

Лекція

з дисципліни: «Комп'ютерна схемотехніка та архітектура комп'ютерів»

на тему: «Суперкомп'ютери. Паралельні обчислювальні системи»

ПЛАН ЛЕКЦІЇ

1. Класифікація систем комп'ютерів.
2. Векторні суперкомп'ютери (SIMD).
3. Кластери (MIMD)
4. Класифікація та архітектурні особливості трансп'ютерів
5. Класифікація та архітектурні особливості кластерних комп'ютерів

ЛІТЕРАТУРА

1. Злобін Г. Г. Архітектура та апаратне забезпечення ПЕОМ : навч. посібник / Г. Г. Злобін, Р. Є. Рикалюк – К.: Каравела, 2012. – 304 с.
2. Мельник А.О. Персональні суперкомп'ютери: архітектура, проектування, застосування: монографія / А.О. Мельник, В.А. Мельник. – Львів: Видавництво Львівської політехніки, 2013. – 516 с.
3. Матвієнко М. П. Архітектура комп'ютера. Навчальний посібник. / Матвієнко М. П., Розен В. П., Закладний О. М.— К: Видавництво Ліра-К, 2016. — 264 с.

1. Класифікація систем комп'ютерів

Існують різні класифікації комп'ютерів – за призначенням, за технічною сумісністю за програмною сумісністю, за розмірами і т. д. Наведемо класифікацію, що використовують провідні виробники комп'ютерів:

- суперкомп'ютери;
- мейнфрейми;
- сервери;
- персональні комп'ютери;
- настільні;
- переносні;

Суперкомп'ютери – це великі комп'ютери, що створюються для задач, які вимагають великих обчислень, таких, як визначення координати далекої зірки або галактики, моделювання клімату, складання карт нафтових і газових покладів і т. д.

Суперкомп'ютери – це штучний продукт, вони створюються для рішення конкретних задач замовника, але елементи, що входять до складу суперкомп'ютерів, є серійними.

Суперкомп'ютери складаються із сотень процесорів, мають велику оперативну пам'ять і високу швидкодію. Вони займають великі зали за площею, рівною 2 – 3 баскетбольним площадкам.

Більшість суперкомп'ютерів створюються за кластерною технологією (Cluster). За цією технологією комп'ютер будується з декількох десятків серверів, що працюють як єдина система. Кластерні суперкомп'ютери легко масштабуються і дозволяють створювати дублюючі обчислювальні лінії, що є необхідним, коли обчислення, наприклад, моделюють процеси в реальному часі і збої неприпустимі.

Мейнфрейми

Мейнфрейми – це великі комп'ютери, з високою швидкодією і великими обчислювальними ресурсами, що можуть обробляти велику кількість даних і виконувати обробку запитів одночасно декількох тисяч користувачів.

Мейнфрейми виконані із надлишковими технічними характеристиками, що робить їх дуже надійними.

Фізично мейнфрейми мають один корпус – системний блок розміром із шафу, до якого можуть підключатися термінали (термінал складається з монітора і клавіатури).

Використовуються мейнфрейми для збереження й обробки великих баз даних, а також великих web-вузлів з великою кількістю одночасних обертів.

Сервери

Сервери – це комп'ютери, що служать центральними вузлами в комп'ютерних мережах. На серверах встановлюється програмне забезпечення, що дозволяє управляти роботою мережі.

На серверах зберігається інформація, якою можуть користуватися всі комп'ютери, підключені до мережі. Від сервера залежить працездатність усієї мережі і схоронність баз даних та іншої інформації, тому сервери мають декілька резервних дублюючих систем збереження даних, електроживлення, можливість заміни несправних блоків без переривання роботи. Сервери можуть містити від декількох десятків процесорів. За технологічною сумісністю сервери бувають IBM-сумісними і Macintosh-сумісними.

Персональні комп'ютери

Персональні комп'ютери (ПК) – це комп'ютери, що можуть використовуватися однією людиною автономно, незалежно від інших комп'ютерів. Персональні комп'ютери можуть бути настільними, переносними і кишеньковими.

За технологічною сумісністю персональні комп'ютери бувають IBM-сумісними і Macintosh-сумісними.

2. Векторні суперкомп'ютери (SIMD)

Серед сучасних суперЕОМ цю архітектуру мають однопроцесорні векторні суперкомп'ютери. Практично всі вони випускаються також у мультипроцесорних конфігураціях, що відносяться до класу MIMD. Однак багато особливостей архітектури векторних суперЕОМ можна зрозуміти, розглядаючи навіть однопроцесорні системи.

Типова схема однопроцесорного векторного суперкомп'ютера представлена на прикладі FACOM VP-200 японської фірми Fujitsu. Схожу архітектуру мають і інші векторні суперкомп'ютери, наприклад, фірми Cray Research і Convex. Загальним для всіх векторних суперкомп'ютерів є наявність у системі команд векторних операцій, наприклад, додавання векторів, що допускають роботу з векторами визначеної довжини, допустимо, 64 елементи по 8 байтів. У таких комп'ютерах операції з векторами звичайно виконуються над векторними регістрами, що, однак, зовсім не є обов'язковим. Наявність регістрів маски дозволяє виконувати векторні команди не над всіма елементами векторів, а тільки над тими, на які вказує маска.

Багатопроцесорні векторні суперкомп'ютери (MIMD)

Усі згадані векторні суперкомп'ютери випускаються в багатопроцесорних конфігураціях, що відносяться вже до класу MIMD.

У архітектурі багатопроцесорних векторних комп'ютерів можна відзначити дві найважливіші характеристики: симетричність (рівноправність) усіх процесорів системи і поділ усіма процесорами загального поля оперативної пам'яті. Подібні комп'ютерні системи називаються сильно зв'язаними. Якщо в однопроцесорних векторних ЕОМ для створення ефективної програми її треба векторизувати, то в багатопроцесорних з'являється задача розпаралелювання програми для її виконання одночасно на декількох процесорах.

Задача розпаралелювання є, мабуть, більш складною, оскільки в ній необхідно організувати синхронізацію паралельно виконуваних процесів. Практика показала можливості ефективного розпаралелювання великого числа алгоритмів для розглянутих сильно зв'язаних систем. Відповідний підхід до розпаралелювання на таких комп'ютерах називається іноді моделлю поділюваної загальної пам'яті.

Багатопроцесорні SMP-сервери на базі мікропроцесорів RISC-архітектури (MIMD)

Продуктивність деяких сучасних мікропроцесорів RISC-архітектури стала порівняною з продуктивністю процесорів векторних комп'ютерів. Як наслідок цього, з'явилися ці досягнення, що використовують суперЕОМ нової архітектури, – сильно зв'язані комп'ютери класу MIMD, які становлять симетричні багатопроцесорні сервери з загальним полем оперативної пам'яті. Цим перспективним системам має сенс приділити більш уваги, ніж іншим комп'ютерним архітектурам.

У модулях пам'яті звичайно використовується технологія DRAM, що дозволяє досягти великих ємностей пам'яті при відносно низькій ціні. Однак швидкість обміну даними між процесорами і пам'яттю в таких серверах набагато нижче, ніж пропускна спроможність аналогічного тракту у векторних суперЕОМ, де оперативна пам'ять складається на більш дорогій технології DRAM. У цьому полягає одна із основних відмінностей у підходах до суперкомп'ютерних обчислень, що застосовуються для багатопроцесорних векторних ЕОМ і SMP-серверів. У перших звичайно є відносно невелике число векторних регістрів, тому, як уже відзначалося, для підтримки високої продуктивності необхідно швидко завантажувати в них дані або, навпаки, записувати з них інформацію в оперативну пам'ять. Таким чином, потрібна висока продуктивність тракту процесор-пам'ять.

4. Кластери (MIMD)

Кластери є найбільш дешевим способом нарощування продуктивності вже інсталюваних комп'ютерів. Фактично кластер становить набір з декількох ЕОМ,

поєднаних через деяку комунікаційну інфраструктуру. Такою структурою може виступати звичайна комп'ютерна мережа, однак з понять підвищення продуктивності бажано мати високошвидкісні з'єднання (FDDI/ ATM/ HiPPi і т. п.). Кластери можуть бути утворені як з різних комп'ютерів (гетерогенні кластери), так і з однакових (гомогенні кластери). Очевидно, що всі такі системи відносяться до класу MIMD. Кластери є класичним прикладом слабо зв'язаних систем.

Перевагою кластерного підходу порівняно з SMP-серверами є поліпшення можливостей масштабування. На відміну від серверів SMP-архітектури, де нарощування конфігурації обмежене пропускнуною спроможністю шини, додавання комп'ютерів у кластер дозволяє збільшувати пропускну спроможність оперативної пам'яті і підсистем вводу-виводу.

У кластерних системах для організації взаємодії між процесами, що виконуються на різних комп'ютерах при рішенні однієї задачі, застосовуються різні моделі обміну повідомленнями (PVM, MPI і т. п.). Однак задача розпаралелювання в таких системах з розподіленою між окремими комп'ютерами пам'яттю в межах цих моделей є більш складною, ніж у моделі загального поля пам'яті, як наприклад, у SMP-серверах. До цього варто додати чисто апаратні проблеми наявності затримок при обмінах повідомленнями і підвищення швидкості передачі даних. Тому спектр задач, що можуть ефективно вирішуватися на кластерних системах, порівняно із симетричними сильно зв'язаними системами досить обмежений. Для паралельної обробки запитів до баз даних у подібних системах також мають свої власні підходи.

4. Класифікація та архітектурні особливості трансп'ютерів

Для підвищення продуктивності звичайно збільшується швидкість послідовних фоннейманівських архітектур шляхом застосування більш сучасної технології, наприклад, переходів Джозефсона або GaAS (галій-арсенід технології). Однопроцесорна фоннейманівська архітектура має межу, що визначається швидкістю розповсюдження електричного сигналу по фізичних лініях зв'язку між компонентами.

Альтернативою послідовних процесів є архітектури, розробка яких заснована на паралелізмі: дані на алгоритми розподіляються між процесорами.

Зокрема, до високопаралельних архітектур класу МКМД відносяться трансп'ютери, які розглянемо пізніше більш докладно.

Серія машин, що реалізувалися на базі трансп'ютерних технологій, складається з набору системних компонентів, кожний з яких об'єднує елемент, що обробляє, систему зв'язку і пам'ять в одній мікросхемі, виготовлений за технологією СВІС. Однією з найважливіших властивостей технології СВІС є та, що час обміну між модулями набагато більший, ніж час обміну між компонентами одного модуля. У комп'ютері при виконанні майже кожної операції процесор звертається до пам'яті, тому трансп'ютер включає і процесор, і пам'ять на одній мікросхемі.

Для забезпечення максимальної швидкості і мінімального монтажу трансп'ютер використовує послідовні двоточкові лінії для прямого зв'язку з іншими трансп'ютерами. Трансп'ютер містить типовий процесор з мікрокодом. У невеликій пам'яті зберігаються 32 інструкції (а в сучасних трансп'ютерах – до 64) і забезпечують виконання простих послідовних програм. Окрім цього, є інші більш спеціалізовані групи інструкцій, що реалізують, наприклад, арифметичні операції з високою точністю, і планувальник.

Процес, що виконується в трансп'ютері, може складатися з ряду паралельних процесів, отже трансп'ютер повинен на апаратному рівні підтримувати модель програмування мови Окам. Таким чином, трансп'ютер має вбудований планувальник, що розподіляє час процесора між паралельними процесами.

Планувальник підтримує два рівня процесів: високопродуктивний процес, якщо він готовий до виконання, і низькопріоритетний процес.

Трансп'ютер складається з блока пам'яті, процесора і системи зв'язку, з'єднаних 32-розрядною шиною. До цієї шини також підключений інтерфейс із зовнішньою пам'яттю, що дозволяє використовувати локальну додаткову пам'ять. Процесор, пам'ять і система зв'язку займають по 25 % всієї площі кремнію; інша площа використовується для відведення тепла, таймерів і зовнішніх з'єднань.

Структурна типова схема трансп'ютера представлена на рис. 1.



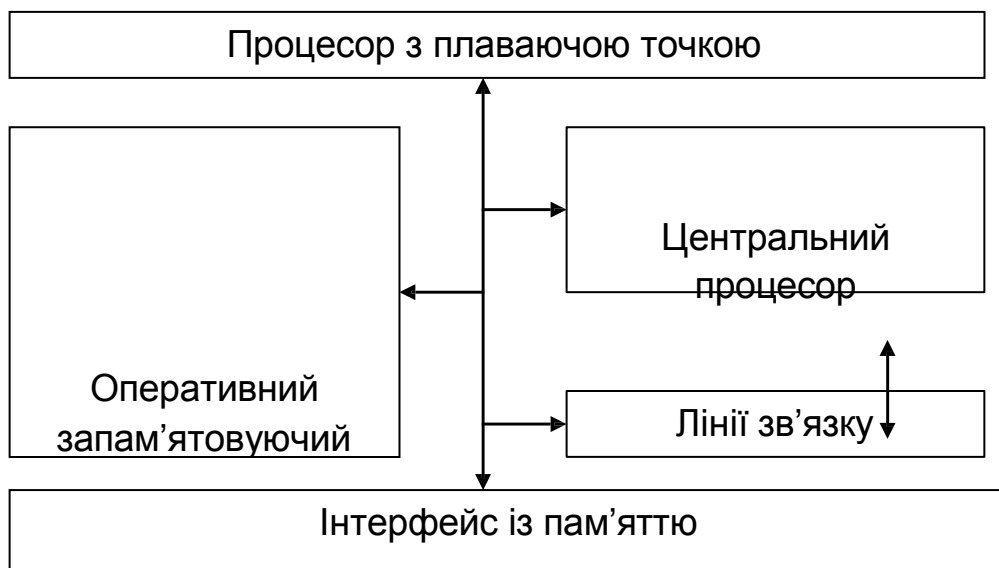


Рис. 1. Внутрішня архітектура трансп'ютера

Центральний процесор (ЦП) трансп'ютера містить три регістри (А, В і С) для роботи з цілими числами і адресами, що формують апаратний стек. При запису в стек вміст регістра В записується в регістр С, вміст регістра А – у В, і лише після цього в регістр А заноситься нове значення. При читанні з стека витягається значення регістра А, а вміст регістра С – у регістр В. Відповідно процесор з плаваючою точкою (ППТ) містить трьохрегістровий стек для зберігання проміжних результатів при роботі з числами з плаваючою точкою (обчислювальний стек). Це регістри АF, ВF, СF. Робота з цим стеком аналогічна роботі з регістрами А, В і С.

Адреси значень змінних з плаваючою точкою формуються в стеку ЦП, і ці значення передаються між осередками, що адресуються пам'яті і стеку ППТ під управлінням ЦП. Стек ЦП використовується тільки для зберігання адреси змінних з плаваючою точкою, розрядність ЦП не залежить від розрядності ПКП. Отже, один і той же ПКП може використовуватися як з 32-розрядним, так і з 16-розрядним процесором.

Планувальник трансп'ютера підтримує два рівня пріоритетів. Регістровий стек дублюється для того, щоб при переключенні трансп'ютера з низького пріоритету на високий не зберігати в пам'яті інформацію про стан ППТ. Це дозволило зменшити час реакції на переривання до 3,7 мкс. Більше того, дублювання регістрового стека дозволяє програмі обробки переривань виконувати операції з плаваючою точкою без зниження продуктивності.

Характеристики трансп'ютера

1. **Коректне функціонування ПЗ.** Мова програмування Окам має безліч властивостей, що дозволяє застосовувати засоби формального

проектування. Зокрема, така властивість, як концепція виконання паралельних процесів, що не поділяють загальну пам'ять і можуть обмінюватися інформацією тільки за допомогою каналів зв'язку, потенційно дозволяє одержувати продукт високої цілісності, бо при цьому підтримується певна форма обмеження розповсюдження помилок. Розподіл процесів можна забезпечити, якщо:

весь код трансп'ютерної програми написаний коректною мовою Окам;

жоден процес не виконує операції запису за адресами поза адресного простору робочої області.

Першу вимогу можна задовольнити, якщо ретельно контролювати проектування ПЗ. Другу вимогу можна задовольнити в принципі шляхом обов'язкового безперервного контролю над границями масивів. У зв'язку з тим, що трансп'ютер не має можливості управління пам'яттю, для того, щоб обмежити область роботи процесу його робоніж простором, такий підхід може бути недостатнім, особливо якщо врахувати, що іонізуюча радіація може призвести до ефекту довільного SEU (Single event upset, одинична зміна стану).

2. Надійність елементів. Оцінка якості виробництва трансп'ютерів показала, що процес виготовлення відповідає такому високому стандарту, який дозволяє ставити позначку "придатні до використання в космосі".

3. Радіаційні ефекти. Радіаційні тести показали, що в трансп'ютерах можна очікувати всі три механізми відмов:

накопичування зарядів. Трансп'ютер витримує досить значну дозу радіації $\sim 50 \text{ Krad/s}$ при інтенсивності $0,5 \text{ Krad/m}$ до настання відмови. Спостерігався також ефект нормалізації, коли накопичений заряд розсіюється, якщо мікросхему виведено із зони радіації;

пробоїві. Для відновлення пробою необхідно знеструмити уражену мікросхему і роз'єднати її лінії зв'язку з іншими ураженими мікросхемами. Більш пізні екземпляри трансп'ютерів, що виготовлені на епітаксіальній кремнієвій шаровій основі, не підвладним пробоям ні за яких умов;

одинична зміна стану (SEU). Виникає з такою ж частотою, що й інші мікросхеми, виготовлені за технологією CMOS. Помилки в одному біті в трансп'ютері можуть відбуватися з частотою в діапазоні від однієї за місяць, на низьких екваторіальних орбітах і до десятків за день у міжпланетному просторі при сонячній активності.

Області використання трансп'ютерів

Існують декілька важливих областей застосування трансп'ютерів:

плати-прискорювачі для персональних комп'ютерів і робочих станцій – для підвищення продуктивності існуючих комп'ютерів;

вбудовані системи, наприклад, у контролерах лазерного принтера, для обробки зображень і в космічних системах;

комп'ютери загального призначення – окремі паралельні комп'ютери загального призначення;

обробка документів – розпізнання друкованих або рукописних символів і зменшення смуги пропускання при передачі факсимільних зображень;

автоматизація виробництва – автоматичне складання компонентів й контроль;

медицина і біологія – розпізнання образів для комп'ютерної томографії, знаходження пухлин, цитології, радіології й огляду пацієнта;

збір даних із супутників – метеорологія, контроль навколишнього середовища, сільське господарство, використання землі;

засоби масової інформації – кодування цифрового телебачення, зменшення смуги пропускання, підвищення якості;

військове застосування – зобов'язання і контроль над їхнім виконанням, знаходження загублених і запам'ятованих снарядів, моделювання польоту і процесу старіння ядерних боєголовок.

5. Класифікація та архітектурні особливості кластерних комп'ютерів

При побудові будь-яких досить потужних обчислювальних систем, що підтримують чутливі до збоїв додатка, необхідно вирішити дві основні проблеми: забезпечення високої продуктивності й тривалого функціонування. Це насамперед стосується критичних до будь-якого роду зупинок додатків, пов'язаних з обробкою транзакцій, базами даних і телекомунікаційних сервісів.

Тривалість функціонування системи залежить від трьох факторів: надійності, готовності й ергономічності обслуговування. Ці фактори визначаються стійкістю системи до виникаючих збоїв, здатністю адекватно реагувати на відмову окремих елементів і системи в цілому, здатністю відновлювати свої властивості після збою або відмов у її роботі. Перераховані фактори тісно пов'язані один з одним, і тому оцінювати їх необхідно в комплексі.

Підвищення надійності засноване на принципі зниження до мінімального рівня несправностей, відмов і збоїв. Ця мета досягається за рахунок застосування сучасної елементної бази, а також за рахунок удосконалювання методів складання апаратури.

Готовність системи до використання – показник її надійності для кінцевого користувача – виражається в процентному співвідношенні між

працездатним станом системи і її простоїв з будь-якої причини, у тому числі й плановими (до них відноситься профілактичні роботи, реконфігурації, зміна версій програмного забезпечення, модернізація встаткування й т. п.).

Підвищення рівня готовності припускає придушення в певних межах впливу відмов і збоїв на роботу системи за допомогою засобів контролю й корекції позаштатних ситуацій, а також засобів автоматичного відновлення обчислювального процесу після прояву не справності, включаючи апаратну й програмну надмірність, на основі якої реалізуються різні варіанти стійких архітектур.

Основні експлуатаційні характеристики системи серйозно залежать від зручності її обслуговування, зокрема, від ремонтпридатності, контролю придатності й т. д.

Термін *"кластеризація"* у комп'ютерній області має багато значень. Найбільш загальне визначення могло б звучати так: "Реалізація об'єднання комп'ютерів, що представляється єдиним цілим для операційної системи, системного програмного забезпечення, прикладних програм і користувачів".

Комп'ютери, що утворюють кластер, – так звані вузли кластера – завжди відносно незалежні, що допускає зупинку або вимикання кожного з них для проведення профілактичних робіт або установки додаткового встаткування без порушення працездатності всього кластера. Крім того, всі вузли кластера виконують корисну роботу – на відміну від систем з реплікацією даних (приміром, Novell SFT III), а також на відміну від FT- Систем (Fault Tolerant), які забезпечують не більш п'яти хвилин простою за рік за рахунок максимальної надмірності – резервування всіх підданих відмовам елементів. При відмові одного з вузлів кластера вузли, що залишилися, можуть взяти на себе завдання, що виконувалися на вузлі, який відмовив, не зупиняючи загальний процес роботи. При цьому збій у роботі кластера виражається лише в деякому зниженні продуктивності системи або в неприступності додатків на короткий час, необхідний для перемикання на інший вузол.

Продуктивність кластерної системи легко масштабується; це значить, що додавання в систему додаткових процесорів, оперативної й дискової пам'яті й нових вузлів може виконуватися в будь-який час, коли це реально потрібно.

Концепція кластерних систем

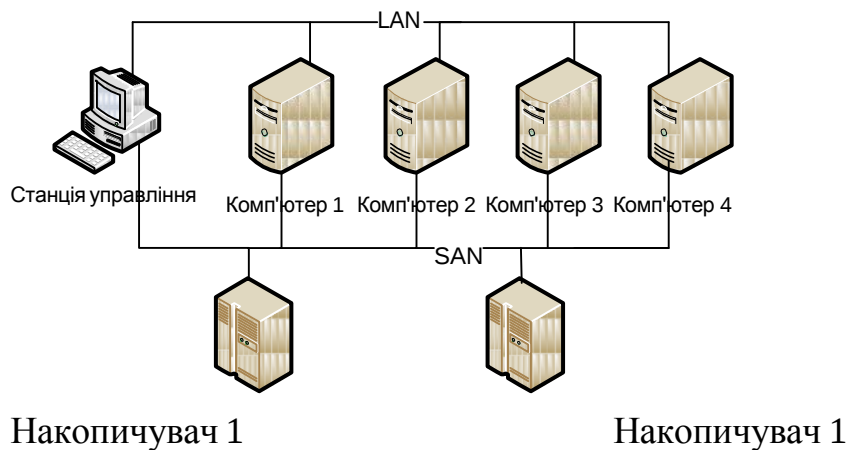


Рис. .2. Кластерна система

LAN – Local Area Network, локальна мережа.

SAN – Storage Area Network, мережа зберігання даних.

Уперше в класифікації обчислювальних систем термін "кластер" визначила компанія Digital Equipment Corporation (DEC).

За визначенням DEC, кластер – це група обчислювальних машин, які пов'язані між собою й функціонують як один вузол обробки інформації.

Кластер функціонує як єдина система, тобто для користувача або прикладного завдання вся сукупність обчислювальної техніки виглядає як один комп'ютер. Саме це і є найважливішим при побудові кластерної системи.

Перші кластери компанії Digital були побудовані на машинах VAX. Ці машини вже не виробляються, але усе ще працюють на площадках, де були встановлені багато років тому. І напевно найважливіше те, що загальні принципи, закладені при їхньому проектуванні, залишаються основою при побудові кластерних систем і сьогодні.

До загальних вимог, що пропонуються до кластерних систем, відноситься:

1. Висока готовність.
2. Висока швидкодія.
3. Масштабування.
4. Загальний доступ до ресурсів.
5. Зручність обслуговування.