ПРОЕКТИРОВАНИЕ И ЗАЩИТА БД

Лекция

Безопасность систем баз данных

Нарушение целостности данных может быть вызвано рядом причин:

- сбои оборудования, физические воздействия или стихийные бедствия;
- ошибки санкционированных пользователей или умышленные действия несанкционированных пользователей;
- программные ошибки СУБД или ОС;
- ошибки в прикладных программах;
- совместное выполнение конфликтных запросов пользователей и др.

ТИПЫ ПОЛЬЗОВАТЕЛЕЙ БД

- администраторы баз данных;
- администраторы по защите данных;
- разработчики приложений;
- администраторы приложений;
- администраторы сети;
- пользователи базы данных.

ОБЯЗАННОСТИ АДМИНИСТРАТОРА БД

- **установка нового программного обеспечения, конфигурация программного и аппаратного обеспечения (вместе с системным администратором);**
- **обеспечение безопасности** : добавление и удаление пользователей, управление квотами, парольной политикой, аудит и разрешение проблем безопасности;
- **настройка производительности и мониторинг;**
- **резервное копирование и восстановление системы - процедура планового обслуживания, тестирование и настройка;**
- **локализация неисправностей, восстановление системы после сбоя.**

Решения по защите БД

- многофакторная аутентификация (подтверждения прав доступа к аккаунту).

Многофакторная аутентификация — это технология контроля доступа в несколько этапов: помимо ввода логина и пароля к аккаунту, пользователь вводит код подтверждения, полученный в SMS-сообщении (one time password — OTP), проходит голосовую верификацию или использует токен.

- разграничение доступа.
- шифрование данных.
- резервирование данных.
- аудит доступа к данным.
- тестирование нагрузки, оптимизация запросов, индексирование.
- мониторинг трафика и защита базы данных от нежелательной активности.
- применение RAID-массивов.

RAID - технология

- На втором месте по ненадежности (после человека) в компьютерных системах - является жесткий диск.
- Для увеличения надежности дисков используются дисковые массивы(**RAID**).
- **RAID-**массивы это высокопроизводительные, устойчивые к отказам подсистемы ввода-выводы.
- **RAID** - это технология для расширения пропускной способности системы ввода/вывода и обеспечения возможности хранения избыточных данных.
- Актуально для “критически важных систем” (“mission critical”), “электронной коммерции” (“e-business”) и “24x7 без простоев”.

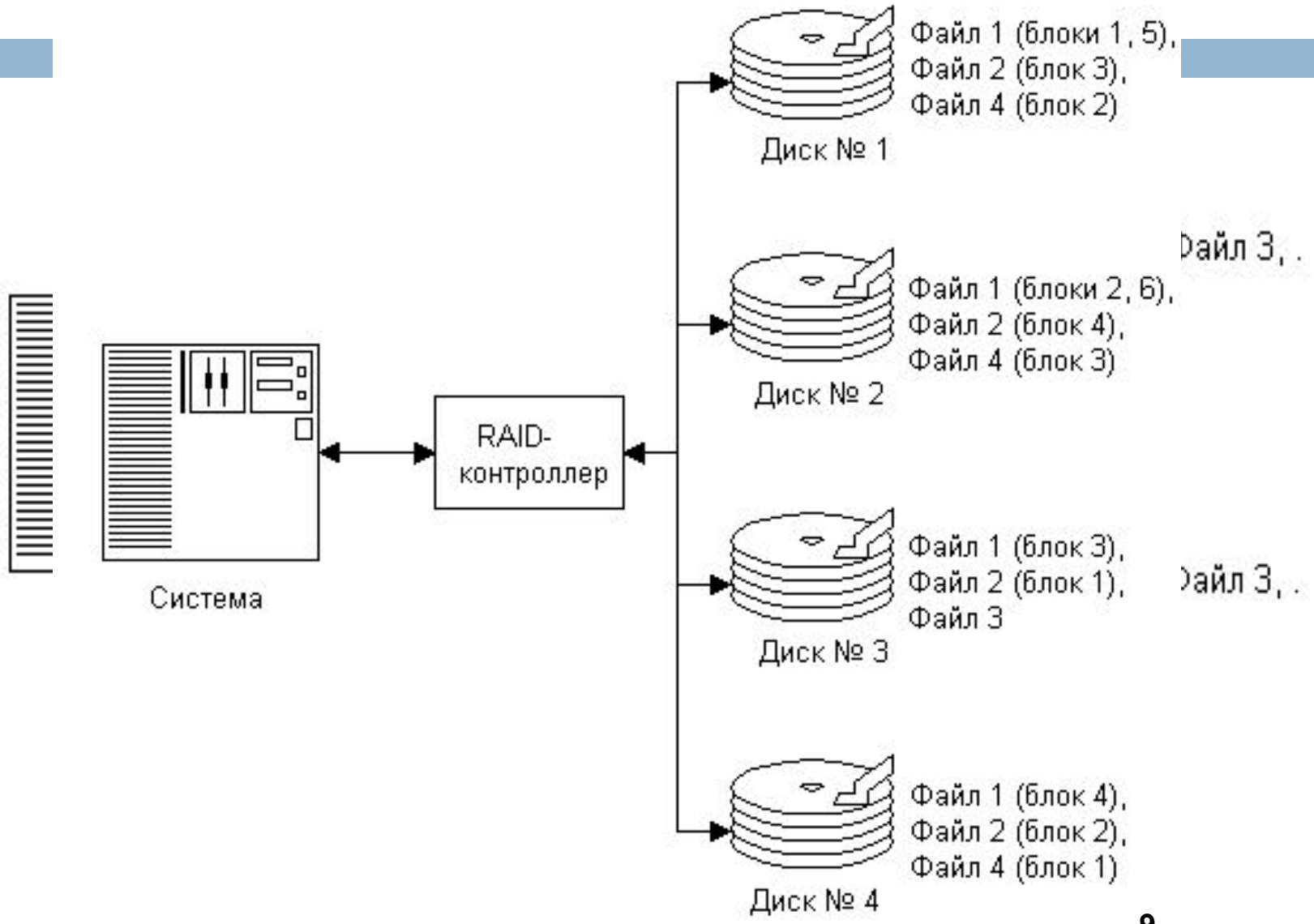
ОСНОВНЫЕ ПРИНЦИПЫ RAID

- **Избыточный массив недорогих дисков** (Redundant Arrays of Inexpensive (Independent" (независимых) Disks – **RAID**)
- Принцип RAID-технологии состоит в использовании нескольких накопителей на жестких дисках в виде **массива**, который во многих отношениях действуют как **один большой и быстрый накопитель.**
- **Массивом** называют несколько накопителей, которые централизованно настраиваются, форматируются и управляются.
- **Логический массив** – более высокий уровень представления, на котором не учитываются физические характеристики системы. Логические диски могут по количеству и объему не совпадать с физическими,

ОСНОВНЫЕ ПРИНЦИПЫ RAID

- **Страйпинг(чередование)** – процесс разделения данных на отдельные части и распределение их на несколько дисков**. RAID 0.**
- **Зеркалирование –** процесс одновременной записи данных на два или более дисков при каждой операции записи, позволяет приложениям работать до тех пор, пока цела хотя бы одна из копий. Может обеспечить и улучшение операций чтения**. RAID 1.**
- **Контроль по четности** связан с обработкой ошибок. Коды, генерируемые алгоритмом контроля по четности, используются для воссоздания данных в случае дискового сбоя(-ев). Чередование на уровне блоков с распределенным паритетом. **RAID 5.**
- **Паритет** - избыточная информация, которая вычисляется из фактических значений данных, обычно вычисляется с помощью логической операции исключающего **ИЛИ**.

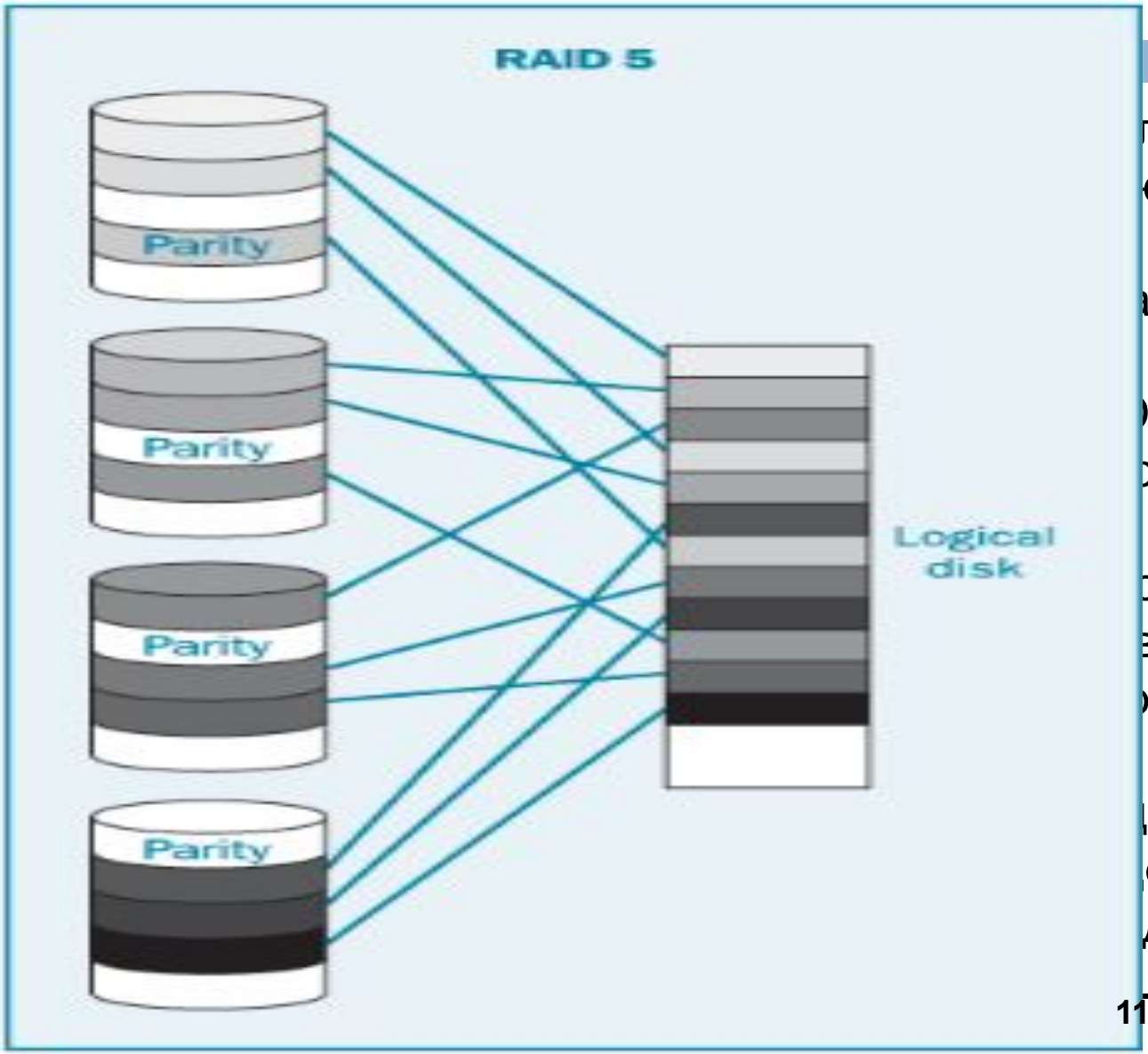
ОСНОВНЫЕ ПРИНЦИПЫ RAID



ОСНОВНЫЕ ПРИНЦИПЫ RAID

- Паритет защищает данные от отказа любого одного накопителя в массиве.
- Недостатки чередования с паритетом относятся к сложности: все байты паритета необходимо вычислять (миллионы байтов в секунду!), а для этого требуется вычислительная мощность.
- Для достижения высокой производительности нужен аппаратный контроллер, поскольку при реализации программного RAID-массива с чередованием и паритетом системный процессор просто "захлебнется" в вычислениях.
- По сравнению с зеркалированием усложняется восстановление при отказе накопителя.
- RAID 10 - это два зеркала в страйпе - т.е. пишется паралельно на два зеркала, т.е. в 2! потока. Читается сразу со всех дисков, т.е. в 4 потока. RAID 10 дает 2*х скорость записи и 4*х скорость чтения.

ОСНОВНЫЕ ПРИНЦИПЫ RAID



- Применяі
- Любой тс
- объема в
- RAID 5 яі
- производ
- Минусы
- - относи
- контроль
- - при вых
- т.к. чтени
- зачастую
- соответст
- Различие
- что мног
- **паритетс**
- приложен

ля чтения.
от 10% от
ачиваетесь
10:
рассчитать
суммы
снижается,
ва. Это же
диска и
дит к тому,
ование с
для таких

Угрозы, от которых стоит защищаться, используя резервное копирование:

- сбой/отказ/порча носителя и их группы по любым причинам. Причинами могут быть как отказ оборудования, так и внешние воздействия: затопление, 2×220 или 1×380 вольт в розетке, фаза на нарушенном контуре заземления, стихийные бедствия и т.д.;
- удаление или ошибочная модификация оригинала самим пользователем (согласно статистике, это главная опасность!);
- сбой ПО, повлекший порчу файлов, в т. ч. по причине нештатного отключения носителей и/или прерывания нормальной работы;
- кража оборудования, вандализм;
- повреждение данных вредоносным ПО. Этот вид угроз и устойчивость к нему различных видов резервного копирования рассматривается отдельно.

Типичные ошибки при резервировании БД и меры устранения:

- **Удаление предыдущей копии бэкапа до того, как будет создана новая копия бэкапа.**

Рекомендация: Не делайте новый бэкап в уже существующий файл.

- **Перезапись существующей БД при восстановлении из бэкапа.**

Рекомендации: никогда не перезаписывайте файл боевой БД, не получив письменного указания руководства.

- **Хранение бэкапа и БД на одном и том же физическом устройстве.**
- **Отсутствие проверки выполнения бэкапа и его целостности.**

Рекомендации: использовать инструменты автоматизации бэкапов, которые умеют отслеживать успешные и неуспешные бэкапы, проверять не повреждены ли бэкапы, сообщать пользователям о проблемах, и имеют обзорные средства контроля (особенно актуально, когда нужно контролировать

Типичные ошибки при резервировании БД и меры устранения:

- **Отсутствие контроля за свободным местом для бэкапа -п**ри недостатке места бэкап занимает все свободное место и аварийно завершается. При размещении бэкапа на одном диске вместе с БД может привести к остановке работы с БД, при размещении на системном диске – к поломке системы.

Рекомендации: использовать инструменты для бэкапа, которые делают прогноз размера бэкапа и предупреждают о возможной нехватке места.

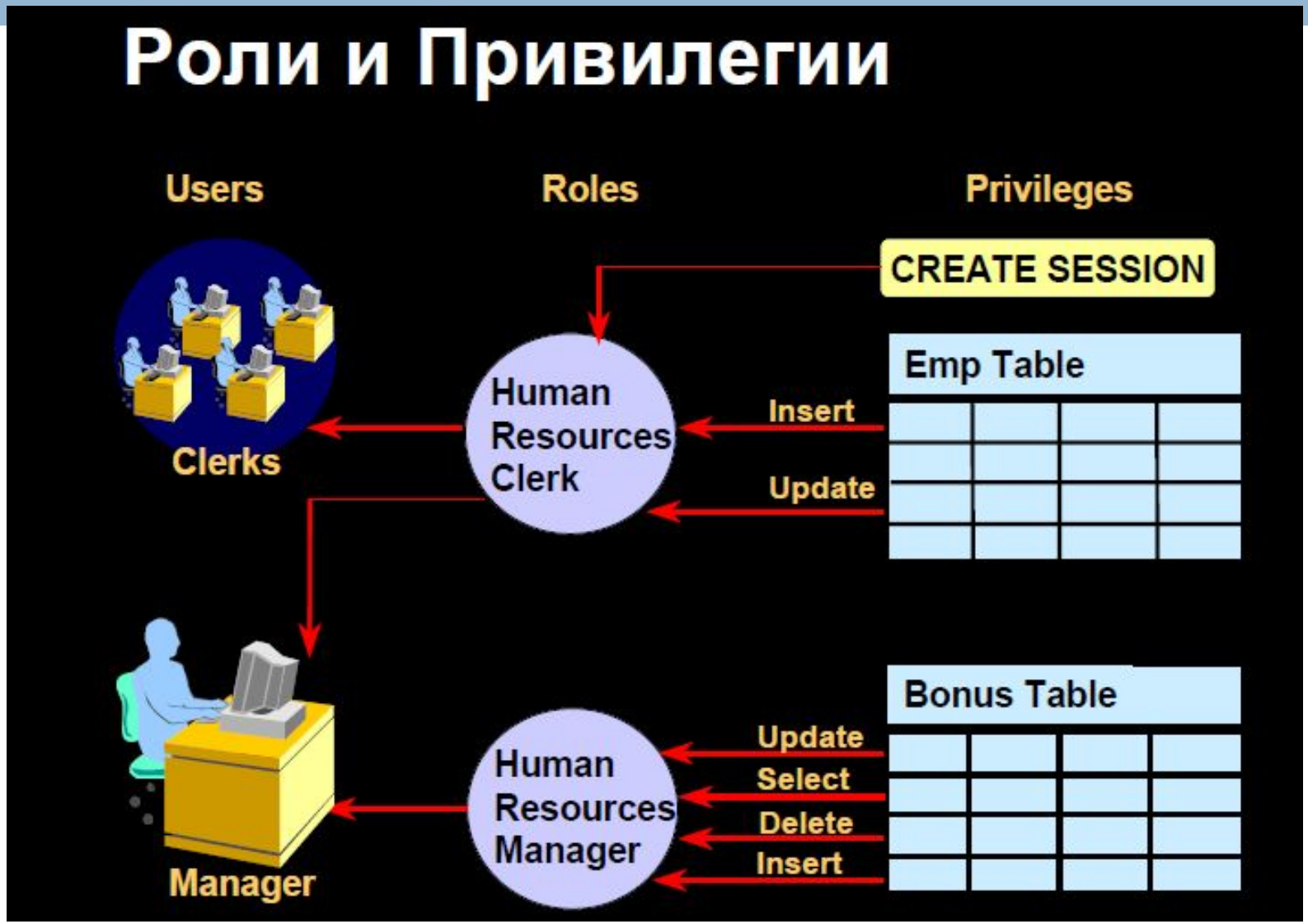
- **Отсутствие контроля времени продолжительности бэкапа – увеличение** времени подготовки бэкапа может связано с тем,что выпал диск из RAID-массива.
- Если не контролировать время исполнения бэкапа, то можно проглядеть возникшую проблему и упустить шанс исправить ее до того, как она станет большой.
  Если система резервного копирования не отслеживает состояния заданий, а запускает их просто по графику, может новый бэкап начаться в момент, пока предыдущий еще не закончился.

Рекомендации: использовать средства контроля продолжительности процесса бэкапа.

- **Исполнение бэкапа БД во время применения апдейтов ОС.**

Рекомендации: Если нельзя отключить, то назначьте апдейты ОС на такое время, когда они не смогут помешать бэкапам.

СИСТЕМА РАЗГРАНИЧЕНИЯ ДОСТУПА



Избирательные политики безопасности

- **Объект доступа** (диск, каталог, файл, системная служба, средства обработки и передачи информации) — любой элемент системы, доступ к которому может быть произвольно ограничен.
- **Субъект доступа** (пользователь, процессы, программные средства, посредством которых осуществляется доступ к объектам) — любая сущность, способная инициировать выполнение операций над объектами.
- **Механизмы разграничения доступа** оперируют с множествами операций, которые субъекты могут инициировать над объектами.
- Для каждой пары «субъект —объект» вводится множество **методов доступа и разрешенных операций,** являющееся подмножеством всего множества допустимых операций.
- Некоторые методы доступа для удобства использования объединяются в группы, называемые **правами доступа.**
- Выделяют такие типы доступа субъекта к объекту как

Избирательные политики безопасности

- Существуют модели разграничения доступа: **дискреционная (одноуровневая) и мандатная (многоуровневая).**
- Большинство ОС реализуют дискреционную модель разграничения доступа.
- Система правил дискреционной модели разграничения доступа формулируется следующим образом:

1.У каждого объекта существует владелец.

2. Владелец объекта может произвольно ограничивать (или разрешать) доступ других субъектов к данному объекту.

3. Для каждой тройки **субъект-объект-метод** возможность доступа определена однозначно.

4. Существует хотя бы один привилегированный пользователь, имеющий возможность обратиться к любому объекту по любому методу доступа.

Дискреционная модель

- Дискреционная модель разграничения доступа может быть представлена в виде **матрицы доступа**, строки которой соответствуют субъектам системы, а столбцы — объектам.
- Элементы матрицы - **права доступа** конкретного субъекта к конкретному объекту.
- Множество объектов и типов доступа к ним субъекта может изменяться в соответствии с некоторыми правилами, существующими в данной системе.
- Доступ субъекта к конкретному объекту может быть разрешен *в определенные дни , часы , в зависимости от других характеристик субъекта (контекстно-зависимое условие) или в зависимости от характера предыдущей работы.*
- Такие условия на доступ к объектам обычно используются в СУБД.

Матрица доступа

Субъект	Объект	Метод	Возможность
Ювченко	C:\ Приказы и распоряжения	Чтение	Разрешено
Ювченко	C:\ Приказы и распоряжения	Изменение	Запрещено

Дискреционная модель

- Субъект с определенными полномочиями может передать их другому субъекту (если это не противоречит правилам политики безопасности).
- Избирательное управление доступом реализует принцип **«ЧТО НЕ РАЗРЕШЕНО, ТО ЗАПРЕЩЕНО» (явное** разрешение доступа субъекта к объекту).
- Матрица доступа может формироваться на основе двух различных принципов**: централизованного и децентрализованного**.
- Возможность доступа субъектов к объектам определяется администратором (**централизованный, принудительный**). При реализации **децентрализованного** (добровольного) принципа доступом управляет владелец объекта.
- На практике часто применяют принудительный принцип управления доступа с элементами добровольного подхода.

Достоинства и недостатки дискреционной политики безопасности

Достоинства:

- простая реализация соответствующих механизмов защиты, самая хорошо изученная модель.

Недостатки:

- нельзя контролировать утечку конфиденциальной информации;
- статичность: после открытия документа права доступа не изменяются; при использовании дискреционной политики безопасности возникает вопрос определения правил распространения прав доступа и анализа их влияния на безопасность АС.
- слабая защита от вредоносных программ.

Пример (сценарий троянской атаки):

Троянская программа – программа, от которой ожидается выполнение некоторого ожидаемого действия, но выполняется недекларированное действие

	O1	**O2**
U1	OWN, R, W	**W,EXE**
U2 - нарушитель		**OWN, R, W**

Нарушитель U2 «подбрасывает» закладку O2 **пользователю U1**. Пользователь U1 запускает закладку, дает ей возможность от своего имени записать информацию из ценного объекта в себя, после чего нарушитель читает ценную информацию в доступном ему объекте- закладке.

Мандатная модель

- Применяется в совокупности с дискреционной.
- Для моделирования полномочного управления доступом используется **модель Белла-Лападулла**, включающая в себя понятия **безопасного (с точки зрения политики) состояния и перехода.**
- Основное назначение **полномочной политики (мандатной) безопасности —** регулирование доступа субъектов системы к объектам с различным уровнем критичности и предотвращение утечки информации с верхних уровней должностной иерархии на нижние, а также блокирование возможных проникновении с нижних уровней на верхние.
- Изначально полномочная политика безопасности была разработана в интересах Минобороны США для обработки информации с различными грифами секретности.

Мандатная модель

Правила мандатной модели разграничения доступа с контролем информационных потоков:

1. У любого объекта системы существует владелец.

2. Владелец объекта может произвольно ограничивать (или разрешать) доступ других субъектов к данному объекту.

3. Для каждой четверки **субъект-объект-метод-процесс** возможность доступа определена *однозначно в каждый момент времени*.

4. Существует хотя бы один привилегированный пользователь, имеющий возможность удалить любой объект.

5. Во множестве объектов выделяются множества объектов полномочного разграничения доступа. Каждый **объект имеет свой уровень конфиденциальности**.

6. Каждый **субъект имеет уровень допуска**.

Мандатная модель

7. **Запрет чтения вверх (Not Read Up — NRU):** запрет доступа по методу «чтение», если уровень конфиденциальности объекта выше уровня допуска субъекта, осуществляющего запрос.

8. Каждый процесс имеет уровень конфиденциальности, равный максимуму из уровней конфиденциальности объектов, открытых процессом.

9. **Запрет записи вниз (Not Write Down — NWD):** запрет доступа по методу «запись», если уровень конфиденциальности объекта ниже уровня конфиденциальности процесса, осуществляющего запрос.

10. Понизить гриф секретности объекта может субъект, который имеет доступ к объекту (по правилу 7) и обладает специальной привилегией.

Достоинства и недостатки МПБ

- **Достоинства** - более высокая степень надежности, нет опасности утечки конфиденциальной информации.
- **Недостатки –** реализация сложна и требует значительных ресурсов вычислительной системы, существенное снижение производительности: права доступа нужно определять для каждой операции записи/чтения; неудобство для пользователя; нельзя создать открытый объект пользователю с большими полномочиями.
- **Основная цель** мандатной модели разграничения доступа с контролем информационных потоков, — это предотвращение утечки информации определенного уровня конфиденциальности к

Использование МПБ

- В широко распространенных ОС реализована дискреционная модель. При условии обработки информации, составляющей государственную тайну, необходимо применение дополнительных средств, реализующих мандатную модель разграничения доступа.
- Программно-аппаратные средства защиты информации «Страж NT», «Dallas Lock», «Secret Net» и «Аккорд-АМДЗ», являются надстройкой над существующей программной средой АС и предназначены, в частности, для внедрения мандатной модели в системы.
- В современных СУБД реализованы механизмы управления доступом на основе уровней (меток) доступа.

Режим изолированной или замкнутой программной среды

- В дополнение в защищенных многопользовательских АС должен применяться **режим изолированной или замкнутой программной среды**.
- Для обработки информации применяется определенный перечень программных продуктов, и политикой безопасности запрещается использование других программ в целях, не имеющих отношение к выполнению функциональных обязанностей пользователями.
- Метод обеспечивает защиту компьютера от создания и запуска на нем вредоносного программного кода.
- Для каждого пользователя формируется перечень исполняемых файлов, которые могут быть им запущены, признаки возможности запуска тем или иным пользователем.
- Целесообразно осуществлять проверку и целостность исполняемых файлов при каждом их запуске.

Модели контроля целостности информации

- *В моделях контроля целостности информации* запись наверх может представлять угрозу в том случае, если субъект с низким уровнем безопасности искажает или уничтожает данные в объекте более высокого уровня.
- Исходя из этого, такая запись должна быть запрещена.
- Аналогично, чтение снизу также необходимо запретить.
- Эти наблюдения сделал в 70 годах Биб, создав три модели целостности Биба.

Модели целостности Биба

1. Мнадатная модель целостности Биба (инверсия модели Белла-Лападулла):

- нет чтения снизу, запрет субъектам на чтение информации из объекта с более низким уровнем целостности;
- нет записи вверх, запрет субъектам на запись информации в объект с более высоким уровнем целостности.
- **Когда к «чистому» субъекту (с высоким уровнем целостности) попадает информация из менее «чистого» источника, субъект «портится» и его уровень целостности, соответственно должен быть изменен.** Модель не предусматривает механизма повышения уровня целостности субъекта.

Модели целостности Биба

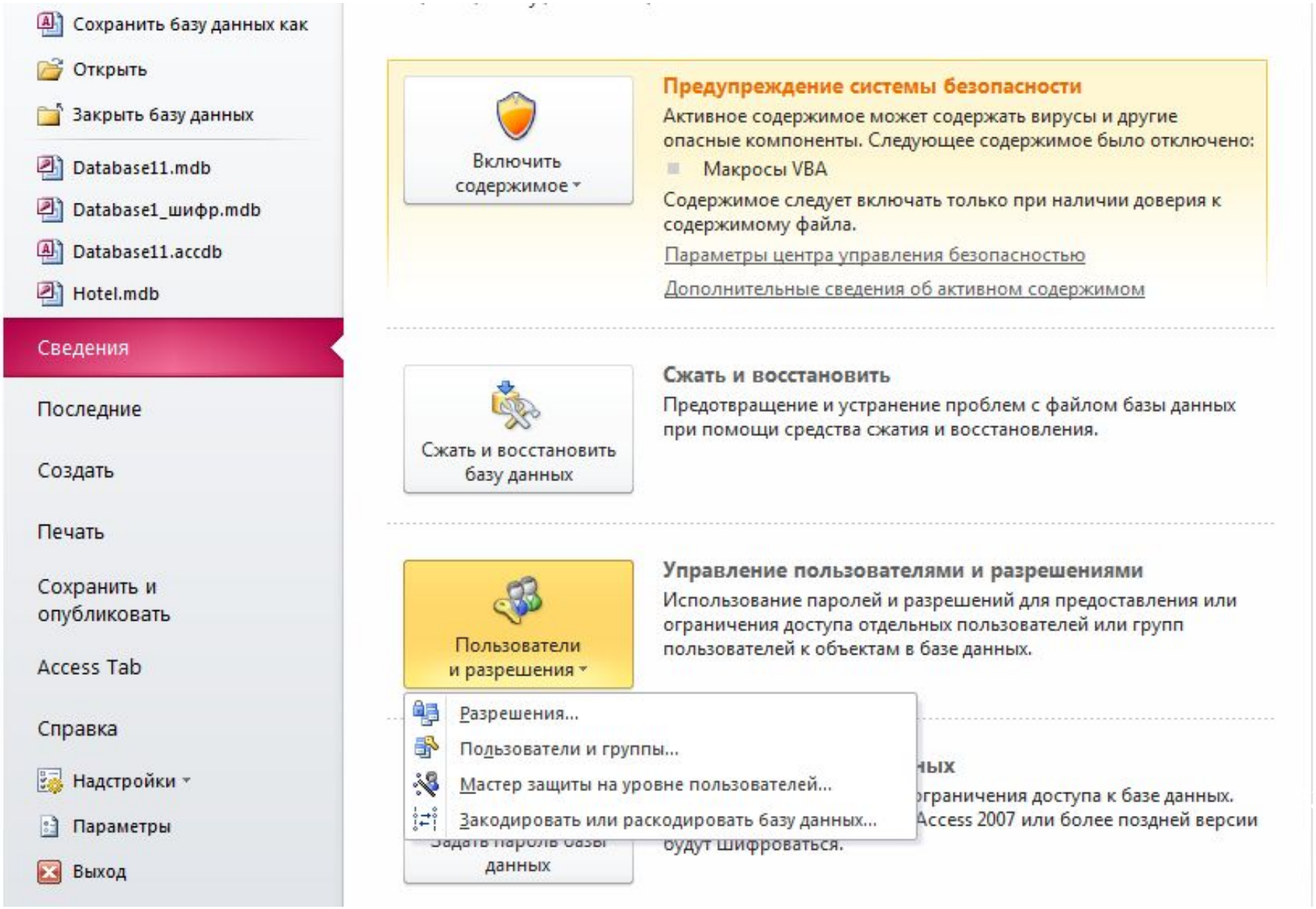
2. Модель понижения уровня субъекта - заключается в небольшом ослаблении правила чтения снизу. В модели позволяется субъекту с более высоким уровнем целостности читать информацию из объектов с более низким уровнем целостности, но при этом уровень целостности субъекта понижается до уровня целостности объекта.

3 . Модель понижения уровня объекта представляет собой ослабление правила для записи наверх, т.е. вместо полного запрета записи наверх модель разрешает такую запись, но снижает уровень целосности объекта до уровня целостности субъекта.

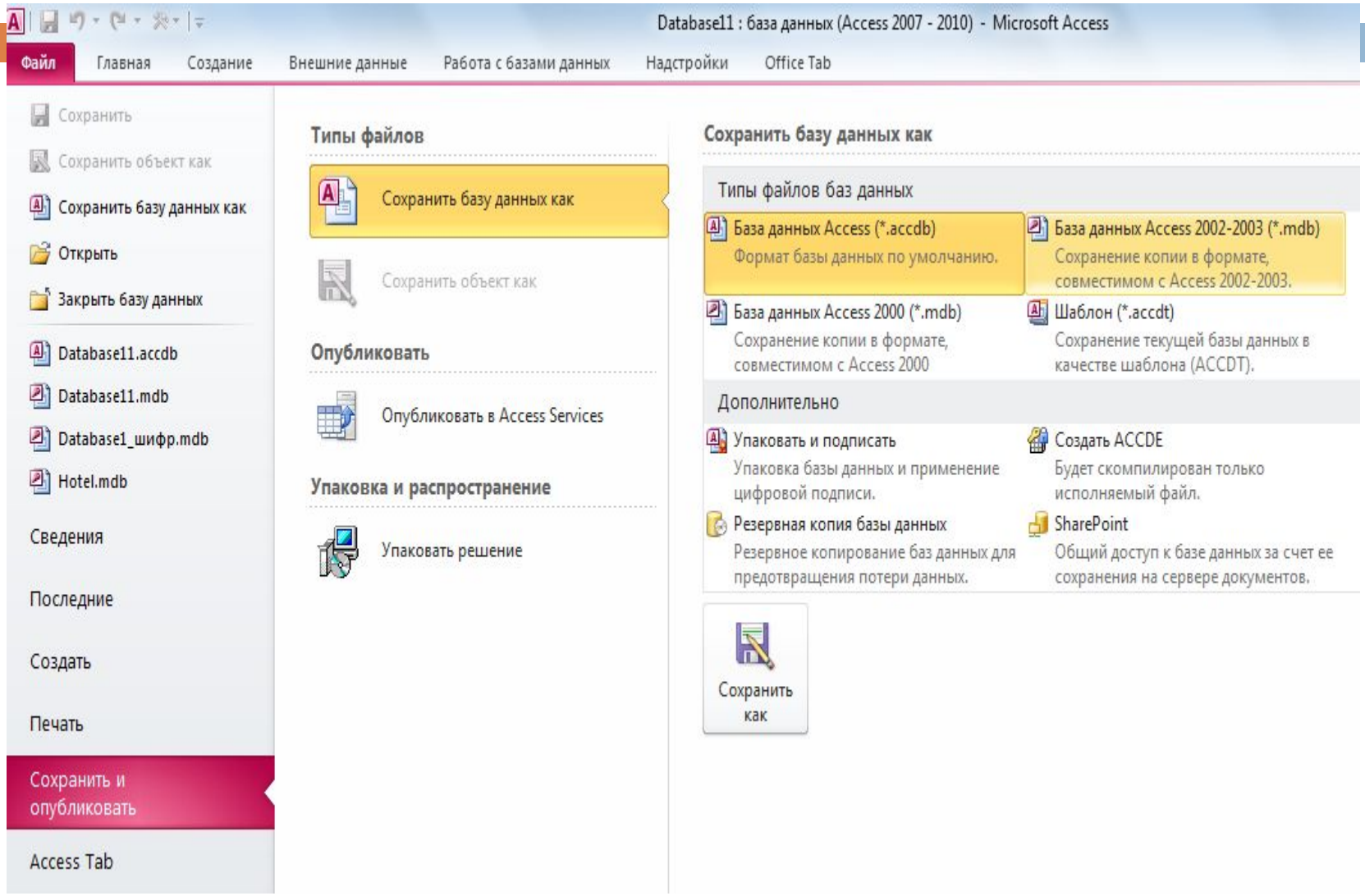
Защита данных в БД

- Для защиты данных предоставьте доступ к файлу БД на уровне пользователей только надежным пользователям, используя **разрешения файловой системы Windows.**
- В Access 2010 не поддерживается защита на уровне пользователя для баз данных, созданных в новом формате (ACCDB и ACCDE-файлы).
- При открытии базы данных из более ранней версии Access 2010, имеющей защиту на уровне пользователя, в Access 2010 эти параметры будут продолжать действовать.

Защита данных в БД



Защита данных в БД



Защита данных в БД

- Защиту базы данных обеспечивает добавление пароля, ограничивающего пользователей, которые могут открывать базу данных.
- **Внимание!** Если Вы потеряли или забыли пароль, восстановить его будет невозможно, и невозможно открыть базу данных.
- Нельзя установить **пароль базы данных (Монопольный режим)**, если для базы данных была определена **защита на уровне пользователей**, а у Вас нет разрешений **администратора для базы данных**.
- Пароль базы данных определяется дополнительно **к защите на уровне пользователей**.

Защита на уровне пользователей

- **Защита на уровне пользователей** предполагает задание администратором БД определенных разрешений отдельным пользователям и группам пользователей на объекты: **таблицы, запросы, формы, отчеты и макросы.**

Причинами использования защиты на уровне пользователей являются:

- Защита приложения от повреждения из-за неумышленного изменения пользователями таблиц, запросов, форм, отчетов и макросов, от которых зависит работа приложения;
- Защита конфиденциальных сведений в БД.

Защита на уровне пользователей

В *Microsoft Access* определены две стандартные группы:

- **администраторы – группа «*Admins*»** (учетная запись группы системных администраторов, имеющих полные разрешения на все базы данных, используемые рабочей группой. Программа установки автоматически добавляет в группу «Admins» стандартную учетную запись **пользователя «*Admin*»**)
- **пользователи – группа «*Users*»** (учетная запись группы, в которую включаются все учетные записи пользователей. *Microsoft Access* автоматически добавляет создаваемые учетные записи пользователей в группу «*Users*»), но допускается определение дополнительных групп.
- Для назначения прав может быть использован **мастер защиты.**

Отчет мастера защиты

Данный отчет содержит все сведения, необходимые для воссоздания файла рабочей группы и восстановления доступа к защищенной базе данных в случае повреждения. Напечатайте или экспортируйте этот отчет и сохраните его в надежном месте.

Незащищенная база данных:

C:\Users\irina\Desktop\Hotel.bak

Защищенная база данных:

C:\Users\irina\Desktop\Hotel.mdb

Файл рабочей группы:

C:\Users\irina\Desktop\Security.mdw

Пользователь:

Microsoft Access

Организация:

Код рабочей группы:

BNyB267oO82PjJe1wbH

Защищенные объекты:

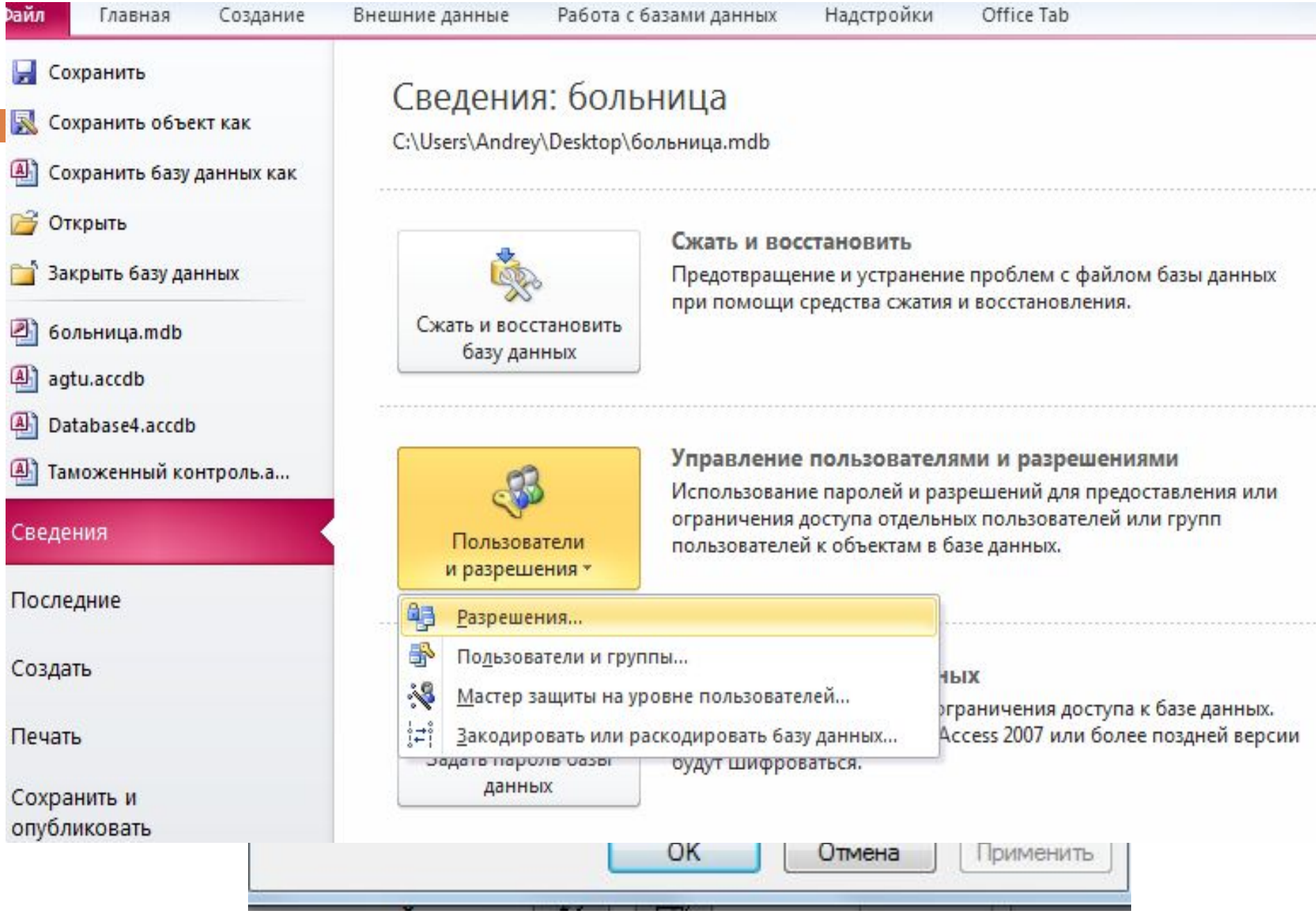
Таблицы:

ГОРОД
ДОПОЛНИТЕЛЬНЫЕ_УСЛУПИ
ЖУРНАЛ_РЕГИСТРАЦИИ
КЛИЕНТ
НОМЕР
ПОЛЬЗОВАТЕЛЬ

Защита на уровне пользователей

Изменять разрешения других пользователей на отдельные объекты базы данных могут следующие пользователи:

- члены ***группы «Admins»***, определенной в файле рабочей группы, который использовался при создании конкретной базы данных;
- ***владелец объекта*** (при активизированной системе защиты это учетная запись пользователя, которая осуществляет управление базой данных или ее объектом. По умолчанию разрешения владельца имеет учетная запись пользователя, создавшего базу данных или ее объект);
- любой пользователь, получивший на этот объект разрешения администратора.


Файл
Главная
Создание
Внешние данные
Работа с базами данных
Надстройки
Office Tab
Сохранить
Сохранить объект как
Сохранить базу данных как
Открыть
Закрыть базу данных
больница.mdb
agtu.accdb
Database4.accdb
Таможенный контроль.a...
Сведения
Последние
Создать
Печать
Сохранить и опубликовать
Сведения: больница
C:\Users\Andrey\Desktop\больница.mdb
Сжать и восстановить базу данных
Сжать и восстановить
Предотвращение и устранение проблем с файлом базы данных при помощи средства сжатия и восстановления.
Пользователи и разрешения
Управление пользователями и разрешениями
Использование паролей и разрешений для предоставления или ограничения доступа отдельных пользователей или групп пользователей к объектам в базе данных.
Разрешения...
Пользователи и группы...
Мастер защиты на уровне пользователей...
Закодировать или раскодировать базу данных...
ограничения доступа к базе данных.
Access 2007 или более поздней версии
будут шифроваться.
Задать пароль базы данных
OK
Отмена
Применить

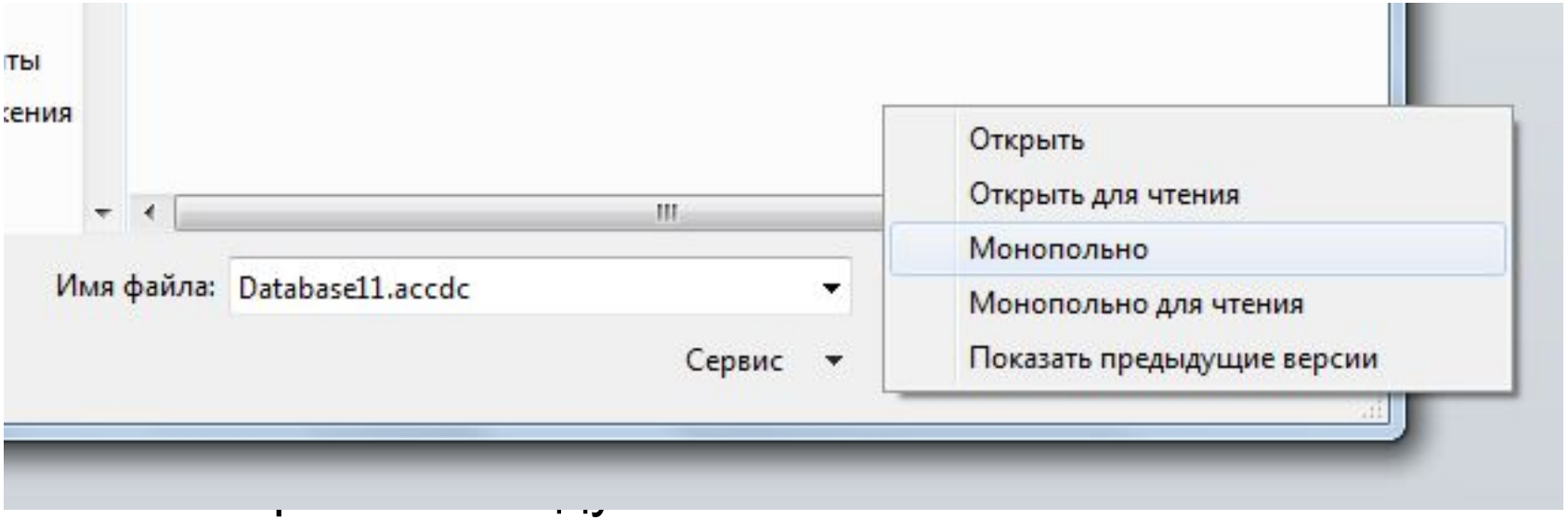
Соглашения о паролях

- Имена учетных записей могут иметь длину от 1 до 20 знаков и могут состоять из букв, цифр, пробелов и символов из расширенных наборов, за исключением следующих:
- знаки " \ [ ] : | < > + = ; , . ? *
- пробелы в начале имени;
- управляющие знаки (с кодами ASCII от 10 до 31).
- В паролях учитывается регистр знаков
- Пароль базы данных сохраняется в базе данных, а не в файле рабочей группы.
- **Файл рабочей группы** – это файл, который *Microsoft Access* считывает при запуске, содержащий **сведения о пользователях, входящих в рабочую группу**. Эти сведения включают **имена учетных записей пользователей, их пароли и имена групп, в которые входят пользователи.**

Соглашения о паролях

- *Microsoft Access* по умолчанию присваивает учетной записи пользователя «***Admin***» и **любым новым учетным записям пользователя**, создаваемым в рабочей группе, **пустые пароли**.

Шифрование с использованием пароля базы данных



- На вкладке **Файл** выберите пункт **Сведения** и нажмите кнопку **Зашифровать паролем**.
- Откроется диалоговое окно **Задание пароля базы данных**.
- Введите пароль в поле **Пароль** и повторите его в поле **Подтверждение**.

ПОДХОДЫ В ПРОЕКТИРОВАНИИ БД

1. Классический подход к проектированию.

- Подход исходит от системы документов -на входе БД имелась одна система документов, которая при использовании БД трансформировалась в другую (выходную) систему документов (таблиц, файлов).

2. Современный подход к проектированию.

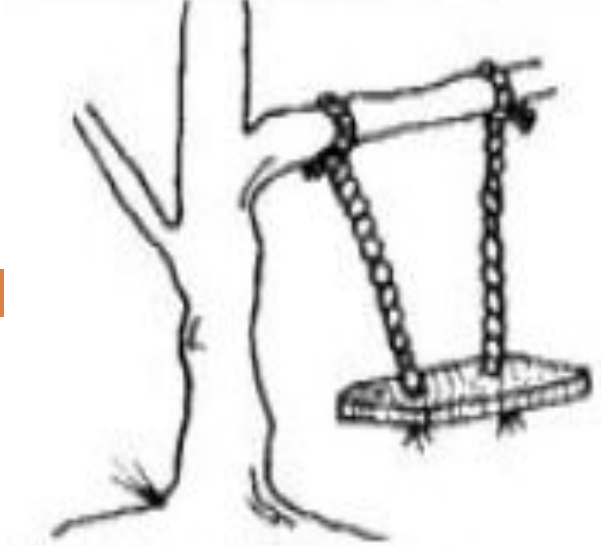
- Современный подход исходит от задач (в терминах АСУ), т.е. от приложений, под которые создается БД. Под приложением понимается программа или группа программ, предназначенных для выполнения определенных однотипных работ.

ЭТАПЫ ПРИ ПРОЕКТИРОВАНИИ БД

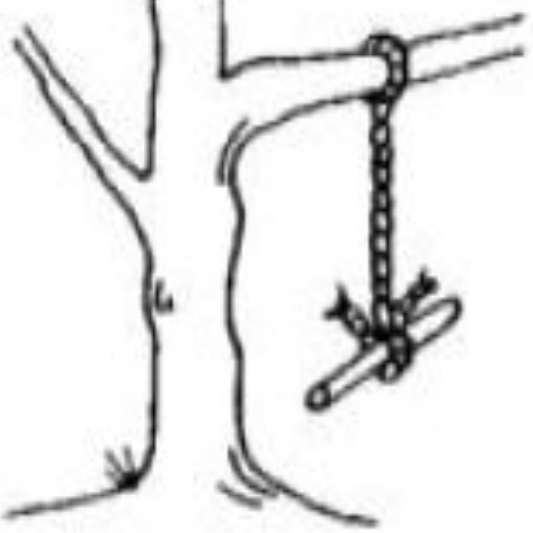
- **КОНЦЕПТУАЛЬНОЕ МОДЕЛИРОВАНИЕ (инфологическое)**
- **ЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ (даталогическое)**
- **ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ**

КОНЦЕПТУАЛЬНОЕ (инфологическое) ПРОЕКТИРОВАНИЕ БАЗЫ ДАННЫХ

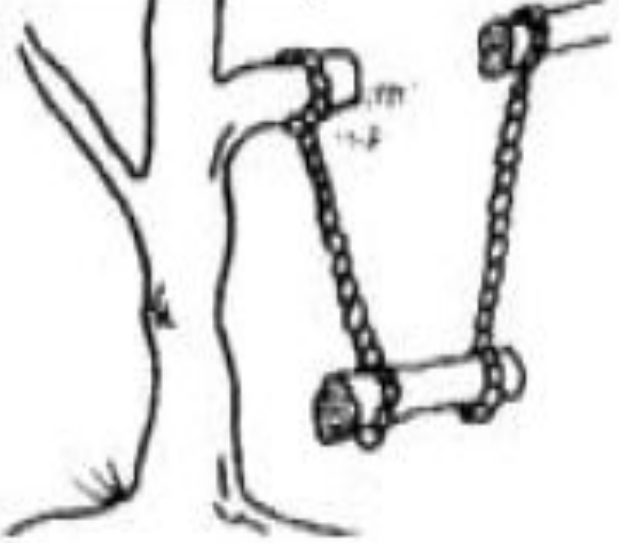
1. Создание локальной концептуальной модели данных исходя из представлений о предметной области каждого из типов пользователей.
2. Определение типов сущностей.
3. Определение типов связей.
4. Определение атрибутов, связывание их с типами сущностей, определение связей.
5. Определение доменов атрибутов.
6. Определение атрибутов, являющихся потенциальными и первичными ключами.
7. Проверка модели на отсутствие избыточности.
8. Проверка соответствия локальной концептуальной модели конкретным пользовательским транзакциям, обсуждение концептуальных моделей данных с конечными пользователями.



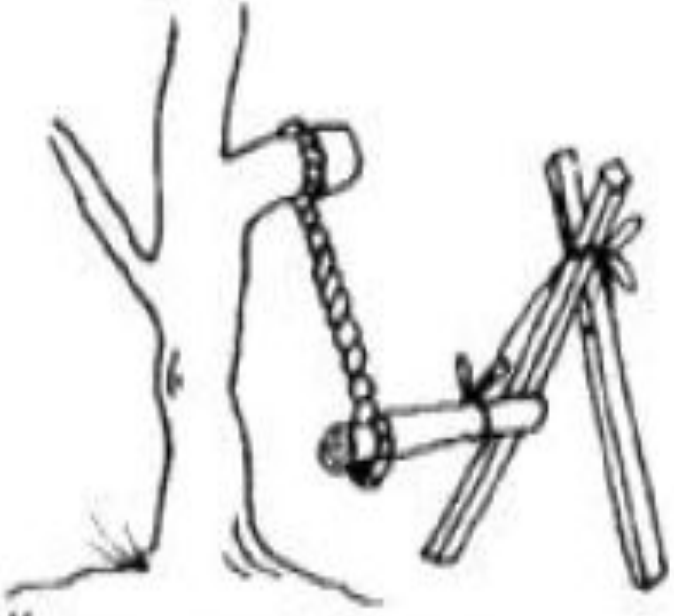
Что хотел потребитель



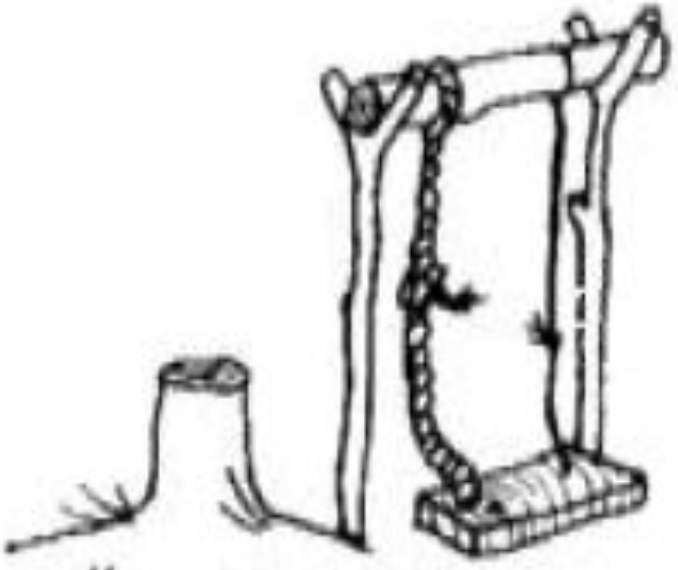
Что попросил пользователь



Как увидел аналитик



Как написал программист



Как реализовано

ПРЕДМЕТНАЯ ОБЛАСТЬ

Предметной областью называется часть реального мира, представляющая интерес для исследования (использования).

Описание предметной области содержит:

- цель, назначение, основные функции предприятия или организации, пользователи;
- описание входных и выходных документов, используемых при выполнении функций;
- описание всех используемых и создаваемых элементов данных;
- определение задач и запросов пользователей и их характеристик;
- направление развития.

ТЕХНИЧЕСКОЕ ЗАДАНИЕ-ТЗ

- В ТЗ должны быть определены основные цели приложения БД, технические требования (ТТ).
- ТТ должны содержать перечень конкретных задач, реализуемых с использованием БД.
- В разработке ТЗ участвуют инициаторы разработки проекта БД (директор или владелец предприятия).

ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ

- "Каковы ваши должностные обязанности?"
- "Какого вида задачи вы повседневно выполняете?"
- "С данными какого рода вы обычно работаете?"
- "Какого типа отчеты вы обычно используете?"
- "Дела какого типа вам необходимо отслеживать?"
- "Какие услуги предоставляет ваша компания своим заказчикам?"

МЕТОДИКИ СБОРА ФАКТОВ О ПРЕДМЕТНОЙ ОБЛАСТИ

СБОР И АНАЛИЗ ТРЕБОВАНИЙ ПОЛЬЗОВАТЕЛЕЙ

- Изучение документации;
- Проведение собеседований;
- Наблюдение за работой предприятия;
- Проведение исследований;
- Проведение анкетирования.

ИЗУЧЕНИЕ ДОКУМЕНТАЦИИ

Назначение документации	Примеры полезных источников
Описывает проблему и необходимость в базе данных	Внутренние служебные записки, электронная почта, протоколы встреч, жалобы служащих/заказчиков и документы, описывающие проблему. Обзоры и отчеты о работе
Описывает задачи предприятия, связанные с рассматриваемой проблемой	Организационный график, техническое задание и стратегический план предприятия. Технические требования для изучаемой части предприятия. Описания задач/заданий. Примеры заполненных рукописных форм и отчетов. Примеры готовых компьютеризированных форм и отчетов.
Описывает существующую систему	Различные виды блок-схем и диаграмм, словарь профессиональных терминов, словарь данных. Проект приложения базы данных. Программная документация. Руководства по обучению пользователей

СОБЕСЕДОВАНИЕ

Преимущества	Недостатки
Позволяет опрашиваемому лицу почувствовать себя участником проекта, позволяет опрашиваемому свободно открыто отвечать на вопросы	Успех зависит от навыков общения лица, проводящего собеседование
Позволяет лицу, проводящему собеседование изменить ход опроса в ответ на неожиданные комментарии со стороны опрашиваемого лица	Успех может зависеть от желания опрашиваемых лиц участвовать в интервью
Позволяет лицу, проводящему собеседование, переформулировать или иначе построить вопросы во время собеседования	Трудоемкий и дорогой, поэтому может быть непрактичным
Позволяет лицу, проводящему собеседование, наблюдать за поведением опрашиваемого лица	

НАБЛЮДЕНИЕ

Преимущества	Недостатки
Позволяет убедиться в достоверности фактов и данных	Люди, находящиеся под наблюдением, могут сознательно или бессознательно вести себя иначе
Наблюдатель может наглядно видеть, что происходит	В процессе наблюдения могут остаться незамеченными действия, выполняемые при решении задач другого уровня сложности или интенсивности, некоторые задачи могут иногда выполняться с помощью способов, отличающихся от наблюдаемых
Наблюдатель может получать данные, описывающие физические условия работы	
Относительно недорогой способ сбора фактов	

ИССЛЕДОВАНИЕ

Преимущества	Недостатки
Позволяет сэкономить время, если решение уже существует	**Может потребовать много времени**
Исследователь может узнать, как другие решают подобные проблемы, или создают системы, удовлетворяющие аналогичным требованиям	**Требует доступа к соответствующим источникам информации**
Позволяет исследователю быть в курсе современных достижений	**Исследователь может, в конечном счете, не решить проблему, поскольку такая проблема еще нигде не описана**

АНКЕТИРОВАНИЕ

Преимущества	Недостатки
Люди могут заполнять и возвращать анкеты в удобное для них время	Не все могут согласиться ответить на вопросы анкеты;
Относительно недорогой способ сбора данных с участием большого количества людей	Анкеты могут возвращать незаполненными (безответственное, равнодушное или скептическое отношение к процессу)
Люди склонны сообщать в ответах действительные факты, если проводится анонимное анкетирование (если нет каких-то особых причин для искажения информации)	Не предоставляют возможность пояснить или переформулировать неправильно понятые вопросы
Ответы могут быть сведены в таблицу и быстро проанализированы	Нет возможности наблюдать и анализировать реакцию респондента на отдельные вопросы. Подготовка опросных листов может потребовать много времени

СБОР ИНФОРМАЦИИ О ПОЛЬЗОВАТЕЛЬСКИХ ПРЕДСТАВЛЕНИЯХ

Участник собеседования	Содержание собеседования
Разработчик БД	Данные какого типа должны храниться о персонале?
Менеджер	Для каждого сотрудника это — имя и фамилия, должность, пол, дата рождения и заработная плата
Разработчик БД	Что вы собираетесь делать с данными о персонале?
Менеджер	✔ вводить данные о новых сотрудниках и удалять их при увольнении сотрудников ✔ хранить данные о текущем персонале и печатать отчеты с именами и фамилиями, должностью и зарплатой каждого сотрудника в моем отделении ✔ распределять сотрудников по инспекторам. ✔ иногда для связи с другим отделением, нужны имена и телефоны менеджеров других отделений

СБОР ИНФОРМАЦИИ О ПОЛЬЗОВАТЕЛЬСКИХ ПРЕДСТАВЛЕНИЯХ

Данные	Тип доступа	Директор	Менеджер	Инспектор	Ассистент
[illegible] отделения	Обработка				
	Запрос	x	x		
	Отчет	x	x		
Одно отделение	Обработка		x		
	Запрос		x		
	Отчет		x		
Весь персонал	Обработка				
	Запрос	x	x		
	Отчет	x	x		
Персонал одного отделения	Обработка		x		
	Запрос		x	x	

Create view as select ……

Grant select on ……..to…….

СБОР ИНФОРМАЦИИ О СИСТЕМНЫХ ТРЕБОВАНИЯХ ДЛЯ ПРИЛОЖЕНИЯ БД

- ✔ "Какие транзакции в базе данных выполняются чаще?"
- ✔ "В какие периоды бывает низкая, нормальная и высокая загрузка по выполнению наиболее важных транзакций?"
- ✔ "Имеются ли конфиденциальные данные, к которым должны иметь доступ только определенные сотрудники?"
- ✔ "За какой прошедший период необходимо хранить данные?"
- ✔ "Какие требования к работе в сети и совместному доступу предъявляются к системе базы данных?"
- ✔ "Какого типа защиту от аварийных ситуаций или потерь данных необходимо обеспечить для приложения базы данных?"

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ ДЛЯ ПРИЛОЖЕНИЯ БД

- ✔ начальный размер базы данных;
- ✔ темп роста базы данных;
- ✔ типы информационного поиска и их распределение по частоте использования;
- ✔ требования к работе в сети и совместному доступу;
- ✔ производительность;
- ✔ защита;
- ✔ резервное копирование и восстановление;
- ✔ юридические вопросы.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Начальный размер базы данных

1. Примерно 2000 сотрудников работают в более чем 100 отделениях компании.

В среднем 20 и максимум 40 сотрудников имеются в каждом отделении.

2. Приблизительно 100 000 объектов недвижимости доступны для аренды во всех отделениях. В среднем 1000 и максимум 3000 объектов недвижимости имеются в каждом отделении.

3. Примерно 60 000 владельцев недвижимости, в среднем 600 и максимум 1000 владельцев недвижимости зарегистрированы в каждом отделении…

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Темп роста базы данных

1. Каждый месяц к базе данных добавляются примерно 500 новых объектов недвижимости и 200 новых владельцев недвижимости.

2. Как только объект недвижимости становится недоступным для сдачи в аренду, соответствующая запись удаляется из базы данных. Каждый месяц удаляются примерно 100 записей об объектах недвижимости.

3. Если владелец недвижимости не предоставляет для аренды объект недвижимости в течение 2 лет, запись о нем удаляется. Каждый месяц удаляются примерно 100 записей о владельцах недвижимости.

4. Каждый месяц в компанию поступают на работу и увольняются из нее приблизительно 20 сотрудников. Запись о сотрудниках удаляется через год после их увольнения. Каждый месяц удаляются примерно 20 записей о сотрудниках…

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Типы информационного поиска и их распределение по частоте использования

1. Поиск сведений об отделении — приблизительно 10 раз в день.

2. Поиск сведений о сотруднике отделения — приблизительно 20 раз в день.

3. Поиск сведений о конкретном объекте недвижимости — приблизительно 5000 раз в день (с понедельника по четверг), приблизительно 10 000 раз в день (с пятницы по субботу). Пик нагрузки — с 12.00 до 14.00 и с 17.00 до 19.00 ежедневно.

4. Поиск сведений о владельце недвижимости — приблизительно 100 раз в день.

5. Поиск сведений о клиенте — приблизительно 1000 раз в день (с понедельника по четверг), приблизительно 2000 раз в день (с пятницы по субботу). Пик нагрузки — с 12.00 до 14.00 и с 17.00 до 19.00 ежедневно…

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Требования к работе в сети и совместному доступу

1. Все отделения должны быть объединены в сеть с централизованной БД, находящейся в головном офисе компании в Глазго, с соблюдением мер защиты.
2. Система должна предоставлять возможность одновременного доступа к ней хотя бы 2 или 3 сотрудникам из каждого отделения.
3. Необходимо предусмотреть приобретение определенного количества пользовательских лицензий для обеспечения одновременного доступа к СУБД такому числу пользователей.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Производительность

1. В утренние часы, но не в часы максимальной нагрузки, время ожидания ответа на поиск одной записи — менее 1 секунды. В часы максимальной загрузки время ожидания ответа на один поиск — менее 5 секунд.

2. В утренние часы, но не в часы максимальной загрузки, время ожидания ответа на поиск множества записей — менее 5 секунд. В часы максимальной загрузки время ожидания ответа на один поиск нескольких записей —менее 10 секунд.

3. В утренние часы, но не в часы максимальной загрузки, время выполнения операции обновления/сохранения — менее 1 секунды. В часы максимальной загрузки время выполнения операции обновления/сохранения — менее 5 секунд.

СИСТЕМНАЯ СПЕЦИФИКАЦИЯ «Риэлторское агентство» (пример)

Защита

1. База данных должна быть защищена паролем.

2. Каждому сотруднику должны быть присвоены привилегии (полномочия) доступа к базе данных согласно его пользовательскому представлению, а именно: **директора**, **менеджера**, **инспектора и ассистента.**

3. Сотруднику можно видеть только данные, необходимые для его работы, и в удобном для этого виде.

Копирование и восстановление

База данных должна копироваться **ежедневно в полночь.**

Юридические вопросы

1. В каждой стране имеются законы, регулирующие способ компьютеризированного хранения личных данных.

2. Так, если база данных содержит данные о персонале, клиентах и владельцах, необходимо изучить и учитывать любые правовые нормы, которым она должна удовлетворять.

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

- Логическая модель данных учитывает особенности выбранной модели организации данных в целевой СУБД (например, реляционная).
- На этом этапе игнорируются остальные характеристики выбранной СУБД, например, любые особенности физической организации ее структур хранения данных и построения индексов.
- Для проверки правильности логической модели данных используется **метод *нормализации***.

НОРМАЛИЗАЦИЯ

- **Нормализация** -процесс реорганизации данных путем ликвидации повторяющихся групп и иных противоречий в хранении данных с целью приведения таблиц к виду, позволяющему осуществлять непротиворечивое и корректное редактирование данных.
- **Нормальная форма** - совокупность требований, которым должно удовлетворять отношение.
- Управление данными становится очень простым, если данные организованы согласно нескольким правилам. Эти правила стали известными как правила нормализации- **правила Кодда.**

НОРМАЛИЗАЦИЯ 1 НФ

Таблица 1. Состав семьи (правило 1 нарушается)

п/п	Член семьи	Степень родства	Иждивенцы
1	Борис	Глава семьи	Екатерина, Александра
2	Тамара	Мать	Нет
3	Петр	Сводный брат	Ирина
4	Михаил	Дядя	Анна, Наталья
5	Иван	Дядя	Елена

Таблица 2. Состав семьи (в первой нормальной форме)

п/п	Член семьи	Степень родства	Иждивенец1	Иждивенец2
1	Борис	Глава семьи	Екатерина	Александра
2	Тамара	Мать	Нет	Нет
3	Петр	Сводный брат	Ирина	Нет
4	Михаил	Дядя	Анна	Наталья
5	Иван	Дядя	Елена	Нет

НОРМАЛИЗАЦИЯ 2НФ

Определение 1. Пусть R - отношение. Множество атрибутов Y функционально зависимо от множества атрибутов X (X функционально определяет Y) тогда и только тогда, когда во всех кортежах, имеющих одинаковые значения атрибутов X, значения атрибутов Y также совпадают в любом состоянии отношения R. Символически функциональная зависимость записывается

$X \to Y$.

Множество атрибутов X называется детерминантом функциональной зависимости, а множество атрибутов Y называется зависимой частью.

- Таблица содержит ключ
- Все неключевые столбцы зависят от полного ключа

Составной ключ

Курс	Студент	Тип обучения	Оценка	ак. часов
10774	Е. Онегин	очно-заочное	отлично	40
10775	В. Ленский	вебинар	хорошо	40
10776	Т. Ларина	очное	отлично	40
10777	В. Ленский	очное	хорошо	40

НОРМАЛИЗАЦИЯ 3 НФ

- В фокусе внимания – неключевые столбцы
- Нет транзитивных зависимостей

▪ Неключевые столбцы не зависят от других неключевых столбцов

№ заказа	Покупатель	Товар	Цена (шт.)	Кол-во	Сумма
1001	Е. Онегин	телефон	100	2	200
1002	В. Ленский	ноутбук	2000	1	2000
1003	Т. Ларина	утюг	50	1	50
1004	В. Ленский	утюг	50	2	100

НОРМАЛИЗАЦИЯ НФ БОЙСА-КОДДА (НФБК)

- Отношение R находится в НФ Б-К тогда и только тогда, когда **детерминанты всех функциональных** зависимостей являются **потенциальными ключами**.
- Если отношение находится в НФБК, то оно автоматически находится и в 3НФ.

ПРИВЕДЕНИЕ К 1 НФ

Накладная № 123

Дата	Покупатель	Адрес
10.01.2001	ТОО "Суперпупс"	г. Кукуевск ул. Большая Трубная д.6

Отпущен товар	Количество	Ед. Изм.	Цена за ед.	Общая стоимость
Банка стеклянная	100	шт.	3,45	345
Стакан граненый	34	шт.	1,34	45,56
Бутылка "чебурашка'	367	шт.	0,45	165,15
Вода минеральная	40	б		314
Водка "Столичная"	25	б		262,5
Пиво "Амур ДВ"	40			182,4

ОТПУСК ТОВАРОВ СО СКЛАДА
Дата
Покупатель
Город
Адрес
Товар
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость
Номер_накладной

ПРИВЕДЕНИЕ К 2 НФ

Накладная № 123

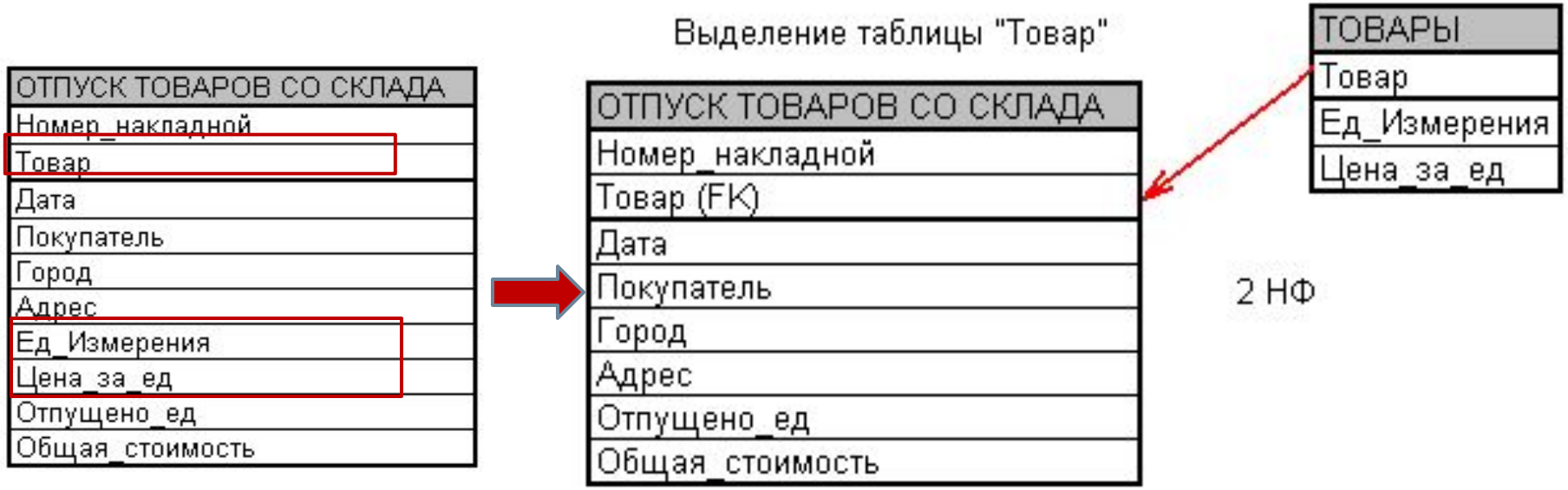
Дата	Покупатель	Адрес
10.01.2001	ТОО "Суперпупс"	г. Кукуевск ул. Большая Трубная д.6

Отпущен товар	Количество	Ед. Изм.	Цена за ед.	Общая стоимость
Банка стеклянная	100	шт.	3,45	345
Стакан граненый	34	шт.	1,34	45,56
Бутылка "чебурашка'	367	шт.	0,45	165,15
Вода минеральная	40	бутылка	7,85	314
Водка "Столичная"	25	бутылка	10,50	262,5
Пиво "Амур ДВ"	40	банка	4,56	182,4

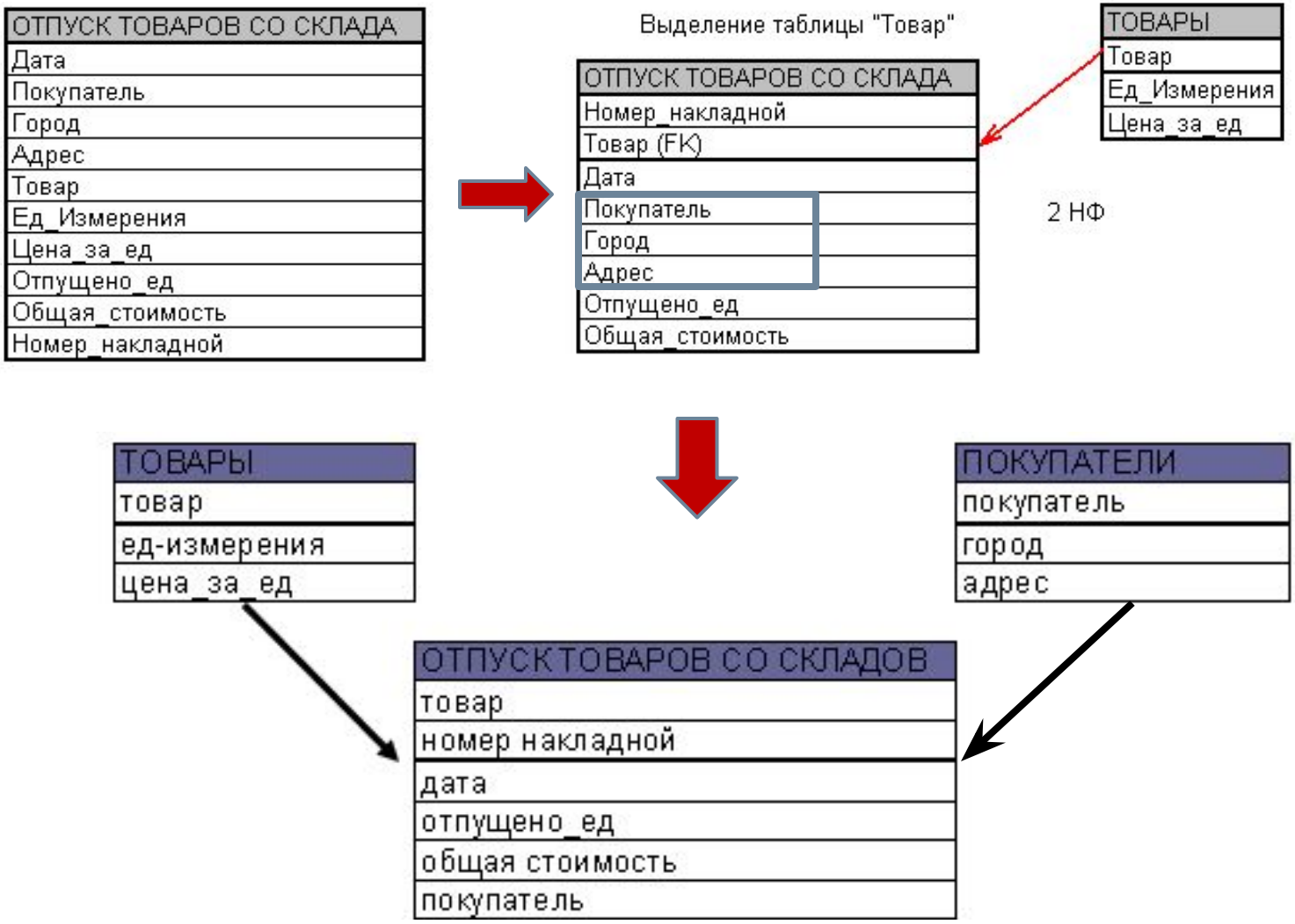
ОТПУСК ТОВАРОВ СО СКЛАДА
Дата
Покупатель
Город
Адрес
Товар
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость
Номер_накладной

ОТПУСК ТОВАРОВ СО СКЛАДА
Номер_накладной
Товар
Дата
Покупатель
Город
Адрес
Ед_Измерения
Цена_за_ед
Отпущено_ед
Общая_стоимость

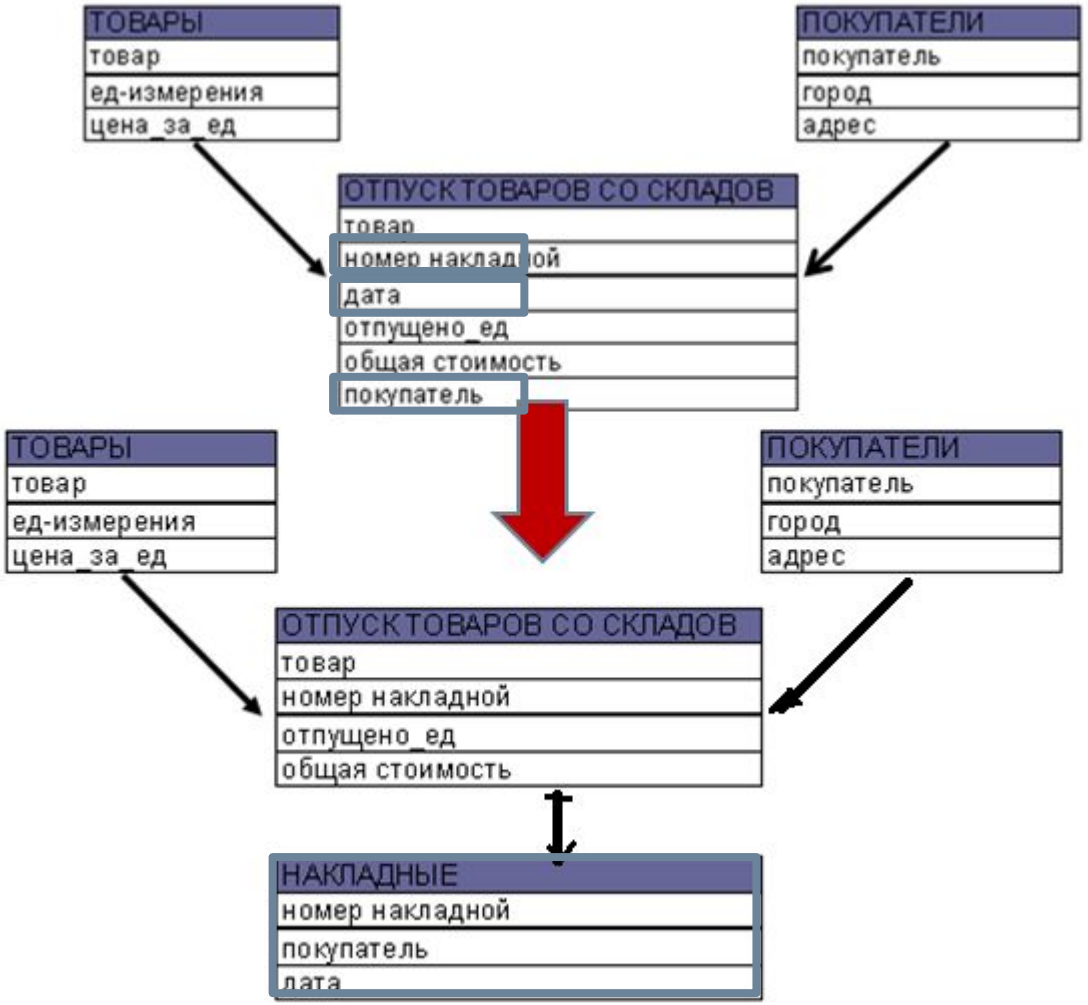
ПРИВЕДЕНИЕ К 2 НФ



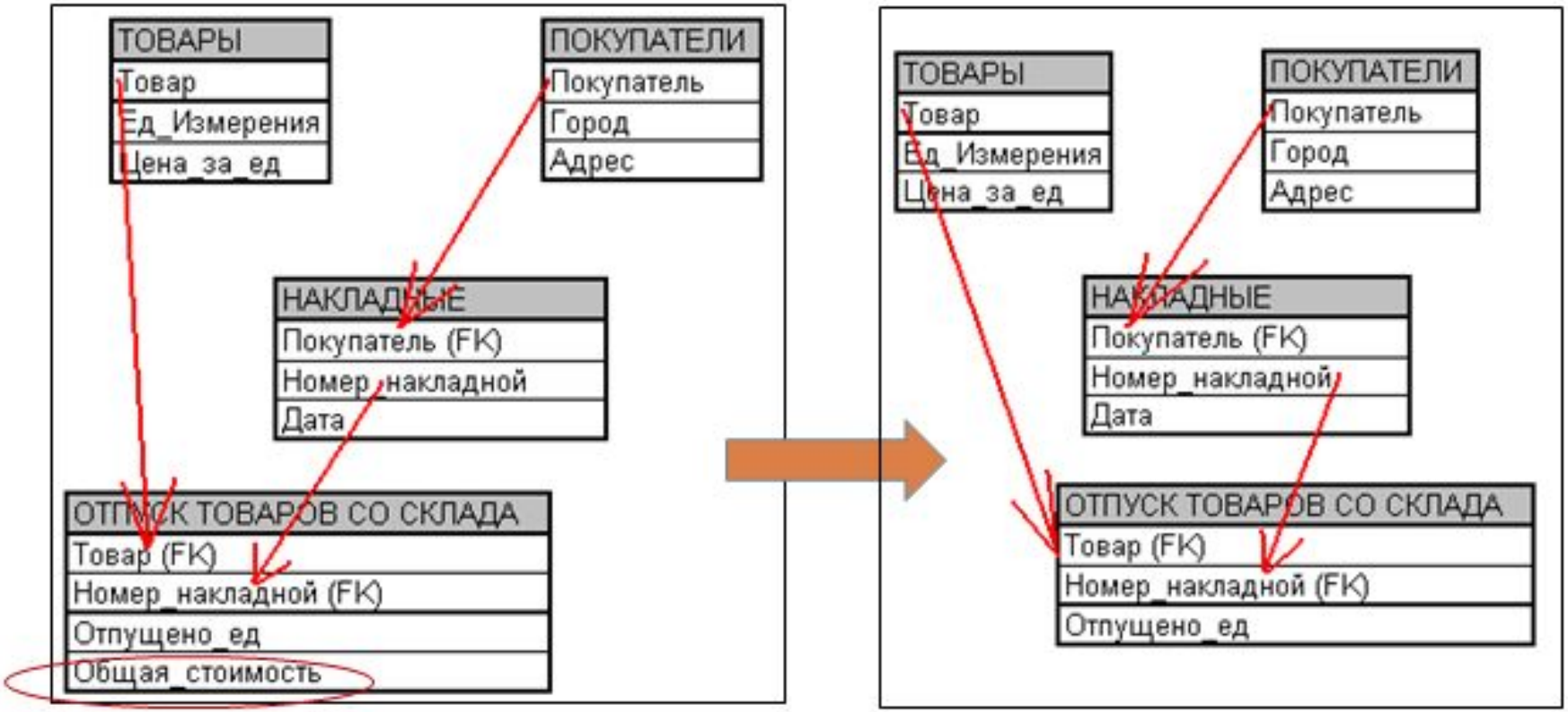
ПРИВЕДЕНИЕ К 3 НФ



ПРИВЕДЕНИЕ К 2 НФ



ПРИВЕДЕНИЕ К 3 НФ



НЕДОСТАТКИ НОРМАЛИЗАЦИИ

- Большее количество сущностей БД. Сопровождение и поддержка такой БД сложна
- Трудности построения запросов к таким БД, так как необходимо связывать несколько таблиц.
- Оперативность выборки данных низкая для высоко нормализованных БД (3 НФ).

OLTP И OLAP-СИСТЕМЫ

- **Сильно нормализованные** модели данных хорошо подходят для OLTP-приложений (On-Line Transaction Processing (OLTP)- оперативная обработка транзакций)
- OLAP-приложения (On-Line Analitical Processing (OLAP) - оперативная аналитическая обработка данных) используют **слабо нормализованные** модели данных

OLTP-ПРИЛОЖЕНИЯ

- Поддерживает большое число пользователей, работающих параллельно
- Большое значение имеет время ответа на запрос
- OLTP-системы сопряжены с интенсивными процессами чтения-записи
- Примеры OLTP-приложений - системы складского учета, системы заказов билетов, банковские системы, выполняющие операции по переводу

OLAP-ПРИЛОЖЕНИЯ

- оперируют с большими массивами данных
- добавление в систему новых данных происходит относительно редко крупными блоками (например, раз в квартал загружаются данные по итогам квартальных продаж из OLTP-приложения)
- добавленные данные в систему обычно никогда не удаляются
- перед загрузкой данные проходят различные процедуры "очистки", связанные с тем, что в одну систему могут поступать данные из многих источников, имеющих различные форматы

OLAP-ПРИЛОЖЕНИЯ

- запросы к системе являются нерегламентированными, достаточно сложными.
- скорость выполнения запросов важна, но не критична.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- Проектирование базовых отношений в среде целевой СУБД, отношений, содержащих производные данные.
- Реализация ограничений предметной области.
- Проектирование физического представления БД
- Анализ транзакций.
- Выбор файловой структуры.
- Определение индексов.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- Определение требований к дисковой памяти.
- Разработка пользовательских представлений.
- Анализ необходимости введения контролируемой избыточности.
- Организация мониторинга и настройка функционирования ОС.
- Разработка средств и механизмов защиты.

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- выбор типа носителя, способа организации данных, методов доступа (**определение пользователей базы данных, их уровней доступа, разработка и внедрение правил безопасности доступа**),
- определение размеров физического блока, управление размещением данных на внешнем носителе,
- управление свободной памятью, определение целесообразности сжатия данных и используемых методов сжатия,

ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ БД

- оценка размеров объектов базы (определение размеров табличных пространств и особенностей их размещения на носителях информации,
- определение спецификации носителей информации для промышленной системы (например, тип raid-массивов, их количество),
- разработка топологии базы данных в случае распределенной базы данных, определение механизмов доступа к удаленным данным.

СПАСИБО ЗА ВНИМАНИЕ!