

6.4. Информационные хранилища

Применение технологии "клиент-сервер" не дает желаемого результата для анализа данных и построения систем поддержки и принятия решений. Это связано с тем, что базы данных, которые являются основой технологии "клиент-сервер", ориентированы на автоматизацию рутинных операций: выписки счетов, оформления договоров, проверки состояния склада и т. д. , и предназначены, в основном, для линейного персонала.

Для менеджеров и аналитиков требуются системы, которые бы позволяли:

- анализировать информацию во временном аспекте;
- формировать произвольные запросы к системе;
- обрабатывать большие объемы данных;
- интегрировать данные из различных регистрирующих систем.

Решением данной проблемы стала реализация технологии информационных хранилищ (складов данных).

Основное назначение информационного хранилища - информационная поддержка принятия решений, а не оперативная обработка данных. Технология информационного хранилища обеспечивает сбор данных из существующих внутренних баз предприятия и внешних источников, формирование, хранение и эксплуатацию информации как единой, хранение аналитических данных в форме, удобной для анализа и принятия управленческих решений.

Внутренние базы - локальные базы функциональных подсистем предприятия:	Внешние базы - базы, содержащие сведения других предприятий и организаций:
<ul style="list-style-type: none">• базы бухгалтерского учета;• базы финансового учета;• базы кадрового учета и т. д.	<ul style="list-style-type: none">• базы предприятий-конкурентов;• базы правительственных и законодательных органов и др.

Основные отличия локальной базы данных от информационного хранилища представлены в [табл. 6.4](#).

Элемент отличия	База данных	Информационное хранилище
Данные, содержащиеся в системе	Оперативные данные организации	Внутренние данные организации, внешние данные других источников
Модели данных	Поддерживается одна модель данных	Поддерживается большое количество моделей данных
Выполняемые запросы	Запросы по оперативным данным предприятия, отражающим ситуацию на настоящий момент времени	Оперативные и ретроспективные запросы, содержащие данные предприятия и внешних организаций как на настоящий момент времени, так и за

Принцип, положенный в основу технологии информационного хранилища, заключается в том, что все необходимые для анализа данные извлекаются из нескольких локальных баз, преобразуются посредством статистических методов в аналитические данные, которые помещаются в один источник данных - информационное хранилище.

В процессе перемещения данных из локальной базы данных в информационное хранилище выполняются следующие преобразования:

- **очищение данных** - устранение ненужной для анализа информации (адреса, почтовые индексы, идентификаторы записей и т. д.);
- **агрегирование данных** - вычисление суммарных, средних, минимальных, максимальных и других статистических показателей;
- **преобразование в единый формат** - производится в том случае, если одинаковые по наименованию данные, взятые из разных внешних и внутренних источников, имеют разный формат представления (например, даты).
- **согласование во времени** - приведение данных в соответствие к одному моменту времени (например, к единому курсу рубля на текущий момент).

Технология помещения данных в информационное хранилище

Данные, содержащиеся в информационном хранилище, обладают следующими свойствами:

1. Предметная ориентация	Данные организованы в соответствии со способом их представления в предметных приложениях
2. Целостность	Данные объединены едиными наименованиями, единицами измерения и т. д.
3. Отсутствие временной привязки	В отличие от локальных баз данных в информационном хранилище содержатся данные, накопленные за большой интервал времени (года и десятилетия)
4. Согласованность во времени	Данные приведены к единому моменту времени
5. Неизменяемость	Данные в информационных хранилищах не обновляются и не изменяются, они считываются из различных источников и доступны только для чтения

Существует три вида информационных хранилищ:

- витрины данных;
- информационные хранилища двухуровневой архитектуры;
- информационные хранилища трехуровневой архитектуры.

Витрины данных - это небольшие хранилища с упрощенной архитектурой. Витрины данных строятся без создания центрального хранилища, при этом информация поступает из локальных баз данных и ограничена конкретной предметной областью, поэтому в разных витринах данных информация может дублироваться. При построении витрин используются основные принципы построения хранилищ данных, поэтому их можно считать хранилищами данных в миниатюре. Принцип построения витрины данных

Информационные хранилища двухуровневой архитектуры характеризуются тем, что данные концентрируются в одном источнике, к которому все пользователи имеют доступ. Таким образом, обеспечивается возможность формирования ретроспективных запросов, анализа тенденций, поддержки принятия решений. Принцип построения информационного хранилища двухуровневой архитектуры

Информационные хранилища трехуровневой архитектуры

На первом уровне располагаются разнообразные источники данных - локальные базы данных, справочные системы, внешние источники (данные информационных агентств, макроэкономические показатели). Второй уровень содержит центральное хранилище, куда стекается информация от всех источников с первого уровня, и, возможно, оперативный склад данных, который не содержит исторических данных и выполняет две основные функции:

- источник аналитической информации для оперативного управления;
- подготовка данных для последующей загрузки в центральное хранилище.

Под подготовкой данных понимают их преобразование и проведение определенных проверок. Наличие оперативного склада данных необходимо при различном регламенте поступления информации из источников. Третий уровень представляет собой набор предметно-ориентированных витрин данных, источником информации для которых является центральное хранилище данных. Именно с витринами данных и работает большинство конечных пользователей.