

# ЛЕКЦІЯ 1

## РОЗПОДІЛЕНІ СИСТЕМИ

### 1.1. Визначення розподілених систем

Розподілена система – це набір незалежних комп'ютерів, що представляється їх користувачам єдиною об'єднаною системою.

Розподілена система – це така система, в якій взаємодія та синхронізація програмних компонентів, що виконується на незалежних мережевих комп'ютерах, виконується за допомогою передачі повідомлень. Тобто це набір незалежних комп'ютерів, що не мають спільної пам'яті або спільного часу та вони взаємодіють через комунікаційну мережу, передаючи повідомлення, де кожен комп'ютер використовує свою власну оперативну пам'ять та свою операційну систему. Ресурси багатьох комп'ютерів об'єднуються для того, щоб виконати єдину задачу.

Є два моменти.

Перший відноситься до апаратури: всі машини автономні.

Другий стосується програмного забезпечення: користувачі думають, що мають справу з єдиною системою.

Ми розглядаємо розподілену систему з апаратної точки зору, як сукупність незалежних процесів. Комп'ютери, процесори чи процеси будемо називати вузлами розподільної системи. Щоб комп'ютер був автономним, він повинен мати своє власне, незалежне керування. Під незалежністю процесів розуміємо той факт, що кожен процес у даний момент часу має свій власний стан, що можна представити у вигляді набору даних, що включають значення лічильника команд, реєстрів та змінних, до яких процес може звертатися та змінювати. Стан кожного процесу є закритим для інших процесів. Швидкості виконання операцій різних процесів у розподільній системі різні та відомі заздалегідь, а он час передання повідомлень може бути непередбачуваним.

Як ми вже сказали, процеси теж можуть бути вузлами розподіленої системи, тобто такою системою не обов'язково повинна бути сукупність

комп'ютерів, а й програма чи скоріше система програм, що є сукупністю діючих процесів, що виконуються на одному пристрої. Але все ж таки більшість розподілених систем представлені комп'ютерною мережею.

## **1.2. Переваги та недоліки розподілених систем**

### **1.2.1. Переваги розподілених систем**

Замість того, щоб намагатися дати визначення розподіленої системи можна перерахувати її характеристики. Визначимо ~~нієть~~ основних характеристик розподілених систем.

- 1) *Спільне використання ресурсів.* Розподілені системи дозволяють спільне використання апаратних та програмних ресурсів, наприклад жорстких дисків, принтерів, файлів, компіляторів та інше, об'єднаних засобами мережі. Очевидно, що розподіл ресурсів можливий і в багатокористувацьких системах, але в цьому випадку за надання ресурсів і їх керування повинен керувати центральний процесор.
- 2) *Відкритість.* Це можливість розширювати систему шляхом додавання нових ресурсів. Розподілені системи – це відкриті системи, до яких приєднується апаратне і програмне забезпечення від різних виробників.
- 3) *Паралельність.* В розподілених системах декілька процесів можуть одночасно виконуватися на різних комп'ютерах в мережі. Ці процеси можуть взаємодіяти один з одним під час їх виконання.
- 4) *Масштабованість.* В принципі всі розподілені системи є масштабованими: щоб система відповідала новим вимогам, її можна нарощувати за допомогою додавання нових обчислювальних ресурсів. Але на практиці нарощування може обмежуватися мережею, яка об'єднує окремі комп'ютери системи. Якщо приєднати багато нових машин, то пропускна здатність мережі може виявитися недостатньою.
- 5) *Відмовостійкість.* Наявність декількох комп'ютерів і можливість дублювання інформації означає, що розподілені системи стійкі до певних апаратних і програмних помилок. Більшість розподілених

систем у випадку помилки, як правило, можуть підтримувати хоча б частково функціональність. Повний збій в системі відбувається тільки у випадку мережових помилок.

- 6) *Прозорість*. Ця властивість означає, що користувачам надано повністю прозорий доступ до ресурсів і в той же час приховано інформацію про розподіл ресурсів у системі. Однак, в багатьох випадках конкретні знання про організацію системи допомагає користувачу краще використовувати ресурси.
- Відсутність єдиного часу для компонентів розподіленої системи. Це важливе припущення для проектування та побудови розподілених систем. Воно характеризує територіальне розташування компонентів системи та з нього випливає, що робота компонентів системи не синхронізована.
  - Відсутність спільної пам'яті. Це ключова характеристика, з якої випливає необхідність обміну повідомленнями між компонентами системи. Якщо ж система представляє собою набір програм, то відсутність спільної пам'яті це не обов'язково фізично різні пам'яті, а скоріше доступ до різних її областей. А також ця характеристика говорить про очевидну відсутність спільного фізичного часу.
  - Географічне розподілення. Можна сказати, що умова великої відстані між компонентами системи є достатньою для того, щоб систему можна було назвати розподільною, але не необхідною (знову ж таки система програм). Останнім часом кластери робочих станцій (Cluster of Workstation, COW), об'єднаний за допомогою локальної обчислювальної мережі все більше розглядається як невелика розподільна система. Всі компоненти системи можуть знаходитися в декількох сусідніх спорудах. Подібні кластери стають все більш популярними завдяки тому, що її компоненти коштують не багато, при

цьому продуктивність висока. Наприклад, ядро пошукової системи Google побудоване на системі COW.

- Незалежність та гетерогенність. Комп'ютери, що входять до розподіленої системи можуть мати різні технічні характеристики та виконувати однакові задачі за різну кількість часу.

### 1.2.2. Недоліки розподілених систем

Очевидно, розподіленим системам властиві певні недоліки.

- 1) *Складність*. Розподілені системи складніші від централізованих. Набагато складніше зрозуміти і оцінити властивості розподілених систем в цілому, а також тестувати ці системи. Наприклад, продуктивність системи залежить від швидкості роботи одного процесора, а від смуги пропускання мережі і швидкодії різних процесорів. Переміщаючи ресурси з одної частини мережі в іншу, можемо радикально вплинути на продуктивність системи.
- 2) *Безпека*. Як правило доступ до системи можемо отримати з декількох різних машин, повідомлення в мережі можуть переглядатися або перехоплюватися. Тому, в розподіленій системі набагато складніше підтримувати безпеку.
- 3) *Керованість*. Система може складатися з різнотипних комп'ютерів, на яких можуть бути встановлені різні версії операційних систем. Помилка однієї машини не розповсюджується на інші машини з непередбачуваними наслідками. Тому необхідно значно більше зусиль, щоб керувати і підтримувати систему в робочому стані.
- 4) *Непередбачуваність*. Як відомо всім користувачам Web-мережі, реакція розподілених систем на певні події непередбачувана і залежить від повного завантаження системи, її організації і мереженого навантаження. Оскільки ці всі параметри можуть постійно змінюватися, час , затрачений на виконання запиту користувача, в той чи інший момент може суттєво різнитися.

### 1.3. Проблеми проектування розподілених систем

При обговоренні переваг та недоліків розподілених систем визначено перелік критичних проблем проектування даних систем.

- 1) *Ідентифікація ресурсів*. Ресурси в розподіленій системі розміщуються на різних комп'ютерах, тому систему імен ресурсів необхідно продумати так, щоб користувачі могли без труднощів відкривати необхідні їм ресурси і посилатися на них. Прикладом може бути система уніфікованого показника ресурсів URL, яка визначає адреси Web-сторінок. Без зручної і універсальної системи ідентифікації більша частина ресурсів виявиться недоступною користувачам системи.
- 2) *Комунікації*. Універсальна працездатність Internet і ефективна реалізація протоколів TCP/IP в Internet для більшості розподілених систем є прикладом ефективного способу організації взаємодії між комп'ютерами. Однак там, де на продуктивність, надійність та інше накладаються спеціальні вимоги, можна скористатися альтернативними способами системних комунікацій.
- 3) *Якість системного сервісу*. Якість сервісу, який надає система, відображає її продуктивність, працездатність і надійність. На якість сервісу впливає цілий ряд факторів: розподіл системних процесів, розподіл ресурсів, системні і мережеві апаратні засоби та можливість адаптації системи.
- 4) *Архітектура програмного забезпечення*. Архітектура програмного забезпечення описує розподіл системних функцій за компонентами системи, а також розподіл даних компонент за процесорами. Якщо необхідно підтримувати високу якість системного сервісу, вибір правильної архітектури виявляється вирішальним фактором

#### **1.4. Цілі побудови розподільних систем**

За останні роки кількість розподілених систем стає все більше. Серед основних причин росту можна виділити наступні:

Географічно розподілена обчислювальна середа. Сьогодні у більшості випадків сама обчислювана середа є ‘розкиданою’ територіально. Як приклад можна навести банківську мережу. Кожен банк обслуговує рахунки своїх клієнтів та оброблює операції з ними. У разі переводу грошей з одного банку в інший потрібна взаємодія банків один з одним. Сама мережа Інтернет також є прикладом географічно розподільної системи.

Потреба збільшувати ефективність обчислювальних систем. Швидкість традиційних однопроцесорних обчислювальних систем досягла своєї межі. Різні архітектури (такі як суперскалярна архітектура, матричні та векторні процесори, однокристальні багатопроцесорні системи) збільшують ефективність обчислювальних систем за рахунок паралельного виконання деяких процесів. Однак такі методи дозволяють збільшити ефективність роботи системи тільки в десятки разів, до того ж, у них існує нагальна проблема з масштабуванням. Для збільшення ефективності у сотні й тисячі разів необхідно звести численні процесори та забезпечити їх ефективну взаємодію. Цей принцип реалізовується за допомогою великих багатопроцесорних систем та машинних комплексів.

Сумісне використання ресурсів. Важливою частиною побудови та використання розподільних систем є надання користувачеві та програмам доступу до віддалених ресурсів та забезпечення їх сумісного використання. В даному контексті до ресурсів ми відносимо як апаратне так і програмне забезпечення, з якими працює обчислювальна система. Наприклад користувач першого комп’ютеру може використовувати дисковий простір іншого комп’ютеру для зберігання інформації. Або програма може використовувати вільну обчислювальну потужність для прискорення власної роботи. Розподільні бази даних та системи об’єктів можуть стати ідеальним прикладом, коли відповідні програмні абстракції розподілені між декількома

комп'ютерами та сумісно обслуговуються декількома процесорами, утворюючи розподільну систему.

Стійкість до падіння системи. У традиційних нерозподільних системах, що побудовані на базі 1 комп'ютера (навіть доволі високоефективного) вихід з ладу одного з компонентів призводить до краху всієї системи. Характерною рисою розподільних систем є те, що вони доволі стійкі до часткових відмов компонентів системи. Тобто система продовжує функціонувати після часткової відмови, тільки ефективність падає. Така можливість досягається за рахунок додавання у систему додаткових апаратних чи програмних ресурсів. Щоб зрозуміти як функціонує така система уявимо футбольні команди. Нехай ми проводимо футбольний матч, де приймають участь 2 команди, одна з них має доволі сильних гравців, але не має заміни. Тобто при будь-якій травмі будь-якого гравця гра закінчується, бо замінити його немає ким. Проти неї грає команда, де є всі заміни, при чому кожні 20 хвилин по 2-3 гравці міняються, щоб запобігнути перевтоми гравців. Це примітивна демонстрація плюсів розподіленої системи.

Завдання розробників розподілених систем – спроектувати програмне або апаратне забезпечення таким чином, щоб реалізувати всі необхідні характеристики розподіленої системи.

В розподіленій системі різні системні компоненти можуть бути реалізовані на різних мовах програмування і виконуватись на різних типах процесорів. Моделі даних, подання інформації і протоколи взаємодії – все це необов'язково буде однотипним в розподіленій системі. Отже для розподілених систем необхідне таке програмне забезпечення, яке могло би керувати цими різнотипними частинами та гарантувати взаємодію і обмін даними між ними. *Проміжне програмне забезпечення* відноситься власне до такого класу ПЗ. Воно знаходиться якби посередині між різними частинами розподілених компонент системи.