**ПРАКТИЧЕСКОЕ ЗАНЯТИЕ № 9**

**Формирование аппаратных требований и схемы банка данных**

**Цель работы:** ознакомиться с принципом разработки информационного банка данных.

**Теоретическая часть**

Сегодня трудно себе представить сколько-нибудь значимую информационную систему, которая не имела бы в качестве основы или важной составляющей базу данных. Концепции и технологии баз данных складывались постепенно и всегда были тесно связанны с развитием систем автоматизированной обработки информации.

Использование баз данных и информационных систем становится неотъемлемой составляющей деятельности современного человека и организаций, шагающих в ногу со временем. В связи с этим большую актуальность приобретает построение и эффективное применение соответствующих технологий и программных продуктов.

Перерабатывать большой объем информации в заданные сроки без специальных средств практически невозможно. К сожалению, большая часть информации еще находится вне ЭВМ, что объясняется отсутствием достаточного количества и номенклатуры технических средств обработки. Но если учесть, что стоимость ЭВМ снижается, то можно предположить, что в перспективе машинная обработка информации будет основной повсеместно. В ЭВМ могут храниться и обрабатываться не только печатные тексты, но и чертежи, фотографии, запись голосов и т. д.

Методы организации процессов обработки информации, реализуемые в концепции банков данных и знаний, позволили по-новому подойти к их реализации в автоматизированных системах. Рассматривая данные как один из ресурсов автоматизированных систем (АС), можно сказать, что банк данных (БнД) централизованно управляет этим ресурсом в интересах всей системы. Наличие централизованного управления данными – главная отличительная черта банка данных.

Таким образом, банк данных – это информационная система, реализующая централизованное управление данными в интересах всех пользователей, АС, в состав которой она входит.

Предметная область – это область применения конкретного БнД. Различают банки данных, применяемые в сфере управления предприятиями и организациями, транспортом, в медицине, научных исследованиях и т. д.

Банки данных возникли в связи с потребностью в интеграции данных. Дальнейшее продвижение в этом направлении потребовало решения проблемы интеграции и процессов обработки, что привело к появлению банков знаний.

В банках знаний решаемые задачи интегрируются как по данным, так и по их обработке, возрастает интеллектуализация этих систем, цель которой – максимальное удовлетворение потребностей пользователей. Использование формальных методов преобразования и интерпретации данных позволяет автоматизировать процессы обработки накопленных в системе знаний, их получение и синтез. Одной из отличительных особенностей банков знаний является наличие в них так называемого интеллектуального интерфейса, в состав которого входят база знаний, процессор общения и программа-планировщик. Интеллектуальный интерфейс решает проблему перевода текста, написанного на естественном языке и содержащего условие задачи, в рабочую программу решения этой задачи на ЭВМ. От пользователя требуется ввести в систему корректную постановку задачи, которая его интересует, на том профессиональном языке, на котором он работает в своей предметной области, а интеллектуальный интерфейс должен выполнить всю работу, которую ранее выполнял программист.

Банки данных и знаний являются одним из основных компонентов автоматизированных систем различных уровней и типов. Их создают для многих отраслей и сфер народного хозяйства: планирования, учета, управления предприятиями, статистики, здравоохранения и др.

Концепция банков данных стала определяющим фактором при создании систем автоматизированной обработки информации. Рассматривая общие вопросы, связанные с функционированием баз данных, обсуждаются основные компоненты банков данных и получивший наибольшее распространение трехуровневый подход к построению банков данных, включающий внешний, концептуальный и внутренний уровни представления данных.

Банк данных является современной формой организации хранения и доступа к информации. Существует много определений банка данных. Мы будем использовать следующее определение: банк данных — это система специальным образом организованных данных (баз данных), программных, технических, языковых, организационно-методических средств, предназначенных для обеспечения централизованного накопления и коллективного многоцелевого использования данных.

Массив данных, хранимый в вычислительной системе, называют базой данных. База данных вместе с системой управления ею является составной частью банка данных.

База данных (БД) – именованная совокупность данных, отражающая состояние объектов и их отношений в рассматриваемой предметной области.

Банк знаний (БнЗ) — это автоматизированная система, содержащая различные виды знаний (например, концептуальные, понятийные знания) о предметной области. Хранящиеся в БнЗ знания используются для вывода новых знаний на основании специальных механизмов, имеющихся в БнЗ.

С БнД и БнЗ в процессе их создания и эксплуатации взаимодействуют пользователи различных категорий, основными из которых являются конечные пользователи. Ими являются специалисты предметных областей, для удовлетворения информационных потребностей которых и создаются БнД и БнЗ. Конечные пользователи различаются сферой интересов, информационными потребностями, квалификацией и т. п. Конечными пользователями могут быть как физические лица, так и различные вычислительные процессы, задачи, а иногда и целые системы, взаимодействующие с БнД и БнЗ. Во всех случаях результатом взаимодействия является информация, данные, знания. Банк данных – основной элемент информационной системы, в нем хранится информация по определенной проблеме в виде, допускающем компьютерную обработку. При этом сами данные образуют базу данных, а банк, наряду с базой, содержит программные средства обработки данных и реализации запросов, т. е. систему управления базой данных (СУБД). Как правило, банки данных являются системами коллективного пользования и информация, хранимая в них, доступна по телекоммуникационным сетям. В современном мире существует огромное число банков данных. В них содержатся сведения коммерческого характера, данные по библиотечным фондам, системам здравоохранения, транспорта и т. д.

Иногда в составе банка данных выделяют архивы. Основанием для этого является особый режим использования данных, когда только часть данных находится под оперативным управлением СУБД. Все остальные обычно располагаются на носителях, не управляемых СУБД. Одни и те же данные в разные моменты времени могут входить как в базы данных, так и в архивы. Банки данных могут не иметь архивов, но если они есть, то состав банка данных может входить и система управления архивами.

Эффективное управление внешней памятью являются основной функцией СУБД. Эти специализированные средства настолько важны с точки зрения эффективности, что при их отсутствии система просто не сможет решать некоторые задачи уже потому, что их выполнение будет занимать слишком много времени. При этом ни одна из таких специализированных функций (построение индексов, буферизация данных, организация доступа и оптимизация запросов) не является видимой для пользователя и обеспечивает независимость между логическим и физическим уровнями системы: прикладной программист не должен писать программы индексирования, распределять память на диске и т. д.

Основными требованиями, предъявляемыми к БнД, являются:

– адекватность отображения предметной области (полнота, целостность и непротиворечивость данных, актуальность информации (т. е. ее соответствие состоянию объекта на данный момент времени));

– возможность взаимодействия пользователей разных категорий и в разных режимах, обеспечение высокой эффективности доступа для разных приложений; дружелюбность интерфейсов и быстрое освоение системы, особенно для конечных пользователей;

– обеспечение секретности и конфиденциальности для некоторой части данных;

– определение групп пользователей и их полномочий;

– обеспечение взаимной независимости программ и данных;

– обеспечение надежности функционирования БнД, защита данных от случайного и преднамеренного разрушения; возможность быстрого и полного восстановления данных в случае их разрушения; технологичность обработки данных, приемлемые характеристики функционирования БнД (стоимость обработки, время реакции системы на запросы, требуемые машинные ресурсы и др.).

Компоненты банка данных.

Банк данных является сложной человеко-машинной системой, включающей в свой состав различные взаимосвязанные и взаимозависимые компоненты.

Информационная компонента. Ядром БнД является база данных (БД). База данных – это поименованная совокупность взаимосвязанных данных, находящихся под управлением системы управления базой данных.

Существует множество определений базы данных. Некоторые из них имеют право на существование. Другие устарели и не соответствуют современным представлениям о БД. Так, в ранних определениях базы данных указывалось на их не избыточность, отсутствие дублирования данных в них. На самом деле это не так. В базах данных может наблюдаться избыточность информации. Она может быть вызвана спецификой используемой модели данных, не позволяющей полностью устранить дублирование, или технологическими причинами (обеспечение большей надежности, сокращение времени реакции системы и др.). Но это должна быть управляемая избыточность, причины и цели возникновения которой известны администратору базы данных и управляются как им, так и системой управления базами данных (СУБД).

Системой управления базой данных называется совокупность языковых и программных средств, облегчающих для пользователей выполнение всех операций, связанных с организацией хранения данных, их корректировкой и доступом к ним.

Программные средства БнД представляют собой сложный комплекс, обеспечивающий взаимодействие всех частей информационной системы при ее функционировании.

Основу программных средств БнД представляет СУБД. В ней можно выделить ядро СУБД, обеспечивающее организацию ввода, обработки и хранения данных, а также другие компоненты, обеспечивающие настройку системы, средства тестирования, утилиты, обеспечивающие выполнение вспомогательных функций (восстановление баз данных, сбор статистики о функционировании БнД и др.). Важной компонентой СУБД являются трансляторы или компиляторы для используемых ею языковых средств.

Управляет базой данных администратор базы данных. Необходимо отметить, что при рассмотрении всего контура управления базой данных, следует учитывать и операционную систему (ОС), поскольку программы управления базой данных выполняются непосредственно под управлением ОС. Подавляющее большинство СУБД работает в среде универсальных операционных систем и взаимодействует с ОС при обработке обращений к БнД.

Для обработки запросов к БД пишутся соответствующие программы, которые представляют прикладное программное обеспечение БнД.

Языковые средства СУБД являются важнейшей компонентой банков данных, так как в конечном счете они обеспечивают интерфейс пользователей разных категорий с банком данных.

В качестве технических средств для БнД используется ЭВМ. В технической документации некоторых СУБД, а также в некоторых литературных источниках в состав БД включаются не только собственно хранимые данные о предметной области, но и описания БД. Более правильно описания баз данных считать самостоятельными компонентами БнД, даже если они и хранятся вместе с самими данными.

Организационно-методические средства представляют собой различные инструкции, методические и регламентирующие материалы, предназначенные для пользователей разных категорий, взаимодействующих с БнД.

Функционирование БнД невозможно без администраторов БнД – специалистов, обеспечивающих создание, функционирование и развитие БнД.

Структура банка данных.

Если в состав БнД входит одна БД, банк принято называть локальным, если БД несколько – интегрированным.

ВС – вычислительная система, включающая технические средства и общее программное обеспечение, базы данных, систему управления базами данных, администратора баз данных (АБД), а также обслуживающий персонал и словарь данных.

В состав любой СУБД входят языки двух типов:

– язык описания данных (с его помощью описываются типы данных, их структура и связи);

– язык манипулирования данными (его часто называют языком запросов к БД), предназначенный для организации работы сданными в интересах всех типов пользователей.

Словарь данных предназначен для хранения единообразной и централизованной информации обо всех ресурсах данных конкретного банка:

– об объектах, их свойствах и отношениях для данной ПО;

– данных, хранимых в БД (наименование, смысловое описание, структура, связи и т. п.);

– возможных значениях и форматах представления данных; источниках возникновения данных;

– кодах защиты и разграничении доступа пользователей к данным и т. п.

Структура банка данных представлена на рисунке **1**.

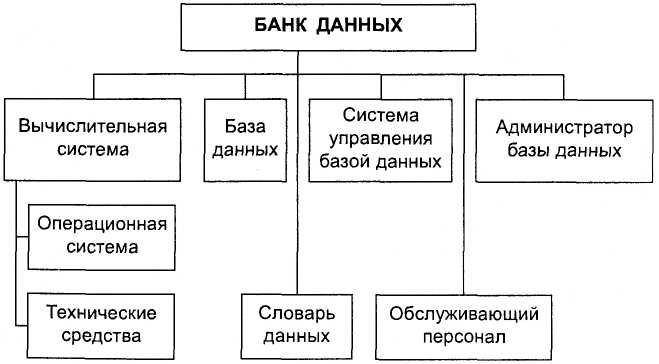


Рисунок 1 – Структура банка данных

**Ход работы:** Сформируйте аппаратные требования и постройте схему банка данных по предметной области, рассмотренной в ПЗ № 3.

Аппаратные требования для базы данных UniversityDB

1. **Требования к серверу**

Для обеспечения стабильной работы базы данных и обслуживания до 1,000 абитуриентов и других пользователей:

**Процессор (CPU):**

* Минимум: 4 ядра, 2.5 ГГц (например, Intel Xeon E3 или AMD Ryzen 5).
* Рекомендуется: 6-8 ядер, 3.0 ГГц (Intel Xeon E5, AMD Ryzen 7 или выше).

**Оперативная память (RAM):**

* Минимум: 16 ГБ (для работы с данными объемом до 300 МБ).
* Рекомендуется: 32 ГБ (для обеспечения резервов на рост объема данных и нагрузки).

**Жесткий диск (HDD/SSD):**

* Минимум: 500 ГБ (HDD) или 256 ГБ (SSD).
* Рекомендуется: 1 ТБ SSD (для высокой скорости работы с данными).

Опционально: RAID 1 для обеспечения отказоустойчивости.

**Сетевое подключение:**

* Минимум: 1 Гбит/с.
* Рекомендуется: 10 Гбит/с для быстрого взаимодействия с клиентами в локальной сети.

**Операционная система:**

* Windows 10

2. **Требования к рабочим станциям пользователей**

Для администраторов и сотрудников, работающих с базой данных через клиентские приложения:

**Процессор (CPU):** 2 ядра, 2.0 ГГц (Intel Core i3 или выше).

**Оперативная память (RAM):** 4 ГБ (рекомендуется 8 ГБ).

**Жесткий диск:** 128 ГБ SSD или 500 ГБ HDD.

**Операционная система:** Windows 10

3. **Резервное копирование и отказоустойчивость**

1. Система резервного копирования:
   * Ежедневные полные резервные копии на внешний диск (минимум 2 ТБ).
   * Автоматическое копирование в облачное хранилище (например, AWS S3 или Google Cloud).
2. Источник бесперебойного питания (UPS) для защиты от сбоев электроснабжения.

4. **Дополнительные требования**

1. Система охлаждения для сервера (с учетом нагрузки и расположения).
2. Возможность масштабирования (расширение оперативной памяти или дискового пространства).

Схема банка данных для **UniversityDB**

1. Вычислительная система

**Операционная система**:

Windows 10

**Технические средства**:

Сервер с минимальными характеристиками:

* + Процессор: 4 ядра, 2.5 ГГц
  + ОЗУ: 16 ГБ
  + Накопитель: 500 ГБ SSD

2. База данных

Имя базы данных: UniversityDB.

Содержит 7 таблиц, которые связаны с помощью первичных и внешних ключей.

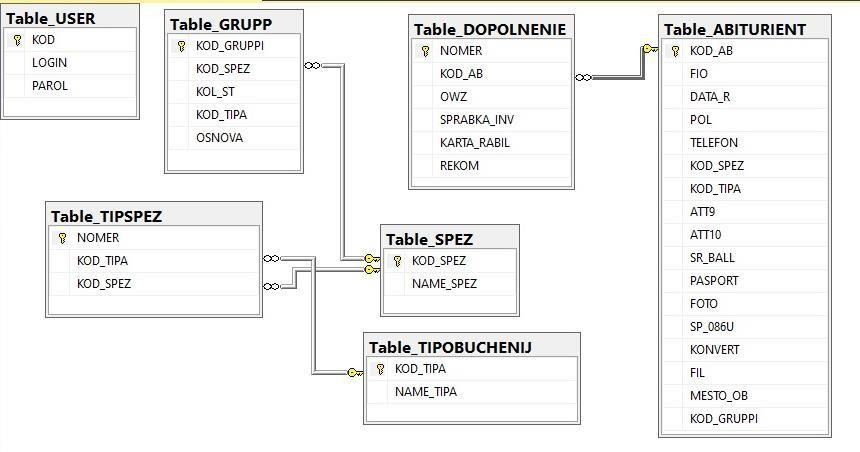


Рисунок 2 – Схема базы данных

3. Словарь данных

Содержит информацию о размерах таблиц, количестве записей и общем объеме данных.

Таблица 1 – Словарь данных

|  |  |  |  |
| --- | --- | --- | --- |
| **Таблица** | **Размер записи (байт)** | **Количество записей** | **Общий объем (байт)** |
| Table\_USER | 104 | 100 | 10,400 |
| Table\_TIPSPEZ | 12 | 50 | 600 |
| Table\_SPEZ | 104 | 100 | 10,400 |
| Table\_TIPOBUCHENIJ | 104 | 20 | 2,080 |
| Table\_GRUPP | 66 | 200 | 13,200 |
| Table\_ABITURIENT | 135 | 1,000 | 135,000 |
| Table\_DOPOLNENIE | 208 | 500 | 104,000 |
| **Итого** |  |  | **275,680** |

4. Система управления базой данных (СУБД)

* Выбранная СУБД: **MS SQL**.
* Поддержка реляционных данных, работы с транзакциями и индексами.

5. Обслуживающий персонал

**Разработчики базы данных**: ответственные за проектирование и создание структуры базы данных.

**Технический персонал**: обеспечивает работоспособность вычислительной системы.

6. Администратор базы данных

Отвечает за управление, настройку, резервное копирование и восстановление данных, а также за контроль доступа к базе.