

ВИКОРИСТАННЯ МАТЕРІАЛІЗОВАНИХ ПРЕДСТАВЛЕНЬ ДЛЯ ПІДВИЩЕННЯ ШВИДКОДІЇ СКБД

¹Вінницький національний технічний університет

Анотація

Проаналізовано застосування матеріалізованих представлень для підвищення швидкості обробки запитів в СКБД. Досліджено підхід інкрементного підтримання представлень за допомогою «дельта-запитів» та використання цілісності даних. Визначено перспективні напрями дослідження з застосуванням адаптивних стратегій оновлення матеріалізованих представлень.

Ключові слова: СКБД, програмна інженерія, матеріалізоване представлення, інкрементне підтримання представлень, оптимізація запитів.

Abstract

The paper analyses the use of materialised views to increase the speed of query processing in DBMS. It investigates the approach of incremental view maintenance using delta queries and data integrity. The paper identifies promising areas of research involving the use of adaptive strategies for updating materialised views.

Keywords: DBMS, software engineering, materialised view, incremental view maintenance, query optimisation.

Вступ

Швидкодія виконання запитів до баз даних залишається одним із ключових показників ефективності сучасних інформаційних систем. Класичні засоби оптимізації швидкодії, як-от індексування, можуть бути недостатньо ефективними в великих сховищах даних або аналітичних системах через значне збільшення розміру самих структур індексів [1]. Крім того, часті оновлення або додавання нових записів можуть зменшити ефективність інших методів, як-от сегментація даних, де швидкодія досягається розподілом БД між окремими вузлами [2], що можуть бути обмежені пропускнуою здатністю мережі при занадто великих об'ємах даних.

Матеріалізовані представлення є різновидом кешування результатів запитів в СКБД, що можуть бути більш компактними та швидкими за індексування [3]. Крім того, вони можуть застосовуватись з методами інкрементного підтримання представлень для вирішення класичної проблеми застарілих даних в кеші. Тому, питання дослідження переваг, недоліків та шляхів розвитку матеріалізованих представлень для оптимізації швидкодії систем керування базами даних є актуальним.

Метою дослідження є аналіз застосування матеріалізованих представлення в системах керування базами даних.

Об'єктом дослідження є процес підвищення швидкодії систем керування базами даних.

Предмет дослідження – матеріалізовані представлення в системах керування базами даних.

Основна частина

Матеріалізоване представлення (англ. Materialized View, MV, іноді – матеріалізований розріз) базується на ідеї кешування результатів комплексних запитів для подальшого їх використання в наступних операціях без потреби повторних обчислень. Крім того, матеріалізовані представлення є компактнішими в порівнянні з індексами та деякими іншими способами оптимізації запитів, так як зберігають інформацію лише для конкретних комплексних запитів, а не усієї БД або таблиці. Швидкодія та компактність MV роблять їх особливо привабливими в аналітичних середовищах [4]. До типових сценаріїв застосування MV в СКБД можна віднести різноманітні сховища даних (англ. Data Warehouses), OLAP-системи або хмарні сервіси на зразок Amazon Redshift, де зниження часу відповіді на складні аналітичні запити є критичним операційним параметром.

Проте, матеріалізовані представлення надають перевагу лише тоді, коли вони актуальні. Повне оновлення кожного MV при кожній зміні бази даних може погіршити їх швидкість або ж навпаки навіть вповільнити виконання запитів. Інкрементне підтримання представлень (англ. Incremental View Maintenance, IVM) розширює концепцію MV та спрямоване на вирішення цієї проблеми. IVM пропонує відслідковувати зміни в БД та реагувати на них окремими оновленнями самого матеріалізованого представлення, а не повторним виконанням запиту та його кешуванням [5].

Класичний підхід до IVM базується на обчисленні «дельта-запитів» – запитів, результати яких при об'єднанні з MV дадуть ті ж самі результати, якби оригінальний запит для побудови MV було виконано повторно [6]. Цей підхід є найбільш популярним рішенням у комерційних СКБД та різноманітних аналітичних системах. Однак, він має очевидне обмеження: для складних запитів із використанням великої кількості таблиць вартість обробки одного оновлення та формування необхідної дельти змін може наближатись до вартості виконання оригінального MV. Зокрема, стандартний IVM для матеріалізованого представлення з k таблицями вимагає UNION ALL запиту з k гілками – по одній для кожної базової таблиці, що призводить до надлишкових обчислень навіть у випадках, коли більшість гілок повертають порожній результат.

Іншим практичним напрямом оптимізації IVM є використання цілісності даних, як-от обмежень зовнішніх ключів. Для матеріалізованих представлень, побудованих на з'єднаннях таблиць через проміжні відношення, знання про зовнішній ключ дозволяє скоротити кількість гілок UNION ALL до мінімуму [7]. Теоретично це базується на логічних зв'язках між даними: нещодавно вставлені рядки однієї таблиці з більшим шансом можуть бути з'єднані з нещодавно вставленими рядками іншої, і аналогічно для видалень. Практичні результати на платформі Amazon Redshift в одному з досліджень показали прискорення дельта-обчислень до 2.7 разів і загального процесу оновлення до 2 разів [7].

Проте, обидві розглянуті стратегії не позбавлені обмежень. Мета-евристичні підходи до вибору матеріалізованих представлень зосереджені на мінімізації вартості обробки запитів, вартості їх підтримання, відповіді тощо, але стикаються зі зростанням часу виконання при збільшенні розміру вибірки даних [4]. Перспективним є вивчення умов, за яких IVM досягає «найкращого можливого» режиму – константного або мінімально можливого часу оновлення та аналогічної затримки перерахування. Цей напрям отримав розвиток у системах Dynamic Yannakakis, F-IVM та Crown, і активно впроваджується в комерційній системі RelationalAI [8].

Подальші дослідження у цій галузі можна структурувати за кількома напрямками. По-перше, поширення теорії оптимального IVM на рекурсивні запити, запити з агрегацією та вкладеними підзапитами, де застосування розглянутих підходів є обмеженим. По-друге, розробка адаптивних стратегій, що динамічно перемикаються між різними схемами підтримання стану MV залежно від статистики оновлень. Це може включати поєднання IVM із машинним навчанням та аналітикою в базах даних (англ. In-Database ML) для оптимального вибору набору матеріалізованих представлень в умовах динамічно змінюваних робочих навантажень.

Висновок

Було досліджено концепцію матеріалізованих представлень як підходу до оптимізації виконання складних запитів за рахунок кешування їх результатів та визначено типові сценарії, у яких MV є найбільш ефективними. Проведено аналіз проблеми актуальності матеріалізованих представлень і розглянуто інкрементне підтримання представлень (IVM) як механізм зменшення вартості оновлення стану MV при змінах у базі даних. Проаналізовано класичний підхід «дельта-запитів» та показано його обмеження для складних представлень із великою кількістю базових таблиць. Досліджено практичні підходи оптимізації IVM із використанням обмежень цілісності даних, як-от зовнішніх ключів, що дозволяють зменшити обсяг дельта-обчислень і прискорити оновлення MV. Дослідження показало, що подальший розвиток IVM доцільно пов'язувати з розробкою адаптивних стратегій, які можуть перемикатися між схемами підтримання MV залежно від статистики оновлень і робочих навантажень.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Mohamed M. A. Challenges and recommendations in big data indexing strategies [Electronic resource] / Mohamed Attia Mohamed, Manal A. Abdel-Fattah, Ayman E. Khedr // International Journal of e-Collaboration. – 2021. – Vol. 17, no. 2. – P. 22–39. – Mode of access: <https://doi.org/10.4018/ijec.2021040102> (date of access: 17.02.2026). – Title from screen.

2. Миргородський А. В. Розвиток методів сегментації даних в розподілених системах керування базами даних [Електронний ресурс] / Андрій Вікторович Миргородський, Оксана Володимирівна Романюк // Наукові праці Вінницького національного технічного університету. – 2025. – № 4. – Режим доступу: <https://doi.org/10.31649/2307-5376-2025-4-87-98>. – Назва з екрана.
3. Rizzi S. View materialization vs. indexing: balancing space constraints in data warehouse design / Stefano Rizzi, Ettore Saltarelli // Proceedings of the 15th international conference on advanced information systems engineering. – Berlin, Heidelberg, 2003. – P. 502–519.
4. Srinivasarao P. A novel hybrid optimization algorithm for materialized view selection from data warehouse environments [Electronic resource] / Popuri Srinivasarao, Aravapalli Rama Satish // Computer systems science and engineering. – 2023. – Vol. 47, no. 2. – P. 1527–1547. – Mode of access: <https://doi.org/10.32604/csse.2023.038951>. – Title from screen.
5. DBSP: automatic incremental view maintenance for rich query languages [Electronic resource] / Mihai Budiu [et al.] // Proceedings of the VLDB Endowment. – 2023. – Vol. 16, no. 7. – P. 1601–1614. – Mode of access: <https://doi.org/10.14778/3587136.3587137>. – Title from screen.
6. Olteanu D. Recent increments in incremental view maintenance [Electronic resource] / Dan Olteanu // Companion of the 43rd symposium on principles of database systems. – [S. l.], 2024. – P. 8–17. – Mode of access: <https://doi.org/10.1145/3635138.3654763>. – Title from screen.
7. Foreign keys open the door for faster incremental view maintenance [Electronic resource] / Christoforos Svingos [et al.] // Proceedings of the ACM on management of data. – 2023. – Vol. 1, no. 1. – P. 1–25. – Mode of access: <https://doi.org/10.1145/3588720>. – Title from screen.
8. Maintaining triangle queries under updates [Electronic resource] / Ahmet Kara [et al.]. – Mode of access: <https://arxiv.org/abs/2004.03716>. – Title from screen.

Миргородський Андрій Вікторович – аспірант групи 121-24а, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e-mail: mirgorodskijav@gmail.com

Романюк Оксана Володимирівна – к.т.н., доцент кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, e-mail: romaniukoksana@gmail.com

Myrhorodskiy Andrii – graduate student of group 121-24a, Faculty for Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: mirgorodskijav@gmail.com

Oksana Romaniuk – Candidate of Technical Sciences, Associate Professor of the Software Chair, Vinnytsia National Technical University, Vinnytsia, e-mail: romaniukoksana@gmail.com