacyjną (35)

Systemy operacyjne



Ochrona pamięci

- Weryfikacja poprawności adresu pod kątem zakresu dopuszczalności
- Weryfikacja praw dostępu:
 - prawa zapisu
 - prawa odczytu
 - prawa wykonania

Zarządzanie pamięcią operacyjną (36)

Systemy operacyjne

Stronicowanie

- Arbitralny podział pamięci fizycznej na ramki, w które ładowane są odpowiednie strony obrazu procesu.
- Podział logicznej przestrzeni adresowej na strony o takim samym rozmiarze, jak ramki w pamięci fizycznej.
- Zalety:
 - brak problemu fragmentacji zewnętrznej,
 - wspomaganie dla współdzielenia i ochrony pamięci.
- Wady:
 - narzut czasowy przy transformacji adresu,
 - narzut pamięciowy (na potrzeby tablicy stron),
 - fragmentacja wewnętrzna (niewielka).

Zarządzanie pamięcią operacyjną (37)

Systemy operacyjne

Stronicowanie — transformacja adresu

- Adres logiczny zawiera numer strony i przesunięcia na stronie (ang. offset), np.:

| | |
|------------------------|----------------------------|
| nr strony (22 bity) | przesunięcie (10 bitów) |
|------------------------|----------------------------|

- Transformacja adresu polega na zastąpieniu numeru strony numerem ramki.
- Odzworowanie numeru strony na numer ramki wykonywane jest za pomocą tablicy stron (ang. page table).

Zarządzanie pamięcią operacyjną (38)

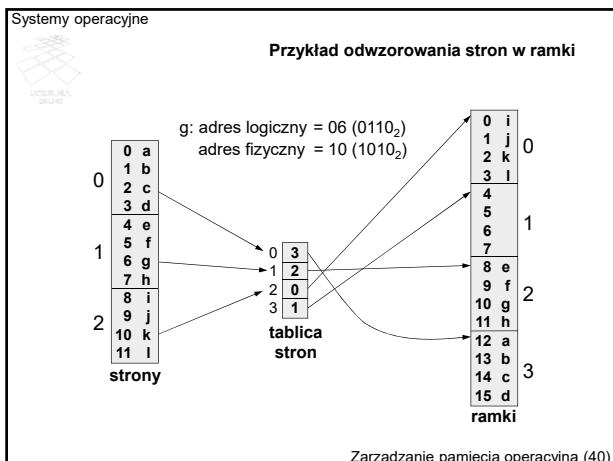
Systemy operacyjne

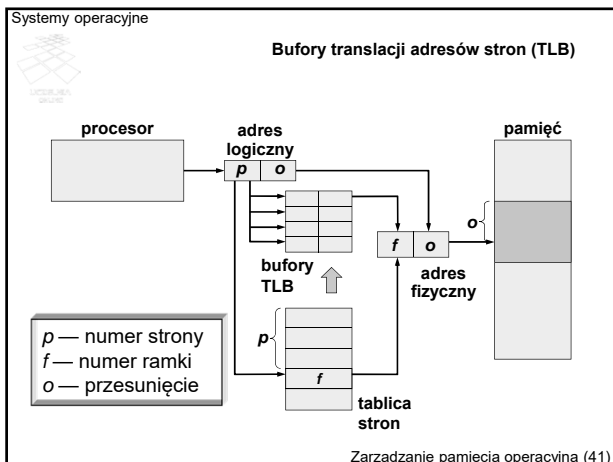
Schemat transformacji adresu w systemie pamięci stronicowanej

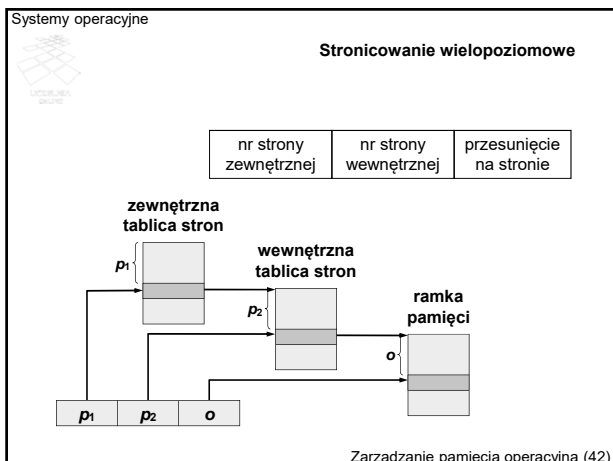
The diagram illustrates the address translation process. On the left, a box labeled 'procesor' sends an 'adres logiczny' consisting of 'p' (page number) and 'o' (offset). This 'p' is used to look up a value 'f' (frame number) in a 'tablica stron' (page table). The 'adres fizyczny' (physical address) is then formed by combining 'f' and 'o'. This physical address is used to access 'pamięć' (memory).

p — numer strony
 f — numer ramki
 o — przesunięcie

Zarządzanie pamięcią operacyjną (39)







Systemy operacyjne

Segmentacja

- Przestrzeń adresów logicznych jest postrzegana jako zbiór segmentów.
- Podstawowe atrybuty segmentu:
 - identyfikator,
 - adres bazowy,
 - rozmiar,
 - informacja o rodzaju zawartości i formach dostępu.
- Adres logiczny składa się z numeru segmentu i przesunięcia wewnątrz segmentu.
- Odwzorowanie adresów logicznych w fizyczne zapewnia tablica segmentów.

Zarządzanie pamięcią operacyjną (43)

