

Plan wykładu	Wieloprocusorowość	Systemy równoległe	Systemy wielokomputerowe	Pytania
	○○○ ○○○○○ ○○○	○○○ ○○○○○ ○	○ ○○○○○	

Architektura komputerów

Systemy wieloprocusorowe

prof. dr hab. inż. Joanna Józefowska

Poznań, rok akademicki 2010/2011

prof. dr hab. inż. Joanna Józefowska

Architektura komputerów

Plan wykładu	Wieloprocusorowość	Systemy równoległe	Systemy wielokomputerowe	Pytania
	●○○ ○○○○○ ○○○	○○○ ○○○○○ ○	○ ○○○○○	

Ograniczenie związane z prędkością światła

- 20 cm/ns w przewodzie miedzianym,
- przy częstotliwości 10 GHz dystans pokonywany przez sygnał nie może przekroczyć 2cm, a przy 100 GHz - 2mm.

prof. dr hab. inż. Joanna Józefowska

Architektura komputerów

Plan wykładu	Wieloprocusorowość	Systemy równoległe	Systemy wielokomputerowe	Pytania
	○○○ ○○○○○ ○○○	○○○ ○○○○○ ○	○ ○○○○○	

Plan wykładu

- 1 Plan wykładu
- 2 Wieloprocusorowość
 - Ograniczenia zwiększania szybkości procesorów
 - Kilka podobnych pojęć
 - Klasyfikacja systemów wieloprocusorowych
- 3 Systemy równoległe
 - Procesory symetryczne
 - Procesory asymetryczne
 - Pamięci systemów wieloprocusorowych
- 4 Systemy wielokomputerowe
 - Definicja
 - Przykłady
- 5 Pytania

prof. dr hab. inż. Joanna Józefowska

Architektura komputerów

Plan wykładu	Wieloprocusorowość	Systemy równoległe	Systemy wielokomputerowe	Pytania
	●○○ ○○○○○ ○○○	○○○ ○○○○○ ○	○ ○○○○○	

Odprowadzanie ciepła

- Moc rozpraszana w procesorze jest (w dużym uproszczeniu) proporcjonalna do iloczynu (a) liczby tranzystorów (b) kwadratu napięcia zasilania (c) częstotliwości taktowania.
- Jeżeli aktualne trendy (stała powierzchnia chipu, a liczba tranzystorów i częstotliwość taktowania rosną wykładniczo) mają się utrzymać, to stosunek rozpraszanej mocy do powierzchni rozpraszającej tę moc całkiem niedługo przekroczy wartość charakterystyczną dla reaktora jądowego i powierzchni Słońca.
- Np. Pentium 4 Prescott rozprasza max ~89W mocy.

WNIOSEK: trend do umieszczania wielu procesorów w jednej strukturze.

prof. dr hab. inż. Joanna Józefowska

Architektura komputerów

Co to jest?

Wieloprocessorowość

jest to skoordynowane wykonywanie programów przez więcej niż jeden procesor w komputerze lub kilka komputerów.

Wieloprocessorowość generalnie może oznaczać:

- dynamiczne przyporządkowanie programu do jednego z dwóch lub więcej komputerów pracujących w parze (lub procesorów pracujących na jednym komputerze),
- zaangażowanie wielu komputerów do pracy nad tym samym programem w tym samym czasie (równoległe).

Wielowątkowość (1)

Wielowątkowość programowa - współbieżność

(ang. multithreading) – cecha systemu operacyjnego, dzięki której w ramach jednego procesu może wykonywać kilka wątków lub jednostek wykonawczych. Nowe wątki to kolejne ciągi instrukcji wykonywane oddzielnie. Wszystkie wątki tego samego procesu współdzielą kod programu i dane. W systemach nie obsługujących wielowątkowości pojęcia procesu i wątku utożsamiają się.

Wniosek: w tym ujęciu wielowątkowość nie jest cechą architektury, tylko systemu operacyjnego.

Wielozadaniowość

Wielozadaniowość

(ang. multitasking) – cecha systemu operacyjnego umożliwiająca mu równoczesne wykonywanie więcej niż jednego procesu. Zwykle za poprawną realizację wielozadaniowości odpowiedzialne jest jądro systemu operacyjnego.

Wniosek: wielozadaniowość nie jest cechą architektury, tylko systemu operacyjnego.

Wielowątkowość (2)

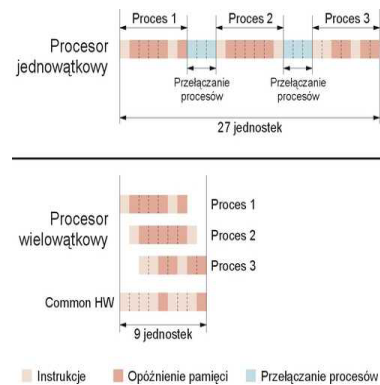
Wielowątkowość sprzętowa

Wielowątkowość może także odnosić się do samych procesorów. W takim wypadku oznacza możliwość jednoczesnego wykonywania wielu wątków sprzętowych na pojedynczej jednostce wykonawczej – rdzeniu (ang. core). W tym celu procesor powinien mieć więcej niż jeden licznik i zbiór programowalnych rejestrów. Powielenie tych zasobów jest dużo mniej kosztowne niż wymiana całej jednostki obliczeniowej.

Wielowątkowość (3)

Jeśli ze względu na opóźnienia w komunikacji między procesorem a pamięcią, pojedyncze zadanie nie będzie w stanie całkowicie wykorzystać zasobów procesora, jednocześnie przetworzenie wielu zadań trwałoby krócej niż wykonanie ich sekwencyjne.

Dla procesora MIPS32 34K, wykorzystującego architekturę wielowątkową, wzrost powierzchni układu scalonego o 14% odpowiada 60-procentowemu wzrostowi przepustowości w stosunku do rdzenia jednowątkowego.



Klasyfikacja systemów wieloprocessorowych

Systemy równoległe to systemy komputerowe wyposażone w wiele procesorów, które mogą równoległe wykonywać obliczenia. Procesory te mogą być ze sobą:

- ściśle powiązane - współdzielą niektóre pozostałe składowe systemu komputerowego (magistrale, pamięć, urządzenia zewnętrzne, itp.).
- luźno powiązane - każdy procesor posiada własną pamięć i magistralę, tworzy niezależny podsystem komputerowy i komunikuje się z pozostałymi podsystemami poprzez szybkie linie komunikacyjne.

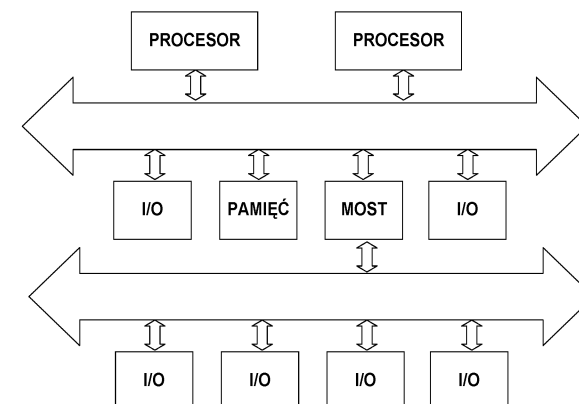
Wielordzeniowość

Procesor wielordzeniowy

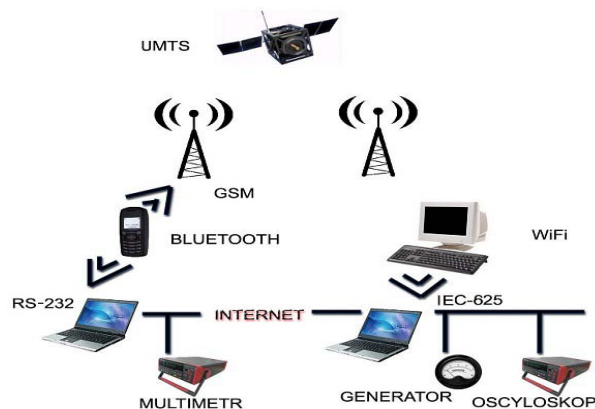
(ang. multicore) procesor, w którym kilka rdzeni zamkniętych w jednej obudowie umożliwia jednocześnie wykonywanie różnych zadań (programów). Nie wymaga to specjalnie napisanych aplikacji.

- Cechą systemów wielordzeniowych jest wykorzystanie przetwarzania współbieżnego do podziału zadań.
- Liczba wątków oprogramowania wykonywanych w tym samym czasie, powinna odpowiadać liczbie dostępnych w systemie procesorów.
- Rdzenie w urządzeniu wielordzeniowym mogą współdzielić pamięć podręczną najwyższego poziomu (np. L2 dla Intel Core 2) lub posiadać odrębne pamięci podręczne (np. AMD).
- Rdzenie współdzielą połączenia z pozostałymi elementami systemu.

Architektura systemu dwuprocessorowego



Architektura systemu rozproszonego



prof. dr hab. inż. Joanna Józefowska Architektura komputerów

Systemy równoległe ściśle powiązane

Wyróżniamy dwa rodzaje równoległych systemów ściśle powiązanych:

- systemy z symetrycznymi procesorami,
- systemy z asymetrycznymi procesorami.

Symetryczne procesory

to równoprawne procesory centralne, które współdzielą zegar, magistralę systemową i pamięć, wraz z jakimś mechanizmem synchronizacji ich działania.

prof. dr hab. inż. Joanna Józefowska Architektura komputerów

Wieloprocusorowość symetryczna

- (ang. Symmetric Multiprocessing), przetwarzanie symetryczne
- Symetryczna wieloprocusorowość jest obecnie najpopularniejszą techniką konstrukcji równoległych komputerów ogólnego przeznaczenia.
- Systemami obsługującymi SMP jest większość Uniksów, Windows NT, BeOS.

prof. dr hab. inż. Joanna Józefowska Architektura komputerów

Wieloprocusorowość symetryczna - własności

- Pozwala na znaczne zwiększenie mocy obliczeniowej systemu komputerowego poprzez wykorzystanie 2 lub więcej procesorów do jednoczesnego wykonywania zadań.
- Szczególną korzyść z przetwarzania równoległego czerpią aplikacje do renderingu (grafika trójwymiarowa) i edycji wideo oraz nowoczesne gry, małą natomiast pakiety biurowe.
- Ważne jest wykorzystanie wielowątkowości przez programy komputerowe (oraz obsługa wielowątkowości przez system operacyjny) - upraszcza to możliwość "podzielenia" procesu dla kilku procesorów.
- Odporność na uszkodzenia - w przypadku awarii jednego z procesorów pozostałe procesory mogą przejąć jego zadania.

prof. dr hab. inż. Joanna Józefowska Architektura komputerów

Wieloprocessorowość asymetryczna

Procesory asymetryczne

poszczególne procesory nie są traktowane jednakowo, jak w architekturze symetrycznej, lecz niektórym z nich przydzielane są zadania specjalne.

Jeden z procesorów może np. obsługiwać wyłącznie operacje wejścia/wyjścia, podczas gdy pozostałe zajmują się obsługą normalnych aplikacji.

Wieloprocessorowość asymetryczna - własności

- (ang. Asymmetric Multiprocessing), przetwarzanie asymetryczne
- Procesory w układzie ASMP mogą mieć dostęp do rozłącznych zestawów urządzeń peryferyjnych. Takie rozwiązanie upraszcza projektowanie systemów, ale ogranicza maksymalną wydajność w porównaniu do architektury symetrycznej.
- Przykładem zastosowania ASMP w komputerach osobistych, był procesor Intel 80286 współpracujący z koprocesorem numerycznym 80287, który był wykorzystywany tylko do przeprowadzania obliczeń.

Wieloprocessorowość asynchroniczna

- (ang. Asynchronous Multiprocessing - AMP)
- odmiana wieloprocessorowości asymetrycznej,
- poszczególne procesory nie są traktowane jednakowo
- poszczególne procesory:
 - mogą być taktowane innymi zegarami lub
 - mogą być na nich uruchomione różne systemy operacyjne
 - pracują niezależnie od siebie.
- Układy AMP wykorzystywane są w systemach czasu rzeczywistego (RTOS), które wymagają ścisłej kontroli nad procesorem, na którym system jest uruchomiony.
- Przykładem zastosowania ASMP w komputerach osobistych, był procesor Intel 80286 współpracujący z koprocesorem numerycznym 80287, który był taktowany zegarem o innej częstotliwości niż procesor główny.

Komputery masowo równoległe

MPP

(ang. Massively Parallel Processors - MPP) – architektura systemu komputerowego zawierającego wiele niezależnych jednostek obliczeniowych (CPU) lub mikroprocesorów, które pracują równoległe. Określenie "masowo" oznacza, że chodzi o setki a nawet tysiące takich jednostek.

Przykłady pierwszych systemów tego typu to Distributed Array Processor, Goodyear MPP, Connection Machine i Ultrakomputer.

Architektura UMA

Uniform Memory Access

jeden z rodzajów systemu wieloprocusorowego, charakteryzuje się wspólną pamięcią operacyjną dla wszystkich procesorów, każdy procesor ma własną pamięć podręczną (Cache). Czas dostępu do pamięci jest dla wszystkich procesorów identyczny.

Charakterystyczna dla SMP.

Architektura NUMA

Non-Uniform Memory Access

charakteryzuje się tym, że udostępnia użytkownikowi spójną logicznie przestrzeń adresową, chociaż pamięć jest fizycznie podzielona. Niejednorodność w dostępie do pamięci polega na tym, że dany procesor szybciej uzyskuje dostęp do swojej lokalnej pamięci, niż do pamięci pozostałych procesorów lub pamięci współdzielonej.

W wieloprocusorach typu NUMA istnieje pojedyncza wirtualna przestrzeń adresowa widoczna przez wszystkie procesory. Wartość zapisana, przez dowolny procesor jest natychmiast widoczna dla wszystkich pozostałych, dlatego też następna operacja odczytu z tej komórki pamięci przez dowolny procesor poda właśnie tę wartość.

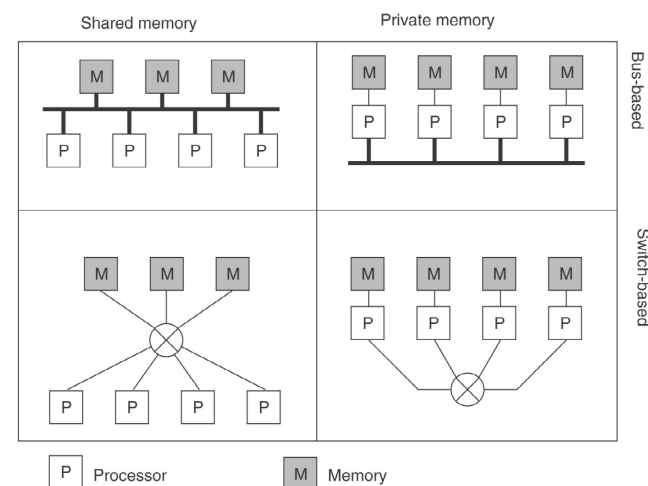
Architektura CCNUMA

Cache Coherent Non-Uniform Memory Access

charakteryzuje się niejednorodnym dostępem do pamięci, ze spójną pamięcią podręczną, udostępnia użytkownikowi spójną logicznie przestrzeń adresową, pomimo tego, że pamięć jest fizycznie podzielona.

W CC-NUMA pamięci podręczne muszą być spójne. Spójność ta jest zrealizowana za pomocą dodatkowego sprzętu oraz protokołu. Wykorzystywany do tego jest katalog czyli dodatkowa pamięć, która służy do zapamiętywania faktu pobrania danych z pamięci i przesłania ich do węzła lub w bardziej skomplikowanych przypadkach miejsca pobytu wszystkich kopii danych.

Systemy z szyną i przełącznikami



System rozproszony

Distributed system

zbiór niezależnych urządzeń technicznych połączonych w jedną, spójną logicznie całość. Zwykle łączonymi urządzeniami są komputery, rzadziej – systemy automatyki.

Systemy rozproszone

- Połączenie komputerów najczęściej realizowane jest przez sieć komputerową, jednak można wykorzystać również inne magistrale komunikacyjne.
- Obecnie w systemach rozproszonych istotną rolę odgrywa warstwa pośrednicząca (ang. middleware), która pośredniczy w komunikacji między komponentami systemów rozproszonych, np.
 - gniazda (ang. sockets),
 - RPC (Remote Procedure Call),
 - DCE (Distributed Computing Environment),
 - CORBA (Common Object Request Broker Architecture),
 - DCOM (Distributed Component Object Model),
 - RMI (Remote Method Invocation).

Cechy systemów rozproszonych

- ① współdzielenie zasobów (ang. resource sharing) – wielu użytkowników systemu może korzystać z danego zasobu (np. drukarek, plików, usług, itp.),
- ② otwartość (ang. openness) – możliwość rozbudowy systemu zarówno pod względem sprzętowym, jak i oprogramowania,
- ③ współbieżność (ang. concurrency) – zdolność do przetwarzania wielu zadań jednocześnie,
- ④ skalowalność (ang. scalability) – zachowanie podobnej wydajności systemu przy zwiększeniu skali systemu (np. liczby procesów),
- ⑤ odporność na błędy (ang. fault tolerance) – zdolność działania systemu mimo pojawiania się błędów,
- ⑥ przezroczystość (ang. transparency) – postrzeganie systemu przez użytkownika jako całości, a nie poszczególnych składowych.

Systemy wielokomputerowe

- klastry,
- farmy komputerów,
- gridy

Klastry



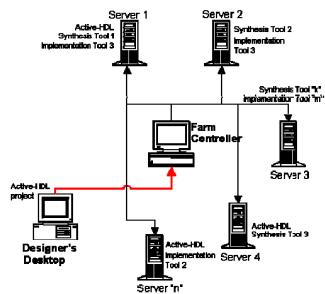
Klaster

Klaster komputerowy (ang. cluster) - grupa połączonych jednostek komputerowych, które współpracują ze sobą w celu udostępnienia zintegrowanego środowiska pracy.

Własności klastrów

- Komputery wchodzące w skład klastra (będące członkami klastra) nazywamy węzłami (ang. node).
- Jedną z najbardziej popularnych implementacji klastrów obliczeniowych jest klaster typu Beowulf gdzie rolą węzłów pełnią wydajne komputery klasy PC, pracujące pod kontrolą GNU/Linuksa oraz z zainstalowanym oprogramowaniem pozwalającym uzyskać przetwarzanie równoległe np. bibliotekami MPI, PVM.
- Klaster tworzą węzły połączone siecią wewnętrzną.
- Są one również połączone zbiorem macierzy dyskowych.

Farma komputerów

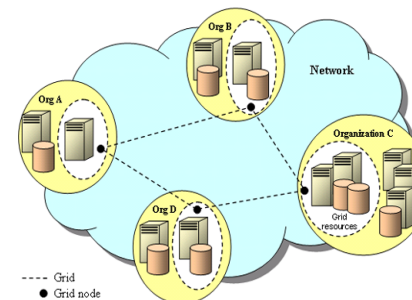


Farma komputerów

Połączenie wielu pracujących jednocześnie serwerów i stacji roboczych.

- Oprogramowanie przydziela obciążenie dla poszczególnych elementów farmy.
- Zadaniom przydzielane są priorytety.

Grid



Grid

Metoda konstruowania rozproszonych sieci obliczeniowych, mogących się składać z systemów zupełnie różnych typów, pracujących w sposób autonomiczny.

Plan wykładu	Wieloprocessorowość	Systemy równoległe	Systemy wielokomputerowe	Pytania
	○○○ ○○○○○ ○○○	○○○ ○	○ ○○○○○	

Pytania

- 1 Podaj klasyfikację systemów wieloprocessorowych.
- 2 Omów architektury systemów wieloprocessorowych ze względu na pamięć.
- 3 Objaśnij różnicę pojęć: wielozadaniowość, wielowątkowość i wielordzeniowość.
- 4 Podaj cechy systemów rozproszonych.