

Лекция 9. Физические файловые системы

Операционные системы

1 января 2017 г.

Рассматриваемые темы

Уровни организации файловых систем

- Физическая организация данных на внешних носителях.
- Разбиение дисков на тома.
- Физические файловые системы.
- Виртуальные файловые системы.

Концепция виртуальной файловой системы

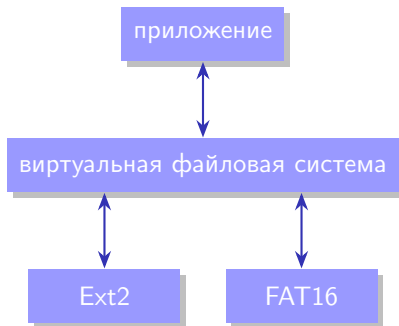


Рис. 1: взаимодействие приложения с файлами

Адресация данных на аппаратном уровне

Определения

Сектор: (*sector*) — аппаратно адресуемый блок данных носителя (512 байт).

Кластер: (*cluster*) — адресуемый блок, используемый файловой системой, содержит несколько секторов.

Метаданные: (*metadata*) — данные, хранящиеся в томе, необходимые для поддержания работы файловой системы (не принадлежат пользовательским данным в файлах).

Способы адресации секторов

- Cylinder-Head-Sector (CHS).
- Logical Block Address (LBA).

Разделы дисков

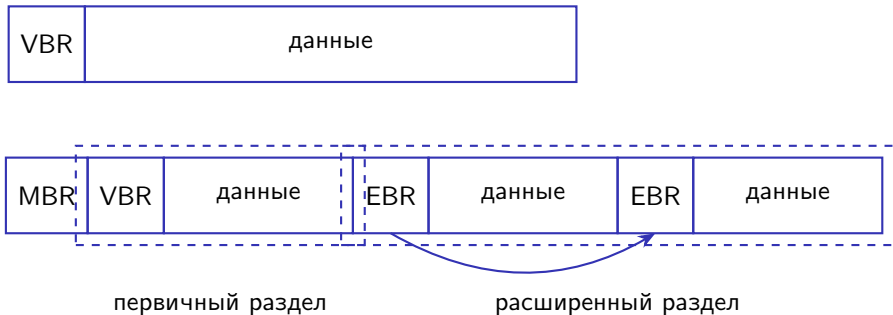


Рис. 2: организация разделов физического диска

Дескрипторы разделов

Определения

Главная загрузочная запись: (*Master Boot Record, MBR*) — находится в первом физическом секторе (секторах), содержит:

- код загрузки (*bootstrap code*);
- таблицу разделов (*partition table*) — до 4;
- сигнатуру;
- ...

Загрузочная запись тома: (*Volume Boot Record, VBR*) — находится в начале неразделённого диска или раздела диска, содержит код загрузки.

Расширенная загрузочная запись: (*Extended Boot Record, EBR*) — находится в начале логического раздела, структура, как у MBR, но запись для *n* раздела указывает на следующую EBR.

Таблица разделов GUID

Определение

Таблица разделов GUID: (*GUID Partition Table, GPT*) — стандарт формата размещения разделов на диске.

Особенности

- Открытый стандарт (часть Unified Extensible Firmware Interface — UEFI);
- Ограничение на размер диска: 8 Збайт - 512 байт (у MBR: 2 Тбайт - 512 байт);
- Заголовок GPT и таблица разделов хранятся в начале и конце диска;
- «Защитный MBR» в начале диска;
- Таблица разделов содержит контрольные суммы (CRC32) и GUID диска;
- Элементы таблицы разделов содержат GUID типов разделов и разделов.

FAT (12, 16, 32)

Система	Кластеров	Размеры кластера	Макс. том	Макс. файл
FAT12	4 096	512 байт – 8 Кбайт	32 Мбайт	—
FAT16	65 536	512 байт – 64 Кбайт	4 Гбайт	4 Гбайт
FAT32	268 435 456	4 Кбайт – 64 Кбайт	16 Тбайт	4 Гбайт
exFAT	268 435 456	4 Кбайт – 32 Мбайт	16 Эбайт	4 Гбайт

Таблица 1: версии файловой системы FAT

Организация FAT

загрузочный сектор (VBR)	сектор информации о ФС (FAT32)	резерв (опционально)	таблица размещения файлов	таблица размещения файлов	корневой каталог (FAT12, 16: 256 элементов)	остальные каталоги и файлы
--------------------------	--------------------------------	----------------------	---------------------------	---------------------------	---	----------------------------

Рис. 3: организация метаданных системы FAT

Запись FAT

Информация записи таблицы размещения файлов (file allocation table)

- Номер следующего кластера цепочки;
- Отметка конца цепочки ($0xFF8 \dots 0xFFF$, $0xFFFF$, $0xFFFFFFFF$);
- Отметка повреждённого кластера ($0xFF7$, $0xFFF7$, $0xFFFFFFFF7$);
- Отметка неиспользуемого кластера (0).

Запись каталога

Информация записи таблицы каталогов (directory table) — 32 байт

- Короткое имя «8.3»
(DATA____.TXT), 1 символ имеет особые значения;
- Атрибуты:
 - «только для чтения» (read only);
 - «скрытый» (hidden);
 - «системный» (system);
 - «метка тома» (volume label);
 - «подкаталог» (subdirectory);
 - «архивный» (archive);
- Дата/время:
 - создания (до 10 мс);
 - последнего доступа (дата);
 - последнего изменения файла (до 2 с);
- Номер кластера начала файла;
- Размер файла в байтах (32 бит).

Таблица размещения файлов и каталог

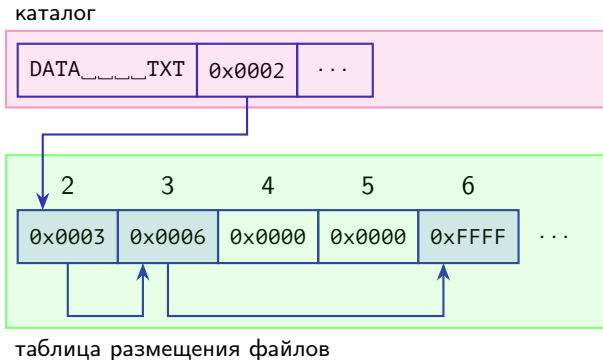


Рис. 4: таблица размещения файлов и каталог (FAT16)

Удаление файла в FAT

Последовательность действий при удалении

- 1 В записи каталога: первому символу имени присваивается $0xE5$ или $0x05$ ($0x00 \Rightarrow$ все записи каталога дальше игнорируются);
- 2 В таблице размещения: цепочка записей файла обнуляется.

Сохраняется в записи каталога

- размер файла;
- Номер кластера начала файла.

Удаление файла в FAT

Последовательность действий при удалении

- 1 В записи каталога: первому символу имени присваивается $0xE5$ или $0x05$ ($0x00 \Rightarrow$ все записи каталога дальше игнорируются);
- 2 В таблице размещения: цепочка записей файла обнуляется.

Сохраняется в записи каталога

- размер файла;
- Номер кластера начала файла.

Условия восстановления файла

- Кластеры располагались последовательно;
- Данные в кластерах не были перезаписаны.

Загрузочный сектор (VBR)

Состав загрузочного сектора

- Загрузочный код;
- Структура BIOS Parameter Block (BPB);
- ...

BIOS Parameter Block (BPB)

Информация в BIOS Parameter Block/Extended BIOS Parameter Block (EBPB)

- Байт/логический сектор (степень 2);
- Логических секторов/кластер;
- Количество зарезервированных логических секторов перед FAT1;
- Количество FAT (обычно 2, RAM Disk — 1);
- Максимальное количество записей корневого каталога, 0 для FAT32;
- Количество логических секторов;
- ...

BIOS Parameter Block (BPB, окончание)

Информация в Extended BIOS Parameter Block (EBPB) для FAT32

- Номер кластера корневого каталога для FAT32, обычно 2;
- Номер логического сектора информации о ФС, обычно 1;
- Логических секторов/FAT;
- Идентификатор тома (серийный номер);
- Метка раздела тома;
- Тип файловой системы.
- ...

FS Information Sector (FAT32)

Состав сектора информации о ФС (FS Information Sector)

- Последнее известное количество свободных кластеров тома (0xFFFFFFFF — неизвестно);
- Последний известный номер выделенного кластера под данные (0xFFFFFFFF — неизвестно \Rightarrow 2);
- ...

Первые две записи FAT

Запись 0: идентификатор FAT;

Запись 1: два старших бита:

- том не был корректно извлечён (\Rightarrow FS Information Sector не нужно использовать);
- обнаружены проблемы ввода/вывода диска при запуске.

Поддержка длинных имён

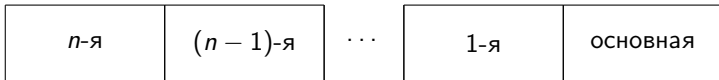


Рис. 5: записи таблицы каталога с длинными именами VFAT ($n \leq 20$, до 255 символов)

Состав дополнительной записи каталога

- 1 символ имени ($0xE5 \Rightarrow$ удалено): флаг «последняя», номер;
- Атрибуты: «метка тома», «системный», «скрытый», «только чтение»;
- Первый кластер: 0;
- Символы в UCS-2 (до 13) — перемежаются с остальными полями.

Особенности системы FAT

Преимущества

- широкая распространённость;
- простота реализации;
- высокая скорость на небольших носителях;
- малый объём метаданных.

Недостатки

- ограничения по размеру томов и файлов;
- невысокая надёжность;
- нагрузка на небольшую часть диска;
- фрагментация;
- отсутствие поддержки безопасности и других возможностей;
- патенты (FAT32, длинные имена).

Внешняя фрагментация

Проблемы

- Поиск файлов в больших подкаталогах;
- Вычисление объёма свободного места.

Способы борьбы

- Карты свободных кластеров (HPFS, exFAT);
- Связные списки свободных кластеров (Unix);
- Хранение количества свободных кластеров;
- Выделение кластеров после последнего выделенного (MS-DOS 3.0+);
- Кэширование позиций последних открытых файлов (MS-DOS 3.3);
- Заблаговременное выделение кластеров (Windows NT);
- Хранение древовидных описаний используемых кластеров в памяти;
- Кэширование чтения/записи.

NTFS

Особенности системы NTFS

- индексирование файлов по атрибутам;
- защита файлов и каталогов;
- дисковые квоты;
- сжатие файлов;
- жёсткие связи;
- символные ссылки (точки соединения);
- шифрование;
- теньевые копии старого содержимого файлов;
- групповые транзакции.

Журналируемая файловая система

Определение

Журналируемая файловая система: (*journaling file system*) — ФС, ведущая журнал транзакций перед их исполнением в выделенной области диска в циклическом буфере.

Примеры

- NTFS;
- Ext3;
- Ext4 (отключается).