Приймак Василь Іванович, доктор економічних наук, професор

Львівський національний університет імені Івана Франка, Львів

ORCID: https://orcid.org/0000-0003-0244-8661

Рибак Юрій Михайлович, магістрант

Львівський національний університет імені Івана Франка

Голубник Ольга Романівна, кандидат економічних наук, доцент

ORCID: https://orcid.org/0000-0003-1211-4614

**АРХІТЕКТУРА РОЗПОДІЛЕНИХ СИСТЕМ ЗБЕРІГАННЯ ДАНИХ**

У сучасному технологічному світі компанії потребують швидкого і точного аналізу для прийняття обґрунтованих рішень. Однак зростаючий обсяг даних і різноманітність типів даних створюють проблеми зі зберіганням і обробкою інформації, на якій ґрунтується прийняття рішень. Тому виникають питання щодо вибору відповідних ефективних систем зберігання та обробки даних.

Дійсно, неефективно обробляти дані, які містяться в базах даних і файлових системах, на одному комп’ютері. Доцільніше розподілити інформацію між багатьма вузлами, які потім можуть оброблятися паралельно. Технічно можливо використовувати не один, а декілька фізичних пристроїв, які можуть бути розташовані в різних місцях. В результаті паралельна обробка не тільки підвищує продуктивність і швидкість роботи системи, але й дозволяє декільком користувачам одночасно працювати з інформацією.

Для паралельної обробки інформація розбивається на менші частини, які можуть обчислюватися незалежно на різних пристроях. Це значно прискорює процес обробки даних і дозволяє гнучко та ефективно організувати обробку запитів. Такий підхід не тільки прискорює процес обробки запитів, але й забезпечує одночасний доступ до інформації та підвищує надійність її зберігання [1].

На даний час існує два принципових підходи до забезпечення необхідного рівня продуктивності систем:

* вертикально масштабований підхід;
* горизонтально масштабований підхід.

Якщо ви просто хочете збільшити навантаження, найпростіший спосіб – купити більш потужну машину. Більше процесорів, оперативної пам’яті і багато дисків можуть бути об’єднані в рамках однієї операційної системи, а високошвидкісні з’єднання дозволяють будь-якому центральному процесору (ЦП) отримати доступ до будь-якої частини пам’яті або диска. Цей тип архітектури, або підхід архітектури спільної пам’яті, дозволяє розглядати всі компоненти як єдине ціле [2, с. 146].

Один з недоліків такої архітектури стосується вартості системи, яка зростає більшими темпами, ніж обсяг технічних пристроїв: машина з подвоєним процесором, оперативною пам’яттю і дисковим простором коштує більш ніж удвічі дорожче. Інший недолік пов’язаний з потужністю машини. Вдвічі більша система може не впоратися з удвічі більшим навантаженням через певні вузькі місця.

Недоліком такого підходу є те, що, архітектура зі спільною пам’яттю є явно менш відмовостійкою, оскільки якщо один компонент виходить з ладу, то вся система може вийти з ладу. Ще однією важливою характеристикою, якій не відповідають такі системи, є забезпечення географічної близькості між клієнтами та даними. Іншими словами, такий архітектурний підхід означає, що сервери розташовуються лише в одному географічному місці, якщо ж сервіси розгорнуті по всьому світу, то неможливо забезпечити близькість між клієнтами та серверами для сервісів.

Альтернативним підходом до shared-memory architecture є горизонтальне масштабування – архітектура без розділення ресурсів (shared-nothing architecture). У цьому підході кожен комп’ютер або віртуальна машина, на якій працює програмне забезпечення бази даних, називається вузлом. Кожен вузол використовує процесор, пам’ять і диски індивідуально. Взаємодія між вузлами здійснюється на програмному рівні, з використанням звичайних мереж [2, с. 146].

Архітектура shared-nothing не потребує спеціального обладнання, а тому може використовувати найкращі комп’ютери з точки зору співвідношення ціни та продуктивності. Географічний розподіл даних дозволяє скоротити час доступу користувачів до них, а в разі необхідності – витримати втрату центру обробки даних.

Слід зазначити, що існує третій архітектурний підхід (shared disk architecture), який використовує загальні диски. Він базується на використанні декількох машин з незалежними центральними процесорами та оперативною пам’яттю. У цьому випадку самі дані зберігаються на низці дисків, до яких надається спільний доступ через високошвидкісну мережу. Такий підхід використовується для деяких завдань зберігання даних у сховищах даних, але конфлікти, накладні витрати на блокування та інші причини обмежують масштабованість архітектур із загальними дисками.

Згідно з нашими дослідженнями, з розвитком інформаційних технологій та збільшенням обсягів використовуваних даних постає проблема оптимального зберігання цих даних з метою забезпечення необхідного рівня продуктивності комп’ютерних систем. На сьогодні існує два основних підходи до забезпечення необхідного рівня продуктивності системи: вертикально масштабований підхід та горизонтально масштабований підхід. Другий з цих підходів, тобто тип архітектури без розділення ресурсів є найбільш придатним для вирішення поставленої задачі.

**Список використаних джерел**

1. M. Klymash, O. Hordiichuk-Bublivska, I.Chaikovskyi and T. Danylchenko. Research of algorithms for parallel processing of information in databases, *Infocommunication technologies and electronic engineering*, vol. 1 (1), pp. 51-62, 2021.

2. M. Kleppmann Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainаble Systems, 590 pp. 2017.