

БАЗЫ ДАННЫХ

Лекция 1
Введение в курс баз данных

Перед тем, как начать

- Лектор – Блинова Евгения Александровна,
eugenia.blinova@gmail.com
- Доклады
- Выступления на конференциях, олимпиадах и проекты
- Книги, ПО и задачи
- <https://diskstation.belstu.by>:5001 student fitfit
- /Для_студентов_ФИТ_БГТУ/ПРЕПОДАВАТЕЛИ/
Блинова/Базы данных - 2 курс
- Посещение
- Сдача лабораторных работ

Описание курса

- Два семестра:
 - Microsoft SQL Server 2012
 - ORACLE 12c
- 36 часов лекций, 36 часов лабораторных работ
- Задания на лабораторные работы - лабораторный практикум
- Самостоятельная работа
- 2 контрольные работы и коллоквиум
- Экзамен – 2 теоретических вопроса и задача

База данных

- База данных – это совокупность взаимосвязанных данных

Требования к информации в БД

- Полезность - уменьшает информационную энтропию системы
- Полнота информации - информации должно быть достаточно, чтобы осуществить качественное управление
- Точность
- Достоверность - заведомо ошибочные данные не должны храниться в базе данных
- Непротиворечивость
- Актуальность

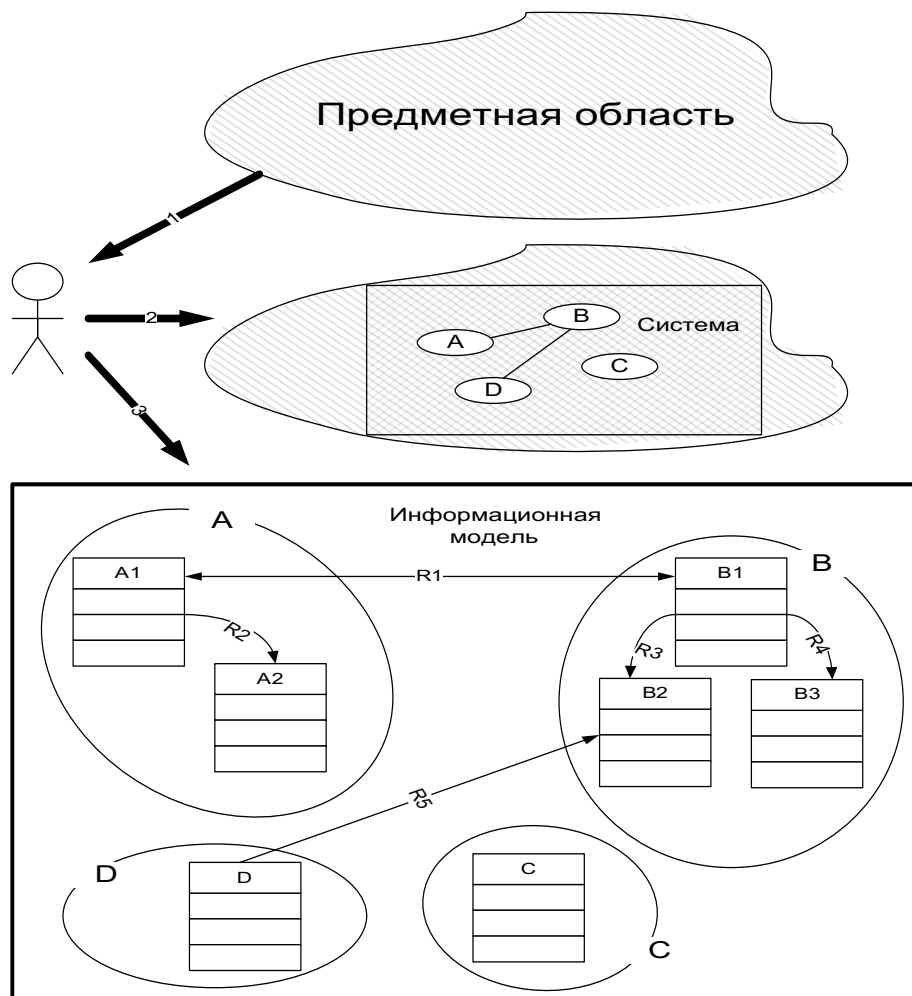
Проектирование БД

- Определение границ исследуемой области – предметной области
- Системный анализ – определение объектов и связей между ними
- Построение логической схемы базы данных в соответствии с определенными правилами – моделью данных
- Реализация базы данных – описание ее в терминах некоторой СУБД

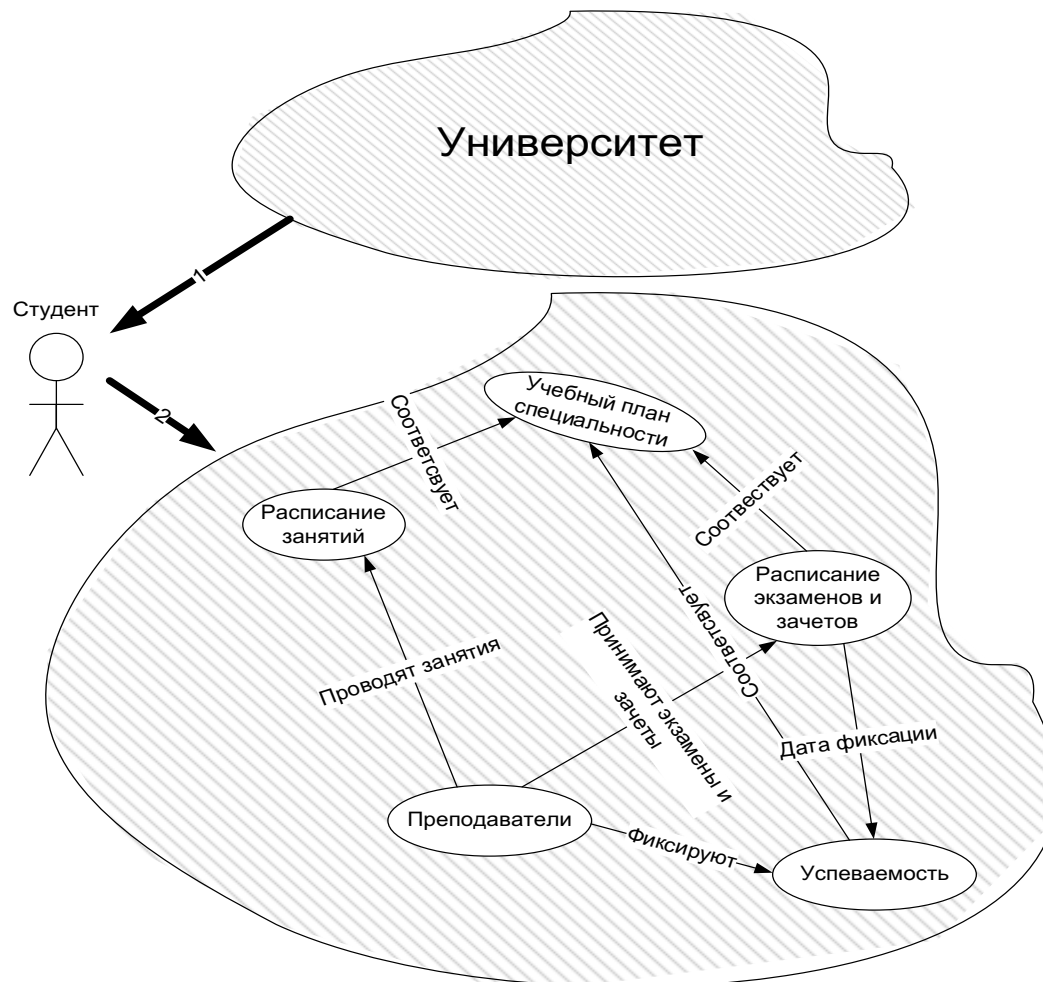
Проектирование БД

- **Предметная область** – часть реального мира, подлежащая изучению, с целью описания и управления
- **Предметная область** – это множество объектов и связей между этими объектами
- **Модель данных** – структурированное представление данных и связей между ними

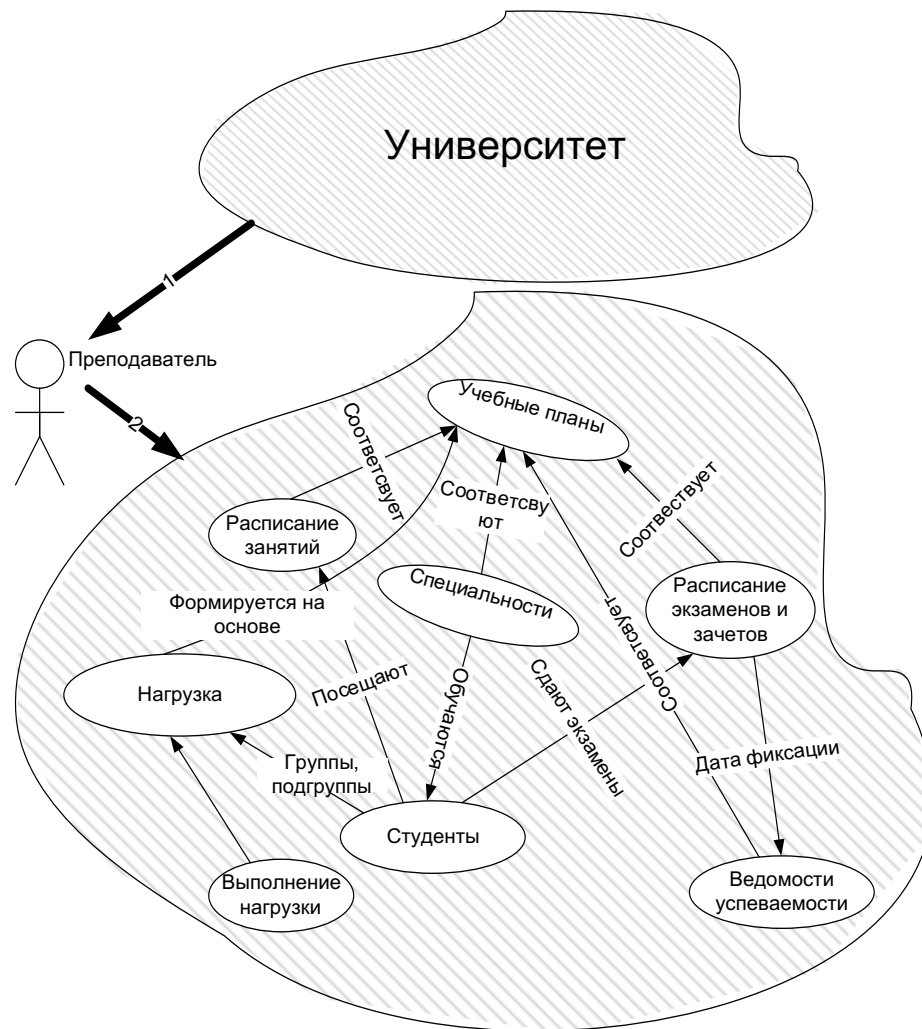
Проектирование БД



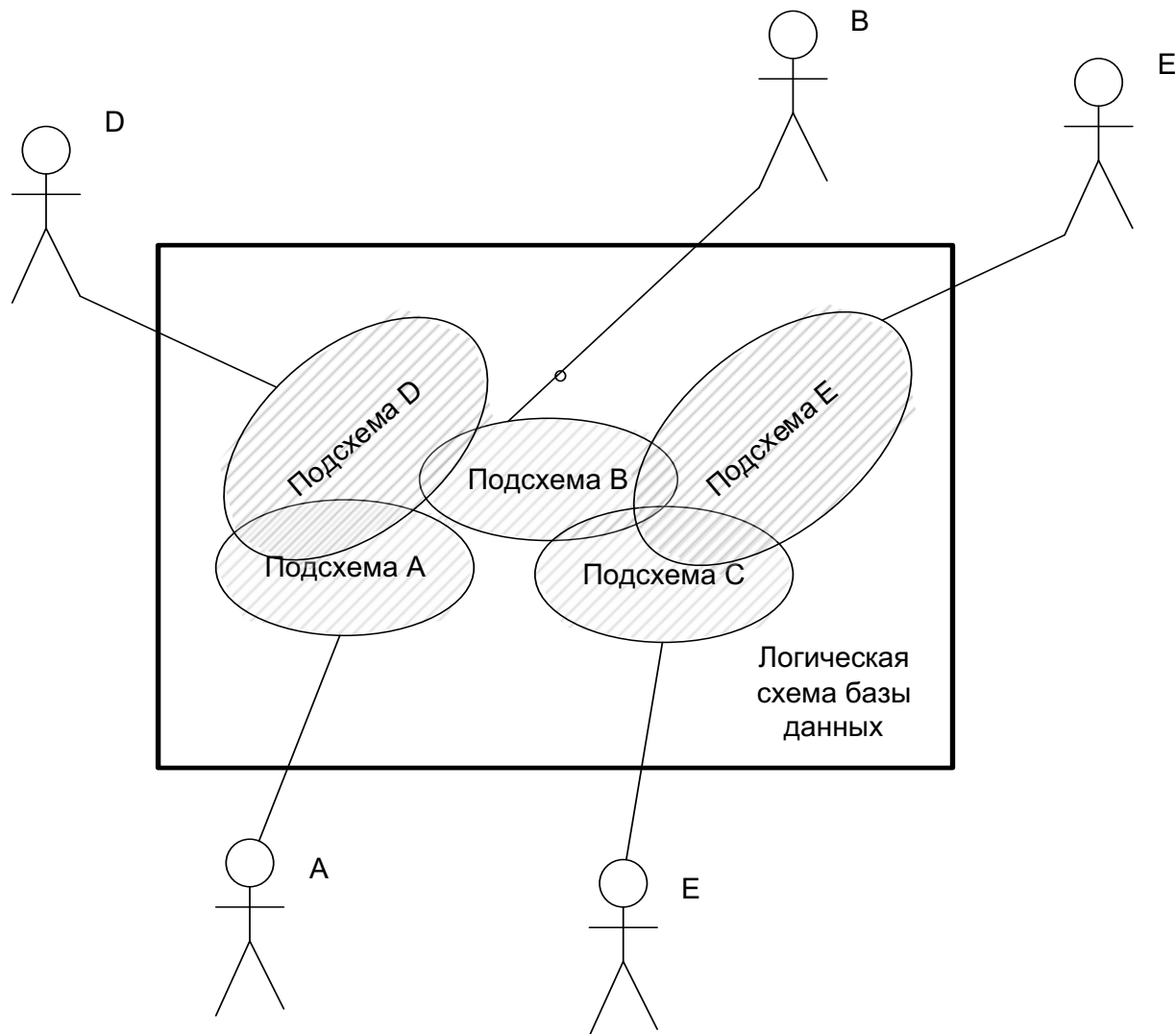
Пример



Пример



Логическая схема данных



База данных

- Хранилище динамически обновляемой информации
- Информация отражает состояние некоторой предметной области (объекта) и должна быть полезной, точной, актуальной и непротиворечивой
- Информация представлена в вид:
 - метаданных (описание модели данных)
 - данных
- Каждый пользователь базы данных знает только о существовании данных, необходимых для решения его задач
- Совокупность всех представлений - это логическая схема данных

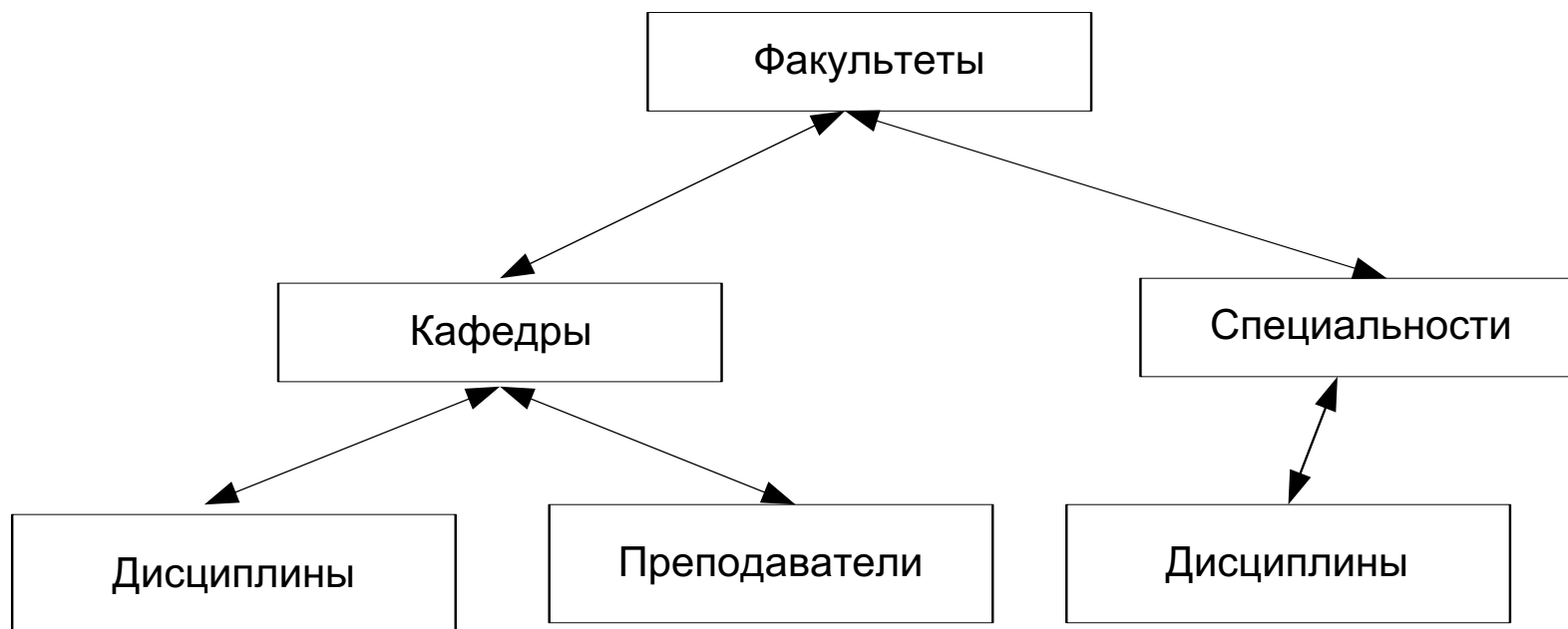
Система управления базами данных

- Программная реализация **технологии** хранения, извлечения, обновления и обработки данных в базе данных

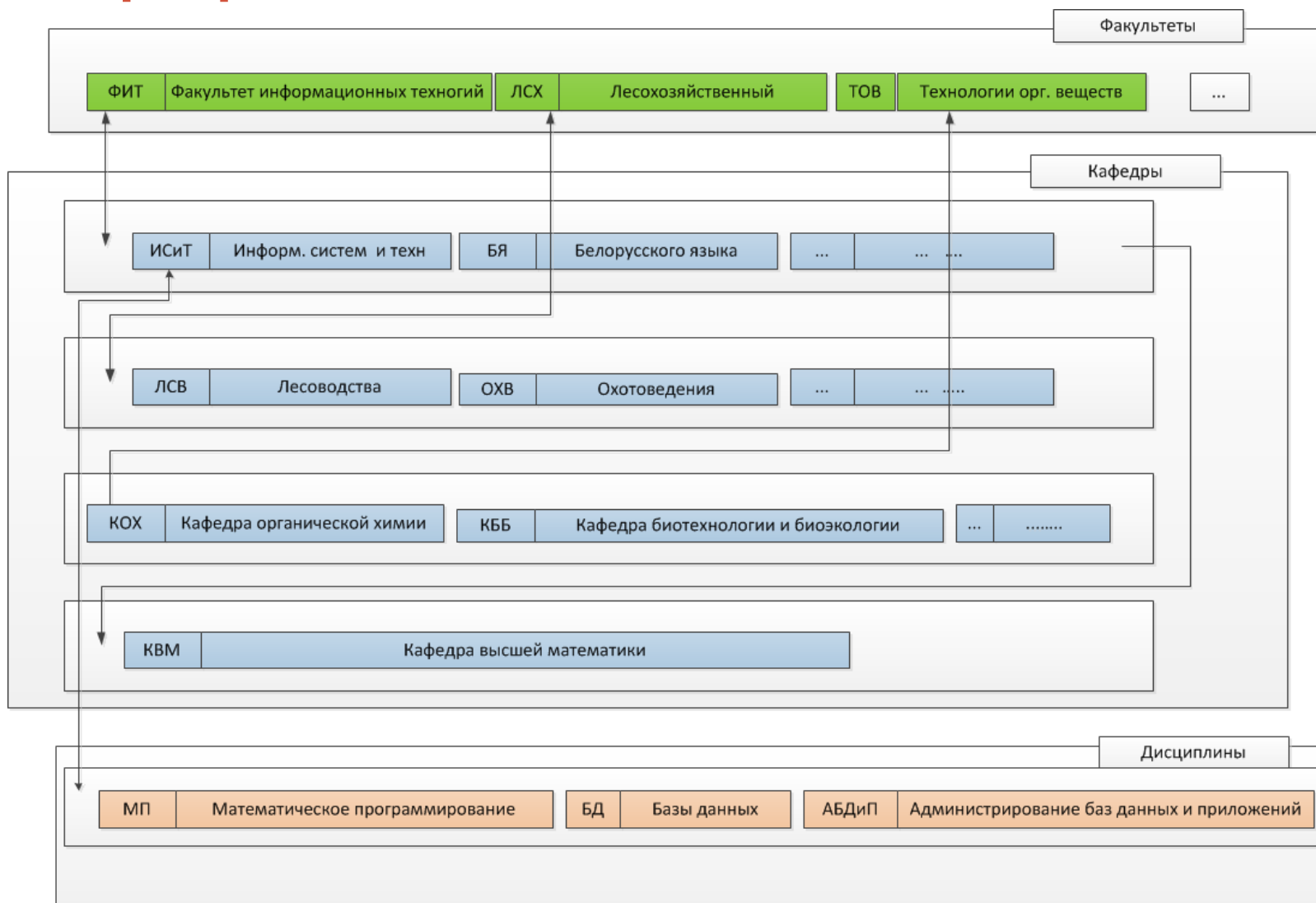
Модели данных

- Иерархическая
- Сетевая
- Реляционная

Иерархическая модель данных



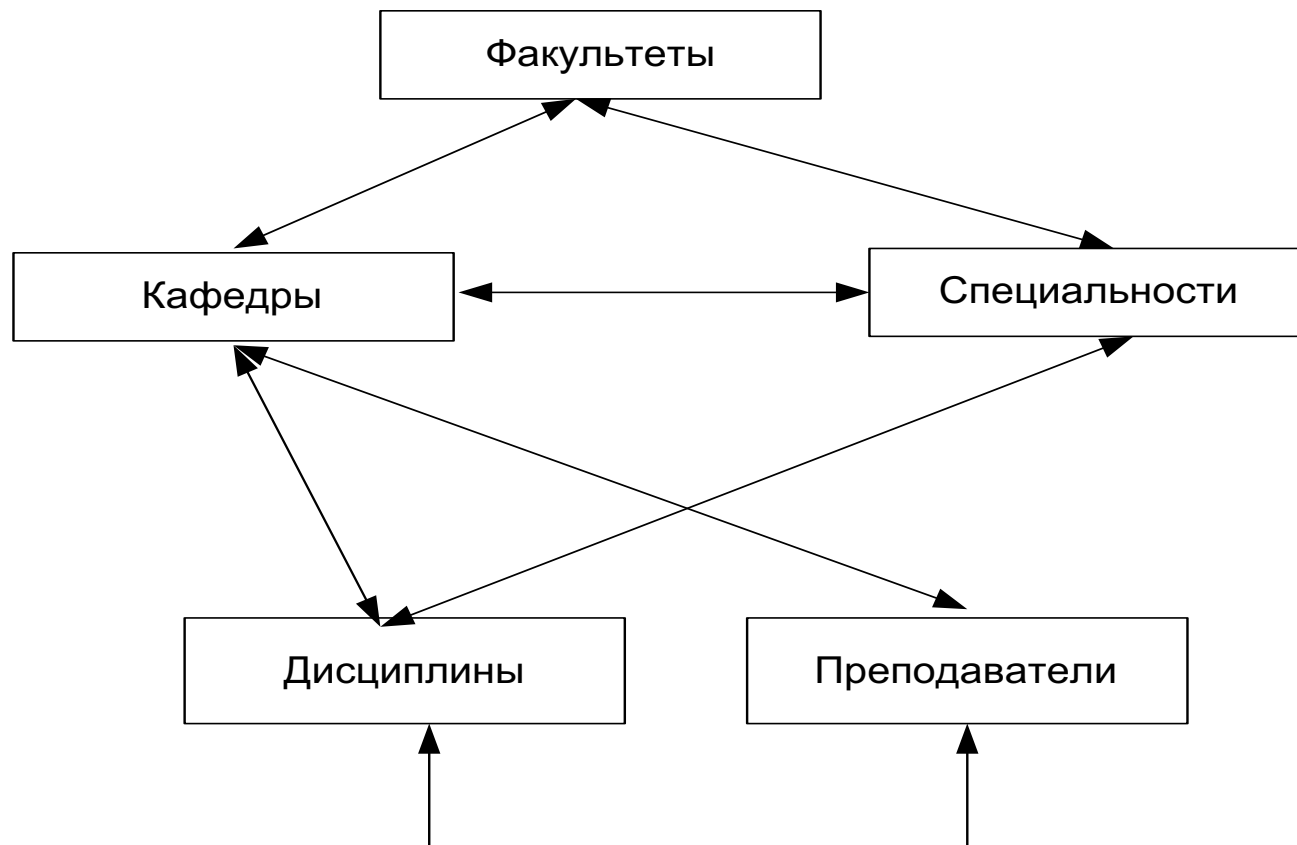
Иерархическая модель данных



Иерархическая модель данных

- Типичный представитель – IBM Information Management System

Сетевая модель данных



Сетевая модель данных

- Integrated Data Store от General Electric
- DMS от UNIVAC

Реляционная модель данных

- Основана на теории множеств
- Реляционная алгебра – Эдгар Франк Кодд (1923-2003)
- Данные имеют собственную природу, независимую от способа их использования
- Определения:
 - - **домен**: множество;
 - - **таблица**: отношение;
 - - **атрибут**: имя столбца таблицы (имя домена);
 - - **заголовок таблицы**: множество всех атрибутов;
 - - **кортеж**: элемент отношения или строка таблицы

Операции реляционной алгебры

- **UNION** (объединение)
- **INTERSECT** (пересечение)
- **MINUS** (разность)
- **TIMES** (декартово произведение)
- **WHERE** (ограничение)
- **PROJECT** (проекция)
- **JOIN** (соединение)
- **DIVIDE BY** (реляционное деление)
- **RENAME** (переименование)
- **:=** (присваивание).

Реляционная модель данных

- Relation – отношение
- Отношение может быть представлено в виде двумерной таблицы
- Реляционная база данных представляет собой набор взаимосвязанных таблиц
- Все объекты разделяются на типы
- Объекты одного и того же типа имеют свой набор атрибутов
- Один из атрибутов однозначно идентифицирует объект в таблице – первичный ключ

Реляционная модель данных

- Структурный аспект — данные в базе данных представляют собой набор отношений
- Аспект целостности — отношения (таблицы) отвечают определенным условиям целостности
- РМД поддерживает декларативные ограничения целостности уровня домена (типа данных), уровня отношения и уровня базы данных
- Аспект обработки — РМД поддерживает операторы манипулирования отношениями (реляционная алгебра, реляционное исчисление)

Нормализация данных

- **Нормализация данных** – процесс преобразования таблиц базы данных к нормальной форме
- **Шесть нормальных форм** – 1NF, 2NF,...6NF.
- Широкое практическое применение имеют формы 1NF, 2NF, 3NF

Первая нормальная форма

- Таблица не должна содержать повторяющихся групп данных
- Атомарность – каждый столбец должен содержать одно неделимое значение
- Пример:
- ФИО – Адрес (город, улица, дом, квартира)
- Фильм – Исполнители (список актеров)

Первая нормальная форма

- Устранить повторяющиеся группы в отдельных таблицах
- Создать отдельную таблицу для каждого набора связанных данных
- Идентифицировать каждый набор связанных данных с помощью первичного ключа

Вторая нормальная форма

- Таблица находится в первой нормальной форме
- Каждый неключевой атрибут полностью функционально зависит от каждого возможного ключа
- Простой и составной ключ
- Пример:
- Студент – Университет – Средний балл – Стипендия

Вторая нормальная форма

- Создать отдельные таблицы для наборов значений, относящихся к нескольким записям
- Связать эти таблицы с помощью внешнего ключа

Третья нормальная форма

- Таблица находится во второй нормальной форме
- Отсутствуют транзитивные зависимости
- Пример:
- Студент – Группа – Факультет – Университет

SQL

- Язык **SQL (Structured Query Language)**, язык структурированных запросов) – специализированный язык, предназначенный для написания запросов к реляционной БД

SQL

- 1986 – первый вариант стандарта
- 1989 – доработан стандарт
- 1992 – внесены значительные изменения (SQL2)
- 1999 – (SQL3) добавлены:
 - поддержка регулярных выражений
 - рекурсивных запросов
 - поддержка триггеров
 - базовые процедурные расширения
 - не скалярные типы данных
 - некоторые объектно-ориентированные возможности
- 2003 - поддержка XML
- 2006 - возможность совместно использовать в запросах SQL и XQuery
- 2008 - улучшены возможности оконных функций

Операторы SQL

- **DDL** - Data Definition Language - язык определения данных
- **DML** - Data Manipulation Language - язык манипулирования данными
- **TCL** - Transaction Control Language - язык управления транзакциями
- **DCL** - Data Control Language - язык управления данными

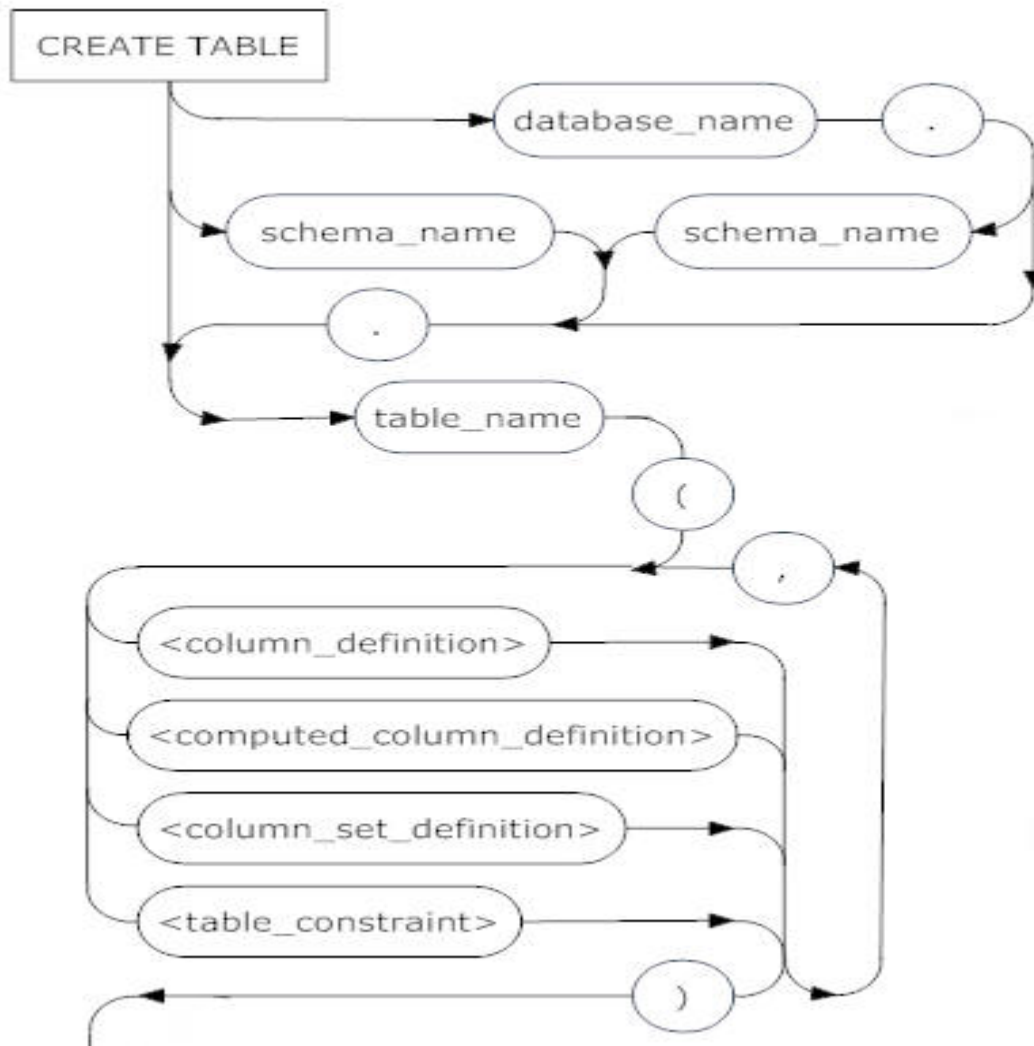
Операторы DDL

- Операторы DDL предназначены для **создания, удаления и изменения** объектов БД или сервера СУБД
- DDL включает операторы:
 - **CREATE**
 - **ALTER**
 - **DROP**

create тип имя дополнение

```
CREATE TABLE STUDENT (NAME nvarchar(50), GROUP_NUM int)
```

SQL Railroad Diagram



Операторы DML

- Операторы DML предназначены для **работы со строками таблиц**
- DML включает операторы:
 - **SELECT**
 - **INSERT**
 - **DELETE**
 - **UPDATE**

select список дополнение

SELECT * FROM STUDENT

SELECT NAME, GROUP_NUM FROM STUDENT

Операторы TCL

- Операторы TCL предназначены для **управления транзакциями**
- Транзакция – это несколько DML-операторов, которые либо **все** выполняются, либо все не выполняются
- TCL SQL включает операторы:
 - **BEGIN TRAN**
 - **SAVE TRAN**
 - **COMMIT TRAN**
 - **ROLLBACK TRAN**

`begin tran дополнение`

Операторы DCL









- Операторы DCL предназначены для **управления процессом авторизации**
- **Авторизация** – это процедура проверки разрешений на выполнение определенных операций
- **Принципал** – это объект сервера или БД, которому может быть выдано разрешение на выполнение операции, а также отобрано или запрещено разрешение
- DCL включает в себя операторы:
 - **GRANT**
 - **REVOKE**
 - **DENY**

`grant список on объект to принципал`

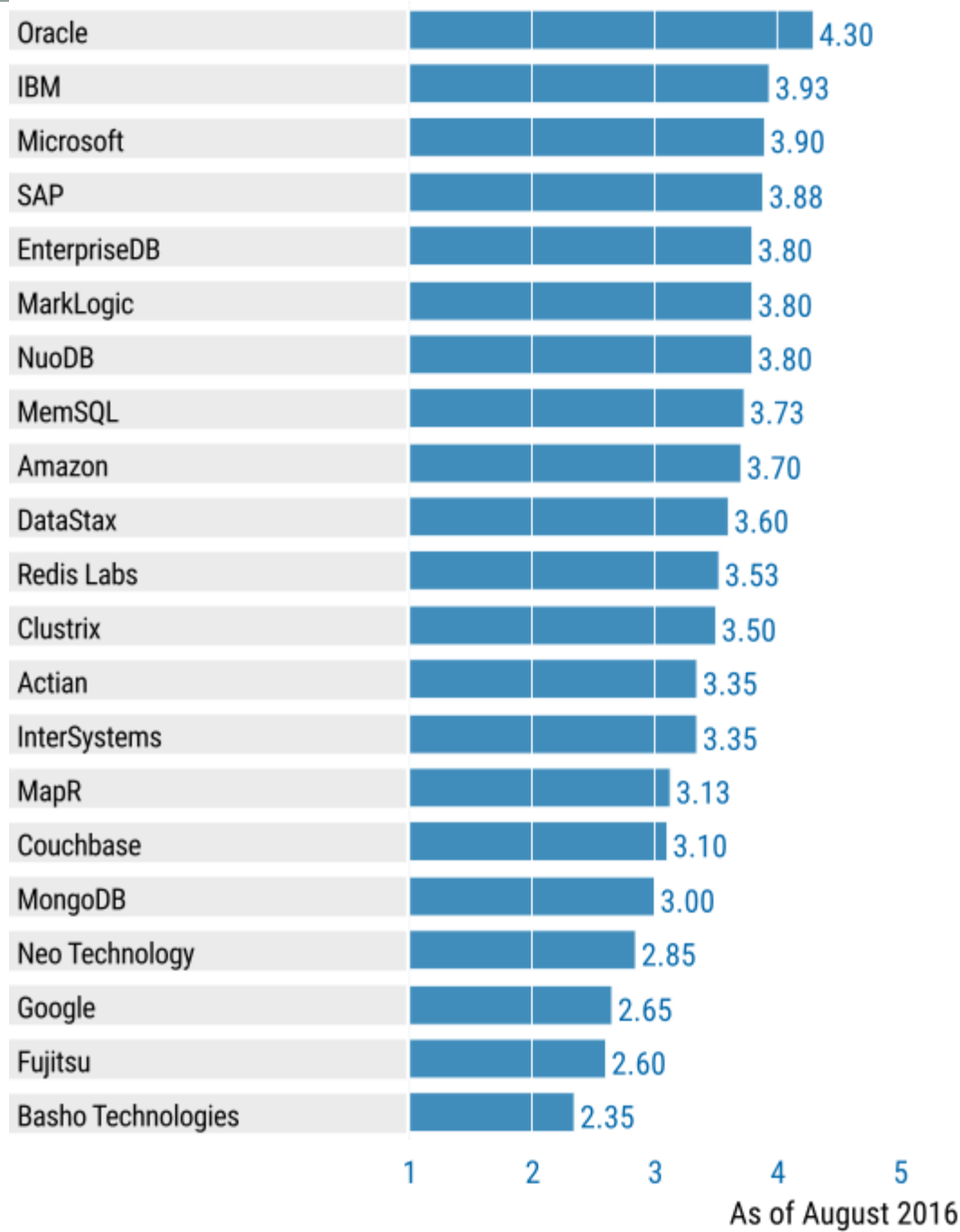
Microsoft SQL Server

- **Microsoft SQL Server** — система управления реляционными базами данных (РСУБД), разработанная корпорацией Microsoft
- Входит в «большую тройку»: Oracle, DB2
- Имеет свой диалект языка запросов SQL
- Имеет свое процедурное расширение языка запросов Transact-SQL

| # | | Название | Год | SQL | Разработчик |
|---------|---|---------------------------------|------|-----|-------------------|
| 1 0 |  | MySQL | 1995 | ✓ | Oracle |
| 2 +1 |  | PostgreSQL | 1995 | ✓ | сообщество |
| 3 -1 |  | MS SQL Server | 1988 | ✓ | Microsoft |
| 4 +1 |  | MongoDB | 2009 | ✗ | MongoDB |
| 5 +1 |  | SQLite | 2000 | ✓ | Нwaci, сообщество |
| 6 -2 |  | Oracle Database | 1979 | ✓ | Oracle |

| | | | | | |
|-----------|--|--------------------------------------|------|---|---|
| 7 +1 |  Firebird | Firebird | 2000 | ✓ | сообщество |
| 8 0 |  CouchDB | CouchDB | 2005 | ✗ | Apache |
| 8 -1 |  IBM | DB2 | 1995 | ✓ | IBM |
| 9 new |  MariaDB | MariaDB | 2009 | ✓ | MariaDB Corporation Ab, MariaDB Foundation, сообщество |
| 10 -1 |  RAVENDB | RavenDB | 2009 | ✗ | Hibernating Rhinos |
| 10 new |  redis | Redis | 2009 | ✗ | Redis Labs, сообщество |
| 10 -1 |  SAP | SAP ASE (ex: Sybase) | 1988 | ✓ | SAP AG |
| 11 new |  PERCONA Server | Percona Server | 2006 | ✓ | Percona |

Product or Service Scores for Traditional Transactions



| Rank | | | DBMS | Database Model | Score | | |
|----------|----------|----------|------------------------------|-------------------|----------|----------|----------|
| Sep 2018 | Aug 2018 | Sep 2017 | | | Sep 2018 | Aug 2018 | Sep 2017 |
| 1. | 1. | 1. | Oracle | Relational DBMS | 1309.12 | -2.91 | -49.97 |
| 2. | 2. | 2. | MySQL | Relational DBMS | 1180.48 | -26.33 | -132.13 |
| 3. | 3. | 3. | Microsoft SQL Server | Relational DBMS | 1051.28 | -21.37 | -161.26 |
| 4. | 4. | 4. | PostgreSQL | Relational DBMS | 406.43 | -11.07 | +34.07 |
| 5. | 5. | 5. | MongoDB | Document store | 358.79 | +7.81 | +26.06 |
| 6. | 6. | 6. | DB2 | Relational DBMS | 181.06 | -0.78 | -17.28 |
| 7. | 8. | 10. | Elasticsearch | Search engine | 142.61 | +4.49 | +22.61 |
| 8. | 7. | 9. | Redis | Key-value store | 140.94 | +2.37 | +20.54 |
| 9. | 9. | 7. | Microsoft Access | Relational DBMS | 133.39 | +4.30 | +4.58 |
| 10. | 10. | 8. | Cassandra | Wide column store | 119.55 | -0.02 | -6.65 |
| 11. | 11. | 11. | SQLite | Relational DBMS | 115.46 | +1.73 | +3.42 |
| 12. | 12. | 12. | Teradata | Relational DBMS | 77.38 | -0.02 | -3.52 |
| 13. | 13. | 16. | Splunk | Search engine | 74.03 | +3.53 | +11.45 |
| 14. | 14. | 18. | MariaDB | Relational DBMS | 70.64 | +2.34 | +15.17 |
| 15. | 15. | 13. | Solr | Search engine | 60.20 | -1.69 | -9.71 |
| 16. | 18. | 19. | Hive | Relational DBMS | 59.63 | +1.69 | +11.02 |
| 17. | 17. | 15. | HBase | Wide column store | 58.47 | -0.33 | -5.87 |
| 18. | 16. | 14. | SAP Adaptive Server | Relational DBMS | 58.04 | -2.39 | -8.71 |
| 19. | 19. | 17. | FileMaker | Relational DBMS | 55.30 | -0.75 | -5.69 |
| 20. | 21. | 22. | Amazon DynamoDB | Multi-model | 53.34 | +1.69 | +15.52 |
| 21. | 20. | 20. | SAP HANA | Relational DBMS | 52.73 | +0.80 | +4.40 |
| 22. | 22. | 21. | Neo4j | Graph DBMS | 40.10 | -0.83 | +1.67 |
| 23. | 23. | 23. | Couchbase | Document store | 34.55 | +1.59 | +1.44 |
| 24. | 24. | 24. | Memcached | Key-value store | 31.54 | -1.38 | +2.60 |
| 25. | 25. | 27. | Microsoft Azure SQL Database | Relational DBMS | 25.25 | -0.85 | +3.65 |





Microsoft SQL Server

| Версия | Год | Название |
|--------|------|---------------------------|
| 1.0 | 1989 | SQL Server 1.0 (16 bit) |
| 1.1 | 1991 | SQL Server 1.1 (16 bit) |
| 4.21 | 1993 | SQL Server 4.21 |
| 6.0 | 1995 | SQL Server 6.0 |
| 6.5 | 1996 | SQL Server 6.5 |
| 7.0 | 1998 | SQL Server 7.0 |
| 8.0 | 2000 | SQL Server 2000 |
| 8.0 | 2003 | SQL Server 2000 64-bit |
| 9.0 | 2005 | SQL Server 2005 |
| 10.0 | 2008 | SQL Server 2008 |
| 10.25 | 2010 | Azure SQL DB |
| 10.50 | 2010 | SQL Server 2008 R2 |
| 11.0 | 2012 | SQL Server 2012 |
| 12.0 | 2014 | SQL Server In-Memory OLTP |
| 13.0 | 2016 | SQL Server 2016 |
| 14.0 | 2017 | SQL Server vNext |

Редакции Microsoft SQL Server 2012

- **SQL Server 2012 Enterprise Edition**

- Не имеет ограничений по количеству поддерживаемых ядер
- Не имеет ограничений по максимальному объему используемой памяти.
- Максимальный размер реляционной базы данных — 524 Пб.

- **SQL Server 2012 Business Intelligence Edition**

- Имеет ограничение — 4 процессора или 16 ядер на экземпляр;
- Максимальный объем используемой памяти — 64 Гб;
- Максимальный размер реляционной базы данных — 524 Пб.

- **SQL Server 2012 Standard Edition**

- Имеет ограничение — 4 процессора или 16 ядер на экземпляр;
- Максимальный объем используемой памяти — 64 Гб;
- Максимальный размер реляционной базы данных — 524 Пб.

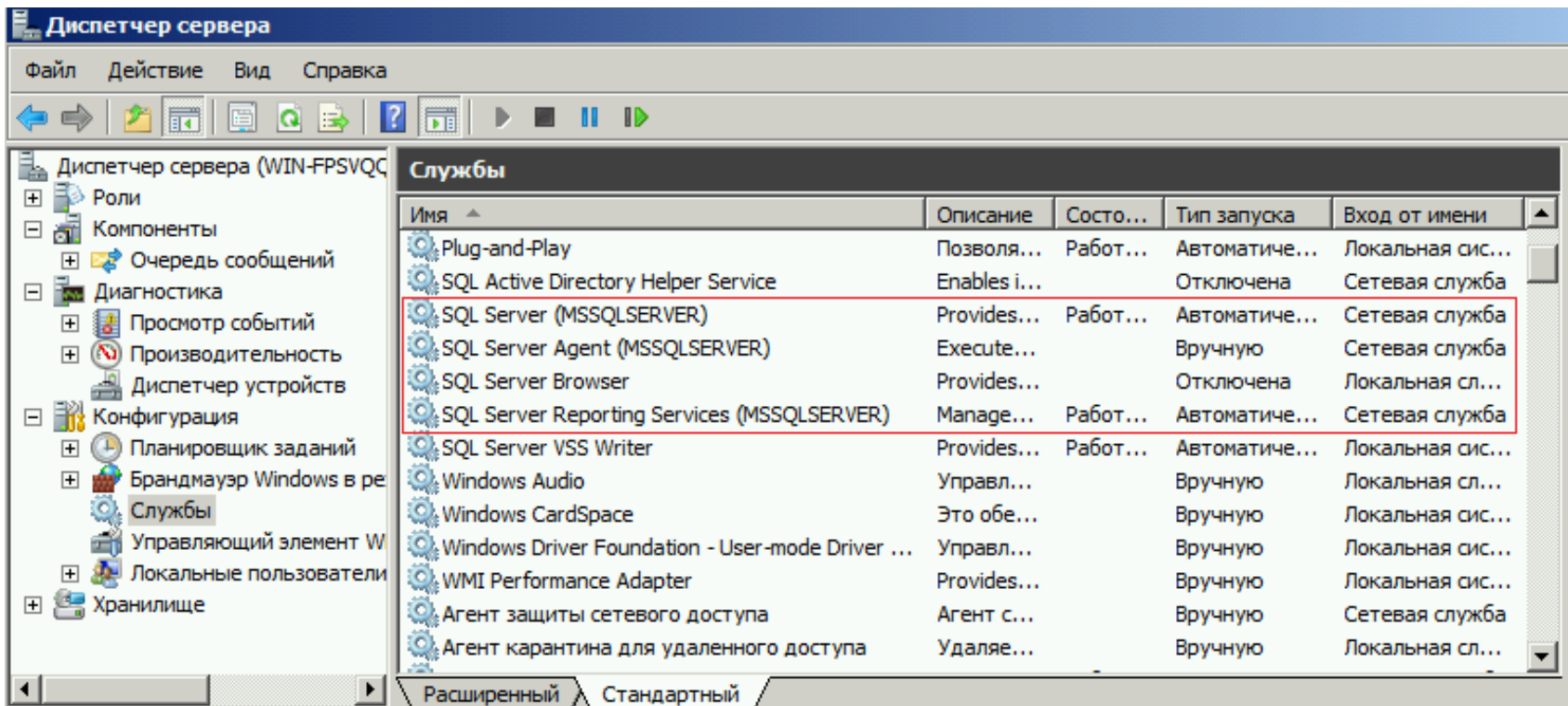
- **SQL Server 2012 Web Edition**

- **SQL Server 2012 Developer Edition .**

- **SQL Server 2012 Express Edition**

- Имеет ограничение — 1 процессор (до 4 ядер);
- Максимальный объем используемой памяти — 1 Гбайт;
- Максимальный размер реляционной базы данных — 10 Гб.

Microsoft SQL Server



Службы Microsoft SQL Server

| Наименование службы | Назначение |
|----------------------|---|
| Database Engine | Управление реляционными БД |
| Analysis Services | Управление OLAP-кубами и интеллектуальный анализ данных |
| Integration Services | Поддержка решений по извлечению, преобразованию и загрузке данных |
| Reporting Services | Управление отчетами, построенными на основе SQL-запросов к реляционным БД |
| Full-Text Search | Управление полнотекстовым поиском |
| SQL Server Agent | Автоматизация административных задач |
| SQL Server Browser | Управление соединениями |

Database Engine

- **Database Engine** является ядром системы управления реляционной БД
- Может быть установлено **несколько** экземпляров службы **Database Engine**
- При этом **только один** экземпляр может быть службой по умолчанию (с именем **MS SQL SERVER**), другие экземпляры должны иметь уникальные имена
- Каждый экземпляр службы **Database Engine** требует отдельной инсталляции, конфигурации и настройки безопасности
- Один **Database Engine** может обеспечить доступ к нескольким БД

Системные базы данных

| Системная база данных | Назначение |
|-----------------------|---|
| master | Хранит все системные данные Database Engine, а также информацию о других БД. |
| msdb | Используется службами SQL Server Agent (выполнение заданий по расписанию), Database Mail (формирование уведомлений по электронной почте), а также хранит информацию о резервном копировании БД. |
| tempdb | Пространство для временных объектов Database Engine и пользовательских временных таблиц. База данных пересоздается при каждой перезагрузке |
| model | Шаблон, используемый при создании всех БД, управляемых экземпляром Database Engine. |
| resource | БД, используемая только для чтения. Содержит системные объекты экземпляра Database Engine. Файлы БД являются скрытыми и не отображаются в MSMS. |

Утилиты Microsoft SQL Server

- SQL Server Management Studio
- SQL Server Books Online
- SQLCMD
- Microsoft SQL Configuration Manager

Вопросы?