

УДК 004.65

І.І. Яремцьо

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

АНАЛІЗ СХОВИЩ ДАНИХ ТА СУБД ДЛЯ РОБОТИ З ЧАСОВИМИ РЯДАМИ

UDC004.65

I.I. Yaremtso

ANALYSIS DATA WAREHOUSE AND DBMS TO WORK WITH TIME SERIES

Для аналізу обмежимося вбудованими СУБД та сховищами типу «ключ-значення». Вся інформація бралася з офіційної документації.

LevelDB. Це сховище «ключ-значення» від Google. Переваги: ключі та значення представляються довільним набором байт; дані зберігаються у відсортованому до ключа вигляді; розробнику надається можливість перевизначити функцію порівняння, щоб задати потрібний порядок сортування; простий інтерфейс для базових операцій – Put (key, value), Get (key), Delete (key); кілька операції можна виконати як одну (batch); можливість зробити знімок БД; стиснення даних через Snappy (бібліотеку для компресії даних від Google); підтримка обходу ітератором в прямому і зворотному порядку. Обмеження: не підтримує реляційну модель даних, індекси і мову SQL; тільки один процес може мати доступ до БД в певний момент часу.

HyperLevelDB. Це форк LevelDB, в якому були збережені всі її переваги і покращено багатопотокову взаємодію з БД. LevelDB влаштована таким чином, що в певний момент часу тільки один потік може щось записувати в БД, а якщо таких потоків декілька, то для їх синхронізації використовується м'ютекс (семафор) і запис проводиться за принципом черги (FIFO – first in, first out – першим увійшов, першим вийшов): доступ до БД отримує потік, що знаходиться в «голові» черги, а всі інші чекають. Це гарантує, що одночасна робота цих потоків з БД буде безпечна і цілісність даних не порушиться. HyperLevelDB дозволяє всім потокам одночасно записувати дані в БД, тим самим роблячи паралельну роботу більш ефективною, а синхронізація досягається за рахунок іншого механізму.

BangDB. Сховище «ключ-значення», котре може перебувати і в оперативній пам'яті, і на диску. Основна перевага – висока продуктивність при багатопотоковій роботі. Дана БД підтримує ACID-транзакції і використовує компоненти, з допомогою яких досягається оптимізація ресурсів, що гарантує високу продуктивність і низьку затримку. BangDB реалізує свій буферний кеш, систему управління буфером, механізм управління пам'яттю і т.д., також має багатий API, який крім стандартних CRUD-операцій забезпечує гнучкий інтерфейс запитів (з використанням В-дерев). Дане сховище може бути серверним або вбудованим.

Sophia. Сучасне сховище «ключ-значення» з підтримкою транзакцій, яке може розташовуватися як в оперативній пам'яті, так і на диску. Розроблене для забезпечення його оптимальної роботи без деградації в часі і гарантує складність порядку $O(1)$ в найгіршому для операцій читання і запису. Переваги: підтримка ACID-транзакцій; управління паралельним доступом за допомогою багатoversійності; оптимістичне блокування; стиснення даних; простий інтерфейс; не має зовнішніх залежностей.

SQLite. Відкрита вбудована реляційна СУБД. Переваги: вбудована (звернення до БД відбувається не за рахунок клієнт-серверної взаємодії, а як виклик бібліотечної функції, що збільшує продуктивність); БД є одним файлом; підтримка ACID-транзакцій; підтримка багатопотокового читання; підтримка динамічно типізованих типів даних; підтримка SQL. Недоліки: відсутність призначеного для користувача управління; бідні можливості додаткового налаштування.