

Лекция № 2

Файловая система и ввод-вывод информации. Организация хранения данных. Архивация данных.

В рамках операционных систем и сред одной из важнейших тем является эффективная организация хранения и обработки данных. Рассмотрим: файловые системы, механизм ввода-вывода информации, а также принципы и методы архивации данных. Все эти компоненты имеют решающее значение для эффективной работы с данными в современных компьютерных системах.

1. Файловая система

Файловая система (ФС) — это механизм для управления данными на устройствах хранения (жестких дисках, SSD, флеш-накопителях и других). Она определяет способ, которым файлы и каталоги (папки) организуются, хранятся и извлекаются на устройствах.

Основные функции файловой системы

- Организация хранения данных: Файловая система определяет, как данные структурируются на устройстве (например, как данные разбиваются на блоки).
- Путеводитель (метаданные): Файловая система управляет метаданными файлов, такими как имя файла, его расположение, права доступа, временные метки и т.д.
- Доступ к данным: обеспечивает механизмы для доступа к данным - чтение, запись, удаление файлов.
- Управление правами доступа: Файловая система регулирует, какие пользователи и процессы могут работать с конкретными файлами и каталогами.

Типы файловых систем

- FAT (File Allocation Table): Простая файловая система, использующая таблицу для отслеживания расположения файлов. Применяется на устройствах с малым объёмом памяти (флешки, карты памяти).
- NTFS (New Technology File System): Более сложная файловая система, используемая в Windows. Поддерживает большие объёмы данных, сложную структуру метаданных и более гибкое управление правами доступа.
- ext3/ext4: Используются в Linux. Они обеспечивают высокий уровень надёжности и поддержку больших объёмов данных.
- HFS+ (Hierarchical File System Plus): Файловая система, используемая в macOS, также с поддержкой метаданных и большого объёма хранения.

Структура файловой системы

Файловая система организована в иерархическую структуру, включающую:

- Файлы: хранят данные.
- Каталоги (папки): хранят ссылки на файлы или другие каталоги.
- Сектора и блоки: на физическом уровне данные разделены на блоки (или кластеры), которые могут иметь фиксированный размер.
- Индексы и дескрипторы: Метаданные, описывающие файлы (например, путь, права доступа, размер, дата последнего изменения).

2. Механизм ввода-вывода

Механизм ввода-вывода отвечает за взаимодействие операционной системы с внешними устройствами, такими как диски, клавиатуры, мыши, принтеры и т.д. Он реализует абстракцию над аппаратным обеспечением и позволяет программам работать с устройствами через стандартные интерфейсы.

Типы устройств ввода-вывода

Блочные устройства: Это устройства, работающие с данными блоками фиксированного размера (например, жёсткие диски, SSD). Данные можно читать и записывать блоками.

Потоковые устройства: Эти устройства обрабатывают данные по одному элементу (например, клавиатура или принтер). Взаимодействие с ними обычно более линейное.

Принципы работы с вводом-выводом

- **Буферизация:** для повышения производительности данные часто сохраняются во временных буферах перед отправкой в устройство или после получения из него.

- **Асинхронность:** Операции ввода-вывода могут выполняться асинхронно, что позволяет процессу продолжать выполнение, не ожидая завершения операции.

- **Системы очередей:** Операции ввода-вывода часто обрабатываются с использованием очередей. Например, несколько процессов могут запросить доступ к диску, и система будет обрабатывать эти запросы в определённом порядке (например, с использованием алгоритмов планирования).

Модели ввода-вывода

- **Блочный ввод-вывод:** Операции с данными происходят с блоками фиксированного размера, например, чтение или запись файлов.

- **Символьный ввод-вывод:** Операции с устройствами, такими как клавиатура, принтеры или последовательные порты.

Технологии ввода-вывода

- Директ- доступ (DMA, Direct Memory Access): Технология, позволяющая устройствам передачи данных напрямую в память, минуя процессор, что ускоряет операцию ввода-вывода.
- I/O Scheduling: Алгоритмы планирования операций ввода-вывода, например, SCAN или C-SCAN, помогают улучшить производительность за счёт оптимизации порядка обработки запросов на доступ к диску.

3. Организация хранения данных

Хранение данных в компьютерных системах требует эффективной структуры, чтобы обеспечить доступность, надежность и производительность.

Логическое и физическое хранение данных

- Логическое хранение: это представление данных на уровне операционной системы, где пользователи работают с файлами и каталогами.
- Физическое хранение: это непосредственно расположение данных на устройствах хранения, таких как жесткие диски или SSD. На этом уровне также важна оптимизация доступа к данным (например, разбиение на блоки).

Управление пространством хранения

Операционная система должна эффективно управлять физическим пространством хранения, выделяя и освобождая блоки данных. Это включает:

- Размещение файлов: как файлы записываются на диске.
- Фрагментация: при повторной записи данных могут возникать разрывы в хранилище, что снижает эффективность доступа к данным. Важным аспектом является дефрагментация.
- Алгоритмы размещения: например, последовательное, цепочечное или индексированное размещение.

Резервное копирование и восстановление данных

Процесс резервного копирования данных крайне важен для обеспечения их сохранности и восстановления в случае потерь или повреждений.

4. Архивация данных

Архивация данных — это процесс сжатия и хранения данных в одном или нескольких файлах для экономии пространства, а также для удобства их переноса или длительного хранения.

Методы архивации

- Без потерь: Сжатие данных без потерь информации. Примеры: ZIP, GZIP, RAR.
- С потерями: Сжатие данных с потерей части информации, например, сжатие изображений или аудио (JPEG, MP3).

Программы и инструменты архивации

- tar и gzip в Linux/Unix системах — популярные инструменты для создания и распаковки архивов.
- WinRAR, 7-Zip — стандартные программы в Windows для работы с архивами.

Задачи архивации

- Экономия пространства: Архивация позволяет значительно уменьшить объём данных.
- Передача данных: Сжатие данных облегчает их передачу по сети.
- Долговременное хранение: Архивы часто используются для хранения данных на длительный срок.