

УДК 004.6

В. В. Юрчак

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ТЕХНОЛОГІЇ ПОБУДОВИ ВІ-СИСТЕМ ДЛЯ АНАЛІЗУ ВЕЛИКИХ ДАНИХ

UDC 004.6

V. V. Yurchak

TECHNOLOGIES OF BUILDING BI-SYSTEMS FOR BIG DATA ANALASYS

При побудові ВІ (Business Intelligence) - систем для аналізу великих даних для різних завдань, зазвичай використовується технологія ROLAP. Реляційна онлайн-аналітична обробка (Relational OLAP, ROLAP) - це форма OLAP (online analytical processing – аналітична обробка у реальному часі), яка виконує динамічний багатовимірний аналіз даних, що зберігаються в реляційній базі даних, а не в багатовимірній базі даних.

Принцип роботи такої моделі досить простий: при активізації аналітичної функції формується SQL – запит, який йде на той чи інший бекенд. При цьому ВІ -система може бути інтегрована з будь-якою СУБД, що підтримує мову SQL (наприклад, реляційні СУБД MySQL, PostgreSQL та інші).

Перевага даної технології у тому, що для роботи ВІ- системи немає обмежень за обсягом даних. Недолік – відносно невисока швидкість виконання запитів, що негативно впливає на продуктивність ВІ -системи.

Для вирішення цієї проблеми використовуються різні підходи, у тому числі застосування спеціалізованих колонкових СУБД, наприклад ClickHouse або Vertica. Однак колонкові СУБД використовують власний обрізаний діалект SQL, що суттєво знижує функціональні можливості такого рішення.

Більш перспективним є рішення, засноване на застосуванні СУБД, в яких реалізована технологія In-Memory, що є технологією виконання комп'ютерних обчислень повністю в пам'яті комп'ютера, наприклад, в оперативному запам'ятовуєчому пристрої (ОЗП).

Цей термін, зазвичай, має на увазі великомасштабні складні обчислення, які вимагають спеціального системного програмного забезпечення для виконання обчислень на комп'ютерах, що працюють разом у кластері. Як кластер комп'ютери об'єднують свою оперативну пам'ять, тому розрахунок, по суті, виконується між комп'ютерами і використовує загальний простір ОЗП всіх комп'ютерів разом. Обробка в пам'яті усуває всі повільні звернення до даних і покладається виключно на дані, що зберігаються в ОЗП [1]. Загальна продуктивність обробки не знижується через затримку, яка зазвичай спостерігається при доступі до жорстких дисків або твердотільних накопичувачів. Обробка в пам'яті часто виконується за допомогою технології відомої як мережі даних у пам'яті.

Такі СУБД мають перевагу у швидкості порівняно з традиційними дисковими СУБД. Багато існуючих реляційних СУБД підтримують механізм In Memory, наприклад, пізні версії MS SQL Server, Oracle Database та інші.

За результатами представленого аналізу вибираємо як технологію побудови ВІ -системи для аналізу великих даних технологію, засновану на застосуванні реляційної СУБД, що підтримує механізм In-Memory.

Література

1. In-Memory Processing [Електронний ресурс]. – Режим доступу: <https://hazelcast.com/glossary/in-memory-processing/> (дата звертання: 04.12.2024).