

Системы хранения данных

Виктор Ашик для «Курса информационных технологий»

Яндекс

Память компьютера: классификация

Основная память

- регистры, кэш ЦПУ, ОЗУ

Дополнительная память

- Жесткие диски, SSD

Съемные накопители

- CD/DVD-RW

Роботизированные библиотеки со съемными накопителями

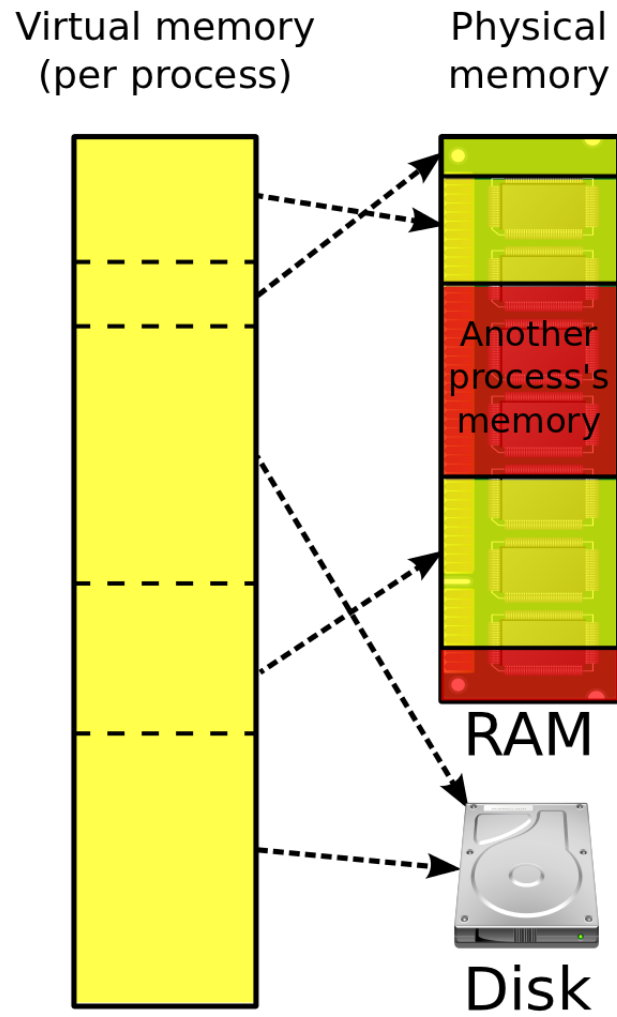
Оперативная память: термины

Назовите известные вам термины, относящиеся к оперативной памяти.

Оперативная память: термины

- Бит
- Байт
- Слово
- Сегмент
- Смещение
- Адрес
- Стек
- Куча
- Защита
- ...

Виртуальная память



Виртуальная память: термины

Назовите известные вам термины, относящиеся к виртуальной памяти.

Виртуальная память: термины

- Page (Страница), Huge Page, Large Page
- Physical Address (физический адрес)
- Page Table (таблица страниц)
- Memory Management Unit (модуль управления памятью)
- Page Fault (промах страницы), Page Fault Exception (исключение промаха страницы)
- Thrashing
- Working Set (рабочий набор)

Типы дополнительной памяти

DAS

NAS

SAN

Интерфейсы подключения накопителей

IDE/ATAPI

SCSI

SATA

SAS

FC

InfiniBand

Интерфейсы SAN

ATA over Ethernet (AoE)

Fibre Channel Protocol (FCP)

Fibre Channel over Ethernet (FCoE)

iSCSI

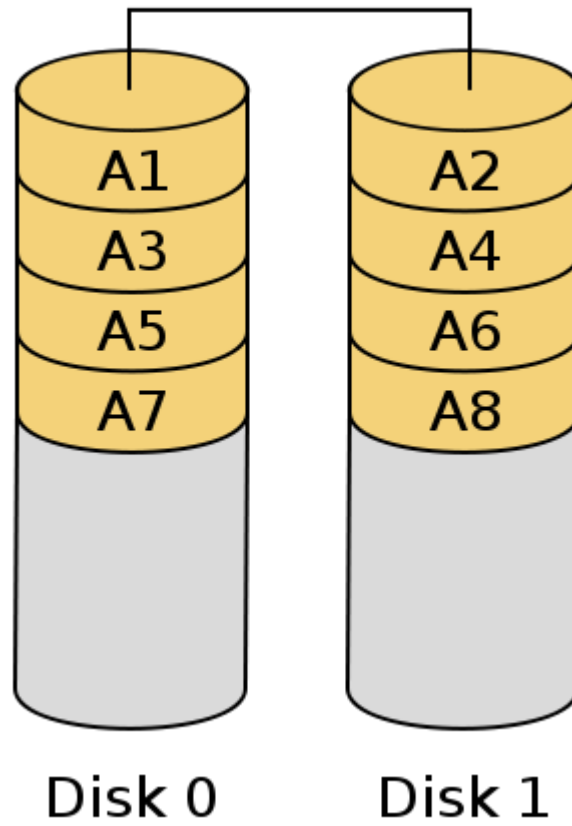
Fibre Channel

InfiniBand

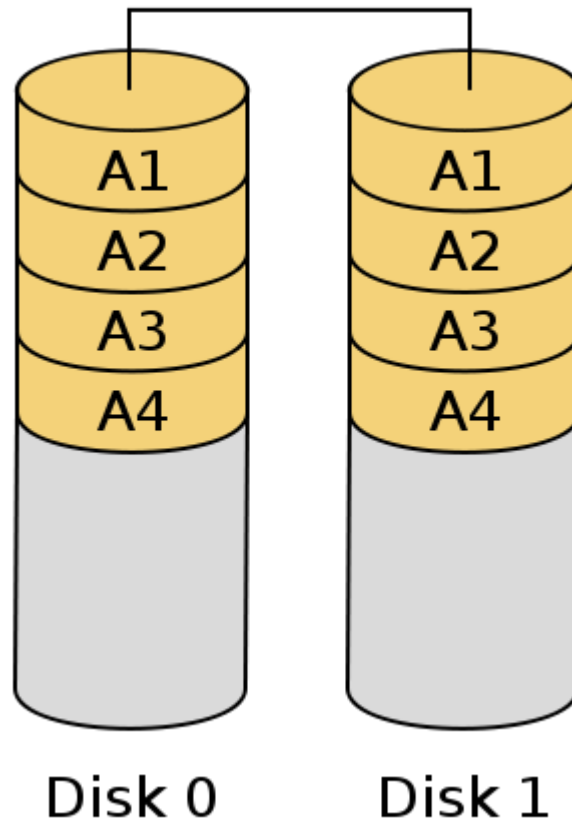
RAID

- 0 — striping — чередование блоков
- 1 — mirroring — зеркалирование
- 2 — чередование битов с кодом Хэмминга
- 3 — контрольный байт на отдельном диске
- 4 — контрольный блок на отдельном диске
- 5 — контрольный блок на одном из дисков (с чередованием)
- 6 — 2 контрольных блока (с чередованием)

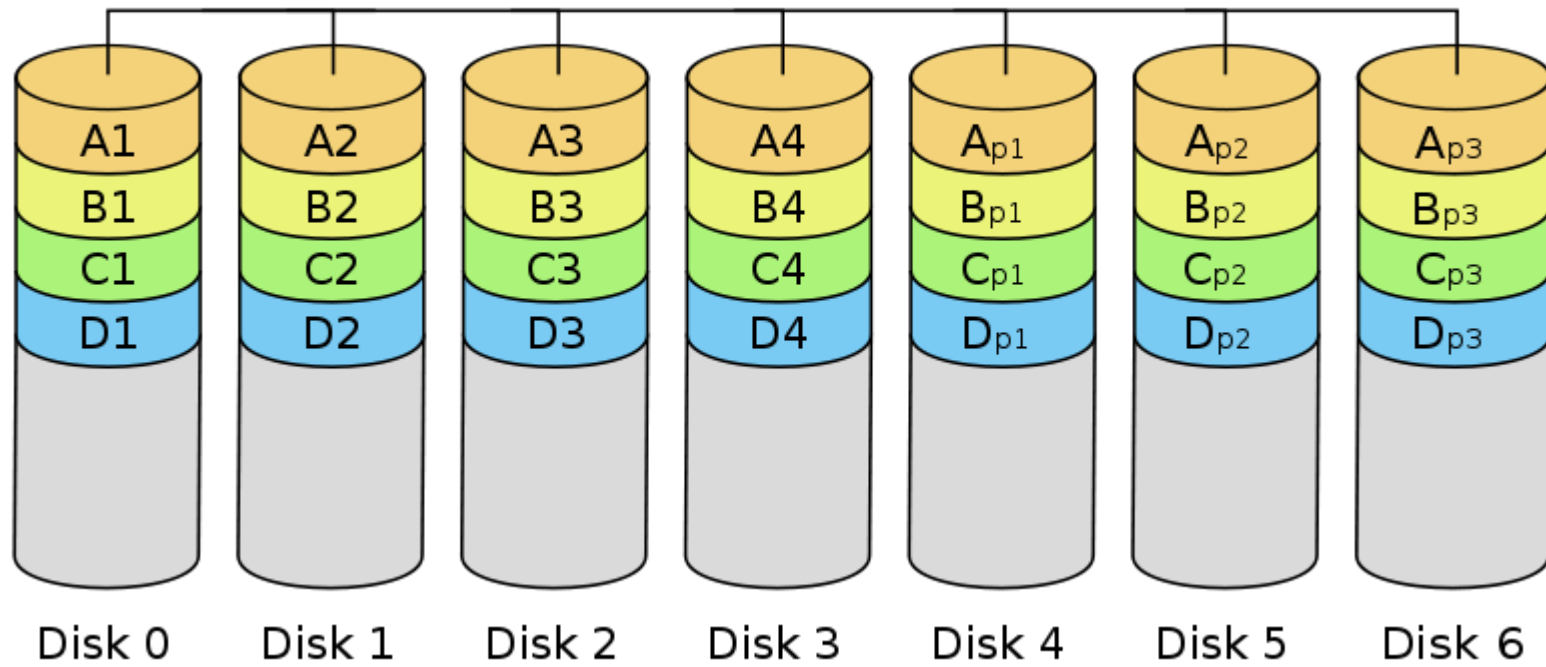
RAID 0



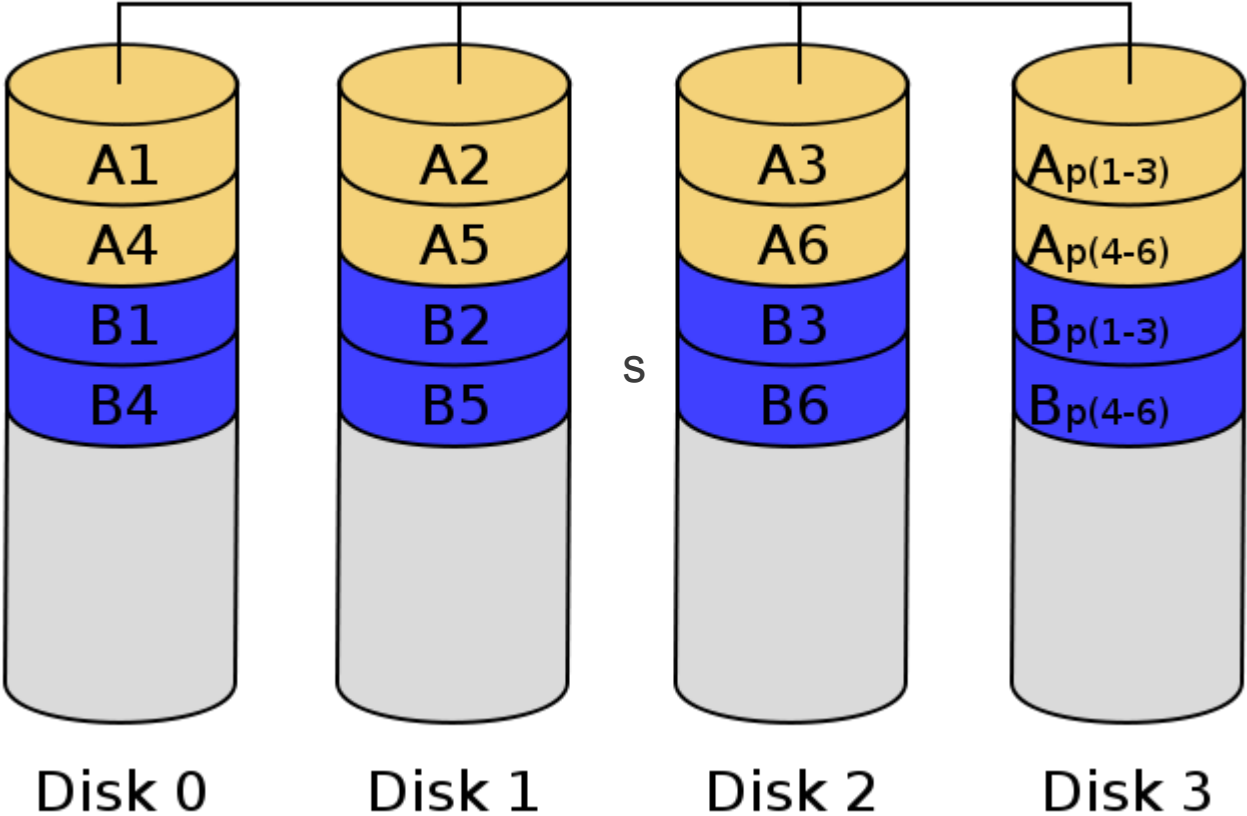
RAID 1



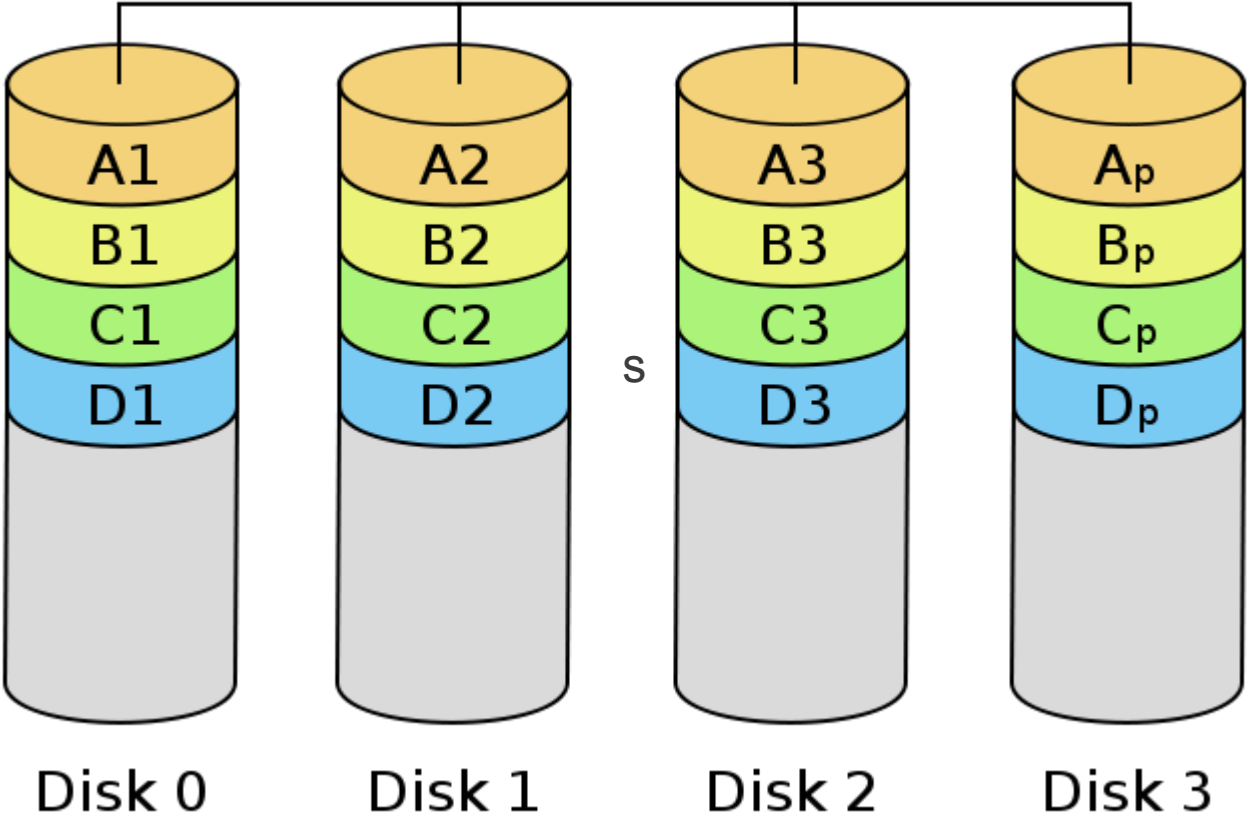
RAID 2



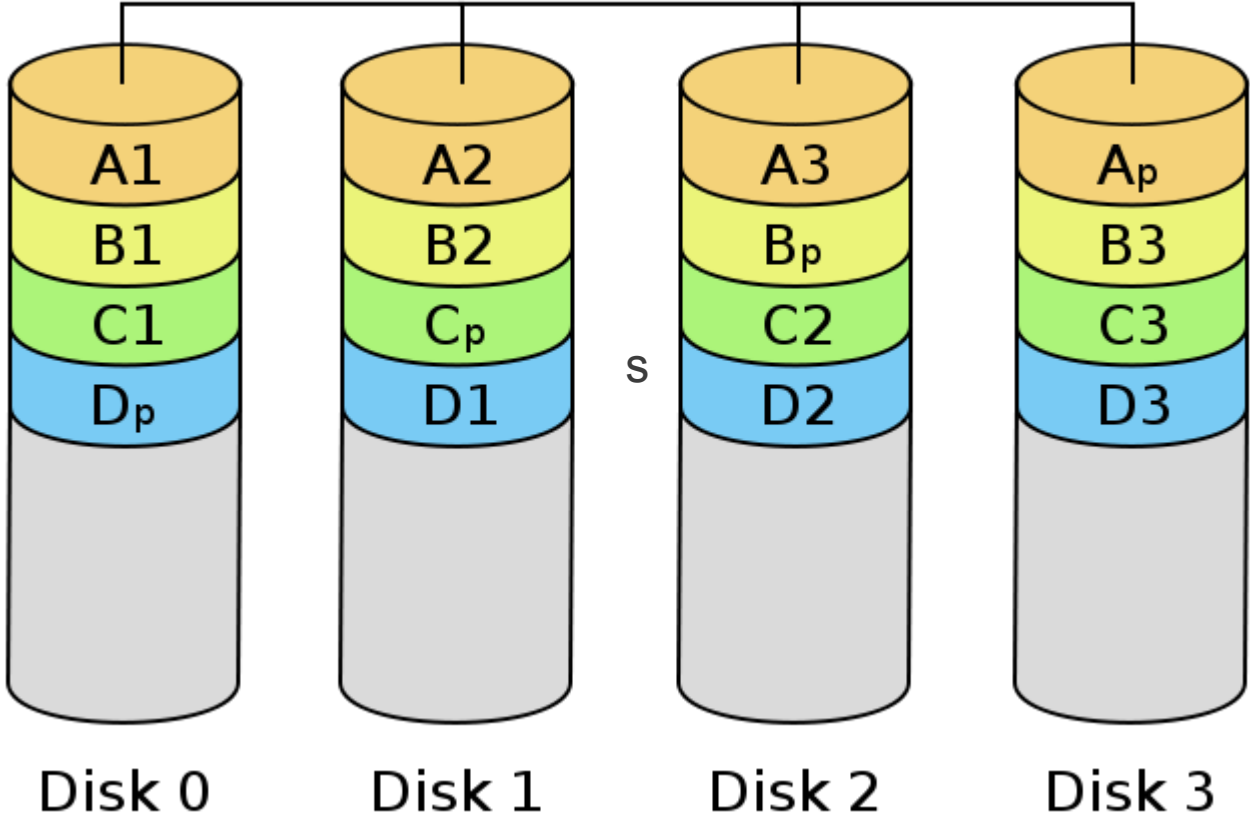
RAID 3



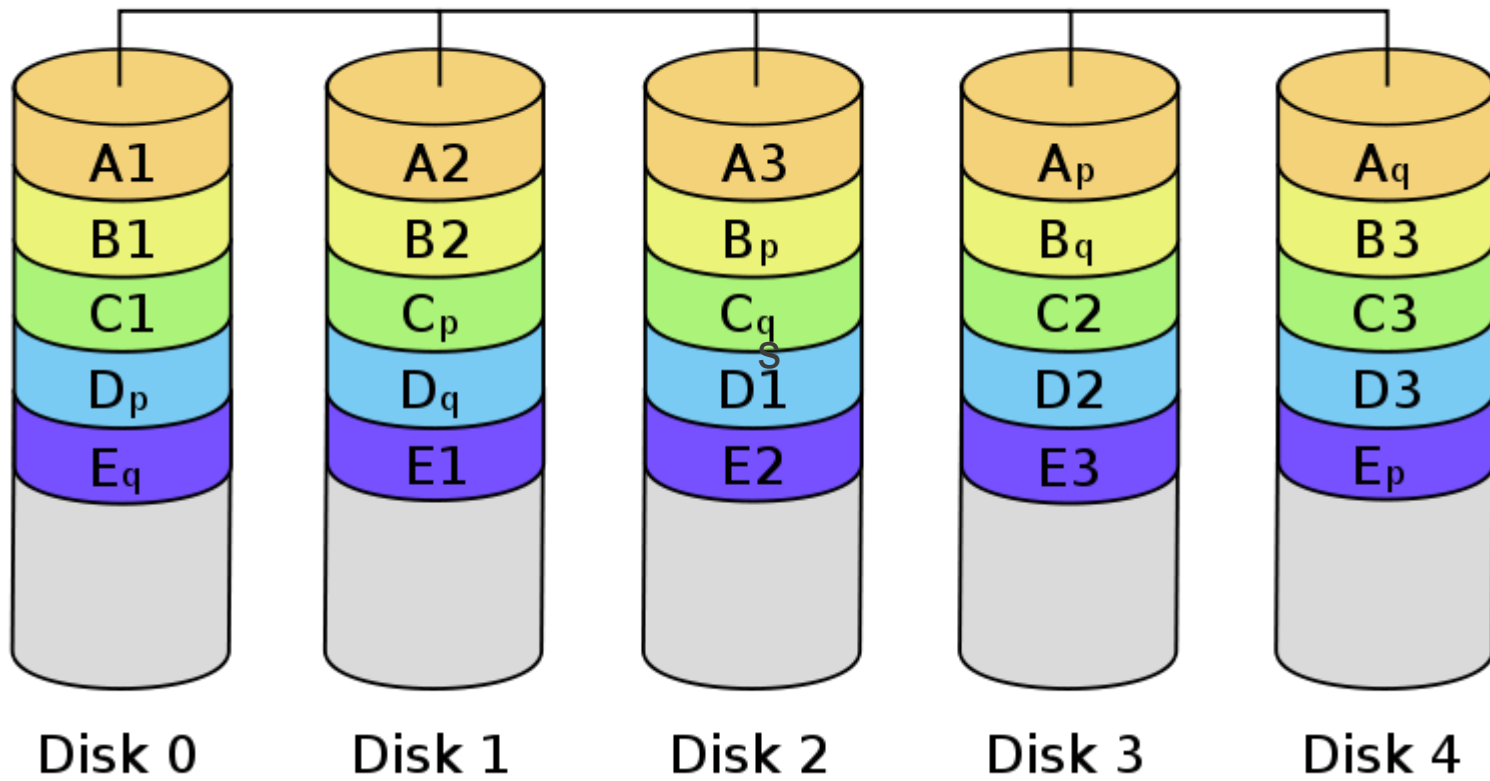
RAID 4



RAID 5



RAID 6



XOR

$$0 \text{ XOR } 0 = 0$$

$$0 \text{ XOR } 1 = 1$$

$$1 \text{ XOR } 0 = 1$$

$$1 \text{ XOR } 1 = 0$$

Пример

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: 11110111 (Data)
Drive #4: 10110101 (Data)
Drive #5: - - - - - (Hot Spare)
Drive #6: - - - - - (Parity)

Пример

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: 1110111 (Data)
Drive #4: 10110101 (Data)
Drive #5: - - - - - (Hot Spare)
Drive #6: 11100110 (Parity)

Пример

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: - - Dead - - (Data)
Drive #4: 10110101 (Data)
Drive #5: - - - - - (Hot Spare)
Drive #6: 11100110 (Parity)

Пример

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: - -Dead - - (Data)
Drive #4: 10110101 (Data)
Drive #5: 11110111 (Hot Spare)
Drive #6: 11100110 (Parity)

Пример

Drive #1: 00101010 (Data)
Drive #2: 10001110 (Data)
Drive #3: - - - - - (Hot Spare)
Drive #4: 10110101 (Data)
Drive #5: 11110111 (Data)
Drive #6: 11100110 (Parity)

Комбинированные RAID

RAID

- 0+1 – зеркало из чередующихся томов
- 1+0 – чередование зеркал
- 5+1 – зеркало из массивов RAID5

Реализация хранения в Linux

Блочные устройства: /dev/sda, /dev/sdb, ...

Разбиение на разделы: /dev/sda1, /dev/sda2

- MBR
 - основные: 1,2,3,4
 - логические: 5,6,...,15
- GPT: 1..128
- ...

Файловая система

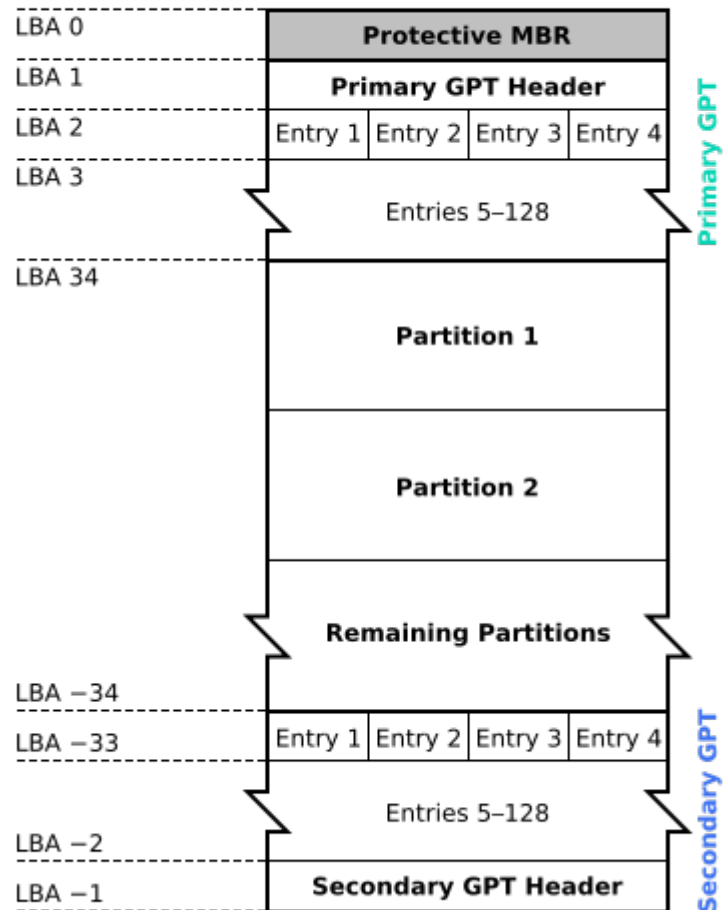
Структура хранения и упорядочения файлов

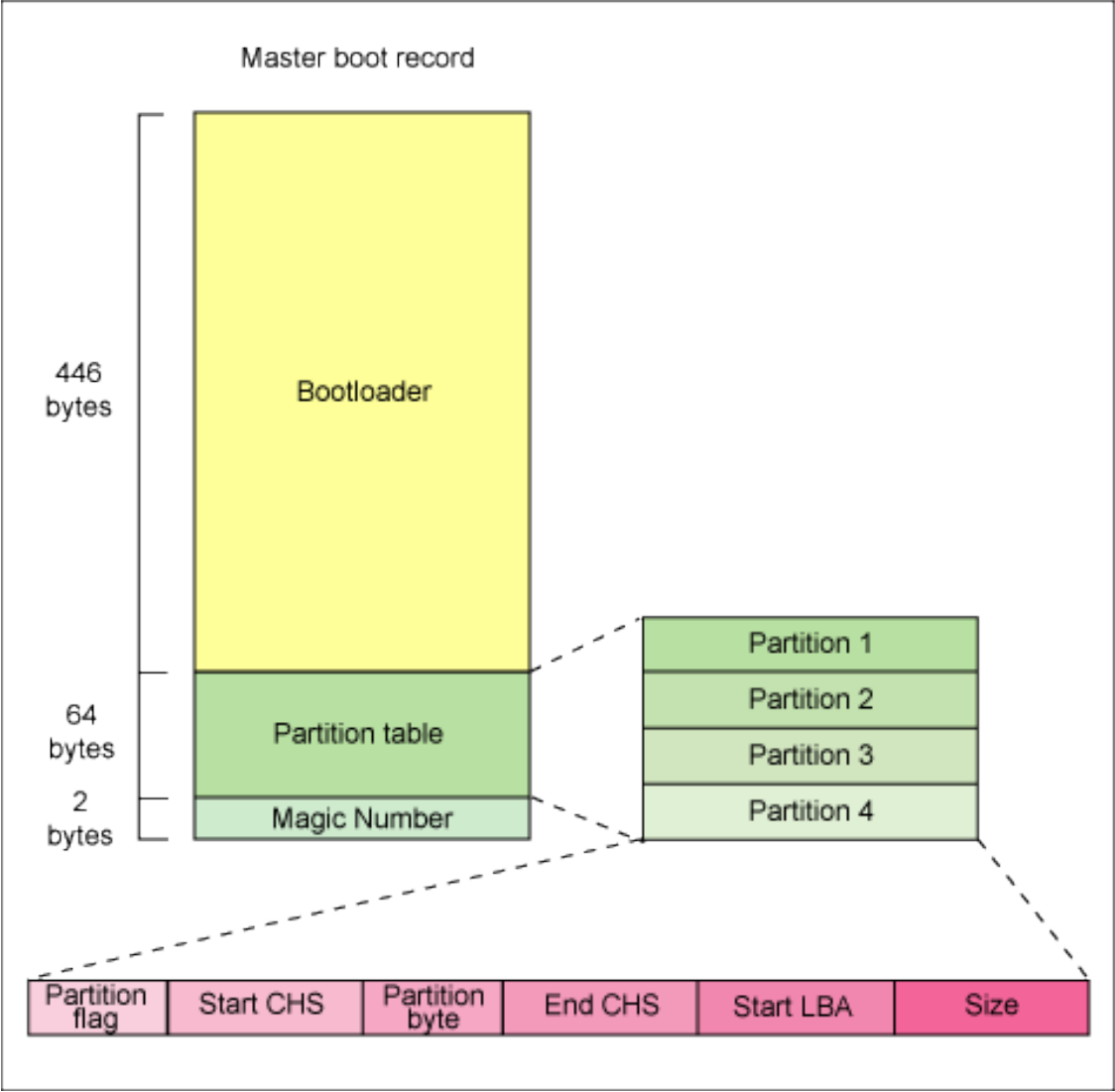
- имена файлов
- метаданные
- ссылки
 - жесткие
 - символические
 - ..
- типы файлов
- разрешения доступа

Реализация хранения в Linux

- Блоковые устройства: /dev/sda, /dev/sdb, ...
- Разбиение на разделы: /dev/sda1, /dev/sda2
 - MBR
 - основные: 1,2,3,4
 - логические: 5,6,...,15
 - GPT: 1..128
 - ...

GUID Partition Table Scheme





Недостатки разделов

Трудно увеличивать и уменьшать

Должны быть непрерывными

Привязаны к диску — неудобно переносить

Менеджер логических томов LVM2

Возможности LVM

Зеркалирование и чередование

Увеличение и уменьшение томов

Перенос между физическими дисками

Моментальные снимки для чтения и записи

Поддержка multipath

Команды LVM2

lvm2

pv(createldisplaylremovelresizelscanlmove...)

vg(createlchangeldisplaylscanlremovelrename...)

lv(createldisplaylextendlchangelrenamel...)

Программный RAID в Linux

mdadm

- RAID 0, 1, 4, 5, 6, 10, mulipath

```
mdadm --create /dev/md_d0 --level=10 --raid-devices=14 --chunk=2048 --assume-clean  
--metadata=1.0 --name=data --auto=mdp /dev/sd[abcdefghijklmn]1
```

```
mke2fs -b 4096 -R stride=512 -j -J size=384 -m 0 -O dir_index /dev/md_d0
```

#yakit2

- 03.10 Что такое быть системным администратором.
- 05.10 Операционные системы, их история, устройство и функционирование.
- 10.10. Архитектура ЭВМ и интерфейсы периферийного оборудования.
- 12.10 Устройство GNU/Linux на примере Debian/Ubuntu и Fedora/RHEL (Экстрополис).
- 17.10 Управление хранением данных: RAID, LVM, резервное копирование и восстановление.
- 19.10 Системы управления базами данных на примере MySQL.
- 24.10 Экскурсия. Сети, протоколы и сетевое оборудование.
- 26.10 Управление хранением данных: файловые системы (Экстрополис).
- 31.10 Информационная безопасность и средства её обеспечения.
- 02.11 Виртуализация вычислительных ресурсов.
- 07.11 Экзамен

Виктор Ашик
куратор практик

vashik@yandex-team.ru
+7 495 739-70-00

ул. Льва Толстого, 16
Москва, Россия, 119021
Яндекс Москва

The logo for Yandex, featuring a red Cyrillic letter 'Я' followed by the word 'ндекс' in black.