Физическая организация и адрес файла

Физическая организация файла описывает правила расположения файла на устройстве внешней памяти, в частности на диске. Файл состоит из физических записей - блоков. Блок - наименьшая единица данных, которой внешнее устройство обменивается с оперативной памятью. Непрерывное размещение - простейший вариант физической организации (**Рис. 1**а), при котором файлу предоставляется последовательность блоков диска, образующих единый сплошной участок дисковой памяти. Для задания адреса файла в этом случае достаточно указать только номер начального блока. Другое достоинство этого метода - простота. Но имеются и два существенных недостатка. Во-первых, во время создания файла заранее не известна его длина, а значит не известно, сколько памяти надо зарезервировать для этого файла, во-вторых, при таком порядке размещения неизбежно возникает фрагментация, и пространство на диске используется не эффективно, так как отдельные участки маленького размера (минимально 1 блок) могут остаться не используемыми.

Следующий способ физической организации - размещение в виде связанного списка блоков дисковой памяти (**Рис. 1**б). При таком способе в начале каждого блока содержится указатель на следующий блок. В этом случае адрес файла также может быть задан одним числом - номером первого блока. В отличие от предыдущего способа, каждый блок может быть присоединен в цепочку какого-либо файла, следовательно фрагментация отсутствует. Файл может изменяться во время своего существования, наращивая число блоков. Недостатком является сложность реализации доступа к произвольно заданному месту файла: для того, чтобы прочитать пятый по порядку блок файла, необходимо последовательно прочитать четыре первых блока, прослеживая цепочку номеров блоков. Кроме того, при этом способе количество данных файла, содержащихся в одном блоке, не равно степени двойки (одно слово израсходовано на номер следующего блока), а многие программы читают данные блоками, размер которых равен степени двойки.

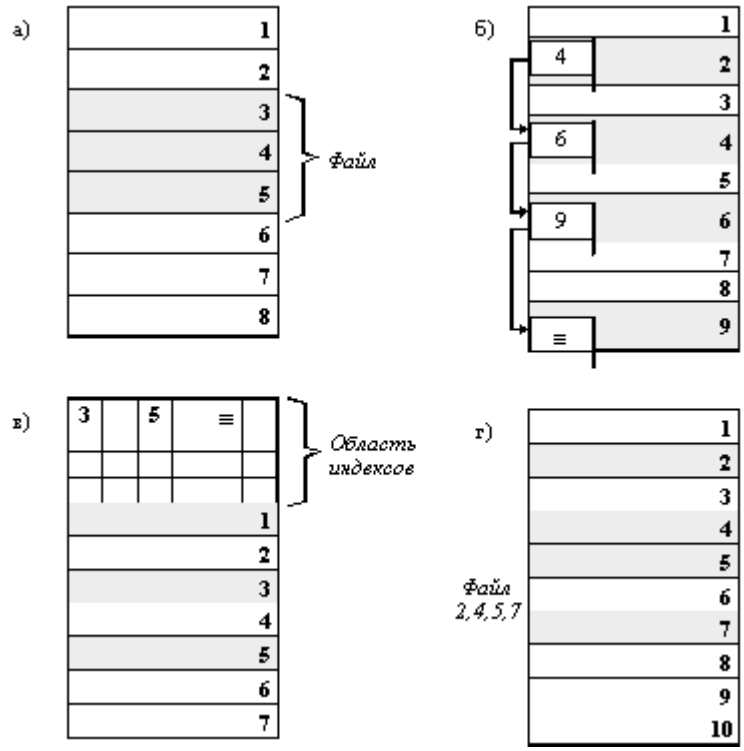


Рис. 1 Физическая организация файла
а - непрерывное размещение; б - связанный список блоков;
в - связанный список индексов; г - перечень номеров блоков

Популярным способом, используемым, например, в файловой системе FAT операционной системы MS-DOS, является использование связанного списка индексов. С каждым блоком связывается некоторый элемент - индекс. Индексы располагаются в отдельной области диска (в MS-DOS это таблица FAT). Если некоторый блок распределен некоторому файлу, то индекс этого блока содержит номер следующего блока данного файла. При такой физической организации сохраняются все достоинства предыдущего способа, но снимаются оба отмеченных недостатка: во-первых, для доступа к произвольному месту файла достаточно прочитать только блок индексов, отсчитать нужное количество блоков файла по цепочке и определить номер нужного блока, и, во-вторых, данные файла занимают блок целиком, а значит имеют объем, равный степени двойки.

В заключение рассмотрим задание физического расположения файла путем простого перечисления номеров блоков, занимаемых этим файлом. ОС UNIX использует вариант данного способа, позволяющий обеспечить фиксированную длину адреса, независимо от размера файла. Для хранения адреса файла выделено 13 полей. Если размер файла меньше или равен 10 блокам, то номера этих блоков непосредственно перечислены в первых десяти полях адреса. Если размер файла больше 10 блоков, то следующее 11-е поле содержит адрес блока, в котором могут быть расположены еще 128 номеров следующих блоков файла. Если файл больше, чем 10+128 блоков, то используется 12-е поле, в котором находится номер блока, содержащего 128 номеров блоков, которые содержат по 128 номеров блоков данного файла. И, наконец, если файл больше 10+128+128^128, то используется последнее 13-е поле для тройной косвенной адресации, что позволяет задать адрес файла, имеющего размер максимум 10+ 128 + 128^128 + 128^128^128.

Права доступа к файлу

Определить права доступа к файлу - значит определить для каждого пользователя набор операций, которые он может применить к данному файлу. В разных файловых системах может быть определен свой список дифференцируемых операций доступа. Этот список может включать следующие операции:

- создание файла,
- уничтожение файла,
- открытие файла,
- закрытие файла,
- чтение файла,
- запись в файл,
- дополнение файла,
- поиск в файле,
- получение атрибутов файла,
- установление новых значений атрибутов,
- переименование,
- выполнение файла,
- чтение каталога,

и другие операции с файлами и каталогами.

В самом общем случае права доступа могут быть описаны матрицей прав доступа, в которой столбцы соответствуют всем файлам системы, строки - всем пользователям, а на пересечении строк и столбцов указываются разрешенные операции (Рис. 2). В некоторых системах пользователи могут быть разделены на отдельные категории. Для всех пользователей одной категории определяются единые права доступа. Например, в системе UNIX все пользователи подразделяются на три категории: владельца файла, членов его группы и всех остальных.

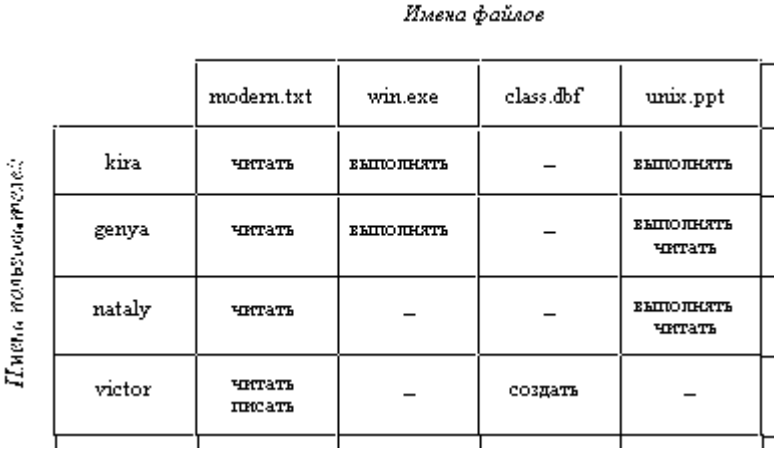


Рис. 2 Матрица прав доступа

Различают два основных подхода к определению прав доступа:

- избирательный доступ, когда для каждого файла и каждого пользователя сам владелец может определить допустимые операции;
- мандатный подход, когда система наделяет пользователя определенными правами по отношению к каждому разделяемому ресурсу (в данном случае файлу) в зависимости от того, к какой группе пользователь отнесен.

Кэширование диска

В некоторых файловых системах запросы к внешним устройствам, в которых адресация осуществляется блоками (диски, ленты), перехватываются промежуточным программным слоем-подсистемой буферизации. Подсистема буферизации представляет собой буферный пул, располагающийся в оперативной памяти, и комплекс программ, управляющих этим пулом. Каждый буфер пула имеет размер, равный одному блоку. При поступлении запроса на чтение некоторого блока подсистема буферизации просматривает свой буферный пул и, если находит требуемый блок, то копирует его в буфер запрашивающего процесса. Операция ввода-вывода считается выполненной, хотя физического обмена с устройством не происходило. Очевиден выигрыш во времени доступа к файлу. Если же нужный блок в буферном пуле отсутствует, то он считывается с устройства и одновременно с передачей запрашивающему процессу копируется в один из буферов подсистемы буферизации. При отсутствии свободного буфера на диск вытесняется наименее

используемая информация. Таким образом, подсистема буферизации работает по принципу кэш-памяти.

Общая модель файловой системы

Функционирование любой файловой системы можно представить многоуровневой моделью (Рис. 3), в которой каждый уровень предоставляет некоторый интерфейс (набор функций) вышележащему уровню, а сам, в свою очередь, для выполнения своей работы использует интерфейс (обращается с набором запросов) нижележащего уровня.

Задачей символьного уровня является определение по символьному имени файла его уникального имени. В файловых системах, в которых каждый файл может иметь только одно символьное имя (например, MS-DOS), этот уровень отсутствует, так как символьное имя, присвоенное файлу пользователем, является одновременно уникальным и может быть использовано операционной системой. В других файловых системах, в которых один и тот же файл может иметь несколько символьных имен, на данном уровне просматривается цепочка каталогов для определения уникального имени файла. В файловой системе UNIX, например, уникальным именем является номер индексного дескриптора файла (i-node).

На следующем, базовом уровне по уникальному имени файла определяются его характеристики: права доступа, адрес, размер и другие. Как уже было сказано, характеристики файла могут входить в состав каталога или храниться в отдельных таблицах. При открытии файла его характеристики перемещаются с диска в оперативную память, чтобы уменьшить среднее время доступа к файлу. В некоторых файловых системах (например, HPFS) при открытии файла вместе с его характеристикам в оперативную память перемещаются несколько первых блоков файла, содержащих данные.

Рис. 3 Общая модель файловой системы

Следующим этапом реализации запроса к файлу является проверка прав доступа к нему. Для этого сравниваются полномочия пользователя или процесса, выдавших запрос, со списком разрешенных видов доступа к данному файлу. Если запрашиваемый вид доступа разрешен, то выполнение запроса продолжается, если нет, то выдается сообщение о нарушении прав доступа.

На логическом уровне определяются координаты запрашиваемой логической записи в файле, то есть требуется определить, на каком расстоянии (в байтах) от начала файла находится требуемая логическая запись. При этом абстрагируются от физического расположения файла, он представляется в виде непрерывной последовательности байт. Алгоритм работы данного уровня зависит от логической организации файла. Например, если файл организован как последовательность логических записей фиксированной длины l, то n-ая логическая запись имеет смещение l((n-1) байт. Для определения координат логической записи в файле с индексно-последовательной организацией выполняется чтение таблицы индексов (ключей), в которой непосредственно указывается адрес логической записи.

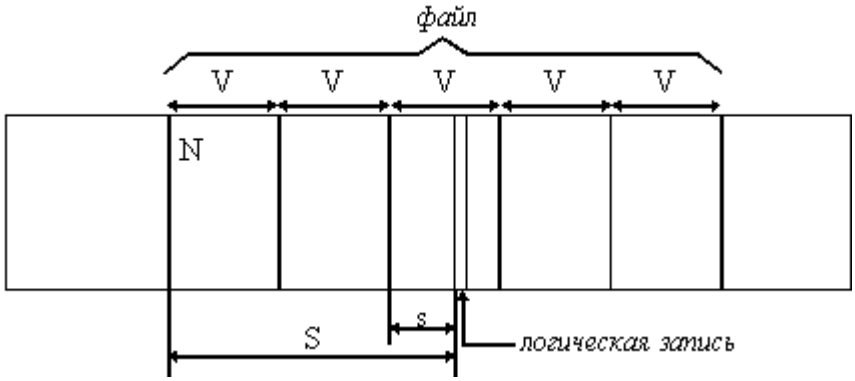


Рис. 4 Функции физического уровня файловой системы

Исходные данные:

V - размер блока

N - номер первого блока файла

S - смещение логической записи в файле

Требуется определить на физическом уровне:

n - номер блока, содержащего требуемую логическую запись

s - смещение логической записи в пределах блока

n = N + [S/V], где [S/V] - целая часть числа S/V

s = R [S/V] - дробная часть числа S/V

На физическом уровне файловая система определяет номер физического блока, который содержит требуемую логическую запись, и смещение логической записи в физическом блоке. Для решения этой задачи используются результаты работы логического

уровня - смещение логической записи в файле, адрес файла на внешнем устройстве, а также сведения о физической организации файла, включая размер блока. Рис. 4 иллюстрирует работу физического уровня для простейшей физической организации файла в виде непрерывной последовательности блоков. Подчеркнем, что задача физического уровня решается независимо от того, как был логически организован файл.

После определения номера физического блока, файловая система обращается к системе ввода-вывода для выполнения операции обмена с внешним устройством. В ответ на этот запрос в буфер файловой системы будет передан нужный блок, в котором на основании полученного при работе физического уровня смещения выбирается требуемая логическая запись.

Отображаемые в память файлы

По сравнению с доступом к памяти, традиционный доступ к файлам выглядит запутанным и неудобным. По этой причине некоторые ОС, начиная с MULTICS, обеспечивают отображение файлов в адресное пространство выполняемого процесса. Это выражается в появлении двух новых системных вызовов: MAP (отобразить) и UNMAP (отменить отображение). Первый вызов передает операционной системе в качестве параметров имя файла и виртуальный адрес, и операционная система отображает указанный файл в виртуальное адресное пространство по указанному адресу.

Предположим, например, что файл *f*имеет длину 64 К и отображается на область виртуального адресного пространства с начальным адресом 512 К. После этого любая машинная команда, которая читает содержимое байта по адресу 512 К, получает 0-ой байт этого файла и т.д. Очевидно, что запись по адресу 512 К + 1100 изменяет 1100 байт файла. При завершении процесса на диске остается модифицированная версия файла, как если бы он был изменен комбинацией вызовов SEEK и WRITE.

В действительности при отображении файла внутренние системные таблицы изменяются так, чтобы данный файл служил хранилищем страниц виртуальной памяти на диске. Таким образом, чтение по адресу 512 К вызывает страничный отказ, в результате чего страница 0 переносится в физическую память. Аналогично, запись по адресу 512 К + 1100 вызывает страничный отказ, в результате которого страница, содержащая этот адрес, перемещается в память, после чего осуществляется запись в память по требуемому адресу. Если эта страница вытесняется из памяти алгоритмом замены страниц, то она записывается обратно в файл в соответствующее его место. При завершении процесса все отображенные и модифицированные страницы переписываются из памяти в файл.

Отображение файлов лучше всего работает в системе, которая поддерживает сегментацию. В такой системе каждый файл может быть отображен в свой собственный сегмент, так что k-ый байт в файле является k-ым байтом сегмента. На Рис. 5а изображен процесс, который имеет два сегмента-кода и данных. Предположим, что этот процесс копирует файлы. Для этого он сначала отображает файл-источник, например, abc. Затем он создает пустой сегмент и отображает на него файл назначения, например, файл ddd.

С этого момента процесс может копировать сегмент-источник в сегмент-приемник с помощью обычного программного цикла, использующего команды пересылки в памяти типа *mov*. Никакие вызовы READ или WRITE не нужны. После выполнения копирования процесс может выполнить вызов UNMAP для удаления файла из адресного пространства, а затем завершиться. Выходной файл ddd будет существовать на диске, как если бы он был создан обычным способом.

Хотя отображение файлов исключает потребность в выполнении ввода-вывода и тем самым облегчает программирование, этот способ порождает и некоторые новые проблемы. Во-первых, для системы сложно узнать точную длину выходного файла, в данном примере ddd. Проще указать наибольший номер записанной страницы, но нет способа узнать, сколько байт в этой странице было записано. Предположим, что программа использует только страницу номер 0, и после выполнения все байты все еще установлены в значение 0 (их начальное значение). Быть может, файл состоит из 10 нулей. А может быть, он состоит из 100 нулей. Как это определить? Операционная система не может это сообщить. Все, что она может сделать, так это создать файл, длина которого равна размеру страницы.

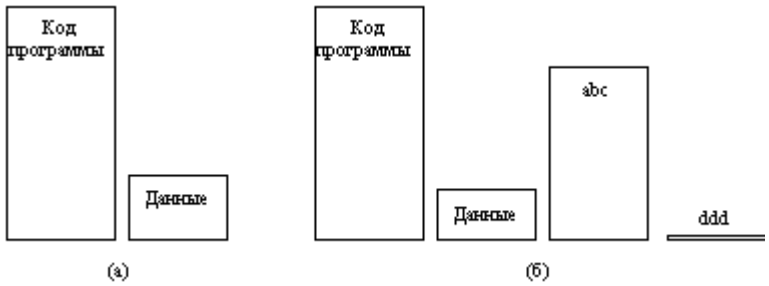


Рис. 5 (а) Сегменты процесса перед отображением файлов в адресное пространство; (б) Процесс после отображения существующего файла abc в один сегмент и создания нового сегмента для файла ddd

Вторая проблема проявляется (потенциально), если один процесс отображает файл, а другой процесс открывает его для обычного файлового доступа. Если первый процесс изменяет страницу, то это изменение не будет отражено в файле на диске до тех пор, пока страница не будет вытеснена на диск. Поддержание согласованности данных файла для этих двух процессов требует от системы больших забот.

Третья проблема состоит в том, что файл может быть больше, чем сегмент, и даже больше, чем все виртуальное адресное пространство. Единственный способ ее решения состоит в реализации вызова MAP таким образом, чтобы он мог отображать не весь файл, а его часть. Хотя такая работа, очевидно, менее удобна, чем отображение целого файла.

Современные архитектуры файловых систем

Разработчики новых операционных систем стремятся обеспечить пользователя возможностью работать сразу с несколькими файловыми системами. В новом понимании файловая система состоит из многих составляющих, в число которых входят и файловые системы в традиционном понимании.

Новая файловая система имеет многоуровневую структуру (Рис. 6), на верхнем уровне которой располагается так называемый переключатель файловых систем (в Windows 95, например, такой переключатель называется устанавливаемым диспетчером файловой системы - installable filesystem manager, IFS). Он обеспечивает интерфейс между запросами приложения и конкретной файловой системой, к которой обращается это приложение. Переключатель файловых систем преобразует запросы в формат, воспринимаемый следующим уровнем - уровнем файловых систем.

Каждый компонент уровня файловых систем выполнен в виде драйвера соответствующей файловой системы и поддерживает определенную организацию файловой

системы. Переключатель является единственным модулем, который может обращаться к драйверу файловой системы. Приложение не может обращаться к нему напрямую. Драйвер файловой системы может быть написан в виде реентерабельного кода, что позволяет сразу нескольким приложениям выполнять операции с файлами. Каждый драйвер файловой системы в процессе собственной инициализации регистрируется у переключателя, передавая ему таблицу точек входа, которые будут использоваться при последующих обращениях к файловой системе.

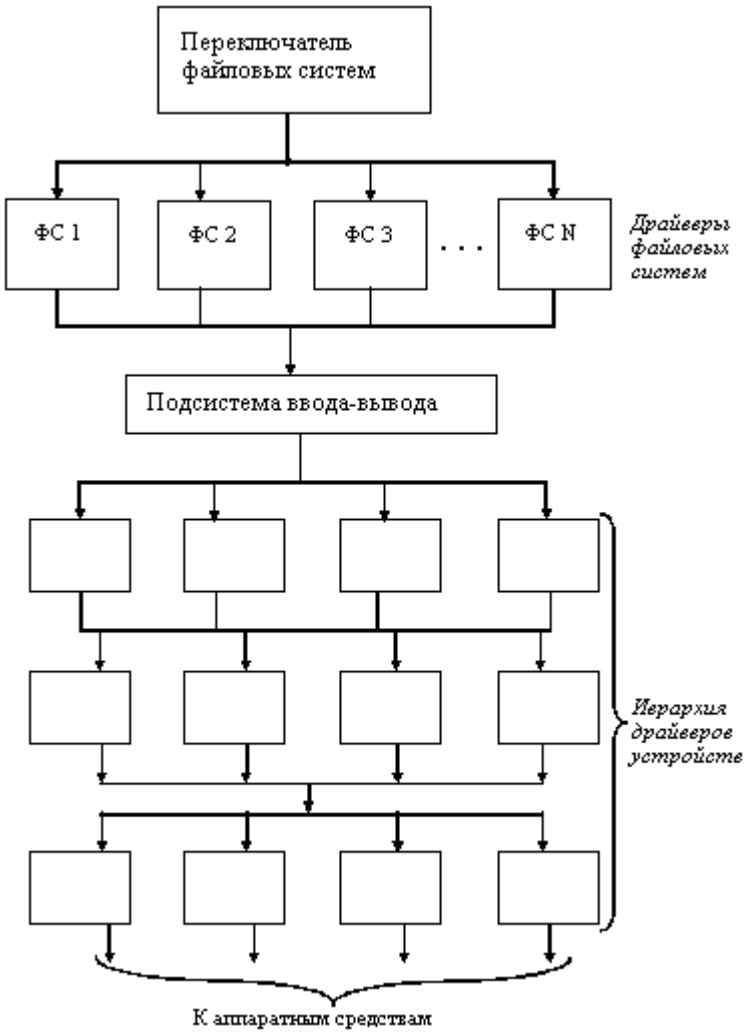


Рис. 6 Архитектура современной файловой системы

Для выполнения своих функций драйверы файловых систем обращаются к подсистеме ввода-вывода, образующей следующий слой файловой системы новой архитектуры. Подсистема ввода вывода - это составная часть файловой системы, которая отвечает за загрузку, инициализацию и управление всеми модулями низших уровней файловой системы. Обычно эти модули представляют собой драйверы портов, которые непосредственно занимаются работой с аппаратными средствами. Кроме этого подсистема ввода-вывода обеспечивает некоторый сервис драйверам файловой системы, что позволяет им осуществлять запросы к конкретным устройствам. Подсистема ввода-вывода должна постоянно присутствовать в памяти и организовывать совместную работу иерархии драйверов устройств. В эту иерархию могут входить драйверы устройств определенного

типа (драйверы жестких дисков или накопителей на лентах), драйверы, поддерживаемые поставщиками (такие драйверы перехватывают запросы к блочным устройствам и могут частично изменить поведение существующего драйвера этого устройства, например, зашифровать данные), драйверы портов, которые управляют конкретными адаптерами.

Большое число уровней архитектуры файловой системы обеспечивает авторам драйверов устройств большую гибкость - драйвер может получить управление на любом этапе выполнения запроса - от вызова приложением функции, которая занимается работой с файлами, до того момента, когда работающий на самом низком уровне драйвер устройства начинает просматривать регистры контроллера. Многоуровневый механизм работы файловой системы реализован посредством цепочек вызова.

В ходе инициализации драйвер устройства может добавить себя к цепочке вызова некоторого устройства, определив при этом уровень последующего обращения. Подсистема ввода-вывода помещает адрес целевой функции в цепочку вызова устройства, используя заданный уровень для того, чтобы должным образом упорядочить цепочку. По мере выполнения запроса, подсистема ввода-вывода последовательно вызывает все функции, ранее помещенные в цепочку вызова.

Внесенная в цепочку вызова процедура драйвера может решить передать запрос дальше - в измененном или в неизмененном виде - на следующий уровень, или, если это возможно, процедура может удовлетворить запрос, не передавая его дальше по цепочке.

Управление вводом-выводом

Одной из главных функций ОС является управление всеми устройствами ввода-вывода компьютера. ОС должна передавать устройствам команды, перехватывать прерывания и обрабатывать ошибки; она также должна обеспечивать интерфейс между устройствами и остальной частью системы. В целях развития интерфейс должен быть одинаковым для всех типов устройств (независимость от устройств).

Физическая организация устройств ввода-вывода

Устройства ввода-вывода делятся на два типа: *блок-ориентированные*устройства и *байт-ориентированные*устройства. Блок-ориентированные устройства хранят информацию в блоках фиксированного размера, каждый из которых имеет свой собственный адрес. Самое распространенное блок-ориентированное устройство - диск. Байт-ориентированные устройства не адресуемы и не позволяют производить операцию поиска, они генерируют или потребляют последовательность байтов. Примерами являются терминалы, строчные принтеры, сетевые адаптеры. Однако некоторые внешние устройства не относятся ни к одному классу, например, часы, которые, с одной стороны, не адресуемы, а с другой стороны, не порождают потока байтов. Это устройство только выдает сигнал прерывания в некоторые моменты времени.

Внешнее устройство обычно состоит из механического и электронного компонента. Электронный компонент называется контроллером устройства или адаптером. Механический компонент представляет собственно устройство. Некоторые контроллеры могут управлять несколькими устройствами. Если интерфейс между контроллером и устройством стандартизован, то независимые производители могут выпускать совместимые как контроллеры, так и устройства.

Операционная система обычно имеет дело не с устройством, а с контроллером. Контроллер, как правило, выполняет простые функции, например, преобразует поток бит в блоки, состоящие из байт, и осуществляют контроль и исправление ошибок. Каждый контроллер имеет несколько регистров, которые используются для взаимодействия с центральным процессором. В некоторых компьютерах эти регистры являются частью физического адресного пространства. В таких компьютерах нет специальных операций ввода-вывода. В других компьютерах адреса регистров ввода-вывода, называемых часто портами, образуют собственное адресное пространство за счет введения специальных операций ввода-вывода (например, команд IN и OUT в процессорах i86).

ОС выполняет ввод-вывод, записывая команды в регистры контроллера. Например, контроллер гибкого диска IBM PC принимает 15 команд, таких как READ, WRITE, SEEK, FORMAT и т.д. Когда команда принята, процессор оставляет контроллер и занимается другой работой. При завершении команды контроллер организует прерывание для того, чтобы передать управление процессором операционной системе, которая должна проверить результаты операции. Процессор получает результаты и статус устройства, читая информацию из регистров контроллера.

Организация программного обеспечения ввода-вывода

Основная идея организации программного обеспечения ввода-вывода состоит в разбиении его на несколько уровней, причем нижние уровни обеспечивают экранирование особенностей аппаратуры от верхних, а те, в свою очередь, обеспечивают удобный интерфейс для пользователей.

Ключевым принципом является независимость от устройств. Вид программы не должен зависеть от того, читает ли она данные с гибкого диска или с жесткого диска.

Очень близкой к идее независимости от устройств является идея единообразного именования, то есть для именования устройств должны быть приняты единые правила.

Другим важным вопросом для программного обеспечения ввода-вывода является обработка ошибок. Вообще говоря, ошибки следует обрабатывать как можно ближе к аппаратуре. Если контроллер обнаруживает ошибку чтения, то он должен попытаться ее скорректировать. Если же это ему не удается, то исправлением ошибок должен заняться драйвер устройства. Многие ошибки могут исчезать при повторных попытках выполнения операций ввода-вывода, например, ошибки, вызванные наличием пылинок на головках чтения или на диске. И только если нижний уровень не может справиться с ошибкой, он сообщает об ошибке верхнему уровню.

Еще один ключевой вопрос - это использование блокирующих (синхронных) и неблокирующих (асинхронных) передач. Большинство операций физического ввода-вывода выполняется асинхронно - процессор начинает передачу и переходит на другую работу, пока не наступает прерывание. Пользовательские программы намного легче писать, если операции ввода-вывода блокирующие - после команды READ программа автоматически приостанавливается до тех пор, пока данные не попадут в буфер программы. ОС выполняет операции ввода-вывода асинхронно, но представляет их для пользовательских программ в синхронной форме.

Последняя проблема состоит в том, что одни устройства являются разделяемыми, а другие - выделенными. Диски - это разделяемые устройства, так как одновременный доступ нескольких пользователей к диску не представляет собой проблему. Принтеры - это выделенные устройства, потому что нельзя смешивать строчки, печатаемые различными пользователями. Наличие выделенных устройств создает для операционной системы некоторые проблемы.

Для решения поставленных проблем целесообразно разделить программное обеспечение ввода-вывода на четыре слоя Рис. 7):

- Обработка прерываний,
- Драйверы устройств,
- Независимый от устройств слой операционной системы,
- Пользовательский слой программного обеспечения.

Обработка прерываний

Прерывания должны быть скрыты как можно глубже в недрах операционной системы, чтобы как можно меньшая часть ОС имела с ними дело. Наилучший способ состоит в разрешении процессу, инициировавшему операцию ввода-вывода, блокировать себя до

завершения операции и наступления прерывания. Процесс может блокировать себя, используя, например, вызов DOWN для семафора, или вызов WAIT для переменной условия, или вызов RECEIVE для ожидания сообщения. При наступлении прерывания процедура обработки прерывания выполняет разблокирование процесса, инициировавшего операцию ввода-вывода, используя вызовы UP, SIGNAL или посылая процессу сообщение. В любом случае эффект от прерывания будет состоять в том, что ранее заблокированный процесс теперь продолжит свое выполнение.

Драйверы устройств

Весь зависимый от устройства код помещается в драйвер устройства. Каждый драйвер управляет устройствами одного типа или, может быть, одного класса.

В операционной системе только драйвер устройства знает о конкретных особенностях какого-либо устройства. Например, только драйвер диска имеет дело с дорожками, секторами, цилиндрами, временем установления головки и другими факторами, обеспечивающими правильную работу диска.

Рис. 7 Многоуровневая организация подсистемы ввода-вывода

Драйвер устройства принимает запрос от устройств программного слоя и решает, как его выполнить. Типичным запросом является чтение n блоков данных. Если драйвер был свободен во время поступления запроса, то он начинает выполнять запрос немедленно. Если же он был занят обслуживанием другого запроса, то вновь поступивший запрос

присоединяется к очереди уже имеющихся запросов, и он будет выполнен, когда наступит его очередь.

Первый шаг в реализации запроса ввода-вывода, например, для диска, состоит в преобразовании его из абстрактной формы в конкретную. Для дискового драйвера это означает преобразование номеров блоков в номера цилиндров, головок, секторов, проверку, работает ли мотор, находится ли головка над нужным цилиндром. Короче говоря, он должен решить, какие операции контроллера нужно выполнить и в какой последовательности.

После передачи команды контроллеру драйвер должен решить, блокировать ли себя до окончания заданной операции или нет. Если операция занимает значительное время, как при печати некоторого блока данных, то драйвер блокируется до тех пор, пока операция не завершится, и обработчик прерывания не разблокирует его. Если команда ввода-вывода выполняется быстро (например, прокрутка экрана), то драйвер ожидает ее завершения без блокирования.

Независимый от устройств слой операционной системы

Большая часть программного обеспечения ввода-вывода является независимой от устройств. Точная граница между драйверами и независимыми от устройств программами определяется системой, так как некоторые функции, которые могли бы быть реализованы независимым способом, в действительности выполнены в виде драйверов для повышения эффективности или по другим причинам.

Типичными функциями для независимого от устройств слоя являются:

- обеспечение общего интерфейса к драйверам устройств,
- именование устройств,
- защита устройств,
- обеспечение независимого размера блока,
- буферизация,
- распределение памяти на блок-ориентированных устройствах,
- распределение и освобождение выделенных устройств,
- уведомление об ошибках.

Остановимся на некоторых функциях данного перечня. Верхним слоям программного обеспечения не удобно работать с блоками разной величины, поэтому данный слой обеспечивает единый размер блока, например, за счет объединения нескольких различных блоков в единый логический блок. В связи с этим верхние уровни имеют дело с абстрактными устройствами, которые используют единый размер логического блока независимо от размера физического сектора.

При создании файла или заполнении его новыми данными необходимо выделить ему новые блоки. Для этого ОС должна вести список или битовую карту свободных блоков диска. На основании информации о наличии свободного места на диске может быть

разработан алгоритм поиска свободного блока, независимый от устройства и реализуемый программным слоем, находящимся выше слоя драйверов.

Пользовательский слой программного обеспечения

Хотя большая часть программного обеспечения ввода-вывода находится внутри ОС, некоторая его часть содержится в библиотеках, связываемых с пользовательскими программами. Системные вызовы, включающие вызовы ввода-вывода, обычно делаются библиотечными процедурами. Если программа, написанная на языке С, содержит вызов

```
count = write (fd, buffer, nbytes),
```

то библиотечная процедура write будет связана с программой. Набор подобных процедур является частью системы ввода-вывода. В частности, форматирование ввода или вывода выполняется библиотечными процедурами. Примером может служить функция printf языка С, которая принимает строку формата и, возможно, некоторые переменные в качестве входной информации, затем строит строку символов ASCII и делает вызов write для вывода этой строки. Стандартная библиотека ввода-вывода содержит большое число процедур, которые выполняют ввод-вывод и работают как часть пользовательской программы.

Другой категорией программного обеспечения ввода-вывода является подсистема спулинга (spooling). Спулинг - это способ работы с выделенными устройствами в мультипрограммной системе. Рассмотрим типичное устройство, требующее спулинга - строчный принтер. Хотя технически легко позволить каждому пользовательскому процессу открыть специальный файл, связанный с принтером, такой способ опасен из-за того, что пользовательский процесс может монополизировать принтер на произвольное время. Вместо этого создается специальный процесс - монитор, который получает исключительные права на использование этого устройства. Также создается специальный каталог, называемый каталогом спулинга. Для того, чтобы напечатать файл, пользовательский процесс помещает выводимую информацию в этот файл и помещает его в каталог спулинга. Процесс-монитор по очереди распечатывает все файлы, содержащиеся в каталоге спулинга.

Что такое ЛВС?

Локальная сеть (ЛВС) представляет собой коммуникационную систему, позволяющую совместно использовать ресурсы компьютеров, подключенных к сети, таких как принтеры, плоттеры, диски, модемы, приводы CD-ROM и другие периферийные устройства. Локальная сеть обычно ограничена территориально одним или несколькими близко расположенными зданиями. Каждый компьютер в составе ЛВС должен иметь следующие компоненты:

1. сетевой адаптер;
2. кабель;
3. сетевая операционная система (сетевые программы).

Сетевые адаптеры

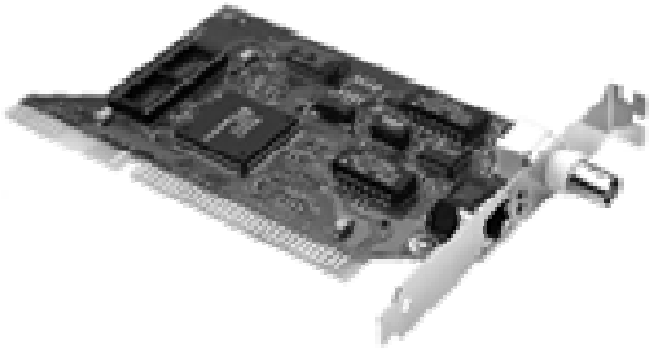


Рис. 8 Сетевой адаптер

Сетевые адаптеры (Рис. 8) и кабели являются аппаратной основой организации компьютерных сетей, их нормальная работа жизненно важна для сети. С кабелями и адаптерами связано обычно 80% неполадок в сети.

В каждом компьютере должен быть установлен сетевой адаптер, обеспечивающий подключение к выбранному типу кабеля.

Функцией сетевого адаптера является передача и прием сетевых сигналов из кабеля. Адаптер воспринимает команды и данные от сетевой операционной системы (ОС), преобразует эту информацию в один из стандартных форматов и передает ее в сеть через подключенный к адаптеру кабель.

Конфигурация адаптера

Каждый адаптер, устанавливаемый в компьютер, должен нормально работать с остальной частью ПК. Нужно всегда обращать внимание на два важнейших параметра каждого устройства, устанавливаемого в компьютер.

I/Obase

Базовый адрес ввода-вывода является "каналом", по которому адаптер взаимодействует с другими компонентами компьютера. Каждое устройство должно использовать уникальный диапазон адресов ввода-вывода.

IRQ

Номер линии запроса прерывания. Запрос прерывания является сигналом, передаваемым устройством для того, чтобы привлечь внимание процессора (прервать его текущую деятельность). Такой сигнал обычно подается при появлении новых данных или завершении той или иной операции. Каждое устройство должно использовать уникальное значение IRQ.

Таблица 1 содержит адреса ввода-вывода и номера IRQ, используемые различными устройствами.

Таблица 1 Адреса ввода-вывода

Устройство	IRQ	IOBase	Адреса памяти	Канал ПДП (DMA)
Системный таймер	0			
Клавиатура	1			
Коммуникационный порт COM1	4	3F8-3FF		
Коммуникационный порт COM2	3	2F8-2FF		
Параллельный (принтерный) порт LPT1	7	378-37F		
Параллельный (принтерный) порт LPT2 [На компьютерах за исключением XT]	5	278-27F		
Дисковый контроллер XT	5	320-32F	C800-CBFF	3
Дисковый контроллер AT	14	1F0-1F8		
Контроллер SCSI	Различные			1,3
Адаптер VGA	2,9 или без IRQ	3C0-3DA	A0000-BFFFF	0
VGA BIOS			C0000-C7FFF	
Адаптер EGA	2,9 или без IRQ	3C0-3CF	A0000-AFFFF B0000-BFFFF	0
EGA BIOS			C0000-C7FFF	
Адаптер MDA		3B0-3BF	B0000-B3FFF	0
Адаптер Hercules		3B4-3BF	B0000-BfFFF	
Адаптер CGA		3D0-3DF	B8000-B3FFF	0
Буфер символов CGA, EGA			B8000-BBFFF	
Игровой порт (джойстик)		200-20F		

Сетевые кабели

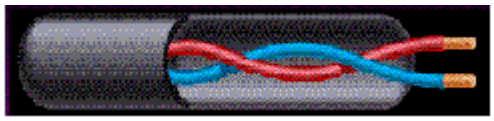


Рис. 9 Сетевой кабель

Кабель (Рис. 9) обеспечивает канал связи компьютера с остальными машинами сети. При установке кабелей нужно точно следовать спецификациям. Пренебрежение этим правилом может принести очень много неприятностей. Отметим разницу между кабелем и кабельным сегментом говоря о кабеле, будем всегда иметь в виду отрезок провода, соединяющего два узла сети; сегментом же будем называть весь комплект кабелей от одного конца сети до другого (между терминаторами). Терминаторы представляют собой резисторы, устанавливаемые на обоих концах сегмента для согласования волнового сопротивления кабеля. Сигнал, дошедший до конца сегмента, поглощается терминатором - это позволяет избавиться от паразитных отраженных сигналов в сети. Если терминаторы не устанавливать, отраженный от конца кабеля сигнал снова попадает в кабель - этот отраженный сигнал будет являться в данном случае помехой и может породить множество проблем вплоть до полной неработоспособности сети.

Кабель на основе скрученных пар (витая пара, TP).

Кабель содержит две или более пары проводов, скрученных один с другим по всей длине кабеля. Скручивание позволяет повысить помехоустойчивость кабеля и снизить влияние каждой пары на все остальные.

Коаксиальный кабель

Рис. 10 Коаксиальный кабель

Состоит из центрального проводника (сплошного или многожильного), покрытого слоем полимерного изолятора, поверх которого расположен другой проводник (экран) (Рис. 10). Экран представляет собой оплетку из медного провода вокруг изолятора или обернутую вокруг изолятора фольгу. В высококачественных кабелях присутствуют и оплетка и фольга. Коаксиальный кабель обеспечивает более высокую помехоустойчивость по сравнению с витой парой, но он дороже. Существуют различные виды коаксиальных кабелей. При установке сети следует выбирать кабель в точном соответствии со спецификацией.

Оптический кабель

Состоит из одного или нескольких кварцевых волокон (иногда полимерных), покрытых защитной оболочкой. Оболочка как правило состоит из нескольких слоев для обеспечения лучшей защиты волокон.

Архитектура сети

Сетевая архитектура сродни архитектуре строений. Архитектура здания отражает стиль конструкций и материалы, используемые для постройки. Архитектура сети описывает не только физическое расположение сетевых устройств, но и тип используемых адаптеров и кабелей. Кроме того, сетевая архитектура определяет методы передачи данных по кабелю.

Топология сети

Топология сети описывает схему физического соединения компьютеров. Существуют 3 основных типа сетевой топологии:

Общая шина.

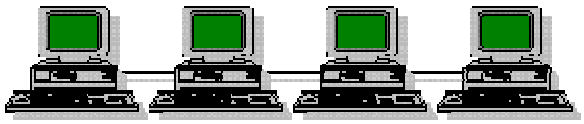


Рис. 11 Топология "Общая шина"

При использовании шинной топологии (Рис. 11) компьютеры соединяются в одну линию, по концам которой устанавливают терминаторы. Преимущества шинной топологии заключаются в простоте организации сети и низкой стоимости. Недостатком является низкая устойчивость к повреждениям - при любом обрыве кабеля вся сеть перестает работать, а поиск повреждения весьма затруднителен.

Звезда.

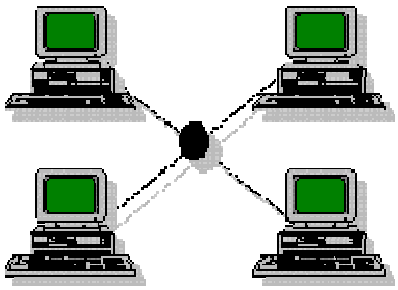


Рис. 12 Топология "Звезда"

При использовании топологии "звезда" Рис. 12), каждый компьютер подключается к специальному концентратору (хабу). Преимуществом этой топологии является ее устойчивость к повреждениям кабеля - при обрыве перестает работать только один из узлов сети и поиск повреждения значительно упрощается. Недостатком является более высокая стоимость.

Кольцо.

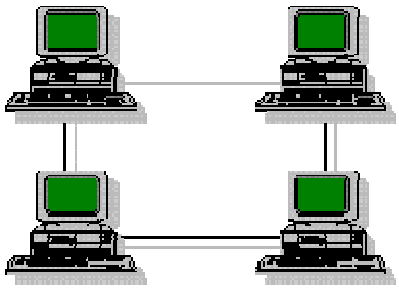


Рис. 13 Топология "кольцо"

При такой топологии узлы сети образуют виртуальное кольцо (концы кабеля соединены друг с другом, Рис. 13). Каждый узел сети соединен с двумя соседними. Эту топологию активно продвигает фирма IBM (сети Token Ring). Преимуществом кольцевой топологии является ее высокая надежность (за счет избыточности), однако стоимость такой сети достаточно высока за счет расходов на адаптеры, кабели и дополнительные приспособления.

Спецификации IEEE

Каждая сеть должна следовать определенным правилам (протоколам) при передачи данных от одного компьютера к другому. Протокол определяет способ доступа узла к передающей среде (кабелю) и способ передачи информации от одного узла к другому.

В настоящее время используется два типа протоколов доступа к среде:

- **передача маркера** (token) используется в сетях IBM Token Ring и FDDI;
- **множественный доступ с детектированием несущей** (CSMA) используется в сетях Ethernet.

Протокол Ethernet был разработан в 1973 году компанией Xerox и развит впоследствии ею совместно с Intel и Digital Equipment Corp. С тех пор этот протокол стал международным стандартом организации компьютерных сетей. Стандарт был документирован и развит институтом IEEE и получил известность как спецификация IEEE 802.3. IEEE представляет собой организацию Международного Комитета по Стандартам (ISO), ответственной за выработку сетевых стандартов.

Серверы и рабочие станции

Сеть представляет собой не просто компьютеры, соединенные вместе кабелем. Сеть - это набор компьютеров, осуществляющих обмен данными между собой с определенными целями. Если компьютер Ивана хочет "разговаривать" с машиной Марии, это обычно означает, что Иван что-то хочет получить от Марии. Для удовлетворения таких потребностей, каждый компьютер должен осуществлять определенные функции. Независимо от используемых на каждом компьютере приложений все машины сети делятся на два класса - серверы и рабочие станции.

Что такое сервер?

Сервером будем называть компьютер, предоставляющий свои ресурсы (например, диски) другим компьютерам сети. Серверы предоставляют свои ресурсы рабочим станциям.

Что такое рабочая станция?

Рабочая станция или клиент использует ресурсы сервера. Рабочие станции имеют доступ к сетевым ресурсам, но своих ресурсов в общее пользование не предоставляют.

С сетевыми ресурсами обычно связывают локальные имена (A-Z для дисков; LPTx, COMx - для портов).

Сети с выделенными серверами и одноранговые сети

- Сети с архитектурой клиент-сервер используют центральный сервер для обслуживания запросов клиентов Рис. 14а), тогда как одноранговые сети позволяют любой рабочей станции функционировать одновременно в качестве сервера, если этого требуют задачи Рис. 14б).
- По сравнению с универсально одноранговой архитектурой сеть клиент-сервер более специализирована.

Рис. 14 Сети с выделенными серверами а) и одноранговые сети б)

Сети с архитектурой клиент-сервер

- Специализированный компьютер (выделенный сервер) используется для установки всех разделяемых ресурсов. Такое решение ускоряет доступ пользователей к централизованным ресурсам сети.
- Сетевое администрирование проще за счет незначительного числа серверов в сети и их узкой специализации.
- Высокие требования к выделенному серверу обеспечение высокой производительности требует установки на сервере большого количества оперативной памяти, диска большого размера и использования в сервере производительного процессора.
- При нарушении работы сервера сеть становится практически неработоспособной.

Одноранговые сети

- Сетевые приложения могут быть распределены по многочисленным серверам для повышения производительности сети и снижения расходов.
- Гибкое разделение ресурсов любого узла сети.
- Администрирование одноранговой сети может быть сложнее за счет большего числа серверов и более развитых возможностей каждого сервера.
- Невыделенные серверы медленнее специализированных.
- В сети LANtastic могут использоваться как выделенные серверы, так и невыделенные серверы/рабочие станции.

Распределенные (глобальные) сети

В общем случае сети, расположенные в одном (или нескольких близко расположенных) здании называется локальной (ЛВС). Сети, объединяющие компьютеры в разных зданиях, городах и странах, она называется распределенной (WAN -Wide Area Network). Распределенные сети очень большого масштаба (например, Internet, EUNET, Relcom, FIDO) часто называют глобальными. Распределенные сети состоят из трех основных компонент:

1. Локальные сети, как узлы распределенной сети
2. Каналы, соединяющие ЛВС.
3. Оборудование и программы, обеспечивающие локальным сетям доступ к каналам связи.

Оборудование распределенных сетей

Для объединения локальных сетей требуется специальное оборудование независимо от того, находятся ли эти ЛВС в одном здании или связаны через распределенную сеть.

- Повторители (Repeater) - усиливают полученный из кабельного сегмента сигнал и передают его в другой сегмент.

- объединяют идентичные ЛВС;
- простое усиление сигналов.

- Мосты (Bridge) передают сообщения на основе записей в таблице пересылки.
  - Возможность фильтрации сетевого трафика;
  - сохраняет информацию о всех узлах;
  - соединяет идентичные или разные сети (например, Ethernet и Token Ring).
- Маршрутизаторы (Router) обеспечивают выбор маршрута обмена данными между узлами сети.
  - Принимает решение о выборе "лучшего пути";
  - Дистанция обычно оценивается в интервалах (hop) - промежутках между двумя соседними маршрутизаторами на пути от отправителя к получателю.

Сетевые операционные системы

Структура сетевой операционной системы

Сетевая операционная система составляет основу любой вычислительной сети. Каждый компьютер в сети в значительной степени автономен, поэтому под сетевой операционной системой в широком смысле понимается совокупность операционных систем отдельных компьютеров, взаимодействующих с целью обмена сообщениями и разделения ресурсов по единым правилам - протоколам. В узком смысле сетевая ОС - это операционная система отдельного компьютера, обеспечивающая ему возможность работать в сети.

Рис. 15 Структура сетевой ОС

В сетевой операционной системе отдельной машины можно выделить несколько частей (Рис. 15):

- Средства управления локальными ресурсами компьютера: функции распределения оперативной памяти между процессами, планирования и диспетчеризации процессов, управления процессорами в мультипроцессорных машинах, управления периферийными устройствами и другие функции управления ресурсами локальных ОС.
- Средства предоставления собственных ресурсов и услуг в общее пользование - серверная часть ОС (сервер). Эти средства обеспечивают, например, блокировку файлов и записей, что необходимо для их совместного использования; ведение справочников имен сетевых ресурсов; обработку запросов удаленного доступа к собственной файловой системе и базе данных; управление очередями запросов удаленных пользователей к своим периферийным устройствам.
- Средства запроса доступа к удаленным ресурсам и услугам и их использования - клиентская часть ОС (редиректор). Эта часть выполняет распознавание и перенаправление в сеть запросов к удаленным ресурсам от приложений и пользователей, при этом запрос поступает от приложения в локальной

форме, а передается в сеть в другой форме, соответствующей требованиям сервера. Клиентская часть также осуществляет прием ответов от серверов и преобразование их в локальный формат, так что для приложения выполнение локальных и удаленных запросов неразличимо.

- Коммуникационные средства ОС, с помощью которых происходит обмен сообщениями в сети. Эта часть обеспечивает адресацию и буферизацию сообщений, выбор маршрута передачи сообщения по сети, надежность передачи и т.п., то есть является средством транспортировки сообщений.

В зависимости от функций, возлагаемых на конкретный компьютер, в его операционной системе может отсутствовать либо клиентская, либо серверная части.

На Рис. 16 показано взаимодействие сетевых компонентов. Здесь компьютер 1 выполняет роль "чистого" клиента, а компьютер 2 - роль "чистого" сервера, соответственно на первой машине отсутствует серверная часть, а на второй - клиентская. На рисунке отдельно показан компонент клиентской части - редиректор. Именно редиректор перехватывает все запросы, поступающие от приложений, и анализирует их. Если выдан запрос к ресурсу данного компьютера, то он переадресовывается соответствующей подсистеме локальной ОС, если же это запрос к удаленному ресурсу, то он переправляется в сеть. При этом клиентская часть преобразует запрос из локальной формы в сетевой формат и передает его транспортной подсистеме, которая отвечает за доставку сообщений указанному серверу. Серверная часть операционной системы компьютера 2 принимает запрос, преобразует его и передает для выполнения своей локальной ОС. После того, как результат получен, сервер обращается к транспортной подсистеме и направляет ответ клиенту, выдавшему запрос. Клиентская часть преобразует результат в соответствующий формат и адресует его тому приложению, которое выдало запрос.

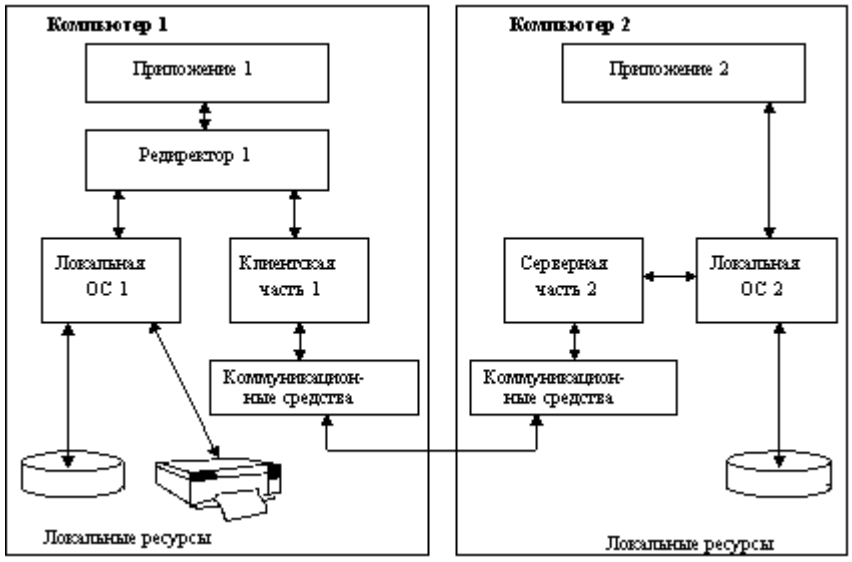


Рис. 16 Взаимодействие компонентов операционной системы при взаимодействии компьютеров

На практике сложилось несколько подходов к построению сетевых операционных систем (Рис. 17).

Первые сетевые ОС представляли собой совокупность существующей локальной ОС и надстроенной над ней*сетевой оболочки*. При этом в локальную ОС встраивался минимум сетевых функций, необходимых для работы сетевой оболочки, которая выполняла основные сетевые функции. Примером такого подхода является использование на каждой машине сети операционной системы MS DOS (у которой начиная с ее третьей версии появились такие встроенные функции, как блокировка файлов и записей, необходимые для совместного доступа к файлам). Принцип построения сетевых ОС в виде сетевой оболочки над локальной ОС используется и в современных ОС, таких, например, как LANtastic или Personal Ware.

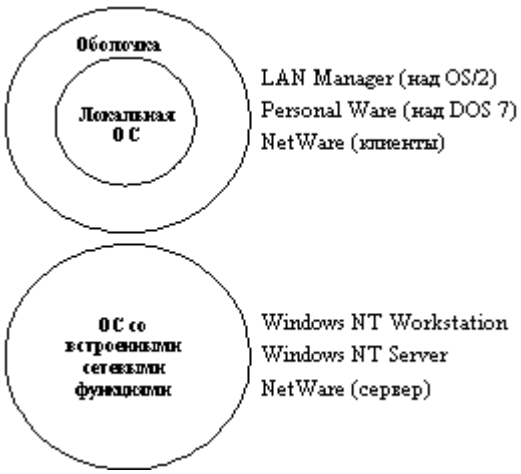


Рис. 17 Варианты построения сетевых ОС

Однако более эффективным представляется путь разработки операционных систем, изначально предназначенных для работы в сети. Сетевые функции у ОС такого типа глубоко *встроены* в основные модули системы, что обеспечивает их логическую стройность, простоту эксплуатации и модификации, а также высокую производительность. Примером такой ОС является система Windows NT фирмы Microsoft, которая за счет встроенности сетевых средств обеспечивает более высокие показатели производительности и защищенности информации по сравнению с сетевой ОС LAN Manager той же фирмы (совместная разработка с IBM), являющейся надстройкой над локальной операционной системой OS/2.

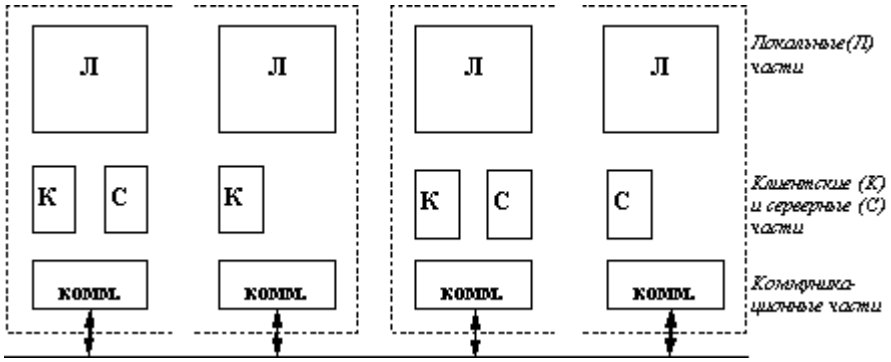
Одноранговые сетевые ОС и ОС с выделенными серверами

В зависимости от того, как распределены функции между компьютерами сети, сетевые операционные системы, а следовательно, и сети делятся на два класса: одноранговые и двухранговые (Рис. 18). Последние чаще называют сетями с выделенными серверами.

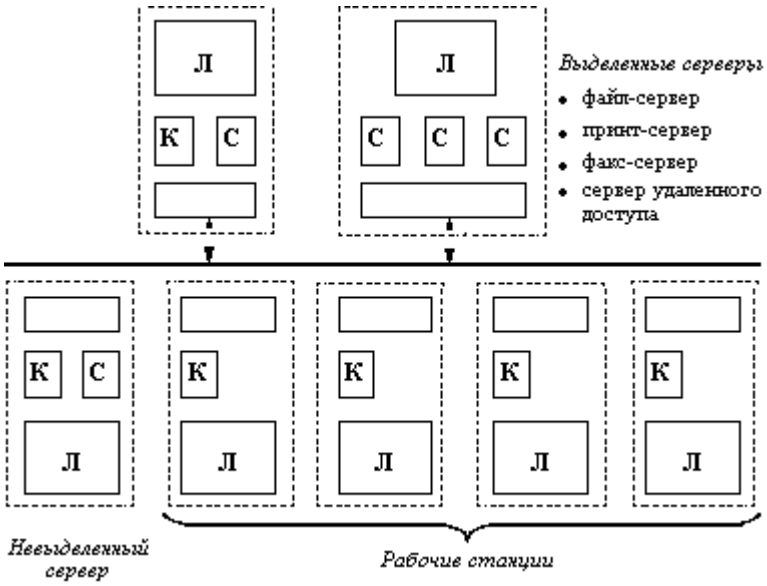
Если компьютер предоставляет свои ресурсы другим пользователям сети, то он играет роль сервера. При этом компьютер, обращающийся к ресурсам другой машины, является клиентом. Как уже было сказано, компьютер, работающий в сети, может выполнять функции либо клиента, либо сервера, либо совмещать обе эти функции.

Если выполнение каких-либо серверных функций является основным назначением компьютера (например, предоставление файлов в общее пользование всем остальным пользователям сети или организация совместного использования факса, или предоставление всем пользователям сети возможности запуска на данном компьютере своих приложений), то такой компьютер называется выделенным сервером. В зависимости от того, какой ресурс сервера является разделяемым, он называется файл-сервером, факс-сервером, принт-сервером, сервером приложений и т.д.

Очевидно, что на выделенных серверах желательно устанавливать ОС, специально оптимизированные для выполнения тех или иных серверных функций. Поэтому в сетях с выделенными серверами чаще всего используются сетевые операционные системы, в состав которых входит нескольких вариантов ОС, отличающихся возможностями серверных частей. Например, сетевая ОС Novell NetWare имеет серверный вариант, оптимизированный для работы в качестве файл-сервера, а также варианты оболочек для рабочих станций с различными локальными ОС, причем эти оболочки выполняют исключительно функции клиента. Другим примером ОС, ориентированной на построение сети с выделенным сервером, является операционная система Windows NT. В отличие от NetWare, оба варианта данной сетевой ОС - Windows NT Server (для выделенного сервера) и Windows NT Workstation (для рабочей станции) - могут поддерживать функции и клиента и сервера. Но серверный вариант Windows NT имеет больше возможностей для предоставления ресурсов своего компьютера другим пользователям сети, так как может выполнять более широкий набор функций, поддерживает большее количество одновременных соединений с клиентами, реализует централизованное управление сетью, имеет более развитые средства защиты.



(а)



(б)

Рис. 18 (а) - Одноранговая сеть, (б) - Двухранговая сеть

Выделенный сервер не принято использовать в качестве компьютера для выполнения текущих задач, не связанных с его основным назначением, так как это может уменьшить производительность его работы как сервера. В связи с такими соображениями в ОС Novell NetWare на серверной части возможность выполнения обычных прикладных программ вообще не предусмотрена, то есть сервер не содержит клиентской части, а на рабочих станциях отсутствуют серверные компоненты. Однако в других сетевых ОС функционирование на выделенном сервере клиентской части вполне возможно. Например, под управлением Windows NT Server могут запускаться обычные программы локального пользователя, которые могут потребовать выполнения клиентских функций ОС при появлении запросов к ресурсам других компьютеров сети. При этом рабочие станции, на которых установлена ОС Windows NT Workstation, могут выполнять функции невыделенного сервера.

Важно понять, что несмотря на то, что в сети с выделенным сервером все компьютеры в общем случае могут выполнять одновременно роли и сервера, и клиента, эта сеть функционально не симметрична: аппаратно и программно в ней реализованы два типа компьютеров - одни, в большей степени ориентированные на выполнение серверных функций и работающие под управлением специализированных серверных ОС, а другие - в основном выполняющие клиентские функции и работающие под управлением соответствующего этому назначению варианта ОС. Функциональная несимметричность, как правило, вызывает и несимметричность аппаратуры - для выделенных серверов используются более мощные компьютеры с большими объемами оперативной и внешней памяти. Таким образом, функциональная несимметричность в сетях с выделенным сервером сопровождается несимметричностью операционных систем (специализация ОС) и аппаратной несимметричностью (специализация компьютеров).

В одноранговых сетях все компьютеры равны в правах доступа к ресурсам друг друга. Каждый пользователь может по своему желанию объявить какой-либо ресурс своего компьютера разделяемым, после чего другие пользователи могут его эксплуатировать. В таких сетях на всех компьютерах устанавливается одна и та же ОС, которая предоставляет всем компьютерам в сети *потенциально*равные возможности. Одноранговые сети могут быть

построены, например, на базе ОС LANtastic, Personal Ware, Windows for Workgroup, Windows NT Workstation.

В одноранговых сетях также может возникнуть функциональная несимметричность: одни пользователи не желают разделять свои ресурсы с другими, и в таком случае их компьютеры выполняют роль клиента, за другими компьютерами администратор закрепил только функции по организации совместного использования ресурсов, а значит они являются серверами, в третьем случае, когда локальный пользователь не возражает против использования его ресурсов и сам не исключает возможности обращения к другим компьютерам, ОС, устанавливаемая на его компьютере, должна включать и серверную, и клиентскую части. В отличие от сетей с выделенными серверами, в одноранговых сетях отсутствует специализация ОС в зависимости от преобладающей функциональной направленности - клиента или сервера. Все вариации реализуются средствами конфигурирования одного и того же варианта ОС.

Одноранговые сети проще в организации и эксплуатации, однако они применяются в основном для объединения небольших групп пользователей, не предъявляющих больших требований к объемам хранимой информации, ее защищенности от несанкционированного доступа и к скорости доступа. При повышенных требованиях к этим характеристикам более подходящими являются двухранговые сети, где сервер лучше решает задачу обслуживания пользователей своими ресурсами, так как его аппаратура и сетевая операционная система специально спроектированы для этой цели.

ОС для рабочих групп и ОС для сетей масштаба предприятия

Сетевые операционные системы имеют разные свойства в зависимости от того, предназначены они для сетей масштаба рабочей группы (отдела), для сетей масштаба кампуса или для сетей масштаба предприятия.

- *Сети отделов* - используются небольшой группой сотрудников, решающих общие задачи. Главной целью сети отдела является разделение локальных ресурсов, таких как приложения, данные, лазерные принтеры и модемы. Сети отделов обычно не разделяются на подсети.
- *Сети кампусов* - соединяют несколько сетей отделов внутри отдельного здания или внутри одной территории предприятия. Эти сети являются все еще локальными сетями, хотя и могут покрывать территорию в несколько квадратных километров. Сервисы такой сети включают взаимодействие между сетями отделов, доступ к базам данных предприятия, доступ к факс-серверам, высокоскоростным модемам и высокоскоростным принтерам.
- *Сети предприятия (корпоративные сети)* - объединяют все компьютеры всех территорий отдельного предприятия. Они могут покрывать город, регион или даже континент. В таких сетях пользователям предоставляется доступ к информации и приложениям, находящимся в других рабочих группах, других отделах, подразделениях и штаб-квартирах корпорации.

Главной задачей операционной системы, используемой в сети масштаба отдела, является организация разделения ресурсов, таких как приложения, данные, лазерные принтеры и, возможно, низкоскоростные модемы. Обычно сети отделов имеют один или два файловых сервера и не более чем 30 пользователей. Задачи управления на уровне отдела

относительно просты. В задачи администратора входит добавление новых пользователей, устранение простых отказов, инсталляция новых узлов и установка новых версий программного обеспечения. Операционные системы сетей отделов хорошо отработаны и разнообразны, также, как и сами сети отделов, уже давно применяющиеся и достаточно отлаженные. Такая сеть обычно использует одну или максимум две сетевые ОС. Чаще всего это сеть с выделенным сервером NetWare 3.x или Windows NT, или же одноранговая сеть, например сеть Windows for Workgroups.

Пользователи и администраторы сетей отделов вскоре осознают, что они могут улучшить эффективность своей работы путем получения доступа к информации других отделов своего предприятия. Если сотрудник, занимающийся продажами, может получить доступ к характеристикам конкретного продукта и включить их в презентацию, то эта информация будет более свежей и будет оказывать большее влияние на покупателей. Если отдел маркетинга может получить доступ к характеристикам продукта, который еще только разрабатывается инженерным отделом, то он может быстро подготовить маркетинговые материалы сразу же после окончания разработки.

Итак, следующим шагом в эволюции сетей является объединение локальных сетей нескольких отделов в единую сеть здания или группы зданий. Такие сети называют сетями кампусов. Сети кампусов могут простираться на несколько километров, но при этом глобальные соединения не требуются.

Операционная система, работающая в сети кампуса, должна обеспечивать для сотрудников одних отделов доступ к некоторым файлам и ресурсам сетей других отделов. Услуги, предоставляемые ОС сетей кампусов, не ограничиваются простым разделением файлов и принтеров, а часто предоставляют доступ и к серверам других типов, например, к факс-серверам и к серверам высокоскоростных модемов. Важным сервисом, предоставляемым операционными системами данного класса, является доступ к корпоративным базам данных, независимо от того, располагаются ли они на серверах баз данных или на миникомпьютерах.

Именно на уровне сети кампуса начинаются проблемы интеграции. В общем случае, отделы уже выбрали для себя типы компьютеров, сетевого оборудования и сетевых операционных систем. Например, инженерный отдел может использовать операционную систему UNIX и сетевое оборудование Ethernet, отдел продаж может использовать операционные среды DOS/Novell и оборудование Token Ring. Очень часто сеть кампуса соединяет разнородные компьютерные системы, в то время как сети отделов используют однотипные компьютеры.

Корпоративная сеть соединяет сети всех подразделений предприятия, в общем случае находящихся на значительных расстояниях. Корпоративные сети используют глобальные связи (WAN links) для соединения локальных сетей или отдельных компьютеров.

Пользователям корпоративных сетей требуются все те приложения и услуги, которые имеются в сетях отделов и кампусов, плюс некоторые дополнительные приложения и услуги, например, доступ к приложениям мейнфреймов и миникомпьютеров и к глобальным связям. Когда ОС разрабатывается для локальной сети или рабочей группы, то ее главной обязанностью является разделение файлов и других сетевых ресурсов (обычно принтеров) между локально подключенными пользователями. Такой подход не применим для уровня предприятия. Наряду с базовыми сервисами, связанными с разделением файлов и принтеров, сетевая ОС, которая разрабатывается для корпораций, должна поддерживать более широкий набор сервисов, в который обычно входят почтовая служба, средства коллективной работы,

поддержка удаленных пользователей, факс-сервис, обработка голосовых сообщений, организация видеоконференций и др.

Кроме того, многие существующие методы и подходы к решению традиционных задач сетей меньших масштабов для корпоративной сети оказались непригодными. На первый план вышли такие задачи и проблемы, которые в сетях рабочих групп, отделов и даже кампусов либо имели второстепенное значение, либо вообще не проявлялись. Например, простейшая для небольшой сети задача ведения учетной информации о пользователях выросла в сложную проблему для сети масштаба предприятия. А использование глобальных связей требует от корпоративных ОС поддержки протоколов, хорошо работающих на низкоскоростных линиях, и отказа от некоторых традиционно используемых протоколов (например, тех, которые активно используют широковещательные сообщения). Особое значение приобрели задачи преодоления гетерогенности - в сети появились многочисленные шлюзы, обеспечивающие согласованную работу различных ОС и сетевых системных приложений.

К признакам корпоративных ОС могут быть отнесены также следующие особенности.

Поддержка приложений. В корпоративных сетях выполняются сложные приложения, требующие для выполнения большой вычислительной мощности. Такие приложения разделяются на несколько частей, например, на одном компьютере выполняется часть приложения, связанная с выполнением запросов к базе данных, на другом - запросов к файловому сервису, а на клиентских машинах - часть, реализующая логику обработки данных приложения и организующая интерфейс с пользователем. Вычислительная часть общих для корпорации программных систем может быть слишком объемной и неподъемной для рабочих станций клиентов, поэтому приложения будут выполняться более эффективно, если их наиболее сложные в вычислительном отношении части перенести на специально предназначенный для этого мощный компьютер - *сервер приложений*.

Сервер приложений должен базироваться на мощной аппаратной платформе (мультипроцессорные системы, часто на базе RISC-процессоров, специализированные кластерные архитектуры). ОС сервера приложений должна обеспечивать высокую производительность вычислений, а значит поддерживать многонитевую обработку, вытесняющую многозадачность, мультипроцессирование, виртуальную память и наиболее популярные прикладные среды (UNIX, Windows, MS-DOS, OS/2). В этом отношении сетевую ОС NetWare трудно отнести к корпоративным продуктам, так как в ней отсутствуют почти все требования, предъявляемые к серверу приложений. В то же время хорошая поддержка универсальных приложений в Windows NT собственно и позволяет ей претендовать на место в мире корпоративных продуктов.

*Справочная служба.*Корпоративная ОС должна обладать способностью хранить информацию обо всех пользователях и ресурсах таким образом, чтобы обеспечивалось управление ею из одной центральной точки. Подобно большой организации, корпоративная сеть нуждается в централизованном хранении как можно более полной справочной информации о самой себе (начиная с данных о пользователях, серверах, рабочих станциях и кончая данными о кабельной системе). Естественно организовать эту информацию в виде базы данных. Данные из этой базы могут быть востребованы многими сетевыми системными приложениями, в первую очередь системами управления и администрирования. Кроме этого, такая база полезна при организации электронной почты, систем коллективной работы, службы безопасности, службы инвентаризации программного и аппаратного обеспечения сети, да и для практически любого крупного бизнес-приложения.

База данных, хранящая справочную информацию, предоставляет все то же многообразие возможностей и порождает все то же множество проблем, что и любая другая крупная база данных. Она позволяет осуществлять различные операции поиска, сортировки, модификации и т.п., что очень сильно облегчает жизнь как администраторам, так и пользователям. Но за эти удобства приходится расплачиваться решением проблем распределенности, репликации и синхронизации.

В идеале сетевая справочная информация должна быть реализована в виде единой базы данных, а не представлять собой набор баз данных, специализирующихся на хранении информации того или иного вида, как это часто бывает в реальных операционных системах. Например, в Windows NT имеется по крайней мере пять различных типов справочных баз данных. Главный справочник домена (NT Domain Directory Service) хранит информацию о пользователях, которая используется при организации их логического входа в сеть. Данные о тех же пользователях могут содержаться и в другом справочнике, используемом электронной почтой Microsoft Mail. Еще три базы данных поддерживают разрешение низкоуровневых адресов: WINS - устанавливает соответствие Netbios-имен IP-адресам, справочник DNS - сервер имен домена - оказывается полезным при подключении NT-сети к Internet, и наконец, справочник протокола DHCP используется для автоматического назначения IP-адресов компьютерам сети. Ближе к идеалу находятся справочные службы, поставляемые фирмой Banyan (продукт Streettalk III) и фирмой Novell (NetWare Directory Services), предлагающие единый справочник для всех сетевых приложений. Наличие единой справочной службы для сетевой операционной системы - один из важнейших признаков ее корпоративности.

Безопасность. Особую важность для ОС корпоративной сети приобретают вопросы безопасности данных. С одной стороны, в крупномасштабной сети объективно существует больше возможностей для несанкционированного доступа - из-за децентрализации данных и большой распределенности "законных" точек доступа, из-за большого числа пользователей, благонадежность которых трудно установить, а также из-за большого числа возможных точек несанкционированного подключения к сети. С другой стороны, корпоративные бизнес-приложения работают с данными, которые имеют жизненно важное значение для успешной работы корпорации в целом. И для защиты таких данных в корпоративных сетях наряду с различными аппаратными средствами используется весь спектр средств защиты, предоставляемый операционной системой: избирательные или мандатные права доступа, сложные процедуры аутентификации пользователей, программная шифрация.

Управление распределенными ресурсами

Базовые примитивы передачи сообщений в распределенных системах

Единственным по-настоящему важным отличием распределенных систем от централизованных является межпроцессная взаимосвязь. В централизованных системах связь между процессами, как правило, предполагает наличие разделяемой памяти. Типичный пример - проблема "поставщик-потребитель", в этом случае один процесс пишет в разделяемый буфер, а другой - читает из него. Даже наиболее простая форма синхронизации - семафор - требует, чтобы хотя бы одно слово (переменная самого семафора) было разделяемым. В распределенных системах нет какой бы то ни было разделяемой памяти, таким образом вся природа межпроцессных коммуникаций должна быть продумана заново.

Основой этого взаимодействия может служить только передача по сети сообщений. В самом простом случае системные средства обеспечения связи могут быть сведены к двум

основным системным вызовам (примитивам), один - для посылки сообщения, другой - для получения сообщения. В дальнейшем на их базе могут быть построены более мощные средства сетевых коммуникаций, такие как распределенная файловая система или вызов удаленных процедур, которые, в свою очередь, также могут служить основой для построения других сетевых сервисов.

Несмотря на концептуальную простоту этих системных вызовов - ПОСЛАТЬ и ПОЛУЧИТЬ - существуют различные варианты их реализации, от правильного выбора которых зависит эффективность работы сети. В частности, эффективность коммуникаций в сети зависит от способа задания адреса, от того, является ли системный вызов блокирующим или неблокирующим, какие выбраны способы буферизации сообщений и насколько надежным является протокол обмена сообщениями.

Способы адресации

Для того, чтобы послать сообщение, необходимо указать адрес получателя. В очень простой сети адрес может задаваться в виде константы, но в более сложных сетях нужен и более изощренный способ адресации.

Одним из вариантов адресации на верхнем уровне является использование физических адресов сетевых адаптеров. Если в получающем компьютере выполняется только один процесс, то ядро будет знать, что делать с поступившим сообщением - передать его этому процессу. Однако, если на машине выполняется несколько процессов, то ядру не известно, какому из них предназначено сообщение, поэтому использование сетевого адреса адаптера в качестве адреса получателя приводит к очень серьезному ограничению - на каждой машине должен выполняться только один процесс.

Альтернативная адресная система использует имена назначения, состоящие из двух частей, определяющие номер машины и номер процесса. Однако адресация типа "машина-процесс" далека от идеала, в частности, она не гибка и не прозрачна, так как пользователь должен явно задавать адрес машины-получателя. В этом случае, если в один прекрасный день машина, на которой работает сервер, отказывает, то программа, в которой жестко используется адрес сервера, не сможет работать с другим сервером, установленном на другой машине.

Другим вариантом могло бы быть назначение каждому процессу уникального адреса, который никак не связан с адресом машины. Одним из способов достижения этой цели является использование централизованного механизма распределения адресов процессов, который работает просто, как счетчик. При получении запроса на выделение адреса он просто возвращает текущее значение счетчика, а затем наращивает его на единицу. Недостатком этой схемы является то, что централизованные компоненты, подобные этому, не обеспечивают в достаточной степени расширяемость систем. Еще один метод назначения процессам уникальных идентификаторов заключается в разрешении каждому процессу выбора своего собственного идентификатора из очень большого адресного пространства, такого как пространство 64-х битных целых чисел. Вероятность выбора одного и того же числа двумя процессами является ничтожной, а система хорошо расширяется. Однако здесь имеется одна проблема: как процесс-отправитель может узнать номер машины процесса-получателя. В сети, которая поддерживает широковещательный режим (то есть в ней предусмотрен такой адрес, который принимают все сетевые адаптеры), отправитель может широковещательно передать специальный пакет, который содержит идентификатор процесса назначения. Все ядра получат эти сообщения, проверят адрес процесса и, если он совпадает с

идентификатором одного из процессов этой машины, пошлют ответное сообщение "Я здесь", содержащее сетевой адрес машины.

Хотя эта схема и прозрачна, но широковещательные сообщения перегружают систему. Такой перегрузки можно избежать, выделив в сети специальную машину для отображения высокоуровневых символьных имен. При применении такой системы процессы адресуются с помощью символьных строк, и в программы вставляются эти строки, а не номера машин или процессов. Каждый раз перед первой попыткой связаться, процесс должен послать запрос специальному отображающему процессу, обычно называемому сервером имен, запрашивая номер машины, на которой работает процесс-получатель.

Совершенно иной подход - это использование специальной аппаратуры. Пусть процессы выбирают свои адреса случайно, а конструкция сетевых адаптеров позволяет хранить эти адреса. Теперь адреса процессов не обнаруживаются путем широковещательной передачи, а непосредственно указываются в кадрах, заменяя там адреса сетевых адаптеров.

Блокирующие и неблокирующие примитивы

Примитивы бывают блокирующими и неблокирующими, иногда они называются соответственно синхронными и асинхронными. При использовании блокирующего примитива, процесс, выдавший запрос на его выполнение, приостанавливается до полного завершения примитива. Например, вызов примитива ПОЛУЧИТЬ приостанавливает вызывающий процесс до получения сообщения.

При использовании неблокирующего примитива управление возвращается вызывающему процессу немедленно, еще до того, как требуемая работа будет выполнена. Преимуществом этой схемы является параллельное выполнение вызывающего процесса и процесса передачи сообщения. Обычно в ОС имеется один из двух видов примитивов и очень редко - оба. Однако выигрыш в производительности при использовании неблокирующих примитивов компенсируется серьезным недостатком: отправитель не может модифицировать буфер сообщения, пока сообщение не отправлено, а узнать, отправлено ли сообщение, отправитель не может. Отсюда сложности в построении программ, которые передают последовательность сообщений с помощью неблокирующих примитивов.

Имеется два возможных выхода. Первое решение - это заставить ядро копировать сообщение в свой внутренний буфер, а затем разрешить процессу продолжить выполнение. С точки зрения процесса эта схема ничем не отличается от схемы блокирующего вызова: как только процесс снова получает управление, он может повторно использовать буфер.

Второе решение заключается в прерывании процесса-отправителя после отправки сообщения, чтобы проинформировать его, что буфер снова доступен. Здесь не требуется копирование, что экономит время, но прерывание пользовательского уровня делает программирование запутанным, сложным, может привести к возникновению гонок.

Вопросом, тесно связанным с блокирующими и неблокирующими вызовами, является вопрос тайм-аутов. В системе с блокирующим вызовом ПОСЛАТЬ при отсутствии ответа вызывающий процесс может заблокироваться навсегда. Для предотвращения такой ситуации в некоторых системах вызывающий процесс может задать временной интервал, в течение которого он ждет ответ. Если за это время сообщение не поступает, вызов ПОСЛАТЬ завершается с кодом ошибки.

Буферизуемые и небуферизуемые примитивы

Примитивы, которые были описаны, являются небуферизуемыми примитивами. Это означает, что вызов ПОЛУЧИТЬ сообщает ядру машины, на которой он выполняется, адрес буфера, в который следует поместить пребывающее для него сообщение.

Эта схема работает прекрасно при условии, что получатель выполняет вызов ПОЛУЧИТЬ раньше, чем отправитель выполняет вызов ПОСЛАТЬ. Вызов ПОЛУЧИТЬ сообщает ядру машины, на которой выполняется, по какому адресу должно поступить ожидаемое сообщение, и в какую область памяти необходимо его поместить. Проблема возникает тогда, когда вызов ПОСЛАТЬ сделан раньше вызова ПОЛУЧИТЬ. Каким образом сможет узнать ядро на машине получателя, какому процессу адресовано вновь поступившее сообщение, если их несколько? И как оно узнает, куда его скопировать?

Один из вариантов - просто отказаться от сообщения, позволить отправителю взять тайм-аут и надеяться, что получатель все-таки выполнит вызов ПОЛУЧИТЬ перед повторной передачей сообщения. Этот подход не сложен в реализации, но, к сожалению, отправитель (или скорее ядро его машины) может сделать несколько таких безуспешных попыток. Еще хуже то, что после достаточно большого числа безуспешных попыток ядро отправителя может сделать неправильный вывод об аварии на машине получателя или о неправильности использованного адреса.

Второй подход к этой проблеме заключается в том, чтобы хранить хотя бы некоторое время, поступающие сообщения в ядре получателя на тот случай, что вскоре будет выполнен соответствующий вызов ПОЛУЧИТЬ. Каждый раз, когда поступает такое "неожидаемое" сообщение, включается таймер. Если заданный временной интервал истекает раньше, чем происходит соответствующий вызов ПОЛУЧИТЬ, то сообщение теряется.

Хотя этот метод и уменьшает вероятность потери сообщений, он порождает проблему хранения и управления преждевременно поступившими сообщениями. Необходимы буферы, которые следует где-то размещать, освобождать, в общем, которыми нужно управлять. Концептуально простым способом управления буферами является определение новой структуры данных, называемой почтовым ящиком.

Процесс, который заинтересован в получении сообщений, обращается к ядру с запросом о создании для него почтового ящика и сообщает адрес, по которому ему могут поступать сетевые пакеты, после чего все сообщения с данным адресом будут помещены в его почтовый ящик. Такой способ часто называют буферизуемым примитивом.

Надежные и ненадежные примитивы

Ранее подразумевалось, что когда отправитель посылает сообщение, адресат его обязательно получает. Но реально сообщения могут теряться. Предположим, что используются блокирующие примитивы. Когда отправитель посылает сообщение, то он приостанавливает свою работу до тех пор, пока сообщение не будет послано. Однако нет никаких гарантий, что после того, как он возобновит свою работу, сообщение будет доставлено адресату.

Для решения этой проблемы существует три подхода. Первый заключается в том, что система не берет на себя никаких обязательств по поводу доставки сообщений. Реализация надежного взаимодействия становится целиком заботой пользователя.

Второй подход заключается в том, что ядро принимающей машины посылает квитанцию-подтверждение ядру отправляющей машины на каждое сообщение. Посылающее

ядро разблокирует пользовательский процесс только после получения этого подтверждения. Подтверждение передается от ядра к ядру. Ни отправитель, ни получатель его не видят.

Третий подход заключается в использовании ответа в качестве подтверждения в тех системах, в которых запрос всегда сопровождается ответом. Отправитель остается заблокированным до получения ответа. Если ответа нет слишком долго, то посылающее ядро может переслать запрос специальной службе предотвращения потери сообщений.

Заключение

Современные концепции и технологии проектирования операционных систем

Требования, предъявляемые к ОС 90-х годов

Операционная система является сердцевиной сетевого программного обеспечения, она создает среду для выполнения приложений и во многом определяет, какими полезными для пользователя свойствами эти приложения будут обладать. В связи с этим рассмотрим требования, которым должна удовлетворять современная ОС.

Очевидно, что главным требованием, предъявляемым к операционной системе, является способность выполнения основных функций: эффективного управления ресурсами и обеспечения удобного интерфейса для пользователя и прикладных программ. Современная ОС, как правило, должна реализовывать мультипрограммную обработку, виртуальную память, свопинг, поддерживать многооконный интерфейс, а также выполнять многие другие, совершенно необходимые функции. Кроме этих функциональных требований к операционным системам предъявляются не менее важные рыночные требования. К этим требованиям относятся:

- *Расширяемость*. Код должен быть написан таким образом, чтобы можно было легко внести дополнения и изменения, если это потребуется, и не нарушить целостность системы.
- *Переносимость*. Код должен легко переноситься с процессора одного типа на процессор другого типа и с аппаратной платформы (которая включает наряду с типом процессора и способ организации всей аппаратуры компьютера) одного типа на аппаратную платформу другого типа.
- *Надежность и отказоустойчивость*. Система должна быть защищена как от внутренних, так и от внешних ошибок, сбоев и отказов. Ее действия должны быть всегда предсказуемыми, а приложения не должны быть в состоянии наносить вред ОС.
- *Совместимость*. ОС должна иметь средства для выполнения прикладных программ, написанных для других операционных систем. Кроме того, пользовательский интерфейс должен быть совместим с существующими системами и стандартами.
- *Безопасность*. ОС должна обладать средствами защиты ресурсов одних пользователей от других.
- *Производительность*. Система должна обладать настолько хорошим быстродействием и временем реакции, насколько это позволяет аппаратная платформа.

Рассмотрим более подробно некоторые из этих требований.

Расширяемость

В то время, как аппаратная часть компьютера устаревает за несколько лет, полезная жизнь операционных систем может измеряться десятилетиями. Примером может служить ОС

UNIX. Поэтому операционные системы всегда эволюционно изменяются со временем, и эти изменения более значимы, чем изменения аппаратных средств. Изменения ОС обычно представляют собой приобретение ею новых свойств. Например, поддержка новых устройств, таких как CD-ROM, возможность связи с сетями нового типа, поддержка многообещающих технологий, таких как графический интерфейс пользователя или объектно-ориентированное программное окружение, использование более чем одного процессора. Сохранение целостности кода, какие бы изменения не вносились в операционную систему, является главной целью разработки.

Расширяемость может достигаться за счет модульной структуры ОС, при которой программы строятся из набора отдельных модулей, взаимодействующих только через функциональный интерфейс. Новые компоненты могут быть добавлены в операционную систему модульным путем, они выполняют свою работу, используя интерфейсы, поддерживаемые существующими компонентами.

Использование объектов для представления системных ресурсов также улучшает расширяемость системы. Объекты - это абстрактные типы данных, над которыми можно производить только те действия, которые предусмотрены специальным набором объектных функций. Объекты позволяют единообразно управлять системными ресурсами. Добавление новых объектов не разрушает существующие объекты и не требует изменений существующего кода.

Прекрасные возможности для расширения предоставляет подход к структурированию ОС по типу клиент-сервер с использованием микроядерной технологии. В соответствии с этим подходом ОС строится как совокупность привилегированной управляющей программы и набора непривилегированных услуг-серверов. Основная часть ОС может оставаться неизменной в то время, как могут быть добавлены новые серверы или улучшены старые.

Средства вызова удаленных процедур (RPC) также дают возможность расширить функциональные возможности ОС. Новые программные процедуры могут быть добавлены в любую машину сети и немедленно поступить в распоряжение прикладных программ на других машинах сети.

Некоторые ОС для улучшения расширяемости поддерживают загружаемые драйверы, которые могут быть добавлены в систему во время ее работы. Новые файловые системы, устройства и сети могут поддерживаться путем написания драйвера устройства, драйвера файловой системы или транспортного драйвера и загрузки его в систему.

Переносимость

Требование переносимости кода тесно связано с расширяемостью. Расширяемость позволяет улучшать операционную систему, в то время как переносимость дает возможность перемещать всю систему на машину, базирующуюся на другом процессоре или аппаратной платформе, делая при этом по возможности небольшие изменения в коде. Хотя ОС часто описываются либо как переносимые, либо как непереносимые, переносимость - это не бинарное состояние. Вопрос не в том, может ли быть система перенесена, а в том, насколько легко можно это сделать. Написание переносимой ОС аналогично написанию любого переносимого кода - нужно следовать некоторым правилам.

Во-первых, большая часть кода должна быть написана на языке, который имеется на всех машинах, куда вы хотите переносить систему. Обычно это означает, что код должен быть написан на языке высокого уровня, предпочтительно стандартизованном, например, на языке С. Программа, написанная на ассемблере, не является переносимой, если только вы не собираетесь переносить ее на машину, обладающую командной совместимостью с вашей.

Во-вторых, следует учесть, в какое физическое окружение программа должна быть перенесена. Различная аппаратура требует различных решений при создании ОС. Например, ОС, построенная на 32-битовых адресах, не может быть перенесена на машину с 16-битовыми адресами (разве что с огромными трудностями).

В-третьих, важно минимизировать или, если возможно, исключить те части кода, которые непосредственно взаимодействуют с аппаратными средствами. Зависимость от аппаратуры может иметь много форм. Некоторые очевидные формы зависимости включают прямое манипулирование регистрами и другими аппаратными средствами.

В-четвертых, если аппаратно зависимый код не может быть полностью исключен, то он должен быть изолирован в нескольких хорошо локализуемых модулях. Аппаратно-зависимый код не должен быть распределен по всей системе. Например, можно спрятать аппаратно-зависимую структуру в программно-задаваемые данные абстрактного типа. Другие модули системы будут работать с этими данными, а не с аппаратурой, используя набор некоторых функций. Когда ОС переносится, то изменяются только эти данные и функции, которые ими манипулируют.

Для легкого переноса ОС при ее разработке должны быть соблюдены следующие требования:

- *Переносимый язык высокого уровня*. Большинство переносимых ОС написано на языке С (стандарт ANSI X3.159-1989). Разработчики выбирают С потому, что он стандартизован, и потому, что С-компиляторы широко доступны. Ассемблер используется только для тех частей системы, которые должны непосредственно взаимодействовать с аппаратурой (например, обработчик прерываний) или для частей, которые требуют максимальной скорости (например, целочисленная арифметика повышенной точности). Однако непереносимый код должен быть тщательно изолирован внутри тех компонентов, где он используется.
- *Изоляция процессора*. Некоторые низкоуровневые части ОС должны иметь доступ к процессорно-зависимым структурам данных и регистрам. Однако код, который делает это, должен содержаться в небольших модулях, которые могут быть заменены аналогичными модулями для других процессоров.
- *Изоляция платформы*. Зависимость от платформы заключается в различиях между рабочими станциями разных производителей, построенными на одном и том же процессоре (например, MIPS R4000). Должен быть введен программный уровень, абстрагирующий аппаратуру (кэши, контроллеры прерываний ввода-вывода и т. п.) вместе со слоем низкоуровневых программ таким образом, чтобы высокоуровневый код не нуждался в изменении при переносе с одной платформы на другую.

Совместимость

Одним из аспектов совместимости является способность ОС выполнять программы, написанные для других ОС или для более ранних версий данной операционной системы, а также для другой аппаратной платформы.

Необходимо разделять вопросы двоичной совместимости и совместимости на уровне исходных текстов приложений. Двоичная совместимость достигается в том случае, когда можно взять исполняемую программу и запустить ее на выполнение на другой ОС. Для этого необходимы: совместимость на уровне команд процессора, совместимость на уровне системных вызовов и даже на уровне библиотечных вызовов, если они являются динамически связываемыми.

Совместимость на уровне исходных текстов требует наличия соответствующего компилятора в составе программного обеспечения, а также совместимости на уровне библиотек и системных вызовов. При этом необходима перекомпиляция имеющихся исходных текстов в новый выполняемый модуль.

Совместимость на уровне исходных текстов важна в основном для разработчиков приложений, в распоряжении которых эти исходные тексты всегда имеются. Но для конечных пользователей практическое значение имеет только двоичная совместимость, так как только в этом случае они могут использовать один и тот же коммерческий продукт, поставляемый в виде двоичного исполняемого кода, в различных операционных средах и на различных машинах.

Обладает ли новая ОС двоичной совместимостью или совместимостью исходных текстов с существующими системами, зависит от многих факторов. Самый главный из них - архитектура процессора, на котором работает новая ОС. Если процессор, на который переносится ОС, использует тот же набор команд (возможно с некоторыми добавлениями) и тот же диапазон адресов, тогда двоичная совместимость может быть достигнута достаточно просто.

Гораздо сложнее достичь двоичной совместимости между процессорами, основанными на разных архитектурах. Для того, чтобы один компьютер выполнял программы другого (например, DOS-программу на Mac), этот компьютер должен работать с машинными командами, которые ему изначально непонятны. Например, процессор типа 680x0 на Mac должен исполнять двоичный код, предназначенный для процессора 80x86 в PC. Процессор 80x86 имеет свои собственные дешифратор команд, регистры и внутреннюю архитектуру. Процессор 680x0 не понимает двоичный код 80x86, поэтому он должен выбрать каждую команду, декодировать ее, чтобы определить, для чего она предназначена, а затем выполнить эквивалентную подпрограмму, написанную для 680x0. Так как к тому же у 680x0 нет в точности таких же регистров, флагов и внутреннего арифметико-логического устройства, как в 80x86, он должен имитировать все эти элементы с использованием своих регистров или памяти. И он должен тщательно воспроизводить результаты каждой команды, что требует специально написанных подпрограмм для 680x0, гарантирующих, что состояние эмулируемых регистров и флагов после выполнения каждой команды будет в точности таким же, как и на реальном 80x86.

Это простая, но очень медленная работа, так как микрокод внутри процессора 80x86 исполняется на значительно более быстродействующем уровне, чем эмулирующие его внешние команды 680x0. За время выполнения одной команды 80x86 на 680x0, реальный 80x86 может выполнить десятки команд. Следовательно, если процессор, производящий эмуляцию, не настолько быстр, чтобы компенсировать все потери при эмуляции, то программы, исполняющиеся под эмуляцией, будут очень медленными.

Выходом в таких случаях является использование так называемых прикладных сред. Учитывая, что основную часть программы, как правило, составляют вызовы библиотечных функций, прикладная среда имитирует библиотечные функции целиком, используя заранее написанную библиотеку функций аналогичного назначения, а остальные команды эмулирует каждую по отдельности.

Соответствие стандартам POSIX также является средством обеспечения совместимости программных и пользовательских интерфейсов. Во второй половине 80-х правительственные агентства США начали разрабатывать POSIX как стандарты на поставляемое оборудование при заключении правительственных контрактов в компьютерной области. POSIX - это "интерфейс переносимой ОС, базирующейся на UNIX". POSIX - собрание международных стандартов интерфейсов ОС в стиле UNIX. Использование стандарта POSIX (IEEE стандарт 1003.1 - 1988) позволяет создавать программы стиле UNIX, которые могут легко переноситься из одной системы в другую.

Безопасность

В дополнение к стандарту POSIX правительство США также определило требования компьютерной безопасности для приложений, используемых правительством. Многие из этих требований являются желаемыми свойствами для любой многопользовательской системы. Правила безопасности определяют такие свойства, как защита ресурсов одного пользователя от других и установление квот по ресурсам для предотвращения захвата одним пользователем всех системных ресурсов (таких как память).

Обеспечение защиты информации от несанкционированного доступа является обязательной функцией сетевых операционных систем. В большинстве популярных систем гарантируется степень безопасности данных, соответствующая уровню С2 в системе стандартов США.

Основы стандартов в области безопасности были заложены "*Критериями оценки надежных компьютерных систем*". Этот документ, изданный в США в 1983 году национальным центром компьютерной безопасности (NCSC - National Computer Security Center), часто называют Оранжевой Книгой.

В соответствии с требованиями Оранжевой книги безопасной считается такая система, которая "посредством специальных механизмов защиты контролирует доступ к информации таким образом, что только имеющие соответствующие полномочия лица или процессы, выполняющиеся от их имени, могут получить доступ на чтение, запись, создание или удаление информации".

Иерархия уровней безопасности, приведенная в Оранжевой Книге, помечает низший уровень безопасности как D, а высший - как A.

- В класс D попадают системы, оценка которых выявила их несоответствие требованиям всех других классов.
- Основными свойствами, характерными для С-систем, являются: наличие подсистемы учета событий, связанных с безопасностью, и избирательный контроль доступа. Уровень С делится на 2 подуровня: уровень С1, обеспечивающий защиту

данных от ошибок пользователей, но не от действий злоумышленников, и более строгий уровень С2. На уровне С2 должны присутствовать *средства секретного входа*, обеспечивающие идентификацию пользователей путем ввода уникального имени и пароля перед тем, как им будет разрешен доступ к системе. *Избирательный контроль доступа*, требуемый на этом уровне позволяет владельцу ресурса определить, кто имеет доступ к ресурсу и что он может с ним делать. Владелец делает это путем предоставляемых прав доступа пользователю или группе пользователей. *Средства учета и наблюдения* (*auditing*) - обеспечивают возможность обнаружить и зафиксировать важные события, связанные с безопасностью, или любые попытки создать, получить доступ или удалить системные ресурсы. *Защита памяти* - заключается в том, что память инициализируется перед тем, как повторно используется. На этом уровне система не защищена от ошибок пользователя, но поведение его может быть проконтролировано по записям в журнале, оставленным средствами наблюдения и аудитинга.

- Системы уровня В основаны на помеченных данных и распределении пользователей по категориям, то есть реализуют *мандатный контроль доступа*. Каждому пользователю присваивается рейтинг защиты, и он может получать доступ к данным только в соответствии с этим рейтингом. Этот уровень в отличие от уровня С защищает систему от ошибочного поведения пользователя.
- Уровень А является самым высоким уровнем безопасности, он требует в дополнение ко всем требованиям уровня В выполнения формального, математически обоснованного доказательства соответствия системы требованиям безопасности.

Различные коммерческие структуры (например, банки) особо выделяют необходимость учетной службы, аналогичной той, что предлагают государственные рекомендации С2. Любая деятельность, связанная с безопасностью, может быть отслежена и тем самым учтена. Это как раз то, что требует С2 и то, что обычно нужно банкам. Однако, коммерческие пользователи, как правило, не хотят расплачиваться производительностью за повышенный уровень безопасности. А-уровень безопасности занимает своими управляющими механизмами до 90% процессорного времени. Более безопасные системы не только снижают эффективность, но и существенно ограничивают число доступных прикладных пакетов, которые соответствующим образом могут выполняться в подобной системе. Например для ОС Solaris (версия UNIX) есть несколько тысяч приложений, а для ее аналога В-уровня - только сотня.

Тенденции в структурном построении ОС

Как уже отмечалось выше, для удовлетворения требований, предъявляемых к современной ОС, большое значение имеет ее структурное построение. Операционные системы прошли длительный путь развития от монолитных систем к хорошо структурированным модульным системам, способным к развитию, расширению и легкому переносу на новые платформы.

Монолитные системы

В общем случае "структура" монолитной системы представляет собой отсутствие структуры (Рис. 19). ОС написана как набор процедур, каждая из которых может вызывать другие, когда ей это нужно. При использовании этой техники каждая процедура системы имеет хорошо определенный интерфейс в терминах параметров и результатов, и каждая вольна вызвать любую другую для выполнения некоторой нужной для нее полезной работы.

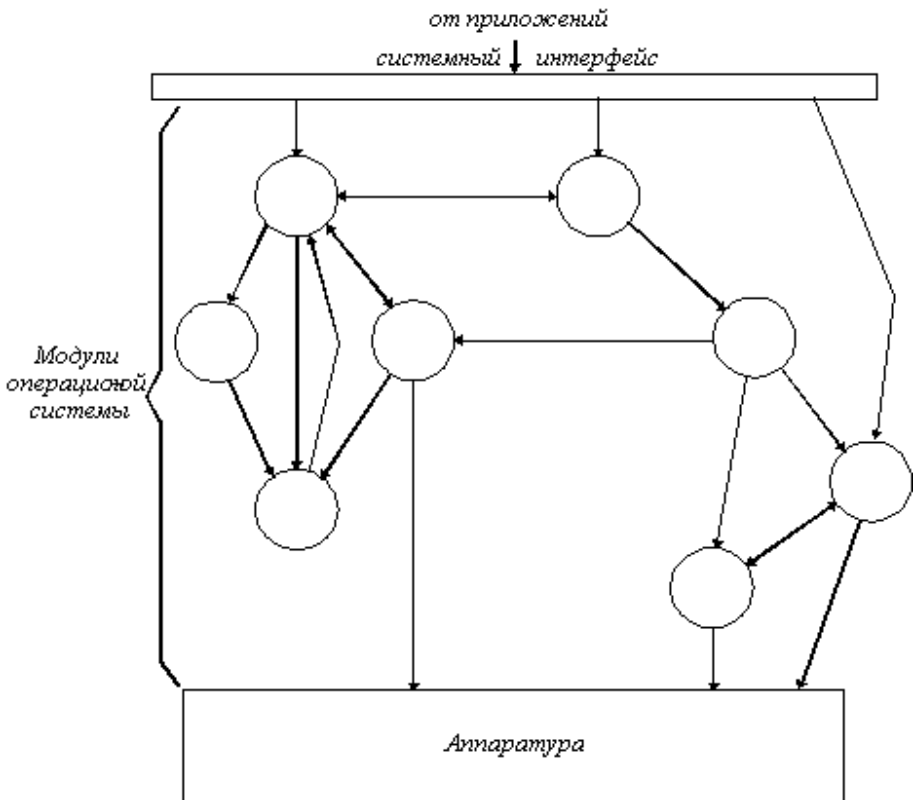


Рис. 19 Монолитная структура ОС

Для построения монолитной системы необходимо скомпилировать все отдельные процедуры, а затем связать их вместе в единый объектный файл с помощью компоновщика (примерами могут служить ранние версии ядра UNIX или Novell NetWare). Каждая процедура видит любую другую процедуру (в отличие от структуры, содержащей модули, в которой большая часть информации является локальной для модуля, и процедуры модуля можно вызвать только через специально определенные точки входа).

Однако даже такие монолитные системы могут быть немного структурированными. При обращении к системным вызовам, поддерживаемым ОС, параметры помещаются в строго определенные места, такие, как регистры или стек, а затем выполняется специальная команда прерывания, известная как вызов ядра или вызов супервизора. Эта команда переключает

машину из режима пользователя в режим ядра, называемый также режимом супервизора, и передает управление ОС. Затем ОС проверяет параметры вызова для того, чтобы определить, какой системный вызов должен быть выполнен. После этого ОС индексирует таблицу, содержащую ссылки на процедуры, и вызывает соответствующую процедуру. Такая организация ОС предполагает следующую структуру:

1. Главная программа, которая вызывает требуемые сервисные процедуры.
2. Набор сервисных процедур, реализующих системные вызовы.
3. Набор утилит, обслуживающих сервисные процедуры.

В этой модели для каждого системного вызова имеется одна сервисная процедура. Утилиты выполняют функции, которые нужны нескольким сервисным процедурам. Это деление процедур на три слоя показано на Рис. 20.

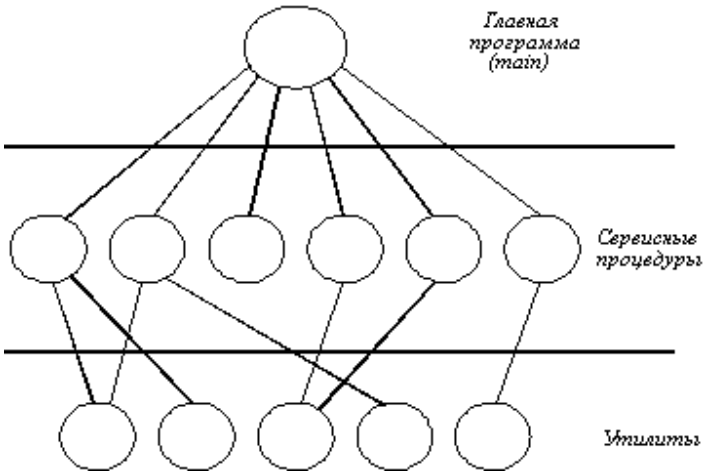


Рис. 20 Простая структуризация монолитной ОС

Многоуровневые системы

Обобщением предыдущего подхода является организация ОС как иерархии уровней. Уровни образуются группами функций операционной системы - файловая система, управление процессами и устройствами и т.п. Каждый уровень может взаимодействовать только со своим непосредственным соседом - выше- или нижележащим уровнем. Прикладные программы или модули самой операционной системы передают запросы вверх и вниз по этим уровням.

Первой системой, построенной таким образом была простая пакетная система THE, которую построил Дейкстра и его студенты в 1968 году.

Система имела 6 уровней. Уровень 0 занимался распределением времени процессора, переключая процессы по прерыванию или по истечении времени. Уровень 1 управлял памятью - распределял оперативную память и пространство на магнитном барабане для тех частей процессов (страниц), для которых не было места в ОП, то есть слой 1 выполнял функции виртуальной памяти. Слой 2 управлял связью между консолью оператора и процессами. С помощью этого уровня каждый процесс имел свою собственную консоль оператора. Уровень 3 управлял устройствами ввода-вывода и буферизовал потоки информации к ним и от них. С

помощью уровня 3 каждый процесс вместо того, чтобы работать с конкретными устройствами, с их разнообразными особенностями, обращался к абстрактным устройствам ввода-вывода, обладающим удобными для пользователя характеристиками. На уровне 4 работали пользовательские программы, которым не надо было заботиться ни о процессах, ни о памяти, ни о консоли, ни об управлении устройствами ввода-вывода. Процесс системного оператора размещался на уровне 5.

В системе THE многоуровневая схема служила, в основном, целям разработки, так как все части системы компоновались затем в общий объектный модуль.

Дальнейшее обобщение многоуровневой концепции было сделано в ОС MULTICS. В системе MULTICS каждый уровень (называемый кольцом) является более привилегированным, чем вышележащий. Когда процедура верхнего уровня хочет вызвать процедуру нижележащего, она должна выполнить соответствующий системный вызов, то есть команду TRAP (прерывание), параметры которой тщательно проверяются перед тем, как выполняется вызов. Хотя ОС в MULTICS является частью адресного пространства каждого пользовательского процесса, аппаратура обеспечивает защиту данных на уровне сегментов памяти, разрешая, например, доступ к одним сегментам только для записи, а к другим - для чтения или выполнения. Преимущество подхода MULTICS заключается в том, что он может быть расширен и на структуру пользовательских подсистем. Например, профессор может написать программу для тестирования и оценки студенческих программ и запустить эту программу на уровне n, в то время как студенческие программы будут работать на уровне n+1, так что они не смогут изменить свои оценки.

Многоуровневый подход был также использован при реализации различных вариантов ОС UNIX.

Хотя такой структурный подход на практике обычно работал неплохо, сегодня он все больше воспринимается монолитным. В системах, имеющих многоуровневую структуру было нелегко удалить один слой и заменить его другим в силу множественности и размытости интерфейсов между слоями. Добавление новых функций и изменение существующих требовало хорошего знания операционной системы и массы времени. Когда стало ясно, что операционные системы живут долго и должны иметь возможности развития и расширения, монолитный подход стал давать трещину, и на смену ему пришла модель клиент-сервер и тесно связанная с ней концепция микроядра.

Модель клиент-сервер и микроядра

Модель клиент-сервер - это еще один подход к структурированию ОС. В широком смысле модель клиент-сервер предполагает наличие программного компонента - потребителя какого-либо сервиса - клиента, и программного компонента - поставщика этого сервиса - сервера. Взаимодействие между клиентом и сервером стандартизуется, так что сервер может обслуживать клиентов, реализованных различными способами и, может быть, разными производителями. При этом главным требованием является то, чтобы они запрашивали услуги сервера понятным ему способом. Инициатором обмена обычно является клиент, который посылает запрос на обслуживание серверу, находящемуся в состоянии ожидания запроса. Один и тот же программный компонент может быть клиентом по отношению к одному виду услуг, и сервером для другого вида услуг. Модель клиент-сервер является скорее удобным концептуальным средством ясного представления функций того или иного программного

элемента в той или иной ситуации, нежели технологией. Эта модель успешно применяется не только при построении ОС, но и на всех уровнях программного обеспечения, и имеет в некоторых случаях более узкий, специфический смысл, сохраняя, естественно, при этом все свои общие черты.

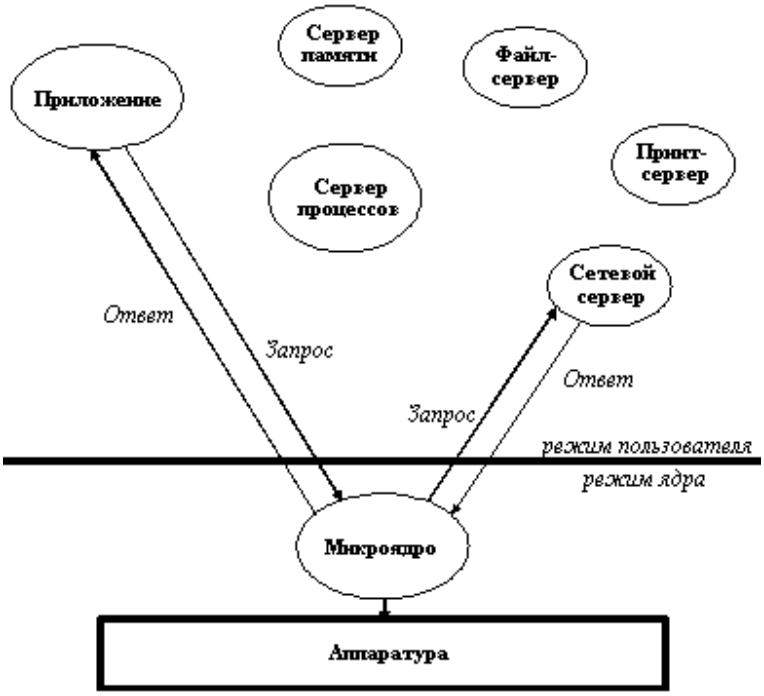


Рис. 21 Структура ОС клиент-сервер

Применительно к структурированию ОС идея состоит в разбиении ее на несколько процессов - серверов, каждый из которых выполняет отдельный набор сервисных функций - например, управление памятью, создание или планирование процессов. Каждый сервер выполняется в пользовательском режиме. Клиент, которым может быть либо другой компонент ОС, либо прикладная программа, запрашивает сервис, посылая сообщение на сервер. Ядро ОС (называемое здесь микроядром), работая в привилегированном режиме, доставляет сообщение нужному серверу, сервер выполняет операцию, после чего ядро возвращает результаты клиенту с помощью другого сообщения (Рис. 21).

Подход с использованием микроядра заменил вертикальное распределение функций операционной системы на горизонтальное. Компоненты, лежащие выше микроядра, хотя и используют сообщения, пересылаемые через микроядро, взаимодействуют друг с другом непосредственно. Микроядро играет роль регулировщика. Оно проверяет сообщения, пересылает их между серверами и клиентами, и предоставляет доступ к аппаратуре.

Данная теоретическая модель является идеализированным описанием системы клиент-сервер, в которой ядро состоит только из средств передачи сообщений. В действительности различные варианты реализации модели клиент-сервер в структуре ОС могут существенно различаться по объему работ, выполняемых в режиме ядра.

На одном краю этого спектра находится разрабатываемая фирмой IBM на основе микроядра Mach операционная система Workplace OS, придерживающаяся чистой микроядерной доктрины, состоящей в том, что все несущественные функции ОС должны выполняться не в режиме ядра, а в непривилегированном (пользовательском) режиме. На другом - Windows NT, в составе которой имеется исполняющая система (NT executive), работающая в режиме ядра и выполняющая функции обеспечения безопасности, ввода-вывода и другие.

Микроядро реализует жизненно важные функции, лежащие в основе операционной системы. Это базис для менее существенных системных служб и приложений. Именно вопрос о том, какие из системных функций считать несущественными, и, соответственно, не включать их в состав ядра, является предметом спора среди соперничающих сторонников идеи микроядра. В общем случае, подсистемы, бывшие традиционно неотъемлемыми частями операционной системы - файловые системы, управление окнами и обеспечение безопасности - становятся периферийными модулями, взаимодействующими с ядром и друг с другом.

Главный принцип разделения работы между микроядром и окружающими его модулями - включать в микроядро только те функции, которым абсолютно необходимо исполняться в режиме супервизора и в привилегированном пространстве. Под этим обычно подразумеваются машиннозависимые программы (включая поддержку нескольких процессоров), некоторые функции управления процессами, обработка прерываний, поддержка пересылки сообщений, некоторые функции управления устройствами ввода-вывода, связанные с загрузкой команд в регистры устройств. Эти функции операционной системы трудно, если не невозможно, выполнить программам, работающим в пространстве пользователя.

Есть два пути решения этой проблемы. Один путь - разместить несколько таких, чувствительных к режиму работы процессора, серверов, в пространстве ядра, что обеспечит им полный доступ к аппаратуре и , в то же время, связь с другими процессами с помощью обычного механизма сообщений. Такой подход был использован, например, при разработке Windows NT: кроме микроядра, в привилегированном режиме работает часть Windows NT, называемая executive управляющей программой. Она включает ряд компонентов, которые управляют виртуальной памятью, объектами, вводом-выводом и файловой системой (включая сетевые драйверы), взаимодействием процессов, и частично системой безопасности.

Другой путь, заключается в том, чтобы оставить в ядре только небольшую часть сервера, представляющую собой механизм реализации решения, а часть, отвечающую за принятие решения, переместить в пользовательскую область. В соответствии с этим подходом, например, в микроядре Mach, на базе которого разработана Workplace OS, размещается только часть системы управления процессами (и нитями), реализующая диспетчеризацию (то есть непосредственно переключение с процесса на процесс), а все функции, связанные с анализом приоритетов, выбором очередного процесса для активизации, принятием решения о переключении на новый процесс и другие аналогичные функции выполняются вне микроядра. Этот подход требует тесного взаимодействия между внешним планировщиком и резидентным диспетчером.

Здесь важно сделать различие - запуск процесса или нити требует доступа к аппаратуре, так что по логике - это функция ядра. Но ядру все равно, какую из нитей запускать, поэтому решения о приоритетах нитей и дисциплине постановки в очередь может принимать работающий вне ядра планировщик.

Кроме уже представленных соображений, перемещение планировщика на пользовательский уровень может понадобиться для чисто коммерческих целей. Некоторые производители ОС (например, IBM и OSF со своими вариантами микроядра Mach) планируют лицензировать свое микроядро другим поставщикам, которым может потребоваться заменить исходный планировщик на другой, поддерживающий, например, планирование в задачах реального времени или реализующий какой-то специальный алгоритм планирования. А вот другая ОС - Windows NT, также использующая микроядерную концепцию - воплотила понятие приоритетов реального времени в своем планировщике, резидентно расположенном в ядре, и это не дает возможности заменить ее планировщик на другой.

Как и управление процессами, управление памятью может распределяться между микроядром и сервером, работающим в пользовательском режиме. Например, в Workplace OS микроядро управляет аппаратурой страничной памяти. Пейджер (система управления страничной памятью), работающий вне ядра, определяет стратегию замещения страниц (т.е. решает, какие страницы следует удалить из памяти для размещения страниц, выбранных с диска в ответ на прерывание по отсутствию необходимой страницы), а микроядро выполняет перемещение выбранных пейджером страниц. Как и планировщик процессов, пейджер является заменяемой составной частью.

Драйверы устройств также могут располагаться как внутри ядра, так и вне его. При размещении драйверов устройств вне микроядра для обеспечения возможности разрешения и запрещения прерываний, часть программы драйвера должна исполняться в пространстве ядра. Отделение драйверов устройств от ядра делает возможной динамическую конфигурацию ОС. Кроме динамической конфигурации, есть и другие причины рассматривать драйверы устройств в качестве процессов пользовательского режима. СУБД, например, может иметь свой драйвер, оптимизированный под конкретный вид доступа к диску, но его нельзя будет подключить, если драйверы будут расположены в ядре. Этот подход также способствует переносимости системы, так как функции драйверов устройств могут быть во многих случаях абстрагированы от аппаратной части.

В настоящее время именно операционные системы, построенные с использованием модели клиент-сервер и концепции микроядра, в наибольшей степени удовлетворяют требованиям, предъявляемым к современным ОС.

Высокая степень переносимости обусловлена тем, что весь машинно-зависимый код изолирован в микроядре, поэтому для переноса системы на новый процессор требуется меньше изменений и все они логически сгруппированы вместе.

Технология микроядер является основой построения множественных прикладных сред, которые обеспечивают совместимость программ, написанных для разных ОС. Абстрагируя интерфейсы прикладных программ от расположенных ниже операционных систем, микроядра позволяют гарантировать, что вложения в прикладные программы не пропадут в течение нескольких лет, даже если будут сменяться операционные системы и процессоры.

Расширяемость также является одним из важных требований к современным операционным системам. Является ли операционная система маленькой, как DOS, или большой, как UNIX, для нее неизбежно настанет необходимость приобрести свойства, не заложенные в ее конструкцию. Увеличивающаяся сложность монолитных операционных систем делала трудным, если вообще возможным, внесение изменений в ОС с гарантией надежности ее последующей работы. Ограниченный набор четко определенных интерфейсов микроядра открывает путь к упорядоченному росту и эволюции ОС.

Обычно операционная система выполняется только в режиме ядра, а прикладные программы - только в режиме пользователя, за исключением тех случаев, когда они обращаются к ядру за выполнением системных функций. В отличие от обычных систем, система построенная на микроядре, выполняет свои серверные подсистемы в режиме пользователя, как обычные прикладные программы. Такая структура позволяет изменять и добавлять серверы, не влияя на целостность микроядра.

Иногда имеется потребность и в сокращении возможностей ОС. На Windows NT или UNIX перешло бы большее число пользователей, если бы для этих операционных систем не требовалось 16 Мб оперативной памяти и 70 Мб и более пространства на жестком диске. Микроядро не обязательно подразумевает небольшую систему. Надстроенные службы, типа файловой системы и системы управления окнами, добавят к ней немало. Конечно же, не всем нужна безопасность класса С2 или распределенные вычисления. Если важные, но предназначенные для определенных потребителей свойства можно исключать из состава системы, то базовый продукт подойдет более широкому кругу пользователей.

Использование модели клиент-сервер повышает надежность. Каждый сервер выполняется в виде отдельного процесса в своей собственной области памяти, и таким образом защищен от других процессов. Более того, поскольку серверы выполняются в пространстве пользователя, они не имеют непосредственного доступа к аппаратуре и не могут модифицировать память, в которой хранится управляющая программа. И если отдельный сервер может потерпеть крах, то он может быть перезапущен без останова или повреждения остальной части ОС.

Эта модель хорошо подходит для распределенных вычислений, так как отдельные серверы могут работать на разных процессорах мультипроцессорного компьютера или даже на разных компьютерах. При получении от процесса сообщения микроядро может обработать его самостоятельно или переслать другому процессу. Так как микроядру все равно, пришло ли сообщение от локального или удаленного процесса, подобная схема передачи сообщений является элегантным базисом для RPC. Однако такая гибкость не дается даром. Пересылка сообщений не так быстра, как обычные вызовы функций, и ее оптимизация является критическим фактором успеха операционной системы на основе микроядра. Windows NT, например, в некоторых случаях заменяет перенос сообщений на коммуникационные каналы с общей памятью, имеющие более высокую пропускную способность. Хотя это и стоит дороже в смысле потребления фиксированной памяти микроядра, данная альтернатива может помочь сделать модель пересылки сообщений более практичной.

Коммерческие версии микроядер

Одной из первых представила понятие микроядра фирма Next, которая использовала в своих компьютерах систему Mach, прошедшую большой путь развития в университете Карнеги-

Меллона при помощи агентства Министерства обороны США DARPA. Теоретически, ее небольшое привилегированное ядро, окруженное службами пользовательского режима, должно было обеспечить беспрецедентную гибкость и модульность. Но на практике это преимущество было несколько уменьшено наличием монолитного сервера операционной системы BSD 4.3, который выполнялся в пользовательском пространстве над микроядром Mach. Однако Mach дал Next возможность предоставить службу передачи сообщений и объектно-ориентированные средства, которые предстали перед конечными пользователями в качестве элегантного интерфейса пользователя с графической поддержкой конфигурирования сети, системного администрирования и разработки программного обеспечения.

Затем пришла Microsoft Windows NT, рекламировавшая в качестве ключевых преимуществ применения микроядра не только модульность, но и переносимость. Конструкция NT позволяет ей работать на системах на основе процессоров Intel, MIPS и Alpha (и последующих), и поддерживать симметричную многопроцессорность. Из-за того, что NT должна была выполнять программы, написанные для DOS, Windows, OS/2 и использующих соглашения POSIX, Microsoft использовала модульность, присущую микроядерному подходу для того, чтобы сделать структуру NT не повторяющей ни одну из существующих операционных систем. Вместо этого NT поддерживает каждую надстроенную операционную систему в виде отдельного модуля или подсистемы.

Более современные архитектуры микроядра были предложены Novell, USL, Open Software Foundation, IBM, Apple и другими. Одним из основных соперников NT на арене микроядер является микроядро Mach 3.0, которое и IBM и OSF взялись привести к коммерческому виду. (Next в качестве основы для NextStep пока использует Mach 2.5, но при этом внимательно присматривается к Mach 3.0.). Основной соперник Mach - микроядро Chorus 3.0 фирмы Chorus Systems, выбранный USL за основу для своих предложений. Это же микроядро будет использоваться в SpringOS фирмы Sun, объектно-ориентированном преемнике Solaris.

Сегодня стало ясно, что имеется тенденция движения от монолитных систем в сторону подхода с использованием небольших ядер. Именно такой подход уже использовался компаниями QNX Software и Unisys, в течение нескольких лет поставляющих пользующиеся успехом операционные системы на основе микроядра. QNX фирмы QNX Software обслуживает рынок систем реального времени, а CTOS фирмы Unisys популярна в области банковского дела.

Объектно-ориентированный подход

Хотя технология микроядер и заложила основы модульных систем, способных развиваться регулярным образом, она не смогла в полной мере обеспечить возможности расширения систем. В настоящее время этой цели в наибольшей степени соответствует объектно-ориентированный подход, при котором каждый программный компонент является функционально изолированным от других.

Основным понятием этого подхода является "объект". *Объект* - это единица программ и данных, взаимодействующая с другими объектам посредством приема и передачи сообщений.

Объект может быть представлением как некоторых конкретных вещей - прикладной программы или документа, так и некоторых абстракций - процесса, события.

Программы (функции) объекта определяют перечень действий, которые могут быть выполнены над данными этого объекта. Объект-клиент может обратиться к другому объекту, послав сообщение с запросом на выполнение какой-либо функции объекта-сервера.

Объекты могут описывать сущности, которые они представляют, с разной степенью детализации. Для обеспечения преемственности при переходе к более детальному описанию разработчикам предлагается механизм наследования свойств уже существующих объектов, то есть механизм, позволяющий порождать более конкретные объекты из более общих. Например, при наличии объекта "текстовый документ" разработчик может легко создать объект "текстовый документ в формате Word 6.0", добавив соответствующее свойство к базовому объекту. Механизм наследования позволяет создать иерархию объектов, в которой каждый объект более низкого уровня приобретает все свойства своего предка.

Внутренняя структура данных объекта скрыта от наблюдения. Нельзя произвольно изменять данные объекта. Для того, чтобы получить данные из объекта или поместить данные в объект, необходимо вызывать соответствующие объектные функции. Это изолирует объект от того кода, который использует его. Разработчик может обращаться к функциям других объектов, или строить новые объекты путем наследования свойств других объектов, ничего не зная о том, как они сконструированы. Это свойство называется инкапсуляцией.

Таким образом, объект предстает для внешнего мира в виде "черного ящика" с хорошо определенным интерфейсом. С точки зрения разработчика, использующего объект, пока внешняя реакция объекта остается без изменений, не имеют значения никакие изменения во внутренней реализации. Это дает возможность легко заменять одну реализацию объекта другой, например, в случае смены аппаратных средств; при этом сложное программное окружение, в котором находятся заменяемые объекты, не потребует никаких изменений.

С другой стороны, способность объектов представать в виде "черного ящика" позволяет упаковывать в них и представлять в виде объектов уже существующие приложения, ничего в них не изменяя.

Использование объектно-ориентированного подхода особенно эффективно при создании активно развивающегося программного обеспечения, например, при разработке приложений, предназначенных для выполнения на разных аппаратных платформах.

Полностью объектно-ориентированные операционные системы очень привлекательны для системных программистов, так как, используя объекты системного уровня, программисты смогут залезать вглубь операционных систем для приспособления их к своим нуждам, не нарушая целостность системы.

Но особенно большие перспективы имеет этот подход в реализации распределенных вычислительных сред. В то время, как сейчас разные пакеты, работающие в данный момент в сети, представляют собой статически связанные наборы программ, в будущем, с использованием объектно-ориентированного подхода, они могут превратиться в единую совокупность динамически связываемых объектов, где каждый объект оперативно устанавливает и разрывает связи с другими объектами для выполнения актуальных в данный

момент задач. Приложения, созданные для такой сетевой среды, основанной на объектах, могут выполняться, динамически обращаясь к множеству объектов, независимо от их местонахождения в сети и независимо от их операционной среды.

Поскольку любое объектно-ориентированное приложение представляет собой набор объектов, разработчику желательно иметь стандартные средства для управления объектами и организации их взаимодействия. При использовании и разработке объектно-ориентированных приложений в неоднородных распределенных средах, нужны также средства, упрощающие доступ к объектам сети. При возникновении запроса к какому-либо объекту распределенной среды, независимо от того, находится требуемый объект на том же компьютере или на одном из удаленных, прозрачным образом должен быть выполнен поиск объекта, передача ему сообщения, и возврат ответа. Для обеспечения прозрачного обнаружения объектов, все они должны быть снабжены ссылками, хранящимися в каталогах. Отсюда вытекает очень сложная проблема организации службы каталогов, позволяющей программистам именовать и искать объекты в сети, которая, вообще говоря, может быть разбросана по всему миру.

Однако, несмотря на упомянутые сложности и проблемы, объектно-ориентированный подход является одной из самых перспективных тенденций в конструировании программного обеспечения.

Коммерческие объектно-ориентированные средства

Объектно-ориентированный подход к построению операционных систем, придающий порядок процессу добавления модульных расширений к небольшому ядру был принят на вооружение многими известными фирмами, такими как Microsoft, Apple, IBM, Novell/USL (UNIX Systems Laboratories) и Sun Microsystems - все они развернули свои операционные системы в этом направлении. Taligent, совместное предприятие IBM и Apple, надеется опередить всех со своей от начала до конца объектно-ориентированной операционной системой. Тем временем Next поставляет Motorola- и Intel-версии NextStep, наиболее продвинутой объектно-ориентированной операционной системы из имеющихся. Хотя NextStep и не имеет объектной ориентированности сверху донизу, как это планируется в разработках Taligent, но она доступна уже сегодня.

Одним из первых применений объектных систем для большинства пользователей станут основанные на объектах прикладные программы. К объектно-ориентированным технологиям этого уровня, уже имеющимся сейчас или доступным в ближайшем будущем относятся:

- Microsoft OLE (Object Linking and Embedding - связывание и внедрение объектов),
- стандарт OpenDoc от фирм Apple, IBM, WordPerfect, Novell и Borland,
- DSOM (Distributed System Object Model - объектная модель распределенных систем) фирмы IBM,
- PDO (Portable Distributed Objects - переносимые распределенные объекты) фирмы Next,

- каркасы (frameworks) фирмы Taligent,
- архитектура CORBA объединения OMG.

Средства OLE

Для пользователей Windows объектно-ориентированный подход проявляется при работе с программами, использующими технологию OLE фирмы Microsoft. В первой версии OLE, которая дебютировала в Windows 3.1, пользователи могли вставлять объекты в документы-клиенты. Такие объекты устанавливали ссылку на данные (в случае связывания) или содержали данные (в случае внедрения) в формате, распознаваемом программой-сервером. Для запуска программы-сервера пользователи делали двойной щелчок на объекте, посредством чего передавали данные серверу для редактирования. OLE 2.0, доступная в настоящее время в качестве расширения Windows 3.1, переопределяет документ-клиент как контейнер. Когда пользователь щелкает дважды над объектом OLE 2.0, вставленным в документ-контейнер, он активизируется в том же самом месте. Представим, например, что контейнером является документ Microsoft Word 6.0, а вставленный объект представляет собой набор ячеек в формате Excel 5.0. Когда вы щелкнете дважды над объектом электронной таблицы, меню и управляющие элементы Word как по волшебству поменяются на меню Excel. В результате, пока объект электронной таблицы находится в фокусе, текстовый процессор становится электронной таблицей.

Инфраструктура, требуемая для обеспечения столь сложных взаимодействий объектов, настолько обширна, что Microsoft называет OLE 2.0 "1/3 операционной системы". Хранение объектов, например, использует docfile, который в действительности является миниатюрной файловой системой, содержащейся внутри обычного файла MS-DOS. Docfile имеет свои собственные внутренние механизмы для семантики подкаталогов, блокировок и транзакций (т.е. фиксации-отката).

Наиболее заметный недостаток OLE - отсутствие сетевой поддержки, и это будет иметь наивысший приоритет при разработке будущих версий OLE. Следующая основная итерация OLE появится в распределенной, объектной версии Windows, называемой Cairo (Каир), ожидаемой в 1995 году.

Стандарт OpenDoc

Apple, совместно с WordPerfect, Novell, Sun, Xerox, Oracle, IBM и Taligent, известными вместе как Component Integration Laboratory (Лаборатория по объединению компонентов), также занимается архитектурой объектно-ориентированных составных документов, называемой OpenDoc. Создаваемый для работы на разных платформах, этот проект значительно отстает по степени готовности от OLE 2.0.

Ключевыми технологиями OpenDoc являются механизм хранения Бенто (названный так в честь японской тарелки с отделениями для разной пищи), технология сценариев (scripting), позаимствованная в значительной степени из AppleSript, и SOM фирмы IBM. В Бенто-документе каждый объект имеет постоянный идентификатор, перемещающийся вместе с ним от

системы к системе. Хранение не только является транзакционным, как в OLE, но и позволяет держать и отслеживать многочисленные редакции каждого объекта. Если имеется несколько редакций документа, то реально храниться будут только изменения от одной редакции к другой. Верхняя граница числа сохраняемых редакций будет определяться пользователем.

Команда Apple планирует сделать OpenDoc совместимым с Microsoft OLE. Если план завершится успехом, система OpenDoc сможет окружать объекты OLE слоем программ трансляции сообщений. Контейнер OpenDoc будет видеть встроенный объект OLE как объект OpenDoc, а объект OLE будет представлять свой контейнер, как контейнер OLE. Утверждается, что будет допустима также и обратная трансляция по этому сценарию, когда объекты OpenDoc функционируют в контейнерах OLE. Слой трансляции разрабатывается WordPerfect при помощи Borland, Claris, Lotus и других.

В основе OLE и OpenDoc лежат две соперничающие объектные модели: Microsoft COM (Component Object Model - компонентная объектная модель) и IBM SOM. Каждая определяет протоколы, используемые объектами для взаимодействия друг с другом. Основное их различие заключается в том, что SOM нейтральна к языкам программирования и поддерживает наследование, тогда как COM ориентирована на C ++ и вместо механизма наследования использует альтернативный механизм, который Microsoft называет агрегацией.

Семейство CORBA

Hewlett-Packard, Sun Microsystems и DEC экспериментируют с объектами уже много лет. Теперь эти компании и много других объединились вместе, основав промышленную коалицию под названием OMG (Object Management Group), разрабатывающую стандарты для обмена объектами. OMG CORBA (Common Object Request Broker Architecture - Общая архитектура посредника обработки объектных запросов) закладывает фундамент распределенных вычислений с переносимыми объектами. CORBA задает способ поиска объектами других объектов и вызова их методов. SOM согласуется с CORBA. Если вы пользуетесь DSOM под OS/2, вы сможете вызывать CORBA-совместимые объекты, работающие на HP, Sun и других архитектурах. Означает ли это возможность редактировать объект OpenDoc, сделанный на Macintosh, в документе-контейнере на RISC-рабочей станции? Вероятно, нет. CORBA может гарантировать только механизм нижнего уровня, посредством которого одни объекты вызывают другие. Для успешного взаимодействия требуется также понимание сообщений друг друга.

Множественные прикладные среды

В то время как некоторые идеи (например, объектно-ориентированный подход) непосредственно касаются только разработчиков и лишь косвенно влияют на конечного пользователя, концепция множественных прикладных сред приносит пользователю долгожданную возможность выполнять на своей ОС программы, написанные для других операционных систем и других процессоров.

И сейчас дополнительное программное обеспечение позволяет пользователям некоторых ОС запускать чужие программы (например, Mac и UNIX позволяют выполнять программы для

DOS и Windows). Но в зарождающемся поколении операционных систем средства для выполнения чужих программ становятся стандартной частью системы. Выбор операционной системы больше не будет сильно ограничивать выбор прикладных программ. Хотя столкновение пользовательских интерфейсов программ для Mac, Windows и UNIX на одном и том же экране и заставит пользователя немного потрудиться, но все равно, множественные прикладные среды операционных систем скоро станут такими же стандартными, как мыши и меню.

Множественные прикладные среды обеспечивают совместимость данной ОС с приложениями, написанными для других ОС и процессоров, на двоичном уровне, а не на уровне исходных текстов. Для пользователя, купившего в свое время пакет (например, Lotus 1-2-3) для MS DOS, важно, чтобы он мог запускать этот полюбившийся ему пакет без каких-либо изменений и на своей новой машине, построенной, например, на RISC-процессоре, и работающей под управлением, например, Windows NT.

При реализации множественных прикладных сред разработчики сталкиваются с противоречивыми требованиями. С одной стороны, задачей каждой прикладной среды является выполнение программы по возможности так, как если бы она выполнялась на "родной" ОС. Но потребности этих программ могут входить в конфликт с конструкцией современной операционной системы. Специализированные драйверы устройств могут противоречить требованиям безопасности. Могут конфликтовать схемы управления памятью и оконные системы. Чисто экономические вопросы (например, стоимость лицензирования программ и угроза судебного преследования) также могут повлиять на дизайн чужих прикладных сред. Но самой большой потенциальной проблемой является производительность - прикладная среда должна выполнять программы с приемлемой скоростью.

Этому требованию не могут удовлетворить широко используемые ранее эмулирующие системы. Для сокращения времени на выполнение чужих программ прикладные среды используют имитацию программ на уровне библиотек. Эффективность этого подхода связана с тем, что большинство сегодняшних программ работают под управлением GUI (графических интерфейсов пользователя) типа Windows, Mac или UNIX Motif, при этом приложения тратят большую часть времени, производя некоторые хорошо предсказуемые вещи. Они непрерывно выполняют вызовы библиотек GUI для манипулирования окнами и для других связанных с GUI действий. И это то, что позволяет прикладным средам возместить время, потраченное на эмулирование команды за командой. Тщательно сделанная прикладная среда имеет в своем составе библиотеки, имитирующие внутренние библиотеки GUI, но написанные на родном коде, то есть она совместима с программным интерфейсом другой ОС. Иногда такой подход называют трансляцией для того, чтобы отличать его от более медленного процесса эмулирования кода по одной команде за раз.

Например, для Windows-программы, работающей на Mac, при интерпретировании команд 80x86 производительность может быть очень низкой. Но когда производится вызов функции открытия окна, модуль прикладной среды может переключить его на перекомпилированную для 680x0 подпрограмму открытия окна. Так как библиотекам GUI не нужно дешифрировать и имитировать каждую команду, то в частях программы, относящихся к вызовам GUI ABI (Application Binary Interface - двоичный интерфейс прикладного программирования), производительность может резко вырасти. В результате на таких участках кода скорость работы программы может достичь (а возможно, и превзойти) скорость работы на своем родном процессоре.

Сегодня в типичных программах значительная часть кода занята вызовом GUI ABI. Apple утверждает, что программы для Mac тратят до 90 процентов процессорного времени на выполнение подпрограмм из Mac toolbox, а не на уникальные для этих программ действия. SunSelect говорит, что программы для Windows тратят от 60 до 80 процентов времени на работу в ядре Windows. В результате при эмуляции программы на основе GUI потери производительности могут быть значительно меньше. SunSelect заявляет, что его новая прикладная среда Windows, WABI (Windows Application Binary Interface - двоичный интерфейс прикладных программ Windows), благодаря сильно оптимизированным библиотекам, на некоторых платформах при исполнении одних и тех же тестов может обогнать настоящий Microsoft Windows.

С позиции использования прикладных сред более предпочтительным является способ написания программ, при котором программист для выполнения некоторой функции обращается с вызовом к операционной системе, а не пытается более эффективно реализовать эквивалентную функцию самостоятельно, работая напрямую с аппаратурой. Отбить у программистов охоту "обращаться к металлу" сможет наличие в библиотеках мощных и сложных программ, к которым гораздо проще обращаться, чем писать самому.

Модульность операционных систем нового поколения позволяет намного легче реализовать поддержку множественных прикладных сред. В отличие от старых операционных систем, состоящих из одного большого блока для всех практических применений, разбитого произвольным образом на части, новые системы являются модульными, с четко определенными интерфейсами между составляющими. Это делает создание дополнительных модулей, объединяющих эмуляцию процессора и трансляцию библиотек, значительно более простым.

К усовершенствованным операционным системам, явно содержащим средства множественных прикладных сред, относятся: IBM OS/2 2.x и Workplace OS, Microsoft Windows NT, PowerOpen компании PowerOpen Association и версии UNIX от Sun Microsystems, IBM и Hewlett-Packard. Кроме того, некоторые компании переделывают свои интерфейсы пользователя в виде модулей прикладных сред, а другие поставщики предлагают продукты для эмуляции и трансляции прикладных сред, работающие в качестве прикладных программ.

Существует много разных стратегий по воплощению идеи множественных прикладных сред, и некоторые из этих стратегий диаметрально противоположны. В случае UNIX, транслятор прикладных сред обычно делается, как и другие прикладные программы, плавающим на поверхности операционной системы. В более современных операционных системах типа Windows NT или Workplace OS модули прикладной среды выполняются более тесно связанными с операционной системой, хотя и обладают по-прежнему высокой независимостью. А в OS/2 с ее более простой, слабо структурированной архитектурой средства организации прикладных сред встроены глубоко в операционную систему.

Использование множественных прикладных сред обеспечит пользователям большую свободу выбора операционных систем и более легкий доступ к более качественному программному обеспечению.

Сетевой пакет DCE фирмы OSF

Распределенные вычисления имеют дело с понятиями более высокого уровня, чем физические носители, каналы связи и методы передачи по ним сообщений. Распределенная

среда должна дать пользователям и приложениям прозрачный доступ к данным, вычислениям и другим ресурсам в гетерогенных системах, представляющих собой набор средств различных производителей. Стратегические архитектуры каждого крупного системного производителя базируются сейчас на той или иной форме распределенной вычислительной среды (DCE). Ключом к пониманию выгоды такой архитектуры является прозрачность. Пользователи не должны тратить свое время на попытки выяснить, где находится тот или иной ресурс, а разработчики не должны писать коды для своих приложений, использующие местоположение в сети. Никто не заинтересован в том, чтобы заставлять разработчиков приложений становиться гуру коммуникаций, или же в том, чтобы заставлять пользователей заботиться о монтировании удаленных томов. Кроме того, сеть должна быть управляемой. Окончательной картиной является виртуальная сеть: набор сетей рабочих групп, отделов, предприятий, объединенных сетей предприятий, которые кажутся конечному пользователю или приложению единой сетью с простым доступом.

Одной из технологий, которая будет играть большое значение для будущего распределенных вычислений является технология DCE (Distributed Computing Environment) некоммерческой организации Open Software foundation (OSF). DCE OSF - это интегрированный набор функций, независимых от операционной системы и сетевых средств, которые обеспечивают разработку, использование и управление распределенными приложениями. Из-за их способности обеспечивать управляемую, прозрачную и взаимосвязанную работу систем различных производителей и различных платформ, DCE является одной из самых важных технологий десятилетия.

Большинство из ведущих фирм-производителей ОС договорились о поставках DCE в будущих версиях своих системных и сетевых продуктов. Например, IBM, которая предлагает несколько DCE-продуктов, базирующихся на AIX, распространила их и на сектор сетей персональных компьютеров. В сентябре 1993 года IBM начала поставлять инструментальные средства для разработки DCE-приложений - DCE SDK для OS/2 и Windows, и в это же время выпустила на рынок свой первый DCE-продукт для пользователей персональных компьютеров - DCE Client for OS/2. Компания IBM достаточно далеко продвинулась в реализации спецификации DCE в своих продуктах, обогнав Microsoft, Novell и Banyan. Сейчас она продает как отдельный продукт клиентскую часть служб DCE для операционных сред MVS, OS/400, OS/2, AIX и Windows. IBM собирается реализовать отсутствующую в LAN Server единую справочную службу и новую службу безопасности в соответствии со спецификацией DCE и выпустить интегрированный вариант DCE/LAN Server в конце этого года.

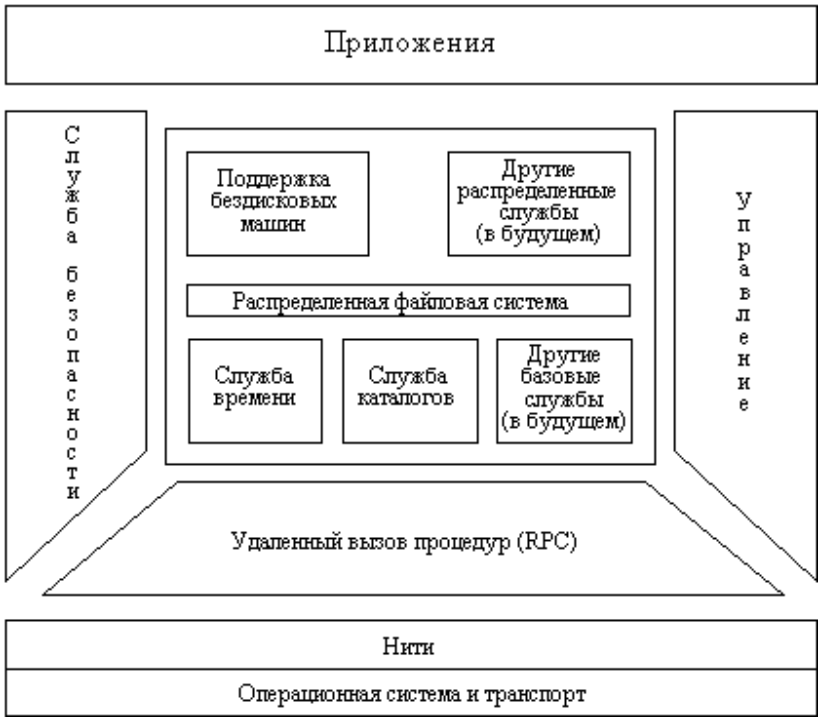


Рис. 22 Архитектура средств OSF DCE

Некоторые крупные фирмы-потребители программных продуктов обращаются непосредственно в OSF за лицензиями на первые версии DCE OSF.

В настоящее время DCE состоит из 7 средств и функций, которые делятся на базовые распределенные функции и средства разделения данных. Базовые распределенные функции включают:

1. нити,
2. RPC,
3. службу каталогов,
4. службу времени,
5. службу безопасности.

Функции разделения данных строятся над базовыми функциями и включают:

1. распределенную файловую систему (DFS),
2. поддержку бездисковых машин.

На Рис. 22 показана архитектура DCE OSF. Эта архитектура представляет собой многоуровневый набор интегрированных средств. Внизу расположены базисные службы, такие как ОС, а на самом верхнем уровне находятся потребители средств DCE приложения. Средства безопасности и управления пронизывают все уровни. OSF резервирует место для функций,

которые могут появиться в будущем, таких как спулинг, поддержка транзакций и распределенная объектно-ориентированная среда.

Нити

Обычно приложения имеют дело с процессами, каждый из которых состоит из одной нити управления.

Нити - это важная модель для выражения параллелизма внутри процесса, особенно для распределенного окружения. Например, возможности многонитевой обработки становятся особенно важными в контексте RPC. RPC является синхронным механизмом по своей природе: клиент делает вызов удаленной функции и ожидает выполнения вызова. При использовании нитей одна нить может сделать запрос, а другая начать обрабатывать данные от другого запроса. Следовательно, использование нитей может существенно улучшить производительность распределенного приложения.

Нитевая модель предъявляет меньшие требования к искусству программиста, чем другие альтернативы параллелизма, такие как асинхронные операции или разделение памяти.

Поскольку поддержка нитей дает большой выигрыш в производительности, большинство современных операционных систем является многонитевыми. Однако многие используемые в настоящее время операционные системы таковыми не являются. DCE OSF предлагает специальный пакет, предназначенный для обеспечения многонитевости для организации распределенных приложений в ОС, не имеющих собственных встроенных средств поддержки многонитевости. Этот пакет выполнен как библиотека функций работы с нитями.

По сравнению со встроенными в ядро ОС функциями поддержки нитей, библиотечная их реализация имеет некоторые функциональные ограничения. С другой стороны, в этом случае предоставляются хорошие шансы для обеспечения совместимости свойств нитей в гетерогенных средах.

Все остальные службы DCE OSF - RPC, службы безопасности, каталогов и времени, распределенная файловая система - используют сервис этого пакета.

Рассмотрим пакет, реализующий нити в DCE OSF, более подробно. Как и большинство программного обеспечения DCE, этот пакет большой и сложный. Он состоит из 51 библиотечной функции, относящейся к нитям, которыми могут пользоваться прикладные программы. Многие из них не являются необходимыми, а разработаны только для удобства. Мы рассмотрим только основные.

Эти функции удобно разделить на 7 категорий, каждая из которых имеет дело с определенным аспектом работы с нитями. Первая категория имеет дело с управлением нитями. В нее входят вызовы:

- *CREATE* - создает новую нить
- *EXIT* - вызывается нитью при завершении
- *JOIN* - подобен системному вызову WAIT в UNIX
- *DETACH* - отсоединение нити-потомка

Эти вызовы позволяют создавать и завершать нити. Родительская нить может ждать потомка, используя вызов join, подобно вызову wait в UNIX'e. Если родительская нить не планирует этого, то она может отсоединиться от потомка, сделав вызов detach. В этом случае, когда нить-потомок завершается, ее память освобождается немедленно, вместо обычного ожидания родителя, который издал вызов join.

Вторая категория функций позволяет пользователю создавать, разрушать и управлять шаблонами для нитей, для семафоров (в данном случае имеются ввиду бинарные семафоры, имеющие два состояния, называемые мьютексами). Шаблоны могут использоваться с соответствующими начальными условиями. При создании объекта один из параметров вызова create является указателем на шаблон. Например, шаблон нити имеет по умолчанию размер стека 8 К. Все нити, созданные с помощью этого шаблона имеют размер стека 8 К. Смысл использования шаблонов в том, что они позволяют не задавать все параметры создаваемого объекта.

Существуют системные вызовы добавления новых атрибутов к шаблонам. Вызовы attr_create и attr_delete создают и удаляют шаблоны нитей. Другие вызовы позволяют программе читать и модифицировать атрибуты шаблона, такие как размер стека или параметры планирования. Имеются также шаблоны для создания и удаления семафоров.

Третья группа вызовов содержит пять функций для управления мьютексами: создание, удаление, блокирование, разблокирование.

При работе с семафорами возникает очевидный вопрос, что случится, если использовать операцию "разблокировать" по отношению к мьютексу, который не заблокирован. Сохранится ли эта операция или будет потеряна? Эта проблема и вынудила Дейкстру ввести в свое время семафоры, при использовании которых порядок выполнения операций блокирования и разблокирования не имеет значения. В результате не возникает эффекта гонок. Логика работы мьютекса в DCE зависит от реализации, что, конечно, не позволяет писать переносимые программы.

Существует два типа мьютексов в DCE - быстрые и дружественные. Быстрый мьютекс является аналогом блокировки в СУБД. Если процесс пытается заблокировать незаблокированную запись, то эта операция выполняется успешно. Однако, если он попытается применить эту блокировку второй раз, то он сам заблокируется, ожидая, пока кто-нибудь не снимет блокировку с мьютекса.

Дружественный мьютекс позволяет нити заблокировать уже заблокированный мьютекс. Предположим, что главная программа блокирует семафор, а затем вызывает функцию, которая также блокирует мьютекс. Чтобы избежать клинча, принимается и вторая блокировка. Если

количество открытий и закрытий мьютекса совпадает, связанные операции блокирования-разблокирования могут иметь произвольную глубину вложения. Разработчики пакета, очевидно не пришли к согласию относительно того, что делать с повторной блокировкой от той же нити, и решили реализовать обе версии.

RPC

Хорошо известный механизм для реализации распределенных вычислений, RPC, расширяет традиционную модель программирования - вызов процедуры - для использования в сети. RPC может составлять основу распределенных вычислений. Теоретически программист не должен становиться экспертом по коммуникационным средствам для того, чтобы написать распределенное сетевое приложение. При использовании RPC программист будет применять специальный язык для описания операций. Данное описание обрабатывается компилятором, который создает код как для клиента, так и для сервера. Для обеспечения такого типа функций средства RPC должны быть простыми, прозрачными, надежными и высокопроизводительными.

Средства RPC пакета DCE OSF обладают простотой. Они приближаются к модели вызова локальных процедур насколько это возможно. Они почти не требуют изменения концептуального мышления программиста, и поэтому уменьшают время его переподготовки. Это особенно важно для коллективов программистов, работающих в фирмах, которые не являются производителями коммерческих программных продуктов.

Неизменность протокола - это другая особенность RPC DCE OSF. Этот протокол четко определен и не может изменяться пользователем (в данном случае разработчиком сетевых приложений). Такая неизменность, гарантируемая ядром, является важным свойством в гетерогенных средах, требующих согласованной работы. В отличие от OSF, некоторые другие разработчики средств RPC полагают, что гибкость и возможность приспосабливать эти средства к потребностям пользователей являются более важными.

Средства RPC DCE OSF поддерживают ряд транспортных протоколов и позволяют добавлять новые транспортные протоколы, сохраняя при этом свои функциональные свойства.

DCE RPC эффективно используется в других службах DCE: в службах безопасности, каталогов и времени, в распределенной файловой системе. DCE RPC интегрируется с системой идентификации для обеспечения защиты доступа и с нитями клиента и сервера для того, чтобы, сохраняя синхронный характер выполнения вызова, обеспечить параллелизм. Способность RPC посылать и получать потоки типизированных данных неопределенной длины используется в распределенной файловой системе.

Распределенная служба каталогов

Задачей службы каталогов в распределенной сети является поиск сетевых объектов, то есть пользователей, ресурсов, данных или приложений. Служба каталогов (или иначе, имен) должна отобразить большое количество системных объектов (пользователей, организаций, групп, компьютеров, принтеров, файлов, процессов, сервисов) на множество понятных пользователю имен. Эта проблема сложна даже для гомогенных сетей, так как персонал и оборудование перемещается, изменяет свои имена, местонахождение и т.д. В гетерогенных глобальных сетях служба каталогов становится намного более сложной, из-за необходимости синхронизации различных баз данных каталогов. Более того, при появлении в сети распределенных приложений служба каталогов должна начать отслеживание всех таких объектов и всех их компонентов.

Хорошая служба каталогов делает использование распределенного окружения прозрачным для пользователя. Пользователям не нужно знать расположение удаленного принтера, файла или приложения.

OSF определяет двухярусную архитектуру для службы каталогов для целей адресации межячеечного и глобального взаимодействия. Ячейка (cell) - это фундаментальная организационная единица для систем в DCE OSF. Ячейки могут иметь социальные, политические или организационные границы. Ячейки состоят из компьютеров, которые должны часто взаимодействовать друг с другом - это могут быть, например, рабочие группы, отделы, или отделения компаний. В общем случае компьютеры в ячейке географически близки. Размер ячеек изменяется от 2 до 1000 компьютеров, хотя OSF считает наиболее приемлемым диапазон от десятков до сотен компьютеров.

Некоторые производители и пользователи агитируют за реализацию X.500 как общей службы каталогов на всех уровнях. Но OSF полагает, что использование X.500 на уровне рабочей группы (то есть ячейки) было бы слишком громоздким из-за требований к программному обеспечению по производительности - особенно, когда более гибкие средства службы каталогов уровня ячейки уже существуют на рынке.

Служба каталогов DCE состоит из 4-х элементов:

- CDS (Cell Directory Service) - служба каталогов ячейки. Ячейка сети - это группа систем, администрируемых как единое целое. CDS оптимизируется для локального доступа. Большинство запросов к службе каталогов относятся к ресурсам той же ячейки. Каждая ячейка сети нуждается по крайней мере в одной CDS.
- GDA (Global Directory Agent) - агент глобального каталога. GDA - это шлюз имен, который соединяет домен DCE с другими административными доменами посредством глобальной службы каталогов X.500 и DNS (domain name service - сервис имен домена). GDA передает запрос на имя, которое он не смог найти в локальной ячейке, службе каталогов другой ячейки, или глобальной службе каталогов (в зависимости от места хранения имени). Для того, чтобы отыскать имя, клиент делает запрос к локальному агенту GDA. Затем GDA передает запрос на междоменное имя службе X.500. Эта служба возвращает ответ GDA, который в свою очередь передает его клиенту. OSF GDA может быть совместим с любой схемой глобального именования.

- GDS (Global Directory Service) - глобальная служба каталогов. Основанная на стандарте X.500, GDS работает на самом верхнем уровне иерархии и обеспечивает связь множества ячеек в множестве организаций.
- XDS (X/Open Directory Service) - обеспечивает поддержку X/Open API функций службы каталогов и позволяет разработчикам писать приложения, независимые от нижележащих уровней архитектуры службы каталогов. XDS-совместимые приложения будут работать одинаковым образом со службами каталогов DCE и X.500.

Распределенная служба безопасности

Имеется две больших группы функций службы безопасности: идентификация и авторизация. Идентификация проверяет идентичность объекта (например, пользователя или сервиса). Авторизация (или управление доступом) назначает привилегии объекту, такие как доступ к файлу.

Авторизация - это только часть решения. В распределенной сетевой среде обязательно должна работать глобальная служба идентификации, так как рабочей станции нельзя доверить функции идентификации себя или своего пользователя. Служба идентификации представляет собой механизм передачи третьей стороне функций проверки идентичности пользователя.

Служба безопасности OSF DCE базируется на системе идентификации Kerberos, разработанной в 80-е годы и расширенной за счет добавления элементов безопасности. Kerberos использует шифрование, основанное на личных ключах, для обеспечения трех уровней защиты. Самый нижний уровень требует, чтобы пользователь идентифицировался только при установлении начального соединения, предполагая, что дальнейшая последовательность сетевых сообщений исходит от идентифицированного пользователя. Следующий уровень требует идентификацию для каждого сетевого сообщения. На последнем уровне все сообщения не только идентифицируются, но и шифруются.

Система безопасности не должна сильно усложнять жизнь конечного пользователя в сети, то есть он не должен запоминать десятки паролей и кодов.

Весьма полезным сетевым средством для целей безопасности является служба прав доступа или, другими словами, авторизация. Служба авторизации OSF базируется на POSIX-совместимых списках прав доступа - ACL.

В то время как система Kerberos основана на личных ключах, в настоящее время широкое распространение получили методы, основанные на публичных ключах (например, метод RSA). OSF собирается сделать DCE-приложения переносимыми из Kerberos в RSA.

Распределенная файловая система DFS OSF

Распределенная файловая система DFS OSF предназначена для обеспечения прозрачного доступа к любому файлу, расположенному в любом узле сети. Главная концепция такой распределенной файловой системы - это простота ее использования.

Распределенная файловая система должна иметь единое пространство имен. Файл должен иметь одинаковое имя независимо от того, где он расположен. Другими желательными свойствами являются интегрированная безопасность, согласованность и доступность данных, надежность и восстанавливаемость, производительность и масштабируемость до очень больших конфигураций без уменьшения производительности и независимое от места расположения управление и администрирование.

Распределенная файловая система DFS OSF базируется на известной файловой системе AFS (The Andrew File System).

Файловая система AFS

AFS была разработана в университете Карнеги-Меллона и названа в честь спонсоров-основателей университета Andrew Carnegie и Andrew Mellon. Эта система, созданная для студентов университета, не является прозрачной системой, в которой все ресурсы динамически назначаются всем пользователям при возникновении потребностей. Несмотря на это, файловая система была спроектирована так, чтобы обеспечить прозрачность доступа каждому пользователю, независимо от того, какой рабочей станцией он пользуется.

Особенностью этой файловой системы является возможность работы с большим (до 10 000) числом рабочих станций.

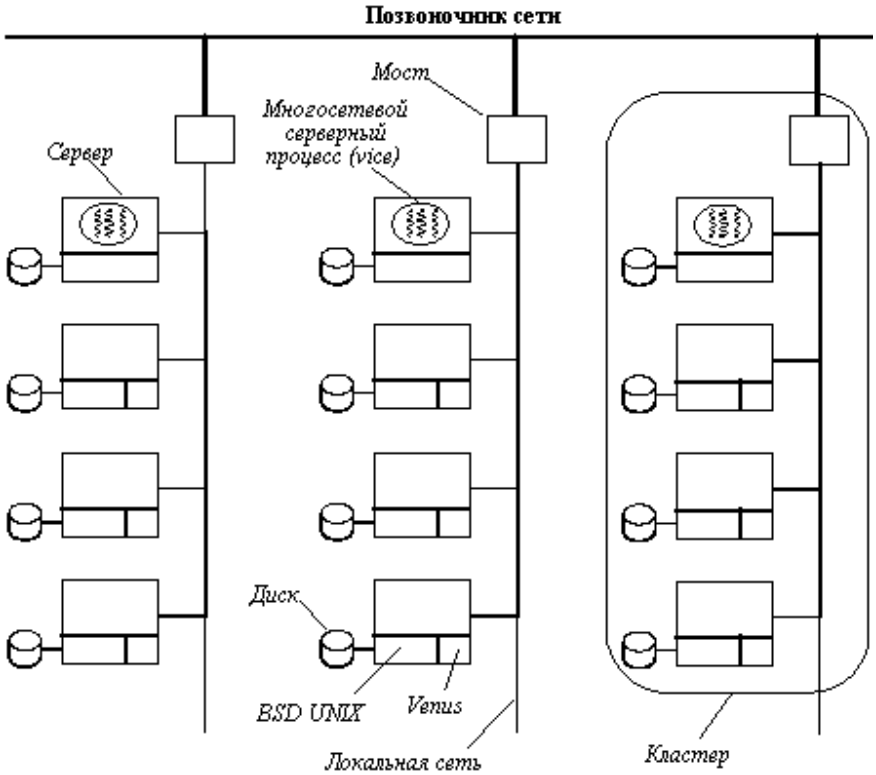


Рис. 23 Конфигурация системы, используемая AFS в университете Карнеги-Меллона

Конфигурация системы показана на Рис. 23. Она состоит из кластеров, каждый из которых включает файловый сервер и несколько десятков рабочих станций. Идея состоит в том, чтобы распределить большую часть трафика в пределах отдельных кластеров и тем самым уменьшить загрузку позвоночника сети.

Так как студенты могли входить в систему и в университете, и в общежитии, то иногда они оказывались далеко от сервера, который содержал их файлы. Несмотря на это, пользователь должен иметь возможность работать на произвольно выбранной рабочей станции, как на своем персональном компьютере.

Физически нет разницы между машиной клиента и сервера, и все они выполняют одну и ту же ОС BSD UNIX с его большим монолитным ядром. Однако над ядром выполняются совершенно различные программы серверов и клиентов. На клиент-машинах выполняются менеджеры окон, редакторы и другое стандартное программное обеспечение системы UNIX. Каждый клиент имеет также часть кода - venus, которая управляет интерфейсом между клиентом и серверной частью системы, называемой vice. Вначале venus выполнялся в пользовательском режиме, но позже он был перемещен в ядро для повышения производительности. Venus работает также в качестве менеджера кэша. В дальнейшем мы будем называть venus просто клиентом, а vice - сервером.

Пространство имен, видимое пользовательскими программами, выглядит как традиционное дерево в ОС UNIX с добавленным к нему каталогом /cmu (рисунок 4.5). Содержимое каталога /cmu поддерживается AFS посредством vice-серверов и идентично на всех рабочих станциях. Другие каталоги и файлы исключительно локальны и не разделяются.

Возможности разделяемой файловой системы предоставляются путем монтирования к каталогу /cmu. Файл, который UNIX ожидает найти в верхней части файловой системы, может быть перемещен символьной связью из разделяемой файловой системы (например, /bin/sh может быть символьно связан с /cmu/bin/sh).

В основе AFS лежит стремление делать для каждого пользователя как можно больше на его рабочей станции и как можно меньше взаимодействовать с остальной системой. При открытии удаленного файла весь файл (или его значительная часть, если он очень большой) загружается на диск рабочей станции и кэшируется там, причем процесс, который сделал вызов OPEN, даже не знает об этом. По этой причине каждая рабочая станция имеет диск.

После загрузки файла на локальный диск он помещается в локальный каталог /cash, так что он выглядит для ОС как нормальный файл. Дескриптор файла, возвращаемый системным вызовом OPEN, соответствует именно этому файлу, так что вызовы READ и WRITE работают обычным путем, без участия клиента и сервера. Другими словами, хотя системный вызов OPEN значительно изменен, реализация READ и WRITE не изменилась.

Безопасность - это главный вопрос в системе с 10 000 пользователей. Так как пользователи вольны перегружать свои рабочие станции, когда захотят и могут выполнять на них модифицированные версии ОС, то главный принцип сервера - не доверять клиентским рабочим станциям. Все сообщения между рабочими станциями шифруются на уровне аппаратуры.

Защита выполнена несколько необычным путем. Каталоги защищаются списками прав доступа (ACL), но файлы имеют обычные биты RWX UNIX'a. Разработчики системы предпочитают механизм ACL, но так как многие UNIX-программы работают с битами RWX, то они оставлены для совместимости. Списки прав доступа могут содержать и отсутствие прав, так что можно, например, потребовать, чтобы доступ к файлу был разрешен для всех, кроме одного конкретного человека.

Диски рабочих станций используются только для временных файлов, кэширования удаленных файлов и хранения страниц виртуальной памяти, но не для постоянной информации. Это существенно упрощает управление системой, в этом случае нужно управлять и архивировать только файлы серверов, а рабочие станции не требуют никаких забот. Концептуально они могут начинать каждый рабочий день с чистого листа.

AFS сконструирована для расширения до масштабов национальной файловой системы. Система, показанная на рисунке, на самом деле представляет собой отдельную ячейку (cell). Каждая ячейка - это административная единица, такая как отдел или компания. Ячейки могут быть соединены друг с другом с помощью монтирования, так что дерево разделяемых файлов может покрывать многие города.

В дополнение к концепциям файла, каталога и ячейки AFS поддерживает еще одно важное понятие - том. Том - это собрание каталогов, которые управляются вместе. Обычно все файлы какого-либо пользователя составляют том. Таким образом поддерево, входящее в /usr/john, может быть одним томом, поддерево, входящее в /usr/mary, может быть другим томом. Фактически, каждая ячейка представляет собой ничто иное, как набор томов, соединенных вместе некоторым, подходящим с точки зрения монтирования, образом. Большинство томов

содержат пользовательские файлы, некоторые другие используются для двоичных выполняемых файлов и другой системной информации. Тома могут иметь признак "только для чтения".

Семантика, предлагаемая AFS, близка к сессионной семантике. Когда файл открывается, он берется у подходящего сервера и помещается в каталог /cash на локальном диске на рабочей станции. Все операции чтения-записи работают с кэшированной копией. При закрытии файла он выгружается назад на сервер. Следствием этой модели является то, что когда процесс открывает уже открытый файл, то версия, которую он видит, зависит от того, где находится процесс. Процесс на той же рабочей станции видит копию в каталоге /cash, при этом выполняется вся семантика UNIX.

В то же время процесс на другой рабочей станции продолжает видеть исходную версию файла на сервере. Только после того, как файл будет закрыт и отослан обратно на сервер, последующая операция открытия увидит новую версию. После того, как файл закрывается, он остается в кэше, на случай, если он скоро будет снова открыт. Как мы видели ранее, повторное открытие файла, который находится в кэше, порождает проблему: "Как клиент узнает, последняя ли это версия файла?" В первой версии AFS эта проблема решалась прямым запросом клиента к серверу. К сожалению эти запросы создавали большой трафик и впоследствии алгоритм был изменен. В новом алгоритме, когда клиент загружает файл в свой кэш, то он сообщает серверу, что его интересуют все операции открытия этого файла процессами на других рабочих станциях. В этом случае сервер создает таблицу, отмечающую местонахождение этого кэшированного файла. Если другой процесс где-либо в системе открывает этот файл, то сервер посылает сообщение клиенту, чтобы тот отметил этот вход кэша как недействительный. Если этот файл в настоящее время используется, то использующие его процессы могут продолжать делать это. Однако, если другой процесс пытается открыть его заново, то клиент должен свериться с сервером, действителен ли все еще этот вход в кэше, а если нет, то получить новую копию. Если рабочая станция терпит крах, а затем перезагружается, то все файлы в кэше отмечаются как недействительные.

Блокировка файла поддерживается с помощью системного вызова UNIX FLOCK. Если блокировка не снимается в течение 30 минут, то она снимается по тайм-ауту. Тома, предназначенные только для чтения, такие как системные двоичные файлы, реплицируются, а пользовательские файлы - нет.

Хотя прикладные программы видят традиционное пространство имен UNIX, внутренняя организация сервера и клиента использует совершенно другую схему имен. Они используют двухуровневую схему именования, при которой каталог содержит структуры, называемые fids (file identifiers), вместо традиционных номеров i-узлов.

Fid состоит из трех 32-х битных полей. Первое поле - это номер тома, который однозначно определяет отдельный том в системе. Это поле говорит, на каком томе находится файл. Второе поле называется vnode, это индекс в системных таблицах определенного тома. Оно определяет конкретный файл в данном томе. Третье поле - это уникальный номер, который используется для обеспечения повторного использования vnode. Если файл удаляется, то его vnode может быть повторно использован, но с другим значением уникального номера, для того, чтобы обнаружить и отвергнуть все старые fids.

Протокол между сервером и клиентом использует fid для идентификации файла. Когда fid поступает в сервер, по значению номера тома производится поиск в базе данных,

управляемой всеми серверами, чтобы обнаружить нужный сервер. Тома могут перемещаться между серверами, но не части томов, так что эта база данных требует периодического обновления, но перемещения томов случается редко, так что трафик обновления невелик. Перемещение тома является неделимым - сначала на сервере назначения делается копия тома, а затем удаляется оригинал. Этот механизм также используется для репликации томов только для чтения за исключением того, что исходный том не удаляется после его копирования. Этот же алгоритм используется для резервного копирования. Когда делается копия, то она помещается в файловую систему как том только для чтения. В течение последующих 24 часов процесс скопирует этот том на ленту. Дополнительное преимущество этого метода - пользователь, который случайно удалил файл, все еще имеет доступ ко вчерашней копии.

Теперь рассмотрим общий механизм доступа к файлам в AFS. Когда приложение выполняет системный вызов OPEN, то он перехватывается оболочкой клиента, которая первым делом проверяет, не начинается ли имя файла с /cmu. Если нет, то файл локальный, и обрабатывается обычным способом. Если да, то файл разделяемый. Производится грамматический разбор имени, покомпонентно находится fid. По fid проверяется кэш, и здесь имеется три возможности:

1. Файл находится в кэше, и он достоверен.
2. Файл находится в кэше, и он не достоверен.
3. Файл не находится в кэше.

В первом случае используется кэшированный файл. Во втором случае клиент запрашивает сервер, изменялся ли файл после его загрузки. Файл может быть недостоверным, если рабочая станция недавно перезагружалась или же некоторый другой процесс открыл файл для записи, но это не означает, что файл уже модифицирован, и его новая копия записана на сервер. Если файл не изменялся, то используется кэшированный файл. Если он изменялся, то используется новая копия. В третьем случае файл также просто загружается с сервера. Во всех трех случаях конечным результатом будет то, что копия файла будет на локальном диске в каталоге /cash, отмеченная как достоверная.

Вызовы приложения READ и WRITE не перехватываются оболочкой клиента, они обрабатываются обычным способом. Вызовы CLOSE перехватываются оболочкой клиента, которая проверяет, был ли модифицирован файл, и, если да, то переписывает его на сервер, который управляет данным томом.

Помимо кэширования файлов, оболочка клиента также управляет кэшем, который отображает имена файлов в идентификаторы файлов fid. Это ускоряет проверку, находится ли имя в кэше. Проблема возникает, когда файл был удален и заменен другим файлом. Однако этот новый файл будет иметь другое значение поля "уникальный номер", так что fid будет выявлен как недостоверный. При этом клиент удалит вход (pass, fid) и начнет грамматический разбор имени с самого начала. Если дисковый кэш переполняется, то клиент удаляет файлы в соответствии с алгоритмом LRU.

Vice работает на каждом сервере как отдельная многонитевая программа. Каждая нить обрабатывает один запрос. Протокол между сервером и клиентом использует RPC и построен непосредственно на API. В нем есть команды для перемещения файлов в обоих направлениях, блокирования файлов, управления каталогами и некоторые другие. Vice хранит свои таблицы в виртуальной памяти, так что они могут быть произвольной величины.

Так как клиент идентифицирует файлы по их идентификаторам fid, то у сервера возникает следующая проблема: как обеспечить доступ к UNIX-файлу, зная его vnode, но не зная его полное имя. Для решения этой проблемы в AFS в UNIX добавлен новый системный вызов, позволяющий обеспечить доступ к файлам по их индексам vnode.

Реализация DFS на базе AFS дает прекрасный пример того, как работают вместе различные компоненты DCE. DFS работает на каждом узле сети совместно со службой каталогов DCE, обеспечивая единое пространство имен для всех файлов, хранящихся в DFS. DFS использует списки ACL системы безопасности DCE для управления доступом к отдельным файлам. Потоковые функции RPC позволяют DFS передавать через глобальные сети большие объемы данных за одну операцию.

Распределенная служба времени

В распределенных сетевых системах необходимо иметь службу согласования времени. Многие распределенные службы, такие как распределенная файловая система и служба идентификации, используют сравнение дат, сгенерированных на различных компьютерах. Чтобы сравнение имело смысл, пакет DCE должен обеспечивать согласованные временные отметки.

Сервер времени OSF DCE - это система, которая поставляет время другим системам в целях синхронизации. Любая система, не содержащая сервера времени, называется клерком (clerk). Распределенная служба времени использует три типа серверов для координации сетевого времени. Локальный сервер синхронизируется с другими локальными серверами той же локальной сети. Глобальный сервер доступен через расширенную локальную или глобальную сети. Курьер (courier) - это специальный локальный сервер, который периодически сверяет время с глобальными серверами. Через периодические интервалы времени серверы синхронизируются друг с другом с помощью протокола DTS OSF. Этот протокол может взаимодействовать с протоколом синхронизации времени NTP сетей Internet.

Многие фирмы-потребители программного обеспечения уже используют или собираются использовать средства DCE, поэтому ведущие фирмы-производители программного обеспечения, такие как IBM, DEC и Hewlett-Packard, заняты сейчас реализацией и поставкой различных элементов и расширений этой технологии.

Одной из главных особенностей и достоинств пакета DCE OSF является тесная взаимосвязь всех его компонентов. Это свойство пакета иногда становится его недостатком. Так, очень трудно работать в комбинированном окружении, когда одни приложения используют базис DCE, а другие - нет. В версии 1.1 совместимость служб пакета с аналогичными средствами других производителей улучшена. Например, служба Kerberos DCE в текущей версии несовместима с реализацией Kerberos MIT из-за того, что Kerberos DCE работает на базе средств RPC DCE, а Kerberos MIT - нет. OSF обещает полную совместимость с Kerberos MIT в версии 1.1. Имеются и положительные примеры совместимости пакета DCE со средствами других производителей, например со средствами Windows NT. Хотя Windows NT и не является платформой DCE, но их совместимость может быть достигнута за счет полной совместимости средств RPC. Поэтому, после достаточно тщательной работы на уровне исходных кодов, разработчики могут создать DCE-сервер, который сможет обслуживать Windows NT-клиентов, и Windows NT-сервер, который работает с DCE-клиентами.

Для того, чтобы стать действительно распространенным базисом для создания гетерогенных распределенных вычислительных сред, пакет DCE должен обеспечить поддержку двух ключевых технологий - обработку транзакций и объектно-ориентированный подход. Поддержка транзакций совершенно необходима для многих деловых приложений, когда недопустима любая потеря данных или их несогласованность. Две фирмы - IBM и Transarc - предлагают дополнительные средства, работающие над DCE и обеспечивающие обработку транзакций. Что же касается объектно-ориентированных свойств DCE, то OSF собирается снабдить этот пакет средствами, совместимыми с объектно-ориентированной архитектурой CORBA, и работающими над инфраструктурой DCE. После достаточно тщательной работы на уровне исходных кодов, разработчики могут создать DCE-сервер, который сможет обслуживать Windows NT-клиентов, и Windows NT-сервер, который работает с DCE-клиентами.

Для того, чтобы стать действительно распространенным базисом для создания гетерогенных распределенных вычислительных сред, пакет DCE должен обеспечить поддержку двух ключевых технологий - обработку транзакций и объектно-ориентированный подход. Поддержка транзакций совершенно необходима для многих деловых приложений, когда недопустима любая потеря данных или их несогласованность. Две фирмы - IBM и Transarc - предлагают дополнительные средства, работающие над DCE и обеспечивающие обработку транзакций. Что же касается объектно-ориентированных свойств DCE, то OSF собирается снабдить этот пакет средствами, совместимыми с объектно-ориентированной архитектурой CORBA и работающими над инфраструктурой DCE.

Обзор сетевых операционных систем

Большое разнообразие типов компьютеров, используемых в вычислительных сетях, влечет за собой разнообразие операционных систем: для рабочих станций, для серверов сетей уровня отдела и серверов уровня предприятия в целом. К ним могут предъявляться различные требования по производительности и функциональным возможностям, желательно, чтобы они обладали свойством совместимости, которое позволило бы обеспечить совместную работу различных ОС.

Сетевые ОС могут быть разделены на две группы: масштаба отдела и масштаба предприятия. ОС для отделов или рабочих групп обеспечивают набор сетевых сервисов, включая разделение файлов, приложений и принтеров. Они также должны обеспечивать

свойства отказоустойчивости, например, работать с RAID-массивами, поддерживать кластерные архитектуры. Сетевые ОС отделов обычно более просты в установке и управлении по сравнению с сетевыми ОС предприятия, у них меньше функциональных свойств, они меньше защищают данные и имеют более слабые возможности по взаимодействию с другими типами сетей, а также худшую производительность.

Сетевая операционная система масштаба предприятия прежде всего должна обладать основными свойствами любых корпоративных продуктов, в том числе:

- масштабируемостью, то есть способностью одинаково хорошо работать в широком диапазоне различных количественных характеристик сети,
- совместимостью с другими продуктами, то есть способностью работать в сложной гетерогенной среде интерсети в режиме plug-and-play.

Корпоративная сетевая ОС должна поддерживать более сложные сервисы. Подобно сетевой ОС рабочих групп, сетевая ОС масштаба предприятия должна позволять пользователям разделять файлы, приложения и принтеры, причем делать это для большего количества пользователей и объема данных и с более высокой производительностью. Кроме того, сетевая ОС масштаба предприятия обеспечивает возможность соединения разнородных систем - как рабочих станций, так и серверов. Например, даже если ОС работает на платформе Intel, она должна поддерживать рабочие станции UNIX, работающие на RISC-платформах. Аналогично, серверная ОС, работающая на RISC-компьютере, должна поддерживать DOS, Windows и OS/2. Сетевая ОС масштаба предприятия должна поддерживать несколько стеков протоколов (таких как TCP/IP, IPX/SPX, NetBIOS, DECnet и OSI), обеспечивая простой доступ к удаленным ресурсам, удобные процедуры управления сервисами, включая агентов для систем управления сетью.

Важным элементом сетевой ОС масштаба предприятия является централизованная справочная служба, в которой хранятся данные о пользователях и разделяемых ресурсах сети. Такая служба, называемая также службой каталогов, обеспечивает единый логический вход пользователя в сеть и предоставляет ему удобные средства просмотра всех доступных ему ресурсов. Администратор, при наличии в сети централизованной справочной службы, избавлен от необходимости заводить на каждом сервере повторяющийся список пользователей, а значит избавлен от большого количества рутинной работы и от потенциальных ошибок при определении состава пользователей и их прав на каждом сервере.

Важным свойством справочной службы является ее масштабируемость, обеспечиваемая распределенностью базы данных о пользователях и ресурсах.

Такие сетевые ОС, как Banyan Vines, Novell NetWare 4.x, IBM LAN Server, Sun NFS, Microsoft LAN Manager и Windows NT Server, могут служить в качестве операционной системы

предприятия, в то время как ОС NetWare 3.x, Personal Ware, Artisoft LANtastic больше подходят для небольших рабочих групп.

Критериями для выбора ОС масштаба предприятия являются следующие характеристики:

- Органичная поддержка многосерверной сети;
- Высокая эффективность файловых операций;
- Возможность эффективной интеграции с другими ОС;
- Наличие централизованной масштабируемой справочной службы;
- Хорошие перспективы развития;
- Эффективная работа удаленных пользователей;
- Разнообразные сервисы: файл-сервис, принт-сервис, безопасность данных и отказоустойчивость, архивирование данных, служба обмена сообщениями, разнообразные базы данных и другие;
- Разнообразные программно-аппаратные хост-платформы: IBM SNA, DEC NSA, UNIX;
- Разнообразные транспортные протоколы: TCP/IP, IPX/SPX, NetBIOS, AppleTalk;
- Поддержка многообразных операционных систем конечных пользователей: DOS, UNIX, OS/2, Mac;
- Поддержка сетевого оборудования стандартов Ethernet, Token Ring, FDDI, ARCnet;
- Наличие популярных прикладных интерфейсов и механизмов вызова удаленных процедур RPC;
- Возможность взаимодействия с системой контроля и управления сетью, поддержка стандартов управления сетью SNMP.

Конечно, ни одна из существующих сетевых ОС не отвечает в полном объеме перечисленным требованиям, поэтому выбор сетевой ОС, как правило, осуществляется с учетом производственной ситуации и опыта. Таблица 2 содержит основные характеристики популярных и доступных в настоящее время сетевых ОС.

Таблица 2 Основные характеристики сетевых операционных систем

Novell NetWare 4.1	Специализированная операционная система, оптимизированная для работы в качестве файлового сервера и принт-сервера Ограниченные средства для использования в качестве сервера приложений: не имеет средств виртуальной памяти и вытесняющей многозадачности, а поддержка симметричного мультипроцессирования отсутствовала до самого недавнего времени. Отсутствуют API основных операционных сред, используемых для разработки приложений, - UNIX, Windows, OS/2 Серверные платформы: компьютеры на основе процессоров Intel, рабочие станции RS/6000 компании IBM под управлением операционной системы

	AIX с помощью продукта NetWare for UNIX Поставляется с оболочкой для клиентов: DOS, Macintosh, OS/2, UNIX, Windows (оболочка для Windows NT разрабатывается компанией Novell в настоящее время, хотя Microsoft уже реализовала клиентскую часть NetWare в Windows NT) Организация одноранговых связей возможна с помощью ОС PersonalWare Имеет справочную службу NetWare Directory Services (NDS), поддерживающую централизованное управление, распределенную, полностью реплицируемую, автоматически синхронизируемую и обладающую отличной масштабируемостью Поставляется с мощной службой обработки сообщений Message Handling Service (MHS), полностью интегрированную (начиная с версии 4.1) со справочной службой Поддерживаемые сетевые протоколы: TCP/IP, IPX/SPX, NetBIOS, Appletalk Поддержка удаленных пользователей: ISDN, коммутируемые телефонные линии, frame relay, X.25 - с помощью продукта NetWare Connect (поставляется отдельно) Безопасность: аутентификация с помощью открытых ключей метода шифрования RSA; сертифицирована по уровню C2 Хороший сервер коммуникаций Встроенная функция компрессии диска Сложное обслуживание
Banyan VINES 6.0 и ENS (Enterprise Network Services) 6.0	Серверные платформы: ENS for UNIX: работает на RISC-компьютерах под управлением SCO UNIX, HP-UX, Solaris, AIX ENS for NetWare: работает на Intel-платформах под управлением NetWare 2.x, 3.x, 4.x VINES работает на Intel-платформах Клиентские платформы: DOS, Macintosh, OS/2, UNIX, Windows for Workgroups, Windows NT Хороший сервер приложений: поддерживаются вытесняющая многозадачность, виртуальная память и симметричное мультипроцессирование в версии VINES и в ENS-версиях для UNIX. Поддерживаются прикладные среды UNIX, OS/2, Windows Поддержка одноранговых связей - отсутствует Справочная служба - Streettalk III, наиболее отработанная из имеющихся на рынке, с централизованным управлением, полностью интегрированная с другими сетевыми службами, распределенная, реплицируемая и автоматически синхронизируемая, отлично масштабируемая Согласованность работы с другими сетевыми ОС: хорошая; серверная оболочка работает в средах NetWare и UNIX; пользователи NetWare, Windows NT и LAN Server могут быть объектами справочной службы Streettalk III Служба сообщений - Intelligent Messaging, интегрирована с другими службами Поддерживаемые сетевые протоколы: VINES IP, TCP/IP, IPX/SPX, Appletalk Поддержка удаленных пользователей: ISDN, коммутируемые телефонные линии, X.25 Служба безопасности: поддерживает электронную подпись (собственный

	алгоритм), избирательные права доступа, шифрацию; не сертифицирована Простое обслуживание Хорошо масштабируется Отличная производительность обмена данными между серверами, хуже - при обмене сервер-ПК
Microsoft LAN Manager	широкая распространенность работает под OS/2 и UNIX поддерживает мощные серверные платформы один сервер может поддерживать до 2 000 клиентов
Microsoft Windows NT Server 3.51 и 4.0	Серверные платформы: компьютеры на базе процессоров Intel, PowerPC, DEC Alpha, MIPS Клиентские платформы: DOS, OS/2, Windows, Windows for Workgroups, Macintosh Организация одноранговой сети возможна с помощью Windows NT Workstation и Windows for Workgroups Windows NT Server представляет собой отличный сервер приложений: он поддерживает вытесняющую многозадачность, виртуальную память и симметричное мультипроцессирование, а также прикладные среды DOS, Windows, OS/2, POSIX Справочные службы: доменная для управления учетной информацией пользователей (Windows NT Domain Directory service), справочные службы имен WINS и DNS Хорошая поддержка совместной работы с сетями NetWare: поставляется клиентская часть (редиректор) для сервера NetWare (версий 3.x и 4.x в режиме эмуляции 3.x, справочная служба NDS поддерживается, начиная с версии 4.0), выполненная в виде шлюза в Windows NT Server или как отдельная компонента для Windows NT Workstation; недавно Microsoft объявила о выпуске серверной части NetWare как оболочки для Windows NT Server Служба обработки сообщений - Microsoft Mail, основанная на DOS-платформе, в этом году ожидается версия для платформы Windows NT - Microsoft Message Exchange, интегрированная с остальными службами Windows NT Server Поддерживаемые сетевые протоколы: TCP/IP, IPX/SPX, NetBEUI, Appletalk Поддержка удаленных пользователей: ISDN, коммутируемые телефонные линии, frame relay, X.25 - с помощью встроенной подсистемы Remote Access Server (RAS) Служба безопасности: мощная, использует избирательные права доступа и доверительные отношения между доменами; узлы сети, основанные на Windows NT Server, сертифицированы по уровню C2 Простота установки и обслуживания Отличная масштабируемость
IBM LAN Server 4.0	Серверные платформы: операционные системы MVS и VM для мейнфреймов; AS/400 с OS/400, рабочие станции RS/6000 с AIX, серверы Intel 486 или Pentium под OS/2 Поставляется с оболочками для клиентов: DOS, Macintosh, OS/2, Windows, Windows NT, Windows for Workgroups Серверы приложений могут быть организованы с помощью LAN Server 4.0 в

	операционных средах MVS, VM, AIX, OS/2, OS/400. В среде OS/2 поддерживаются: вытесняющая многозадачность, виртуальная память и симметричное мультипроцессирование Организация одноранговых связей возможна с помощью ОС Warp Connect Справочная служба - LAN Server Domain, то есть основа на доменном подходе Поддерживаемые сетевые протоколы: TCP/IP, NetBIOS, Appletalk Безопасность - избирательные права доступа, система не сертифицирована Служба обработки сообщений - отсутствует Высокая производительность Недостаточная масштабируемость
IBM и NCR LAN Manager	LAN Manager for UNIX хорошо распространена (15% объема мировых продаж сетевых ОС) LAN Manager for AIX поддерживает RISC компьютеры System/6000 в качестве файлового сервера Работает под UNIX, имеет все преимущества, связанные с использованием этой ОС