Красавин Евгений Васильевич доцент кафедры ИУК4

Цвет учебника – виртуальный

06.09

Введение в операционные системы

ОС обеспечивает удобство посредством представления расширенной машины и рациональное управление ресурсами компьютера

Управление ресурсами включает планирование ресурса и отслеживание состояний

Эволюция ОС

1945 – 1955 – использовались реле и электронные лампы

1955 – 1965 – появление полупроводниковых элементов, произошло разделение персонала на программистов и операторов, эксплуатационников и разработчиков вычислительных машин, появились первые алгоритмические языки и компиляторы, первые системы пакетной обработки, которые просто автоматизировали запуск одной программы за другой, по итогам пакетной обработки был разработан язык управления заданиями, с помощью которой программист сообщал системе и оператору, какую работу он хочет выполнить, совокупность нескольких заданий получило название пакетозаданий, компьютеры использовались для научных и технических вычислений, появились языки программирования Fortran и Assembler и первые ОС

1965 – 1980 – переход к интегральным микросхемам, создание семейств программно совместимых машин, а программная совместимость требовала программной совместимости операционных систем, важнейшим достижением являлась реализация мультипрограммирования (одновременного выполнения нескольких программ), другое нововведение – спулинг – способ организации вычислительного процесса, в соответствии с которым задания считывались с перфокарт на диск в том темпе, в каком они появлялись в вычислительном центре, а затем после завершения очередного задания новое загружалось в освободившийся раздел, появился новый тип ОС с разделением времени

1980 – наше время – появление больших интегральных схем, персональных компьютеров, доминирование на рынке двух ОС – однопрограммной однопользовательской MS DOS и мультипрограммной мультипользовательской UNIX, – с развитием сетей бурно развивались сетевые ОС

13.09

Классификация ОС

По размеру:

* ОС мейнфреймов – в основном ориентированы на обработку множества одновременных заданий, большинству из которых требуется огромное количество операций ввода-вывода. Предлагают 3 вида обслуживания:
  + Пакетная обработка
  + Обработка транзакций
  + Разделение во времени
* Серверные ОС – работают на серверах, рабочих станциях и некоторых мейнфреймах. Одновременно обслуживают множество пользователей и позволяют делить им между собой программные и аппаратные ресурсы
* Многопроцессорные ОС
  + Мультипроцессорные ОС
    - Ассиметричные – ОС целиком выполняется только на одном процессоре, распределяя прикладные задачи по остальным
    - Симметричные – ОС полностью децентрализована, используется весь пул процессоров, разделяемых между системными и прикладными
  + Многомашинные системы и кластеры
  + Глобальные распределенные системы
* ОС для ПК – предоставляют удобный интерфейс для одного пользователя
* ОС реального времени – главным параметром является время. Существует гибкая система реального времени, в которой допустимы редкие пропуски сроков выполнения операции
* Встроенные ОС
* ОС для смарт-карт

В зависимости от выбранной архитектуры организации различают:

* Монолитные – параметры перемещаются в строго определенные места – регистры или стек, – после чего выполняется команда прерывания – вызов ядра, – которая переключает машину из режима пользователя в режим ядра и передает управление операционной системе. Структура такой ОС:
  + Главная программа, которая вызывает служебную процедуру
  + Набор служебных процедур, выполняющих системные вызовы
  + Набор утилит, обслуживающих служебные процедуры
* Многоуровневые системы
* Виртуальные машины. Экзо-ядро
* Модель клиент-сервер. Микроядро

20.09

Управление процессами

Процесс (задача) – абстракция, описывающая выполняющуюся программу

Подсистема управления процессами планирует выполнение процессов, распределяет процессорное время между несколькими одновременно существующими в системе процессами, а также занимается создание и уничтожением процессов, обеспечивает процессы необходимыми системными ресурсами и поддерживает взаимодействие между процессами

Реальный процессор переключается с процесса на процесс, но рассматривается набор процессов, идущих параллельно, то есть процессор переключается от программы к программе, и это переключение называется многозадачностью.

Состояние процессов

В многозадачной многопроцессорной системе процесс может находиться в одном из трех основных состояний:

* Выполнение – активное состояние процесса, во время которого процесс обладает всеми необходимыми ресурсами и непосредственно выполняется процессором
* Ожидание – пассивное состояние процесса, процесс заблокирован, он не может выполняться по своим внутренним причинам, он ждет осуществления некого события
* Готовность – пассивное состояние процесса, но в этом случае процесс заблокирован в связи с внешним по отношению к нему обстоятельством, то есть процесс имеет все требуемые для него ресурсы, он готов выполняться, однако процессор занят выполнением другого процесса

В состоянии выполнения в однопроцессорной системе может находиться только один процесс, чего не скажешь о процессах готовности и ожидания. Жизненный цикл процесса начинается с состояния готовности. При активизации процесс переходит в состояние выполнения и находится в нем до тех пор, пока либо сам не освободит процессор, перейдя в состояние ожидания, либо будет насильно вытеснен, возвращаясь в состояние готовности. В состояние готовности процесс переходит из состояния ожидания, когда это событие произойдет.

Создание и завершение процесса

Для того, чтобы возобновить выполнение процесса, необходимо восстановить состояние его операционной среды. Состояние операционной среды отображается состоянием регистров, программного счетчика, режимом работы процессора, указателем на открытые файлы, информацией о незавершенных операциях ввода-вывода и др. Эта информация называется контекстом процесса. Операционной системе для реализации планирования процессов требуется дополнительная информация: идентификатор процесса, состояние процесса, данные о степени привилегированности процесса и др. Такую информацию называют дескриптором процесса. Очереди процессов представляют собой дескрипторы отдельных процессов, объединенных в списки.

Создать процесс – значит:

* Создать информационный структуры, описывающие данный процесс
* Включить дескриптор нового процесса в очередь процессов
* Загрузить кодовый сегмент процесса в оперативную память или область свопинга

Фоновые процессы называются демонами. Процессы могут создаваться не только в момент загрузки, но и по просьбе текущего процесса. Процесс завершается вследствие следующих событий:

* Обычный выход, преднамеренно
* Выход по ошибке, преднамеренно
* Выход по неисправимой ошибке, непреднамеренно
* Уничтожение другим процессом, непреднамеренно

Алгоритмы планирования процессов

Планирование включает:

* Определение момента времени для смены выполняемого процесса
* Выбор процесса для выполнения из очереди готовых процессов
* Переключение контекста старого и нового процессов

Алгоритмы планирования можно разделить на 2 группы:

* Основанные на квантовании
* Основанные на приоритетах

При квантовании смена активного процесса происходит, если:

* Процесс завершился и покинул систему
* Произошла ошибка
* Процесс перешел в состояние ожидания
* Исчерпан квант процессорного времени, отведенный данному процессу

Кванты, выделяемые процессором, могут быть одинаковыми или изменяться в разные периоды жизни. Очередь готовых процессов может быть организована:

* Циклически по правилу FIFO
* Циклически по правилу LIFO

Приоритет – число, характеризующее степень привилегированности процесса при использовании ресурсов вычислительной машины. Чем выше приоритет – тем выше привилегии. Приоритет может назначаться директивно администратором системы либо операционной системой по определенному правилу. Приоритет может оставаться фиксированным на протяжении жизни процесса либо изменяться во времени в соответствии с некоторым законом и быть динамическим. Существует 2 разновидности приоритетных алгоритмов:

* Алгоритмы, использующие относительные приоритеты
* Алгоритмы, использующие абсолютные приоритеты

В системах с относительным приоритетом активный процесс выполняется до тех пора, пока сам не покинет процессор, переходя в состояние ожидания. В системах с абсолютным приоритетом выполнения активного процесса прерывается при еще одном условии – если в очереди готовых процессов появился процесс с более высоким приоритетом.

Вытесняющий и невытесняющие алгоритмы планирования

Существует два типа процедур планирования: вытесняющие и невытесняющие. Невытесняющая многозадачность – это способ планирования процессов, при котором активный процесс выполняется до тех пор, пока сам не отдаст управление планировщику ОС. Вытесняющая – способ, при котором решение о переключении процессора с выполнения одного процесса на выполнение другого принимается планировщиком ОС.

При невытесняющей многозадачности механизм планирования распределен между системой и прикладными программами. Преимущество невытесняющих систем – более высокая скорость переключения с задачи на задачу. Почти все системы написаны на вытесняющей многозадачности.

27.09

Планирование в системах реального времени

Системы реального времени делятся на жесткие системы реального времени, что означает наличие жестких сроков для каждой задачи, и гибкие системы реального времени, в которых нарушение временного графика нежелательно, но допустимо.

Системы реального времени, удовлетворяющие этому условию, называются поддающимися планированию

Событие с номером i поступает с периодом Pi и на его обработку уходит ci времени

Алгоритмы для систем реального времени могут статическими и динамическими

Статические – решения о планировании принимаются до запуска система

Динамические – решения о планировании принимаются по ходу

Статические алгоритмы заранее назначают каждому процессу фиксированный приоритет, после чего выполняют приоритетное планирование с переключением. У динамических алгоритмов нет фиксированных приоритетов.

Работа алгоритмов планирования RMS и EDF

Алгоритм планирования – планирование с приоритетом пропорциональном частоте

Алгоритм используется при следующих условиях:

1. Каждый периодический процесс должен быть завершен за время его периода
2. Ни один процесс не должен зависеть от любого другого процесса
3. Каждому процессу требуется одинаковое процессорное время на каждом интервале
4. У непериодических процессов нет жестких сроков
5. Прерывание процесса происходит мгновенно без накладных расходов

Алгоритм RMS работает, назначая каждому процессу фиксированный приоритет, равный частоте возникновения процесса

Алгоритм EDF – процесс с ближайшим сроком завершения выполняется в первую очередь

Средства синхронизации и взаимодействия процессов

Критическая секция

Критическая секция – часть программы, в которой осуществляется доступ к разделяемым данным. Чтобы исключить эффект «гонок» в отношении к некотором ресурсу, необходимо обеспечить, чтобы в каждый момент в критической секции, связанной с этим ресурсом, находился максимум один процесс. Этот прием называют взаимным исключением.

Простейший способ взаимного исключение – запрещение прерывания. Другой способ – использование блокирующих переменных. Операция проверки и установки блокирующей переменной должна быть неделимой, желательно иметь единую команду «проверка-установка» либо реализовывать программные примитивы, запрещающие прерывания на протяжении всей операции «проверки-установки».

Обобщающее средство синхронизации процессов вводит два примитива, обозначаемые P и V, оперирует целыми неотрицательными переменными, называемыми семафорами. В частном случае семафор может принимать значение только 0 или 1. Он превращается в блокируемую переменную, называемую Мьютекс.

Тупики

Существует проблема синхронизации – взаимные блокировки – называемые дедлоками, клинчами или тупиками. Проблема тупиков включает предотвращение тупиков, распознавание тупиков и восстановление системы после тупиков. Для написания корректных программ используется высокоуровневое средство синхронизации, называемое монитором. Монитор – набор процедур, переменных и структур данных. Имеет свойство: только один процесс может быть активным по отношению к монитору. Существует несколько ЯП, поддерживающих мониторы.

Потоки

Современные ОС предлагают использовать новый механизм многопоточной обработки – мультитрединг. Потоки имею собственные программный счетчик, стек, регистры состояния, разделяют адресное пространство, глобальные переменные, открытые файлы, таймеры, семафоры, статистическую информацию, дочерние процессы. Поток реализуются в пространстве пользователя или ядре.

04.10

Управление памятью

Распределению подлежит вся память, не занятая системой. ОС обычно располагается в самых младших адресах. Функциями ОС по управлению памятью является:

* Отслеживание свободной и занятой памяти
* Выделение памяти процессам и освобождение их при завершении
* Вытеснение процессов из оперативной памяти на диск и возвращение их в оперативную память
* Настройка адресов программы на конкретную область физической памяти

Типы адресов

* Символьные имена
* Виртуальные адреса
* Физические адреса



Совокупность адресов процесса называется виртуальным адресным пространством

Физические адреса соответствуют номерам ячеек оперативной памяти, где в действительности расположены или будут расположены переменные и команды

Переход от виртуальных адресов к физическим осуществляется двумя способами:

1. Специальная системная программа, перемещающая загрузчик. Он на основании исходных у него данных о начальном адресе физических данных и информации транслятора выполняет загрузку программы с заменой адресов с виртуальных на физические
2. Динамическое преобразование аппаратными средствами заключается в том, что программа загружается в неизменном виде в виртуальных адресах. ОС фиксирует смещение действительного расположение программного кода относительно виртуального адресного пространства, а во время выполнения программы при каждом обращении к виртуальной памяти выполняется преобразование виртуального адреса в физический. Отображением виртуальных адресов на физические занимается диспетчер памяти, являющийся частью процессора.

Второй способ более гибкий, так как допускает перемещение программ во время их выполнения.

Методы распределение памяти без использования дискового пространства

1. Распределение памяти фиксированными разделами. Самый простой способ управления оперативной памятью – разделение на несколько разделов фиксированной величины (вручную оператором). Очередная задача помещается либо в общую очередь, либо в очередь к какому-то разделу. Подсистема управления памятью, сравнивая размер программы и разделов, выбирает подходящий раздел, выполняет настройку программы и адресов. Достоинства: простота. Недостатки: неэффективное использование памяти, уровень мультипрограммирования ограничен числом разделов.
2. Разделение памяти разделами переменной величины. Сначала память свободна, каждой поступающей задаче выделяется необходимая ей память. После завершения задачи память освобождается и на это место может быть загружена другая задача. Задачами ОС при этом являются:
   1. ведение таблиц свободных и занятых областей, в которых указываются начальные адреса и размеры участков памяти;
   2. при поступлении новой задачи – анализ запроса, просмотр таблицы свободных областей и выбор раздела, размер которого достаточен для поступившей задачи;
   3. загрузка задачи в выделенный ей раздел, корректировка таблиц свободных и занятый областей;
   4. после завершения задачи – корректировка таблиц свободных и занятых областей

Недостаток: фрагментация памяти.

1. Перемещаемые разделы. Для борьбы с фрагментации происходит перемещение участков либо в сторону старших, либо в сторону младших адресов. Поэтому у ОС добавляется еще одна задача – периодически сжимать память. Процесс дефрагментации называется сжатием. Если процесс может иметь 2 увеличивающихся сегмента: сегмента данных и стека. Предлагается предоставлять вверху области для стека, который расширяется вниз, и сегмента данных, который расширяется вверх.

Методы распределение памяти с использованием дискового пространства

Виртуальная память – совокупность программно-аппаратных средств, позволяющих пользователям писать программы, размер которых превосходит имеющуюся оперативную память. Виртуальная память решает задачи:

* Размещает данные в запоминающих устройствах разного типа
* Перемещает при необходимости между запоминающими устройствами разного типа
* Преобразует виртуальные адреса в физические

Эти действия выполняются автоматически без участия программиста.

Страничное разделение памяти

11.10

Виртуальное адресное пространство каждого процесса делится на части одинакового фиксированного для данной системы размера, называемого виртуальными страницами. Размер виртуального адресного пространства не кратен размеру страницу, поэтому последняя страница дополняется фиктивной областью. Оперативная память делится на части, называемые физическими страницами, размер обычно выбирается равным степени двойки.

При загрузке процесса часть его виртуальных страниц помещается в оперативную память, а остальные – на диск. При загрузке ОС создает для каждого процесса таблицу страниц, в которой устанавливается соответствие между номерами виртуальных и физических страниц для страниц, загруженных в оперативную память, или делается отметка о том, что виртуальная страница выгружена на диск. Кроме того, содержится признак модификации страницы, признак невыгружаемости, признак обращения к странице и др. При каждом обращении к памяти происходит чтение из таблицы страниц информации. Если эта страница находится в оперативной памяти, выполняется преобразование виртуального адреса в физический, если выгружена на диск – происходит страничное прерывание.

Алгоритм замещения страниц

* NRU – очень грубый – делит страницы на 4 класса в зависимости от состояния битов R и M обращения и модификации (0 – не было обращений и изменений, 1 – не было обращений, страница изменена, 2 – было обращение, страницы не изменена, 3 – произошло обращение и изменение). Для вытеснения выбирается страница с наименьшим номером.
* FIFO – может выгрузить важные страницы – отслеживает порядок загрузки страниц в память, храня их в связанном списке
* Вторая попытка – значительное усовершенствование FIFO – аналогичен FIFO, но проверяет, используется ли страница или нет
* Часы – реалистичный – модификация второй попытки, но более быстродействующий
* LRU – отличный алгоритм, но его можно осуществить целиком – страница, не использовавшаяся дольше всего, требует специального аппаратного обеспечения, поддерживающего связанный список страниц, содержащихся в памяти, где последняя использованная страница находится в начале списка, а дольше всех неиспользуемая – в конце, список должен обновляться при каждом обращении к памяти
* NFU – довольно грубое приближение алгоритма NRU – редко использовавшиеся страницы, модификация LRU, создается программный счетчик, связанный с каждой страницей в памяти и изначально равный нулю, во время каждого прерывания по таймеру исследуются страницы в памяти, бит R обращения прибавляется к счетчику
* Старение – эффективный алгоритм, хорошо аппроксимирующий алгоритм LRU
* WSClock – хороший рациональный алгоритм
* Рабочий набор – немного дорог для реализации

Два последних алгоритма используют рабочий набор, многие системы со страничной организацией пытаются отслеживать рабочий набор каждого процесса и обеспечивают его нахождения в памяти до запуска процесса. Загрузка страниц перед тем, как разрешить процессу работу, называется опережающей подкачкой страниц.

Механизм преобразования виртуального адреса в физический при страничной организации памяти

При каждом обращении к оперативной памяти аппаратными средствами выполняется:

* На основании начального адреса таблицы страниц (регистр), номеру виртуальной страницы (старшие разряды виртуального адреса) и длины записи в таблице страниц (системная константа) определяется адрес нужной записи в таблице
* Из этой записи извлекается номер физической таблицы
* К номеру физической таблицы присоединяется смещение

Простейший алгоритм распределения страничного пространства на диске заключается в поддержке специальной области подкачки – свопинга – на диске. При загрузке системы эта область является пустой, имеет единую запись с начальным адресом и размером. Когда запускается первый процесс, резервируется участок области подкачки размером с этот процесс. Новым процессам предоставляются участки области подкачки, равные по размерам их областям памяти. Когда процессы завершаются, дисковое пространство освобождается. Область подкачки управляется как список свободных участков. Вычисление адреса для записи страницы: прибавляется смещение страницы внутри виртуального адреса пространства к области начала подкачки. При запуске процесса 2 способа:

1. Копируется полный образ процесса в область подкачки и при необходимости переносится в память
2. Загружается весь процесс в память и постранично выгружается на диск, когда это требуется

Сегментное распределение

Виртуальное адресное пространство процесса делится на сегменты, размер которых определяется программистом с учетом смыслового значения хранящейся в ней информации. При загрузке процесса часть сегментов помещается в оперативную память, часть размещается в дисковой памяти. Система с сегментной организацией функционирует аналогично системе со страничной организацией.

Сегментно-страничное распределение

Виртуальное адресное пространство процесса делится на сегменты, а каждый сегмент – на виртуальные страницы, которые нумеруются в пределах сегмента. Загрузка процесса выполняется постранично, для каждого сегмента создается своя таблица страниц.

Свопинг

При свопинге весь процесс перемещается между процессором и диском целиком.

Иерархия запоминающих устройств, принцип кэширвоания

Кэш-память – это способ организации совместного функционирования двух типов запоминающий устройств. Алгоритм запроса с кэш-памятью:

1. Рассматривается содержимое кэш-памяти
2. Если данные обнаруживаются в кэш-памяти, они считываются из нее
3. Если данных нет, они копируются в кэш-память, а затем передаются в процессор

Объективные свойства:

1. Пространственная локальность – если было обращение по некоему адресу, то велика вероятность обращения к соседнему адресу
2. Временная локальность – если произошло обращение по некоему адресу, то велика вероятность обращения по тому же адресу

Для ускорения процесса преобразования виртуального адреса в физический используется аппаратное решение – таблица TLB.

18.10

Средства аппаратной поддержки управления памятью

Средства поддержки сегментации памяти

Физическое адресное пространство х86 процессора составляет для 32-раряздного процессора 4 Гб, что определяется 32-разрядной шиной адреса.

Физическая память является линейной. Виртуальный адрес, используемый в программе, представляет собой пару: номер сегмента и смещение внутри сегмента. Смещение хранится в поле команды, а номер сегмента – в одной из 6 сегментных регистров процессора, каждый из которых является 16-битным. Средства сегментации образуют верхний уровень управления памятью, а средства страничной организации – нижний уровень. Средства страничной организации могут быть как включены, так и выключены за счет установки определенного бита управляющим регистром процессора.

Количество сегментов определяется размером поля, отведенном в сегментном регистре под номер сегмента. Селектор (структура данных в сегментном регистре) предназначен для выбора дескриптора определенного сегмента из таблиц дескрипторов сегмента. Дескриптор сегмента описывает все характеристики сегмента, необходимые для проверки правильности доступа к нему и нахождении его в физическом адресном пространстве. Процессор х86 поддерживает 2 таблицы дескрипторов сегмента: глобальную и локальную. Глобальная предназначена для описания сегментов операционной системы и сегментов межзадачного взаимодействия, а локальная – бля остальных задач. Глобальная таблица одна, локальных – столько, сколько в системе выполняется задач, при этом активной в каждый момент времени может быть только одна локальная таблица.

Селектор состоит из 3 полей: 13 бит – номер сегмента в таблицах глобальной и локальной, 1 бит – указатель типа используемой таблицы дескрипторов, 2 бит – текущие права доступа задачи. Разрядность поля индекса определяет максимальное число глобальных и локальных сегментов по 8 Кб (всего 16). С учетом максимального размера сегмента 4 Гб каждая задача при чистой сегментной организации виртуальной памяти работает в виртуальном адресном пространстве 64 Тб.

Когда задаче необходимо получить доступ к ячейке физической памяти, то для выбора дескриптора виртуального сегмента используется значение селектора из соответствующего сегментного регистра процессора. Значение поля типа таблицы указывается на то, какую таблицу использовать. Для хранения этих таблиц используется оперативная память. Для того, чтобы процессор смог найти в физической памяти глобальную таблицу, ее полный физический адрес начала таблицы (32 бита), также размер (16 бит) хранится в специальном регистре процессора.

Каждый дескриптор GDT и LDT имеет размер 8 байт, поэтому максимальный размер этих страниц – 64 Кб (8к дескрипторов по 8 байт). Для извлечения нужного дескриптора из таблицы процессора складывает базовый адрес таблицы GDT со сдвинутым на 3 разряда влево (умножение на 8) значением поля индекса из сегментного регистра и получает физический линейный адрес нужного дескриптора физической памяти. Таблица GT постоянно присутствует в физической памяти, поэтому процессор извлекает по этому адресу нужный дескриптор сегмента и помещает его во внутренний программно недоступный регистр процессора.

Процессор сначала определяет правильность адреса, сравнивая смещение и размер сегмента. В случае выхода за границы сегмента происходит прерывание типа исключение. Потом процессор проверяет права доступа задачи к данному сегменту, сравнивая значения полей CPL селектора и DPL дескриптора. Самый высокий уровень – нулевой – используется ядром. В случае, когда селектор указывает на LDT, виртуальный адрес преобразуется в физический аналогичным образом, но для доступа к самой LDT добавляется дополнительный этап: в процессоре регистр указывает на размещение таблицы LDT не прямо, а косвенно. Сначала по значению селектора определяется физический адрес дескриптора из таблицы GDT, описывающего начало расположения таблицы в физической памяти, а затем с помощью селектора задачи определяется физический адрес нужного дескриптора. Далее процесс аналогичен таблице GDT.

25.10

Дескриптор сегмента содержит поля:

* Поле G – однобитное поле, определяет единицу измерения сегмента (0 – в байтах, 1 – в страницах по 4 Кб)
* Поле D – определяет тип адресации сегмента (0 – 16-битный, 1 – 32-разрядный)
* Поле S – определяет, является ли системным (1) или пользовательским (0). Пользовательские делятся на сегменты данных (E = 0) и кода (E = 1). Для сегмента данных определяются однобитные поля.
  + ED – направление распространения сегмента (0 – в сторону увеличения адресов, 1 – стек, вниз).
  + W – Поле разрешения записи в сегментах (1 – разрешена, 0 – запрещена).
  + A – поле доступа к сегменту.

Сегмент кода:

* + А – поле доступа к сегменту,
  + R – разрешение на чтение для кодового сегмента,
  + C – бит подчинения, разрешает доступ из сегмента с более низким приоритетом

Для использования чисто сегментного механизма процессора операционной системе необходимо сформировать таблицы GDT и LDT, загрузить их в память, загрузить указатели на них в регистры и выключить страничную поддержку.

Сегментно-страничный механизм

При включенной системе управления страницами работает как сегментный механизм, так и механизм управления страницами. В этом случае виртуальное адресное пространство имеет 4 Гб, в котором размещаются все сегменты. Селектор задачи определяет номер виртуального сегмента, а смещение в команде задачи – смещение внутри этого сегмента. Все сегменты разделяют одно адресное пространство. Задачи исключения, наложения сегментов решает ОС.

Первый этап преобразования виртуального адреса аналогичен сегментному. Структуры данных и таблиц те же, но дескриптор содержит базовый адрес сегмента в виртуальном адресном пространстве. В результате его сложения со смещением получается линейный виртуальный адрес, который на страничном этапе преобразуется в номер физической страницы. Для реализации механизма управления страницами физическое и виртуальное адресное пространство разбиты на страницы по 4К (всего 1 Н страниц). Нумерация страниц сквозная.

Линейный виртуальный адрес содержит в своих старших 20 разрядах номер виртуальной страницы, а в младших 20 – смещение внутри страницы. Для отображения виртуальной страницы в физическую достаточно построить таблицу страниц, каждый элемент которой – дескриптор виртуальной страницы – содержал бы номер физической страницы и ее атрибуты.

* P – бит присутствия страницы в физической памяти
* W – бит разрешения записи в страницу
* U – бит пользователя, супервайзер
* A – признак, был ли доступ к странице
* D – признак, была ли модификация
* PCD и PWT – управляют механизмом кэширования страниц, разрешают попадание страницы в кэш при записи и чтении

При небольшом размере страницы таблица страниц занимает в памяти место около 4 Мб. Поэтому используется деление всей таблицы страниц на разделы по 1024 дескриптора (чтобы раздел занимал одну физическую страницу). Всего получается 1024 раздела. Для того, чтобы не хранить все разделы страниц одновременно в физической памяти, используется каталог разделов таблиц страниц, который имеет аналогичную структуру. Для его хранения необходима одна физическая страница. Совокупность дескрипторов, описывающих состояние и характеристики виртуальных страниц разделов таблиц и страниц, называется каталогом разделов или таблиц.

Виртуальная страница, хранящая содержимое каталога, всегда находится в физической памяти, и номер ее физической страницы указан в специальном управляющем регистре CR3 процессора. Преобразование виртуального адреса в физический: поле номера виртуальной страницы (старшие 20 разрядов) делится на равные части по 10 разрядов (поле номера раздела и поле номера страницы в разделе), с помощью номера физической страницы, хранящей каталог и смещение в этой странице, задаваемого поля номера раздела, процессор находит дескриптор виртуальной станицы раздела. В соответствии с атрибутами этого дескриптора определяются права доступа к этой странице, а также ее наличие в физической памяти. В случае отсутствия происходит страничное прерывание, и ОС должна загрузить ее в память. Далее для определения адреса элемента данных используется смещение, определяемое полем номера страницы линейного виртуального адреса.

Двухуровневая система адресации замедляет преобразование, но уменьшает объем памяти. Организация виртуальной памяти позволяет защитить адресное пространство различных процессов за счет: изоляции адресных пространств процесса физической памяти, назначением им различных страниц и сегментов, и защита сегментов с помощью привилегий 4 уровней.

Средства вызова подпрограмм и задач

Вызов подпрограммы отличается от запуска задачи тем, что адресное пространство остается тем же, а при вызове задачи меняется полностью. Для вызова подпрограммы, код которой находится в другом сегменте, используется защита с помощью прав доступа и 2 варианта вызова. Непосредственное указание в поле команды селектора сегмента, содержащего код вызываемой программы и смещение в этом сегменте адреса начала подпрограммы. Этот способ непригоден для вызова функций операционной системы, имеющей нулевой уровень прав. Поэтому используется набор точек входа в привилегированные кодовые сегменты, описываемые с помощью дескрипторов шлюзов подпрограмм.

Возникает проблема передачи параметров между различными стеками. Селекторы этих сегментов хранятся в контексте задачи сегменте TSS. Структура этого сегмента достаточно сложная. Контекст задачи должен содержать все данные для того, чтобы можно было восстановить выполнение прерванной в любой момент задачи (то есть значение регистров процессора, указатели на открытые файлы и другие зависящие от ОС переменные).

01.11

Особенности 64-разрядной архитектуры

Наиболее известными и распространенными в настоящее время являются 2 архитектуры микропроцессоров: IA-64 и Intel64. Intel64 представляет собой расширение архитектуры x86 с полной обратной совместимостью. Процессоры широко применяются в ПК.

Архитектура Intel64 (AMD64)

Архитектура обеспечивает поддержку устаревшего 16- и 32-разрядного кода приложений в ОС без их модификаций и перекомпиляций.

Достоинства:

* 64-битное адресное пространство
* Расширенный набор регистров
* Привычный для разработчиков набор команд
* Возможность запуска старых 32-битных приложений в 64-битной ОС

64-битные ОС

Адресное пространство

Размер страниц составляет 4 Кб. Первые 64 Кб адресного пространства никуда не отображаются. Системные библиотеки загружаются выше 4 Кб. Параметры передаются через регистры, что повышает производительность.

Трехуровневая иерархия кэша

TLB

При каждом доступе к памяти виртуальный адрес надо переводить в физический. Для этого используется таблица. В связи с ее большим объемом она размещается в основной памяти и сбрасывается на жесткий диск. Обращение к диску слишком медленно, поэтому вернулись к принципу физической адресации, добавив небольшую кэш-память в процессоре, которая хранит соответствие нескольких недавно запрошенным адресов. Кэш-память называется TLB.

08.11

Управление вводом-выводом

ОС должна передавать устройству команды, обрабатывать прерывания и ошибки, а также обеспечивать интерфейс между устройствами и остальной частью системы.

Устройства ввода-вывода делятся на 2 типа: блок- и байт-ориентированное устройства. Блок-ориентированные хранят информацию в блоках фиксированного размера. Байт-ориентированные устройства не адресуемые, не позволяют проводить операцию поиска, генерируют или потребляют последовательность байт. Некоторые внешние устройства не относятся ни к одному из этих классов.

Устройства ВВ обычно состоят из механической и электронной частей (контроллер или адаптер). Механический компонент находится в самом устройстве. Работа контроллера заключается в конвертировании последовательного потока битов в блок байтов и выполнение коррекции ошибок, если это необходимо. Обычно байтовый блок собирается бит за битом в буфере контроллера, затем проверяется контрольная сумма блока, и если она совпадает с указанной в заголовке сектором, то блок объявляется считанным без ошибок, после чего копируется в оперативную память.

Отображаемый на адресное пространство вывод-ввод

У каждого контроллера есть несколько регистров для коммуникации с центральным процессором. Помимо управляющих регистров у многих устройств есть буфер данных, из которого ОС может читать данные и писать в него. Существует 2 альтернативных способа реализации доступа к управляющим регистрам и буферам данных устройств ВВ.

1. Каждому управляющему устройству назначается номер порта ВВ. При такой схеме адресации пространства оперативной памяти и устройств ВВ не пересекаются.
2. Отображение всех управляющих регистров периферийных устройств на адресное пространство памяти. Такая схема называется отображаемым на адресное пространство памяти вводом-выводом. Обычно адреса регистров в памяти отводятся вверху адресного пространства.

Также существуют гибридные схемы с отображаемым на адресное пространство памяти буфера данных и отдельными портами ввода-вывода.

Достоинства второго варианта:

1. Не требуются специальные команды процессора
2. Не требуется специального механизма защиты от пользовательских процессов, пытающихся обращаться к внешним устройствам, каждая команда процессора, обращаясь к памяти, аналогично обращается к устройству

Недостатки первого варианта:

1. Управляющий регистр устройства должен быть сначала считан в регистр процессора, а затем произведены действия

Недостатки второго варианта:

1. В большинстве компьютеров используется кэширование памяти. Кэширование управляющих регистров приводит к ошибкам.
2. В едином адресном пространстве все модули памяти и УВВ должны изучать все обращения процессора к памяти, чтобы определить, на какие им реагировать. Многие современные компьютеры имеют выделенную скоростную шины обращения к памяти. Сложность применения выделенной шины памяти в машинах в том, что УВВ не видит эту шину и не реагирует на запрос. Поэтому сначала все обращения к памяти посылаются процессором по выделенной быстрой шине. Если память не отвечает, запрос повторяется по другим шинам. Второе решение – установка на шину памяти специального следящего устройства, передающего все адреса потенциально заинтересованным УВВ. Однако они не всегда могут быстро отработать эти запросы. Третье решение состоит в фильтрации адресов микросхемой моста PCI. Эта микросхема содержит регистры диапазона, заполняемые во время загрузки компьютера.

Прямой доступ к памяти

На практике часто применяется схема под названием прямой доступ к памяти. ОС может воспользоваться ПДП только при наличии аппаратного контроллера. DMA-контроллер может получать доступ к центральной шине независимо от ЦП. Центральный процессор программирует контроллер, устанавливая его регистры, указывая, какие данные и кулан ало переместить. Затем процессор дает команду дисковому контроллеру прочитать данные во внутреннем буфере и проверить контрольную сумму. Когда данные получены и проверены контроллер диска, DMA-контроллер начинает перенос данных, посылая дисковому контроллеру запрос чтения. Запись в память является следующим стандартным циклом. По окончании записи контроллер диска по шине посылает сигнал контроллеру DMA. Контроллер DMA увеличивает используемый адрес памяти и уменьшает значение счетчика байтов. Шаги повторяются, пока значение счетчика не станет равным нулю. По завершении копирования контроллер DMA инициирует прерывание процессора. Данный режим называется сквозным режимом. В некоторых контроллерах используется режим, при котором контроллер посылает слово данных контроллеру DMA, который затем выставляет на шину еще один запрос для передачи, куда это надо передать, то есть требуется лишний цикл. Большинство контроллеров DMA используют для передачи физические адреса. В некоторых контроллерах DMA применяется схема, при которых контроллер записывает сразу виртуальный адрес. В этом случае контроллер DMA использует менеджер памяти для преобразования адреса. Виртуальный адрес может быть выставлен на адресную шину только в том случае, когда менеджер памяти является частью памяти, а не частью ЦП.

15.11

Прерывания

На аппаратном уровне прерывания работают следующим образом. Когда устройство вывода заканчивает свою работу, оно инициирует прерывание при условии, что прерывания разрешены ОС. Для этого устройство выставляет сигнал на выделенную устройству специальную линию шины. Этот сигнал распознается микросхемой контроллера прерываний, расположенной на материнской плате. Контроллер прерываний принимает решение о дальнейших действиях.

Для обработки прерывания контроллер выставляет на адресную шину номер устройства, требующего к себе внимания, и устанавливает сигнал прерывания на соответствующий контакт процессора. Номер, выставленный на адресную шину, используется в качестве индекса в таблице, называемой вектором прерываний, из которой извлекается новое значение счетчика команд. Расположение вектора может быть жестко прошито на аппаратном уровне либо располагаться в произвольном месте в памяти, на которое указывает специальный регистр процессора, загружаемый операционной системой.

До начала обработки процедуры прерывания сохраняется определенная информация (как минимум счетчик команд). Большинство ЦП сохраняет информацию в стеке. Проблема сохранения информации вызвана тем, что современные процессоры широко используют конвейеры.

Прерывания, оставляющие машину в строго определенном состоянии, называются точными.

Свойства:

1. Счетчик команд сохраняется в известном месте
2. Все команды до той, на которую указывает счетчик, выполнены полностью
3. Ни одна команда после той, на которую указывает счетчик, не была выполнена
4. Состояние команды, на которую указывает счетчик команд, известно

Прерывание, не удовлетворяющее данным требования, называется неточным. Машина с неточным прерыванием выгружает огромное количество данных, что сильно замедляет обработку прерывания и восстановление после нее.

Принципы программного обеспечения ввода-вывода

Ключевая концепция – обеспечение ВВ независимо от устройства. Следствие – принцип единообразного наименования.

Способы переноса данных: синхронный (блокирующий) и асинхронный (управляемый прерываниями). Большинство операций ВВ на физическом уровне являются асинхронными, т.е. ЦП запускает процесс переноса данных и выполняет другой процесс, пока не придет прерывание.

Для обеспечения ВВ используется буферизация. При ВВ используется понятие коллективных устройств.

Существует 3 способа осуществления ВВ:

1. Программный ВВ
2. ВВ, управляемый прерываниями
3. ВВ с использованием DMA

ПО ВВ обычно организуется в виде четырех уровней:

1. ПО ВВ уровня пользователя
2. Независимое от устройств ПО ОС
3. Драйверы устройств
4. Обработчики прерываний

Когда происходит прерывание, начинает работать обработчик прерываний:

1. Сохраняются все регистры несохраненные аппаратурой
2. Устанавливается контекст для процедуры обработчика прерываний
3. Устанавливается указатель стека для процедуры обработки прерываний
4. Выдается подтверждение контроллеру прерываний
5. Копируется содержимое регистров из места сохранения в таблицу процессов
6. Запуск процедуры обработки прерываний
7. Выбирается процессора для передачи управления
8. Устанавливается контекст для следующего рабочего процесса
9. Загружаются регистры нового процесса
10. Выполняется новый процесс

Драйверы устройств

Чтобы получить доступ к аппаратной части устройства, драйвер должен быть частью ядра ОС. Драйверы устройств обычно располагаются под остальной ОС. ОС классифицирует их по нескольким категориям в зависимости от типа устройств. В большинстве ОС определен стандартный интерфейс, который должны обеспечивать драйверы.

Независимое от устройства ПО ВВ, функции:

1. Единообразный интерфейс для драйверов устройств
2. Буферизация
3. Сообщение об ошибках
4. Захват и освобождение выделенных устройств
5. Размер блока, не зависящий от устройства

22.11

Особенности дисковых устройств ввода-вывода

Магнитные диски вращаются с постоянной скоростью.

Магнитные диски

Магнитные диски организованы в цилиндры, каждый из которых содержит столько дорожек, сколько есть у устройства головок, установленных вертикально. Дорожки делятся на сектора. Магнитные диски обладают свойством – контроллер способен производить поиск дорожки на двух и более дисках. Многие контроллеры жестких дисков могут совмещать операцию чтения или записи на одном диске с поиском на другом или нескольких дисках.

Однако между контроллером и оперативной памятью в каждый момент может происходить только одна операция по переносу данных. Современные диски разделены на зоны с большим числом секторов на внешних дорожках и меньшим на внутренних. При этом используется виртуальная геометрия диска. Чтобы преодолеть физическое ограничение, многие диски поддерживают логическую адресацию блоков. При этом максимальная емкость может быть 131 000 ТБ.

Система RAID 0-2

1. способов организации дисков для улучшения производительности и надежности.

RAID – на компьютер устанавливается массив дисков, использующий специальный RAID-контроллер. Весь набор с точки зрения ОС выглядит как один большой дисковый накопитель.

* Система RAID уровня 0. Единый виртуальный диск, эмулируемый контроллером как разбитый на полосы, состоящий из одинакового числа секторов. Первый блок записывается на первый диск, второй – на второй и т.д. Способ хранения называется чередующимся набором.
* RAID массив уровня 1. Дублируются все диски. Система называется зеркальным набором, каждая полоса записывается дважды. При чтении используется любая копия. Недостаток – снижение используемой общей емкости дисков вдвое.
* RAID уровня 2. Работает не с полосами, а на уровне слов и байтов. Использует разбиение каждого байта на пару 4 битовых полубайтов, затем добавление к каждому из них кода Хейминга с образованием 7-битного слова, в котором 1, 2, 4 являются битами четности. Недостаток – требуется синхронизация вращения дисков и большие накладные расходы.
* RAID 3 уровня. Упрощенная версия уровня 2. Для каждого слова данных считается 1 бит четности, которые записывается на отдельный диск четности.
* RAID массивы 4 и 5 уровней. Работают с полосами, поэтому не требуют синхронизации дисков. RAID 4 уровня аналогичен уровню 0, но с дополнительным диском четности, содержащем сумму по модулю 2 всех данных с остальных дисков. Проблема повышенной нагрузки на диск в RAID 4 уровня решается в системе 5 уровня, в которой биты четности равномерно распределены по всем дискам.