

## ***Лекція 2\_дод. Паралельні обчислювальні системи та їх класифікація.***

### *План лекції.*

1. Структура сучасних векторноконвеєрних систем.
2. Структурні особливості SMP-систем.
3. Системи масового паралелізму (MPP-системи).
4. Кластерні системи.
5. Деякі нетрадиційні підходи до побудови паралельних обчислювальних систем.
6. Приклади класифікацій паралельних систем (за Фліном, за Хокні, класифікація суперкомп'ютерів, V-класифікація). Переваги та недоліки цих класифікацій.

**1.** Як було зазначено у попередній лекції, основними принципами, закладеними в архітектуру **векторноконвеєрних систем**, є:

- конвеєрна організація оброблення потоку команд;
- введення в систему команд набору векторних операцій, які дозволяють оперувати з масивами даних.

Довжина оброблюваних векторів у сучасних векторноконвеєрних системах складає, зазвичай, 128 або 256 елементів. Основне призначення векторних операцій полягає в розпаралелюванні виконання операторів циклу, в яких зосереджена велика частина обчислювальної роботи.

Сучасні **векторноконвеєрні** системи мають **ієрархічну структуру**:

- на нижньому рівні ієрархії розташовані **конвеєри операцій** (наприклад, конвеєр додавання дійсних чисел, конвеєр множення таких чисел тощо);
- деяка сукупність конвеєрів операцій об'єднується в **конвеєрний функціональний пристрій**;
- векторноконвеєрний **процесор** складається з низки конвеєрних функціональних пристроїв;
- декілька векторноконвеєрних процесорів (від 2 до 16), об'єднаних спільною пам'яттю, утворюють обчислювальний **вузол**;
- декілька таких вузлів об'єднуються з допомогою комутаторів, утворюючи або NUMA-систему, або MPP-систему.

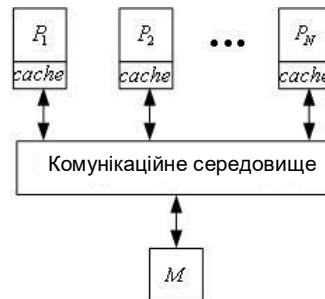
Типовими представниками такої архітектури є обчислювальні системи:

CRAY J90/T90; CRAY SV1; CRAY X1;  
NEC SX-4/SX-5/SX-6/SX-7/SX-8/SX-8R/SX-9/SX-ACE;  
Earth Simulator / 2 / 3.

**2.** Всі процесори **SMP-системи** мають симетричний доступ до пам'яті, тобто пам'ять такої системи є UMA (Unified Memory Access)-пам'яттю з однаковим часом доступу. Під **симетричністю** розуміємо: рівні права всіх процесорів на доступ до пам'яті, одну і ту ж адресацію для всіх елементів пам'яті, однаковий час

доступу всіх процесорів системи до пам'яті (без врахування взаємного блокування).

Загальна структура SMP-системи має вигляд:



Тут  $P_i$  ( $i = \overline{1, N}$ ) – процесори системи, а  $M$  – блок пам'яті.

**Комунікаційне середовище** SMP-системи будується на основі якої-небудь високошвидкісної системної шини або високошвидкісного комутатора. Окрім **однакових процесорів** і блоку спільної пам'яті  $M$  до цієї ж шини або комутатора підключаються **пристрої вводу-виводу**.

У таких системах виникають значні проблеми, пов'язані загалом з оперативною пам'яттю. Проблема конфліктів під час звертання до спільної шини була частково усунута завдяки **розділенню пам'яті на блоки**, підключення до яких за допомогою комутаторів дозволило розпаралелити звертання від різних процесорів. За такого підходу SMP-системи змогли об'єднувати вже до **32 процесорів**.

Щоб усунути розрив між швидкістю роботи оперативної пам'яті та швидкістю роботи процесора, сучасні процесори облаштовуються високошвидкісною буферною пам'яттю (**кеш-пам'яттю**). Швидкість доступу до цієї пам'яті в **декілька десятків разів** є більшою за швидкість доступу до **основної пам'яті** процесора. Однак, наявність кеш-пам'яті порушує принцип рівноправного доступу до будь-якої точки пам'яті, оскільки дані, що знаходяться в кеш-пам'яті одного процесора, недоступні для інших процесорів. Тому після кожної модифікації копії змінної, що знаходиться в кеш-пам'яті якого-небудь процесора, необхідно здійснювати синхронну модифікацію самої цієї змінної, що розташована в основній пам'яті. У сучасних SMP-системах **когерентність** кеш-пам'яті підтримується апаратно або операційною системою.

Із-за обмеженої пропускної здатності комунікаційного середовища SMP-системи **погано масштабуються**. Відомим **недоліком** їх є те, що їх вартість зростає швидше, ніж продуктивність при збільшенні кількості процесорів у системі. Зауважимо, що SMP-системи появились як **альтернатива** для дорогих **багатопроекторних систем** на базі векторноконверсних процесорів.

Найбільш відомими SMP-системами є, наприклад,

SMP-сервери IBM, HP, Compaq, Dell, Fujitsu

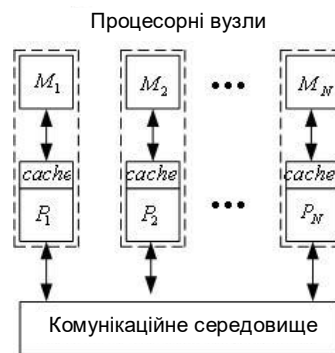
тощо. Така система функціонує під керуванням єдиної операційної системи (найчастіше – Unix і подібної до неї).

**3. Обчислювальні системи масового паралелізму (MPP-системи)** є багато-процесорними обчислювальними системами з розподіленою пам'яттю, в яких з допомогою деякої комунікаційної мережі об'єднуються однорідні обчислювальні вузли. Кожен **вузол** складається з одного або декількох процесорів, власної оперативної пам'яті, комунікаційного обладнання, підсистеми вводу/виводу. Процесори в таких системах мають прямий доступ лише до своєї локальної пам'яті.

Порівняно з SMP-системами, архітектура MPP-системи **усуває** одночасно як **проблему конфліктів** під час звертання до пам'яті, так і **проблему когерентності** кеш-пам'яті. **Головною перевагою** MPP-систем є висока степінь масштабування. Успішно функціонують такі системи з **сотнями і тисячами** процесорів, зокрема, в

ASCI Blue Mountain – 6144 процесори (запуск в 1998 р.; зупинена в 2004 р.);  
 ASCI White – 8192 процесори (запуск в 2001 р.; списана в 2006 р.);  
 CRAY XC40 – 185088 ядер.

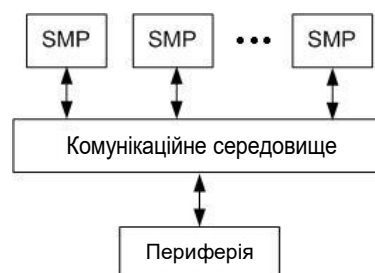
**Структура MPP-системи** має вигляд:



Тут кожен процесор  $P_j$  має свій блок пам'яті  $M_j$  для всіх  $j = \overline{1, N}$ .

Практично всі **рекорди з продуктивності** на сьогоднішній день встановлені саме на MPP-системах. Останнім часом для з'єднання обчислювальних вузлів у них все частіше використовується ієрархічна система **високошвидкісних комунікаторів**, яка вперше була реалізована в системах **IBM SP2**. Ця топологія дає можливість прямого обміну даними між будь-якими вузлами без залучення проміжних вузлів.

Далі розглянемо **NUMA-систему**, яка зазвичай будується на базі однорідних процесорних вузлів, які складаються з невеликої кількості процесорів і блоку пам'яті. Ці вузли об'єднуються з допомогою деякого високошвидкісного комунікаційного середовища. **Структура** такої системи має вигляд:



При цьому підтримується єдиний адресний простір. Апаратно підтримується доступ до віддаленої пам'яті, тобто до пам'яті інших вузлів. Доступ до локальної пам'яті здійснюється в декілька разів швидше, ніж до віддаленої.

По суті NUMA-система є MPP-системою, де як окремі обчислювальні елементи використовуються SMP-вузли. Логічна загальнодоступність пам'яті в NUMA-системах, з одного боку дозволяє працювати з **єдиним адресним простором**, а, з іншого боку, дозволяє досить просто забезпечити високу масштабованість системи. Сеймур Крей (Seymour R. Cray) додав до такої системи новий елемент – когерентний кеш, створивши так звану архітектуру cc-NUMA (Cache Coherent Non-Uniform Memory Access), яка розшифровується як «неоднорідний доступ до пам'яті із забезпеченням когерентності кешів». Найбільш відомими системами архітектури cc-NUMA є:

HP 9000 V-class в SCA-конфігураціях,  
SGI Origin3000,  
Sun HPC 15000,  
IBM/Sequent NUMA-Q 2000.

На сьогодні максимальна кількість процесорів у cc-NUMA-системах може перевищувати 1000 одиниць.

**4. Кластерні системи** – це логічне продовження розвитку ідей, закладених в архітектурі MPP-систем, фактично кластери є більш дешевим варіантом таких систем. Кластерні технології дозволяють для досягнення необхідної продуктивності обчислень об'єднувати в єдині обчислювальні системи засоби різного типу, починаючи від ПК і закінчуючи потужними суперкомп'ютерами. Такі технології набули широкого розповсюдження як засіб створення систем суперкомп'ютерного класу із складових частин масового виробництва, що значно здешевлює вартість обчислювальної системи загалом.

Кластерна система складається з окремих обчислювальних вузлів, об'єднаних єдиним комунікаційним середовищем. Кожен вузол має свою локальну оперативну пам'ять. При цьому **спільної фізичної оперативної пам'яті** для вузлів зазвичай **не існує**. На склад та архітектуру вузлів не накладаються жодні обмеження і кожен з них може функціонувати під керуванням власної операційної системи. Комунікаційне середовище кластера зазвичай дозволяє вузлам **взаємодіяти** між собою лише за допомогою **передачі повідомлень**.

Зараз розроблено цілу низку **технологій об'єднання вузлів** кластера. Найбільш **поширеною** (бо проста у використанні та має **низьку вартість**) є

Fast Ethernet (100 Mbit/s).

Більш ефективними є технології

Gigabit Ethernet (1 Gbit/s),  
10/40 Gigabit Ethernet,

а також швидкісні мережі, такі як

100 Gigabit Ethernet,

InfiniBand (100 Gbit/s, після 2023 р. – 250 Gbit/s),  
Intel Omni-Path (100 Gbit/s, у 2019 р. відміна створення технології з продуктивністю 200 Gbit/s)

тощо.

У середині 70-х років XX ст. за кількістю застосувань лідерами були **векторноконвеєрні системи**, а після 2002 р. лідерами стали **кластери**. Векторноконвеєрні системи, обладнані спеціалізованими процесорами для оброблення одразу наборів даних, зберегли переваги у швидкодії та **простоті програмування** порівняно з кластерами. Зате останні використовують у вузлах **стандартні процесори** і є значно **дешевшими**, хоча потребують більш складного програмного забезпечення.

Зазначимо, що пік розробок паралельних обчислювальних систем у колишньому Радянському Союзі припав на 1986 р. Серед проєктів, які фінансувалися тоді, були:

- матричний спецпроцесор ЕС2700, м. Єреван;
- мультипроцесор з динамічною архітектурою ЕС2704, м. Ленінград;
- мультипроцесор ЕС2706, м. Таганрог;
- сімейство мультипроцесорів ПС ІПУ АН СРСР, м. Москва;
- системи НДІ «Квант», м. Москва.

Після розпаду Радянського Союзу майже 10 років на колишніх його теренах не приділялась належна увага дослідженням, пов'язаним із високопродуктивними обчисленнями. І лише у 2000 р. на науковій конференції в Черноголовці (Росія) було засвідчено значний інтерес до цієї проблеми.

Світова тенденція стосовно розробок та поширення кластерних систем спостерігається і в країнах СНД. Зокрема, в Росії кластерні системи почали створювати на базі систем сімейства МВС (МВС-100, МВС-1000); у Білорусі (спільно з Росією) розроблено кластери сімейства СКИФ; у Вірменії (спільно з Росією) розроблено Armcluster; в Україні створено ціле сімейство кластерів СКІТ (ІК ім. В.М. Глушкова НАН України). Однак це був лише початок...

Зараз триває **процес масової модернізації** кластерів на підставі переходу на **багатоядерні процесори** виробництва компаній IBM, Intel, AMD, Sun та залучення графічних процесорів. В Україні встановлено кластерні системи у багатьох вузах та НДІ НАН України, зокрема і у Львові (ІФКС НАН України, ФМІ НАН України, ІППММ НАН України тощо).

Одним із шляхів **створення значних обчислювальних ресурсів** та головним інтеграційним напрямком для всієї суперкомп'ютерної галузі є обчислювальні мережі **grid computing**. Також розвивається концепція **розподілених обчислювальних середовищ** для розв'язання складних обчислювальних задач на доступних розподілених неоднорідних засобах.

На підставі викладеного можна зробити **висновок**, що за останні більш як 20 років паралельні обчислювальні системи перетворились із засекречених засобів, які використовувались виключно з оборонною метою, у могутній високоінтелектуальний ресурс для відкритих застосувань.

5. Крім розглянутих вище **широко розповсюджених** підходів до побудови досить універсальних паралельних обчислювальних систем знаходять своє застосування і інші, можливо **менш поширені**, але перспективні підходи, наприклад:

- розроблення **нейрокомп'ютерів** – апаратних засобів для реалізації нейромережних методів і алгоритмів;
- створення **VLIW (Very Long Instruction Word)-систем** з дуже довгим командним словом, які дозволяють одночасне виконання набору простих команд;
- розроблення **систем із структурно-процедурною** організацією обчислень;
- побудова **систолічних та квазісистолічних** обчислювальних структур.

6. Із викладеного випливає, що існує велика кількість обчислювальних систем і, очевидно, виникає бажання ввести для них яку-небудь **класифікацію**. Класифікація потрібна для того, щоб легше орієнтуватися під час вибору обчислювальної системи (якщо є така можливість) для реалізації заданого алгоритму розв'язання задачі або програми, що описує цей алгоритм. Загалом відомо дуже багато різних класифікацій паралельних обчислювальних систем. Ми розглянемо деякі із них, які є **найбільш відомими**:

- класифікація за М.Д. Фліном;
- класифікація за Р. Хокні;
- класифікація суперкомп'ютерів;
- V-класифікація.

**Найбільш розповсюдженою** стала класифікація обчислювальних систем, запропонована в 1966 р. професором Стенфордського університету **М.Д. Фліном** (M.J. Flynn). Ця класифікація **охоплює** лише **дві ознаки** – тип потоку команд і тип потоку даних.

В **одинарному потоці** команд в один момент часу може виконуватися лише одна команда. У цьому випадку ця єдина команда визначає в даний момент часу роботу всіх або хоча б багатьох пристроїв обчислювальної системи.

У **множинному потоці** команд в один момент часу може виконуватися багато команд. У цьому випадку кожна із таких команд визначає в даний момент часу роботу лише одного або лише декількох (**але не всіх**) пристроїв обчислювальної системи.

В одинарному потоці послідовно виконуються окремі команди, а в множинному – групи команд.

Одинарний **потік даних** обов'язково передбачає наявність в обчислювальній системі лише одного пристрою оперативної пам'яті і одного процесора. Однак при цьому **процесор** може бути якзавгодно **складним**, так що процес обробки кожної одиниці інформації в потоці може вимагати виконання багатьох команд.

Множинний потік даних складається із багатьох одинарних потоків даних.

Відповідно до сказаного всі обчислювальні системи **за Фліном** діляться на чотири типи:

- SISD (Single Instruction Single Data);

- MISD (Multiple Instruction Single Data);
- SIMD (Single Instruction Multiple Data);
- MIMD (Multiple Instruction Multiple Data).

Обчислювальна система **SISD** є звичайною **послідовною** (однопроцесорною) ЕОМ фон-неймановської архітектури.

На обчислювальні системи **MISD** існують різні погляди. За одним із них – за всю історію розвитку обчислювальної техніки такі системи **не були створені**. За іншим (менш поширеним, ніж перший) **MISD**-система – це система, в якій кілька процесорів виконують різні потоки команд і типовим представником цих систем є векторноконвеєрні обчислювальні системи. Ми будемо дотримуватися першого погляду.

У системах **SIMD** кілька процесорів одночасно (синхронно) виконують **одну і ту ж команду**, але над **різними потоками даних**. Ці системи діляться на два великі класи:

- векторноконвеєрні системи;
- векторнопаралельні або матричні обчислювальні системи.

Представниками **SIMD**-систем є:

ICL DAP, STARAN, BSP,  
ILLIAC IV, Connection Machine, PC-2000 тощо.

Принцип **SIMD** реалізується в сучасних графічних процесорах.

Обчислювальні **системи MIMD** складаються із багатьох процесорів, які (звичай, асинхронно) виконують **різні команди** над **різними даними**. Зараз це найпоширеніший клас архітектур паралельних систем. Ці системи ще називають **багатопроеесорними**. Представниками **MIMD**-систем є:

Intel Paragon, Cray T3D, CM (Connection Machine) -5,  
Tianhe-2A - TH-IVB-FEP Cluster, Pioneer-EUS - NDv4 cluster,  
Amazon EC2 Instance Cluster us-east-1a тощо.

Розглянута **класифікація Фліна** дозволяє за приналежністю обчислювальної системи до класу **SIMD** або **MIMD** зрозуміти базовий принцип її роботи. Досить часто цього є достатньо. **Недоліком** цієї класифікації є «**перевантаженість**» класу **MIMD**. Крім цього, вона є неповною бо, наприклад, вона не включає систолічні обчислювальні структури – набори зв'язаних однотипних процесорів (іноді дуже простих), що **синхронно виконують різні програми**.

Пізніше, у 1985 р., появилася класифікація **Р. Хокні** (R. Hockney), яка детально класифікує обчислювальні системи, які за Фліном потрапляють у клас **MIMD**. У цьому випадку виділяється три групи систем:

- конвеєрні системи;
- з комутатором (системи зі спільною та розподіленою пам'яттю);
- мережеві системи (регулярні решітки, ієрархічні структури, системи зі змінною конфігурацією, гіперкуби).

Зараз **статус суперкомп'ютерів** у світі визначається лише за обчислювальними системами, що увійшли до останньої версії випуску рейтингу top500

([www.top500.org](http://www.top500.org)). Згідно з редакцією рейтингу за листопад 2021 року найбільше таких систем встановлено в Китаї (173), далі йдуть США (149), Японія (32), Німеччина (26), Франція (19).

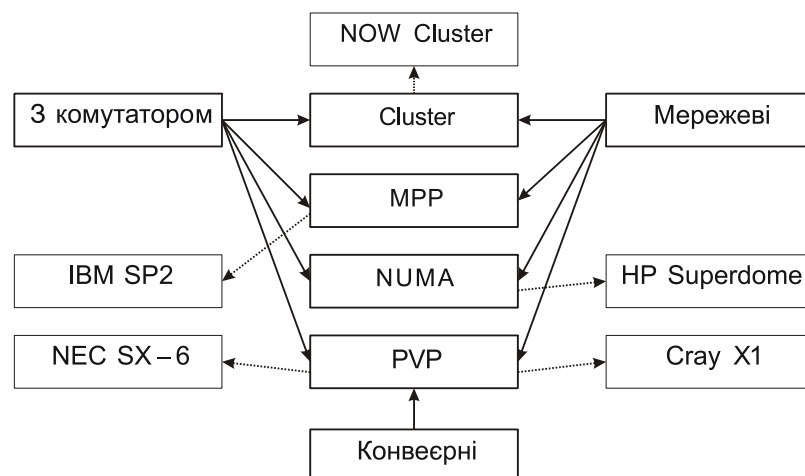
Для класифікації суперкомп'ютерних систем прийнято такі терміни:

- SMP-системи;
- NUMA-системи;
- MPP-системи;
- Cluster Systems – кластерні системи;
- PVP (Parallel Vector Processing) – паралельні векторні системи.

Наведені класи суперкомп'ютерних систем мають особливості, властиві різним класам систем за Хокні. До того ж всі обчислювальні системи, що потрапляють в список надпотужних систем світу, мають властивості як **мережевих** систем, так і систем з **комутатором**.

У 2004 р. була введена V-класифікація для сучасних суперсистем і їх відповідності класифікації Хокні. На цю класифікацію відображаються найбільш розповсюджені системи, наприклад, IBM SP2, NEC SX-6, HP Superdome, Cray X1, NOW Cluster тощо.

Згідно V-класифікації Cluster-, MPP-, NUMA-, PVP-системи мають властивості як систем з комутатором, так і мережевих і крім цього PVP-системи мають властивості конвеєрних систем. Прикладом MPP-системи з комутатором є IBM SP2; PVP-системи з комутатором – NEC SX-6; NUMA-системи з мережею – HP Superdome; PVP-мережевої системи – Cray X1; Cluster-системи – NOW Cluster.



В архітектурі сучасних суперсистем, як і понад 10 років назад, переважає **напрямок**, пов'язаний з побудовою масовопаралельних систем (**MPP-систем**). Використання стандартних обчислювальних процесорів під час розробки таких систем є економічно доцільним і найбільш ефективним для досягнення **оптимального** співвідношення **ціна/швидкодія**. У цьому разі використання саме процесорів Intel і AMD дозволяє досягти високої обчислювальної потужності за мінімальних витрат, а **проблема** збільшення **швидкодії** ефективно вирішується лише шляхом **нарощування** кількості **процесорів** у системі.