



11010010



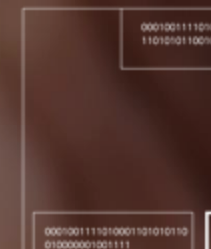
00010011



POCATOM

Файловые системы

Карапетьянц Николай



2022

СОДЕРЖАНИЕ

- Устройства хранения
- MBR vs GPT
- Раздел файловой системы
- Монтирование
- Swap
- Команды

УСТРОЙСТВА ХРАНЕНИЯ



Жесткий диск (HDD)

Магнитные диски



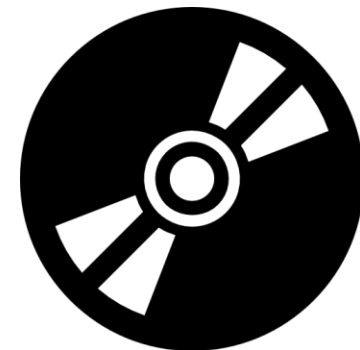
Твердотельный накопитель (SSD)

Флэш память



Оптический диск (CD, DVD)

Светочувствительный материал



ЖЕСТКИЙ ДИСК

Параметры:

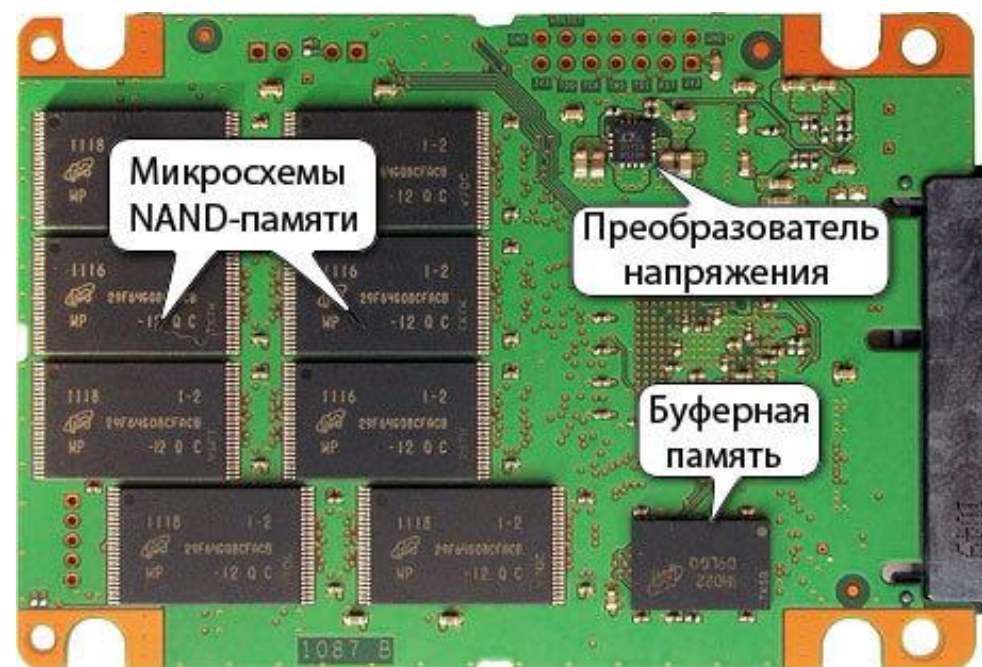
- Объем кэша
- Скорость вращения шпинделя
- Тип записи: PMR, SMR
- Ёмкость диска
- Уровень шума
- Скорость передачи данных
- Время доступа
- Ударостойкость
- Интерфейс: SATA(6 Гбит/с), SAS(12 Гбит/с)



ТВЕРДОТЕЛЬНЫЙ НАКОПИТЕЛЬ

Параметры:

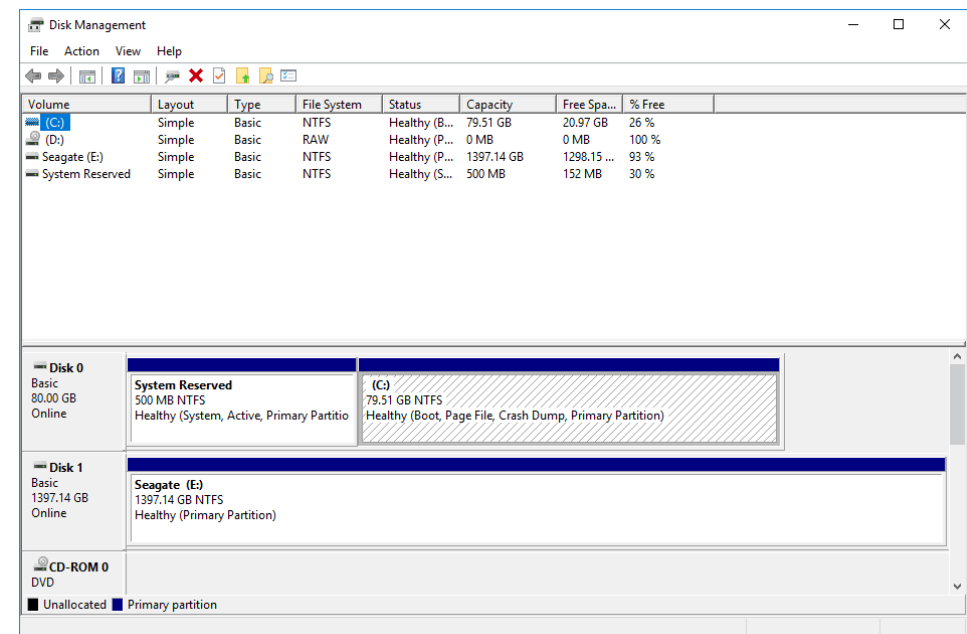
- Емкость накопителя
- Объем буфера
- Скорость чтения/записи(IOPS)
- Ресурс – TBW
- Контроллер
- Тип памяти: SLC, MLC, TLC, QLC
- Форм-фактор: M.2, U.2, 2,5"
- Интерфейс: PCI-E, SATA, SAS



СТРУКТУРА РАЗДЕЛОВ

Таблица разделов – таблица содержащая информацию о структуре распределения разделов на накопителе.

- GPT – GUID Partition Table – UEFI
- MBR – Master Boot Record – BIOS



The screenshot shows the Windows Disk Management console. At the top, a table lists the volumes on the system. Below this, the details for Disk 0 and Disk 1 are shown, including their partitions and their status.

Volume	Layout	Type	File System	Status	Capacity	Free Space	% Free
(C:)	Simple	Basic	NTFS	Healthy (B...	79.51 GB	20.97 GB	26 %
(D:)	Simple	Basic	RAW	Healthy (P...	0 MB	0 MB	100 %
Seagate (E:)	Simple	Basic	NTFS	Healthy (P...	1397.14 GB	1298.15 ...	93 %
System Reserved	Simple	Basic	NTFS	Healthy (S...	500 MB	152 MB	30 %

Disk	Partition	File System	Status
Disk 0 Basic 80.00 GB Online	System Reserved 500 MB NTFS Healthy (System, Active, Primary Partitio	NTFS	Healthy
	(C:) 79.51 GB NTFS Healthy (Boot, Page File, Crash Dump, Primary Partition)	NTFS	Healthy
Disk 1 Basic 1397.14 GB Online	Seagate (E:) 1397.14 GB NTFS Healthy (Primary Partition)	NTFS	Healthy

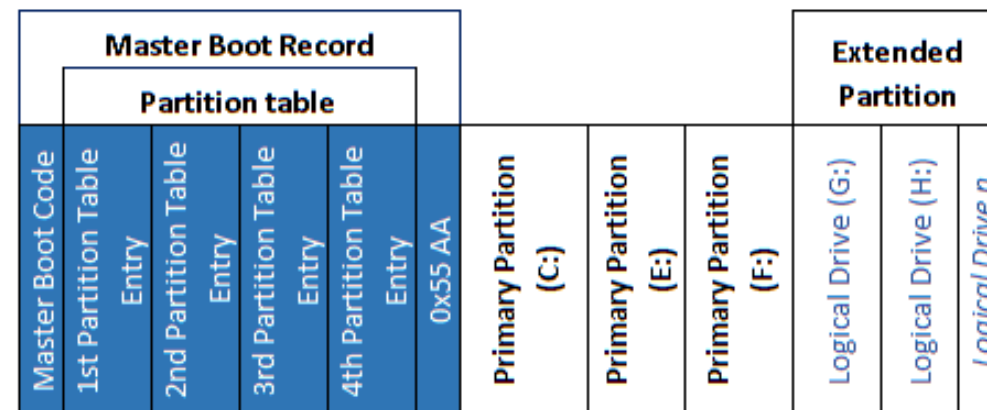
MBR

Преимущества:

- Совместим со старыми системами

Недостатки:

- Количество основных разделов: до 4 шт.
- Размер раздела: до 2 ТБ
- При повреждении главной загрузочной записи диск перестает быть читаемым



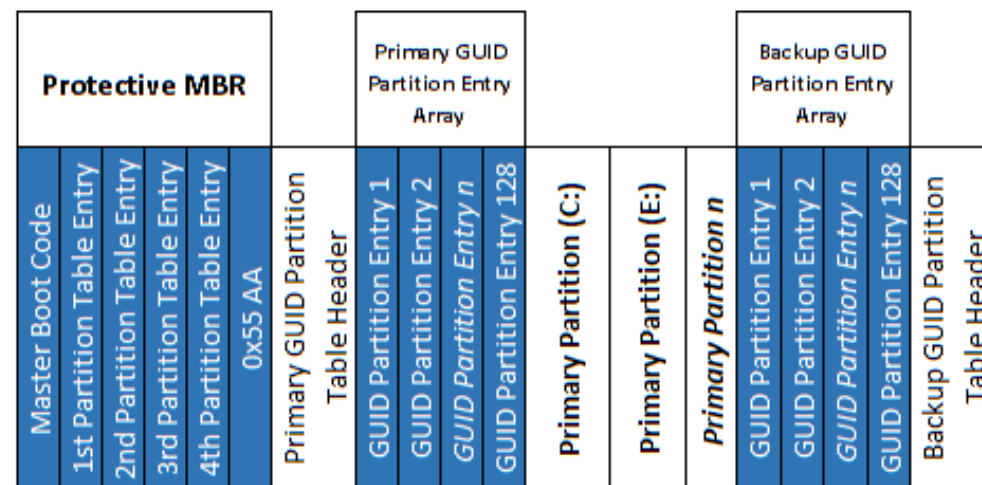
GPT

Преимущества:

- Неограниченное количество разделов
- Размер раздела: 9,4 зеттабайт
- Наличие возможности восстановить загрузочный раздел и данные
- Наличие контроля целостности
- Наличие поддержки Unicode

Недостатки:

- Несовместимость со старыми системами

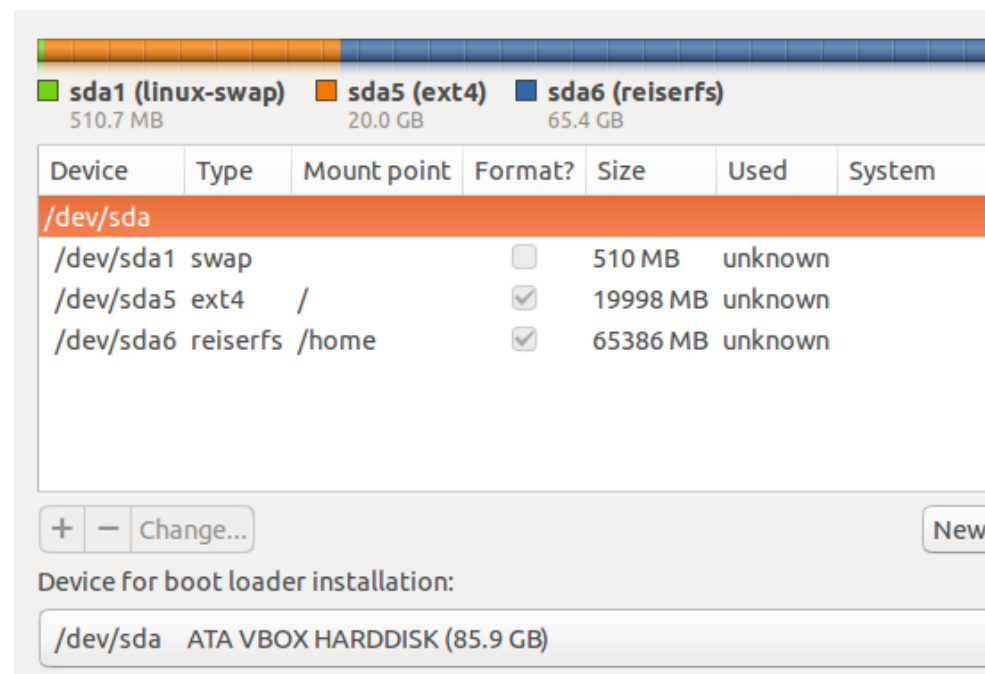


РАЗДЕЛЫ

Часть долговременной памяти жёсткого диска или флеш-накопителя, выделенная для удобства работы, и состоящая из смежных блоков.

Типы:

- Основной
- Расширенный – позволяет создавать внутри основного раздела несколько логических



The screenshot shows a disk partitioning interface. At the top, a progress bar indicates the status of three partitions: sda1 (linux-swap) in green, sda5 (ext4) in orange, and sda6 (reiserfs) in blue. Below this, a table lists the partitions with their device names, types, mount points, formats, sizes, and usage. The table is as follows:

Device	Type	Mount point	Format?	Size	Used	System
/dev/sda						
/dev/sda1	swap		<input type="checkbox"/>	510 MB	unknown	
/dev/sda5	ext4	/	<input checked="" type="checkbox"/>	19998 MB	unknown	
/dev/sda6	reiserfs	/home	<input checked="" type="checkbox"/>	65386 MB	unknown	

Below the table, there are buttons for '+', '-', and 'Change...', and a 'New' button. At the bottom, a section labeled 'Device for boot loader installation:' shows the selected device as '/dev/sda ATA VBOX HARDDISK (85.9 GB)'.

ФАЙЛОВЫЕ СИСТЕМЫ

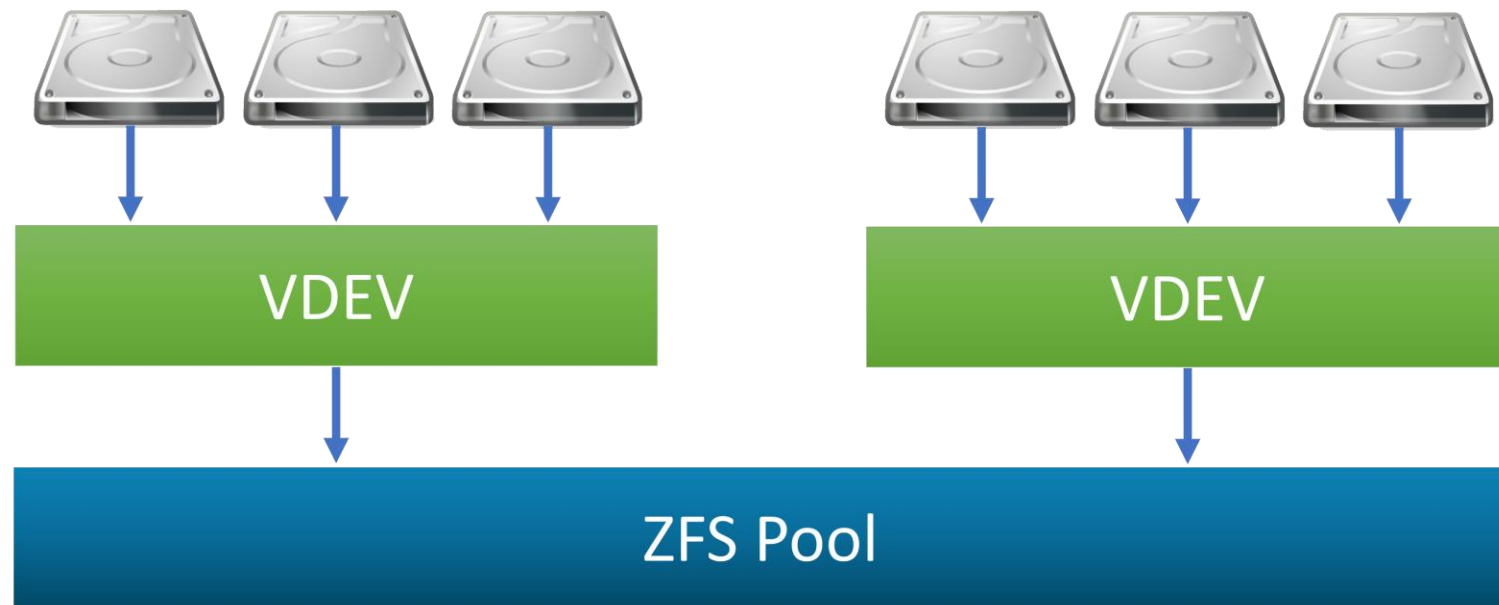
Локальные

EXT4
XFS
BTRFS
JFS
ReiserFS
ZFS

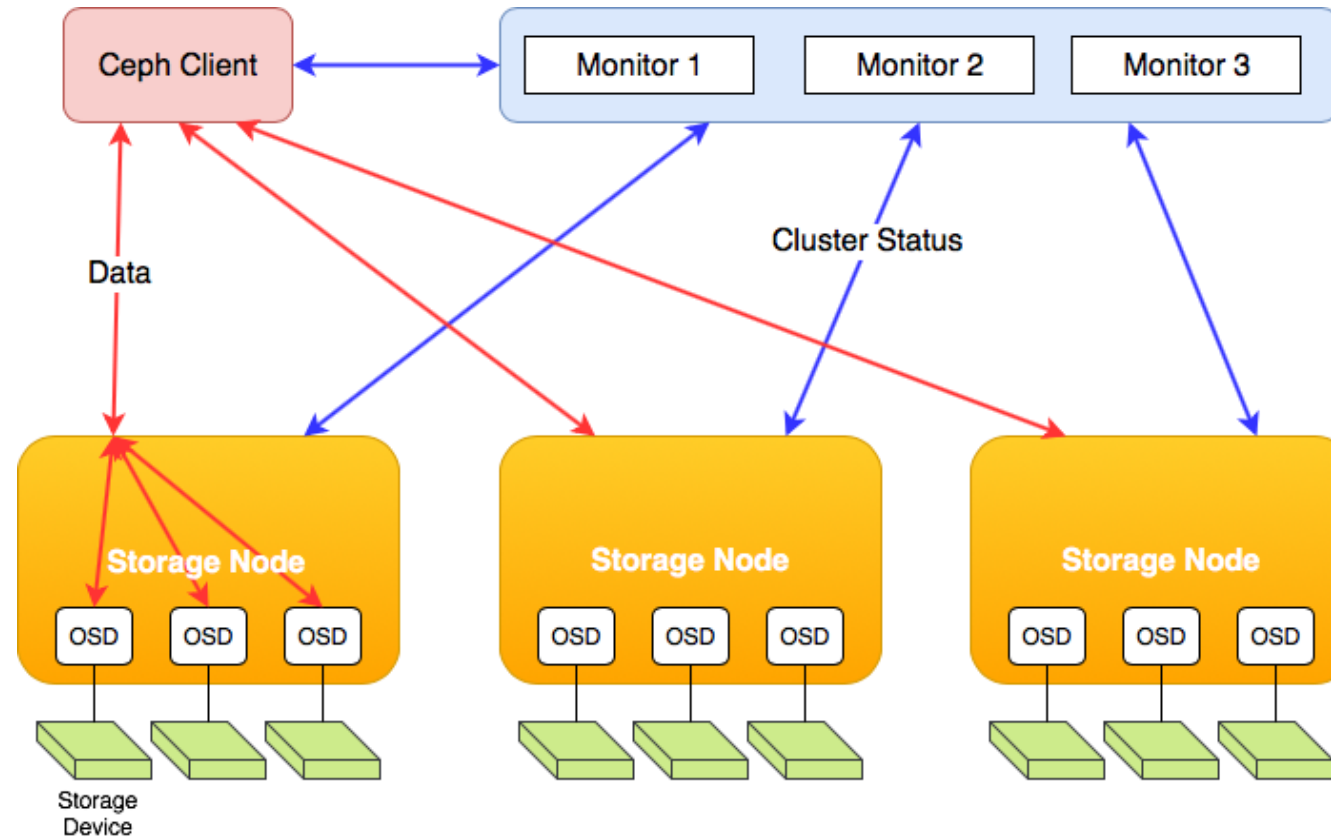
Удаленные

GlusterFS
CephFS
NFS
SMB

ZFS



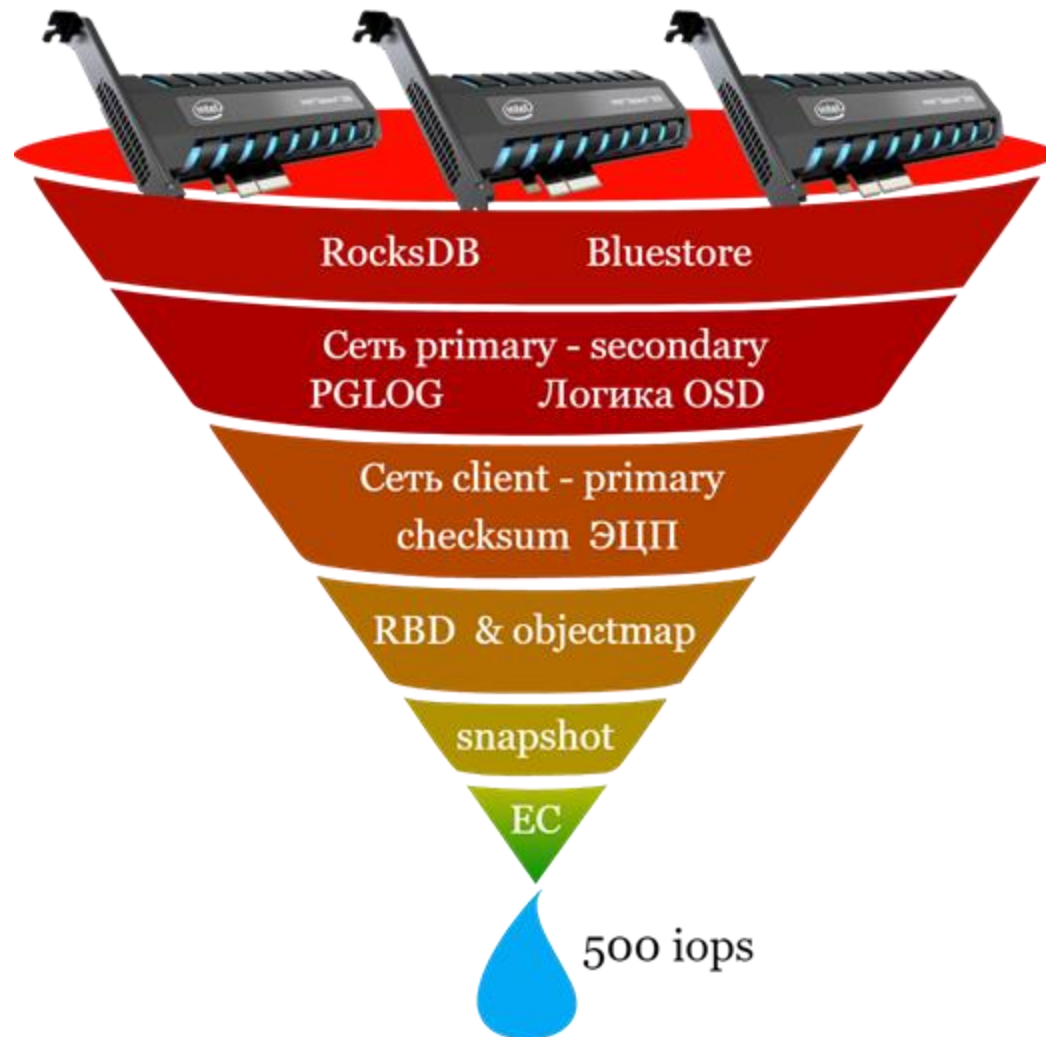
CEPHFS CLUSTER



ПРОБЛЕМЫ CEPHFS



Как это происходит в Ceph
 $300000 \text{ iops} * 3$

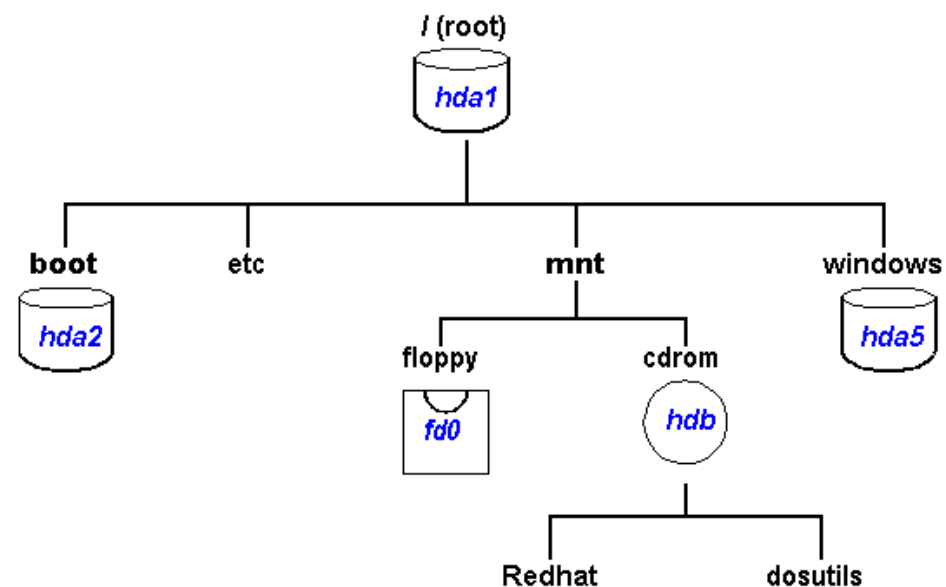


МОНТИРОВАНИЕ

Процесс, подготавливающий раздел диска к использованию операционной системой.

Этапы:

- Определение типа монтируемой системы
- Проверка целостности монтируемой системы
- Считывание системных структур и инициализация соответствующего модуля
- Включение файловой системы в общее пространство имен



FSTAB



/etc/fstab – содержит параметры монтирования блочных устройств, разделов на диске и удаленных файловых систем.

<device> - UUID или имя блочного устройства

<dir> - каталог в который будет перемонтирована файловая система

<type> - тип файловой системы

<option> - опции файловой системы

<dump> - поддержка создания резервной копии данных

<fsck> - поддержка проверки целостности файловой системы

```
# <device>                                <dir> <type> <options> <dump> <fsck>
UUID=0a3407de-014b-458b-b5c1-848e92a327a3 /      ext4   noatime   0        1
UUID=f9fe0b69-a280-415d-a03a-a32752370dee none    swap     defaults 0         0
UUID=b411dc99-f0a0-4c87-9e05-184977be8539 /home  ext4     noatime   0         2
```


SWAP

Пространство подкачки – расширение оперативной памяти за счет использования дискового пространства.

Представляет собой **раздел** или **файл**, в котором хранятся неиспользуемые данные из оперативной памяти.

Преимущества:

- Не допускает переполнения ОЗУ
- Включает поддержку гибернации

Недостатки:

- Занимает место на накопителе
- Уменьшает срок службы накопителя
- В некоторых случаях снижает производительность.

КОМАНДЫ



• Управление разделами:

- **fdisk** – разметка диска (поддерживает интерактивный режим) только MBR;
- **gdisk** - разметка диска GPT;
- **lsblk** – отображение информации об устройствах хранения;
- **blkid** – вывод информации о блочных устройствах.

• Управление файловыми системами:

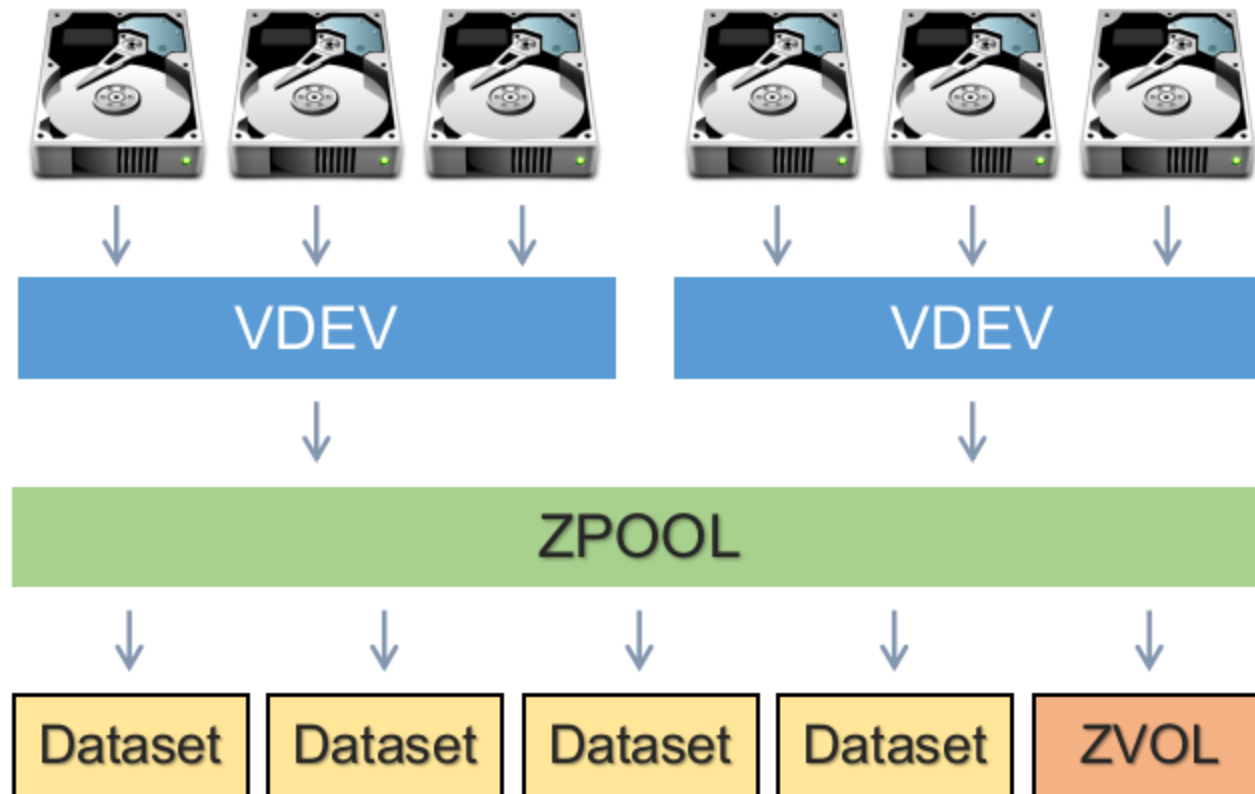
- **mkfs** – создание файловой системы;
- **mk2fs** – создание файловой системы ext2/ext3/ext4;
- **tune2fs** – изменение параметров файловой

системы ext2/ext3;

- **mount** – монтирование файловой системы;
- **umount** – размонтирование файловой системы;
- **findmnt** – список смонтированных файловых систем;
- **findfs** – найти файловую систему по метке или UUID;
- **dump** – резервная копия файловой системы: ext2/ext3;
- **dd** - низкоуровневое копирование файлов;
- **fsck**- проверка и восстановление файловой системы.



ZFS



ZFS VDEV

Single



VDEV

Mirror



VDEV

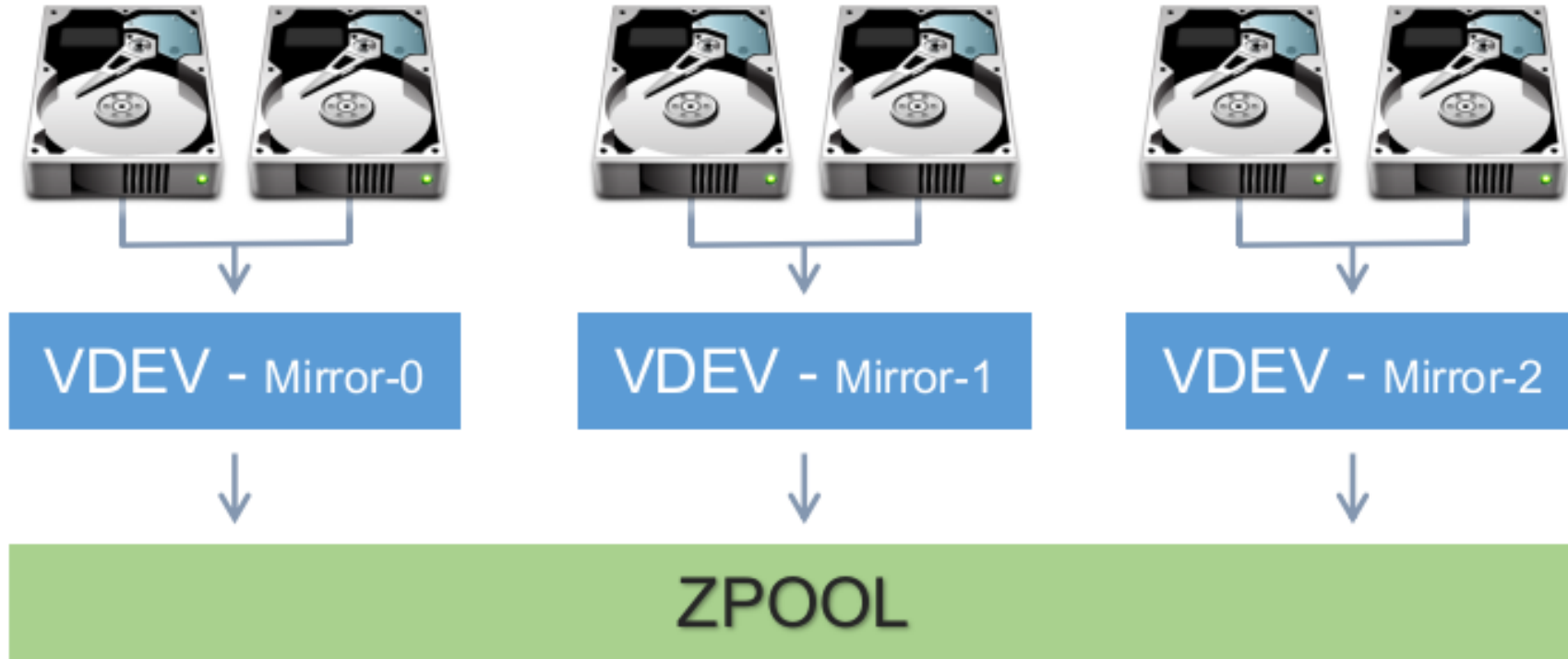
RAIDZ-1
















VDEV

ZFS RAID-10

RAID-10 (1+0)
Stripe of mirrors



ZFS RAID-10

RAID	Minimum Drives	Parity	Acceptable Drive Failure
Dynamic Stripe			
Mirror			
RAIDZ			
RAIDZ-2			
RAIDZ-3			

ССЫЛКИ



- Производительность Ceph:
https://yourcmc.ru/wiki/%D0%9F%D1%80%D0%BE%D0%B8%D0%B7%D0%B2%D0%BE%D0%B4%D0%B8%D1%82%D0%B5%D0%BB%D1%8C%D0%BD%D0%BE%D1%81%D1%82%D1%8C_Ceph
- Fstab:
[https://wiki.archlinux.org/title/Fstab_\(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9\)](https://wiki.archlinux.org/title/Fstab_(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9))
- Dump: <https://linux.die.net/man/8/dump>
- Fsck:
[https://wiki.archlinux.org/title/Fsck_\(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9\)](https://wiki.archlinux.org/title/Fsck_(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9))
- Swap:
[https://wiki.archlinux.org/title/Swap_\(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9\)](https://wiki.archlinux.org/title/Swap_(%D0%A0%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9))
- Резервное копирование файловой системы:
https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/storage_administration_guide/ext4backup
- ZFS RAID:
<https://blog.victormendonca.com/2020/11/03/zfs-for-dummies/>
- Quota:
<https://www.digitalocean.com/community/tutorials/how-to-set-filesystem-quotas-on-ubuntu-20-04>