10. Распределенные вычислительные системы и сети

10.1 Понятие о вычислительных сетях. Классификация и свойства, топология, примеры.

Компьютерная сеть (вычислительная сеть, сеть передачи данных) — система связи компьютеров или вычислительного оборудования (серверы, маршрутизаторы и другое оборудование).

Классификация вычислительных сетей.

По способу организации сети подразделяются на реальные и искусственные.

Искусственные сети (псевдосети) позволяют связывать компьютеры вместе через витую пару, а ранее использовались последовательные или параллельные порты, и не нуждаются в дополнительных устройствах. Искусственные сети используются, когда необходимо перекачать информацию с одного компьютера на другой.

Реальные сети позволяют связывать компьютеры с помощью специальных устройств коммутации и физической среды передачи данных.

Все многообразие компьютерных сетей можно классифицировать по группе признаков:

1. По территориальной распространенности сети могут быть локальными, глобальными, и региональными.

-Локальные (LAN) - это сети, расположенные в пределах одного здания.

-Региональные (MAN) - расположенные на территории города или области.

-Глобальные (WAN) - на территории государства или группы государств, например, всемирная сеть Internet.

Локальные сети являются сетями закрытого типа, доступ к ним разрешен только ограниченному кругу пользователей, для которых работа в такой сети непосредственно связана с их профессиональной деятельностью. Глобальные сети являются открытыми и ориентированы на обслуживание любых пользователей.

2. По архитектуре:

- Ethernet

- TokenRing

- AppleTalk

- ARCnet

3. Ведомственная принадлежность

По принадлежности различают ведомственные и государственные сети.

- Ведомственные принадлежат одной организации и располагаются на ее территории.

- Государственные сети - сети, используемые в государственных структурах.

4. По скорости передачи

По скорости передачи информации компьютерные сети делятся на низко-, средне- и высокоскоростные.

- низкоскоростные (до 10 Мбит/с),

- среднескоростные (до 100 Мбит/с),

- высокоскоростные (свыше 100 Мбит/с);

5. По типу среды передачи

По типу среды передачи сети разделяются на:

- проводные –коаксиальные, на витой паре, оптоволоконные;

- беспроводные - с передачей информации по радиоканалам, в инфракрасном диапазоне.

6. Топологии компьютерных сетей

- общая шина;

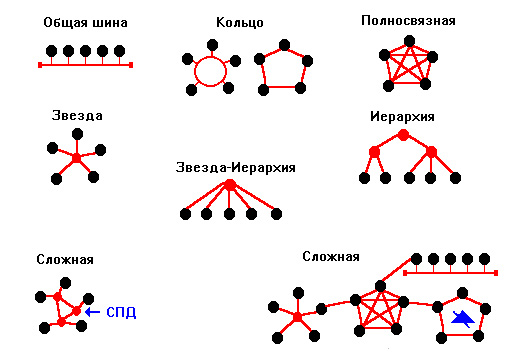
- кольцо;

- звезда;

- древовидная;

- ячеистая или сетчатая;

- смешанная (гибридная);



Компьютерная сеть состоит из информационных систем и каналов связи.

Под информационной системой следует понимать объект, способный осуществлять хранение, обработку или передачу информация. В состав информационной системы входят: компьютеры, программы, пользователи и другие составляющие, предназначенные для процесса обработки и передачи данных. В дальнейшем информационная система, предназначенная для решения задач пользователя, будет называться – рабочая станция (client). Рабочая станция в сети отличается от обычного персонального компьютера (ПК) наличием сетевой карты (сетевого адаптера), канала для передачи данных и сетевого программного обеспечения.

Под каналом связи следует понимать путь или средство, по которому передаются сигналы. Средство передачи сигналов называют абонентским, или физическим, каналом.

Каналы связи (data link) создаются по линиям связи при помощи сетевого оборудования и физических средств связи. Физические средства связи построены на основе витых пар, коаксиальных кабелей, оптических каналов или эфира. Между взаимодействующими информационными системами через физические каналы коммуникационной сети и узлы коммутации устанавливаются логические каналы.

Логический канал – это путь для передачи данных от одной системы к другой. Логический канал прокладывается по маршруту в одном или нескольких физических каналах. Логический канал можно охарактеризовать, как маршрут, проложенный через физические каналы и узлы коммутации.

Информация в сети передается блоками данных по процедурам обмена между объектами. Эти процедуры называют протоколами передачи данных.

Протокол – это совокупность правил, устанавливающих формат и процедуры обмена информацией между двумя или несколькими устройствами.

Загрузка сети характеризуется параметром, называемым трафиком. Трафик (traffic) – это поток сообщений в сети передачи данных. Под ним понимают количественное измерение в выбранных точках сети числа проходящих блоков данных и их длины, выраженное в битах в секунду.

Примеры. В сеть были объединены разработанные Институтом точной механики и вычислительной техники АН СССР компьютеры «Диана I» и «Диана II» (создатели С. А. Лебедев, В. С. Бурцев).

Корпоративная сеть.

10.2 Архитектура Internet

Случай, когда две машины включены в одну локальную сеть понятен. В случае удаленных машин, не имеющих между собой прямого соединения данные проходят через промежуточные машины, называемые мостами или маршрутизаторами. Эти машины обычно являются мощными серверами с сетевыми операционными системами (например, Unix) и постоянно включенным питанием. Множество компьютеров может быть представлено в виде графа, в котором длине ребра можно сопоставить время передачи единицы данных между двумя машинами. Тогда в общем случае задача передачи данных между двумя компьютерами сводится к задаче нахождения кратчайшего пути в графе. Процесс выбора следующей машины на пути следования данных называется процессом маршрутизации и будет рассмотрен ниже. Большим достоинством Интернет, является то, что пути могут выбираться разными способами и, например, в случае поломки одного из маршрутизаторов, данные могут идти в обход его.

TCP/IP

Для возможности работы приложений на разных компьютерах, использующих разные сетевые соединения используется протокол TCP/IP. Если Интернет не имеет определенной спецификации с точки зрения физических соединений его можно четко определить на логическом уровне. Как правило считается, что компьютер имеет полноценное подключение к Интернет, если на нем поддерживается протокол передачи данных TCP/IP. Драйвера этого протокола могут быть установлены на любой операционной системе. Использование унифицированных функций делает задачу программирования передачи данных в Интернет крайне простой. Все что нужно программисту, это вызвать функцию, передав ей массив с данными и адрес машины, куда их нужно передать.

Протокол TCP является надстройкой над протоколом IP и использует его в своей работе. Он обеспечивает гарантированную доставку данных, однако, из-за этого более медлителен. Кроме TCP может использоваться протокол UDP, также являющейся надстройкой над IP.

IP адрес, порт

Адрес машины в Интернет представляет собой 4 байтовых числа, разделенных точками, например 193.124.214.99 и называется IP адресом. Каждая машина, имеющая подключение к Интернет имеет собственный уникальный адрес, из чего следует, что таких машин может быть не более 4 млрд. В связи с этим в последнее время встает вопрос о нехватке адресов и переходу к новому протоколу с более длинным адресом. Хотя имеются программные средства использование целого сегмента машин, с выделением им всего одного IP адреса. В IP адресе определенное количество младших бит (обычно восемь) соответствует номеру компьютера, остальные задают адрес сети. Это разбиение задается т.н. маской сети, представляющей собой 4 байтовых числа, где вместо адреса сети выставляются 1-цы, а биты, задающие номер машины устанавливаются в 0. Для случая, когда адрес машины задается последним байтом, маска сети рана 255.255.255.0. Это означает, что две соседние машины в сети могут иметь, например, такие адреса: 193.124.214.99 и 193.124.214.100 (ws01 и ws02). Кроме того существуют специальные зарезервированные адреса, например, адрес 127.0.0.1 всегда соответствует адресу локальной машины.

Порт является понятием логическим и представляет собой двухбайтовое целое число (0-65535). Он является частью адреса и нужен для распределения пакетов между разными приложениями в пределах одного компьютера.

Доменная Система Имен (DNS)

Числовые IP адреса не слишком удобны для использования человеком и, поэтому, кроме них широко используются символьные или т.н. доменные адреса. Такой адрес представляет собой последовательность строк, разделенных точками, например cctelcom.nsu.ru.

Прежде чем данные пойдут к указанному доменному адресу, его необходимо преобразовать к числовому IP адресу. Эта задача решается посредством DNS серверов. Обычно каждая крупная подсеть имеет свой собственный DNS сервер, обрабатывающий запросы всех близлежащих машин. Все приложения автоматически разрешают эту задачу, что делает возможным работу с символьными адресами. Понятно, что на DNS сервере не может храниться информация по адресам всех машин в Интернет. Когда DNS сервер не может разрешить адрес, он отправляет запрос на другой DNS сервер отвечающий за адреса соответствующие самому верхнему домену (в данном случае, отвечающий за адреса всех компьютеров в России). Тот в свою очередь перенаправляет запрос рекурсивно на следующий сервер более низкого уровня.

10.3 Распределенные ВС. Определение, архитектурные принципы, классификация. Примеры реализации распределенных ВС.

Распределенные ВС: Все ресурсы представляются в виде композиции функционально-завершенных элементарных машин.

Распределенная ВС - объединение пространственно удаленных друг от друга сосредоточенных ВС, основанное на принципах:

• параллельности функционирования сосредоточенных ВС (т. e. способности нескольких или всех сосредоточенных систем совместно и одновременно решать одну сложную задачу, представленную параллельной программой);

• программируемости структуры (т. e. возможности автоматически настраивать сеть связи между сосредоточенными ВС);

• гомогенности состава (т. e. программной совместимости различных сосредоточенных ВС и однотипности элементарных машин в каждой из них).

На сегодняшний день можно выделить две основные классификации распределенных систем. Во-первых, это классификация по размерам систем и способам администрирования:

• Кластер - это несколько десятков компьютеров, объединенных с помощью локальной сети. Администрирование осуществляется вручную;

• Распределенная система корпоративного уровня – десятки и даже сотни компьютеров, при работе которых необходимо устанавливать правила совместного использования ресурсов. Масштаб таких систем, как правило, небольшой, и можно обходиться прямыми административными мерами для организации работы ресурсов и пользователей;

• Глобальная система (грид-система) – огромное количество компьютеров, число которых может достигать нескольких миллионов, распределенных по миру и объединенных глобальной сетью. Административное программное обеспечение встроено в промежуточное программное обеспечение.

Во-вторых, распределенные системы можно рассматривать с точки зрения их функциональности, т.е. по типу предоставляемых ресурсов и видам прикладных задач, под решение которых они оптимизированы. Таким образом, можно выделить [2]:

• вычислительные системы (Computational Grid) – это тип сис-ем, в которых основным компьютерным ресурсом является мощность процессора;

• информационные системы (Data Grid) – это тип систем, в которых основным компьютерным ресурсом является объем памяти данных. Эти системы рассматриваются как огромные хранилища данных.