Лекция 10

Понятие об администрировании БД

Администрирование данных – это управление информационными ресурсами, включая планирование базы данных, разработку и внедрение стандартов, определение ограничений и процедур, а также концептуальное и логическое проектирование БД.

Администратор данных отвечает за корпоративные информационные ресурсы в целом, включая не компьютеризированные.

Администрирование БД – это управление физической реализацией приложений, в том, числе физическое БД и её реализация, организация поддержки целостности и защиты данных, наблюдение за текущем уровнем производительности, а также реорганизация базы по мере необходимости.

Классификация подходов к построению БД

DM = <D, R, A>

D – носитель структуры (заданное множество)

R – конечный набор отношений (типовая характеристика структуры)

A – ограничение, накладываемое на отношение (аксиома структуры)

Понятие модели данных позволяет построить основную классификацию БД по взаимодействию концептуально модели данных (DM) и системы БД.

1. Модель данных отсутствует= невозможно априорно(заранее) построить концептуальную модель для данного массива. Примеры – тексты деловой, научно-технический или художественной прозы, фотографии, другие медиа-данные.

Разумеется, какая-то структура есть, но пока она не поддаётся конструктивному математическому описанию

Варианты хранения:

1.1 Из текста предварительно извлекаются мета-данные, то есть выполняется индексирование, аннотирование, рубрицирование и т.д. Из пункта 1.1 получаются Информационно-поисковые и информационно-справочные системы.

1.2 Мета-данных нет, тексты хранятся целиком \*тексто-графические БД\*

2. Модель данных существует, она мало информативна для системы БД. Пример – гипертекстовые и гипермедийные БД. Здесь имеется некоторая информация о частичном порядке внесённом в совокупность массивов данных, но нет информации об отношениях между элементарными данными внутри массива. Такая модель поддерживается с помощью базы слабоструктурированных данных, например на языках разметки, в частности XML.

3. Модель данных существует, но описана частично на синтаксическом уровне, а частично на семантическом, т.е. требует для машинной реализации креативного участия программиста. Структуры хранения – файловые системы.

4. На синтаксическом уровне существует полноценная модель данных, которая известна поддерживающей структуре системе БД. Это значит, что между реалиями предметной области установлено по крайней мере отношения частичного порядка.

Модель данных поддерживается фактографическими (или дата-логическими) БД.

4.1 Иерархические БД

4.2 Сетевые БД

4.3 Реляционные БД

5. Модель данных построена по объектному принципу.

Если является объектной, то два типа.

5.1 Объектные БД

5.2 Реляционные БД

6. Синтаксическая модель данных есть, но она описана не только энтоксионально (набор фактов), но и интосионально (заданы процедуры логического вывода).

7. Модель данных существует, её основной реалией является отдельный факт. Под фактом понимается привязанная ко времени событие в предметной области, определяемое сравнительно небольшим набором параметров.

7.1 Факты являются справочными данными, доступными только для чтения. Тогда таблицы фактов имеют огромный размер по сравнению с таблицами параметров. Такая конструкция БД называется хранилищем данных. В нём поддерживается небольшой список операции выборки.

7.2 Поддерживается оперативное агрегирование фактических данных по любым параметрам или их набором. В этом случае структура хранения – многомерная модель данных (или OLAP).

Классификация БД по размерности.

1) Однопользовательские и многопользовательские

2) Территориально- централизованные и территориально-распределённые

3) По истории ведения в практику информационных систем.

А) Эпоха до СУБД, только файловая система

Б) БД первого поколения – двухуровневые, иерархические, сетевые

В) БД второго поколения - реляционные

Г) БД третьего поколения – это системы с семантическим моделированием данных, объектные БД, объектно-реляционные БД.

Проблема реляционной модели данных:

• Реляционная модель прекрасно зарекомендовала себя в традиционных бизнес-приложениях \*обработка заказов, учёт складских запасов, банковское дело и т.д.\*. Работает с проблемами в приложениях следующих типов: \*автоматизирование проектирование, автоматизированное производство, автоматизированная разработка ПО, мультимедийные системы, цифровое издательское дело, геоинформационные системы\*

Причины:

1) неадекватное представление сущности реального мира. Нормализация создаёт псевдо-сущности, не поддерживаются разные типы связи.

2) Слабая поддержка ограничений целостности и корпоративных ограничений.

3) Требование однородных структур данных. Классический пример – лавинообразное увеличение количества компонентов при описании сложных технических конструкций.

4) Сложности с крупными неструктурированными объектами (BLOB) Как правило они хранятся в БД как ссылка на внешний файл => невозможно модифицировать внутреннюю структуры и использовать защиту.

5) Ограниченный набор операций. Недоступны операции определения расстояния, нахождения пересечений и оценки включения.

6) Проблема разсогласования типов данных

7) Отсутствие поддержки длительных, уникальных и динамически изменяемых транзакций.

8) Не гибкая структура БД.

Основные пути разрешения проблем:

1) Использование чисто объектной модели данных

2) Расширение реляционной модели объектными возможностями