

**База данных** представляет собой совокупность специальным образом организованных данных, хранимых в памяти вычислительной системы и отображающих состояние объектов, и их взаимосвязей в рассматриваемой предметной области.

**Предметная область** — часть реального мира, подлежащая изучению в целях организации управления и автоматизации. Предметная область представляется множеством фрагментов (например, предприятие — цехами, дирекцией, бухгалтерией и т.д.). Каждый фрагмент предметной области характеризуется множеством объектов и процессов, использующих объекты, а также множеством пользователей, характеризующихся различными взглядами на предметную область.

**Структурирование** — это введение соглашений о способах представления данных. Неструктурированными называют данные, записанные, например, в текстовом файле.

*Пользователями базы данных могут быть различные прикладные программы, программные комплексы, а также специалисты предметной области, выступающие в роли потребителей или источников данных, называемые конечными пользователями.*

Из приведенных определений становятся очевидными следующие аспекты:

- БД — это единое хранилище, которое определяется один раз;
- доступ к базе данных могут иметь одновременно несколько пользователей;
- БД содержит не только данные, но и описания этих данных;
- в БД информация должна быть организована так, чтобы обеспечить минимальную долю ее избыточности;
- в БД должны храниться данные, логически связанные между собой.

*Чтобы обеспечить все эти требования, база данных должна обладать определенными свойствами: независимостью данных от использующих их прикладных программ; целостностью; восстанавливаемостью; безопасностью; эффективностью; предельными размерами и эксплуатационными ограничениями.*

1. *Независимость данных от использующих их прикладных программ* означает, что изменения в данных не приводят к изменению программ обработки.
2. *Целостность* базы данных — свойство базы данных, подразумевающее, что БД содержит полную и непротиворечивую информацию, необходимую для корректного функционирования приложений. Для обеспечения целостности накладывают ограничения

в части определения диапазона допустимых значений, соотношения между значениями в полях, особенностей написания формата, а также операций по добавлению, удалению и модификации данных.

3. *Восстанавливаемость* базы данных предполагает возможность восстановления БД после сбоя системы.
4. *Безопасность* базы данных означает защиту данных от преднамеренного и непреднамеренного доступа, модификации или разрушения. Для обеспечения данного свойства используются методы защиты от несанкционированного доступа, защиты от копирования и криптографической защиты.
5. *Эффективность* опирается на два основных критерия: минимальное время реакции и выполнения запроса пользователя и минимальные потребности в памяти.
6. *Предельные размеры и эксплуатационные ограничения.* Предельные размеры базы данных и эксплуатационные характеристики оказывают существенное влияние на проектные решения в части выбора архитектуры, информационной системы, СУБД, программных и технических средств поддержки и сопровождения.

*Создание базы данных, поддержка ее в целостном, непротиворечивом состоянии, предоставление различных информационных услуг пользователям, обеспечение безопасности функционирования и сохранности данных с возможностью восстановления после различных видов сбоев обеспечивается системами управления базами данных.*

**Система управления базами данных** — комплекс программно-аппаратных средств, обеспечивающих создание, поддержку и доступ к БД и управление данными.

Организация типичной СУБД и состав ее компонентов соответствует следующему набору функций СУБД:

- управление данными во внешней памяти;
- управление буферами оперативной памяти;
- управление транзакциями;
- журнализация и восстановление БД после сбоев;
- поддержание языков БД.

Логически в современной реляционной СУБД можно выделить наиболее внутреннюю часть — ядро СУБД (часто его называют *Data Base Engine*), компилятор языка БД (обычно SQL), подсистему поддержки времени выполнения, набор утилит. В некоторых системах эти части выделяются

явно, в других — нет, но логически такое разделение можно провести во всех СУБД.

*Ядро СУБД* отвечает за управление данными во внешней памяти, управление буферами оперативной памяти, управление транзакциями и журнализацию.

Под *транзакцией* понимается неделимая с точки зрения воздействия на БД последовательность операторов манипулирования данными (чтения, удаления, вставки, модификации), приводящая к одному из двух возможных результатов: либо последовательность выполняется, если все операторы правильные, либо вся транзакция откатывается, если хотя бы один оператор не может быть успешно выполнен.

### **Классификация СУБД:**

#### ***1) по модели данных:***

- иерархические;
- сетевые;
- реляционные;
- пост реляционные;
- многомерные;
- объектно-ориентированные.

*Иерархические* СУБД поддерживают иерархическую модель базы данных, которая основана на древовидной структуре хранения информации. В этом смысле иерархические базы данных очень напоминают файловую систему компьютера.

Типовая *сетевая* модель данных была предложена рабочей группой по базам данных (*Data Base Task Group — DBTG*) системного комитета CODASYL (*Conference of Data System Languages*), основными функциями которого были анализ известных фирменных систем обработки управленческих данных с единых позиций и в единой терминологии, обобщение опыта организации таких систем и разработка рекомендаций по созданию соответствующих систем.

*Реляционная* СУБД — это СУБД, управляющая реляционными базами данных. Понятие реляционный (от англ. *relation* — отношение) связано с разработками известного математика Эдгара Кодда (Edgar Codd).

В постреляционной модели СУБД сняты ограничения неделимости данных, хранящихся в записях таблиц, т.е. допускается наличие многозначных полей

В многомерной модели СУБД данные рассматриваются как кубы, которые являются обобщением электронных таблиц на любое число измерений.

*Объектно-ориентированная СУБД* реализует объектно-ориентированный подход. Эта система управления обрабатывает данные как абстрактные объекты, наделенные свойствами, в виде неструктурированных данных, и использующие методы взаимодействия с другими объектами окружающего мира.

**2) по архитектуре организации хранения данных:**

- *централизованные СУБД* (все части СУБД размещаются на одном компьютере);
- *распределенные СУБД* (части СУБД могут размещаться на двух и более компьютерах)

**3) по способу доступа к БД:**

- файл-серверные (Microsoft Access, Visual FoxPro, MySQL версии ниже 5.0);
- клиент-серверные (Interbase, MS SQL Server, PostgreSQL, MySQL версии выше 5.0);
- встраиваемые (OpenEdge, SQLite, BerkeleyDB).

Перечень требований к СУБД, используемых при анализе той или иной информационной системы, может изменяться в зависимости от поставленных целей. Факторы, влияющие на выбор СУБД, характеризуют саму СУБД, программные средства ее окружения, инфраструктуру и особенности предполагаемого применения СУБД.

***Тем не менее можно выделить несколько групп критериев:***

- моделирование данных;
- особенности архитектуры и функциональные возможности;
- контроль работы системы;
- особенности разработки приложений;
- производительность;
- надежность;
- требования к рабочей среде;
- смешанные критерии.

## **Тенденции развития СУБД:**

- 1) поддержка разных стилей проектирования (традиционный, объектно-ориентированный, визуальный);
- 2) использование мультимедийных систем; обработка графической информации; управление пространственной информацией;
- 3) повышение уровня языковых средств;
- 4) включение в СУБД разнообразных языковых средств, ориентированных на разные категории пользователей;
- 5) поддержка решений для реализации хранилищ данных для обеспечения процессов принятия решений;
- 6) наличие разнообразных сервисных средств;
- 7) расширение функциональных возможностей;
- 8) многоплатформенность — наличие нескольких вариантов «одноименных» СУБД, реализованных для разных операционных систем и разных типов компьютеров;
- 9) рост производительности;
- 10) использование более развитых и разнообразных средств обеспечения целостности и безопасности данных;
- 11) «интернационализация» — использование СУБД ведущих производителей в разных странах и поддержка национальных языков;
- 12) работа в многопользовательской среде; распределенность; масштабируемость;
- 13) открытость систем;
- 14) преобладание реляционной модели данных;
- 15) унифицированный язык запросов — SQL;
- 16) развитие объектно-ориентированных и объектно-реляционных СУБД;
- 17) развитие средств упрощения эксплуатации баз данных;
- 18) широкое использование графических интерфейсов; визуализация процессов проектирования и управления системой.

## Жизненный цикл.

**Жизненный цикл БД** – концепция, в рамках которой рассматривается развитие БД во времени. Жизненный цикл БД определяет и жизненный цикл всей информационной системы, поскольку БД является фундаментным компонентом информационной системы. Главной составляющей в жизненном цикле БД является создание единой БД и программ, необходимых



для ее работ жизненный цикл базы данных представляет собой концепцию, в рамках которой рассматривается развитие базы данных во времени.

Жизненный цикл базы данных определяет и жизненный цикл всей информационной системы, поскольку база данных является фундаментальным компонентом информационной системы. Главной составляющей в жизненном цикле базы данных является создание единой базы данных и программ, необходимых для ее работы.

Жизненный цикл базы данных включает в себя следующие этапы (рис. 2.1):

- 1) анализ:
  - а) планирование разработки базы данных;
  - б) определение требований к системе;
  - в) сбор и анализ требований пользователей;

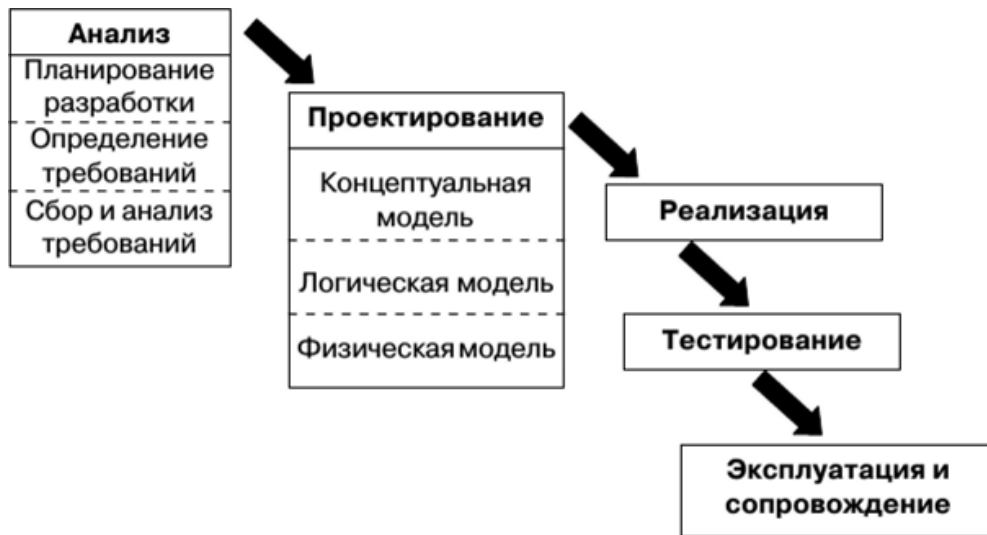


Рис. 2.1. Этапы жизненного цикла базы данных

- 2) проектирование базы данных:
  - а) концептуальное проектирование;
  - б) логическое проектирование;
  - в) физическое проектирование;
- 3) реализация;
- 4) тестирование;
- 5) эксплуатация и сопровождение.

**Планирование разработки базы данных.** На этом этапе, тесно связанном с общей стратегией построения информационной системы организации, определяют три основных компонента: объем работ, объем ресурсов, стоимость проекта. Важной частью разработки стратегического плана является проверка осуществимости проекта, которая включает в себя:

- проверку технологической осуществимости, т.е. наличия соответствующего оборудования и программного обеспечения;
- проверку операционной осуществимости, т.е. оценку достаточности штата экспертов и персонала для работы с базой данных;
- проверку экономической целесообразности осуществления проекта.

**Определение требований к системе.** Этот этап включает в себя выбор целей БД, выяснение информационных потребностей организации и ее структурных подразделений, определение требований к аппаратному и программному обеспечению. На этом этапе важно спрогнозировать архитектуру разрабатываемой информационной системы.

**Сбор и анализ требований пользователей.** Этот этап является предварительным этапом концептуального проектирования базы данных. Необходимо собрать и структурировать информацию по всем информационным потребностям организации. Структурирование информации приводит к созданию модели движения потоков данных и процесса документооборота.

### **Проектирование базы данных.**

Основными целями проектирования базы данных являются:

- представление данных и связей между ними, необходимых для всех основных областей применения разрабатываемой базы данных и любых существующих групп пользователей;
- создание модели данных, способной поддерживать выполнение любых требуемых процессов обработки данных;
- разработка предварительного варианта проекта с учетом требований, предъявляемых к производительности системы.

Данные, используемые для описания предметной области, представляются в виде трехуровневой схемы (так называемая модель ANSI/SPARC) (см. рис. 2.2). Внешнее представление (внешняя схема) данных является совокупностью требований к данным со стороны некоторой конкретной функции, выполняемой пользователем. Концептуальная схема — это полная совокупность всех требований к данным, полученная из пользовательских представлений о реальном мире. Внутренняя схема — это сама база данных.

Трехуровневая модель представления данных

*Рис. 2.2. Трехуровневая модель представления данных*



**Концептуальное проектирование.** На этом этапе проводится глубокий анализ предметной области. Целью анализа является определение: принципов функционирования предметной области; задач, которые будет решать пользователь с помощью базы данных; потоков информации, ее



структуры и взаимосвязей, источников и приемников; форм входных и выходных документов.

### **Формализованное описание** предметной области

называется *концептуальной моделью*. Построение концептуальной модели может выполняться и вручную, и с использованием автоматизированных средств проектирования. Разработанная концептуальная модель данных является источником для этапа логического проектирования базы данных.

**Логическое проектирование.** Цель этого этапа заключается в построении логической модели данных, которая описывает понятия предметной области, их взаимосвязь, а также ограничения на данные, налагаемые предметной областью. Логическая модель данных является начальным прототипом будущей базы данных.

**Физическое проектирование.** Целью проектирования на этом этапе является построение физической модели базы данных. Физическая модель данных описывает данные средствами конкретной СУБД.

**Реализация.** На данном этапе осуществляется физическая реализация базы данных и разработка пользовательских приложений, позволяющих пользователю манипулировать данными в базе данных и выполнять запросы к БД.

**Тестирование.** Тщательное тестирование должен проходить любой программный продукт, но особые требования выдвигаются к тестированию информационных систем в силу специфики их применения.

**Эксплуатация и сопровождение.** Этот этап является самым продолжительным в жизненном цикле базы данных. Основные действия, связанные с этим заключительным этапом, сводятся к наблюдению за созданной системой и поддержкой ее нормального функционирования.

## **Модели данных**

В классической теории баз данных модель данных характеризуется как формальная теория представления и обработки данных в системе управления базами данных, которая включает по меньшей мере три аспекта: аспект структуры (методы описания типов и логических структур данных), аспект манипуляции (методы манипулирования данными), аспект целостности (методы описания и поддержки целостности базы данных). *Аспект структуры* определяет, что из себя логически представляет база данных, *аспект целостности* определяет средства описаний корректных состояний базы данных, *аспект манипуляции* определяет способы перехода между состояниями базы данных и способы извлечения данных из базы данных.