

ВИКОРИСТАННЯ МЕТОДУ МАТЕРІАЛІЗОВАНИХ ПРЕДСТАВЛЕНЬ В РЕІНЖИНІРИНГУ БАЗ ДАНИХ

Горішня К.О., Шебанов Є.О., Хацько Н.Є.

Харківський національний університет радіоелектроніки, м. Харків

Національний технічний університет

«Харківський політехнічний інститут», м. Харків

Метою даної роботи є дослідження методу матеріалізації даних як одного з методів реінжинірингу БД для визначення його ефективності та перспективи подальшого застосування на практиці, зокрема для документообігу [1].

Необхідні ефективні рішення для трьох проблем: визначення того, які представлення матеріалізувати, включно з тим, як їх зберігати та індексувати; ефективне оновлення матеріалізованих представлень під час оновлення базових таблиць; ефективне використання матеріалізованих представлень для прискорення обробки запитів.

Пропонується план дослідження методу матеріалізованих представлень, який передбачає відправку агрегованих запитів до БД зі зміною кількості інформації, що запитується, від 100 до 1000 рядків. Для визначення впливу методу матеріалізованих представлень на час виконання запиту проводились серії однакових запитів без під'єднання метода та з під'єднанням. Для проведення експерименту обрана БД «AdventureWorks sample databases» компанії Microsoft з відкритим доступом, розмір цієї бази становить 178.75 Мб. Наведемо приклад агрегованого запиту зі вказанням створення матеріалізованого представлення:

```
CREATE MATERIALIZED VIEW SalesSummary AS
SELECT c.CategoryName, YEAR(soh.OrderDate) AS OrderYear, SUM(sod.OrderQty * sod.UnitPrice) AS TotalSales
FROM Sales.SalesOrderDetail sod
JOIN Sales.SalesOrderHeader soh ON sod.SalesOrderID = soh.SalesOrderID
JOIN Production.Product p ON sod.ProductID = p.ProductID
JOIN Production.ProductCategory c ON p.ProductCategoryID = c.ProductCategoryID
WHERE soh.OrderDate BETWEEN '2017-01-01' AND '2017-12-31'
GROUP BY c.CategoryName, YEAR(soh.OrderDate);

SELECT * FROM SalesSummary ORDER BY TotalSales DESC;
```

За результатами експериментів, використання матеріалізованих представлень дозволяє майже вдвічі скоротити час отримання відповіді на агрегаційні запити при розмірі відповідей до 600 рядків. Після цього зростання швидкості повільно спадає. Таким чином, зменшення часу відповіді на запити забезпечує більш ефективну роботу системи в цілому, зазначений метод доцільно використовувати для запитів, що часто робляться на відповіді на них не перевищують 1000 рядків. Цей метод має недолік – потрібна додаткова пам'ять для зберігання представлень, тому використовувати його потрібно з урахуванням доцільності підвищення швидкості роботи певної системи.

Література:

1. Vlasenko, P., Khatsko, N. Difficulties in the implementation of electronic document management by small businesses related to personnel in Ukraine. // Питання сучасної модернізації науки та освіти : зб. наук. ст. / Харків. нац. автомоб.-дор. ун-т. – Харків : ХНАДУ, 2023. – С. 179–183.