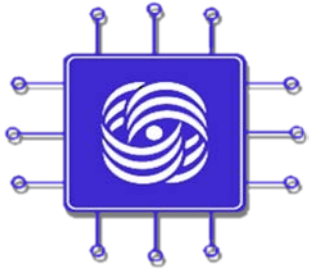


Сетевые Хранилища Данных – Storage Area Network

Проф. Смелянский Р.Л. Доп. главы компьютерных сетей
Сетевые Хранилища Данных

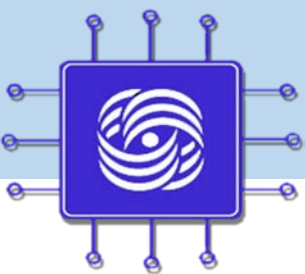
Storage Networks Explained: Basics and Application of Fibre Channel SAN, NAS, iSCSI, InfiniBand and FCoE, Second Edition.

U. Troppens, W. Müller-Friedt, R. Wolafka, R. Erkens and N. Haustein © 2009 John Wiley & Sons Ltd. ISBN: 978-0-470-74143-6

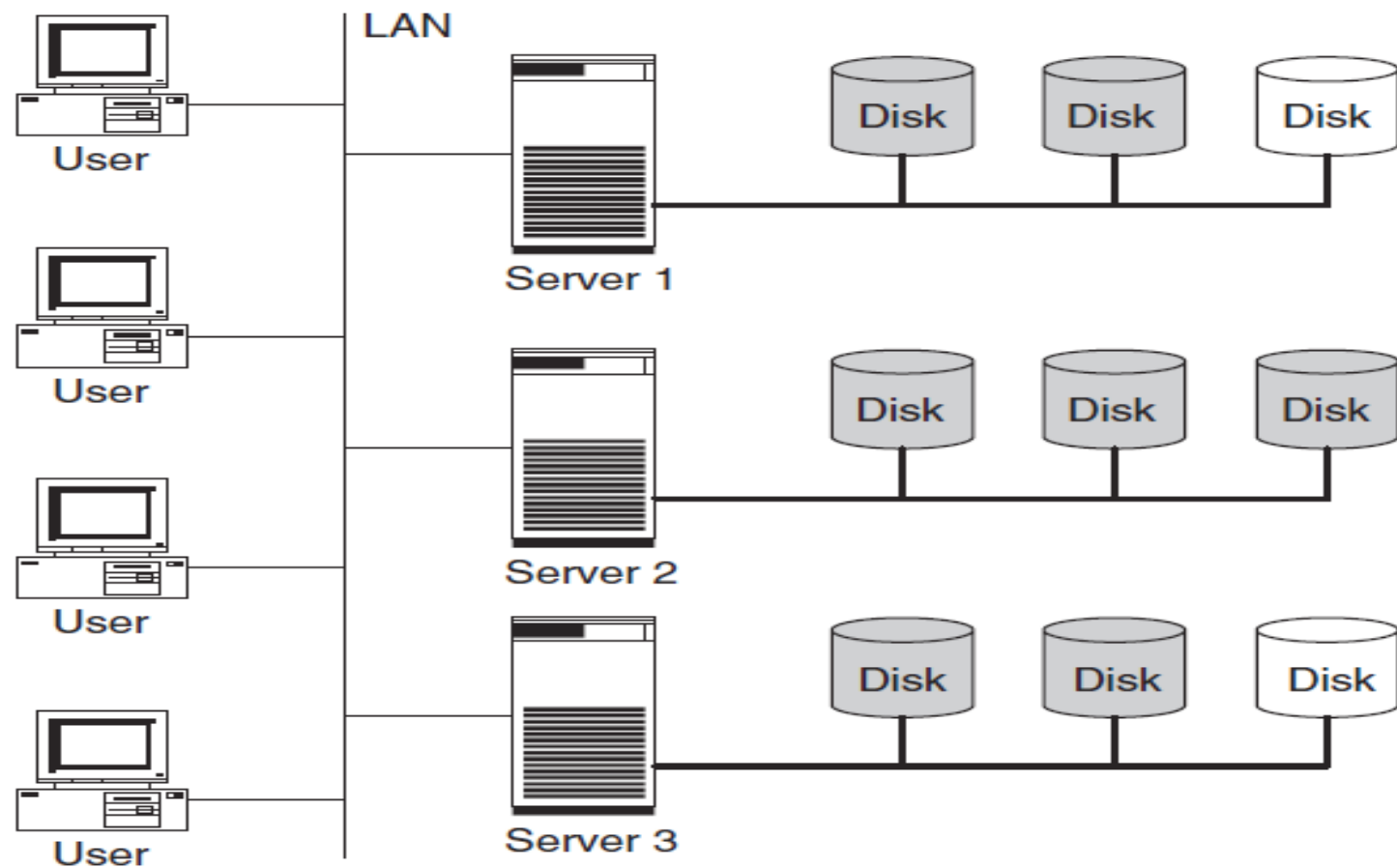


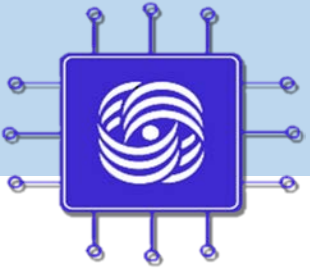
Содержание

- Архитектура информационной инфраструктуры
- Дисковые подсистемы и их организация
- JBOD
- RAID
- Интеллектуальные ДПС
- От CPU до ДПС
- SCSI
- Fibre Channel
- Заключение

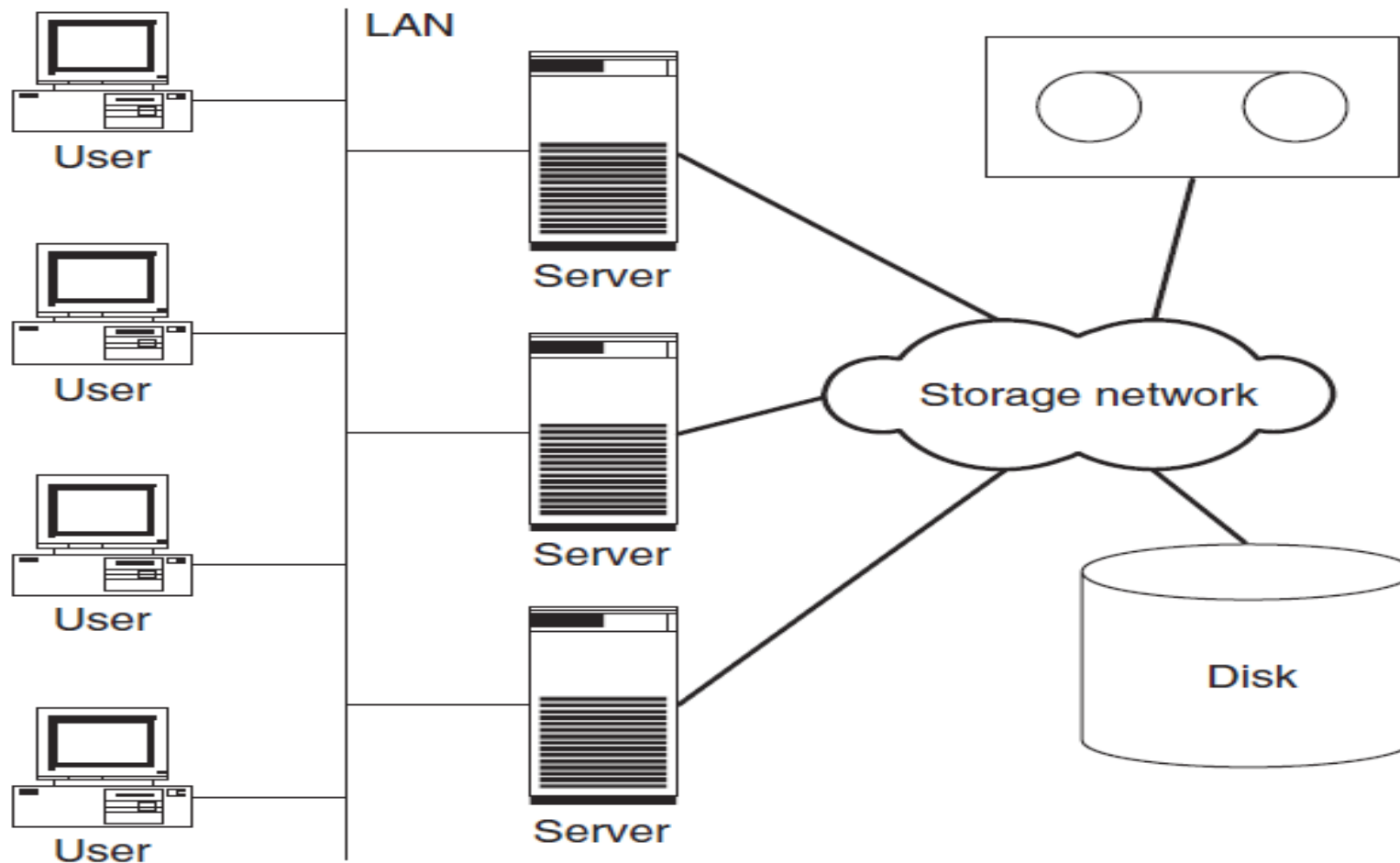


Server Centric Architecture

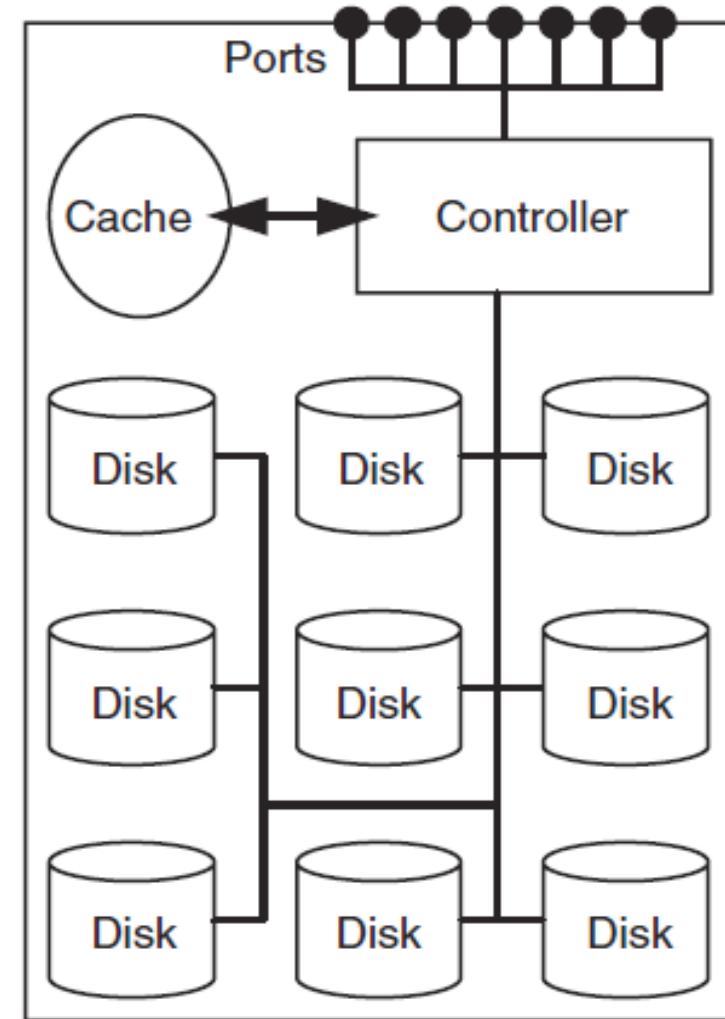
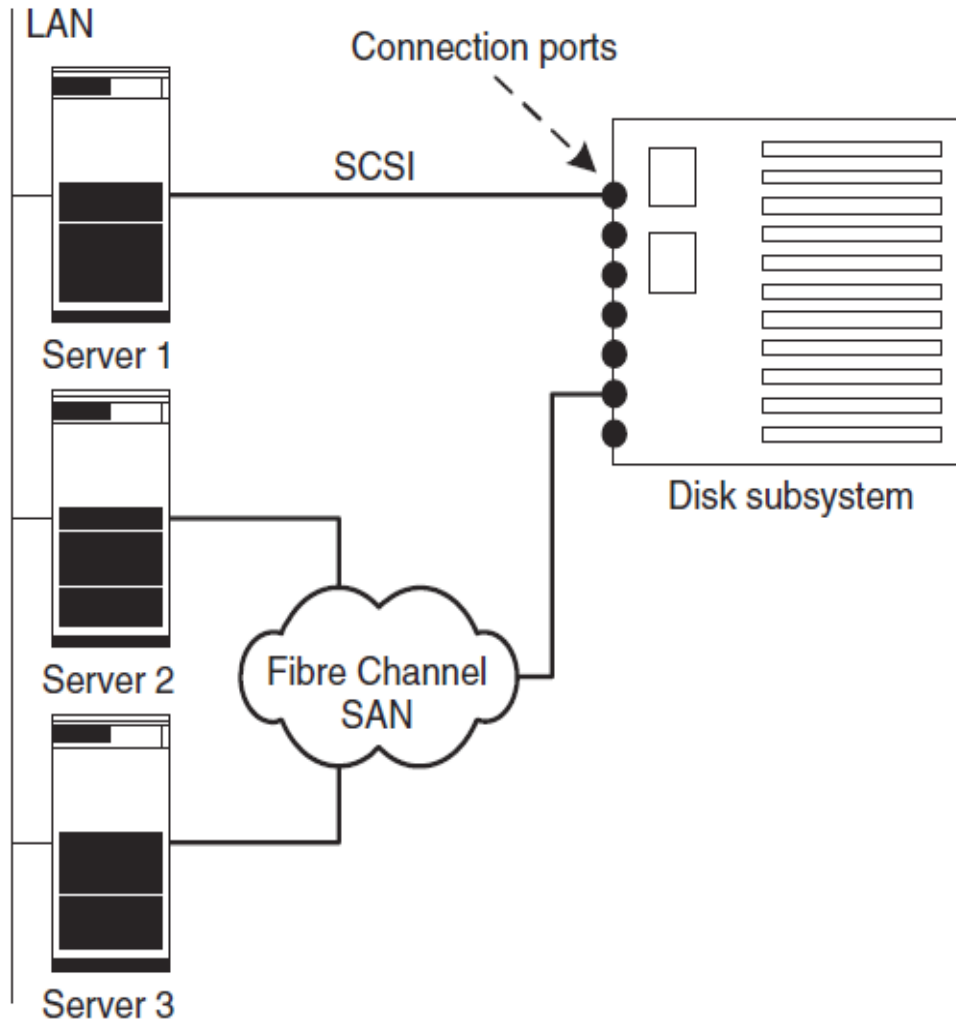
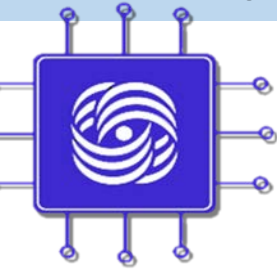


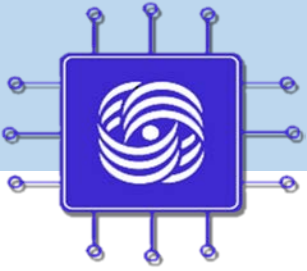


Storage Centric Architecture

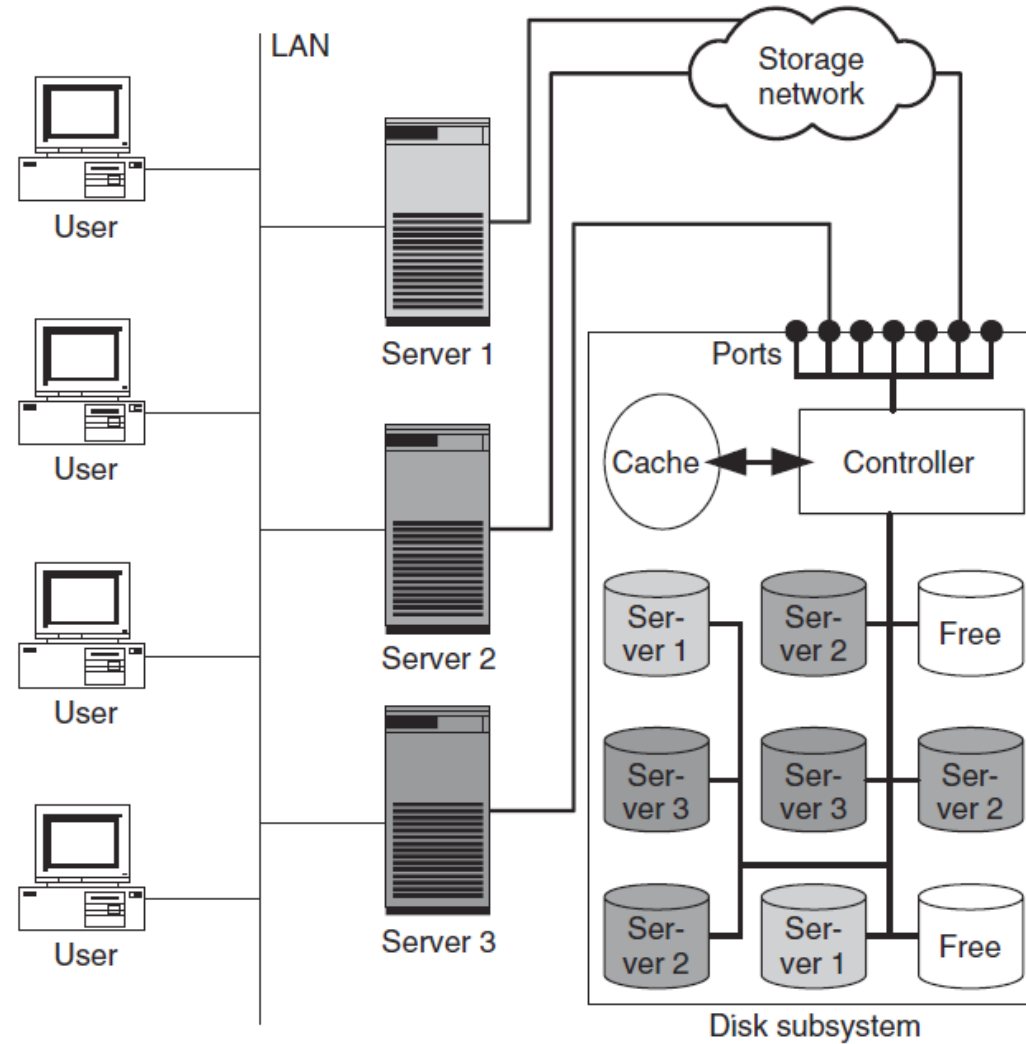


Архитектура Дисковой подсистемы (ДПС)

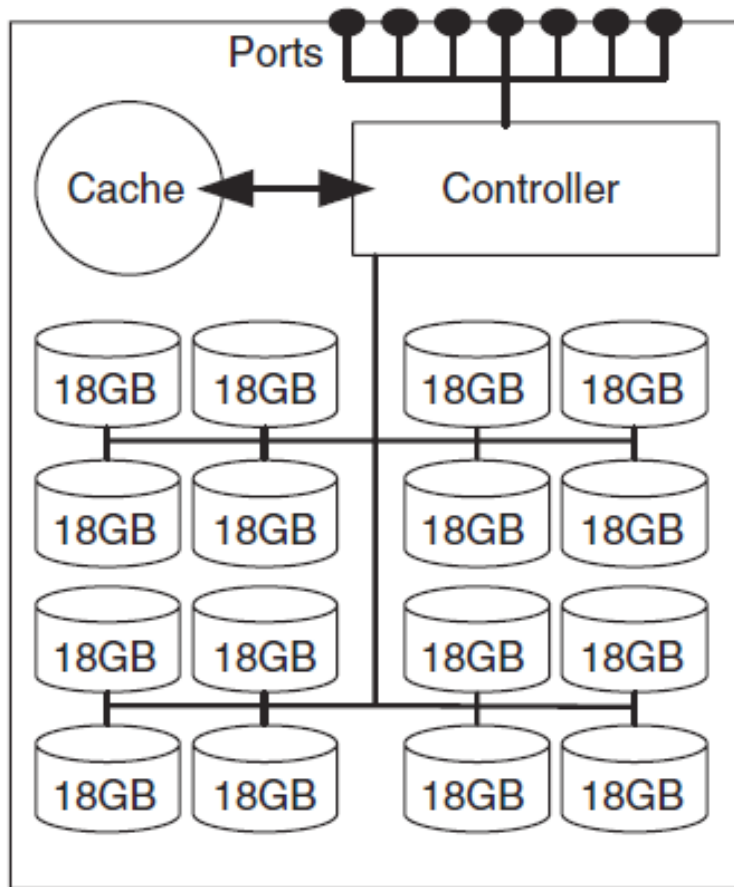
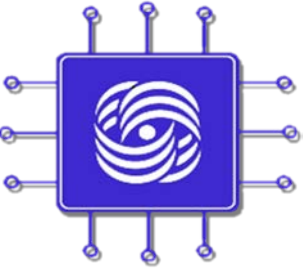




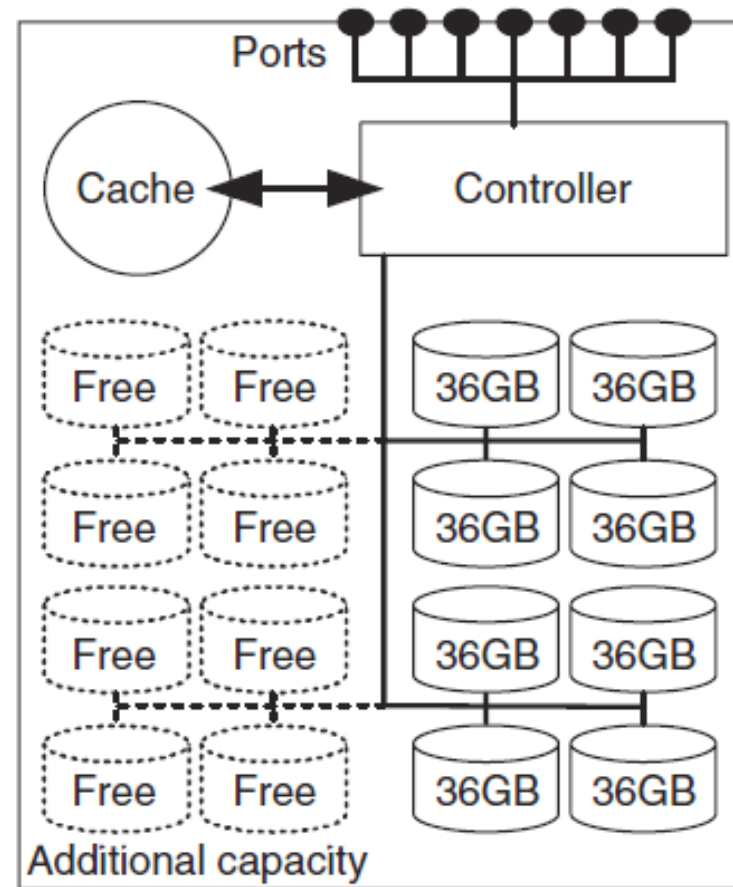
Пример использования ДПС



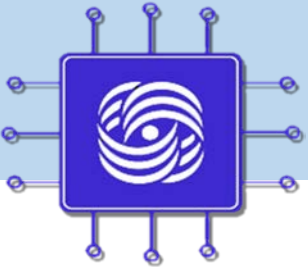
Внутренняя организация и емкость дисков



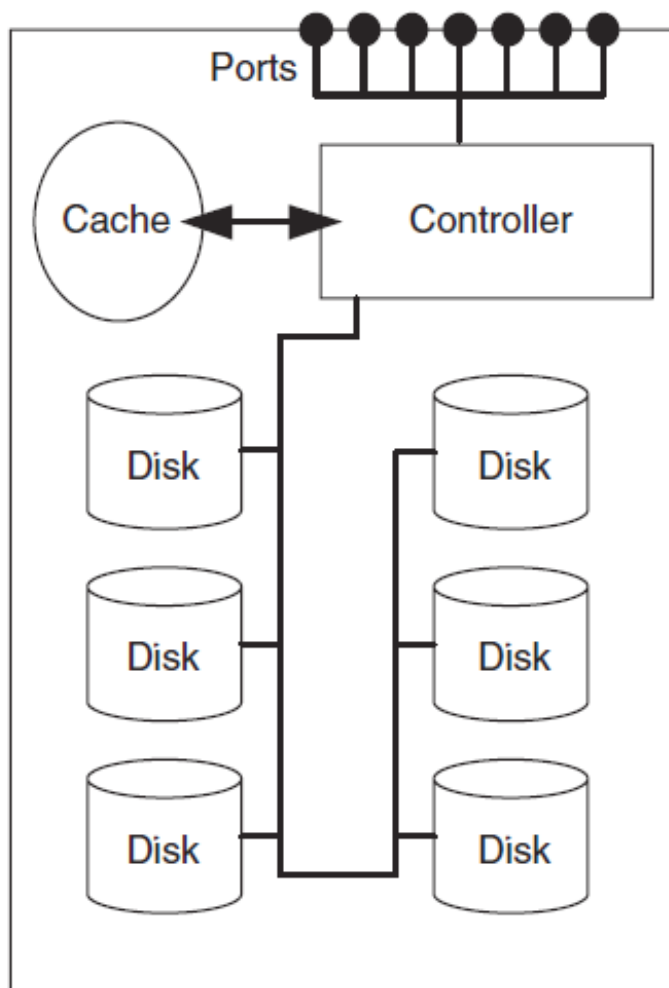
Маленькие диски



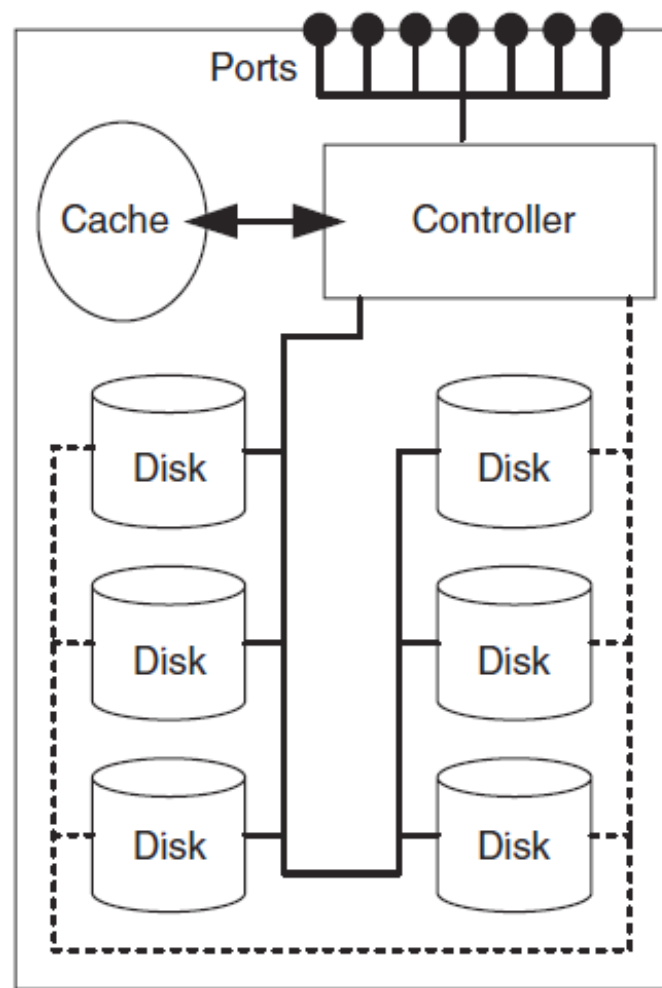
Большие диски



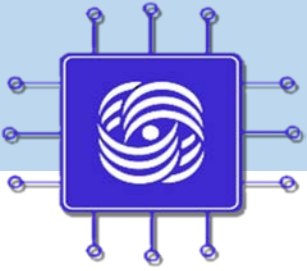
Внутренняя организация ДПС



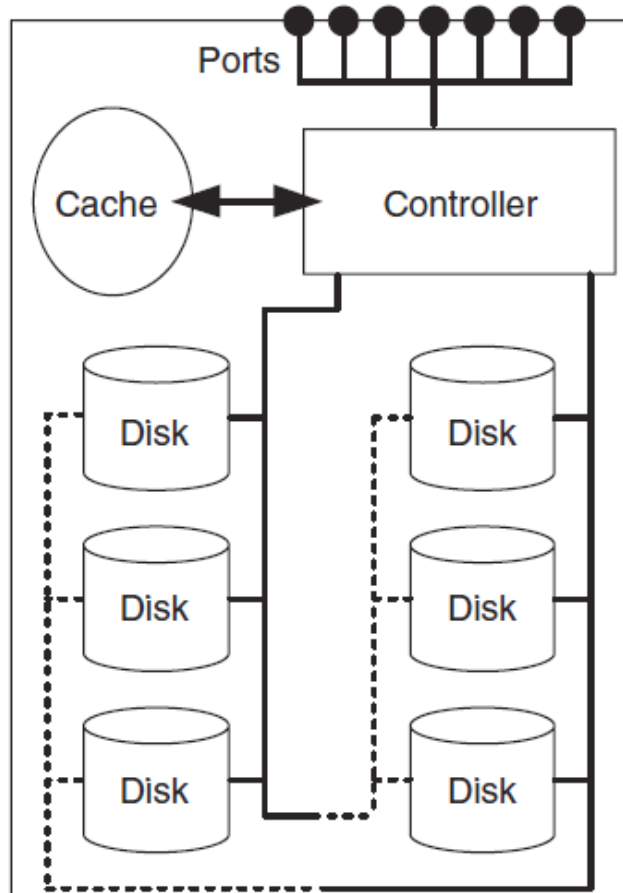
Одиная



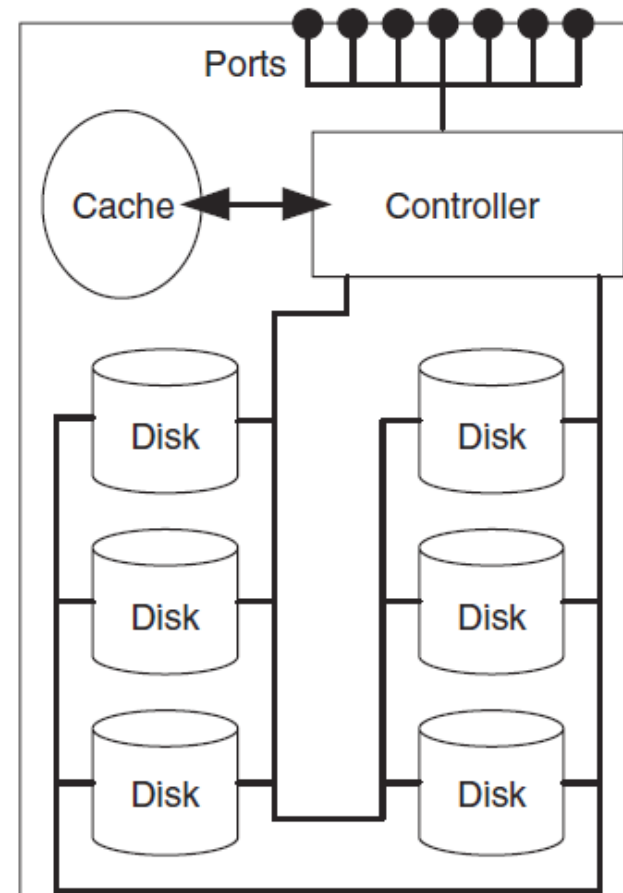
Дублированная



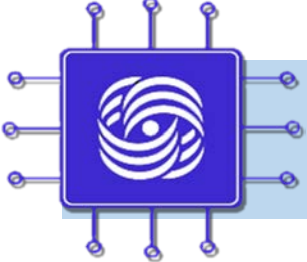
Активное дублирование



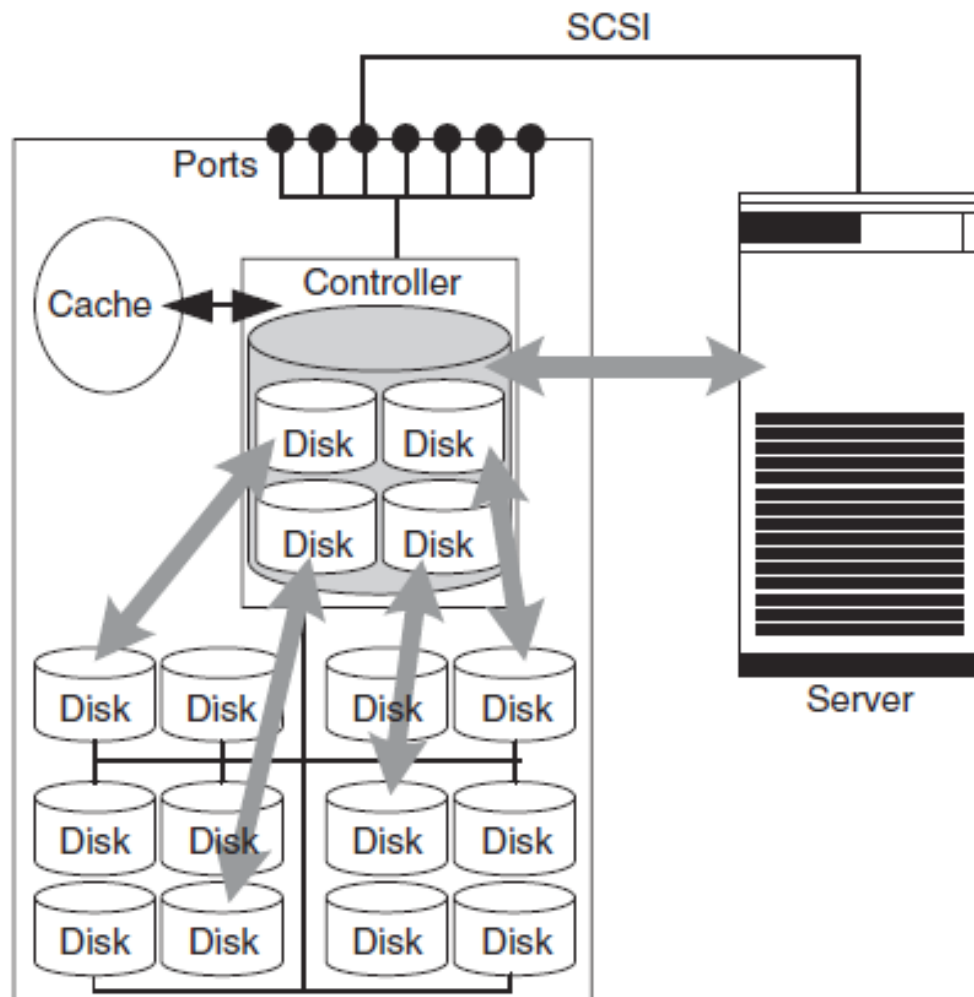
С фиксированным распределением дисков

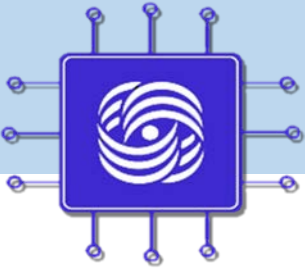


С динамическим распределением дисков

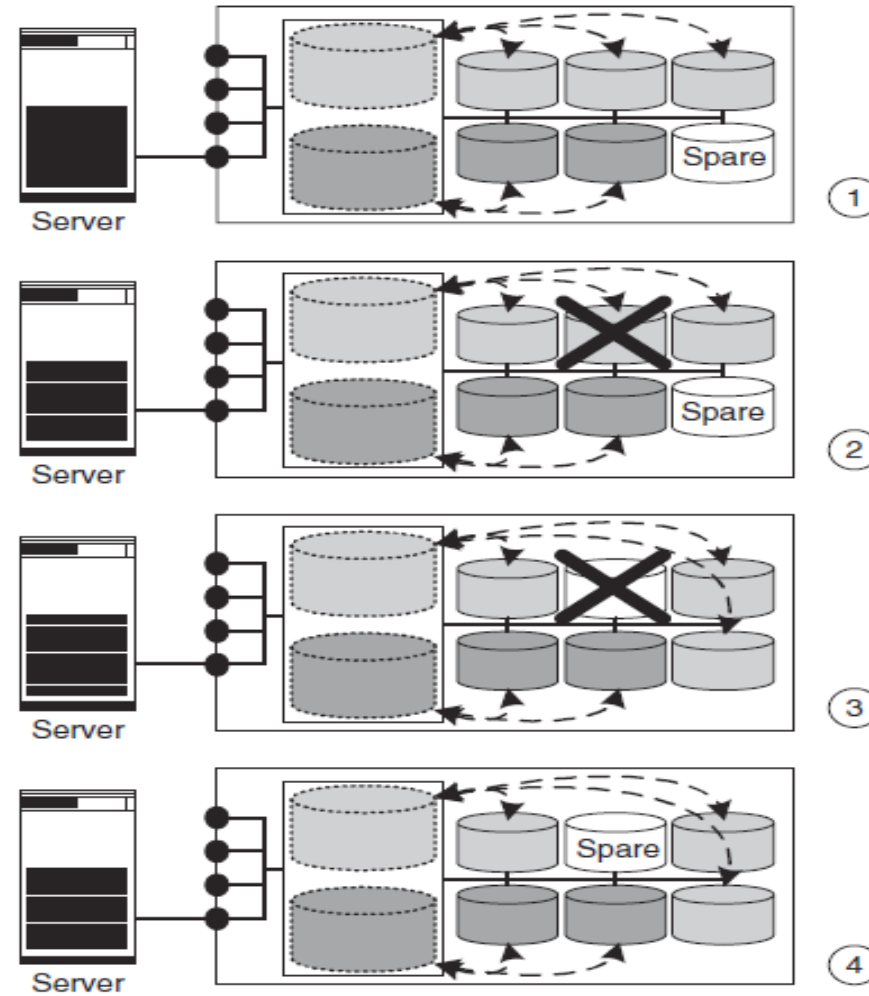


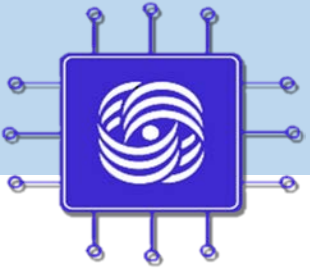
RAID – Избыточность массива независимых дисков



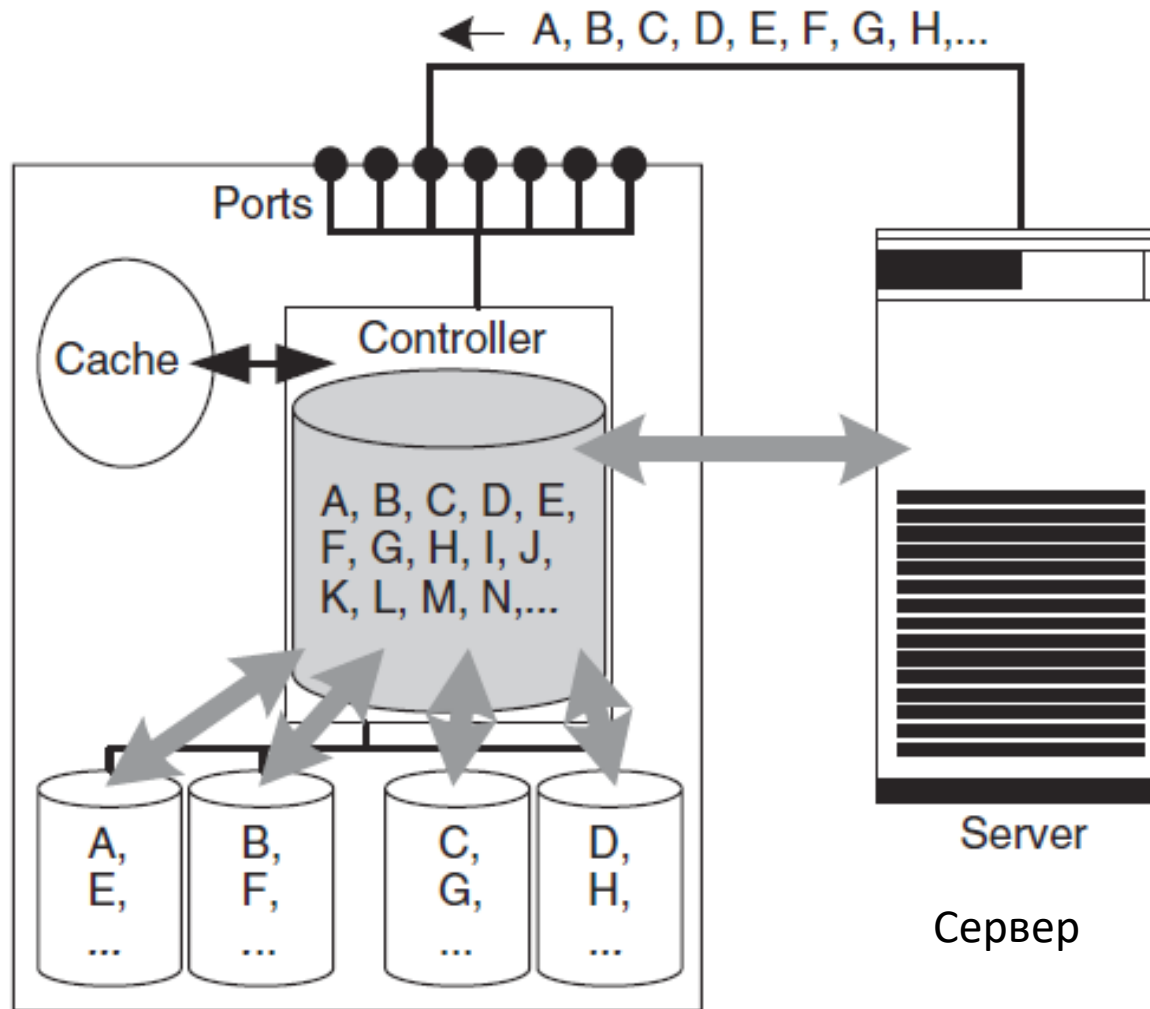


Горячий дисковый резерв

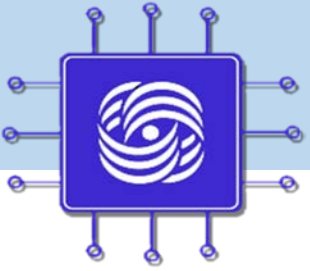




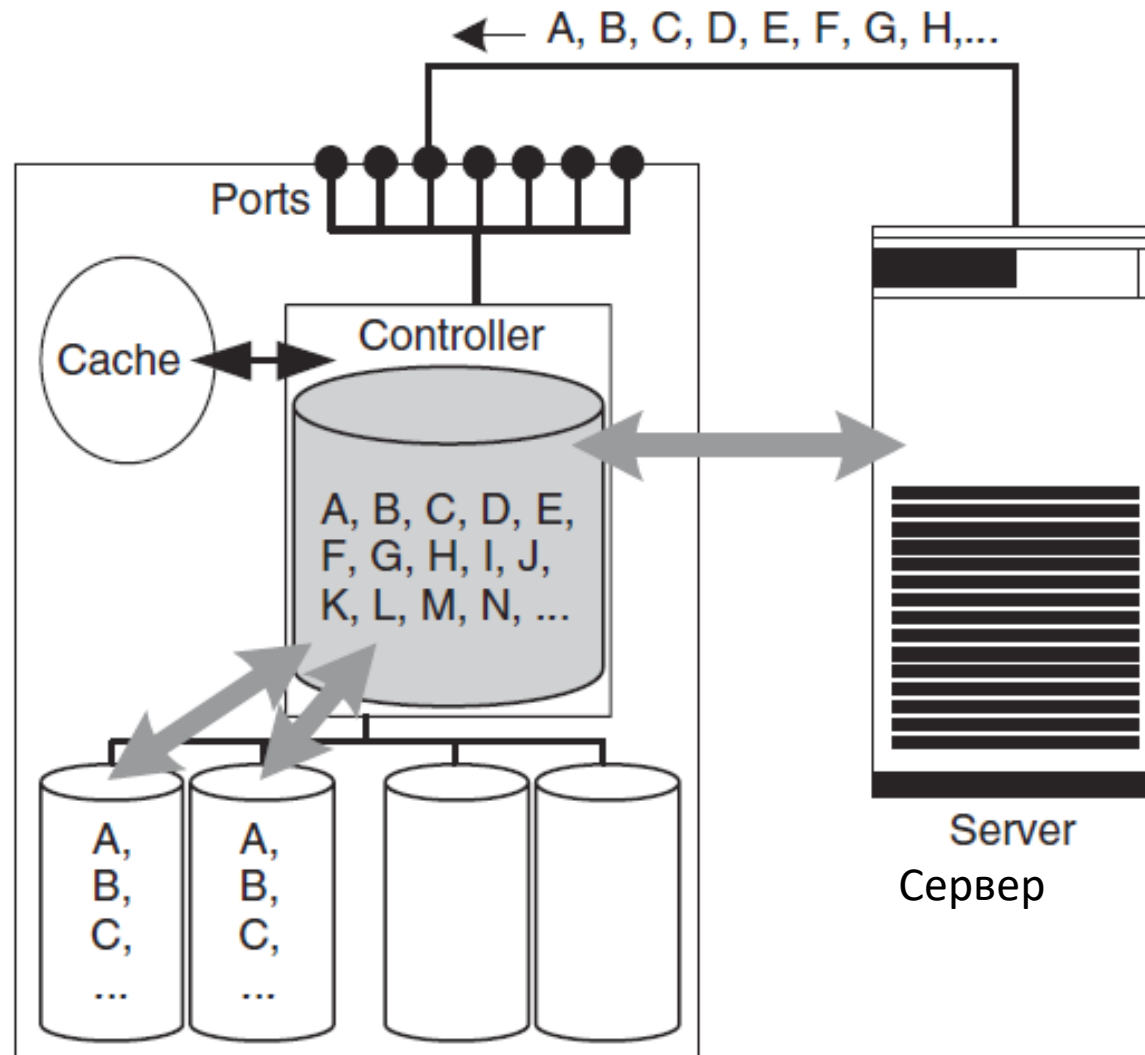
Уровни RAID = 0

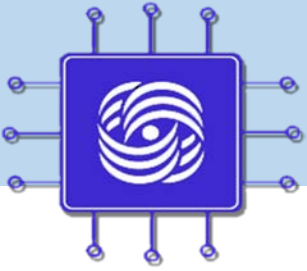


Выход из строя любого физического диска нарушит целостность данных.

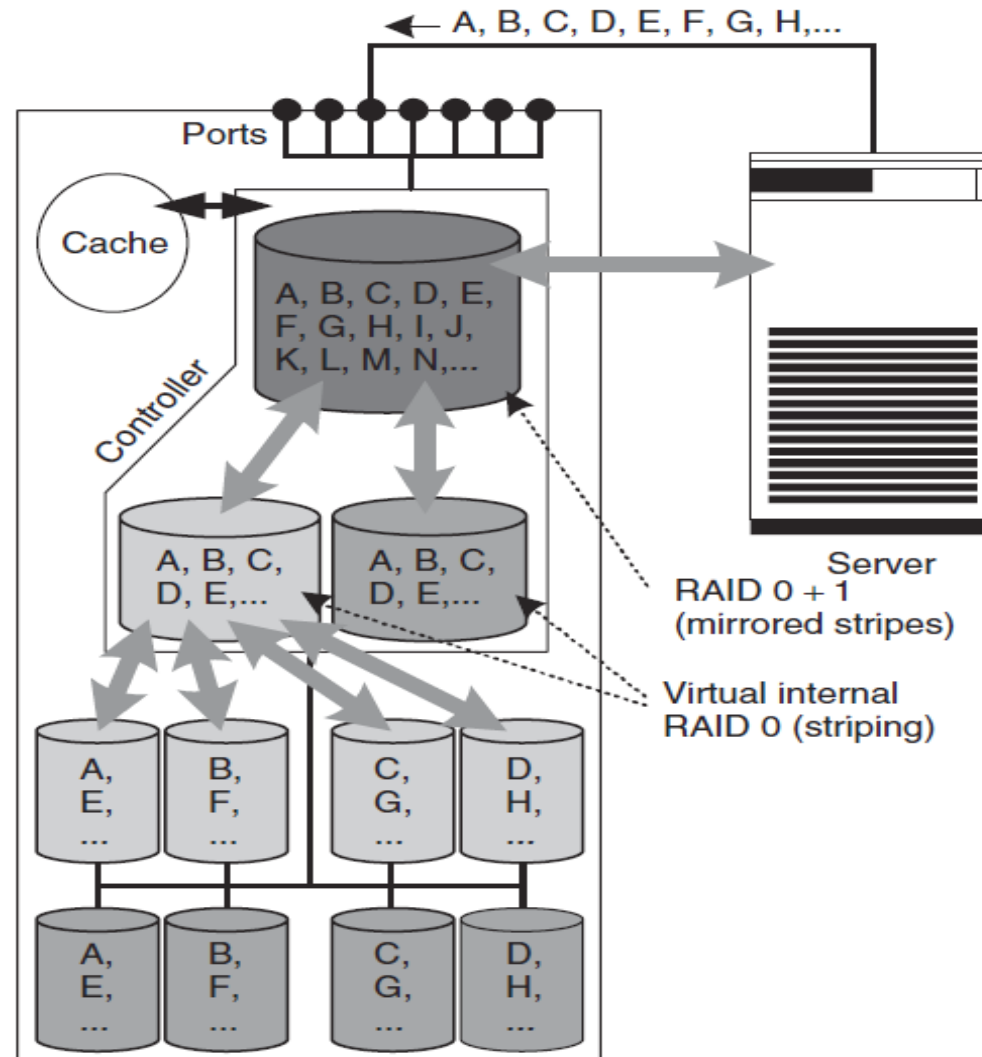


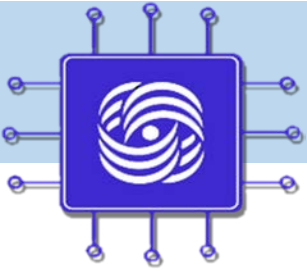
RAID = 1



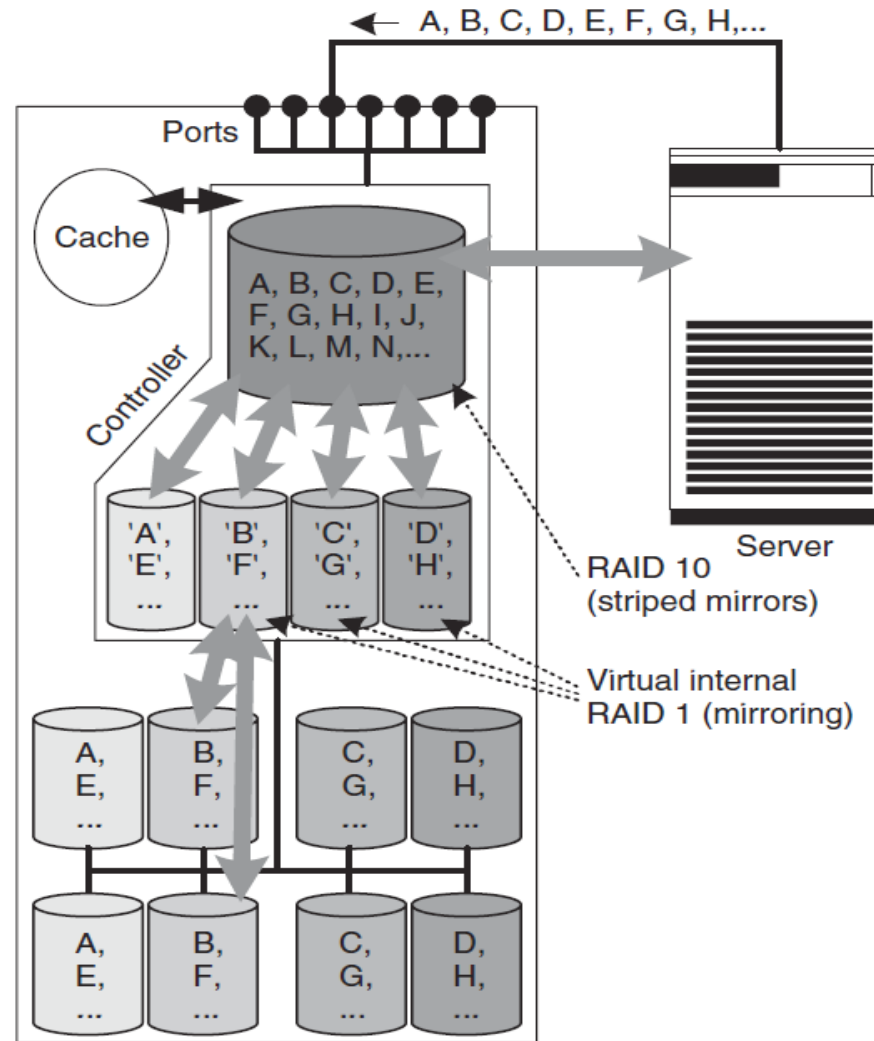


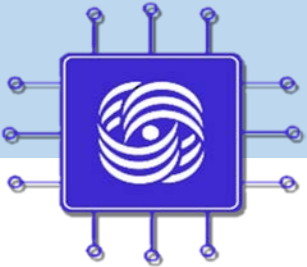
RAID = 0+1



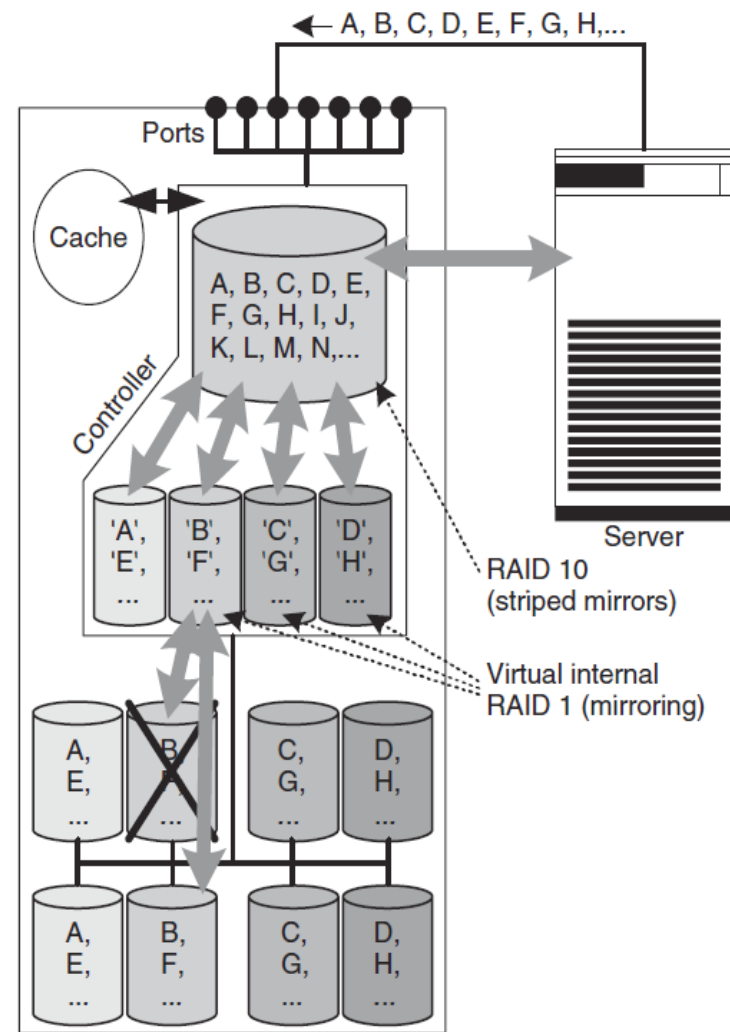
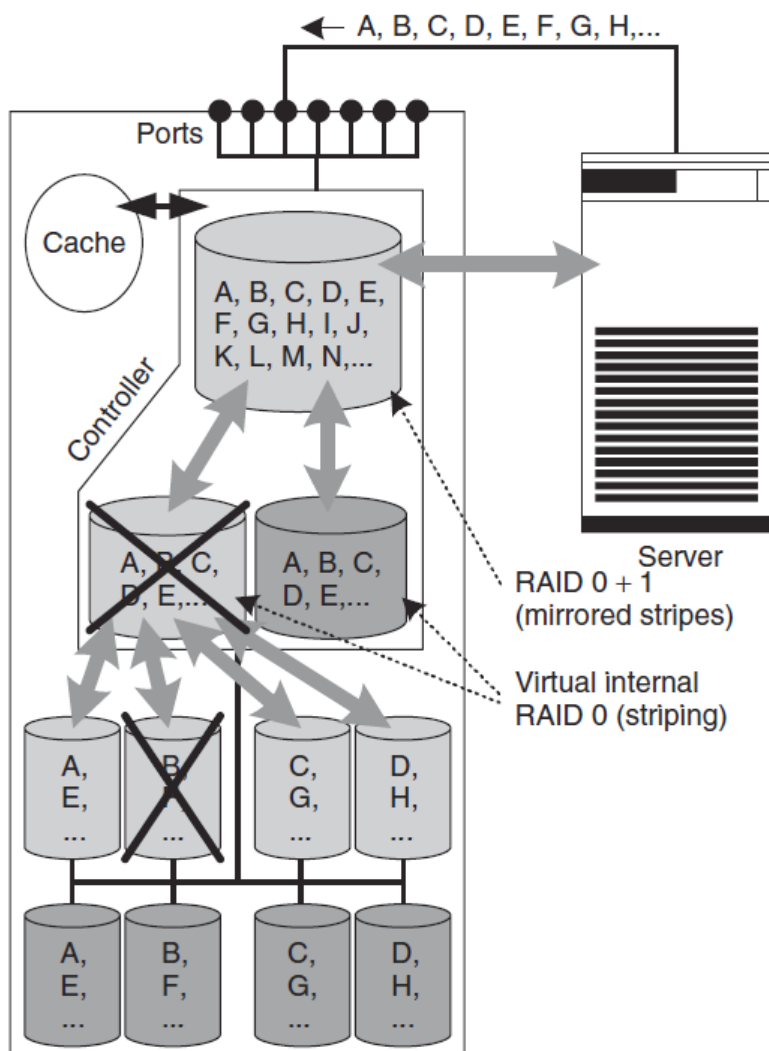


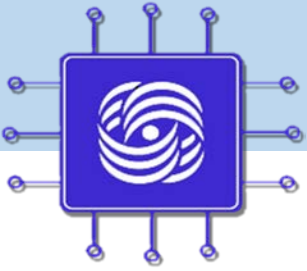
RAID = 1+0



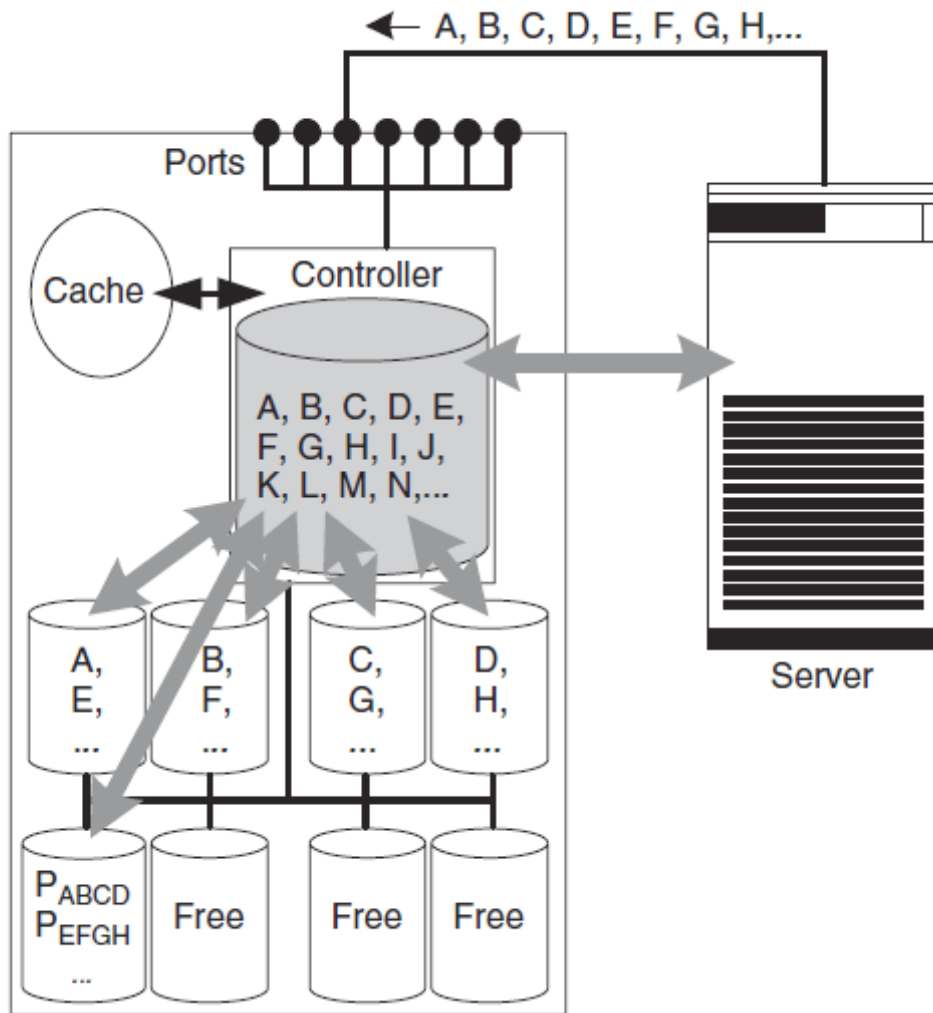


Сравнение RAID 0+1 и 1+0





RAID = 4 vs RAID = 5

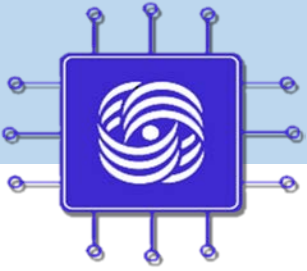


А изменили на $\sim A$

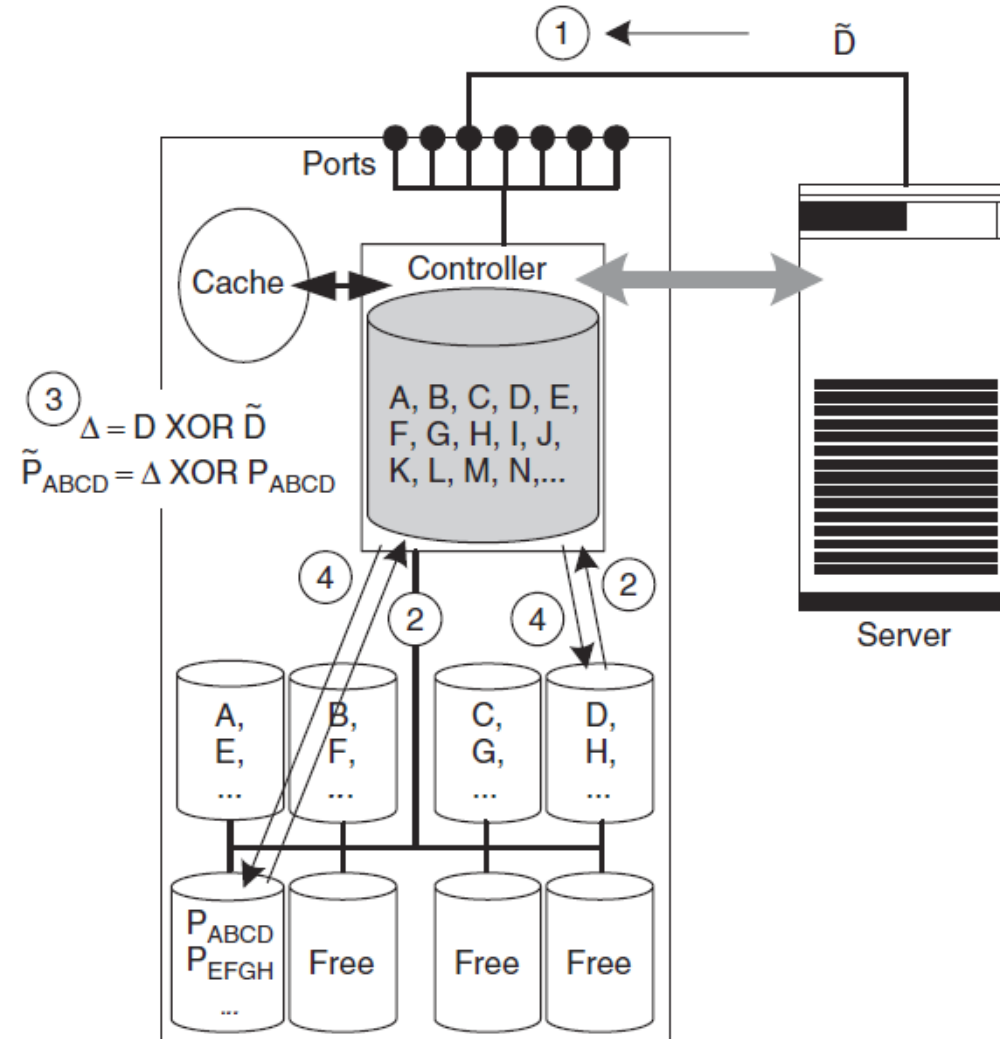
$$\Delta = A \text{ xor } \sim A$$

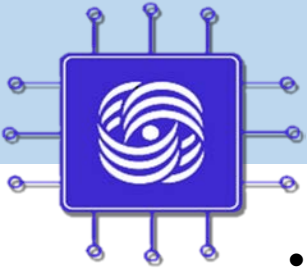
$$\sim P = \Delta \text{ xor } P$$

Если изменился только блок А,
То легко пересчитать P_{ABCD} , не читая BCD,
но нужен старый блок А, чтобы рассчитать Δ



Накладные расходы на запись в RAID = 4 и 5



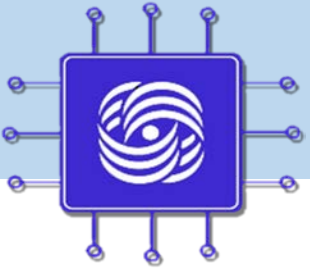


RAID = 6

- Современные диски 1TB с BER 10^{-15} => 100 TB одним сектором без ошибок не считать
- 10 дисковых массивов по 16X1TB будут терять один массив 1 раз в год+
- Режим эксплуатации теперь 7X24
- RAID 6 использует дополнительный диск четности
 - Увеличение затрат
 - Увеличение времени операции записи и коррекции

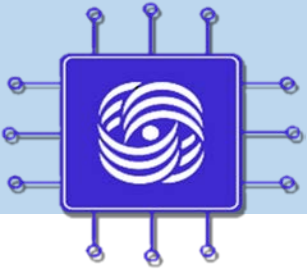
Сравнение схем RAID массивов

RAID level	Fault-tolerance	Read performance	Write performance	Space requirement
RAID 0	None	Good	Very good	Minimal
RAID 1	High	Poor	Poor	High
RAID 10	Very high	Very good	Good	High
RAID 4	High	Good	Very very poor	Low
RAID 5	High	Good	Very poor	Low
RAID 6	Very high	Good	Very very poor	Low

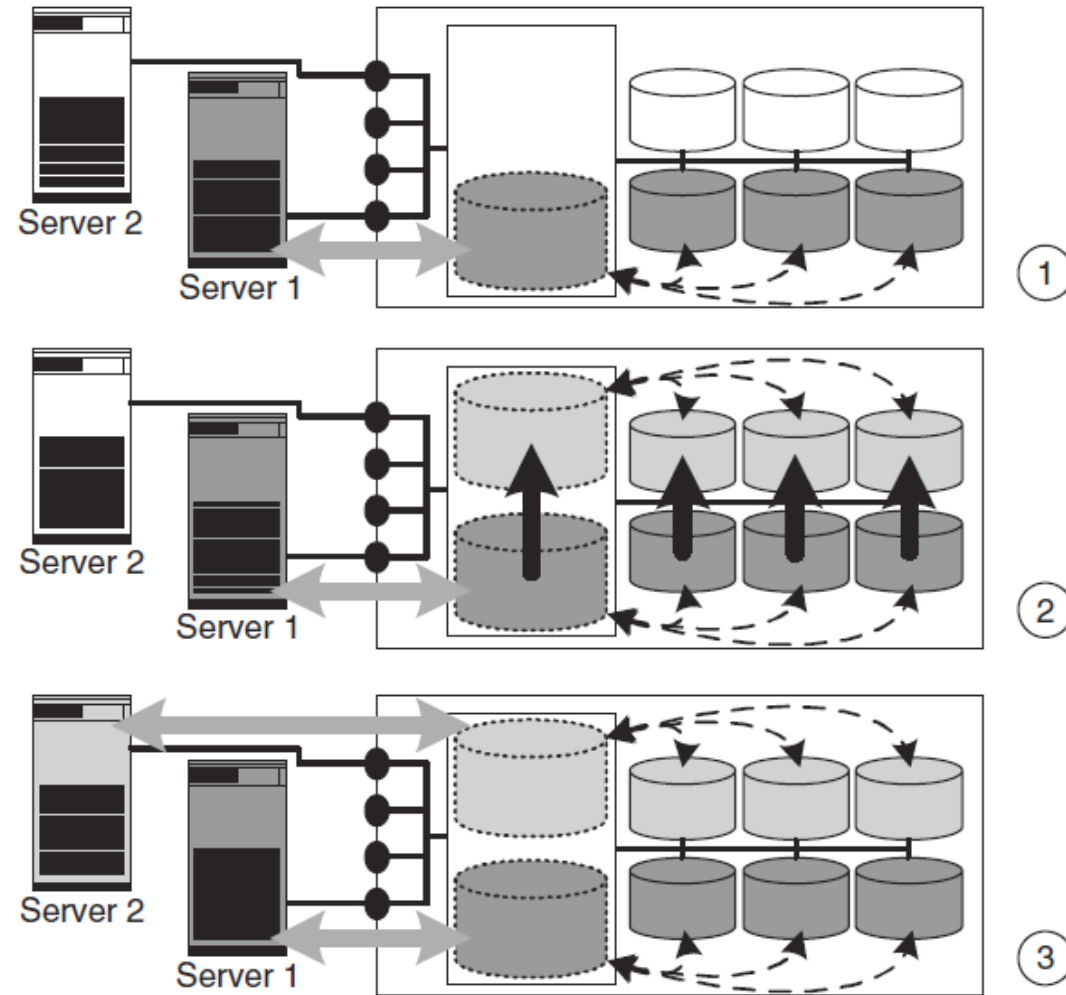


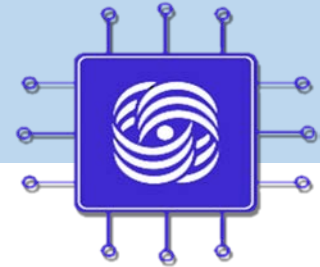
Caching – ускорение дисковых хранилищ

- Кэш на уровне HD
- Кэш на уровне контроллера ДС при записи
 - ГБ кэш
 - Здесь главное сохранить данные в кэш даже при отключении питания (UPS)
 - Важно для блочных приложений
- Кэш для ускорения чтения контроллером ДС
- Ускорение записи
- Ускорение чтения

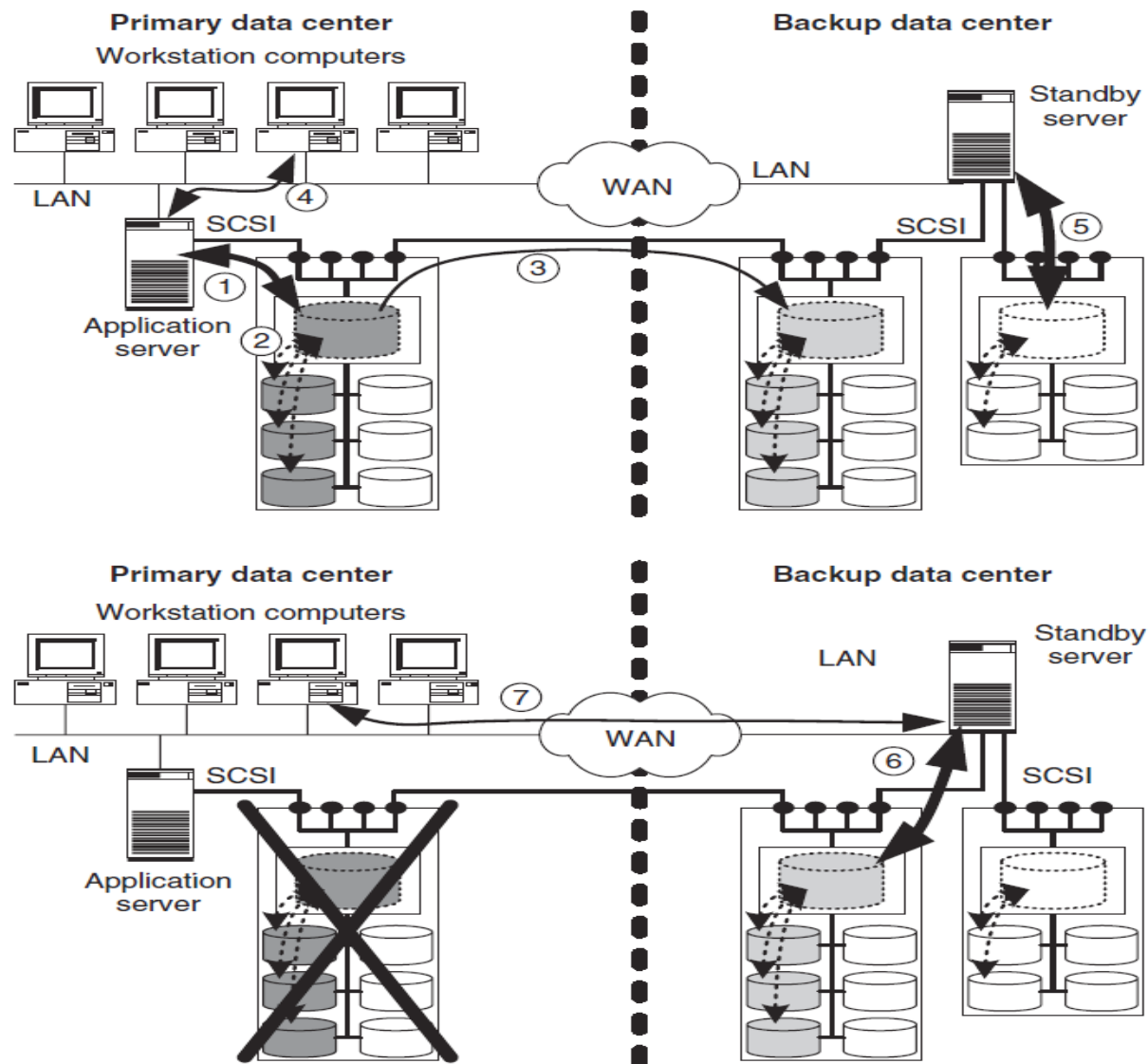


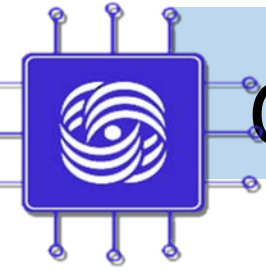
Интеллектуальная дисковая подсистема (моментальное копирование)



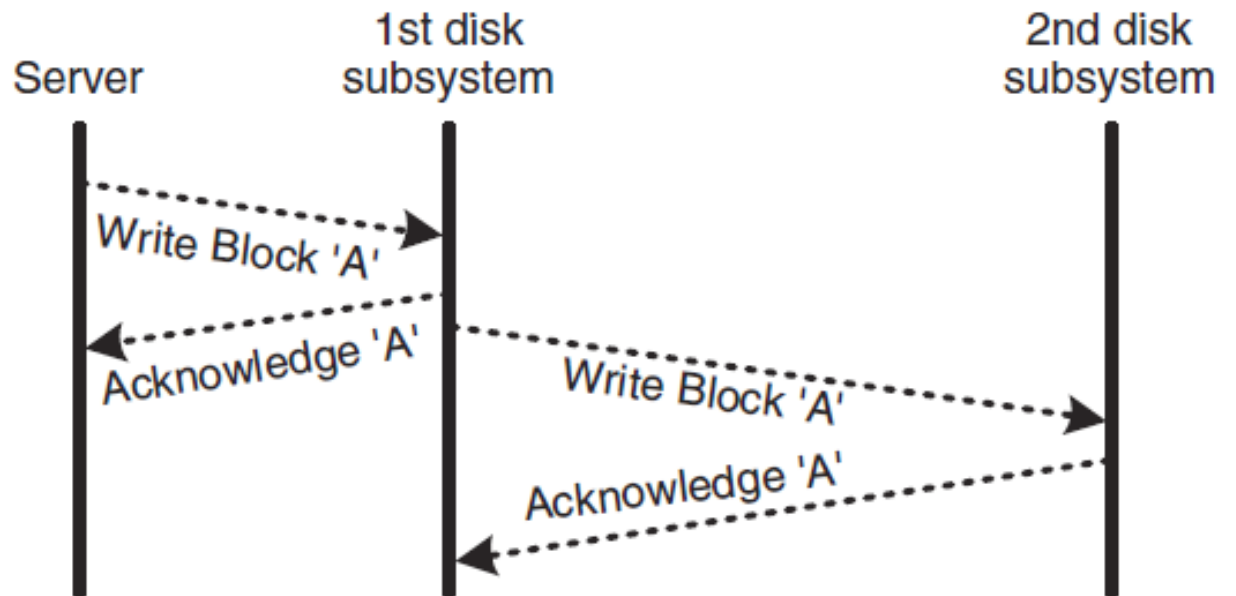
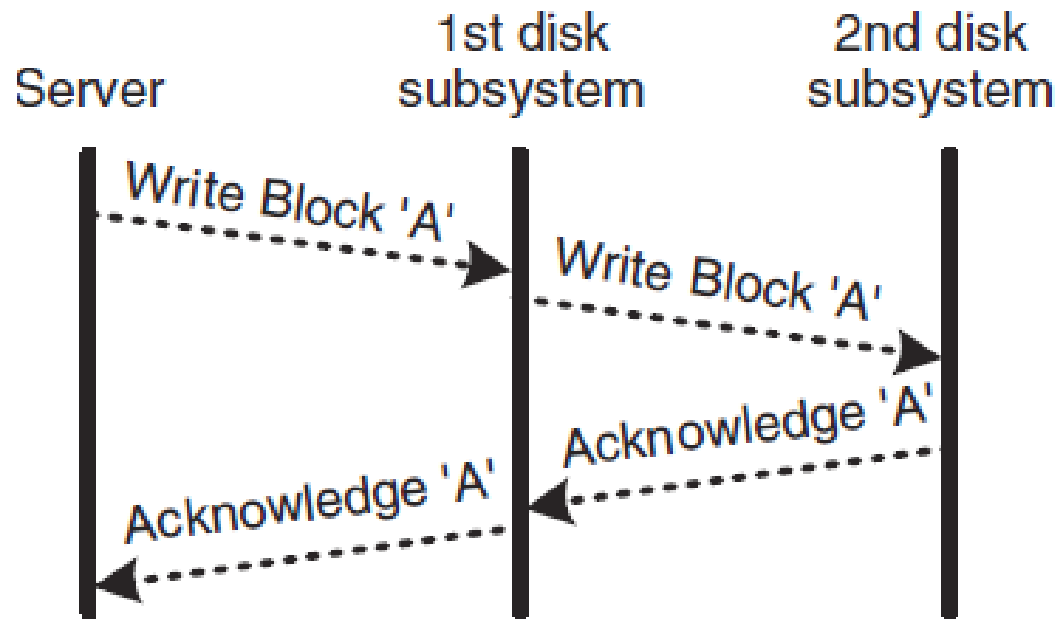


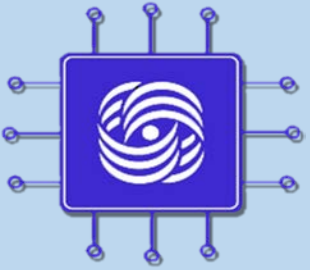
Удаленное зеркалирование



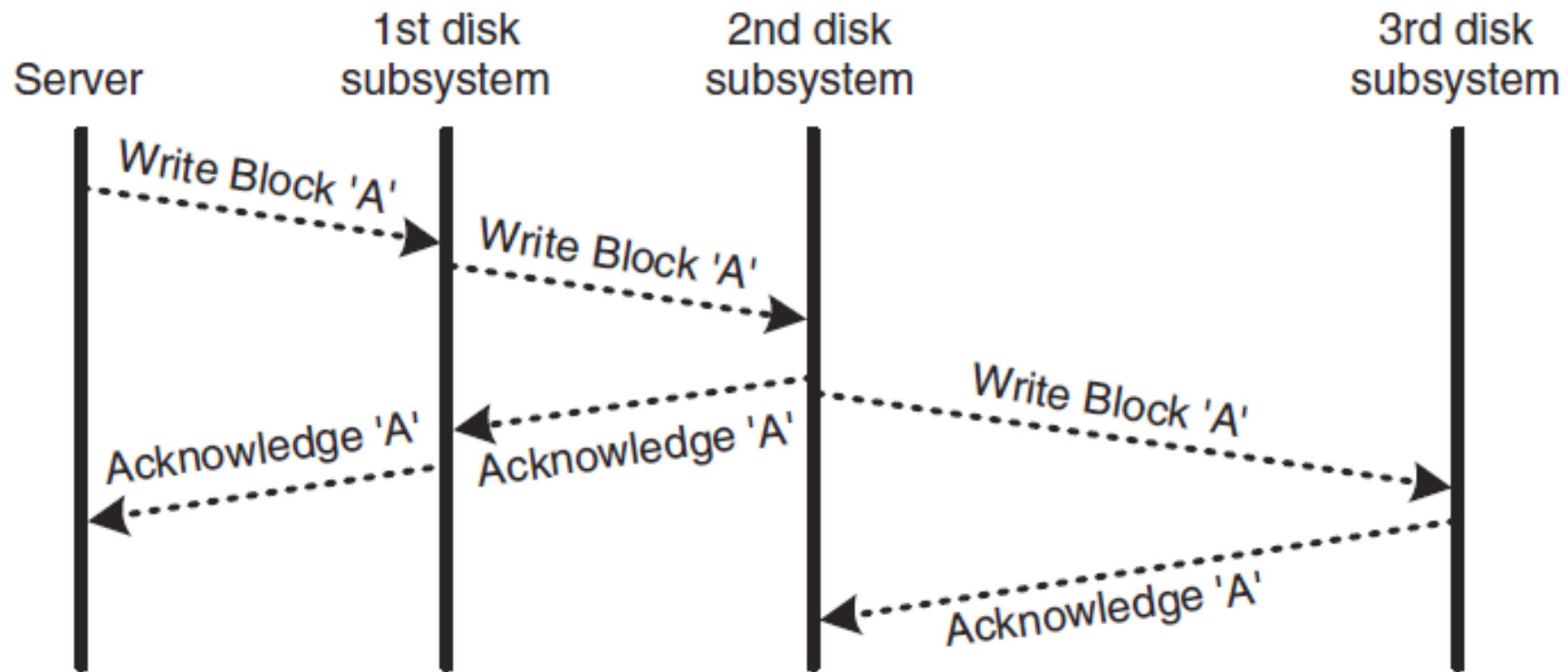


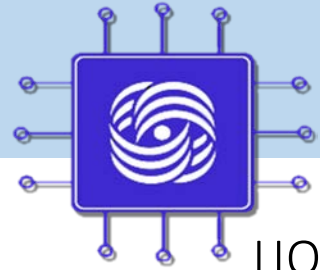
Синхронное и асинхронное удаленное зеркалирование



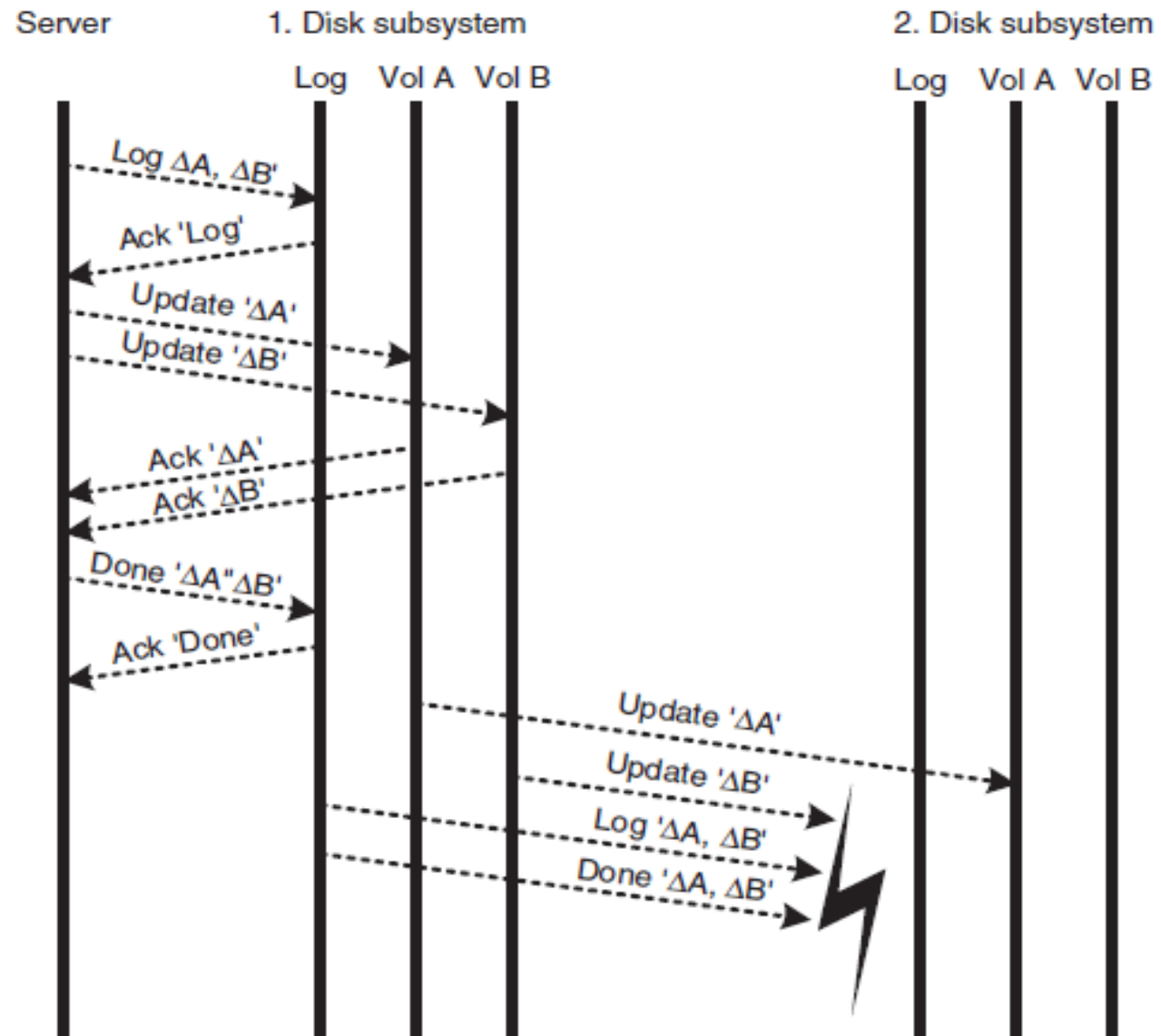
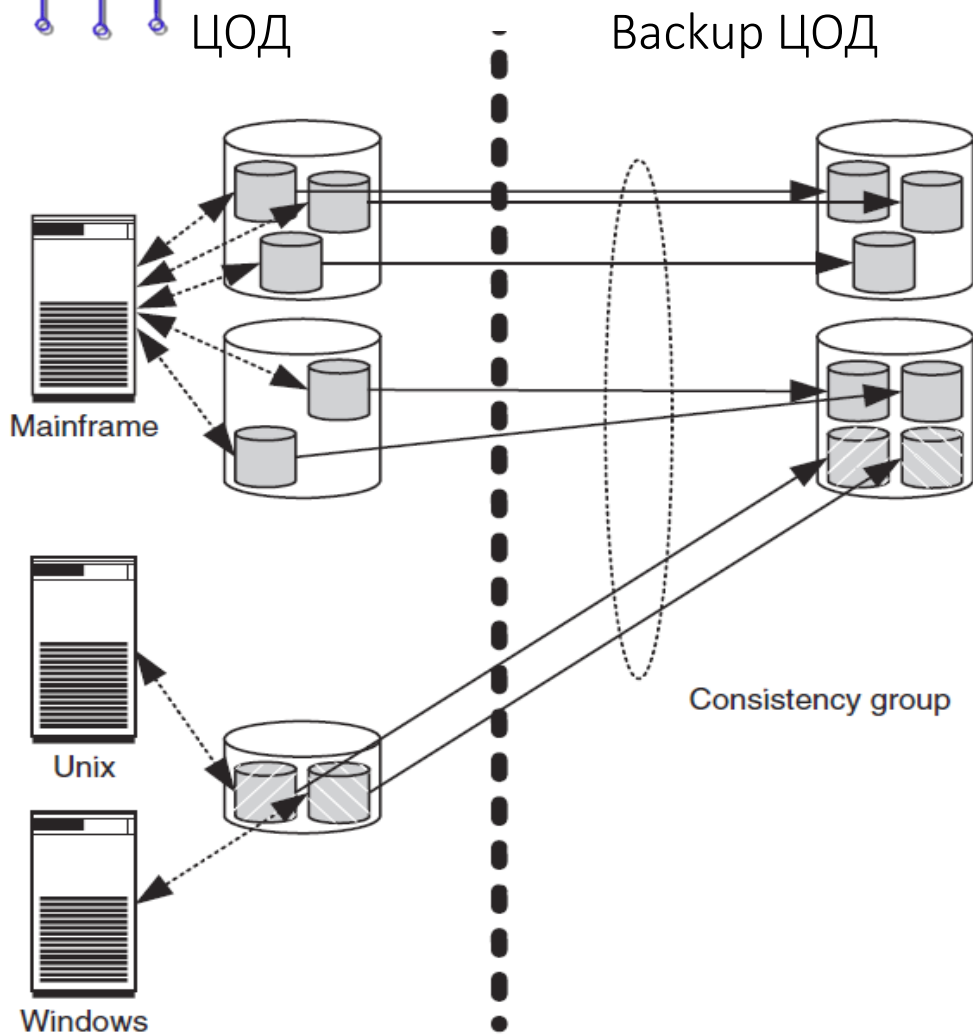


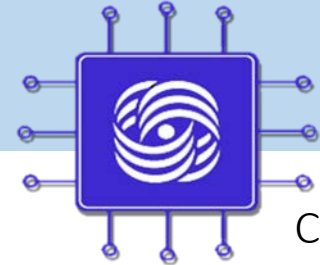
Комбинированная схема удалённого зеркалирования



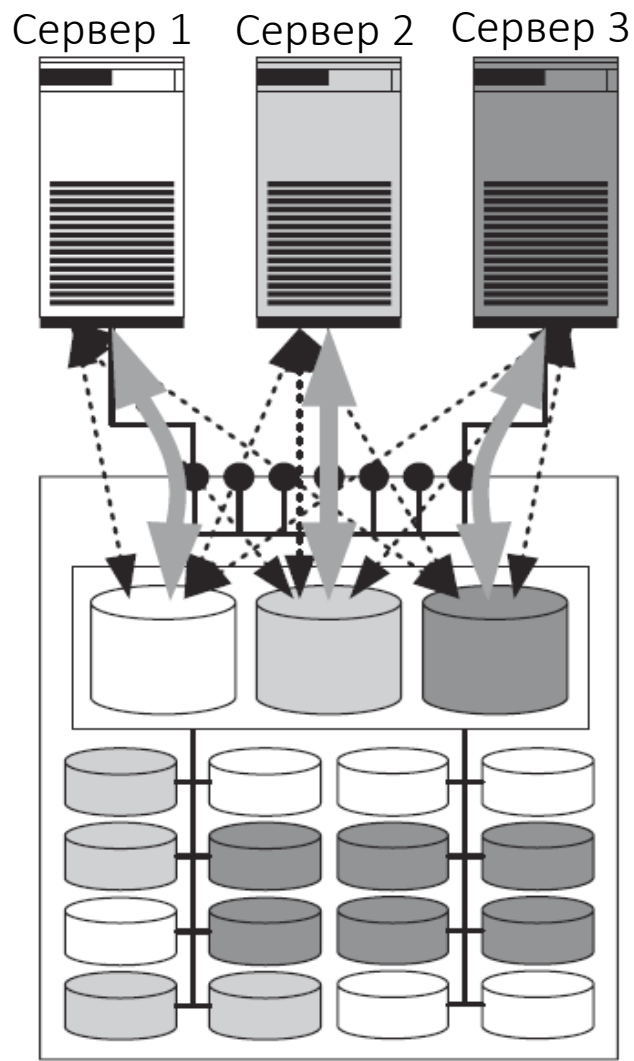


Групповая консистентность





LUN маскирование

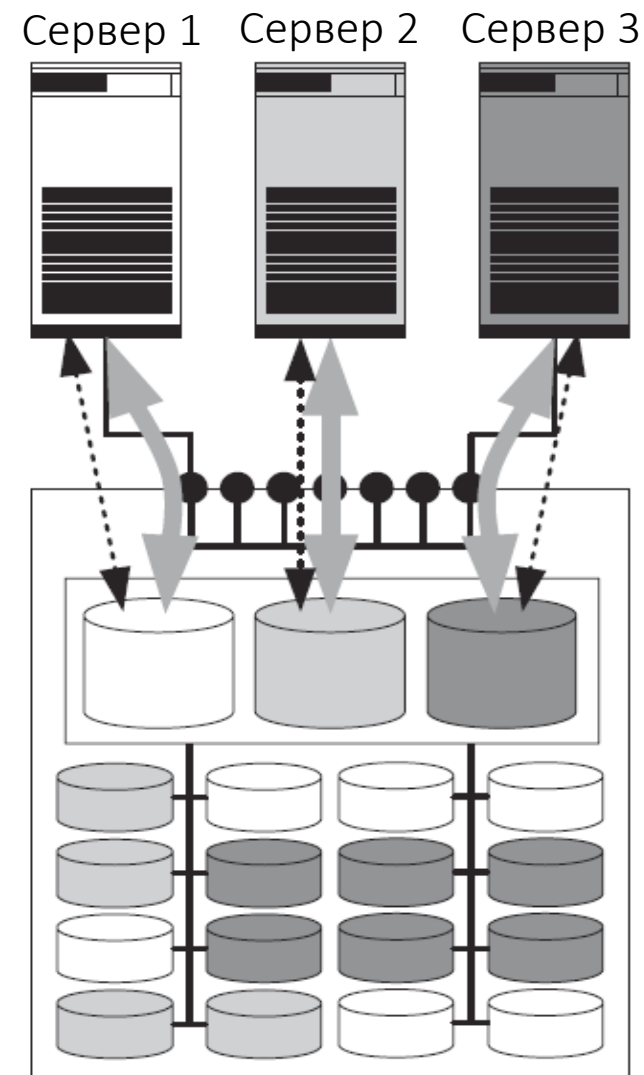


Logical
Unit
Number

Сервер использует LUN



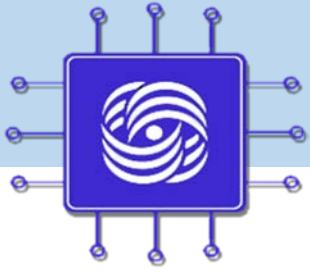
Сервер видит LUN



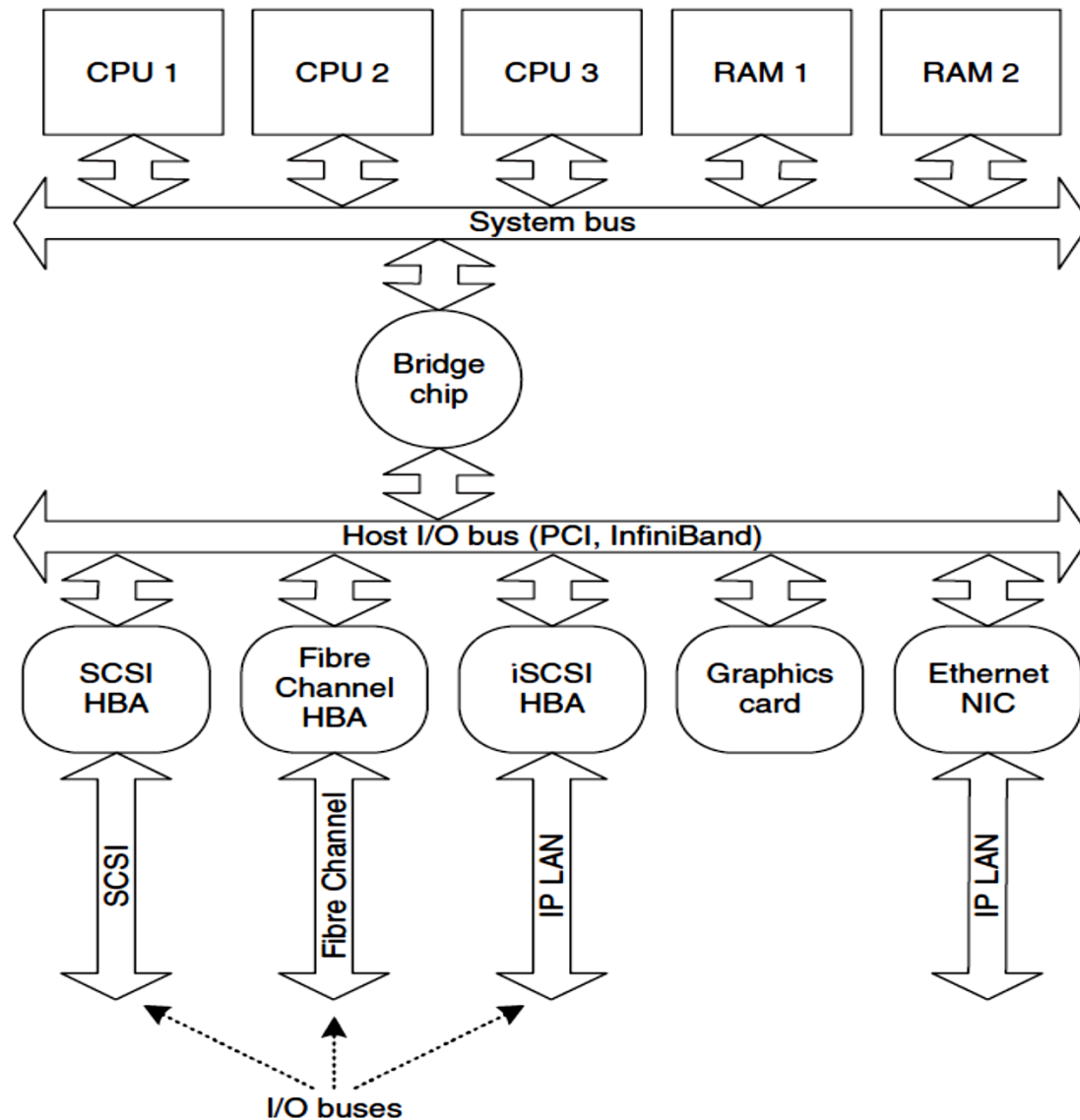


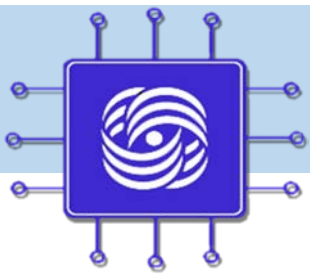
Устойчивость работоспособности ДПС

- Данные распределяют по нескольким дискам с помощью механизмов RAID и снабжают избыточными данными (блоки четности).
- На каждом физическом диске данные закодированы кодом Хемминга. Кроме этого диск оснащен подсистемой самодиагностики, которая контролирует частоту ошибок, вибрацию шпинделя и т.д. Это позволяет проактивно прогнозировать отказы диска.
- Каждый диск подсоединен к контроллеру хотя бы через две внутренние шины.
- Контроллер дисковой подсистемы может быть продублирован. Выход одного экземпляра, автоматически будет активизировать следующий экземпляр. Схема Active-Standby.
- Дублируются UPS, системы охлаждения.
- ДС подключают к разным электрическим сетям
- Сервер соединяют с ДС через несколько линий.
- Используют периодическое мгновенное копирование для защиты от логических ошибок. Например, создание мгновенной копии данных через каждый час. Тогда в случае сбоя и уничтожения какой-то таблицы, она может быть восстановлена.
- Удаленное зеркалирование используют от физического уничтожения или повреждения оборудования (катастрофоустойчивость). В сочетании с мгновенным копированием эти сервисы гарантируют сохранение и консистентность данных даже для нескольких виртуальных дисков или дисковых подсистем.
- LUN маскирование защищает от несанкционированного доступа, упрощает работу системного администратора, защищает от случайных сбоев в работе приложений серверов и их оборудования.

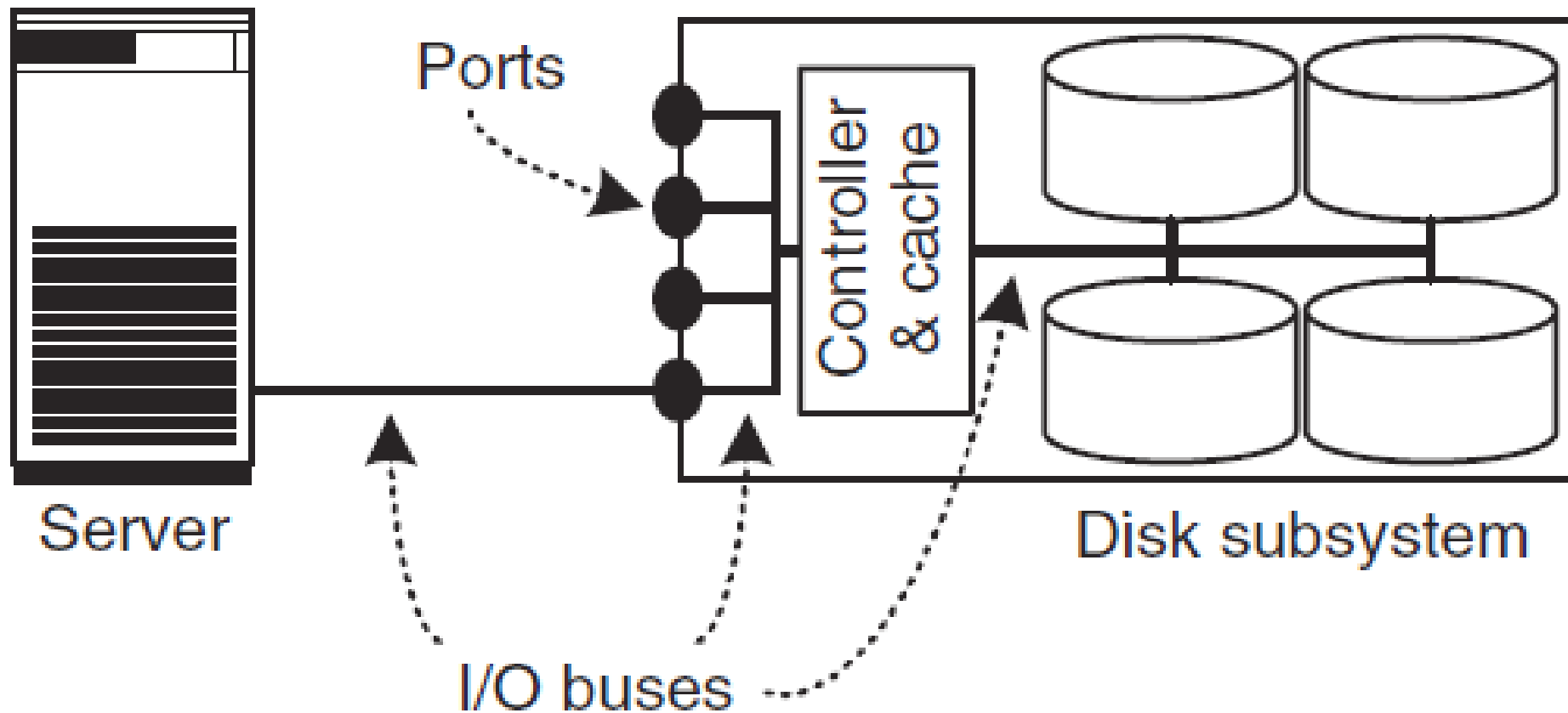


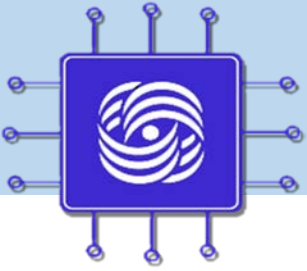
Тракт от CPU до СХД



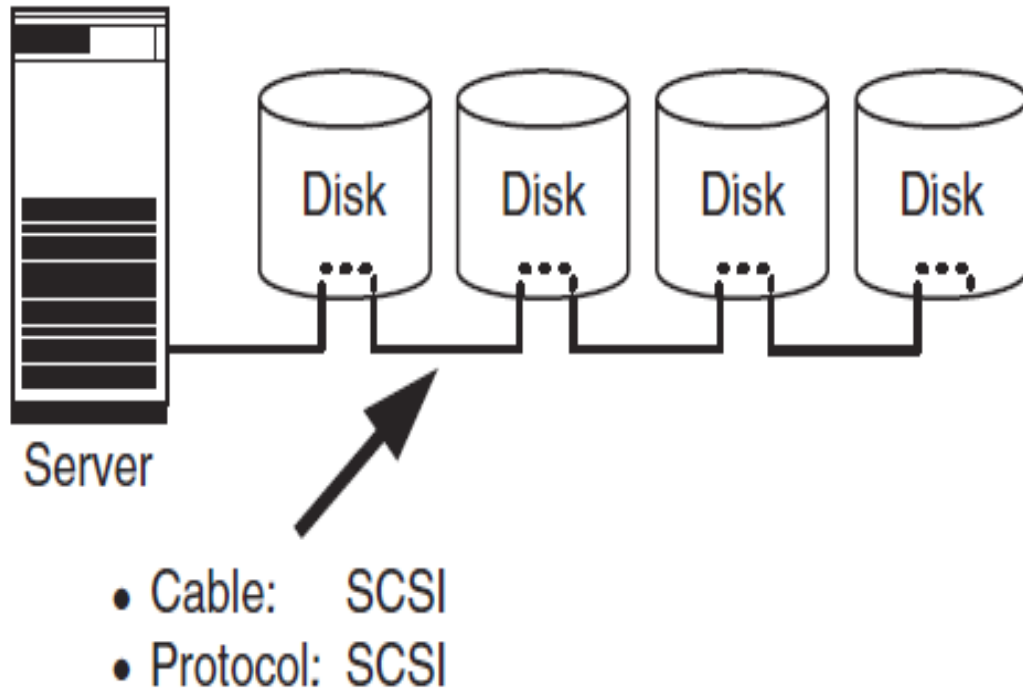


Тракт от CPU до СХД (2)



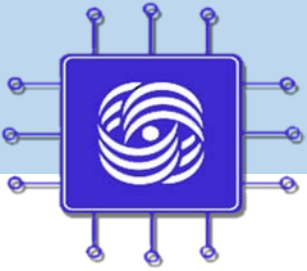


Small Computer System Interface (SCSI)

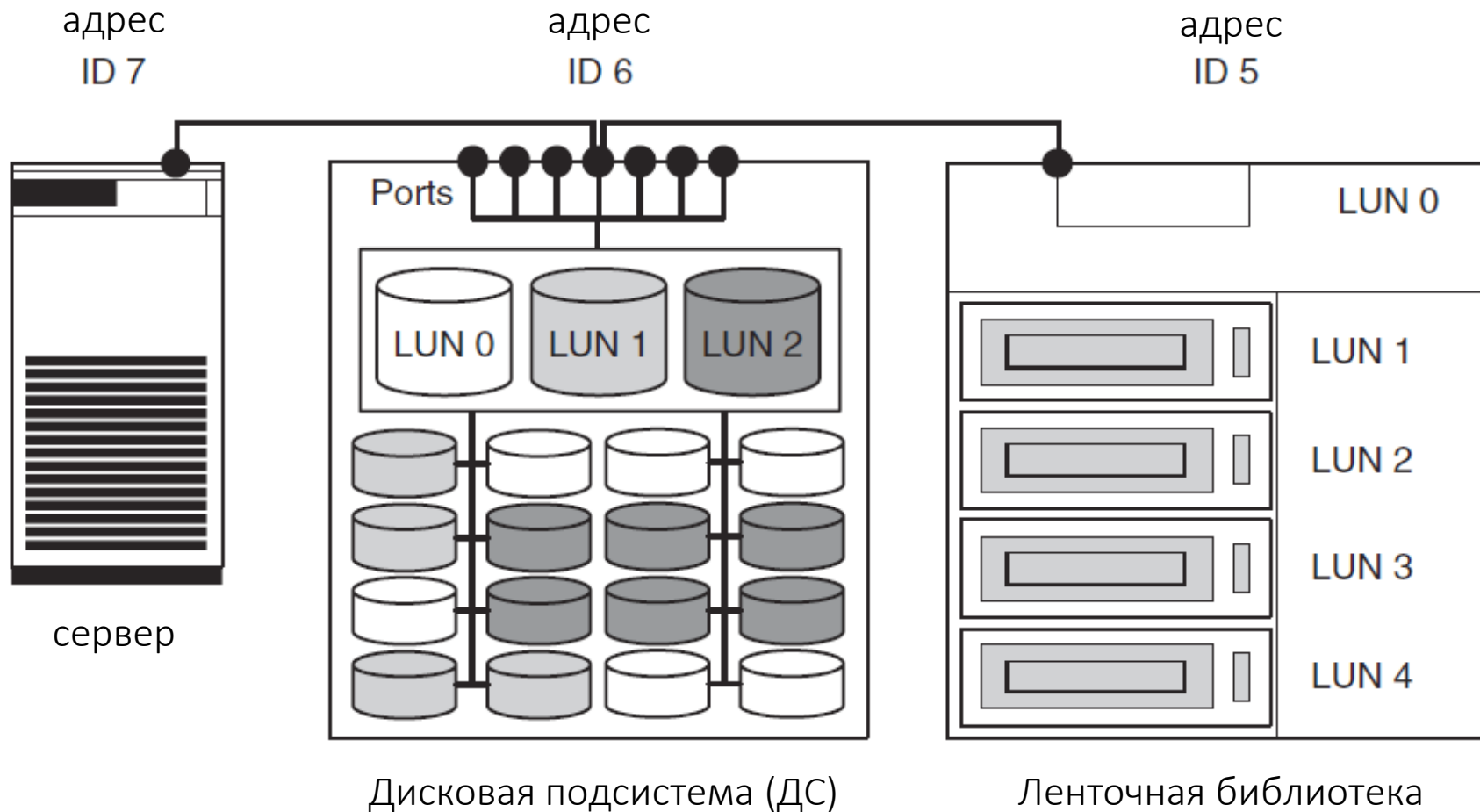


SCSI version	MByte/s MByte/s	Bus width	Max. no. of devices
SCSI-2	5	8	8
Wide Ultra SCSI	40	16	16
Wide Ultra SCSI	40	16	8
Wide Ultra SCSI	40	16	4
Ultra2 SCSI	40	8	8
Wide Ultra2 SCSI	80	16	16
Ultra3 SCSI	160	16	16
Ultra320 SCSI	320	16	16

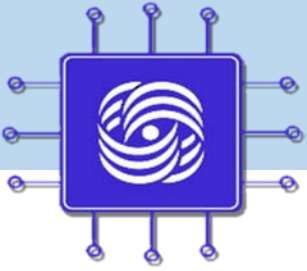
Макс. Длина шины ~ 25 м.



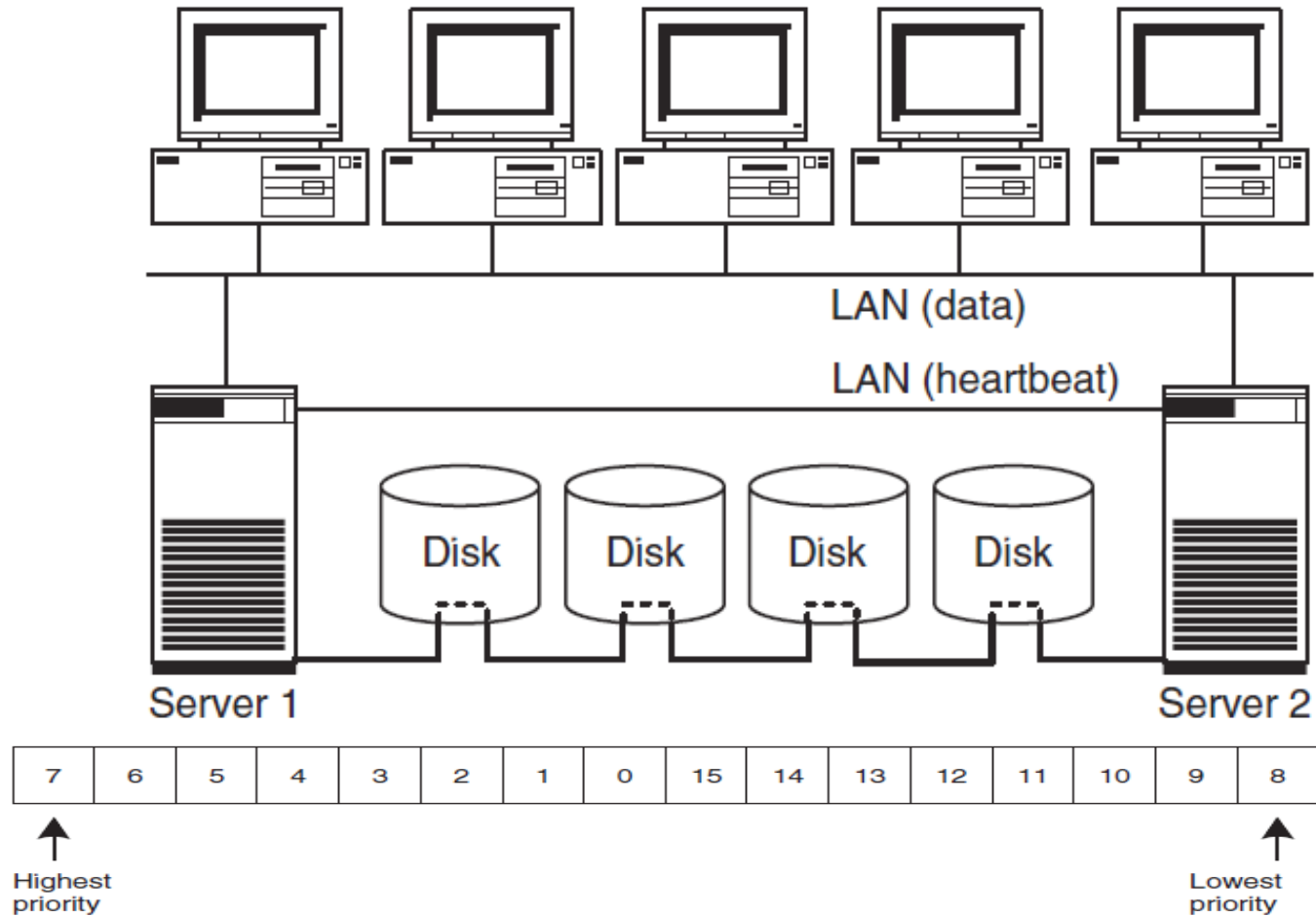
Адресация устройств на шине SCSI

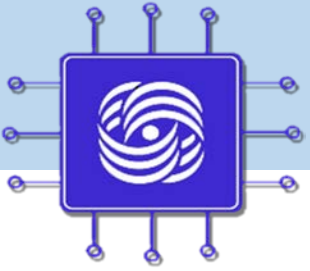


The SCSI arbitration procedure is therefore 'unfair' !

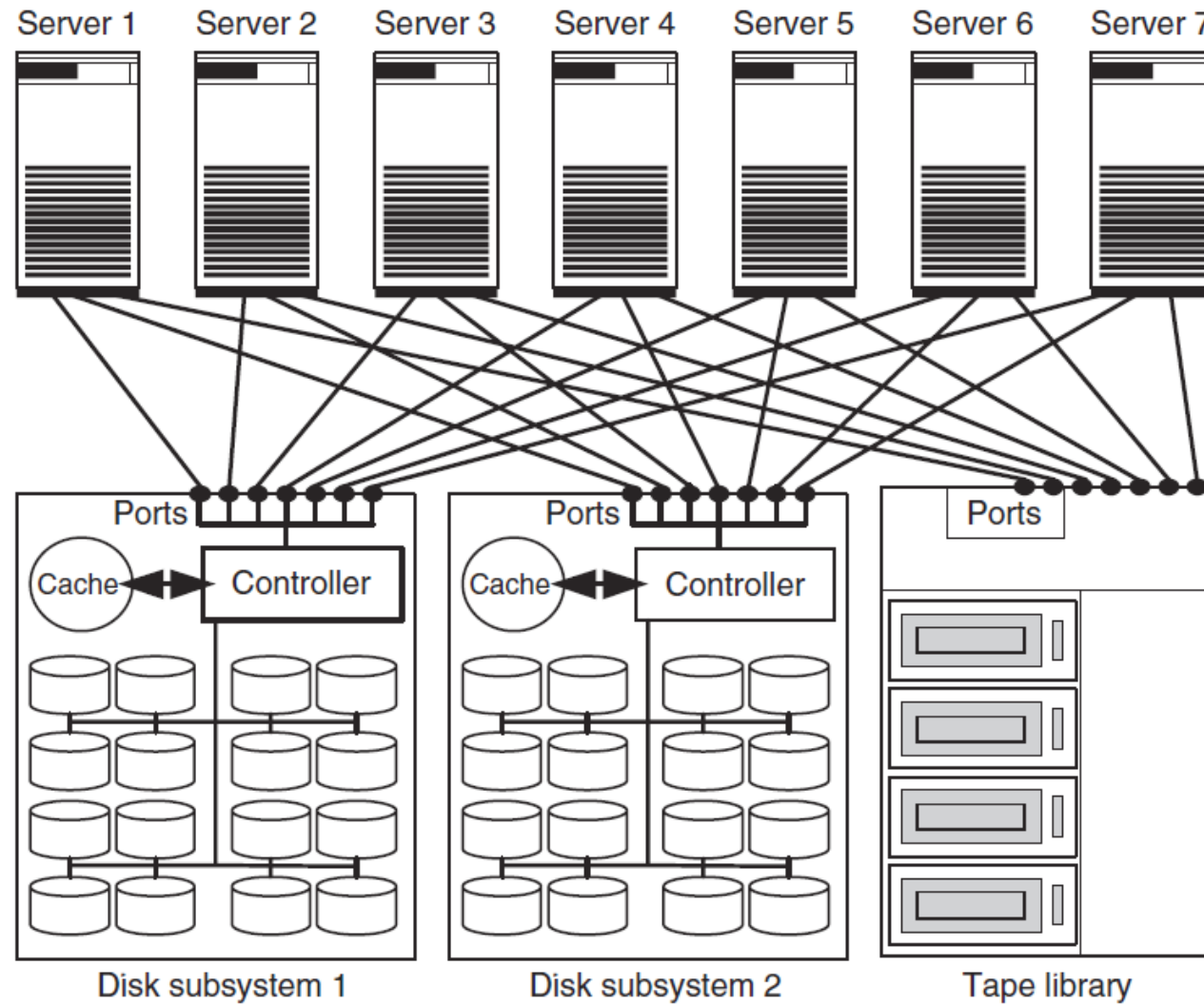


SCSI сеть хранения





SCSI и сеть хранения

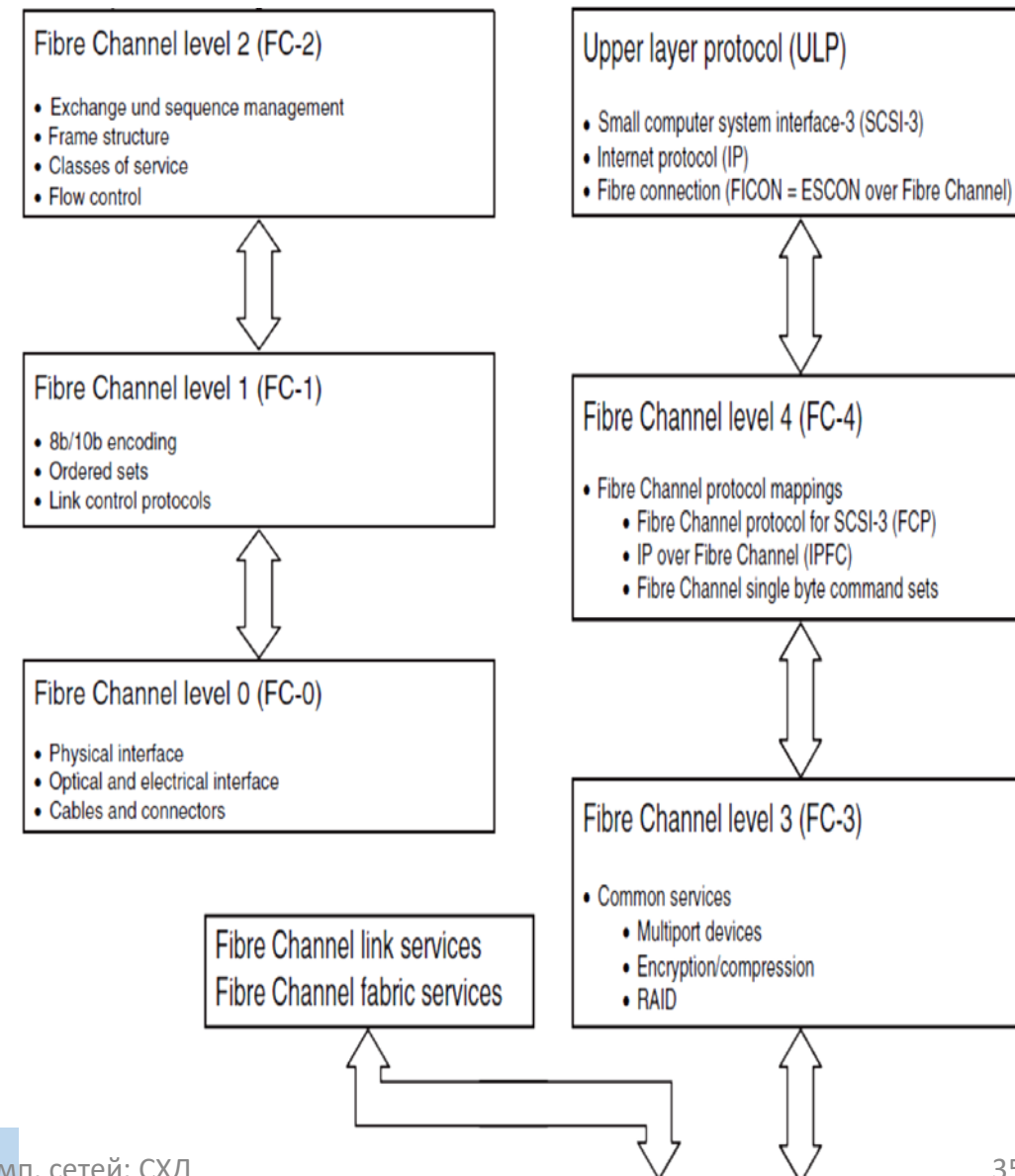


Fibre Channel

- I/O канал для LAN
 - Последовательной передачи на большие расстояния и высокой скорости.
 - Низким уровнем ошибок передачи
 - Малой задержкой передачи
 - Дешевизной реализации FC на картах HBA на серверах
- Fiber Channel сопрягаем с IPI (Intelligent Peripheral Interface), SCSI, HIPPI (High Performance Parallel Interface), ATM, IP и 802.2 (Ethernet).
- Fibre Channel - $n \times 100 \text{ МБ/с}$ при длинах канала 10 км и более, где n – число каналов. Предельная скорость передачи - 4,25 Гбод.
- В качестве физической среды может использоваться одномодовое или многомодовое оптическое волокно. Допускается применение медного коаксиального кабеля и витых пар (при скоростях до 200 МБ/с).

Стек Fiber Channel

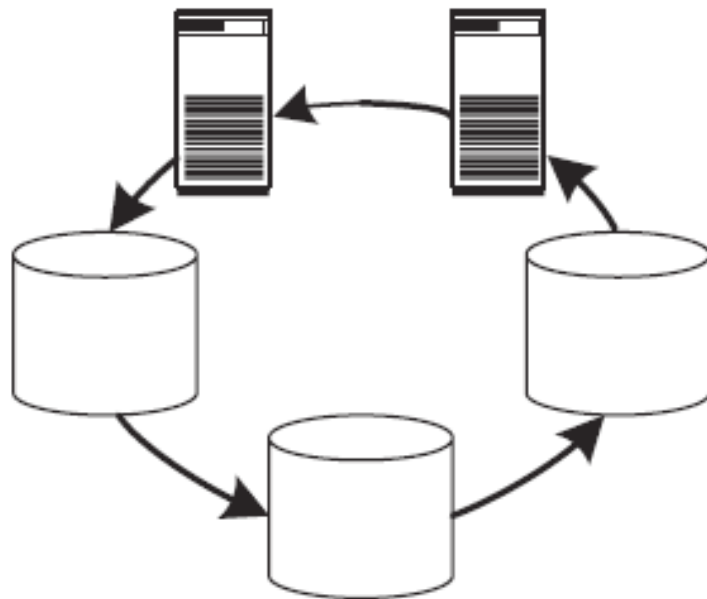
- FC - 0 определяет физические характеристики интерфейса и среды, включая кабели, разъемы, драйверы (ECL, LED, лазеры), передатчики и приемники. Вместе с FC-1 этот уровень образует физический слой.
- FC - 1 определяет метод кодирования/декодирования (8B/10B) и протокол передачи, где объединяется пересылка данных и синхронизирующей информации.
- FC - 2 (управление передачей) определяет правила сигнального протокола, классы услуг, топологию, методику сегментации, задает формат кадра и описывает передачу информационных кадров.
- FC - 3 (адресация) определяет работу нескольких портов на одном узле и обеспечивает общие виды сервиса.
- FC - 4 обеспечивает реализацию набора прикладных команд и протоколов вышележащего уровня (например, для SCSI, IPI, IEEE 802, SBCCS, HIPPI, IP, ATM и т.д.)



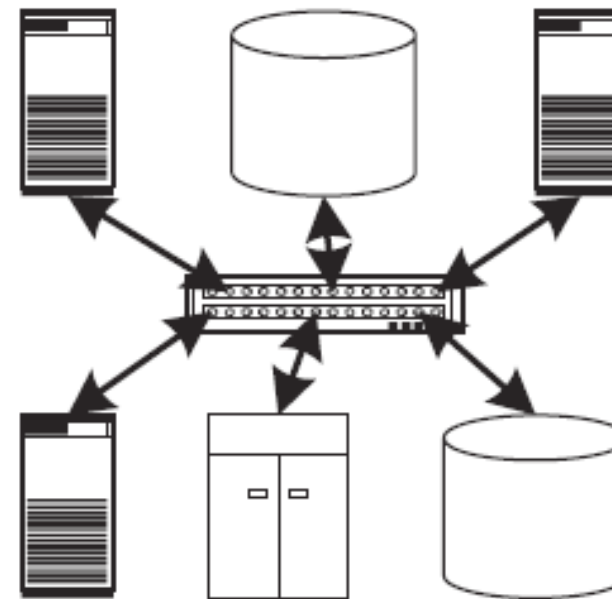
Топологии FC



Точка-Точка



Кольцо с арбитражем



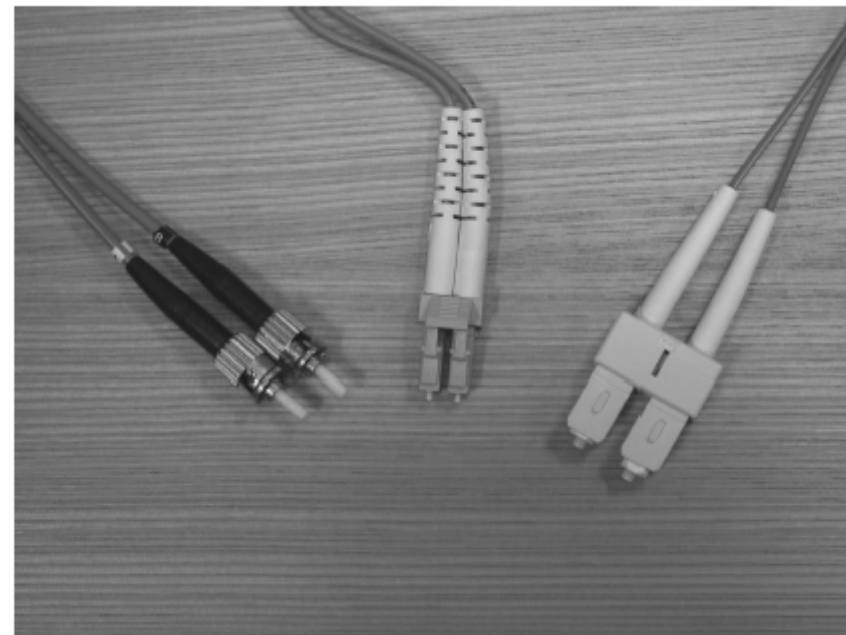
Коммутатор

Fibre Channel: типы портов

- **N-Port** — порт определяет характеристики конечного узла для соединения через коммутатор или P2P
- **F-Port** — порт для подключения к коммутатору
- **L-Port** — порт для KcA
- **NL-Port** — работает как N Port либо как L-Port.
- **FL-Port** — для подключения коммутатора в KcA
- **E-Port** — для соединения двух FC коммутаторов
- **G-Port** — может настраиваться как E либо как FL в зависимости от подключения.
- **B-Port** — для соединения двух FC коммутаторов через ATM, SDH, Ethernet или IP. Например две FC SAN могут быть соединены через WAN.

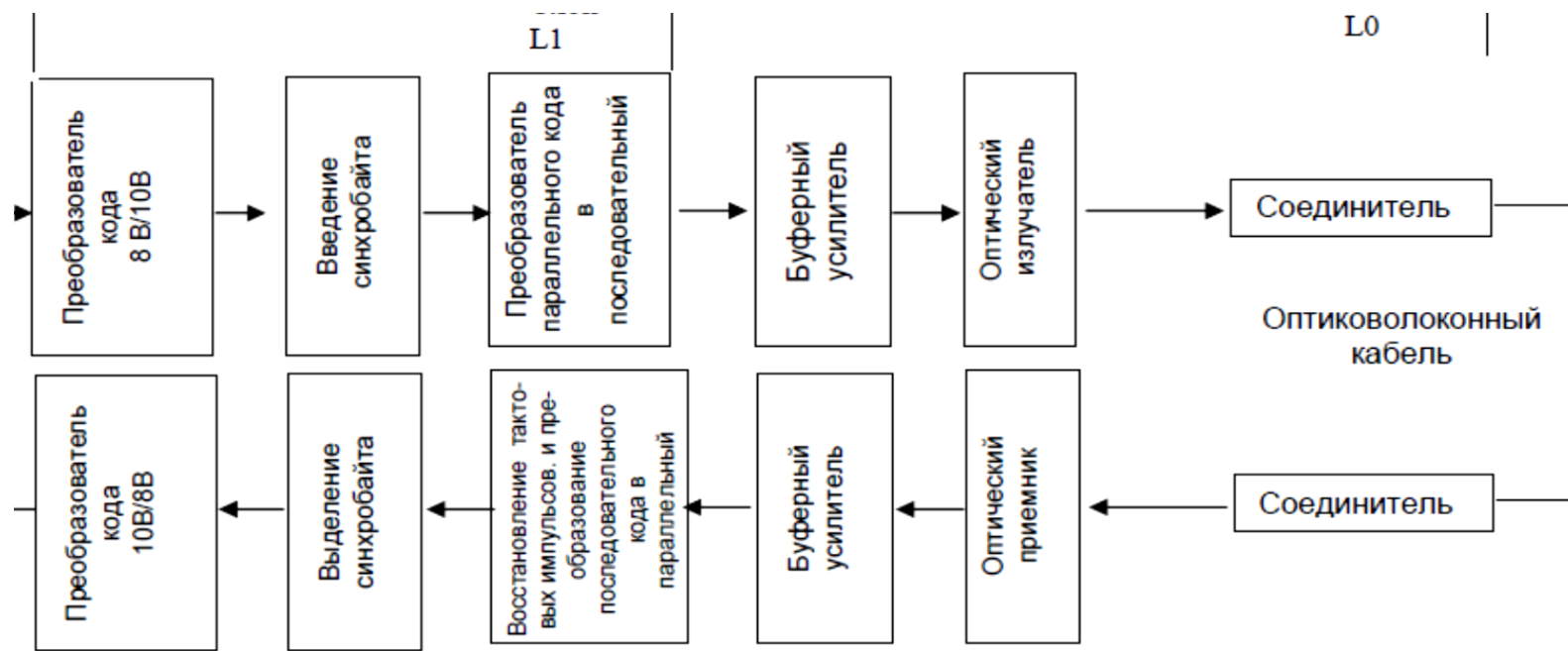
FC-0: разъемы, кабели и кодировка

- Скорость от 100 МБ/с до 10ГБ/с ожидается до 20ГБ/с в одном направлении
- Есть два вида оборудования Base 2 и Base 10
- Передача последовательная
- $BER = 10^{-12}$, т.е. для линии 100Мб/с ошибка не чаще чем раз в 16,6 мин.
Протоколы верхнего уровня для обнаружения ошибок оптимизированы под эту величину BER
- Преимущества оптического кабеля
 - Большая дистанция по сравнению с медью;
 - Нечувствительность к электромагнитному излучению;
 - Нет к электромагнитного излучения;
 - Нет электромагнитной связи между устройствами;
 - Нет проблем 'cross-talking';
 - Высокая скорость передачи.



FC-1: кодировка, упорядоченные наборы, управление линией

- 8b/10b кодирование (https://en.wikipedia.org/wiki/8b/10b_encoding)
- Transmission words
 - Data word: SOF, 4 bytes, EOF
 - Ordered set: EOF, K28.5, SOF
- Управление линией



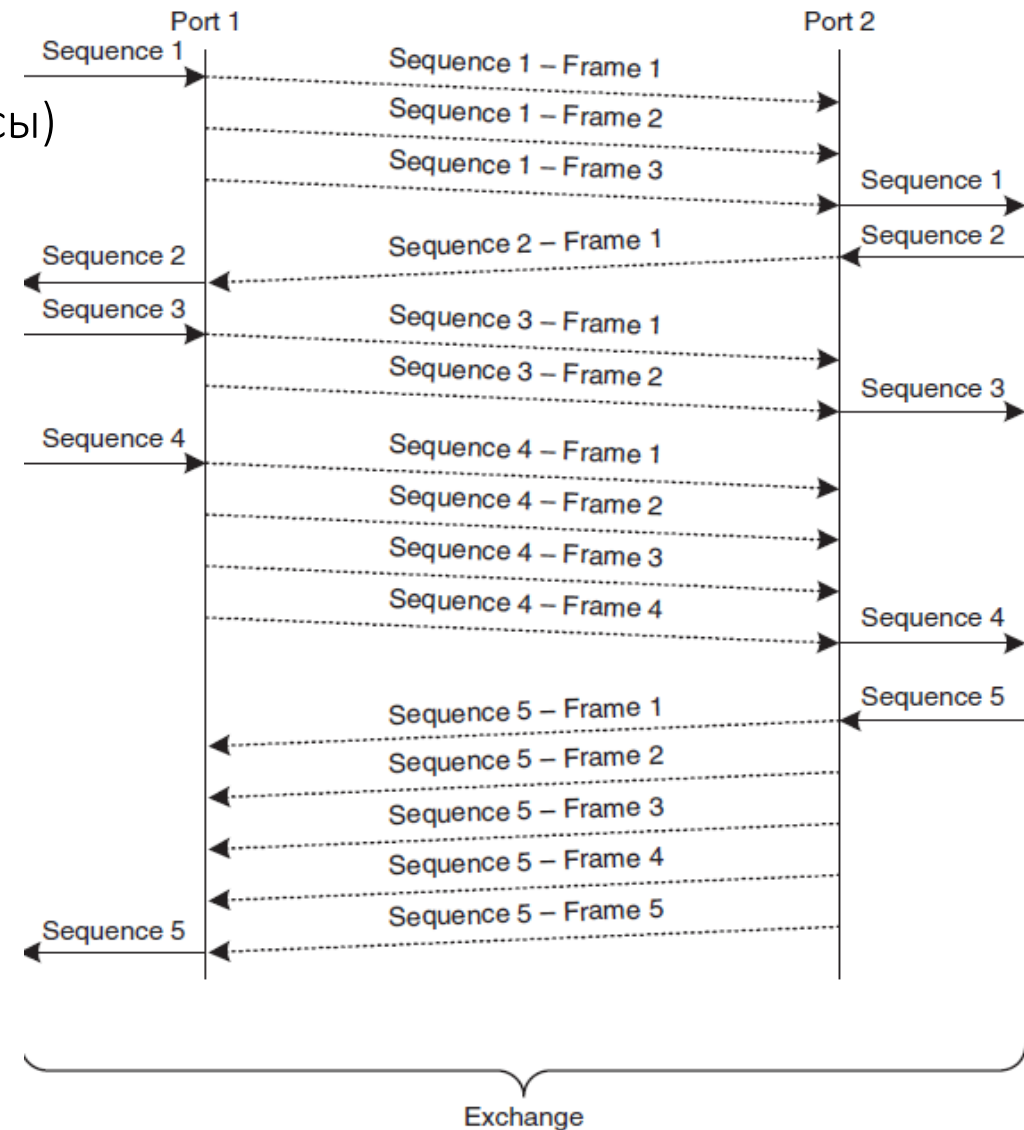
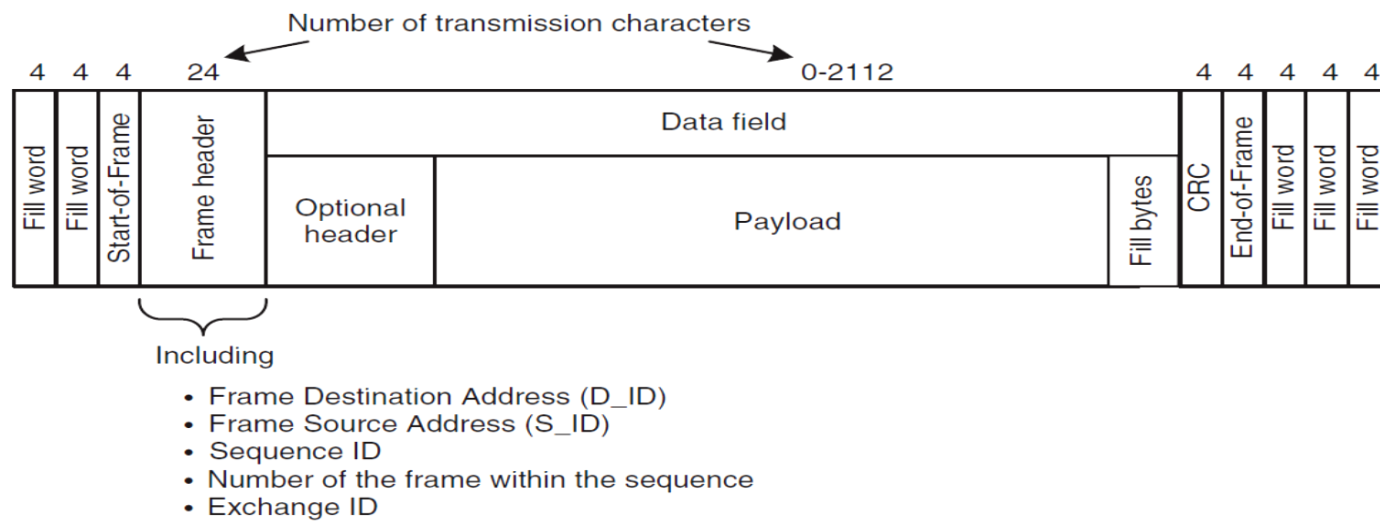
Т а б л и ц а А.1 –Специальные коды

Обозначения специальных кодов	TP –	TP+
	abcdei fghj	abcdei fghj
K28.0	001111 0100	110000 1011
K28.1	001111 1001	110000 0110
K28.2	001111 0101	110000 1010
K28.3	001111 0011	110000 1100
K28.4	001111 0010	110000 1101
K28.5	001111 1010	110000 0101
K28.6	001111 1000	110000 1001
K23.7	111010 1000	000101 0111
K27.7	110110 1000	001001 0111
K29.7	101110 1000	010001 0111
K30.7	011110 1000	100001 0111

Асинхронная линия последовательной передачи (АЛПИ)

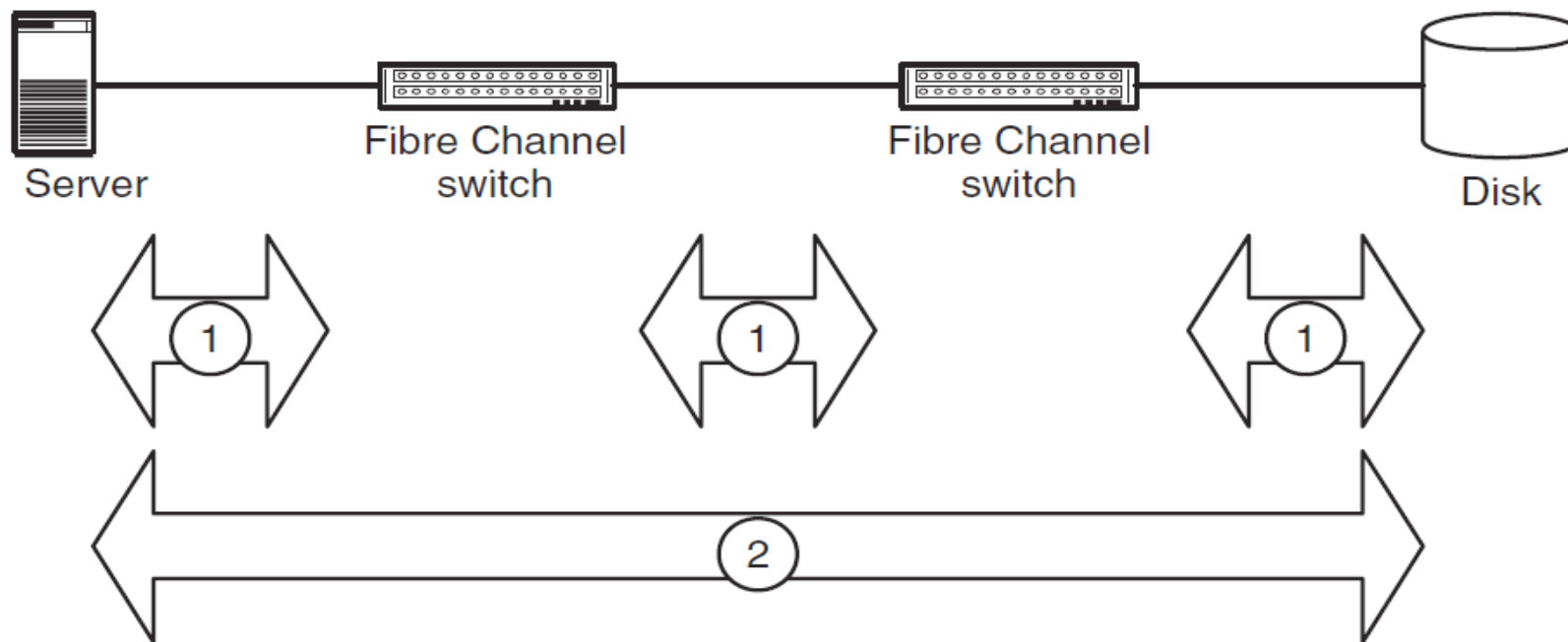
FC-2: передача данных

- Структура передачи данных данных
 - Exchange – сессия между логическими сущностями (процессы)
 - Sequence – последовательность кадров
 - Frame – управления и данных (2 112 Б)
- Управление потоком
- Классы сервиса



FC-2: Управление потоком

- Кредитная схема
- E2E flow control (2)
- Link flow control (1)



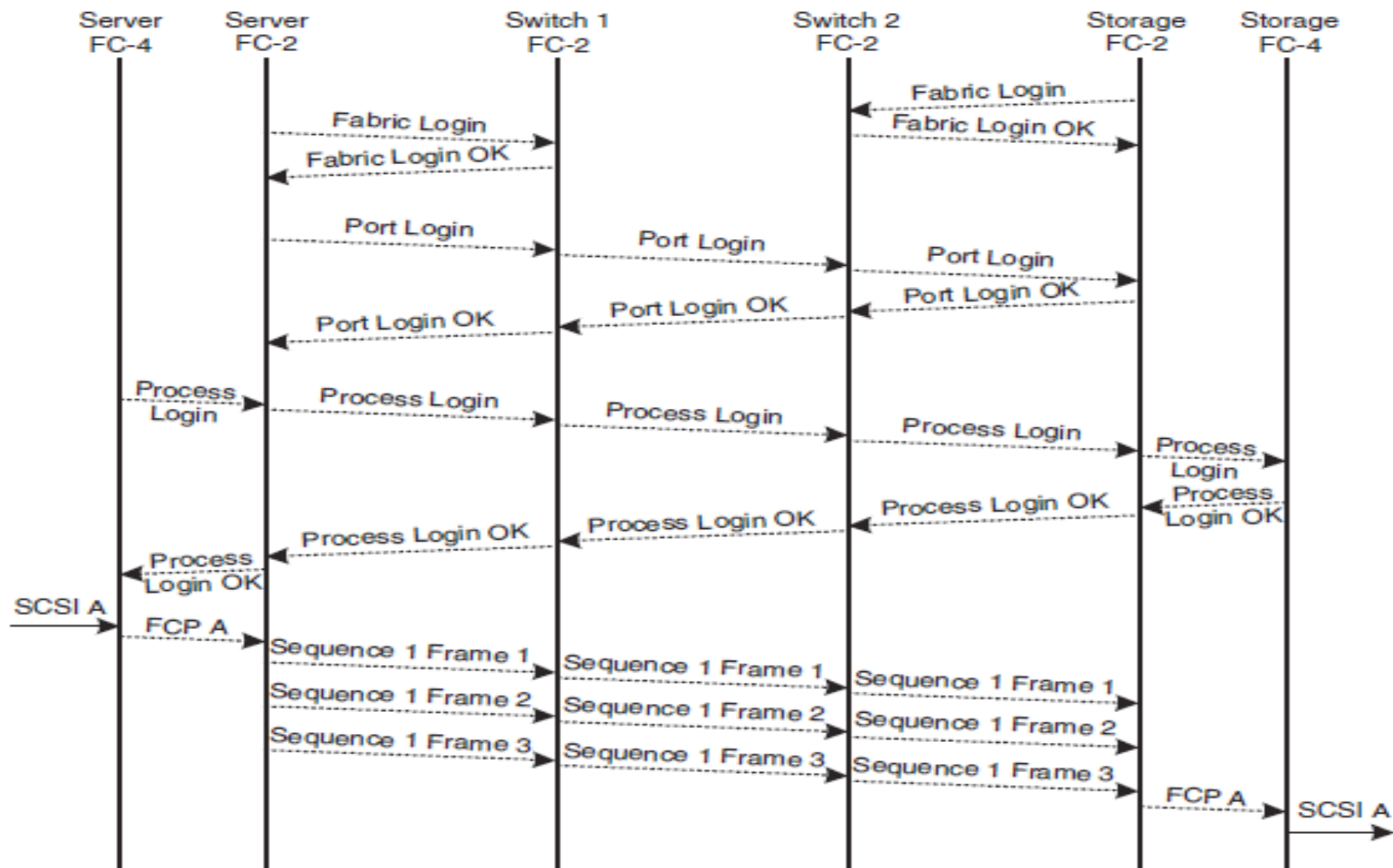
FC-2: Классы обслуживания

- Класс 1** Соединение точка-точка (end-to-end) между портами типа `n_port` через коммутацию каналов. Класс удобен для аудио и видео приложений, например, видеоконференций. После установления соединения используется вся доступная полоса пропускания канала. При этом гарантируется, что кадры будут получены в том же порядке, в каком они были посланы. Есть управление потоком.
- Класс 2** Без установления соединения с коммутацией пакетов, гарантирующий доставку данных. Порт может взаимодействовать одновременно с любым числом портов типа `n_port` в режиме дуплекс. Не гарантируется порядок доставки кадров, кроме соединения R2P или KcA. Есть управление потоком. Этот класс характерен для локальных сетей, где время доставки данных не является критическим.
- Класс 3** Обмен дейтограммами без установления соединения и без гарантии доставки. Есть управление потоком. Применяется для каналов SCSI.
- Класс 4** Обеспечивает выделение определенной доли пропускной способности канала с заданным качеством обслуживания (QoS). Только для топологии матрица с `n_port`. Гарантируется порядок доставки кадров.
- Класс 5** Регламентирующие документы находятся в процессе подготовки.
- Класс 6** Предусматривает групповое-обслуживание с коммутацией.

FC-3: сервисы

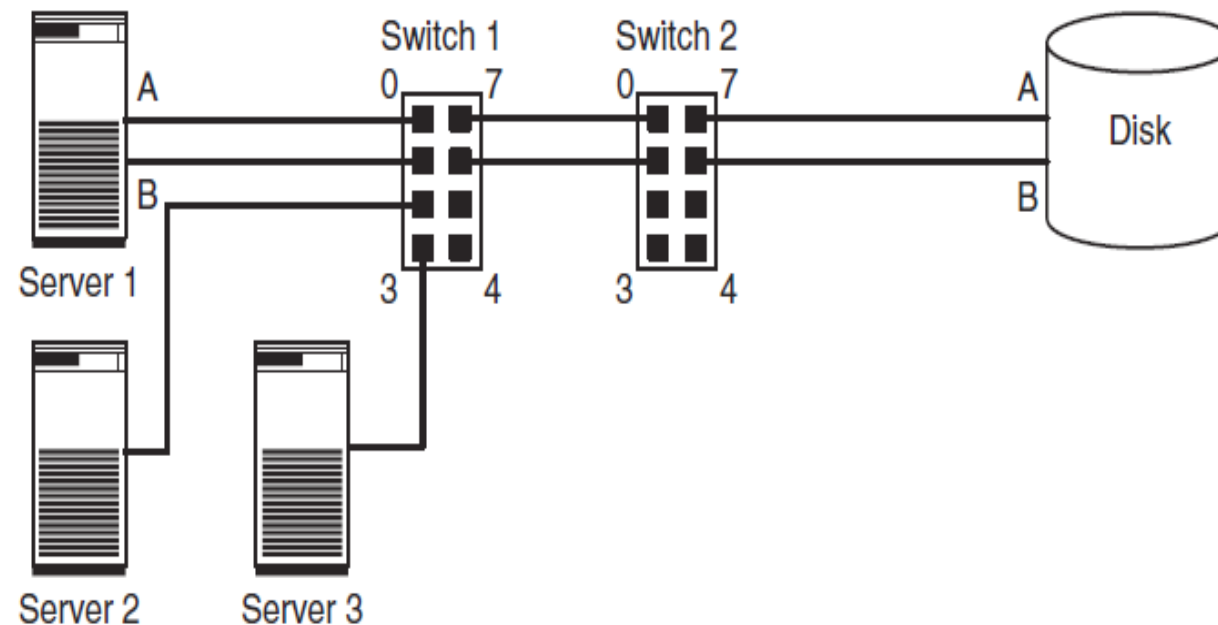
- Кадры не подтверждают (не надежный сервис), есть только управление потоком на уровне линии, но не e2e.
- Распределение кадров по маршрутам между много портовыми устройствами для увеличения пропускной способности
- Формирование логических групп маршрутов для управления переполнением на маршруте или сбоя, чтобы не загружать этими проблемами верхние уровни стека.
- Компрессия передаваемых данных (на HBA)
- Шифрование данных (на HBA)
- Зеркалирование (пока в планах)

Службы линии: идентификация и адресация



Адресация

- Имена и адреса в FC
- У всех устройств FC сети есть 64 бит. имена
- WWN vs FCN
- WWN: WWPN. WWNN
- FLOG – 24 bit port address
- S_ID vs D_ID
- KcA 8 bit AL_PA (Arbitrated Loop Physical Address)



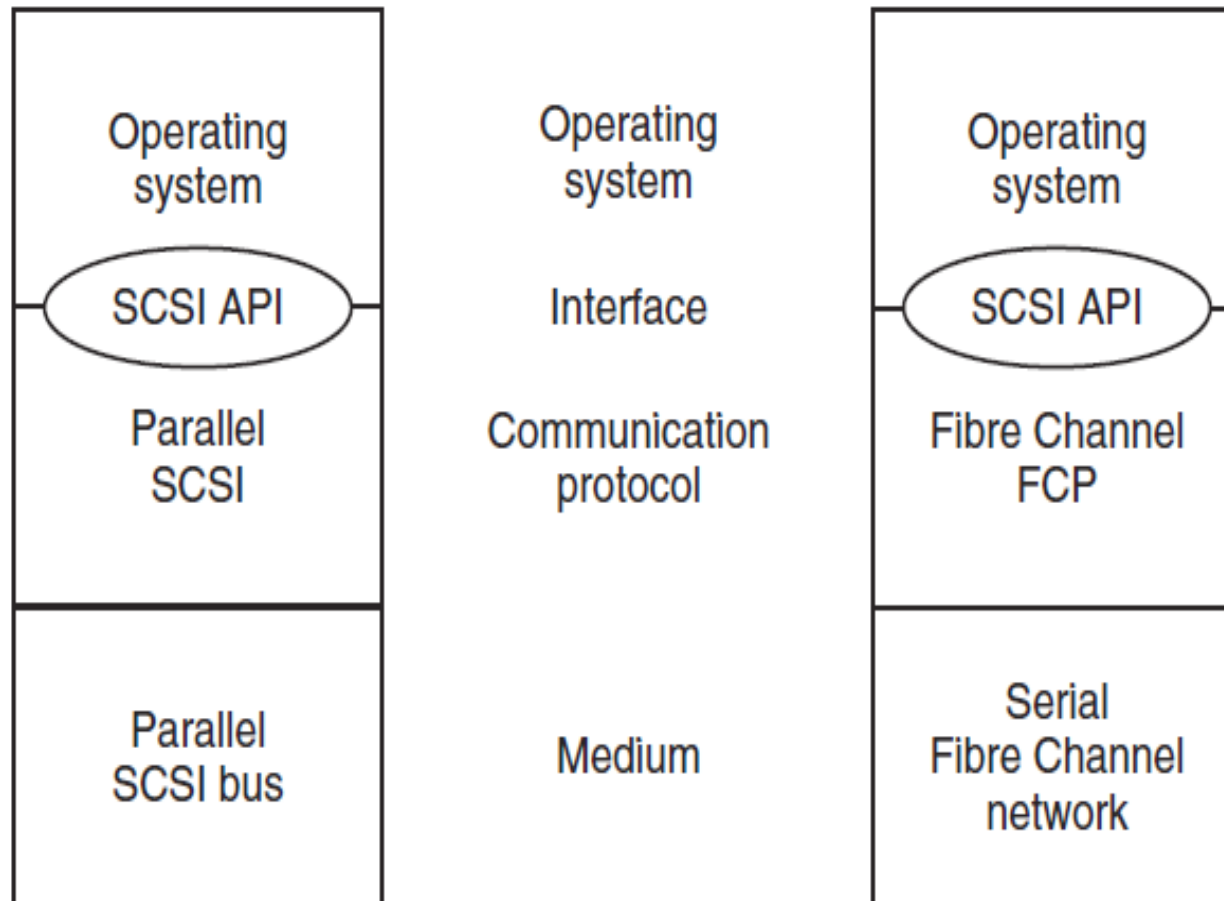
Port_ID	WWPN	WWNN	Device
010000	20000003 EAFE2C31	2100000C EAFE2C31	Server 1, Port A
010100	20000003 C10E8CC2	2100000C EAFE2C31	Server 1, Port B
010200	10000007 FE667122	10000007 FE667122	Server 2
010300	20000003 3CCD4431	2100000A EA331231	Server 3
020600	20000003 EAFE4C31	50000003 214CC4EF	Disk, Port B
020700	20000003 EAFE8C31	50000003 214CC4EF	Disk, Port A

Сервисы коммутационной среды

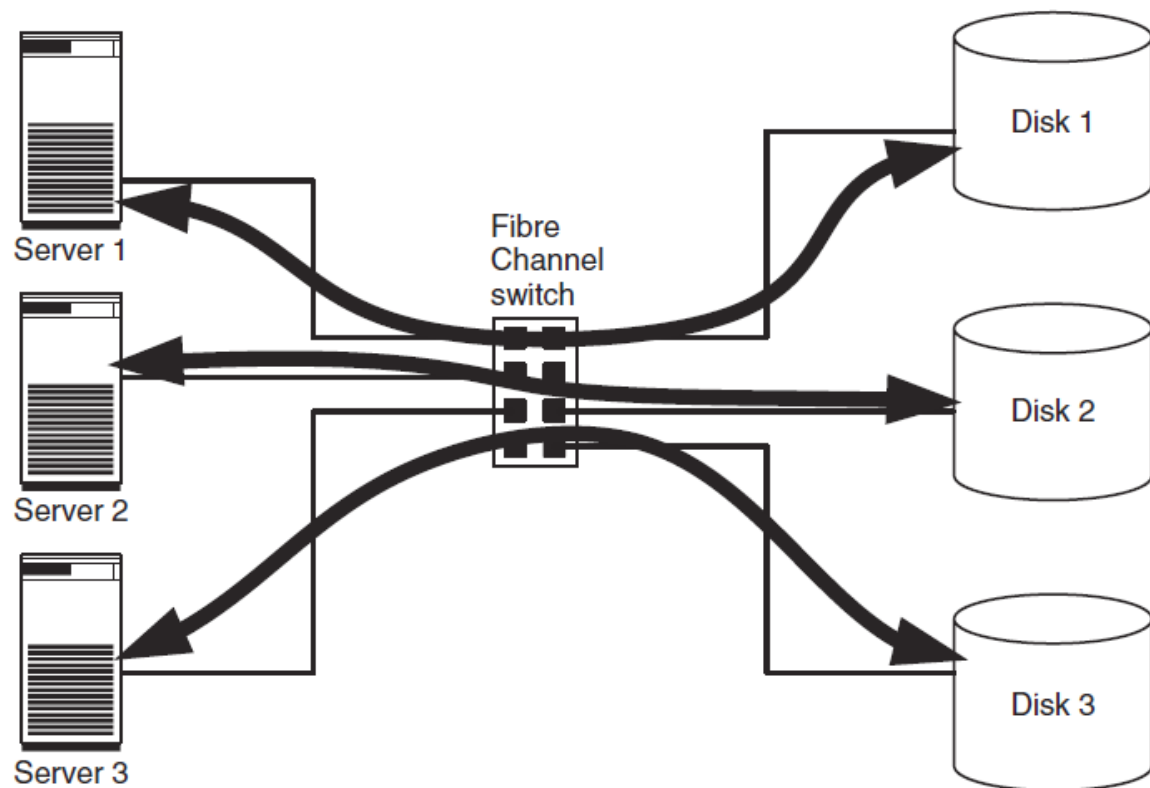
- Сервисы КС нужны для управления инфраструктурой и потоками в FC
- Все сервисы реализуют определённые сервера, которые имеют строго определённые адреса.
- FLOG сервер отвечает за обработку всех входящих fabric login request
- За всеми изменениями в FC сети следит fabric controller
- Name server – отвечает за БД все имен N_Port'ов

Address	Description
0xFF FF FF	Broadcast addresses
0xFF FF FE	Fabric Login Server
0xFF FF FD	Fabric Controller
0xFF FF FC	Name Server
0xFF FF FB	Time Server
0xFF FF FA	Management Server
0xFF FF F9	Quality of Service Facilitator
0xFF FF F8	Alias Server
0xFF FF F7	Security Key Distribution Server
0xFF FF F6	Clock Synchronisation Server
0xFF FF F5	Multicast Server
0xFF FF F4	Reserved
0xFF FF F3	Reserved
0xFF FF F2	Reserved
0xFF FF F1	Reserved
0xFF FF F0	Reserved

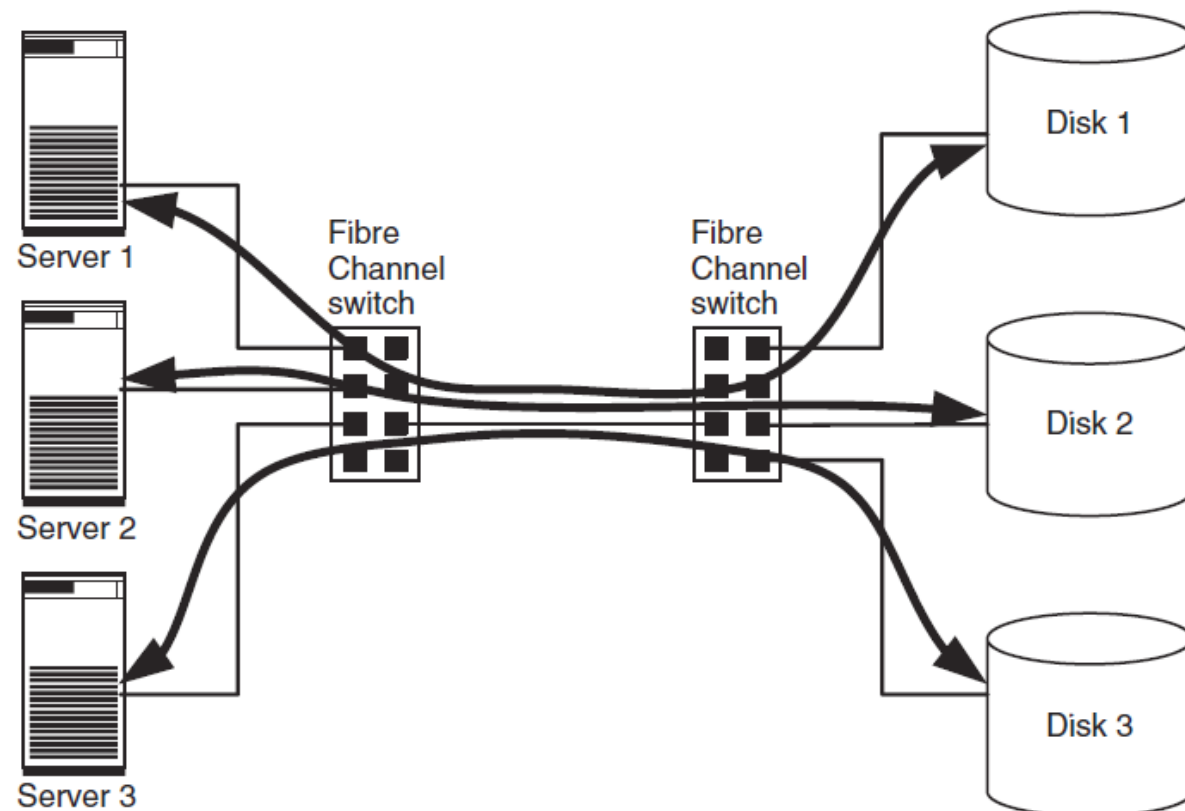
FC-4: ULP



Топология коммутационной среды (fabric)

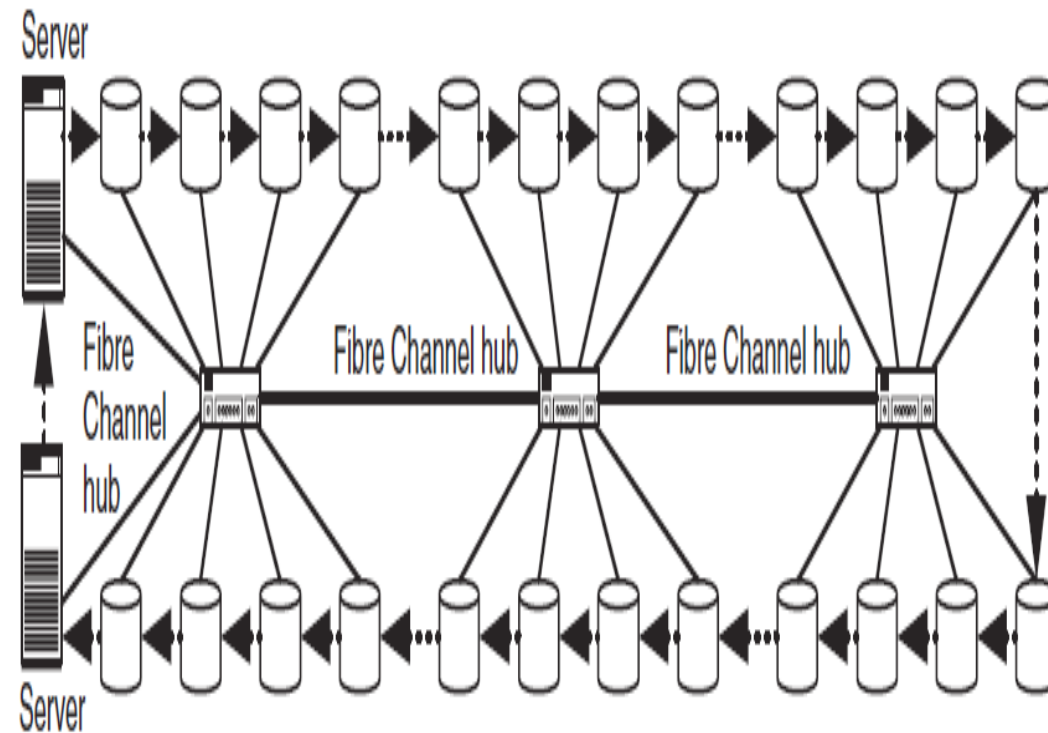
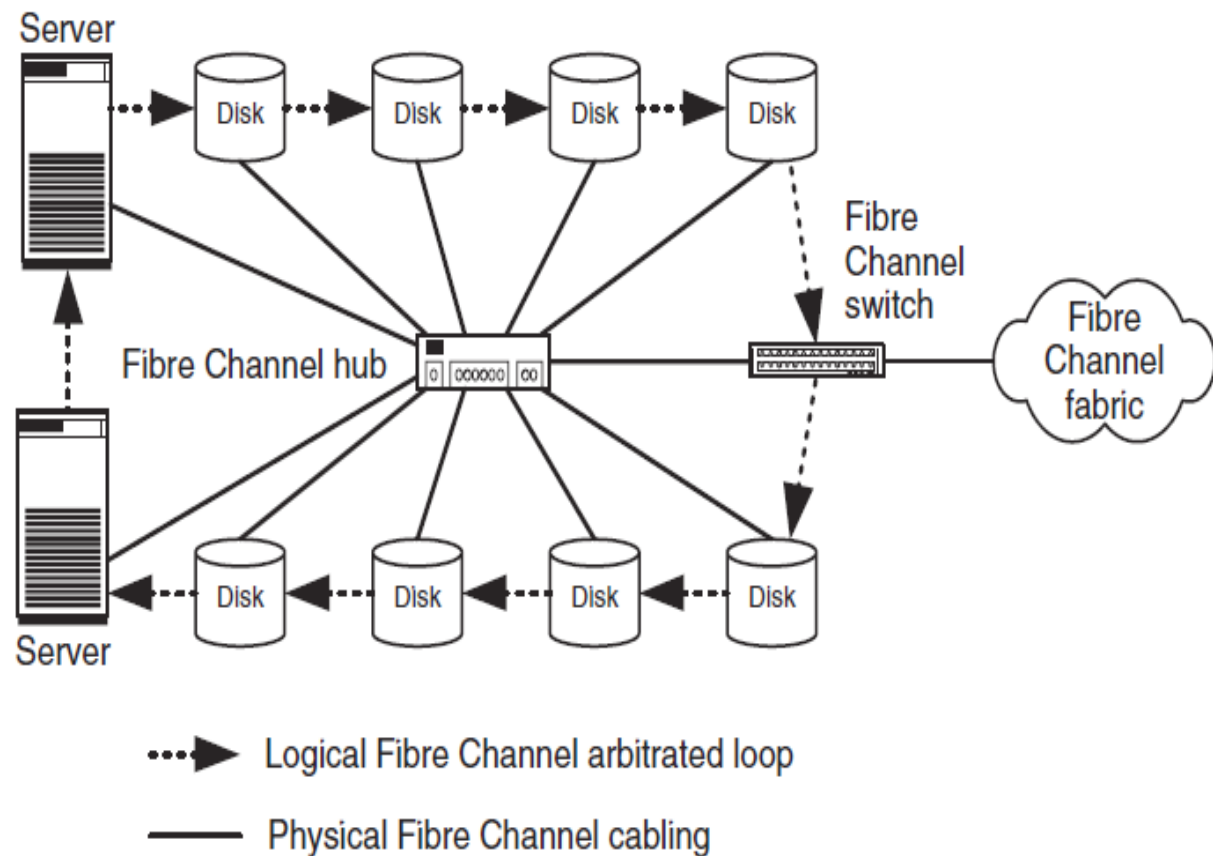


Несколько соединений на полной скорости.

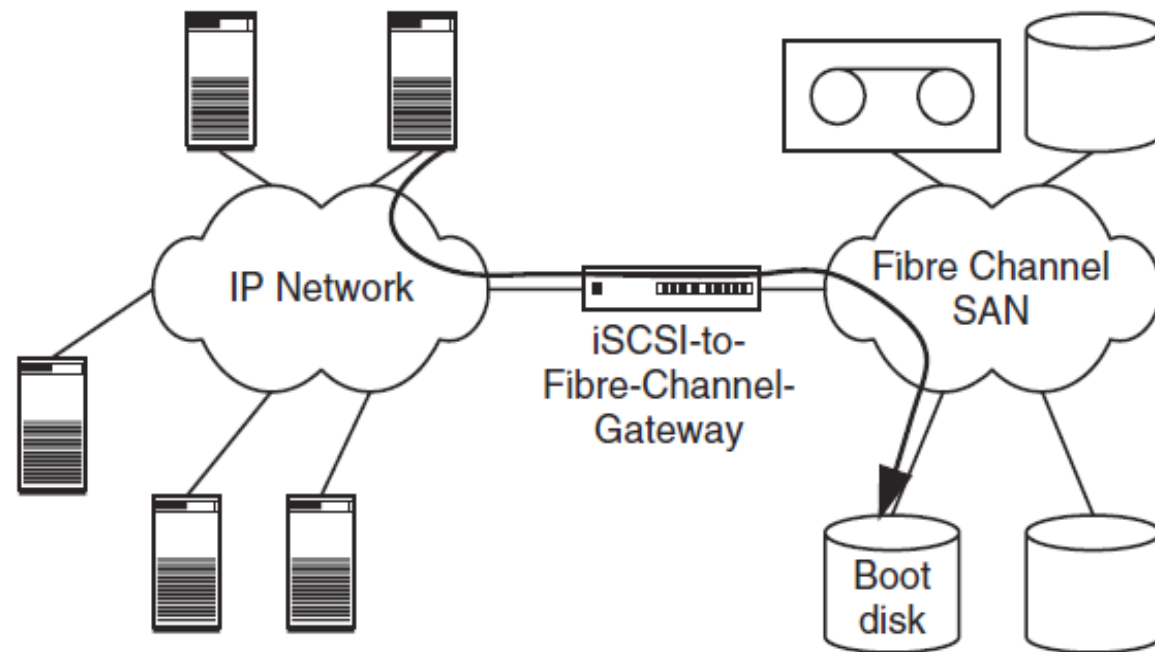
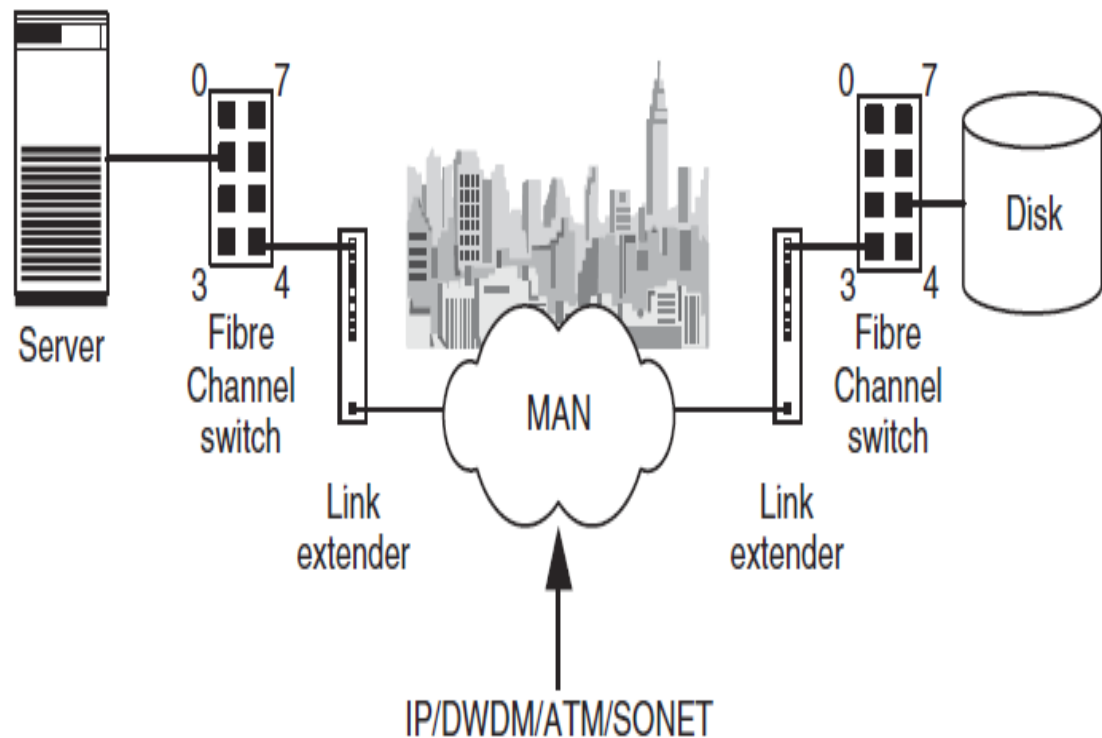


Несколько соединений через один ISL.

КсА топология



Metro SAN и IP_FC SAN



В качестве заключения: Новые горизонты



Тенденции рынка информационных технологий



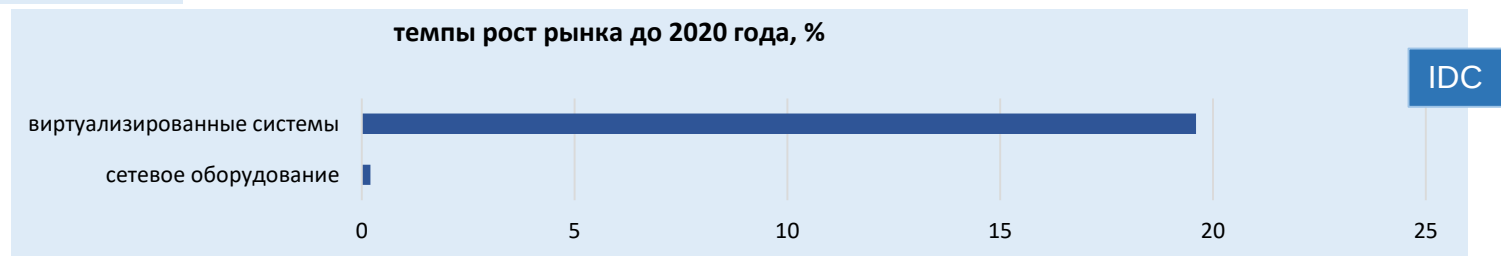
Мобильность



Виртуализация

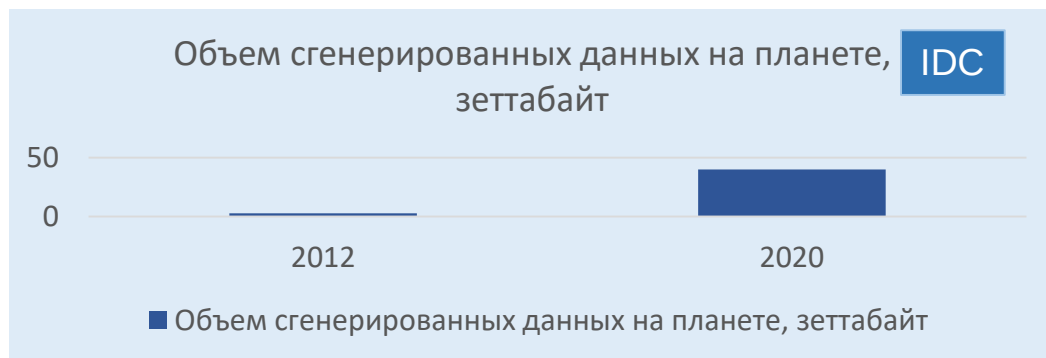


Консолидация инфраструктуры



Тенденции рынка телекоммуникационных технологий

Big Data



Всего с начала 2010 г. объем хранимых данных вырос в 50 раз

Центры обработки данных

- В 2014 году объем мирового рынка колокации в ЦОДах составил \$22,8 млрд. Общая площадь размещения оборудования достигла 10,13 кв. км. (451 Research)
- Общее число дата-центров всех типов в 2017 г. вырастет до 8,6 млн (IDC)

Телеком

- Каждый из пользователей глобальной сети генерирует больше трафика, чем вся Всемирная паутина 30 лет назад
- В 2014 году интернет-трафик вырос, по сравнению с 1984 годом, в 2,7 миллиардов раз (Cisco)

Изменение бизнес-модели

РЕНТАБЕЛЬНОСТЬ БИЗНЕСА



Промышленное производство
Энергодобыча
Тяжелая и легкая промышленность

РЕНТАБЕЛЬНОСТЬ БИЗНЕСА



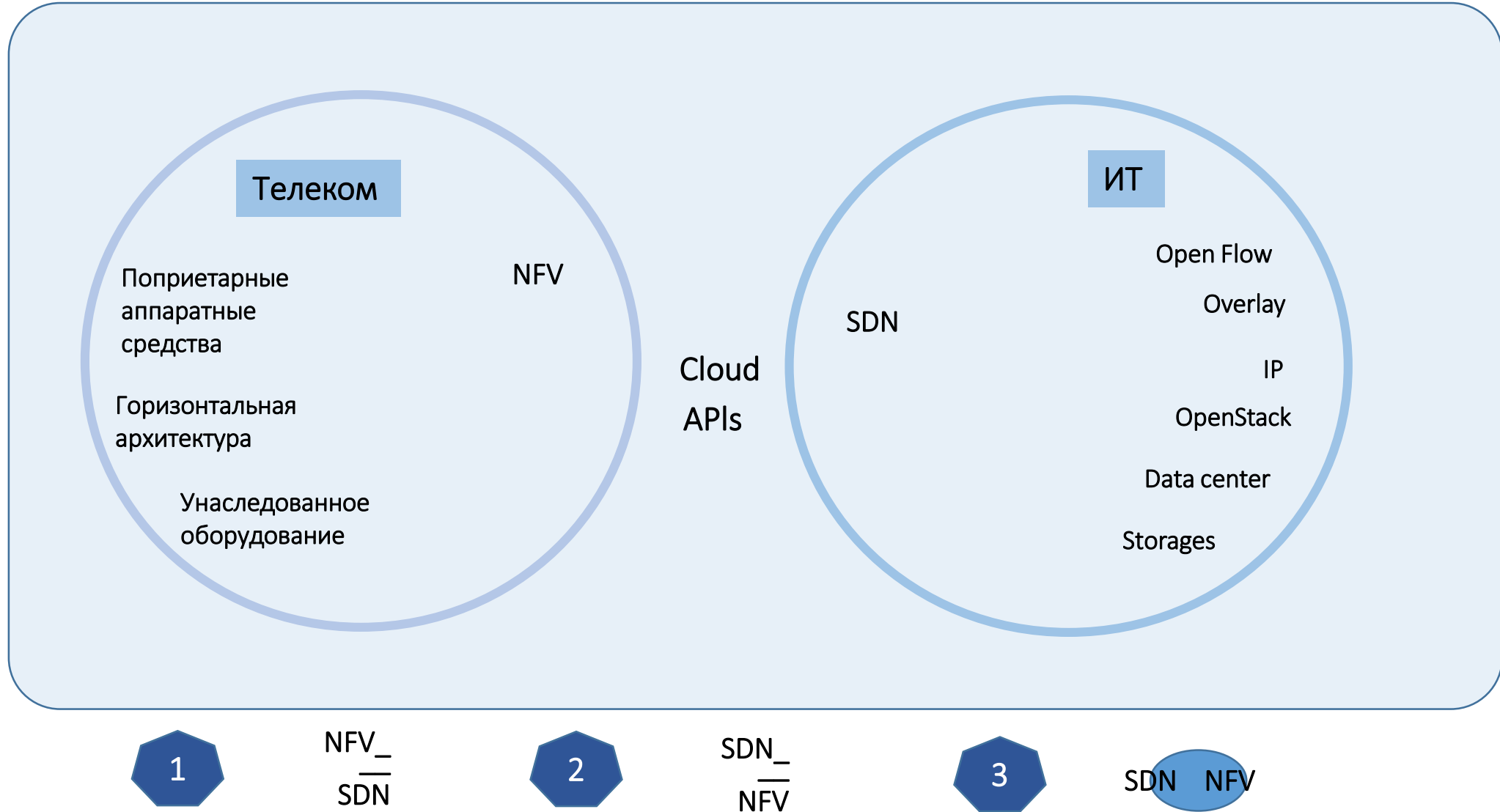
Оператор связи
Интернет провайдер

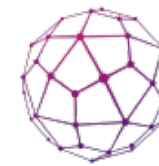
**Доступ к «транспорту» должен быть бесплатным,
платным должен стать контент и услуги.**

Информация о пользователе:

- Использование: посещаемые сайты, звонки и сообщения (включая тип сообщений и их частоту);
 - География: где находится мобильное устройство в конкретный момент (уровень точности может разниться от района к району);
 - Демография: доход домохозяйства, число и возраст проживающих детей;
 - Уровень дохода: тарифный план, история платежей, паттерн совершения покупок;
 - Мультиплатформенность: использование данных на разных устройствах и типах подключения к сети (3G, WiFi и т.п.).
- 2011 - AT&T – запуск подразделения AdWorkds: поддержка целевой рекламы в web, мобильной среде и ТВ.
 - 2013 – AdWorks открывает доступ к анализу данных 70 млн.пользователей.
 - 2012 – Verizon - запуск инициативы Precision Market Insights – доступ к мобильным данным пользователей для маркетинговых и рекламных компаний.

Симбиоз SDN и NFV





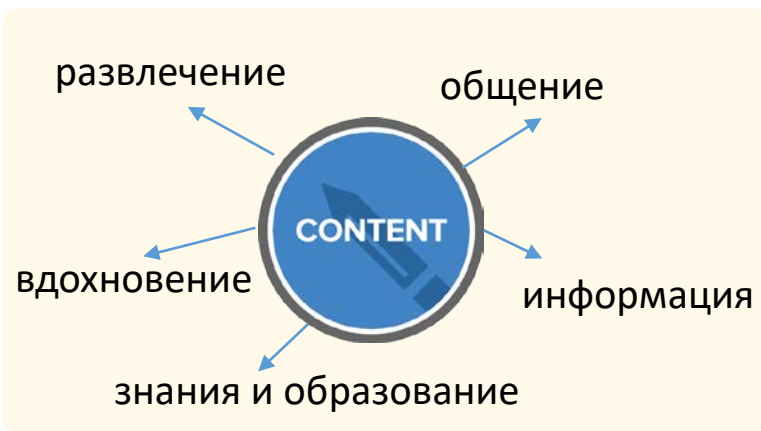
ЦЕНТР
ПРИКЛАДНЫХ
ИССЛЕДОВАНИЙ
КОМПЬЮТЕРНЫХ
СЕТЕЙ

Сети CDN следующего поколения

Руслан Смелянский
МГУ, ЦПИКС

Роль контента в коммуникациях

Любое определение контента покрывает одну или несколько зон



Девид Эриксон (Facebook):

Важные, убедительные, своевременные и ценные **знания** и /или **развлечение**.

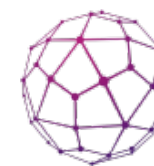
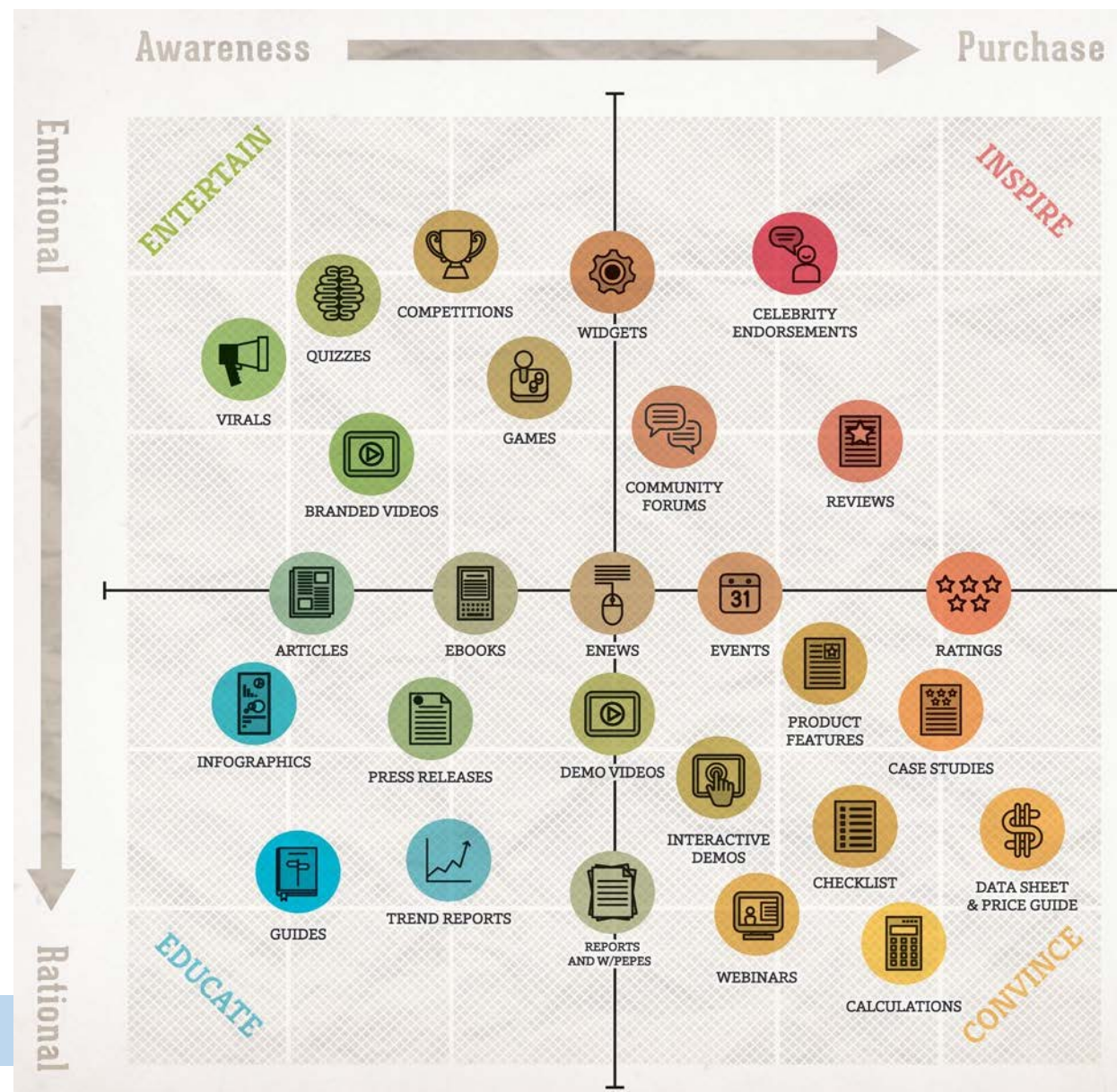
Джо Пулицци (G+):

Важная информация, которая информирует, вовлекает или развлекает.

Интересно:

Любой контент может быть использован для создания другого контента.

Типы контента и их распространение



ЦЕНТР
ПРИКЛАДНЫХ
ИССЛЕДОВАНИЙ
КОМПЬЮТЕРНЫХ
СЕТЕЙ

Роль контента в коммуникациях

Шок контента: Экспоненциально увеличивающиеся объемы контента превышают наши ограниченные способности для его потребления



Средний уровень внимания уменьшился с 12 секунд в 2000 до 8 секунд, при этом средний уровень внимания золотых рыбок составляет 9 секунд

Как технологии влияют на контент

- Новые типы контента:



Знания и образование



Coursera, e-learning



Информация

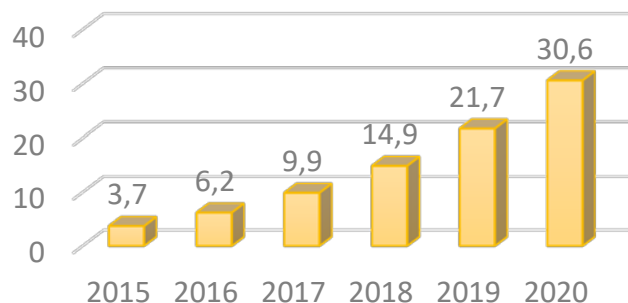


Мобильные устройства

- Новые объемы данных:

Рост мобильного трафика

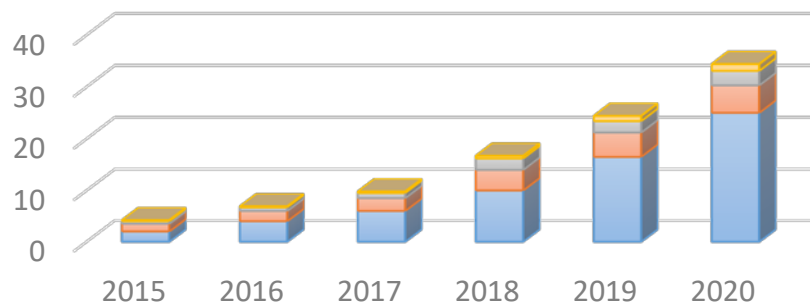
Экзабайт в месяц



Source: Cisco VNI Mobile, 2016

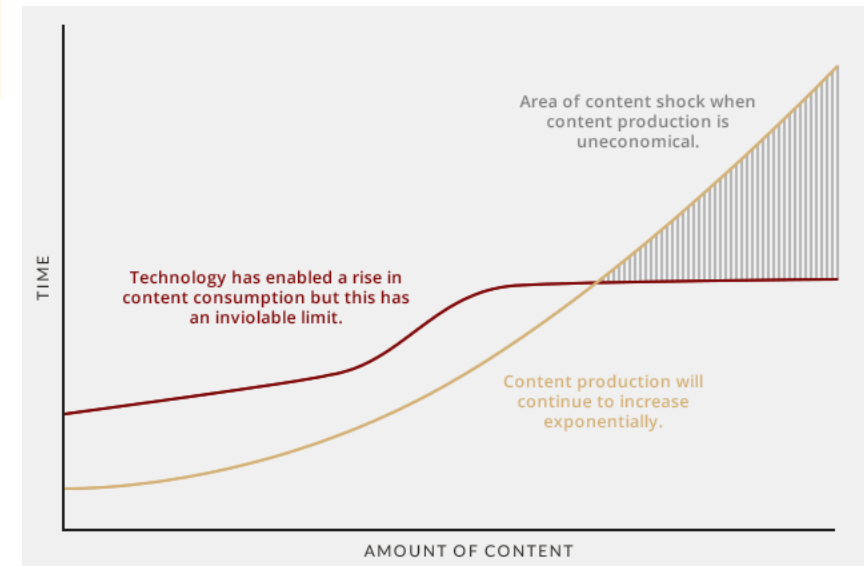
Рост видео трафика

Экзабайт в месяц



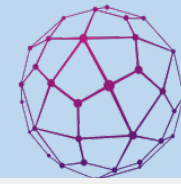
Mobile Video (55%, 75%)
Mobile Web/Data/VoIP (36%, 17%)
Mobile Audio (8%, 6%)
Mobile File Sharing (1%, 2%)

Figures in parentheses refer to 2015 and 2020 traffic shares



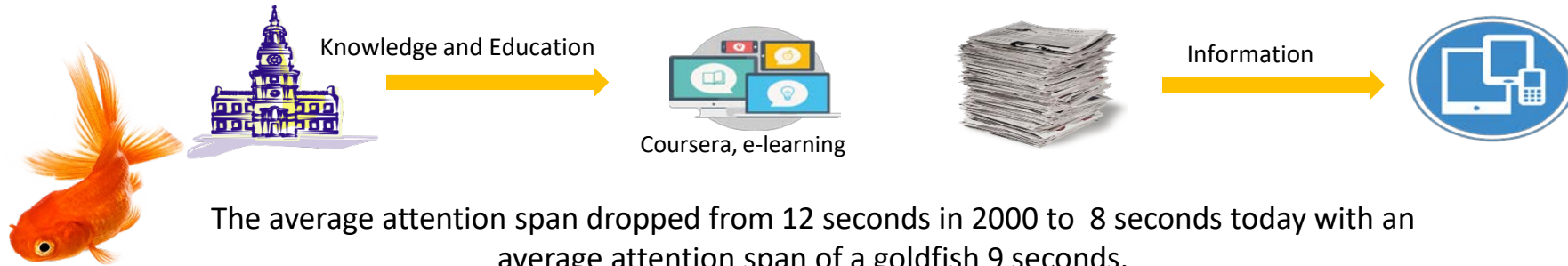
- 5000 маркетинговых сообщений в день
- 86% людей пропускают ТВ рекламу
- 90% электронных писем никогда не открывают
- 99.5% ссылок в электронных письмах никогда не открываются
© 2013 SAP AG
- 60-70% маркетингового контента остается невостребованным
© Sirius Decisions

Влияние технологии на контент: From Greece Academy to Coursera

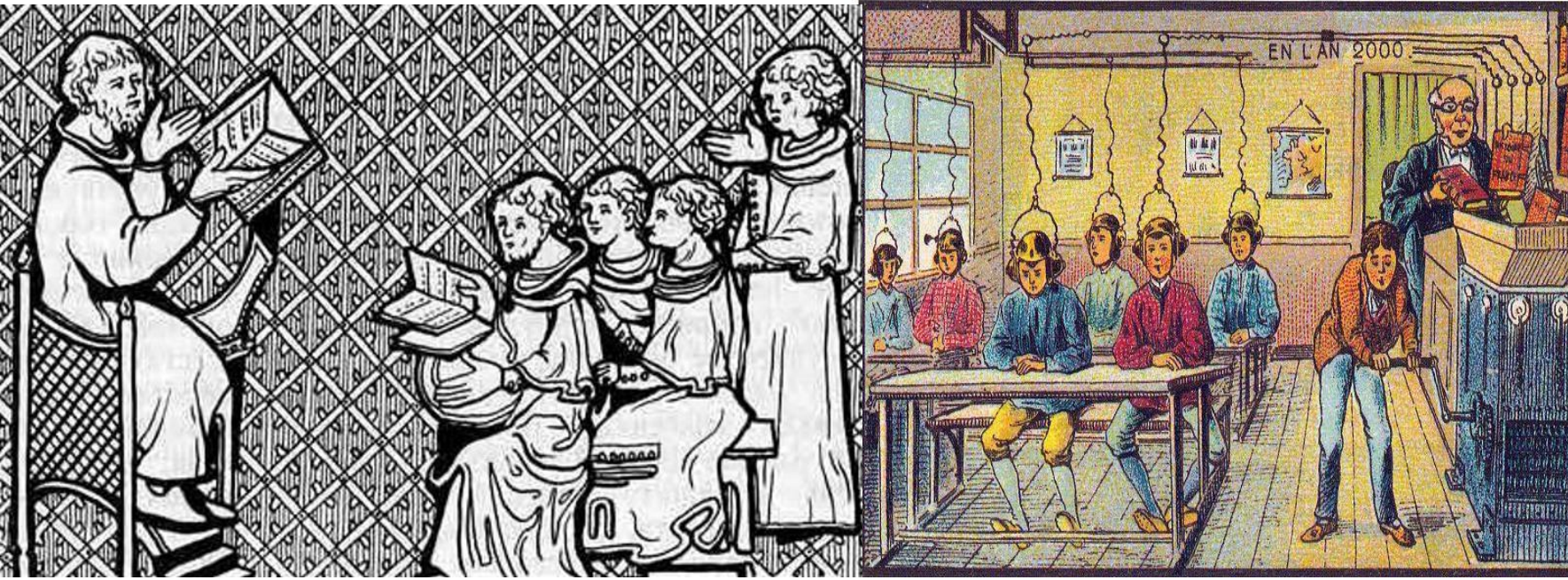
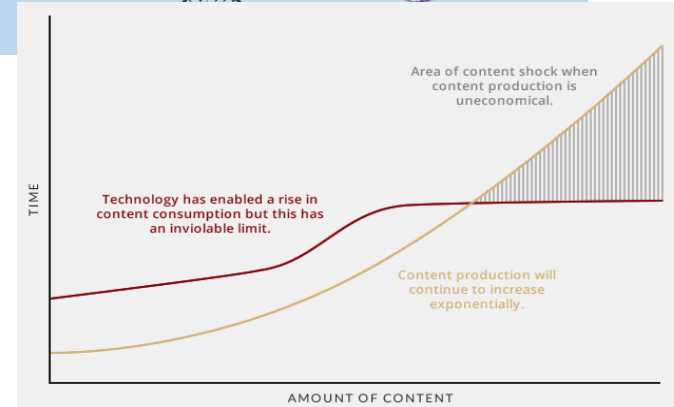


APPLIED
RESEARCH
CENTER FOR
COMPUTER
NETWORKS

- New types of content:

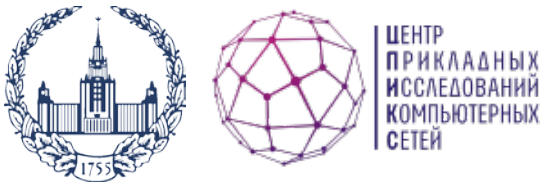


The average attention span dropped from 12 seconds in 2000 to 8 seconds today with an average attention span of a goldfish 9 seconds.



Content Shock: Exponentially increasing volumes of content intersect our limited human capacity to consume it.

Передача видео



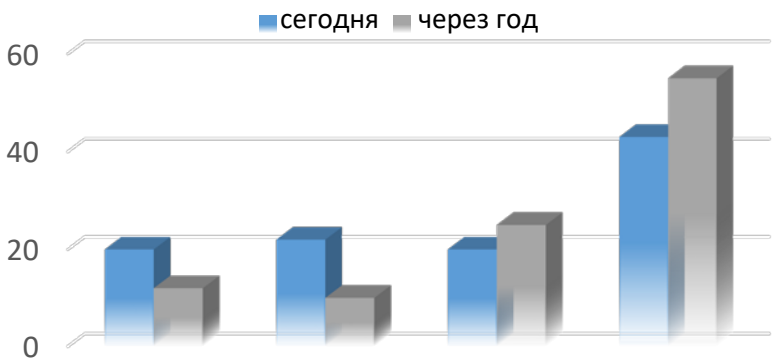
Как операторы планируют доставлять видео контент
Проценты



- **33%** Доля немедленных отказов от видео плохого качества © Akamai
- **50%** Доля отказа от контента с задержкой в 5 секунд (включая рекламу) © Akamai
- **66%** Потребителей разочарованы буферизацией видео © Research Now
- **21%** Сильно раздражаются, если видео не загружается быстро © Research Now
- **35%** Пользователей ощущают повышение пульса во время буферизации © Erickson

- Доля видео в глобальном интернет трафике:
2015: 63% / 2020: 79%
- Доля видео в формате HD и ultra HD
2015: 53% / 2020: 82%
- Минут интернет видео в месяц:
2020: 3 триллиона
= 5 миллионов лет видео в месяц
= 1 миллион минут видео каждую секунду
(c) Cisco Visual Networking Index Complete Forecast for 2015 to 2020

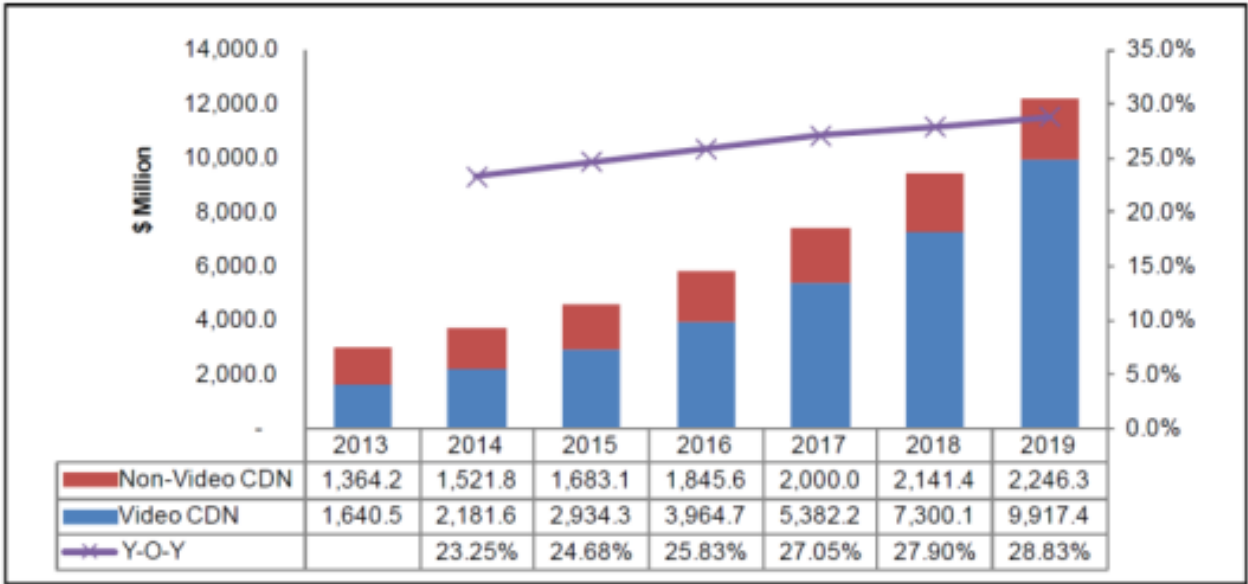
Профиль видео библиотек в формате HD операторов



Процент видео в формате HD

© nScreenMedia, 2016

Размер рынка CDN, по типам, 2013-2019 (\$ млн, Y-O-Y %)



Source: MarketsandMarkets Analysis

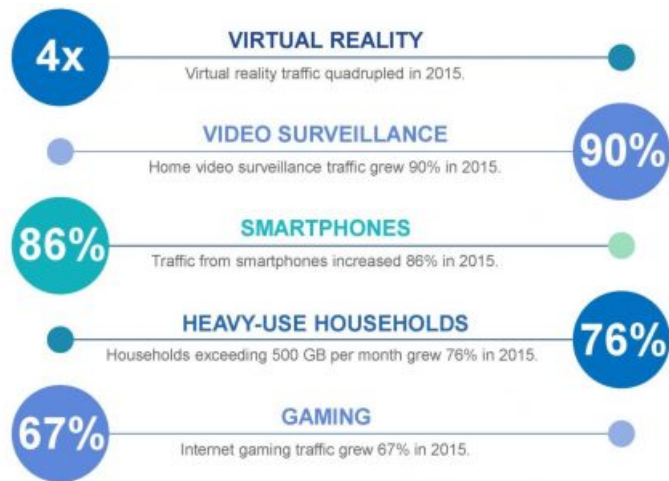
*OVSP - OTT video service provider

*Y-o-Y = year-over-year

Будущее контента

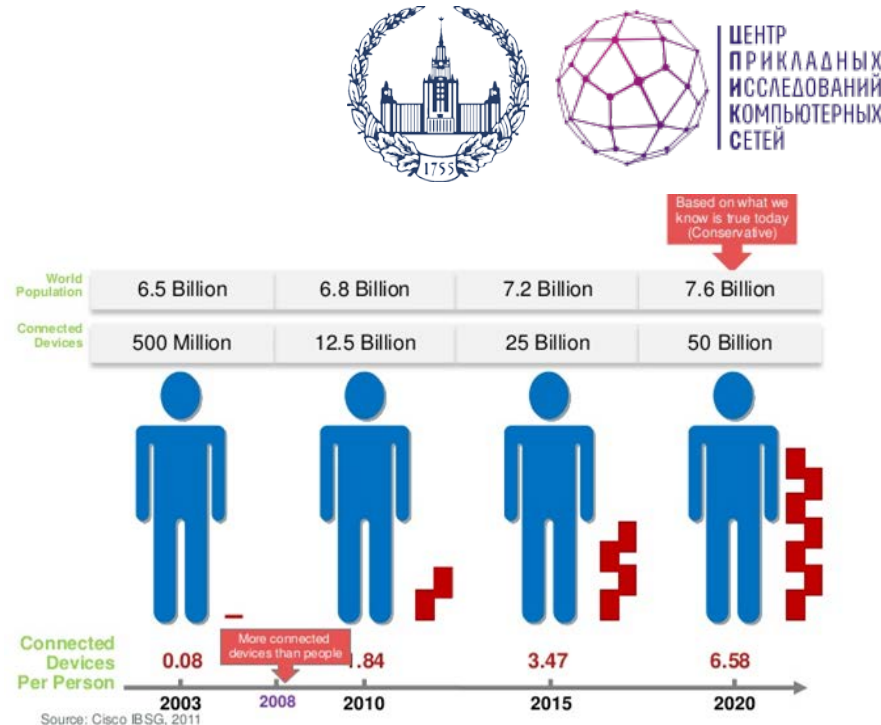
Общие тренды:

- Годовой рост глобального IP трафика: > 1000 EB к концу 2016, 2.3 ZB в год в 2020
- Интернет трафик в часы пик растет быстрее, чем средний интернет трафик
- Трафик смартфонов превысит трафик стационарных компьютеров к 2020
- Трафик от беспроводных и мобильных устройств будет составлять 2/3 глобального IP трафика к 2020
- CDN будут передавать около 2/2 интернет трафика к 2020
- Скорость широкополосного доступа почти удвоится к 2020: фиксированная ШПД достигнет 47.7 Мбайт/с
- Количество устройств с доступом в интернет превысит население земли в три раза к 2020



Глобальные мобильные тренды:

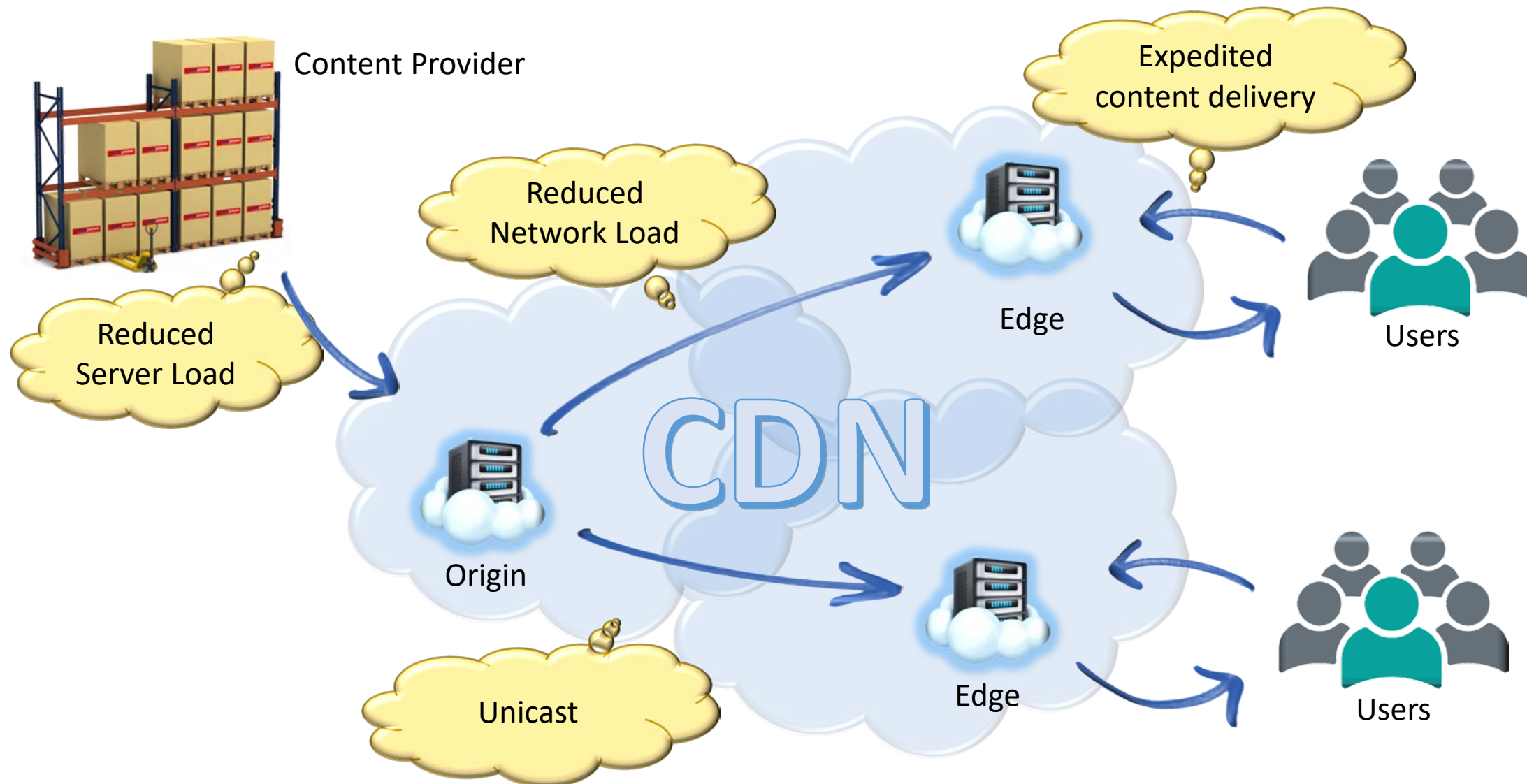
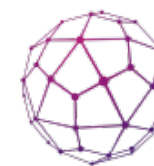
- Мобильный трафик возрастет в 8 раз с 2015 по 2020
- Мобильный трафик будет расти в три раза большими темпами, чем фиксированный IP трафик с 2015 по 2020



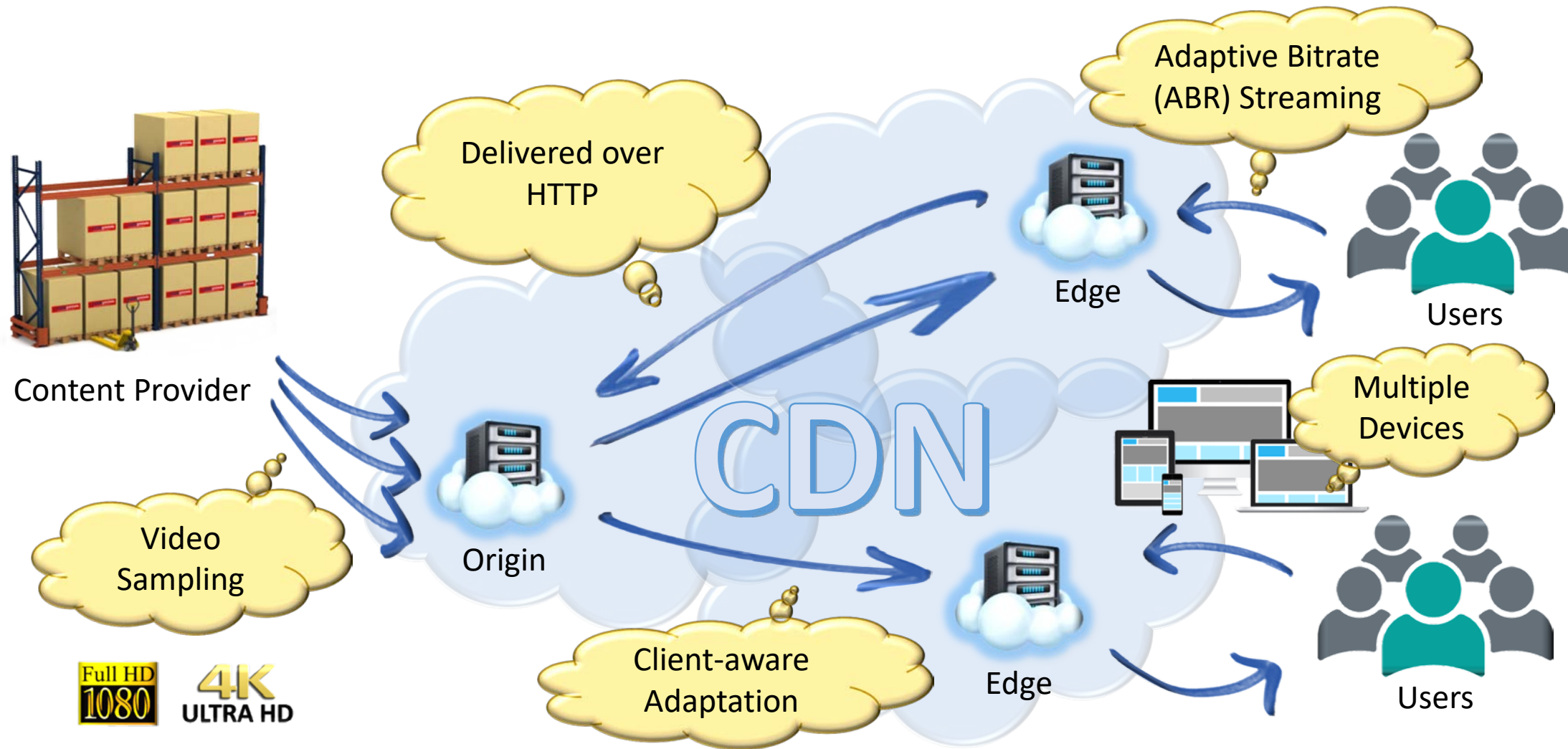
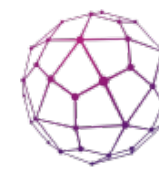
Global Internet Video and Gaming Highlights:

- В 2020 просмотр всего видео, прокачанного через Интернет за месяц, займет 5 миллионов лет
- VR трафик удвоился в 2016; Ожидается, что он возрастет в 61 раз к 2020 со скоростью по 127% в год.
- Игровой трафик возрастет 7-кратно к 2020
- Трафик video-on-demand (VoD) удвоится к 2020
- Уже в 2015 50% пользователей смотрели IPTV

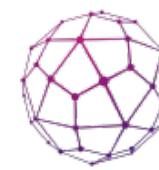
Статический контент



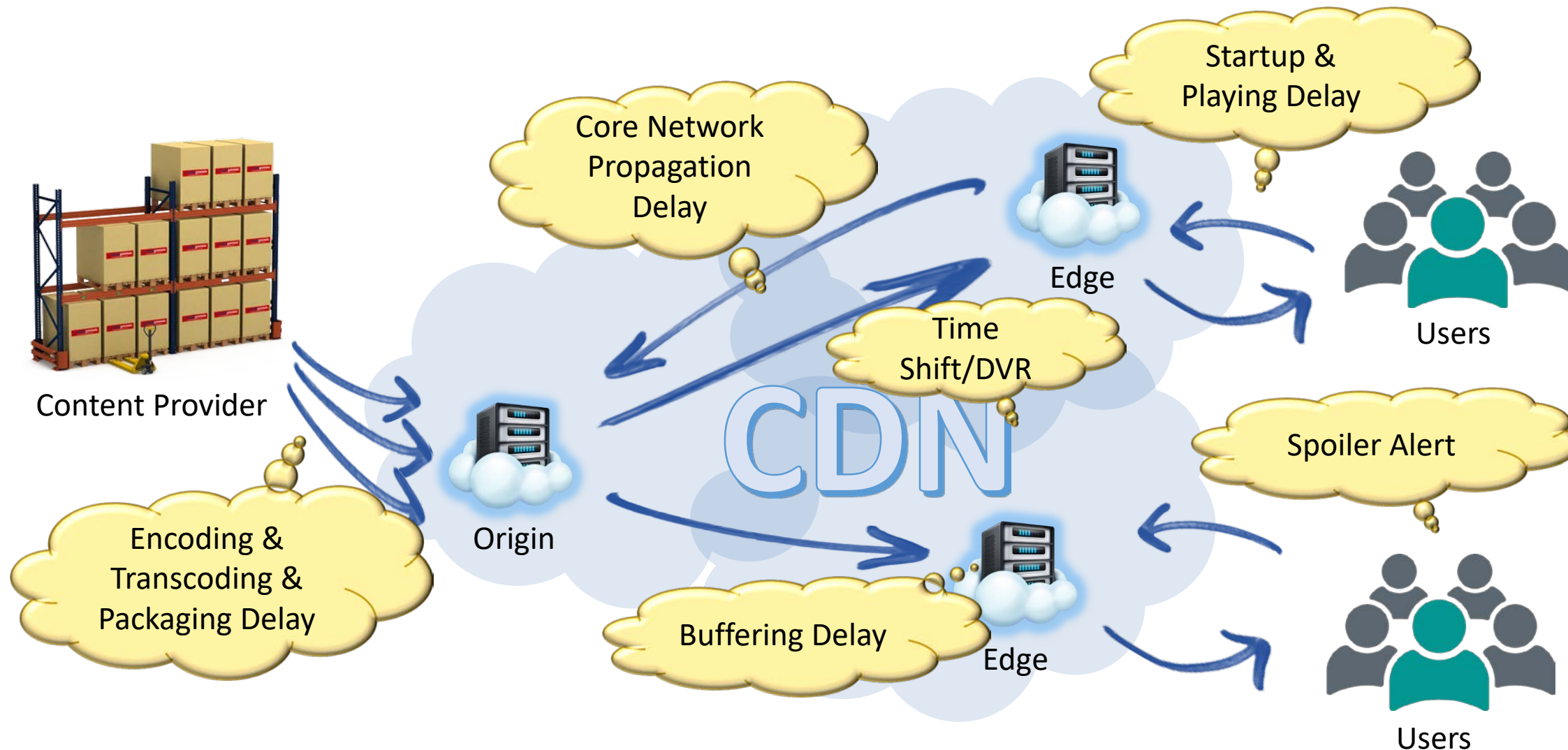
Видео по запросу (VoD)



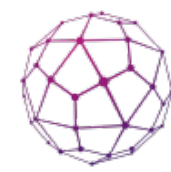
Живая трансляция



ЦЕНТР
ПРИКЛАДНЫХ
ИССЛЕДОВАНИЙ
КОМПЬЮТЕРНЫХ
СЕТЕЙ



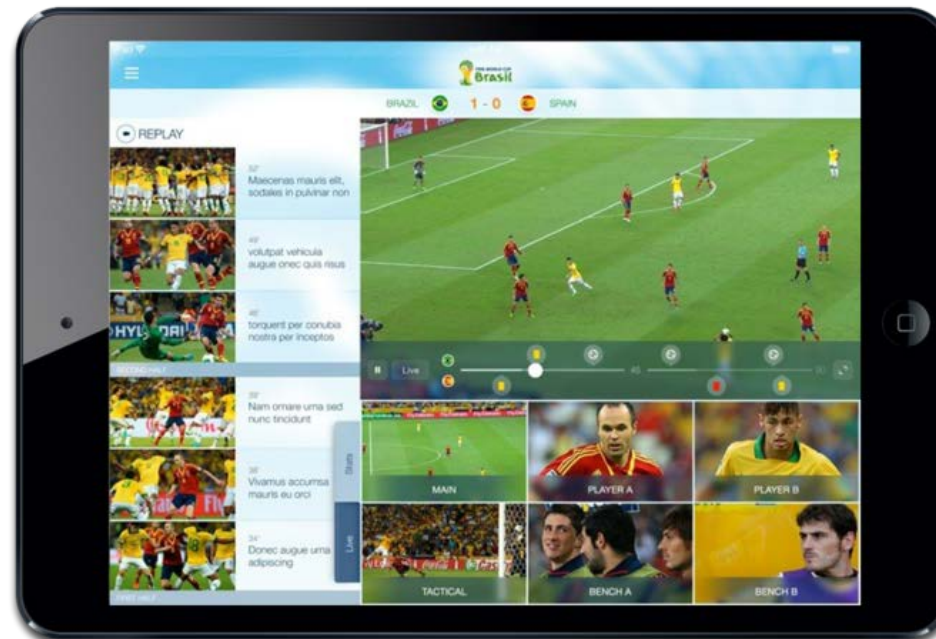
CDN должны предоставлять кастомизированный опыт



ЦЕНТР
ПРИКЛАДНЫХ
ИССЛЕДОВАНИЙ
КОМПЬЮТЕРНЫХ
СЕТЕЙ

Чего хотят пользователи:

- Смотреть Олимпийские велосипедные гонки с максимальным качеством
- Наложение инфографики дорожного ландшафта
- Показывать локацию, относительные позиции и время для сборных команд
- Отображать видео в режиме мягкого цвета (ночной режим)
- Сделать скроллинг текста для олимпийских соревнований по спортивной гимнастике
- Смена канала по попыткам сборной
- Делать другие корректировки по своему усмотрению



**CDN должны
предоставить
приложение для
потребления контента.**

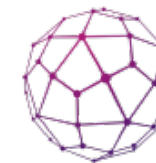
Кастомизация доставки контента, разработанная индустрией:

- Elemental Cloud для облачной обработки потока в реальном времени
- Aspera для передачи файлов с высокой пропускной способностью от места действия до облачной инфраструктуры
- Облачная платформа перекодирования мультимедийных данных Brightcove
- Amazon S3 storage
- Akamai content delivery network (CDN)
- NETCO Sports second-screen app designer

**Люди готовы оплачивать дополнительные
расходы на кастомизацию**



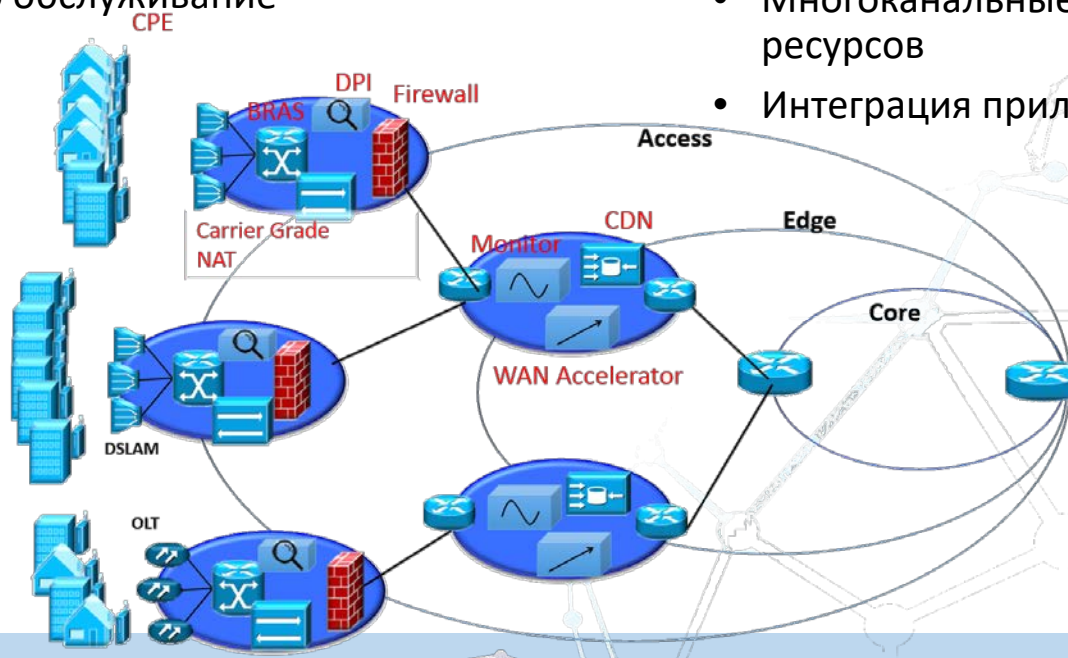
Конвергенция между сетями CDN и IP-сетей



- «Чистые» сервера должны выдерживать пиковые нагрузки
 - В большинстве случаев сильно не доиспользованы
- Виртуализированные серверы привязываются к крупным центрам обработки данных
 - Не могут менять локацию по запросу
- Туннели «точка-точка» приводят к ненужной репликации
 - Мультикастинг имеет лучшую эффективность
- CDN не взаимодействуют с используемыми сетями
 - Не влияет на качество обслуживания

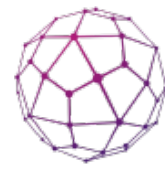
- Адаптивные технологии BitRate на базе на HTTP (и TCP)
- Потокным игрокам не нужен TCP общего назначения :
 - Не может обеспечить достаточную скорость на дальних дистанциях
 - Невозможно составить многоадресную рассылку
 - Не удастся удалить пропущенные кадры
- Существует потребность в новых сетевых протоколах на основе контента
 - Многоканальные протоколы для использования пула ресурсов
 - Интеграция приложений на транспортном уровне

Инфраструктура CDN не обладает эластичностью



Сетевой стек не оптимизирован для доставки контента

Концепция CCN как основа для конвергенции сетей CDN и IP



Конвергенция сетей CDN и IP требует:

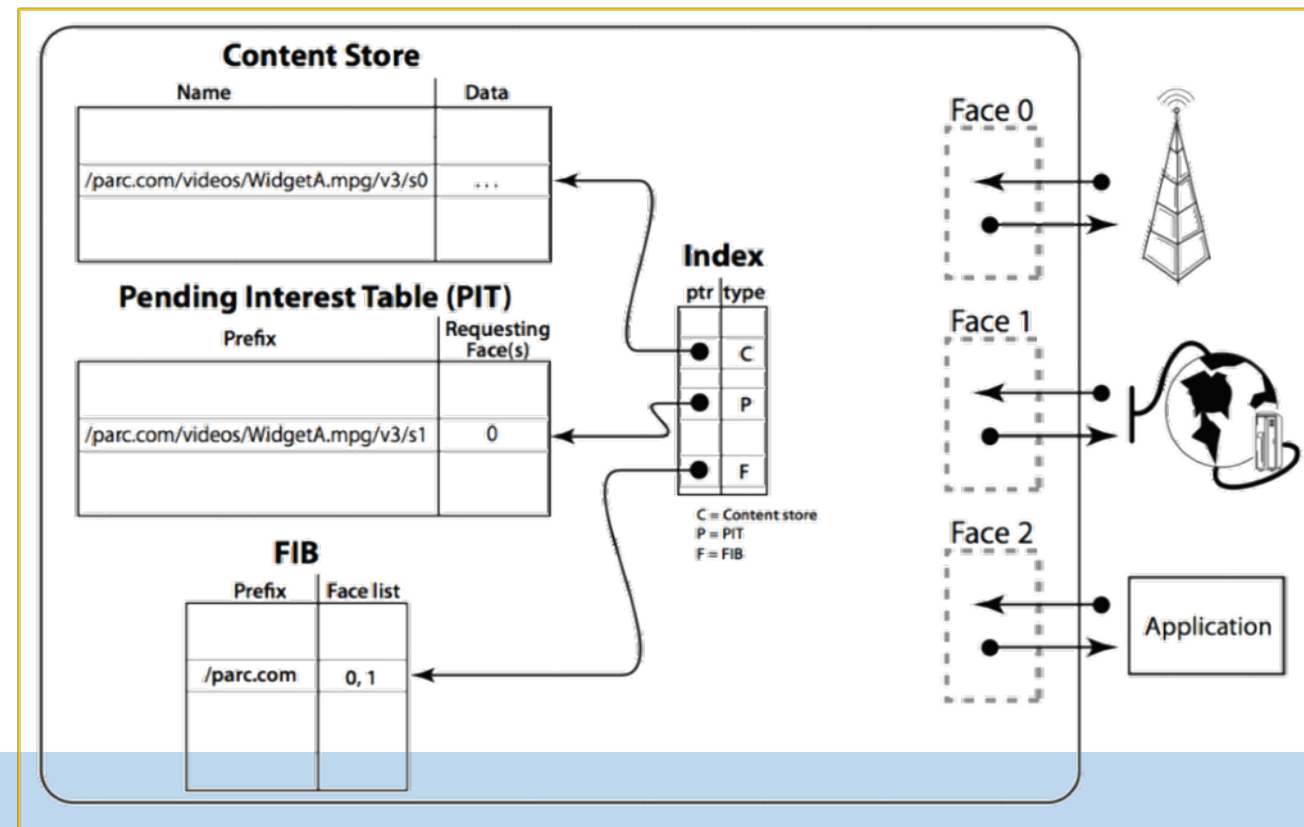
- Присвоение имен контенту и его оптимальное размещение
- Получение контента по имени
- Разрешение имен
- Показывать ожидаемое QoE на примере
- Доставку запрашиваемого контента с требуемым QoE

о Доступ к контенту, а не к устройству

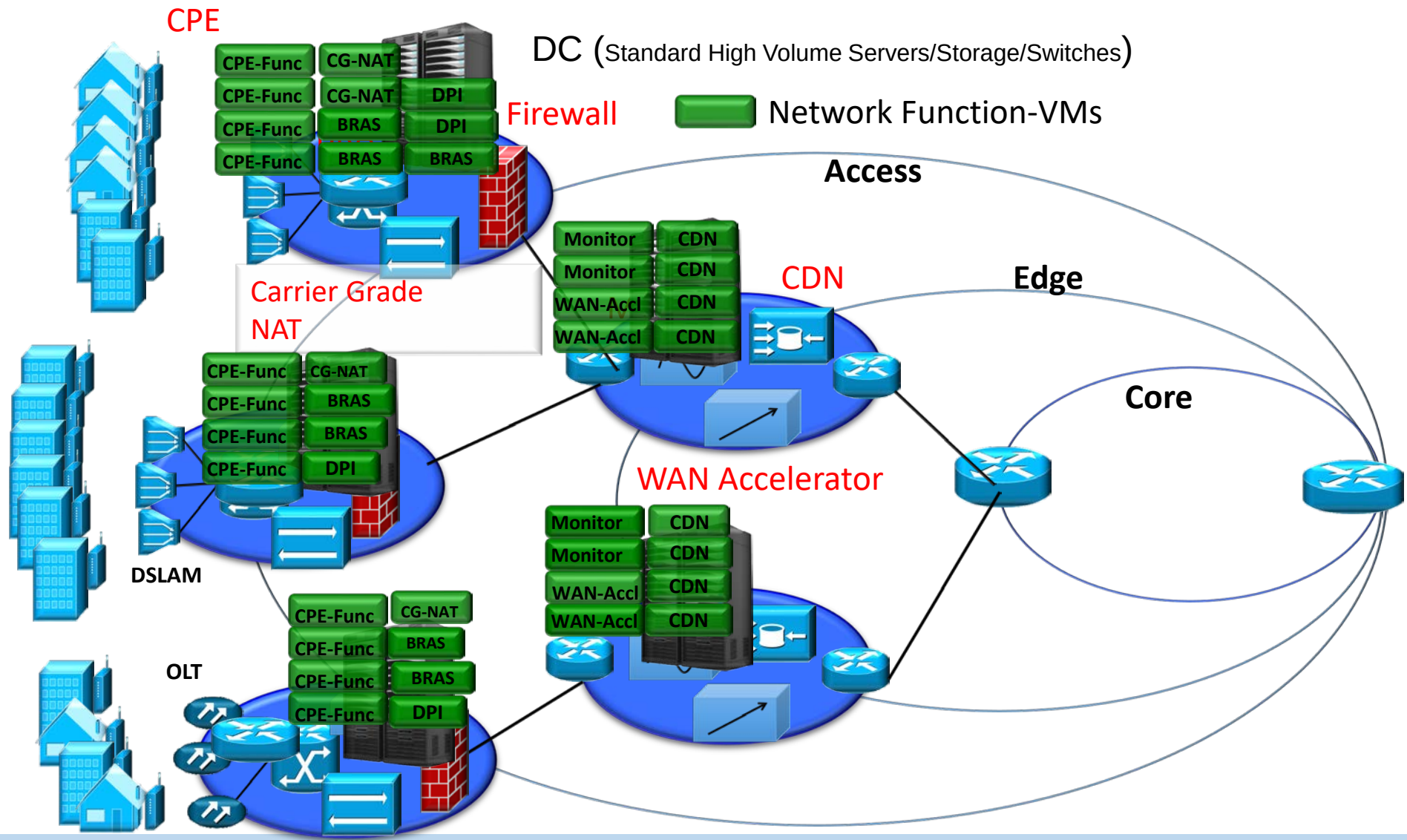
о Доступ к контенту по имени, а не по адресу

- ✓ Присвоение имен
- ✓ Разрешение имен
- ✓ Мобильность
- ✓ Безопасность

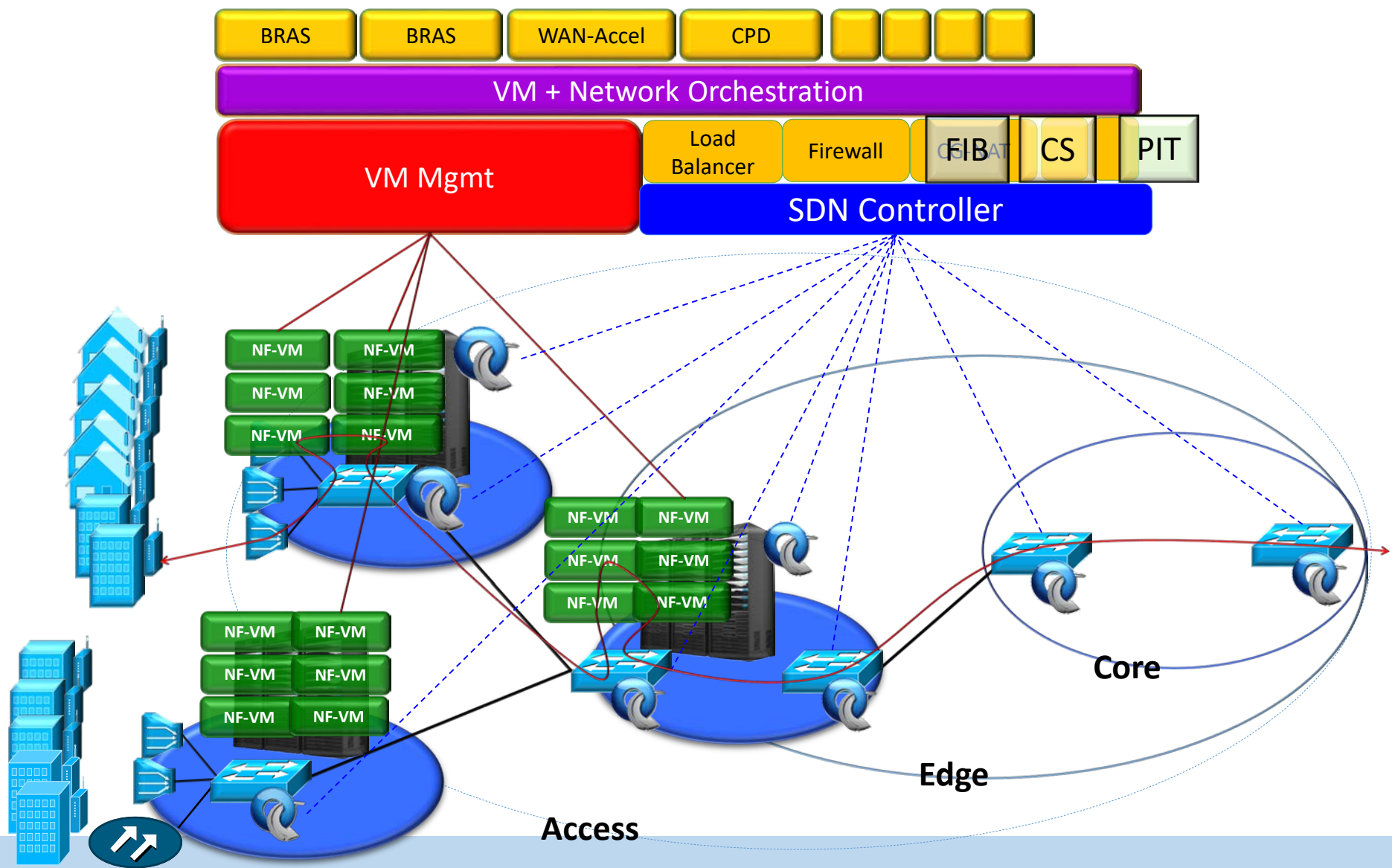
Forwarding Process in CCN:



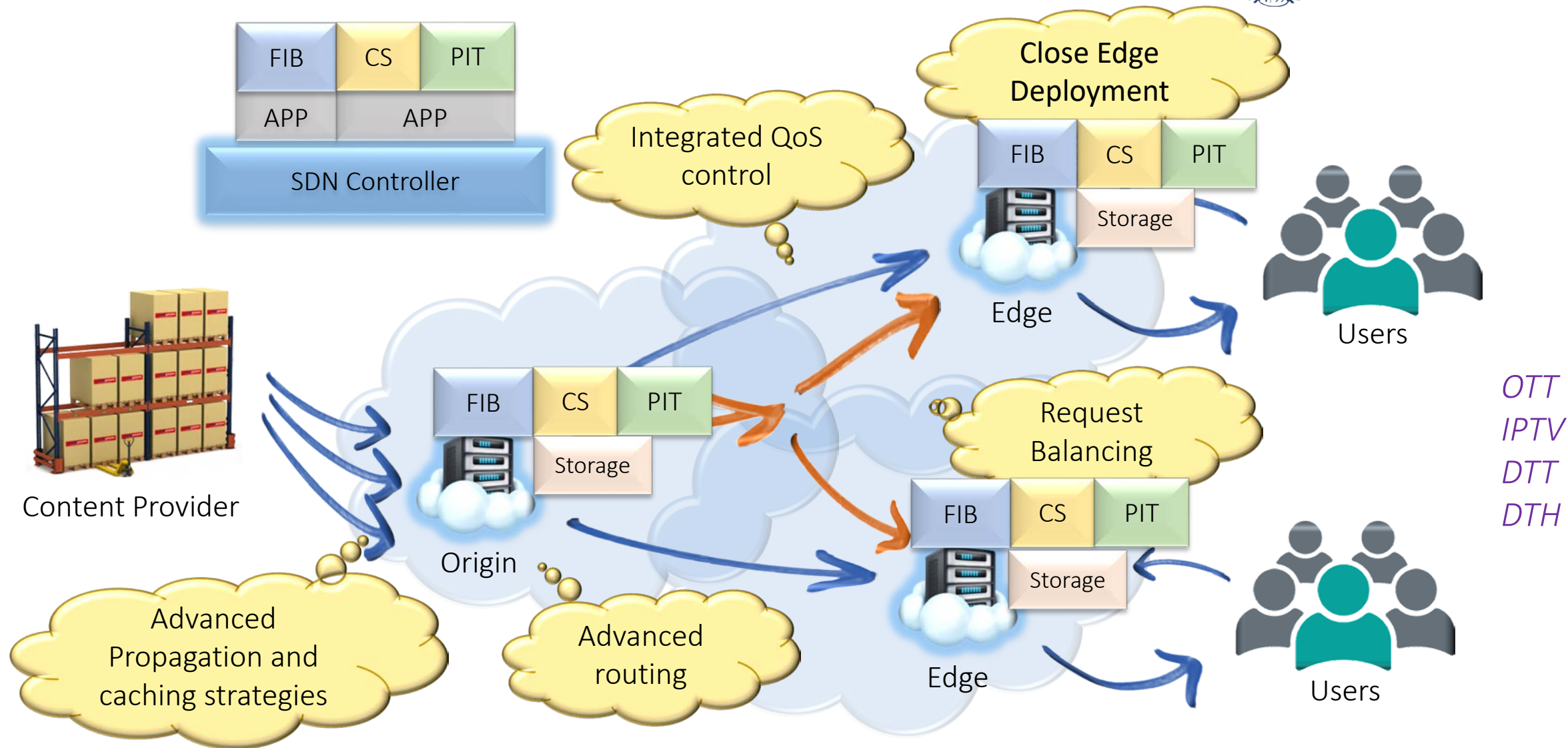
Мягкая эволюция на основе NFV



Конфигурация SDN сети и контроль QoS

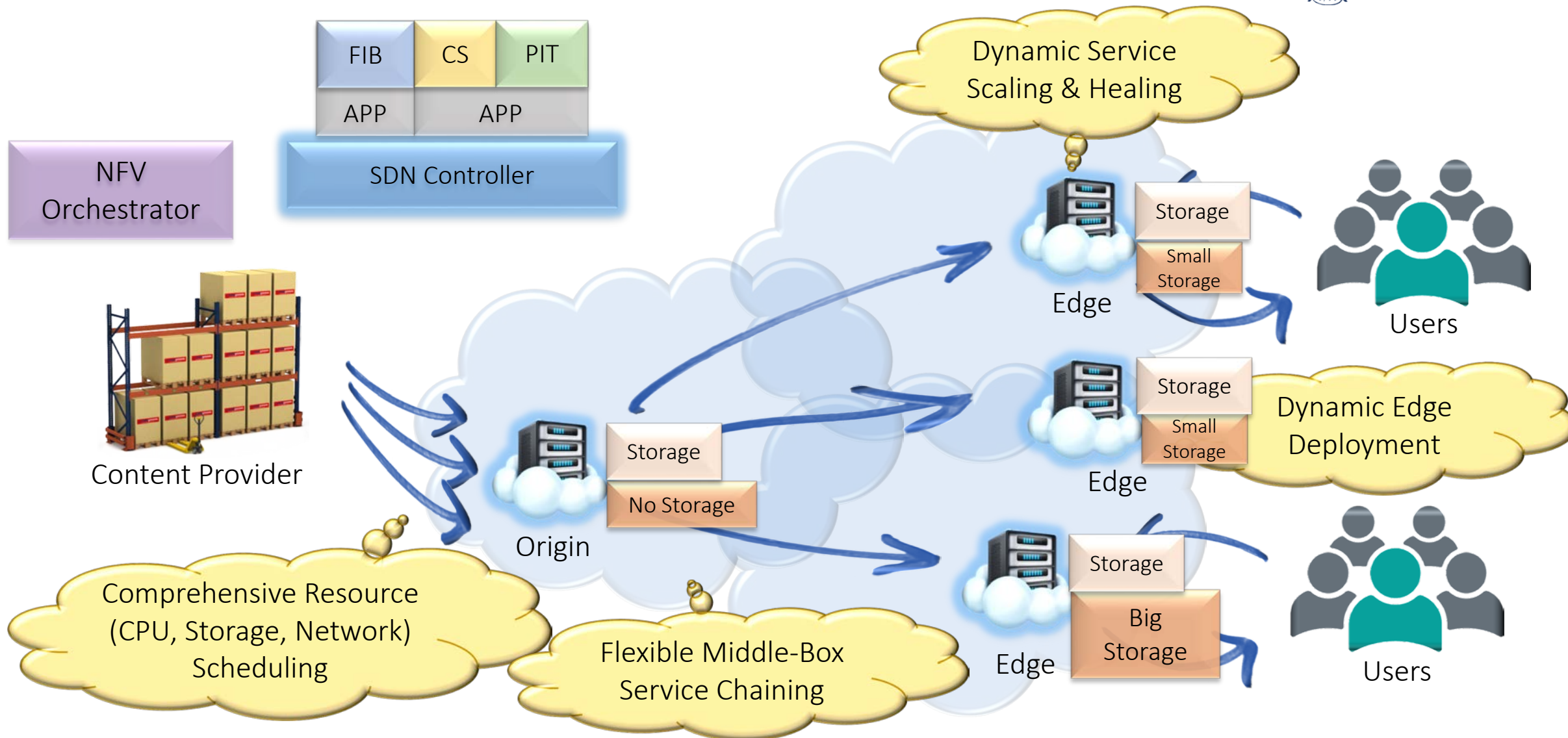
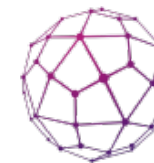


NG-CDN = SDN ?

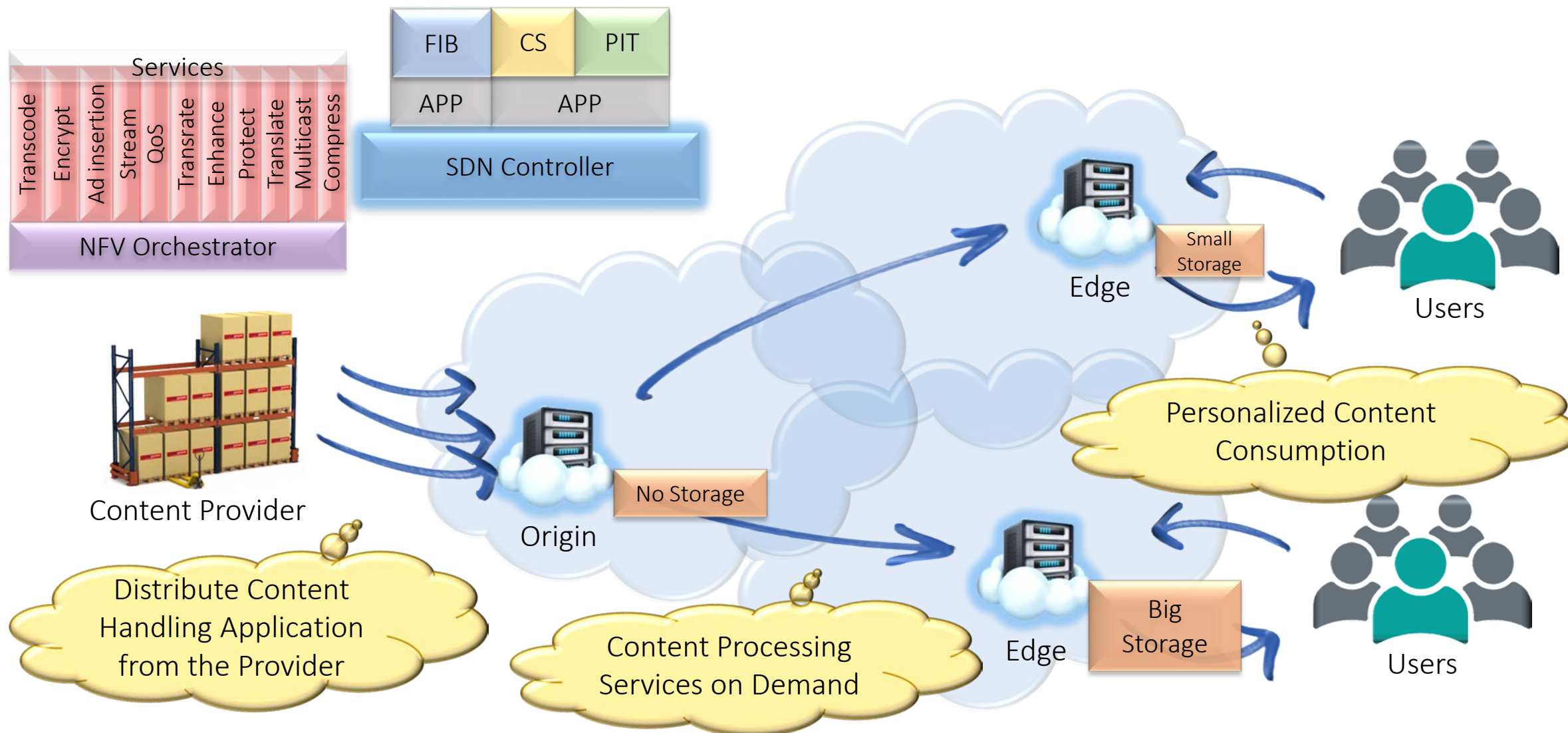
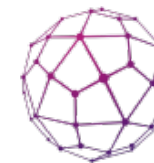


OTT
IPTV
DTT
DTH

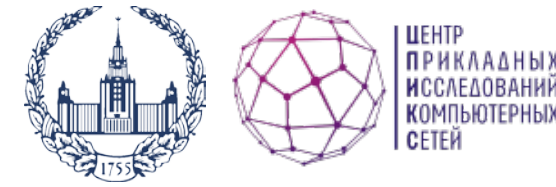
NG-CDN = SDN + NFV ?



NG-CDN = SDN + NFV + vCaching + Naming + vTranscoding...?



NG-CDN and Cache Policy



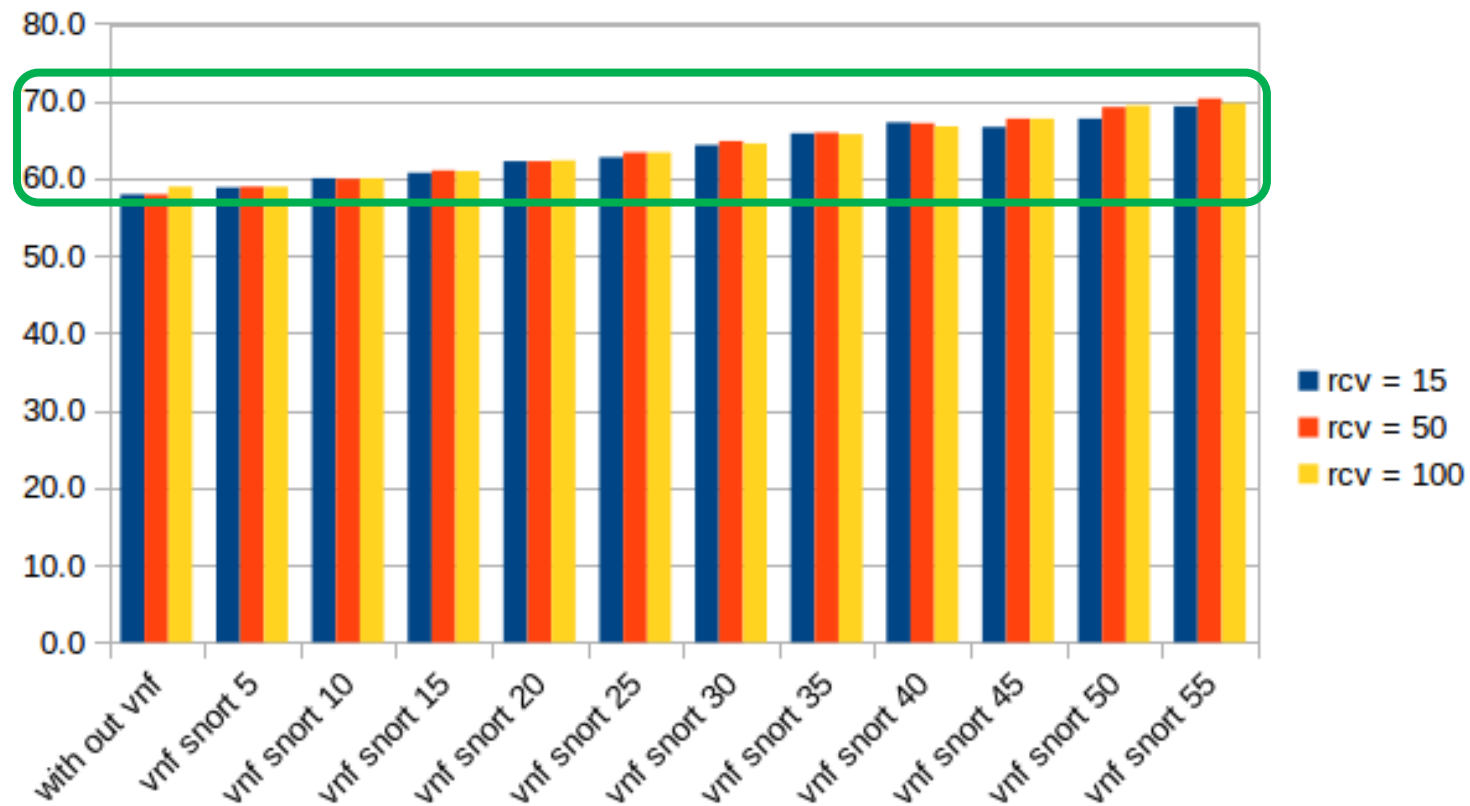
- Стратегии развертывания кэша:

- Прокси
- Резервирование прокси
- Прямой
- Адаптивный
- Активный

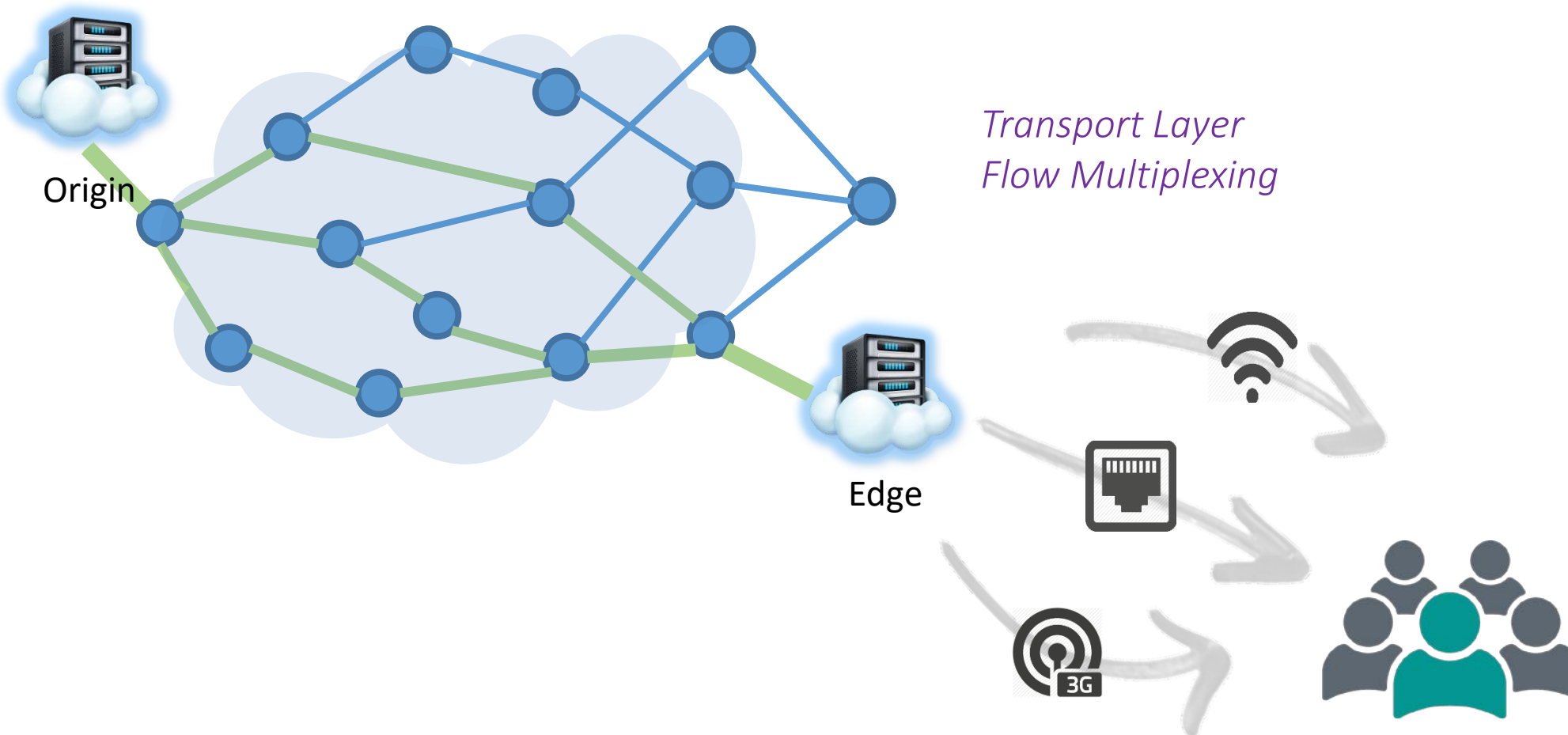
Метрики:

- Частота - в числовых терминах, сколько запросов размещены или, как часто запрашивается объект?
- Периодичность - время, в течение которого на объект или контент ссылались или требовались
- Размер – размер контента
- Стоимость извлечения - затраты, понесенные для извлечения содержимого или объекта
- Время обновления – модификация к кеше
- Замена - лучшее время, когда контент становится менее актуальным

Service chaining overheads



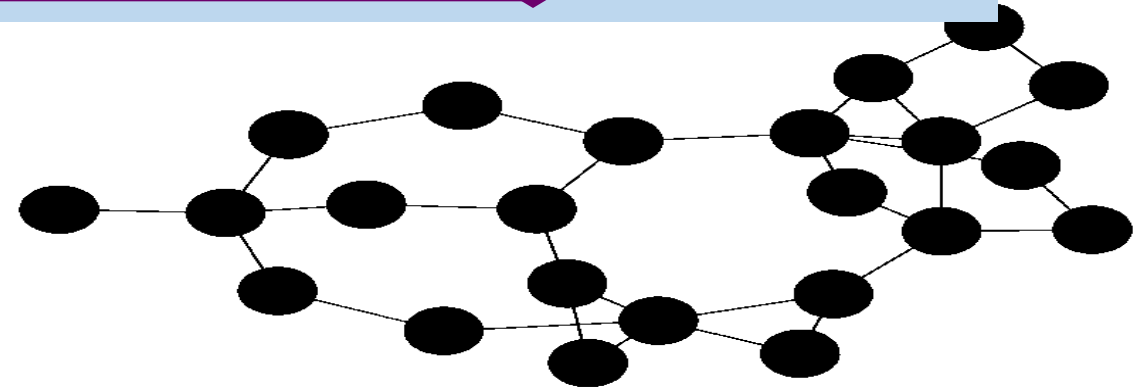
NG-CDN and Path Diversity



Flow (De)Multiplexing Protocol

Routes:

- 1.rate 2 Gbit/s, delay 50 ms, loss 0.05%
- 2.rate 1 Gbit/s, delay 50 ms, loss 0.005%
- 3.rate 1 Gbit/s, delay 50 ms, loss 0.1%



TCP protocol & congestion control	Interval peak for average aggregate rate (Mbps) over 100 players			
	2 seconds	10 seconds	30 seconds	60 seconds
Single Path Baseline	206.988	195.985	194.042	192.655
MP TCP: LIA	218.991	215.066	211.573	207.948
MP TCP: OLIA	228.747	226.022	224.382	221.541
MP TCP: BALIA	219.445	214.949	212.916	210.408
FDMP	450.477	414.202	403.952	378.816

FDMP provide throughput growth capacity: 96% over single-flow optimized TCP and 70% compared to non-optimized multi-path TCP



Сервер маршрутизации (Route Server, RS)

Чтобы начать использовать сервер маршрутизации (RS), участник должен выполнить следующие шаги для установки сеанса пиринга BGP с автономной системой (AS):

- Написать политику описания для взаимодействия с сервером маршрутизации AS;
- Конфигурироваться сессии BGP со всеми серверами маршрутизации;
- Сервер маршрутизации ретранслирует сообщение BGP между связанными участниками, таким образом, просмотр с помощью RS означает установление пиринговых отношений со всеми другими участниками IXP, связанными с RS
- AS передает описание политики маршрутизации серверу маршрутизации AS.

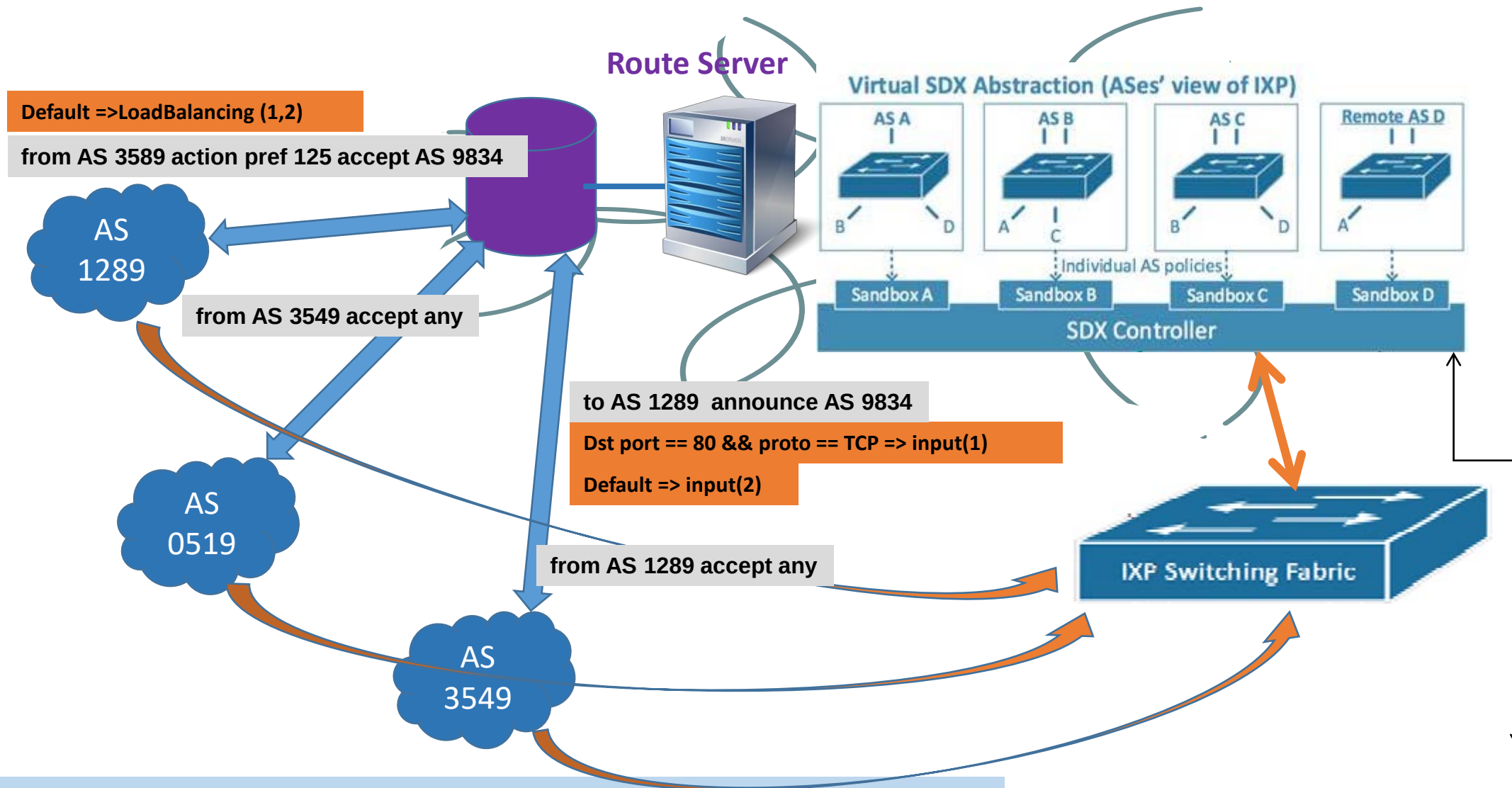
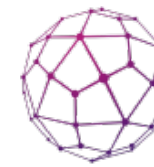
Недостатки RS:

- ❑ Обновления конфигурации сервера маршрута занимают часы. Обновления конфигурации включают проверки и обновления или фильтры политики маршрутизации, применение изменений в конфигурации RS.
- ❑ График обновления конфигурации - один раз в день. Процедура занимает от нескольких минут до одного часа

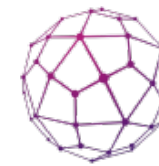
Преимущества RS:

- ❑ Сервер маршрутизации (RS) упрощает пиринг между участниками IXP и позволяет им сократить количество индивидуальных сеансов пиринга.

IXP with SDN & Route Server (SDX)



Темы для исследований



- Проблема присвоение имени для различных видео сервисов
 - Схема присвоения имен
 - Разрешение имен
- Виртуализация хранения кеша Cache Storage Virtualization and Resource Scheduling
 - Хранение кеша как сервис управления и оркестрации
 - Адаптивные алгоритмы планирования ресурсов
- Политика кеширования и управления хранением
 - On-Path и Off-Path кеширование
- Стратегия переадресации трафика контента
 - Комбинация уникаст и мультикаст техник
 - Использование специализированных протоколов, например, FDMP
- “On-Fly” CDN трансформация трафика
 - Транскодирование
 - Сжатие
 - Семплинг
- Управление QoS: multipath routing, transcoding, балансировка нагрузки
- CDN маршрутизация трафика в мультидоменных сетях (SDX)



APPLIED
RESEARCH
CENTER FOR
COMPUTER
NETWORKS



www.arccn.ru



+7 (495) 240-50-63



smel@arccn.ru



@ArccnNews