

Spis treści

1 System Operacyjny	2
1.1 Co robi?	2
1.2 Jaką rolę pełni system operacyjny?	2
1.3 Zasoby	2
1.4 Wywołania systemowe	2
1.5 Tryb jądra	2
1.6 Rozruch komputera	2
2 Procesor	3
2.1 Cykl rozkazowy	3
2.2 Przerwania	3
2.2.1 Przerwania zegarowe	3
2.3 Rejestry	3
2.4 Wielowątkowość	3
2.5 Potoki	3
3 Urządzenia I/O	4
4 Magistrale	4
5 Procesy	4
5.1 Tworzenie procesu	4
5.2 Wieloprogramowość	4
5.3 Zarządcy	4
5.4 Struktury danych	4
5.5 Przełączenie kontekstu	5
5.6 Kolejki procesów	5
5.7 Planiszczyki	5
5.8 Wątki	5
6 Przydział procesora	5
6.1 Funkcja priorytetu	5
6.2 Reguła arbitrażu	5
6.3 Tryb decyzji	5
6.4 Ocena algorytmu planowania przez procesor	6
6.5 Ocena algorytmu planowania przez użytkownika	6
6.6 Podstawowe algorytmy	6
7 Pamięć	6
7.1 Przestrzeń adresowa	6
7.2 Fragmentacja	6
7.3 Podział	6
7.4 Wybór odpowiedniego bloku	7
7.5 Ładowanie kodu do pamięci	7
7.6 Stronicowanie	7
7.7 Przechowywanie Tablicy Stron	7
7.8 Segmentacja	7
7.9 System hybrydowy	7
8 Problemy współbieżności	7
9 System plików	7
9.1 Dostęp do pliku	8
9.2 Układ systemu plików na dysku	8

1 System Operacyjny

System operacyjny jest warstwą oprogramowania operującą bezpośrednio na sprzęcie, której celem jest zarządzanie zasobami systemu komputerowego i stworzenie użytkownikowi środowiska łatwiejszego do zrozumienia i wykorzystania.

1.1 Co robi?

- Pośredniczy między programami a sprzętem
- Tworzy środowisko w którym operują programy użytkowników (low level interfejsy)
- Steruje komputerem i nadzoruje jego działanie
- Zarządza dostępem do zasobów

1.2 Jaką rolę pełni system operacyjny?

- System operacyjny dostarcza abstrakcji programom
- System operacyjny zarządza wszystkimi składnikami złożonego systemu
- System operacyjny zapewnia kontrolowany i uporządkowany przydział zasobów

1.3 Zasoby

- Czas procesora
- Pamięć operacyjna
- System plików
- Urządzenia wejścia wyjścia
- ...

1.4 Wywołania systemowe

Wywołania systemowe oddają kontrolę systemowi operacyjnemu za prośbą procesu i obsługują np.: dostęp do plików, tworzenie wątku itp.

1.5 Tryb jądra

Wszystkie współczesne systemy operacyjne rozróżniają tryb jądra i użytkownika. W trybie jądra operacje oraz dostęp przez procesor do zasobów nie jest ograniczony. Jest to w pewnym sensie jak najniższa warstwa jaką może osiągnąć kod. Aby operować na tej warstwie kod musi być częścią systemu operacyjnego. W trybie użytkownika dostęp kodu do zasobów jest kontrolowany przez system operacyjny.

1.6 Rozruch komputera

1. BIOS weryfikuje urządzenia i ilość pamięci
2. BIOS określa urządzenie rozruchowe z pamięci CMOS
3. BIOS wczytuje program rozruchowy
4. Program rozruchowy wczytuje system operacyjny
5. System operacyjny odczytuje listę urządzeń z konfiguracji BIOS i sterowniki

2 Procesor

Współczesne procesory operują zgodnie z architekturą von Neumanna. Ta architektura zakłada:

- Program jak i dane znajdują się w pamięci operacyjnej
- Rozkazy wykonuje się dokładnie w tej kolejności w jakiej znajdują się w pamięci (odstępstwa: instrukcje skoku, wywołania, powroty)
- Procesor przechowuje w rejestrze adres komórki pamięci zawierającej kolejną do wykonania instrukcję. W celu pobrania tej instrukcji, procesor wystawia odpowiedni adres na magistrali adresowej.

2.1 Cykl rozkazowy

1. Pobranie kodu rozkazu
2. Pobranie operandów
3. Składowanie wyniku
4. Rozpoznania przerwania

2.2 Przerwania

- Przerwania pozwalają na zatrzymanie cyklu rozkazowego procesora, dzięki czemu system operacyjny może odzyskać kontrolę lub zareagować na zdarzenie asynchroniczne
- Przerwania mogą pochodzić z trzech źródeł:
 - Zewnętrzne – I/O, zegary, układy bezpośredniego dostępu do pamięci
 - Diagnostyczne – w wyniku wykrycia nieoczekiwanego zdarzenia
 - Programowe – w wyniku wykonania instrukcji przerwania (niezależnie od trybu)
- Procesor może decydować czy przyjąć nowe przerwanie podczas obsługi innego, lub wyłączyć obsługę przerwania zupełnie
- Kolejność obsługi przerwania kontroluje kontroler przerwania

2.2.1 Przerwania zegarowe

Zwykle czasomierz generuje około 100 przerwania na sekundę. Umożliwiają one systemowi operacyjnemu analizę stanu systemu i reakcję na zmiany

2.3 Rejestry

Rejestr to znajdująca się niedaleko procesora pamięć. Procesor wykorzystuje je do przechowywania danych oraz wyników operacji. W momencie zmiany kontekstu system operacyjny zachowuje stan rejestrów w pamięci operacyjnej.

2.4 Wielowątkowość

Dzięki wielowątkowości procesor może przechować stan procesora, i potem do niego wrócić. W ten sposób możemy przechowywać stan kilku procesów (wątków), dając iluzję jednoczesności z wystarczająco szybkim przełączaniem. Wielowątkowość też pozwala zachować czas procesora w momentach w których jakiś wątek czeka.

2.5 Potoki

Współczesne procesory dzięki specjalnych architekturach mogą jednocześnie wykonywać etapy cyklu rozkazowego. Np. w trakcie wykonywania polecenie pobrać kod nowego polecenia

3 Urządzenia I/O

Na urządzenie składa się kontroler i samo urządzenie. System operacyjny kontaktuje się przy pomocy sterownika z kontrolerem dzięki czemu ma dostęp do urządzenia. Realizację operacji I/O można przeprowadzić na trzy sposoby. Procesor może czekać aktywnie, robiąc figuratywnie nic. Lepszym rozwiązaniem jest wykorzystanie przerwania, gdzie kontroler na końcu operacji generuje przerwanie. Alternatywnie przy pomocy systemu DMA, gdzie kontroler zwraca dane do DMA, a DMA po zakończeniu czytania generuje przerwanie.

4 Magistrale

Magistrale w komputerze służą do dystrybucji danych w komputerze. Zazwyczaj wyróżniamy trzy równoległe kanały:

- magistrala adresowa - przekazuje adres komórek z operacjami
- magistrala sterująca - przekazuje rodzaj wykonywanej operacji
- magistrala danych - przekazuje dane dot. operacji

5 Procesy

Proces to istota systemu operacyjnego, reprezentująca program wykonywujący się. Każdy proces ma swoją przestrzeń adresową, czyli wydzielony zakres w pamięci. Przestrzeń adresowa przechowuje program, dane, zbiór zasobów oraz blok kontrolny który opisuje stan procesu.

5.1 Tworzenie procesu

Na żądanie procesu nowy proces może zostać stworzony. Nowy proces może współdzielić część zasobów z procesem rodzicem. Usunięcie procesu oznacza zwolnienie jego zasobów.

5.2 Wieloprogramowość

W idealnym świecie żaden z programów nie musiałby czekać na operacje I/O. Lecz niestety tak nie jest. Obecność innych programów i obsługa innych programów umożliwia systemowi zmianę kontekstu w momencie blokady I/O.

5.3 Zarządcy

Procesy współdzielą zasoby dlatego konieczne jest wdrożenie procedur zarządzania procesami i zasobami. Wyróżniamy dwa rodzaje zarządców: zasobów i procesów.

5.4 Struktury danych

Koniecznym jest, aby gdzieś zarządcy zapisywali swoje dane. W tym celu istnieje kilka struktur wykorzystywanych konkretnie do zapisywania wiedzy o procesie i zasobach.

- Deskryptor procesu - używany przez zarządcę procesów w celu rejestrowania stanu procesu
- Tabela procesów - Zbiór deskryptorów
- Deskryptor zasobu - przechowuje informacje o dostępności zasobu

W większość systemów wyróżniamy kilka stanów procesu, z reguły przechowywanych w deskrytorze procesu.

- Nowy
- Wykonywany
- Oczekujący - wykonywanie zostało zakończone, ale wciąż czeka na coś
- Gotowy - oczekiwanie zostało zakończone
- Zakończony - zabity
- Zombie - nie może być zabity, bo ma dzieci

5.5 Przełączenie kontekstu

Zadania wykonywane są na procesorze w krótkich cyklach z czasem przypisanym do 1 zadania. Po końcu cyklu następuje przełączenie kontekstu. Przełączenie kontekstu polega na zachowaniu stanu i kontekstu procesu i załadowaniu nowego kontekstu. Na kontekst składa się stan tych zasobów z których proces korzysta.

5.6 Kolejki procesów

Procesy są organizowane w kolejki. Wyróżniamy kilka rodzajów tych kolejek.

- Kolejka zadań
- Kolejka procesów gotowych
- Kolejka do urządzenia
- Kolejka oczekujących na sygnał synchronizacji

5.7 Planiści

Planiści to programy szeregujące wybierające procesu z kolejek w taki sposób aby dążyć do optymalizacji. Większość systemów wyróżnia następujących planistów:

- Planista krótkoterminowy - zajmuje się przydziałem do procesora procesów gotowych
- Planista średnioterminowy - Zajmuje się wymianą procesów pomiędzy pamięcią główną a pamięcią zewnętrzną. Proces w pamięci operacyjnej nazywamy aktywnym a ten w przestrzeni wymiany zawieszonym
- Planista długoterminowy – Zajmuje się ładowaniem nowych programów do pamięci i kontrolą liczby zadań w systemie

5.8 Wątki

Wątki to lekkie procesy, znajdujące się w dużym procesie. Współdzielą przestrzeń adresową i większość zasobów. Wątki mogą być realizowane przez system operacyjny, lub na poziomie procesu, poprzez przełączanie egzekucji. Obydwa te rozwiązania mają swoje wady i zalety, gdzie realizacja wątków przez system jest bardziej sprawiedliwe, a przez proces szybsze.

6 Przydział procesora

Procesor to cenny zasób dlatego decyzja który proces będzie wykonywany następnie jest niezwykle istotna. Planiści podejmują decyzje, a ekspedytor realizuje ich decyzje i przełącza kontekst.

6.1 Funkcja priorytetu

To funkcja zwracająca numeryczny priorytet procesu. Może mieć różne argumenty w zależności od rodzaju systemu operacyjnego. Argumentami mogą być np.: czas oczekiwania, czas obsługi, priorytet zewnętrzny etc.

6.2 Reguła arbitrażu

Określa co w przypadku remisu. Istnieje wiele podejść do reguły arbitrażu, szczególnie: losowy wybór, cykliczne wykonywanie i chronologiczny – wygrywa starszy proces.

6.3 Tryb decyzji

Określa czy decyzja jest ostateczna czy nie. Rozróżniamy dwa tryby decyzji: wyłączeniowy lub nie.

6.4 Ocena algorytmu planowania przez procesor

- Wykorzystanie procesora – procent czasu przez który procesor przetwarza dane
- Przepustowość – liczba procesów zakończonych w czasie
- Sprawiedliwość
- Respektowanie priorytetów
- Równoważne obciążenie zasobów

6.5 Ocena algorytmu planowania przez użytkownika

- Czas odpowiedzi – czas pomiędzy przedłożeniem zadania a odpowiedzią
- Czas opóźnienia – czas od linii krytycznej do momentu zakończenia wykonywania
- Czas cyklu przetwarzania – ile czasu trwa skończenie konkretnego procesu
- Przewidywalność – czas realizacji żądania niezależnie od obciążenia

6.6 Podstawowe algorytmy

- FIFO
- LIFO
- SJF - shortest job first
- RR - round robin - karuzela
- SRT - shortest remaining time

7 Pamięć

Pamięć ma wiele rodzajów i typów. Na podstawie tych charakterystyk system operuje na hierarchii pamięci, z najszybszą, a jednocześnie najmniejszą pamięcią jak najbliższej procesora.

7.1 Przestrzeń adresowa

Rozróżniamy przestrzeń fizyczną i logiczną. Procesor i system operuje na przestrzeni fizycznej. Z kolei system operacyjny wydziela fragmenty pamięci fizycznej procesom, tworząc im wirtualne pełne przestrzenie adresowe.

7.2 Fragmentacja

W wyniku wielokrotnych operacji na pamięci może dochodzić do fragmentacji, kiedy pamięć nie jest równomiernie i ciągle wykorzystywana, nie pozwalając na nowe alokacje pamięci. Fragmentacja może być wewnętrzna, kiedy proces wykorzystuje więcej pamięci niż potrzebuje, oraz zewnętrzna, kiedy między blokami są luki.

7.3 Podział

Bloki pamięci muszą być jakoś przydzielane procesom. W jaki sposób zapamiętywać, i dzielić pamięć aby ten proces minimalizował fragmentację, i aby struktura danych wewnętrzna dot. bloków była optymalna? Jest kilka głównych strategii:

- Podział pamięci na partycje o stałym rozmiarze
- Podział pamięci na partycje o różnym rozmiarze
- Podział dynamiczny w którym pamięć jest przydzielana z pewną dokładnością (np w blokach 1024 B)
- Podział na bloki bliźniacze, polegający na połowieniu zbyt dużych obszarów pamięci

7.4 Wybór odpowiedniego bloku

- First fit – przydziela się pierwszy znaleziony wystarczająco duży obszar
- Best fit – znajduje się najmniejszy pasujący obszar i się przydziela procesowi
- Next fit – przydziela się pierwsze pasujące miejsce po ostatnim przydzieleniu
- Worst fit – przydziela się początkowy fragment największego wolnego obszaru

7.5 Ładowanie kodu do pamięci

- Ładowanie absolutne – na etapie tworzenia programu ładowalnego znane było dokładne miejsce w pamięci; zawiązanie adresów fizycznych zostało wykonane na etapie konsolidacji
- Ładowanie relokowalne – fizyczna lokalizacja procesu w pamięci ustalana jest na etapie ładowania. Konieczne jest przeliczenie adresów
- Ładowanie dynamiczne w czasie wykonania – adresy są ustalane przez MMU przy odwołaniu się do pamięci

7.6 Stronicowanie

Pamięć fizyczna jest dzielona na ramki o pewnym rozmiarze. Ramkom fizycznym odpowiadają logiczne strony. Adres jest podzielony na dwie części: adres strony i przesunięcie.

7.7 Przechowywanie Tablicy Stron

Może się okazać że tablica stron jest za duża na ramkę. Dlatego wprowadza się system hierarchiczny gdzie adres jest dzielony na części. Pierwsza część adresu adresuje do strony tablic-tablic stron, itd.

7.8 Segmentacja

Segmentacja to dynamiczny wariant stronicowania. Obszar procesu jest dzielony na spójne segmenty o określonym rozmiarze. Rozmiar segmentu jest dynamiczny. Tak jak przy stronach segmenty też mają swoją tablicę.

7.9 System hybrydowy

Aby czerpać zalety stronicowania (małą fragmentacją zewnętrzną i szybkość działania) i segmentacji (mniejsza fragmentacja wewnętrzną) można zbudować system hybrydowy. Pamięć składa się z segmentów. Segmenty te składają się z ramek. W adresie logicznym przesunięcie w segmencie jest traktowane jako adres ramki. Zamiast adresu bazowego segmentu jest tablica ramek.

8 Problemy współbieżności

Jednoczesne działanie kilku procesów, które mogą chcieć mieć dostęp do tych samych zasobów, tworzy nową klasę problemów. Można je rozwiązać przy pomocy następujących narzędzi:

- Semafor – int który ogranicza ilość procesów mających dostęp do zasobu. 1 zasób 1 semafor
- Algorytm Petersona – algorytm zarządzający dwoma procesami tak aby tylko jeden mógł mieć dostęp
- Monitor – zbiór zmiennych i procedur do których tylko jeden proces ma dostęp. Zmienne widzi każdy proces ale procedury tylko proces wewnątrz. Kolejnością wejścia zarządza kolejka

9 System plików

Plik to abstrakcyjny obraz informacji gromadzonej i udostępnianej przez system komputerowy. System operacyjny ma zapewnić identyfikację pliku, interfejs do pliku oraz zapewnić bezpieczeństwo i efektywność operacji na pliku. Struktura fizyczna pliku ma określać jak dane są przechowywane i jest narzucana przed nośnik. Struktura logiczna pliku z kolei ma określać informację wewnątrz pliku.

9.1 Dostęp do pliku

- Sekwencyjny – informacje z pliku są przetwarzane rekord po rekordzie
- Bezpośredni – informacje z pliku są przetwarzane w sposób dowolny
- Indeksowy – rekordy są identyfikowane przez hash, co umożliwia zczytywanie z pliku jak z struktur danych

9.2 Układ systemu plików na dysku

Dysk składa się z partycji. Na początku dysku jest MBR i tabela partycji. Potem są partycje. Rekord MBR odpowiada za wyszukanie aktywnej partycji i odczytanie jej bloku rozruchowego. Za blokiem rozruchowym jest superblok który odpowiada za najważniejsze parametry systemu plików. Potem występują informacje o blokach wolnych np. tabela FAT. Następnie są i-węzły czyli struktury po jednej na plik. I na koniec tuż przed samymi danymi plików jest opis katalogu root.