

# Распределенные системы хранения и обработки данных

Владислав Белогрудов, EMC

[vlad.belogrudov@gmail.com](mailto:vlad.belogrudov@gmail.com)

# Лекция 4

## Архитектура систем хранения данных

# Содержание лекции

- Основные компоненты
- Управление кешированием
- Защита от сбоев
- Классы СХД

# Интеллектуальные системы хранения данных

- Простые СХД (диски, RAID массивы) не обеспечивают возрастающие требования
  - доступности
  - быстродействия
  - надежности
  - масштабируемости
- Интеллектуальные СХД оптимизируют операции ввода-вывода, обеспечивая вышеперечисленные требования

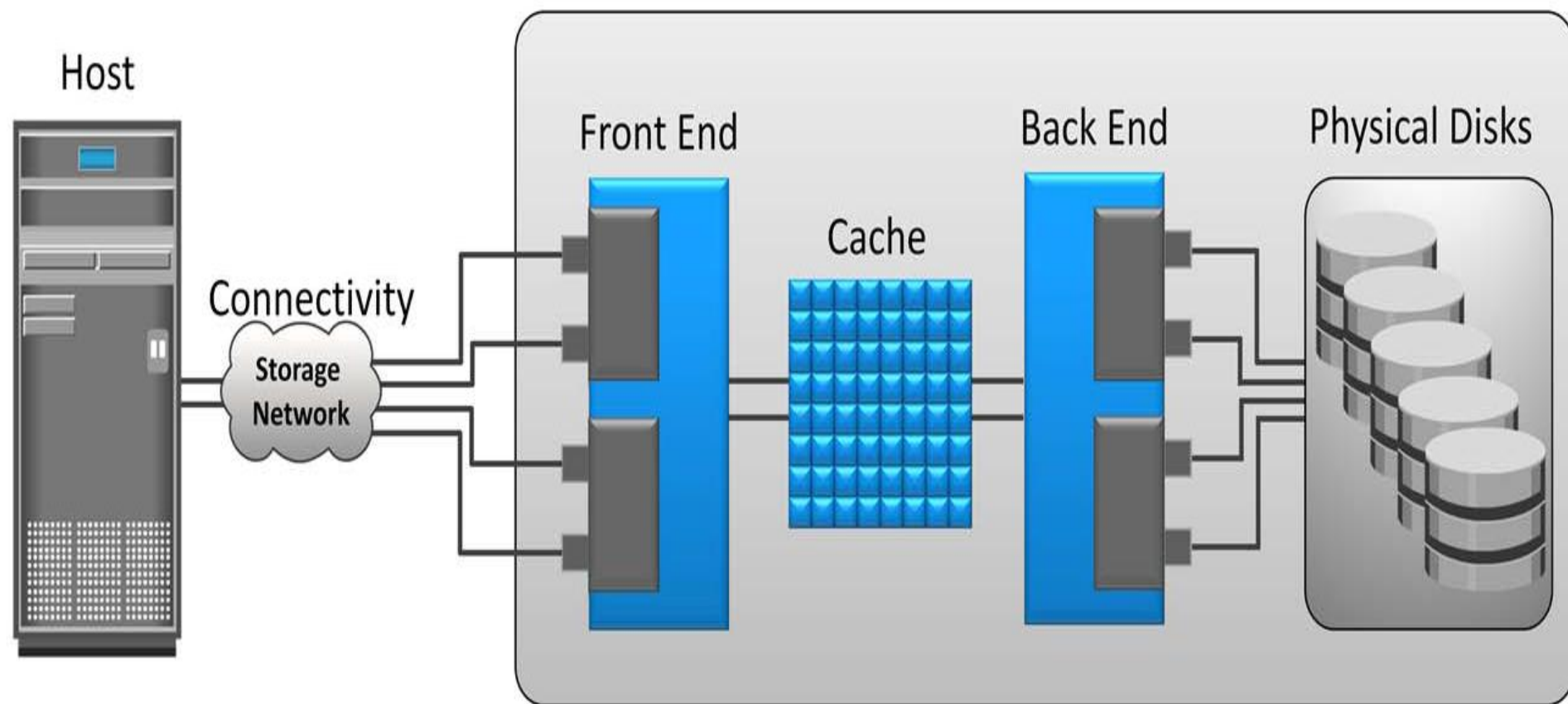


# Интеллектуальные СХД - свойства

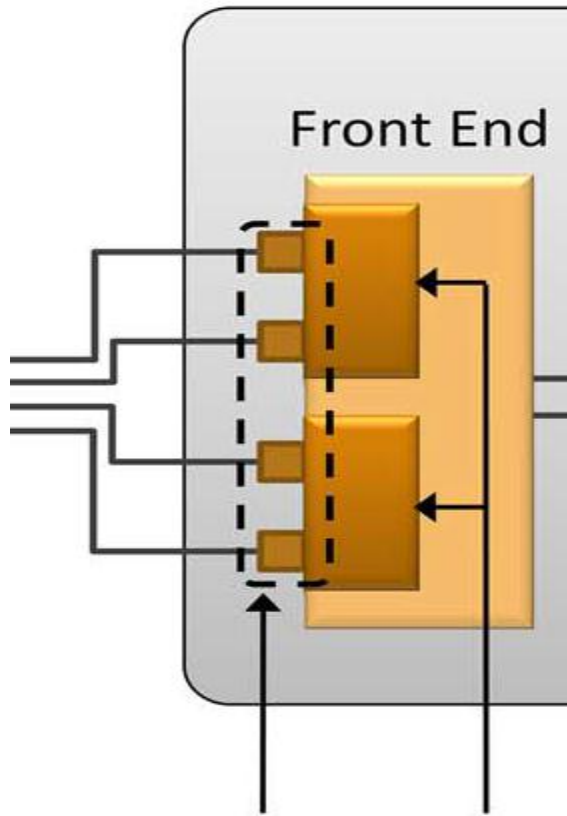
- Оптимизация IO за счет большой кеш-памяти, множественных каналов передачи данных
- Наличие операционной среды
  - управление кеш-памятью
  - множественный доступ хостов
  - управление ресурсами (дисками, каналами,..)
- Поддержка виртуализации, иерархического хранения, ILM, ..

# Интеллектуальные СХД - компоненты

Intelligent Storage System



# Front-end

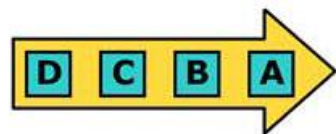


порты    контроллеры

- Интерфейс к хосту (FC, iSCSI, FCoE, TCP/IP)
- Контроль доступа
- Контроль скорости
- Оптимизация очереди команд

# Оптимизация очереди IO

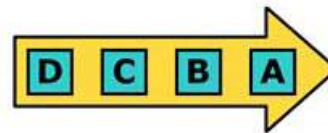
без оптимизации  
(FIFO)



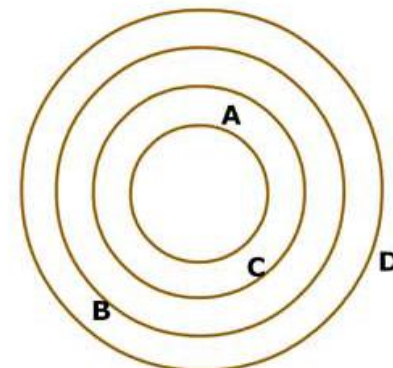
I/O Requests



Front-End  
Controller

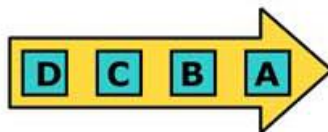


I/O Processing  
Order



Cylinders

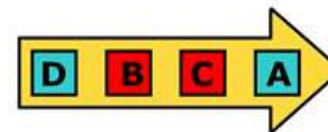
с оптимизацией  
времени поиска



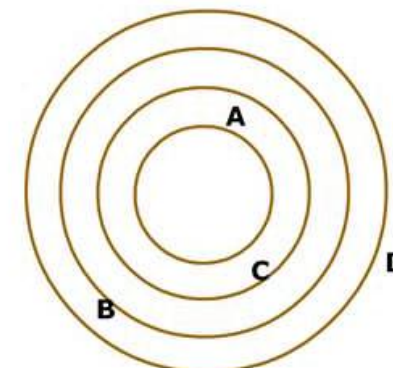
I/O Requests



Front-End  
Controller



I/O Processing  
Order



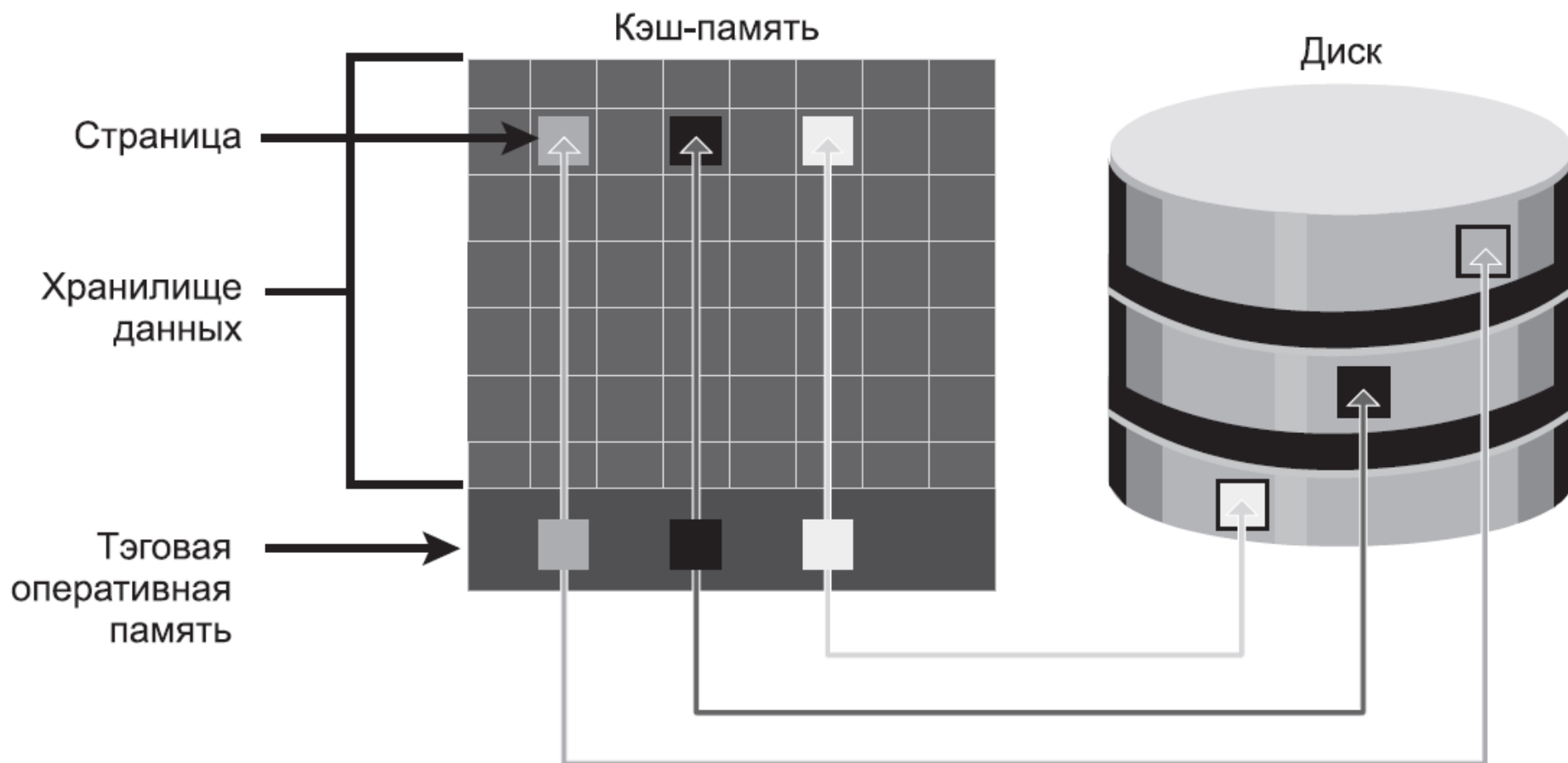
Cylinders



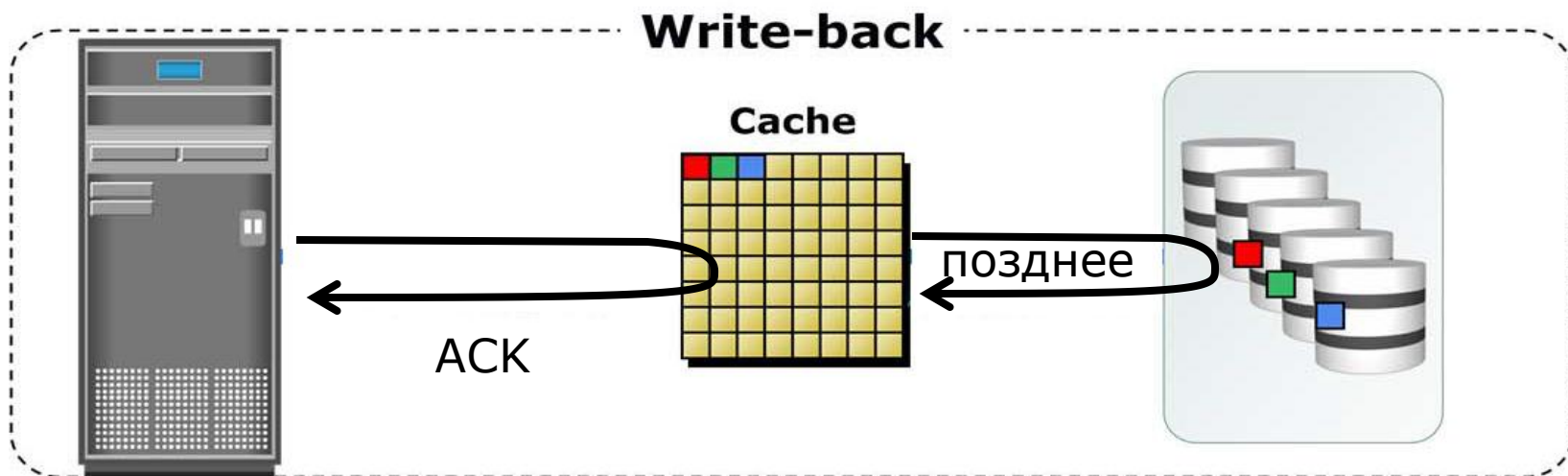
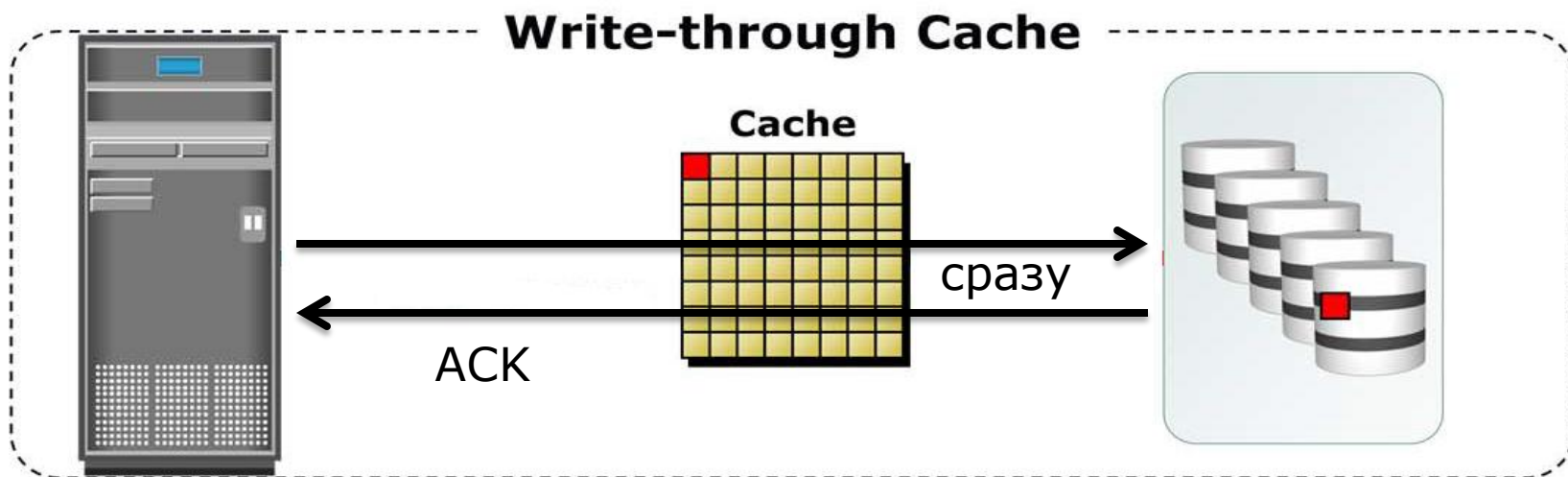
# Способы оптимизации

- Время поиска – оптимизация радиального движения головок чтения
- Время доступа – оптимизация времени поиска и ожидания нужного сектора

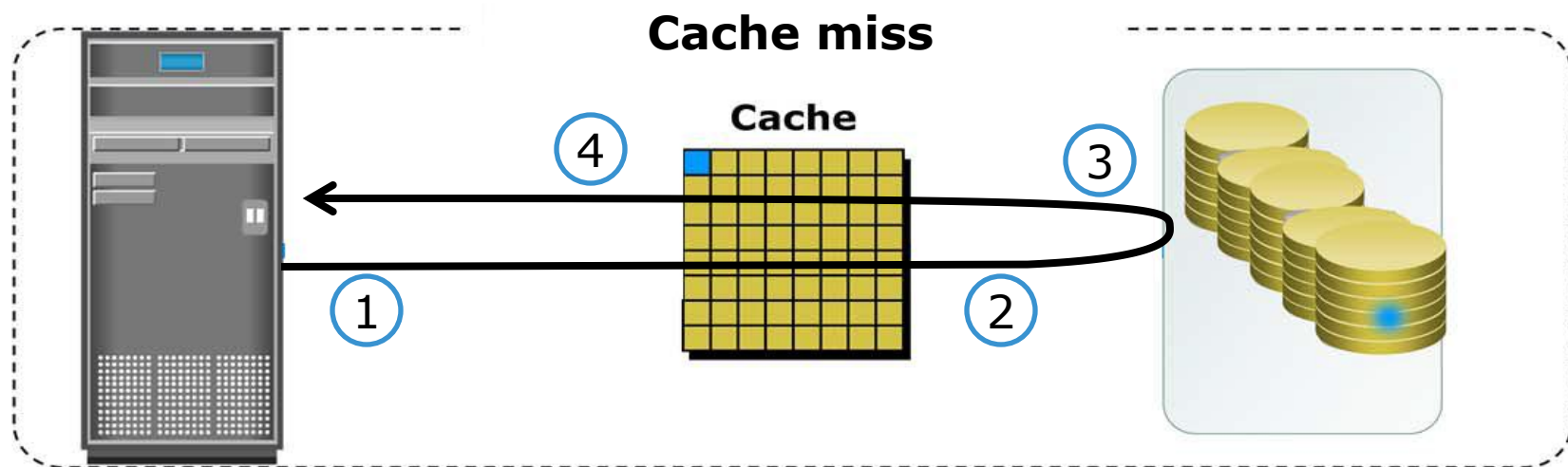
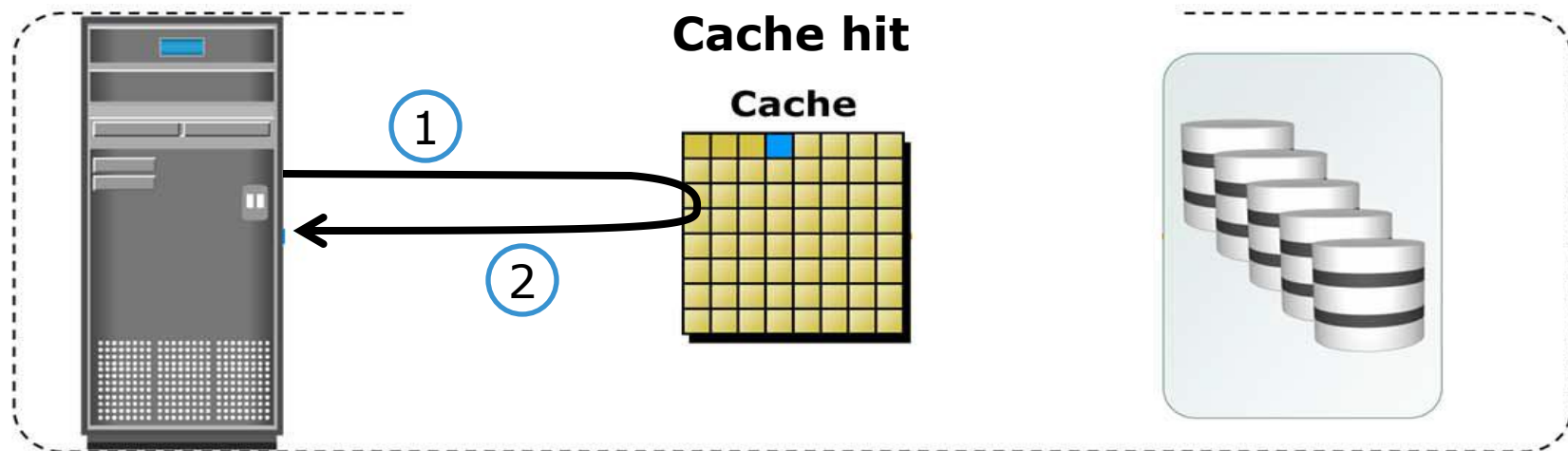
# Кэш-память



# Запись данных



# Чтение данных



# Коэффициент «попаданий» чтения

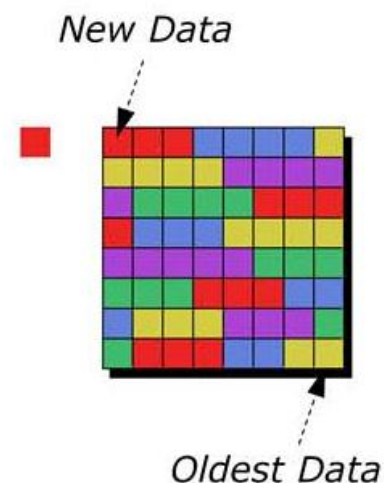
$$\text{Read Hit Ratio} = \frac{\text{Число попаданий при чтении}}{\text{Общее число запросов на чтение}}$$

Чтение «про запас» - с дисков читаются последующие блоки в надежде, что они потребуются в ближайших запросах

- фиксированный размер
- динамический размер, в зависимости от скорости запросов и размера запросов

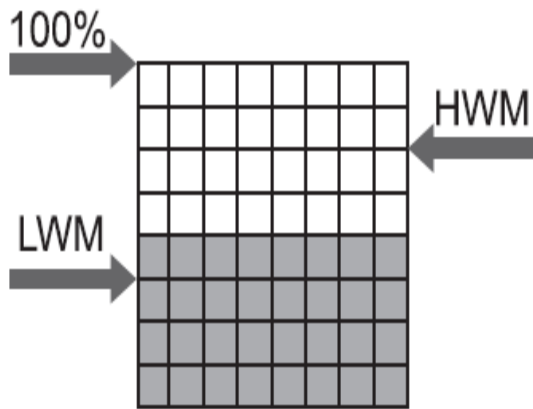
# Управление кэш-памятью

- LRU – least recently used – замещение давно неиспользованных страниц
- MRU – most recently used - замещение последних использованных страниц

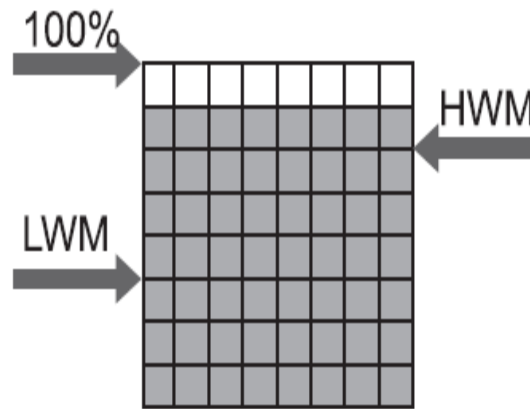


# Освобождение кэш-памяти

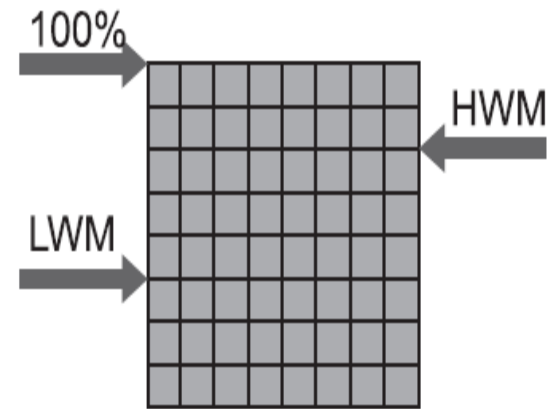
- Кэш не «резиновый», нужен сброс страниц
- Watermarks – low (LWM), high(HWM)



Сброс в фоновом режиме



Сброс на высоком уровне заполнения



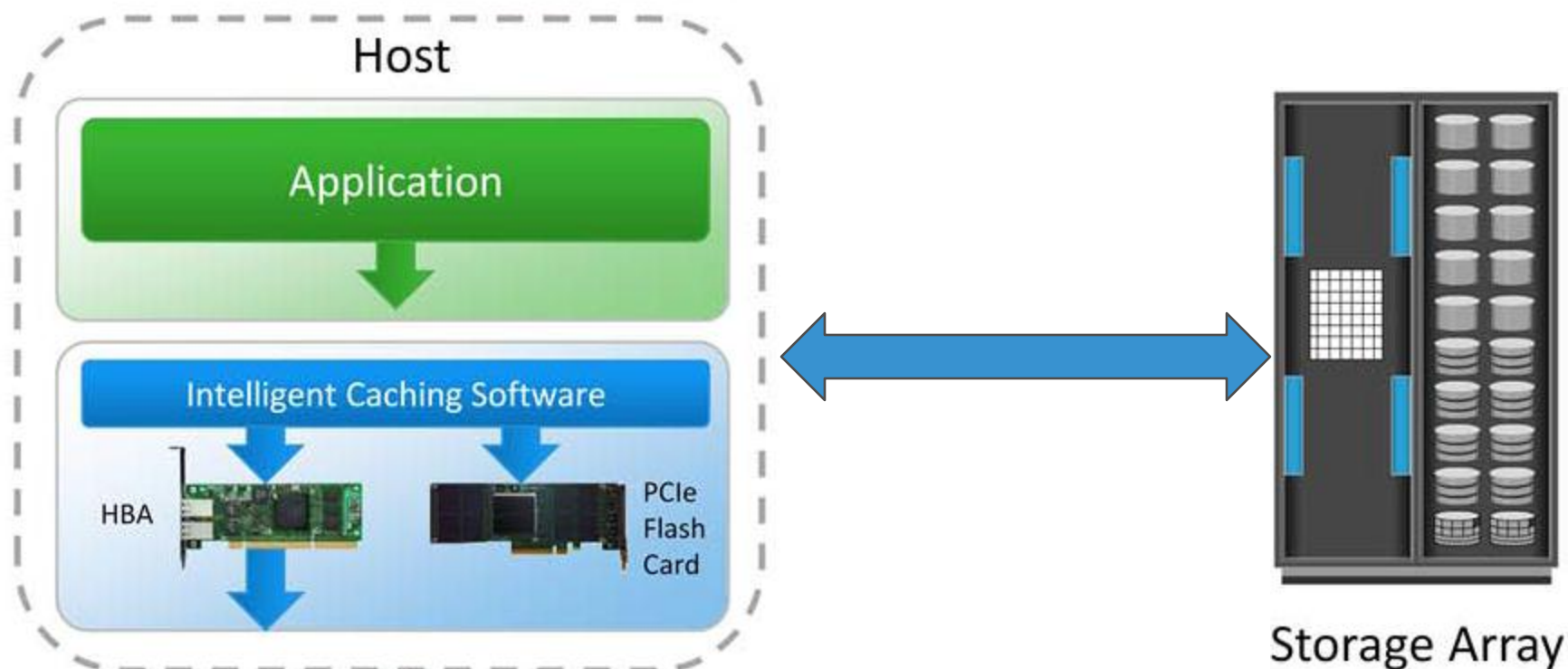
Принудительный сброс

# Защита кэш-памяти

- Зеркалирование (Cache Mirroring)
  - две независимых карты памяти
  - необходима синхронизация
  - дублируются только данные на запись
- Резервирование (Cache Vaulting)
  - защита от перебоев в питании
  - кэш-память копируется на специальный диск
  - При восстановлении питания vault копируется в кэш-память

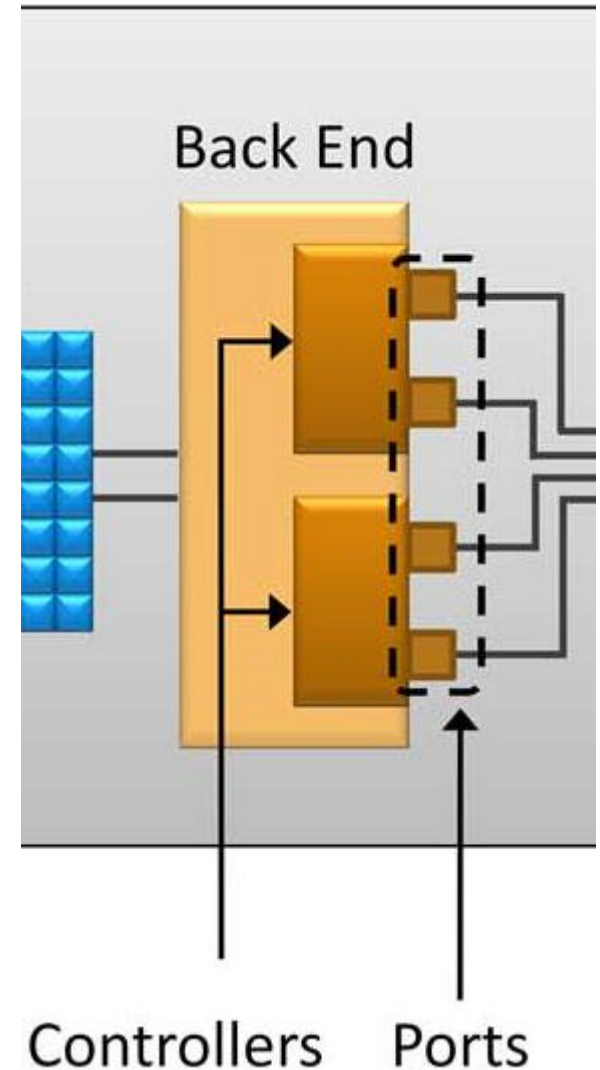


# Кэширование на стороне хоста

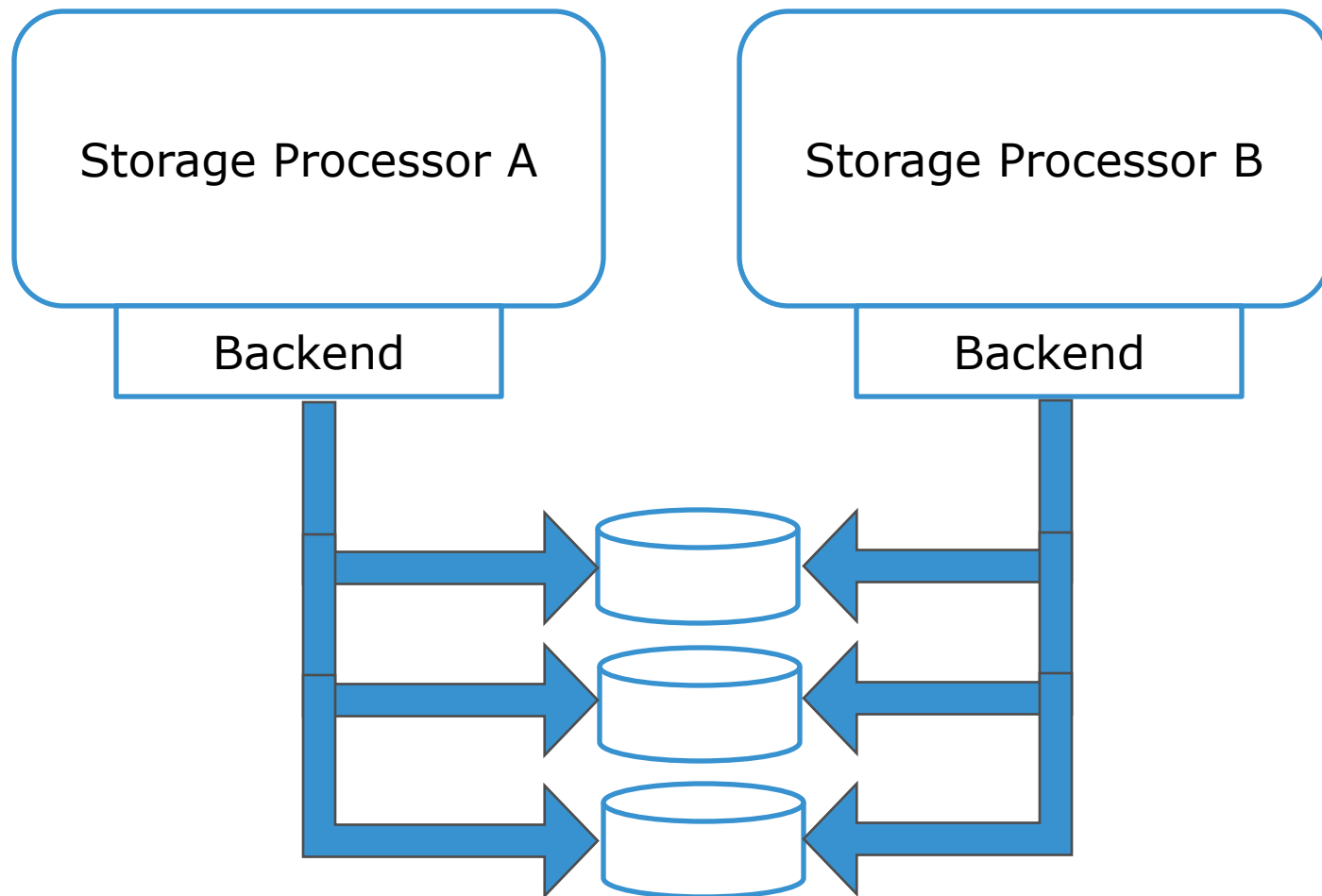


# Back-end

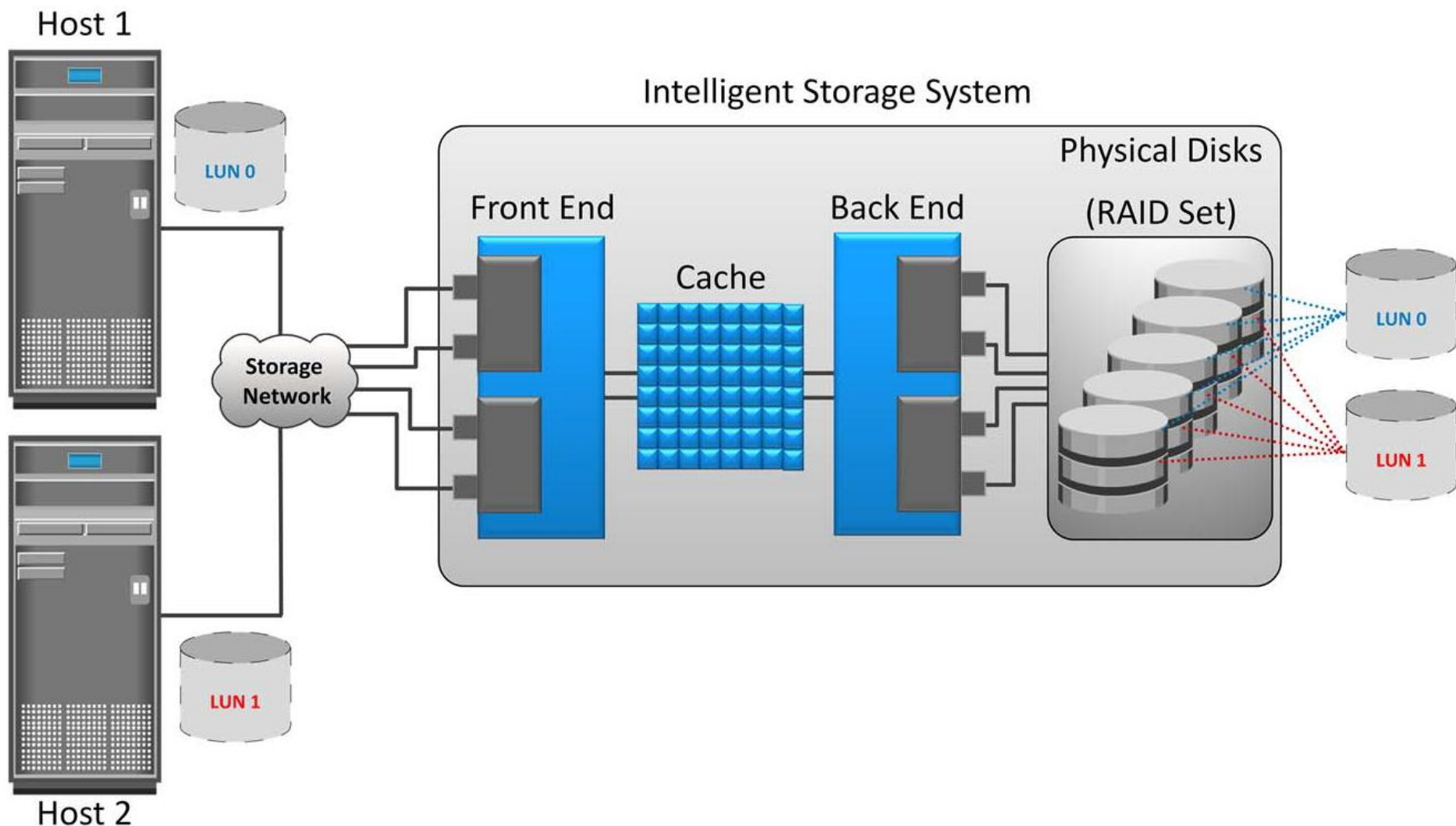
- Интерфейс к дискам
- Различные уровни RAID
- Балансировка нагрузки
- Несколько контроллеров и множество портов для повышения доступности



# Дисковый массив



# Логические диски

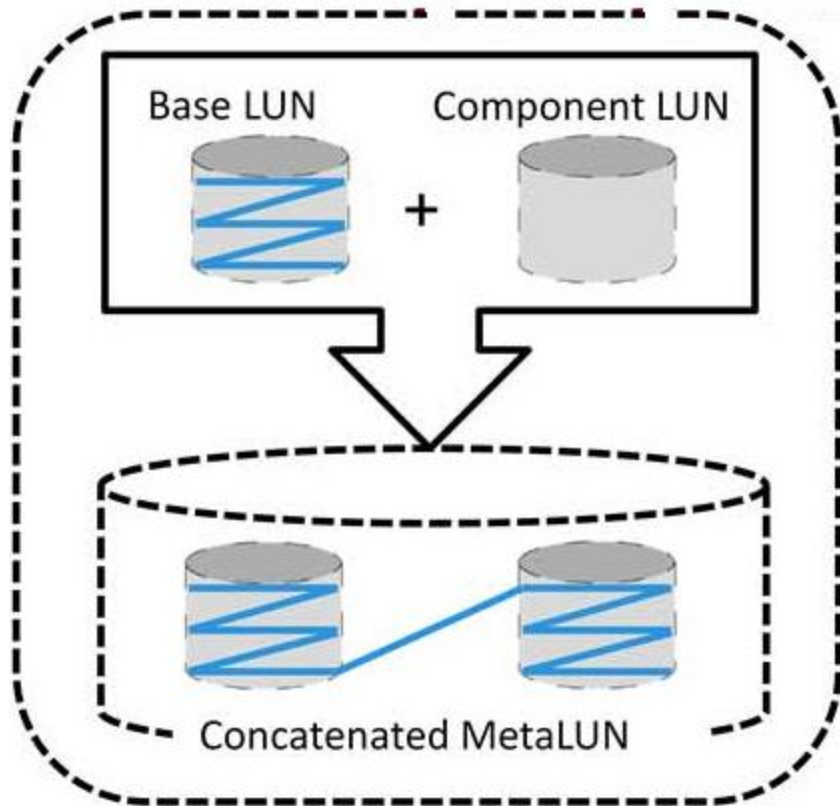


# LUN Masking

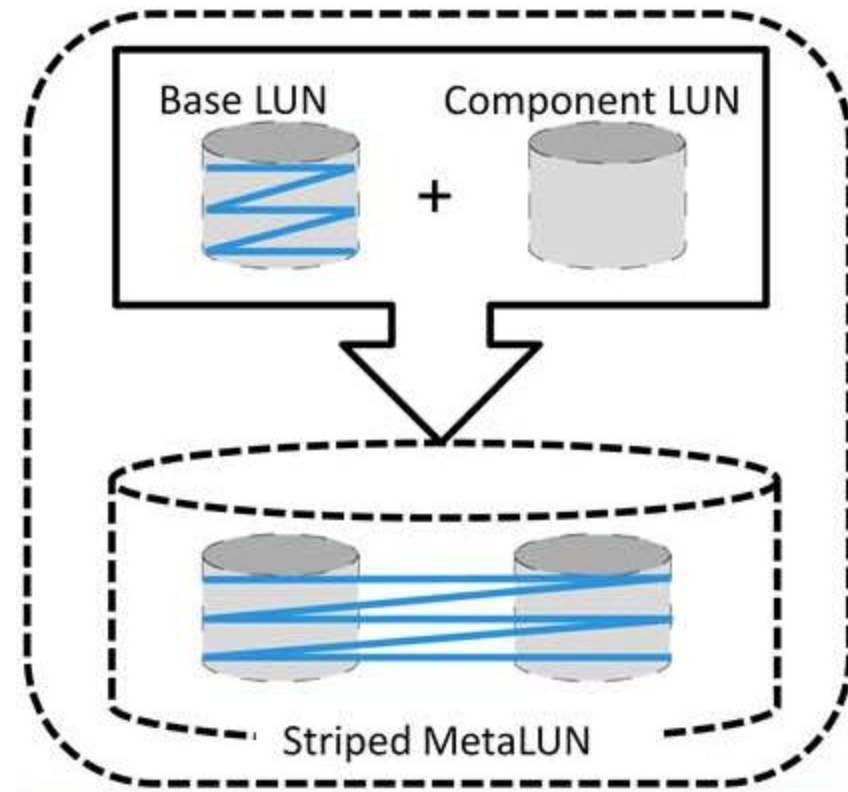
- Наложение маски на LUN
  - управление доступом
  - LUN назначается одному хосту (за исключением серверной виртуализации и резервного копирования)



# MetaLUN

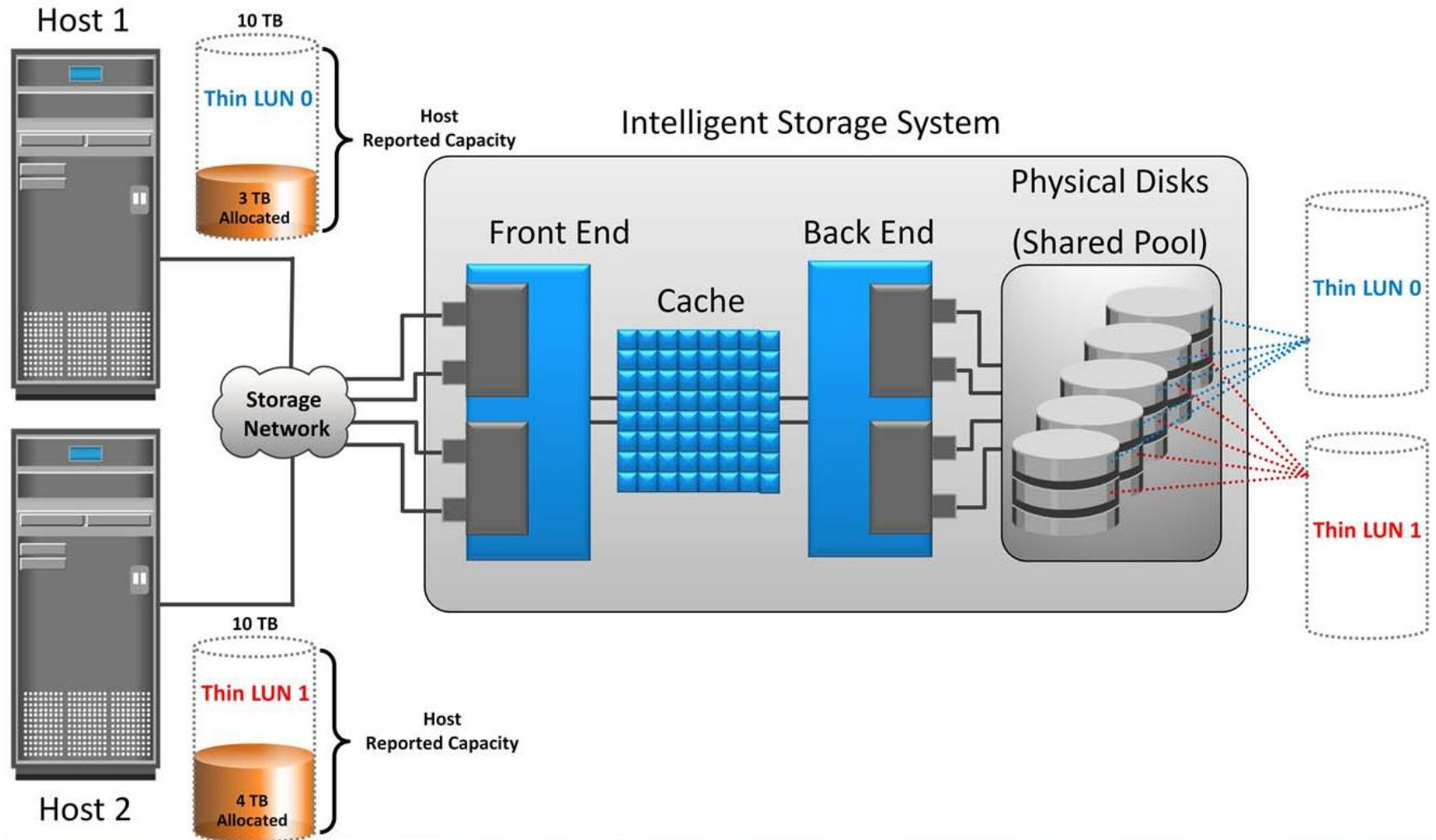


конкатенация



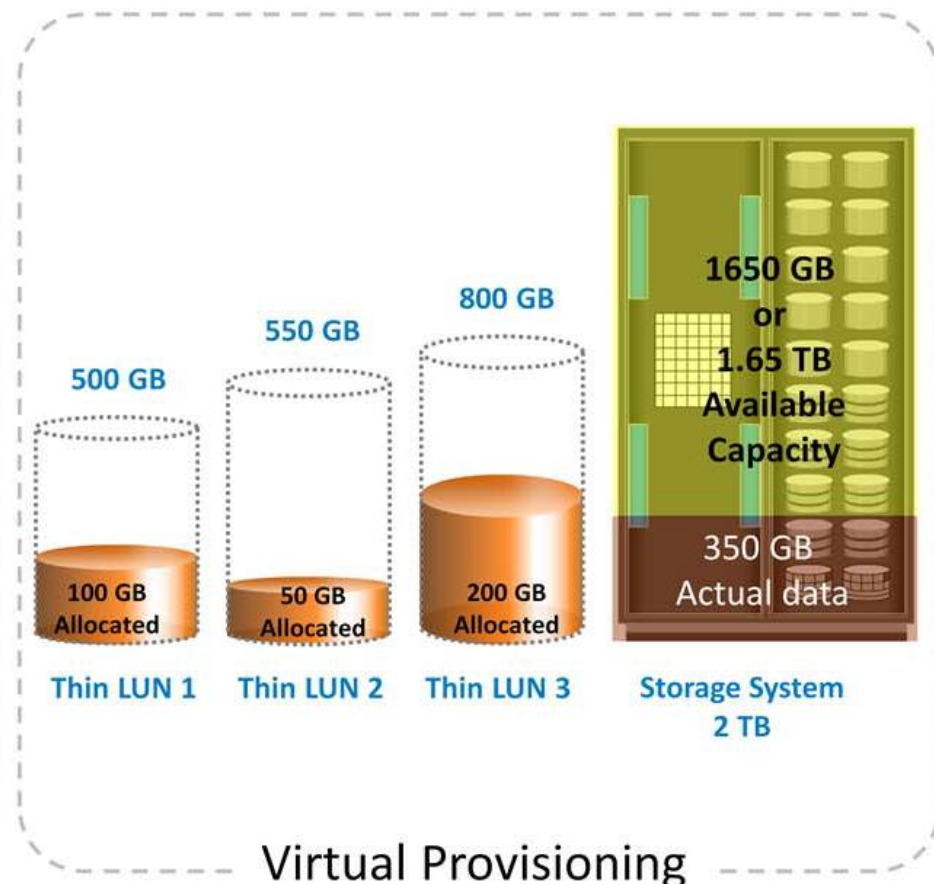
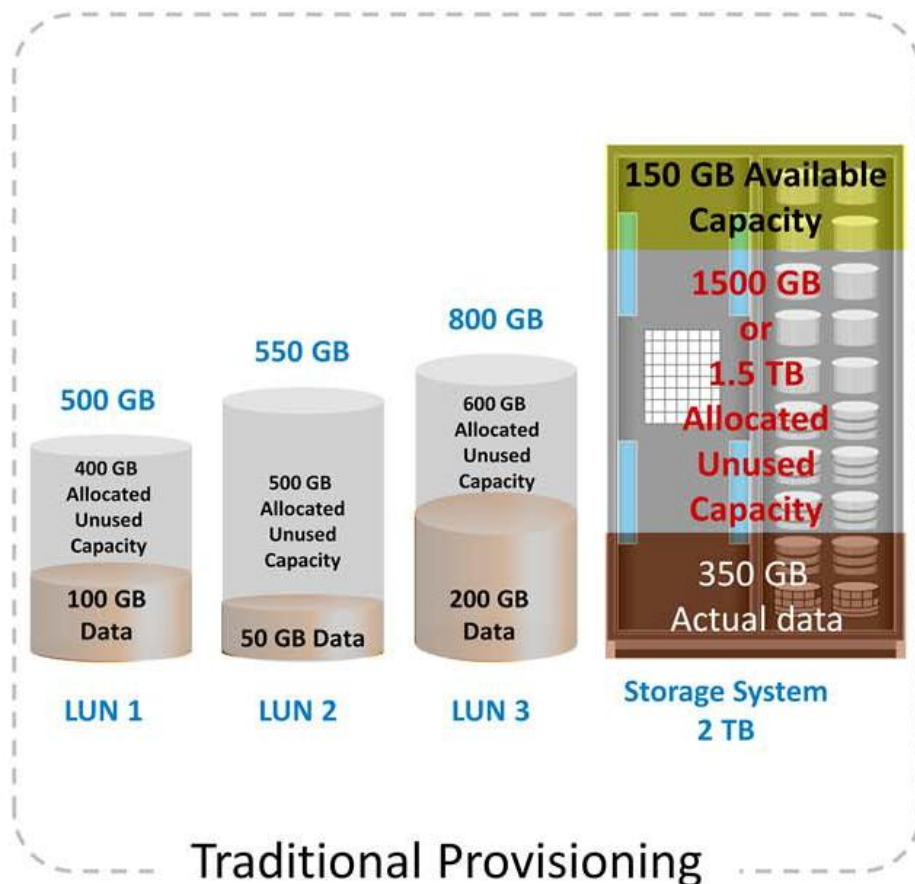
сегментирование

# Thin LUN (Virtual Provisioning)





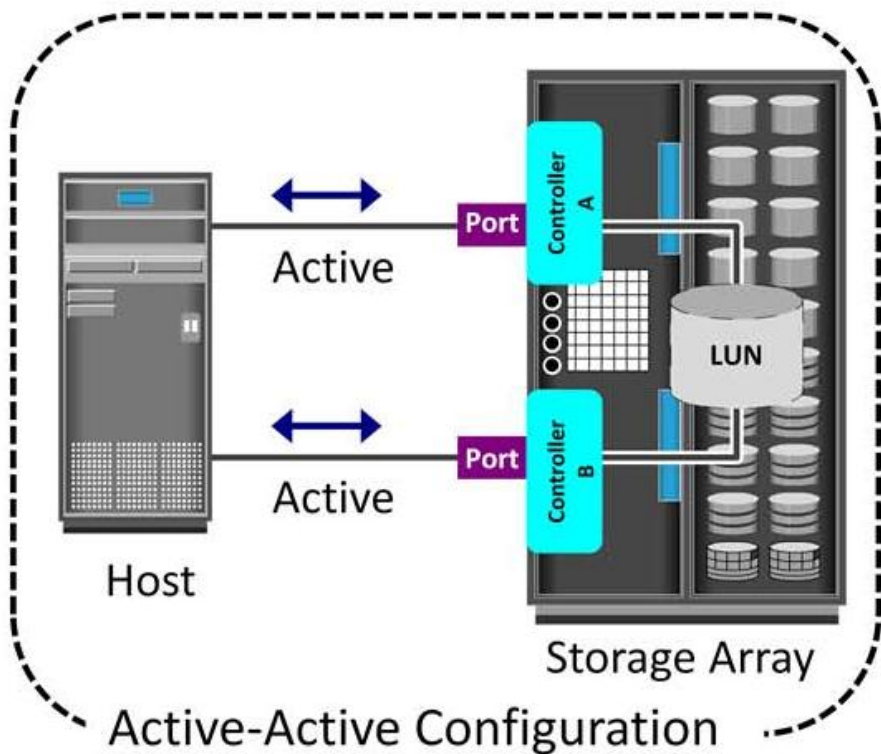
# Thin vs Thick



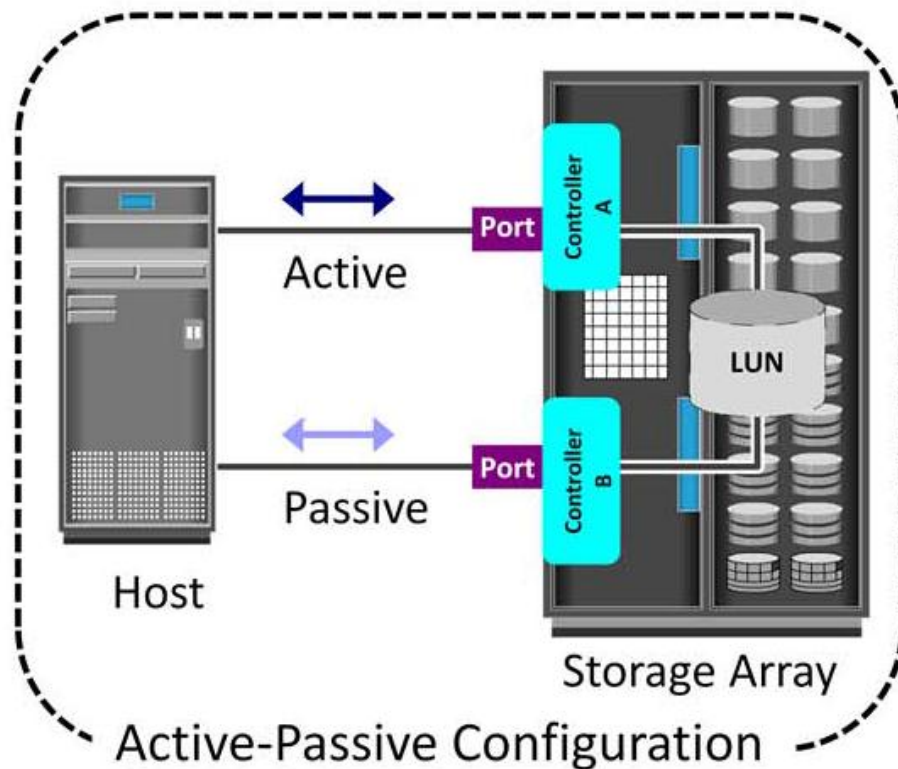


# Классы СХД

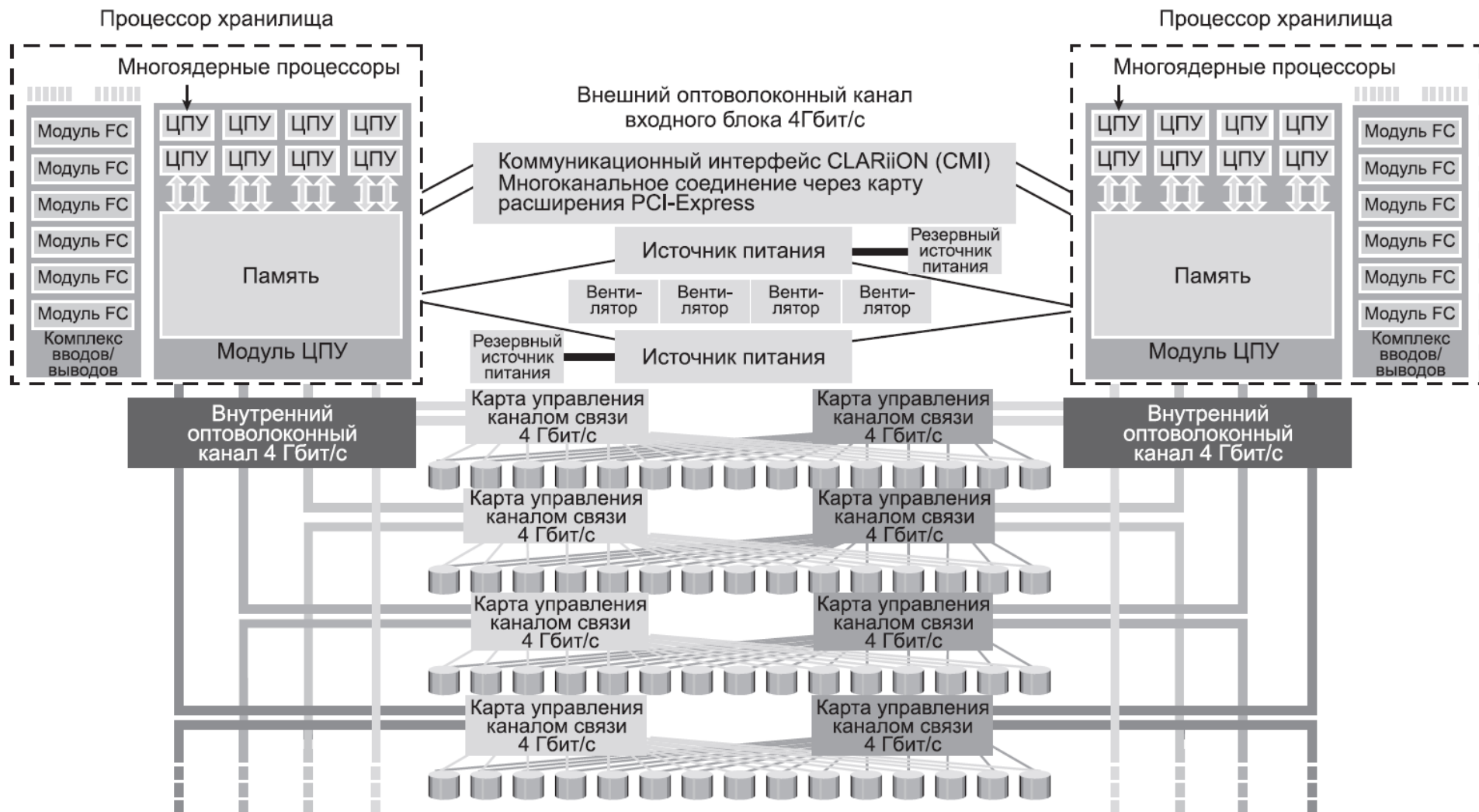
High end



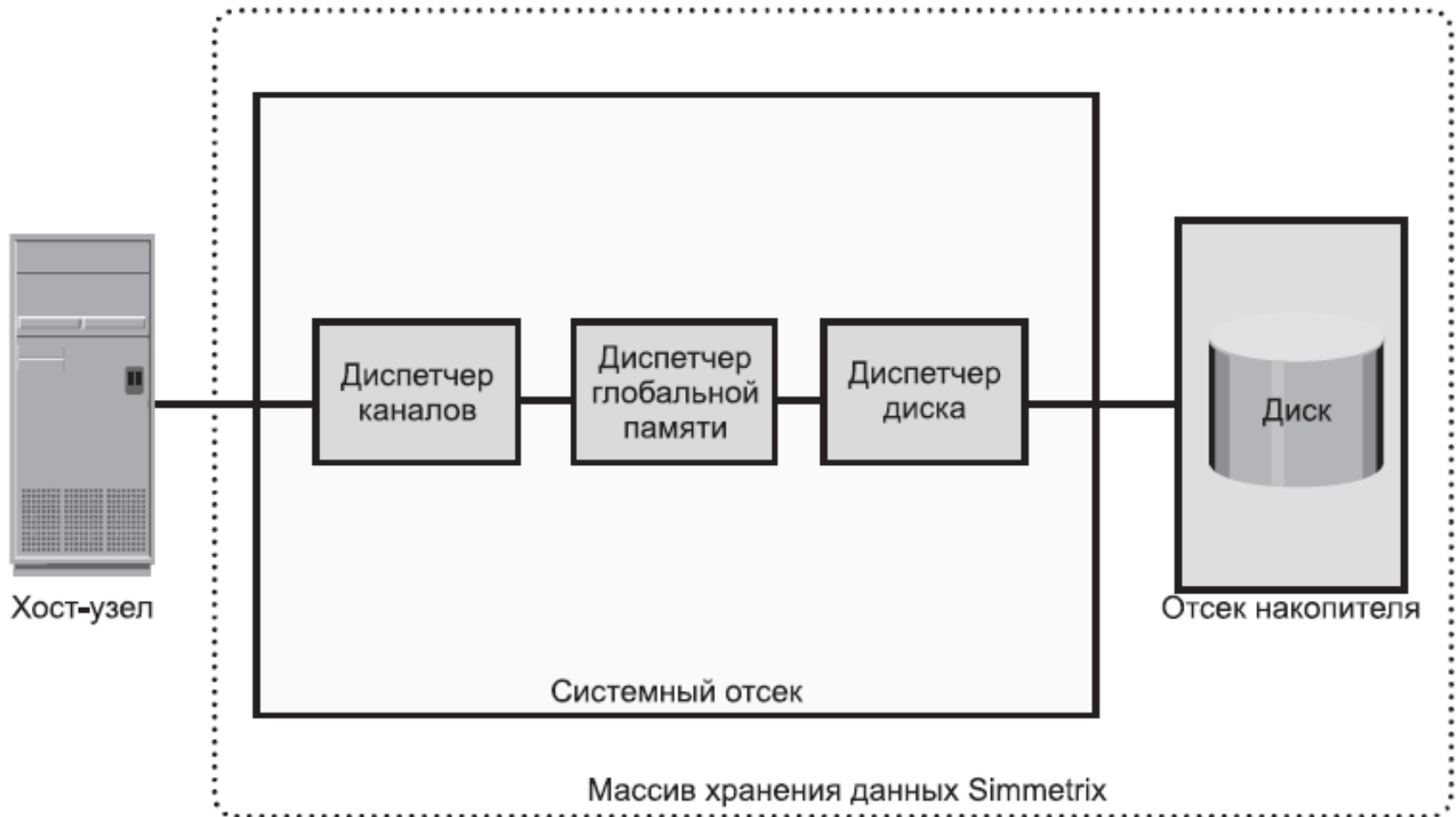
Mid range



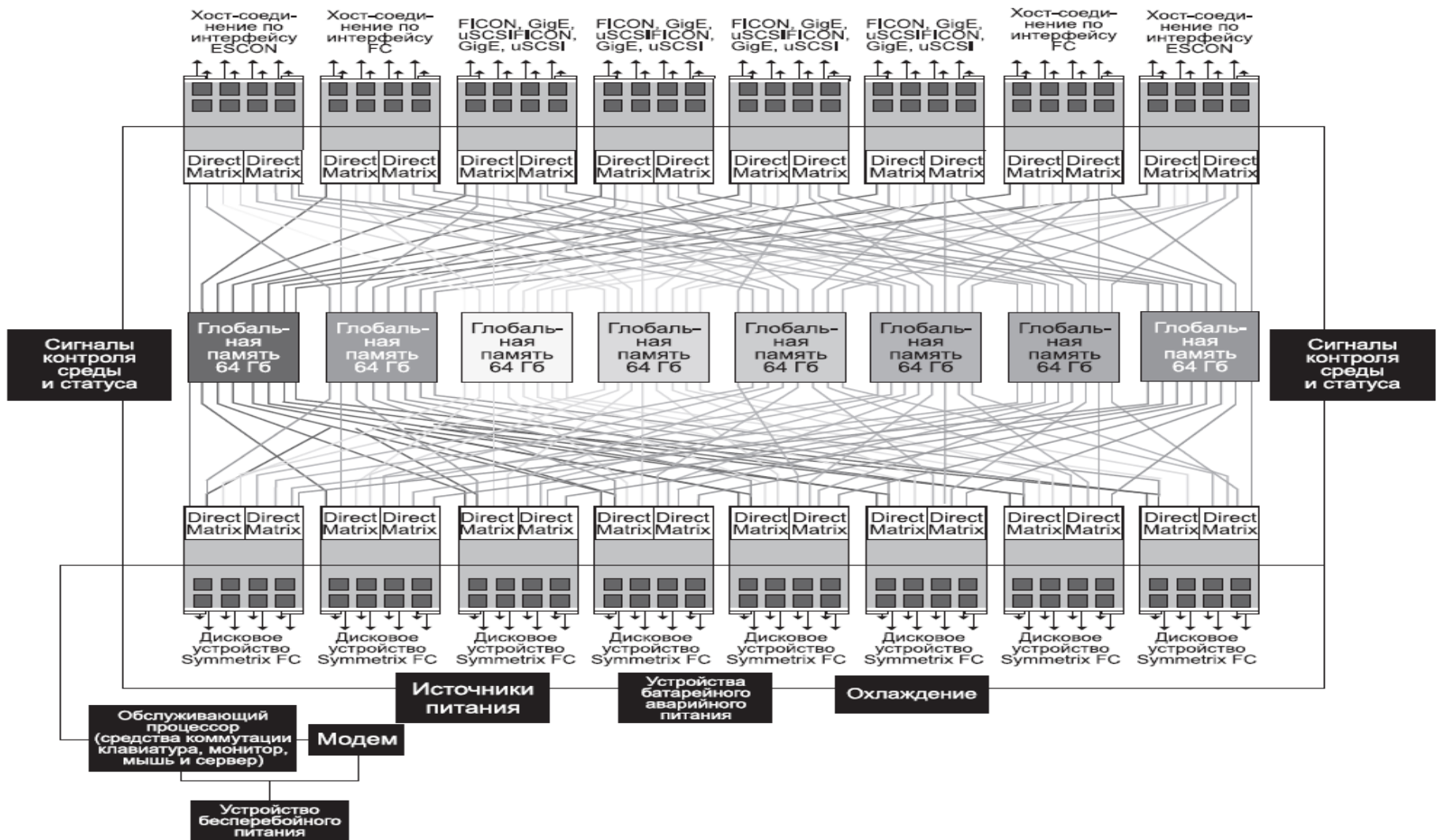
# CLARiiON (mid-range)



# Simmetrix (High-end)



# Direct Matrix



# Спасибо!

EMC<sup>2</sup>®