

Распределенные системы хранения и обработки данных

Владислав Белогрудов, EMC

vlad.belogrudov@gmail.com

Лекция 4

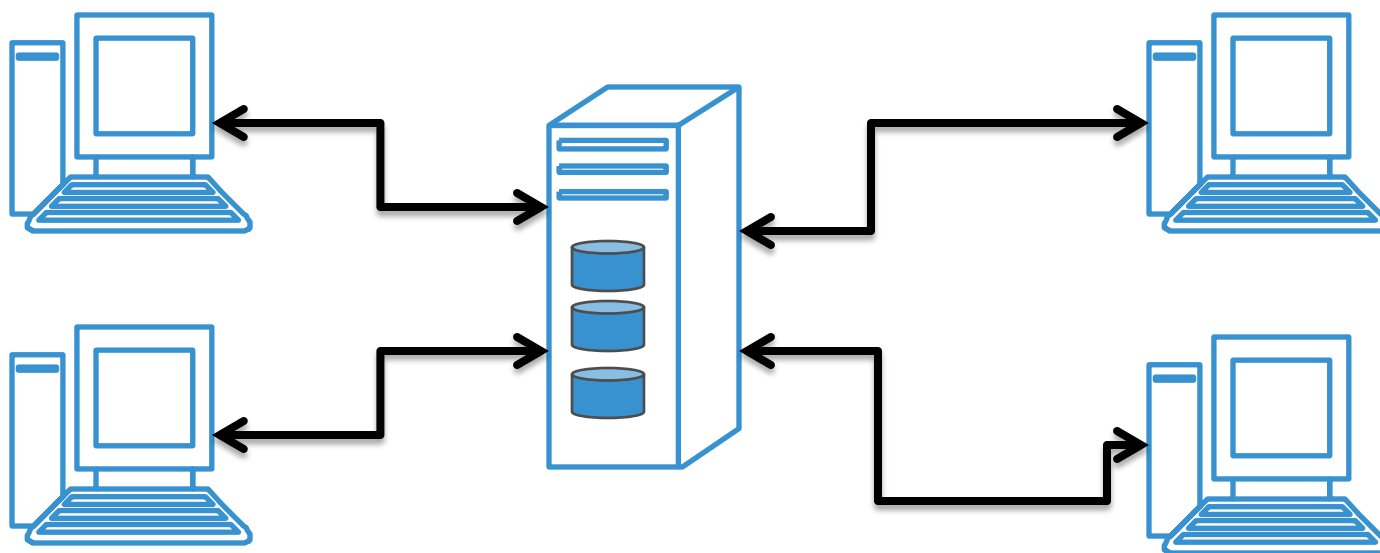
Сети хранения данных FC SAN

Содержание лекции

- DAS
- SCSI
- SAN
 - компоненты
 - архитектура
 - топологии
 - зонирование

DAS – система хранения прямого подключения

- Внешние и внутренние
- Внешние – SCSI или FC



Преимущества DAS

- Низкий порог «входа»
- Простота и быстрота развертывания
- Меньше знаний и времени для администрирования
- Нет задержек сетей, лучше производительность
- Может быть как простым диском так и интеллектуальной СХД

Недостатки DAS

- Плохо масштабируется (диски, хосты)
- Ограничение производительности
- Ограничение расстояния до хостов
- Утилизация ресурсов не оптимальна



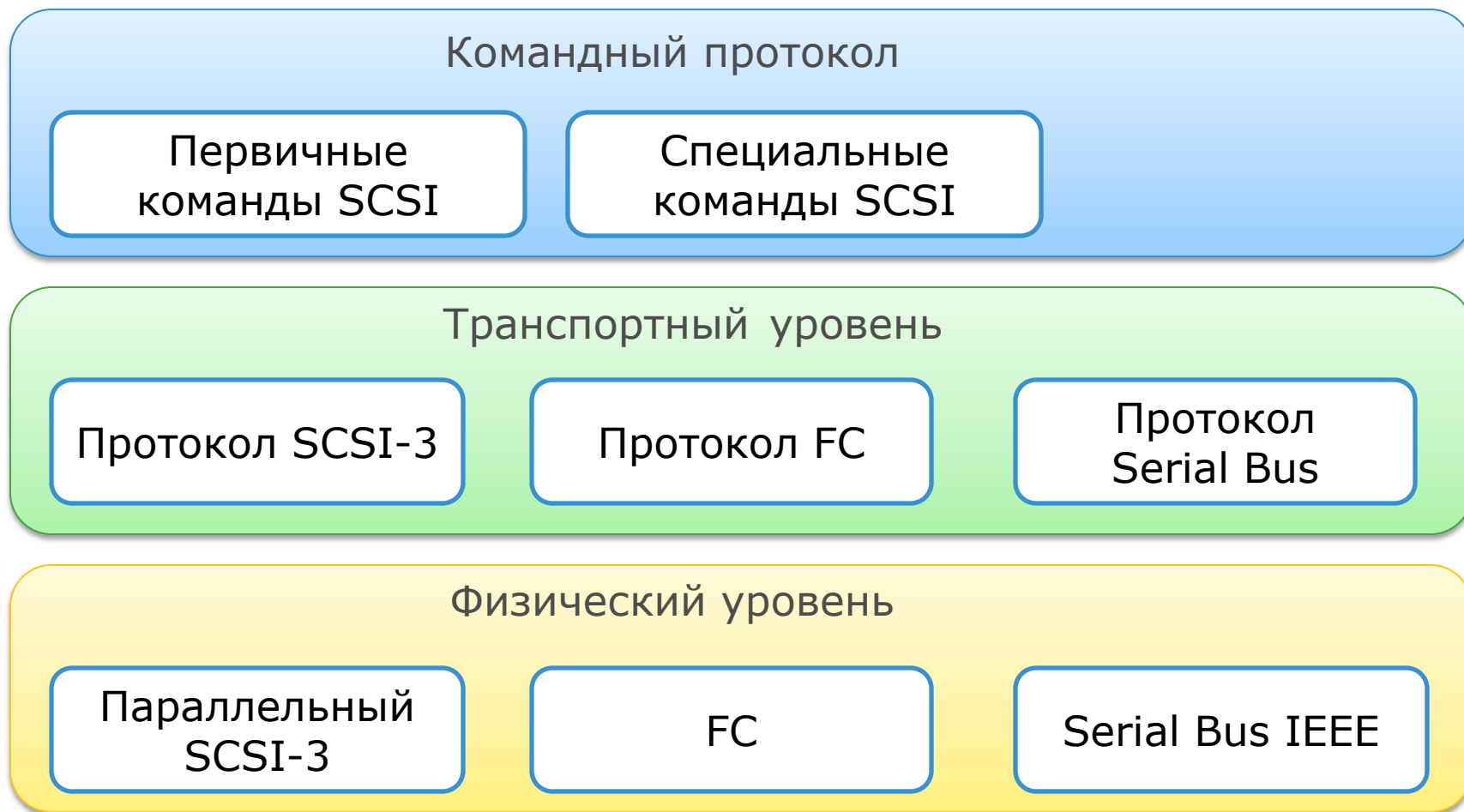
SCSI

- 1981 – SASI (Shugart Associates System Interface)
- 1986 – ANSI SCSI
- Аппаратная независимость (диски, сканеры, ленты..)
- Поддержка нескольких устройств (до 16)
- Последний стандарт – SCSI-3

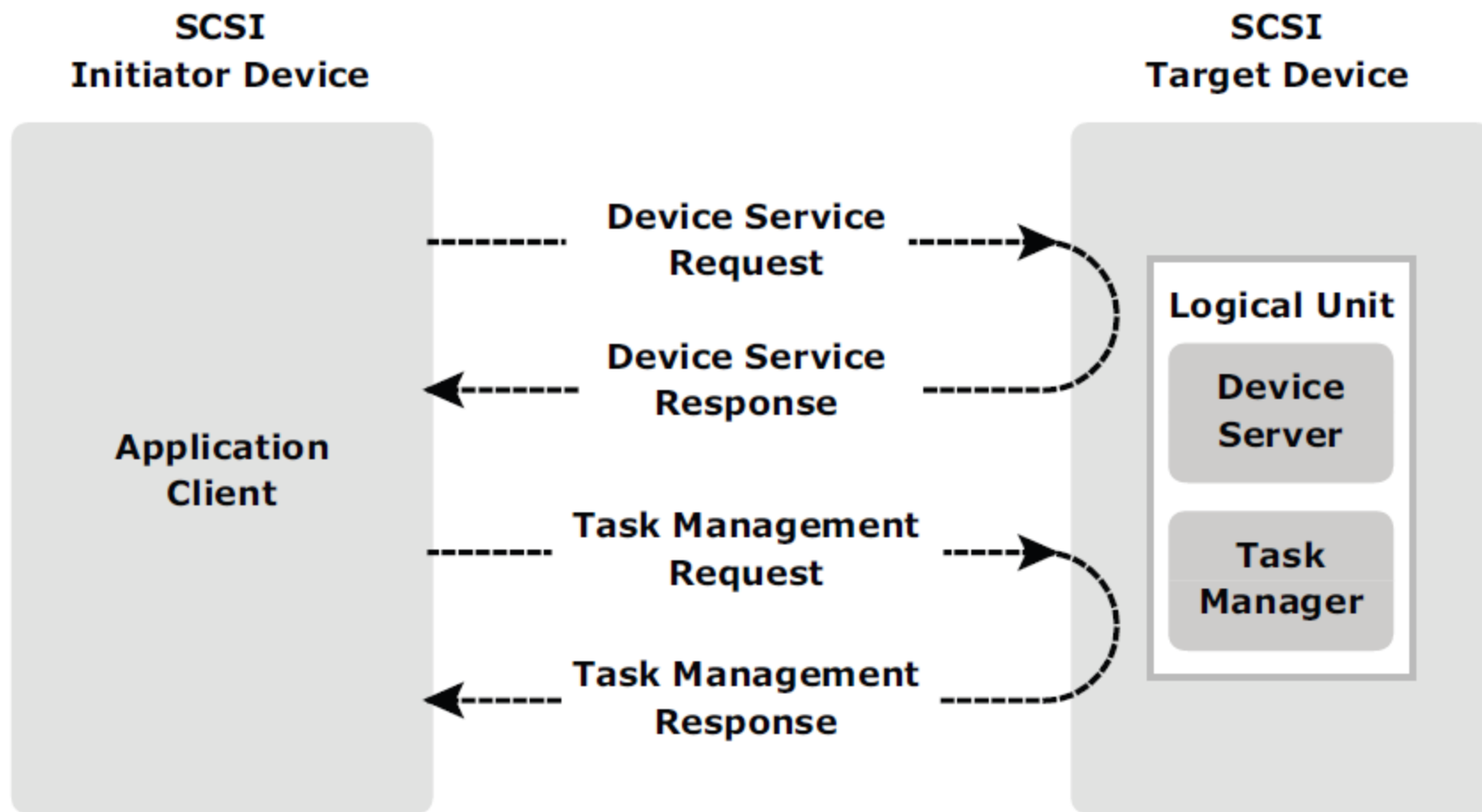
SCSI-3

1. Командный протокол: общие команды для устройств, специальные команды
2. Протоколы транспортного уровня: правила передачи данных
3. Физический интерфейс: электрические сигналы, режимы

Архитектурная модель SCSI



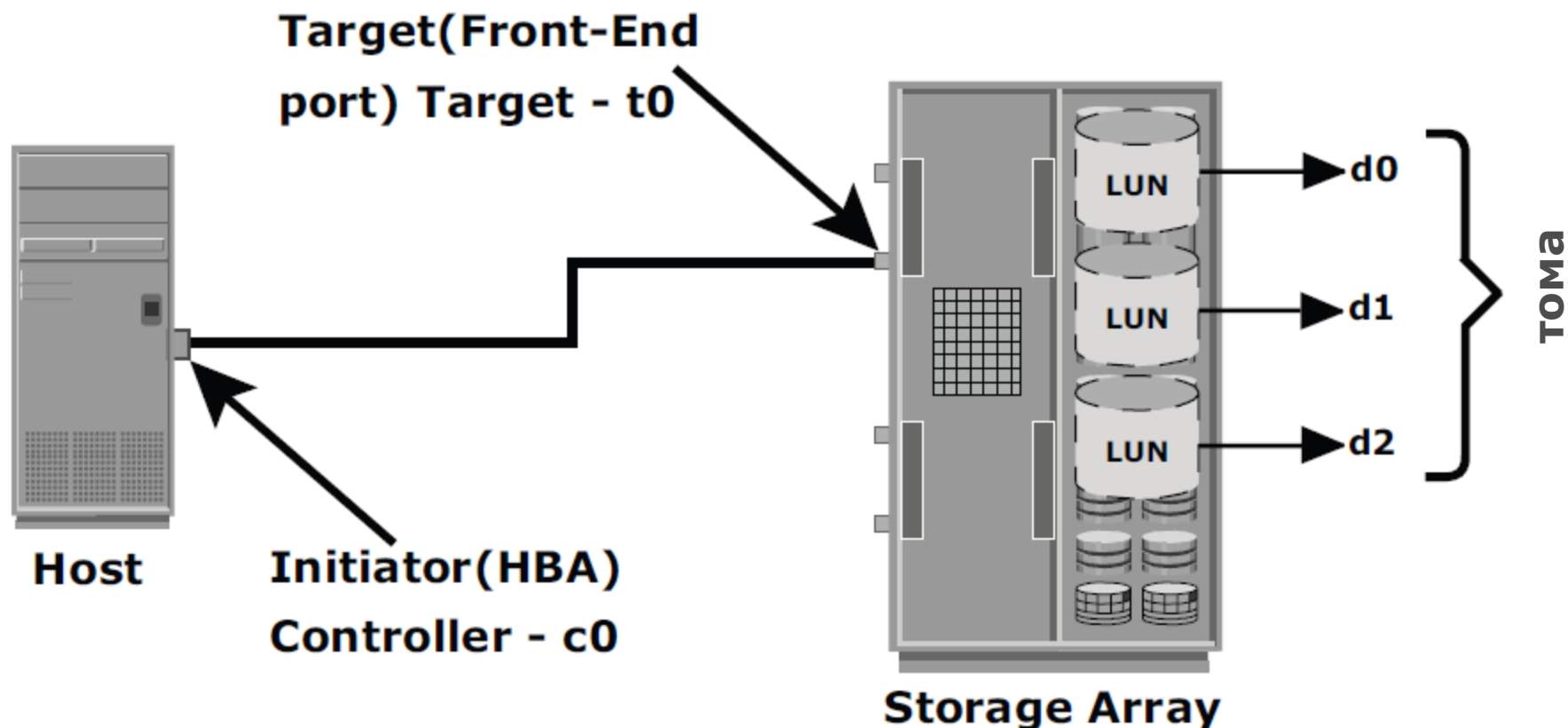
Модель клиент-сервер



Идентификаторы устройств SCSI

- 0..15
- Назначается автоматически
- Приоритет по возрастанию
 - 8, 9, 10, 11, 12, 13, 14, 15,
0, 1, 2, 3, 4, 5, 6, 7

Адресация SCSI интерфейса



Диск 1 = c0 t0 d1

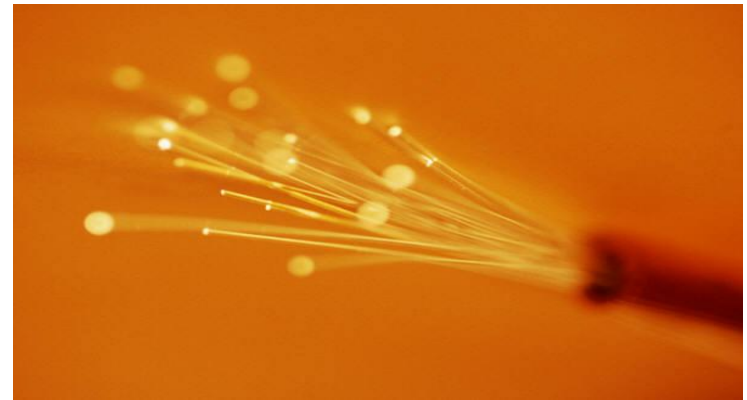
Сеть хранения данных (SAN)

- SAN – высокоскоростная выделенная сеть серверов и СХД
- Как правило на основе FC (в будущем - FCoE)
- Решаемые задачи:
 - доступность
 - безопасность
 - масштабируемость
 - гибкость
 - консолидация
 - управляемость ..



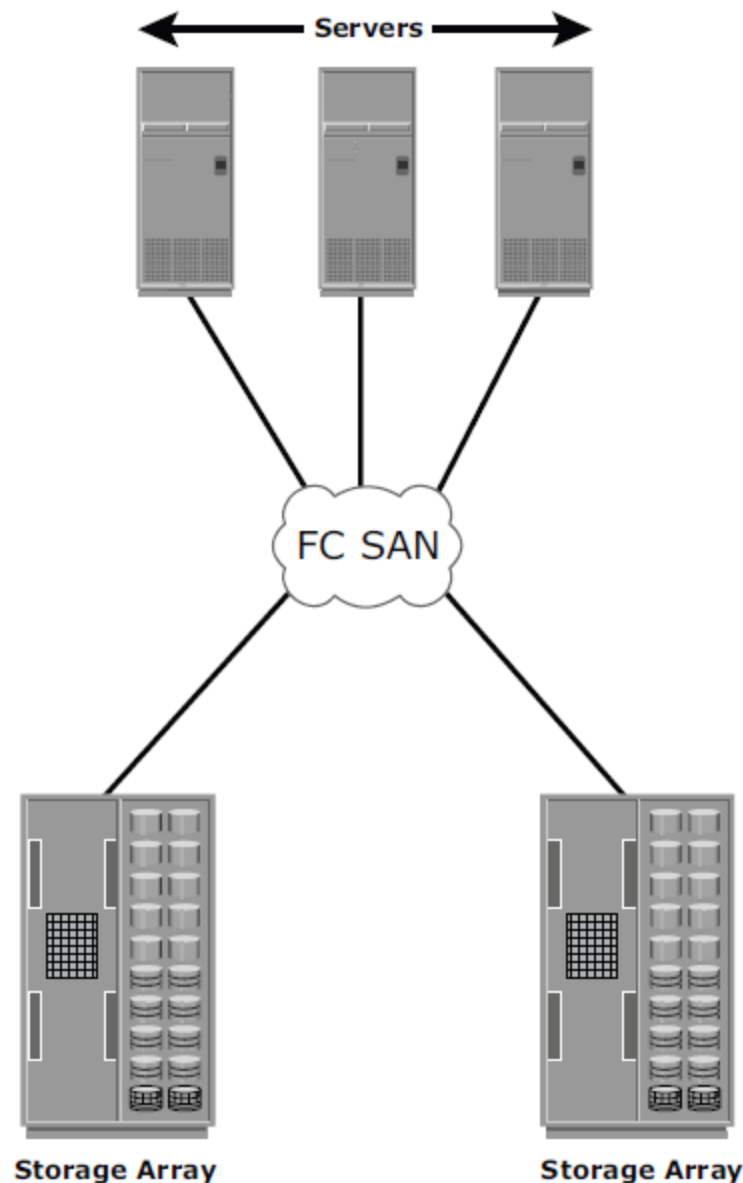
Fibre Channel

- Появление - 1988, стандарт – 1994
- Fibre – технология, Fiber – оптика
- Оптоволокно для frontend-а, медь для дисковых массивов
- 1600 МБ/сек (интерфейс SCSI дисков – 320 МБ/с)
- До 15 миллионов узлов

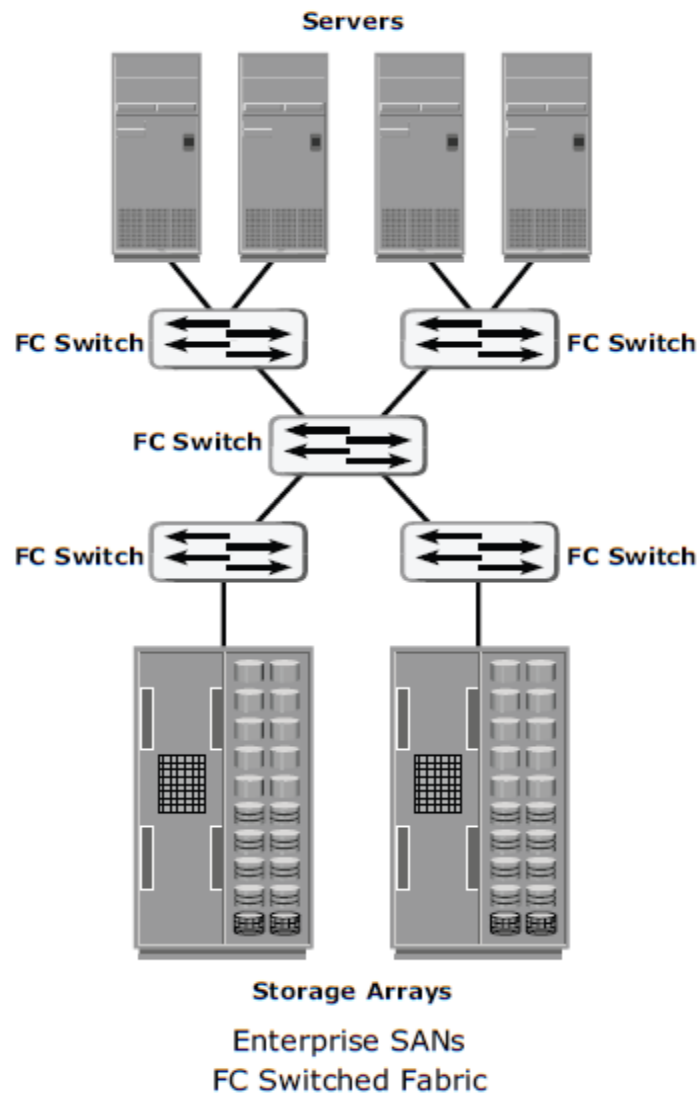
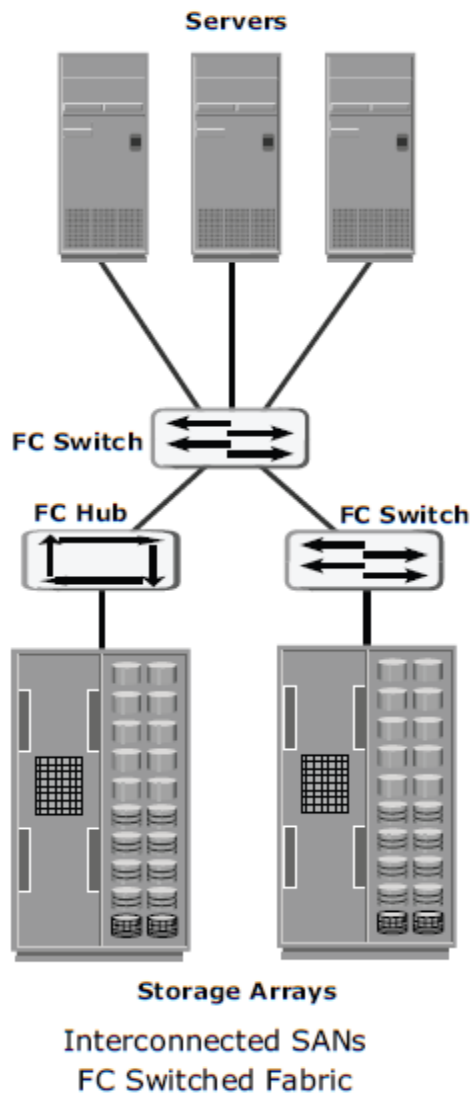
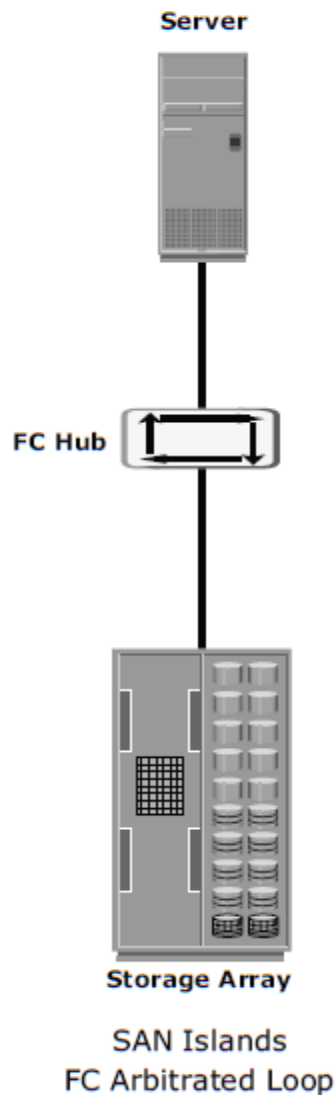


Реализация SAN

- Коммутаторы:
 - FC-AL hub
 - FC switch
 - FC director
- Удаленные участки:
 - iSCSI
 - FCIP

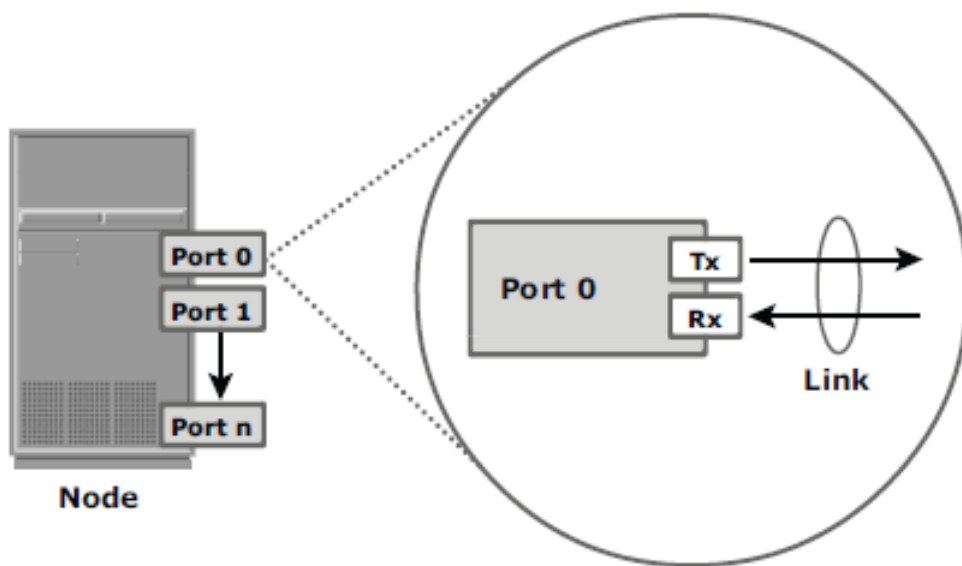


Эволюция FC SAN



Порты узлов

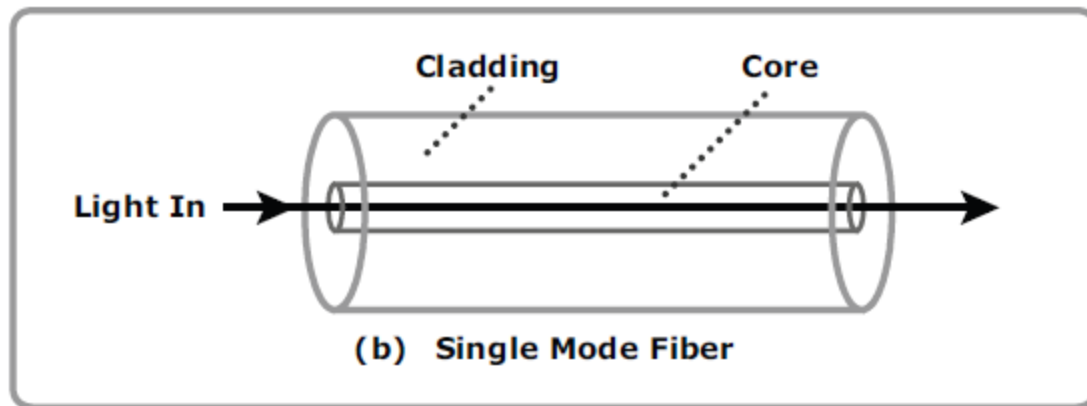
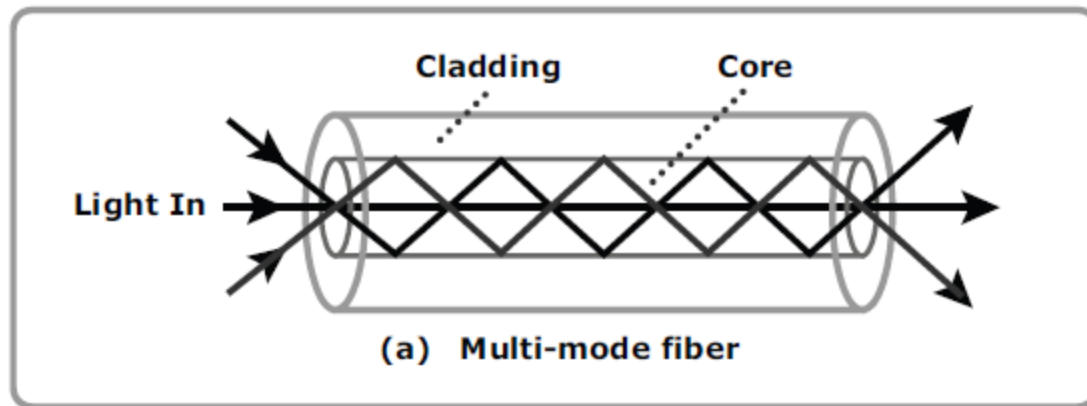
- Узлы – все конечные устройства сети (хосты, СХД)
- Порт – 2 соединения (передача/прием)



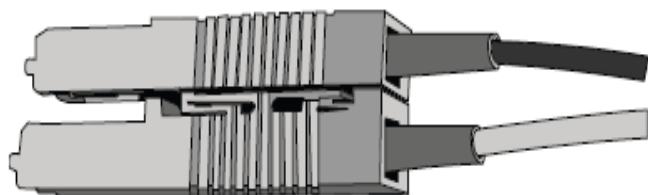
Кабели

- Один луч – Single Mode Fiber
 - дорого
 - надежно
 - большие расстояния (километры)
- Много лучей – Multi-mode Fiber
 - рассеивание
 - малые расстояния (500 метров)
 - дешево

MFM и SMF



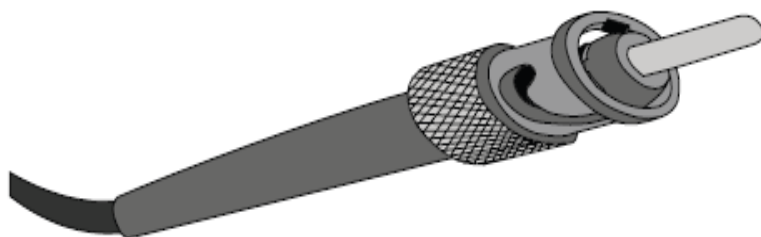
Разъемы



(a) Standard Connector



(b) Lucent connector

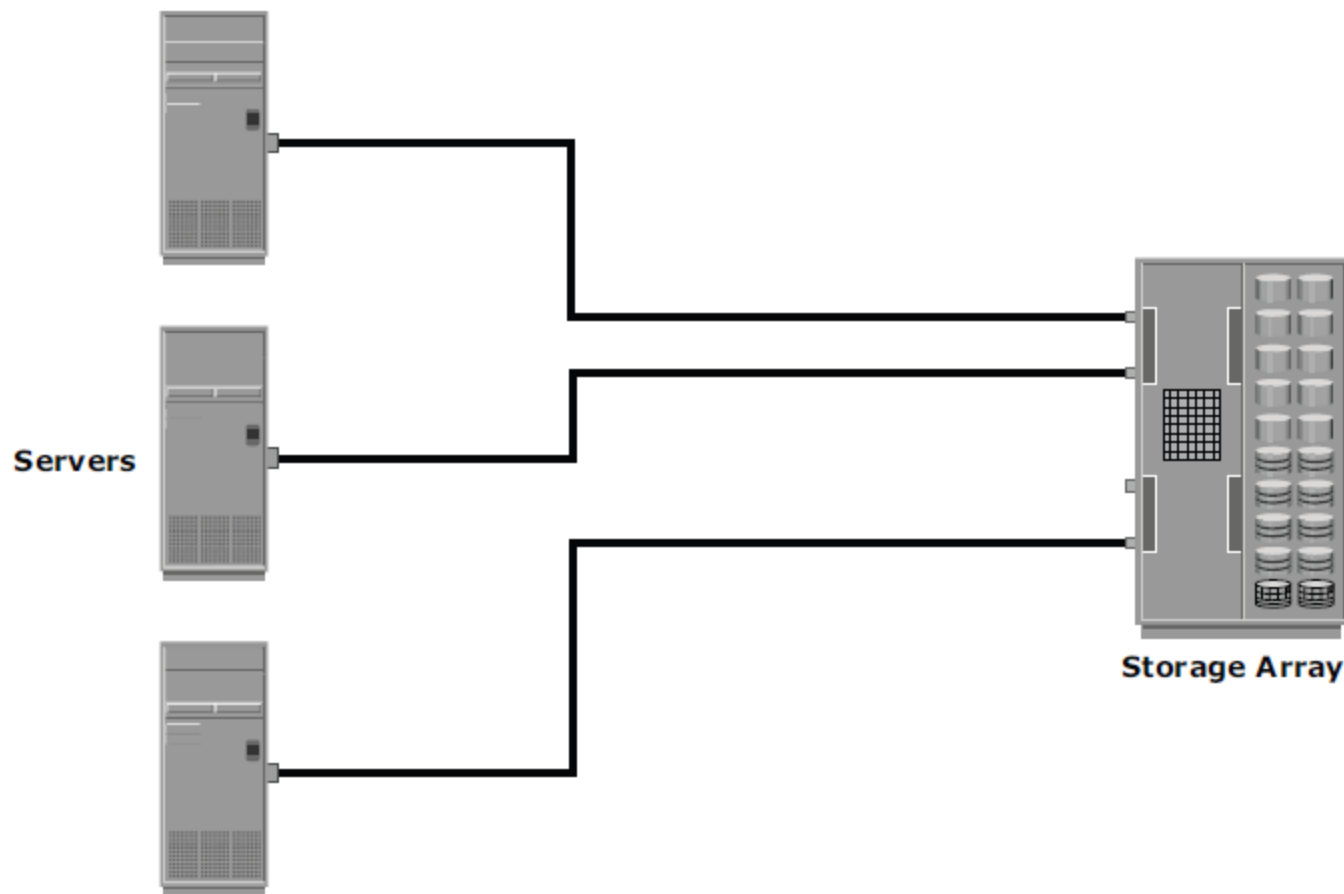


(c) Straight Tip Connector

Коммутационные устройства

- Hub (концентратор)
 - FC-AL
 - топология: физическая звезда, логическая петля (круг)
 - данные передаются через все точки
- Switch (коммутатор)
 - передача данных только между нужными узлами
 - одновременный обмен
 - выделенный канал связи
- Director (модульный коммутатор)

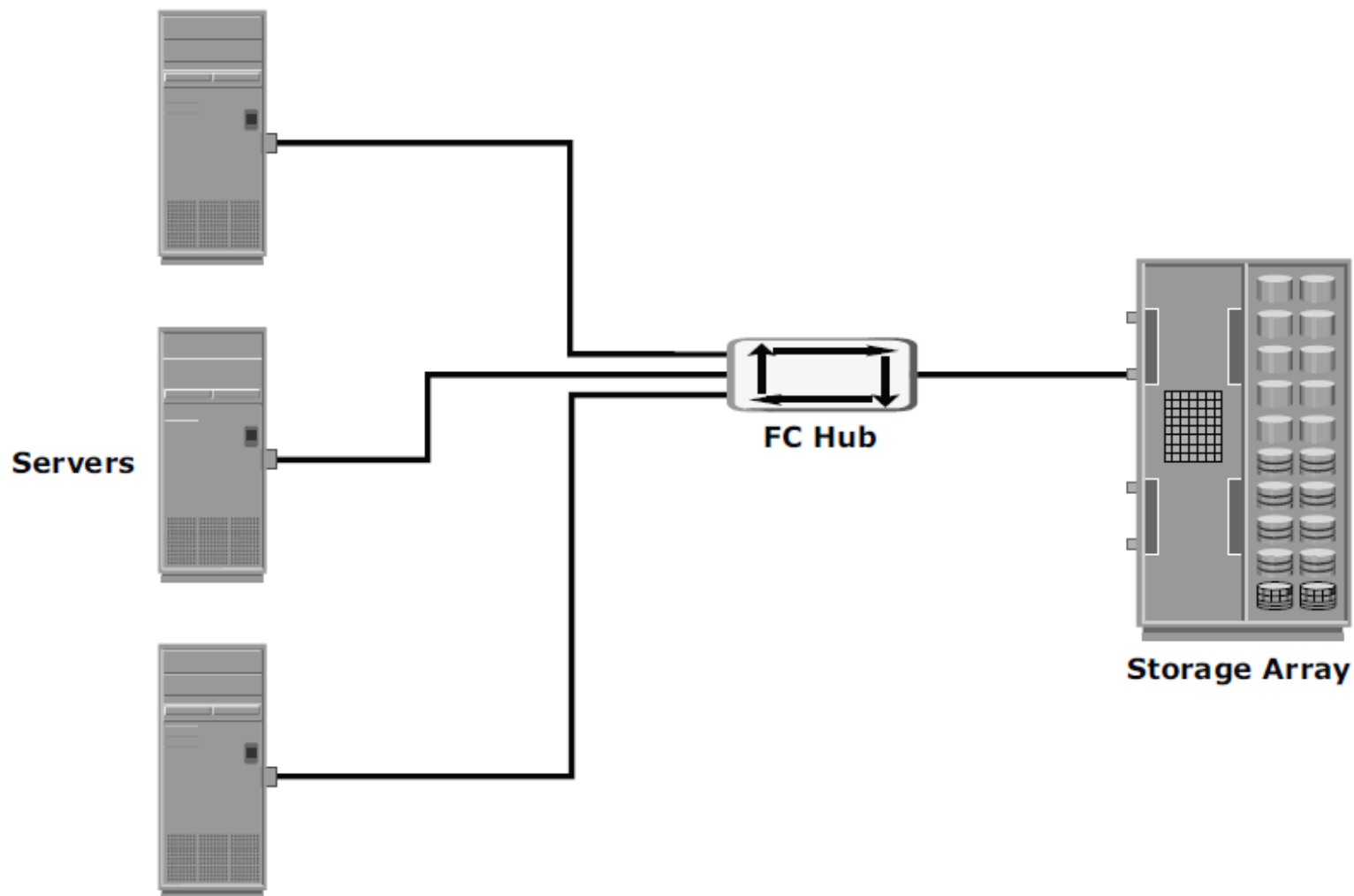
Подключение FC – точка-точка



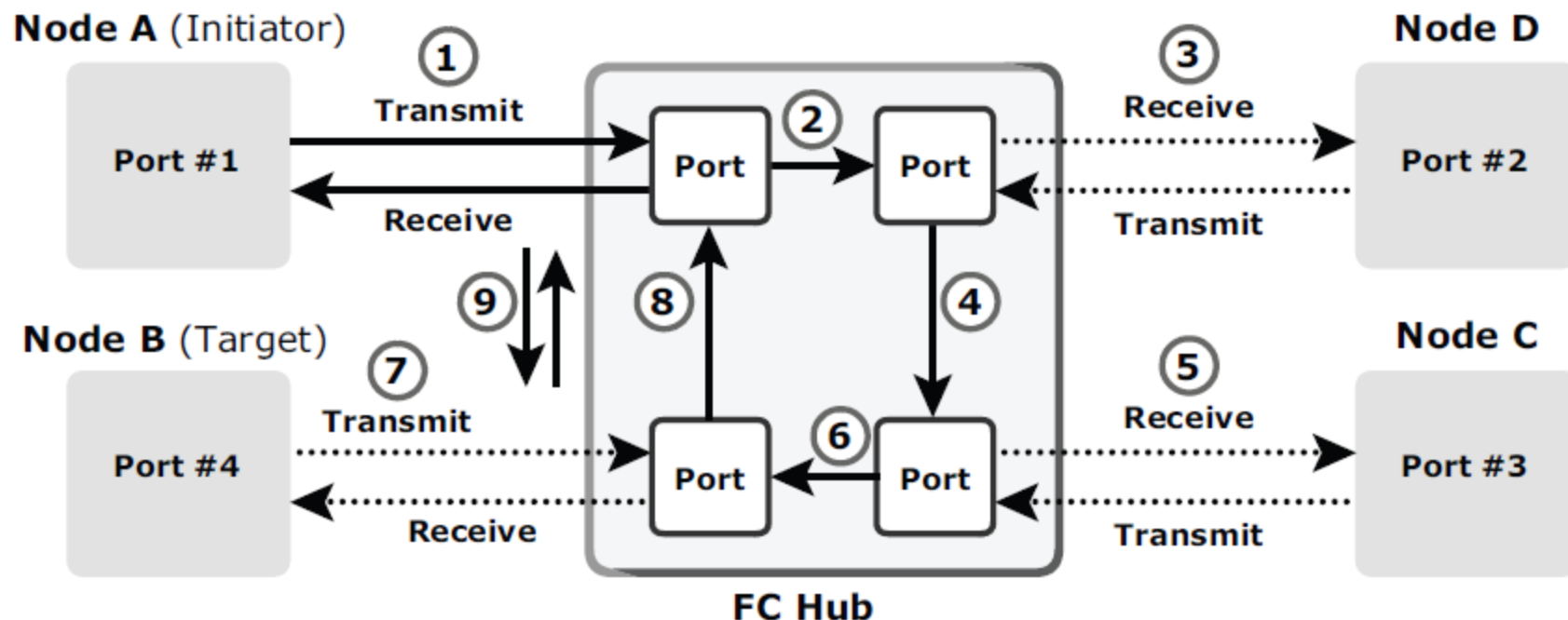
FC-AL (Arbitrated Loop)

- Кольцевая топология (физически - звезда)
- До 126 узлов
- Эстафетный доступ
- Каждый момент времени – только одна передача данных
- Добавление или удаление узла из кольца (петли) – реинициализация и задержки

Подключение FC-AI



Передача данных в FC-AE



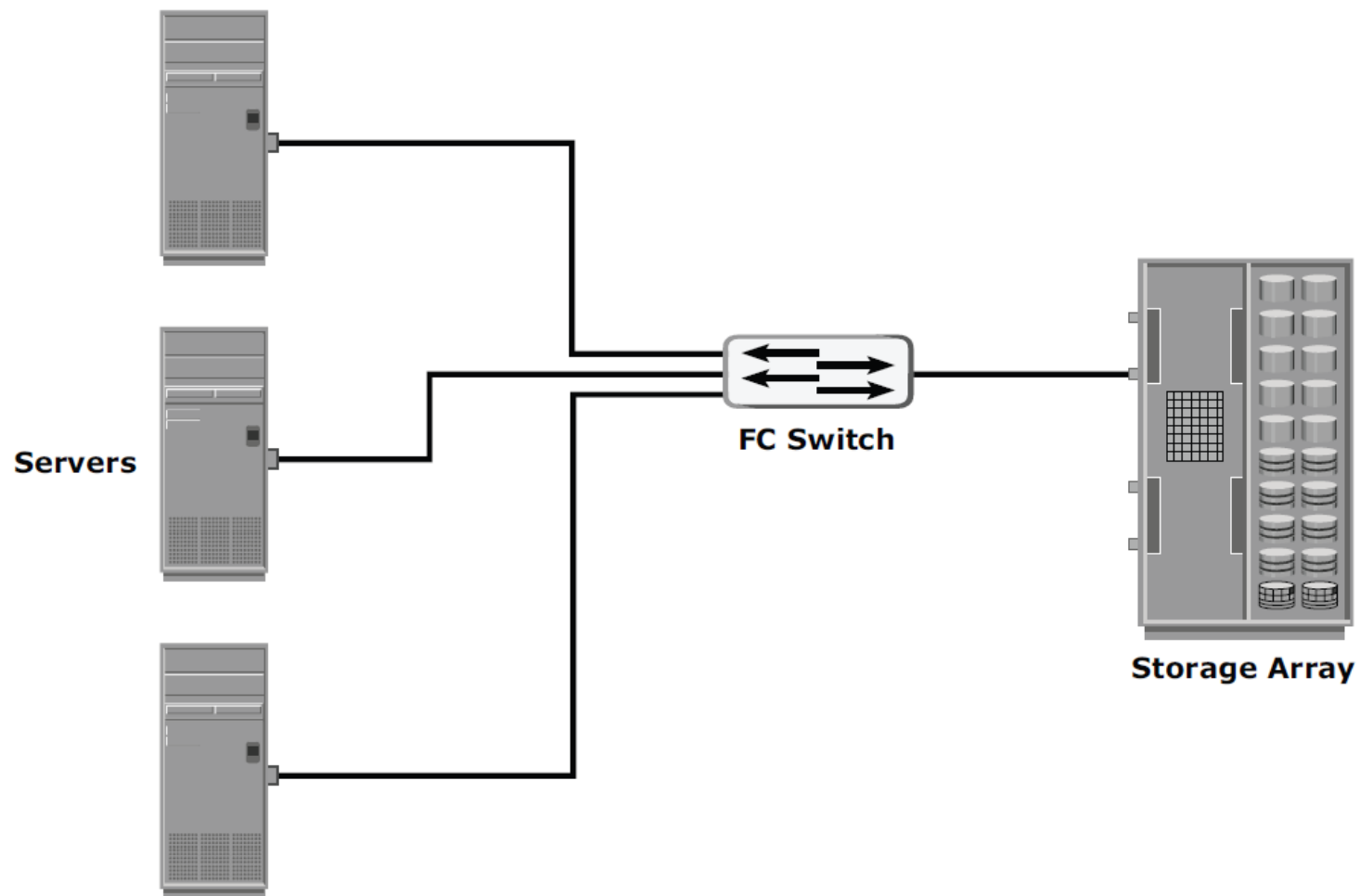
Узел А договаривается со всеми о передаче данных узлу В

FC-SW (FC Switched Fabric)

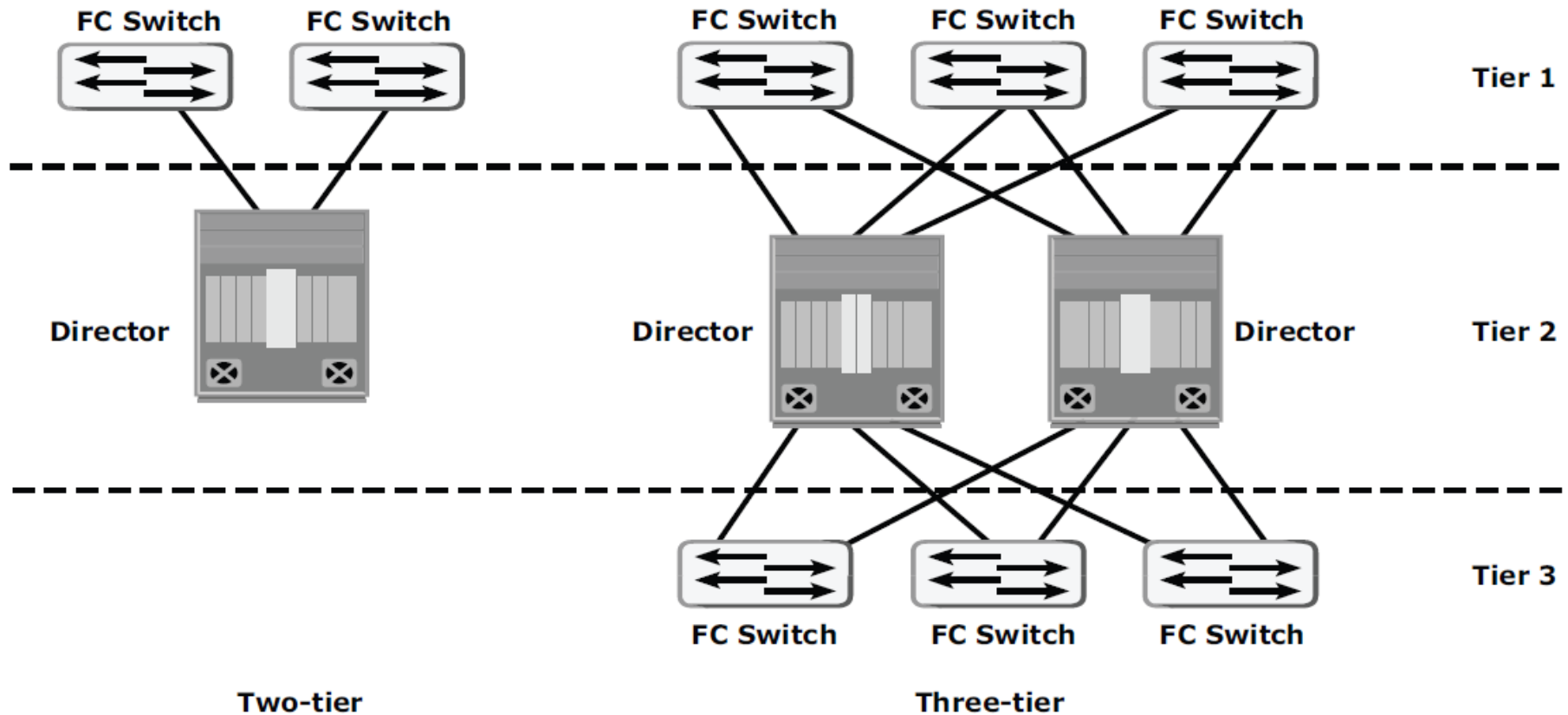
- Коммутируемая сеть
- Масштабируемость
- Выделенный канал каждому
- До 15 миллионов узлов (24-битный адрес узлов)
- Каждый коммутатор в «фабрике» имеет свой «домен»



Подключение FC-SW



Многоуровневые FC-SW сети



Передача данных в FC-SW

Node A

Node D

Port #1

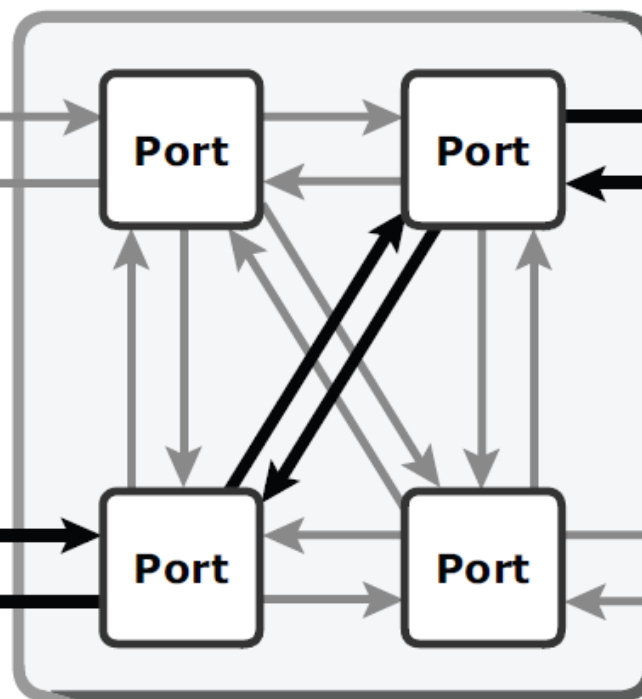
Port #2

Node B

Node C

Port #4

Port #3



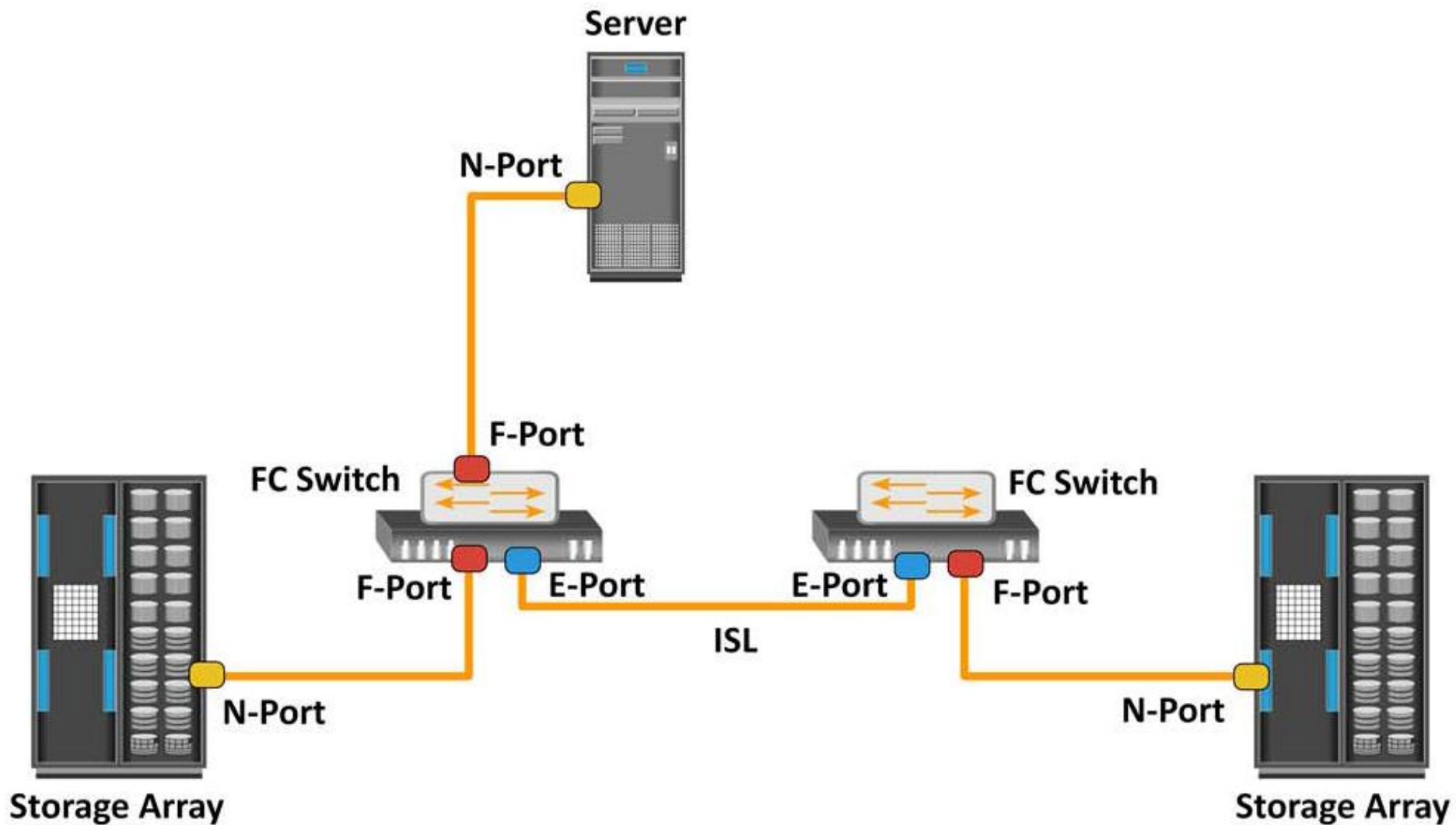
FC Switch

Хост В обменивается данными с СХД D

FC порты

- N_Port – узловой порт хоста или СХД
- F_Port – порт фабрики, идет к N_Port
- E_Port – порт расширения, между коммутаторами. Соединение между коммутаторами – ISL (Inter Switch Link)
- NL_Port, FL_Port – порты FC-AL
- G_Port – общий, может быть N_ и F_ портом

Порты в FC-SW



Регистрация устройств в FC

- FLOGI

- вход в фабрику
- между N_Port и F_Port
- узел посылает свой WWN с FC адресом FFFFFFFE (broadcast), коммутатор назначает ему уникальный FC адрес
- узел регистрируется в сервере имен WWN<->FC

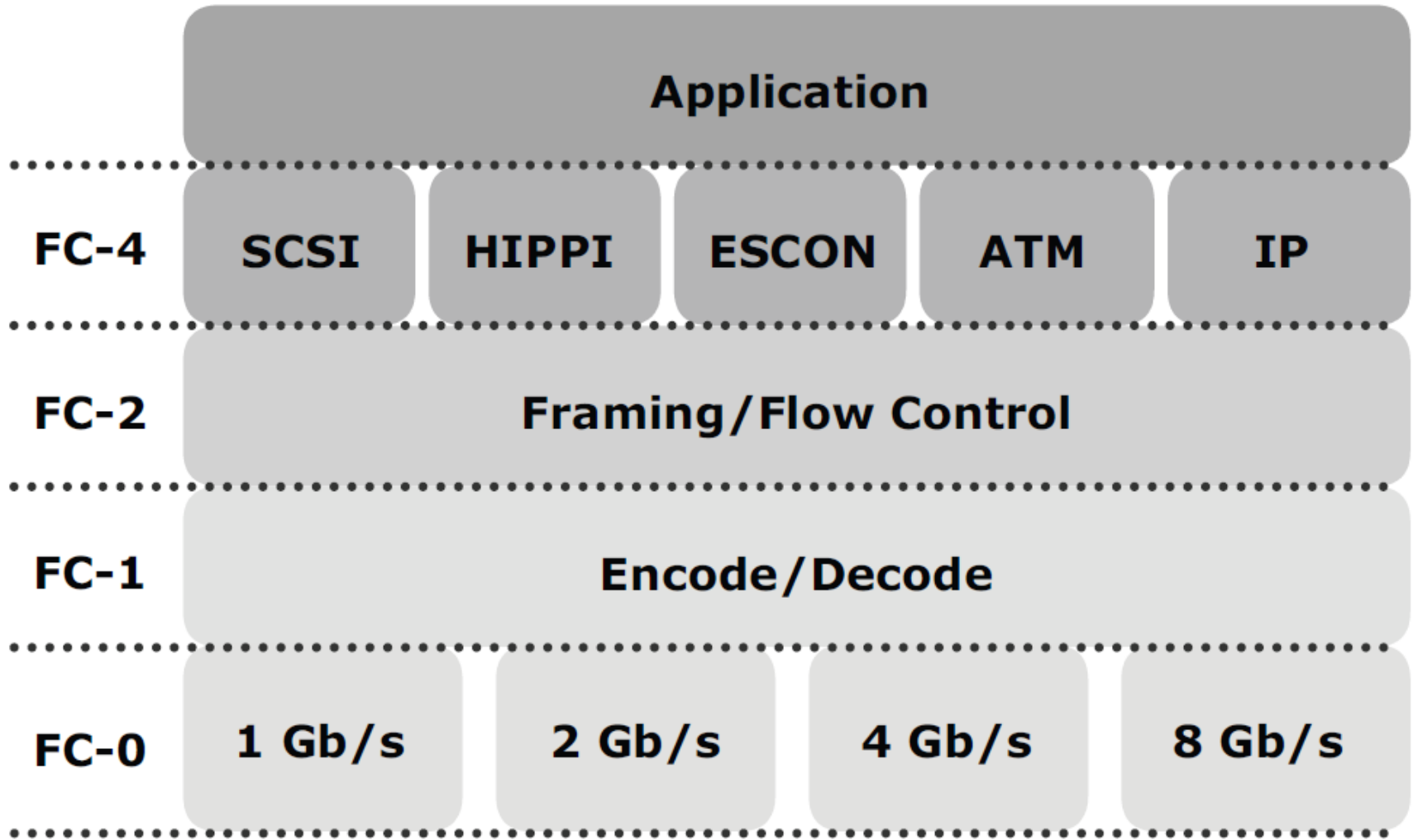
- PLOGI

- узел «знакомится» с другим узлом
- установка сессии

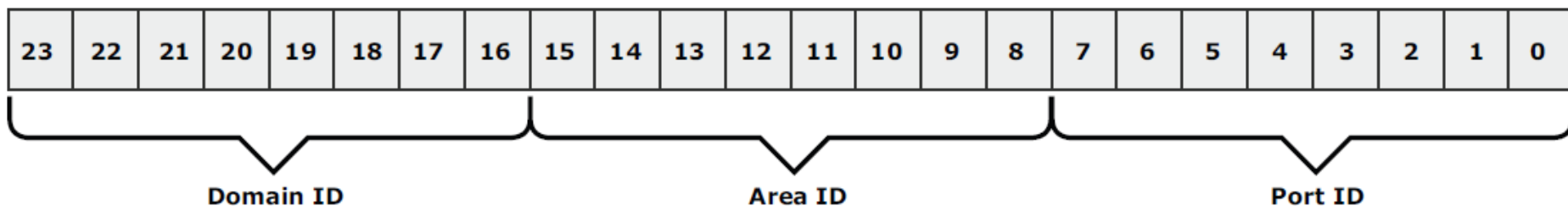
- PRLI

- регистрация по протоколам верхнего уровня

Пакет протоколов FC (FCP)



Адресация FC



- FFFFFFFC – сервер имен
- FFFFFFFE - broadcast

Глобальные имена устройств в FC

- WWN (World Wide Name)
 - 64 бита
 - имя узла WWNN или имя порта (WWNP)
 - аналог MAC в Ethernet, «прожигаются»

World Wide Name - Array															
5	0	0	6	0	1	6	0	0	0	6	0	0	1	B	2
0101	0000	0000	0110	0000	0001	0110	0000	0000	0000	0110	0000	0000	0001	1011	0010
Company ID 24 bits							Port	Model Seed 32 bits							

World Wide Name - HBA															
1	0	0	0	0	0	0	0	c	9	2	0	d	c	4	0
Reserved 12 bits				Company ID 24 bits						Company Specific 24 bits					

Управление потоком данных в FC

- BB_Credit

- контроль максимального числа кадров в канале
- приемная сторона сообщает о количестве свободных буферов при «знакомстве»
- отправляющая сторона считает свободное количество буферов (счетчик + ACKs приемника)

- EE_Credit

- аналогично BB_Credit, для сквозной передачи (инициатор-целевое устройство)

Классы сервиса

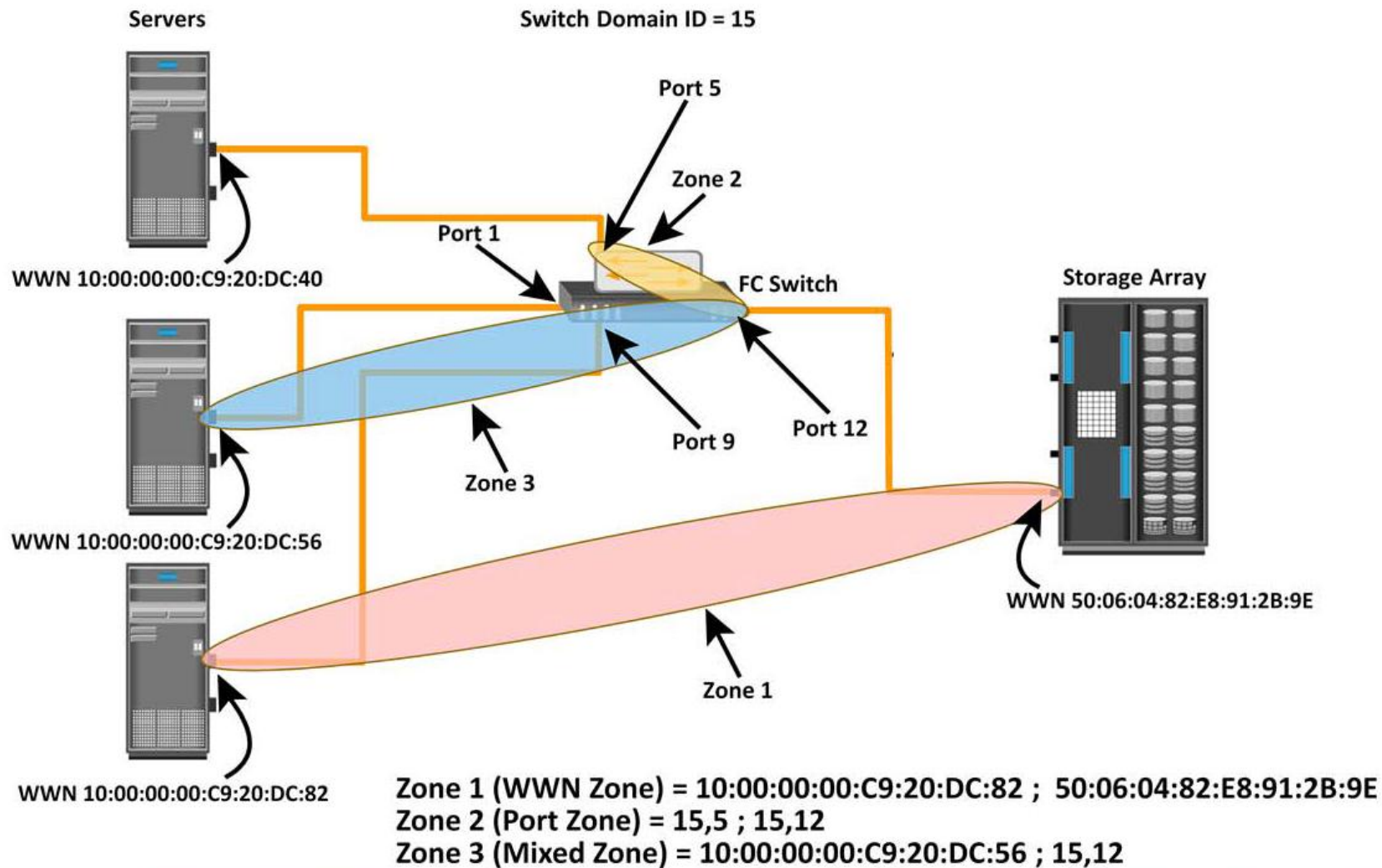
	Класс 1	Класс 2	Класс 3
Тип коммуникации	выделенное соединение	невыделенное соединение	невыделенное соединение
Управление потоками	EE_Credit	EE_Credit BB_Credit	BB_Credit
Очередность кадра	по очереди	не гарантирована	не гарантирована
Подтверждение	да	да	нет
Коэффициент загрузки	низкий	средний	высокий

Зонирование

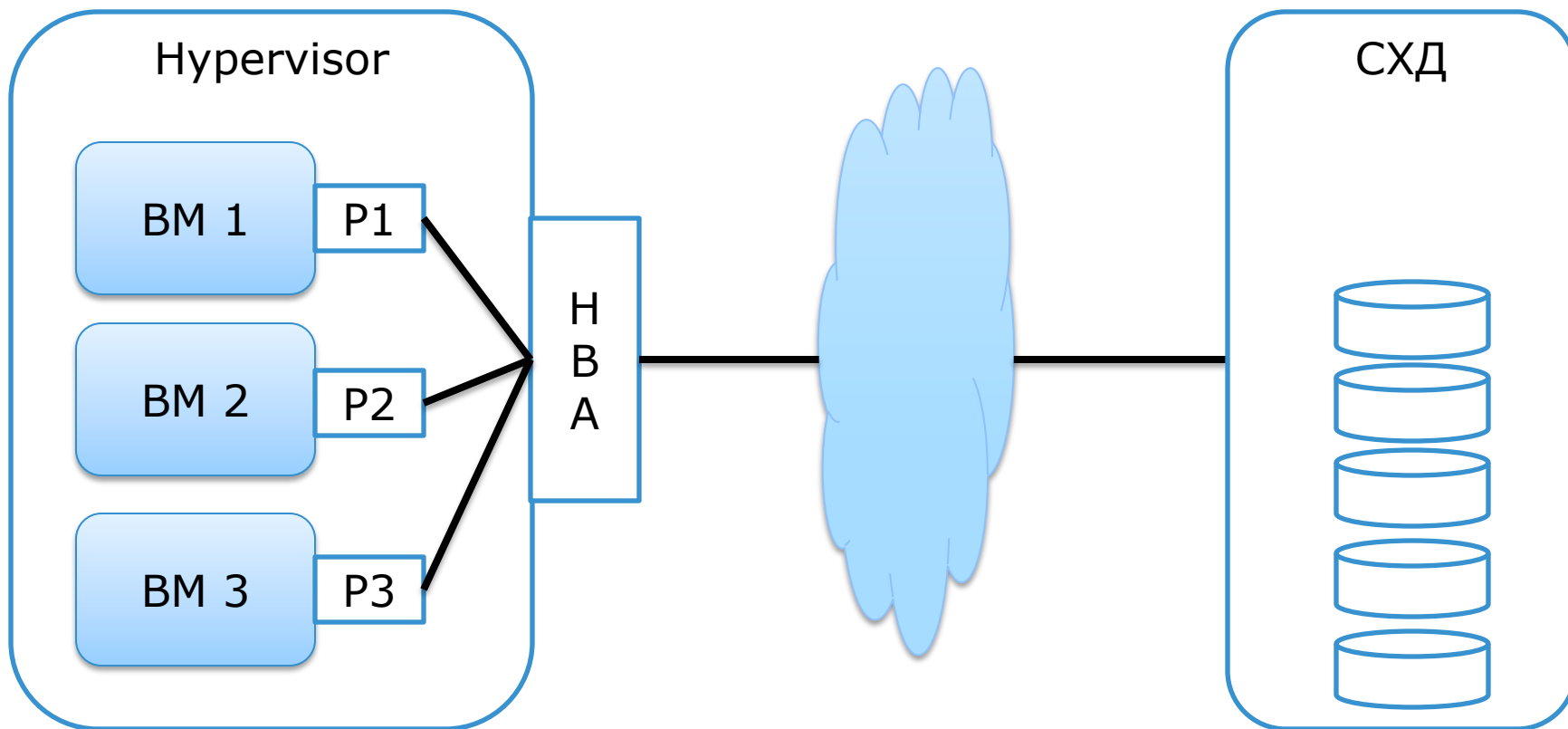
- Объединение узлов в группы на уровне сети
- Типы
 - по WWN
 - по FC портам
- Узел или порт может входить в несколько зон



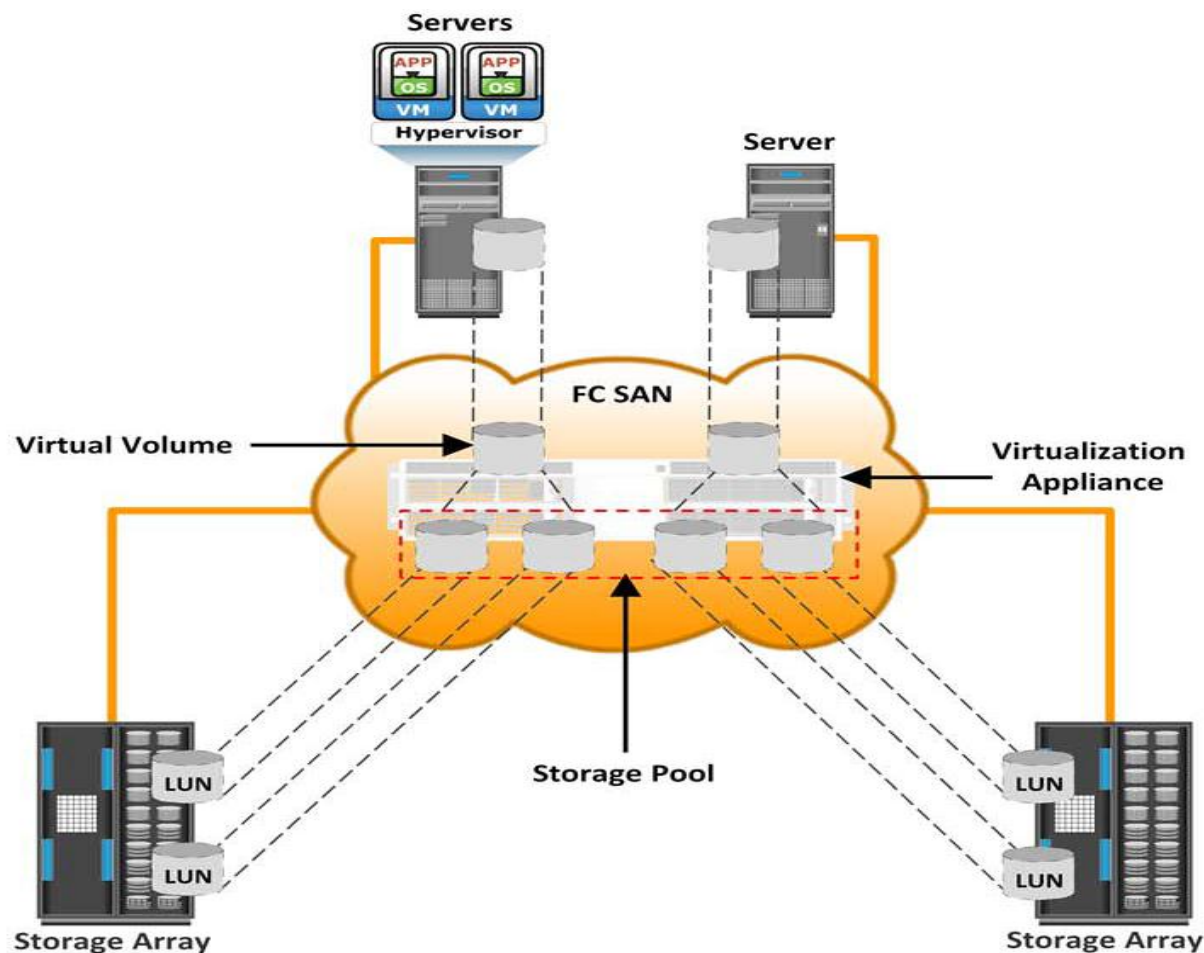
Типы зонирования



N_Port ID Virtualization

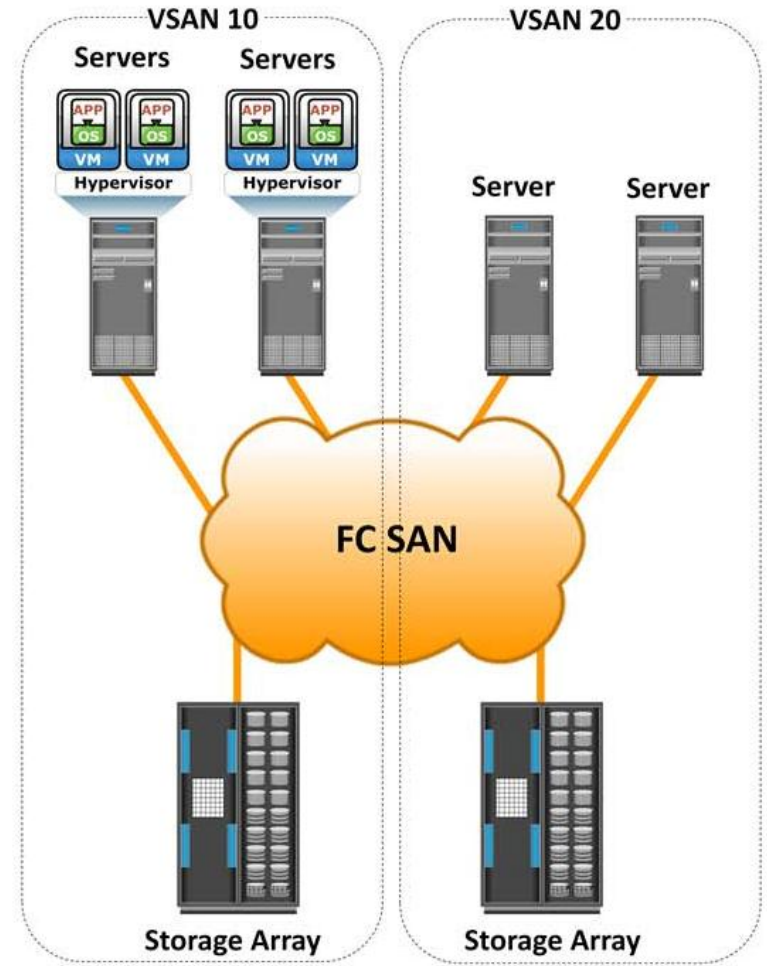


Виртуализация блочных устройств



Виртуальные SAN (VSAN)

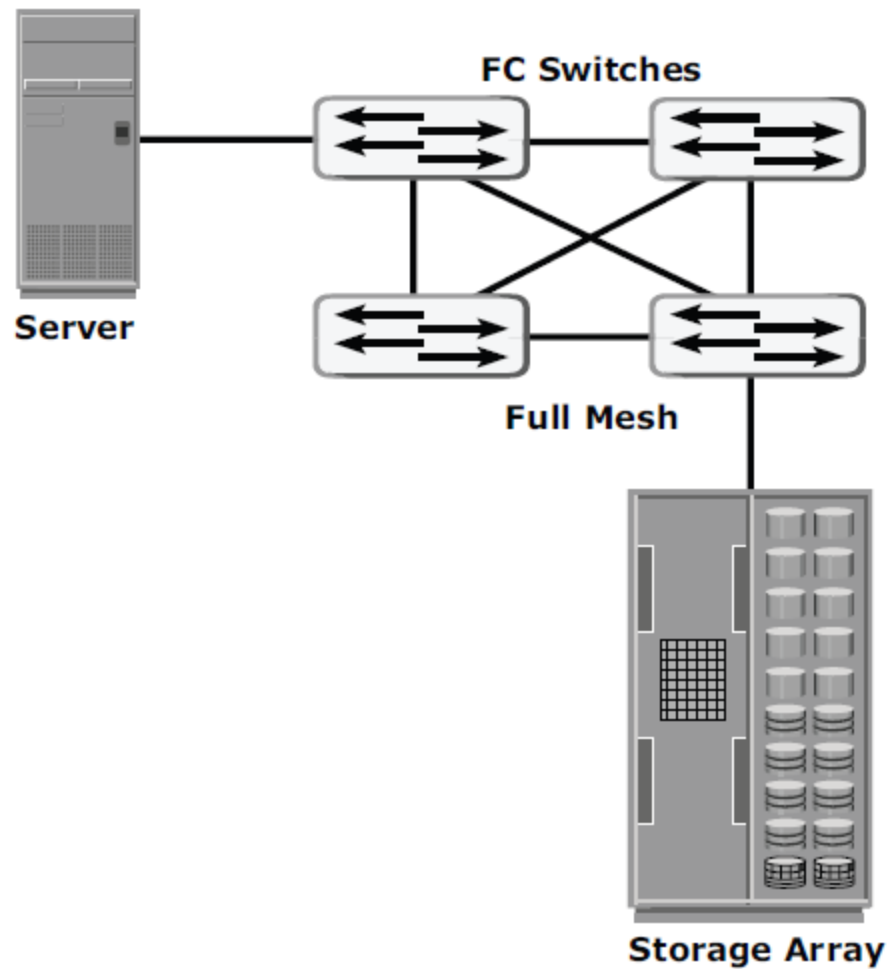
- VSAN объединяет группы узлов независимо от расположения
- Каждый VSAN имеет собственные зоны, FC адреса, сервер имен и т.п.
- Повышение безопасности, масштабируемости, управляемости



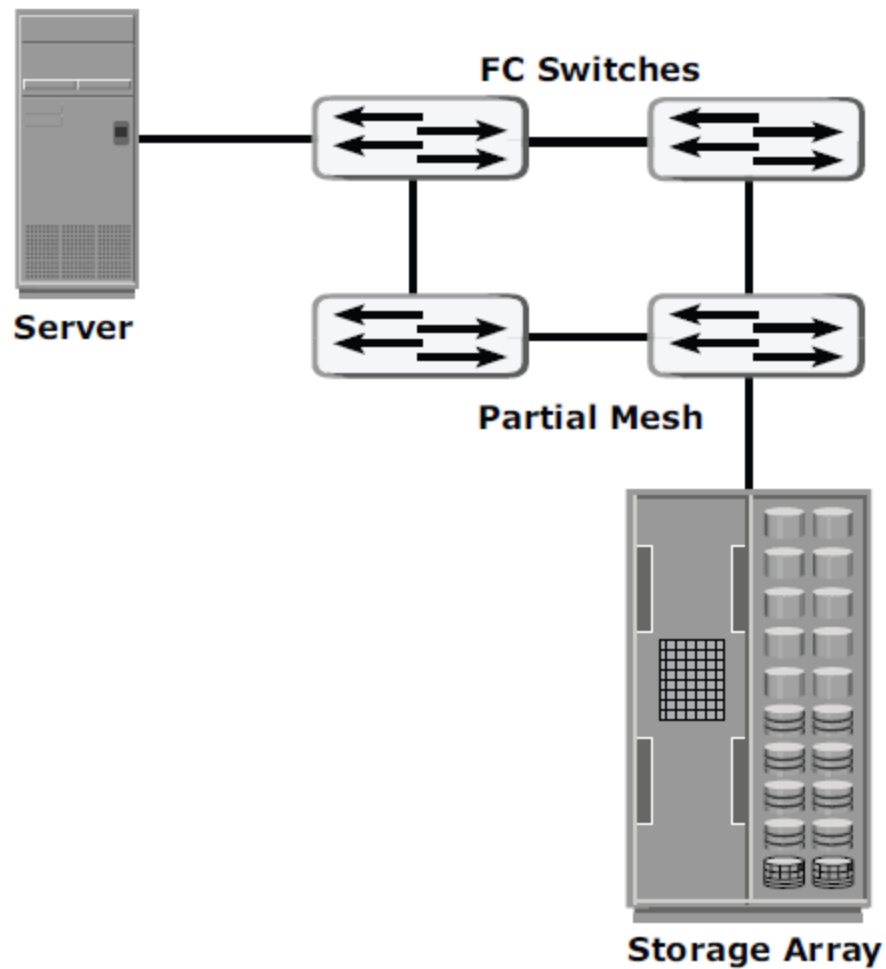
Fan-in и Fan-out

- Виртуализация ресурсов
- Коэффициент масштабирования
- Fan-out
 - несколько хостов подсоединены к одному порту СХД
- Fan-in
 - хост подсоединен одним портом к нескольким СХД
- Типичный Fan-out = 4:1

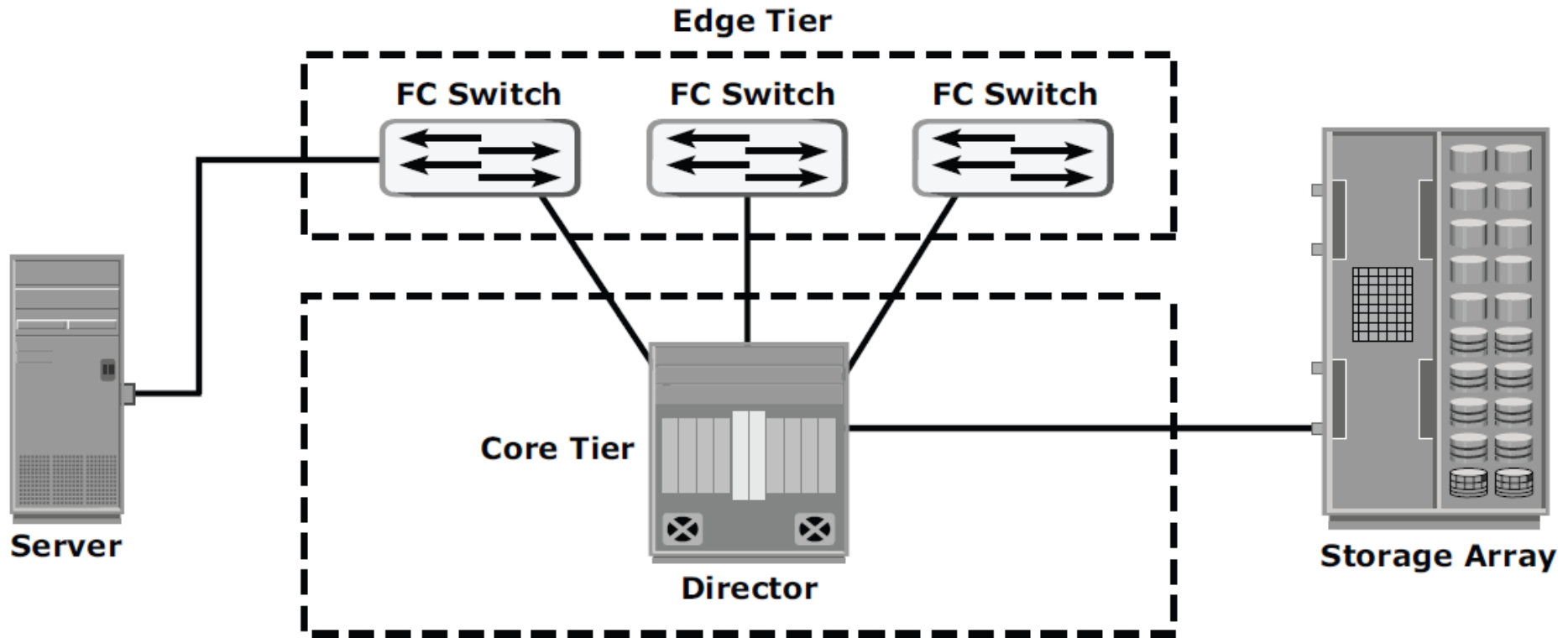
Топология Full Mesh



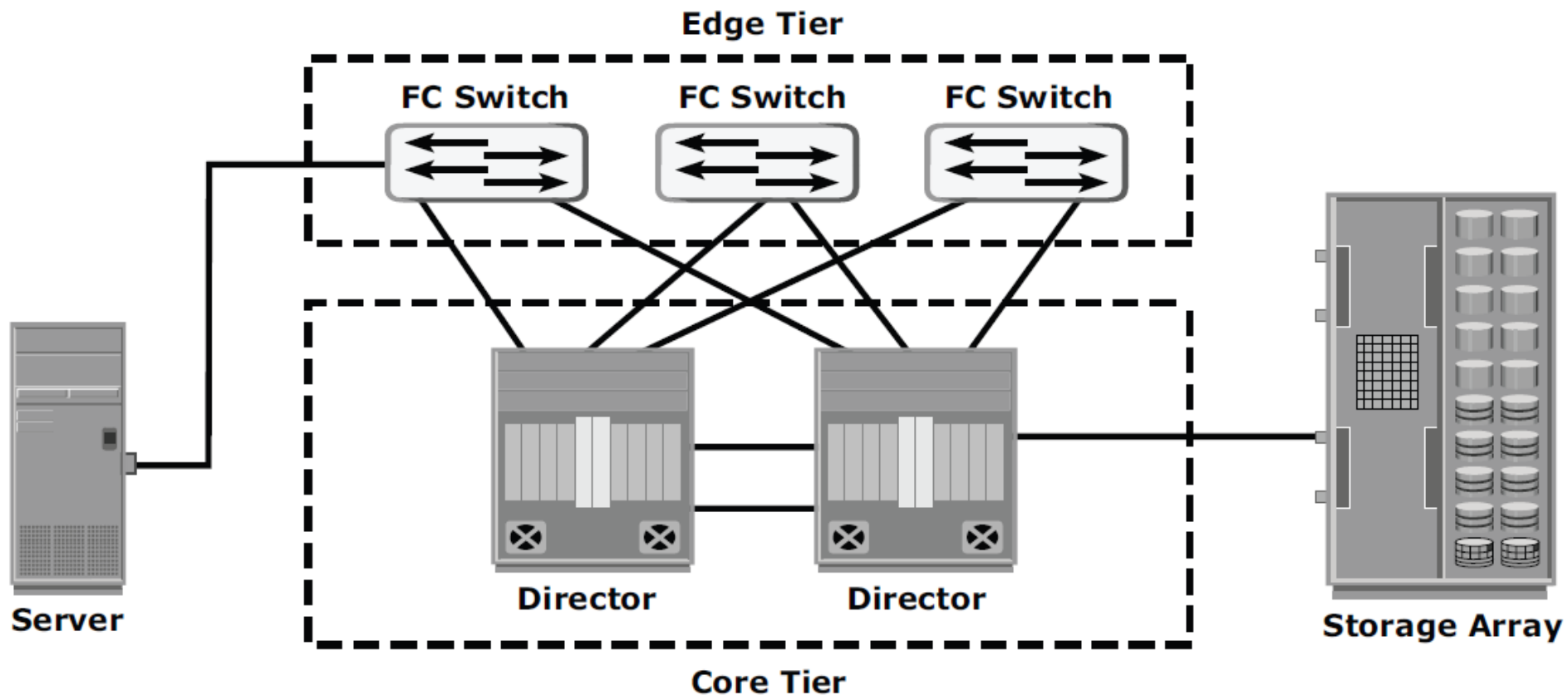
Топология Partial Mesh



Топология Core-Edge - Single Core



Топология Core-Edge - Dual Core



Спасибо!

EMC²®