

Положительное влияние на экономику иностранных инвестиций, привлеченных в форме кредитов, носит ограниченный характер, так как кредиты необходимо возвращать с процентами. Гораздо важнее привлечение прямых и портфельных инвестиций в различные отрасли экономики Беларуси.

Полагаем, что притоку в инвестиционную сферу иностранного капитала препятствует несовершенство законодательства, особенно в части защиты инвестиций, недостаточная экологическая безопасность, неразвитость системы страхования инвестиций от рисков.

УДК 292

Пароменков В.О., Огур М.В.

OLAP ТЕХНОЛОГИЯ

БНТУ, Минск

Научный руководитель: Астапчик Н.И.

Трудно найти в компьютерном мире человека, который хотя бы на интуитивном уровне не понимал, что такое базы данных и зачем они нужны. В отличие от традиционных реляционных СУБД, концепция OLAP не так широко известна. Что же такое технология комплексного многомерного анализа данных получившая название OLAP (Online Analytical Processing)?

OLAP – это не отдельно взятый программный продукт, не язык программирования и даже не конкретная технология. Если постараться охватить OLAP во всех его проявлениях, то это совокупность концепций, принципов и требований, лежащих в основе программных продуктов, облегчающих анализам доступ к данным.

Централизация и удобное структурирование – это далеко не все, что нужно аналитику. Ему ведь еще требуется инструмент для просмотра, визуализации информации. Традиционные

отчеты, даже построенные на основе единого хранилища, лишены одного – гибкости. Их нельзя «покрутить», «развернуть» или «свернуть», чтобы получить желаемое представление данных. Инструментом, позволяющим разворачивать и сворачивать данные просто и удобно выступает OLAP.

Концепция OLAP была описана в 1993 году Эдгаром Коддом, известным исследователем баз данных и автором реляционной модели данных. В 1995 году на основе требований, изложенных Коддом, был сформулирован так называемый тест FASMI (Fast Analysis of Shared Multidimensional Information – быстрый анализ разделяемой многомерной информации), включающий следующие требования к приложениям для многомерного анализа:

- Fast (Быстрый) – предоставление пользователю результатов анализа за приемлемое время (обычно не более 5 с), пусть даже ценой менее детального анализа;

- Analysis (Анализ) – возможность осуществления любого логического и статистического анализа, характерного для данного приложения, и его сохранения в доступном для конечного пользователя виде;

- Shared (Разделяемой) – многопользовательский доступ к данным с поддержкой соответствующих механизмов блокировок и средств авторизованного доступа;

- Multidimensional (Многомерной) – многомерное концептуальное представление данных, включая полную поддержку для иерархий и множественных иерархий (это – ключевое требование OLAP);

- Information (Информации) – приложение должно иметь возможность обращаться к любой нужной информации, независимо от ее объема и места хранения.

Конечной целью использования OLAP является анализ данных и представление результатов этого анализа в виде,

удобном для восприятия и принятия решений. Основная идея OLAP заключается в построении многомерных кубов, которые будут доступны для пользовательских запросов. Однако исходные данные для построения OLAP-кубов обычно хранятся в реляционных базах данных. Нередко это специализированные реляционные базы данных, называемые также хранилищами данных (Data Warehouse). В отличие от так называемых оперативных баз данных, с которыми работают приложения, модифицирующие данные, хранилища данных предназначены исключительно для обработки и анализа информации, поэтому проектируются они таким образом, чтобы время выполнения запросов к ним было минимальным. Обычно данные копируются в хранилище из оперативных баз данных согласно определенному расписанию.

Типичная структура хранилища данных существенно отличается от структуры обычной реляционной СУБД. Как правило, эта структура денормализована (это позволяет повысить скорость выполнения запросов), поэтому может допускать избыточность данных.

Хотя OLAP и не представляет собой необходимый атрибут хранилища данных, он все чаще и чаще применяется для анализа накопленных в этом хранилище сведений.

Оперативные данные собираются из различных источников, очищаются, интегрируются и складываются в реляционное хранилище. При этом они уже доступны для анализа при помощи различных средств построения отчетов. Затем данные (полностью или частично) подготавливаются для OLAP-анализа. Они могут быть загружены в специальную БД OLAP или оставлены в реляционном хранилище. Важнейшим его элементом являются метаданные, то есть информация о структуре, размещении и трансформации данных. Благодаря им обеспечивается эффективное взаимодействие различных компонентов хранилища.

Основными составляющими структуры хранилищ данных являются таблица фактов (fact table) и таблицы измерений (dimension tables).

Следует отметить, что OLAP-функциональность может быть реализована различными способами, начиная с простейших средств анализа данных в офисных приложениях и заканчивая распределенными аналитическими системами, основанными на серверных продуктах.

Все, что говорилось выше про OLAP, по сути, относилось к многомерному представлению данных. То, как данные хранятся, грубо говоря, не волнует ни конечного пользователя, ни разработчиков инструмента, которым клиент пользуется.

Многомерность в OLAP-приложениях может быть разделена на три уровня:

- Многомерное представление данных – средства конечного пользователя, обеспечивающие многомерную визуализацию и манипулирование данными; слой многомерного представления абстрагирован от физической структуры данных и воспринимает данные как многомерные.

- Многомерная обработка – средство (язык) формулирования многомерных запросов (традиционный реляционный язык SQL здесь оказывается непригодным) и процессор, умеющий обработать и выполнить такой запрос.

- Многомерное хранение – средства физической организации данных, обеспечивающие эффективное выполнение многомерных запросов.

Первые два уровня в обязательном порядке присутствуют во всех OLAP-средствах. Третий уровень, хотя и является широко распространенным, не обязателен, так как данные для многомерного представления могут извлекаться и из обычных реляционных структур; процессор многомерных запросов в этом случае транслирует многомерные запросы в SQL-запросы, которые выполняются реляционной СУБД.

Конкретные OLAP-продукты, как правило, представляют собой либо средство многомерного представления данных, OLAP-клиент (например, Pivot Tables в Excel 2000 фирмы Microsoft или ProClarity фирмы Knosys), либо многомерную серверную СУБД, OLAP-сервер (например, Oracle Express Server или Microsoft OLAP Services).

Слой многомерной обработки обычно бывает встроен в OLAP-клиент и/или в OLAP-сервер, но может быть выделен в чистом виде, как, например, компонент Pivot Table Service фирмы Microsoft.

Подводя итоги выше изложенного можно отметить, что:

- Назначение хранилищ данных – предоставление пользователям информации для статистического анализа и принятия управленческих решений.

- Хранилища данных должны обеспечивать высокую скорость получения данных, возможность получения и сравнения, так называемых срезов данных, а также непротиворечивость, полноту и достоверность данных.

- OLAP (On-Line Analytical Processing) является ключевым компонентом построения и применения хранилищ данных. Эта технология основана на построении многомерных наборов данных – OLAP-кубов, оси которого содержат параметры, а ячейки – зависящие от них агрегатные данные.

- Приложения с OLAP-функциональностью должны предоставлять пользователю результаты анализа за приемлемое время, осуществлять логический и статистический анализ, поддерживать многопользовательский доступ к данным, осуществлять многомерное концептуальное представление данных и иметь возможность обращаться к любой нужной информации.