

Лекция 4

Файлы. Файловая структура

Файл и файловая система

Файл – это упорядоченная совокупность данных, занимающая именованную область памяти на внешнем носителе информации.

Файл является логической единицей хранения информации и в процессе обработки рассматривается как единое целое.

Имя файла = Собственно имя . Расширение (тип файла)

В операционной системе Windows имя файла может иметь не более 255 символов (в имени файла могут использоваться латинские и русские буквы, пробелы, тире, символ подчёркивания, точка, восклицательный знак и некоторые другие символы). Расширение файла в системе Windows, как правило, не отображается.

Тип файла	Расширение
Исполняемые файлы (программы, приложения)	exe, com
Текстовые файлы	doc, txt, docx
Графические файлы	bmp, jpeg, jpg, gif
Звуковые файлы	wav, mp3
Видео файлы	avi, mov, mp4, mpg
Коды программ на языках программирования	pas, bas, html, js

Файл.

Имя и расширение (тип) файла

Примеры записи имён файлов в системе Windows:

text.txt, проба.doc, документ.docx, game.zip.com
фильм.avi, ghfd.exe, index.html, excel.exe, foto.jpg,
клип.mov, видео.avi.exe, таблица.xls,
песня.mp3.com

Правила записи имён файлов в других операционных системах могут отличаться от приведённых примеров.

Файл. Каталоги, атрибуты файла

Информация о файлах, размещённых на внешних носителях хранится в **каталогах**. Кроме имени файла (и его типа), в каталоге записывается и другая информация о файле (размер файла, дата и время создания, место хранения на диске и т.д.), а также хранится набор **атрибутов** (свойств) файла, например:

«только для чтения» (read only) – файл доступен ОС и приложениям только для чтения, т.е. в нём нельзя сохранять исправления.

«скрытый» (hidden) – файл не видим при обычном просмотре папок, для отображения файла необходимо изменить настройки отображения папок.

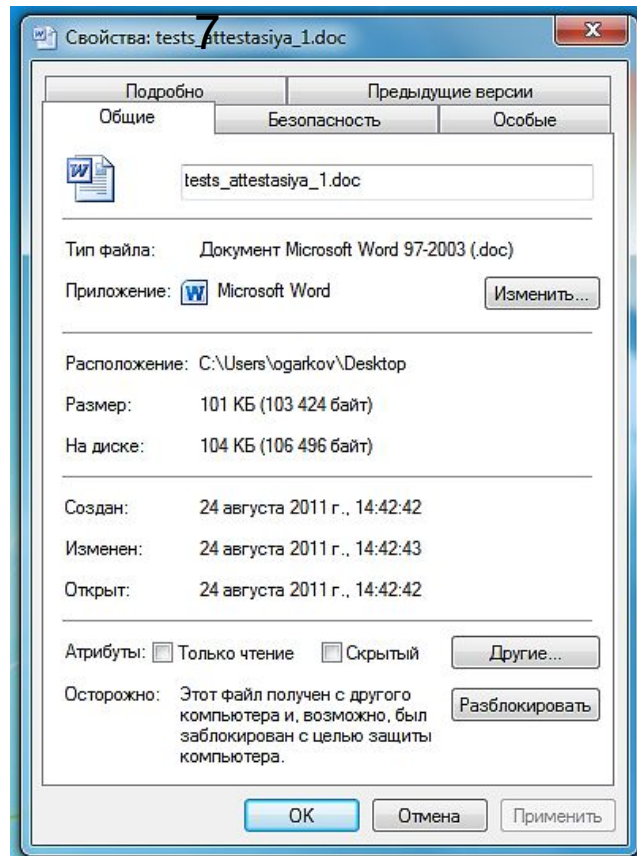
«системный» (system) – файлы относятся к системным файлам ОС, которые нельзя изменять или удалять.

«архивный» (archive) – атрибут указывает программам-архиваторам файлы, предназначенные для резервного копирования.

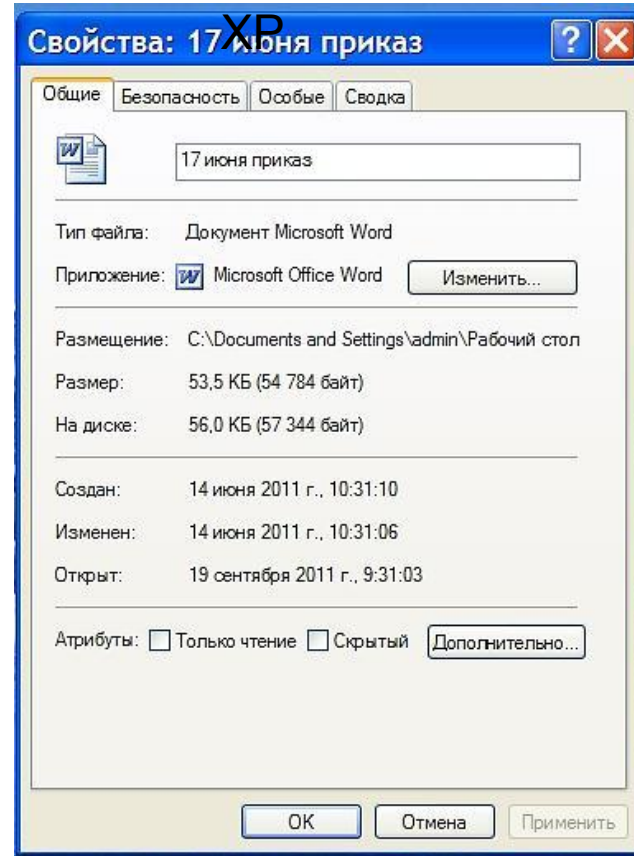
Файл. Установка атрибутов файла

Для установки (снятия) атрибутов файла, необходимо щёлкнуть по нему правой кнопкой мыши для вызова контекстного меню. Далее выбрать пункт меню Свойства. В нижней части окна Свойства установите (снимите) атрибуты файла. Для установки (снятия) других атрибутов щёлкните по кнопке Дополнительно.

Windows



Windows



Файл. Операции с файлами

В процессе работы на компьютере чаще всего приходится выполнять следующие операции с файлами:

- **копирование** - при выполнении этой операции физически создаются новый файл и новая запись в файловой системе;
- **перемещение** - при выполнении этой операции файл физически остается на диске на прежнем месте, но меняется его «адрес» в файловой системе;
- **удаление** - физически файл остается на диске (перемещается в папку Корзина), но запись о нём удаляется из файловой системы;
- **переименование** - изменяется только имя файла в файловой системе, а сам файл остается неизменным;
- **создание** - чаще всего файлы создаются программно: автоматически или при сохранении пользователем введенной в ту или иную программу информации, кроме того файл можно создать в выбранном каталоге.

Выполнять операции с файлами можно как непосредственно при помощи графического интерфейса операционной системы (контекстного меню),

так и с помощью специализированных приложений – различных *файловых менеджеров*: приложения «Проводник», программ Total Commander, Far, Windows Commander и многих других.

Файл. Поиск-сортировка файлов. Маска.

В процессе работы на компьютере часто необходимо найти (выделить) группу однотипных файлов (для копирования, перемещения, удаления, преобразования и .т.п.). При поиске (выделении) файлов можно воспользоваться **масками имен**, или **подстановочными символами**.

Поиск с использованием маски удобен также в том случае, если вы не помните точно имя файла, который надо найти.

В масках, кроме «обычных» символов можно использовать символ «?» (означает один любой символ) и символ «*» (означает любое количество или отсутствие любых символов).

Следовательно для поиска всех документов с расширением **.doc** можно указать в строке для поиска следующую комбинацию: ***.doc**.


















Маска **???.jpg** – означает, что в имени графического файла используется три любых символа. В результате поиска по такой маске будут выделены все графические файлы типа jpg с именем из трёх любых символов.

Для поиска файлов по заданному параметру (имени, типу, размеру, дате создания файла и т.п.) можно также применить сортировку файлов (по возрастанию-убыванию) в любом файловом менеджере.

















Файл. Поиск-сортировка файлов.

Для поиска файлов по заданному параметру (имени, типу, размеру, дате создания файла и т.п.) можно также применить сортировку файлов (по возрастанию-убыванию) в любом файловом менеджере в режиме Таблица.

Сортировка по дате (по возрастанию)

Имя	Дата	Тип	Размер
 1форум.jpg	21.11.2007 16:25	JPEG-рисунок	66 КБ
 zdor_semiya.jpg	14.05.2009 11:59	JPEG-рисунок	95 КБ
 zdor_semiya1.jpg	14.05.2009 11:59	JPEG-рисунок	23 КБ
 stroi-2jpg.jpg	21.05.2010 12:12	JPEG-рисунок	19 КБ
 100_3206.jpg	08.04.2011 12:49	JPEG-рисунок	23 КБ
 pp10.jpg	12.04.2011 14:17	JPEG-рисунок	3 КБ
 1300030720sred7.jpg	12.04.2011 16:10	JPEG-рисунок	4 КБ
 2410.jpg	12.04.2011 16:14	JPEG-рисунок	12 КБ
 a_b7c984b8.jpg	12.04.2011 16:17	JPEG-рисунок	3 КБ
 kosmos2.gif	25.04.2011 16:21	GIF-рисунок	38 КБ
 hochuvseznat.gif	25.04.2011 16:33	GIF-рисунок	5 КБ
 ses-1297286646-562...	25.04.2011 16:34	JPEG-рисунок	5 КБ
 forum.jpg	29.04.2011 11:09	JPEG-рисунок	3 КБ
 459135571.jpg	29.04.2011 13:44	JPEG-рисунок	41 КБ
 Ильиных Кристина...	13.05.2011 12:11	JPEG-рисунок	47 КБ
 Ильиных Кристина...	13.05.2011 12:11	JPEG-рисунок	7 КБ
 Lokeri.jpg	18.05.2011 12:41	JPEG-рисунок	806 КБ

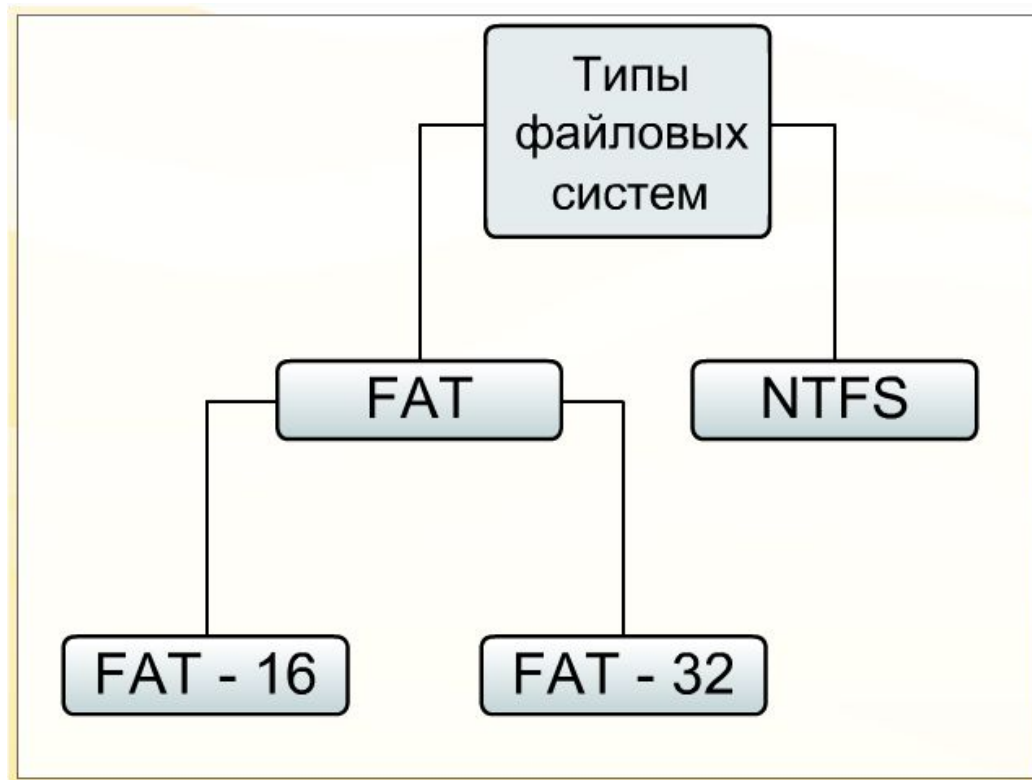
Сортировка по размеру (по убыванию)

Имя	Дата	Тип	Размер
 Lokeri.jpg	18.05.2011 12:41	JPEG-рисунок	806 КБ
 zdor_semiya.jpg	14.05.2009 11:59	JPEG-рисунок	95 КБ
 1форум.jpg	21.11.2007 16:25	JPEG-рисунок	66 КБ
 95787159-964c-4fd5...	10.06.2011 15:48	JPEG-рисунок	63 КБ
 Ильиных Кристина...	13.05.2011 12:11	JPEG-рисунок	47 КБ
 459135571.jpg	29.04.2011 13:44	JPEG-рисунок	41 КБ
 kosmos2.gif	25.04.2011 16:21	GIF-рисунок	38 КБ
 100_3206.jpg	08.04.2011 12:49	JPEG-рисунок	23 КБ
 zdor_semiya1.jpg	14.05.2009 11:59	JPEG-рисунок	23 КБ
 stroi-2jpg.jpg	21.05.2010 12:12	JPEG-рисунок	19 КБ
 untitled.png	10.06.2011 15:34	PNG-рисунок	17 КБ
 1743img.jpg	10.06.2011 16:09	JPEG-рисунок	16 КБ
 pic-all.gif	21.06.2011 14:12	GIF-рисунок	14 КБ
 08d961b75297.jpg	19.05.2011 10:14	JPEG-рисунок	13 КБ
 апель.jpg	21.06.2011 13:52	JPEG-рисунок	13 КБ
 2410.jpg	12.04.2011 16:14	JPEG-рисунок	12 КБ

Файловая система

Файловая система – функциональная часть ОС, организует работу с файлами, отвечает за хранение данных на внешних носителях и обмен данными между внешними носителями.

Устройство файловой системы зависит от операционной системы, установленной на компьютере. операционные системы Windows могут работать с файловыми системами **FAT** и **NTFS**.



Файловая система. Кластеры

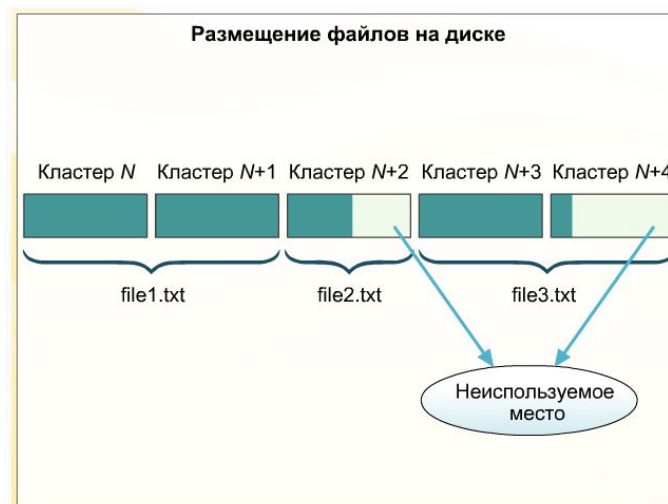
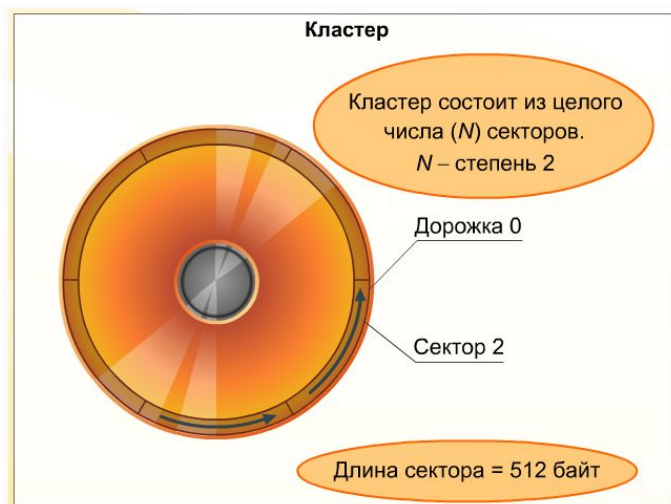
Жесткий диск состоит из дорожек, каждая из которых делится на секторы.

Кластер – это минимальный адресуемый блок дисковой памяти для записи/чтения данных на дисковом накопителе (жёстком диске).

Кластер может занимать один или несколько секторов.

Номера занятых кластеров определяют «координаты» файла для ОС.

Все файлы на диске вне зависимости от своего объема **имеют размер, кратный размеру кластера**. Любой самый маленький файл не может занимать размер на диске меньше кластера.



Файловая система. Кластеры

Размер кластера зависит от типа используемой файловой системы – 512 байт - 64 Кбайта.

Кластеры нумеруются в линейной последовательности – от первого кластера нулевой дорожки до последнего кластера последней дорожки.

Файловая система организует кластеры в файлы и каталоги (каталог – файл содержащий список файлов в данном каталоге).

Файловая система отслеживает состояние кластеров:

- кластер свободен;
- bad-кластер, который по каким-то причинам использовать уже нельзя;
- кластер занят каким-либо файлом;
- последний кластер файла;
- «зарезервированный» кластер.

На незаполненном диске файлы записываются последовательно в свободные кластеры. После удаления файлов кластеры освобождаются.

В дальнейшем файлы записываются в произвольные свободные кластеры, что часто приводит к фрагментации файлов и замедлению их чтения (части файла хранятся в удалённых друг от друга кластерах).



Файловая система. Кластеры

Пример фрагментации файлов.



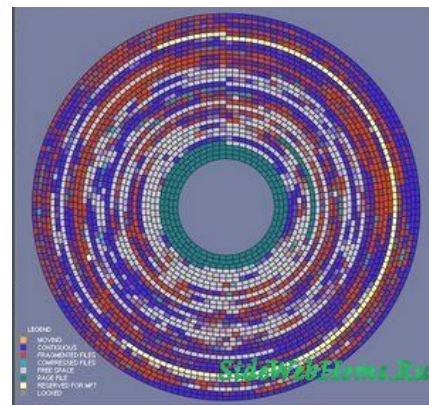
Исходное состояние (диск не фрагментирован)



Файл В удален



Файл D записан на ближайшее свободное место



Так как на диске могут храниться сотни тысяч файлов, то фрагментированность файлов, возникающая при активной работе (запись-удаление файлов), будет существенно замедлять доступ к файлам и приводить к износу жёсткого диска – магнитным головкам придётся постоянно перемещаться с дорожки на дорожку.

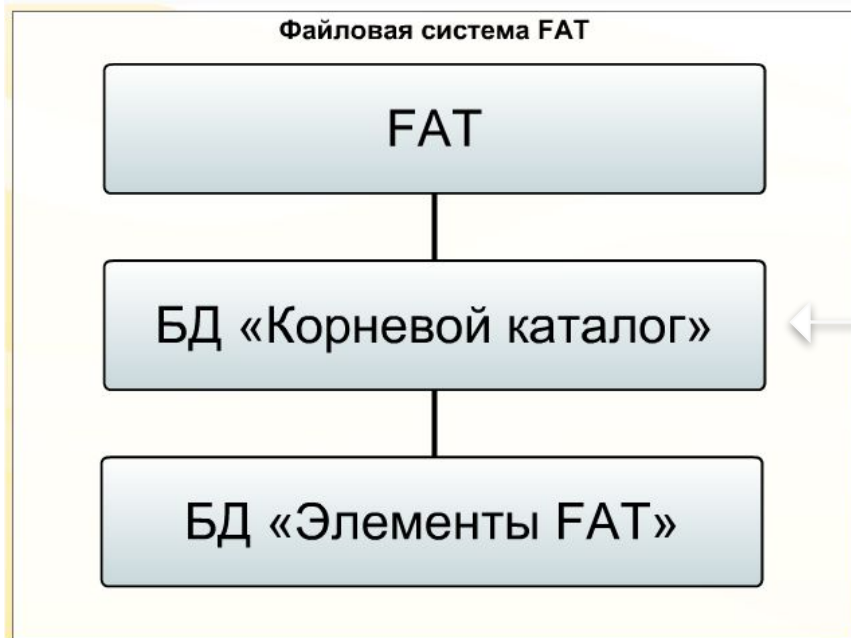
Для устранения данной проблемы рекомендуется периодически производить **дефрагментацию диска**.



Файловая система. FAT-16, FAT-32.

Файловая система **FAT**

(**File Allocation Table** – «таблица размещения файлов»).



БД «Корневой каталог» – это база данных, в которой хранится информация о файлах и папках, находящихся на диске.

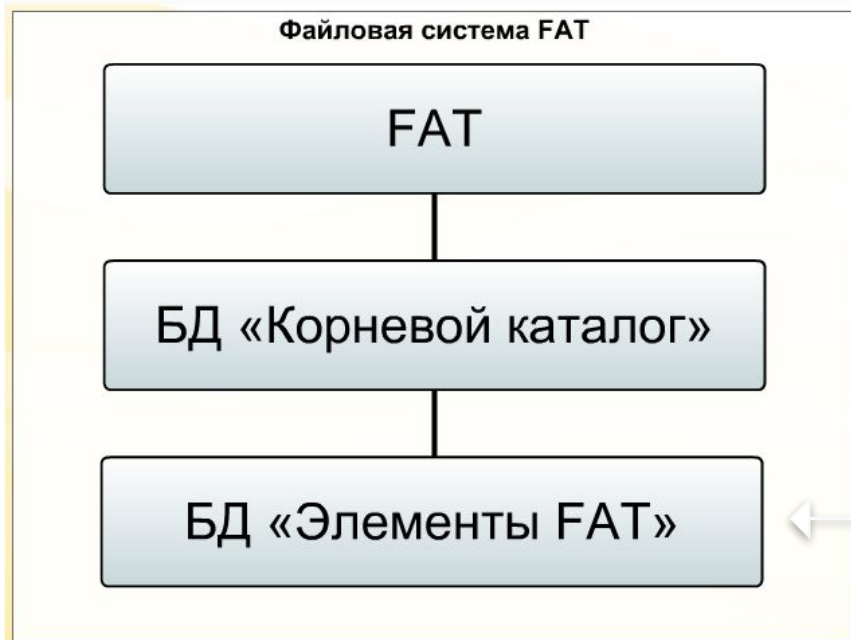
Упрощенную структуру БД «Корневой каталог» можно представить в виде таблицы:

8 байт	3 байта	1 байт	2 байта	2 байта	2 байта	4 байта
Имя файла	Расширение	Атрибут	Время создания	Дата	Номер начального кластера	Размер файла

Файловая система. FAT-16, FAT-32.

Файловая система **FAT**

(**File Allocation Table** – «таблица размещения файлов»).



БД «Элементы FAT» содержит информацию о кластерах диска.

Упрощенно структуру этой базы данных можно представить следующим образом:

Номер кластера/признак
Номер кластера/признак
Номер кластера/признак

Информация о кластере 1
Информация о кластере 2
Информация о кластере 3



Файловая система. FAT-16.

В файловой системе **FAT-16** выделяется 16 бит для хранения адреса кластера.

$2^{16} = 65\,536$ кластеров.

Макс. объём кластера 64 Кбайта, объём сектора 512 байт, следовательно объём кластера не может превышать 128 секторов ($65\,536 : 512 = 128$).

$64 \text{ Кбайта} \times 65\,536 \text{ кластеров} = 4\,194\,304 \text{ Кбайта} = 4 \text{ Гбайта}$,
поэтому FAT-16 нельзя использовать для носителей информации более 4 Гб!

Недостатки файловой системы FAT16:

- ограничения на объем диска и размер файла;
- ограничения длины имени файла;
- фрагментация файлов, приводящая к снижению быстродействия и износу оборудования;
- потери памяти диска, вызванные большими размерами кластера.

Файловая система. FAT-32.

В файловой системе **FAT-32** выделяется 32 бита для хранения адреса кластера.

$2^{32} = 4\,294\,967\,296$ кластеров.

Объём кластера по умолчанию – 4 Кбайта.

4 Кбайта × 4 294 967 296 кластеров = 16 384 Гбайта = 16 Тбайт, поэтому FAT-32 можно использоваться для носителей информации до 16 Тб!

Недостатки файловой системы FAT32:

В файловой системе FAT32 были сняты ограничения на длину имени и размер корневого каталога, но остальные ограничения, присущие FAT, остаются в силе

- ограничения на размер файла;
- фрагментация файлов, приводящая к снижению быстродействия и износу оборудования;
- потери памяти диска, вызванные большими размерами кластера.

Файловая система. NTFS.

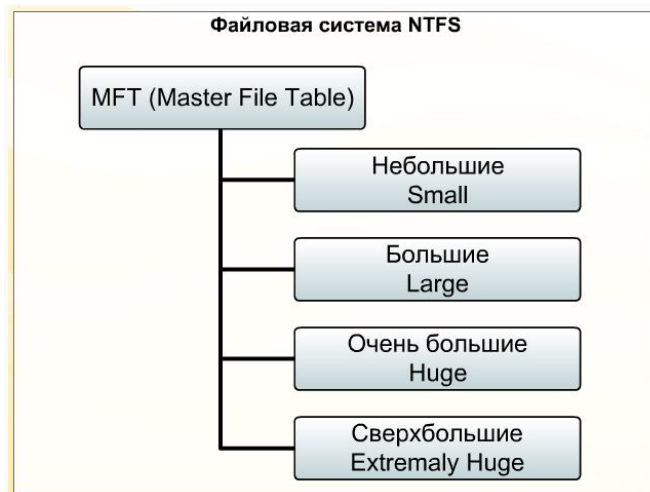
Файловая система **NTFS** (New Technology File System – «файловая система по новой технологии») поддерживает любые размеры кластеров от 512 байт до 64 Кбайт, но по умолчанию используется кластер размером 4 Кбайта.

По сравнению с FAT-32 файловая система NTFS увеличивает надёжность и эффективность использования дискового пространства.

В NTFS для повышения надёжности используется система журналирования – перед фактическими изменениями в файловой системе сохраняется список этих изменений в специальной части файловой системы, называемой «журналом» или «логом». Система восстановления позволяет восстановить работу ОС после критических ошибок.

В файловой системе NTFS

все файлы
подразделяются
по размеру на
следующие
категории:



Основные файловые системы.

- **FAT** - наиболее простая из файловых систем. Диск, отформатированный в файловой системе FAT, делится на кластеры, размер которых зависит от размера тома. Основой системы FAT является таблица размещения файлов, которая помещена в самом начале тома.
- **HPFS** - в этой файловой системе поддерживается структура каталогов FAT и добавлена сортировка файлов по именам. Файл состоит из «данных» и специальных атрибутов, помимо их так же присутствуют сведения о создании и внесении изменений, а также дата и время доступа.

Файловая система. Подготовка жёсткого диска к работе.

Новый жёсткий диск не может быть сразу использован для хранения информации и установки программ.

Вначале на диске необходимо создать **разделы (тома)** и произвести процесс форматирования (разметки) диска. На одном физическом носителе информации можно создать один или несколько разделов (логических дисков). Логическим дискам ставятся в соответствие заглавные буквы латинского алфавита (C:, D:, E: и т.д.).

Форматирование – создание логической структуры диска в соответствии с определённой файловой системой. Разделы на одном физическом носителе могут быть отформатированы в различных файловых системах, соответственно на один физический носитель могут быть установлены различные операционные системы.

При установке операционной системы на новый диск процедура создания разделов и форматирование диска выполняется средствами установочного (инсталляционного) диска.

При установке дополнительного диска, создание разделов и форматирование диска выполняется средствами операционной системы.

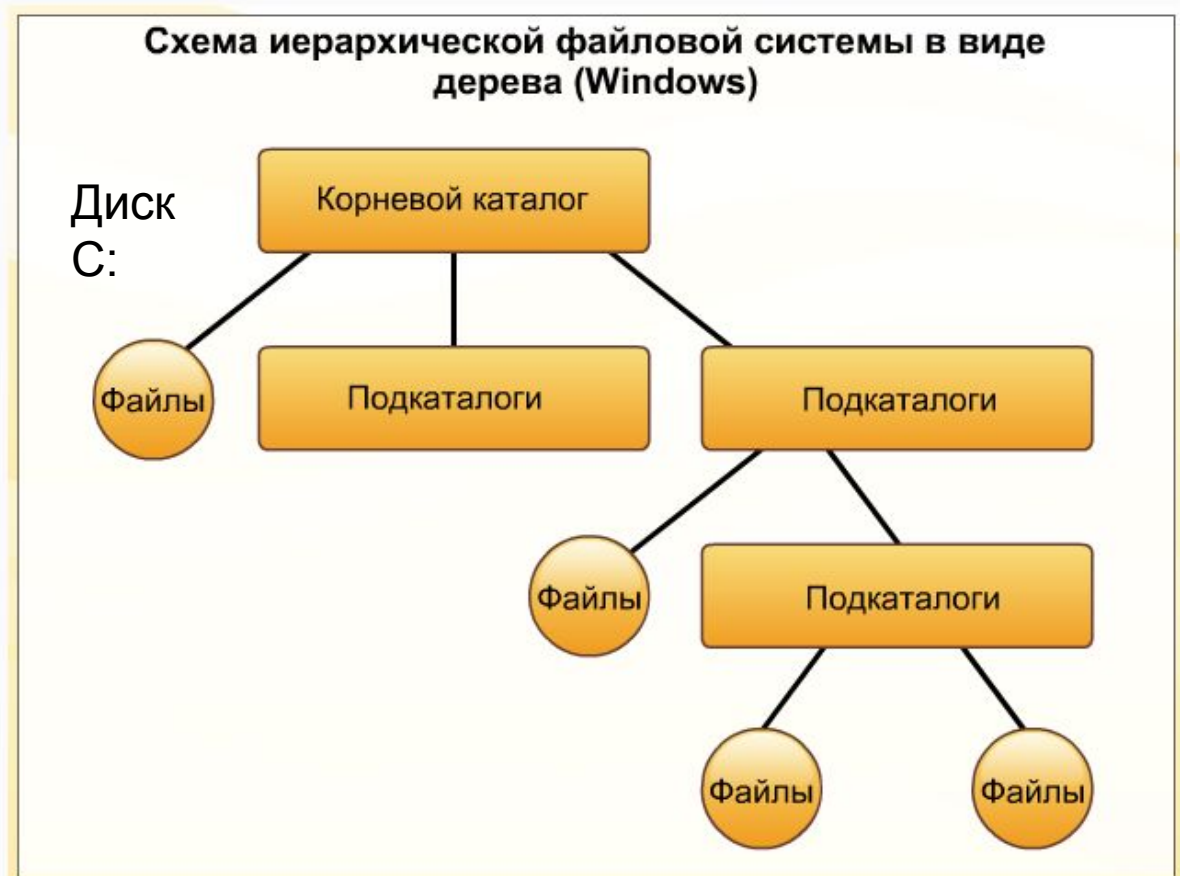
Файловая система. Иерархическая.

В иерархической файловой системе в процессе форматирования создаётся **корневой каталог**. В нем могут храниться не только файлы, но и другие каталоги более низкого уровня, называемые подкаталогами, или поддиректориями.

В операционной системе Windows подкаталоги принято называть **папками**.

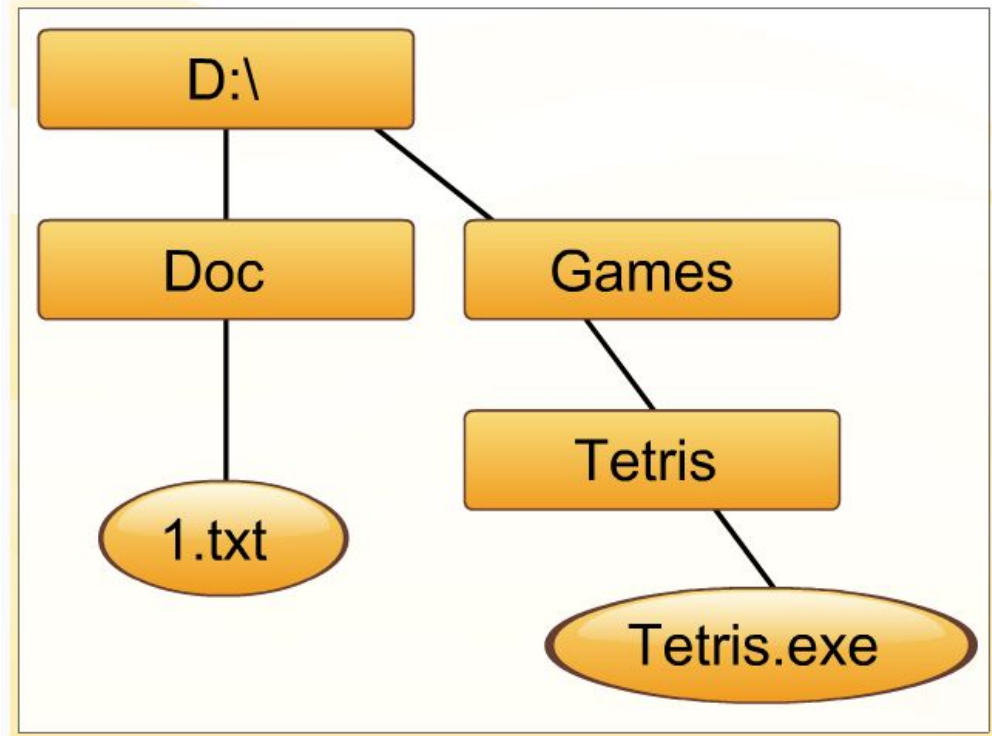
Каталоги, организованные в многоуровневую иерархическую структуру, называют также «**деревом каталогов**».

Каждый каталог, кроме корневого имеет один единственный «родительский» каталог, внутри которого он находится.



Файловая система. Иерархическая.

Чтобы найти нужный файл в указанной иерархической структуре, необходимо указать *путь к этому файлу*. Для этого нужно указать логическое имя диска, на котором хранится файл, а потом через знак «\» указать последовательно все подкаталоги, вложенные друг в друга, где последним будет тот подкаталог, в котором расположен файл.

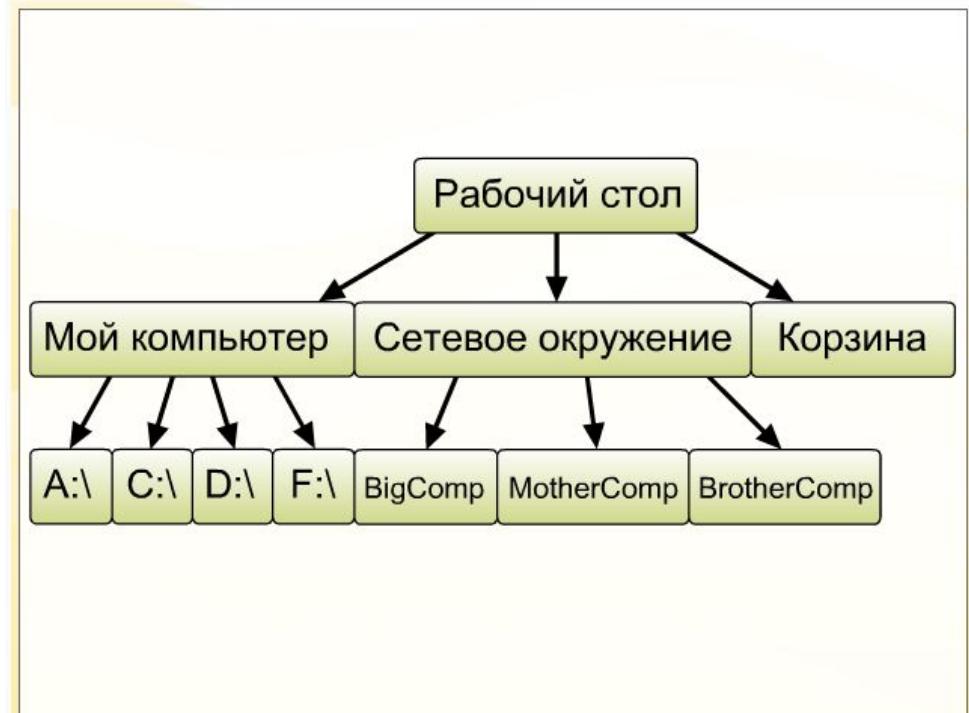


Например, в показанной выше иерархии путь к файлу 1.txt записывается как **D:\Doc**, а путь к файлу tetris.exe – как **D:\Games\Tetris**. Если указаны и путь к файлу, и его имя, то такая конструкция называется *полным именем файла*. Так, **полное имя файла tetris.exe** – это **D:\Games\Tetris\tetris.exe**.

Файловая система. Иерархическая.

Для ОС Windows корневым каталогом можно считать папку «Рабочий стол», в которой располагаются подкаталоги «Мой компьютер», «Корзина» и «Сетевое окружение». Эти подкаталоги носят стандартные названия и служат для заранее определенных целей.

Папка «**Мой компьютер**» в качестве подкаталогов содержит все диски, имеющиеся в данном компьютере.

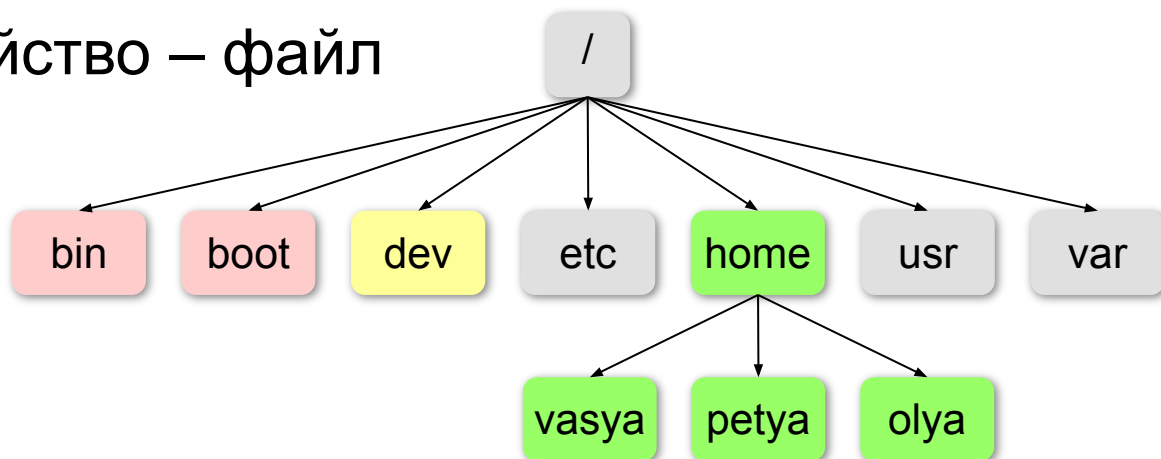


Папка «**Сетевое окружение**» содержит папки всех компьютеров, которые подключены в этот момент к локальной сети и к которым можно обратиться с данного компьютера.

Папка «**Корзина**» временно хранит все удаленные на данном компьютере папки и файлы.

Файловые системы в *Linux*

- один корневой каталог «/»
- любое устройство – файл



Путь к файлу: `/home/petya/qq.txt`

bin – команды операционной системы;

boot – ядро ОС и данные для загрузки;

dev – файлы устройств

etc – файлы с настройками ОС и некоторых программ

home – домашние каталоги пользователей

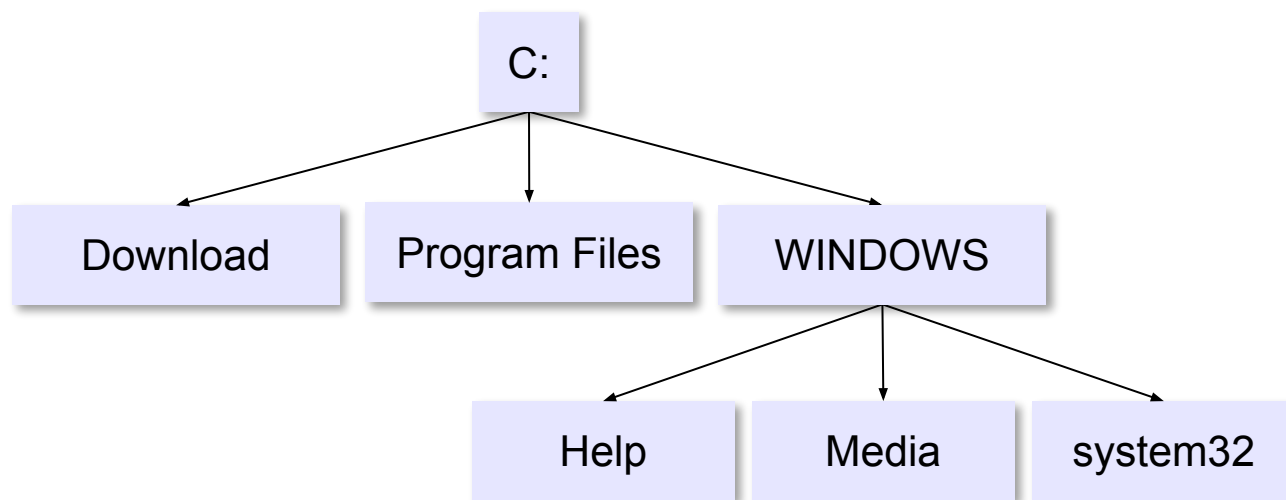
usr – установленные пакеты программ

var – часто меняющиеся данные, например, журналы ОС

Файловые системы в *Windows*

- каждый логический диск имеет свой корневой каталог
- разделитель в записи пути к файлу – «\»
- заглавные и строчные буквы **не** различаются:

`Вася.txt` \equiv `вася.txt` \equiv `ВаСя.ТхТ`



Путь к файлу: `C:\WINDOWS\System32\shell32.dll`