

Распределенные системы хранения и обработки данных

Владислав Белогрудов, EMC

vlad.belogrudov@gmail.com

Лекция 3

Защита данных с помощью RAID

Содержание лекции

- Методы и техники
- Типы
- Производительность
- Сравнение и области применения

Немного истории

- 1987(8) - David A. Patterson, Garth Gibson, and Randy H. Katz: [A Case for Redundant Arrays of Inexpensive Disks \(RAID\)](#)
 - «кризис производительности»
 - среднее время безотказной работы (MTBF)
 - избыточный массив недорогих дисков

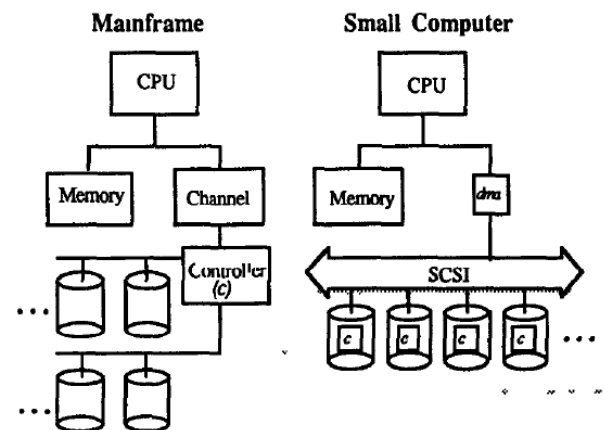
$$MTTF \text{ of a Disk Array} = \frac{MTTF \text{ of a Single Disk}}{\text{Number of Disks in the Array}}$$

$$MIPS = 2^{Year-1984}$$

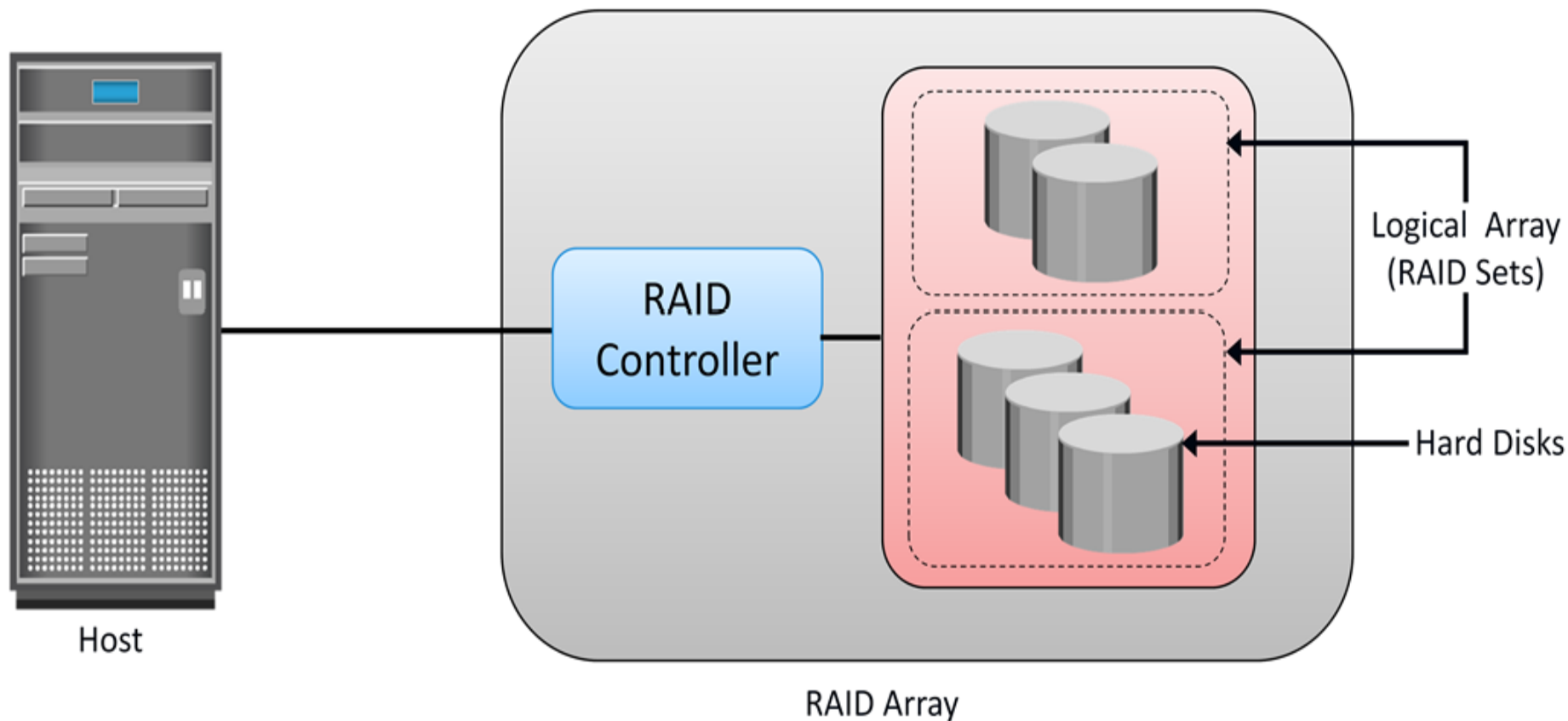
$$MAD = 10^{(Year-1971)/10}$$

Redundant Array of Independent Disks

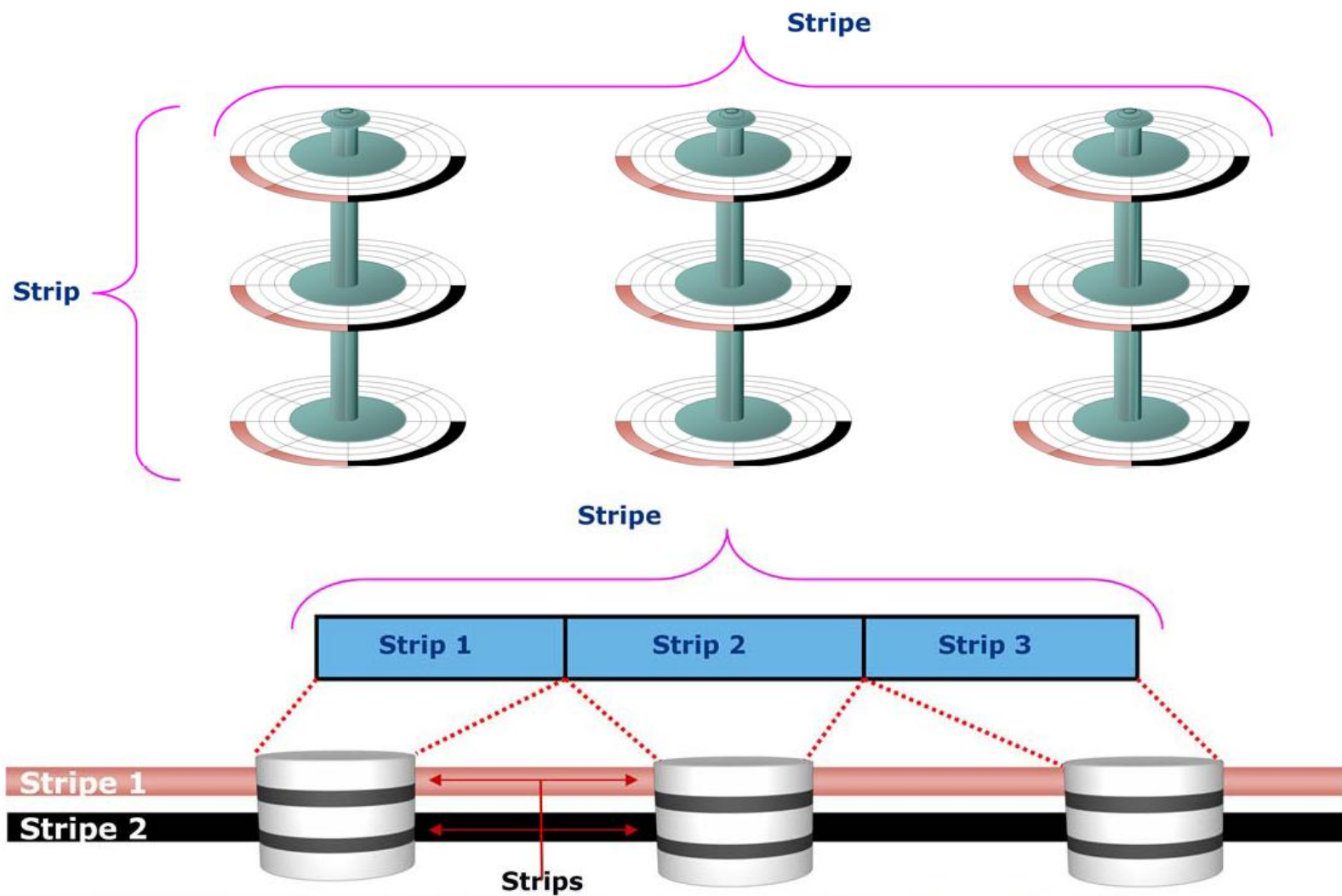
- Техника комбинирования физических дисков в логические диски
- Улучшение доступности и производительности за счет:
 - распределения
 - зеркалирования
 - контроля четности
- Реализации:
 - программная
 - аппаратная (контроллер или внешнее устройство)



Компоненты RAID-массива



Техники RAID - Распределение



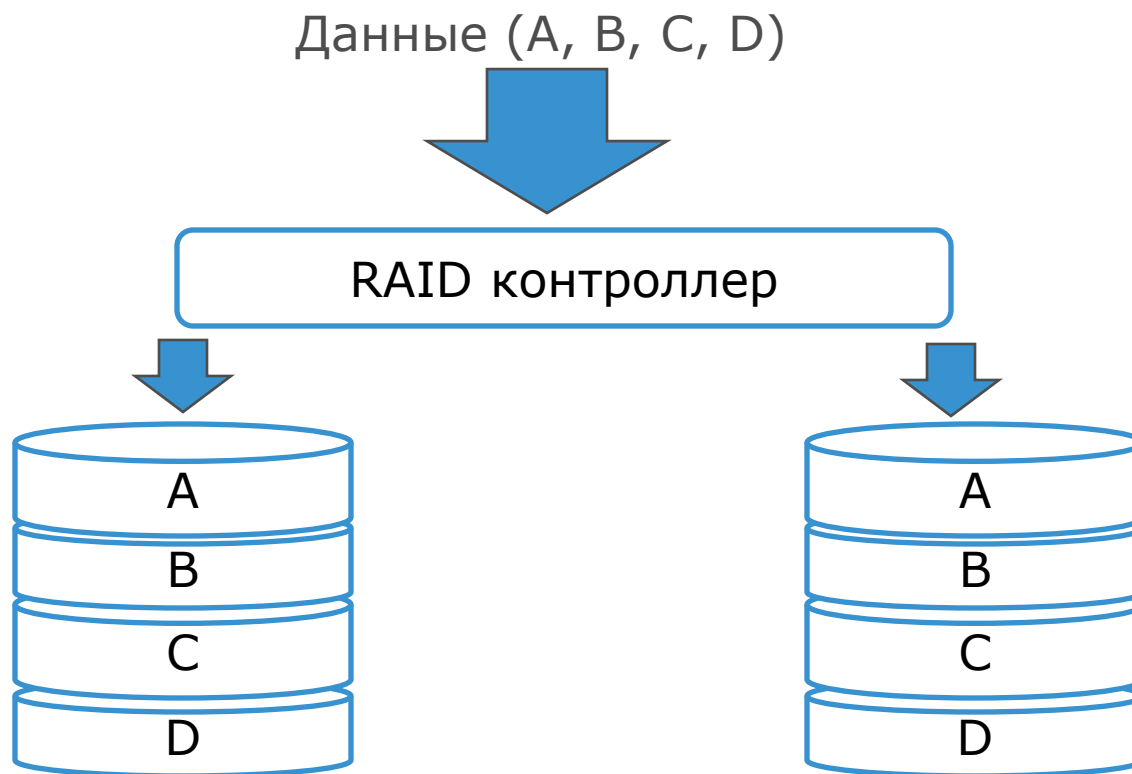
Распределение - Термины

- Strip (стрип, полоса) – блок данных со смежной адресацией
- Stripe (страйп, дорожка) – группа полос, охватывающая все диски
- Объем полосы или глубина дорожки – кол-во блоков в полосе на одном диске
- Объем дорожки – суммарный объем полос
- Ширина дорожки – кол-во полос в дорожке

Распределение – преимущества и недостатки

- Данные записываются по частям на разные диски – IOPS увеличивается пропорционально количеству дисков
- Выход из строя одного диска уничтожает данные на всех остальных

Техники RAID - Зеркалирование

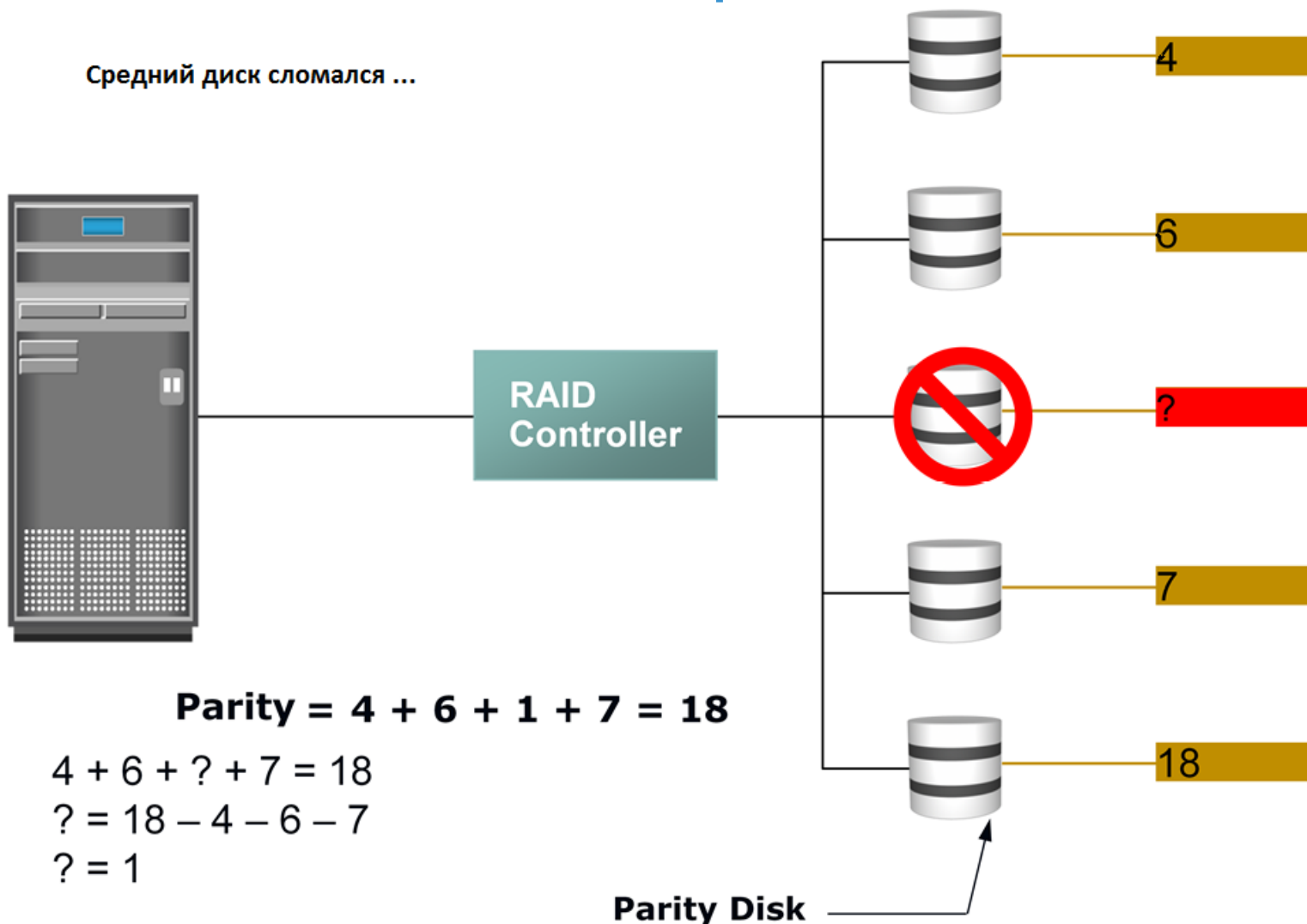


Зеркалирование – преимущества и недостатки

- При выходе из строя одного диска данные остаются целыми на втором
- Поврежденный диск меняется, новый диск получает копию с уцелевшего (без участия хоста)
- Улучшает чтение, замедляет запись
- Не заменяет резервных копий
- Больше дисков (2х, 3х)

Техники RAID – Контроль четности

Средний диск сломался ...



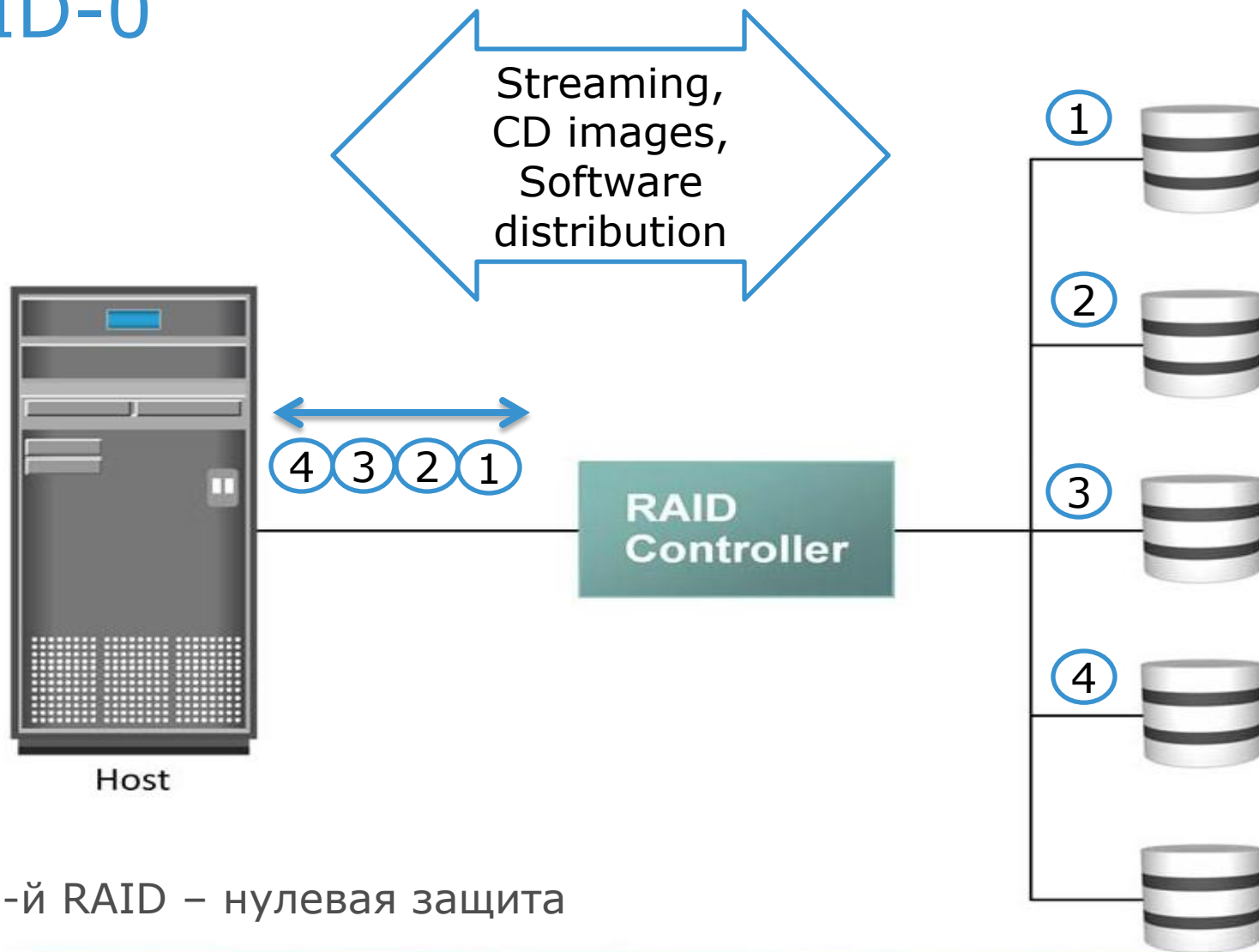
Контроль четности

- Обеспечивает полную защиту от сбоя без дублирования
- XOR – функция четности
- Стоимость защиты гораздо ниже – из 5 дисков – 4 для данных, 1 для контрольных сумм
- Медленная запись – перерасчет контрольных значений
- Хорошая скорость чтения

Уровни RAID

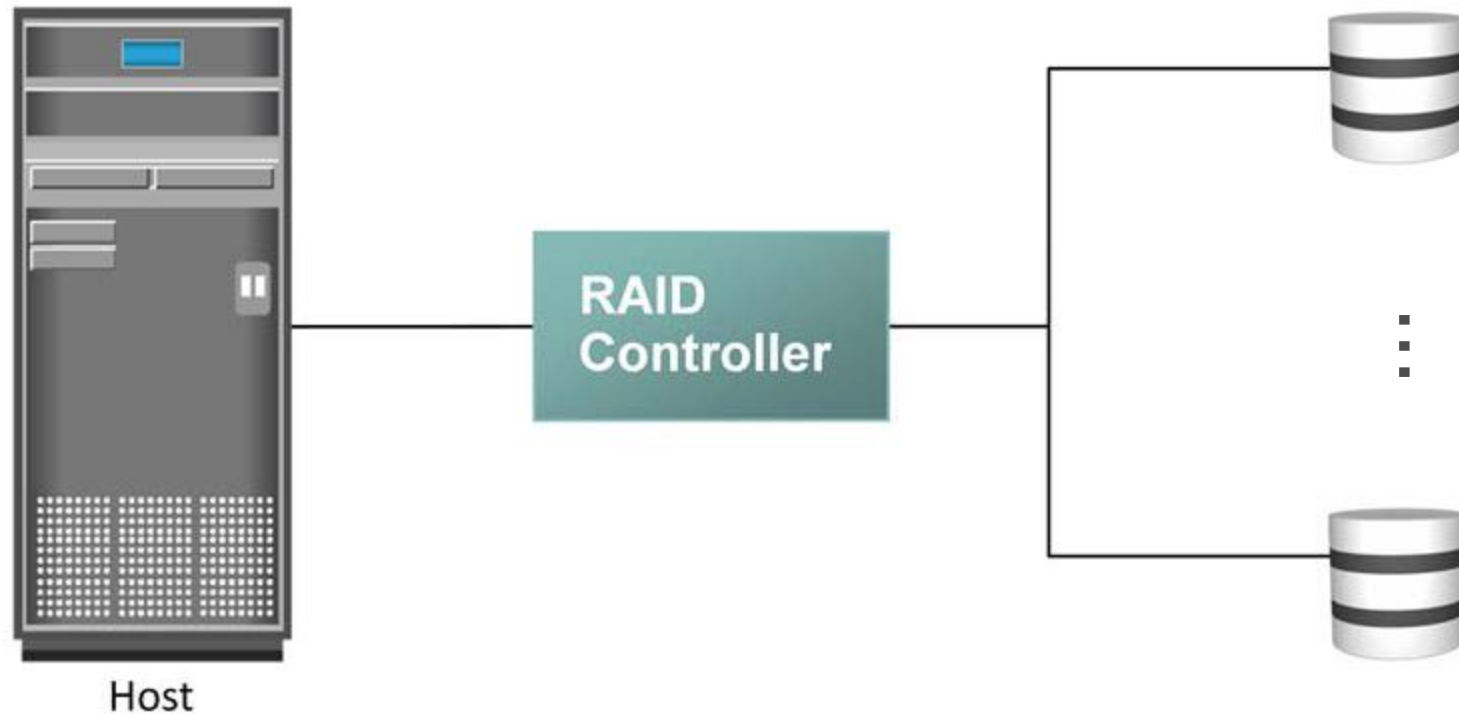
Уровни	Описание
RAID 0	Распределенный массив без отказоустойчивости
RAID 1	Зеркалирование диска
RAID 3	Параллельный (синхронный) доступ с диском контроля четности
RAID 4	Распределенный массив с независимыми дисками и с диском контроля четности
RAID 5	Распределенный массив с независимыми дисками и с распределенным контролем четности
RAID 6	Распределенный массив с независимыми дисками и с двойным распределенным контролем четности
Вложенные	Комбинация уровней, 1+0, 5+1, ..

RAID-0



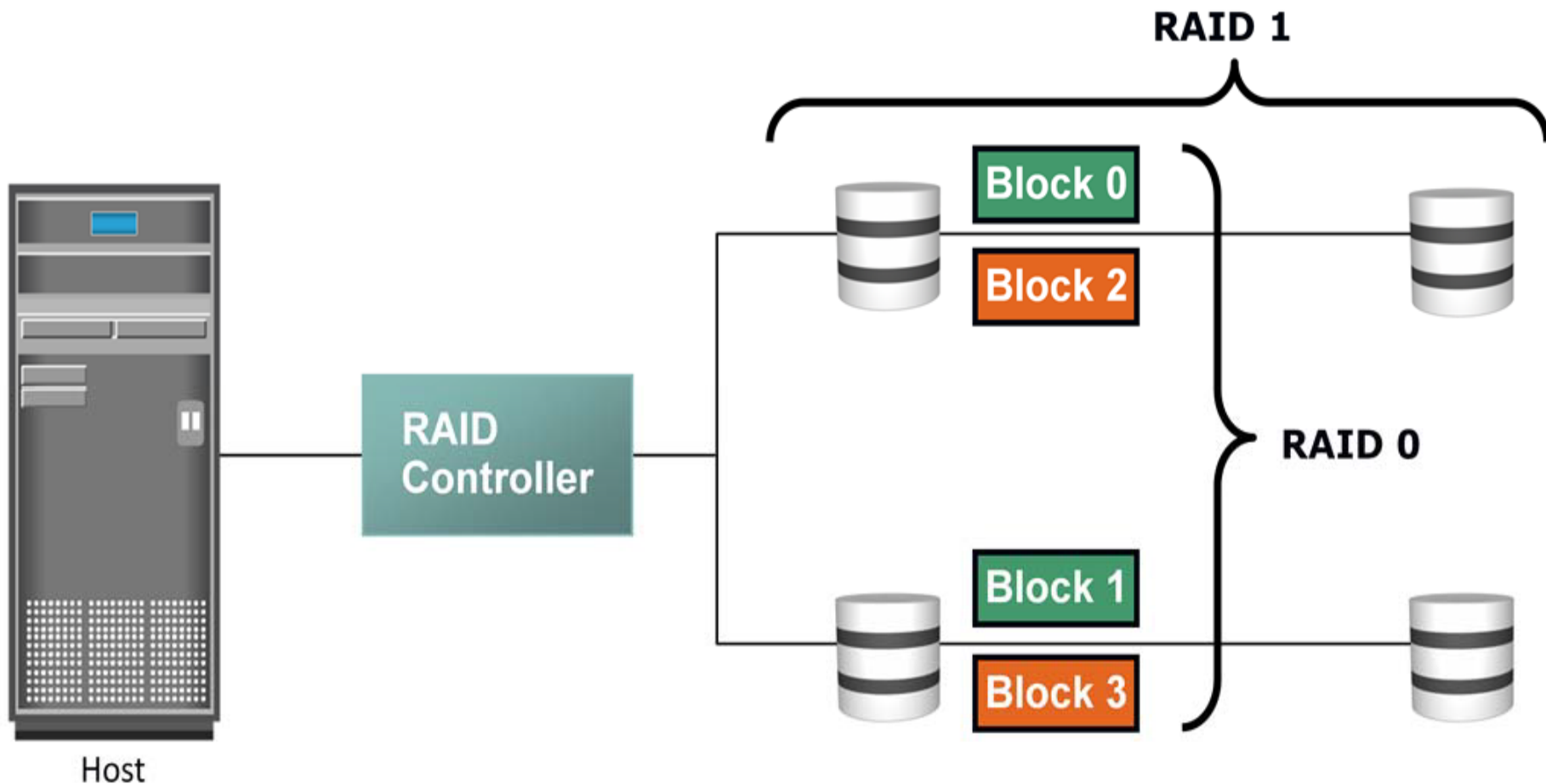
0-й RAID – нулевая защита

RAID-1



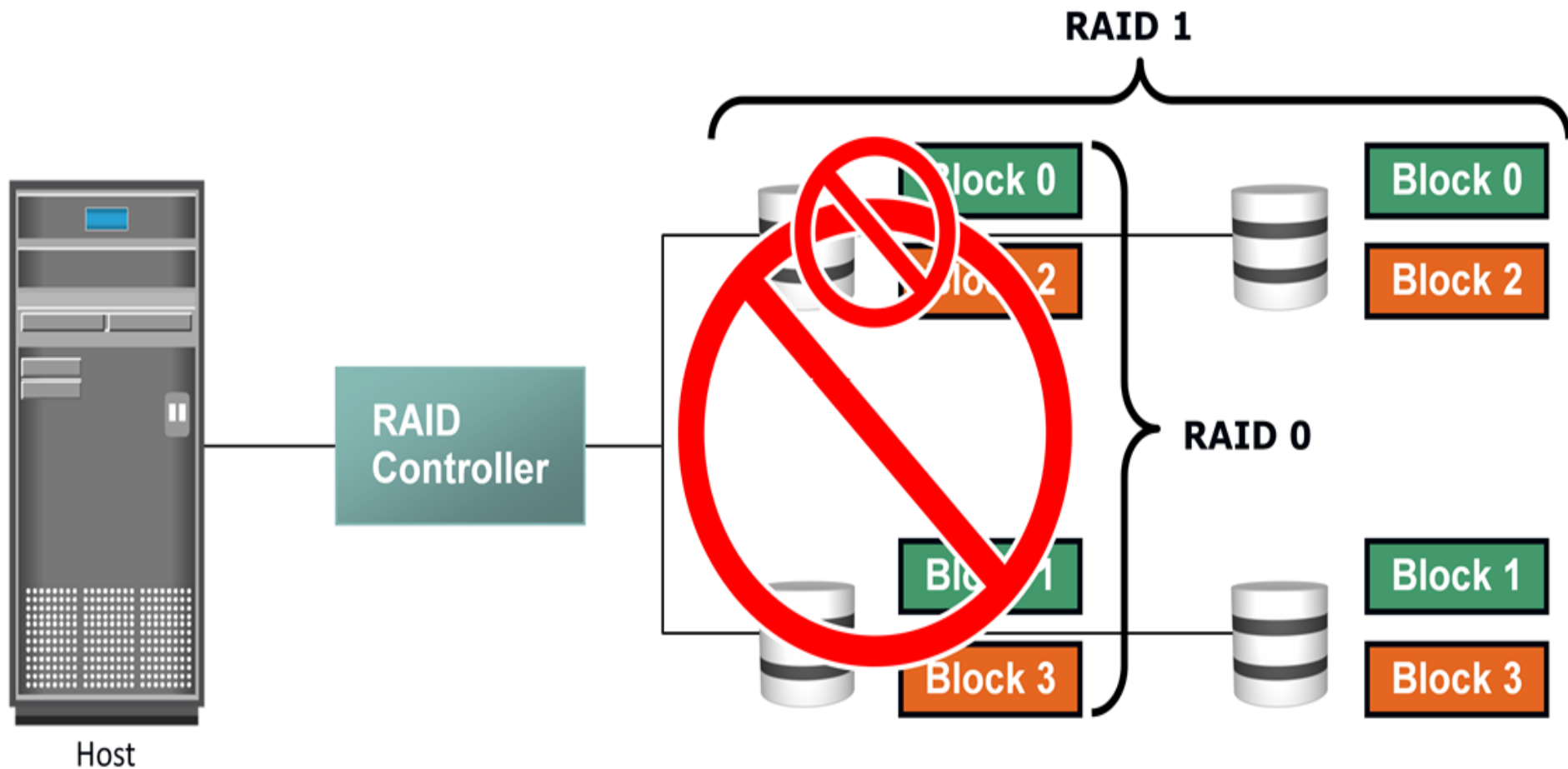
“Mission critical” apps, данные на дисках копируются 1:1

RAID 0+1

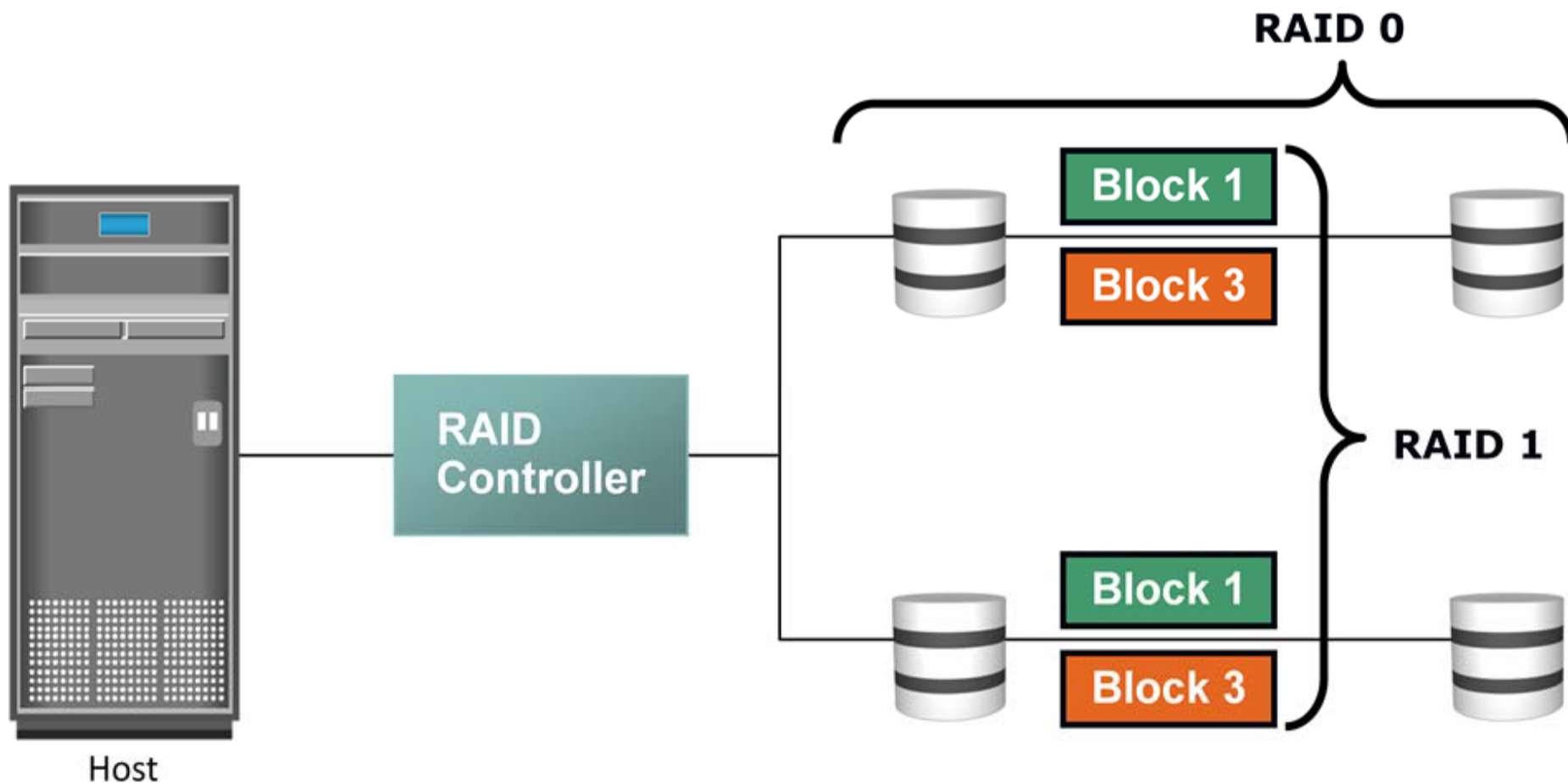


Зеркалированная дорожка

RAID 0+1, потеря диска

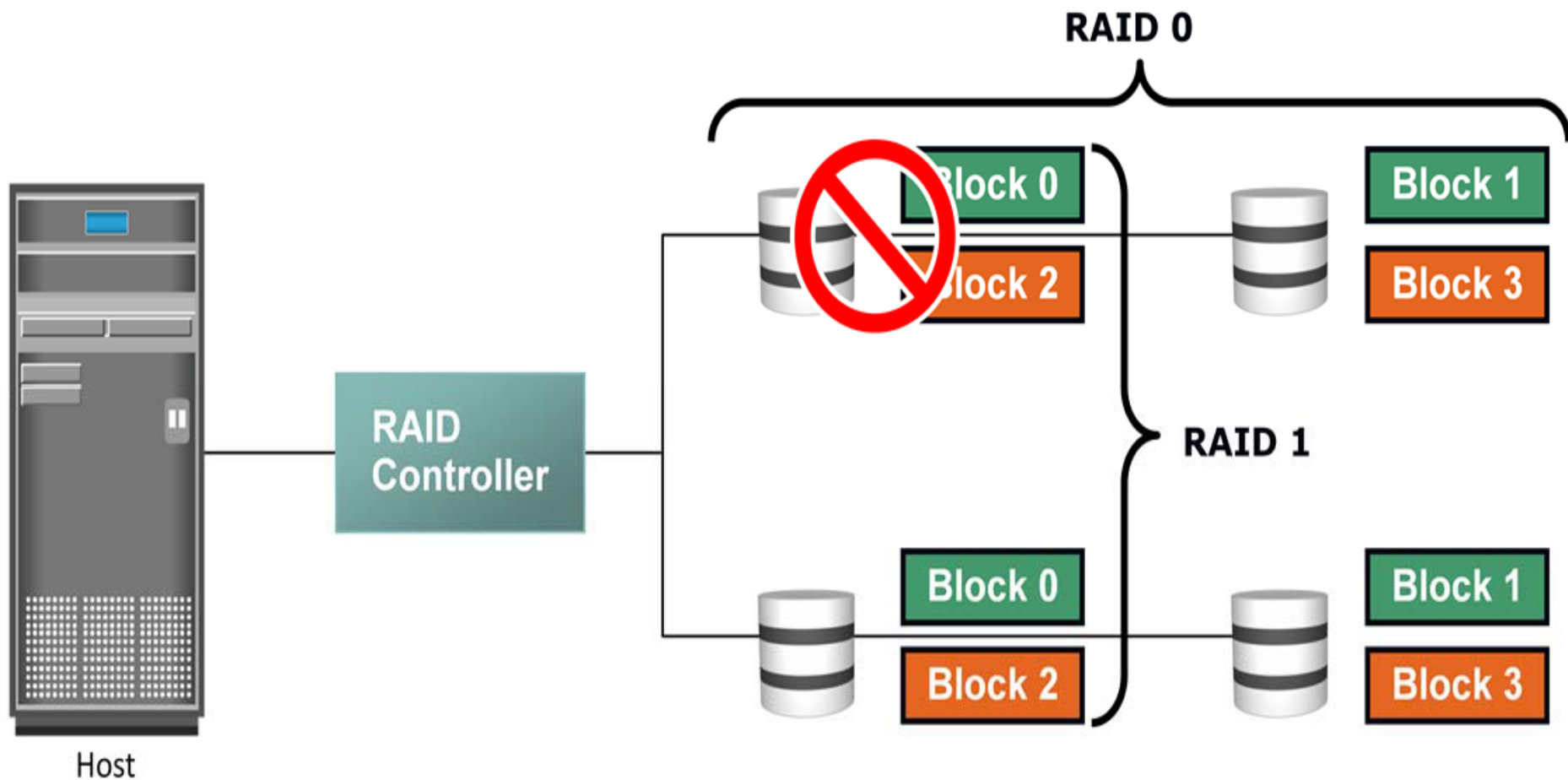


RAID 1+0



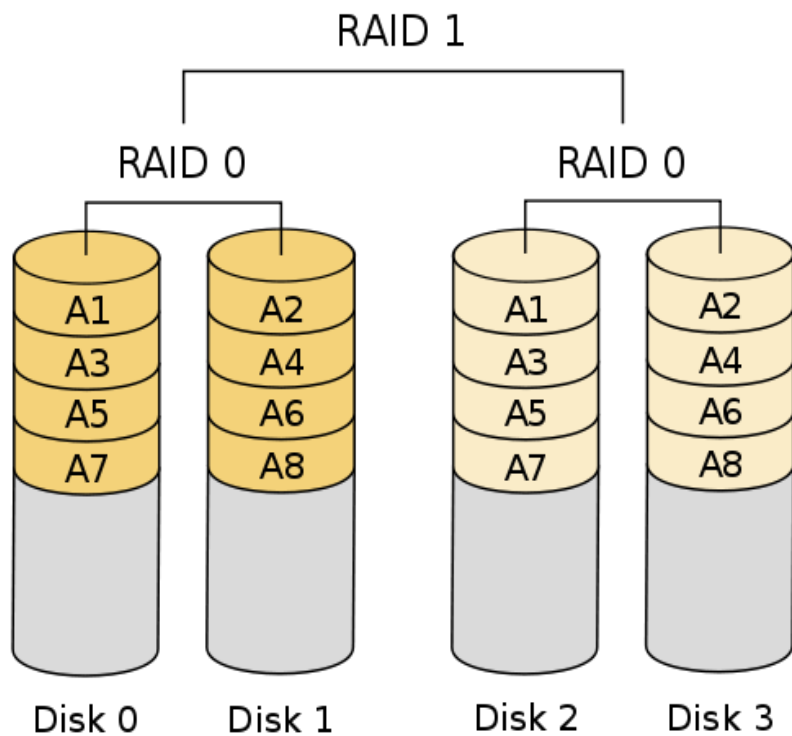
Сегментированное зеркало

RAID 1+0, потеря диска

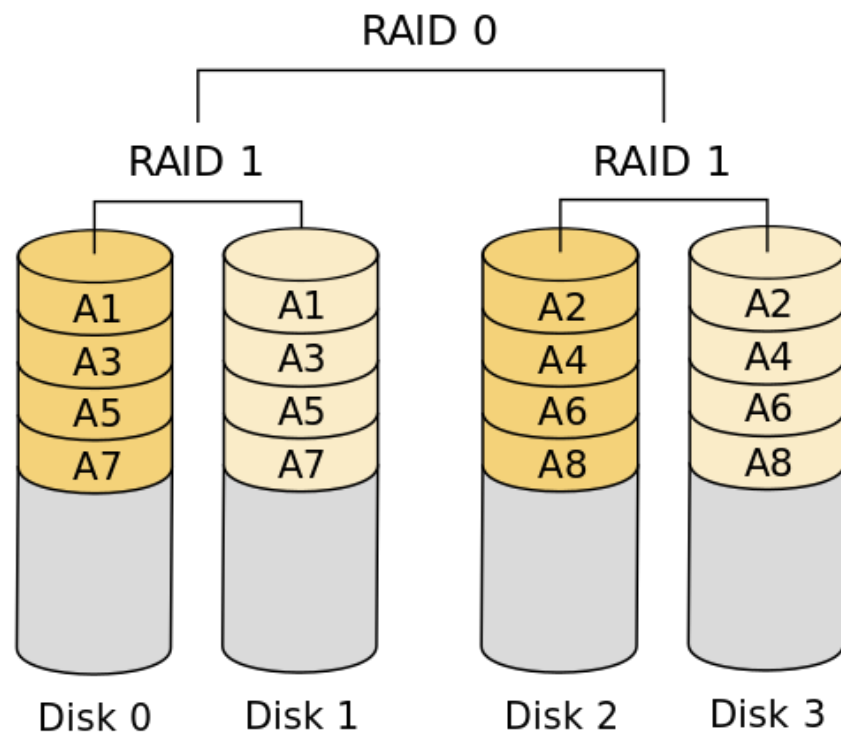


RAID 0+1, 1+0 - сравнение

RAID 0+1



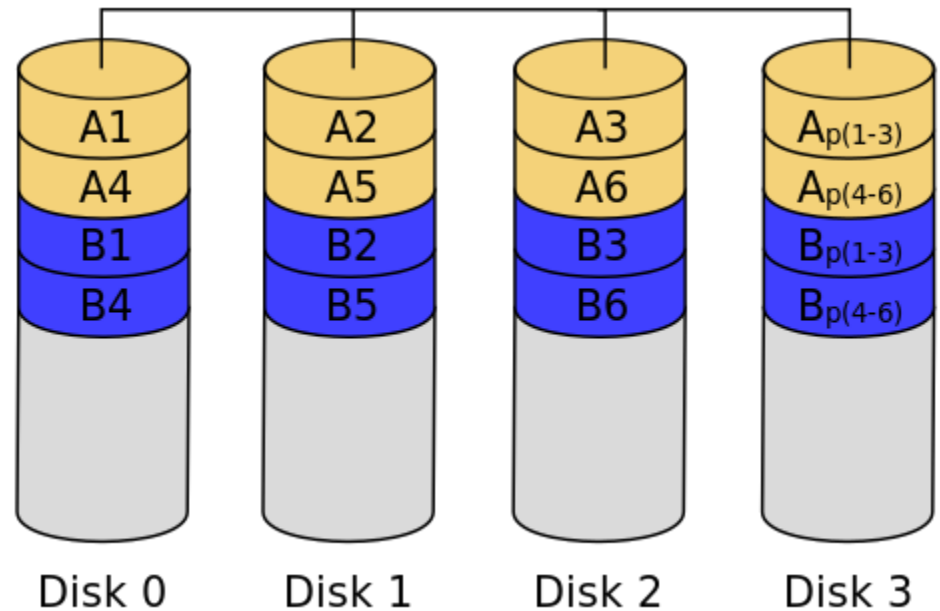
RAID 1+0



RAID 1+0 - применение

- OLTP
- Базы данных
- Центры приема и отправки сообщений
- Приложения с произвольным чтением и записью данных, требовательные к производительности и надежности

RAID-3

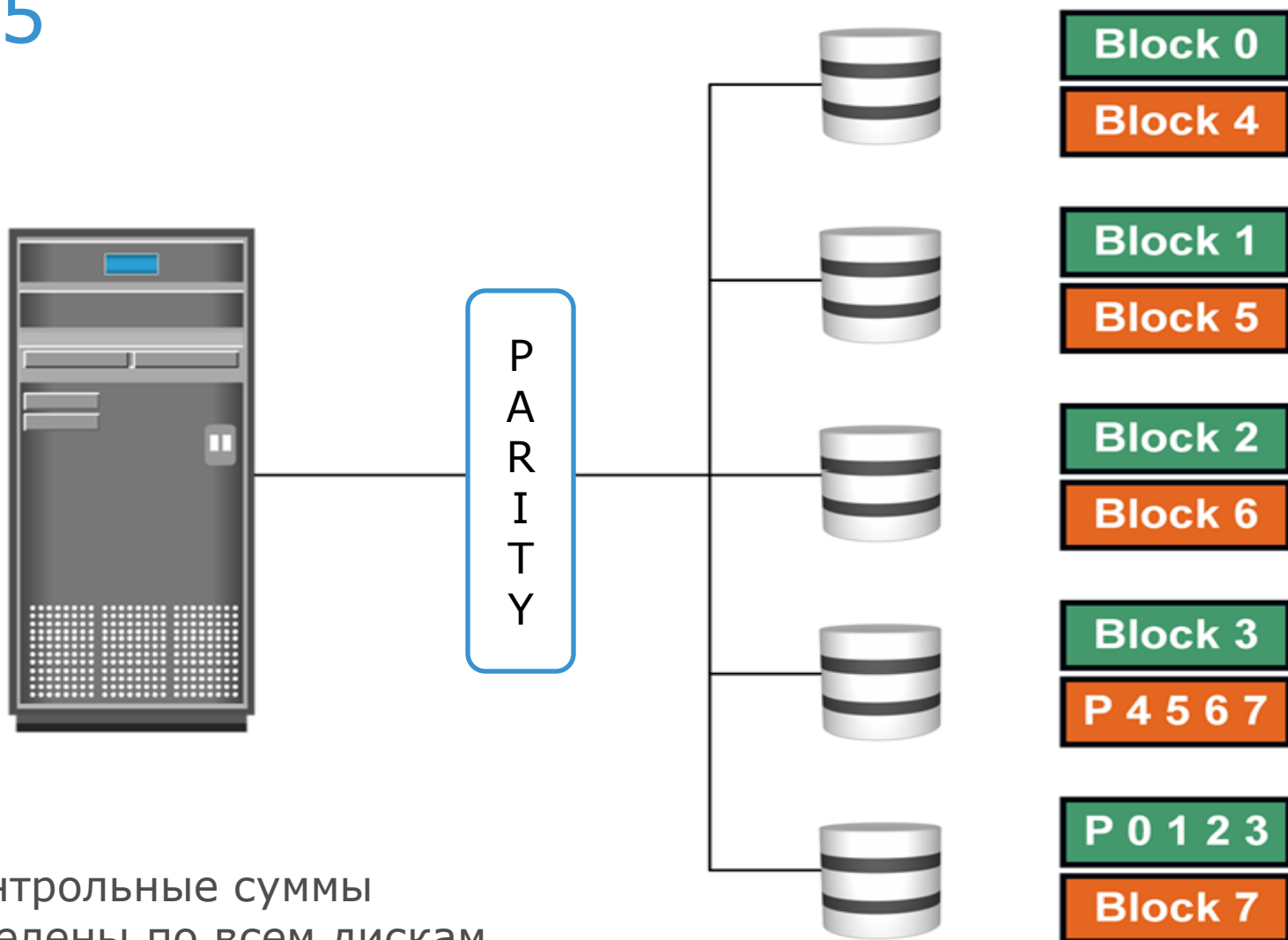


- $D3 = \text{XOR}(D0+D1+D2)$
- Запись и чтение - целыми дорожками
- Полезный объем = $N-1$

RAID-4

- RAID-3 + произвольные чтение и запись
- Узкое место – диск с контрольными суммами
- Используется в основном в СХД NetApp (в связке с кэшированием и WAFL)

RAID-5

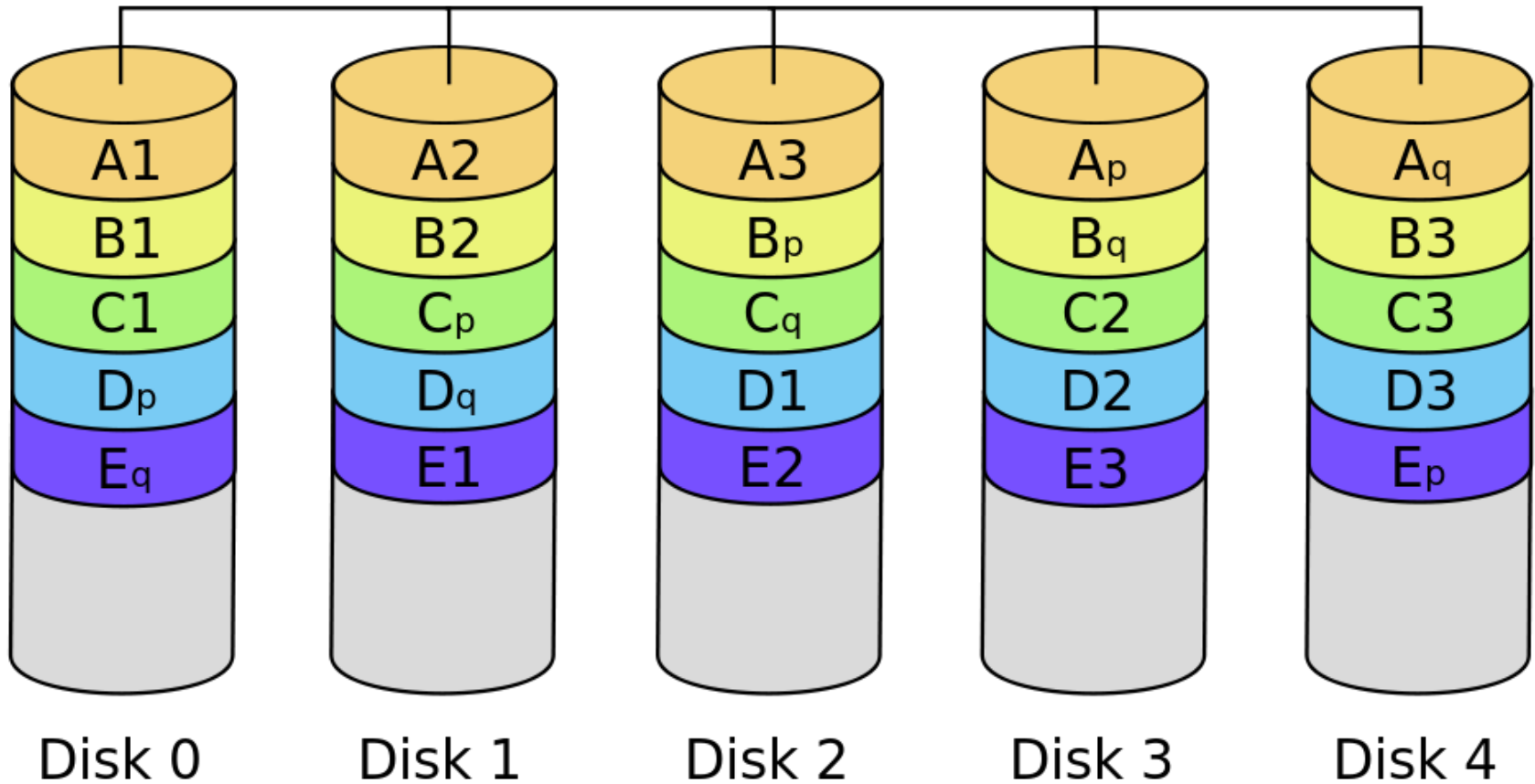


Контрольные суммы
распределены по всем дискам

RAID-6

- Две контрольные суммы на дорожку
- Защита данных от поломки двух дисков
- Контрольные суммы
 - горизонтальная + диагональная
 - $Q+P$
 - ...

RAID-6, P+Q parity



Эффективность хранения

RAID	Мин. дисков	Эффективность, %
0	2	100
1	2	50
3, 4	3	$100 \times (N-1) / N$
5	3	$100 \times (N-1) / N$
6	4	$100 \times (N-2) / N$
1+0, 0+1	4	50

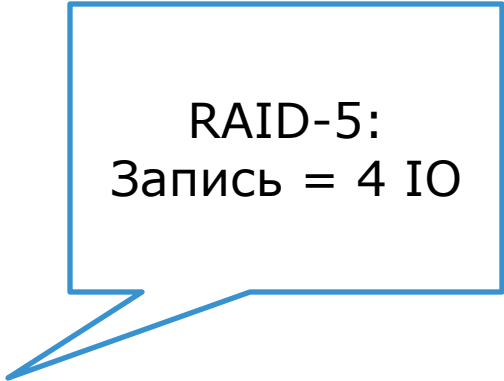
Производительность RAID

- Ускорение чтения \sim количеству дисков
- Запись медленнее из-за
 - дублирования
 - подсчета контрольной суммы

1. $D_p = D_1 + D_2 + D_3 + D_4$

2. $D_4 = \text{new}$

3. $D_{p \text{ new}} = D_{p \text{ old}} - D_{4 \text{ old}} + D_{4 \text{ new}}$



RAID-5:
Запись = 4 IO

Зависимость IOPS от RAID

Задача: сервер производит 5200 IOPS, из них 60% - чтение. Рассчитать загрузку на диски в случае RAID-1 и RAID-5

1. RAID-1: $0.6 \times 5200 + 2 \times 0.4 \times 5200 = 7280$

2. RAID-5: $0.6 \times 5200 + 4 \times 0.4 \times 5200 = 11440$

В RAID-5 необходимо больше дисков, чтобы обеспечить требуемый уровень IO. Его следует использовать в приложениях с малым отношением операций записи к чтению (1:2).

Спасибо!

EMC²®