

Лекция 1. ЭВОЛЮЦИЯ вычислительной сети: ОТ МАШИНЫ ЧАРЛЬЗА

Бэбидж К первой глобальной сетей

Два корня сетей передачи данных

История любой отрасли науки или техники позволяет не только удовлетворить естественное интерес, но и глубже понять сущность основных достижений в этой области, а также выявить тенденции и правильно оценить перспективность тех или иных направлений.

Сети передачи данных, называемые также вычислительными или компьютерными сетями, является результатом эволюции двух важнейших научно-технические отраслей современной цивилизации - компьютерных и телекоммуникационных технологий:

1. С одной стороны, сети передачи данных представляют собой частный случай распределенных вычислительных систем, в которых группа компьютеров согласованно выполняет набор взаимосвязанных задач, обмениваясь данными в автоматическом режиме.

2. С другой стороны, компьютерные сети могут рассматриваться как средство передачи информации на большие расстояния, для чего в них применяются методы кодирования и мультиплексирования данных, получили развитие в различных телекоммуникационных системах.

Итак, компьютерная сеть - это набор компьютеров, связанных коммуникационной системой и снабженных соответствующим программным обеспечением, которое предоставляет пользователям сети доступ к ресурсам этого набора компьютеров; сеть могут образовывать компьютеры различных типов - небольшие микропроцессоры, рабочие станции, мини-компьютеры, персональные компьютеры или суперкомпьютеры; передачу сообщений между любой парой компьютеров сети обеспечивает коммуникационная система, которая может включать кабели, повторители, коммутаторы, маршрутизаторы и другие устройства; компьютерная сеть позволяет пользователю работать с компьютером, как с автономным, и добавляет к этому возможность доступа к информационным и аппаратным ресурсам других компьютеров сети.

Появление первых вычислительных машин

Идея компьютера была предложена английским математиком Чарльзом Бэбиджем (Charles Babbage) в середине девятнадцатого века. Однако его механическая "аналитическая машина" по-настоящему так и не заработала.

Настоящее рождение цифровых вычислительных машин произошло вскоре после окончания второй мировой войны. В середине 40-х были созданы первые ламповые вычислительные устройства. Для этого периода характерно следующее:

- компьютер представлял собой скорее предмет исследования, а не инструмент для решения каких либо практических задач по других областей;
- та же группа людей участвовала и в проектировании, и в эксплуатации, и в программировании вычислительной машины;
- программирования осуществлялось исключительно на машинном языке;

- не было никакого системного программного обеспечения, кроме библиотек математических и служебных подпрограмм;

- операционные системы еще не появились, все задачи организации вычислительного процесса решались вручну каждым программистом с пульта управления.

С середины 50-х годов начался следующий период в развитии вычислительной техники, связан с появлением новой технической базы - полупроводниковых элементов. В этот период:

- выросла быстродействие процессоров, увеличились объемы оперативной и внешней памяти;
- компьютеры стали более надежными;
- появились первые алгоритмические языки, и, таким образом, в библиотеки математических и служебных подпрограмм добавился новый тип системного программного обеспечения - трансляторы;
- были разработаны первые системные управляющие программы - мониторы, которые автоматизировали всю последовательность действий оператора по организации вычислительного процесса.

Программные мониторы - первые операционные системы

Программные мониторы явились прообразом современных операционных систем, они стали первыми системными программами, предназначенными не для обработки данных, а для управления вычислительным процессом.

В ходе реализации мониторов была разработана формализованный язык управления заданиями, по помощи которого программист сообщал системе и оператору, какие действия и в какой последовательности он хотел бы выполнить на вычислительной машине. Типичный набор директив конечно включала признак начала отдельной работы, вызов транслятора, вызов загрузчика, признаки начала и конца выходных данных.

Оператор составлял пакет заданий, которые в дальнейшем без его участия последовательно запускались на выполнения монитором. Кроме того, монитор был способен самостоятельно обрабатывать самые распространенные аварийные ситуации, возникающие при работе пользовательских программ, такие как отсутствие выходных данных, переполнение регистров, деление на ноль, обращение к несуществующей области памяти и т.д.

Мультипрограммирование

Следующий важный период развития операционных систем относится к 1965-1975 годам. В это время в технической базе вычислительных машин произошел переход от отдельных полупроводниковых элементов типа транзисторов к интегральным микросхемам, что открыло путь к появлению следующего поколения компьютеров, представителем которого является, например, IBM / 360.

В этот период были реализованы практически все основные механизмы, присущие современным ОС: мультипрограммирование, мультипроцессирование, поддержка многопользовательского режима, виртуальная память, файловые системы, разграничение доступа и сетевая работа. В эти годы начинается расцвет системного программирования. С направления прикладной математики, представляет интерес для узкого круга специалистов, системное программирование превращается в отрасль индустрии, оказывает непосредственное влияние на практическую деятельность миллионов человек.

В условиях резко возросших возможностей компьютера, связанных с обработкой и хранением данных, выполнение только одной программы в каждый момент времени оказалось крайне неэффективным.

Начались разработки в области мультипрограммирования.

Мультипрограммирование - способ организации вычислительного процесса, при котором в памяти компьютера находится одновременно несколько программ, попеременно выполняющихся на одном процессоре.

Мультипрограммирование было реализовано в двух вариантах:

- пакетная обработка;
- разделение времени.

Системы пакетной обработки предназначались для решения задач в основном вычислительного характера, не требующих быстрого получения результатов. Главной целью и критерием эффективности систем пакетной обработки является максимальная пропускная способность, то есть решение максимального числа задач в единицу времени.

Для достижения этой цели в системах пакетной обработки используется следующая схема функционирования: в начале работы формируется пакет заданий, каждое **Задание** содержит требование к системным ресурсам; с этого пакета заданий формируется мультипрограммный набор, то есть множество одновременно выполняемых задач. Для одновременного выполнения выбираются задачи, предъявляют к ресурсам разные требования, так, чтобы обеспечивалось сбалансированное загрузки всех строений вычислительной машины. Например, в мультипрограммном наборе желательно присутствие и вычислительных задач, и задач по интенсивным вводом-выводом. Таким образом, выбор новой задачи из пакета заданий зависит от внутренней ситуации, складывающейся в системе, то есть выбирается "выгодное" задание. Итак, в вычислительных системах, работающих под управлением **пакетных ОС**, невозможно гарантировать выполнение того или иного задания в течение определенного периода времени.

В системах пакетной обработки переключение процессора с одной задачи на другую происходит по инициативе самой активной задачи, например, когда она "отказывается" от процессора через необходимость выполнить операцию ввода-вывода. Поэтому существует высокая вероятность того, что одна задача может надолго занять процессор, и выполнение интерактивных задач станет невозможным. Взаимодействие пользователя с вычислительной машиной, на которой установлена система пакетной обработки, сводится к тому, что пользователь приносит задание, отдает его диспетчеру-оператору, а в конце дня после выполнения всего пакета заданий получает результат. Очевидно, что такой порядок повышает эффективность функционирования аппаратуры, но снижает эффективность работы пользователя.

В системах разделения времени пользователям (или одному пользователю) предоставляется возможность интерактивной работы сразу с несколькими приложениями. Для этого каждое приложение должно регулярно взаимодействовать с пользователем. Понятно, что в пакетных системах возможности диалога пользователя с приложением ограничены.

В системах разделения времени эта проблема решается за счет того, что ОС принудительно периодически прекращает приложения, не дожидаясь, когда они сами освободят процессор. всем приложениям попеременно выделяется квант процессорного времени, таким образом, пользователи, запустили программы на выполнение, получают возможность поддерживать с ними диалог.

Системы разделения времени призваны исправить основной недостаток систем пакетной обработки - изоляцию пользователя-программиста от процесса выполнения заданий. Каждому пользователю в этом случае предоставляется терминал, с которого он может вести диалог со своей программой. Потому что в системах разделения времени каждой задаче выделяется только квант процессорного времени, ни одна задача не занимает процессор надолго, и время ответа оказывается приемлемым. если квант небольшой, то у всех пользователей, одновременно работающих на одной и той же машине, складывается впечатление, что каждый из них использует машину единолично.

Ясно, что системы разделения времени обладают меньшей пропускной способности, что системы пакетной обработки, так как на выполнение принимается каждая запущенная пользователем задача, а не та, которая "Выгодна" системе. Кроме того, производительность системы снижается из-за дополнительный расход вычислительной мощности на более частое переключение процессора с задачи на задачу. это вполне соответствует тому, что критерием эффективности систем разделения времени является не максимальный пропускная способность, а удобство и эффективность работы пользователя. Вместе с тем, мультипрограммное выполнения интерактивных приложений повышает и пропускную способность компьютера (Пусть и не в такой степени, как пакетные системы). Аппаратура загружается лучше, поскольку пока одно приложение ждет сообщения, другие приложения могут обрабатываться процессором.

Многотерминальные системы - прообраз сети

Терминалы, выйдя за пределы вычислительного центра, рассредоточились по всему предприятию. Многотерминальные режим использовался не только в системах разделения времени, но и в системах пакетной обработки. При этом не только оператор, но и все пользователи получали возможность формировать свои задачи и управлять их выполнением со своего терминала. такие операционные системы получили название систем удаленного ввода задач.

Терминальные комплексы могли располагаться на большом расстоянии от процессорных стоек, соединяясь с ними с помощью различных глобальных связей - модемных соединений телефонных сетей или выделенных каналов. Для поддержки удаленной работы терминалов в операционных системах появились специальные программные модули, реализующие различные (в то время, как правило, нестандартные) протоколы связи. Такие вычислительные системы с удаленными терминалами сохраняя централизованный характер обработки данных, в какой-то степени были прообразом современных компьютерных сетей, а соответствующее системное программное обеспечение - прообразом сетевых операционных систем.

Многотерминальные централизованные системы уже имели все внешние признаки локальных вычислительных сетей, однако по сути ими не были, так как сохраняли сущность централизованной обработки данных автономно работающего компьютера.

Действительно, рядовой пользователь работу за терминалом мэйнфрейма воспринимал примерно так же, как сейчас воспринимает работу за подключенным к сети персональным компьютером. Пользователь мог получить доступ к общим файлам и периферийным устройствам, при этом в него создавалась полная иллюзия единоличного владения компьютером, потому что он мог запустить нужную ему программу в любой момент и почти сразу же получить результат. (Некоторые далеки от вычислительной техники пользователи даже были уверены, что все вычисления выполняются внутри их дисплея.)

Первые сети - глобальные

Хотя теоретические работы по созданию концепций сетевого взаимодействия велись почти с момента появления вычислительных машин, значимые практические результаты по объединению компьютеров в сети были получены лишь в конце 60-х, когда с помощью глобальных связей и техники коммутации пакетов удалось реализовать взаимодействие машин класса мэйнфреймов и суперкомпьютеров. Эти дорогие компьютеры хранили уникальные данные и программы, обмен которыми позволил повысить эффективность их использование.

Но еще до реализации связей "компьютер-компьютер", было решено более простая задача - организация связи "изъят терминал-компьютер". Терминалы, находящиеся от компьютера на расстоянии многих сотен, а то и тысяч километров, соединялись с компьютерами через телефонные сети с помощью модемов. Такие сети позволяли многочисленным пользователям получать удаленный доступ к разделяемым ресурсам нескольких мощных компьютеров класса супер-ЭВМ.

И только потом были разработаны средства обмена данными между компьютерами в автоматическом режиме. На основе этого механизма в первых сетях были реализованы службы обмена файлами, синхронизации баз данных, электронной почты и другие, ставшие теперь традиционными сетевые службы.

В 1969 году министерство обороны США инициировало работы по объединению в общую сеть суперкомпьютеров оборонных и научно-исследовательских центров. Эта сеть, получившая название ARPANET послужила отправной точкой для создания первой и самой известной ныне глобальной сети - Internet. Сеть ARPANET объединяла компьютеры разных типов, работавших под управлением различных ОС с дополнительными модулями, реализующими коммуникационные протоколы, общие для всех компьютеров сети. Такие ОС можно считать первыми сетевыми операционными системами.

Сетевые ОС в отличие от многотерминальные позволяли не только рассредоточить пользователей, но и организовать распределенное хранение и обработку данных между несколькими компьютерами, связанными электрическими связями. Любая сетевая операционная система, с одной стороны, выполняет все функции локальной операционной системы, а с другой стороны, имеет некоторые дополнительные средства, позволяющие ей взаимодействовать по сети с операционными системами других компьютеров. Программные модули, реализующие сетевые функции, появлялись в операционных системах постепенно, по мере развития сетевых технологий, аппаратной базы компьютеров и возникновения новых задач, требующих сетевой обработки.

В 1974 году компания IBM объявила о создании собственной сетевой архитектуры для своих мейнфреймов, получившей название SNA (System Network Architecture, системная сетевая архитектура). В это же время в Европе активно велись работы по созданию и стандартизации сетей X.25.

Таким образом, хронологически первыми появились глобальные сети (Wide Area Networks, WAN), то есть сети, объединяющие территориально рассредоточенные компьютеры, возможно, что находятся в разных городах и странах. Именно при построении глобальных сетей были впервые предложены и отработаны многие основные идеи и концепции современных вычислительных сетей, такие, например, как многоуровневое построение коммуникационных протоколов, технология коммутации пакетов и маршрутизация пакетов в составных сетях.

Наследие телефонных сетей

Глобальные компьютерные сети очень многое унаследовали от других, гораздо более старых и глобальных сетей - телефонных.

Главным результатом создания первых глобальных компьютерных сетей был отказ от принципа коммутации каналов, в течение многих десятков лет, успешно использовался в телефонных сетях.

Выделяемый на все время сеанса связи составной канал с постоянной скоростью не мог эффективно использоваться пульсирующим трафиком компьютерных данных, у которого периоды интенсивного обмена чередуются с продолжительными паузами. Эксперименты и математическое моделирование показали, что пульсирующий и в значительной степени не чувствителен к задержкам компьютерный трафик гораздо эффективнее передается по сетям, использующим принцип коммутации пакетов, когда данные разделяются на небольшие порции, которые самостоятельно перемещаются по сети с учетом встраивания адреса конечного узла в заголовок пакета.

Потому что прокладка высококачественных линий связи на большие расстояния обходится очень дорого, в первых глобальных сетях часто использовались уже существующие каналы связи, изначально предназначенные совсем для других целей. Например, в течение многих лет глобальные сети строились на основе телефонных каналов тональной частоты, способных в каждый момент времени вести передачу только одного разговора в аналоговой форме. Поскольку скорость передачи дискретных компьютерных данных по таким каналам была очень низкой (десятки килобит в секунду), набор предоставляемых услуг в глобальных сетях такого типа обычно ограничивался передачей файлов, преимущественно в фоновом режиме, и по электронной почте.

Кроме низкой скорости такие каналы имеют и другой недостаток - они вносят значительные искажения в передаваемые сигналы. Поэтому протоколы глобальных сетей, построенных с использованием каналов связи низкого качества, отличаются сложными процедурами контроля и восстановления данных. Типичным примером таких сетей являются сети X.25, разработанные еще в начале 70-х, когда низкоскоростные аналоговые каналы, арендуемые в телефонных компаниях, были преобладающим типом каналов, соединяющих компьютеры и коммутаторы глобальной вычислительной сети.

Развитие технологии глобальных компьютерных сетей во многом определялось прогрессом телефонных сетей. С конца 60-х годов в телефонных сетях все чаще стала применяться

передача голоса в цифровой форме, что привело к появлению высокоскоростных цифровых каналов, соединяют АТС, и позволяют одновременно передавать десятки и сотни разговоров. была разработана специальная технология плезиохронной цифровой иерархии (Plesiochronous Digital Hierarchy, PDH), предназначена для создания так называемых первичных, или опорных, сетей. Такие сети не предоставляют услуг конечным пользователям, они являются фундаментом, на котором строятся скоростные цифровые каналы "точка-точка", соединяющих оборудование другой (так называемой наложенной) сети, которая уже работает на конечного пользователя.

Изначально технология PDH, что поддерживает скорости до 140 Мбит / с, была внутренней технологии телефонных компаний. Однако впоследствии эти компании стали сдавать часть своих каналов PDH в аренду предприятиям, которые использовали их для создания собственных телефонных и глобальных компьютерных сетей.

Появившийся в конце 80-х годов технология синхронной цифровой иерархии (Synchronous Digital Hierarchy, SDH) расширила диапазон скоростей цифровых каналов до 10 Гбит / с, а технология спектрального мультиплексирования DWDM (Dense Wave Division Multiplexing) - до сотен гигабит и даже нескольких терабит в секунду.