

Моделирование предметной области

Г. Гарсия –Молина, Дж. Ульман, Дж. Уидом,
Системы баз данных, 2003г, Глава 1

Этапы разработки базы данных в информационной системе

- Проектирование базы данных
 - Логическое проектирование
 - Физическое проектирование
- Планирование базы данных
 - Аппаратных средств
 - Программных средств
- Реализация базы данных
- Сопровождение
- Реинжиниринг

Инфологическое проектирование

- Модель «Сущность – Связь», Entity Relationship Model (ERM), Питер Чен, 1976 г., США, профессор
- Реляционная модель данных
- Object Definition Language – стандарт средств описания объектно-ориентированных баз данных
- UML. Диаграммы классов
- Полуструктурированные данные (XML)
- Семантические сети

Модель «Сущность-связь»

- Основные компоненты ER модели
 - Множество сущностей
 - Атрибуты сущностей
 - Связи
- Этапы проектирования ER модели
 - Выделение сущностей-объектов
 - Определение атрибутов сущностей
 - Выделение идентифицирующих атрибутов-ключей
 - Определение связей
 - Вид связи: 1:1, 1:m, n:m
 - Обязательность связей

Сущность

- Сущность – абстрактный объект определенного вида, моделирующий объекты приложения

Объект материального мира, явление, событие, и т.д.

Примеры: Студент, Форум, Накладная, Ведомость

- Сходство с понятием объект в ОМД
- Сущность – статические объекты, определяет структуры данных
- Нет определения операций
- Множество сущностей (класс)

Атрибуты

- Атрибут является свойством сущности

сотрудник (таб_номер, фамилия, имя,
 отчество, код_должности,
 дата_приема, ...)
- Идентифицирующий атрибут - ключ
- Именованние атрибутов (наиболее
точное соответствие смыслу свойства)

СВЯЗИ

- Связи – соединения (соотношения) между двумя или большим числом сущностей
- ER модель допускает наличие связей, включающей произвольное количество сущностей.
- Наиболее распространена *бинарная связь* между 2-мя множествами сущностей

ТИПЫ связей

- Если каждый член множества E_1 посредством связи R может быть соединен не более чем с одним членом E_2 , то R **связь типа M:1**

Сотрудник \rightarrow Отдел, (m:1)

- Если связь R в обоих направлениях от E_1 к E_2 и от E_2 к E_1 относится к типу M:1, то R **связь типа 1:1**

Человек \longleftrightarrow Паспорт, полис страхования

Компания \longleftrightarrow Президент, директор

Связи (2)

- Если связь R ни в одном из направлений ни от E_1 к E_2 и ни от E_2 к E_1 не относится к типу $M:1$,
то R **связь типа $M:N$**

Студент \longleftrightarrow Преподаватель, (m:n)

- Необязательные связи

Человек $\leftarrow \dots \rightarrow$ Авиабилет

- Многосторонние связи – роли.

«Контракт» – связь между персоной, предприятием, проектом

Обычно многосторонние связи преобразуют в совокупность бинарных, образуя новые сущности вместо атрибутов связи.

Сотрудники \leftarrow Контракты \rightarrow Проекты

Учреждения

- Подклассы – специальные множества сущностей.
Для соединения «полного» множества с его подклассами применяется связь «есть» ($IS\ A$)

Нотация

Сущности
изображаются
прямоугольниками

Имя сущности

Свойства
изображаются
овалами

Имя
атрибута

Связи изображаются
ромбами

Имя связи

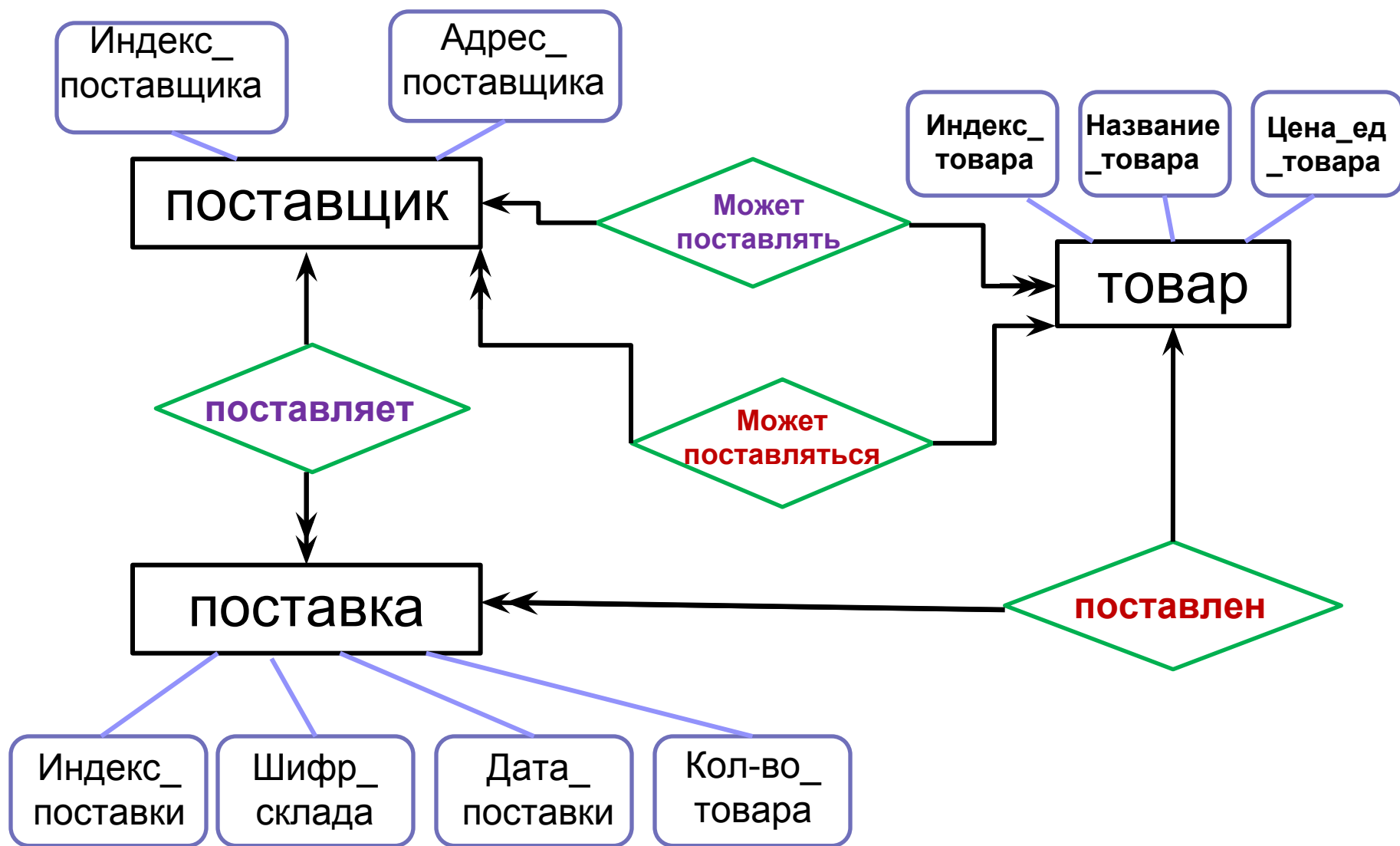
- ❖ Индекс поставки,
- ❖ Индекс поставщика,
- ❖ Адрес поставщика
- ❖ Индекс товара
- ❖ Название товара
- ❖ Количество товара
- ❖ Цена единицы товара
- ❖ Шифр склада
- ❖ Дата поставки

- ❖ Пусть необходимо описать данные возникающие в процессе **поставки материалов** на склад.
- ❖ Данные берутся из документов, сопровождающих поставки – **накладных**.
- ❖ Пусть из накладных мы выяснили, что в процессе поставки следует зафиксировать следующую информацию о поставках:

Первое представление



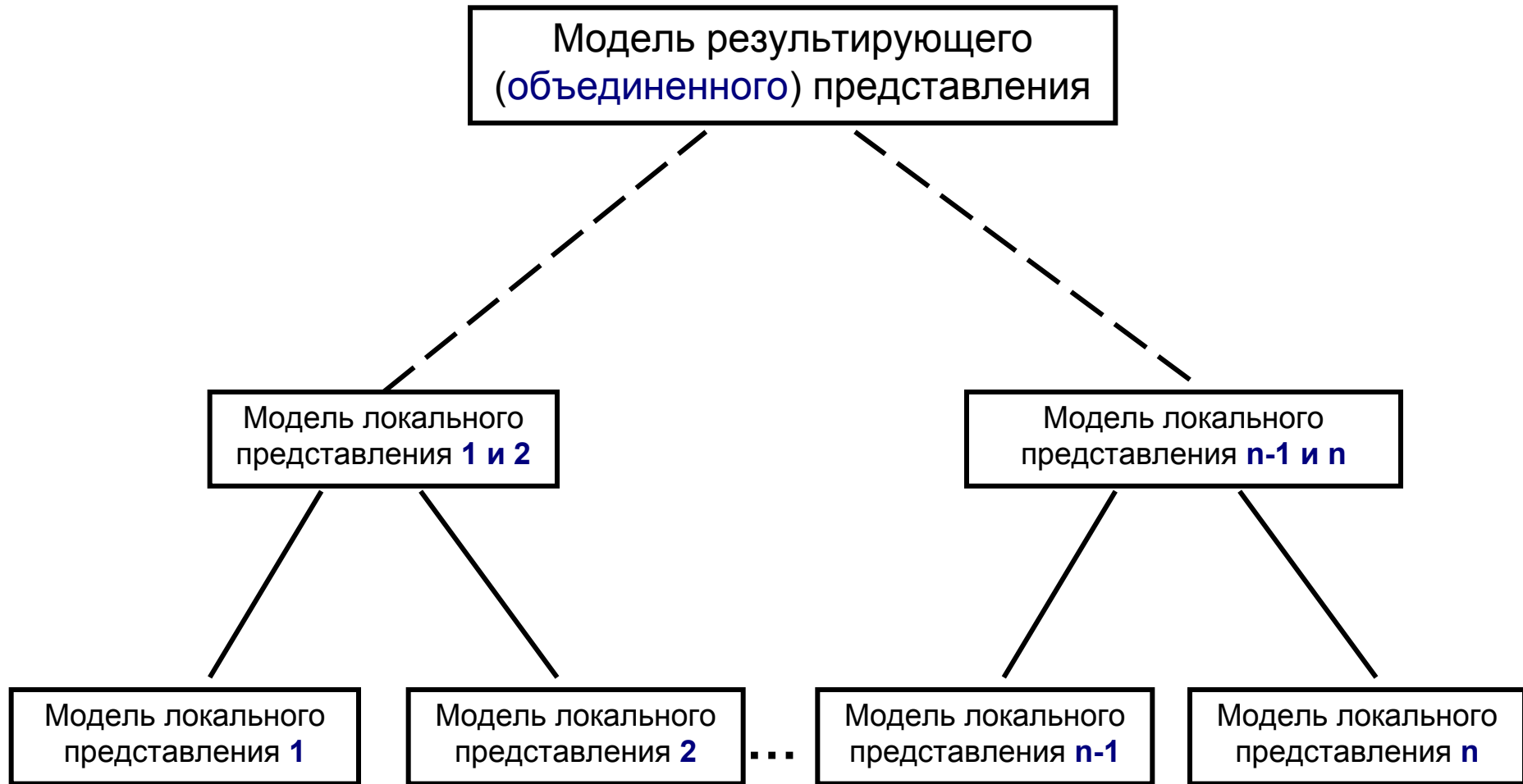
Модификация исходного представления



Этапы разработки ER модели

- Разработка локальных представлений (частей информационной системы)
 - ❖ Формулирование сущностей (имена, содержание)
 - ❖ Выбор идентифицирующего атрибута
 - ❖ Спецификация связей (имена, типы)
 - ❖ Добавление описательных атрибутов
- Объединение представлений пользователей
 - ❖ Идентичность
 - ❖ Агрегация
 - ❖ Обобщение

Схема процесса объединения ЛП



Идентичность

- Два или более элементов являются идентичными, если они имеют **одинаковое семантическое значение**
- В силу абстрагирования при выявлении сущностей идентичность элементов **устанавливается экспертом**
- Идентичность и подобие
- Персона, Служащий, Сотрудник, Работник, Персонал

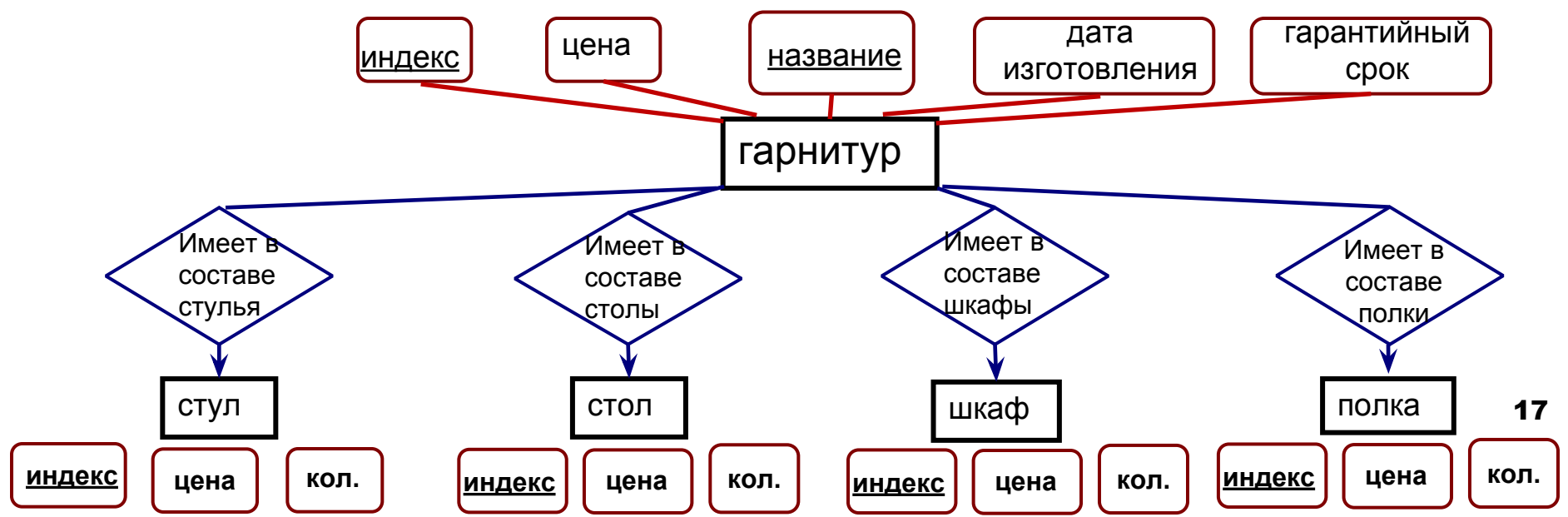
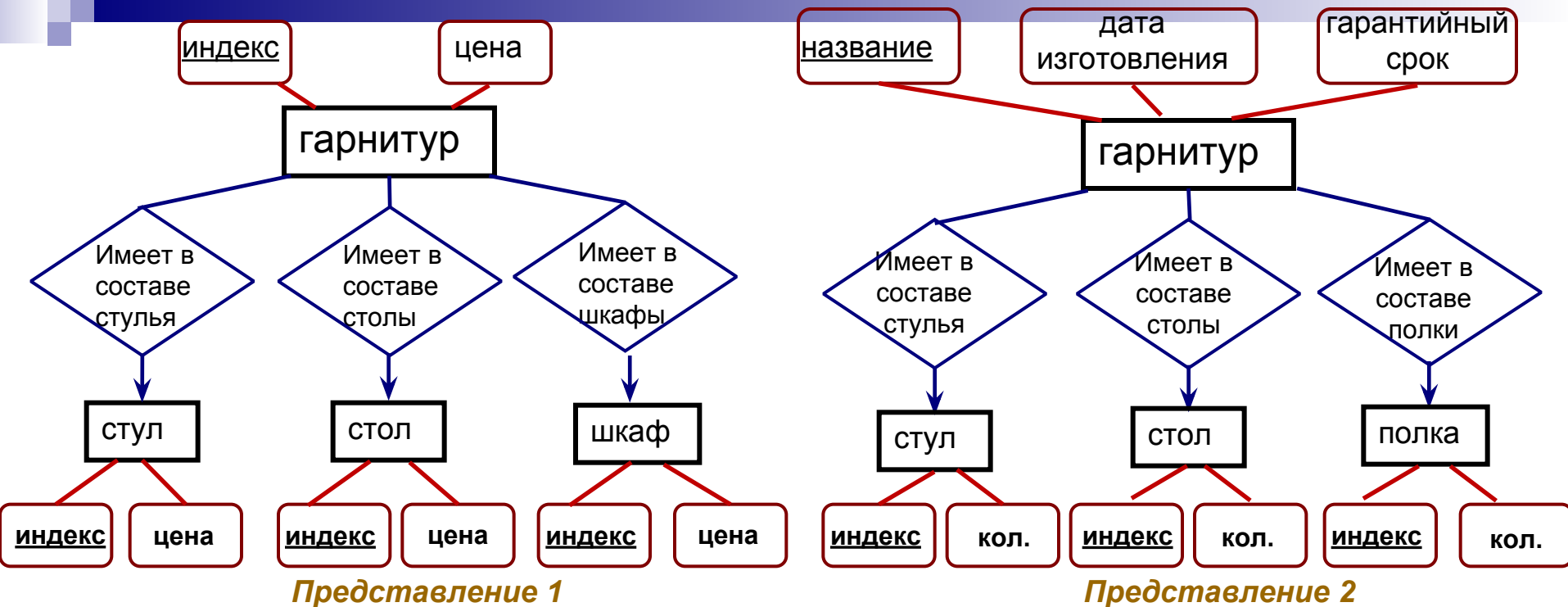
Агрегация

- Агрегация – декартово произведение сущностей
- Новая сущность формируется на основе данных о частях объектах
 - Имя, Должность, Ном_отдела = Сотрудник
 - Имя, Паспорт, Дата приема = Сотрудник
 - Имя, Номер-страхового полиса, Адрес = Сотрудник

Имя, Номер-страхового полиса, Паспорт, Должность,
Ном_отдела, Дата приема, Адрес = СОТРУДНИК

Обобщение

- Обобщение – это абстракция данных, позволяющая трактовать **класс объектов** как **ОДИН объект**
- При агрегации **части соединяются в целое**
- Обобщение фиксирует **РОДО-ВИДОВЫЕ отношения**
- **ВИД** есть **РОД** в совокупности с видовым отличием
- Ваз 2109, Ваз 2101, Ваз 2107 = **Автомобили марки ВАЗ**
- Автомобили марки ВАЗ, автомобили марки Хонда, ... = **Автомобили**
- Автомобиль, велосипед, мотоцикл, трактор = **Колесное средство передвижения**



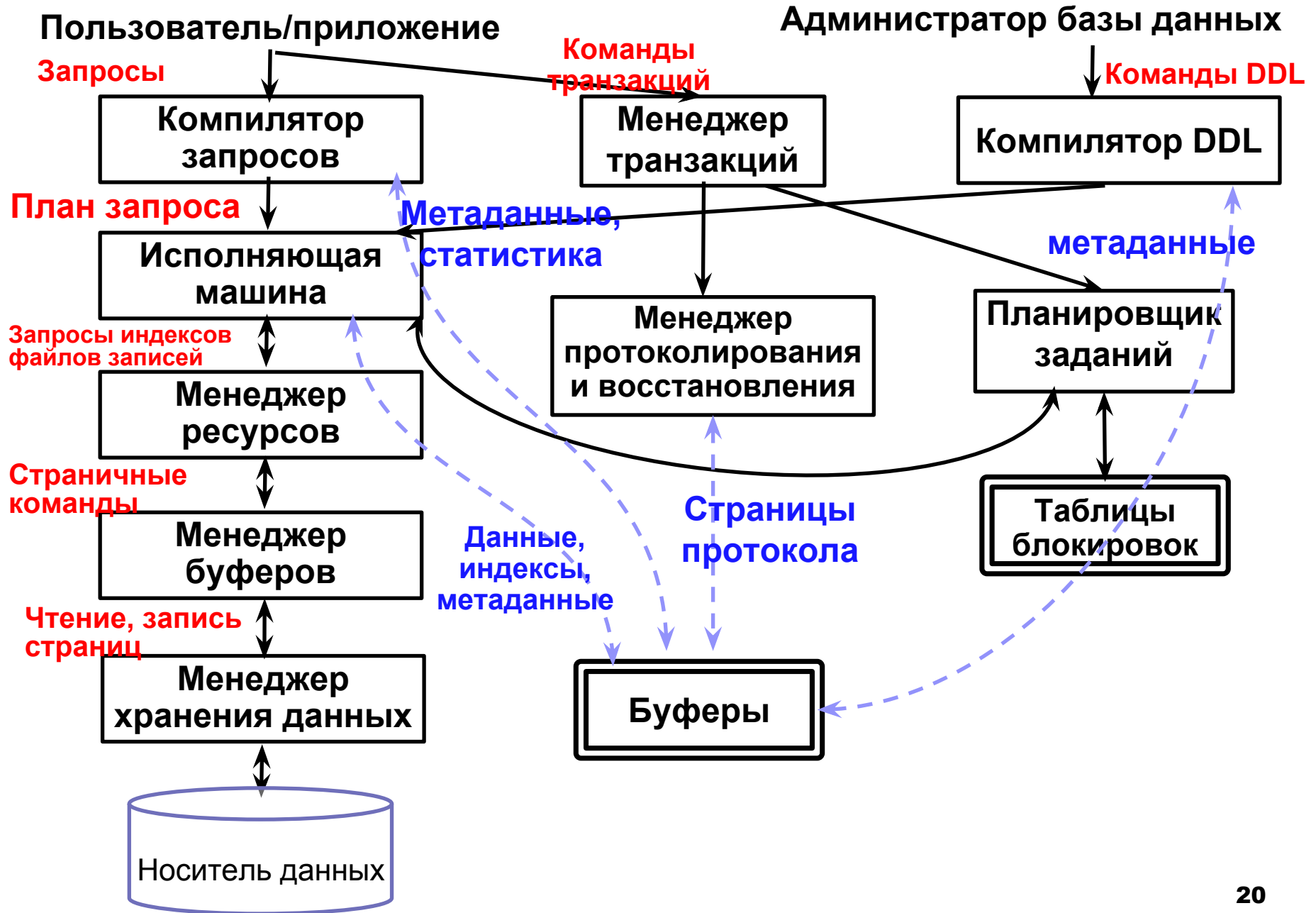
Итоговая схема



Принципы проектирования

- Достоверность
 - ❖ Привлечение экспертов
 - ❖ Детальное изучение предметной области
- Отсутствие избыточности
- Простота
 - ❖ Включайте в проект только те структурные элементы, без которых нельзя обойтись
- Выбор подходящих связей
 - ❖ Множества сущностей можно соединить разными связями. Если брать любые – возможна избыточность.
- Использование элементов адекватных типов

Структура СУБД



Компоненты СУБД

- **Компилятор запросов.** Транслирует запрос во внутренний формат системы – план запроса. Часто инструкции плана запроса представляют собой реализацию операций «реляционной алгебры»
 - ❖ Лексический анализатор. Преобразует запрос во внутреннюю древовидную структуру.
 - ❖ Препроцессор запросов: семантический анализ, функции преобразования дерева, построенного ЛА, в дерево алгебраических операций
 - ❖ Оптимизатор запросов. Преобразует запрос в наиболее эффективную последовательность фактических операций над данными
- **Исполняющая машина.** Выполняет операции запроса. Взаимодействует с большинством других компонентов. Интерпретатор.

Компоненты СУБД

- **Менеджер транзакций (МТ).**
- Запросы и другие команды группируются в транзакции
- МТ реализует выполнение транзакций так, чтобы они выполнялись
 - ❖ атомарным образом
 - ❖ изолировано.
 - ❖ сохраняя целостность БД
 - ❖ обладая свойством устойчивости.
- МТ состоит из:
 - ❖ Планировщика заданий: управление параллельными заданиями
 - ❖ Менеджера протоколирования: протоколирование действий транзакций и восстановление после сбоев
- МТ разрешает проблемы взаимных блокировок

Компоненты СУБД

- **Менеджер буферов** ответственен за разбиение ОП на буферы – участки-страницы, куда помещается содержимое дисковых блоков
- Все компоненты СУБД, обращаются к данными через буфера и менеджера буферов или через исполняющую машину
- Данные могут быть следующих категорий:
 - ❖ Собственно данные – содержимое БД
 - ❖ Метаданные – описание логической структуры БД (схема БД)
 - ❖ Статистика – информация о свойствах данных (размер отношения, сведения о значениях в колонках и т.д.)
 - ❖ Индексы – структуры, повышающие скорость доступа к данным