

Лекция 24

**Многомерная и реляционная модель
хранилища. Кубы фактов.**

Схемы «звезда», «снежинка».

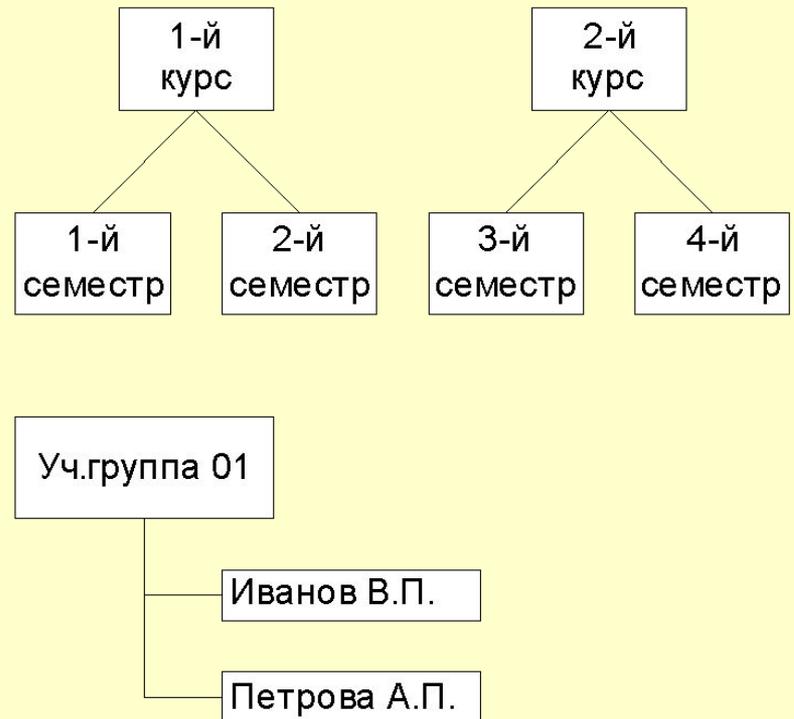
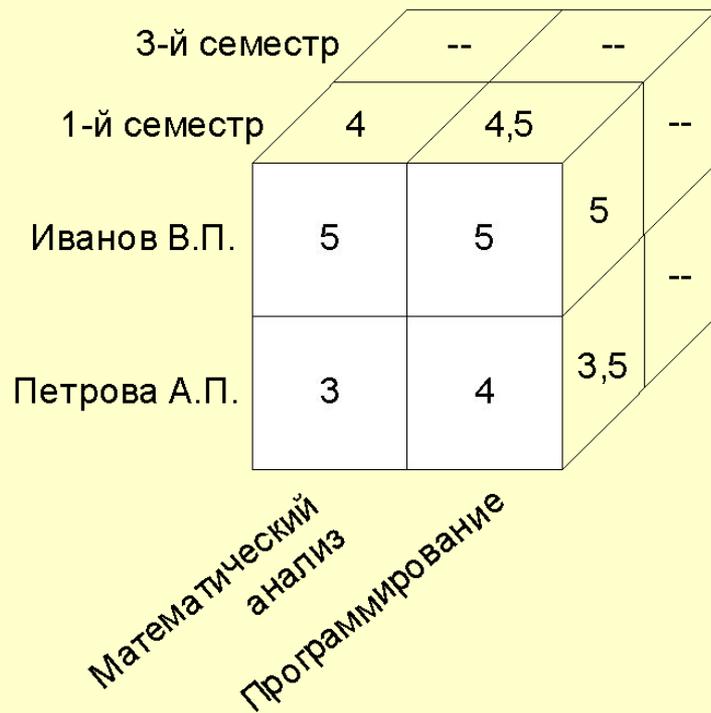
Основные термины

- OLAP (Online Analytical Processing) - автоматизированные технологии сложного (многомерного) анализа данных
- Data mining - извлечение данных, интеллектуальный анализ данных

Базовые понятия OLAP

- Факт (например, сданный экзамен)
- Параметр (например, оценка)
- Измерение (например, ФИО студента, семестр, дисциплина)

OLAP-куб (на примере задачи «Сессия»)



Способы хранения многомерных данных

- Реляционный (Relational OLAP – ROLAP)
- Многомерный (Multidimensional OLAP – MOLAP)
- Гибридный (Hybrid OLAP – HOLAP)

Частный случай ROLAP – оперативный R-ROLAP (Relational ROLAP)

Relational OLAP (ROLAP)

Данные и их представление - в реляционной БД. Ее структуру необходимо специально проектировать. Обработке запроса предшествует подготовка данных (в оперативном или в «регулярном» режиме).

Масштабируемое решение.

Сложность проектирования структуры, низкая эффективность SQL-запросов.

R-ROLAP (Real-time ROLAP)

Перед обработкой запроса делается «снимок» реляционной базы данных.

Данные структурируются и представляются в виде OLAP-куба.

Multidimensional OLAP (ROLAP)

Данные и их многомерное представление хранятся в специальной многомерной базе данных в оптимизированном многомерном формате.

Хорош в условиях небольшого набора данных.

Дороговизна решения и возможная неэффективность выполнения «обычных» запросов к базе данных.

Hybrid OLAP (ROLAP)

Хранение данных в реляционной базе данных, а их многомерное представление – в многомерной базе данных.

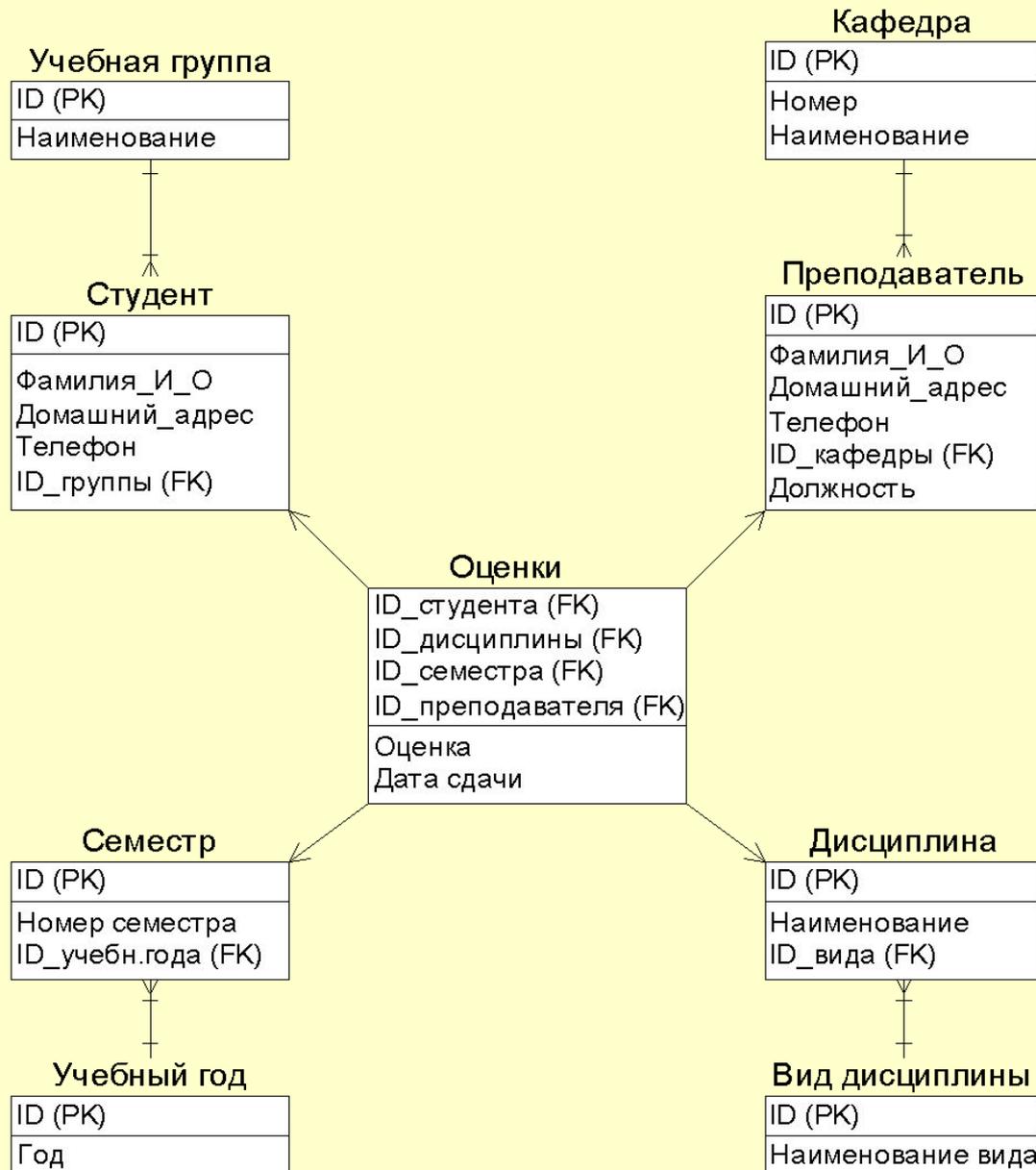
Промежуточное положение между ROLAP и MOLAP.

Необходимость синхронизации обеих баз.

Схема «звезда» (на примере задачи «Сессия»)



Схема «снежинка» (на примере задачи «Сессия»)



Виды запросов к многомерной базе данных

- Slice-and-dice. Выбор, сокращающий OLAP-куб (например, сечение).
- Drill-down. Декомпозиция на основе иерархии измерения.
- Roll-up. Свертка на основе иерархии измерения.
- Drill-across. Комбинирование (соединение) OLAP-кубов, имеющих общие измерения.
- Ranking. Выбор ячеек, появляющихся в верхней или, наоборот, нижней части отсортированного списка.
- Rotating. «Поворот» OLAP-куба, чтобы увидеть данные, сгруппированные по другим измерениям.