Базы данных.

Ульман Основы систем бд

Видосы по бд мехмат мгу

Предметная область – это область реального мира, использующая конкретную инф систему.

Информация – абсолютно любые сведения о каком-либо событии, сфере, явлении, характерные для рассматриваемой предметной области.

Данные – это информация, зафиксированная в определенной форме, пригодной для последующей хранения, обработки и передачи

Файловая система – набор программ, которые выполняют для пользователя необходимые операции.

Проблемы:

1. Разделение и изоляция данных.
2. Дублирование данных.
3. Зависимости от данных.
4. Не совместимость форматов файлов.
5. Фиксированные запросы.

Факторы:

1. Хранение данных в самой программе
2. Не было инструментов для работы с данными

БД – это совокупность специальным образом организованных данных хранимых в памяти ВС и отображающих состояние объектов, и их взаимосвязей в рассматриваемой предметной области.

БД не может существовать сама по себе нужна СУБД

СУБД – программное обеспечение, с помощью которого пользователи могут определять, создавать и поддерживать БД, а также осуществлять к ней контролируемый доступ.

Строилась на ней две системы

Файл - сервер

Клиент – сервер

Архитектура БД

Фотка 2

Внешний и концептуальный – являются логическими

Внутренний – является физическим

Целью данной 3х уровневой архитектуры является отделение пользовательского преставления БД от ее физического.

Концептуальная модель включает:

* Сущности, атрибуты и связи между собой
* Накладываемы на данные ограничения
* Схематически информация о данных

На внутреннем уровне БД представляется в виде совокупности хранимых файлов, для которых известна структура хранимых записей, определены служебные поля реализующие связи между записями определены методы доступа СУБД к этим записям.

Внешний уровень, который разделяется на несколько преставлений для каждого пользователя имеется свое собственное представление БД и содержит только те сущности, атрибуты и связи между ними, которые нужны конкретному пользователю.

Логическая независимость от данных, то есть полную защищенность внешних схем от изменений концептуальной схемы.

Физическая независимость от данных, то есть защищенность концептуальной модели от изменений вносимых во внутреннею модель.

Функции СУБД

1. Непосредственное управление данными во внешней памяти;
2. Управление буферами ОЗУ;
3. Управление транзакциями;

Транзакция – это последовательность операций над БД рассматриваемых СУБД как единое целое.

1. Журнализация – это особая часть БД не доступная пользователям СУБД и поддерживаемая особой тщательностью, в которую поступают записи обо всех изменениях БД;

Под надёжностью хранения данных будем понимать, что СУБД должна быть в состоянии восстановить последнее согласованное состояние БД после любого программного или аппаратного сбоя.

1. Поддержка языков БД;
2. Поддержка словаря данных – метаданные.

Администратор БД – лицо или группа лиц, отвечающая за выработку требований к БД, ее проектирование, реализацию, эффективное использование и сопровождение. Включая управление учетными записями пользователей БД и защиту от несанкционированного доступа.

Задачи администратора БД:

1. Определение концептуальной модели;
2. Определение внутренней модели;
3. Взаимодействие с пользователями;
4. Определение правил безопасности и целостности данных;
5. Координация всех действий по проектированию БД;
6. Определение процедур резервного копирования и восстановления;
7. Ведение словаря данных;
8. Управление производительностью и реагирование на изменяющиеся потребности.

Лекция 2.

Инфологический аспект проектирования – какой-то логический аспект данных, не зависимый от физического расположения данных.

На этапе инфологического проектирования:

1. О каких объектах, событиях реального мира требуется накапливать и обрабатывать информацию в системе .
2. Какие их хар-ки и взаимосвязи необходимо учитывать в системе
3. Уточняются все вводимые в информационную систему понятия об объектах их характеристиках

Датологической аспект проектирования БД употребляется при рассмотрении вопросов представления данных в памяти инф системы.

При датологическом проектировании учитывая возможности имеющийся средств ввода, хранения и обработки информации разрабатываются соответствующие формы представления информации в системе, а также приводятся модели и методы представления и преобразования данных, формулируются правила их смысловой интерпретации.

Процесс проектирования

Модель сущность-связь — это неформальная модель предметной области, которая используется на инфологическом этапе.

Основным назначением модели является семантическое описание предметной области и представление информации для обоснования выбора видов моделей и структур данных.

Сущность – собирательное понятие, некоторая абстракция реально существующего объекта, процесса или явления, о котором необходимо хранить информацию в системе.

Виды сущностей – слабые и сильные

Слабые – сущность существование которой зависит от какой-то другой. - прямоугольник

Сильная – существование которой не зависит от другой. - Прямоугольник в прямоугольнике

Схема 1 – виды сущностей

Атрибут – поименовав хар-ка сущности, которая принимает значения из некоторого множества.

Доме - Множество значений, которое может принимать атрибут

Атрибуты могут быть:

* Однозначный – атрибут для данной конкретной сущности может принимать только одно значение. (Овал)
* Многозначный - атрибут для данной конкретной сущности может принимать множество значений. (Двойной овал)
* Производный – атрибут, который может принимать значения от другого связанного с ним атрибута. (Пунктирный овал)
* Ключи

Ключ – это элемент данных позволяющий уникально идентифицировать отдельные экземпляры некоторой сущности.

* Потенциальный ключ – уникальный атрибут или набор атрибутов, который идентифицирует сущность
* Составной ключ – потенциальный ключ, который состоит из двух или более атрибутов.
* Первичный ключ – потенциальный ключ, который выбран в виде первичного ключа.

Связь – это осмысленна ассоциация между двумя сущностями. (Ромбик) Если соединяет слабую сущность, то ромбик двойно.

Степень связи – кол-во связей между сущностями.

Показатель кардинальности связи – описывает кол-во возможных связей для каждой из сущности участниц. (1:1 1:N N:N)

1:1 – определяет такой тип связи между двумя сущностями, когда каждому экземпляру сущности А соответствует один и только один экземпляр сущности Б и наоборот.

1:N – определяет такой тип связи между сущностями А и Б, когда каждому экземпляру сущности А может соответствовать один или несколько экземпляров сущности Б.

N:N – определяет такой тип связи сущностей А и Б, при котором каждому экземпляру сущности может соответствовать один или более экземпляров сущности Б.

Модель данных – можно представить в виде:

1. Логическая структура данных – способ организации связей между элементами данных.
2. Набор допустимых операций по манипулированию данными для принятой логической структуры.
3. Правила поддержки целостности или не противоречивости данных для выбранной логической структуры.

Модели данных

1. Иерархическая – на основе древовидная
2. Сетевая – на основе графов
3. Реляционная – на основе таблиц
4. Постреляционная – таблицы с возможностью вложения таблиц в таблицы
5. ООПешная – на основе принципов ООП
6. Документно-ориентированные

**Иерархическая модель**.

В СУБД иерархическая типа информация представлена в виде деревьев, узлами которых являются записи. Корневая запись для каждого экземпляра дерева обязательно должная содержать ключ с уникальным значением. Ключи не корневых записей должны иметь уникальные значения только в рамках данного экземпляра дерева. Каждая запись идентифицируется полным сцепленным ключом, то есть совокупностью ключей всех записей начиная от корневой по иерархическом пути.

Каждый экземпляр дерева называется групповым отношение.

Корневая – владельцем группового отношения, остальные дочерние

Эта модель реализует отношение 1:N, между исходной и дочерней записью.

В этих БД поддерживается целостность связей между владельцами и членами группового отношения.

Недостатки:

* Возникает дублирование информации при реализации отношения N:N
* Не эффективный поиск от узла к корню

**Сетевая модель.**

Состоит из множества записей, которые могут быть владельцами или членами групповых отношений.

Основное различие от иерархической модели состоит в том, что запись может быть членом более одного группового отношения.

Достоинство данной модели является отсутствие дублирования информации.

Недостаток — это ее чрезмерная сложность, что приводит к сложностям администрирования. Например, сложность восстановления поврежденных данных. Медленные некоторые запросы.

**Реляционная модель.**

В реляционной модели достигается более высокий уровень абстракции данных чем сетевой или иерархической.

Реляционная модель представляет средства описания данных на основе только их естественной структуры без потребности введения какой-либо дополнительной структуры для целей машинного представления. То есть представление данных в ней не зависит от способа их физической организации.

Основным логическим объектом для хранения данных является таблица. Для организации связей между данными различных таблиц используются их общие столбцы.

**Пост реляционная модель.**

Содержит в основе своей реляционную модель дополненную возможность создания вложенных таблиц.

Однако использование используется довольно редко т. к. СУБД не гарантирует быструю работу с такими таблицами.

ОО-модель.

Отсутствие строгой математической модели ОО-БД.

Наличие огромного наличие данных в реляционных БД и существенных затраты на их конвертацию.

Многие СУБД на сегодняшний день позиционируются как объектно-реляционные. В их основе лежит реляционная модель, но дополненная возможность создания пользовательских типов столбцов с поддержки принципов инкапсуляцией и наследования.