

# Хранилища данных и аналитические системы

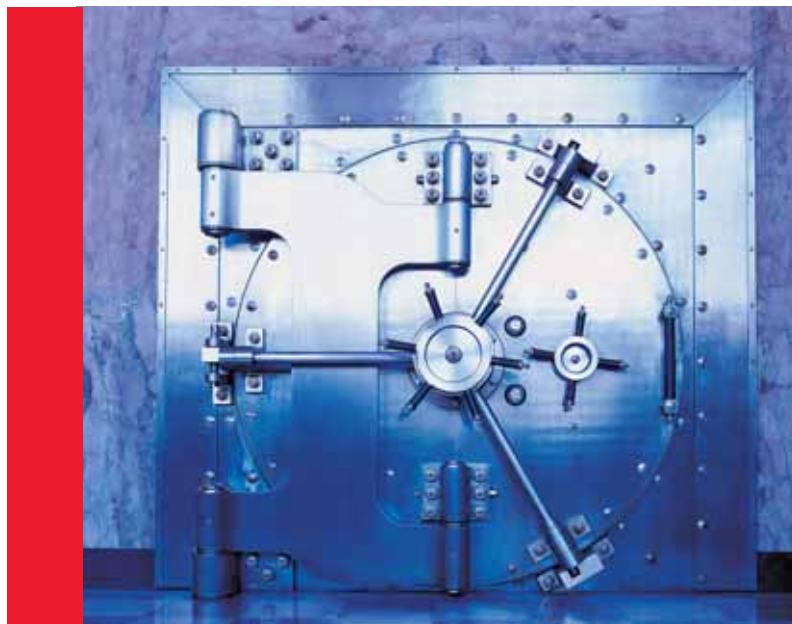
Технологии и инструментальные средства  
корпорации Oracle





**ORACLE IS THE INFORMATION COMPANY**

<b>Введение</b>	<b>2</b>
<b>Архитектура и технология функционирования</b>	<b>3</b>
Общие сведения	3
Извлечение, преобразование и загрузка данных	3
Хранение данных	4
Анализ данных	4
<b>Инструментальные средства</b>	<b>5</b>
Продукты Oracle для хранилищ данных и аналитических систем	5
Oracle Database для реализации хранилища данных	6
Сбор данных из разнородных источников	7
Oracle Warehouse Builder – инструментальная среда создания хранилища данных	9
Бизнес–анализ данных	10
Oracle Business Intelligence Standard Edition	10
Интегрированная среда Oracle Business Intelligence Suite Enterprise Edition	13
<b>Методология создания хранилищ данных</b>	<b>14</b>
<b>Заключение</b>	<b>16</b>



## Введение

К настоящему времени во многих организациях накоплены колоссальные объемы данных, на основе которых можно решать самые разнообразные аналитические и управленческие задачи в любой сфере деятельности. Проблемы хранения и обработки аналитической информации становятся все более актуальными и привлекают внимание специалистов и фирм, работающих в области информационных технологий.

В идеале работа аналитиков и руководителей различных уровней должна быть организована так, чтобы они могли – иметь доступ ко всей интересующей их информации – пользоваться удобными и простыми средствами представления и работы с этой информацией.

Именно на достижение этих целей и направлены информационные технологии, объединяющиеся под общим названием хранилища данных и бизнес-анализа.

Для предоставления необходимой для принятия решений информации обычно приходится собирать данные из нескольких транзакционных баз данных различной структуры и содержания. Основная проблема при этом состоит в несогласованности и противоречивости этих баз-источников, отсутствии единого логического взгляда на корпоративные данные. Решением этой проблемы является **хранилище данных**. В основе концепции хранилищ данных лежит важная идея интеграции ранее разьединенных детализированных данных, содержащихся в исторических архивах, накапливаемых в традиционных системах транзакционной обработки, поступающих из внешних источников, в единую базу данных, их предварительное согласование и, возможно, агрегация.

Автор концепции хранилищ данных (Data Warehouse) является Б.Инмон, который определил хранилища данных, как: *«предметно ориентированные, интегрированные, неизменяемые, поддерживающие хронологию наборы данных, организованные для целей поддержки управления»*, призванные выступать в роли *«единого и единственного источника истины»*, обеспечивающего руководителей и аналитиков достоверной информацией, необходимой для оперативного анализа и принятия решений.

Кроме возможности работать с единым источником информации, руководители и аналитики должны иметь удобные средства визуализации данных, агрегирования, поиска тенденций, прогнозирования. Несмотря на многообразие аналитической деятельности можно выделить типовые технологии анализа данных, каждой из которых соответствует определенный набор инструментальных средств. Вместе с хранилищем данных эти средства обеспечивают полное решение для автоматизации аналитической деятельности и создания корпоративной информационно-аналитической системы.

Направление хранилищ данных и аналитических систем уже несколько лет является для Oracle одним из самых приоритетных. Ориентируясь на общепринятые стандарты в этой области, Oracle предлагает новые решения и технологии для построения эффективных информационно-аналитических систем.

Полномасштабная информационно-аналитическая система должна выполнять сложные и разнообразные функции, включающие сбор данных из различных источников, их согласование, преобразование и загрузку в хранилище, хранение аналитической информации, регламентную отчетность, поддержку произвольных запросов, многомерный анализ и др. Обычно для выполнения этих функций используются различные продукты, что приводит к усложненной архитектуре системы, необходимости интегрировать разнородные инструментальные среды, дополнительным затратам на администрирование, проблемам согласования данных и метаданных на различных серверах.

Для решения этих проблем корпорация Oracle предлагает новый подход к созданию аналитических систем – единую и функционально полную платформу для решения всех перечисленных задач. Основой решения является система управления базами данных Oracle Database, с помощью которой можно не только надежно хранить огромные объемы аналитической информации, но и эффективно выполнять процедуры извлечения данных из разнородных источников, согласовывать, агрегировать и преобразовывать эти данные в аналитическую информацию, загружать ее в хранилище. Кроме того, средствами этого же продукта поддерживаются различные методы анализа данных, включающие многомерный анализ, прогнозирование, автоматический поиск закономерностей.

## Архитектура и технология функционирования

### Общие сведения

В настоящее время существуют фактические стандарты построения корпоративных информационно-аналитических систем, основанных на концепции хранилища. Эти стандарты опираются на современные исследования и общемировую практику создания хранилищ данных и аналитических систем.

В общем виде архитектура корпоративной информационно-аналитической системы описывается схемой с тремя выделенными слоями (рис.1):

- Извлечение, преобразование и загрузка данных
- Хранение данных
- Анализ данных (рабочие места пользователей)

Технология функционирования системы состоит в следующем. Данные поступают из различных внутренних транзакционных систем, от подчиненных структур, от внешних организаций в соответствии с установленным регламентом, формами и макетами отчетности. Вся эта информация проверяется, согласуется, преобразуется и помещается в хранилище и витрины данных. После этого пользователи с помощью специализированных инструментальных средств получают необходимую им информацию для построения различных табличных и графических представлений, прогнозирования, моделирования и выполнения других аналитических задач.

Рассмотрим подробнее каждый из слоев.

### Извлечение, преобразование и загрузка данных

В качестве источников информации для хранилища могут использоваться базы данных внутренних транзакционных систем, информационные системы подчиненных организаций, данные, поступающие из внешних организаций.

С организационной точки зрения, данный слой включает подразделения и структуры организации всех уровней, поддерживающие базы данных оперативного доступа. Он представляет собой низовой уровень генерации информации, уровень внутренних и внешних информационных источников, вырабатывающих “сырую” информацию. Эта информация является рабочей для повседневной деятельности различных подразделений, которые ее вырабатывают и используют.

С системно-технической точки зрения данный слой представлен ЛВС всех подразделений всех уровней, к которым подключены специализированные технические комплексы, хранящие информацию. В качестве таких технических комплексов могут выступать, во-первых, серверы реляционных (SQL-ориентированных) баз данных на базе компьютеров под управлением Windows NT, Unix и др. Во-вторых, это могут быть файловые серверы, на которых установлена какая-либо система обработки данных (например, Btrieve) или сетевая версия СУБД класса персональных (например, Paradox, FoxPRO и т.д.). В-третьих, это могут быть персональные компьютеры с локальными персональными базами данных или файлами.

Из источников данных информация перемещается на основе некоторого регламента в централизованное хранилище. Как правило, необходимые для хранилища данные не хранятся в окончательном виде ни в одной из

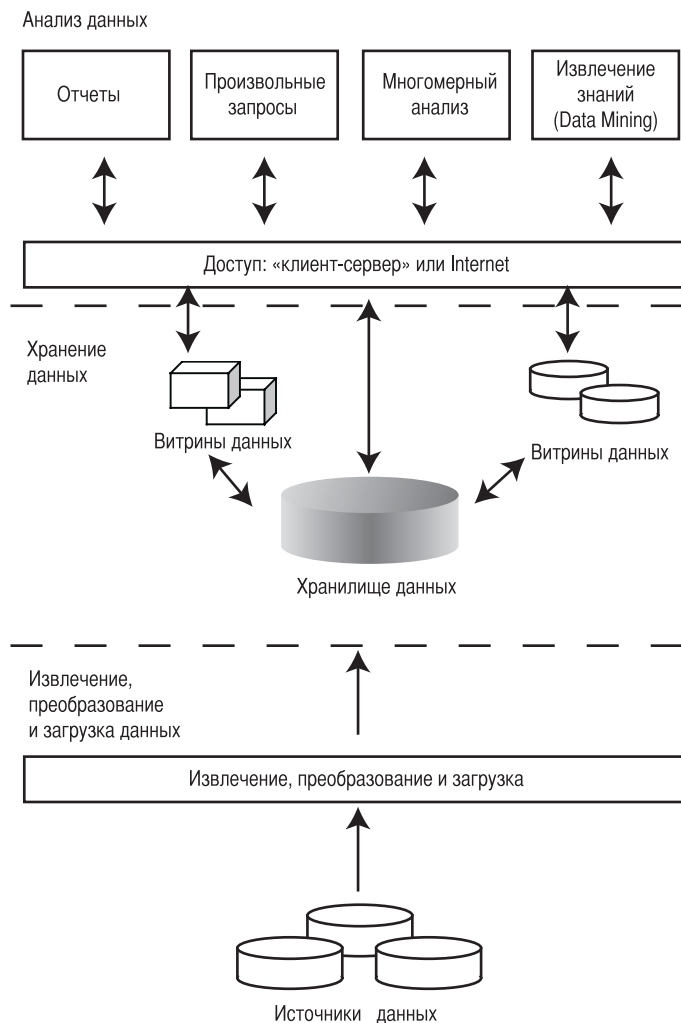


Рис.1 Архитектура корпоративной информационно-аналитической системы

транзакционных систем. Эти данные обычно можно получить из исходных баз данных путем специальных преобразований, вычислений и агрегирования.

Кроме того, несмотря на различную функциональную направленность исходные транзакционные системы часто «пересекаются» по данным, т.е. их локальные базы данных содержат однотипную по смыслу информацию. Это прежде всего касается нормативно-справочной информации, которая используется в том или ином виде в любой оперативной системе. При этом существенно, что одинаковые по смыслу данные обычно имеют в разных системах различный формат, вид представления, идентификацию, единицы измерения и т.п. Перед загрузкой в хранилище вся эта информация должна быть согласована, чтобы обеспечить целостность и непротиворечивость аналитических данных.

Согласование данных необходимо и при загрузке данных из одного источника. Дело в том, что в хранилище хранятся исторические данные, т.е. данные за достаточно большой промежуток времени. В оперативной системе данные хранятся в целостном виде за ограниченный промежуток, после чего они отправляются в архив. При изменениях в структуре или собственно данных архивы не подвергаются никакой дополнительной обработке, а хранятся в исходном виде. Следовательно, при необходимости иметь данные за достаточно большой период времени необходимо согласовывать архивную информацию с текущей.

Таким образом, загрузка данных из источников в хранилище осуществляется специальными процедурами, позволяющими

- извлекать данные из различных баз данных, текстовых файлов;
- выполнять различные типы согласования и очистки данных;
- преобразовывать данные при перемещении их от источников к хранилищу;
- загружать согласованные и «очищенные» данные в структуры хранилища

Для разработки, поддержки и выполнения таких процедур рекомендуется использовать специализированный инструментарий, предназначенный для автоматизации процессов извлечения данных из источников, их преобразования и загрузки в целевое хранилище. Такой инструментарий должен обеспечивать автоматическое формирование процедур загрузки на основе декларативной информации об источниках, правилах согласования и преобразования. Вся эта информация вводится администратором перемещения данных и хранится в виде метаданных в репозитории системы.

Извлечение, преобразование и загрузка данных должна осуществляться как непосредственно вызовом соответствующих процедур, так и в автоматическом режиме, на основе скриптов и расписаний, составленных на этапе разработки системы.

#### **Хранение данных**

Второй слой предназначен непосредственно для хранения значимой, проверенной, согласованной, непротиворечивой и хронологически целостной информации, которую с достаточно высокой степенью уверенности можно считать достоверной.

Собственно хранилище данных не ориентировано на решение какой-либо определенной функциональной аналитической задачи. Цель хранилища – обеспечить целостность и поддерживать хронологию всевозможных корпоративных данных, и с этой точки зрения оно нейтрально по отношению к приложениям. В связи с этим в большинстве случаев для выполнения определенного комплекса функционально замкнутых аналитических задач рационально создавать витрины данных, в основе которых может быть как многомерная, так и реляционная модель данных. По существу витрина представляет собой относительно небольшое, но что самое важное, функционально-ориентированное хранилище, в котором информация хранится специальным образом, оптимизированным с точки зрения решения конкретных аналитических задач некоторого подразделения или группы аналитиков.

Обычно информация попадает в витрины из хранилища и в этом случае витрины называются зависимыми. Возможна также ситуация, когда источником информации для пополнения витрин служат непосредственно оперативные и внешние транзакционные системы. Такие витрины, получившие название независимых, как правило, рассматриваются как временное решение, позволяющее достаточно быстро и с небольшими затратами решить наиболее важные задачи, оценить преимущества нового подхода, сформулировать некоторые рекомендации для более масштабного проекта разработки общего хранилища.

Хранилище реализуется в виде реляционной базы данных, работающей под управлением достаточно мощной реляционной СУБД. Такая СУБД должна поддерживать эффективную работу с терабайтными объемами информации, иметь развитые средства ограничения доступа, обеспечивать повышенный уровень надежности и секретности, соответствовать необходимым требованиям по восстановлению и архивации и т.п.

Витрины данных могут строиться на основе как реляционной, так и многомерной технологии баз данных. Обычно для достаточно большой части аналитических приложений оказывается удобной и эффективной технология интерактивного многомерного анализа и в этом случае витрина представляет собой многомерную базу данных, реализованную в архитектуре OLAP, ROLAP или HOLAP.

#### **Анализ данных**

Для организации доступа аналитиков к данным хранилища и витрин используются специализированные рабочие места, поддерживающие необходимые технологии как оперативного, так и долговременного анализа. Результаты работы аналитиков оформляются в виде отчетов, графиков, рекомендаций и сохраняются как на локальном компьютере, так и в общедоступном узле локальной сети.

Аналитическая деятельность в рамках корпорации достаточно разнообразна и определяется характером решаемых задач, организационными особенностями компании, уровнем и степенью подготовленности аналитиков. В связи с этим современный подход к инструментальным средствам анализа не ограничивается использованием какой-то одной технологии. В настоящее время принято различать четыре основных вида аналитической деятельности (рис. 1): стандартная отчетность, нерегламентированные запросы, многомерный анализ (OLAP) и извлечение знаний (data mining)

Каждая из этих технологий имеет свои особенности, определенный набор типовых задач и должна поддерживаться специализированной инструментальной средой.



## Инструментальные средства

### Продукты Oracle для хранилищ данных и аналитических систем

Инструментальные средства корпорации Oracle обеспечивают полное интегрированное решение для создания хранилищ данных и эффективного использования накопленной в нем информации.

Общий перечень продуктов Oracle, необходимых для реализации технологии хранилищ данных и аналитических приложений, приводится в таблице 1 в соответствии с выделенными в предыдущем разделе компо-

нентами – извлечение и загрузка данных, хранение, анализ (рис.2).

В качестве среды хранения информации в реляционных хранилищах и витринах данных используется сервер Oracle Database. Центральным инструментальным средством создания хранилищ и витрин является Oracle Warehouse Builder, построенный на базе современной архитектуры Common Warehouse Metadata. Он предназначен для описания структуры хранилища и витрин, проектирования и создания процедур извлечения, согласования и загрузки данных, а также генерации метаданных для средств доступа, например таких, как Discoverer.

Таблица 1

Тип средств	Продукт	Комментарий
Извлечение, преобразование и загрузка	Oracle Warehouse Builder	Поддержка процессов извлечения, преобразования и загрузки данных в хранилище
	ETL-средства Oracle Database	
	Oracle Workflow	
Хранение данных	Oracle Database	СУБД для хранилища данных и реляционных витрин данных
	Oracle OLAP Option	Опция СУБД для многомерных витрин данных
Анализ данных Oracle Business Intelligence Standard Edition	Oracle Reports	Регламентированная отчетность
	Oracle Discoverer	Произвольные запросы
	Oracle Discoverer for OLAP OLAP Spreadsheet Jdeveloper + BI Beans	Многомерный анализ
	Oracle Data Mining Option + Oracle Data Miner	Опция СУБД для Извлечения знаний (data mining) и графический интерфейс
Анализ данных Oracle Business Intelligence Suite Enterprise Edition	Reporting and Publishing, Reporting Workbench	Регламентированная отчетность
	Answers	Произвольные запросы
	Interactive Dashboard	Интерактивные информационные панели (dashboards)
	Delivers	Уведомления в реальном времени
	Office Plug-In	Интеграция с MS Office
	Disconnected Analytics	Анализ данных в автономном (off-line) режиме
Business Intelligence Server	Сервер бизнес-анализа	



Проектировать хранилище можно и с помощью стандартного инструмента Oracle Designer, а затем автоматически перенести описание проекта в репозиторий метаданных Oracle Warehouse Builder.

Средства анализа данных охватывают весь спектр аналитических задач и поставляются в виде двух редакций Oracle Business Intelligence Standard Edition и Oracle Business Intelligence Enterprise Edition. Стандартная редакция ориентирована на небольшие и средние организации. В этом случае для стандартной отчетности используется Reports, для генерации нерегламентированных отчетов и запросов — Discoverer, для сложного многомерного анализа — Discoverer OLAP, Spreadsheet Add-In, BI Beans и Jdeveloper, а для задач “извлечения знаний” — Oracle Data Mining. Для крупных корпоративных систем используется Oracle Business Intelligence Suite Enterprise Edition – интегрированная платформа для бизнес-анализа, включающая, кроме стандартных средств отчетности и нерегламентированных запросов, инструменты построения интерактивных информационных панелей, решения для анализа данных в офлайн режиме, средства уведомления и алертов в реальном режиме времени и др. Кроме того, существуют готовые приложения для решения специализированных задач — Enterprise Planning and Budgeting (бюджетирование и планирование), Balanced Scorecard (сбалансированная система показателей), Demand Planning, Value Based Management и другие.

Важнейшей чертой аналитических инструментальных средств и приложений Oracle является их готовность к работе в среде Internet. Менеджеры и аналитики, где бы они ни находились, могут получать информацию из хранилищ и витрин данных в защищенной Интранет-архитектуре с помощью сервера приложений Oracle Application Server.

#### Oracle Database для реализации хранилища данных

Основой хранения аналитической информации является хранилище данных, которое представляет собой базу данных, содержащую достоверную согласованную информацию, предназначенную для решения разнообразных аналитических задач. С точки зрения СУБД, под управлением которой работает хранилище данных, наиболее существенным является тот факт, что

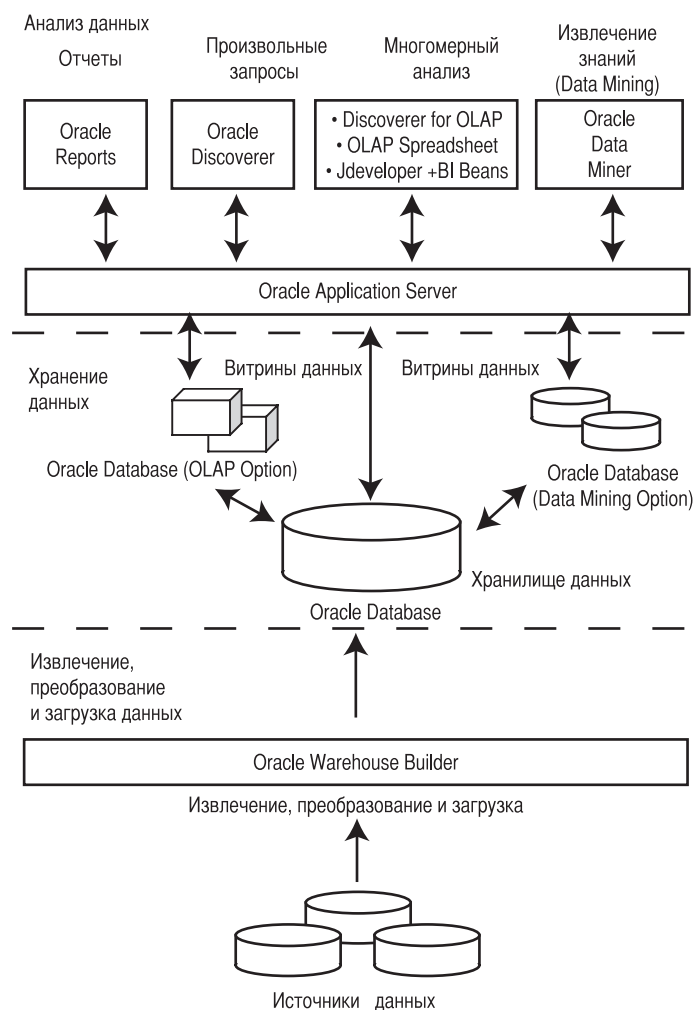


Рис.2 Инструментальные средства Oracle для построения хранилищ данных и аналитических систем

режимы функционирования базы данных для аналитических задач коренным образом отличаются от ситуации в обычных системах транзакционной обработки. Они требуют специальных настроек параметров, методов индексирования и обработки запросов.

СУБД Oracle предоставляет широкий спектр средств, направленных на работу базы в режиме хранилищ и витрин данных. К их числу относятся параллельная обработка запросов, позволяющая наиболее полно использовать возможности многопроцессорных аппаратных платформ, эффективные битовые (bitmap) индексы и специализированные алгоритмы выполнения запросов, которые многократно повышают производительность обработки аналитических запросов, секционирование данных (partitioning), облегчающее управление и значительно ускоряющее обработку очень больших таблиц и индексов и многое другое.

**Битовые индексы.** Использование битовых индексов (Bitmap Index) позволяет резко ускорить выполнение определенных типов запросов, характерных для аналитических приложений. В случае если разброс значений какого-либо параметра или группы параметров небольшой (например: пол - мужской/женский; семейное положение - женат (замужем)/разведен(а)/вдов(а)ец, тип клиента - физ. лицо/юр. лицо и т.д.) возможно создать битовые индексы, которые позволяют очень быстро производить выборку по такому рода параметрам, так как значения хранятся в виде битовой карты и сравнение в запросе происходит на уровне команд процессора сервера.

**Параллельное выполнение.** Параллельное выполнение (Parallel Execution) позволяет значительно ускорить выполнение длительных операций, которые часто встречаются при построении и использовании хранилищ данных и аналитических систем. СУБД Oracle позволяет распараллелить такие операции, как выполнение запросов к очень большим таблицам, связкам таблиц, создание больших индексов, материализованных представлений, вставку и изменение записей в больших таблицах при загрузке данных в хранилище. Параллельное выполнение позволяет полнее раскрыть преимущества многопроцессорных систем и кластеров.

**Секционирование.** Вместе с параллельным выполнением еще одним механизмом увеличения производительности больших хранилищ данных и аналитических систем является секционирование (Partitioning). Этот механизм позволяет физически разбивать таблицы на «горизонтальные» части или секции по логическому условию, например, по дате или по значениям каких-либо других полей. При этом логически таблица остается единой, но на физическом уровне с каждой такой секцией система работает независимо. Это позволяет автоматически заменять запросы к большой исходной таблице запросами к отдельным ее секциям и, таким образом, уменьшить время обработки запросов.

**Материализованные представления.** Запросы аналитиков обычно бывают достаточно сложными, содержат большое количество связей между таблицами, агрегирующие выражения и т.п., что при больших объемах таблиц хранилища приводит к неприемлемым временным затратам на получение результатов. Материализованные представления позволяют хранить в явном виде результаты часто встречающихся запросов и не выполнять их заново каждый раз, а использовать готовый результат. Материализованные представления напоминают обычные представления (view),

но, в отличие от них, в базе данных хранится не только текст запроса, но и результат его выполнения в виде некоторой промежуточной таблицы. Данные в исходных таблицах, участвующих в запросе, могут изменяться и в этом случае необходимо обновлять и промежуточную таблицу. Существенно, что эта синхронизация материализованного представления с исходными таблицами выполняется системой автоматически в соответствии с определенным регламентом, задаваемым разработчиком. Может оказаться, что запрос пользователя не хранится непосредственно в виде материализованного представления, но выполнить его гораздо быстрее не на исходных таблицах, а используя некоторое существующее материализованное представление. Например, если исходная таблица содержит обороты за каждый день и создано материализованное представление, содержащее обороты по месяцам, то было бы разумно использовать его, а не исходную таблицу при вычислении оборотов за каждый квартал. Такое «переписывание» запросов производится также автоматически, что освобождает разработчика от необходимости знать о всех промежуточных результатах, которые могут оказаться полезными при оптимизации запросов, и работать непосредственно с исходными таблицами.

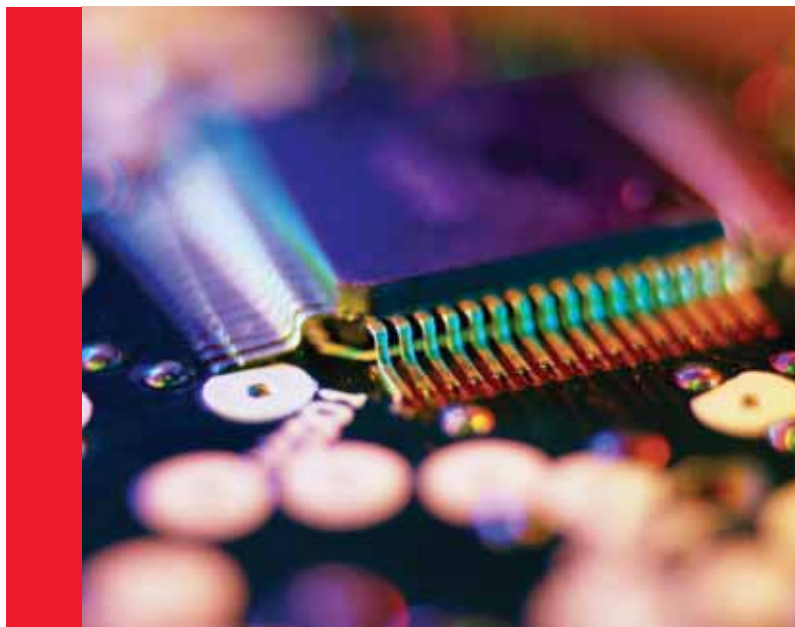
**Управление суммарными данными.** Одним из важнейших усовершенствований в области оптимизации выполнения аналитических запросов является технология управления суммарными данными на основе материализованных представлений (materialized views). Анализируя статистику работы системы, СУБД рекомендует администратору необходимые агрегаты, автоматически их создает и периодически обновляет. Затем при выполнении запросов с агрегированием система автоматически переписывает их таким образом, чтобы они обращались к суммарным данным, хранящимся в материализованных представлениях. Такой подход резко, иногда на несколько порядков, повышает производительность хранилища данных для конечных пользователей.

**Аналитические функции.** Для облегчения программирования сложных аналитических запросов в SQL Oracle встроено большое количество аналитических функций. Это такие функции, как функции ранжирования, вычисления лагов, линейной регрессии, вычисления скользящих, и многие другие. Аналитические функции призваны облегчить программирование и уменьшить нагрузку на клиентские приложения. Следует отметить, что существует возможность создавать свои сложные агрегирующие функции, в том числе не только на языке PL/SQL.

В состав Oracle Database входит Enterprise Manager — мощное графическое средство, специально разработанное для эффективного администрирования. С его помощью можно управлять всеми объектами базы данных и автоматизировать основные административные задачи.

#### **Сбор данных из разнородных источников**

Задачи извлечения данных из разных источников, их согласования, проверки, агрегирования и выполнения других преобразований обычно решаются с использованием специальных инструментальных средств, объединяемых под общим названием ETL-средства. Как правило, такое средство предоставляет возможность визуального проектирования процессов сбора и преобразования данных с помощью специальных графических редакторов, сохраняет эти спецификации в репозитории и затем выполняет их на специально выделенном сервере преобразования и загрузки.



В отличие от этого, основной принцип Oracle в отношении ETL-технологии – максимально использовать возможности сервера Oracle Database для выполнения ETL-процедур. Это позволяет избежать проблем с интеграцией различных средств и повысить производительность преобразования и загрузки данных в хранилище. Для реализации этого подхода стандартные средства СУБД Oracle расширены дополнительными командами и механизмами, полезными для задач извлечения, преобразования и потоковой загрузки данных и повышающими эффективность их реализации. К таким средствам относятся внешние таблицы, табличные функции, одновременный ввод и корректировка данных, ввод данных в несколько таблиц и др.

**Внешние таблицы.** Для облегчения работы с внешними источниками данных при загрузке в хранилище можно воспользоваться механизмом внешних таблиц, представляющих собой виртуальные таблицы, которые воспринимаются внутри СУБД как таблицы Oracle, но физически являются таблицами в других СУБД или в текстовых файлах. Встроенные механизмы позволяют очень эффективно работать с внешними таблицами и достичь значительного увеличения производительности при загрузке данных.

**Табличные функции.** Табличными называются такие функции, которые возвращают не отдельное значение, а множество строк таблицы. При выборке данных с такой функцией можно обращаться как с обычной таблицей, т.е. использовать в операторе SELECT. С этой точки зрения она подобна представлению (view), поскольку результирующие данные также не хранятся, а вычисляются динамически во время запроса в соответствии с некоторым описанием. Но если для представления такое описание всегда ограничено некоторым SELECT-оператором, то в табличной функции формирование данных может задаваться любым алгоритмом. Использование табличных функций при сложных согласованиях и преобразованиях данных позволяет обходиться без многократного сохранения промежуточных результатов и тем самым существенно сократить время обработки данных перед загрузкой в хранилище.

#### **Одновременное обновление и вставка данных.**

Процесс загрузки данных в хранилище часто включает как добавление новой информации, так и корректировку уже существующей. Например, в случае обновления информации о клиентах необходимо для уже существующего клиента произвести обновление соответствующей записи таблицы, а если такого клиента нет, – то вставку новой записи. Обычно это осуществляется двумя отдельными операциями – с помощью команды INSERT производятся все добавления новых данных, а для корректировки существующей информации выполняется UPDATE. В Oracle имеется возможность совместить эти действия с помощью команды MERGE, которая выполняет одновременно и обновление и вставку на уровне сервера, что, естественно, повышает производительность загрузки.

**Одновременная вставка в несколько таблиц.** Еще одно полезное средство Oracle для потоковой загрузки данных – одновременная вставка в несколько таблиц. В ряде случаев при загрузке таблицы в хранилище ее содержимое необходимо «разбросать» по нескольким таблицам. Так, например, при загрузке данных о клиенте может оказаться, что информация обо всех физических лицах должна помещаться в одну таблицу хранилища, а о юридических лицах в другую. Стандартными средствами реляционных СУБД это приходится делать в несколько проходов, что в случае больших таблиц занимает значительное время. В Oracle Database с помощью команды INSERT ALL за один проход по таблице-источнику можно выполнить вставку одновременно в несколько целевых таблиц.

При построении корпоративной информационно-аналитической системы в организациях со сложной многоуровневой филиальной структурой и большим числом территориально распределенных подразделений, не всегда связанных единой локальной сетью, возникает проблема организации процессов сбора информации от различных подразделений в соответствии с существующими регламентами и техническими требованиями. В этом случае, кроме собственно процессов ETL, требуются дополнительные средства и решения, обеспечивающие эффективную реализацию собственно процессов передачи данных между узлами распределенной структуры, пред-

варительного контроля данных, модификации входных форм и т.п. В качестве такого решения можно использовать систему сбора данных, разработанную на основе продуктов Oracle компанией Leaves, являющейся сертифицированным партнером корпорации Oracle.

### Oracle Warehouse Builder – инструментальная среда создания хранилища данных

Oracle Warehouse Builder (OWB) — это многофункциональная расширяемая CASE-среда для разработки и развертывания корпоративных хранилищ и витрин данных. Построенный на базе открытой архитектуры Common Warehouse Metamodel (CWM), OWB – интегрированное средство, позволяющее решать различные задачи, не прибегая к помощи нескольких узкоспециализированных продуктов. К этим задачам относятся проектирование, создание и администрирование хранилища данных, разработка и генерация процедур извлечения, преобразования и загрузки данных из различных источников, управление метаданными и интеграция инструментальных средств доступа.

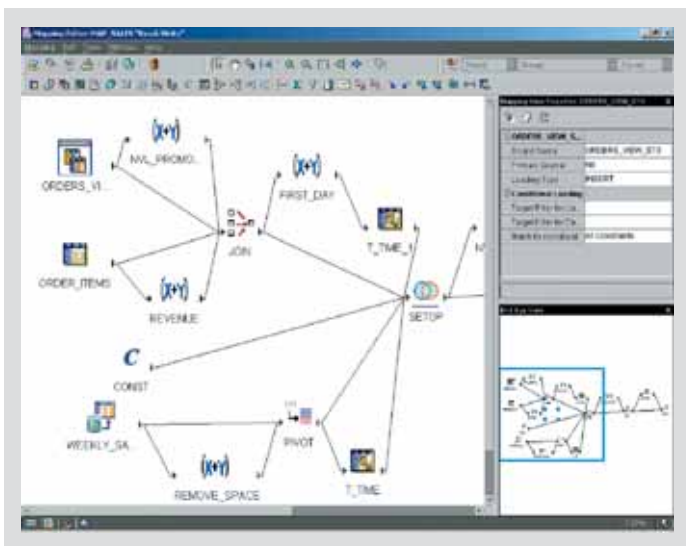


Рис. 3. Редактор отображений Oracle Warehouse Builder

Warehouse Builder использует все новые возможности СУБД Oracle Database, связанные с технологиями хранилищ данных, такие как материализованные представления, объекты типа “измерение” (“dimension”) и т.д. При решении задач администрирования и управления процессами загрузки данных в хранилище Warehouse Builder опирается на Oracle Enterprise Manager и Oracle Workflow, интегрированные с базой данных Oracle Database.

Реализованный на Java и изобилующий Мастерами (Wizards) пользовательский интерфейс Warehouse Builder существенно облегчает создание, развертывание и поддержку Хранилища Данных. Продукт позволяет визуально моделировать схему Хранилища либо импортировать метаданные из репозитория Oracle Designer. Разработчик может визуально определять отображения между источниками данных и Хранилищем, и Warehouse Builder затем автоматически генерирует на их основе модули загрузки (в виде процедур PL/SQL или скриптов SQL\*Loader). Продукт имеет встроенную библиотеку функций преобразования данных, которую при необходимости можно расширять собственными процедурами на PL/SQL.

Источниками данных для OWB, помимо СУБД Oracle различных версий, могут быть плоские файлы, СУБД

других производителей (доступ через шлюзы и ODBC), а также приложения ERP (Oracle E-Business Suite, SAP R/3 и др.).

Открытая архитектура Common Warehouse Metadata позволяет интегрировать метаданные Warehouse Builder и инструментов доступа к информации — Discoverer и OLAP Option. С помощью специального Мастера можно перенести бизнес-описания из репозитория Warehouse Builder в соответствующие слои метаданных средств анализа. Благодаря этому резко снижаются затраты на организацию доступа конечных пользователей к данным Хранилища и обеспечивается целостность метаинформации.

В качестве инструментальной среды для выполнения всех задач по проектированию хранилища можно дополнительно использовать продукт Oracle Designer.

Центральным компонентом Oracle Designer является репозиторий – специальная база данных, в которой хранится вся метаинформация о структуре и объектах хранилища. Различные “клиентские” средства, входящие в состав инструментальной среды, обеспечивают доступ к репозиторию и ориентированы на выполнение различных задач, возникающих в процессе проектирования.

Для изучения и анализа баз данных источников используется процедура реинжиниринга (обратного проектирования), которая автоматически восстанавливает в репозитории всю метаинформацию о схеме и объектах любого источника данных. Для работы с полученными спецификациями, выделения интересующего подмножества таблиц и представлений, служит диаграмма схем базы данных.

Это средство включает в себя удобный графический редактор для отображения в виде диаграмм структур данных, позволяет вводить в репозиторий дополнительную информацию о семантике и особенностях таблиц, столбцов и других объектов.

На основе результатов анализа данных источников и информационных потребностей пользователей проектируется хранилище данных. Эту задачу во многом облегчают удобные средства визуального проектирования, входящие в состав диаграммера схем баз данных, широкий спектр тестов на полноту и согласованность разрабатываемой схемы, автоматическая генерация разнообразной документации и т.п.

Важной особенностью этапа, связанного с созданием схемы данных хранилища, является использование специальной техники проектирования, так называемых схем типа звезда или снежинка. Каждая такая схема состоит из выделенной главной таблицы (таблица фактов), содержащей значения одного или нескольких показателей, и нескольких небольших по числу строк таблиц-справочников, на которые ссылаются строки таблицы фактов.

Для каждой таблицы в репозитории хранится много различной метаинформации. Здесь можно указать различные тонкости хранения таблицы, задать характеристики столбцов, причем перечень этих характеристик шире, чем то, что может быть включено в соответствующую программу создания базы данных. Примером служит поддержка аппарата доменов или пользовательских типов данных. Для таблиц указываются также ограничения целостности.

Важной особенностью средств проектирования Oracle Designer, которая выгодно отличает его от подобных продуктов других фирм, является возможность специфицировать в репозитории не только стандартные объекты логического уровня – таблицы и представления, но и различные специальные объекты как логического, так и фи-



зического уровня. Например, при проектировании хранилища можно заранее предусмотреть различные структуры оптимизации – индексы, кластеры и т.п., определить характеристики табличных пространств. Из объектов логического уровня полезными являются спецификации пользователей и групп пользователей.

После того, как основная работа по проектированию хранилища завершилась, с помощью специальной утилиты генерируется программа создания структуры хранилища. Программы хорошо структурированы: каждому типу объектов отводится свой файл, при этом перед запуском процедуры можно уточнить, какие именно объекты данного типа должны быть созданы.

Инструментальные средства Oracle Designer помогают не только создавать хранилища данных. С их помощью впоследствии можно поддерживать и развивать хранилище, корректировать метаинформацию о старых объектах, вводить новые. Для выполнения этих задач в состав Oracle Designer входит целый ряд специальных утилит и средств, начиная с генерации многочисленных отчетов, обеспечивающих точную техническую документацию, и кончая утилитами, сравнивающими реальные объекты хранилища с их метаописанием. В результате этого сравнения можно сгенерировать программу, осуществляющую реструктуризацию хранилища и приведение реальных объектов к их метаописанию.

### Бизнес-анализ данных

Для реализации различных методов анализа данных используется продукт Oracle Business Intelligence, включающий широкий спектр различных инструментальных средств для отчетности, нерегламентированных запросов, многомерного анализа, построения и использования интерактивных информационных панелей, поддержки режима off-line и др. Этот пакет поставляется в нескольких редакциях – Oracle Business Intelligence Standard Edition One, Oracle Business Intelligence Standard Edition, Oracle Business Intelligence Suite Enterprise Edition.

### Oracle Business Intelligence Standard Edition

Стандартная редакция Oracle Business Intelligence представляет собой интегрированный пакет различных средств анализа данных, основанных на единых стандартах представления и использования данных и использующих возможности СУБД Oracle для повышения производительности аналитической обработки.

### Стандартная отчетность и Oracle Reports

Oracle Reports – это инструмент создания и публикации стандартных форматированных отчетов, который позволяет реализовать технологию регламентированной отчетности в масштабах предприятия. Отчеты Oracle Reports могут иметь сложную структуру, содержать результаты нескольких запросов, автоматически формировать итоги и подитоги, а также включать в себя разнообразную графическую и ссылочную информацию. Как правило, такие отчеты готовятся квалифицированными специалистами и впоследствии выполняются конечными пользователями.

Встроенные в Oracle Reports Мастера помогают новичкам и существенно ускоряют создание отчетов. “Живой” предварительный просмотр дает возможность оценивать полученный результат и делать необходимые изменения.

Масштабируемая многоуровневая архитектура Oracle Reports идеально подходит для оптимизации нагрузки при обработке большого количества запросов. Установив сервер отчетов Reports Server на мощном центральном компьютере, можно повысить общую производительность и освободить ресурсы клиентских рабочих станций. Отчеты могут выполняться по расписанию в периоды низкой загрузки (например, в ночное время), после чего конечные пользователи в любой момент могут просмотреть полученные результаты.



Oracle Reports позволяет легко распространять информацию по организации, используя Web-архитектуру. Используя Мастер для Web, можно за несколько секунд сделать любой отчет доступным в корпоративной сети Интранет. Для вызова таких динамических отчетов и просмотра результатов в формате PDF, HTML или XML пользователям потребуется только браузер Web.

Reports Developer предоставляет коллективу разработчиков среду, в которой можно получить и обработать данные, спроектировать внешний вид отчетных документов. Визуальное проектирование и генерация отчетов могут вестись на разных программных платформах. Пользователи получают результат работы Reports либо в печатном виде, либо в виде электронных файлов стандартного формата HTML, PDF, RTF или XML, доставленных по электронной почте или записанных на файл-сервер. В оперативном режиме доступ пользователей к отчетам может быть организован через Web во внутренней или внешней сети. В этом случае достаточно обычного навигатора, чтобы получить готовый или сгенерированный «на лету» отчет. Отчеты также могут формироваться при возникновении какого-либо, заранее определенного события, например, поступление новой информации в базу данных.

Reports применяет технологию JSP (Java Server Pages) для представления информации в удобном для пользователей виде. Ссылку на созданный отчет можно разместить на любой Web-странице. Важно отметить, что Reports Developer не требует от разработчиков применения нескольких подходов для проектирования отчетов, выполняемых в различных средах. Единжды созданный отчет будет выполняться на большинстве платформ, а его выходная форма представлена в разных форматах - от текстовых файлов до динамических Web-страниц. Если несколько пользователей запрашивают одни и те же данные практически одновременно, то отчеты не создаются при каждом обращении снова. Пользователям выдается копия первого отчета из буфера на сервере приложений. В соответствии с традициями средств разработки Oracle, Reports Developer тесно интегрирован с Oracle Database, использует одинаковые с ним языки разработки. В тоже время, Reports позволяет получать данные из сервера аналитической обработки данных Express Server, а также из других информационных источников - из плоских файлов или через стандартные интерфейсы (ODBC и JDBC).

### Произвольные запросы с помощью Oracle Discoverer

Oracle Discoverer — это инструмент для получения произвольных отчетов, формирования нерегламентированных запросов и анализа данных. Он обеспечивает быстрый и удобный доступ к информации, содержащейся в реляционных Хранилищах и Витринах Данных, а также в OLTP-системах (в том числе не обязательно реализованных на СУБД Oracle).

В отличие от Oracle Reports, где конечный пользователь является только потребителем информации, в Oracle Discoverer ему предоставлена возможность самостоятельного получения необходимых данных. Естественно, для этого нужно скрыть от пользователя внутреннюю структуру Хранилища или Витрины и представить имеющуюся там информацию в понятных ему бизнес-терминах, таких как заказчик, продукт, объем продаж и т.д. Для этого в Discoverer используется семантический слой метаданных — Слой Конечного Пользователя (End User Layer). Он содержит всю описательную и другую метаинформацию, необходимую для эффективной работы конечных пользователей. Этот слой может либо создаваться непосредственно разработчиком, либо автоматически генерироваться в Warehouse Builder на основе метаданных Хранилища или Витрины.

Для повышения производительности в Discoverer реализован ряд уникальных технологических возможностей. Среди них — прогнозирование времени выполнения за-

проса до его начала, возможность создания и последующего автоматического использования суммарных таблиц, интеллектуальный механизм кэширования. Эти технологии обеспечивают приемлемые времена отклика даже при работе с очень большими хранилищами данных.

Существует несколько редакций Oracle Discoverer для разных категорий пользователей. Discoverer Administrator предназначен для создания и администрирования Слоя Конечного Пользователя, а также для определения прав доступа конечных пользователей к информации и функциональным возможностям. С помощью встроенных Мастеров разработчик может легко создавать объекты метаданных (категории, иерархии и другие), редактировать их свойства, управлять суммарными данными и т.д. Мощь и удобство Discoverer Administration Edition обеспечивают высочайшую продуктивность разработки.

Discoverer Desktop — это генератор интерактивных отчетов для конечного пользователя, работающих в архитектуре «клиент-сервер». С помощью этого инструмента пользователь может формулировать запросы в терминах бизнес-области. Результаты выполнения запросов оформляются в виде отчетов или графиков, внешний вид которых определяется также конечным пользователем. Отчеты организуются в рабочие книги, которые могут храниться как локально в виде файлов, так и в базе данных, что делает результаты анализа доступными для других сотрудников. Условия — фильтры, итоги и подитоги, вычисляемые поля и т.д. — все это конечный пользователь может определять с помощью встроенных Мастеров. Возможности Discoverer Desktop по разнообразной детализации данных и получению любых аналитических срезов делают его мощным и удобным инструментом динамического доступа к хранилищам и витринам данных.

Discoverer Plus представляет собой WEB-версию Desktop Edition. Этот продукт функционально совпадает с Discoverer Desktop, но позволяет конечным пользователям работать, имея на своем компьютере только web-браузер. Discoverer Viewer — это функционально ограниченная версия Discoverer Plus, ориентированная на тех конечных пользователей, которым достаточно только просматривать готовые отчеты и представления информации, созданные средствами Desktop Edition или Discoverer Plus. Discoverer Plus и Discoverer Viewer входят в состав Internet Application Server 10g Enterprise Edition.

Таким образом, Oracle Discoverer обеспечивает уникальное сочетание простоты использования, производительности и простоты администрирования. Его применение при небольших затратах дает ощутимый результат — после быстрого внедрения организация получает немедленную выгоду от упрощившегося доступа пользователей к бизнес-информации. Возможность переноса метадан-

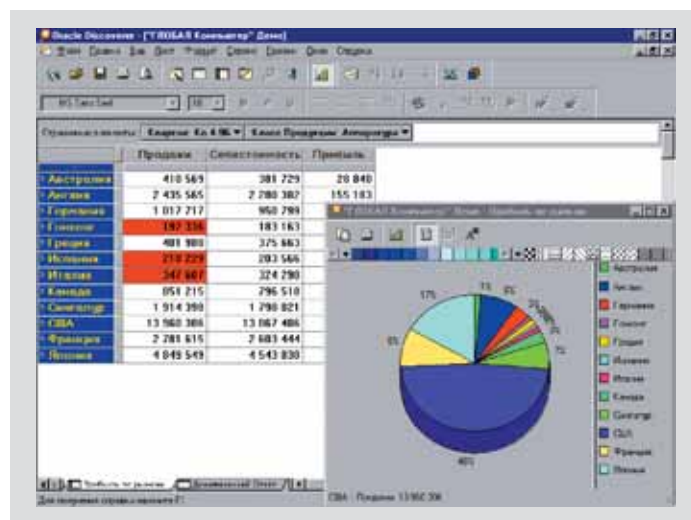


Рис. 4. Анализ данных в Oracle Discoverer

ных из Warehouse Builder сводит к минимуму усилия по внедрению Discoverer над хранилищем данных.

### Многомерный анализ данных на основе Oracle OLAP

Для решения аналитических задач, связанных со сложными расчетами, прогнозированием, моделированием сценариев “что-если” и т.д., применяется особая технология многомерного анализа. Эта технология реализуется специальной компонентой Oracle OLAP, входящей в состав СУБД Oracle Database. Опция Oracle OLAP позволяет хранить и обрабатывать многомерную информацию на том же сервере баз данных, где находится реляционное хранилище. По функциональным возможностям эта опция сравнима с многомерной СУБД Oracle Express и по-существу завершает процесс интеграции технологии Oracle Express с реляционным сервером Oracle Database. Средства OLAP поддерживают в полном объеме основной язык сервера Express (сейчас этот язык переименован в DML и получил ряд функциональных расширений для использования возможностей СУБД Oracle), а для существующих баз данных Express обеспечивается их миграция в СУБД Oracle.

Основой OLAP технологии является логическая многомерная модель данных, с помощью которой можно определять и работать с такими понятиями, как измерения, иерархии, многомерные показатели и т.п. Все определения и описания таких многомерных объектов хранятся в репозитории OLAP Catalog, представляющем собой специально выделенную схему Oracle.

Каждый логический объект в каталоге описывается набором параметров. Например, для измерения указывается его имя, атрибуты, определяются возможные иерархии, а при определении куба задается имя, состав показателей, набор измерений, от которых зависят все эти показатели, типы агрегирования и др. Кроме этого для каждого логического объекта задается ссылка на «физический» объект, содержащий собственно данные. В простейшем случае такими «физическими» объектами являются столбцы обычных реляционных таблиц. Это соответствует ситуации, когда логическая многомерная модель реализуется обычной схемой звезда или снежинка. Предусмотрена и другая возможность хранения многомерных объектов – в виде объектов аналитического пространства. Аналитическое пространство – это многомерная база Oracle Express, хранящаяся внутри базы данных Oracle Database в виде LOB-полей (рис. 5).

Для работы с описанными в каталоге OLAP многомерными объектами в состав опции OLAP включена компонента Oracle9i OLAP API – набор Java-классов, реализующих

все операции по созданию, модификации и манипулированию многомерными объектами. Этот интерфейс позволяет разработчикам на Java создавать различные приложения, работающие с многомерной информацией. Для повышения эффективности процесса разработки фирма Oracle предоставляет инструментальную среду JDeveloper, дополненную специальными компонентами BI Beans. Эти компоненты, оформленные в соответствии со стандартами технологии Java beans, используя Oracle OLAP API для доступа к многомерной информации и позволяют легко и удобно разрабатывать приложения любой степени сложности, работающие с многомерной информацией как в архитектуре “клиент-сервер”, так и в среде Internet.

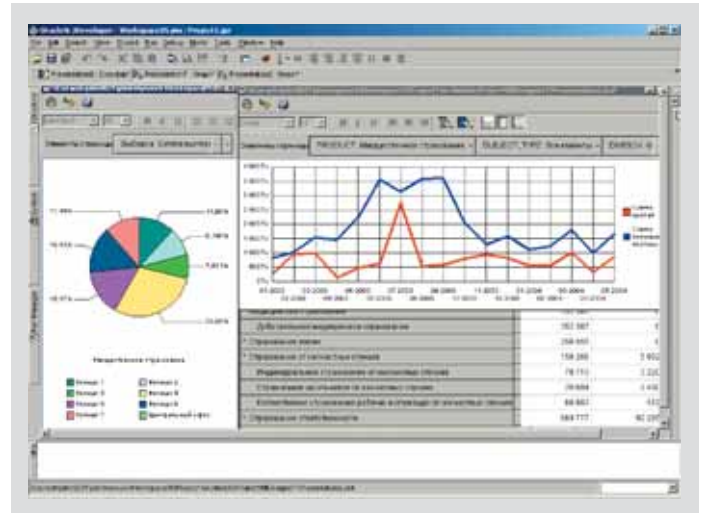


Рис. 6. Многомерный анализ с помощью Jdeveloper и BI Beans

Доступ к многомерной информации можно осуществлять не только из Java-приложений, но и из приложений, ориентированных на SQL. Если данные физически находятся в реляционных таблицах, то с ними можно работать непосредственно из любого языка, включающего SQL. В случае аналитического пространства существует возможность создать набор реляционных представлений (view), реализующих доступ к объектам аналитического пространства. Такие представления автоматически генерируются с помощью специальных хранимых процедур, использующих объектно-ориентированные возможности СУБД Oracle и технологию табличных функций.

Конечные пользователи могут получать динамический доступ к многомерным данным, используя в качестве интерфейса Microsoft Excel. Компонента Oracle Spreadsheet Add-In дополняет стандартные возможности этой электронной таблицы, позволяя с помощью простого Мастера строить в среде Excel интерактивные многомерные отчеты. Пользователям доступны те же основные манипуляции с данными, что и в других инструментах OLAP, – получение различных срезов, детализация и Селектор. Для графического представления данных применяются соответствующие возможности Excel.

### Oracle Data Mining для автоматического поиска закономерностей

В настоящее время наряду с традиционными средствами анализа данных, такими как регламентная отчетность, выполнение нерегламентированных запросов, широко используются современные методы углубленного исследования данных, получившие название извлечение знаний (data mining).

Основная задача технологии извлечения знаний состоит в выявлении в больших наборах данных скрытых закономерностей, зависимостей и взаимосвязей, полезных при принятии решений на различных уровнях управления. Такие закономерности представляются в виде моделей

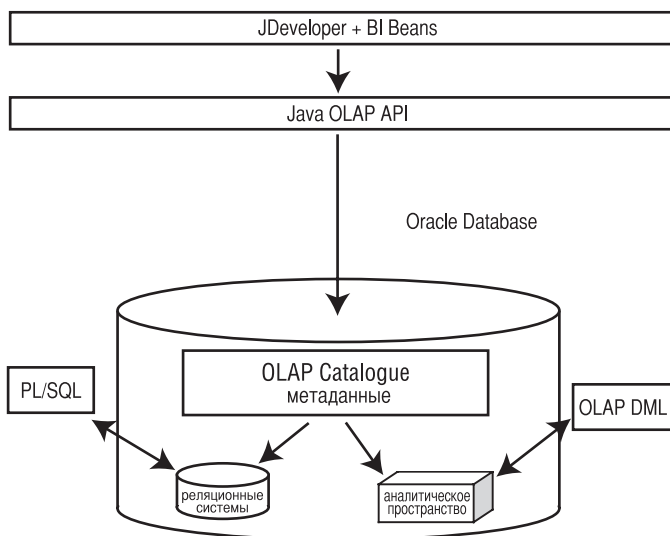


Рис.5 Основные компоненты опции Oracle OLAP

различного типа, позволяющих проводить классификацию ситуаций или объектов, прогнозировать их поведение, выявлять группы сходных объектов и т.п. Существенно, что модели строятся автоматически в процессе анализа имеющихся данных об объектах, наблюдениях и ситуациях с помощью специальных алгоритмов, основанных на различных математических и статистических методах. В банках и финансовых организациях средства извлечения знаний можно использовать для выявления и исследования факторов, влияющих на доходность клиентов, для определения и анализа причин оттока клиентов, сегментации базы физических и юридических лиц, прогнозирования результатов маркетинговой компании по предложению новой услуги и т.п.

Для реализации технологии извлечения знаний в состав СУБД Oracle Database включена специальная опция Oracle Data Mining, которая состоит из двух компонент: Data Mining Server и Oracle Data Mining API.

Алгоритмы, реализованные в Oracle Data Mining	
Классификационные модели	Naive Bayes, Adaptive Bayes Network
Классификации и регрессионные модели	Support Vector Machine
Поиск существенных атрибутов	Minimal Descriptor Length
Кластеризация	Enhanced K-means, O-cluster
Поиск ассоциаций	Apriory Algorithm
Выделение признаков	Non-Negative Matrix Factorization

Основу компоненты Data Mining Server составляют процедуры, реализующие различные алгоритмы построения моделей. Например, классификационный алгоритм строит на основе имеющейся архивной информации о доходности клиентов классифицирующее дерево, позволяющее по значениям различных параметров клиента прогнозировать его доходность. К числу других методов относятся алгоритмы кластеризации, выявляющие группы «похожих» объектов, поиск ассоциаций, с помощью которых можно определять устойчивые взаимосвязи между различными факторами и другие алгоритмы. Ниже в таблице приводится список алгоритмов, реализованных в Oracle Database 10g. Все они вместе с различными метаданными, описывающими параметры моделей, особенности их тестирования и применения и другие характеристики, хранятся в специально выделенной схеме Oracle – репозитории Oracle Data Mining.



Рис.8 Информационная панель в среде Oracle Business Intelligence Suite EE

Важная особенность алгоритмов состоит в том, что все они работают непосредственно с реляционными базами данными и не требуют выгрузки и сохранения данных в специальных форматах. Кроме собственно алгоритмов, в опцию ODM входят средства подготовки данных, оценки результатов, применения моделей к новым наборам данных. Использовать все эти возможности можно как на программном уровне с помощью Java API или PL/SQL API, так и с помощью графической среды Oracle Data Miner, ориентированной на работу аналитиков, решающих задачи прогнозирования, выявления тенденций, сегментации и др.

#### Интегрированная среда Oracle Business Intelligence Suite Enterprise Edition

Корпоративная редакция Oracle Business Intelligence представляет собой интегрированную платформу для реализации различных методов анализа данных, основанную на передовых технологиях Siebel Analytics. Важной отличительной особенностью этой редакции является возможность одновременной работы с разнородными источниками данных без необходимости предварительной загрузки их в единое хранилище данных. Это позволяет единым универсальным способом получать отчеты из любых функционирующих в организации информационных систем. Наряду с традиционными технологиями анализа, такими как регламентная отчетность, произвольные запросы, инструментальная среда позволяет быстро и легко сформировать интерактивные информационные панели, управлять доступом к информационным отчетам, анализировать данные в автономном режиме, т.е. без связи с информационными источниками, распространять в реальном масштабе времени различные уведомления или алерты о состоянии ключевых показателей или выхода некоторого показателя за заданные границы, работать с результатами бизнес-анализа с помощью популярных средств MS Office, включая MS Excel и MS Word. Все эти возможности реализуются не набором отдельных продуктов, а единой инструментальной средой с единым репозиторием метаданных и средством администрирования. Вся деятельность по бизнес-анализу данных осуществляется в среде Интернет под управлением специального сервера Business Intelligence Server, использующего различные механизмы оптимизации для повышения эффективности и производительности аналитической обработки.

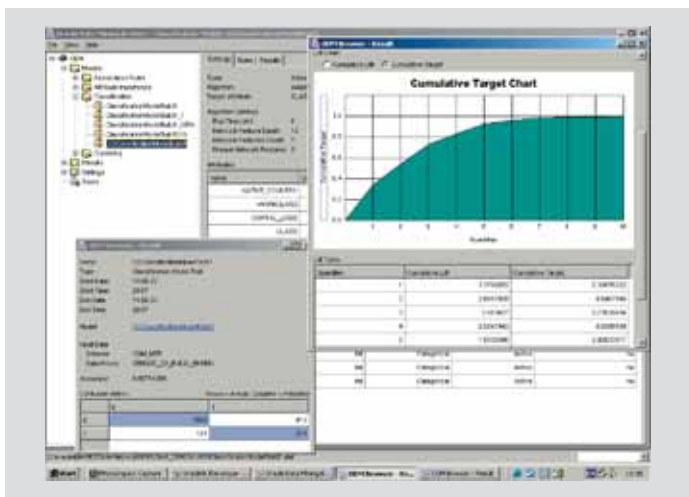


Рис.7 Инструментальная среда Oracle Data Miner



## Методология создания хранилищ данных

Кроме собственно продуктов, обеспечивающих полное решение для корпоративной информационно-аналитической системы, корпорация Oracle предлагает оригинальную методологию выполнения проекта по созданию и сопровождению таких систем. Эта методология называется Data Warehouse Method и является частью общего подхода Oracle к проектированию и реализации различных проектов.

Методология DWM использует подход пошаговой разработки. Достоинством такого подхода является создание масштабируемых решений, то есть решений с постепенным непротиворечивым наращиванием функциональности.

В основе методологии построения систем, основанных на концепции хранилищ данных, лежит «двухмерный» подход ко всему проекту: с одной стороны проект разбивается на процессы или задачи, а с другой – состоит из этапов выполнения.

С точки зрения времени и последовательности выполнения различных задач проект делится на следующие этапы:

- **Стратегия** - определение целей, приоритетов пошаговой разработки и инфраструктуры сетевого графика для построения хранилища данных, основанного на стратегии деловых инициатив клиента. Оценивается организационная структура клиента, критические факторы успеха, главные ограничения, риски и показатель стоимость/прибыль хранилища данных. Кроме этого определяется техническая архитектура предприятия и архитектура хранилища данных. Основными задачами этапа являются сбор и документирование информации по постановке задачи, согласование постановки задачи в минимально возможные сроки. Во время этапа стратегии необходимо также определить проектные стратегии в таких областях как обучение, конвертация данных, тестирование, документация и внедрение системы.

- **Постановка задачи** - определение области и целей для оценки затрат на поэтапную разработку. Документирование источников данных и определение области качества данных. Создание технической архитектуры и архитектуры хранилища данных для целевого решения.

- **Анализ** - составление бизнес требований для целевого решения. Создание логических моделей, сбор подробных требований для используемых источников, данных и документирование требований конечного пользователя к доступу данных.

- **Проектирование** - использование информации, полученной при анализе для разработки технических требований, удовлетворяющих детализованным требованиям. Проверяется удовлетворение и поддержка требований разработкой.

- **Реализация** - создание компонентов хранилища данных, включая базы данных, модули сбора данных и механизмы доступа к данным, а также тестирование конфигурации системы.

- **Внедрение** – установка системного и прикладного программного обеспечения для работы хранилища данных, подготовка персонала клиента для использования и администрирования хранилища данных, переход к промышленной версии решения.

- **Эксплуатация** – оценка предшествующего шага для планирования следующих поэтапных затрат. Анализ оценки требований, которые выявились во время жизненного цикла разработки и оценка использования хранилища данных. Проверка плана проекта, которая станет основанием для следующих действий. Рассмотрение задач, которые выполнены и отсутствует необходимость для их повторного решения.

С точки зрения выполняемых работ проект разбивается на процессы. Каждый процесс может быть представлен как некоторый подпроект, имеющий определенную цель и состоящий из набора задач и соответствующих каждой задаче выходных результатов.

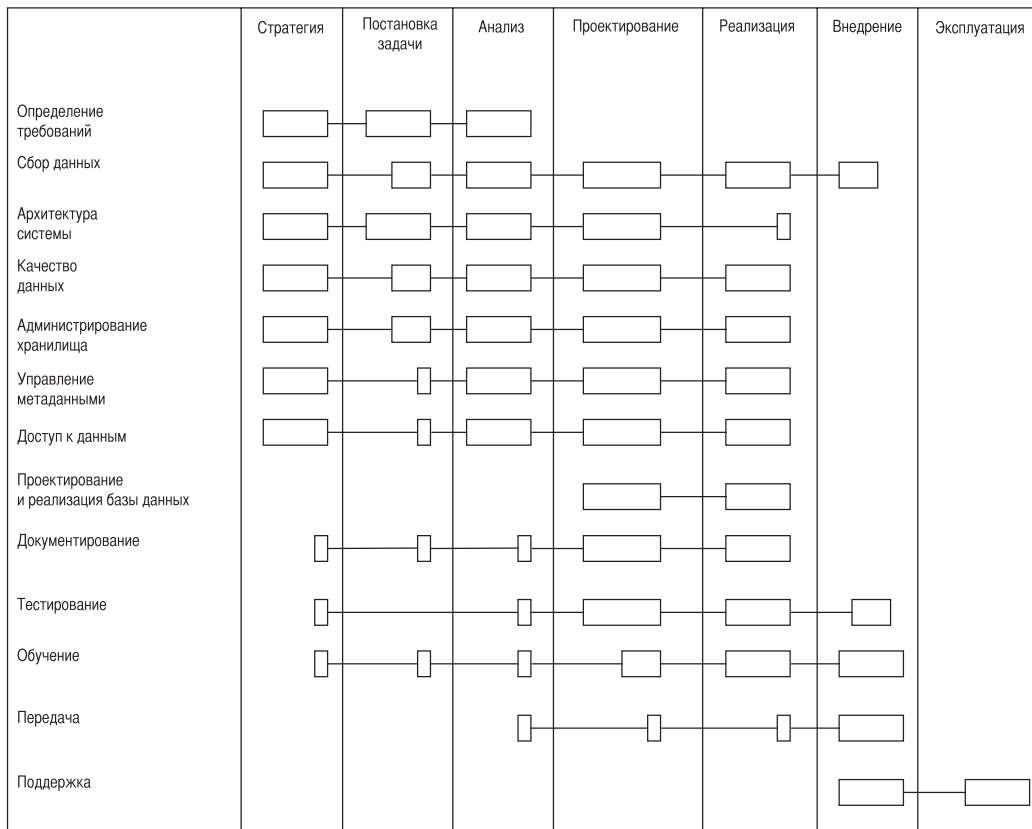


Рис. 8 Процессы и этапы методологии Oracle Data Warehouse Method

При построении хранилищ данных выделяются следующие процессы:

- Определение общих требований
- Сбор данных
- Архитектура системы
- Качество данных
- Администрирование хранилища
- Управление метаданными
- Доступ к данным
- Проектирование и реализация базы данных
- Документирование
- Тестирование
- Обучение
- Передача
- Поддержка

Все указанные выше процессы выполняются на протяжении всех этапов (рис.9). Для каждого процесса четко определена последовательность решаемых задач, представлены шаблоны выходных документов и формы представления результатов.

## Заключение

Направление Хранилищ Данных и систем бизнес-анализа является сегодня для Oracle одним из самых приоритетных. Будучи поставщиком полного технологического решения в данной области, Oracle выпускает новые продукты и постоянно совершенствует существующие.

Важнейшая характеристика систем поддержки принятия решений Oracle — комплексность подхода и его реализации. Можно пытаться строить Хранилище Данных, используя СУБД от одного поставщика, инструментарий создания — от другого, средства доступа к информации — от третьего и т.д. Однако нельзя не учитывать тот факт, что при этом неизбежно возникнет проблема интеграции различных продуктов в единое решение.

Выбирая корпорацию Oracle, организация получает все необходимое для построения Хранилища Данных. Программные продукты Oracle полностью покрывают спектр задач, стоящих в процессе внедрения. Техническая поддержка, обучение и консалтинг, предоставляемые корпорацией и ее партнерами, образуют комплекс услуг, которые востребованы в любом комплексном проекте. Методология Oracle Data Warehouse Method содержит в концентрированной форме многолетний опыт построения Хранилищ Данных и гарантирует успешное внедрение. Свяжитесь с консультантами Oracle или обратитесь в Центры компетенции по направлению Fusion Middleware : Business Intelligence and Data Warehousing.

### Дополнительная информация

Дополнительную информацию о решениях и технологиях Oracle Вы можете найти на сайте корпорации по адресу: <http://www.oracle.com/ru/>

## Центры компетенции по внедрению решений Fusion Middleware: Business Intelligence and Data Warehousing

Успех сложных информационных проектов зависит от экспертизы и опыта специалистов. Статус Центра компетенции является подтверждением опыта, экспертизы, надежной репутации и профессионализма партнера, проявленных в результате реализации проектов по внедрению решений Oracle. Он также подтверждает наличие высококвалифицированных специалистов, проводимую маркетинговую политику и активное продвижение новейших технологий Oracle.



### Консалтинговая Группа Борлас

Консалтинговая группа «Борлас» обладает высшим статусом партнерства с корпорацией Oracle: Мастер-партнер (Oracle Certified Advantage Partner) и с 2005 года является первым в России Центром компетенции по интеграционным технологиям (Fusion Middleware: Integration), а с 2002 года – Центром компетенции по созданию аналитических систем и хранилищ данных (Fusion Middleware: Business Intelligence). Такое сочетание компетенций позволяет реализовывать на базе современных SOA-технологий комплексные проекты, интегрирующие в единый информационный ресурс любые бизнес-приложения предприятия-заказчика.

В компании накоплен уникальный многолетний опыт внедрения проектов, основанный на знании интеграционных технологий Oracle и глубоком понимании бизнес-задач предприятий различных отраслей экономики. В числе заказчиков – ОАО «СИБУР Холдинг», Федеральное агентство по атомной энергии, АО «Народный банк Казахстана».

**Подробнее о Консалтинговой группе «Борлас»:**  
<http://www.borlas.ru>

**Адрес:** 117105, г. Москва, Новоданиловская наб., д. 4а

**Тел.:** +7 (495) 545-5930, **факс:** +7 (495) 545-5931

**E-mail:** [integration@borlas.ru](mailto:integration@borlas.ru)

### «Форс – центр разработки»

Компания «ФОРС – Центр разработки» – Сертифицированный мастер-партнер (Certified Advantage Partner) и официальный дистрибьютор (Value Added Distributor) корпорации Oracle. «ФОРС – Центр разработки» является Центром компетенции по направлениям «Oracle Fusion Middleware: Business Intelligence and Data Warehousing» и «Oracle Fusion Middleware: Portal».

Выполнение проектов, связанных с построением хранилищ данных и аналитических систем, является стратегическим направлением деятельности «ФОРС – Центр разработки». Специалистами компании реализован целый ряд проектов для крупных государственных и коммерческих организаций (в том числе, энергетического и нефтедобывающего секторов экономики).

В компании работают выделенные подразделения по технологическому направлению Business Intelligence (Лаборатория Решений ФОРС, отдел аналитических систем), что позволяет разрабатывать эффективные и качественные решения на основе технологий Oracle Database, Oracle Business Intelligence (Standard Edition, Enterprise Edition), OLAP, Data Mining.

**Подробнее о компании «Форс – центр разработки»:**

**<http://www.fdc.ru>**

**Россия, 129272 Москва, Трифоновский тупик, дом 3.**

**Тел.: (495) 787–7040, факс: (495) 787–7047**

**E-mail: [develop@fors.ru](mailto:develop@fors.ru)**



### Компания TopS Business Integrator (TopS BI)

TopS BI предоставляет комплексные услуги построения информационно-аналитических систем (ИАС) поддержки принятия управленческих решений на основе продуктов Oracle Business Intelligence, включая управленческий консалтинг в части разработки BSC, KPI и систем управленческой отчетности; создание систем формирования управленческой отчетности и KPI; разработку и внедрение систем класса «APM Руководителя»; разработку и внедрение ИАС на базе технологии хранилищ данных; интеграцию с существующими приложениями заказчика; консультации по проектированию, разработке и внедрению ИАС, настройке и установке продуктов Oracle; проведение независимой экспертизы ИТ-решений в области технологий хранилищ данных и аналитических систем.

TopS BI имеет собственный Демонстрационный центр, сертифицированных специалистов и множество наработок по созданию информационно-аналитических систем на базе продуктов Business Intelligence.

TopS BI имеет опыт создания информационно-аналитических систем в следующих отраслях: черная металлургия, пищевая промышленность, торговля и дистрибуция.

**Подробнее о компании TopS Business Integrator можно найти по адресу <http://www.topsbi.ru/>**

**Главный офис: Россия, 117342, Москва, ул. Новорязанская, 31/7  
Тел.: (495) 797–9966, факс: (495) 797–9967**

**Департамент электронного бизнеса и заказных разработок:  
Россия, 115598, Москва, ул. Загорьевская, 10/4  
Тел.: (495) 777–66–08, факс: (495) 329–85–64**

**E-mail: [e-business@topsbi.ru](mailto:e-business@topsbi.ru)**



## BUSINESS INTELLIGENCE SUITE

### КОРПОРАЦИЯ ORACLE

Oracle Россия  
119435, Москва  
Саввинская набережная, 15  
Тел.: +7 (495) 641 1400  
Факс: +7 (495) 641 1414  
Email: oracle\_ru@oracle.com  
Internet: www.oracle.com/ru/

191186, Санкт-Петербург  
Невский пр., 25  
Тел.: +7 (812) 363 3257  
Факс: +7 (812) 363 3258  
Email: oracle\_ru@oracle.com  
Internet: www.oracle.com/ru/

Oracle Украина  
04070, Киев  
ул. Фроловская, 911  
офисный центр «Swiss House»  
Тел.: +380 (44) 490 9050  
+380 (44) 490 9051  
Факс: +380 (44) 490 9052

Oracle Казахстан  
480099, Алматы  
микрорайон Самал2,  
Самал Тауэрс, оф. 97, блок А2, 6-й этаж  
Тел.: +7 (727) 258 4748  
Факс: +7 (727) 258 4744

Copyright © 2007 Oracle Corporation. Все права защищены.

Данный документ предоставлен исключительно в информационных целях и его содержание может быть изменено без уведомления. Этот документ не гарантирует отсутствие ошибок и не подразумевает никаких гарантий или условий, выраженных явно или подразумеваемых законом, включая косвенные гарантии и условия окупаемости или пригодности для решения конкретной задачи. Мы отказываемся от любой ответственности, связанной с этим документом, и никакие договорные обязательства не могут быть оформлены, прямо или косвенно, на основании данного документа. Этот документ не может быть воспроизведен или передан в любой форме и любыми средствами, электронными или механическими, для любых целей, без нашего письменного разрешения. Oracle, JD Edwards, PeopleSoft и Retek являются зарегистрированными товарными знаками корпорации Oracle и/или входящих в нее компаний. Другие наименования могут быть товарными знаками соответствующих владельцев.