

DAS, NAS, SAS

Обзор дисковых интерфейсов и архитектур

Предпосылки

- Необходимость получения больших объемов хранения
- Необходимость доступа к хранилищам множества вычислительных узлов
- Отношение к хранению данных как к отдельному инфраструктурному сервису

Задачи

- создание RAID-массивов, контроль SMART, HotPlug и HotSpare,
- обработка метаданных, позволяющих интерпретировать двоичные данные в виде файлов и записей,
- предоставление данных приложению.

DAS, NAS, SAS

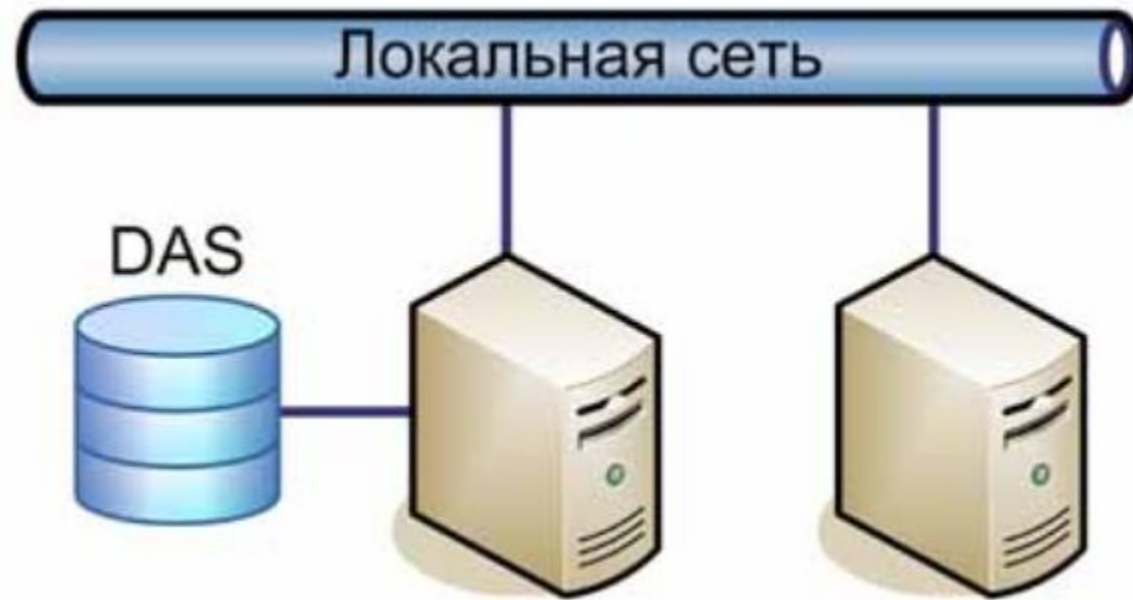
- Direct Attached Storage (DAS)
- Network Attached Storage (NAS)
- Storage Area Network (SAN)

DAS

- Direct Attached Storage (DAS)
 - СХД реализует лишь возможность хранения и доступа к данным.
 - Подключение к вычислительному узлу непосредственно.
 - Все остальное делается на стороне узла.

DAS

- Внешний доступ через сервисы сервера



Интерфейсы DAS

- USB, eSATA, FireWire, SAS, TunderBolt



DEPO Storage 2312 (DAS SAS3)



Интерфейсы DAS

- USB (до 10-20 Гбит/с)
- eSATA (как у SATA 2)
- SAS (соответствует версии SAS)
- FireWare - IEEE 1394 (400–3200 Мбит/с)
- TunderBolt (до 20Гбит / с)



Плюсы и минусы DAS

Плюсы

- Достаточно низкая стоимость.
- Простота развертывания и администрирования.
- Высокая скорость обмена между дисковым массивом и сервером.

Минусы

- Низкая надежность системы в целом – при возникновении проблем в сети или аварии сервера данные становятся недоступны всем сразу.
- Высокая латентность, обусловленная обработкой всех запросов одним сервером и используемым транспортом.
- Высокая загрузка сети, часто определяющая пределы масштабируемости путём добавления клиентов.
- Плохая управляемость – вся ёмкость доступна одному серверу, что снижает гибкость распределения данных.
- Низкая утилизация ресурсов – трудно предсказать требуемые объёмы данных, у одних устройств DAS в организации может быть избыток ёмкости (дисков), у других её может не хватать – перераспределение часто невозможно или трудоёмко.
- Невозможность совместно использоваться СХД несколькими серверами

NAS

- Network Attached Storage (NAS)
 - создание RAID-массивов, контроль SMART, HotPlug и HotSpare,
 - обработка метаданных, позволяющих интерпретировать двоичные данные в виде файлов и записей,
 - предоставление данных приложению через высокоуровневые сетевые протоколы.
 - Непосредственного доступа к дискам у клиентов нет.
 - Интерфейс – сетевой прокол прикладного уровня.
- NAS отличается тем, что в СХД переносятся к тому же и метаданные для обеспечения файлового доступа, здесь клиенту остается только лишь поддерживать сервисы данных.

NAS

- Внешний доступ непосредственно через сеть



NAS. Операционные системы

SOHO

- FreeNAS
Наиболее распространённая ОС на базе FreeBSD.
- OpenMediaVault
Основа на Debian Linux.
- XPEology
Synology DSM. Используется в коммерческих NAS от Synology.
- Rockstor
Основа на CentOS.

NAS. Операционные системы

SOHO \ Enterprise

Общего назначения:

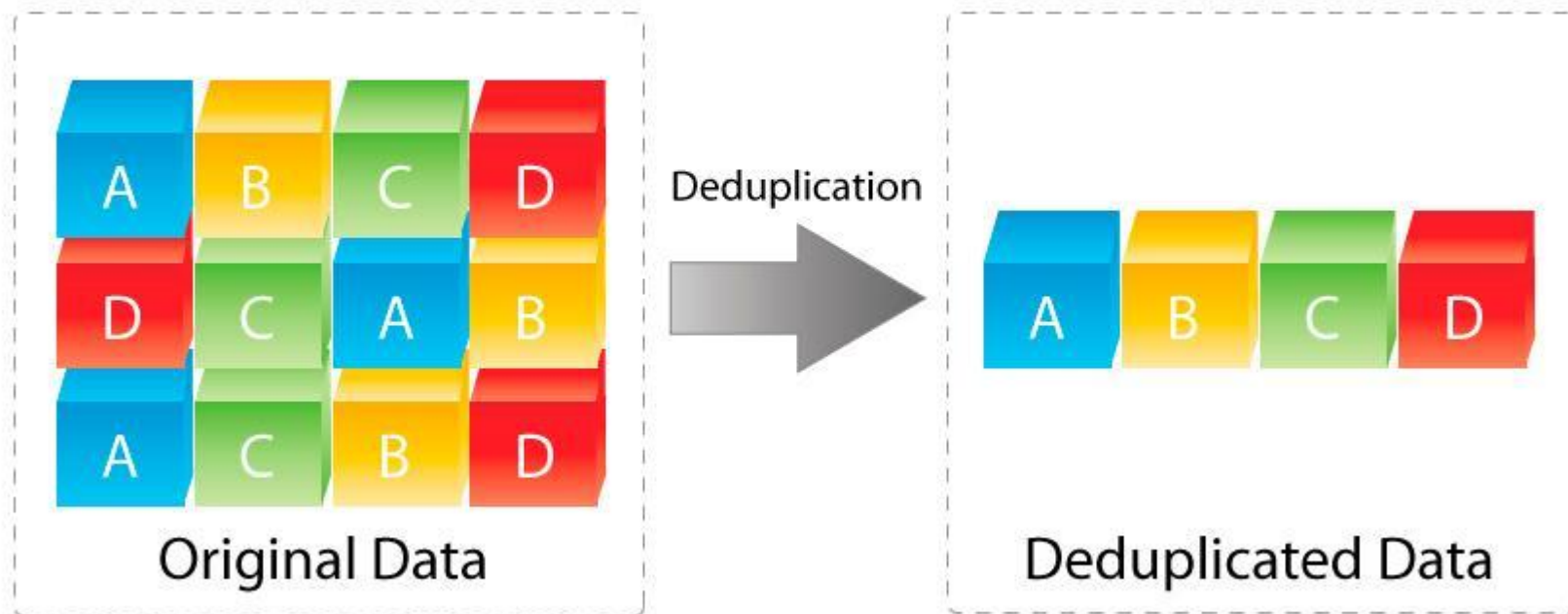
- Linux \ FreeBSD
- Windows Storage Server

Специализированные:

- Open-E DSS
- RAIDIX 4.X
- И д.р.

NAS. Задачи ОС

- Доступ к данным
- Каталогизация
- Безопасность
- Репликация
- Дедупликация



NAS



- 1 Гб\с GigabitEthernet

NAS



Система хранения данных DEPO Storage

3516

До 96 Тб

Доступ через сеть (1/10/40 Гбит / с)

NAS (иногда)

Приводы	(96) Fibre Channel/FATA/SSD; Поддерживается в зависимости от модели и наличия отсека расширения
Функции обеспечения доступности	Полностью симметричный режим Active/Active
Размер файловой системы (макс.)	128 ТБ; Linux; 32 ТБ Windows
Количество файловых систем на кластер (макс.)	512; Linux; Windows 256
Размер кластера (макс.)	16 узлов
Поддерживаемые стоечные серверы	2 контроллера файловых серверов (стандартно)
Интерфейс массива	Fibre Channel
Поддерживаемые платформы BladeSystem	HP ProLiant BladeSystem
Сетевые протоколы	NFS, CIFS, HTTP, FTP, HTTPS, iSCSI
Поддержка RAID	0, 1, 5, 6, 0+1
Поддержка Fibre channel (SAN)	MSA, EVA, XP
Поддержка iSCSI	MSA, EVA
Гарантия (детали-качество сборки-поддержка на месте)	3/3/3



Протоколы доступа

- CIFS
 - SMB1
 - SMB2
 - SMB3
- NFS
- FTP и TFTP
- Rsync
- WebDAV

NAS

Достоинства NAS:

- доступность и низкая стоимость;
- простое совместное использование информации;
- простота подключения и управления;
- гибкость, возможность быстро увеличить объём для хранения данных;
- универсальность клиентов (компьютер под управлением любой операционной системы может получить доступ к файлам).

Недостатки NAS:

- хранение данных только в виде файлов;
- медленный доступ к информации по сетевым протоколам (по сравнению с локальной системой);
- высокая загрузка ЛВС;
- невозможность работы некоторых приложений с сетевыми дисками.

SAN (Storage Area Network)

Storage Area Network (SAN)

- создание RAID-массивов, контроль SMART, HotPlug и HotSpare,
- обработка метаданных, позволяющих интерпретировать двоичные данные в виде файлов и записей,
- предоставление доступа к устройствам хранения (дискам или их логическому эквиваленту) непосредственно вычислительным
- Интерфейс – для узлов стандартные дисковые протоколы.
- Физический интерфейс – компьютерная сеть
 - Общего назначения
 - Специализированная
- Дисковые команды инкапсулируются в сетевые сообщения.
- Нужна специальная инфраструктура

SAN vs NAS

Способ организации обмена данными между устройствами хранения и серверами:

- NAS – высокоуровневые прикладные протоколы.
- SAN – низкоуровневые протоколы. Устройства «видят» удаленные диски как локальные.

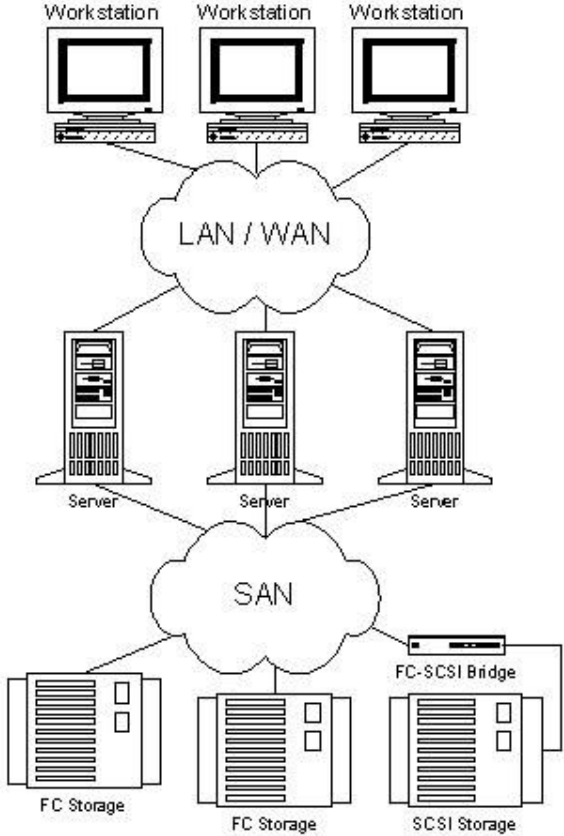
Масштабируемость:

- NAS – выше, чем у DAS
- SAN – выше, чем у NAS

Пропускная способность:

- SAN – выше.

SAN



Основные виды SAN

На сетях общего назначения:

- iSCSI

Специализированные каналные протоколы:

- FiberChanel
- InfiniBand

Основные компоненты

Технологически SAN состоит из следующих компонентов:

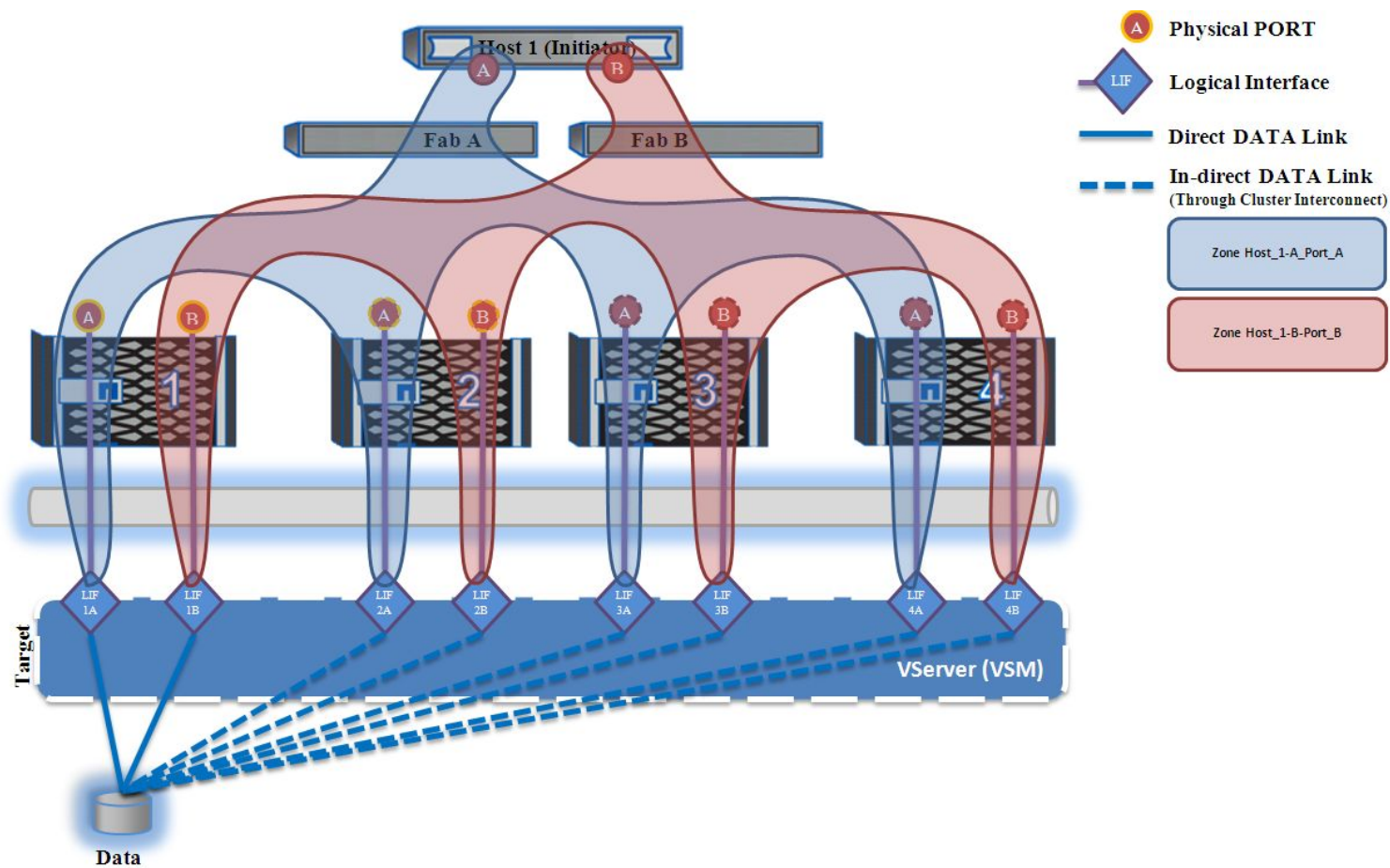
- Узлы, ноды (nodes)
 - Дисковые массивы (системы хранения данных) — хранилища (таргеты - targets)
 - Серверы — потребители дисковых ресурсов (инициаторы - initiators).

- Сетевая инфраструктура
 - Коммутаторы (и маршрутизаторы в сложных и распределённых системах)
 - Кабели

Дополнительные понятия

- WWN-адрес (World Wide Name (WWN) или World Wide Identifier (WWID)) — уникальный идентификатор, который определяет конкретное целевое устройство (таргет) Fibre Channel, Advanced Technology Attachment (ATA) или Serial Attached SCSI (SAS). Каждый WWN представляет собой 8-байтное число, производное от IEEE OUI и информации, предоставляемой производителем.
- Fabric (Фабрика) — совокупность коммутаторов, соединённых между собой межкоммутаторными линками (*ISL — InterSwitch Link*.)
- Zoning – технология, обеспечивающая доступность target C определенных initiators.
 - выбранные пары добавляются в предварительно созданные на коммутаторе зоны (zones);
 - зоны помещаются в наборы зон (zone set, zone config), созданные там же;
 - наборы зон активируются в фабрике.

Дополнительные понятия



Топологии SAN

- Каскад — коммутаторы соединяются последовательно. Если их больше двух, то ненадёжно и непроизводительно.
- Кольцо — замкнутый каскад. Надёжнее просто каскада, хотя при большом количестве участников (больше 4) производительность будет страдать. А единичный сбой ISL или одного из коммутаторов превращает схему в каскад со всеми вытекающими.
- Сетка (mesh). Бывает Full Mesh — когда каждый коммутатор соединяется с каждым. Характерно высокой надёжностью, производительностью и ценой. Количество портов, требуемое под межкоммутаторные связи, с добавлением каждого нового коммутатора в схему растёт экспоненциально. При определённой конфигурации просто не останется портов под узлы — все будут заняты под ISL. Partial Mesh — любое хаотическое объединение коммутаторов.
- Центр/периферия (Core/Edge) — близкая к классической топологии LAN, но без уровня распределения. Нередко хранилища подключаются к Core-коммутаторам, а серверы — к Edge. Хотя для хранилищ может быть выделен дополнительный слой (tier) Edge-коммутаторов. Также и хранилища и серверы могут быть подключены в один коммутатор для повышения производительности и снижения времени отклика (это называется локализацией). Такая топология характеризуется хорошей масштабируемостью и управляемостью.

Примеры оборудования SAN (Fiber Channel)

- Коммутатор HP StorageWorks 4/8 Base SAN Switch

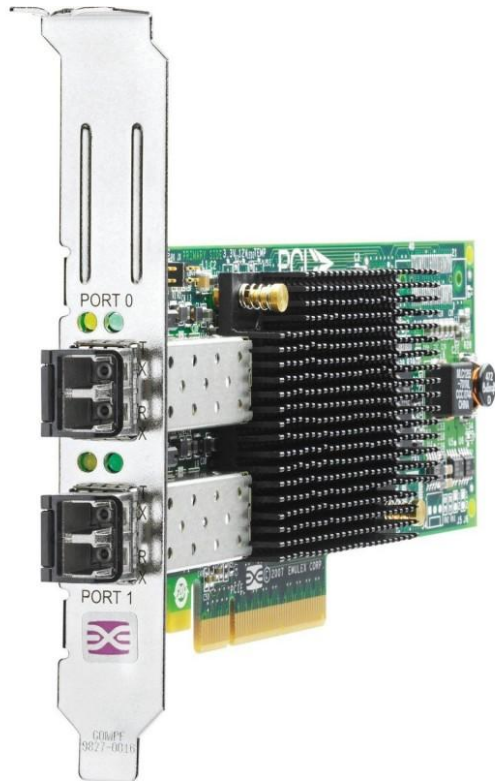


Примеры оборудования SAN (Fiber Channel)

- **Контроллер P9D94A HP StoreFabric SN1100Q 16Gb DP Fibre Channel**



Примеры оборудования SAN (Fiber Channel)



- ▣ **Контроллер 489193-001 HP 82E 8Gb 2-port PCIe Fibre Channel Host Bus Adapter**

Примеры оборудования SAN (Fiber Channel)

- Система хранения Dell PowerVault ME4024 24x2.5/No HDD, 4 x SFP+ FC16

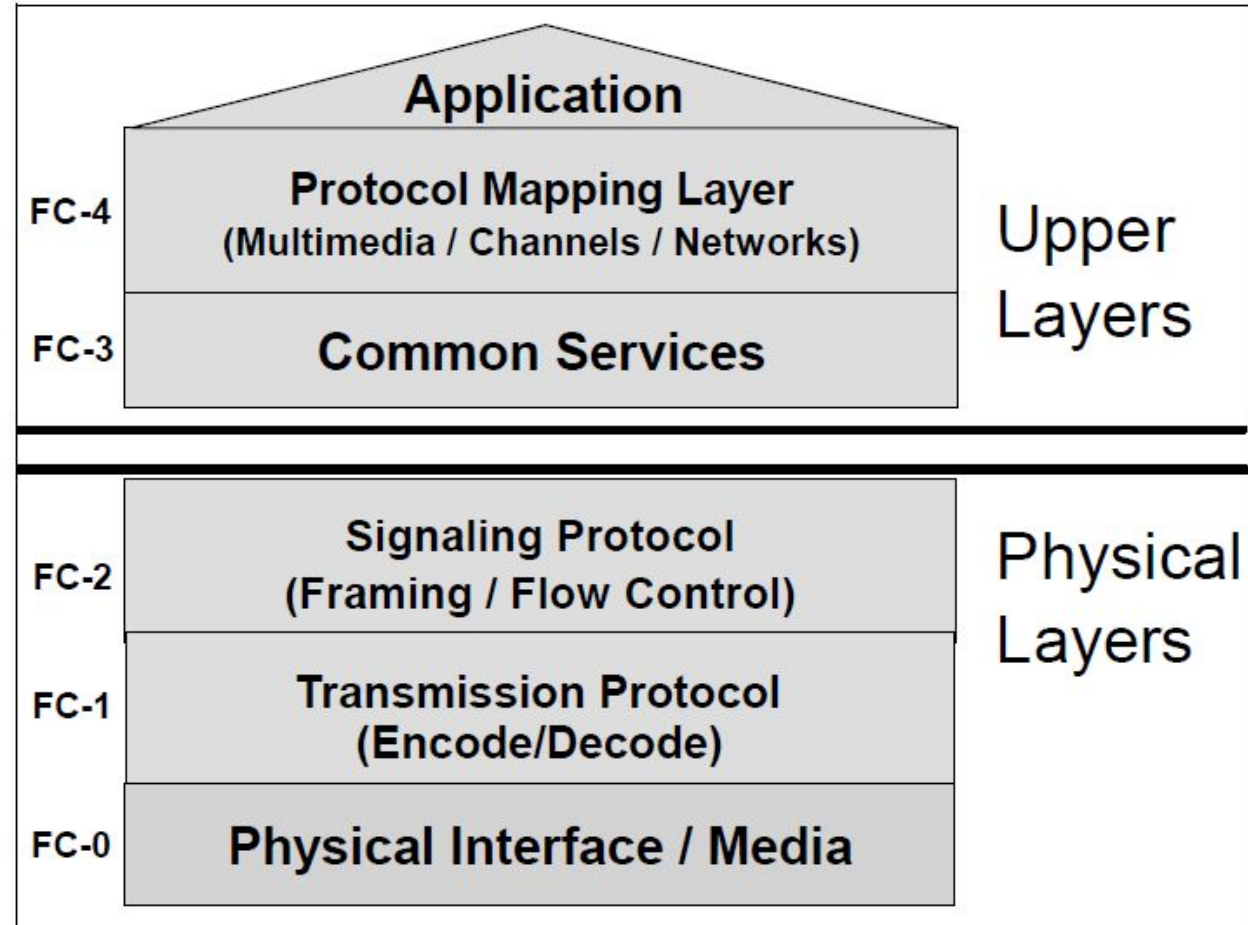


NAS (иногда)

Приводы	(96) Fibre Channel/FATA/SSD; Поддерживается в зависимости от модели и наличия отсека расширения
Функции обеспечения доступности	Полностью симметричный режим Active/Active
Размер файловой системы (макс.)	128 ТБ; Linux; 32 ТБ Windows
Количество файловых систем на кластер (макс.)	512; Linux; Windows 256
Размер кластера (макс.)	16 узлов
Поддерживаемые стоечные серверы	2 контроллера файловых серверов (стандартно)
Интерфейс массива	Fibre Channel
Поддерживаемые платформы BladeSystem	HP ProLiant BladeSystem
Сетевые протоколы	NFS, CIFS, HTTP, FTP, HTTPS, iSCSI
Поддержка RAID	0, 1, 5, 6, 0+1
Поддержка Fibre channel (SAN)	MSA, EVA, XP
Поддержка iSCSI	MSA, EVA
Гарантия (детали-качество сборки-поддержка на месте)	3/3/3



Fiber Channel



Особенности Fiber Channel

Интерфейс	Пропускная способность	Линейная скорость	Схема кодирования	Хост-адаптер
1Gb FC	100 MB/s	1.0625 GBaud	8b/10b	PCI-X
2Gb FC	200 MB/s	2.125 GBaud	8b/10b	PCI-X
4Gb FC	400 MB/s	4.25 GBaud	8b/10b	PCI-X 2.0 или PCIe 1.0 x4
8Gb FC	800 MB/s	8.5 GBaud	8b/10b	PCI-X 1.0 x8 или PCIe 2.0 x4
16Gb FC	1600 MB/s	14.025 GBaud	64b/66b	PCI-X 2.0 x8 или PCIe 3.0 x4
32Gb FC	3200 MB/s	28.05 GBaud	64b/66b	PCIe 3.0 x8
64Gb FC	6400 MB/s	28.9 GBaud	64b/66b	PCIe 4.0

Пропускная способность InfiniBand

	1X	4X	12X	Схема кодирования	Хост-адаптер
SDR	2 Gb/s	8 Gb/s	24 Gb/s	8b/10b	PCIe 1.0 x8
DDR	4 Gb/s	16 Gb/s	48 Gb/s	8b/10b	PCIe 1.0 x16 или PCIe 2.0 x8
QDR	8 Gb/s	32 Gb/s	96 Gb/s	8b/10b	PCIe 2.0 x8
FDR-10*ТОЛЬКО Mellanox	10.31 Gb/s	41.25 Gb/s	123.75 Gb/s	64b/66b	PCIe 3.0 x8
FDR	13.64 Gb/s	54.55 Gb/s	163.64 Gb/s	64b/66b	PCIe 3.0 x8
EDR	25 Gb/s	100 Gb/s	300 Gb/s	64b/66b	PCIe 3.0 x16

SAN

Достоинства

- Высокая надёжность доступа к данным, находящимся на внешних системах хранения. Независимость топологии SAN от используемых СХД и серверов.
- Централизованное хранение данных (надёжность, безопасность).
- Удобное централизованное управление коммутацией и данными.
- Перенос интенсивного трафика ввода-вывода в отдельную сеть – разгрузка LAN.
- Высокое быстродействие и низкая латентность.
- Масштабируемость и гибкость логической структуры SAN
- Географические размеры SAN, в отличие от классических DAS, практически не ограничены.
- Возможность оперативно распределять ресурсы между серверами.
- Возможность строить отказоустойчивые кластерные решения без дополнительных затрат на базе имеющейся SAN.
- Простая схема резервного копирования – все данные находятся в одном месте.
- Наличие дополнительных возможностей и сервисов (снапшоты, удаленная репликация).
- Высокая степень безопасности SAN.

Недостатки

- Все минусы сводятся только к высокой стоимости подобного рода решений.

Задача

- CIFS
 - SMB1
 - SMB2
 - SMB3
- NFS
- FTP и TFTP
- Rsync
- WebDAV
- iSCSI

- 1) Согласовать набор данных
 - 1) 1 Гб не сжатых текстовых данных
 - 2) 1 Гб сжатых бинарных данных (один файл)
 - 3) 1 Гб сжатых бинарных данных (100 файлов)
- 2) Установить сервисы в тестовую среду
- 3) Провести сравнения скоростей передачи разных данных для разных файловых сервисов
- 4) Подготовить доклад о архитектуре, истории, настройке сервиса и результатах исследований.