

# Базы данных и СУБД

## Основные понятия

**База данных** – это набор порций информации, существующий в течении длительного времени.

Данное понятие очень обширное поэтому термин будет немного усложнен. Будем говорить только о компьютерных базах данных, т.е. таких, взаимодействие с которыми осуществляется посредством специального управляющего программного обеспечения – СУБД.

Для создания нормальной базы данных нужны три основные составляющие:

- сами данные;
- аппаратное обеспечение;
- программное обеспечение.

Под программным обеспечением подразумеваются все те средства, которые позволяют конечным пользователям получать доступ к данным и редактировать их. Кроме того это программное обеспечение может решать и другие задачи, такие как например: обеспечение безопасности данных, одновременный доступ и т.д. Весь этот комплекс программ обычно называют системой управления базами данных – СУБД.

## Требования, предъявляемые к современным СУБД

Требования, предъявляемые к СУБД, можно рассматривать как состав функций, которые подлежат реализации идеальной системой управления базами данных. Помимо основной функции системы управления базами данных – организации хранения и доступа к данным – подлежат реализации следующие функции:

**целостность** и **непротиворечивость** – обеспечение невозможности внесения в базу данных различных значений одной и той же характеристики одного и того же объекта;

**восстанавливаемость** – обеспечение сохранности информации базы данных при любых сбоях аппаратных и программных средств (наиболее распространенными способами реализации данной функции являются автоматическое создание резервных копий баз данных и ведение журналов транзакций);

**согласованность** – обеспечение блокировки доступа к информации в момент ее корректировки в целях невозможности получения частично измененной информации;

**контролируемость** – регистрация всех операций выполняемых с информацией базы данных (возможно обобщение функций контролируемости и восстанавливаемости ведением журнала транзакций);

**безопасность и секретность** – обеспечение невозможности непреднамеренного или умышленного получения, модификации и (или) разрушения информации базы данных;

**перемещение данных (архивация** базы данных) – обеспечение оптимального хранения данных, отвечающего реальной потребности в информации в текущий момент времени;

**совместимость** – обеспечение возможности информационного взаимодействия с другими системами, а также автоматизированного ввода информации внемашинной базы данных;

**сервис** администрации, разработчиков, пользователей – обеспечение удобных средств работы различных категорий пользователей баз данных (дополнительные языковые средства, дизайнеры, генераторы и т.п.).

## Функции СУБД

Современная СУБД должна обеспечивать очень широкий набор функций. Вот некоторые из них:

- поддержка логической модели данных (определение данных и оперирование с ними);
- восстановление данных (транзакции, журналирование, контрольные точки);
- управление одновременным доступом;

- конфиденциальность данных (безопасность с точки зрения несанкционированного доступа);
- самостоятельная оптимизация выполнения операции;
- другие функции (администрирование, статистика, распределение данных и т.д.).

### **Модели данных**

Большинство объектов физического мира невероятно сложны по своей организации. Когда мы пытаемся описать какой-либо из таких объектов мы на самом деле придумываем модель, соответствующую ему в нашем понимании. Если объекты можно поделить на некоторые группы, удовлетворяющие одинаковым моделям, то мы получаем ситуацию, когда внутри базы данных хранятся две группы сущностей:

- описания моделей объектов;
- записи, удовлетворяющие какой-либо из модели и соответствующие различным представителям объектов.

Но бывают ситуации, когда объекты настолько различны, что их нельзя классифицировать. Тогда база данных представляет из себя набор из одних лишь моделей.

### **Структуры баз данных**

В настоящее время наиболее популярны три типа структур баз данных: *иерархические, сетевые и реляционные*.

*Иерархическая модель*, при которой элементы данных включаются в контейнеры различных уровней, а контейнеры также могут быть включены в контейнеры более высокого уровня. Примером иерархической базы данных может служить файловая система MS DOS и Windows: файлы вложены в папки, папки могут быть вложены в другие папки, высшая ступень иерархии - магнитный или компакт-диск. Файлы и папки могут быть созданы заново или скопированы. Поиск файлов и папок осуществляется по их названиям или фрагментам названий, или же визуально по древовидной иерархической структуре, но возможен и поиск файлов по контексту - фрагментам текста.

*Сетевая модель* дополняет иерархическую. В файлы сетевой базы данных вставлены гиперссылки, содержащие адреса файлов или закладок, что позволяет легко открывать документы и их разделы. Для вставки гиперссылок используются редакторы Word или Web-страниц. Гиперссылки широко используются в электронных версиях документов и в электронных информационно-консультационных системах.

В последнее время наибольшее распространение получили *реляционные базы данных*, (от английского relation – отношение, связь), состоящие из *двумерных взаимосвязанных таблиц*; в данном случае таблица определяет взаимные отношения составляющих ее элементов. Однородные данные в таблице (например, оценки студентов по одному из предметов – столбцы таблицы) образуют *поля*, а совокупность данных, относящихся к одному объекту (например, фамилия и все оценки в ведомости – строка таблицы) называется *записью*. Теоретической основой реляционных БД является реляционная алгебра, или алгебра отношений, разработанная математиком Ф.Коддом, основные положения которой были опубликованы им в 1970-х годах.

Наиболее известные и продаваемые реляционные серверные СУБД, используемые в информационных системах больших организаций: Oracle, SQL Server, MySQL, SyBase, Informix.

В первых персональных компьютерах широко использовались различные версии разработанной фирмой Microsoft СУБД dBase, таблицы которой можно было использовать с помощью операционной системы MS DOS и программ на языке Basic. Затем та же фирма выпустила СУБД FoxPro, которая затем была модифицирована под Windows и используется до сих пор. Также известна СУБД Paradox, разработанная фирмой Borland. Поскольку эта фирма разработала средства быстрой разработки приложений Windows – Delphi и C++Builder, то формат данных Paradox стал для них основным.

В пакет MS Office, разработанный фирмой Microsoft, наряду с Word и Excel входит *СУБД Access*. Ее характерное отличие от старых СУБД — полное использование возможностей Windows и хранение в одном файле БД нескольких таблиц, запросов, форм, отчетов, макросов и модулей. *Запрос* позволяет выбрать из одной или нескольких таблиц необходимые данные, соответствующие заданным критериям, и расположить их в заданной последовательности. *Форма* позволяет расположить на экране данные,

относящиеся к одной записи или строке запроса, в наиболее наглядном виде; при этом форму можно использовать для ввода или корректировки табличных данных. Записи можно просматривать через форму последовательно. Представление данных в *отчете* аналогично форме, но в отчет включаются данные из всех записей, он разбит на страницы и полностью подготовлен для печати. *Макрос* представляет из себя набор команд, позволяющих объединить разрозненные операции обработки данных (аналогично BAT-файлу). *Модули* содержат программы на языке Visual Basic, которые могут разрабатываться пользователем для реализации нестандартных процедур при создании приложения.

Для таблицы может быть установлен *Первичный ключ* – поле или совокупность полей, значения которых однозначно идентифицируют записи, хранящиеся в таблице. Вторичными ключами называются конкретные параметры, по которым происходит поиск записей.

Реляционные базы данных, к которым принадлежит Access, состоят из *логически связанных* двумерных таблиц. Типы связей таблиц:

*Один-ко-многим* – связь ключевого поля с неключевым;

*Один-к-одному* – связь ключевого поля с ключевым;

*Многие-ко-многим* – связь неключевых полей, обычно ее ликвидируют, вставляя промежуточную таблицу с двумя связями Один-ко-многим.

Таблица со связанным ключевым полем называется главной, с неключевым – зависимой. Связи обеспечивают:

поиск и выборку информации по разным таблицам;

*целостность данных* — невозможность вносить в зависимую таблицу данные, противоречащие данным в соответствующих ключевых полях главной таблицы;

*каскадное обновление* — изменение данных во всех зависимых таблицах при их обновлении в ключевом поле главной таблицы;

*каскадное удаление* — удаление данных во всех зависимых таблицах при их удалении в ключевом поле главной таблицы.

## Персонал, работающий с базами данных

Основными категориями персонала, выполняющего работу с базами данных являются:

**пользователи** – лица, в интересах которых создается база данных, и потребляющие информацию в процессе своей работы;

**прикладные программисты** – лица, осуществляющие разработку прикладных программ, то есть средств, позволяющих пользователям в автоматизированном режиме получать и обрабатывать информацию, необходимую в процессе работы;

**администратор базы данных** – лицо, несущее ответственность за реализацию требований пользователей по составу, содержанию и качеству информации базы данных, а также за обеспечение защиты информации, в том числе и за разграничение прав доступа.

Отдельной группой лиц, участвующих в работах с базами данных, можно считать **системных программистов** и специалистов по техническому обслуживанию, выполняющих функции по обеспечению корректной работы средств общего программного обеспечения и технических средств системы.