

Основы операционных систем.

Андрей Тищенко @AndrewTGk

2024/2025

Содержание

Лекция 2 сентября.

1 Структура вычислительной системы.

Вычислительная система состоит из нескольких частей:

- Пользователь. (алгоритмы и алгоритмические языки)
- Прикладные программы. (ЦГ, matlab и т.п.)
- Системные программы. (системное программирование)
- Операционная система. (основы операционных систем)
- Техническое обеспечение. (архитектура ЭВМ и языки ассемблера)

Прекрасная притча про слепцов и слона.

1.1 Что такое операционная система?

Различные точки зрения (что такое ОС?):

- Распорядитель ресурсов.
- Защитник пользователей и программ.
- Виртуальная машина (фокусник, виртуальная память).
- Кот в мешке (попросил загрузить ОС, значит всё загружено и есть ОС).
- Постоянно функционирующее ядро.

Проще сказать не что такое операционная система, а для чего она нужна и чем занимается. Современные ОС это продукт эволюции вычислительных систем, поэтому стоит эту эволюцию рассмотреть.

1.1.1 Эволюция вычислительных систем

Удобство, стоимость и производительность — самые главные факторы отбора в эволюции операционных систем.

Условные этапы развития вычислительных систем:

период. (1945–1955гг.) Научно-исследовательская работа в области вычислительной техники.

- Ламповые машины.
- Нет разделения персонала.
- Вход программы коммутацией или перфокартами.

- Одновременное выполнение только одной операции.
- Появление прообразов первых компиляторов.
- Нет операционных систем.

период. (1955–начало 60-х гг.) Начало использования вычислительных машин в научных и коммерческих целях.

- Транзисторные машины
- Происходит разделение персонала (появление малочисленной касты программистов)
- Бурное развитие алгоритмических языков (напиши 10 000 строк на асме и отладь, это стало причиной появления первых языков программирования)
- Ввод заадния колодой перфокарт
- Вывод результатов на печать
- Пакеты заданий и системы пакетной обработки.

период. (начало 60-х годов–1980 гг.) Конец этого периода указан точно. Тактовая частота значительно повысилась из-за появления микросхем.

- Машины на интегральных схемах.
- Использование спулинга spooling. (появление процессоров ввода/вывода, которые включаются вместо центрального процессора)
- Планирование заданий (из-за появления магнитных дисков вместо ленты).
- Мультипрограммные пакетные системы. (в память загружается несколько программ, пока происходит ввод/вывод на в одной программе ЦП передаётся другой программе)
- Системы разделения времени (time-sharing).
- Интерактивная отладка программ.
- Виртуальная память.
- Появление семейств ЭВМ.

Мультипрограммирование и эволюция вычислительных систем.

Software	Hardware
Планирование заданий	Защита памяти
Управление памятью	Сохранение контекста
Сохранение контекста	Механизм прерывания
Планирование использования процессора	Привилегированные команды
Системные вызовы	
Средства коммуникации	
Средства синхронизации	

Интерактивная отладка стала возможна благодаря изобретению терминала, но возникла новая проблема: нужно хранить информацию разных пользователей так, чтобы они не могли портить данные других пользователей, также потребовался единый стандарт хранения данных в компьютере.

Хотелось разрешить как можно большему числу пользователей одновременно пользоваться вычислительной машиной. Для этого нужно было экономить оперативную память. С целью решения этой проблемы программы подгружаются в машину по частям (только нужные в данный момент куски). На этой концепции основана виртуальная память.

Семейства ЭВМ были нужны для того, чтобы маленькие компании могли покупать маломощные компьютеры, писать свой код на них, по мере роста компании покупать более мощные машины того же семейства и исполнять такой же код без переписывания и перекompиляции.

й период (1980–2005 гг.)

- Машины на больших интегральных схемах (БИС).

- Персональные ЭВМ.
- Дружелюбное программное обеспечение.
- Сетевые и распределённые операционные системы.

Поскольку теперь любой человек мог стать владельцем персонального ЭВМ, было необходимо создать дружелюбный интерфейс для предоставления человеку без специального образования пользоваться персональной машиной.

Из-за увеличения количества компьютеров увеличилась важность компьютерных сетей. Пользователи, работающие на сетевой ОС должны были понимать как доставать и обрабатывать файлы из сервера.

Владелец распределённой операционной системы не должен понимать, где находится его файл, обращение к файлу на локальной машине и к файлу на сервере, управляемом распределённой ОС, должно выглядеть для пользователя одинаково.

Широкое использование ЭВМ в быту, образовании и на производстве.

й период (2005-?? гг.)

- Машины на многоядерных процессорах.
- Мобильные компьютеры.
- Высокопроизводительные вычислительные системы.
- Облачные технологии.
- Виртуализация выполнения программ.

Можно виртуализовать программы и операционные системы.

Период Глобальной компьютеризации.

1.2 Основные функции ОС

- Планирование заданий и использования процессора. (кто и после кого будет выполняться)
- Обеспечение программ средствами коммуникации и синхронизации.
- Управление памятью.
- Управление файловой системой.
- Управление вводом-выводом.
- Обеспечение безопасности.

Операционные системы существуют потому что на данный момент их существование — это разумный способ использования вычислительных систем.

2 Архитектурные особенности построения ОС

Внутреннее строение ОС (разновидности)

Монолитное ядро.

- Каждая процедура (функция) может вызвать каждую.
- Все процедуры работают в привилегированном режиме.
- Ядро совпадает со всей операционной системой.
- Пользовательские программы взаимодействуют с ядром через системные вызовы.

Многоуровневые (Layered) системы.

- Процедура уровня K может вызывать только процедуры уровня $K - 1$
- Все или почти все уровни работают в привилегированном режиме
- Ядро совпадает или почти совпадает со всей операционной системой
- Пользовательские программы взаимодействуют с ОС через интерфейс пользователя

Самый низкий уровень — Hardware. Самый высокий — интерфейс пользователя. В самой первой такой машине выглядело так:

1. Интерфейс пользователя.
2. Управление ввода вывода.
3. Связь с консолью.
4. Управление памятью.
5. Планирование заданий и программ.

Работает медленнее, чем система монолитного ядра, но отладка и модификация сильно упрощается.

Микроядерная (microkernel) архитектура.

Функции микроядра:

- Взаимодействие между программами.
- Планирование использования процессора.
- Первичная обработка прерываний и процессов ввода вывода
- Управление памятью.

Устройство системы:

- Микроядро составляет лишь малую часть ОС.
- В привилегированном режиме работает только микроядро.
- Взаимодействие частей ОС между собой и с программами пользователей путем передачи сообщений через микроядро.

Все части ОС, кроме микроядра, могут быть заменены без прерывания работы ОС. Микроядро сразу же сможет обратиться к заменённой части системы.

Микроядро в таком концепте работает весьма медленно, поэтому на практике для повышения быстродействия функционал микроядра расширили.

Виртуальные машины

Каждому пользователю предоставляется своя копия виртуального hardware

Новая микроядерная архитектура (экзоядерная архитектура).

- Взаимодействие между программами
- Выделение и высвобождение физических ресурсов
- Контроль прав доступа

Вся остальная функциональность выделяется на поддержание абстракций, которые разделены на библиотеки. Считается спорным подходом и на практике не используется (только в учебных ОС).

Монолитное ядро - необходимость перекомпиляции при каждом изменении, сложность отладки, высокая скорость работы.

Многоуровневые системы - необходимость перекомпиляции при изменениях, отлаживается только изменённый уровень, меньшая скорость работы.

Микроядро - простота отладки, возможность замены компонент, без перекомпиляции и остановки системы, очень медленные.

Все три подхода используются в современных ОС.

3 Понятие процесса. Операции над процессами.

Уточнение терминологии

Термин *программа* - не может использоваться для описания происходящего внутри ОС.

Термин *задание* - не может использоваться для описания происходящего внутри ОС.

Изначально эти термины были придуманы для статических объектов, а нам нужно описывать динамические объекты.

Введём термин *процесс* для описания динамических объектов.

Теперь под *процессом* будем понимать совокупность:

- набора исполняющихся команд
- ассоциированных с ним ресурсов
- текущего момента его выполнения

Вся эта совокупность находится под управлением операционной системы.

Важно: Процесс \neq программа, которая выполняется.

- для исполнения одной программы может организовываться несколько процессов
- В рамках одного процесса может исполняться несколько программ
- В рамках процесса может исполняться код, отсутствующий в программе.

Состояния процесса

1. Исполняется
2. Не исполняется

Новые процессы не исполняются. Переход из состояний происходит в случае выбора его для исполнения или приостановки. После исполнения процесс выходит из ОС.

Примитивная система, которая не является верной, так как процесс может "не исполняться разными способами".

1. Готовность
2. Исполнение
3. Ожидание

Новые процессы попадают в состояние готовности. Выбор готового процесса для исполнения прерывает в состояние исполнения. В случае прерывания исполняемого процесса он переходит в состояние готовности. Если для исполнения нужно дождаться какого-то события, то он переходит в состояние ожидания, после исполнения ожидаемого, он переходит в состояние готовности.

Однако эту схему тоже стоит дополнить:

1. Рождение

2. Готовность
3. Исполнение
4. Ожидание
5. Закончил исполнение

Новые процессы попадают в состояние рождения, после прохождения допуска к планированию процесс попадает в состояние готовности. При завершении работы процесса он попадает в состояние "закончил исполнение".

Важно: переход между состояниями процесса осуществляет ОС.

Лекция 16 сентября

Операции над процессами:

- Создание процесса - завершение процесса
- Запуск процесса - приостановка процесса
- блокировка процесса - разблокирование процесса
- изменение приоритета

Чтобы ОС могла взаимодействовать с процессами, она должна иметь какое-то представление об этом объекте. Эта информация хранится в PCB (process control block).

Содержимое PCB

- Состояние процесса
- Программный счётчик (следующая команда в процессе)
- содержимое регистров
- данные для планирования использования процессора и управления памятью
- учётная информация
- сведения об устройствах ввода-вывода, связанных с процессом (возможно файлы, с которыми взаимодействует).

Регистровый контекст: программный счётчик, содержимое регистров.

Всё остальное входит в системный контекст.

Всё, что лежит в PCB лежит в ядре ОС, код и данные программы называются пользовательским контекстом, входят в адресное пространство процесса.

Создание процесса:

- Присвоение идентификационного номера
- Порождение нового PCB с состоянием процесса "рождение"
- Выделение ресурсов (из ресурсов родителя и ОС)
- Занесение в адресное пространство кода и установка значения программного счётчика. Может создаваться дубликат родителя, владеющий всей информацией родителя до момента желания создать дочерний процесс. Информация процесса может быть получена из файла (пользовательский контекст убирается, вместо него через системный вызов открывается исполняемый файл и подгружается как пользовательский контекст).
- Окончание заполнения PCB
- Изменение состояния процесса на "готовность"

Завершение процесса: Это состояние необходимо для того, чтобы родитель мог узнать, по какой причине ребёнок завершил процесс.

- Изменение состояния процесса на "закончил исполнение"
- Освобождение ресурсов (остаётся только PCB процесса, в нём остаётся учётная информация, родитель)
- Очистка соответствующих элементов в PCB
- Сохранение в PCB информации о причинах завершения.

После окончания родительского процесса некоторые ОС убивают всех его детей, а некоторые (например Windows и Unix) позволяют детям жить.

Но в таком случае возникает вопрос с новым родителем. Например у процесса 251 был ребёнок 1000, после окончания 251 1000 продолжил работать, ОС создала новый процесс 251, который не должен являться родителем 1000. Поэтому осиротевшие процесса усыновляются процессами, которые существуют, пока ОС не прекратит работу.

Если процессы используют какой-то ресурс (например файл), ему присваивается счётчик, равный количеству процессов, использующих этот ресурс. Освобождение происходит если один из процессов явно это попросил или если счётчик стал равен нулю.

Запуск процесса:

- Изменение состояния процесса на "исполнение"
- Обеспечение наличия в оперативной памяти информации, необходимой для его выполнения
- Восстановление значений регистров
- передача управления по адресу, на который указывает программный счётчик.

PCB меняется только при переключении состояния процесса. Во время работы процесса ничего в PCB не заносится (так как это было бы непроизводительно). Приостановка процесса:

- Автоматическое сохранение программного счётчика и части регистров (работа hardware)
- Передача управления по специальному адресу (hardware)
- Сохранение динамической части... UNFINISHED

Блокирование процесса

- Сохранение контекста процесса в PCB.
- Обработка системного вызова (разбираемся, чего мы ждём)
- Перевод процесса в состояние "ожидание"

Разблокирование процесса:

- Уточнение того, какое именно событие произошло
- Проверка наличия процесса, ожидающего этого события
- Перевод ожидающего процесса в состояние "готовность"
- Обработка произошедшего события

4 Кооперация процессов и основные аспекты ее логической организации

Основные причины объединения усилий:

- Повешение скорости решения задач.
- Совместное использование данных.
- Модульная конструкция какой-либо системы.
- Для удобства работы пользователя (например, отладчик).

Кооперативные или взаимодействующие процессы - это процессы, которые влияют на поведение друг друга путем обмена информацией.

Категории средств взаимодействия:

- Сигнальные (переданная информация зачастую ограничивается битами).
- Канальные (создание логического канала связи. Данные из одного процесса могут быть получены при желании другого, тогда произойдёт передача запрошенных данных)
- Разделяемая память (доступна обоим процессам).

Как устанавливается связь

- Нужна или не нужна инициализация?
- Способы адресации
 1. Прямая адресация
 - симметричная
 - асимметричная
 2. Непрямая или косвенная адресация

Информационная валентность процессов и средств связи

- Сколько процессов может быть ассоциировано с конкретным средством связи?
- Сколько идентичных средств связи может быть задействовано между двумя процессами?
- Направленность связи
 - Симплексная связь (только в одну сторону)
 - Полудуплексная связь (двусторонняя, но только по очереди, как рация)
 - Дуплексная связь (двусторонняя)

Лекция 23 сентября.

Основные аспекты логической организации передачи информации

Особенности канальных средств связи

Буферизация

- Буфера нет (нулевая ёмкость). Процесс-передатчик всегда обязан ждать приёма
- Буфер конечной ёмкости. Процесс-передатчик обязан ждать освобождения места в буфере перед продолжением работы.

- Буфер неограниченной ёмкости (нереализуемо!). Процесс-передатчик никогда не ждёт.

Модели передачи данных

- Поточковая модель. Операции приёма/передачи не интересуются содержимым данных и их происхождением. Данные не структурируются.
- Модель сообщений. На передаваемые данные накладывается определённая структура.

Потоковая модель - pipe

Источники вкладывают информацию порциями любого размера и могут считывать эти порции в любом размере. Через pipe могут общаться только процессы, имеющие общего предка, создавшего pipe. Это происходит потому что о положении начала и конца pipe знает только его создатель (и дети). Без средств синхронизации это строго односторонняя связь.

Потоковая модель - FIFO

Представляет собой pipe с помеченным входом и выходом, что позволяет неродственным процессам общаться через него.

Модель сообщений

Сообщения в pipe имеют чёткие границы, прочитать сообщение можно только целиком. У принимающего процесса пропадает возможность получать данные в произвольном размере.

Надёжность средств связи

Средство связи считается надёжным, если:

- Нет потери информации
- Нет повреждения информации
- Нет нарушения порядка поступления информации
- Не появляется лишняя информация

Современные машины считаются надёжными.

Как завершается связь

- Нужны ли специальные действия для прекращения использования средства связи?
- Как влияет прекращение использования средства связи одним процессом на поведение других участников взаимодействия?

Нити исполнения (threads)

Для реализации многопоточности нужно из основного процесса создать ещё один процесс. Переключить контекст, запросить доступ к общей памяти (или другому средству связи). Такие накладные расходы нужны для реализации простейшей многопоточности. На практике из-за накладных расходов это работает ещё медленнее. Поэтому придумали thread'ы.

На удивление хорошо поведение процесса можно сравнить с игрушечной железной дорогой. Шпалы принимаются за инструкции, развилки — условные переходы. Станции — данные со стека или input/output.

Сам поезд можно принимать за регистры и данные в стеке.

Если поставить две такие дороги рядом, то можно получить аналог второго процесса.

А если поставить на те же рельсы второй поезд, можно получить аналог thread. Так как поезда разные, у них разные данные на регистрах и в стеке, всё остальное у них общее, так же должна присутствовать информация, позволяющая различать поезда между собой.

Процесс: системный контекст, регистровый контекст, код, данные вне стека, стек.

Нить исполнения: системный контекст нити, регистровый контекст, стек.

При появлении процесса нить всего одна (называется главной нитью или master нитью). Нити также могут находиться в состоянии готовности, ожидания, исполнения и т.д.

Нить находится в состоянии

- Готовности, если нет ни одной нити в состоянии исполнения и есть хотя бы одна в состоянии готовности
- Ожидания, если нет ни одной нити в состоянии готовности и исполнения.
- Исполнения, если хотя бы одна из нитей находится в состоянии исполнения.
- Закончил исполнения, если все его нити находятся в состоянии закончил исполнение.

По-прежнему нужно тратить ресурсы на создание новой нити (нить создаётся легче, чем процесс). Между нитями нет нужды создавать дополнительные средства связи, так как внестековое пространство у них и так общее. Переключение контекста между нитями происходит быстрее, чем между процессами. Везде выиграли.

Алгоритмы синхронизации

Активности и атомарные операции

Активность — последовательное выполнение ряда действий, направленных на достижение определённой цели.

Активность: приготовление бутерброда.

- Отрезать ломтик хлеба
- Отрезать ломтик колбасы
- Намазать хлеб маслом
- Положить колбасу на хлеб

Каждый элемент активности является атомарным (неделимым). Во время операции никуда отвлекаться нельзя, между ними - можно.

Пусть активность P: a b c, Q: d e f. (какие-то операции)

При последовательном выполнении PQ: a b c d e f.

Псевдопараллельное выполнение

(режим деления времени): a d b e c f. В таком случае атомарные операции одного процесса расположены в правильном порядке, поэтому на итоговый результат не влияет.

Детерминированность выбора

P: $x = 2$, $y = x - 1$, Q: $x = 3$, $y = x + 1$

При последовательном выполнении PQ: $(x, y) = (3, 4)$.

Однако перестановка $x = 2, x = 3, y = x - 1, y = x + 1 \Rightarrow (x, y) = (3, 3)$, приводит к недетерминированности.

Условия Бернштейна

Пусть P: 1. $x = u + V$, 2. $y = x \cdot w$. Входные данные: $R_1 = \{u, v\}$, $R_2 = \{x, w\}$. Выходные данные: $W_1 = \{x\}$, $W_2 = \{y\}$.

Вход для активности $R(P) = \{u, v, x, w\}$. Выход для активности $W(P) = \{x, y\}$

Тогда достаточные условия детерминированности Бернштейна:

1. $W(P) \cap W(Q) = \emptyset$
2. $W(P) \cap R(Q) = \emptyset$
3. $R(P) \cap W(Q) = \emptyset$

Если все условия выполнены, то набор активностей $\{P, Q\}$ является детерминированным. В недетерминированных наборах всегда встречается race condition (состояние гонки). Избегать недетерминированного поведения при неважности очередности доступа можно с помощью *взаимоисключения* (mutual exclusion).

Критическая секция

Пример про трёх студентов и пиво в пятницу. Каждый из студентов приходит в комнату, видит, что пива нет и уходит в магазин покупать. В итоге каждый из них купил пива на троих и получилась славная пьянка. Поэтому операцию уйти за пивом, купить на всех и вернуться с пивом стоит заменить атомарной операцией. Тогда первый студент выполнит эту операцию и двум другим ничего делать не придётся.

Структура кооперативного процесса

```
while (some condition) {  
  entry section  
    critical section  
  exit section  
    remainder section  
}
```

Требования к программным алгоритмам

1. Программный алгоритм должен быть программным (проблемы решаются в коде, а не на уровне hardware)
2. Нет предположений об относительных скоростях выполнения и числе процессоров
3. Выполняется условие взаимоисключения (mutual exclusion) для критических участков.
4. Выполняется условие прогресса (progress).
 - Решения принимают только те, кто готов войти в критическую секцию
 - Решение должно приниматься за конечное время
5. Выполняется условие ограниченного ожидания (bound waiting). Для каждого процесса можно сказать, какое максимальное (конечное) количество процессов он может пропустить, после чего гарантированно пройдёт в критическую секцию.

Лекция 30 сентября

Подробнее о пункте 3. Мы хотим заставить участки процесса выполняться атомарно по отношению процессам, входящим в набор взаимодействующих процессов. Никакие два процесса из набора не могут оказаться в критическом участке в состоянии исполнения или готовности одновременно.

История придумывания алгоритма синхронизации

Запрет прерываний

Существует команда Client Disable Interrupt, она говорит процессору отключить прерывания (кроме критических ошибок (nonmaskable interrupt) и прерываний, созданных пользователем). Также есть команда, возвращающая процессору возможность видеть прерывания. Если перед критической секцией поставить запрет на прерывания, а после этой секции включить прерывания, то процессы внутри будут выполняться атомарно. Минусы: бесконечный цикл внутри критической секции может быть прерван только выключением вычислительной машины. Поэтому такой метод используется только при программировании на уровне ядра.

Переменная “замок”

Оборачивать код в нечто похожее на

```
shared int lock = 0;
while (lock == 1);
lock = 1;
critical section
lock = 0;
```

Минусы: оба процесса хотят войти в критическую секцию, первый проходит проверку на lock, заходит внутрь, но сделать lock = 1 он не успевает, так как у него забирают управление. Второй процесс входит в секцию, проверяет lock, заходит в критическую секцию, ставит lock = 1, после чего у него забирают управление. В результате мы имеем два процесса, один из которых выполняется в критической секции, а другой - в готовности в этой же критической секции.

Строгое чередование

Пусть у нас два процесса. Оборачиваем критическую секцию следующим образом:

```
shared int turn = 1;
```

Process₀

```
while (condition) {
    while (turn != 0);
    critical section
    turn = 1;
    remainder section
}
```

Process₁

```
while (condition) {
    while (turn != 1);
    critical section
    turn = 0;
    remainder section
}
```

Процессы не могут одновременно находиться в критической секции, так как тогда переменная turn должна быть 0 и 1 одновременно, что невозможно.

Однако условие прогресса нарушается. Пусть у *Process₀* долгая remainder section, а у *Process₁* обе секции короткие. Тогда первый процесс может прокрутиться внутри цикла и встать в очередь на вход в критическую секцию. Тем временем нулевой процесс всё ещё в remainder section и не может поставить флаг для первого процесса.

Флаги готовности

Пусть у нас два процесса. Оборачиваем критическую секцию следующим образом:

```
shared int ready[2] = {1, 0};
```

Process₀

```
while (condition) {
    ready[0] = 1;
    while (ready[1]);
    critical section
}
```

```
    ready[0] = 0;
    remainder section
}
```

Process₁

```
while (condition) {
    ready[1] = 1;
    while (ready[0]);
        critical section
    ready[1] = 0;
    remainder section
}
```

Здесь нарушается вторая часть условия прогресса, так как оба процесса могут одновременно быть готовыми ко входу в критическую секцию, тогда они оба будут крутить бесконечный цикл по очереди.

Алгоритм Петерсона

Совмещаем флаг и переменную готовности.

```
shared int ready[2] = {1, 0};
shared int turn;
```

Process₀

```
while (condition) {
    ready[0] = 1;
    turn = 1;
    while (ready[1] && turn == 1);
        critical section
    ready[0] = 0;
    remainder section
}
```

Process₁

```
while (condition) {\
    ready[1] = 1;
    turn = 0;
    while (ready[0] && turn == 0);
        critical section
    ready[1] = 0;
    remainder section
}
```

В этом случае для двух процессов все условия выполняются.

Bakery algorithm

Если бы алгоритм придумали в России, то назвали бы алгоритм регистратуры в поликлинике (так как пекарни на западе выпекают заказы, а не предлагают готовый ассортимент).

Основные идеи

1. Процессы можно сравнивать по именам.
2. Перед входом в критическую секцию процессы получают талончик с номером.
3. Первым в критическую секцию входит процесс с наименьшим номером на талоне.
4. Выписывание талона - неатомарная операция. Могут быть талоны с одинаковыми номерами

5. При совпадении номеров на талоне первым входит процесс с меньшим значением имени процесса.

Последние два алгоритма прекрасно работали до 2005 года. В один роковой день появились многоядерные процессоры. Теперь разные ядра могли писать в одну и ту же память, в результате чего одно ядро могло записать в память одно число, после чего считать оттуда же другое, ведь с этой памятью поработало другое ядро. Для решения проблемы было принято решение ослабить консистентность памяти.

Аппаратная поддержка

Команда Test-And-Set

Выполняет нечто похожее на это:

```
int Test_And_Set (int* a) {
    int tmp = *a;
    *a = 1;
    return tmp;
}
```

Используется для сокращения прологов и эпилогов:

```
shared int lock = 0
while (condition) {
    while(Test_And_Set(lock));
        critical section
    lock = 0;
        remainder section
}
```

Однако это нарушает условие ограниченного ожидания (исправление ошибки оставлено лектором как несложное упражнение).

Команда Swap

```
void Swap(int* a, int* b) {
    int tmp = *a;
    *a = *b;
    *b = tmp;
}
```

Можно использовать так:

```
shared int lock = 0
int key = 0;
while (condition) {
    key = 1;
    do Swap(&lock, &key);
    while(key);
    critical section
    lock = 0
    remainder section
}
```

Однако это нарушает условие ограниченного ожидания (исправление ошибки оставлено лектором как несложное упражнение).

Механизмы синхронизации

Недостатки программных алгоритмов

Непроизводительная трата процессорного времени в циклах пролога

Посмотрим

```
while (condition) {
    entry section
        critical section
    exit section
        remainder section
}
```

Хотелось бы сократить убрать цикл активного ожидания (busy wait) в пассивное ожидание. Это не всегда оправдано, но зачастую является таковым. То есть просим процесс не крутиться в цикле, а перейти в состояние ожидания (обратное переключение с синхронизацией могут занимать больше времени, чем просто прокрутка по циклу).

Задачи с разными приоритетами

Задача с высоким приоритетом с более важным видом ждёт очереди на вход в критическую секцию.

Семафоры Дейкстры

S — семафор. Целая разделяемая переменная с неотрицательными значениями. При создании может быть инициализирована любым неотрицательным значением.

Допустимые атомарные операции

- P(S): пока $S == 0$ процесс блокируется;
Также выполняется $S = S - 1$
- V(S): $S = S + 1$

Проблема Producer-Consumer

Producer:

```
while(1) {
    produce_item();
    put_item();
}
```

Consumer:

```
while(1) {
    get_item();
    consume_item();
}
```

Синхронизируем их работу с помощью семафоров.

```
Semaphore mut_ex = 1;
Semaphore full = 0;
Semaphore empty = N;
```

Producer:

```
while(1) {
    produce_item();
    P(empty);
    P(mut_ex);
    put_item();
    V(mut_ex);
    V(full);
}
```

Consumer:

```
while(1) {
    P(full);
    P(mut_ex);
    get_item();
    V(mut_ex);
    V(empty);
    consume_item();
}
```

Интересный факт: если поменять местами P(full) и P(mut_ex) у Consumer, то можно словить deadlock. Эта ошибка происходит не всегда, найти её сложно, поэтому с семафорами нужно работать осторожно.

Мониторы Хора

Структура

```
Monitor monitor_name {
    params
    void m_1(...) {...}
    void m_2(...) {...}
    ...
    void m_n(...) {...}
    variables initialization
}
```

Условные переменные (condition variables)

Condition C;

- C.wait
Процесс, выполнивший операцию wait над условной переменной, всегда блокируется
- C.signal
Выполнение операции signal приводит к разблокированию только одного процесса, ожидающего этого (если он не существует). Процесс, выполнивший операцию signal, немедленно покидает монитор.

Попытаемся решить задачу Producer-Consumer через мониторы Хора:

```
Monitor Prod_Cons {
    Condition full, empty;
    int count;
    void put() {
        if (count == N) full.wait;
        put_item(); count++;
        if (count == 1) empty.signal;
    }
    void get() {
```



```
        if (count == 0) empty.wait;
        get_item(); count--;
        if (count == N - 1) full.signal;
    }
    Prod_Cons(): count(0) {}
}
```

Producer:

```
while(1) {
    produce_item();
    Prod_Cons.put();
}
```

Consumer:

```
while(1) {
    Prod_Cons.get();
    consume_item();
}
```

Лекция 7 октября

Механизмы синхронизации

Очереди сообщений

- Для передачи данных: `send(address, message)`
Блокируется при попытке записи в заполненный буфер.
 - Для приёма данных: `receive(address, message)`
Блокируется при попытке чтения из пустого буфера.
- Операционная система гарантирует атомарность операций.

С помощью этих инструментов реализуем Produces-Consumer:

Producer

```
while(1) {
    produce_item();
    send(address, item)
}
```

Consumer:

```
while(1) {
    receive(address, item);
    consume_item();
}
```

```
Semaphore mut_ex = 1;

void mon_enter(void) {
    P(mut_ex);
}

// quit monitor with return
void mon_exit(void) {
    V(mut_ex);
}

Semaphore c[n] = 0; // for every conditional variable
int f[n] = 0;

// wait operation
void wait(i) { // i - index of conditional variable
    f[i] += 1;
    V(mut_ex);
    P(c[i]);
    f[i] -= 1;
}

// quit monitor with signal
void signal_exit(i) {
    if (f[i]) {
        V(c[i]);
    } else {
        V(mut_ex);
    }
}
```

```
Semaphore Amut_ex = 1;
Semaphore Afull = 0;
Semaphore Aempty = N; // N = buffer size

put_item(A, msg);
get_item(A, msg);

send(A, msg) {
    P(Aempty);
    P(Amut_ex);
    put_item(A, msg);
    V(Amut_ex);
    V(Afull);
}

receive(A, msg) {
    P(Afull);
    P(Amut_ex);
    get_item(A, msg);
    V(Amut_ex);
    V(Aempty);
}
```

```
Monitor sem {  
    unsigned int count;  
    Condition c;  
    void P(void) {  
        if (count == 0) {  
            c.wait;  
            count = count - 1;  
        }  
    }  
    void V(void) {  
        count = count + 1;  
        c.signal;  
    }  
    {count = N;}  
};
```

```
void Sem_init(int N) {
    int i;
    // create message queue M
    for (i = 0; i < N; i++) {
        send(M, msg);
    }
}

void Sem_P() {
    receive (M, msg);
}

void Sem_V() {
    send(M, msg);
}
```

Задача

В лесу стоит бочка для меда. Изначально она пуста. К бочке прилетают пчелы и кидают в бочку по капле меда. Одновременно положить в бочку мед две пчелы не могут. Бочка вмещает N капель. Если бочка заполнена, пчела кружат вокруг нее. Рядом с бочкой бродит или спит медведь, иногда подходит к бочке. Если бочка заполнена полностью, он отгоняет пчёл, съедает весь мед и отправляется бродить дальше. Пчела, положившая мёд в бочку последней находит медведя и кусает его, чтобы уведомить его о заполнении бочки.

Процесс пчелы:

1. Приносит мёд
2. Если места нет — ждёт
3. Кладёт мёд
4. Если это последняя капля — жалит медведя
5. Улетает за новым мёдом

Процесс медведя:

1. Бродит или спит
2. Если бочка не полная — спит
3. Съедает весь мёд
4. Приклеивает дно обратно
5. Уходит спать

Решение через семафоры

```
Shared int Count = 0;
Semaphore bee = 1, bear = 0;
```

Пчела:

```
while(1) {  
    P(bee);  
    Count++;  
    if (Count == N) {  
        V(bear);  
    } else {  
        V(bee);  
    }  
    \\ go to find new honey  
}
```

Медведь:

```
while(1) {  
    P(bear);  
    Count = 0;  
    V(bee);  
    // bear falling asleep  
}
```

Решение через очереди сообщений

```
Message queue bee, bear;  
Shared int Count = 0;
```

Пчела:

```
while(1) {  
    receive(bee, m);  
    Count++;  
    if (Count == N) {  
        send(bear, m);  
    } else {  
        send(bee, m)  
    }  
    // find honey  
}
```

Медведь:

```
while(1) {  
    receive(bear, m);  
    if (Count == N) {  
        Count = 0;  
        send(bee, m);  
    }  
    // to be continued  
}
```

```
Monitor BeeBear {
    Condition cbee, cbear;
    int Count;
    void bee() {
        if (Count == N) {
            cbee.wait;
        }
        Count++;
        if (Count == N) {
            cbear.signal;
        } else {
            cbee.signal;
        }
    }

    void bear() {
        if (Count != N) {
            cbear.wait;
        }
        Count = 0;
        cbee.signal;
    }

    {Count = 0;}
}
```

Пчела

```
while(1) {
    BeeBear bee();
    // new honey
}
```

Медведь

```
while(1) {
    BeeBear bear();
    // walking
}
```

Планирование процессов

Уровни планирования

- Долгосрочное планирование — планирование заданий.
- Среднесрочное планирование — swapping (некоторые процессы выгружаются из оперативной памяти, по необходимости подгружаются)
- Краткосрочное планирование — планирование использования процессора.

Цели планирования

- Справедливость. Процессам отводится примерно одинаковое количество процессорного времени.
- Эффективность. Добиться максимальной загрузки центрального процессора. В реальности загрузка в 80-90 процентов считается эффективным.

- Сокращение полного времени выполнения (turnaround time). Полное время выполнения — от момента старта процесса до момента завершения.
- Сокращение времени ожидания. Из полного время исполнения вычести время, в течение которого процесс что-то делал.
- Сокращение времени отклика (response time). С точки зрения пользователя: время от момента нажатия клавиши до появления реакции.

Желаемые свойства алгоритмов

- Предсказуемость. Время исполнения программы и её поведение не должно отличаться при одинаковых данных.
- Минимизация накладных расходов.
- Равномерность загрузки вычислительной системы.
- Масштабируемость. При подъёме числа процессов, время исполнения не должно сильно увеличиваться.

Параметры планирования

Статические	Динамические
Статические параметры вычислительной системы — например, предельные значения её ресурсов	Динамические параметры вычислительной системы — например, количество свободных ресурсов в данный момент.

Статические параметры процесса — кем запущен, степень важности, запрошенное процессорное время, какие требуются ресурсы.

Важные динамические параметры процесса

$$\text{CPU burst} \begin{cases} a = 1 \\ b = 2 \\ \text{read } c \end{cases}$$

$$\text{IO burst} \begin{cases} \text{Ожидание окончания ввода} \end{cases}$$

$$\text{CPU burst} \begin{cases} a = a + c + b \\ \text{print } a \end{cases}$$

$$\text{IO burst} \begin{cases} \text{Ожидание окончания вывода} \end{cases}$$

Вытесняющее и невытесняющее

Вынужденное принятие решения:

- Перевод процесса из состояния “исполнение” в состояние “закончил исполнение”
- Перевод процесса из состояния “исполнение” в состояние “ожидание”

Невынужденное принятие решения:

- Перевод процесса из состояния “исполнение” в состояние “готовность”
- Перевод процесса из состояния “ожидание” в состояние “готовность”

Вытесняющее планирование использует вынужденное и невынужденное планирование. Невытесняющее только вынужденное.

Алгоритмы планирования

FCFS (First Come First Served)

Процессы	P_0	P_1	P_2
Продолжительность CPU burst	13	4	1

В этом случае время работы будет равно $13 + 4 + 1 = 18$

Время ожидания: $2 \cdot 13 + 4 = 30$

Если поменять процессы местами, то время работы не изменится, но время ожидания будет $2 \cdot 1 + 4 = 6$. То есть предсказуемость алгоритма очень плохая.

Более сложные схемы управления памятью.

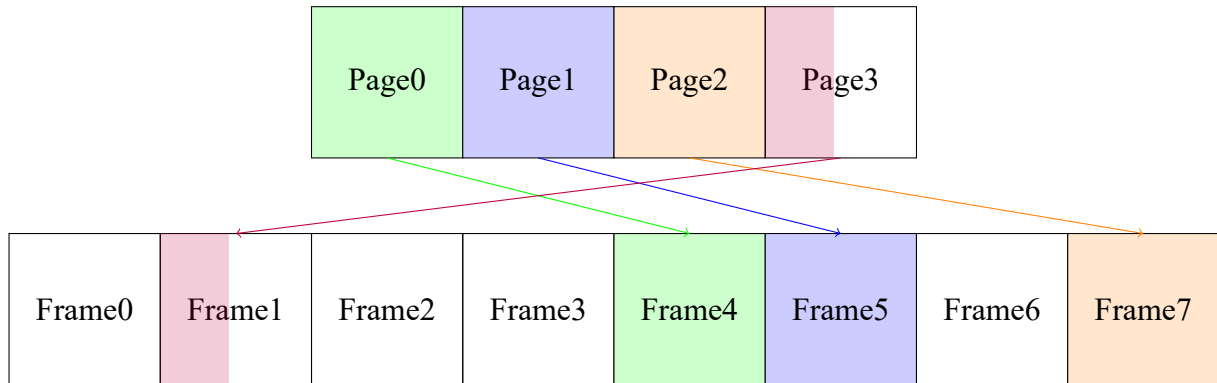
Страничная организация памяти

Имеем логическое адресное пространство и физическое адресное пространство. Делим логическое пространство на пронумерованные страницы определённого (одинакового) размера.

Теперь логический адрес = $N_{page} \cdot size + offset$. То есть позиция в логическом пространстве определяется парой (N_{page} , $offset$).

Физическое пространство делим на пронумерованные кадры одинакового размера.

Теперь физический адрес = $N_{frame} \cdot size + offset$. Определяем парой (N_{frame} , $offset$).



Получаем таблицу страниц (находится в PCB у каждого процесса):

Страница	0	1	2	3
Кадр	4	5	7	1

Кусочно-непрерывное отображение

Процессор -> логический адрес (N_{page} , $offset$) -> таблица страниц -> физический адрес (N_{frame} , $offset$) -> Память

MMU (БУП) - memory management unit (блок управления памятью)

Сегментно-страничная организация

Теперь логический адрес двумерный:

(N_{seg} , $soffset$), где $soffset = N_{page} \cdot size + poffset$. Задаётся кортежем (N_{seg} , N_{page} , $poffset$).

Физический адрес линейный = $N_{frame} \cdot size + offset$. Задаётся парой (N_{frame} , $offset$).

Процессор -> (N_{seg} , $soffset$) -> ... -> (N_{frame} , $offset$)

В ... входят:

1. Таблица сегментов

Адрес	Размер

Проверяем, что $soffset$ не больше размера сегмента, иначе кидает segmentation fault. Если всё хорошо, получаем пару (N_{page} , $poffset$).

2. Таблица страниц

Страница	Кадр

Из неё получаем искомую пару (N_{frame} , $offset$).

Многоуровневая таблица страниц

Таблица страниц процесса располагается в физическом адресном пространстве. При больших размерах таблицы страниц её размещение в последовательных кадрах памяти проблематично.

Решение: разбить таблицу страниц на страницы, сделать таблицу страниц для таблицы страниц и поместить страницы в память в любом порядке (повторять, пока не залезет в память).

При двухуровневой организации таблицы страниц логический адрес процесса описывается тройкой (p_1, p_2, d) , где p_1 — номер страницы в таблице страниц процесса, p_2 — смещение внутри страницы p_1

Хэшированная таблица страниц

Номер страницы хэшируется, из хэш таблицы получаем номер кадра и переходим к физической памяти.

Ассоциативная память (TLB)

(page, offset) -> TLB -> (frame, offset) TLB (translation lookaside buffer):

страница	кадр

Если запись с нужной страницей есть, то мы её быстро получаем. Если нет, то нужно залезать в таблицу страниц.

Расчёт среднего времени доступа

Обозначения:

t_0 — среднее время доступа к оперативной памяти. Пусть 100 наносекунд

t_1 — среднее время доступа к TLB. Пусть 10 наносекунд

h — вероятность наличия информации в TLB (hit ratio). Пусть это 0.8.

Среднее время доступа к данным при двухуровневой страничной схеме — это:

$$T = t_1 + ht_0 + (1 - h) \cdot 3t_0 = 10 + 80 + 60 = 150$$

Без TLB:

$$T = 3t_0 = 300$$

В среднем требуется в два раза меньше времени.

Концепция виртуальной памяти

1. Логическое адресное пространство процесса разбито на участки и линейно кусочно-непрерывно отображается. Недописал
2. Недописал
3. В оперативной физической памяти одновременно размещаются не все участки логического адресного пространства, а только их часть, остальные находятся во вторичной памяти.
4. При обращении к участку логического адресного пространства, находящемуся во вторичной памяти, он подкачивается в оперативную память, возможно, с вытеснением из нее некоторых неиспользуемых в данный момент участков.

Преимущества виртуальной памяти

1. Процесс не ограничен объемом физической памяти. Упрощается разработка программ.
2. Повышается степень мультипрограммирования.
3. Выгрузка во вторичную память части процесса происходит быстрее, чем выгрузка всего процесса. Повышается эффективность работы системы, использующей среднесрочное планирование.

Изменения в таблице страниц.

Добавляем два бита к номеру страницы. Один из них отвечает за наличие страницы, второй (необязательный) — бит модификации страница.

При обращении к странице с нулевым битом наличия происходит исключительная ситуация page fault:

1. Выполнение команды прекращается — hardware.
2. Сохраняется часть контекста перед выполнением команды. Управление передается по заранее определенному адресу — hardware
3. Сохраняется оставшийся контекст — hardware
4. Страница подкачивается в память, возможно, с выталкиванием из памяти другой страницы — software + hardware.
5. Восстановление контекста. Повторное выполнение команды — software + hardware.

Бит модификации нужен для того, чтобы не закачивать немодифицированную страницу памяти на диск при необходимости её выкидывания из оперативной памяти (страница и так уже лежит на диске в неизменённом виде, то есть сохраняем только изменённые страницы).

Стратегии управления

1. Стратегия выборки — когда подкачивать страницу?
 - По запросу (page fault \Rightarrow подгрузили).
 - С упреждением (по запросу + подгружаем соседние страницы)
2. Стратегия размещения — куда подкачивать страницу?
3. Стратегия замещения — что из оперативной памяти убрать?

Алгоритмы замещения страниц

Виды алгоритмов

- Локальные: процессу выделяется определённое количество кадров памяти, и только в этих кадрах он работает.
- Глобальные: при работе процесса можно использовать кадры других процессов.

Для анализа работы алгоритмов используется строка обращений процесса к памяти (строка запросов).

Для локальных алгоритмов обычно используются сокращённые строки обращений. Если к некоторой странице обращаются несколько раз подряд, то записывается только первое такое обращение.

Локальные алгоритмы замещения

FIFO

Аномалия Belady

Для некоторых строк запросов работа алгоритма при увеличении количества выделенных кадров приводит к увеличению числа страничных нарушений.

Определение

Стековые алгоритмы — алгоритмы, для которых при одной и той же строке запросов в один и тот же момент её обработки набор страниц в памяти для n кадров всегда есть подмножество набора страниц в памяти для $n + 1$ кадра, называются стековыми алгоритмами.

Стековые алгоритмы не проявляют аномалии Belady.

4.1 Локальные алгоритмы замещения

4.1.1 OPT оптимальный алгоритм

Идея заключается в том, что мы выгружаем из памяти ту страницу, к которой больше всего не будем обращаться.

На практике не реализуется, потому что для его выполнения мы должны наперёд знать строку запросов.

4.1.2 LRU (Least Recently Used)

Выгружаем из памяти ту страницу, к которой дольше всего не обращались.

Может работать быстрее FIFO, может и медленнее.

4.1.3 NFU (Not Frequently Used)

Выгружаем ту страницу, к которой меньше всего обращались. Есть ещё несколько локальных алгоритмов замещения. На лекции было сказано, что рассказывать про них нет времени, но стоит прочитать про них в учебнике.

4.2 Глобальные алгоритмы замещения

4.2.1 Трешинг (Thrashing)

Процесс (вычислительная система) находится в состоянии трешинга, если при его работа больше времени уходит на подкачку страниц, нежели на выполнение команд. Может произойти при создании большого количества процессов (чем их больше, тем меньше каждому даётся кадров), поэтому каждое обращение их к новой странице может вызывать page fault.

4.2.2 Концепция рабочих множеств

Основывается на принципе локальности.

Рабочее множество — это набор страниц, активно использующихся вместе длительное время.

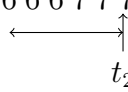
Концепция предполагает, что во время выполнения процесс перемещается от одного рабочего множества к другому.

Рабочие множества определяются структурой программы и организацией данных.

4.2.3 Модель рабочего множества

Использует параметр Δ — окно рабочего множества и полную строку прошедших обращений к памяти.

... 1 6 2 5 5 5 3 3 3 1 2 6 7 2 7 6 6 6 7 7 7 1 3 ...



The diagram shows a horizontal sequence of page numbers: ... 1 6 2 5 5 5 3 3 3 1 2 6 7 2 7 6 6 6 7 7 7 1 3 A horizontal arrow points from the left towards a vertical line. Below the vertical line is the label t_2 . The vertical line is positioned at the first '1' of the last group of numbers (1 3), and the arrow indicates the window of size Δ extending to the left from this point.

4.2.4 Алгоритм границ

Строим график зависимости частоты page fault от числа выделенных кадров.

5 Файлы и файловые системы (часть 1)

5.1 Основные проблемы изложения темы

1. Смешивание понятий различных уровней абстракции (например, логического и физического) для файлов и файловых систем.
2. Практическое игнорирование сходства и различия управления коллекцией файлов и управления оперативной памятью в мультипрограммных вычислительных системах.

5.2 Смешивание понятий. Файлы

1. Файл как математический объект.
2. Файл как абстрактный объект в программировании — логический объект.
3. Файл как набор связанной информации, обычно расположенный во вторичной памяти — физический файл.

5.3 Смешивание понятий. Целые данные

Целое число как математический объект.

1. Множество допустимых значений: от $-\infty$ до $+\infty$
2. Мощность множества — счётное множество.
3. Основные операции — сложение, вычитание, умножение, деление нацело.

Целые числа как абстрактные объекты в программировании — логические целые данные

1. Последовательность битов ограниченного размера, интерпретируемая как некоторая запись целого числа в двоичной системе счисления.
2. Множество допустимых значений от -2^{n-1} до $2^{n-1} - 1$, где n — количество отведённых битов.
3. Мощность множества — конечное множество.
4. Основные операции над целыми данными — сложение, вычитание, умножение, деление, битовые операции.

Целые данные как физические данные в памяти вычислительной системы.

1. Интерпретация старшего бита как знакового бита. Представление отрицательных чисел в дополнительном коде.
2. Отбрасывание старших битов при переполнении при выполнении арифметических операций.
3. Кратность количества битов представления числа должно быть кратно количеству битов в байте.
4. Каждое целое данное в оперативной памяти располагается в непрерывной не успел...

5.4 Файл как математический объект

Файл — это упорядоченный набор информации в цифровом виде, у которого есть начало и конец. Набор снабжён указателем текущей позиции, установленным в начале, в конце или между элементами набора.

Над файлом допустимы следующие операции:

1. Чтение (read). Считываем элемент перед указателем текущей позиции и переходим на следующий элемент. Если читать нечего, возвращаем \emptyset .
2. Запись (write). Записываем элемент перед указателем текущей позиции и переходим на следующий элемент (на записанный).
3. Усечение (truncate). Убираем последний бит в файле, если там стоял указатель текущей позиции, двигаем его на новый конец файла.
4. rewind/seek. Смотри ниже.

Бывают файлы *последовательного* доступа, для них определена ещё операция rewind (переместить указатель в начало).

Бывают файлы *прямого* доступа, для них определена операция seek (поиск указанного элемента, установить указатель перед ним).

5.5 Структурирование информации. Записи

Бывают записи постоянной длины и переменной длины. все операции переходят на работу с этими записями (читать запись, добавить запись, убрать последнюю запись, rewind/seek), указатель может находиться строго между записями.

5.6 Логический файл — файл в программе

1. Файл может храниться на диске или на другом внешнем носителе информации и не прекращать существование после завершения процесса или завершения работы операционной системы.
2. Существуют ограничения на размер файла
3. Минимальной единицей информации для операций чтения и записи является байт
4. Не успел

Операции с логическим файлом

1. Размещение файла во вторичной памяти требует добавления операций create file, delete file.
2. Файл может храниться долго. Появляются атрибуты, содержащие времена создания, модификации, последнего обращения к файлу.
3. Не успел
4. Наличие атрибутов требует операций get attributes, put attributes
5. Появляются операции для безопасного использования файла несколькими пользователями.

5.7 Записи в логическом файле

Файл последовательной длины с записями переменной длины. Необходимо в начале каждой записи указывать её размер, иначе невозможно итерироваться по файлу.

Поскольку в современных вычислительных системах обычно на внешних носителях хранится не единственный файл, а некоторый набор (или коллекция) файлов, то рассмотрение физического файла в отдельности не имеет смысла, необходимо смотреть файловую систему.

6 Файловые системы

Файловая система как способ организации коллекции логических файлов — логическая файловая система. Файловая система как структурированная совокупность файлов. Не успел.

6.1 Логическая файловая система

6.1.1 Куча

При переходе от отдельного логического файла к коллекции логических файлов, файлы требуется именовать (разными именами).

По сути куча является одноуровневой директорией.

Директория — отдельная структура данных, состоящая из записей. Каждая запись содержит имя файла и информацию о том, как до него добраться.

Эта структура тоже требует постоянного хранения. В логической файловой системе появляются 2 различных типа файлоа: директории и (не успел). Двухуровневая директория:

Директория Master files, в которой хранится информация о том, как добраться до user files директории, в которой хранится информация о файлах.

Обобщением предыдущих двух идей является дерево директорий.

Для ориентации по дереву директорий ввелось понятие *полного имени файла* и *относительного имени файла*.

Полное имя файла — название директорий на пути из корневой директории и название файла. Относительное — директории по пути из рабочей директории до выбранного файла и имя этого файла.

Удаление файла требует оповещения об этом его директории. Также нужна операция создания пустой директории.

Ациклический граф.

Появляется операция “создать жёсткую связь”. Операция удаления меняется. Удаляется только запись о файле, а сам он остаётся, пока на него указывает хотя бы одна связь.

Граф “общей структуры”

Появляется операция “создать символическую связь” (по сути является относительным путём).

Операционная система знает только три вида файлов: регулярный файл, директория, файл типа связь.

6.2 Монтирование файловых систем

Совмещение нескольких файловых систем. В одной из систем ищется свободная запись, её делают точкой монтирования и она начинает ссылаться на другую систему. Смонтированные системы можно размонтировать обратно.

Лекция 25 ноября

7 Файлы и файловые системы (часть 2)

Физические файлы и файловая система

Понятие о File Control Block

Для корректной работы с файлом необходимо иметь информацию о:

1. атрибутах файла: тип файла (директория или регулярный файл), времена создания, модификации, последнего обращения к файлу, кто является хозяином файла, какие права доступа имеют к файлу другие пользователи, размер файла и т.д.
2. Расположении файла на внешнем носителе информации.

Эти данные мы будем считать лежащими в блоке управления файлами (File Control Block).

Где может храниться блок управления файлом:

1. В записи директории, содержащей имя файла.
2. Частично в директории, а частично совместно с файлом.
3. Частично в директории, а частично в отдельной структуре.
4. Отдельно от директории и от файла.

Способ хранения определяется реализацией физической файловой системы.

7.1 Организация вторичной памяти

По физическим причинам обмен информацией со вторичной памятью не может осуществляться побайтно. Единицей обмена данными служит сектор или страница, содержащие большое количество байт (минимальное количество информации, которое можно прочитать или записать за раз).

Некоторая часть секторов (с известным расположением, обычно в начале вторичной памяти) выделяется под заголовок физической файловой системы. Оставшиеся сектора объединяются в логические блоки, имеющие последовательную нумерацию.

В заголовке хранятся: тип файловой системы, размер заголовка, количество логических блоков в файловой системе, размер блока в байтах, начальный блок корневой директории и другое.

Информация, содержащаяся в файлах, хранится в логических блоках вторичной памяти.

Логические блоки могут использоваться и для хранения информации из FCB.

7.2 Методы выделения блоков

7.2.1 Непрерывное выделение

FCB хранится в записи директории, содержащей имя файла. Директория “/”.

a	размер файла и другие атрибуты	1
b	размер файла и другие атрибуты	4
c	размер файла и другие атрибуты	7

Всем методам выделения блоков присуща внутренняя фрагментация. При непрерывном выделении блоков чтение информации из файла происходит очень быстро (с точки зрения физики сделана удобная реализация).

Недостатки:

- Проблемы с размещением файла при увеличении его размеров.
- Непрерывному выделению свойственна внешняя фрагментация.
- Метод аналогичен схеме управления памятью с динамическими разделами. Для выделения блоков есть стратегии first fit, best fit, worst fit. (При выборе места для нового размещения файла)

Чаще всего применяется для создания read-only файловых систем, ведь данные в ней не меняются и их не нужно перемещать.

7.2.2 Связный список

В конце сектора хранятся указатели на начало следующего сектора, в конце последнего сектора находится end of file (какая-то пометка).

FCB хранится частично совместно с файлом, а частично — в директории.

a	размер файла и другие атрибуты	1
b	размер файла и другие атрибуты	4
c	размер файла и другие атрибуты	2

Метод свободен от внешней фрагментации.

Недостатки:

- Использование связного списка по сути сводит прямой доступ к файлу к последовательному доступу.
- Дополнительные потери дисковой памяти из-за хранения указателей в блоках.
- В случае порчи устройства памяти (размагничивание и т.п.) при потере начала файла мы теряем сразу всю информацию о файле, ведь мы не знаем, как читать оставшееся.

Не имеет прямых аналогов в схемах управления оперативной памятью. В прямом виде метод связного списка почти не применяется, однако применяется его модификация.

7.2.3 Список с FAT

Введём некую структуру FAT (File Allocation Table) в ней столько же полей, сколько находится логических блоков. (существуют разновидности FAT-8, FAT-16, FAT-32 — количество байт, выделяющихся под одну запись). Эта структура (по сути таблица) хранится в заголовке.

В директории хранится размер файла и другие атрибуты, указатель на начало файла и на первую его запись в FAT. Преимущество в том, что FAT можно целиком загрузить в оперативную память.

Получается, что FCB хранится частично в директории, а частично — в отдельной структуре.

Метод свободен от внешней фрагментации. Однако, если потеряется FAT, то мы потеряем доступ ко всем файлам в файловой системе. Для повышения надёжности хранится несколько копий FAT. Если читаемых копий остаётся мало, система кидает грозные предупреждения о возможной потере всех данных диска.

Не имеет прямых аналогов в схемах управления оперативной памятью.

7.2.4 Метод прямой индексации

Создаём условный массив, в котором по индексу i хранится номер логического блока, в котором находится i -я часть файла.

В директории хранятся размер файла и другие атрибуты, а также ссылку на индексы (называется индексным блоком или блоком косвенной адресации).

В схемах управления оперативной памятью аналогом является страничная память.

Однако ловкие программисты придумали в начало индексного блока записывать размер файла и другие атрибуты, назвали эту штуку индексным узлом. Хранятся эти узлы в заголовке

Теперь в директории достаточно хранить только ссылку на индексы.

7.2.5 Многоуровневая индексация

Храним ссылку на блок двойной косвенной адресации, в индексном блоке храним адреса индексных блоков и уже в них храним ссылки на блоки данных.

7.2.6 Смешанная индексация

У каждого файла есть свой индексный узел, размер которого фиксирован.

В индексном узле хранятся:

Атрибуты файла, некоторое количество байт выделяем под записи прямой адресации, они напрямую ссылаются на блоки с данными (для удобного хранения маленьких файлов до 40 КБ). Далее хранится адрес блока косвенной адресации (позволяет хранить файлы размером 4 МБ + 40 КБ с предыдущего). Далее — адрес блока двойной адресации (файлы до 4 ГБ + предыдущие). Далее — адрес блока тройной адресации (файлы до 4 ТБ + предыдущие).

8 Файловая система как часть ОС

Файловая система как часть ОС - это программные средства, обслуживающие физическую файловую систему.

8.1 Функции файловой подсистемы ОС

8.1.1 Распределение внешней памяти между файлами. Учёт занятого и свободного пространства внешней памяти.

1. Файловая подсистема ОС осуществляет первоначальную разметку внешнего носителя информации — форматирование диска.
2. Файловая подсистема ОС при создании или модификации файла осуществляет выделение (при необходимости) свободных логических блоков и модификацию содержимого FCB.
3. Файловая подсистема ОС ведёт учёт свободных логических блоков:
 - С помощью битового вектора.
 - С помощью связного списка. (ссылка на него в заголовке, в нём адреса свободных блоков, в конце — адрес следующего блока с адресами свободных и т.д.)

8.1.2 Идентификация файлов. Связывание имени файла с его расположением на внешнем устройстве

Для связывания имени файла с его расположением файловая подсистема ОС последовательно разбирает полное или относительное имя файла.

Пример: /a/b.

1. Читаем FCB корневой директории “/” и определяем расположение файла корневой директории на диске.
2. Считываем содержимое корневой директории, находим запись, соответствующую имени “a”, читаем FCB файла “a”, убеждаемся, что это директория, и определяем расположение ее файла на диске.

3. Считываем содержимое директории “/a”, находим запись, соответствующую имени “b”, читаем FCB файла “/a/b”, определяем расположение этого файла на диске.

Требуется большое количество обращений к диску. Для повышения эффективности появляются операции open и close. Эти операции далеко не базовые, они появляются только сейчас для ускорения получения доступа к файлу.

8.1.3 Обеспечение выполнений операций над файлами и логическими файловыми системами

8.1.4 Защита от несанкционированного доступа

В основном используются права доступа к файлам, группы пользователей и права пользователей.

8.1.5 Предоставления возможности совместного использования файлов

ОС может заблокировать файл целиком или его конкретный блок во время использования его одним из пользователей

8.1.6 Обеспечение надежности и отказоустойчивости

Решается на уровне ОС, информация может дублироваться несколько раз для защиты от потери информации из-за дефектов носителя.

9 Система управления вводом-выводом

9.1 Виды деятельности компьютера

1. Обработка информации.
2. Операции ввода-вывода.

С точки зрения программиста:

Обработка информации — выполнение команд процессора над данными, находящимися в памяти, независимо от уровня иерархии.

Ввод-вывод – обмен данными между памятью и устройствами, внешними по отношению к ней и процессору.

С точки зрения ОС:

Обработка информации — выполнение команд процессора над данными, лежащими в памяти на уровнях не ниже основной памяти.

Ввод-вывод — всё остальное.

Положим, что мы рассматриваем систему с бесконечной оперативной памятью (чтобы понятия программиста и ОС об обработке информации совпадали).

1. Обработка информации
 - Что делается? Курс “Алгоритмы и алгоритмические языки”
 - Как делается? Темы 3-7 этого курса
2. Операции ввода-вывода
 - Что делается?
 - Как делается?

Эта тема

9.2 Общие сведения об архитектуре компьютера

Используются шина данных, шина управления, шина адреса.

Они соединяют процессор, память, диски, клавиатуру и монитор. Соединено это локальной магистралью (по сути проводники)

Ширина шины — количество линий проводников в шине.

9.2.1 Передача данных из процессора в память

1. На адресной шине выставить сигналы для адреса памяти.
2. На шине данных выставить сигналы для данных.
3. На шине управления выставить сигналы работы с памятью и операции записи.

9.2.2 Память и устройства I/O

Память:

- Локализована в пространстве
- Ячейки взаимно однозначно отображаются на линейное адресное пространство памяти.

Устройства I/O:

- Пространственно разнесены и подключаются к локальной магистрали (ЛМ) через порты (замечание лектора, ударение на первый слог, иначе это элемент мужской одежды) ввода-вывода.
- Порты ввода-вывода взаимно однозначно отображаются на линейное адресное пространство ввода-вывода (иногда на линейное адресное пространство памяти).

9.2.3 Передача данных из процессора в порт (адресное пространство ввода-вывода)

1. На адресной шине выставить сигналы для адреса порта ввода-вывода
2. На шине данных выставить сигналы для данных
3. На шине управления выставить сигналы работы с устройствами ввода-вывода и операции записи.

9 декабря

10 Функции базовой подсистемы

10.1 Обработка прерывания и ошибок

1. Определение устройства, выдавшего прерывание.
2. Взаимодействие с устройством.
3. Проверка успешности выполнения операции.
4. Попытка устранения возможных ошибок.
5. Определение процесса, ожидающего этого прерывания. Не успел...

Действия по обработке прерывания и компенсации ошибок могут быть частично делегированы драйверу устройства — функция `intr` (от `interrupt`) в интерфейсе драйвера.

10.2 Планирование запросов

Для блокирующихся и асинхронных системных вызовов

1. При занятости устройства запрос ставится в очередь к данному устройству.
2. После освобождения устройства необходимо принять решение: какой из запросов в очереди инициировать следующим — планирование запросов.

Действия по планированию запросов могут быть частично или полностью делегированы драйверу устройства — функция `strategy` в интерфейсе драйвера.

Строение жёсткого диска

Красивая картинка в презентации.

Параметры планирования

- Запрос полностью характеризуется: типом операции, номером цилиндра, номером дорожки, номером сектора.
- Цель планирования — минимизировать время, необходимое для выполнения запросов.
Время выполнения запроса = $\text{transfer time} + \text{positioning time}$.
 $\text{positioning time} = \text{seek time} + \text{positioning latency}$.

Единственным параметром планирования остаётся `seek time` — время пропорциональное разнице между номером цилиндра в запросе и номером текущего цилиндра.

Алгоритмы планирования запросов к жёсткому диску.

Пусть в диске 100 цилиндров (нумерация с 0), очередь запросов: 23, 67, 55, 14, 31, 7, 84, 10. Текущий цилиндр — 63.

1. Алгоритм FCFS (First Come First Served).
23 → 67 → 55 → 14 → 31 → 7 → 84 → 10. Всего перемещение на 329 цилиндров.
2. Алгоритм SSTF (Short Seek Time First).
63 → 67 → 55 → 31 → 23 → 14 → 10 → 7 → 84. Перемещений меньше (однако он не всегда лучше FCFS).
3. Алгоритм SCAN (головки постоянно двигаются).
63 → 55 → 31 → 23 → 14 → 10 → 7 → 0 → 67 → 84. Перемещение на 147 цилиндров.
4. Алгоритм LOOK (модификации SCAN, если по направлению движения нет запросов, головка останавливается).
63 → 55 → 31 → 23 → 14 → 10 → 7 → 67 → 84. Перемещение на 133 цилиндра.
5. Алгоритм C-SCAN
63 → 55 → 31 → 23 → 14 → 10 → 7 → 0 → 99 → 67 → 84. Скачок от 0 до 99 без чтения происходит быстро (особенность диска, можно быстро откуда угодно переместиться в парковочную зону).

11 Сети и сетевые операции

Для чего компьютеры объединяют в сети?

- Совместное использование ресурсов (как физических, так и информационных).
- Ускорение вычислений.
- Повышение надёжности работы.
- Общение пользователей.

Сетевые ОС

- Пользователь должен знать, что компьютер находится в сети, где находятся удалённые ресурсы и как до них добраться.
- От автономной операционной системы отличается наличием дополнительной сетевой части.

Распределённые ОС

- Пользователь работает как в автономной операционной системе.
- Операционная система только отличается от автономных ОС и функционирует (условно) на всех компьютерах.

Особенности взаимодействия удалённых процессов

- В основе взаимодействия локальных процессов лежит использование разделяемой памяти, а в основе взаимодействия удалённых процессов — передача сообщений.
- Информация между удалёнными процессами зачастую передаётся через процессы-посредники, обитающие на компьютерах, отличных от компьютеров получателя и отправителя.
- Удалённое взаимодействие должно строиться исходя из первоначального предположения о ненадёжности связи.
- Адреса взаимодействующих процессов должны быть уникальны в рамках всей сети.
- Необходимо обеспечить взаимоисключения процессов при обращении к линии связи между компьютерами.

Логическая организация связи

- Как прокладывать физические линии связи (стратегия размещения или топология, а не буквально местоположение)?
- Как исключить возникновение race condition при их использовании?
- Какие виды интерфейсов могут быть предоставлены пользователю?
- Как организована адресация удалённых процессов?
- Как информация передаётся от отправителя к получателю через узлы посредники?

11.1 Понятие протокола

Для обмена почтовыми сообщениями между людьми соответствующие службы связи должны договориться о:

- Виде почтовых сообщений
- Содержании служебной информации
- Формате представления служебной информации

Для взаимодействия удалённых процессов сетевые части операционных систем также должны руководствоваться определёнными правилами.

11.2 Многоуровневая модель

Самый высокий уровень — пользовательский интерфейс, самый низкий — Hardware. Каждый уровень пользуется функционалом уровня ниже и предоставляет свои функции вышележащему уровню.

Формальный перечень правил, определяющих последовательность и формат сообщений, которыми обмениваются сетевые компоненты вычислительных систем, лежащие на одном уровне называется **сетевым протоколом**.

Вся совокупность вертикальных и горизонтальных протоколов (интерфейсов и сетевых протоколов) называется **семейством протоколов** или **стеком протоколов**.

11.3 Модель OSI/ISO

Уровни передачи информации:

1. Физический
2. Канальный
3. Сетевой
4. Транспортный
5. Сеансовый (не везде есть)
6. Представления
7. Приложений