

УДК 621.398

С 374

Информационные технологии оперативного анализа данных

Н.П. Сидорова, к.т.н.,

Н.В. Логачева, к.т.н.,

Государственное бюджетное образовательное учреждение
высшего профессионального образования Московской области
«Финансово–технологическая академия», г. Королев, Московская область,

В.Ю. Добродеев, магистр Информатики и вычислительной техники,
Московский энергетический институт, г. Москва

Оперативный анализ данных является широко используемой информационной технологией, которая находит применение в различных сферах деятельности. Для решения задач анализа создаются специальные структуры – центры обработки данных, на базе которых функционируют OLAP-системы. На рынке программных продуктов представлено большое количество средств создания OLAP-систем. В статье проводится обзор возможностей наиболее развитых средств разработки OLAP-систем.

OLAP-система, Microsoft Analysis Services, Oracle OLAP, IBM Cognos.

Information Technology for effective data analysis

N.P. Sidorova, Ph.D,

N.V. Logacheva, Ph.D,

Moscow region state–financed educational institution of higher vocational training
«Finance and technology academy», Korolev, Moscow region,

V.Y. Dobrodeev, master of Computer Science and Engineering
Moscow Power Engineering Institute, Moscow

Operational analysis is a widely used information technology, which finds application in various fields. To solve the problems of the analysis are special structures - data centers, which operate on the basis of OLAP-system. In the software market is represented by a large number of tools for building OLAP-systems. The article gives an overview of the most advanced development tools OLAP-systems.

OLAP-system, Microsoft Analysis Services, Oracle OLAP, IBM Cognos.

Растущий объем данных, возникающий в процессе деятельности, все больше осмысливается менеджментом как важный ресурс для развития бизнеса. Поэтому на первый план выдвигается проблема быстрого и корректного анализа больших объемов данных. Во многих сферах успешно развиваются центры обработки данных (ЦОД). По прогнозам аналитического центра CNews Analytics [5] объем рынка средств управления инфраструктурой ЦОД (рис. 1) к 2015 г. может превысить \$1,27 млрд., а оценка среднегодового роста рынка в период с 2011 по 2015 г. определяется в 30%-40%.

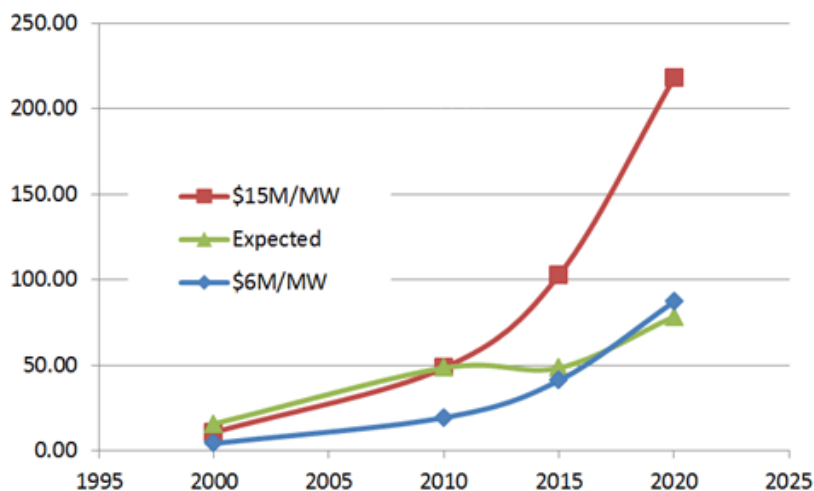


Рисунок 1 – Динамика строительства ЦОД в мире 1995-2025
Источник: Microsoft, 2011

В России по итогам 2013 г. насчитывалось 175 коммерческих ЦОДов (данные iKS-Consulting). Объем этого рынка, по данным iKS-Consulting, составил 4,4 млрд руб. (144 млн долл.). По данным CNews Analytics (рис. 2) в 2011 г. объем рынка составит 178,5 млн долл. при росте в 27%.

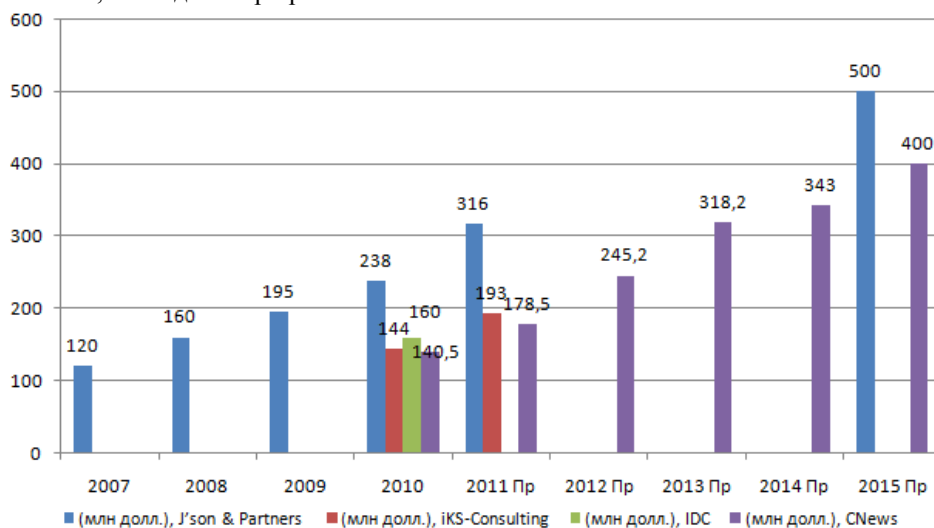


Рисунок 2 – Оценки объема и динамики рынка ЦОД в России по данным разных агентств (млн долл.)

Источник: CNews Analytics, 2011

Данные, поступающие в ЦОД, как правило, являются данными из транзакционных систем, к которым относятся такие системы как ERP, CRM и др. Эффективное принятие решений с использованием таких данных предполагает использование средств их оперативного анализа, которые позволяют целостную модель решения анализируемой проблемы. Анализ данных используется в любой компании для реше-

ния достаточно широкого круга задач: управление продажами, выработка маркетинговых стратегий, формирование корпоративной отчетности и др. В настоящее время использование информационных технологий для анализа данных является важным аспектом управления бизнесом. Для анализа данных применяют различные математические методы: статистические, например, кластерный анализ, который позволяет оценить текущую ситуацию на основе выделения в массиве объектов однородных групп; деревья анализа, которые определяют правила разбиения объектов на группы. Для целей анализа данных в настоящее время широко используют интеллектуальный анализ данных на основе систем Data Mining или Knowledge Discovery in Databases. Они реализуют совокупность методов классификации, прогнозирования и моделирования, общим для которых является использование методов искусственного интеллекта: нейронных сетей, нечеткой логики. Такие методы предназначены для использования в больших базах данных и позволяют отыскивать скрытые закономерности в данных. В данной статье рассматриваются информационные технологии на основе OLAP (online analytical processing), которые в настоящее время расширяют сферы своего применения. Основное преимущество OLAP заключается в его скорости. OLAP-системы позволяют выполнять обработку запросов в несколько раз быстрее, чем реляционные СУБД.

Значение OLAP для пользователей состоит в возможности обеспечить высокую скорость создания отчетов в рамках специализированного хранилища данных, получить данные для выполнения различных видов анализа, разграничить возможности работы для разных групп пользователей в соответствии с политиками безопасности. Основная группа пользователей OLAP-акционеры, руководители предприятий и владельцы бизнес-процессов, руководители подразделений, бизнес-аналитики. OLAP-системы обеспечивают пользователям оперативный доступ к информации, которая представляется в объеме и форме, необходимой менеджеру для принятия решения.

В настоящее время на рынке представлен достаточно широкий выбор инструментов и технологий. Наиболее эффективные из них, по результатам анализа авторитетной аналитической компании Гартнер [6], отражены на магическом квадранте Гартнера (рис. 3).



Рисунок 3 – OLAP-системы

В статье проводится анализ наиболее развитых решений OLAP-систем с точки зрения используемых методов создания моделей данных и средств их обработки.

Microsoft Analysis Services

Microsoft Analysis Services – это часть системы управления базами Microsoft SQL Server, включающая в себя набор средств для работы с OLAP и интеллектуальным анализом данных. Впервые Microsoft включила данную систему в SQL Server 7. Система могла использовать все три разновидности технологии OLAP (MOLAP, ROLAP и HOLAP). В версии продукта Microsoft Analysis Services 2000, включала уже не только собственно OLAP, но и интеллектуальный анализ данных (data mining). Эта система обладала большей гибкостью в проектировании размерности, улучшенным движком вычислений и рядом других преимуществ. Как уже говорилось, система поддерживает три вида хранения многомерной модели данных. При использовании MOLAP обрабатываются сами данные и их агрегаты, которые сохраняются и оптимизируются при помощи специального формата, оптимизированного под многомерные данные. Если применяется ROLAP, то данные и их агрегаты остаются в реляционном источнике данных, благодаря чему не требуется специальной обработки. Если использовать HOLAP, то сами данные будут храниться в реляционном источнике данных, а агрегаты и индексы в специальном формате, оптимизированном для многомерных данных.

Analysis Services [1, 4] поддерживает также наборы программных интерфейсов (API) и объектных моделей. Для извлечения данных могут применяться: ML for Analysis – API нижнего уровня, который может быть использован любой платформой и с любым языком программирования, поддерживающим HTTP и XML; OLEDB for OLAP – расширение OLEDB, основанное на COM и предназначенный для использования C/C++ программах только на Windows-платформе.

Для администрирования системы используются средства: DSO (Decision Support Objects) – для Analysis Services 2000, основанный на COM Automation и предназначенный для программ на Visual Basic для Windows-платформы (реализован в виде набор библиотек, содержащих COM-объекты, которые позволяют создавать и модифицировать многомерные базы данных и содержащиеся в них объекты); AMO – для Analysis Services 2005, основанный на .NET-технологии для программ на платформе .NET

При работе с системой поддерживаются следующие языки запросов:

- 1) Data Definition Language (DDL) – язык определения данных в Analysis Services, основан на XML и поддерживает такие команды как <Create>, <Alter>, <Delete>, <Process> и т.д.
- 2) Data Manipulation Language (DML) – язык манипуляций с данными, включающий в себя следующие разделы: MDX (для реализации запросов OLAP-кубов), SQL (ограниченное подмножество SQL-инструкций для запроса OLAP-кубов), DMX (для запроса моделей интеллектуального анализа данных).

С точки зрения пользователя разницы между способами хранения MOLAP, ROLAP и HOLAP в Analysis Services не существует: обращение к данным при всех этих типах хранения происходит одинаковым образом, аналитические запросы пишутся одинаково. Разница может возникнуть только в производительности при обращении к данным.

Analysis Services используется как компонент корпоративной информационной системы, разработанными Microsoft, например, Microsoft Dynamics AX. В этом случае все необходимые данные предприятия хранятся в рамках корпоративной системы, реализованной на основе реляционной СУБД. Служба OLAP Analysis Services извлекает данные для формирования аналитического отчёта. В этом случае корпоративная система выступает в качестве клиента для OLAP-сервера. Впрочем, благодаря

OLEDB for OLAP клиентами для OLAP-сервера Microsoft могут выступать также и другие Windows-приложения, обладающие поддержкой технологии COM. В любом случае, при использовании Analysis Services эта система выступает сервером в клиент-серверной архитектуре.

Analysis Services строится на основе MS SQL Server, который предусматривает возможность многопользовательского доступа к данным. Соответственно те пользователи, которые работают с SQL Server, могут работать и с Analysis Services. В рамках данной системы реализован графический интерфейс для создания и работы с OLAP-кубами. В Analysis Services 2008 для этих целей используется Business Intelligence Development Studio (BIDS). По сути BIDS представляет собой версию Visual Studio, программного продукта для разработки приложений в операционной системе Windows. Однако в BIDS присутствует ряд дополнительных типов проектов для разработки решений в сфере бизнес-аналитики. Для создания OLAP-куба предназначен проект Analysis Services. При создании проекта имеется возможность определить источники данных, определить измерения, иерархию измерений и другие необходимые свойства создаваемого куба. При этом процесс разработки проекта оформлен в виде мастера. На основе описания куба при создании формируется его визуальное представление. Если пользователь ранее имел опыт работы с приложениями Microsoft (например, программами из пакета Office), то ему будут понятны и функции работы в среде BIDS.

Так как Analysis Services входит в состав Microsoft SQL Server, то и служит для работы именно с этим хранилищем данных. Однако имеется возможность для работы со сторонними источниками данных. Для этого используются приложения от сторонних разработчиков. Условием для работы таких приложений является то, что у Analysis Services имеется возможность создавать запросы OLAP или запросы интеллектуального анализа данных, необходимые решению бизнес-аналитики. Программа, которая используется для соединения и работы с источником данных, называется поставщиком данных. В настоящее время [3] поддерживается работа со следующими источниками данных:

- SQL Server 7.0 с помощью поставщика SQL OLE DB или собственного поставщика данных .NET OLE DB;
- SQL Server 2000 с помощью поставщика SQL OLE DB или собственного поставщика данных .NET OLE DB;
- SQL Server 2005 с помощью поставщика OLE DB для собственного клиента SQL Server или управляемого поставщика SqlClient.NET;
- SQL Server 2008 с помощью поставщика OLE DB для собственного клиента SQL Server или управляемого поставщика SqlClient.NET;
- SQL Server 2008 R2 с помощью поставщика OLE DB для собственного клиента SQL Server или управляемого поставщика SqlClient.NET;
- Параллельное хранилище данных SQL Server 2008 R2 (PDW) с помощью поставщика OLE DB для SQL Server PDW или управляемого поставщика для SQL Server PDW;
- База данных SQL Azure с помощью поставщика OLE DB для собственного клиента SQL Server или управляемого поставщика SqlClient.NET;
- Oracle 9.0 с помощью поставщика Microsoft OLE DB для Oracle или собственного поставщика данных .NET OLE DB;
- IBM DB2 8.1 с помощью поставщика Microsoft OLE DB для DB2 (x86, x64, ia64), доступного только для MicrosoftSQL Server 2005 Enterprise Edition или

MicrosoftSQL Server 2005 Developer Edition и загружаемого как часть пакета дополнительных компонентов для MicrosoftSQL Server с пакетом обновления 1 (SP1);

- Access с помощью поставщика Microsoft Jet 4.0 OLE DB;
- Teradata v2R6 с помощью поставщика OLE DB 1.3 от NCR;
- Управляемый поставщик Teradata.

Таким образом, Analysis Services способна работать с различными источниками данных, основанных на использовании реляционного подхода к их построению.

Модели анализа данных определяют, как будут использоваться данные для решения конкретной задачи. В Analysis Services модели анализа можно строить на основе целого ряда алгоритмов. Например, если планируется применять модель для прогнозирования дискретных значений, то (согласно рекомендации Microsoft) следует применять алгоритмы кластеризации и Байеса. Модель анализа позволяет определить правила группировки фактов, осуществлять прогнозирование.

Результаты обработки в Analysis Services хранятся отдельно от данных. Например, сами данные могут храниться в реляционной структуре, а вот результаты в специальном формате, оптимизированном для многомерных данных. В дальнейшем для доступа к многомерным данным используется специальный язык запросов. Благодаря этому манипуляции с OLAP-кубом отделены от действий с ячейками таблиц данных.

В итоге можно отметить, что в целом Microsoft Analysis Services является полнофункциональной OLAP-системой. Для неё разработчики Microsoft предусмотрели интеграцию с целым рядом своих продуктов. Имеется возможность для работы не только с MS SQL, но и программами из пакета Office, с SQL Azure (то есть с облачными средствами), с корпоративными информационными системами. Это позволяет использовать всю имеющуюся информацию для оперативного анализа. Для построения и визуализации куба используется специальная среда BIDS. Для тех, кто часто работает с продуктами Microsoft (в частности MS Access), этот интерфейс является интуитивно понятным (так как компания придерживается единого стандарта при проектировании интерфейса). При этом для данной аналитической службы существует поставщик данных и для Oracle. Удобство для разработчиков состоит ещё и в том, что имеется возможность работы с кубом через .NET Framework а также в наличии ряда программных интерфейсов. Также существуют наборы объектов для работы с многомерными базами данными. Однако при всех этих достоинствах следует отметить ряд недостатков этой системы.

Во-первых, Analysis Services работает только на компьютерах с операционной системой Windows. В Linux его уже нельзя запускать. Впрочем, хотя эта служба будет запущена на Windows-сервере, возможно реализовать программу для Linux, которая сможет при помощи нужного интерфейса взаимодействовать с этой аналитической службой.

Oracle OLAP

Исторически первой многомерной базой данных считается Express компании IRI, разработанный в 1970 г. Впоследствии права на этот продукт были приобретены компанией Oracle. Эта система под названием OLAP Option стала частью Oracle Database. Следует заметить, что требования к OLAP-системам были выдвинуты Коддом значительно позже. Таким образом, многомерные данные впервые появляются в составе сервера Oracle.

Во многих версиях Oracle Database многомерный движок был реализован в виде Oracle Express. Однако в версии 11g многомерный движок включён непосредственно в систему управления базами данных. Теперь эта система является опцией Oracle Database [2,8]. Она включает в себя набор приложений для планирования, прогнозирования, составления бюджетов. Например, существует достаточно распространённое приложение для финансового анализа Oracle Financial Analyzer.

Для повышения производительности используется специальный тип хранения данных. Дело в том, что OLAP-куб представляет собой достаточно большую структуру (чем больше измерений, тем больше куб). При этом тратится значительное время на чтение его с диска. Благодаря сжатым композитам информация хранится в блоках. При запросе нужного блока он передаётся в оперативную память, где затем распаковывается. Это значительно более быстрый процесс, чем чтение несжатого блока с диска. Однако у этого метода есть и свои недостатки. Например, если требуется обновить какую-либо ячейку памяти, то придётся сначала распаковать блок, затем изменить данные и запаковать обратно. При этом размер блока может измениться и записать его в то же место на диске, где он находился до этого, уже не получится. Ещё один важный момент состоит в том, что все агрегаты на диске хранить практически невозможно, потому что они бы занимали большое количество памяти. Часть агрегатов конечно можно рассчитывать, но на это уже тратятся вычислительные ресурсы. В Oracle OLAP существует такое понятие как карта агрегации, где описано, какие агрегаты нужно вычислять, а затем хранить на диске, а какие вычислять непосредственно при запросе к кубу. В версии 11g появился новый тип расчёта агрегатов: Cost-Based Aggregation. Эта система рассчитывает наиболее сложные агрегаты и хранит их в кубе. При этом есть возможность определять процент хранимых агрегатов. По умолчанию это 20 – 30% агрегатов. Если установить это значение равным 100%, то все агрегаты будут рассчитываться заранее и сохраняться на диск. Все эти средства позволяют ускорить обработку данных.

По сравнению с предыдущими версиями упрощена работа с кубом через SQL. Если раньше требовалось проделать достаточно много дополнительной работы. Теперь механизм стал более понятным. Запрос направляется непосредственно к движку OLAP, имеется возможность объединять в результатах выполнения запроса данные из куба с обычными реляционными данными. Это в частности позволяет добавлять к агрегатам детальные сведения. Для этих целей служит конструкция CUBE_TABLE, которая эффективно работает в связке с SQL. Впрочем, если нужно добавить свой объект для вычислений, он не будет доступен при использовании SQL.

Перечень измерений находится в таблице, что позволяет добавить их достаточно большое количество. Если требуется добавить измерение, то запись о нём делается в таблицу. Затем уже происходит обработка куба для добавления в него нового измерения.

Сам OLAP-куб хранится в базе данных в виде BLOB (Binary Large Object) – больших двоичных объектах. Это в частности даёт преимущества при резервном копировании, восстановлении и безопасностью.

Таким образом, для Oracle OLAP можно выделить следующие возможности:

- опция OLAP встроена в ядро Oracle Database и выполняется в той же службе, что и реляционная база данных, так что нет необходимости в управлении отдельной, дополнительной службой базы данных;
- OLAP-кубы и измерения безопасно и надёжно хранятся в файлах данных Oracle поэтому нет необходимости в управлении отдельных файлов данных;

- благодаря функциональным возможностям Oracle Database опция OLAP максимально эффективно использует крупномасштабные системы компьютерного аппаратного обеспечения;
- безопасность OLAP-кубов также гарантируется функциями безопасности Oracle Database;
- простой доступ к OLAP-кубам предоставляется через SQL: отправлять запросы в один и тот же куб с легкостью могут как многомерные, так и реляционные приложения.

То есть свойство прозрачности в целом поддерживается. Механизмы работы с многомерными данными во многом похожи на работу с обычной SQL базой данных. Это же позволяет сделать работу с данными интуитивно понятными. Так как многомерный движок встроен непосредственно в СУБД, то все возможности этой системы применимы и к OLAP. Это в первую очередь относится к многопользовательской работе.

В рамках клиент-серверной архитектуры OLAP выступает в качестве сервера. Для разработки приложений, которые будут обращаться к многомерному движку, существует API на языке Java (Java OLAP API). Для создания представления (VIEW) куба используется Analytic Workspace Manager (AWM). Работа с разреженными данными поддерживается за счёт механизма секционирования. При этом разные части одной таблицы ведут себя как таблицы меньшего размера со своими собственными локальными индексами. При запросе к основной таблице оптимизатор определяет, к каким секциям нужно обращаться при выборке. Также в Oracle представлен ряд функций для выполнения различных аналитических вычислений (например, ранжирования).

Oracle OLAP в целом отвечает требованиям, предъявляемым к OLAP-системам, и благодаря включению в систему Business Intelligence достаточно широко используется при решении задач бизнес-анализа.

IBM Cognos

IBM Cognos TM1 – это OLAP система от компании IBM. Кубы, которые в этой системе являются самостоятельными и не основаны на базовой таблице. Вычисления производятся автоматически и с помощью специального набора правил, для описания которых существует специальный язык. Это позволяет осуществлять достаточно сложные расчёты. TM1 интегрирована в Microsoft Excel. При этом имеется веб-интерфейс, разработанный на ASP.NET, позволяющий проводить анализ данных в табличном или графическом виде. TM1 используется для автоматизации процессов бюджетирования, планирования, построения отчётов и анализа данных, составления прогнозов.

TM1 представляет собой MOLAP-систему [3, 7]. При этом рассчитана на пользователей без навыков программирования. Может работать в режиме реального времени, то есть данные передаются в куб, как только их вводит пользователь. TM1 позволяет проводить анализ данных и на основе этого анализа моделировать бизнес-процессы организации и составлять прогнозы.

Многомерный 64-разрядный механизм данной OLAP-службы обеспечивает высокую производительность при работе со сложными моделями, большими наборами данных и потоковой информацией. Имеется возможность производить по требованию финансовый анализ, гибкое моделирование и участие в процессе планирования всех подразделений предприятия. Для отдельных пользователей, рабочих групп и компаний можно создавать отдельные сценарии аналитики. Это увеличивает гибкость за счёт более быстрого реагирования на изменение условий.

Из методов планирования и прогнозирования представлены, например, факторное планирование, составление скользящих прогнозов. При проектировании моделей и доступе к данным учитываются особенности бизнес-процесса, получаемая информация представлена в распространённых форматах. Решения для планирования быстро компонуется и внедряются. Данные из систем и от персонала в различных подразделениях собираются автоматически и согласованно на основе этих решений.

Для построения отчётности в системе Cognos 8 BI реализованы следующие инструменты: Query Studio и Report Studio. Query Studio - конструктор отчетов для начинающих пользователей, позволяющий строить сводные отчеты с помощью визуальных методов. Полученный на его помощью отчет можно сохранить для последующего использования, распечатать, добавить форматирование и графики. С помощью этого инструмента можно формировать сводные таблицы и плоские отчеты, не прибегая к помощи ИТ-специалистов. В создаваемых отчётах представлена картина на всех этапах: от постановки целей и планирования до измерения показателей выполнения.

Report Studio - инструмент создания профессиональных многостраничных отчетов с произвольной структурой представления информации, который позволяет комбинировать данные из различных источников.

Cognos 8 BI [7] поддерживает широкий спектр источников данных, а именно:

- реляционные (IBM, Oracle, SQL, Teradata, Sybase, ODBC и т. д.);
- на основе Интернет-технологий (XML, JDBC, LDAP, WSDL);
- хранилища данных OLAP (Cognos PowerCubes, Cognos Planning Contributor Cubes и Cognos Finance Cubes, Microsoft OLAP, IBM DB2/OLAP, Hyperion Essbase, SAP Business Warehouse (BW), ROLAP).

Расширенная возможность сопоставления данных из разных источников для одного отчета - одно из наиболее важных нововведений Cognos. Теперь разработчики отчетов могут комбинировать и связывать данные из произвольного набора источников (реляционная база данных, XML, Excel, OLAP и другие источники) в едином представлении, доступном для пользователей. Такая связь может существовать как в режиме реального времени, так и на основе создания хранилищ или витрин данных с помощью предлагаемых механизмов извлечения и преобразования информации (ETL-средства).

В интегрированной системе Cognos 8 реализован компонент Analysis Studio, предназначенный для анализа данных, сформированных с применением OLAP-технологий. Данный компонент расширяет возможности Cognos TM1 за счет следующих возможностей:

- поддержка прямого доступа к большому количеству OLAP-платформ: Cognos PowerCubes, Cognos Planning Contributor Cubes and Cognos Finance Cubes, Microsoft OLAP, IBM DB2/OLAP, Hyperion Essbase;
- полная интеграция с другими инструментами Cognos 8 BI (возможность доработки и использования в Report Studio отчетов, полученных в Analysis Studio, передача данных и управления в другие компоненты системы, внесение подготовленных отчетов в персональные страницы);
- поддержка ROLAP-моделей;
- расширенные возможности работы с измерениями (поддержка условного отображения содержимого измерений по различным критериям, например, независимые Top 10 позиций в каждой подгруппе для многоуровневых отчетов);
- фильтрация и сортировка данных и измерений по условиям «И» и «ИЛИ»;

- создание, сохранение и повторное использование настраиваемых подмножеств, позволяющих комбинировать данные из произвольных измерений, лежащих на различных уровнях иерархии, - это дает возможность проводить анализ данных, не ограничиваясь заданной структурой измерений;
- улучшенный экспорт отчетов в Excel – поддержка форматирования данных и графического представления информации.

Результаты сравнения рассмотренных OLAP-систем приведены в таблице 1.

Таблица 1 – Результаты сравнения рассмотренных OLAP-систем

Характеристика	Microsoft Analysis Services	Oracle OLAP	IBM Cognos
Методы реализации OLAP-куба	MOLAP, ROLAP, HOLAP	MOLAP, ROLAP	MOLAP, ROLAP
Поддержка различных платформ	нет	есть	есть
Наличие средств администрирования	есть	есть	есть
Поддержка различных источников данных	есть	есть	есть

В заключение хотелось бы отметить, что кроме рассмотренных имеется целый ряд других систем как коммерческих, так и свободно распространяемых, например, OLAP-продукты компаний SAP, SAS и др. Впрочем, использование OLAP имеет свои недостатки. В первую очередь это размер самого куба, что вызывает большие объемы обрабатываемой памяти и необходимость применения специальных средств оптимизации выполнения аналитических запросов. Чем больше измерений, тем больше агрегатов. Существует два варианта работы для OLAP-систем: это хранение агрегатов в памяти (на диске) или вычисление агрегатов по запросу. Первый вариант является затратным с точки зрения использования дискового пространства. Второй вариант является более медленным, так как процесс чтения информации с диска всё-таки является достаточно длительным (по сравнению с чтением информации из оперативной памяти). При этом вовсе не обязательно, что данные, необходимые для вычисления агрегата будут находиться рядом. Стоит отметить также и тот факт что, как правило, куб строится на основе баз данных той СУБД, для которой разрабатывалась OLAP-служба. Однако может возникнуть ситуация, когда куб потребуется строить на основе группы баз данных, которые к тому же могут быть созданы в разных СУБД. При этом эти базы данных могут находиться на удалённых компьютерах. То есть OLAP система в этом случае должна иметь возможность работать с другими СУБД и распределёнными системами. Достаточно близко к решению подобной задачи подходят разработчики фирмы Microsoft Analysis Services за счёт того что для других СУБД имеются поставщики данных, IBM Cognos.

Однако, хотя Microsoft и является крупнейшим производителем программного обеспечения, на рынке СУБД лидерство остаётся за Oracle. Это объясняется тем, что Oracle лучше работает на кластерных системах, более производительная и масштабируемая. В крупных распределённых проектах всё же чаще применяется именно эта СУБД. Для Oracle Database имеется поставщик данных, позволяющий обращаться к ней из Analysis Services. Для Oracle OLAP такая возможность не указана. Тем не менее, возможна интеграция продуктов Oracle с многомерной СУБД Essbase. Ранее Essbase разрабатывалась компанией Hyperion. Однако в настоящее время (так как

Нурегон является частью Oracle) эта СУБД интегрируется со всеми продуктами этой компании.

Что касается развития подобных систем, то аналитики выделяют ряд направлений. Во-первых, это доступ к OLAP мобильных устройств. В принципе это возможно, так как все вычисления происходят на сервере, а клиенту передаётся только результат. Также OLAP-системы должны будут иметь возможность работать с потоковыми данными (присутствует, например, в TM1), в режиме реального времени (Oracle OLAP). Ещё одной тенденцией является персонализация работы с данными в рамках корпоративной аналитической системы. То есть пользователь работает со своими локальными данными, но у него есть возможность работы и с данными из хранилища корпоративной информации. Также интересным направлением представляется интеграция OLAP и социальных сетей. Это может открыть большие возможности для маркетинговой аналитики.

Развивается и такое направление как in-memory OLAP. Идея состоит в том, что в оперативной памяти компьютера расположена промежуточная база данных, которая считывает нужную информацию с диска, а все построения и вычисления выполняет в ОЗУ. Это действительно может позволить ускорить работу за счёт уменьшения количества обращений к диску. Такая технология в частности реализована в TM1. В связи с тем, что в современных компьютерах заметно растёт количество оперативной памяти, такой подход имеет перспективы.

В связи с возросшим интересом бизнеса к использованию таких систем, в настоящее время появилось ещё одно направление, связанное с анализом данных – Big Data (большие данные). Однако, на взгляд авторов, это направление в большей степени затрагивает технические и технологические аспекты хранения и скорости обработки больших объёмов данных, не внося пока новых методов их обработки.

В общем, нужно отметить, что существует тенденция к увеличению использования OLAP. Это объясняется тем, что всё большее число систем предлагают эффективное решение для задач аналитики, особенно аналитики финансовой. Автоматизированные системы позволяют учитывать огромные объёмы данных, которые сложно обрабатывать вручную. Программные комплексы, которые предназначены для этих целей, становятся всё сложнее. Они интегрируются с корпоративными информационными системами, применяются в бизнес-планировании и управлении. Эти сложные комплексы идут по пути повышения производительности. Проблему хранения больших объёмов OLAP-куба решают балансом между вычисляемыми и хранимыми агрегатами. Возможно, определённые результаты может дать совмещение OLAP с искусственным интеллектом (а OLAP всё же является одним из методов Data Mining). В развитии OLAP существует множество тенденций и направлений. Со временем станет ясно, какие из них позволят значительно упростить и ускорить процесс анализа информации. Ведь не проанализировав должным образом поступающую информацию, сложно принимать качественные решения.

Литература

1. Microsoft SQL Server Analysis Services 2008 и MDX для профессионалов. М.: Вильямс, 2010. – 1072 стр.
2. Армстронг-Смит, М. Oracle Discoverer. М.: Лори, 2013. – 483 с.
3. Бергер, А. Б. OLAP и многомерный анализ данных. СПб.: BHV-СПб. – 2007. – 928 с.
4. Санджай, Сони, Уэйн, Курц. Повышение производительности куба, используя Microsoft SQL Server 2000 Analysis Services. М.: Intersoft Lab, 2008 г. – 426 с.
5. Электронный ресурс. Режим доступа: <http://www.cnews.ru/reviews/free/datacenter/> (Дата обращения 10 мая 2014 г.).

6. Электронный ресурс. Режим доступа: <http://www.computerra.ru/95074/bi-magic-quadrant-2014/> (Дата обращения 7 мая 2014 г.).
7. Электронный ресурс. Режим доступа <http://www-03.ibm.com/software/products/ru/business-intelligencem/> (Дата обращения 12 мая 2014 г.).
8. Электронный ресурс. Режим доступа: <http://www.oracle.com/technetwork/database/options/olap/index.html> (Дата обращения 20 мая 2014 г.).