

Управление файлами

Задачи ОС:

- Обеспечить пользователю удобный интерфейс доступа к данным;
- Обеспечить совместное использование данных;

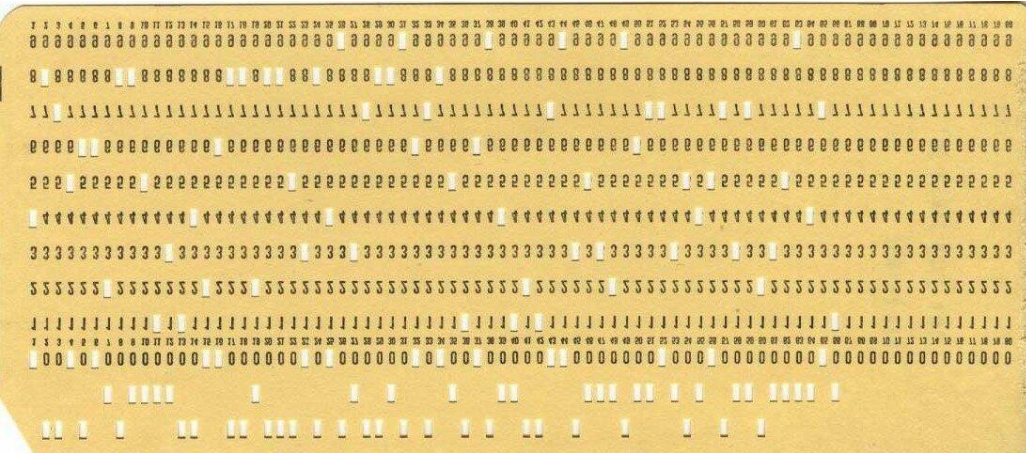
Будет рассмотрено:

1. Физическая структура
2. Логическая организация
3. Хранение конфигурации ОС
4. Механизм защиты файлов
5. Многопоточные файловые системы
6. Специальные файлы (файлы устройств)
7. Дополнительные возможности файловых систем

1. Долговременное хранение – после выключения завершения процесса
 2. Хранение больших объемов
 3. Обеспечить возможность одновременного доступа к информации
- Совокупность файлов/каталогов на диске
 - Наборы структур данных для управления файлами
 - Комплекс программных средств для управления файлами

Упоминание файловых систем – 1964г, DECtape для DEC PDP-6 и OS/360 FS для IBM OS/360, подробнее wiki https://en.wikipedia.org/wiki/Comparison_of_file_systems

Устройства долговременного хранения данных



Перфокарты
(1920-1980гг)



НГМД
(1970-1990гг)
80кб-2,88мб



НЖМД
(1956-н.вр.)
До 22ТБ и более



Магнитные ленты
(1951-н.вр)
Рекорд хранения >50лет
Гарантированный
срок хранения 30-40лет
Объем до 18Тб/картридж

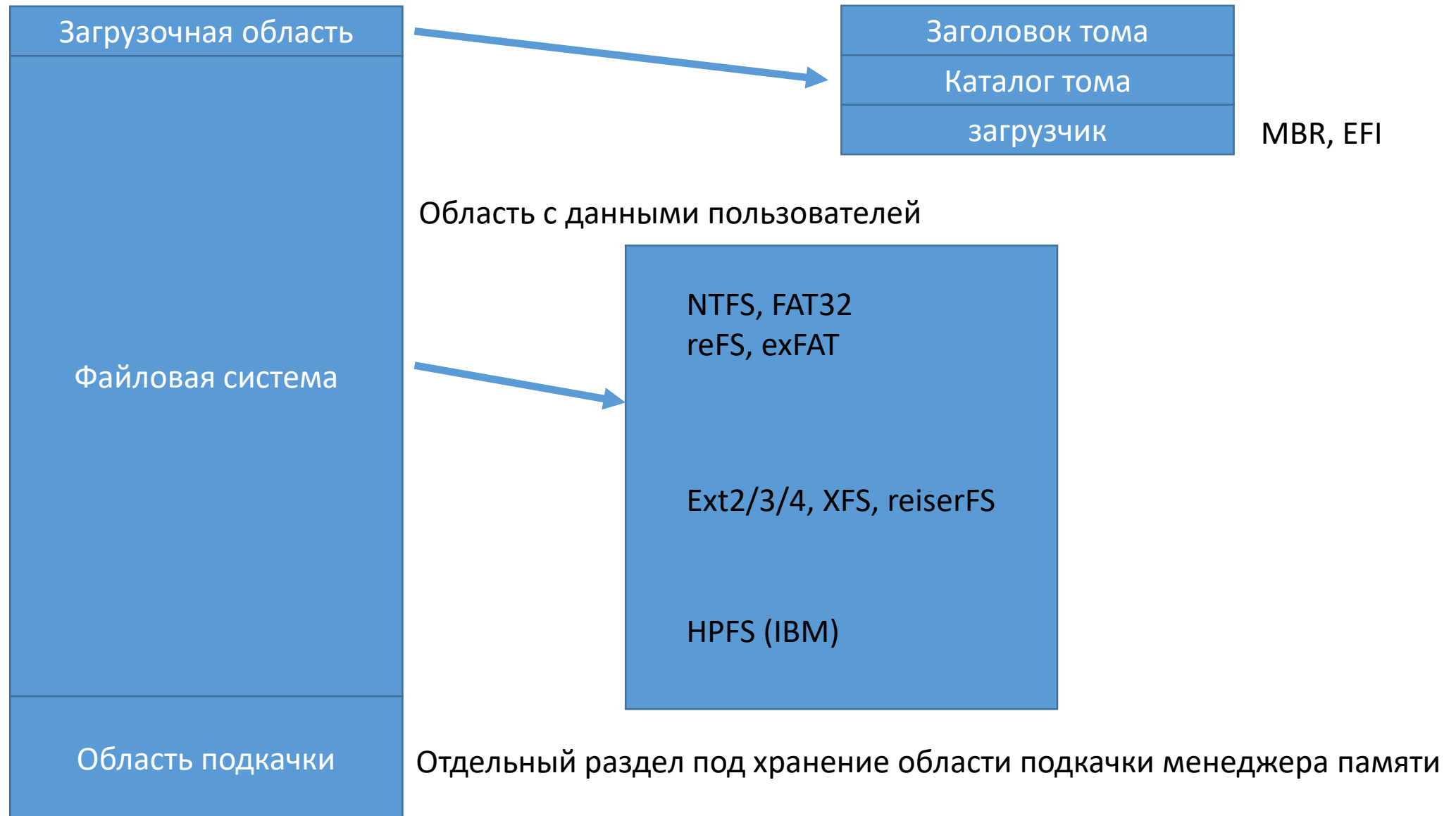
Flash память
(1984-н.вр)
Срок хран 10-20лет
Гарантир. срок
хранения ~5лет



Интерфейсы

- SA1000 – для первого диска 5,25” ST-506 объемом 6Мб, использован в IBM/PC, основа дисков 1980-х
- SCSI
- IDE
- SATA
- SAS
- FC (Fibre channel)
- iSCSI

Физическая организация файловых систем



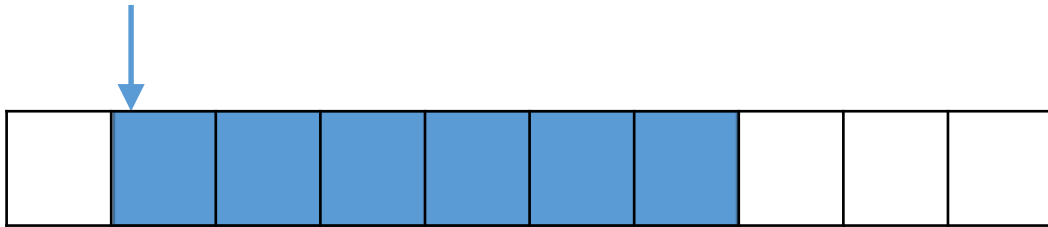
Область подкачки для виртуальной памяти

Раздел накопителя (на HDD/SSD/...) Файл подкачки

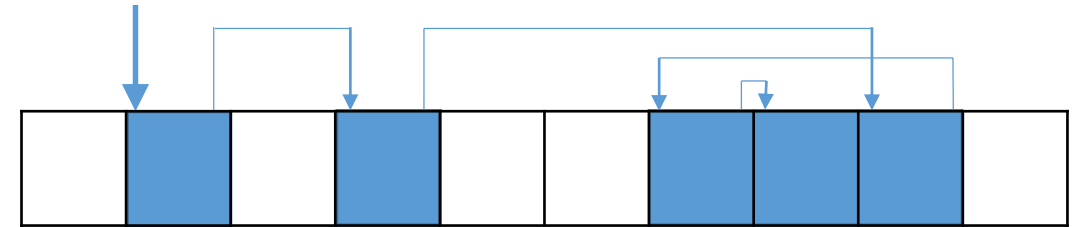
- Скорость работы максимальна за счет прямого доступа к информации на накопителе
- Фиксированный объем
- Динамическое изменение размера
- Доступ к информации через файловую систему

Способы организации хранения информации в файловых системах

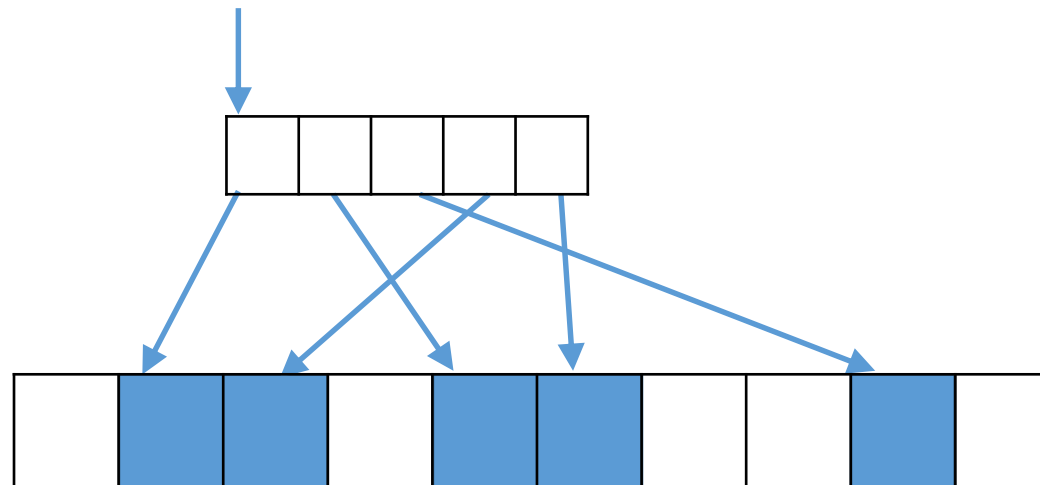
Непрерывными блоками



Связанный индекс



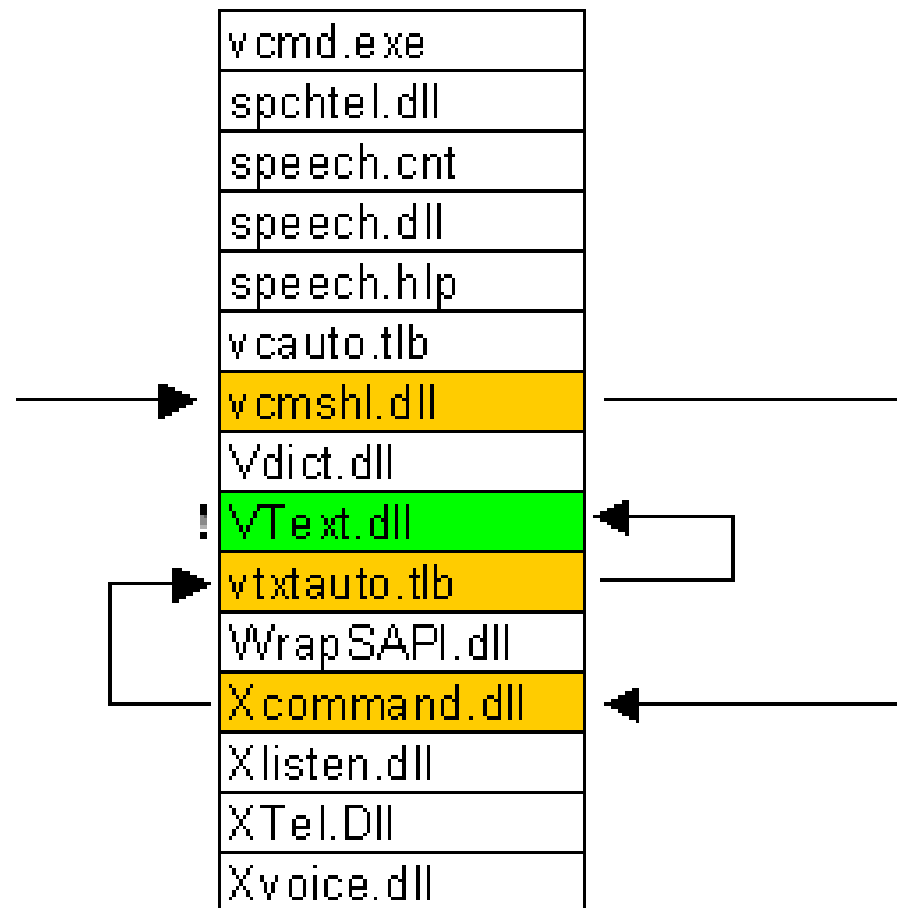
Связанный список блоков



Организация доступа к листингу каталога

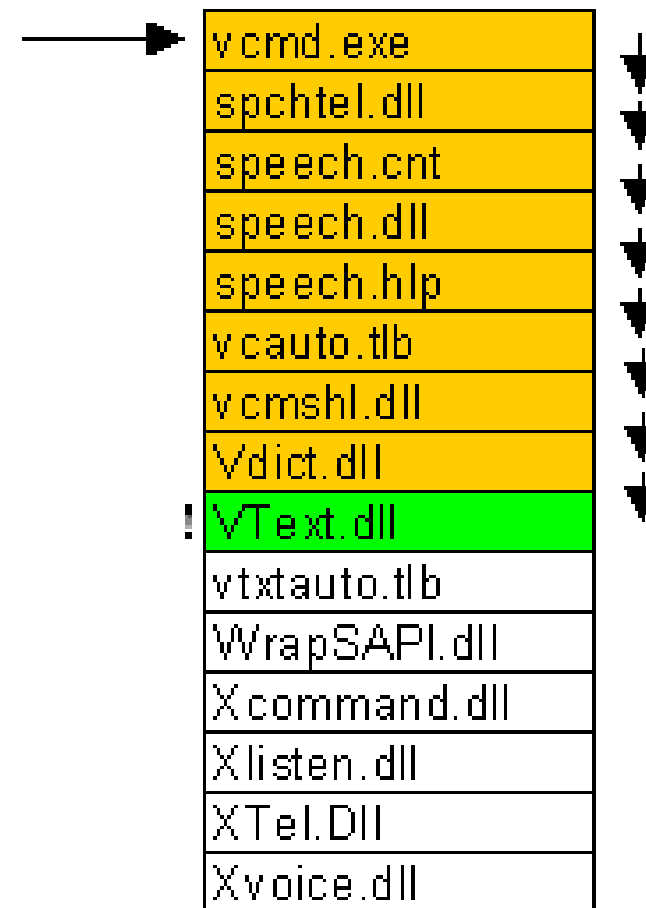
“B-Trees” (двоичные деревья)

Поиск в дереве

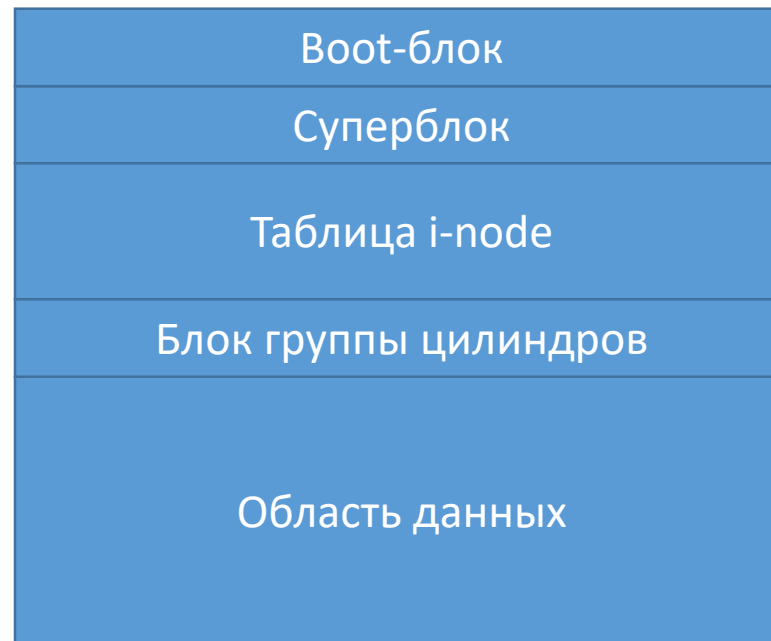
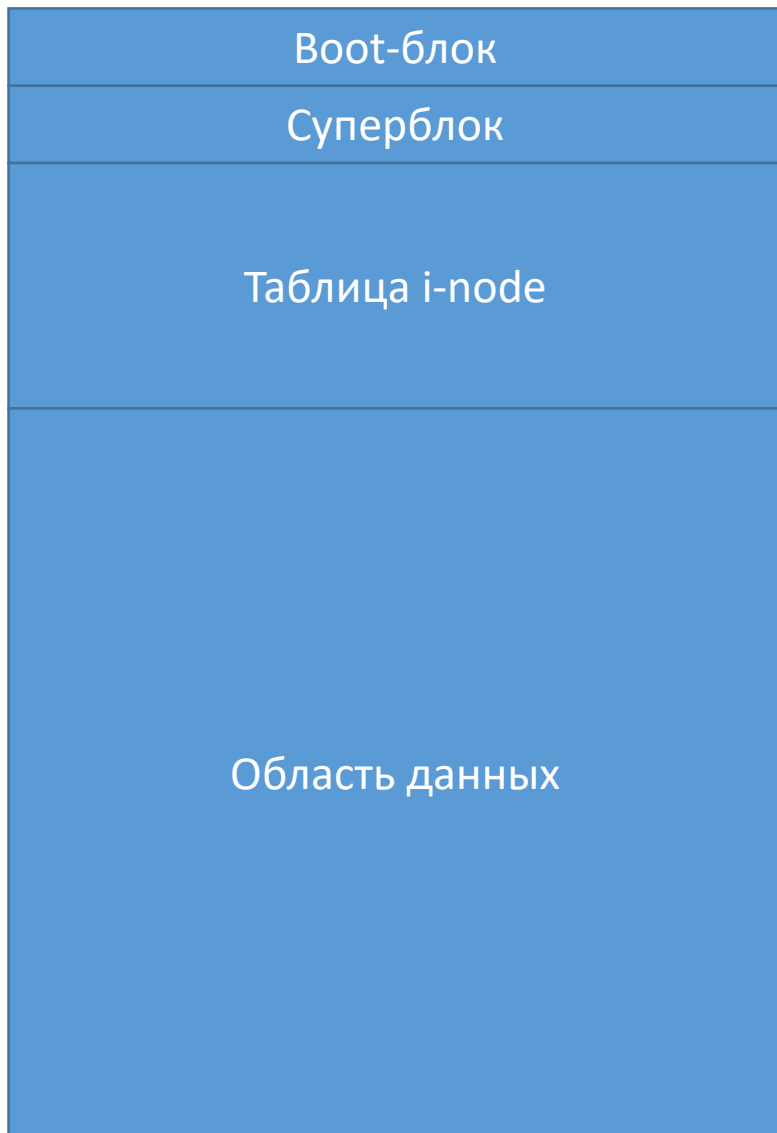


Неупорядоченный перечень
содержимого каталога

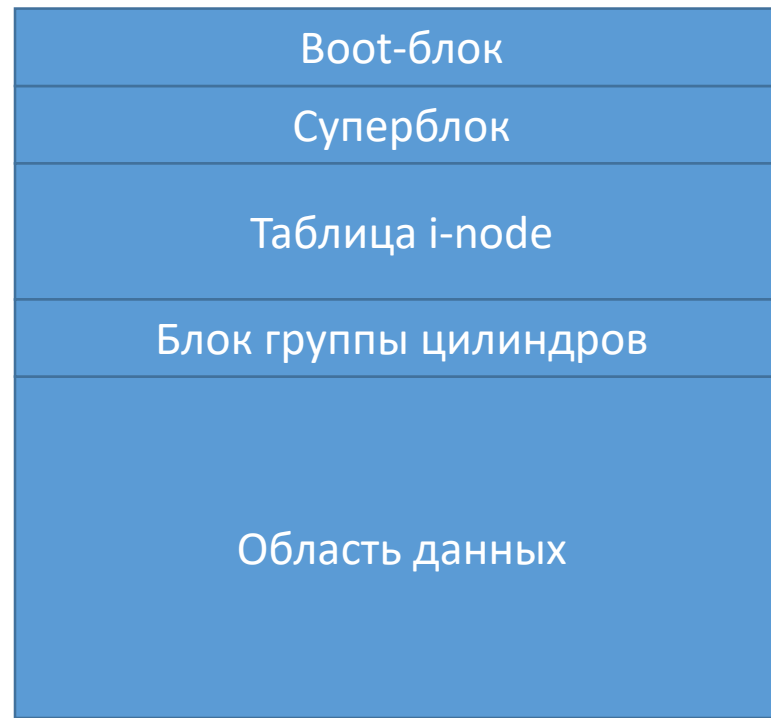
Поиск перебором



Unix, s5



.....

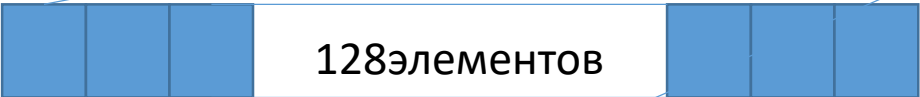


I-node: Реализация адресации к блокам с данными на диске

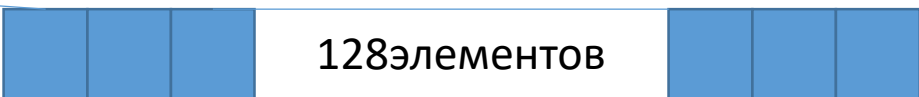
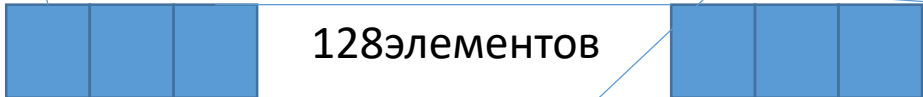
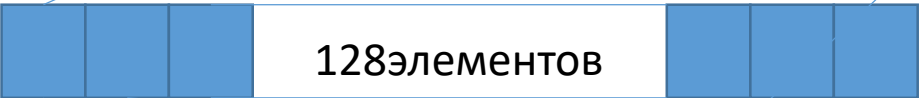
I-node



Первые 10 элементов адресуют
блоки с данными на диске
11,12 и 13 элемент адресуют блоки,
находящиеся в области данных с
следующими 128 адресами



Адреса блоков данных на диске



3-х уровневая схема организации

I-node: реализация соответствия имени физическому расположению на диске

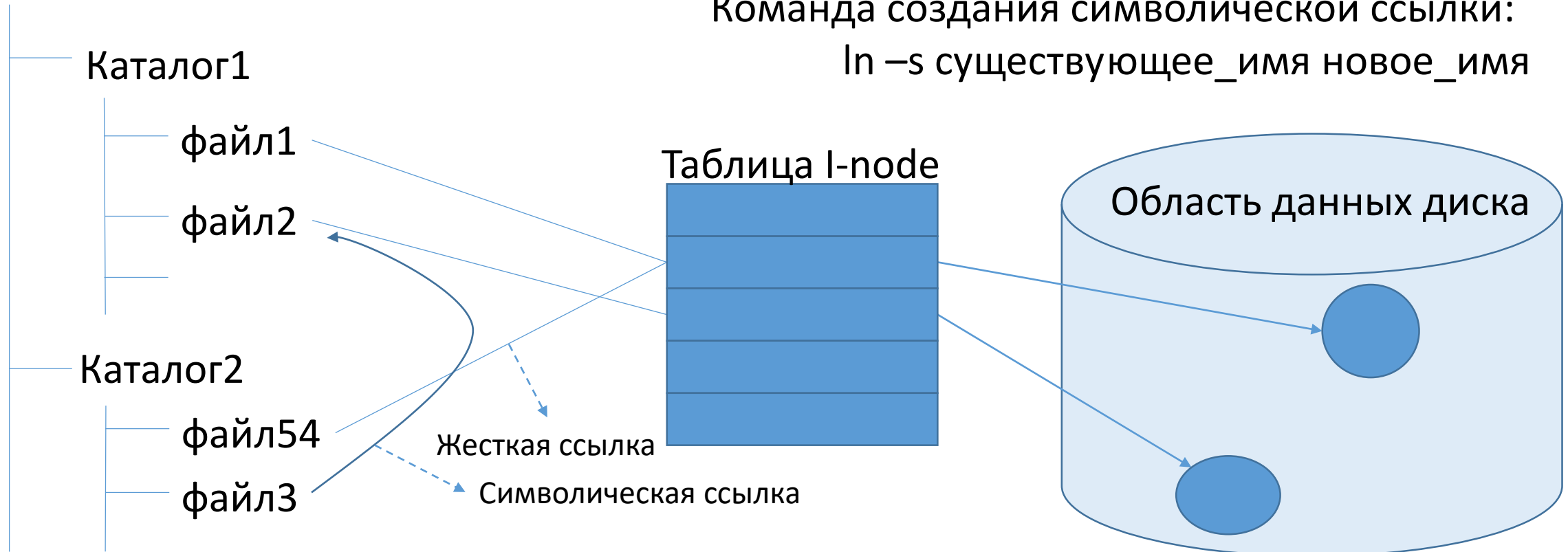
Корень файловой системы

Команда создания жесткой ссылки:

`ln существующее_имя новое_имя`

Команда создания символической ссылки:

`ln -s существующее_имя новое_имя`



Виртуальная файловая система VFS

UNIX System V Release 3

File System Switch, FSS

UNIX System V Release 4 (Sun Microsystems)

Virtual File System (VFS)

IBM HPFS (High Performance File System)

Super block

Spare block

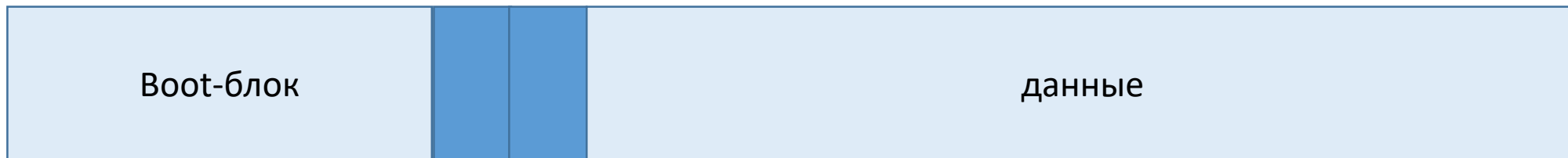
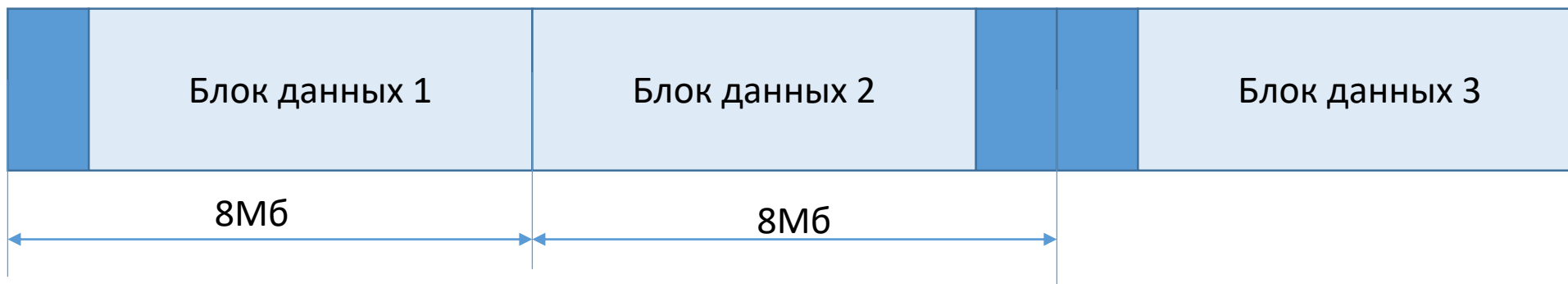


Таблица
свободного
пространства



Для IBM OS/2 разработка ~1989г, поддержка в WindowsNT до 3.5 (4.0 неофиц)

Microsoft

FAT - DOS

vFAT – для Windows95 (особые случаи)

FAT32 – Windows95

exFAT (Extended File Allocation Table)

NTFS WindowsNT (win10/server)

ReFS Windows2012/win8...

	FAT	FAT32	exFAT
р-р ФС	2Гб	8Тиб	
Р-р файла		2(4)Гб	16млн Тб
Назначение	НГМД	НЖМД	Флэш-накопители
Особенности	MSDOS	Win95...	Поддержка прав доступа, при наличии возможности - транзакций
	8имя+3расш	255симв	

FAT (File Allocation Table)

1. Размер кластера до 32К
2. Имя файла 8+3 XXXXXXXX.XXX
3. Нет поддержки прав доступа
4. Есть поддержка атрибутов
5. Определение типов файлов по расширению
 1. EXE (DOS/Windows)
 2. COM (DOS)

Плюсы

1. Простота
2. Ориентирована на небольшие объемы накопителей

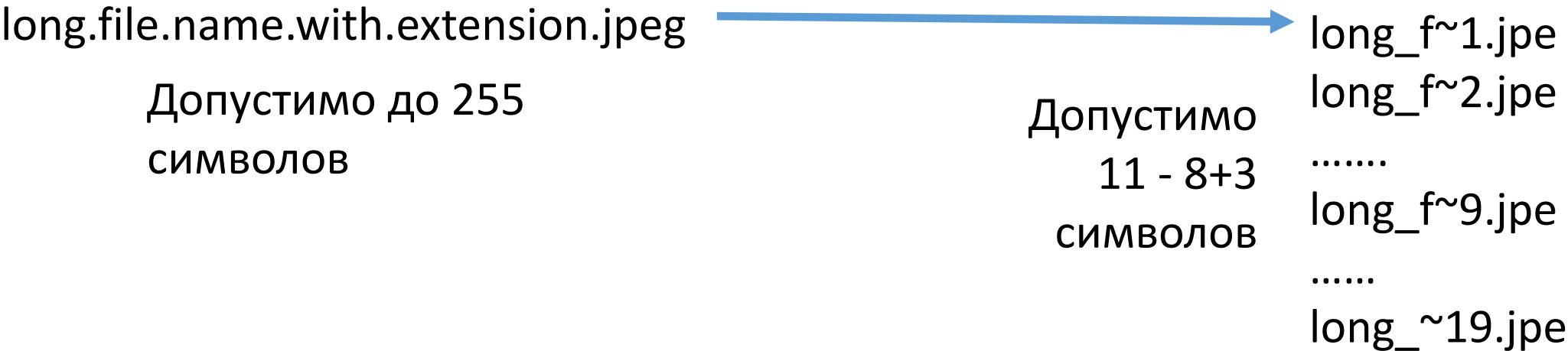
Были FAT12 и FAT16

FAT32

1. нет обратной совместимости с FAT
 2. Размер файла 2гб
 3. Имя файла 255 символов
 4. Нет поддержки прав доступа
-
1. Ориентирована на HDD

Поддержка длинных имен для обратной совместимости

Алгоритм преобразования длинного имени в старый формат

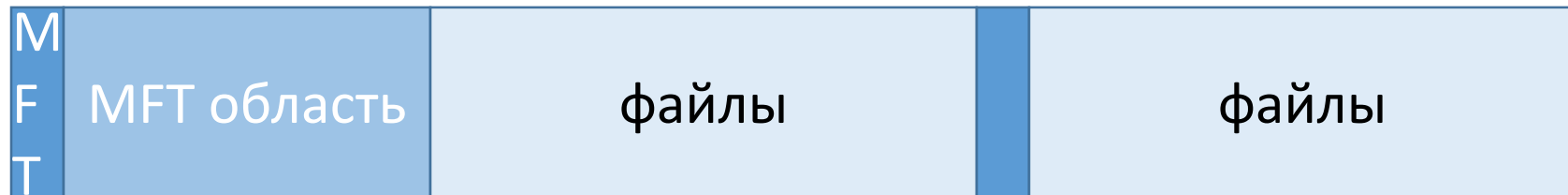


Реализация хранения длинного имени в ФС FAT

FAT
LONGFILE.NAM
LONGFIL1.NAM
LONGFIL2.NAM

VFAT
LONG FILE NAME.EXTEN
LONG F~1 EXT
E.EXTEN
LONG FILE NAM

Microsoft NTFS



Копия первых 3 записей в MFT

\$MFT	Структура MFT
\$MFTmirr	копия первых 16 записей MFT, размещенная посередине диска
\$LogFile	файл поддержки журналирования
\$Volume	служебная информация — метка тома, версия файловой системы, т. д.

- Права доступа - ACL (Access List)
- Шифрование данных
- Сжатие данных
- Функционал дефрагментации файлов

Microsoft ReFS (Resilient File System)

2012г

	ReFS	NTFS
Максимальная длина имени файла	255 символов Юникода	255 символов Юникода
Максимальная длина имени пути	32 КБ символов Юникода	32 КБ символов Юникода
Максимальный размер файла	35 PB (петабайтов)	256 ТБ
Максимальный размер тома	35 ПБ	256 ТБ
	Виртуализация, BigData	

Файловые системы специального назначения

1. Файловые системы `procfs` `sysfs`
2. Файловые системы `devfs`
3. Файловые системы в памяти (ОЗУ) `tmpfs`
4. Сжатые файловые системы, например SquashFS (!НЕ путать с режимом сжатия при записи “на лету”)

Типы объектов файловых систем

1. Обычные файлы – основной тип объектов

- **Текстовые файлы** – в кодировке ASCII, можно просмотреть простейшим текстовым редактором
- **Двоичные файлы** – необходима специальная программа для просмотра содержимого, например JPEG

2. Каталоги – объекты особого формата, доступны только ОС

3. Специальные файлы (файлы устройств) – специальный файл, сопоставленный с драйвером устройством вв/вывода

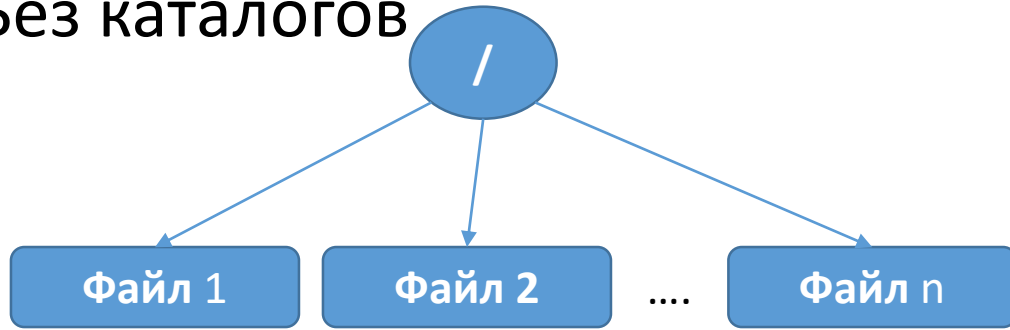
4. Именованные каналы – механизм взаимодействия двух приложения

5. Сокеты – механизм взаимодействия приложений по схеме «клиент-сервер»

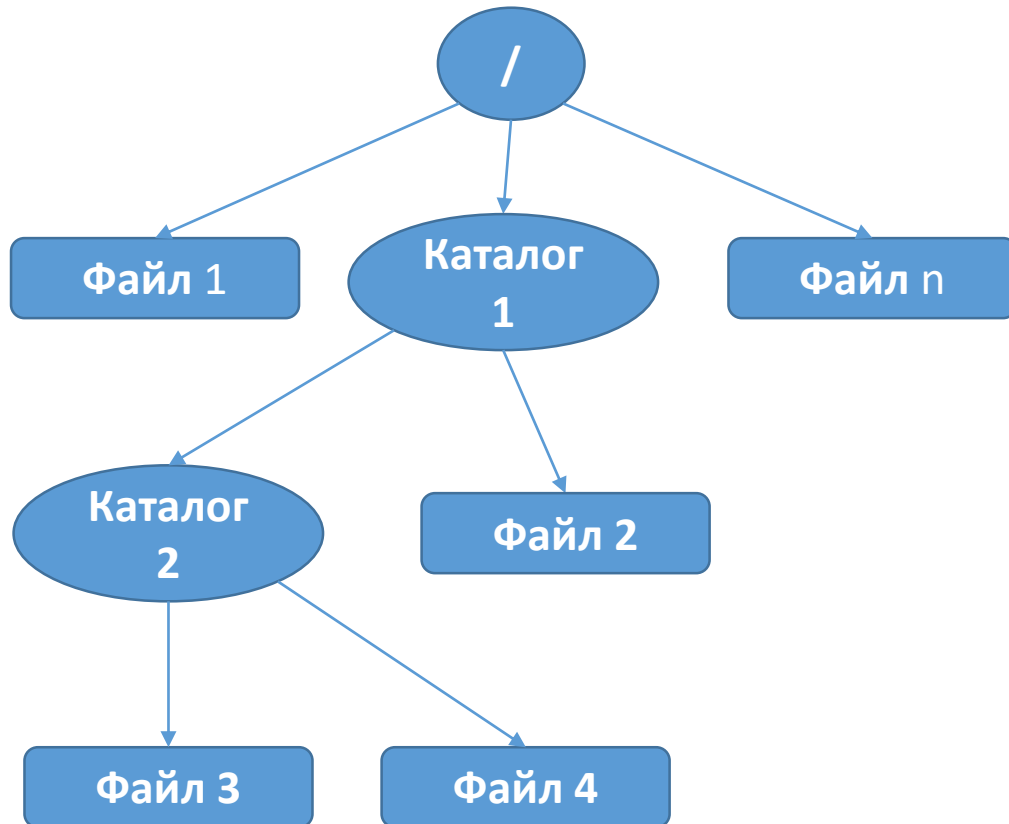
6. Символические ссылки – особые файлы, указывающие на другие объекты (т.н. «ярлыки»)

Способы организации (группировки) хранения информации

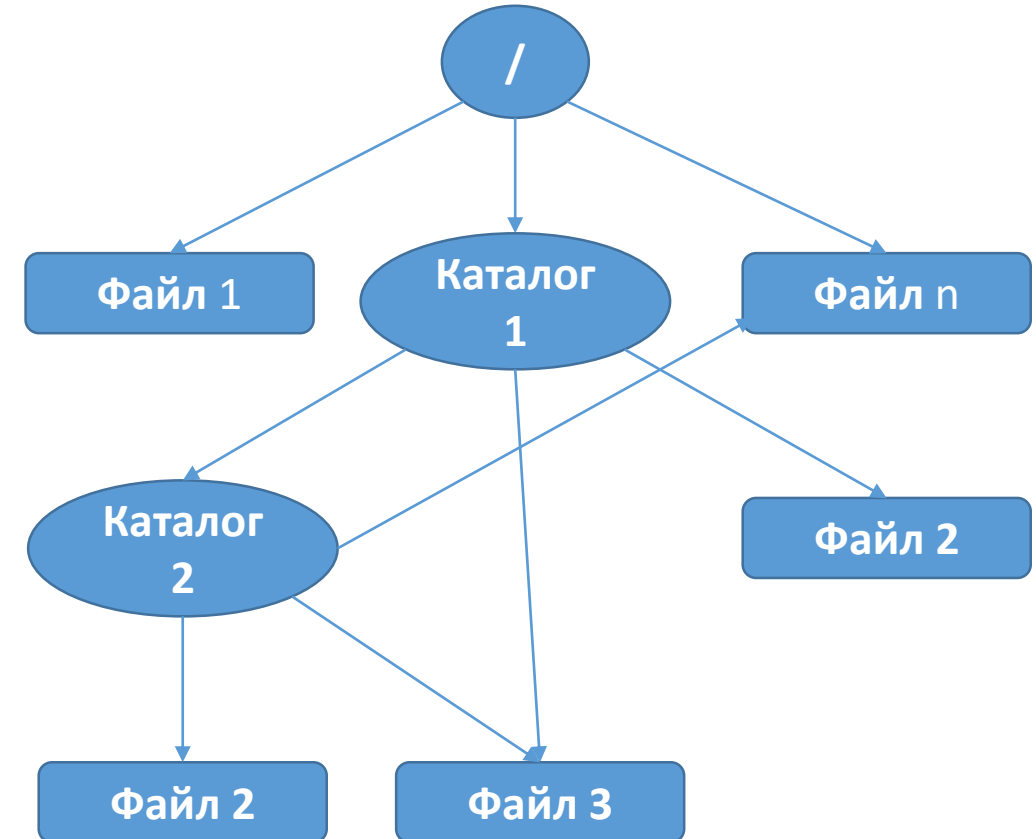
Без каталогов



Иерархическая структура – «Дерево»



Иерархическая структура –
«произвольный граф»



Полные и относительные пути файловой системы

Полный путь к объекту файловой системы

/usr/local/sbin/http



каталог



разделитель

имя объекта

Относительный путь к объекту файловой системы

sbin/http

./sbin/http

../local/sbin/http

Логическая организация файловой системы Windows

Файлы операционной системы:

C:\winnt - windowsNT (4..2000)

C:\windows – windows2k.... Н.В.

C:\windows – windows9X

Установленное программное обеспечение:

C:\program files

WindowsNT .. Windows XP: Каталог с пользовательскими данными

C:\document and settings

Windows Vista ... Н.В.:

Каталог с пользовательскими данными C:\users

Каталог с данными системных приложений C:\ProgramData

Логическая организация файловой системы Unix

ОС Linux – Filesystem Hierarchy Standard (FHS) (<https://refspecs.linuxfoundation.org/fhs.shtml>)

Корневая файловая система состоит из:

- / - содержит основные каталоги системы
- /boot – файлы для загрузки ОС, в т.ч. ядро ОС
- /sbin – системное ПО, доступно системному администратору
- /bin – системное ПО, доступное всем пользователям
- /lib – базовые разделяемые библиотеки
- /root – домашний каталог системного администратора
- /etc – каталог конфигурационных файлов
- /dev – с парой технических спецфайлов *

Виртуальные файловые системы

- /proc – отображение/настройка параметров ядра ОС
- /sys (linux) – отображение/настройка параметров ядра ОС
- /dev – каталог с специальными файлами устройств

Логическая организация файловой системы Unix

Прикладное ПО /usr

/usr/sbin – Прикладное ПО, доступное системному администратору

/usr/bin – прикладное ПО, доступное всем пользователям

/usr/lib – разделяемые библиотеки прикладного ПО

/usr/include – библиотеки языка C

Прочее

/home – каталог

/var – изменяемые файлы – журналы, файлы баз данных и т.д.

/tmp – каталог временных файлов

/mnt – обычный каталог для подключения дополнительных файловых систем

Механизмы разграничения доступа к файловой системе

«Код защиты» - двоичная маска

ACL – «списки доступа»

Sstrwxrwxrwx UID GID

Rwx rwx rwx

R = read

W – write

X – eXecute

Rw- r-- ---

110 100 000

640

chmod 640 file chmod +w file

chown

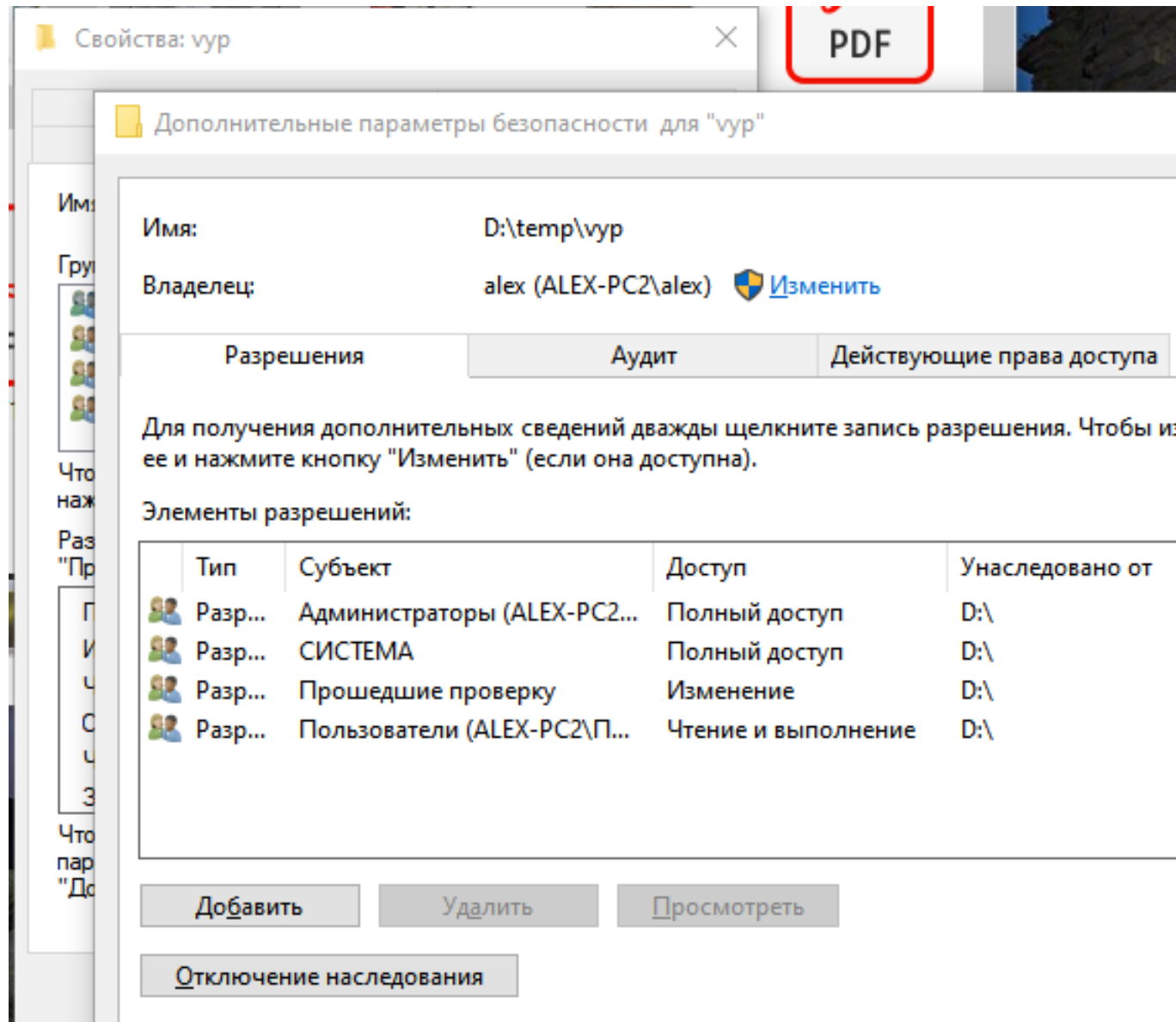
chgrp

SST

SetUID (SUID) (на файл)

SetGID (SGID) (на файл)

sTicky (на каталог)



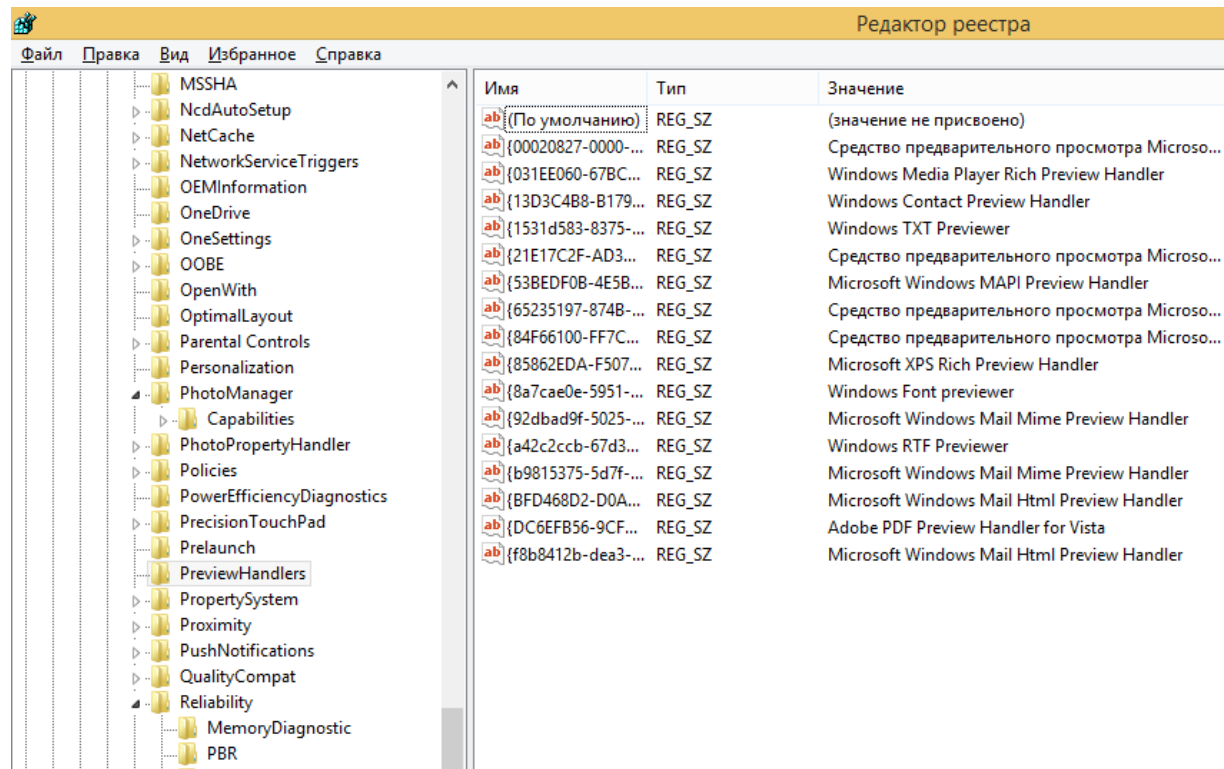
Способы хранения конфигурации ОС и прикладного ПО

В виде отдельных конфигурационных файлов

```
/etc/apache2/httpd.conf 13302/20222 65%  
  
#  
# LogLevel: Control the number of messages logged to the error_log.  
# Possible values include: debug, info, notice, warn, error, crit,  
# alert, emerg.  
#  
#LogLevel warn  
LogLevel debug  
  
<IfModule log_config_module>  
#  
# The following directives define some format nicknames for use with  
# a CustomLog directive (see below).  
#  
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\""" common  
LogFormat "%h %l %u %t \"%r\" %>s %b" common  
  
<IfModule logio_module>  
# You need to enable mod_logio.c to use %I and %O  
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\" %I  
</IfModule>  
  
#
```

+ проще и легче забрать конфигурацию ПО
с неработающих ОС

В виде БД



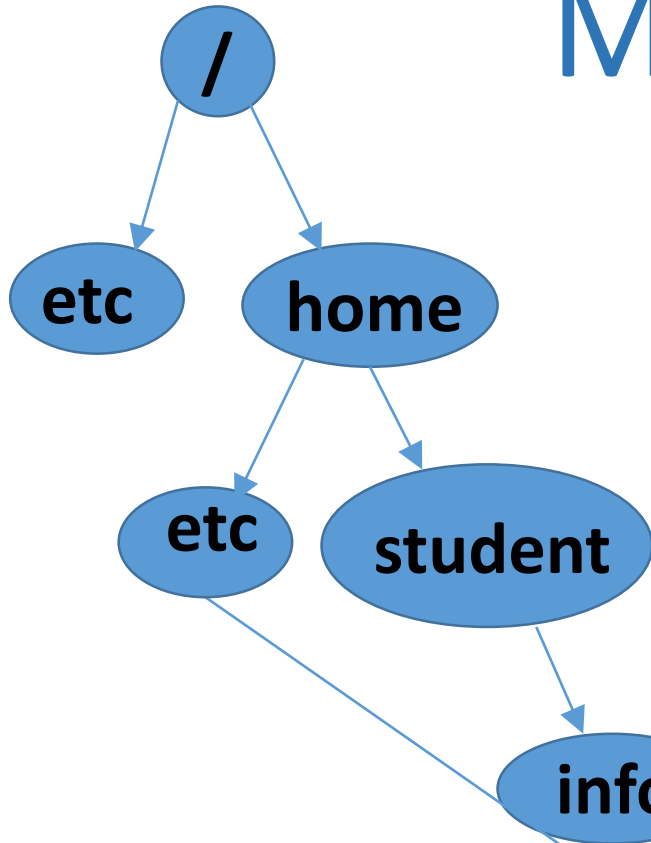
Редактор реестра

Имя	Тип	Значение
(По умолчанию)	REG_SZ	(значение не присвоено)
{00020827-0000-...	REG_SZ	Средство предварительного просмотра Microso...
{031EE060-67BC...	REG_SZ	Windows Media Player Rich Preview Handler
{13D3C4B8-B179...	REG_SZ	Windows Contact Preview Handler
{1531d583-8375-...	REG_SZ	Windows TXT Previewer
{21E17C2F-AD3...	REG_SZ	Средство предварительного просмотра Microso...
{53BEDF0B-4E5B...	REG_SZ	Microsoft Windows MAPI Preview Handler
{65235197-874B-...	REG_SZ	Средство предварительного просмотра Microso...
{84F66100-FF7C...	REG_SZ	Средство предварительного просмотра Microso...
{85862EDA-F507...	REG_SZ	Microsoft XPS Rich Preview Handler
{8a7cae0e-5951-...	REG_SZ	Windows Font previewer
{92dbad9f-5025-...	REG_SZ	Microsoft Windows Mail Mime Preview Handler
{a42c2ccb-67d3...	REG_SZ	Windows RTF Previewer
{b9815375-5d7f-...	REG_SZ	Microsoft Windows Mail Mime Preview Handler
{BFD468D2-D0A...	REG_SZ	Microsoft Windows Mail Html Preview Handler
{DC6EFB56-9CF...	REG_SZ	Adobe PDF Preview Handler for Vista
{f8b8412b-dea3-...	REG_SZ	Microsoft Windows Mail Html Preview Handler

+ предоставляются системные вызовы для
работы с настройками

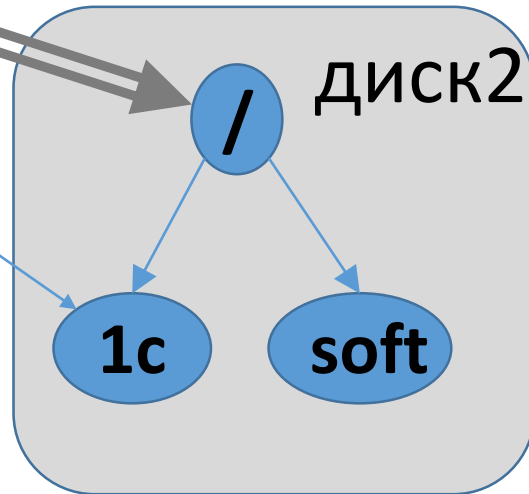
Монтирование

VFS Sun Unix System V release 5

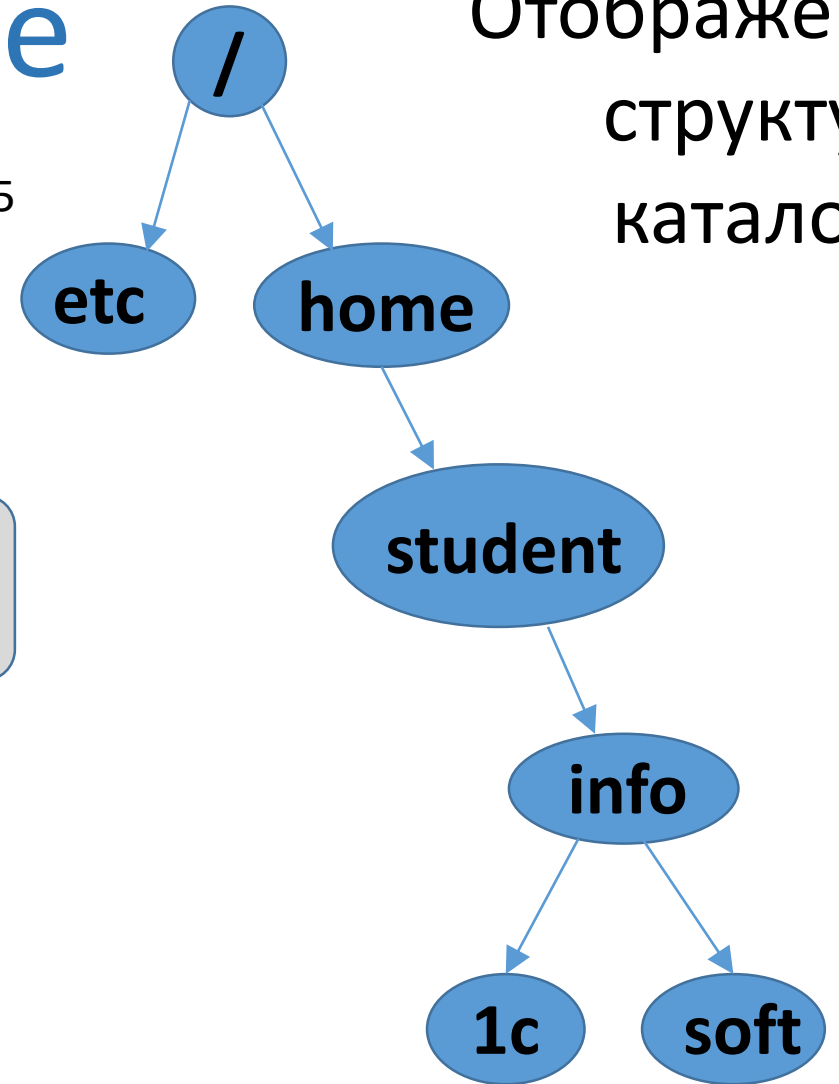


Монтирование

Каталог – точка
монтирования



Отображение
структуры
каталогов



Типы файловых систем

- Локальные ФС – подключаются к данному компьютеру
- Сетевые ФС – доступ к файлам через компьютерную сеть, работа с информацией похожа на работу с локальной файловой системой, информация находится на определенных серверах
- Распределенные ФС – аналогично сетевым, но [в общем случае] пользователь не знает где именно находятся его данные
- Кластерные ФС – применяется в ситуациях прямого соединения системы хранения данных (СХД, в частном случае накопителя) с двумя и более серверами для обеспечения одновременного доступа к одной файловой системе

UNC Universal Naming Conversation

\\workgroup\\hostname\sharename\path-to-object

Альтернатива сетевым/распределенным ФС – FTP/rsync

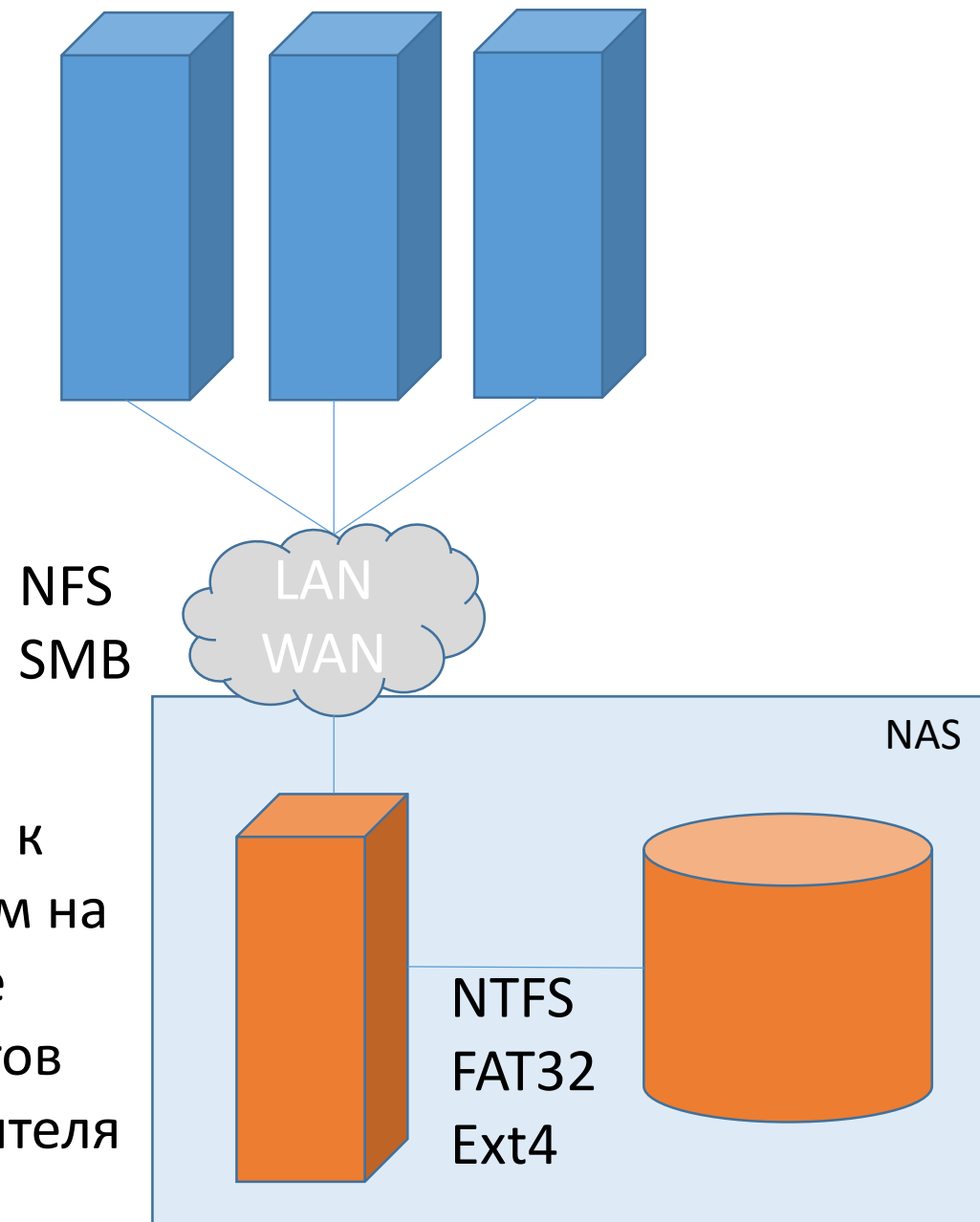
Сетевые файловые системы

Сетевые файловые системы

- NFS – Network File System (Unix)
- Novell Netware
- SMB Lan Manager samba.org

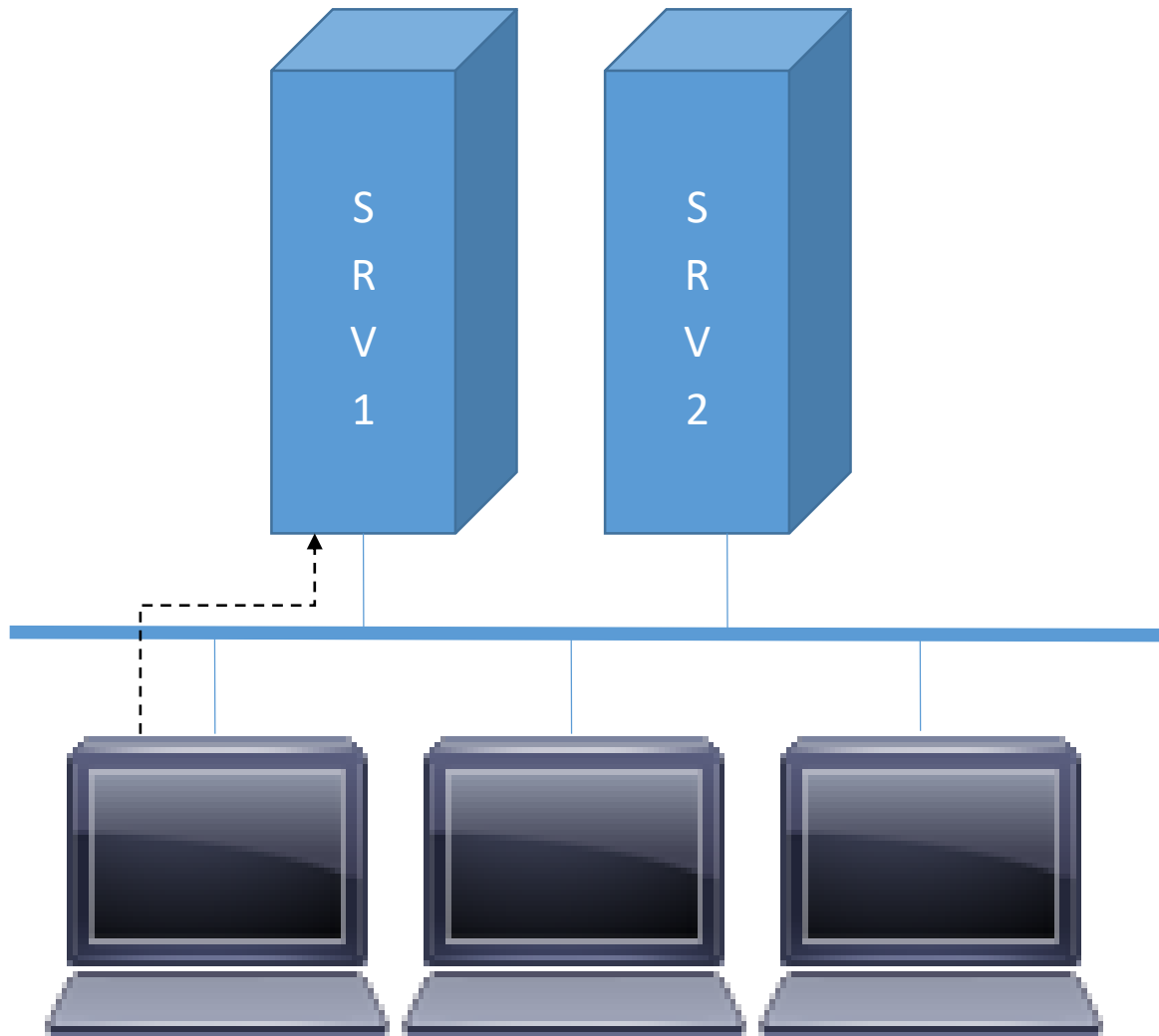


Доступ к
данным на
уровне
объектов
накопителя



Доступ к информации в сетевых и распределенных файловых системах

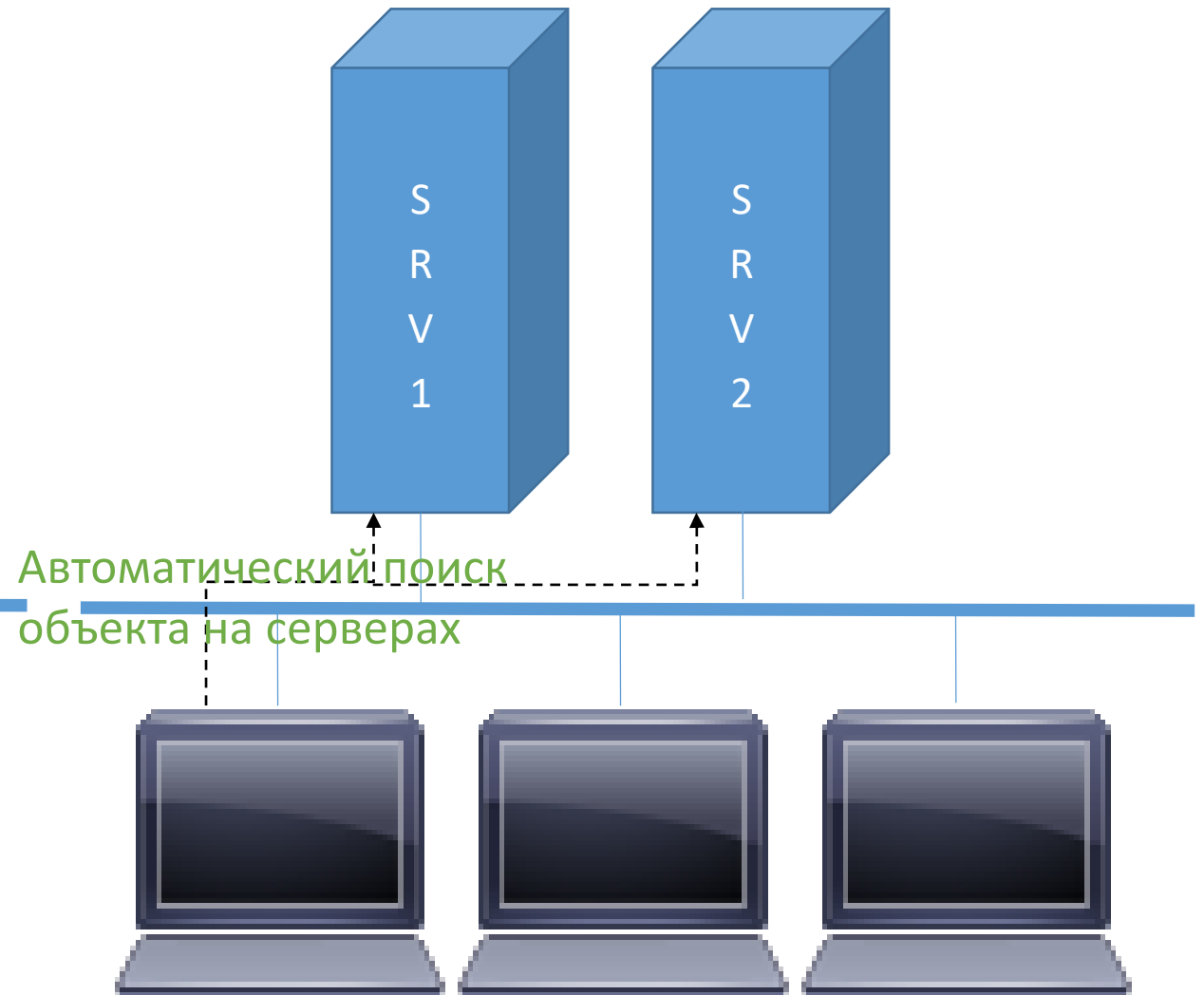
Сетевая файловая система



//SRV1/FOLDER/FILE

Явное указание местоположения объекта

Распределенная файловая система

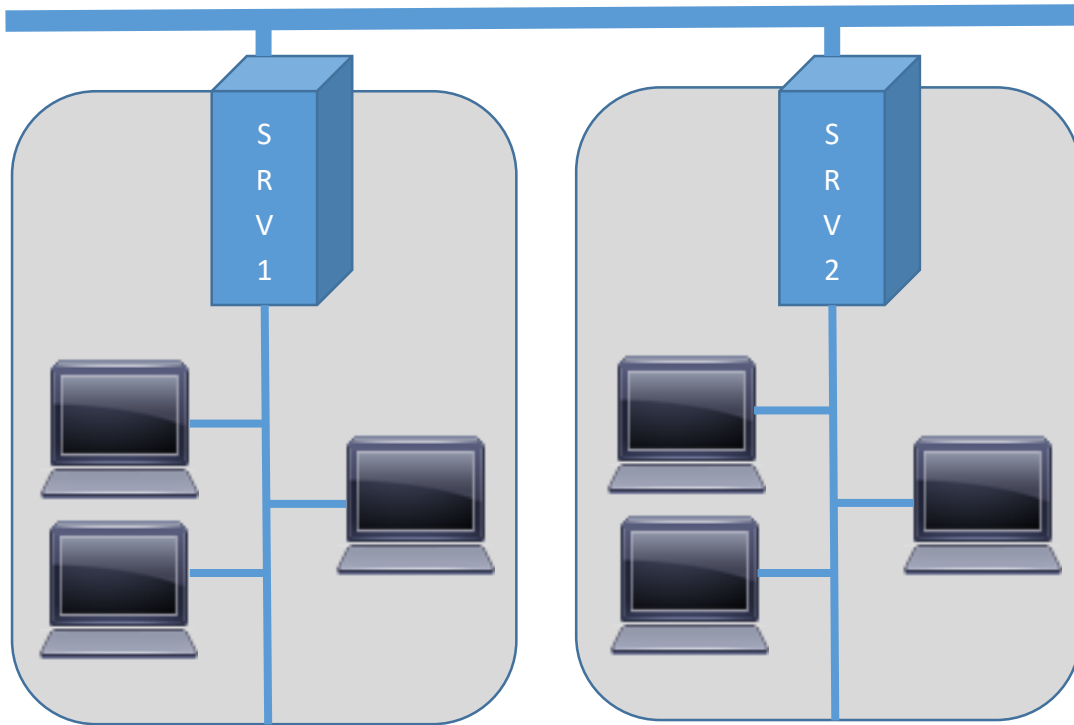


Автоматический поиск
объекта на серверах

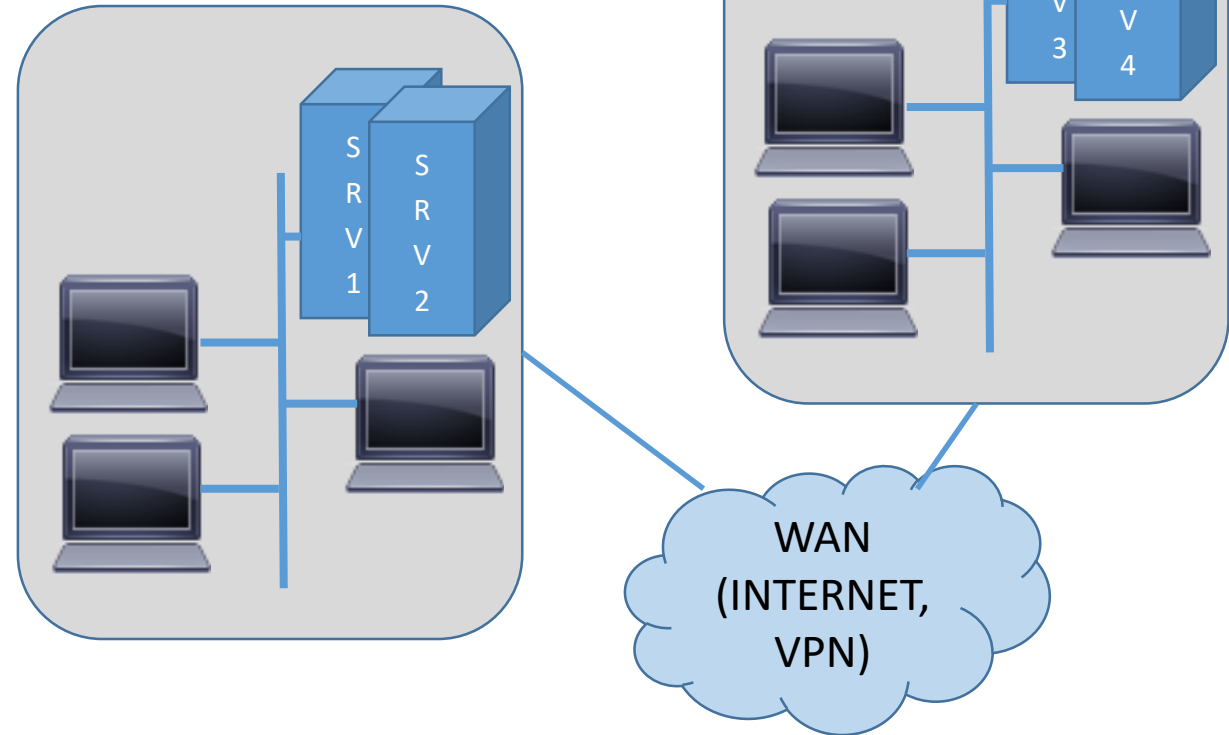
//NET/FOLDER/FILE

Распределенные файловые системы

AFS Andrew File System



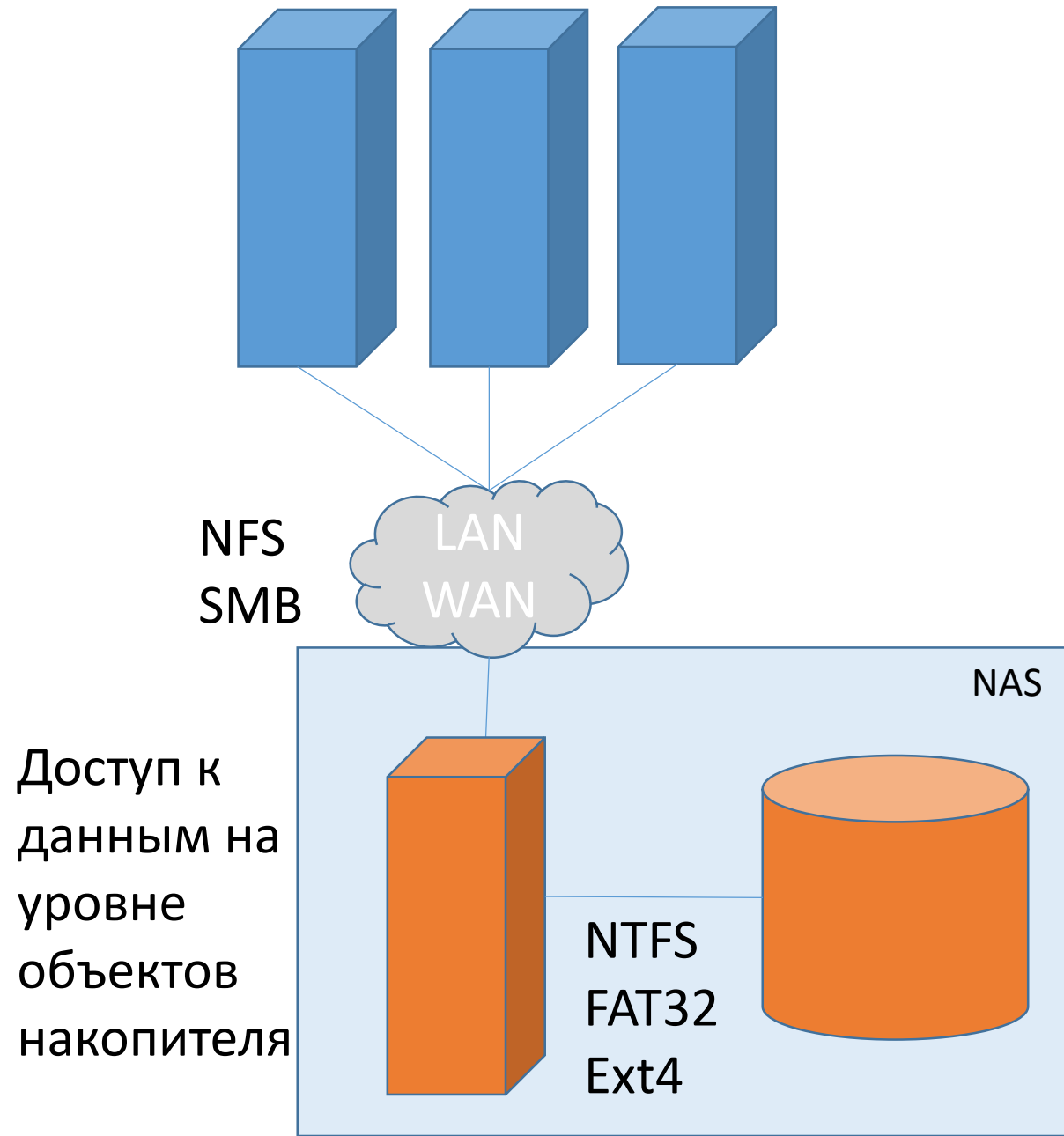
Distributed File System (DFS)



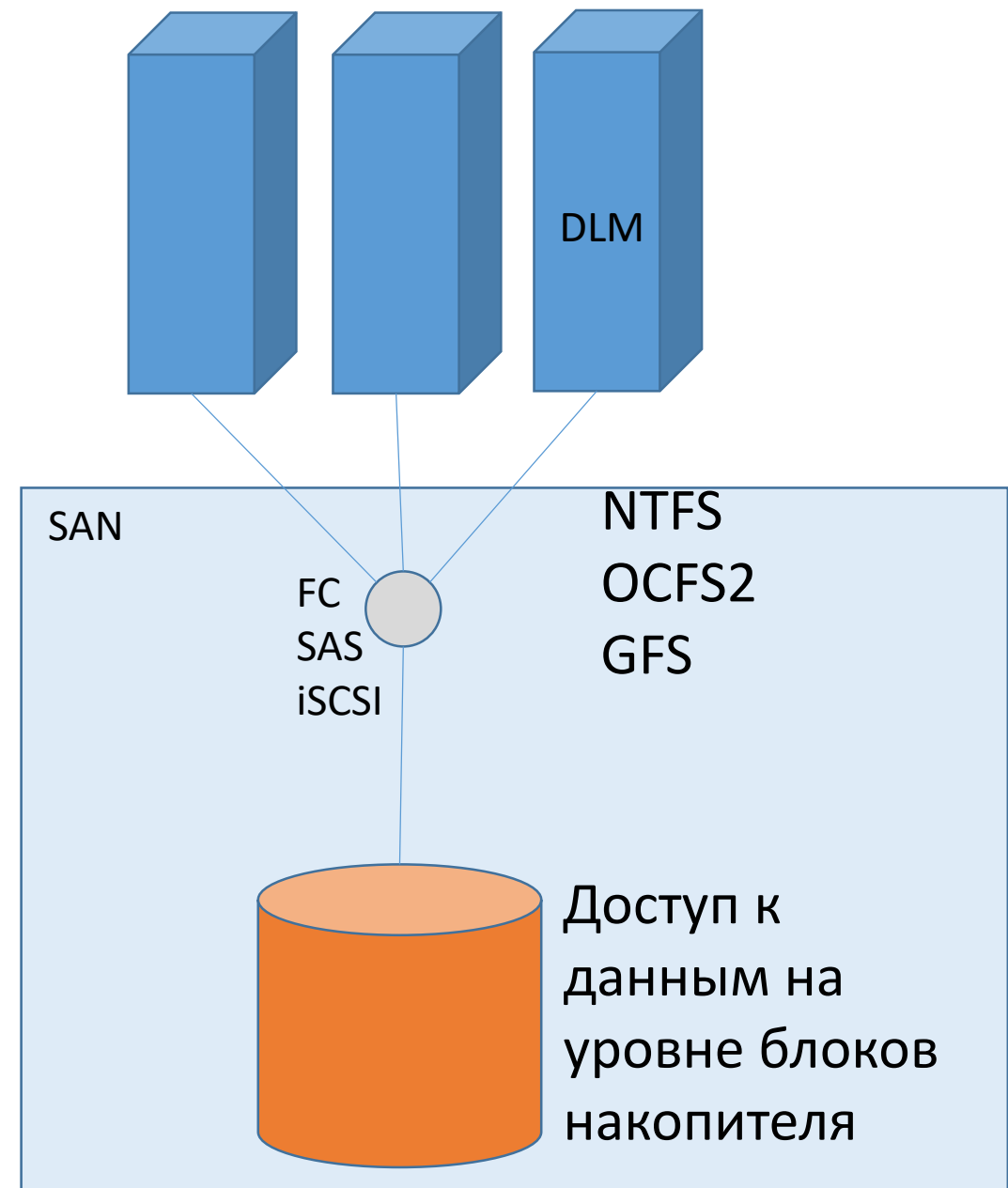
1. Отказоустойчивость
2. Балансировка нагрузки
3. Репликация файлов

NTFRS DFSR (Windows 2003R2)

Сетевые файловые системы



Кластерные файловые системы



Спецфайлы



0 способ – прямое программирование

1 способ – системные вызовы

2 способ – спецфайлы устройств

/dev

B (block) блок ориентированный

C (char) байт ориентированный

`mknod /dev/hhh c 99 1`

/dev/hhh – название

C – тип (c/b)

99 – major

1 - minor

«Особые» драйвера для /dev/null

Udev eudev + механизмы в ядре ОС

Первая необходимость plug-and-play

PCMCIA

CardBus



«Домашние каталоги»/профили пользователя

1. Локальный профиль
2. Перемещаемый профиль (Roaming)
 1. Необходима MS Active Directory
 2. При входе в ОС – профиль синхронизируется с сервером (скачивается)
 3. При выходе из ОС - профиль синхронизируется с сервером (передается на сервер)
 4. Преимущества:
 1. Пользователи могут переходить между компьютерами с автоматическим переносом личной информации
 2. Всегда есть копия на сервере (на случай поломки компьютера)
 5. Недостатки:
 1. На синхронизацию нужно время
 2. На сервере нужно МНОГО места
 3. Необходимо на всей сети иметь типизированное ПО
3. Обязательный
4. временный

Дисковые квоты

Используются на серверах

Жесткие квоты – нельзя превысить

Гибкие квоты – временно превысить

Windows2000 – квоты появились, NTFS v5

WindowsNT4, service pack6

Unix

- Ограничение по количеству объектов (количество i-node)
- Ограничение по суммарному объему объектов
 - Гибкие лимиты
 - Жесткие лимиты
- ☐ Либо по пользователям либо по группам пользователей

Неприменимо для каталогов перемещаемых профилей

Резервное копирование



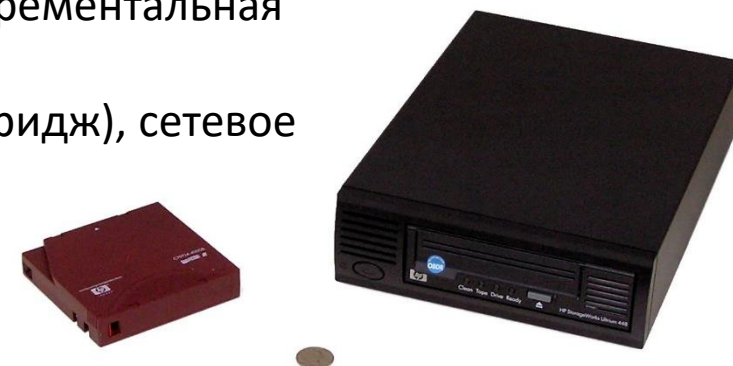
Для сохранения копии информации на определенный момент времени

Вопросы:

1. Срок хранения
2. Где хранить – в другом помещении/другом офисе, вопросы сохранности, доступа к копиям
3. Что архивировать – полная копия или конкретная информация (каталоги, базы данных); поблочная копия или логическое(по объектам) рез.копирование
4. Схема резервного копирования
 1. Полная копия + скорость восстановления, - требуется много места на устр.рез.копирования
 2. Дифференциальная копия + скорость восстановления + экономия места – с течением времени размер дифф. Копии стремится к размеру полной копии, - для восстановления нужна ближайшая полная + ближайшая дифф
 3. Инкрементальная копия + высокая скорость копирования – сложность восстановления – полная копия, ближайшая дифф и копии с момента полной(дифф)
 4. Вариант1 – 1р/нед «полная копия», ежедневно «инкрементальная копия»
 5. Вариант2 – 1р/мес «полная копия», еженедельно дифф копия, ежедневно инкрементальная
5. Промежуток резервного копирования – моменты наименьшей активности
6. Устройства резервного копирования – стримеры (Аннотс2021г – LTO9 – 45Тб/картридж), сетевое хранение в специализированном ПО/серверах

По резервного копирования: ArcServ, TapeWare, Acronis, Seagate backup (Veritas)

Не все можно восстановить. Например сервер-контроллер MS Active Directory



Журнализируемые файловые системы

Не журналируемые ФС:

- Много времени на восстановление после нештатной перезагрузки
- + нет накладных расходов на поддержку журналирования

Журналируемые ФС:

- + Практически моментальное восстановление после нештатной перезагрузки
- Чем подробнее ведутся журналы тем медленнее работают файловые системы, но тем выше вероятность успеха восстановления
 - Помогают восстановить актуальность метаданных о файлах
 - Помогают повысить вероятность успеха восстановления целостности (актуальности) файловой системы
 - При записи информации сначала записываем действия в журнал, а потом физически записываем данные

MS NTFS (\$LogFile), MS ReFS, MS ExFAT

Ext3/ext4

ReiserFS

IBM XFS

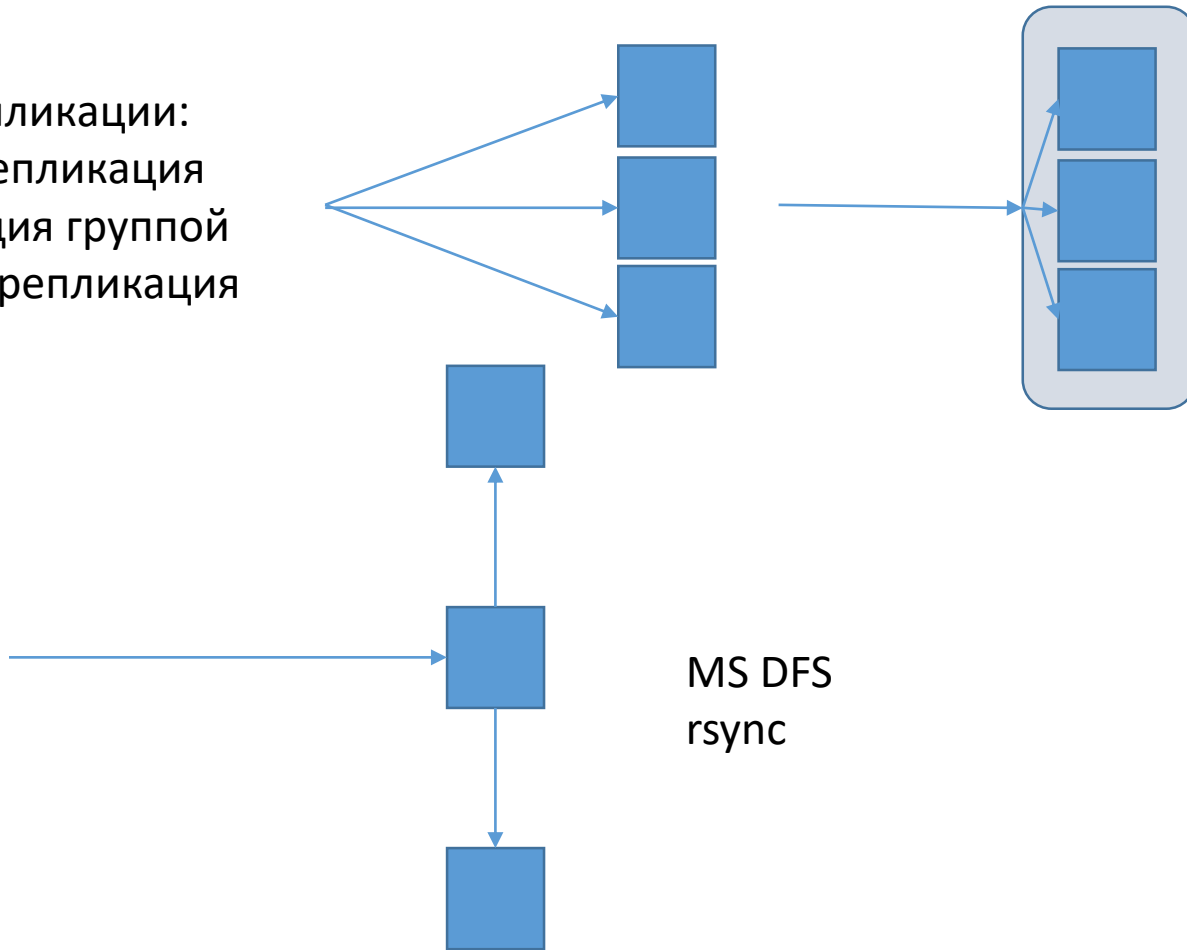
Репликация

Для повышения доступности информации:

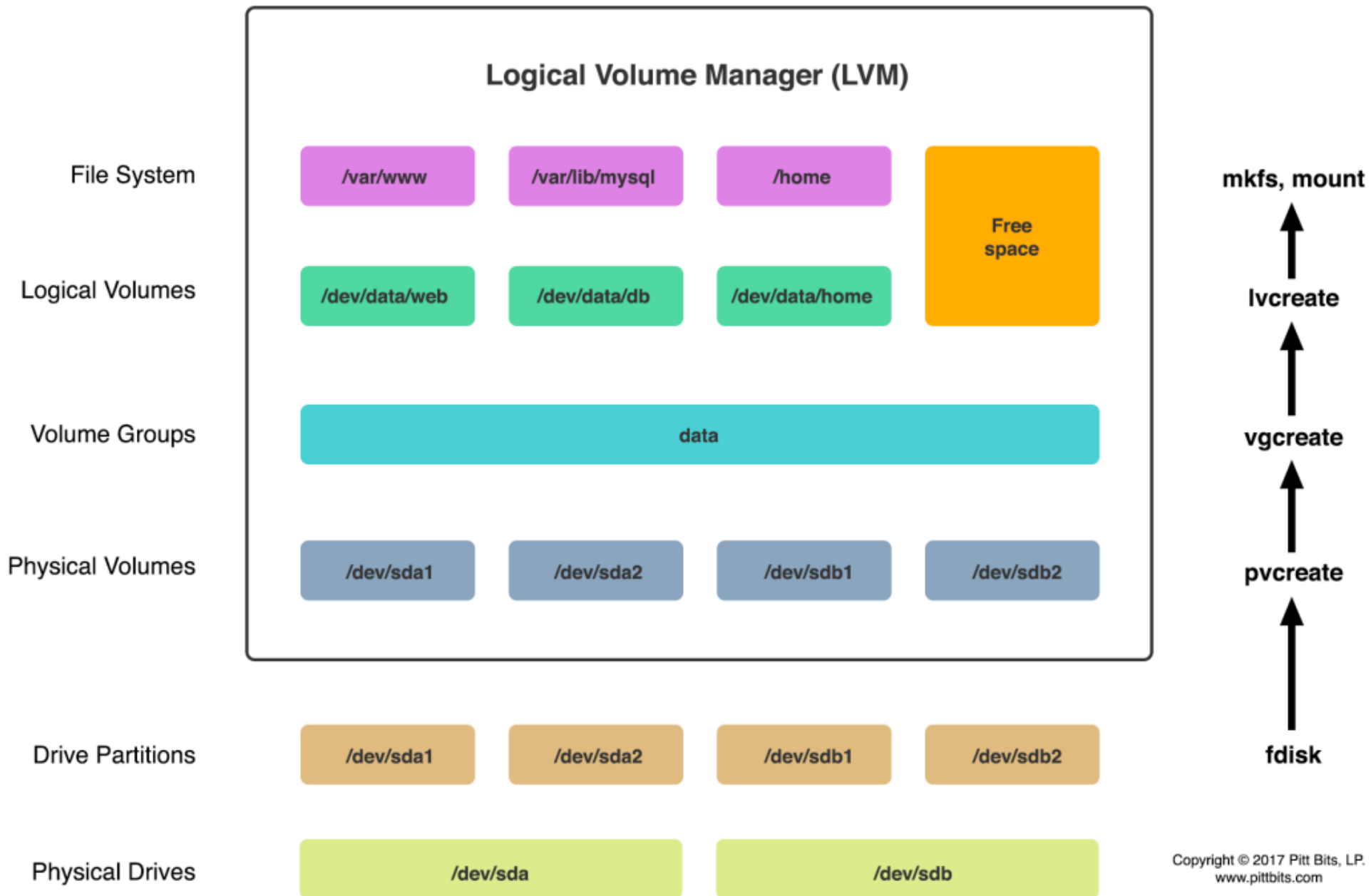
- Сбой одного сервера
- Распределение нагрузки
- Для больших объемов часть вычислений переносится на копию (ReadOnly)

Способы репликации:

1. Точная репликация
2. Репликация группой
3. Ленивая репликация



Менеджер логических томов (LVM, Logical Volume Manager)

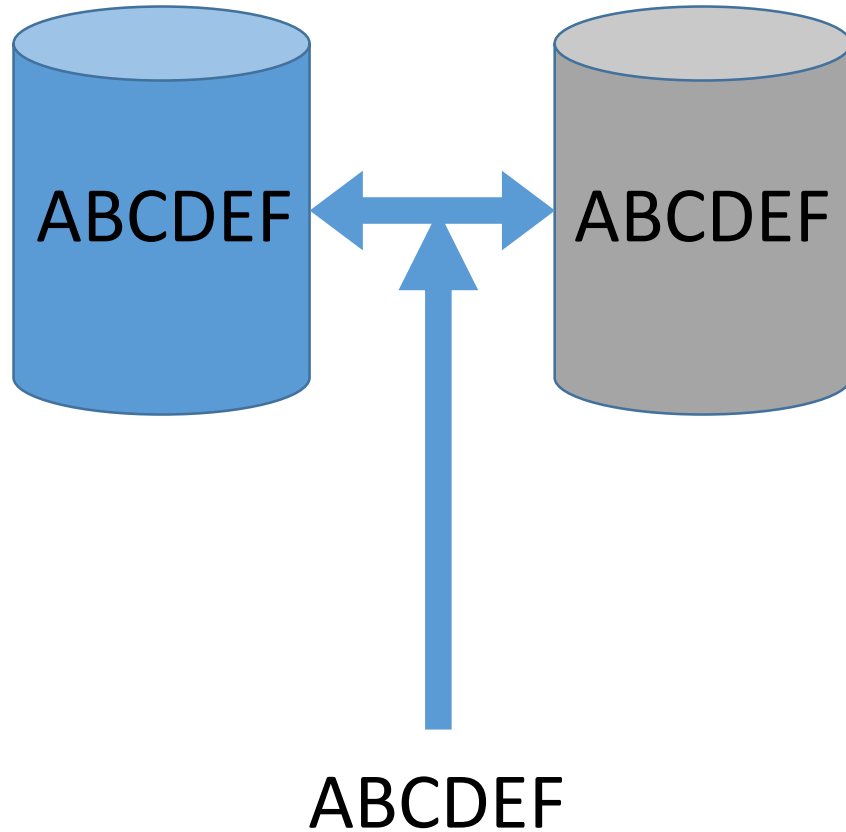


LVM добавляет
уровень
абстракции
между
физическими/
логическими
дисками и
файловой
системой

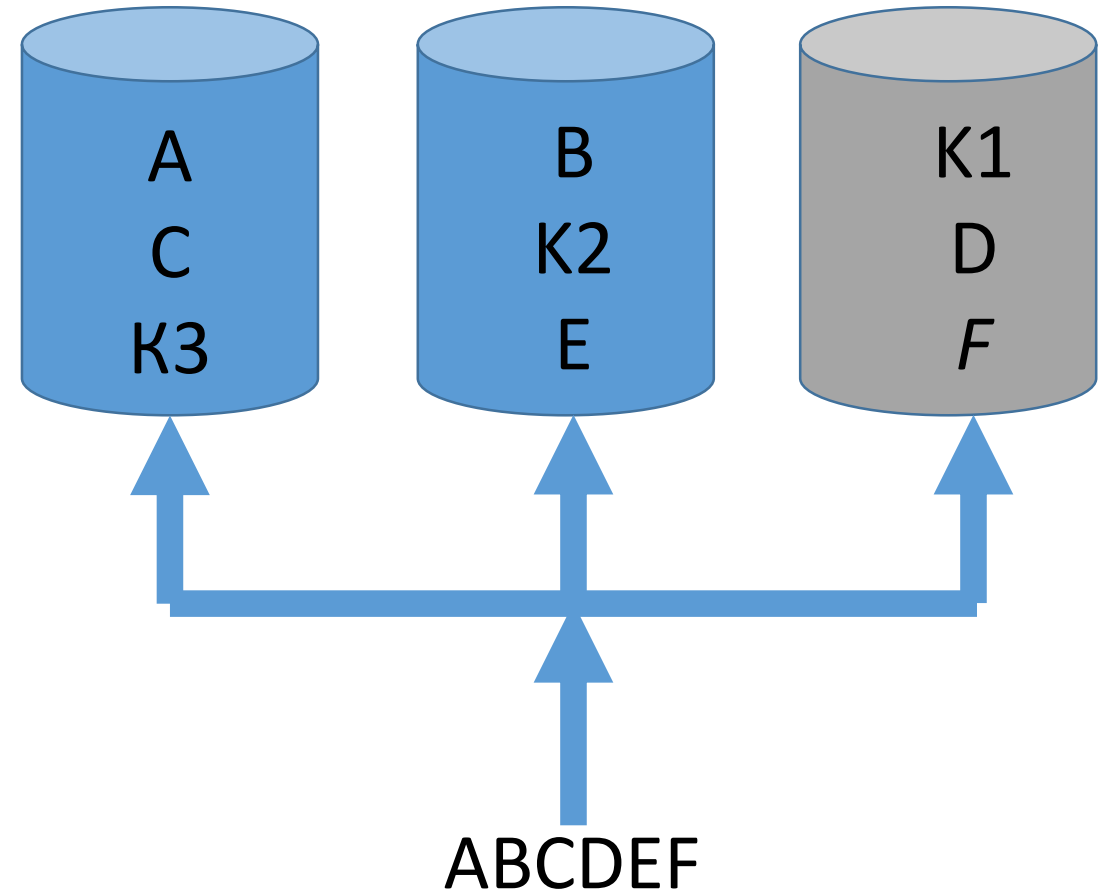
Отказоустойчивые массивы дисков (RAID)

Создаются за счет применения дополнительных дисков под хранение информации

RAID-1 (Mirror)



RAID-5

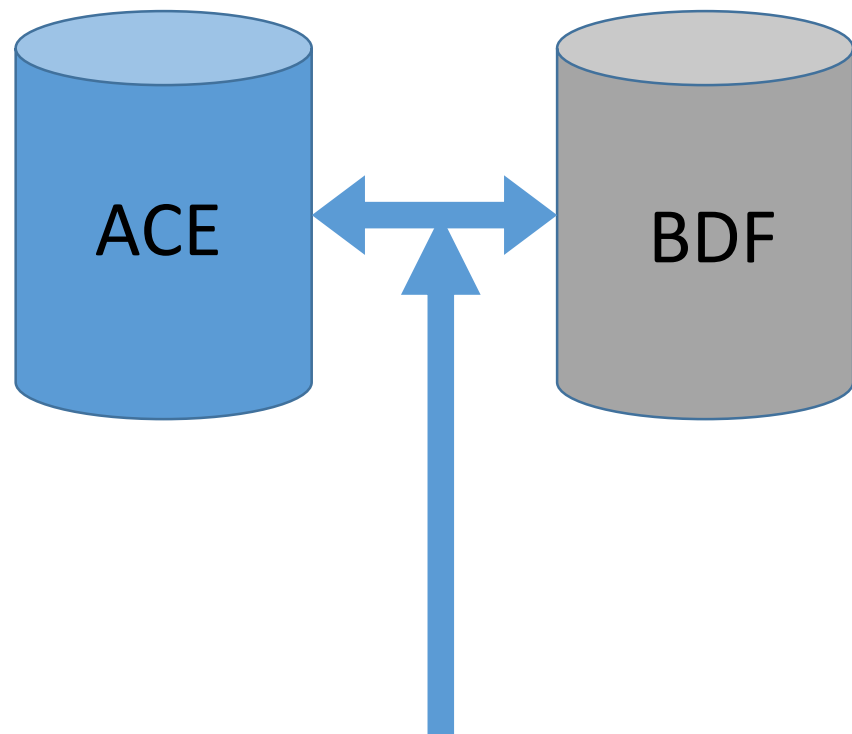


Raid-1..raid-5 защищают от одиночного сбоя
Raid-6 защищает от двойного сбоя

Отказоустойчивые массивы дисков (RAID)

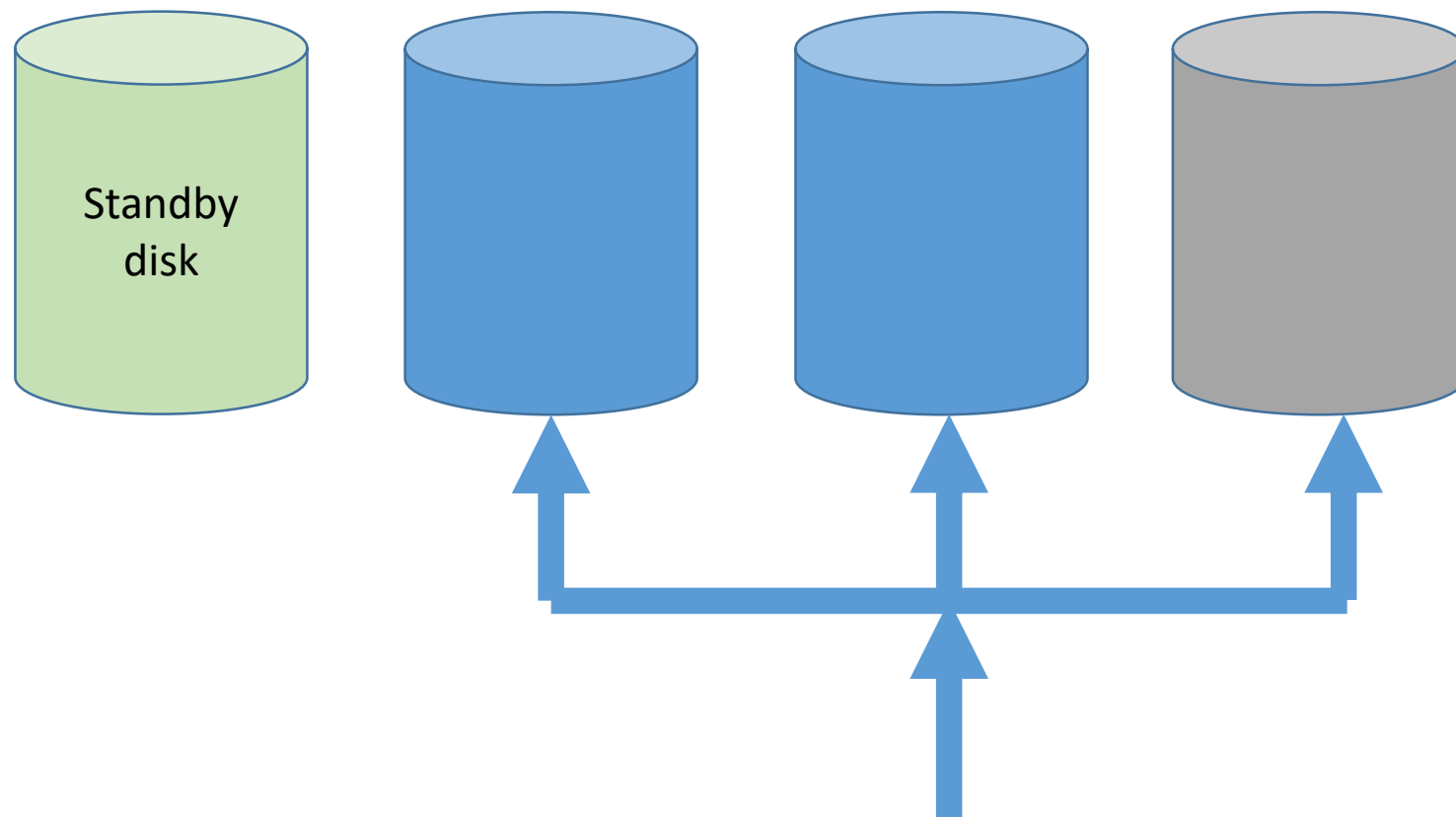
RAID10 = RAID-1 + RAID0

RAID-0



Данные не защищены

RAID с диском Standby



ABCDEF