

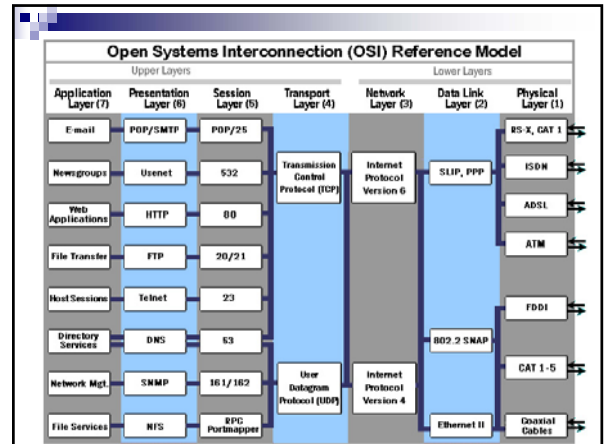
Лекция 17:

Сетевые технологии за полчаса

Распределенные системы Distributed systems

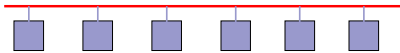
Алексей Линёв
Александр Мощук
Кирилл Погорельский

some slides are adapted from the OS course at the University of Washington



Data link layer (Уровень канала данных): Ethernet

- Сеть с общей средой передачи



- CSMA-CD: Carrier Sense Multiple Access with Collision Detection (Множественный доступ с прослушиванием несущей и обнаружением коллизий)
 - напоминает обмен анекдотами в теплой компании
- Технология с коммутацией пакетов
- Каждый сетевой адаптер имеет уникальный физический адрес (MAC-адрес)
 - 00-08-74-C9-C8-7E



- Формат пакета

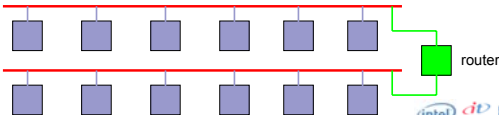


- Сетевой адаптер принимает сигнал в ожидании пакета, в котором будет указан его MAC-адрес в поле получателя
- Когда такой пакет получен, сетевой адаптер инициирует прерывание, и управление получает операционная система

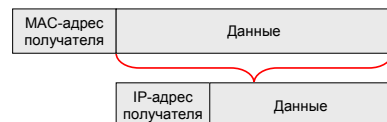


Network layer (Сетевой уровень): IP

- Internet Protocol (IP)
 - выполняет доставку пакетов от источника к получателю через несколько сетей (маршрутизация/routing)
- Каждое устройство имеет уникальный IP-адрес
 - 172.30.192.251
- Отдельные сети соединяются между собой через **маршрутизаторы (routers)**, которые подключены к нескольким сетям (имеют несколько сетевых адаптеров и, соответственно, MAC-адресов)



- Достаточно сложный протокол позволяет любому узлу сети определить физический адрес маршрутизатора в этой сети, который может передать пакет на один шаг в направлении получателя
- Формат пакета:

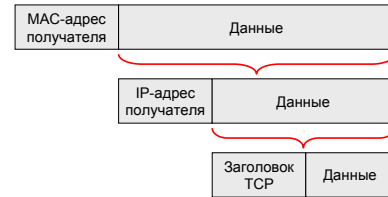


- Еще один сложный протокол, DNS (the Domain Name Service), отображает естественные имена (www.unn.ru) на IP-адреса (62.76.250.90)
- Таким образом, при доставке пакета получателю
 - используем DNS для преобразования доменного имени в IP-адрес
 - формируем IP-пакет из IP-адреса получателя и данных
 - определяем MAC-адрес нужного маршрутизатора
 - инкапсулируем IP-пакет в Ethernet-пакет с определенным на предыдущем шаге MAC-адресом получателя
 - отправляем!
- Добавим еще один тип адресов: **номера портов** представляют собой дополнительное адресное пространство в пределах одной системы



Transport layer (Транспортный уровень): TCP

- TCP: Transmission Control Protocol (протокол управления передачей)
 - организует надежную доставку потоков данных, базируясь на (ненадежном) механизме датаграммной передачи одиночных пакетов
 - аналог: пересылка книги на почтовых открытках
 - как бы это сделали?



Выводы

- Используя TCP/IP и нижележащие уровни, мы можем получить **надежную** доставку **потоков данных** из адресного пространства **A** машины **B** в адресное пространство **C** машины **D**, при этом машины **B** и **D** соединены через несколько гетерогенных сетей, и детали соединения не имеют значения
- Протоколы вышележащих уровней уже предоставляют конкретные сервисы
 - E-mail: smtp
 - web: http
 - file transfer: ftp
 - remote login: telnet



Что такое "распределенная система"?

- Очень широкое определение
 - от слабо-связанных до сильно-связанных систем
- В настоящее время почти все системы являются распределенными в том или ином смысле
 - они используют электронную почту
 - они получают доступ к файлам через сеть
 - они получают доступ к принтерам через сеть
 - их резервная копия снимается через сеть
 - они предоставляют свои аппаратные и другие ресурсы в совместное использование
 - они позволяют общаться людям, сидящим за разными машинами
 - они имеют доступ к www
 - они получают аудио- и видеопотоки данных
 - ...



Распределенные системы – требование времени

- Экономические соображения заставляют нас приобретать небольшие компьютеры
- Всем нужно общение
- Необходимо разделять аппаратные ресурсы (например, принтеры) и данные (например, файлы)
- Множество приложений являются распределенными по своей природе (банкоматы, системы бронирования авиабилетов,...)
- Для решения больших задач мы должны организовать совместную работу множества небольших компьютеров (параллельное программирование)



Слабо-связанные системы Loosely-coupled systems

- В ранних системах просто использовались конкретные сетевые программы
 - ftp (rcp): file transfer program
 - telnet (rlogin/rsh): remote login program
 - mail (SMTP)
- Каждая система была полностью автономной и независимой, соединенной с другими системами через сеть



- Даже в настоящее время большинство распределенных систем являются слабосвязанными
 - каждый ЦП выполняет автономную и независимую ОС
 - компьютеры не доверяют друг другу
 - часть ресурсов предоставлена в совместное использование, большинство – нет
 - система может выглядеть различно с точки зрения различных узлов
 - как правило, скорость передачи между узлами невысока



Сильно-связанные системы Closely-coupled systems

- Распределенная система становится все более и более сильно-связанной по мере того, как
 - выглядит более однородной по своей природе
 - выполняет "единую" операционную систему
 - имеет единый домен безопасности
 - разделяет все логические ресурсы (например, файлы)
 - разделяет все физические ресурсы (центральные процессоры, память, жесткие диски и т.д.)
- В идеале распределенная система представляется пользователю в виде единой системы разделения времени, хоть и состоит из распределенного множества аппаратного и программного обеспечения



Плотно-связанные системы Tightly-coupled systems

- Под "плотно-связанной" системой обычно понимают многопроцессорные системы
 - выполняет единственную копию ОС с единой очередью задач
 - имеет единое адресное пространство
 - как правило, имеет единую шину или системную плату, к которой подключены все процессоры и оперативная память
 - характеризуется очень низкой латентностью передачи
 - процессоры взаимодействуют через разделяемую память



Некоторые замечания на тему распределенных систем

- Прозрачность (насколько видна распределенная природа)
- Безопасность
- Надежность
- Производительность
- Масштабируемость
- Модели программирования
- Модели взаимодействия

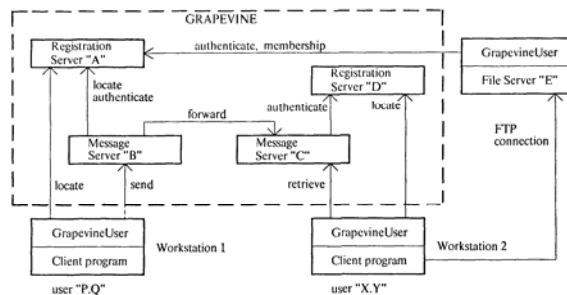


Grapevine distributed mail service (распределенная почтовая служба)

- Xerox PARC (исследовательский центр корпорации Xerox в Пало-Альто), 1980
 - сравнима с Microsoft Outlook/Exchange!
- Цели/задачи
 - нельзя полагаться на клиента (надежность, целостность,...)
 - если система приняла письмо, оно должно быть доставлено
 - сбой любого из компьютеров Grapevine не должен приводить к прекращению обслуживания клиентов (как приема, так и передачи сообщений)
- Компоненты
 - Модуль GrapevineUser на каждой рабочей станции
 - Сервера Регистрации
 - Сервера Сообщений
- Реализация: через Remote Procedure Call (удаленный вызов процедур)
 - RPC – следующая лекция



Grapevine: функциональная диаграмма



Grapevine: Отправление сообщения

- Пользователь подготавливает сообщение в почтовом клиенте
- Почтовый клиент связывается с модулем GrapevineUser, работающем на той же рабочей станции, для отправки сообщения
- Модуль GrapevineUser
 - соединяется с любым Сервером Регистрации и получает список Серверов Сообщений
 - соединяется с произвольным Сервером Сообщений для передачи сообщения
 - предоставляет идентификаторы пользователя-отправителя и пользователя-получателя, и пароль отправителя для аутентификации
 - Сервер Сообщений для аутентификации пользуется любым Сервером Регистрации
 - отправляет тело сообщения Серверу Сообщений
 - Сервер Сообщений размещает сообщение в надежном хранилище и подтверждает отправителю получение запроса



Grapevine: доставка и буферизация

- Для каждого получателя сообщения Сервер Сообщений соединяется с Сервером Регистрации и получает список Серверов Сообщений, на которых хранится почта этого пользователя
- Копия сообщения для этого получателя отправляется на один из серверов из списка



Grapevine: получение почты

- Почтовый клиент пользователя связывается с модулем GrapevineUser, работающем на той же рабочей станции, для получения сообщений
- Модуль GrapevineUser
 - соединяется с любым Сервером Регистрации и получает список всех Серверов Сообщений, хранящих почту данного пользователя
 - соединяется с каждым Сервером Сообщений из списка и получает с него почтовые сообщения
 - предоставляет идентификационные данные пользователя-получателя (например, идентификатор и пароль)
 - Сервер Сообщений для аутентификации пользуется любым Сервером Регистрации
 - подтверждает получение сообщения, таким образом, Сервер Сообщений может удалить их из своего хранилища



Grapevine: Масштабируемость

- Можно увеличивать число Серверов Регистрации
- Можно увеличивать число Серверов Сообщений
- Единственный процесс, не поддающийся масштабированию – обработка списков рассылки
 - принимающий Сервер Сообщений отвечает за разбор списка (при необходимости – рекурсивный) и доставку на соответствующие Сервера Сообщений для каждого получателя
 - некоторые списки рассылки могут содержать практически все множество пользователей
- По словам Джеффа Дина (Jeff Dean, Google), "они даже не думали о масштабах, больших чем сотни"
 - необходимы изменения в фундаментальных положениях архитектуры
 - современные технологические достижения позволяют использовать достаточно сложные решения

