

Технология разработки баз данных

*СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ НАУКИ И
ТЕХНОЛОГИЙ ИМЕНИ АКАДЕМИКА М.Ф. РЕШЕТНЕВА
АЭРОКОСМИЧЕСКИЙ КОЛЛЕДЖ*

Организация данных в базах данных

Причины, обусловившие появление баз данных

В истории вычислительной техники можно проследить развитие двух основных областей ее использования.

Первая область – применение вычислительной техники для выполнения сложных расчетов.

Развитие этой области способствовало:

- интенсификации методов численного решения сложных математических задач;
- появлению языков программирования, ориентированных на удобную запись численных алгоритмов;
- становлению обратной связи с разработчиками новых архитектур ЭВМ.

Характерной особенностью данной области является наличие сложных алгоритмов обработки информации, которые применяются к простым по структуре данным, объем которых сравнительно невелик.

Вторая область – это использование средств вычислительной техники в автоматизированных информационных системах (АИС).

Организация данных в базах данных

Информационная система представляет собой программно-аппаратный комплекс, обеспечивающий выполнение следующих функций:

- надежное хранение информации в памяти компьютера;
- выполнение специфических для данного приложения преобразований информации и вычислений;
- предоставление пользователям удобного и легко осваиваемого интерфейса.

Обычно такие системы работают с большими объемами информации, имеющими достаточно сложную структуру.

Примеры информационных систем:

- банковские системы,
- автоматизированные системы управления предприятиями,
- системы резервирования мест в гостиницах, авиационных и железнодорожных билетов,
- информационно-поисковые системы и т.п.

Организация данных в базах данных

Файловая организация данных в АИС, ее недостатки

В 60-е годы до появления ЭВМ третьего поколения обработка данных осуществлялась в основном при помощи операций ввода-вывода. Файлы при такой обработке были организованы последовательным способом. При этом физическая структура данных и логическая структура файла совпадали.

Пользователям файл в этой ситуации представляется как линейная последовательность записей, где структура записи файла известна только программе, которая с ним работает, поскольку эта структура содержится только в ней.

Для того чтобы извлечь некоторую информацию из файла, необходимо было знать структуру его записи с точностью до бита. При изменении структуры файла требовалось изменять программу. Этот недостаток файловых систем явился толчком к созданию новых систем хранения и обработки данных.

Организация данных в базах данных

Файловая организация данных в АИС, ее недостатки

Так как с одной программой мог работать не один пользователь, то следовало организовывать авторизацию доступа к файлу. Авторизация доступа к файлам (какие действия с файлом может производить тот или иной пользователь) в основном выполняется его создателем-владельцем (программистом).

Для множества файлов, отражающих информационную модель одной предметной области, такой децентрализованный принцип управления доступом вызывал трудности и послужил еще одной причиной разработки систем управления базами данных.

Важным шагом в развитии информационных систем является переход к использованию *централизованных систем управления файлами*. Необходимость обеспечения эффективной работы многих пользователей с одним и тем же файлом стала следующей причиной создания нового подхода к управлению данными.

Организация данных в базах данных

Файловая организация данных в АИС, ее недостатки

В файловых системах одновременная работа нескольких пользователей, связанная с модификацией данных в файле, либо вообще не реализовывалась, либо была замедленной.

Эти недостатки послужили тем толчком, который заставил разработчиков информационных систем предложить новый подход к управлению информацией. Он был реализован в рамках новых программных систем, названных впоследствии **системами управления базами данных (СУБД)**.

Системы управления базами данных

Современные технологии баз данных (БД) являются одним из определяющих факторов успеха в любой отрасли бизнеса, обеспечивая хранение корпоративной информации, представление данных для пользователей и поддержку многих других процессов. Помимо того, БД составляют основу разнообразных научных проектов.

Мощь БД основывается на результатах исследований и технологических разработок, полученных на протяжении нескольких последних десятилетий, и заключена в специализированных программных продуктах, которые принято называть **системами управления базами данных (СУБД) (*Database Management Systems – DBMS*)**.

СУБД – это эффективный инструмент сбора больших порций информации и действенного управления ими, позволяющий сохранять данные в целостности и безопасности на протяжении длительного времени.

СУБД относятся к категории наиболее сложных программных продуктов, имеющих на рынке в настоящее время.

Функциональные возможности СУБД:

1. *Средства постоянного хранения данных.* СУБД, подобно файловым системам, поддерживают возможности хранения огромных фрагментов данных. Однако, они значительно превосходят файловые системы в отношении гибкости представления информации, предлагая структуры, обеспечивающие эффективный доступ к большим порциям данных.
2. *Интерфейс программирования.* СУБД позволяет пользователю обращаться к данным и изменять их посредством команд развитого языка запросов.
3. *Управление транзакциями.* СУБД поддерживают параллельный доступ к данным, то есть возможность единовременного обращения к одной и той же порции данных со стороны нескольких различных процессов, называемых **транзакциями** (*transactions*).

Эволюция СУБД

Что такое база данных? По существу, это набор порций информации, существующий в течение длительного периода времени.

Термином **база данных (database)** обозначают набор данных, находящийся под контролем СУБД.

СУБД обязана обеспечить реализацию следующих требований:

1. Позволять пользователям создавать новые БД и определять их схемы (логические структуры данных) с помощью некоторого специализированного языка, называемого **языком определения данных (Data Definition Language – DDL)**.
2. Предлагать пользователям возможности задания запросов (queries) и модификации данных средствами соответствующего **языка управления данными (Data Manipulation Language – DML)**.
3. Поддерживать способность сохранения больших объемов информации на протяжении длительных периодов времени, предотвращая опасность несанкционированного доступа к данным и гарантируя эффективность операций их просмотра и изменения.
4. Управлять единовременным доступом к данным со стороны многих пользователей, исключая возможность влияния действий одного пользователя на результаты, полученные другим, и запрещая совместное обращение к данным, чреватое их порчей.

Первые СУБД

Появление первых коммерческих СУБД датируется концом 1960-х годов. Непосредственными предшественниками таких систем были файловые системы.

К числу первых серьезных программных приложений СУБД относились те, в которых предполагалось, что данные состоят из большого числа элементов малого объема; для обработки таких элементов требовалось выполнять множество элементарных запросов и операций модификации.

Системы бронирования авиабилетов

Элементы данных такой системы:

1. Сведения о резервировании конкретным пассажиром места на определенный авиарейс, включая информацию о номере места;
2. Информация о полетах – аэропортах отправления и назначения для каждого рейса, сроках отправки и прибытия и т.д.
3. Данные о ценах на авиабилеты, поступивших заявках и наличии свободных мест.

В типичных запросах требовалось выяснить, какие рейсы из одного заданного аэропорта в другой близки по времени отправления к указанному календарному периоду, имеются ли свободные места.

К числу характерных операций по изменению данных относились бронирование места на рейс, определение номера места и выбор обеденного меню. В любой момент к одним и тем же элементам данных могли обращаться несколько операторов-кассиров. СУБД обязана была обеспечить подобную возможность, но исключить любые потенциальные проблемы, связанные, например, с продажей нескольких билетов на одно место, а также предотвратить опасность потери записей данных, если система внезапно выйдет из строя.

Первые СУБД оказывались способными представлять результаты запросов практически только в том виде, который соответствовал структуре хранения данных. Одной из основных проблем, препятствовавших распространению и использованию таких систем, было отсутствие поддержки высокоуровневых языков запросов.

Системы реляционных баз данных

После опубликования в 1970 году статьи Э.Ф. Кодда СУБД претерпели существенные изменения. Кодд предложил схему представления данных в виде таблиц, называемых **отношениями** (*relations*).

В отличие от ранних систем, пользователю **реляционной базы данных (РБД)** вовсе не требуется знать об особенностях организации хранения информации на носителе.

Запросы к такой БД выражаются средствами высокоуровневого языка, позволяющего значительно повысить эффективность работы программиста.

Понятие базы данных

В широком смысле слова **база данных** – это совокупность сведений о конкретных объектах реального мира в какой-либо предметной области.

Объектом может быть предмет, вещество, событие, лицо, явление, т.е. все то, что может характеризоваться набором значений некоторой совокупности атрибутов.

Например, телефонный справочник, личные дела сотрудников, библиотечные каталоги.

Предметная область – часть реального мира, например, предприятие, библиотека, вуз и т.д.

Данные могут представляться разными способами: *не структурированно и структурированно.*

Структурирование – это введение соглашений о способах представления данных

Пример 1

Иванов Иван, группа 91001, оценка по дисциплине ОП – 5; Петров Сергей, группа 91001, оценка по дисциплине ОП – 3; Иванова Ира, группа 91001, оценка по дисциплине ОП – 4.

осуществлять поиск необходимых данных, практически невозможно упорядочить информацию.

Пример 2

Фамилия Имя	Группа	Оценка по дисциплине ОП
Иванов Иван	91001	5
Петров Сергей	91001	3
Иванова Ира	91001	4

База данных – это именованная совокупность структурированных данных, относящихся к определенной предметной области.

БД должна обладать определенными свойствами:

- многократное использование;
- быстрый поиск и получение информации по запросам;
- простота обновления;
- уменьшение избыточности данных;
- защита от несанкционированного доступа;
- защита от уничтожения данных;
- максимальная независимость от прикладных программ.

Обеспечивает выполнение этих и других свойств, предоставляет различные информационные услуги пользователям **система управления базами данных.**

Обслуживанием базы данных занимается группа лиц, называемая **администратором БД**.

Функции администратора:

- защита данных;
- сохранение БД;
- разрешение доступа и т.д.

По технологии обработки данных БД подразделяются на централизованные и распределенные:

Централизованная БД хранится в памяти одной вычислительной системы. Если эта вычислительная система является компонентом сети ЭВМ, возможен распределенный доступ к такой базе. Такой способ часто применяется в ЛС (локальных сетях).

Распределенная БД состоит из нескольких пересекающихся или даже дублирующих друг друга частей, хранимых в различных ЭВМ компьютерной сети.

По способу доступа к данным БД разделяются на БД с локальным доступом и БД с удаленным (сетевым) доступом.

Существуют различные архитектуры систем централизованных БД с сетевым доступом: файл-сервер, клиент-сервер.

Файл-сервер. На сервере файлов (центральной машине) хранится БД. Обработка файлов БД производится на рабочих станциях.

Клиент-сервер. На центральной машине (сервере БД) хранятся файлы БД и здесь же выполняется обработка данных. На рабочие станции передаются только извлеченные данные, а не файлы.

Спецификой данной архитектуры является использование языка запросов, например **SQL (*Structured Query Language* – язык структурных запросов)**.



Преподаватель: Яскина А.Г.

© Сибирский государственный университет
науки и технологий имени академика М. Ф.
Решетнева, 2017