

КОМПОНЕНТЫ ТИПИЧНОГО ХРАНИЛИЩА ДАННЫХ

Горегляд В. В.

*Белорусский государственный университет информатики и радиоэлектроники, Минск,
Республика Беларусь*

В статье рассмотрено разделение хранилища данных на уровни для более удобного и правильного понимания процессов при проектировании системы хранения и обработки данных.

Ключевые слова: хранилище данных, ETL, процессы, данные

Система Extract-Transform-Load (ETL) является основой хранилища данных. Правильно спроектированная система ETL извлекает данные из исходных систем, обеспечивает стандарты качества и согласованности [1]. Система ETL создает или разбивает хранилище данных на два уровня: уровень подготовки данных и уровень доступа к данным.

Уровни хранилища физически, логически и административно разделены. Другими словами, в большинстве случаев они находятся на разных машинах, зависят от разных структур данных и управляются разными ИТ-специалистами [1]. Для конечных пользователей доступ к данным разрешен только на втором уровне.

На рисунке 1 показаны два отдельных компонента типичного хранилища данных.

Подготовка данных, часто называемая управлением данными, предназначена для извлечения информации из источников данных и подготовки ее к процессу преобразования.

Процедуру извлечения данных можно реализовать двумя способами:

- извлечение данных с помощью специализированных программных средств;
- извлечение данных средствами той системы, в которой они хранятся.

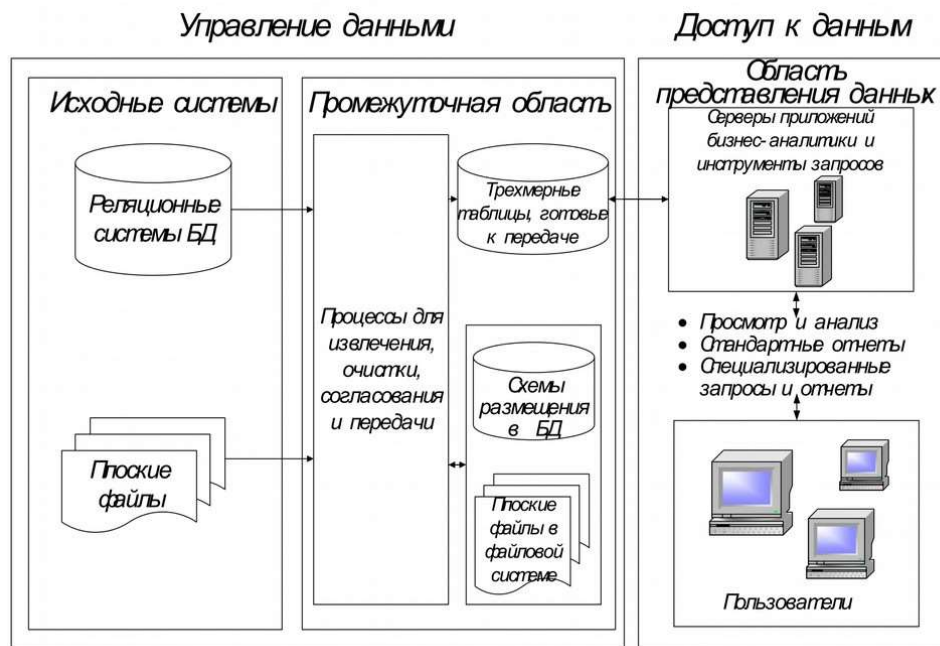


Рисунок 1. Уровень подготовки и уровень доступа к данным

После извлечения данные помещаются в так называемую промежуточную область, где они очищаются, обрабатываются и готовятся к загрузке в область представления. Доступ к данным осуществляется через приложения бизнес-анализа и конкретные запросы конечных поль-

зователей.

Проектирование и разработка ETL-процесса является одной из самых важных задач проектировщика хранилища данных [2].

Список цитируемой литературы:

1. Ralph Kimball. The data warehouse ETL toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data / Ralph Kimball, Joe Caserta // John Wiley & Sons. – 2004. – P. 30-36.
2. Туманов, В.Е.. Проектирование хранилищ данных для систем бизнес-аналитики. Учебное пособие / В. Е. Туманов // Интернет-Университет Информационных Технологий: Бином. Лаборатория знаний. – 2010. - С. 328-333.

COMPONENTS OF A TYPICAL DATA WAREHOUSE

Goreglyad V. V.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

The article considers the division of the data warehouse into layers for a more convenient and correct understanding of the processes in the design of the data storage and processing system.

Keywords: data warehouse, ETL, processes, data