



УНИВЕРСИТЕТ ИТМО

# Технологии хранения больших данных

Лекция №4. OLTP vs. OLAP, аналитические базы  
данных и эффективное представление данных  
для выполнения аналитических задач

Ирина Алексеевна Радченко  
[iradche@gmail.com](mailto:iradche@gmail.com)



# Введение

## Обзор темы

- Различия между OLTP и OLAP
- Введение в аналитические базы данных
- Важность эффективного представления данных

## Цели лекции

- Понимание различий между OLTP и OLAP
- Ознакомление с аналитическими базами данных
- Изучение методов представления данных для аналитических задач



# OLTP (Online Transaction Processing)

## Определение

- Системы, ориентированные на выполнение транзакций

## Характеристики

- Высокая производительность
- Поддержка множества коротких онлайн-транзакций
- Реальное время ответа

## Примеры

- Банковские системы, системы бронирования, онлайн-магазины



# Пример OLTP-системы

## Структура базы данных

- Нормализация данных

## Пример запроса на SQL

```
INSERT INTO Orders (OrderID, CustomerID, OrderDate) VALUES (12345, 67890, '2024-07-19');
```



# OLAP (Online Analytical Processing)

## Определение

- Системы, ориентированные на выполнение аналитических запросов

## Характеристики

- Анализ больших объемов данных
- Сложные запросы, объединения, агрегации
- Долгое время выполнения запросов допустимо

## Примеры

- Системы бизнес-аналитики, системы отчетности



# Пример OLAP-системы

## Структура базы данных

- Денормализация данных (звездообразная и снежинка схемы)

## Пример запроса на SQL

```
SELECT ProductID, SUM(SalesAmount) FROM Sales GROUP BY ProductID;
```



# Сравнение OLTP и OLAP

## Ключевые различия

- OLTP: Транзакционные операции, реальное время
- OLAP: Аналитические операции, долгосрочные данные

## Критерии выбора

- Тип нагрузки, объем данных, требования к производительности



# Аналитические базы данных

## Определение и цель

- Оптимизированы для аналитических запросов и отчетности

## Примеры

- Amazon Redshift, Google BigQuery, Snowflake

## Особенности

- Масштабируемость, высокая производительность, поддержка сложных запросов





# Архитектура аналитических баз данных

## Компоненты

- Хранилище данных (Data Warehouse)
- Хранилище данных в реальном времени (Real-Time Data Warehouse)
- Инструменты ETL (Extract, Transform, Load)

## Процессы

- Загрузка данных, трансформация данных, создание отчетов



# Пример аналитической базы данных

## Структура и схема

- Пример звездообразной схемы

## Пример запроса на SQL

```
SELECT Customer.Region, SUM(Sales.Amount)
FROM Sales
JOIN Customer ON Sales.CustomerID = Customer.CustomerID
GROUP BY Customer.Region;
```



# Методы представления данных для аналитических задач

## Таблицы и графики

- Таблицы, линейные графики, гистограммы

## Интерактивные дашборды

- Использование инструментов BI (Business Intelligence)
- Примеры: Tableau, Power BI



# Примеры представления данных

## Таблица данных

- Пример таблицы продаж по регионам

## Линейный график

- Пример графика роста продаж за год



# Интерактивные дашборды

## Преимущества

- Удобство использования, интерактивность, наглядность

## Инструменты

- Tableau: создание интерактивных визуализаций
- Power BI: интеграция с различными источниками данных



# Пример дашборда в Tableau

## Демонстрация

- Продажи по регионам, динамика за год, топ-продукты

## Возможности

- Фильтрация, детализация, экспорт отчетов



# Эффективные методы визуализации данных

## Принципы

- Ясность, точность, наглядность

## Лучшие практики

- Использование соответствующих типов графиков, минимизация перегрузки информацией, цветовая кодировка



# Заключение

## Подведение итогов

- Важность выбора правильной модели для задач OLTP и OLAP
- Роль аналитических баз данных и визуализации в бизнес-аналитике

## Дополнительные материалы и литература



# Вопросы и обсуждение

Вопросы от студентов

Обсуждение практических примеров

Дальнейшие шаги и изучение



# Благодарность за внимание

## Контактная информация

- Email: [iradche@gmail.com](mailto:iradche@gmail.com)
- Телеграм: [@dadaistka](https://t.me/dadaistka)

## Следующая лекция

- Тема следующей лекции:  
«Поисковые индексы и эффективное хранение и применение индексов на диске – хэш-таблицы, деревья поиска, пространственные индексы, полнотекстовый поиск»

# Спасибо за внимание!

[www.ifmo.ru](http://www.ifmo.ru)

IT'sMO *re than a*  
UNIVERSITY