

streamlined, and understandable for managers. The data warehouse is an assembly line for preparing information in an integrated, consistent, visual form to support management decision making [6, p. 26].

Creation of information and analytical systems that really meet the goals and objectives of organizations is a rather complex process, including the stages of the formation of concepts, design, development, implementation and maintenance. Thus, a general method of creating information and analytical systems is needed, containing the composition and sequence of work and tasks, the composition of role functions and the documents generated.

It is obvious that intellectual technologies reveal new ways to improve the quality of services in the conditions of the modern information society. So an adaptive presentation provides an individual approach, support in solving problems and intelligent analysis of solutions with interactive communication processing can save considerable time, model selection technologies can enhance managerial and communicative aspects.

### References

1. The history of the development of data mining methods/ Data Mining. Internet resource. Access mode. URL: <http://azfor.ucoz.ru/publ/3-1-0-3>.
2. Marina Chapot. Intelligent data analysis in decision support systems. Internet resource. - Access mode. URL: <http://www.osp.ru/os/1998/01/179360/>
3. Dyuk V. A., Samoilenko A. P. DataMining: training course. SPb. : Peter. 2011.
4. B. de Ville. Microsoft Data Mining. DigitalPress. 2011.
5. Han J., Kamber M. Data Mining: Concepts and Techniques // Morgan Kaufmann, 2010.
6. Mikhail Petrovskiy. Fuzzy Kernel-based Method for Real-time Network Intrusion Detection // Springer-Verlag, Lecture Notes in Computer Science. 2010. Vol. 2887. P. 189–200.
7. Igor Mashechkin, Mikhail Petrovskiy and AndreyRozinkin. Enterprise Anti-spam Solution Based on Machine Learning Approach // Proceedings of 7th International Conference on Enterprise Information Systems. USA. Miami. 2010. Vol. 2. P. 188–193.

УДК 334.78

## **ОБЗОР РЕШЕНИЙ ОСНОВНЫХ ПРОИЗВОДИТЕЛЕЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РАЗРАБОТКИ И ОБРАБОТКИ БОЛЬШИХ ДАННЫХ**

*Т. Л. Тен, В. Г. Дрозд, Б. Ж. Спанова*

*Карагандинский экономический университет Казпотребсоюза  
(Республика Казахстан)*

В работе рассмотрены особенности использования технологии Microsoft Reporting Services для предоставления аналитической информации конечным пользователям на базе реализации модуля формирования и доступа к выходным документам посредством возможностей служб Reporting Services. В данной публикации дается обзор решений основных производителей программного обеспечения для разработки ХД. При изложении материала используется, по возможности, следующая схема: название проекта компании и его цель; архитектурные решения; СУБД и используемая модель данных; возможности языка обработки данных; степень охвата жизненного цикла (анализ – проектирование – реализация – поддержка); возможные конкурентные преимущества.

**Ключевые слова:** консолидация информации, технологии анализа данных, формирование отчета, обработка данных, шаблон.

The paper discusses the specifics of using Microsoft Reporting Services technology to provide analytical information to end users based on the implementation of the module for generating and accessing output documents through the capabilities of Reporting Services. This publication provides an overview of the solutions of the main software vendors for CD development. In the presentation of the material, the following scheme is used, if possible: the name of the project of the company and its purpose; architectural solutions; DBMS and the data model used; the capabilities of the data processing language; degree of coverage of the life cycle (analysis - design - implementation - support); possible competitive advantages.

**Keywords:** information consolidation, data analysis technologies, report generation, data processing, template.

На сегодняшний день во многих организациях накапливается огромное количество информации. Для обработки этих данных необходимо применять различные методы, подходы, в зависимости от поставленной задачи и типа имеющихся данных. Универсальных методов и технологий для обработки больших объемов данных и получения из них знаний не существует.

Применение программно-аппаратных комплексов и специальных видов оборудования для обеспечения хранения больших объемов данных.

Компания IBM реализует свое решение под названием DWP - Data Warehouse Plus. В области создания, разработки и поддержки систем складирования данных, целью данной компании является обеспечение пользователя в рамках единой архитектуры интегрированным набором программных сервисов и продуктов.

Компания IBM предоставляет встроенную поддержку трех типовых направлений архитектурных решений для хранилищ данных:

- тип взаимосвязанных киосков данных;
- тип независимых киосков данных;
- тип глобального ХД.

Для ХД несущей СУБД выступает семейство объектно-реляционных СУБД DB2, где SQL является языком манипулирования данными.

Преимущество решений компании IBM демонстрирует себя и в том случае, когда и ХД и системы оперативной обработки данных находятся на ПО IBM, т.е. реализуется так называемое замкнутое типовое решение.

На данном основании основными функциями информационно-аналитической системы выступают:

- Обеспечение хранения данных.
- Нахождение и извлечение данных с разных источников, их загрузка в хранилище и преобразование.
- Проведение анализа данных, включая создание регламентированных отчетов, формирование произвольных запросов, извлечение знаний (data mining) и многомерный анализ (OLAP).

Для реализации этих функций, обычно, применяются различные продукты, что в свою очередь ведет к усложнению архитектуры системы, по-

требности к интеграции различных инструментальных сред, излишним дополнительным затратам на работы по администрированию, проблемам согласования метаданных и данных на различных серверах.

Новый подход к созданию и применению аналитических систем предлагает корпорация Oracle. Она основывается на единой и функционально полной платформе для предоставления решения всех перечисленных ранее задач [4].

Oracle9i Database, как система управления базами данных, выступает основой решения, с помощью которой можно не только надежно и безопасно обеспечивать хранение больших объемов аналитической информации, но также более эффективно реализовывать процедуры нахождения и извлечения данных из различных источников данных, согласовывать, преобразовывать и агрегировать эти данные в аналитическую информацию, обеспечивать ее загрузку в хранилище данных. Средствами данного же продукта, кроме того, могут быть поддержаны другие различные методы анализа данных, включая прогнозирование, многомерный анализ, поиск закономерностей. Все эти функции могут быть реализованы специальными компонентами Oracle9i, описание которых приведено ниже:

Компонент ETL, выступает как расширение стандартных средств СУБД Oracle, обеспеченное дополнительными средствами и командами, полезными для реализации задач сбора и преобразования данных. К подобным средствам следует отнести автоматическую фиксацию изменения данных (СНС - change data capture), наличие внешних таблиц, присутствие табличных функций, ввод данных в несколько таблиц, выполнение одновременного ввода и корректировки данных и др. [5].

Технология data mining реализуется на базе средств опции Oracle9i Data Mining, где в имеющихся больших объемах информации становится возможным автоматически находить взаимосвязи и полезные решения для выполнения управленческих задач.

Предлагаемая Oracle концепция построения систем поддержки принятия решений, объединяет все требуемые компоненты, которые необходимы для создания, поддержания работоспособности и управления Хранилищем Данных, а также для применения собранной в хранилище информации.

На изображенном рисунке 1 приведен набор программных средств Oracle, реализующих вышеописанные задачи.

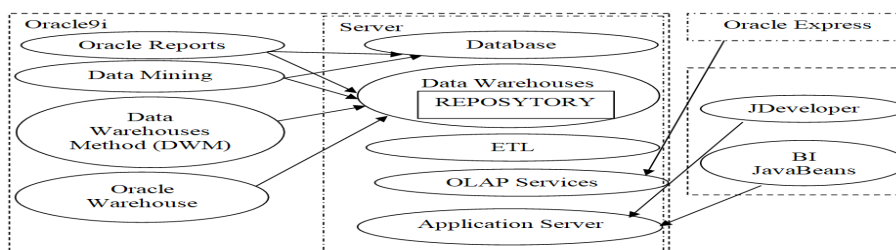


Рис. 1. Набор программных средств Oracle, реализующих технологию работы с ХД

Компания SAS Institute позиционирует себя как поставщик полного решения по предоставлению организации хранилищ данных (ХД). Она распространяет методологию Rapid Data Warehousing (RDW) для обеспечения быстрого создания и быстрого наполнения ХД. В качестве базовой основы этой методологии положено:

- манипулирование и преобразование данными в рамках Data Step (4GL);
- предоставление доступа к данным в хранилище данных с возможностью интероперабельности, т.е. их извлечения из различных источников данных;
- наличие огромного спектра программных продуктов компании для статистического анализа и аналитической обработки данных;
- обладание компанией собственным сервером многомерных БД.

Компания NCR ориентирует свое решение в области хранения данных в направлении самих организаций, у которых присутствует необходимость в системах OLAP и системах поддержки и принятия решений (DSS). Предлагаемая архитектура называется виртуальное предприятие (EIF - Enterprise Information Factory).

Самой «интеллектуальной» из аналитических платформ выступает - Oracle BI Suite EE. Аналитики Gartner, еще совсем недавно, платформу SA (Siebel Analytics) включали лишь в группу «Провидцы» своего «магического квадрата» - MQBIP (Magic Quadrant for Business Intelligence Platforms). Тем самым, отмечая технологические достоинства этой платформы, они указывали на низкую оценку стратегии данной компании по ее продвижению.

В целом же, следует отметить, что положенные в основу принципы, заложенные в архитектуре Oracle BI EE, дают возможность разработчику получить единый взгляд и единую модель представления всей корпоративной информации, которая может содержаться в различных системах. Разработка всего BI-решения, в соответствии с этим, упрощается, а главное снижаются финансовые издержки. Также еще одной важной для разработчика стороной архитектуры выступает доступ к информации в режиме реального времени или через многоуровневую систему кеширования. Для сопровождения системы и администрирования важным является то, что она обладает общими инструментарием администрирования и построена на единой инфраструктуре.

#### Список литературы

1. Ларри Урман. MySQL. Перевод с английского Слинкина А. А. ДМК Пресс. СПб. : Питер. 2004.
2. Миронов А. А., Мордвинов В. А., Скуратов А. К. Семантико-энтропийное управление OLAP и модели интеграции OLAP в SemanticNET (ONTONET) // Информатизация образования и науки. 2009. № 2. С. 21–30.
3. Кудрявцев Ю. А. OLAP технологии: обзор решаемых задач и исследований // Бизнес-информатика. 2008. № 1. С. 66–70.
4. Reporting Services technology, Reporting Services Overview, samples. 2006. Mode of access. URL: <http://www.microsoft.com/sql/reporting/default.asp>
5. Архипенков С. Аналитические системы на базе Oracle Express OLAP. – М. : Диалог-МИФИ, 1999. 320 с.