

Содержание

1 Экзамен	2
1.1 Базы данных и системы управления базами данных. Определения, основные функции и классификация	2
1.2 Семантическое моделирование данных	4
1.3 Реляционная модель данных: структурная, целостная, манипуляционная части. Реляционная алгебра. Исчисление кортежей	4
1.4 Теория проектирования реляционных баз данных: функциональные зависимости, нормальные формы	4
1.5 Теория проектирования хранилищ данных. Основные принципы построения. ETL и ELT процессы	4
1.6 Транзакции. Определение, свойства и уровни изоляции транзакций. Неблагоприятные эффекты, вызванные параллельным выполнением транзакции, и способы их устранения. Управление транзакциями и способы обработки ошибок	4
1.7 Блокировки. Определение, свойства, иерархии, гранулярность и взаимоблокировки, алгоритмы обнаружения взаимоблокировок	4
1.8 Журнализация. Операции журнала транзакций и его логическая и физическая архитектуры. Модели восстановления. Метаданные	4
1.9 Безопасность и Аудит. Ключевые понятия и участники системы безопасности. Модели управления доступом	4
1.10 МРР системы. Распределенное и колоночное хранение. Распределенные вычисления, модель MapReduce. Обеспечение отказоустойчивости.	4
1.11 In-Memory базы данных. Преимущества и недостатки. Примеры использования	4
1.12 Инструкции языка описания данных, инструкции языка обработки данных, инструкции безопасности, инструкции управления транзакциями	5
1.13 Объекты базы данных: функции, процедуры, триггеры и курсоры	5
1.14 Оптимизация запроса: индексы, партиционирование, сегментирование	5
1.15 План запроса. Этапы выполнения запроса	5

1 Экзамен

1.1 Базы данных и системы управления базами данных. Определения, основные функции и классификация

База данных — совокупность хранимых операционных данных, используемых прикладными системами некоторого предприятия.

Операционные данные — быстрые данные (чтобы их можно было быстро прочитать).

Можно читать, писать, обновлять

Данные делятся на две большие категории

- 1) Читать быстро.
- 2) Быстро писать.

OLAP — online analytic processing (операционные данные).

OLTP — online transaction processing (перманентные данные).

Таблица 1

OLAP	OLTP
чтение	вставка, удаление, обновление
минимальное время отклика	минимальное время вставки, удаление, обновление

Транзакции — либо все действия, либо никакие действия.

База данных — это самодокументированное собрание интегрированных записей.

- 1) Самодокументированное — из нее понятно, что в ней хранится. Понятно какие объекты и принципы действий. -> журналы.
- 2) Запись — это события которые надо где-то хранить.
- 3) Интегрированных — записи которые имеют некоторую структуру и имеют некоторую структуру.

Любая бд хранит:

- 1) метаданные;
- 2) файлы данных,
- 3) Индексы (indexes), которые представляют связи между данными, а также служат для повышения производительности приложений базы данных.
- 4) Может содержать метаданные приложений (application metadata).

Основные характеристики, требования

- 1) **Неизбыточность данных** — каждое данное присутствует в БД в единственном экземпляре.
- 2) **Совместное использование данных** многими пользователями.
- 3) **Эффективность доступа** к БД - высокое быстродействие, т. е. малое время отклика на запрос.
- 4) **Целостность данных** — соответствие имеющейся в БД информации её внутренней логике, структуре и всем явно заданным правилам.
- 5) **Безопасность данных** — защита данных от преднамеренного или непреднамеренного искажения или разрушения данных.
- 6) **Восстановление данных** после программных и аппаратных сбоев.
- 7) **Независимость данных** от прикладных программ.

Система управления базами данных (СУБД) — приложение, обеспечивающее создание, хранение, обновление и поиск информации в базах данных.

Зачем СУБД

- 1) Управление данными во внешней памяти.
- 2) Управление буферами оперативной памяти.
- 3) Управление транзакциями.
- 4) Журнализация.
- 5) Поддержка языка или языкового пакета (-ов).

Классификация СУБД

- 1) Дореляционные
 - Инвертированные списки (файлы) (Как дерево)
 - Иерархические (что-то похожее на хэш)
 - Сетевые

- 2) Реляционные
- 3) Постреляционные

Архитектура хранения данных

- 1) Локальные.
- 2) Распределенные.
- 3) По способу обращения к данным.
 - Файл серверные.
 - Клиент серверные (PostGress, MSSQL, Oracle, MySQL, Mongo).
 - Встраиваемые (SQLlite).
 - Сервисно-ориентированные (KafkaBD).

- Прочее - time series.

1.2 Семантическое моделирование данных

Определение: Дискретная

1.3 Реляционная модель данных: структурная, целостная, манипуляционная части. Реляционная алгебра. Исчисление кортежей

Свойства

1.4 Теория проектирования реляционных баз данных: функциональные зависимости, нормальные формы

Пусть

1.5 Теория проектирования хранилищ данных. Основные принципы построения. ETL и ELT процессы

Определение

1.6 Транзакции. Определение, свойства и уровни изоляции транзакций. Неблагоприятные эффекты, вызванные параллельным выполнением транзакций, и способы их устранения. Управление транзакциями и способы обработки ошибок

Свойства для $n = 2$

1.7 Блокировки. Определение, свойства, иерархии, гранулярность и взаимоблокировки, алгоритмы обнаружения взаимоблокировок

Определение

1.8 Журнализация. Операции журнала транзакций и его логическая и физическая архитектуры. Модели восстановления. Метаданные

Нет вопроса

1.9 Безопасность и Аудит. Ключевые понятия и участники системы безопасности. Модели управления доступом

Этого вопроса нет

1.10 MPP системы. Распределенное и колоночное хранение. Распределенные вычисления, модель MapReduce. Обеспечение отказоустойчивости.

Пусть

1.11 In-Memory базы данных. Преимущества и недостатки. Примеры использования

Учитывая равенство $P\{Y < y\} = F_Y(y)$, приходим к формуле ??.

1.12 Инструкции языка описания данных, инструкции языка обработки данных, инструкции безопасности, инструкции управления транзакциями

Когда X_1, X_2 являются *независимыми случайными величинами*, то есть их *двумерная плотность распределения*

1.13 Объекты базы данных: функции, процедуры, триггеры и курсоры

1.14 Оптимизация запроса: индексы, партиционирование, сегментирование

Дисперсией случайной величины X называют число

1.15 План запроса. Этапы выполнения запроса

Пусть X — случайная величина.