

Управление данными –

управление сбором, накоплением, обработкой,
хранением и передачей данных

История развития

На заре развития вычислительной техники возможности компьютеров по хранению информации были очень ограниченными. Надежное и долговременное сохранение информации после выключения электрического питания реализовалось двумя видами физических устройств внешней памяти — магнитными лентами и барабанами.

Емкость магнитных лент была достаточно велика, но они обеспечивали только последовательный доступ к данным.

Магнитные же барабаны давали возможность произвольного доступа к данным, но имели ограниченный объем хранимой информации. Оба устройства достаточно медленные.

Магнитные диски (сменные) обеспечили повышение объема хранимой информации и скорости ее обработки.

Одна из проблем работы с данными - именование частей данных и структуризации данных во внешней памяти.

Файлы и файловые системы

Файл - линейная последовательность записей над которой необходимо выполнять ряд стандартных операций:

- создать файл (требуемого типа и размера);
- открыть ранее созданный файл;
- прочитать из файла некоторую запись (текущую, следующую, предыдущую, первую, последнюю);
- записать в файл на место текущей записи новую, добавить новую запись в конец файла

Для того, чтобы извлечь некоторую информацию из файла, необходимо точно знать структуру записи файла с точностью до бита. Каждая программа, работающая с файлом, должна иметь у себя внутри структуру данных, соответствующую структуре этого файла. Изменение структуры файла, требовало исправления, перекомпиляции и дополнительной отладки всех остальных программ, работающих с этим файлом

Файловые системы являются общим хранилищем файлов, принадлежащих разным пользователям, поэтому системы управления файлами должны обеспечивать авторизацию доступа к файлам, т.е. для каждого существующего файла указываются действия, которые разрешены или запрещены данному пользователю, иначе каждому зарегистрированному пользователю соответствует пара целочисленных идентификаторов; идентификатор группы, к которой относится этот пользователь, и его собственный идентификатор в группе. Впервые такой подход к защите файлов был реализован в ОС UNIX. Для множества файлов, отражающих информационную модель одной предметной области, такой децентрализованный принцип управления доступом вызывал дополнительные трудности.

Отсутствие централизованных методов управления доступом к информации послужило одной из причин разработки СУБД Следующей причиной стала необходимость обеспечения эффективной **параллельной** работы многих пользователей с одними и теми же файлами.

В системах управления файлами при операции открытия файла указывался режим работы (чтение или изменение). Если к моменту выполнения этой операции некоторым пользовательским процессом PR1 файл был уже открыт другим процессом PR2 в режиме изменения, то в зависимости от особенностей системы процессу PR1 либо сообщалось о невозможности открытия файла, либо он блокировался до тех пор, пока в процессе PR2 не выполнялась операция закрытия файла.

Новый подход к управлению информацией был реализован в рамках программных систем, названных впоследствии Системами Управления Базами Данных (СУБД), а сами хранилища информации, которые работали под управлением данных систем, назвали базами или банками данных (БД и БнД).

История развития баз данных

В истории развития вычислительных систем можно проследить развитие двух основных областей ее использования: численные расчеты и создание автоматических и автоматизированных информационных систем.

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

- ввод, надежное хранение и поиск структурированной информации в памяти компьютера;
- выполнение специфических преобразований информации и вычислений;
- предоставление хранимой информации пользователям в нужном виде;
- предоставление пользователям удобного интерфейса для управления данными ИС.

ИС характеризуется большими объемами информации и сложной структурой.

Этапы развития СУБД

История развития СУБД насчитывает более 40 лет.

В **1968** году была введена в эксплуатацию первая промышленная СУБД система IMS фирмы IBM. В 1975 году появился первый стандарт ассоциации по языкам систем обработки данных Conference of Data System Languages (CODASYL), который определил ряд фундаментальных понятий в теории систем баз данных, которые и до сих пор являются основополагающими для сетевой модели данных.

В **1981** году Э. Ф. Кодд получил за создание реляционной модели и реляционной алгебры престижную премию Тьюринга. Можно выделить четыре этапа в развитии данного направления в обработке данных.

Первый этап связан с организацией БД на больших машинах типа IBM 360/370, ЕС-ЭВМ и мини-ЭВМ типа PDP11 (фирмы Digital Equipment Corporation — DEC), разных моделях HP (фирмы Hewlett Packard).

- Особенности этого этапа развития выражаются в следующем:
- все СУБД базируются на мощных мультипрограммных операционных системах (MVS, SVM, RTE, OSRV, RSX, UNIX), поэтому поддерживается работа с централизованной базой данных в режиме распределенного доступа в пакетном режиме.
 - функции управления распределением ресурсов осуществляются операционной системой (ОС).
 - поддерживаются языки низкого уровня манипулирования данными (ЯМД), ориентированные на навигационные методы доступа к данным.
 - значительная роль отводится администрированию данных.
 - обоснование и формализация реляционной модели данных, создание первой системы (System R).
 - разработка теории оптимизации запросов и управления распределенным доступом к централизованной БД, вводится понятие транзакции.
 - внедрение результатов теоретических исследований в коммерческие СУБД.
 - появление языков высокого уровня для работы с реляционной моделью данных (не стандартизированных).

Этап персональных компьютеров

Особенности этого этапа следующие:

- Все СУБД были рассчитаны на создание БД с монопольным доступом.
- большинство СУБД имели развитый и удобный пользовательский интерфейс. В большинстве существовал интерактивный режим работы с БД как в рамках описания, так и в рамках проектирования запросов.
- большинство СУБД имели развитый и удобный инструментарий для разработки готовых приложений без программирования.
- инструментальная среда состояла из готовых элементов приложения в виде шаблонов экранных форм, отчетов, этикеток (Labels), графических конструкторов запросов,.
- во всех настольных СУБД поддерживался только внешний уровень представления реляционной модели, то есть только внешний табличный вид структур данных.
- при наличии высокоуровневых ЯМД типа реляционной алгебры и SQL в настольных СУБД поддерживались низкоуровневые ЯМД на уровне отдельных строк таблиц.

В настольных СУБД отсутствовали средства поддержки ссылочной и структурной целостности базы данных. Эти функции должны были выполнять приложения, однако скудость средств разработки приложений иногда не позволяла это сделать, и в этом случае эти функции должны были выполняться пользователем, требуя от него дополнительного контроля при вводе и изменении информации, хранящейся в БД.

Наличие монопольного режима работы фактически привело к вырождению функций администрирования БД и в связи с этим — к отсутствию инструментальных средств администрирования БД.

И, наконец, последняя и в настоящий момент весьма положительная особенность — это сравнительно скромные требования к аппаратному обеспечению со стороны настольных СУБД. Вполне работоспособные приложения, разработанные, например, на Clipper, работали на РС 286.

В принципе, их даже трудно назвать полноценными СУБД.

Яркие представители этого семейства — очень широко использовавшиеся до недавнего времени СУБД Dbase (DbaseIII+, DbaseIV), FoxPro, Clipper, Paradox.

Распределенные базы данных

Особенности данного этапа:

- поддержка полной реляционной модели, а именно:
 - структурной целостности — допустимыми являются только данные, представленные в виде отношений РМ;
 - языковой целостности, то есть ЯМД высокого уровня (в основном SQL);
 - ссылочной целостности, контроля за соблюдением ссылочной целостности в течение всего времени функционирования системы, и гарантий невозможности со стороны СУБД нарушить эти ограничения.
- многоплатформенная архитектура, то есть работа на компьютерах с разной архитектурой и под разными операционными системами;
- поддержка многопользовательской работы с базой данных и децентрализованное хранение данных;
- развитие средств администрирования БД с реализацией общей концепции средств защиты данных.

- оптимизация реализаций распределенных БД и работа с распределенными транзакциями и запросами
- все современные СУБД имеют средства подключения клиентских приложений, разработанных с использованием настольных СУБД, и средства экспорта данных из форматов настольных СУБД второго этапа развития.
- разработка ряда стандартов в рамках ЯОД и ЯМД: SQL89, SQL92, SQL99 и технологий по обмену данными между различными СУБД: протокол ODBC (Open DataBase Connectivity), предложенный фирмой Microsoft.
- Разработка концепции объектно-ориентированных БД — ООБД.
- СУБД второго этапа: MS Access 97 и все современные серверы баз данных Oracle 7.3, Oracle 8.4 MS SQL 6.5, MS SQL 7.0, MS SQL 2000, System 10, System 11, Informix, DB2, SQL Base и другие современные серверы БД (несколько десятков).

Перспективы развития СУБД

- новая технология доступа к данным — *интранет*.

Основное его отличие от технологии клиент-сервер состоит в том, что не нужно специализированного клиентского программного обеспечения. Для работы с удаленной базой данных используется стандартный браузер Интернета,

При этом встроенный в загружаемые пользователем HTML-страницы код, написанный обычно на языке Java, Java-script, Perl и других, отслеживает все действия пользователя и транслирует их в низкоуровневые SQL-запросы к базе данных,

-Этот подход используется не только для удаленного доступа к базам данных, но и в локальной сети предприятия.

- для подключения нового пользователя не требуется установка дополнительного клиентского программного обеспечения.

- алгоритмически сложные задачи рекомендуется реализовывать в архитектуре «клиент-сервер» с разработкой специального клиентского ПО.

У каждого из подходов к работе с данными есть свои достоинства и свои недостатки в зависимости от области применения того или иного метода

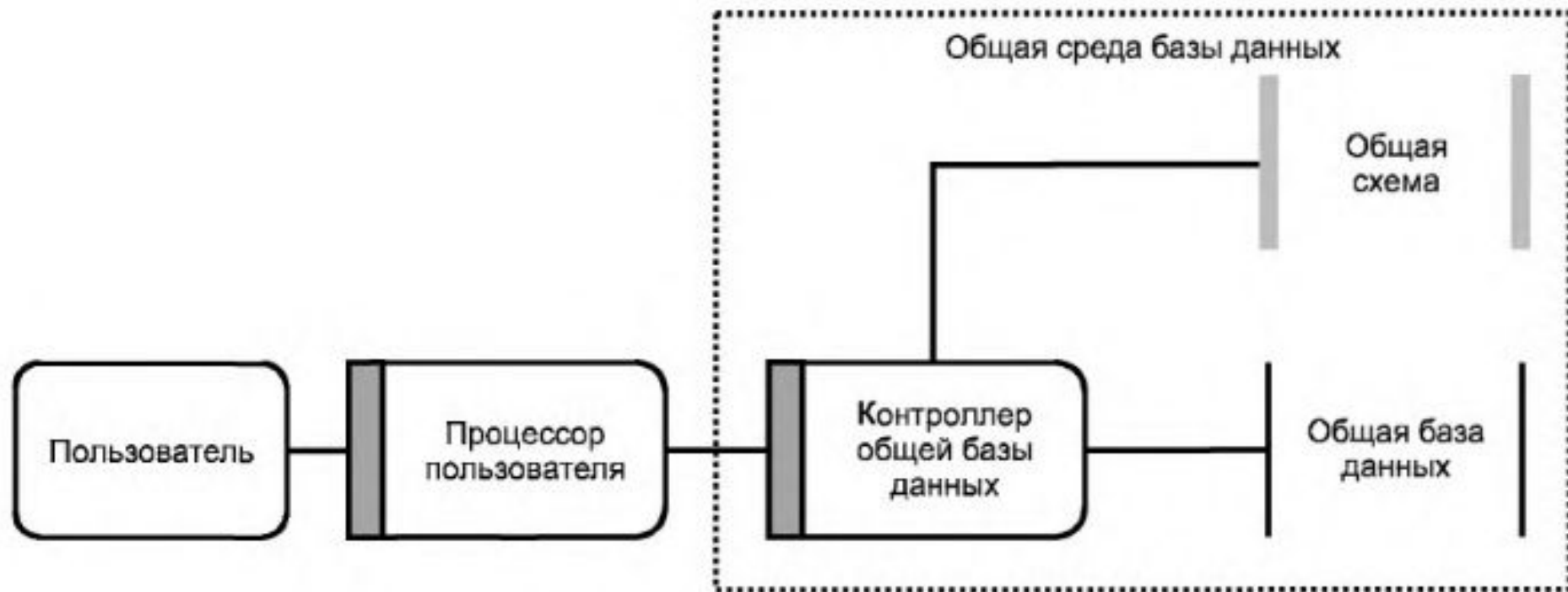


Рис. Общая модель управления данными
(п.6.3 Федеральное агентство по техническому
регулированию и метрологии
Национальный стандарт РФ
ГОСТ Р исо/мэк то 100)

Взаимодействие с базой данных в технологии интранет

