

УДК 004.94

К.Я.Бортник

Луцький національний технічний університет

АНАЛІЗ ОПЕРАТИВНИХ ДАНИХ ЗА ДОПОМОГОЮ OLAP-КУБІВ

В роботі викладено проблеми аналізу оперативних даних. Запропоновано структуру аналітичної системи за допомогою OLAP-кубів та запропоновано розробку модулю вивантаження даних з облікової системи у сервер з використанням технології DSO, вбудованої в Microsoft SQL Server 2008 Enterprise/Standart Edition.

Ключові слова: *OLAP-куби, сервер - OLAP, OLAP-звіти, Microsoft SQL Server 2008.*

Постановка проблеми. Аналіз оперативних даних за допомогою OLAP-кубів в загальному випадку зводиться до пошуку оптимальної структури аналітичної системи, яка забезпечила б значне скорочення часу на обробку даних і формування звітів та унікальну гнучкість в побудові звітності.

Мета дослідження. Дослідити поточний стан галузі, наявні на ринку програмні продукти та технології, виявити найбільш перспективні з них для використання, як в роботі українських підприємств, так і з точки зору використання у навчальному процесі з підготовки спеціалістів-аналітиків даних на основі використання OLAP-кубів.

Основні результати дослідження. Велику частку ринку ПЗ для автоматизації середніх торгових фірм займають продукти, облікові системи, побудовані, як на реляційних СУБД (SQL-версії) так і на СУБД з файловою структурою.

Яскравим прикладом цьому є фірма "1С" зі своїми програмами системи "1С: Підприємство" ("1С: Торгівля і Склад", конфігурації "Торгівля + Склад + Бухгалтерія" і "Бухгалтерія + Торгівля + Склад + Зарплата + Кадри").

Зазвичай такого класу програми включають набір документів для відображення типових торговельних операцій та отримання звітності. Але виникає проблема: при певній інтенсивності операцій, що реєструються в програмі, формування звітів (таких як "Аналіз продажів", "Партії товарів") триває протягом помітних проміжків часу - від хвилини і довше, досягаючи години і більше у компаній з великою кількістю товару і проведених операцій. Дуже сильно впливає на час формування звітності період вибірки даних. Так, звіт про рух товару за останній день робиться кілька секунд, за останній же рік - більше 10 хвилин. Навіть в SQL - версіях звіти за великими обсягами даних формуються хоч і в рази швидше DBF-версій, але далеко не в режимі реального часу.

Друга проблема, яка актуальна для користувачів облікових програм - досить обмежений набір аналітичних звітів. Так, отримати необхідні дані про закупівлі, продажу, оплати й заборгованості цілком реально, але форми звітів, що поставляються в початкових конфігураціях, не можна вважати достатніми, а часто навіть зручними для аналізу того, що відбувається на фірмі. Наприклад можлива ситуація, коли в початкових конфігураціях немає звітів, що дозволяють аналізувати динаміку складських залишків (система координат "Залишок / День") - справа часто в тому, що такий звіт, виконаний штатними засобами, буде формуватися дуже довго.

Користувачі вирішують цю проблему або за допомогою власних програмістів, що підтримують програму, або, як у випадку 1С, звертаючись до спеціалізованих фірм мережі "1С: Франчайзинг".

Завдяки вбудованим в "1С: Підприємство" засобів розробки чи наявності якісної підтримки розробників інших програм, будь-які придумані звіти можуть бути реалізовані за час від годин до днів. Така ситуація є гальмуючим фактором при прийнятті управлінських рішень, а в умовах високої конкуренції, своєчасне їх прийняття може дуже вплинути на кінцевий результат.

В процесі висвітлення основних проблем аналізу оперативних даних була поставлена задача розробити єдину систему звітності, що надала б можливість оптимізувати усі бізнес-потоки і при цьому дозволила б відмовитися від рутинної ручної роботи. При цьому, були сформовані основні вимоги до аналітичної системи:

- можливість отримати будь-який аналітичний звіт без виклику програміста;

- час отримання звіту в межах секунд за будь-який період;
- наскрізний аналіз даних за кілька років;
- наявність єдиного сховища даних і системи аналізу інформації.

Вирішення всіх цих проблем було знайдено за допомогою OLAP-кубів (оперативний багатовимірний аналіз). Бази даних Online Analytical Processing (OLAP) полегшують створення запитів для оброблення й аналізу інформації. OLAP — це технологія баз даних, оптимізована для створення запитів та звітів, а не для проведення транзакцій. Вихідними даними для OLAP є бази даних Online Transactional Processing (OLTP), які часто зберігаються в інформаційних сховищах. Дані OLAP створюються на основі цих накопичених даних та впорядковуються в структурах, що дозволяють проводити складний аналіз. Дані OLAP також впорядковані ієрархічно та зберігаються в кубах, а не в таблицях. Це складна технологія, яка використовує багатовимірні структури, що надають швидкий доступ до даних для аналізу. Така організація полегшує відображення зведень високого рівня у звітах зведених таблиць або зведених діаграм, наприклад, обсягів продажу в країні або в регіоні, а також відображення відомостей про місця, де рівень продаж особливо високий або низький.

Бази даних OLAP створені для прискорення витягнення даних. Оскільки підрахунок сумарних значень здійснює сервер OLAP, а не Microsoft Office Excel, менше даних потрібно надсилати до Excel під час створення або змінення звіту. Такий підхід надає можливість працювати з набагато більшою кількістю вихідних даних, ніж у разі, якщо б дані було оформлено в традиційному стилі, коли Excel завантажує усі окремі звіти, а тоді підраховує сумарні значення.

Бази даних OLAP містять два основних типа даних: показники, які є числовими даними, кількості та середні значення, які використовуються для прийняття ділових рішень, та виміри, які є категоріями, використовуваними для організації цих показників. Бази даних OLAP допомагають організувати дані за багатьма рівнями деталізації з використанням тих самих категорій, за допомогою яких здійснюється аналіз даних.

Куб – це структура даних, яка об'єднує показники за рівнями та ієрархіями кожного з вимірів, які потрібно проаналізувати. Куби поєднують кілька вимірів — наприклад, час, розташування та лінію продуктів — з сумарними даними, такими як показники продажів або резерву.

OLAP-аналіз даних є життєвою необхідністю при управлінні сучасним підприємством. Він дозволяє керівникові узагальнювати різносторонню інформацію про роботу компанії й ухвалювати обґрунтовані тактичні й стратегічні рішення.

Для вирішення вищезазначених проблем залишалось вибрати технологічну платформу. За основу було взято структуру типової аналітичної системи:

- 1) Сервер OLAP (наприклад MS Analyses Services 2005/2008);
- 2) Реляційна СУБД сумісна з обраним сервером OLAP (наприклад Microsoft SQL Server 2005/2008);
- 3) Модуль вивантаження даних з облікової системи у сервер OLAP (виконується програмістами на замовлення);
- 4) Типові OLAP-куби "Залишки", "Рух грошових коштів", "Продаж" та "Закупівлі" і т.д. (формулюються програмістами на початок роботи);
- 5) Клієнт OLAP для роботи менеджерів (наприклад Microsoft reporting Services 2005/2008).

На базі цього зроблено ґрунтовний аналіз розробок провідних компаній.

Microsoft Analysis Services (Microsoft SQL Server) не був оригінальною розробкою корпорації. Технологія, розроблена в 1996 році, була куплена у компанії Panorama Software Systems. Продукт Microsoft, створений на основі цієї технології, швидко завоював ринок OLAP-рішень за рахунок вдосконаленої архітектури зберігання даних із застосуванням ROLAP / MOLAP / HOLAP, а також демпінгових цін і агресивного маркетингу. Analysis Services 2005 виграв в ціні у конкурентів, крім того, корпорація спочатку вбудовували зачатки OLAP-системи в MS Office, а в Excel 2000 спеціальний додаток PivotTables допомагало інтегруватися у вигляді OLAP-клієнта до Microsoft Analytic Services.

У Microsoft SQL Server 2005 присутня єдина модель OLAP і реляційної звітності. Моделі можна будувати з довільних розрізах куба. При виборі одного виміру інші, пов'язані з ним, можна розглядати як вкладені таблиці, причому до кожного допускається застосування незалежних розрізів. Продукт дозволяє застосовувати OLAP-модель до глобального кубу, ділити його на групи,

створювати кластери в режимі реального часу. Сервер дозволяє постійно оновлювати куби, зводячи різницю між реляційними і багатовимірними базами до мінімуму.

Microsoft завжди намагалася залучати інших розробників до створення інструментарію, тому окремо варто згадати продукти, інтегровані з Microsoft Analysis Services. Один з них - ProClarity Analytics Platform, OLAP-клієнт від компанії Proclarity (колишньої Knosys), що дозволяє взаємодіяти з іншими MS додатками, такими як Excel, PowerPoint, Outlook Digital Dashboards і з web-додатками. Пакет дає можливість віддаленої роботи з базами даних. Серед користувачів продуктів Proclarity такі компанії, як Compaq, Ericsson, Hewlett-Packard, L'Oreal, Pennzoil, Siemens, а також Swiss Stock Exchange.

Ще одна платформа для швидкого розгортання систем аналітики та звітності, інтегруються з Microsoft Analysis Services, - Panorama NovaView компанії Panorama Software. Цей OLAP-клієнт забезпечує зв'язку між Microsoft Reporting Services, Microsoft Excel і Microsoft Share Point Portal Server. Як операційні джерел в Panorama NovaView можуть бути використані будь-які бази даних, у тому числі Oracle, Teradata, DB2 і Microsoft.

Другим гравцем на ринку OLAP-систем є компанія Hyperion Solutions Corporation, яка пропонує одразу кілька рішень. Hyperion Performance Suite включає в себе два програмних продукту: Hyperion Intelligence, систему формування запитів і аналізу отриманої інформації, а також Hyperion SQR, яка формує презентаційні звіти. Рішення забезпечує взаємодію з реляційними, багатовимірними і операційними базами даних усіх відомих розробників, включаючи Oracle, DB2, MS SQL Server та інші. Звіти можна формувати практично у всіх форматах, включаючи XML, PDF, HTML, текст, PostScript, і т.д. Ще один OLAP-сервер від цього виробника, Hyperion Essbase, що використовує технологію HOLAP, підтримує обсяги даних, що вимірюються в терабайтах, доступ до яких можуть отримати тисячі користувачів одночасно. Середній час виконання запиту при цьому, за твердженням розробників, не перевищує одну секунду.

Цікаво, що продукти Hyperion Solutions в тій чи іншій мірі містять технології, розроблені одинадцятьма компаніями в різний час. Тому при створенні нових версій своїх рішень, корпорація приділила особливу увагу більш глибокої інтеграції компонентів. Крім того, був розроблений більш дружній інтерфейс.

Дещо поступається OLAP-гігантам на ринку компанія Cognos, в числі найбільш популярних її рішень у цій сфері - Cognos Enterprise Planning і Cognos PowerPlay. Перше є інструментом, що дозволяє вести деталізований план доходів і витрат, здійснювати управління активами з урахуванням фінансових ризиків і погоджувати діяльність всіх підрозділів компанії. Це рішення, випущене після придбання Cognos в 2003 році компанії Adaytum, забезпечує обмін інформацією між різноманітними джерелами і одержувачами. Cognos PowerPlay - засіб OLAP-аналізу, що дозволяє будувати багатовимірний куб, що містить понад 2 мільярдів різних даних і категорій. Користувачі можуть отримати доступ до кубів, використовуючи веб-клієнти, засоби операційних систем або Excel.

В останніх версіях своїх продуктів Cognos перейшов на нову платформу. Раніше основний акцент компанія робила саме на OLAP-звітах, зараз користувач може працювати одночасно з реляційними базами, а також з спрощеним ROLAP і повним MOLAP серверами. У цьому плані Cognos краще справляється з OLAP, ніж інші конкуренти, такі як Business Objects. При цьому нова версія надає просунуті можливості та інструменти OLAP-сервера для користувачів.

А тому, на сьогоднішній день для малих та середніх фірм запропоновано структуру аналітичної системи за допомогою OLAP-кубів (див.табл.1)

Таблиця 1.

Рекомендоване рішення для впровадження OLAP-аналізу

№ п/п	Тип продукту	Виробник	Назва продукту
1	2	3	4
1	Сервер OLAP	Microsoft	Microsoft Analyses Services 2005/2008 в складі Microsoft SQL Server 2008 Enterprise/Standart Edition
2	Модуль вивантаження даних	Власна розробка	Модуль вивантаження даних з облікової системи у сервер з використанням технології DSO вбудованої в Microsoft SQL Server 2008 Enterprise/Standart Edition

1	2	3	4
3	Типові OLAP-куби	Microsoft + власна доводка	Типові шаблони аналітичних кубів та зрізів в складі Microsoft SQL Server 2008 Enterprise/Standart Edition. Настройка і доводка власними силами
4	Клієнт OLAP	Microsoft	Microsoft Reporting Services 2005/2008 в складі Microsoft SQL Server 2008 Enterprise/Standart Edition
5	Сервер для аналізу (апаратна платформа)	Будь-який	Мінімальні орієнтовні вимоги : Два процесори Intel Xeon E5504, 16Gb пам'яті DDR3-1333 ECC, 4-и диски 146GB SAS hot-swap, RAID, Windows Server 2003 Standart R2 64-bit

Висновки

В роботі викладено проблеми аналізу оперативних даних. Зроблено ґрунтовний аналіз новітніх технологій аналізу оперативних даних за допомогою OLAP-кубів та запропоновано структуру аналітичної системи на сьогоднішній день для малих та середніх бізнес-структур.

За прогнозами The OLAP Report, в 2010 році сегмент OLAP перетне бар'єр в \$ 6 млрд. Тим не менш, він поки не став мейнстрімом на ринку корпоративного ПЗ. Але, на відміну від деяких інших сегментів, більшість компаній, які вже впровадили OLAP-системи, в основному задоволені результатом. За результатами опитування OLAP Surveys, проведеного Survey.com, 73% підприємств задоволені впровадженням OLAP.

1. Артемов А. Microsoft SQL Server для професіоналов. – М.: Издат. дом "Вильямс", 2002. — 576 с.
2. Бергер А., Горбач И., Меломед Э., Щербинин В., Степаненко В. «Microsoft SQL Server 2005 Analyses Services. OLAP и многомерный анализ данных» СПб «БХВ-Петербург», 2007, ISBN 978-5-94157-158-1
3. Баргесян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. «Методы и модели анализа данных : OLAP и Data Mining» СПб «БХВ-Петербург», 2004, ISBN 5-94157-522-X
4. www.microsoft.com/ukraine/casestudies.