**BC/NW 2017 № 1 (30):8.2**

**ВИДЫ СИСТЕМ ИНТЕРАКТИВНОЙ АНАЛИТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ СРАВНИТЕЛЬНЫЙ АНАЛИЗ OLAP-КЛИЕНТОВ И СЕРВЕРОВ**

Тумкина А.И.

Технология аналитической обработки в реальном времени основана на многомерном иерархическом представлении данных. OLAP-инструменты обеспечивают пользователя возможностями гибкого просмотра аналитических сведений в различных срезах.

Отчёты, полученные посредством такого инструмента, хорошо структурированы, обладают гибкостью, то есть поддерживают выполнение операции по переходу от детального представления данных к агрегированному (консолидация) и наоборот (детализация), перемене мест строк и столбцов таблицы (вращение). [1]

Исходная информация представлена в виде многомерного куба с возможностью проведения над ним манипуляций, позволяющих выполнить необходимые срезы. При этом куб представлен пользователю в виде многомерной динамической таблицы, в которой данные в различных срезах суммируются автоматически. За выполнение таких операций отвечает OLAP-машина.

Построение многомерных кубов осуществляется на основе исходных и агрегатных данных. Данные могут храниться и в реляционных, и в многомерных базах данных. Поэтому все OLAP-продукты принято разделять на несколько категорий: [1]

1. многомерную OLAP, являющуюся классической формой OLAP и использующую выделенный сервер, на стороне которого выполняется предварительная обработка данных. Данные формируются в многомерные кубы.

MOLAP-система имеет возможность структурировать данные под любые запросы пользователей, поэтому данная система является самой эффективной при обработке данных.

Поскольку MOLAP системы имеют возможность реорганизовать и структурировать данные под запросы пользователей, они являются довольно эффективными для обработки данных. Кроме того, имеющиеся в подобных системах инструменты позволяют производить сложные расчёты, быстро формировать запросы и предоставлять результаты.

Среди недостатков таких систем выделяют ограничение на объём обрабатываемых данных и их избыточность, так как в ходе формирования многомерных кубов имеется необходимость в дублировании данных.

1. реляционную OLAP - вид системы, работающий с реляционной базой данных. Данные в базе хранятся в виде набора реляционных таблиц, а обращение к данным выполняется напрямую. Благодаря использованию инструментов SQL и запросов, с точки зрения пользователя многомерный анализ выполняется как в традиционной системе-OLAP.

Главным преимуществом реляционной системы является возможность эффективной обработки большого объёма как числовых, так и текстовых данных.

К недостаткам ROLAP относят низкую по сравнению с традиционными системами производительность, возникающую по причине обработки данных сервером OLAP, а также ограничение функциональных возможностей, связанных с использованием языка SQL.

1. гибридную OLAP, которая объединяет в себе возможности реляционной и многомерной систем. За счёт использования многомерной базы данных и управления реляционной базой данные хранятся в реляционных таблицах, а обработанные данные размещаются в заранее выстроенных многомерных кубах. Преимуществом гибридной системы является масштабируемость данных, быстрая их обработка и гибкий доступ к источникам.

Другие виды OLAP систем в большей степени являются маркетинговым ходом производителей, чем самостоятельным видом OLAP системы: [2]

1. WOLAP (Web OLAP). Вид OLAP системы с поддержкой web интерфейса. В этих системах OLAP есть возможность обращаться к базам данных через web интерфейс.
2. DOLAP (Desktop OLAP). Этот вид OLAP системы дает возможность пользователям загрузить на локальное рабочее место базу данных и работать с ней локально.
3. MobileOLAP. Это функция OLAP систем, которая позволяет работать с базой данных удаленно, с использованием мобильных устройств.
4. SOLAP (Spatial OLAP). Этот вид OLAP систем предназначен для обработки пространственных данных. Он появился как результат интеграции географических информационных систем и OLAP системы. Эти системы позволяют обрабатывать данные не только в буквенно-цифровом формате, но и в виде визуальных объектов и векторов.

Многообразие представленных на рынке продуктов аналитической обработки информации позволяет каждому пользователю сделать выбор в пользу того продукта, который будет отвечать всем предъявляемым требованиям.

Каждый OLAP-продукт состоит из сервера и клиентской части. OLAP-сервер является хранилищем информации и позволяет производить выборку данных, основанную на пользовательском запросе. Клиентская часть продукта представлена пользовательским интерфейсом, позволяющим аналитику самостоятельно проектировать многомерные кубы.

В проведенном анализе были рассмотрены наиболее известные и свободно распространяемые OLAP-сервера и OLAP-клиенты. [1]

Критерия, по которым был выполнен анализ нескольких хранилищ данных, основан на требованиях, которые предъявляются к OLAP-серверу, характеристиках, отвечающих за быструю работу сервера и безопасность его среды. Результаты анализа приведены в таблице 1.

Таблица 1 Сравнительный анализ OLAP-серверов

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Критерий/OLAP-сервер** | **MSAS** | **Essbase** | **TM1** | **Mondrian** | **Palo** |
| Работа с PostgreSQL | - | - | + | + | + |
| Форма хранения информации (MOLAP,ROLAP,HOLAP) | (+,+,+) | (+,-,-) | (+,-,-) | (-,+,-) | (+,-,-) |
| Создание пользователем агрегирующей функции | + | + | + | - | - |
| Масштабируемость | + | + | + | + | + |
| User friendly интерфейс | + | + | + | + | + |
| Решение проблемы разреженности данных | + | + | + | + | + |
| Использование языка XMLA и MDX | + | + | + | + | + |
| Русскоязычная документация | + | + | + | - | - |
| Минимальная стоимость | 4000 | 40000 | 50000 | бесплатно | бесплатно |

По результатам проведенного анализа можно отметить, что все рассмотренные программные продукты в полной мере не только отвечают особенностям процесса аналитической обработки данных, но и соответствуют принципам, опубликованным Э.Коддом. Однако большими преимуществами обладает Pentaho Mondrian – бесплатный продукт с имеющейся возможностью его доработки, написанный на кроссплатформенном языке Java и имеющий удобный графический интерфейс для работы с многомерным кубом.

Анализ OLAP-клиентов был проведен на основании предъявляемых к клиентскому приложению требований, среди которых основным является требование по поддержке выполнения операций над многомерным кубом, а также доступность фильтрации данных. Результаты отображены в таблице 2.

Таблица 2 Сравнительный анализ OLAP-клиента

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Критерий/OLAP-клиент** | **MicroStrategy** | **PowerPlay** | **JPalo Pivot** | **Saiku** |
| Выполнение основных операций над многомерным кубом | + | + | + | + |
| Фильтрация данных | + | + | - | - |
| Русскоязычная документация | + | + | - | + |
| Использование языка XMLA и MDX | + | + | + | + |
| Минимальная стоимость | 1000 | 1000 | бесплатно | бесплатно |

На основе полученных результатов видно, что все исследуемые продукты отвечают требованиям, предъявляемым к процессу аналитической обработки данных в реальном времени, изложенные Э.Коддом. Кроме того, все OLAP-клиенты поддерживают работу с XMLA и MDX. [2]

В качестве клиентского приложения фаворитом является продукт Saiku, отличающийся интуитивно понятным интерфейсом и возможностью создания и публикации отчётов с минимальными финансовыми затратами.

Таким образом, технология аналитической обработки упрощает, а вместе с тем и ускоряет процесс принятия решений на основе аналитической информации.

Широкий ассортимент OLAP-инструментов предоставляет каждому пользователю право выбора того продукта, который в наибольшей степени удовлетворяет предъявляемым требованиям, среди которых главным является дружественный интерфейс, точность и скорость выполнения запросов.

**Литература**

1. «Введение в OLAP-технологии Microsoft». Наталия Елманова, Алексей Фёдоров. Издательство Диалог-МИФИ, 2002г.

2. «Архитектура OLAP». Найгель Пендс, перевод Шамиля Абушаева. High Quality Content, 2013г.