



ЛОГИЧЕСКАЯ ОРГАНИЗАЦИЯ ФАЙЛОВОЙ СИСТЕМЫ

Операционные системы и среды

Логическая организация файловой системы

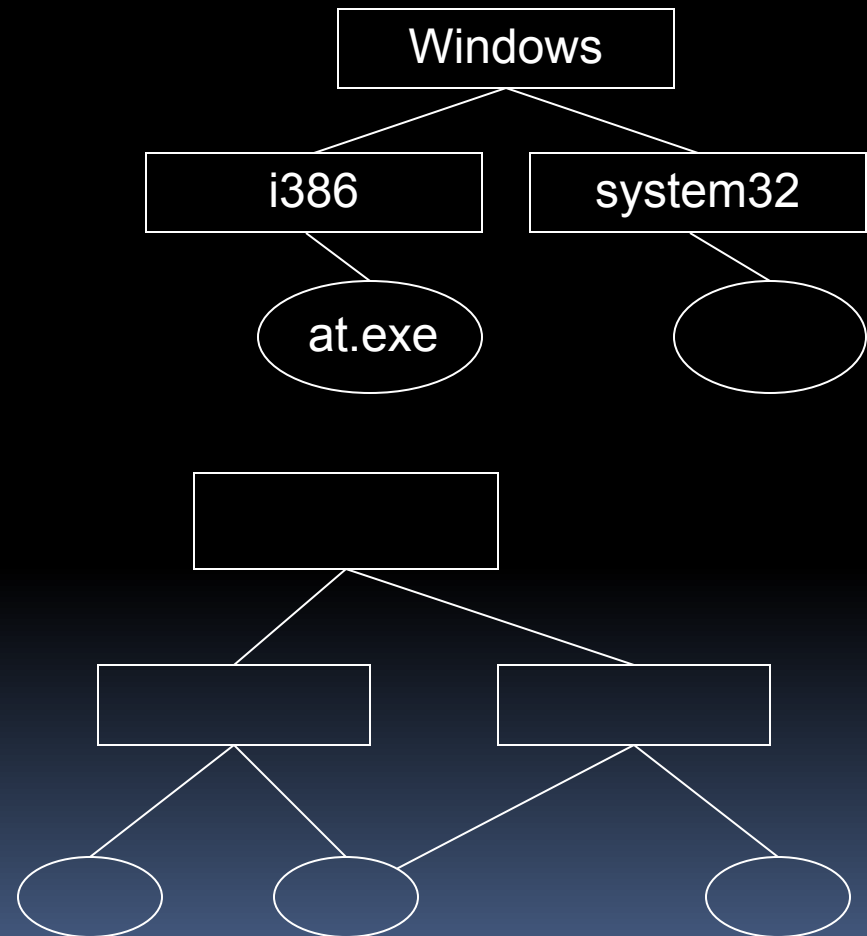
- Одна из основных задач ОС – предоставление удобного пользовательского интерфейса при работе с данными, хранящимися на носителях. Логическая модель в рамках ОС подменяет физическую модель размещения данных на носителях.
- **Файл** – именованная область внешней памяти, в которую могут записываться и откуда могут считываться данные. Применение файлов позволяет решить следующие задачи:
 - долговременное хранение информации;
 - совместное использование информации.

Файловая система

- **Файловая система** – часть ОС, включающая:
 - совокупность всех файлов на дисках;
 - наборы структур данных, используемых для управления (каталоги, дескрипторы файлов, таблицы распределения свободного и занятого пространства);
 - комплекс системных программных средств, реализующих операции над файлами (создание, удаление, чтение, запись, именование и поиск файлов).
 - В многопользовательских системах добавляются функции по обеспечению защиты данных от несанкционированного доступа.
- Файловые системы поддерживают несколько функционально различных типов файлов:
 - обычные файлы;
 - каталоги;
 - ссылки;
 - именованные конвейеры;
 - файлы, отображаемые на память и т.д.

Иерархическая структура файловой системы

- Пользователи обращаются к файлам по их символьным именам. Для удобства пользователя логическая структура хранения данных представляет иерархическую структуру.
- Граф, описывающий структуру файловой системы может представлять собой **дерево** или **сеть**.
- В Windows используется древовидная организация, в UNIX – сетевая.



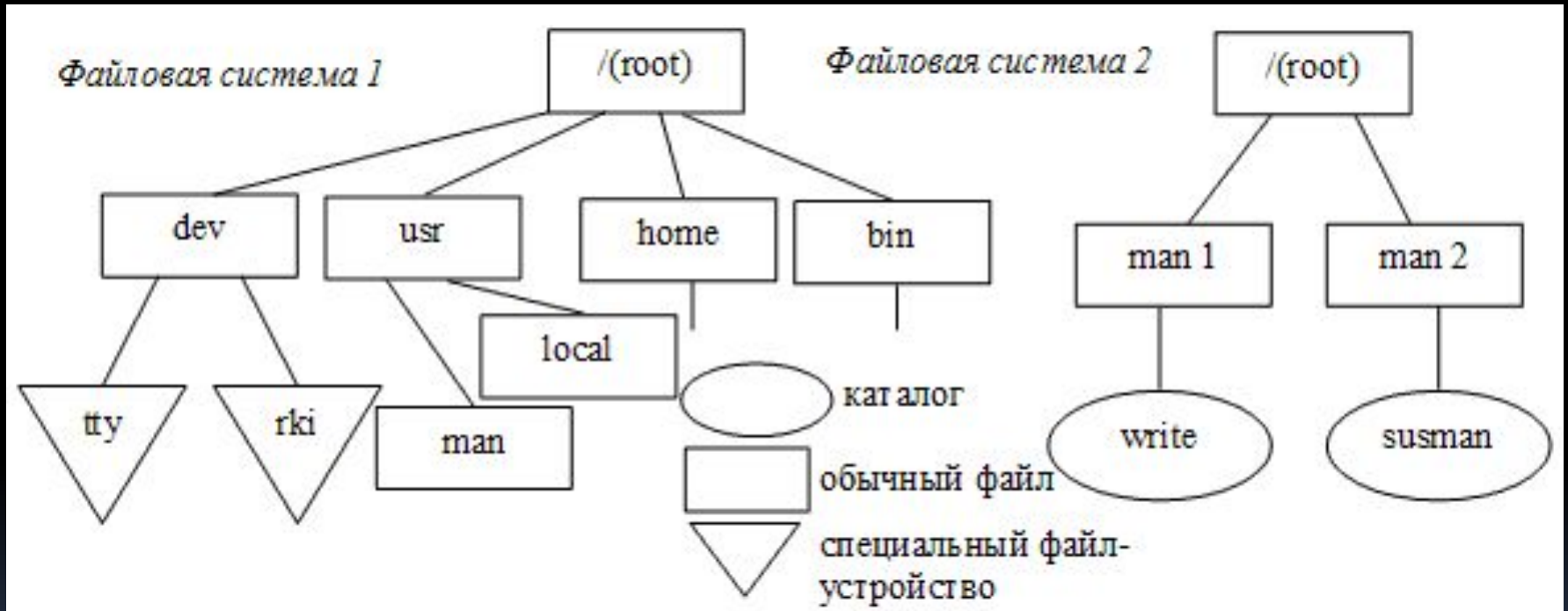
Имена файлов

- Каждый файл имеет некоторое символьное имя. В иерархических системах выделяют три типа имен файлов:
 - простое (имя файла в пределах одного каталога);
 - полное (цепочка простых символьных имен всех каталогов, через которые проходит путь от корня до файла);
 - относительное (имя включает имена каталогов, через которые проходит маршрут от текущего каталога к искомому файлу).
- В различных операционных системах есть свои ограничения на использование символов при присвоении имени, а также на длину относительного и полного имени файла.

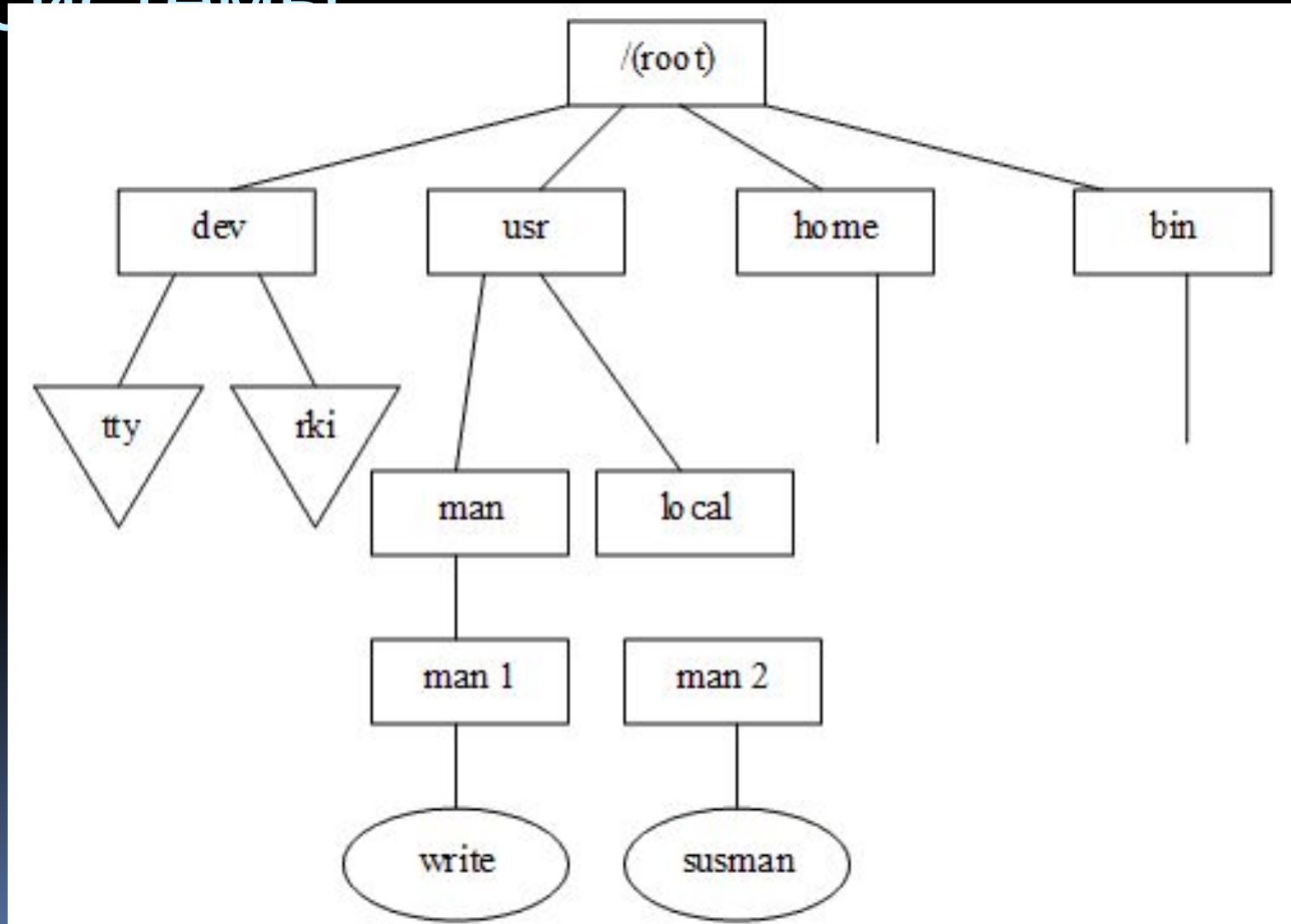
Монтирование файловой системы

- В общем случае вычислительная система может иметь несколько устройств внешней памяти. Для обеспечения доступа к данным, хранящимся на разных носителях используются два подхода.
 - На каждом устройстве размещается автономная файловая система, со своим деревом каталогов (например, в MS-DOS накопители нумеруются a:, c: и т.д.).
 - Монтирование файловой системы – операция объединения файловых систем в единую файловую систему (например, в операционных системах семейства UNIX).

Монтирование файловой системы



Монтирование файловой системы



Атрибуты файлов

Атрибут – информация, описывающая некоторые свойства файла, например:

- Тип файла
- Владелец файла
- Создатель файла
- Пароль для доступа к файлу
- Информация о разрешенных операциях к файлу
- Время создания, последнего доступа и модификации файла
- Признак «только для чтения»
- Признак «скрытый файл»
- Признак «системный файл»
- Признак «двоичный/символьный файл»
- Признак «временный файл»
- Признак блокировки
- Длина записи в файле
- Др.

Логическая организация файла

- В общем случае данные, хранящиеся в файле, имеют некоторую логическую структуру (формат хранения данных). Поддержание структуры данных в файле возлагается либо целиком на приложение, либо часть функций на файловую систему.
- Неструктурированная модель файла широко используется в большинстве современных ОС.
- Структурированный файл рассматривается ОС как упорядоченная совокупность логических записей. Развитием данного подхода являются системы управления базами данных (СУБД).

Логическая организация файла

- **Логическая запись** является наименьшим элементом данных, которым может оперировать программист при организации обмена с внешним устройством.

Файловая система может использовать два способа доступа к логическим записям:

- **последовательный доступ** - читать или записывать логические записи последовательно ;
- **прямой доступ** - позиционировать файл на запись с указанным номером.

Логическая организация файла

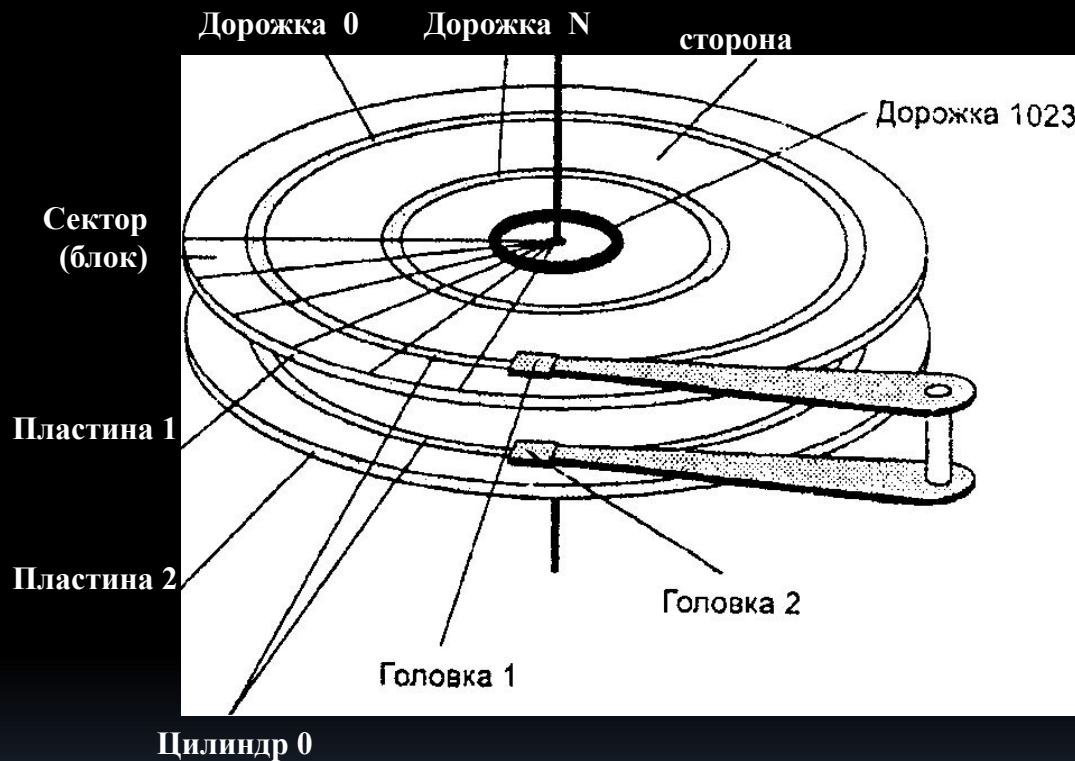
Файлы, доступ к записям которых осуществляется последовательно, по номерам позиций, называются неиндексированными, или последовательными.

- Индексированные файлы допускают более быстрый прямой доступ к отдельной логической записи, которая имеет одно или более ключевых (индексных) полей и могут адресоваться путем указания значений этих полей.

ФИЗИЧЕСКАЯ ОРГАНИЗАЦИЯ ФАЙЛОВОЙ СИСТЕМЫ

Операционные системы и среды

Схема устройства жесткого диска



Диск – одна или несколько стеклянных или металлических пластин, образующих пакет.
Дорожки – тонкие концентрические кольца (от 0 до N).

Цилиндр – совокупность дорожек одного радиуса на всех поверхностях всех пластин.

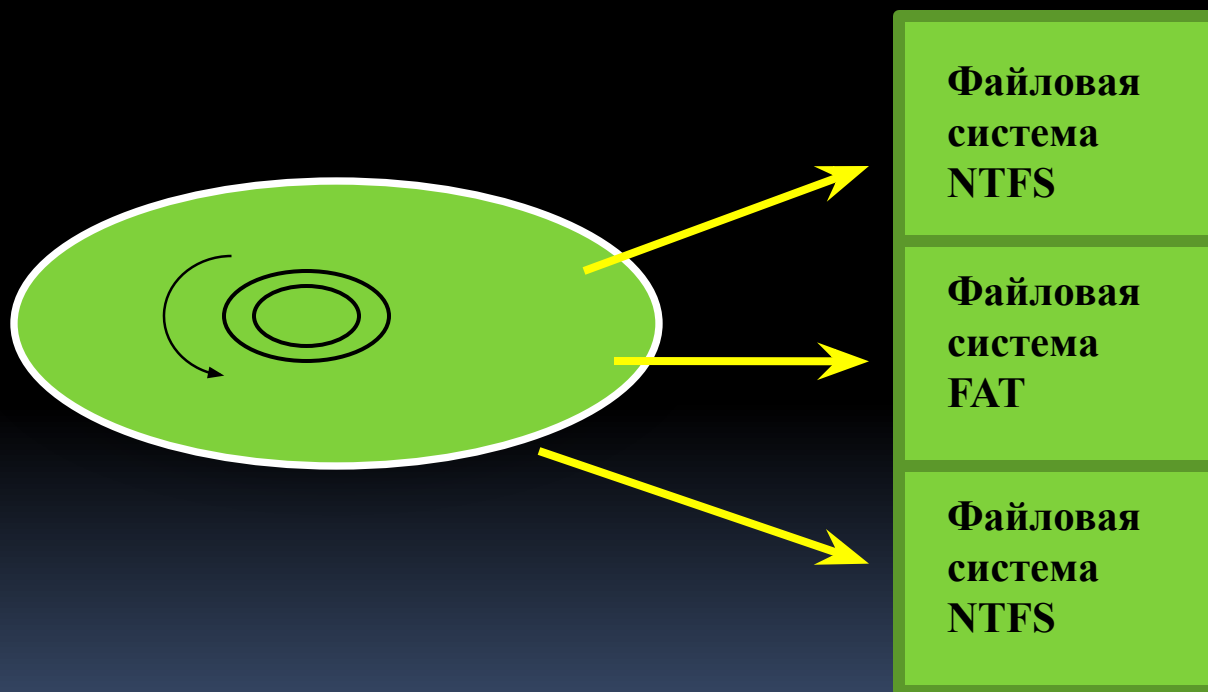
Сектор (блок) – фрагмент дорожки. Все дорожки имеют равное число секторов, в которые можно максимально записать одно и то же число байт.

Размер сектора – 512 байт.

Сектор – наименьшая адресуемая единица обмена данными дискового устройства с оперативной памятью. Для того чтобы контроллер мог найти на диске нужный сектор, необходимо задать ему все составляющие адреса сектора: номер цилиндра, номер поверхности и номер сектора.

Операционная система при работе с диском использует собственную единицу дискового пространства, называемую **кластером**. При создании файла место на диске ему выделяется кластерами. Дорожки и секторы создаются в результате выполнения процедуры физического, или низкоуровневого форматирования диска.

Разметку диска под конкретный тип файловой системы выполняют процедуры **высокоуровневого**, или **логического, форматирования**.



Раздел – это непрерывная часть физического диска, которую ОС представляет пользователю как **логическое устройство (логический диск)**. На разных логических устройствах одного и того же физического устройства могут располагаться файловые системы разного типа.

Разбиение диска на разделы

Физическая организация и адресация файла

Важным компонентом физической организации файловой системы является *физическая организация файла*, то есть способ размещения файла на диске. Основными критериями эффективности физической организации файлов являются:

- ✓ скорость доступа к данным;
- ✓ объем адресной информации;
- ✓ степень фрагментированности дискового пространства;
- ✓ максимально возможный размер файла.

Физическая организация файла

1. Непрерывное размещение



Файлу предоставляется последовательность кластеров диска, образующих непрерывный участок дисковой памяти.

Достоинства: высокая скорость доступа, так как затраты на поиск и считывание кластеров файла минимальны. Минимален объем адресной информации - достаточно хранить только номер первого кластера и объем файла.

Недостатки: большие трудности в реализации, фрагментация

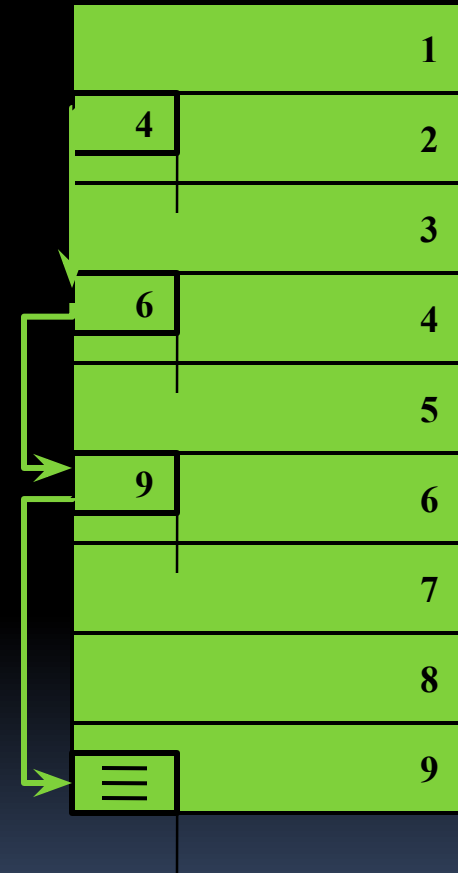
Физическая организация файла

2. Связанный список кластеров

В начале каждого кластера содержится указатель на следующий кластер. Адресная информация минимальна: расположение файла может быть задано одним числом - номером первого кластера. Каждый кластер может быть присоединен к цепочке кластеров какого-либо файла, а значит, фрагментация на уровне кластеров отсутствует.

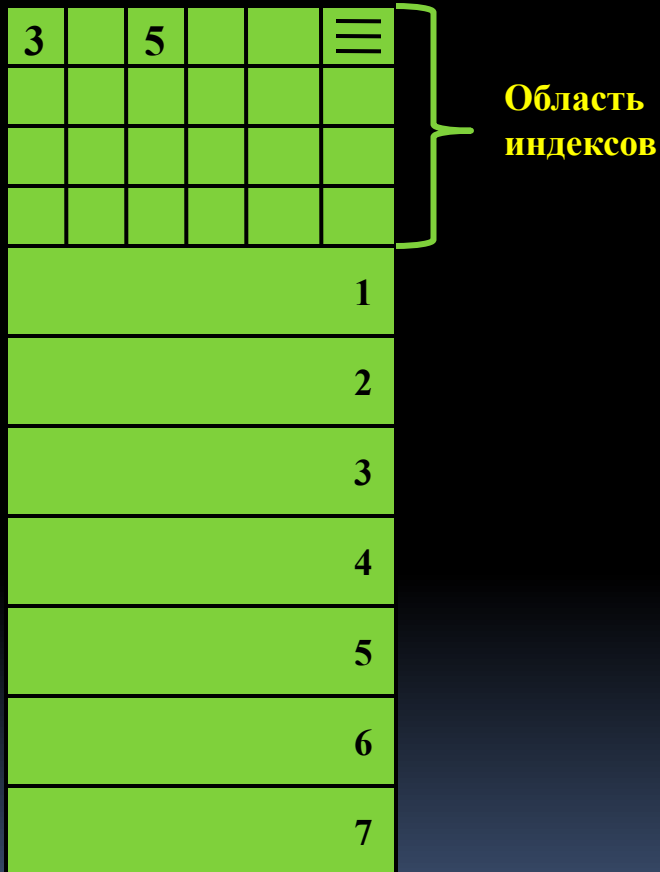
Достоинства: адресная информация минимальна.

Недостатки: сложность реализации доступа к произвольно заданному месту файла – чтобы прочитать пятый по порядку кластер файла, необходимо прочитать четыре первых кластера, проследив цепочку номеров кластеров.



Физическая организация файла

3. Связанный список индексов



Файлу выделяется память в виде связанного списка кластеров. Номер первого кластера запоминается в записи каталога, где хранятся характеристики этого файла. Остальная адресная информация отделена от кластеров файла. С каждым кластером диска связывается некоторый элемент – индекс. Индексы располагаются в отдельной области диска - в MS DOS это таблица FAT, занимающая один кластер.

Достоинства: минимальность адресной информации, отсутствие фрагментации. Доступ к произвольному кластеру файла.

Физическая организация файла

4. Перечень номеров кластеров

1
2
3
4
5
6
7
8
9
10

Файл
2,4,5,7

Перечень номеров кластеров и служит адресом файла.

Достоинства: высокая скорость доступа к произвольному кластеру файла, так как применяется прямая адресация. Отсутствует фрагментация.

Недостатки: длина адреса зависит от размера файла и для большого файла может составлять большую величину.

Физическая организация FAT



Области FAT:

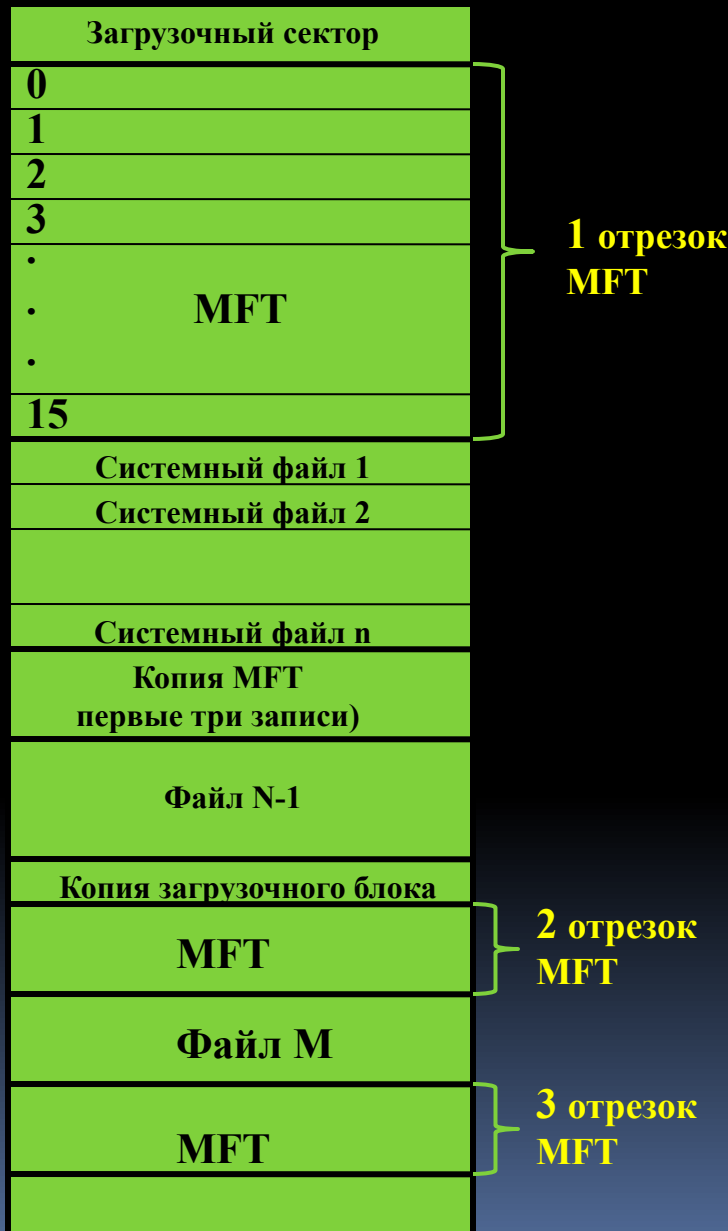
- ✓ **загрузочный сектор** содержит программу начальной загрузки ОС;
- ✓ **основная копия FAT** содержит информацию о размещении файлов и каталогов на диске;
- ✓ **резервная копия FAT**;
- ✓ **корневой каталог** занимает фиксированную область размером в 32 сектора (16 Кбайт), что позволяет хранить 512 записей о файлах и каталогах, так как каждая запись каталога состоит из 32 байт;
- ✓ **область данных** предназначена для размещения всех файлов и всех каталогов, кроме корневого каталога.

Физическая структура файловой системы FAT

Физическая организация NTFS

Основные отличительные свойства NTFS:

- ✓ поддержка больших файлов и больших дисков объемом до 2^{64} байт;
- ✓ восстанавливаемость после сбоев и отказов программ и аппаратуры управления дисками ;
- ✓ высокая скорость операций, в том числе и для больших дисков;
- ✓ низкий уровень фрагментации , в том числе и для больших дисков;
- ✓ гибкая структура, допускающая развитие за счет добавления новых типов записей и атрибутов файлов с сохранением совместимости с предыдущими версиями ФС;
- ✓ устойчивость к отказам дисковых накопителей;
- ✓ поддержка длинных символьных имен;
- ✓ контроль доступа к каталогам и отдельным файлам.



Структура тома NTFS

Основой структуры тома NTFS является **главная таблица файлов (MFT)**, которая содержит по крайней мере одну запись для каждого файла тома, включающая одну запись для самой себя. Каждая запись MFT имеет фиксированную длину, равную 2 Кбайт. Весь том NTFS состоит из последовательности кластеров, порядковый номер которого называется **логическим номером кластера (LCN)**. Файл NTFS также состоит из последовательности кластеров, порядковый номер которого внутри файла называется **виртуальным номером кластера (VCN)**.

Отрезок – базовая единица распределения дискового пространства. **Адресом отрезка NTFS** является пара **LCN** и **k** – количество кластеров в отрезке k.