

Лекция №8

Интеграция информационных технологий.

Учебные вопросы:

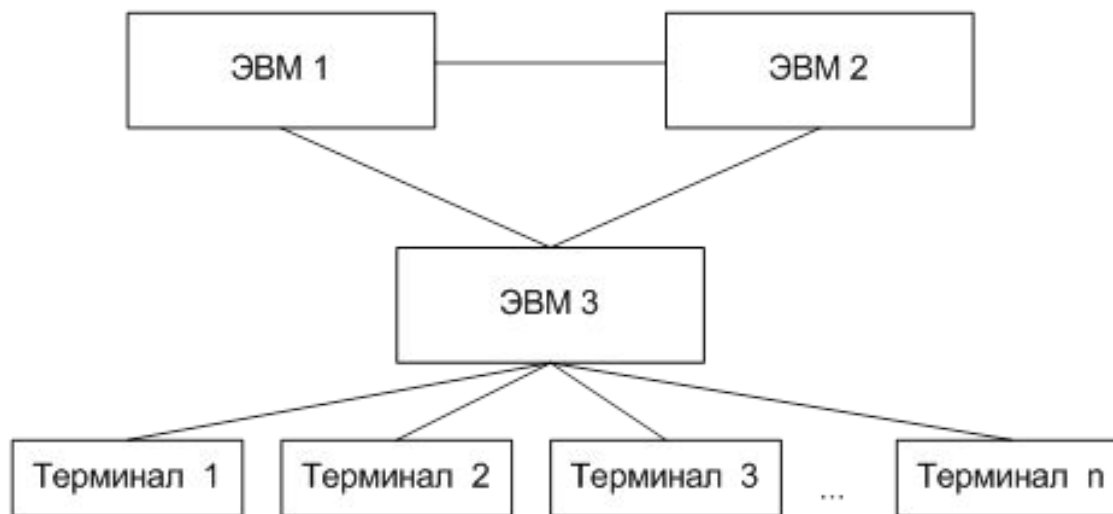
1. Технологии распределенных вычислений (РВ)
2. Распределенные базы данных
3. Технологии и модели "Клиент-сервер"

1. Технологии распределенных вычислений (РВ)

Принцип централизованной обработки данных (рис. 5.1) не отвечал высоким требованиям к надежности процесса обработки, затруднял развитие систем и не мог обеспечить необходимые временные параметры при диалоговой обработке данных в многопользовательском режиме. Кратковременный выход из строя центральной ЭВМ приводил к роковым последствиям для системы в целом.



Появление персональных компьютеров потребовало нового подхода к организации систем обработки данных, к созданию новых информационных технологий. Возникло логически обоснованное требование перехода от использования отдельных ЭВМ в системах централизованной обработки данных к распределенной обработке данных (рис. 5.2).



Распределенная обработка данных - обработка данных, выполняемая на независимых, но связанных между собой компьютерах, представляющих распределенную систему

В основе распределенных вычислений лежат две основные идеи:

- много организационно и физически распределенных пользователей, одновременно работающих с общими данными - общей базой данных (пользователи с разными именами, которые могут располагаться на различных вычислительных установках, с различными полномочиями и задачами);
- логически и физически распределенные данные, составляющие и образующие тем не менее, общую базу данных (отдельные таблицы, записи и даже поля могут располагаться на различных вычислительных установках или входить в различные локальные базы данных).

Для реализации распределенной обработки данных были созданы многомашинные ассоциации, структура которых разрабатывается по одному из следующих направлений:

-многомашинные вычислительные комплексы (МВК);

-компьютерные (вычислительные) сети.

Многомашинный вычислительный комплекс - группа установленных рядом вычислительных машин, объединенных с помощью специальных средств сопряжения и выполняющих совместно единый информационно-вычислительный процесс. Под процессом понимается некоторая последовательность действий для решения задачи, определяемая программой.

Многомашинные вычислительные комплексы могут быть:

- ***локальными***, при условии установки компьютеров в одном помещении, не требующих для взаимосвязи специального оборудования и каналов связи;

- ***дистанционными***, если некоторые компьютеры комплекса установлены на значительном расстоянии от центральной ЭВМ и для передачи данных используются телефонные каналы связи.

Компьютерная (вычислительная) сеть - вычислительная система, включающая в себя несколько компьютеров, терминалов и других аппаратных средств, соединенных между собой линиями связи, обеспечивающими передачу данных

Терминал - устройство, предназначенное для взаимодействия пользователя с вычислительной системой или сетью ЭВМ. Состоит из устройства ввода (чаще всего это клавиатура) и одного или нескольких устройств вывода (дисплей, принтер и т.д.).

2. Распределенные базы данных

Распределенная база данных - это совокупность логически взаимосвязанных баз данных, распределенных в компьютерной сети.

Система управления распределенной базой данных - это программная система, которая обеспечивает управление распределенной базой данных и прозрачность ее распределенности для пользователей.

Распределенная база данных может объединять базы данных, поддерживающие любые модели (иерархические, сетевые, реляционные и объектно-ориентированные базы данных) в рамках единой глобальной схемы.

Основные принципы создания и функционирования распределенных баз данных:

- прозрачность расположения данных для пользователя (иначе говоря, для пользователя распределенная база данных должна представляться и выглядеть точно так же, как и нераспределенная);
- изолированность пользователей друг от друга (пользователь должен "не чувствовать", "не видеть" работу других пользователей в тот момент, когда он изменяет, обновляет, удаляет данные);
- синхронизация и согласованность (непротиворечивость) состояния данных в любой момент времени.

Из основных вытекает ряд дополнительных принципов:

- локальная автономия (ни одна вычислительная установка для своего успешного функционирования не должна зависеть от любой другой установки);
- отсутствие центральной установки (следствие предыдущего пункта);

- независимость от местоположения (пользователю все равно, где физически находятся данные, он работает так, как будто они находятся на его локальной установке);
- непрерывность функционирования (отсутствие плановых отключений системы в целом, например для подключения новой установки или обновления версии СУБД);
- независимость от фрагментации данных (как от горизонтальной фрагментации, когда различные группы записей одной таблицы размещены на различных установках или в различных локальных базах, так и от вертикальной фрагментации, когда различные поля-столбцы одной таблицы размещены на разных установках);

3. Технологии и модели "Клиент-сервер"

В технологиях "Клиент-сервер" отступают от одного из главных принципов создания и функционирования распределенных систем - отсутствия центральной установки. Поэтому можно выделить две основные идеи, лежащие в основе клиент-серверных технологий:

- общие для всех пользователей данные на одном или нескольких серверах;
- много пользователей (клиентов), на различных вычислительных установках, совместно (параллельно и одновременно) обрабатывающих общие данные.

Под сервером в широком смысле понимается любая система, процесс, компьютер, владеющие каким-либо вычислительным ресурсом (памятью, временем, производительностью процессора и т. д.).

Клиентом называется также любая система, процесс, компьютер, пользователь, запрашивающие у сервера какой-либо ресурс, пользующиеся каким-либо ресурсом или обслуживаемые сервером иным способом.

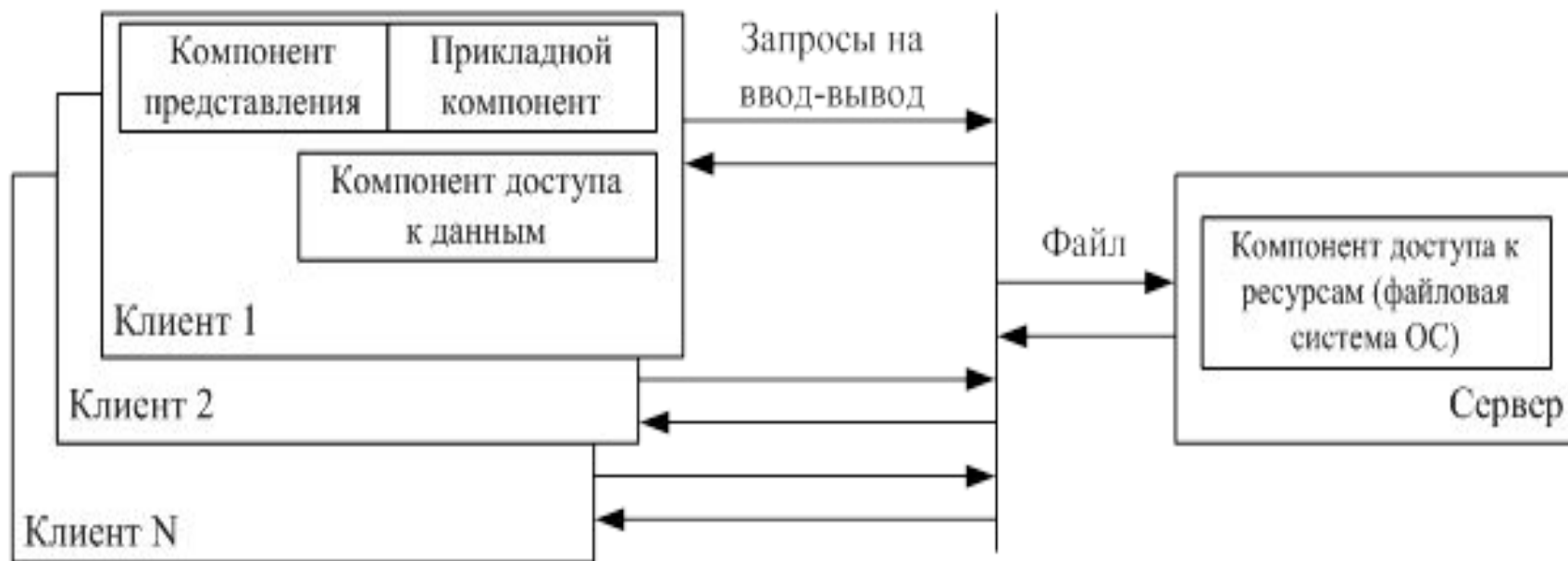


Рис 5.3 - Модель файлового сервера



Рис 5.4. Модель удаленного доступа к данным (RDA-модель)

Интероперабельность (многопротокольность) СУБД - способность СУБД обслуживать прикладные программы, первоначально ориентированные на разные типы СУБД. Иначе говоря, специальный компонент ядра СУБД на сервере (так называемый драйвер ODBC) способен воспринимать, обрабатывать запросы и направлять результаты их обработки на клиентские установки, функционирующие под управлением реляционных СУБД других, не "родных" типов.

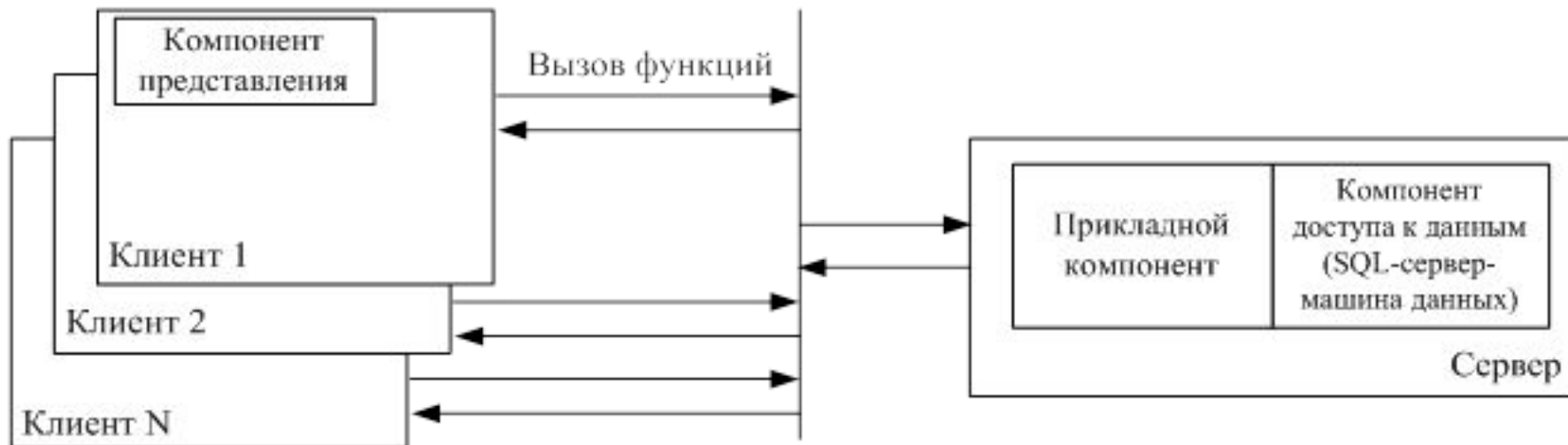


Рис. 5.5 Модель сервера базы данных (DBS-модель)

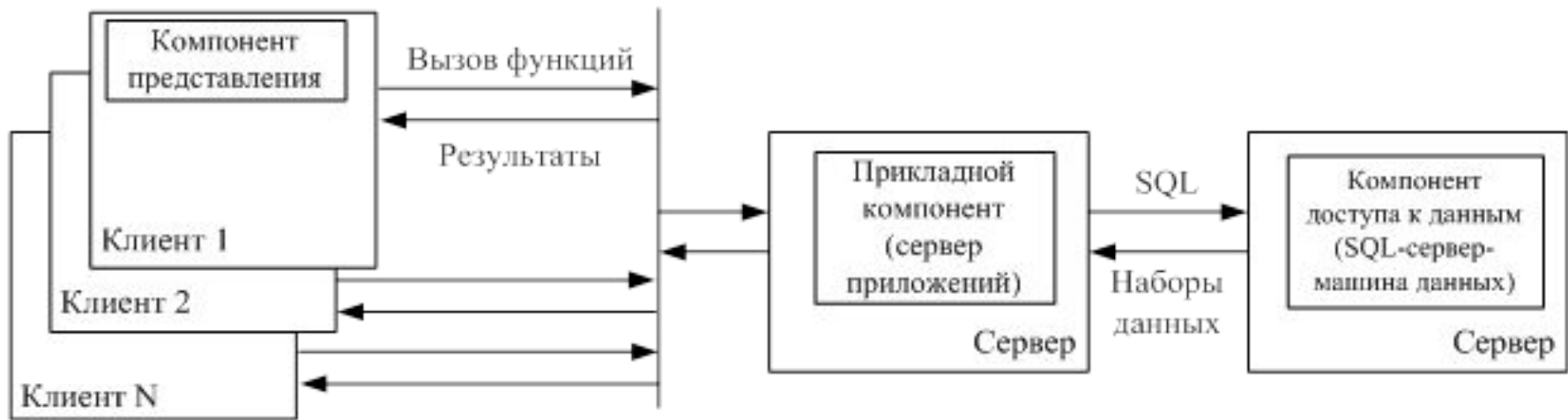


Рис. 5.6. Модель сервера приложений (AS-модель)

Репликой называют особую копию базы данных для размещения на другом компьютере сети с целью автономной работы пользователей с одинаковыми (согласованными) данными общего пользования

Тиражирование (или репликация,) - создание дублирующих копий (репликатов) объектов данных на разных узлах с целью повышения доступности и/или сокращения времени доступа к критически важным данным.

При этом, однако, возникают две проблемы обеспечения одного из основополагающих принципов построения и функционирования распределенных систем (а именно, - непрерывности согласованного состояния данных):

- обеспечение согласованного состояния во всех репликах количества и значений общих данных;

- обеспечение согласованного состояния во всех репликах структуры данных.

Обеспечение согласованного состояния общих данных, в свою очередь, основывается на реализации одного из двух принципов:

- принципа непрерывного размножения обновлений (любое обновление данных в любой реплике должно быть немедленно размножено);
- принципа отложенных обновлений (обновления реплик могут быть отложены до специальной команды или ситуации).