



SYSTEMY OPERACYJNE

POJĘCIE I BUDOWA SYSTEMU KOMPUTEROWEGO

SYSTEM KOMPUTEROWY

System komputerowy (ang. computer system) – układ współdziałania dwóch składowych: sprzętu komputerowego oraz oprogramowania, działających coraz częściej również w ramach sieci komputerowej. Można mówić o następujących poziomach takiego systemu: sprzęt komputerowy, system operacyjny (oprogramowanie systemowe), oprogramowanie użytkowe (aplikacje). W pełni zautomatyzowany system komputerowy działa bez udziału człowieka.

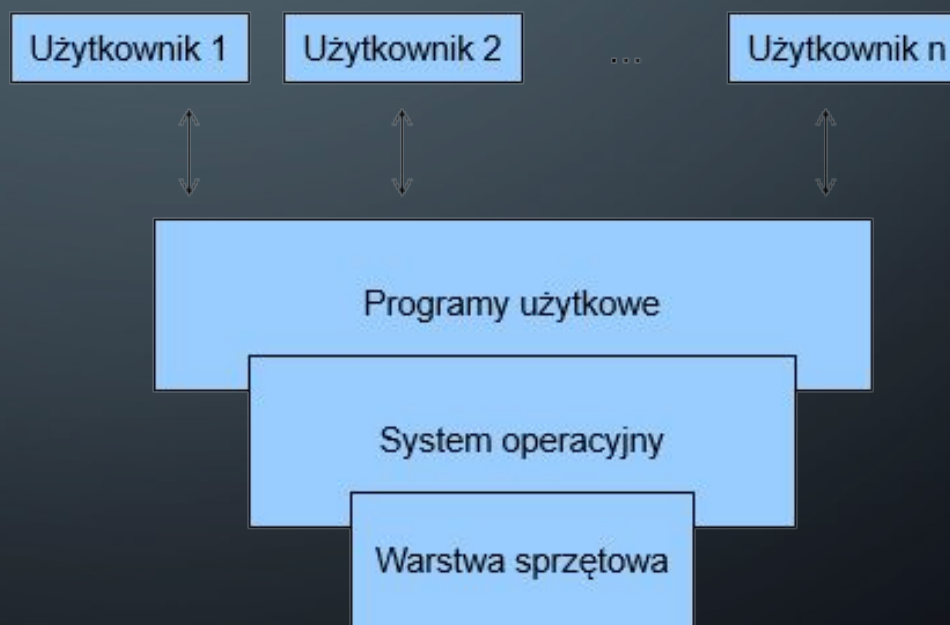
SYSTEMY KOMPUTEROWE

Organizacja systemu komputerowego to opis zależności sprzętowych, przedstawienie poszczególnych podzespołów komputera który funkcjonuje według pewnych reguł i zasad, współpracuje ze sobą – by osiągnąć określony cel. Organizacja systemu komputerowego określa zasady, reguły, cele oraz sposób wspomaganie działań poszczególnych podzespołów.

WARSTWY SYSTEMU KOMPUTEROWEGO

Struktura systemu komputerowego podzielona została na warstwy. Każda z warstw realizuje odmienne zadania. Wyróżnia się następujące warstwy:

- Warstwa sprzętowa
- System operacyjny
- Programy narzędziowe
- Programy użytkowe
- Użytkownicy



WARSTWA SPRZĘTOWA

Warstwa sprzętowa zapewnia podstawowe możliwości obliczeniowe. Zestaw komputerowy składa się z:

- Jednostki centralnej
- Urządzeń wejścia
- Urządzeń wyjścia

Ze względu na przeznaczenie, urządzenia zestawu dzielimy na urządzenia wejścia (np. klawiatura, mysz) i urządzenia wyjścia (np. drukarka, monitor).

W jednostce centralnej znajdują się najważniejsze elementy odpowiedzialne za prawidłową pracę komputera, między innymi płyta główna, na której umieszczone są elementy niezbędne do prawidłowej pracy komputera: procesor, pamięć i wiele innych.

OPROGRAMOWANIE SYSTEMOWE

Oprogramowanie systemowe – kontroluje i koordynuje użycie zasobów sprzętowych poprzez różne programy użytkowe. Warstwa tworzona poprzez twórców systemu operacyjnego.

OPROGRAMOWANIE NARZĘDZIOWE

Oprogramowanie narzędziowe – wspomaga zarządzanie zasobami sprzętowymi poprzez dogodne interfejsy użytkowe oraz usprawnia, modyfikuje oprogramowanie systemowe, zazwyczaj pisane przez niezależnych programistów którzy mają na celu usprawnienia wykonywania programów w bardziej wygodny i wydajny sposób, a przy tym często eliminują błędy czy też niedociągnięcia oprogramowania systemowego.

OPROGRAMOWANIE UŻYTKOWE

Oprogramowanie użytkowe – określają sposoby, w jakie zostają użyte zasoby systemowe do rozwiązywania problemów obliczeniowych zadanych przez użytkownika (kompilatory, systemy baz danych, gry, oprogramowanie biurowe).

UŻYTKOWNICY

Użytkownicy – ludzie, maszyny, inne komputery, które mają bezpośredni kontakt z oprogramowaniem użytkowym. Realizują różne zadania za pomocą programów użytkowych na sprzęcie komputerowym pod nadzorem systemu operacyjnego.

HARDWARE

Sprzęt komputerowy (ang. hardware) – materialna część komputera. Ogólnie hardwarem nazywa się sprzęt komputerowy jako taki i odróżnia się go od software'u – czyli oprogramowania.

Podział ten jest nieostry, gdyż współcześnie wiele elementów sprzętu komputerowego posiada "wszyte" weń na stałe oprogramowanie, stanowiące jego integralną część, bez którego elementy te nie mogłyby funkcjonować. Np. większość drukarek komputerowych posiada w swojej pamięci zestaw komend, przy pomocy których realizuje proces drukowania i których odpowiednik znajduje się w pamięci komputera stanowiąc programowy sterownik tego urządzenia. Wiele urządzeń – typu karty graficzne, płyty główne posiada własne oprogramowanie nazywane BIOS-em. W stosunku do oprogramowania niektórych urządzeń używa się słowa firmware.

RODZAJE PAMIĘCI

W komputerze stosowane są różne rodzaje pamięci. Podstawowe znaczenie ma pamięć RAM, która umożliwia zapisywanie i odczytywanie informacji. Oprócz niej jest stosowana pamięć ROM, służąca jedynie do odczytywania informacji.

Pamięć ROM – pamięć nieulotna, w której umieszczane są informacje stałe. W pamięci ROM zapisane są fabrycznie wszystkie procedury potrzebne do uruchomienia komputera i jego obsługi. Procedury te są zapisane w BIOS-ie.

Pamięć RAM – pamięć używana do ładowania programów oraz wszystkich danych wprowadzanych za pomocą klawiatury lub innych urządzeń wejściowych przez użytkownika. Każde wyłączenie zasilania komputera powoduje jej oczyszczenie. Jest to pamięć ulotna.

RODZAJE PAMIĘCI

Pamięci masowe – pamięci trwałe, do długotrwałego przechowywania danych na nich zapisanych. Należą do nich nośniki magnetyczne, np. dyski twarde, taśmy magnetyczne, napędy optyczne oraz pamięci flash.

Pamięć podręczna (cache) – dodatkowa pamięć, która jest schowkiem danych podręcznych dla procesora. Najczęściej przechowywane są w niej dane używane do obliczeń, co przyczynia się do zwiększenia prędkości pracy komputera.

STRUKTURA PAMIĘCI

Programy muszą znajdować się w pamięci operacyjnej, aby mogły być wykonywane. Pamięć operacyjna jest jedynym wielkim obszarem pamięci dostępnym dla procesora bezpośrednio. Tworzy ona tablicę słów lub bajtów, których liczba waha się od setek tysięcy do setek milionów. Każde słowo ma własny adres. Współpraca z pamięcią operacyjną odbywa się za pomocą ciągu rozkazów **load** (pobierz) lub **store** (przechowaj) odnoszących się do określonych adresów. Rozkaz **load** powoduje pobranie słowa z pamięci operacyjnej do wewnętrznego rejestru jednostki centralnej, natomiast rozkaz **store** powoduje umieszczenie zawartości rejestru procesora w pamięci operacyjnej. Oprócz jawnych pobrań i przechowań jednostka centralna automatycznie pobiera z pamięci operacyjnej rozkazy do wykonania.

STRUKTURA PAMIĘCI

Typowy cykl rozkazowy w systemie o architekturze von Neumanna zaczyna się od pobrania rozkazu z pamięci i przesłania go do rejestru rozkazów. Rozkaz jest następnie dekodowany i realizowany. Po wykonaniu rozkazu na argumentach jego wynik można z powrotem przechować w pamięci. Zauważmy, że jednostka pamięci „widzi” tylko strumień adresów pamięci. Nie jest jej znany sposób, w jaki one powstały (licznik rozkazów, indeksowanie, modyfikacje pośrednie, adresy literalne itp.) ani czemu służą (rozказы lub dane). Z uwagi na to możemy zaniedbać sposób generowania adresu pamięci przez program. Interesujemy się tylko ciągiem adresów pamięci wytwarzanych przez wykonywany program.

STRUKTURA PAMIĘCI

W idealnych warunkach moglibyśmy sobie życzyć, aby program i dane stale pozostawały w pamięci operacyjnej. Nie jest to możliwe z dwu powodów:

- Pamięć operacyjna jest zazwyczaj za mała, aby przechowywać na stałe wszystkie potrzebne programy i dane.
- Pamięć operacyjna jest tzw. pamięcią ulotną (nie trwałą). Traci ona swoją wartość po odłączeniu zasilania.

Wobec tego większość systemów komputerowych jest wyposażona w pamięć pomocniczą, która rozszerza pamięć operacyjną. Od pamięci pomocniczej wymaga się przede wszystkim, aby mogła trwale przechowywać duże ilości danych.

STRUKTURA PAMIĘCI

Najpopularniejszym urządzeniem pamięci pomocniczej jest dysk magnetyczny umożliwiający zapamiętywanie zarówno programów, jak i danych. Większość programów przechowuje się na dysku, zanim nie nastąpi ich umieszczenie w pamięci operacyjnej. Wiele programów używa potem dysku zarówno jako źródła, jak i miejsca przeznaczenia przetwarzanych przez siebie informacji. Dlatego też właściwe zarządzanie pamięcią dyskową ma zasadnicze znaczenie w systemie komputerowym. W szerszym jednak rozumieniu przedstawiona przez nas struktura pamięci, składająca się z rejestrów, pamięci operacyjnej i dysków magnetycznych, jest tylko jednym z wielu możliwych systemów pamięci. Istnieją także pamięci podręczne, pamięci na płytach kompaktowych, taśmy magnetyczne itd. Każdy system pamięci spełnia podstawowe funkcje przechowywania danych do czasu, gdy zostaną one z niego odzyskane. To, co przede wszystkim różni te systemy pamięci, to szybkość działania, koszt, rozmiar i ulotność.

PAMIĘĆ OPERACYJNA

Pamięć operacyjna oraz rejestry wbudowane w procesor są jedynymi rodzajami pamięci dostępnej dla jednostki centralnej bezpośrednio.

Istnieją rozkazy, których argumentami są adresy pamięci operacyjnej, lecz nie ma rozkazów posługujących się adresami dyskowymi. Z tego powodu każdy wykonywany rozkaz i wszystkie używane przez niego dane muszą znajdować się w jednym z tych urządzeń pamięci o dostępie bezpośrednim. Jeżeli danych nie ma w pamięci operacyjnej, to należy je do niej sprowadzić, zanim jednostka centralna zacznie je przetwarzać.

PAMIĘĆ OPERACYJNA

Rejestry wbudowane w jednostkę centralną są na ogół dostępne w jednym cyklu jej zegara. Większość procesorów może dekodować rozkazy i wykonywać proste działania na zawartości rejestrów z szybkością jednej lub więcej operacji na jeden impuls zegara. Nie można tego powiedzieć o pamięci operacyjnej, do której dostęp odbywa się za pośrednictwem transakcji z szyną pamięci. Dostęp do pamięci może zajmować wiele cykli, a wtedy procesor zazwyczaj musi utykać (ang. stall), gdyż brakuje mu danych do zakończenia rozkazu, który właśnie wykonuje. Jest to sytuacja nie do przyjęcia, zważywszy na częstość kontaktów z pamięcią. Ratunkiem jest wstawienie między jednostkę centralną a pamięć operacyjną jakiejś szybkiej pamięci. Bufor pamięci stosowany do niwelowania różnic w szybkości nazywa się pamięcią podręczną (ang. cache).

DYSKI MAGNETYCZNE

Dyski magnetyczne stanowią zdecydowaną większość pamięci pomocniczych współczesnych systemów komputerowych. Zasada działania dysków jest stosunkowo prosta. Każda płyta dysku ma kształt kolisty, jak płyta kompaktowa. Średnice popularnych płyt wahają się w przedziale od 1,8 do 5,25 cala. Obie powierzchnie płyty są pokryte materiałem magnetycznym, podobnym do stosowanego na taśmach magnetycznych. Informacje przechowuje się przez odpowiednie namagnesowanie warstwy magnetycznej. Głowice unoszą się tuż nad powierzchnią każdej płyty. Są one przymocowane do ramienia dysku, które przemieszcza je wszystkie jednocześnie. Powierzchnia płyty jest logicznie podzielona na koliste ścieżki, które z kolei dzielą się na sektory. Zbiór ścieżek przy danym położeniu ramienia tworzy cylinder. Na dysku mogą być tysiące koncentrycznych cylindrów, a każda ścieżka może zawierać setki sektorów.

DYSKI MAGNETYCZNE

Podczas pracy dysk wiruje z dużą prędkością, wprawiany w ruch przez silnik jego napędu. Szybkość dysku jest określana przez dwa czynniki. Tempo przesyłania oznacza szybkość, z jaką dane przepływają między napędem dysku a komputerem. Na czas ustalania położenia głowicy, niekiedy nazywany czasem losowego dostępu składają się: czas przesuwania głowicy do odpowiedniego cylindra, nazywany czasem wyszukiwania, oraz czas, w którym potrzebny sektor, obracając się, przejdzie pod głowicą. Czas ten jest nazywany opóźnieniem obrotowym. Typowe dyski mogą przesyłać kilka megabajtów danych na sekundę, a ich czasy wyszukiwania i opóźnienia obrotowe wynoszą kilka milisekund. Ponieważ głowica dysku unosi się na niezwykle cienkiej poduszce powietrznej (mierzonej w mikronach), istnieje niebezpieczeństwo zetknięcia się jej z powierzchnią dysku. Choć płyty dysku są powleczone warstwą ochronną, czasami głowica uszkadza powierzchnię magnetyczną.

HIERARCHIA PAMIĘCI



HIERARCHIA PAMIĘCI

Różne rodzaje pamięci w systemie komputerowym można zorganizować w hierarchię zależnie od ich szybkości i kosztów. Na najwyższych poziomach pamięci są drogie, za to szybkie. W miarę przemieszczania się w dół hierarchii, ogólnie biorąc maleje cena jednego bitu, natomiast wydłuża się czas dostępu. Ten kompromis jest zrozumiały; gdyby dany system pamięci był zarówno szybszy, jak i tańszy od innych, przy takich samych innych właściwościach, to nie byłoby sensu używać pamięci wolniejszej i droższej. W rzeczywistości wiele wczesnych urządzeń pamięci, wliczając w to taśmę papierową i pamięci rdzeniowe, trafiło do muzeów z chwilą, gdy taśma magnetyczna i pamięci półprzewodnikowe stały się szybsze i tańsze.

HIERARCHIA PAMIĘCI

Oprócz szybkości i kosztu różnych systemów pamięci uwzględnia się także jej ulotność. Pamięć ulotna traci zawartość po odłączeniu od niej zasilania. W razie braku drogich, rezerwowych źródeł zasilania (baterii lub generatorów) w celu bezpiecznego przechowywania dane należy zapisywać w pamięciach nieulotnych. W hierarchii pokazanej na rys. 2.6 systemy pamięci leżące powyżej różnych typów dysków są ulotne, a te, które znajdują się poniżej pamięci operacyjnej, są nieulotne. Projektując pełny system pamięci, należy równoważyć wszystkie te czynniki. Drogiej pamięci używa się tylko w niezbędnych ilościach, natomiast pamięci taniej i nieulotnej dostarcza się w ilościach możliwie dużych. W celu niwelowania różnic w wydajności, tam, gdzie występują długie czasy dostępu lub dysproporcje w szybkości przesyłania między dwoma składowymi, można instalować pamięci podręczne.

PRZECHOWYWANIE PODRĘCZNE

Stosowanie pamięci podręcznej jest ważną zasadą przy projektowaniu systemów komputerowych. W normalnych warunkach informacje są przechowywane w jakimś systemie pamięci (np. w pamięci operacyjnej). Przed ich użyciem są kopiowane do szybszego systemu pamięci - tj. do pamięci podręcznej - na okres przejściowy. Gdy jest potrzebny jakiś fragment informacji, wtedy sprawdza się najpierw, czy nie ma go w pamięci podręcznej. Jeśli jest, informacje pobiera się wprost z pamięci podręcznej; jeśli zaś nie, to korzysta się z informacji w głównym systemie pamięci, umieszczając ich kopię w pamięci podręcznej przy założeniu, że istnieje duże prawdopodobieństwo, że będą one znów potrzebne.

PRZECHOWYWANIE PODRĘCZNE

Zarządzanie pamięcią podręczną jest ważnym zagadnieniem projektowym ze względu na ograniczone rozmiary tych pamięci. Staranny dobór wielkości pamięci podręcznej i polityki zastępowania w niej informacji może spowodować, że 80 do 99% wszystkich dostępow będzie się odnosić do pamięci podręcznej, co w dużym stopniu usprawni działanie systemu.

Przemieszczanie informacji między poziomami hierarchii pamięci może być jawne lub niejawne - zależnie od konstrukcji sprzętu i nadzoru ze strony oprogramowania systemu operacyjnego. Na przykład przesyłanie danych z pamięci podręcznej do jednostki centralnej i rejestrów jest zwykle funkcją sprzętową, nie wymagającą żadnej interwencji ze strony systemu operacyjnego. Z kolei przesyłanie danych z dysku do pamięci operacyjnej jest zazwyczaj nadzorowane przez system operacyjny.

JEDNOSTKI PAMIĘCI

Informacja w komputerze jest zapamiętywana inaczej niż zapamiętują ludzie. Układy pamięci i inne układy komputera rozróżniają jedynie wartości, oznaczone symbolami 1 i 0. Taki sposób interpretowania umownych wartości 0 i 1 został powszechnie przyjęty i z tego względu ilość pamięci odpowiadającą jednej z tych wartości jest jednostką i nazywa się **bitem**. Oznacza się ją symbolem **b**.

Bit to najmniejsza jednostka informacji. Jednostką większą od bitu jest **bajt** – liczba złożona z 8 bitów. Oznacza się ją symbolem **B**.

Do wyrażania pojemności pamięci używa się większych jednostek takich jak: kilobajty (kB), megabajty (MB), gigabajty (GB) czy terabajty (TB).

JEDNOSTKI TRANSMISJI DANYCH

Istnieją różne typy transmisji danych. Transfer może odbywać się pomiędzy poszczególnymi komponentami wewnątrz komputera lub między różnymi komputerami za pośrednictwem sieci, linii telefonicznych, modemu itp. Podstawową jednostką prędkości przesyłania danych jest bit na sekundę. Jednostka ta może być zapisywana jako **bps** lub **b/s**.

kb/s (kbps) – kilobity na sekundę

Mb/s (Mbps) – megabity na sekundę

Gb/s (Gbps) – gigabity na sekundę

Ponieważ wielkość plików zwykle podawana jest w bajtach, należy uwzględnić różnice w jednostkach. Aby obliczyć prędkość pobierania plików, których rozmiar podany jest w bajtach, prędkość przesyłania należy podzielić przez 8, czyli transfer 256 kb/s – jest równoznaczny pobieraniu 32 kB/s.

The image features a dark blue background with white, stylized circuit board traces in the corners. These traces consist of straight lines and right-angle turns, ending in small circles that represent components or connection points. The patterns are located in the top-left, top-right, bottom-left, and bottom-right corners, framing the central text.

DZIĘKUJĘ ZA UWAGĘ