Виртуализация на сториджа

проф. д-р инж. Христо Вълчанов

http://cs.tu-varna.bg

Достъп до устройство за съхранение

- **Storage** – устройство за съхранение на данни

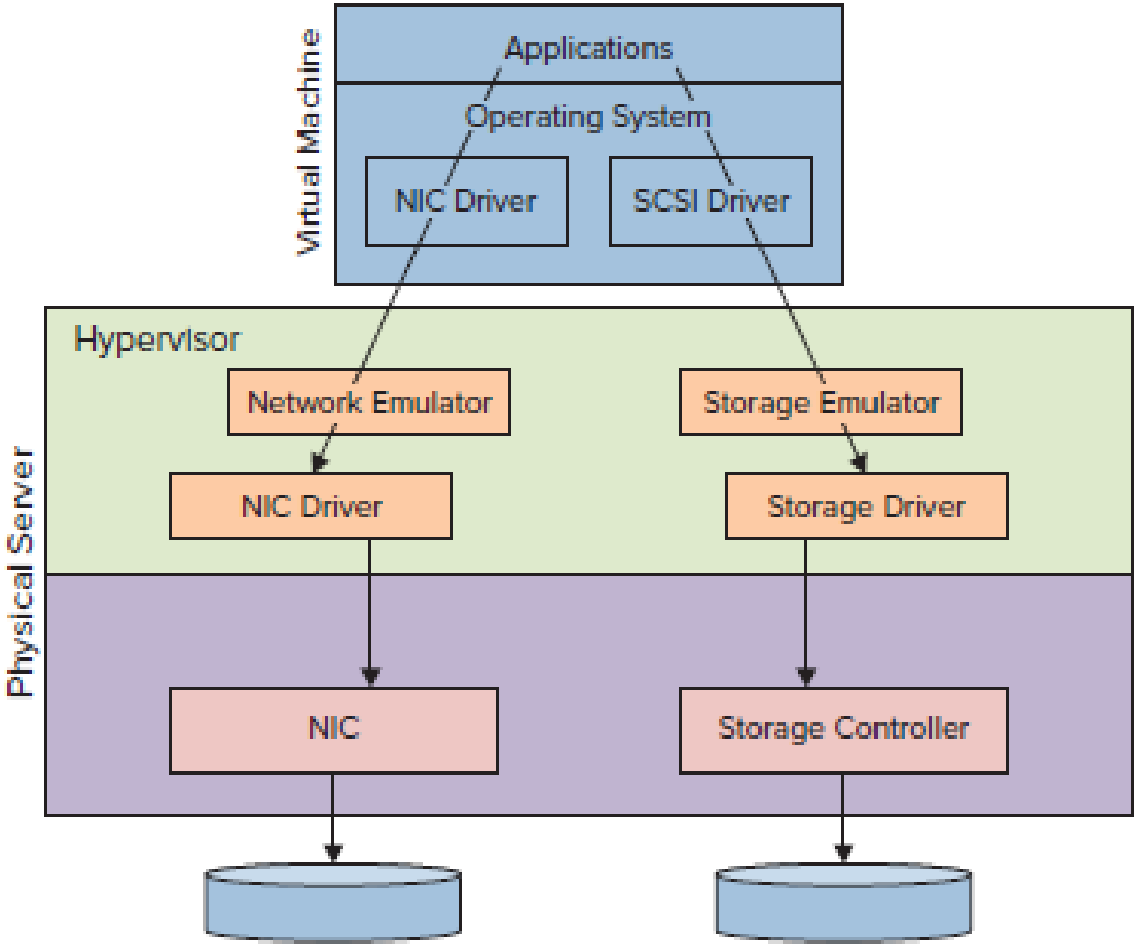
- Процес на получаване на данни:
  - Приложението изпраща заявка към ОС;
  - ОС изпраща заявката към подсистемата за съхранение;
  - Подсистемата открива информацията и я изпраща обратно към ОС като блокове данни;
  - ОС прехвърля тези блокове към приложението.

Виртуализация на сторидж

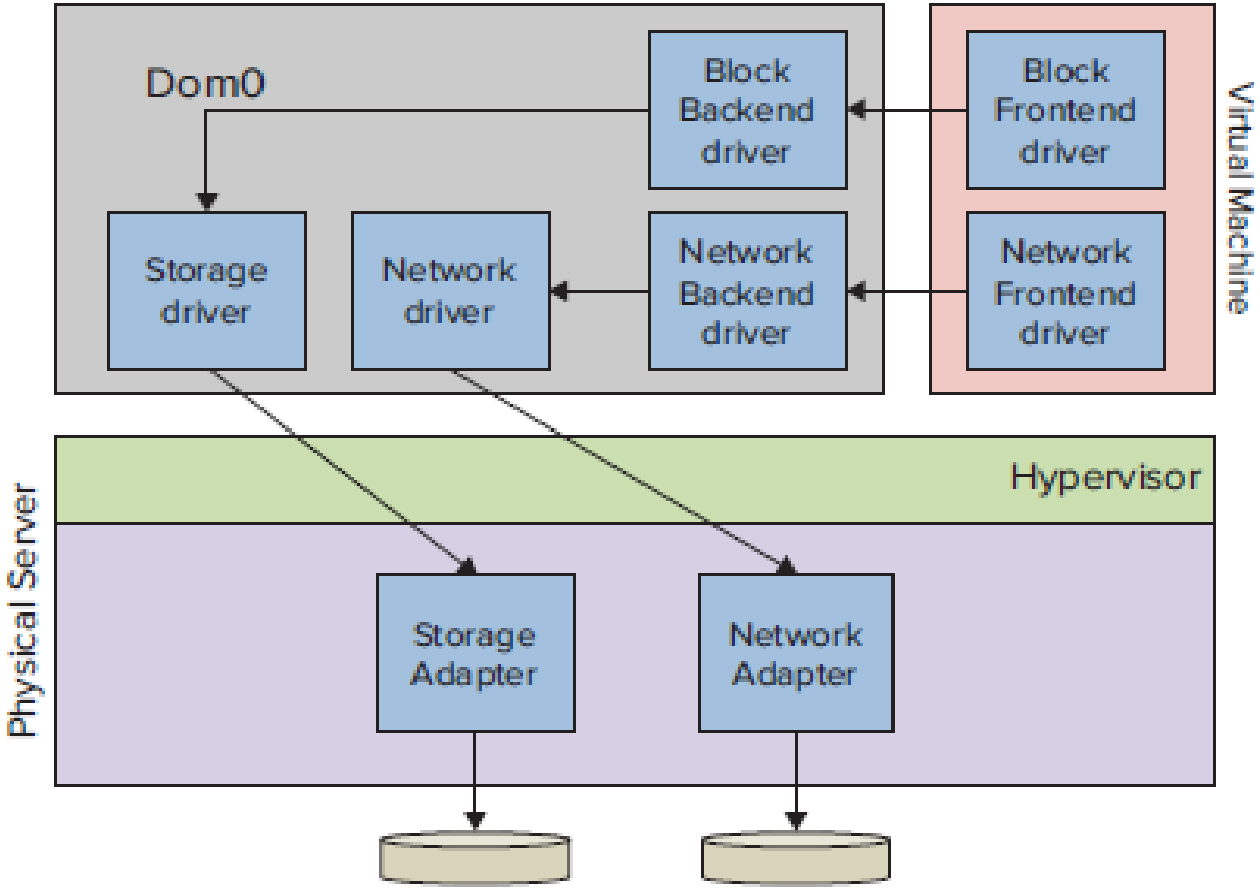
Rick F. Van der Lans:

„Виртуализацията на сторидж означава, че приложенията могат да използват сторидж без да се грижат къде той се намира, какъв интерфейс се използва и как е реализиран, каква платформа има и колко е наличен като обем.“

Достъп до устройство за съхранение във виртуална среда

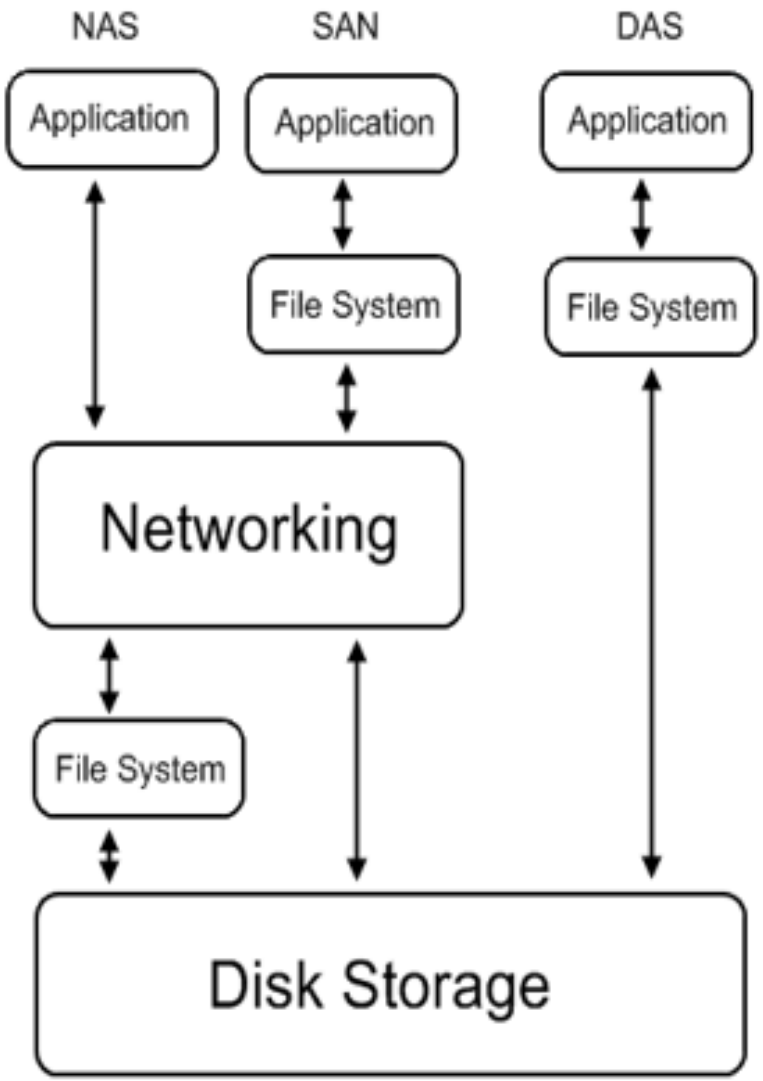


Достъп до устройство за съхранение във виртуална среда



Архитектури на сторидж

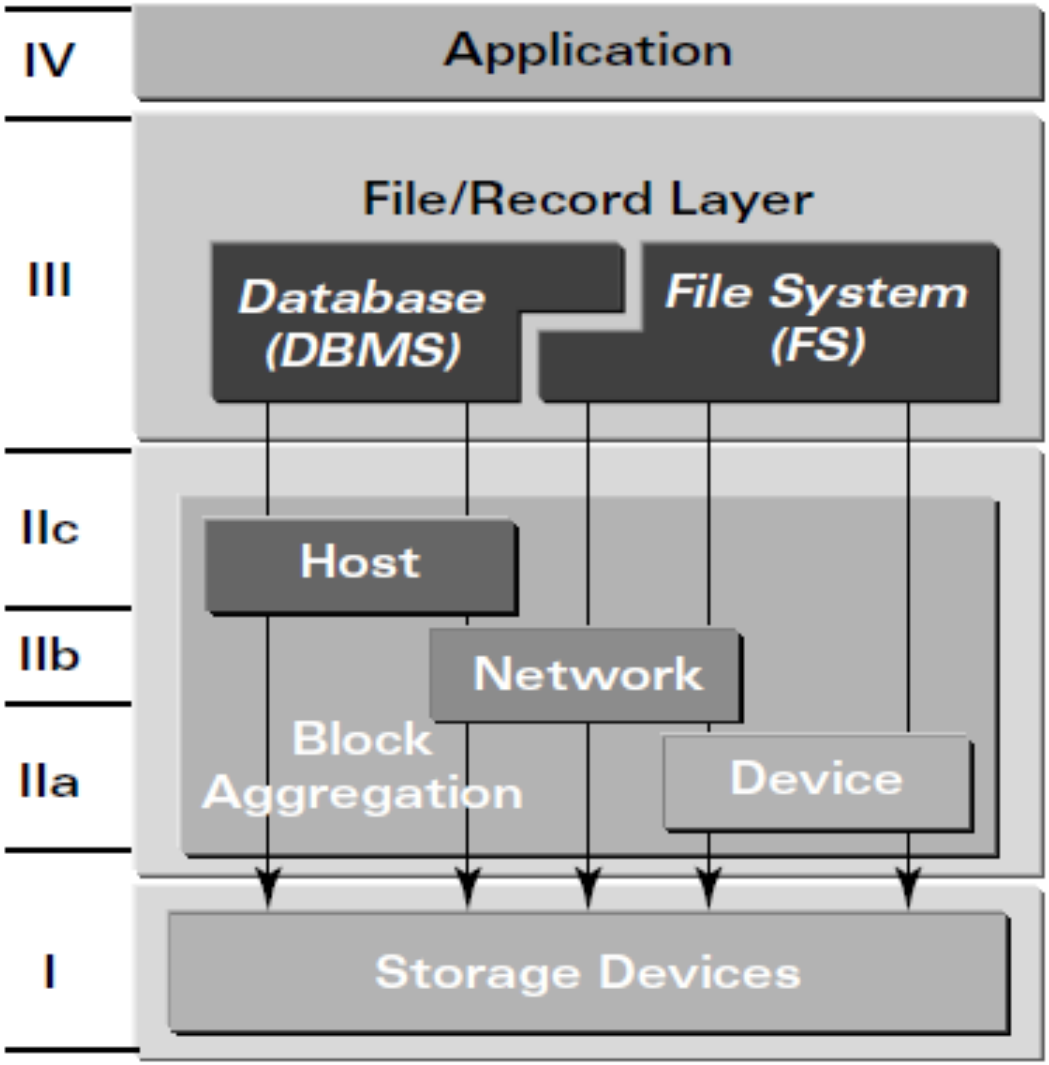
- DAS – Direct Attached Storage
  - Устройството е директно свързано към машина
- NAS – Network Attached Storage
  - Устройство за съхранение, свързано през мрежа към машина
- SAN – Storage Area Network
  - Свързване на отдалечени устройства към сървър, които изглеждат като едно локално устройство за ОС



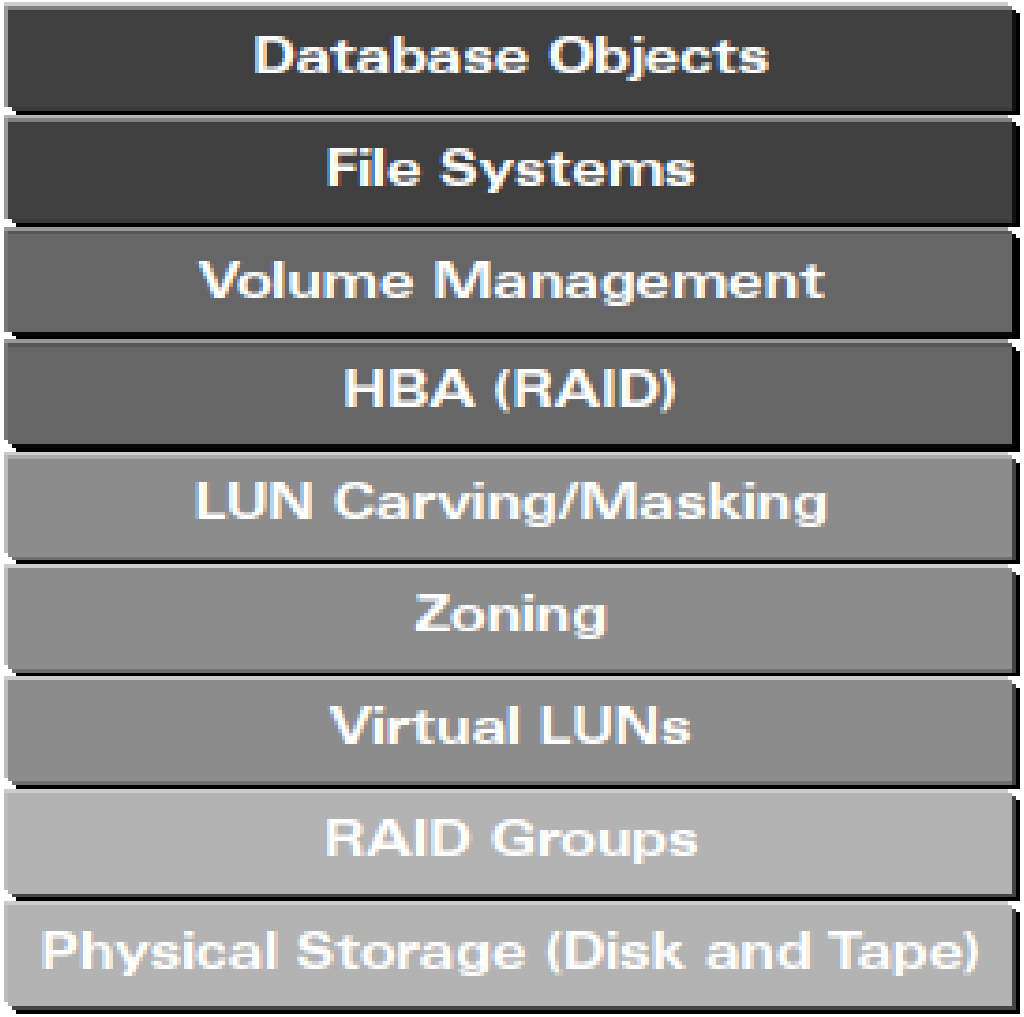
Изисквания към виртуализацията на сторидж

- Управляемост
- Наличност
- Разширяемост
- Сигурност

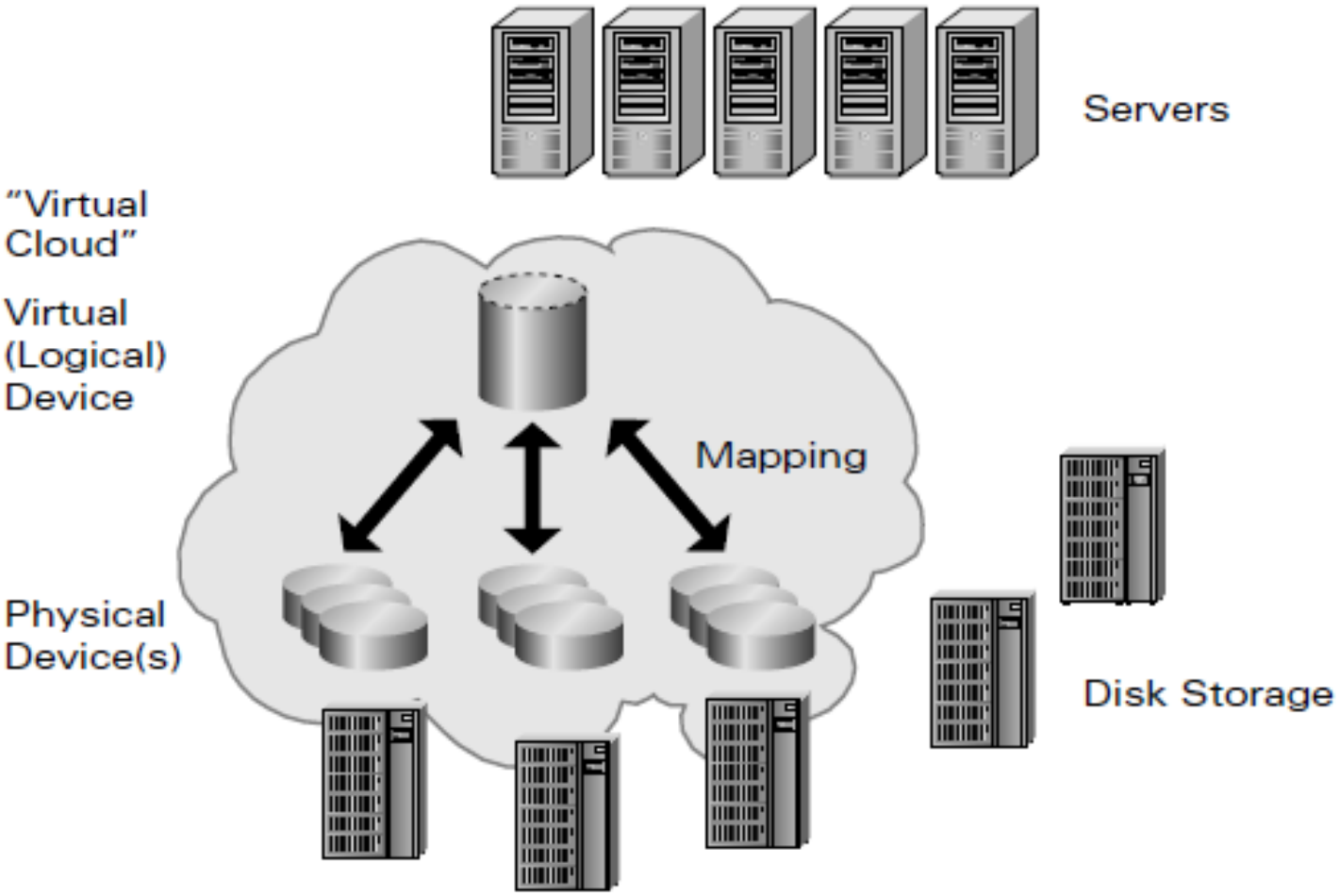
SNIA базов сторидж модел



Модел на функциите на сторидж

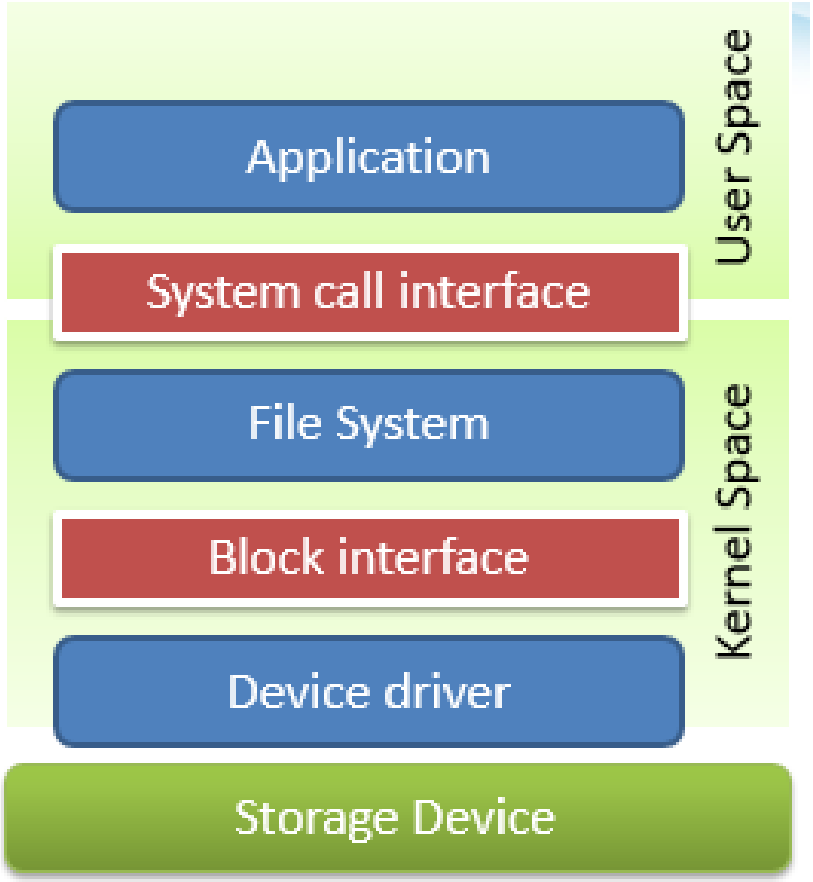


Виртуален сторидж



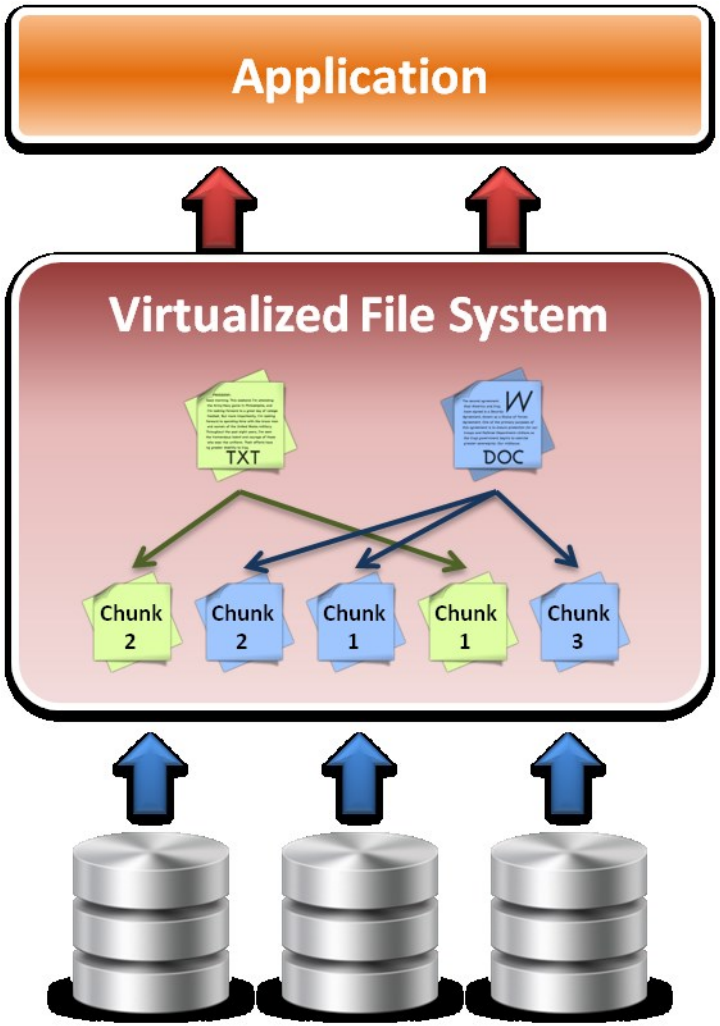
Слоеве за виртуализиране

- Файлова система
- Блокови устройства



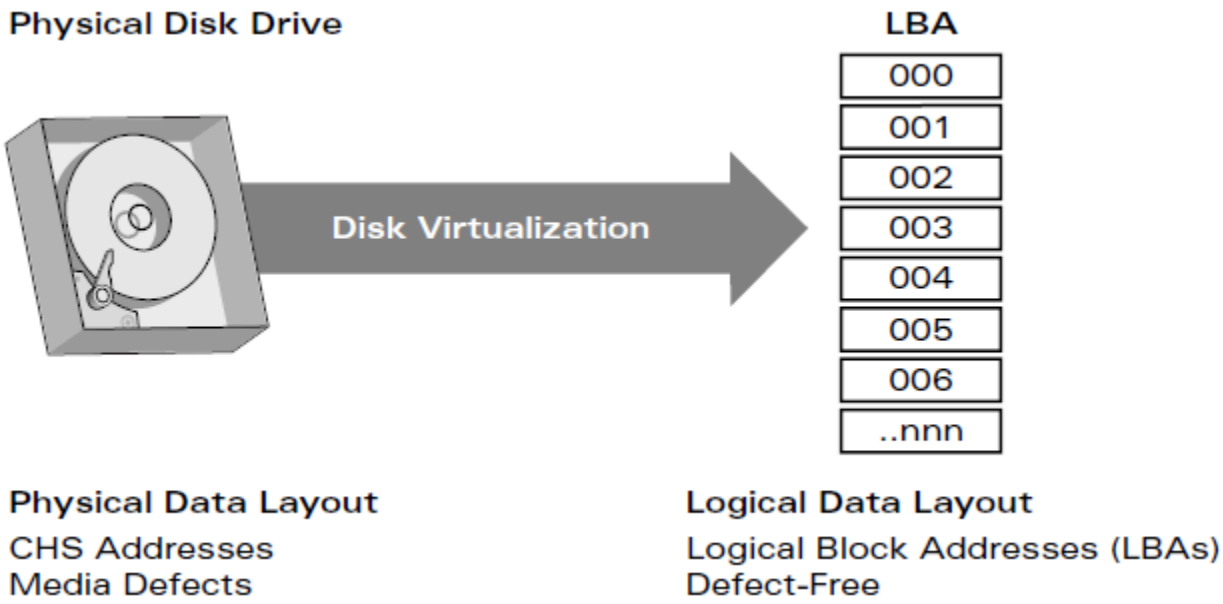
Виртуализиране на файлова система

- Данни
- Файл
- Мета данни
- Файлова система



Виртуализиране на блокови устройства

- Данни на диска – цилиндър, глава, сектор
- LBA – Logical Block Address



Виртуализиране на блокови устройства

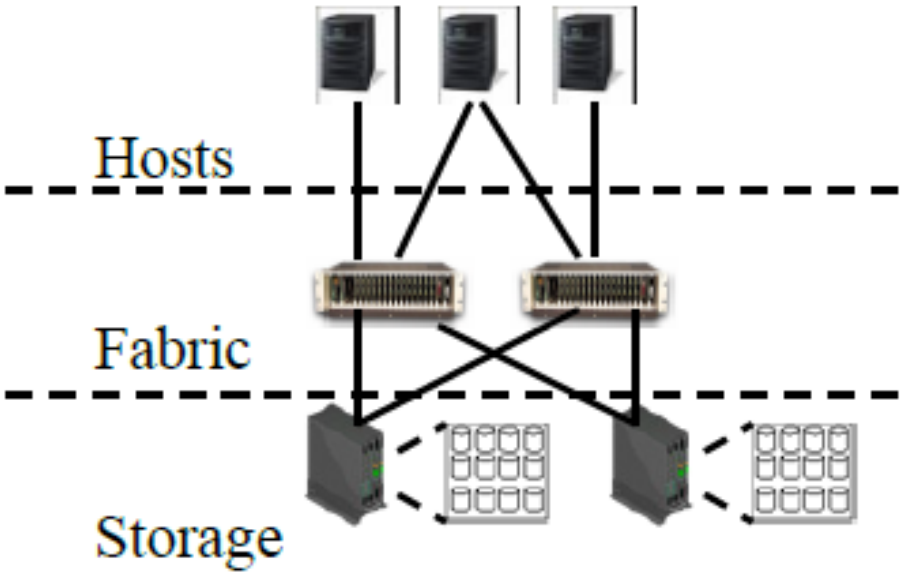
- Logical Unit (LU) – обработвана единица в сториджа със собствен номер (LUN)
- Logical Volume (LV) – том, представящ капацитета на едни или няколко дискове. Съдържат и метаданни.
- Единичен блок от информация се адресира чрез LUN и отместването в него - LBA

Протоколи за връзка със сторидж

- Fibre Channel (FC) – поддържа комуникация от вида point-to-point. Използва се при високи изисквания за производителност
- iSCSI (internet SCSI) – базиран на TCP/IP. Използва се при средни изисквания за производителност.
- FCIP (Fibre Channel over IP) – тунелен протокол за предаване на FC фреймове през TCP/IP.

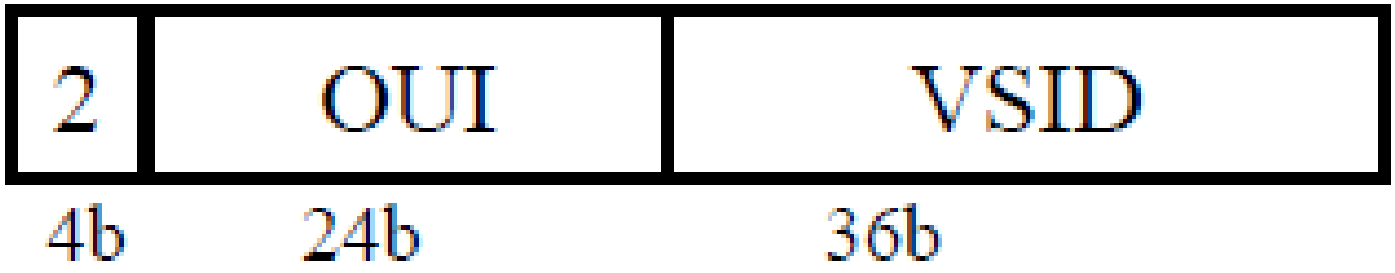
Fibre Channel

- ANSI T11 стандарт за високоскоростни SAN
- Скорости от 2, 4, 8, 16, 32 Gbps
- Топологии: point-to-point, ring, switched fabric



Fibre Channel

- Host Bus Adapter (HBA) – мрежово устройство за свързване на машина с SAN
- Всеки адаптер има уникален World Wide Number (WWN) – 64 бита



Подходи за виртуализация на сторидж

- Host-based – реализирана като софтуер, изпълняващ се на машина
- Network-based – реализирана на мрежови устройства
- Storage-based – реализирана на подсистемата на целевия сторидж

Host-based виртуализация

- Ниво файл – виртуализирана файлова система на хоста, съпоставяща файлове към блокове данни, разпределени между много сторидж устройства
- Ниво блок – изпълняващ се на хост софтуер за управление на логически том на хоста, прехващаш заявките за I/O и прехвърлящ ги към сторидж устройствата
- Предоставя софтуерен RAID
- Пример: NFS

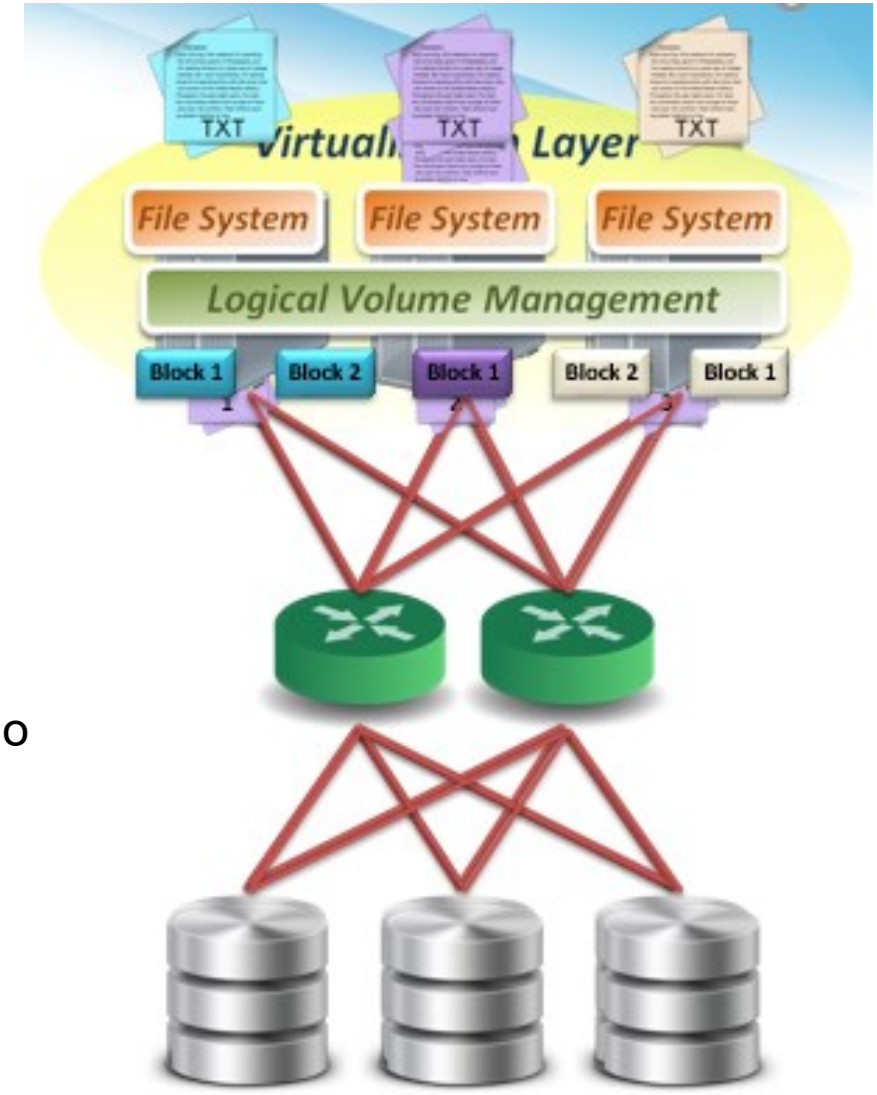
Host-based виртуализация

Предимства:

- Няма нужда от допълнителен хардуер
- Лесна реализация

Недостатъци:

- Използването на сторидж е само по отношение на хост
- Реализацията зависи от ОС
- Използва CPU цикли за виртуализацията



Storage-based виртуализация

- Ниво файл – изпълнява се софтуер на сторидж устройството, предоставящ базирани на файлове услуги на хост през мрежата
- Ниво блок – вградено в технологията на целевия сторидж
- Предоставя Storage Pooling, репликации и RAID, споделяне на данни
- Пример: дискови масиви

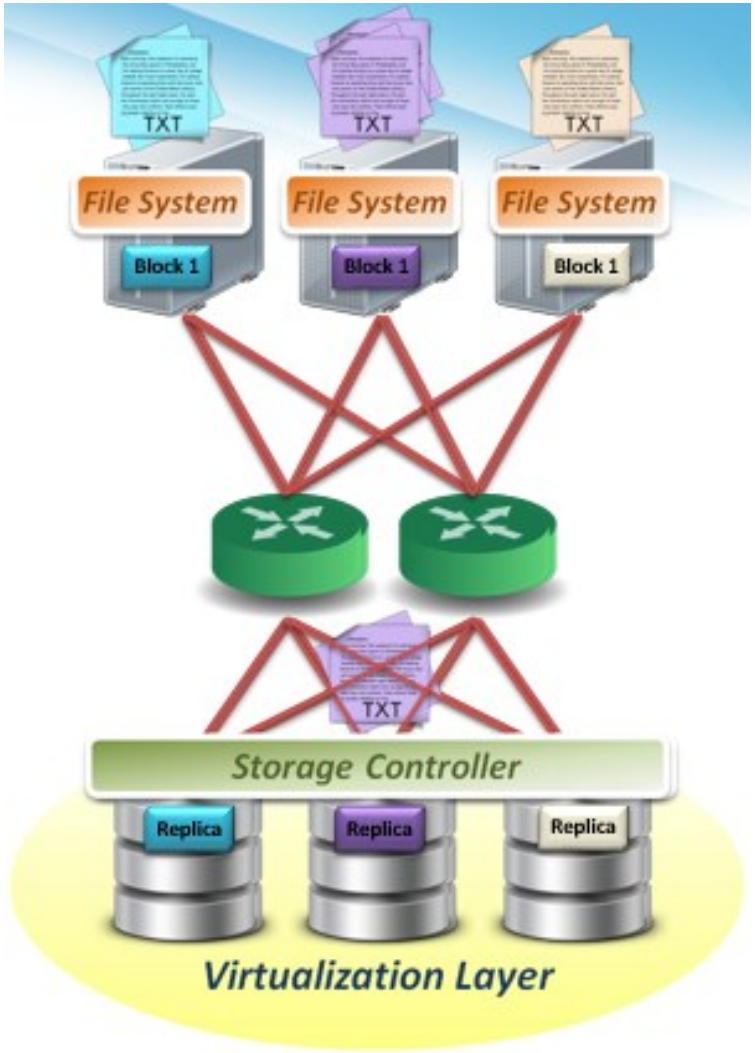
Storage-based виртуализация

Предимства:

- Редуцира латентността на индивидуален вход-изход

Недостатъци:

- Оптимизиране на сториджа само през свързани контролери
- Репликация и миграция на данни са възможни само през свързани контролери от един и същ производител



Network-based виртуализация

- Ниво файл – рядко се реализира на това ниво
- Ниво блок – софтуер, изпълняващ се на отделни appliance или интелигентни комутатори и рутери
- Предоставя Multi-path и Storage pooling
- Пример: IBM SVC (SAN Volume Controller)

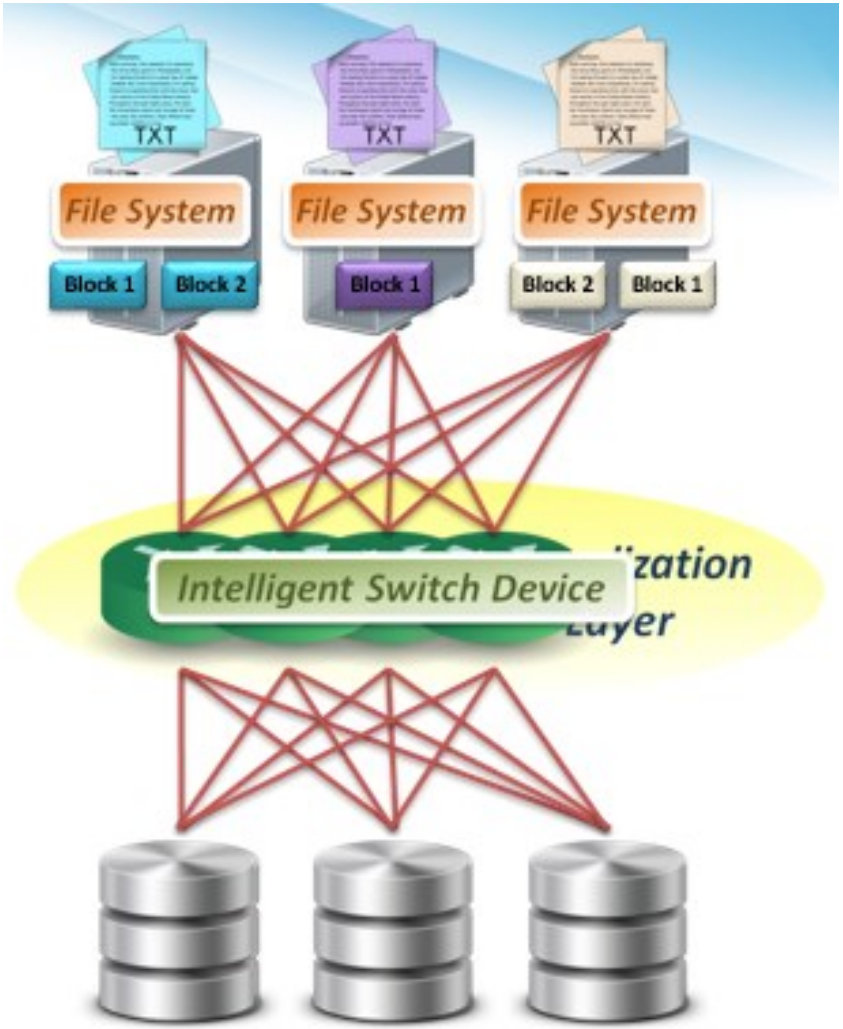
Network-based виртуализация

Предимства:

- Истинска хетерогенна сторидж виртуализация
- Няма нужда от модифициране на хост или сторидж системата
- Multi-path техники за производителност

Недостатъци:

- Сложна организация на взаимодействието, ограничена от поддръжката на производителя
- Изисква специфична мрежа (Fibre Channel)



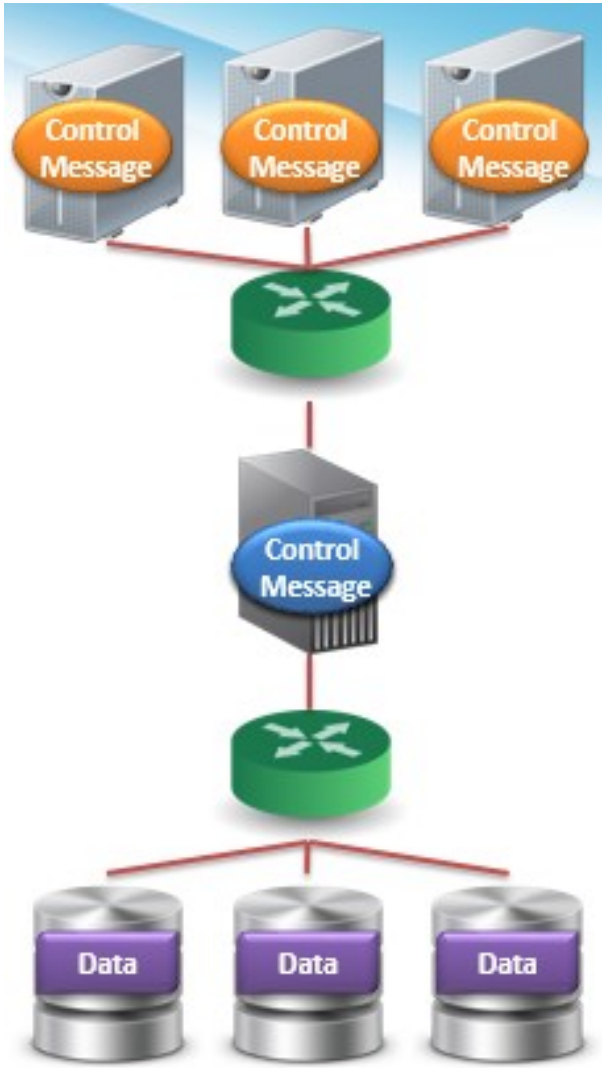
Реализиране на виртуализацията

- In-band
- Out-of-band

In-band виртуализация

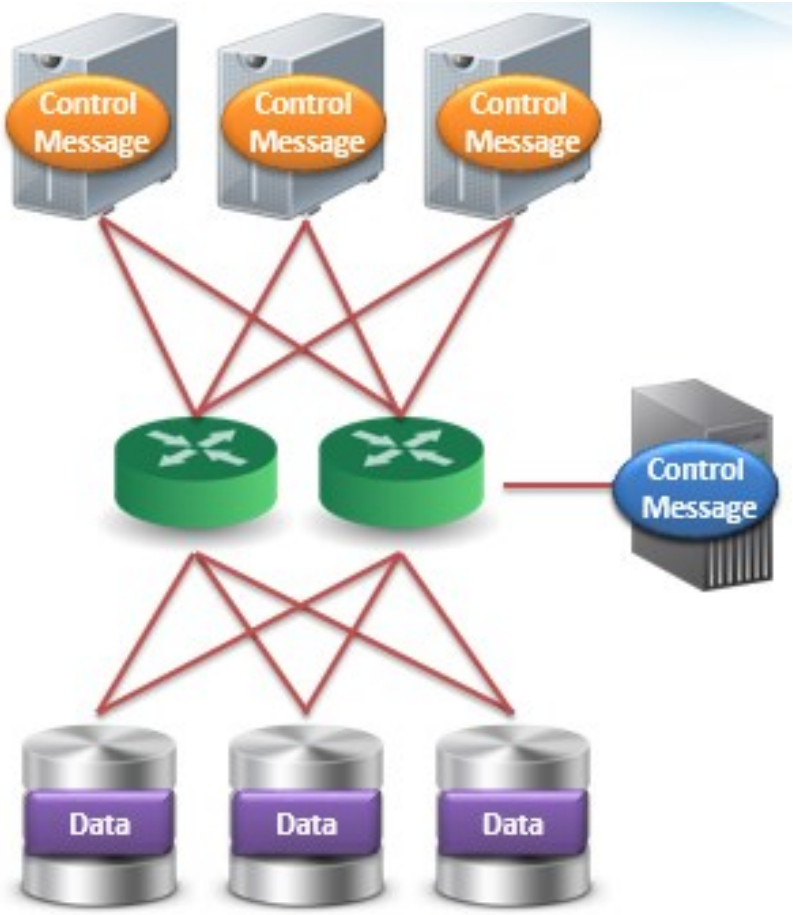
- Симетрична
- Виртуализационното устройство (ВУ) е разположено между хоста и сториджа
- Хостът извършва вход-изход към ВУ без директно да комуникира със сторидж устройството

- Лесно за реализация
- Неудобно за разширяване, проблемна точка



Out-of-band виртуализация

- Асиметрична
- Виртуализационното устройство (ВУ) – сървър за метаданни
- Изисква допълнителен софтуер в хоста, чрез който да се получи местоположението на актуалните данни



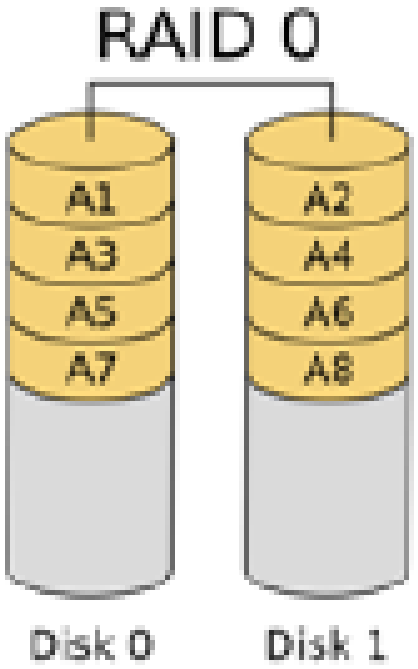
- Възможност за разширяване, висока производителност
- Трудно за реализация

Оптимизация на сториджа - RAID

- RAID (Redundant Array of Inexpensive Disks)
- Предоставя баланс между:
  - Надеждност
  - Наличност
  - Производителност
  - Капацитет

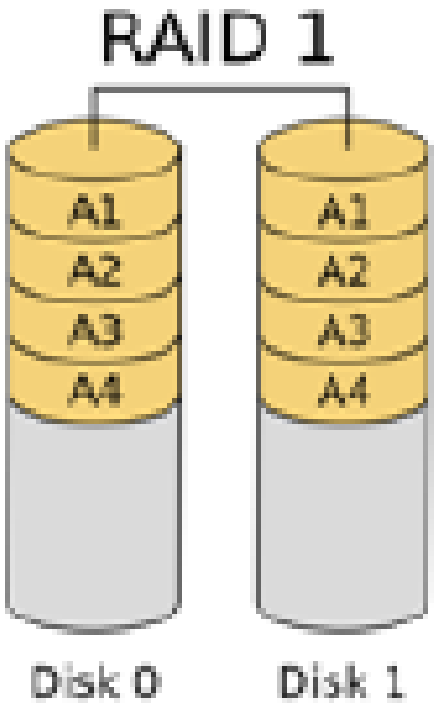
RAID 0

- Разцепване на блоково ниво без контрол по четност (parity) или дублиране (mirroring)
- Без резервираност
- Без отказоустойчивост
- Висока производителност и капацитет



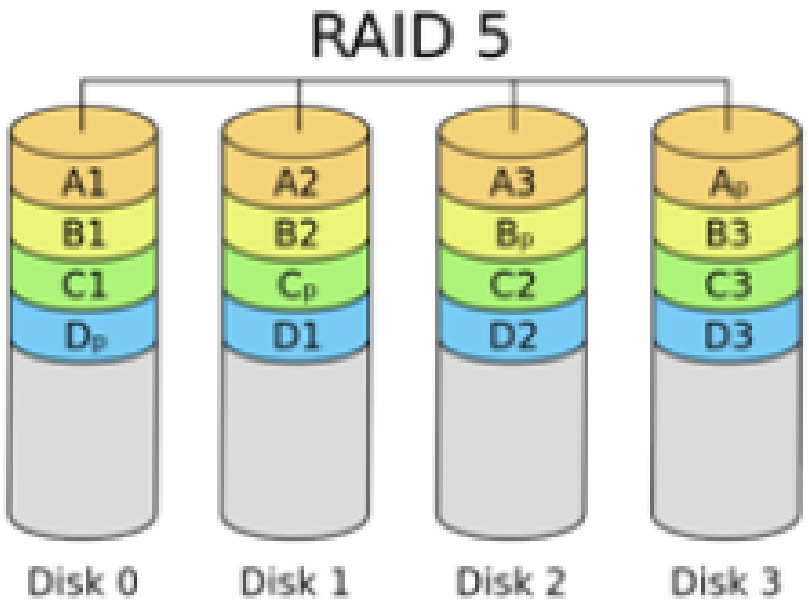
RAID 1

- Дублиране (mirroring) без контрол по четност
- Две или повече дублирания
- Висока производителност при четене – чете се от диск с по-малко време за преместване и латентност на въртене
- Записът зависи от по-бавния диск



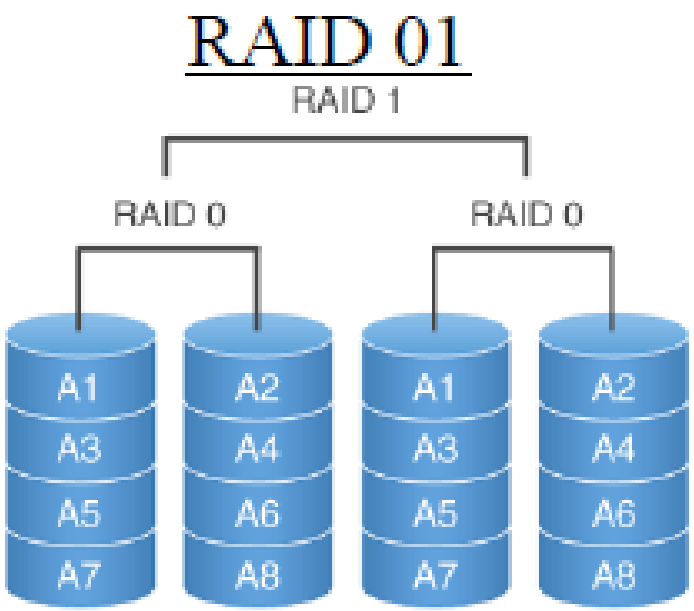
RAID 5

- Разцепване на блоково ниво с разпределен контрол по четност
- Разпределя паритета на отделен диск
- Изисква най-малко 3 диска



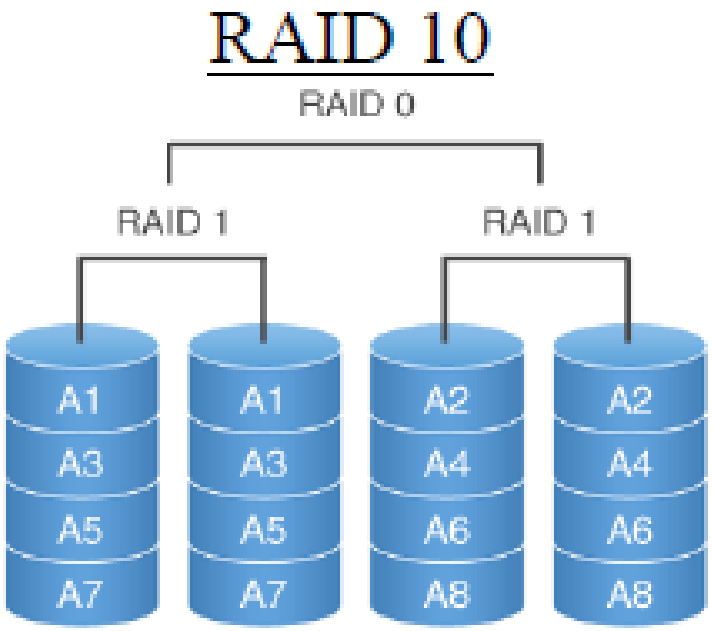
RAID 01

- Разцепване и после дублиране.
- Данните се разцепват между главните дискове, дублирането – на огледалните

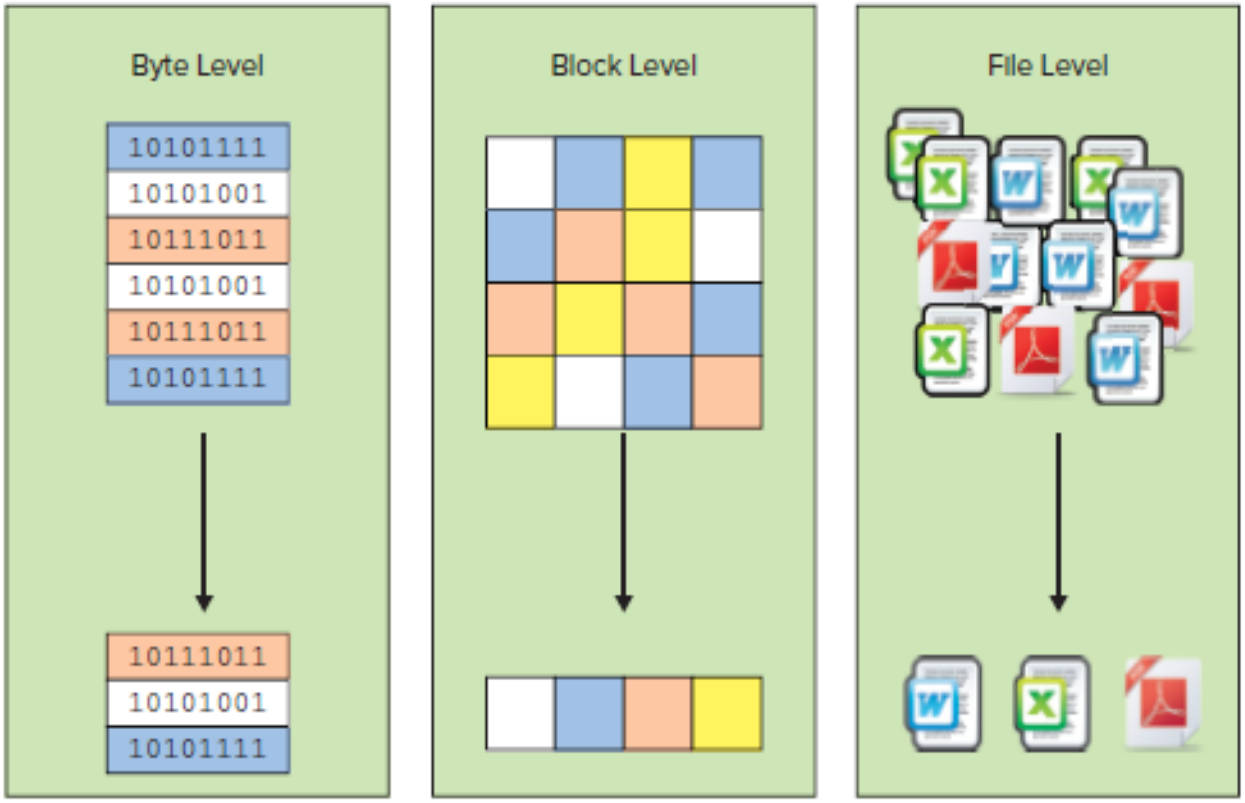


RAID 10

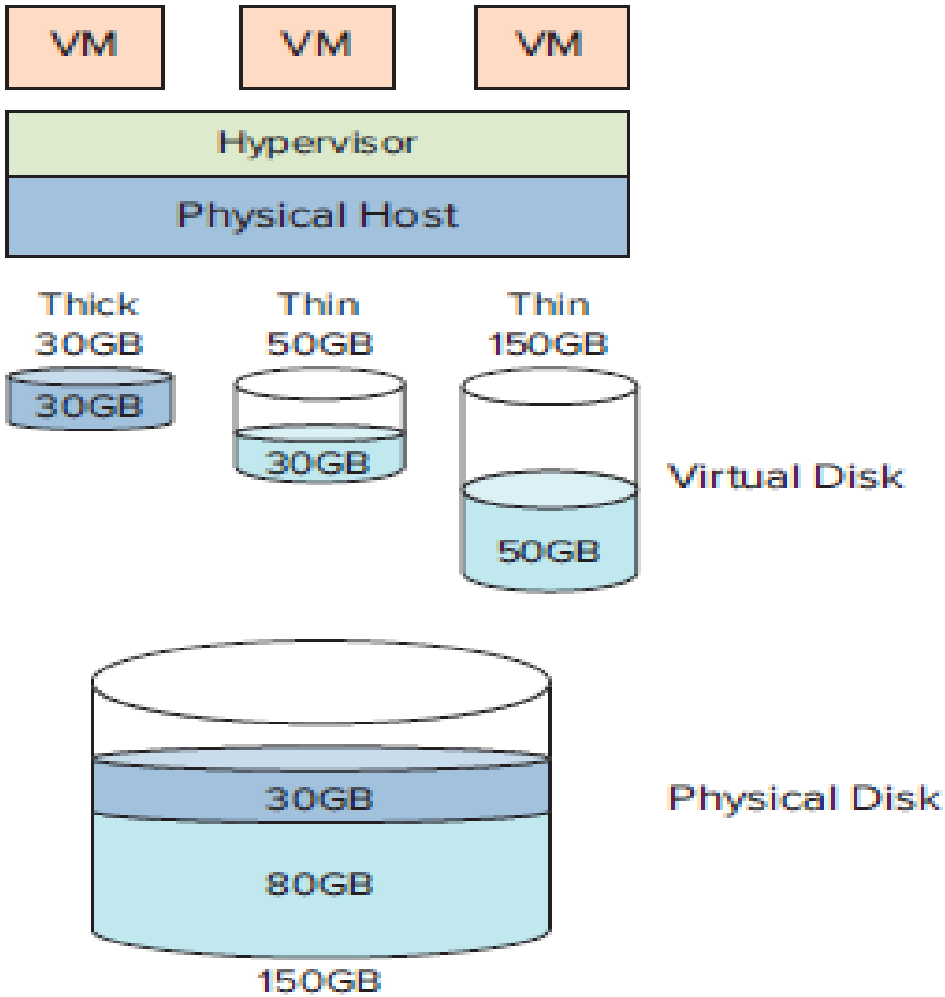
- Дублиране и после разцепване



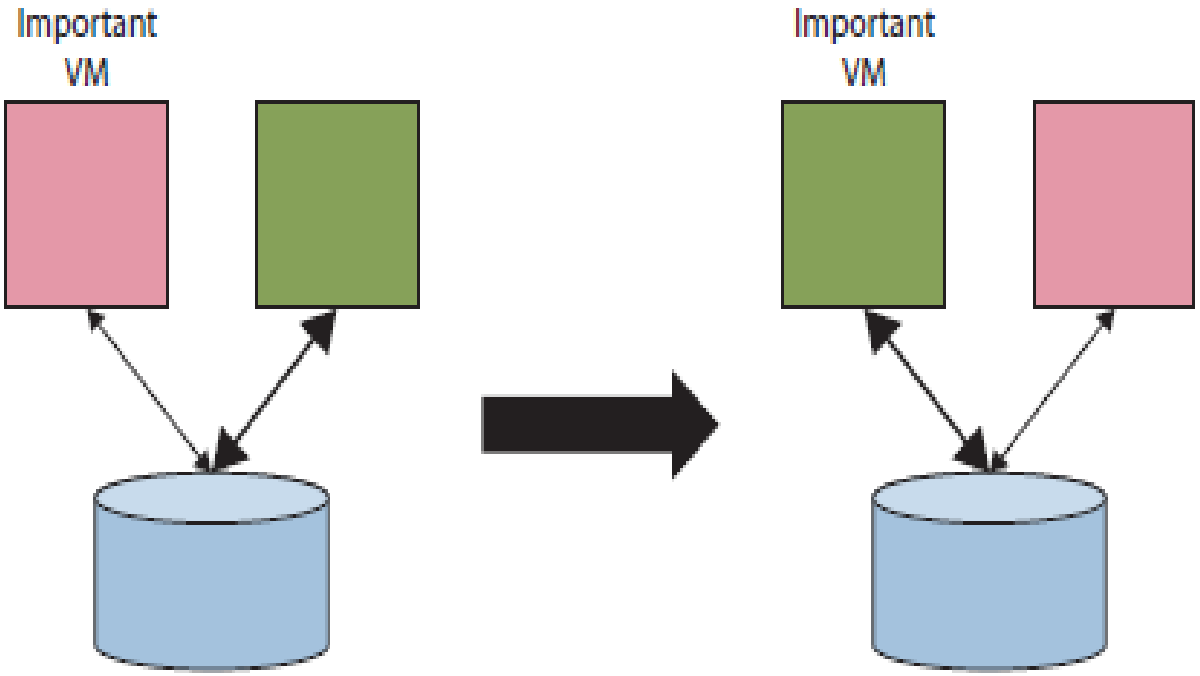
Оптимизация на сториджа – дедублиране на данни



Оптимизация на сториджа – гъвкав размер на диска



Оптимизация на сториджа – приоритети на достъп до диска



Въпроси?