

УДК 681.3.01

Ю.А.Костів

Тернопільський національний технічний університет імені Івана Пулюя, Україна

КОМП'ЮТЕРНА СИСТЕМА БАГАТОВИМІРНОЇ ОБРОБКИ ДАНИХ НА ОСНОВІ OLAP-ТЕХНОЛОГІЇ

Yu.A. Kostiv

COMPUTER SYSTEM OF MULTIDIMENSIONAL DATA PROCESSING BASED ON OLAP- TECHNOLOGIES

Сьогодні важливим компонентом успіху в бізнесі є розумне використання СППР. Системи аналізу даних, побудовані на основі OLAP (Online Analytical Processing - оперативний аналіз даних), дозволяють виконувати швидкий і ефективний аналіз над великими об'ємами даних. Дані зберігаються в багатовимірному вигляді, що найбільш близько відображає природний стан реальних бізнес-даних. Засоби OLAP-аналізу можуть добувати дані безпосередньо з реляційних систем і з багатовимірних СУБД [1].

Основними методами аналітичної обробки даних в СППР є: підтримка прийняття рішень через аналіз даних – OLAP; Knowledge Discovery in Databases - виявлення знань у базах даних; Data Mining – добування даних.

СБЕАД (система багатовимірного експрес-аналізу даних), на відміну від статичних систем, дозволяє користувачу динамічно оперувати даними, швидко створювати нові звіти і переглядати їх результати. Основна мета, яку необхідно досягнути при розробці СБЕАД – швидкість опрацювання даних. Це дозволяє використовувати їх в процесі інтерактивної роботи аналітика з інформацією. Тому, необхідно використовувати OLAP-системи [1].

Першим кроком при розробці архітектури СБЕАД є аналіз запитів і вимог, пропонує користувачами до СБЕАД. Було проведено декілька консультацій зі співробітниками підприємства, які планують використати СБЕАД і з'ясовано перелік вимог до системи. Було визначено первісний перелік питань, відповіді на які необхідно буде одержати. Найбільше значення мають не самі формулювання питань, а перелік фактів, подій, об'єктів, які будуть потрібні для відповідей на такі питання, і які необхідно зберігати в сховищі даних.

OLAP-клієнт може бути реалізований декількома методами. По-перше, можна використати ActiveX-елемент, що входить до складу Microsoft Office і забезпечує можливість аналізу багатовимірних даних та відображення різних звітів. По-друге, мовою високого рівня може бути написаний клієнтський додаток, що виконує віддалений доступ до сховища даних, реалізуючи при цьому будь-які специфічні потреби користувачів. По-третє, можна використати інтерфейс ADO MD (Active Data Objects MultiDimensional), завдяки якому можна створювати Веб-орієнтовані додатки.

В роботі досліджено OLAP-технологію для реалізації СБЕАД, зокрема проаналізовано поняття сховища даних, основні елементи OLAP, описано архітектуру OLAP-додатків. Також проведено аналіз реалізації СБЕАД в розподілених середовищах та запропоновано вирішення проблеми проектування високоефективних паралельних архітектур серверів БД (симетричні багатопроцесорні системи (SMP) - найбільш часто використовується форма сильно зв'язаних багатопроцесорних систем, тобто систем, що розділяють єдину оперативну пам'ять і найбільш часто - дискову підсистему; слабозв'язані багатопроцесорні системи - сукупність самостійних комп'ютерів, об'єднаних у єдину систему швидкодіючою мережею й, можливо, що мають загальну дискову підсистему, як, наприклад, кластерні інсталяції; системи з масовим паралелізмом (MPP) - системи із сотнями й навіть тисячами процесорів, деталі їхньої реалізації можуть значно розрізнятися). Виділено чотири групи вимог, які визначають з технічної точки зору споживачів якості сучасної СУБД: масштабування; продуктивність; можливість змішаного

завантаження різними типами завдань; забезпечення постійного доступу даних. Для досягнення максимальної продуктивності СУБД при змішаному завантаженні необхідно використати оптимізатор запитів, керування ресурсами й паралельною обробкою запитів.

Процес проектування багатовимірної БД складається з декількох етапів: визначення запитів та вимог до системи; визначення структури БД (виміри й факти, їхні взаємозв'язки й рівні агрегації збережених даних); визначення реляційної схеми («сніжинка» - дозволяє заощадити дисковий простір, збільшується швидкість перегляду вимірів); реалізація багатовимірної БД в MS OLAP Services. Останній етап повинен бути виконаний у такій послідовності: створення БД - New Database; створення джерела даних - Data Sources; створення кубів – Cube Wizard; редагування створеної схеми куба (за потреби) – Cube Editor; проектування фізичного сховища - Design Storage; визначення ступені «агрегації» куба – Storage Design Wizard; завантаження даних в куб - Process Now. Після цього їх можна переглянути засобом Cube Browser.

Проблема впровадження СБЕАД у середовище WWW стоїть досить гостро з кількох причин. По-перше, використання в якості програми-клієнта звичайного Веб-браузера дозволяє знизити вимоги до системних ресурсів клієнтського робочого місця. По-друге, якщо буде потреба внесення будь-яких змін в інтерфейс або в програмну частину, необхідно змінити програму лише на Веб-сервері. Після цього всі користувачі відразу одержать доступ до оновленої версії СБЕАД. По-третє, використання мережі Internet дозволить мати доступ до СБЕАД, хоча це, у свою чергу, має на увазі необхідність створення надійних засобів забезпечення безпеки. По-четверте, WWW-технологія створювалася для забезпечення багатокористувацької роботи, що також є однією з основних вимог до СБЕАД.

Реалізація Веб-клієнта для доступу до СБЕАД можлива, наприклад, із застосуванням інтерфейсу ADO MD, що використовує технологію Active. ASP-програма, написана мовою VB Script, дозволяє будь-якому клієнтові мережі Internet з'єднатися з OLAP сервером, що перебуває за тисячі кілометрів від нього, і за допомогою інтерфейсу ADO через спеціальний OLE DB Provider для OLAP переглядати таблиці з аналітичними даними. Дана програма під'єднується до OLAP-сервера й по заданому MDX-запиті показує результати в HTML-таблиці. Для виконання програми з HTML-браузера, треба попередньо скопіювати її в каталог WEB-сервера. Потім потрібно за допомогою OLAP-менеджера встановити права читання інформаційного куба для всіх користувачів, що підключаються, у режимі читання. Цю програму можна виконати із браузера з будь-якого комп'ютера локальної мережі. Наступний варіант реалізації www-клієнта для OLAP-сервера простіший, ніж попередній. У веб-сторінку встановлюється ActiveX-компонент Pivot Table, що на основі інтерфейсу OLE DB for OLAP здійснює взаємодію з OLAP-сервером. Для реалізації такого клієнта необхідно мати на серверному комп'ютері MS Office із установленим Microsoft Front Page. Однак це призводить до необхідності установки даного ActiveX-компонента на кожному із клієнтських місць. Це є обмеженням, але тільки лише у випадку, якщо надається доступ клієнтів через мережу Internet. У випадку роботи в мережі Intranet, що є більше ймовірним для СППР, ця проблема знімається, тому що не є складно встановити ActiveX-компонент на декількох клієнтських комп'ютерах, на яких користувачі будуть робити аналіз даних.

Література

1. Джеффри Д. Ульман. Анализ больших наборов данных / Джеффри Д. Ульман, Ананд Раджараман, Юре Лесковец. – М.: Литрес, 2016. – 500 с.