

Cisco Expo 2011



# Архитектура и решения Cisco для современного ЦОД

Скороходов Александр  
Системный инженер – консультант

[askorokh@cisco.com](mailto:askorokh@cisco.com)

+7(495)789-8615

innovate *together*

# План потока по центрам обработки данных

День 1	День 2	День 3
Тесно интегрированные решения Cisco и NetApp для построения современного ЦОД. Компания NetApp	ЦОД будущего: универсальные строительные блоки с единым управлением. Компания EMC	Программные комплексы Struxureware для контроля инженерных инфраструктур. Компания Schneider Electric
Архитектура и решения Cisco для современного ЦОД	Принципы построения катастрофоустойчивых ЦОД. Связь распределённых ЦОД с использованием технологий Cisco OTV и LISP (2 часа)	Объединенная вычислительная система Cisco UCS. Совместно с компанией I-Teco (2 часа)
Развитие технологий унифицированной сети ЦОД (FCoE, FabricPath)	Эволюция технологий оптимизации приложений - Cisco WAAS	Коммутаторы сетей хранения Cisco MDS 9000
Объединенные сетевые сервисы для виртуализованных вычислений	Преимущества внедрения Unified Communications на платформе Cisco UCS. Компания СТИ	Облачные решения компании BMC Software на платформе Cisco UCS
Развитие технологий коммутации трафика виртуальных машин	Опыт внедрения решений Cisco UCS на инфраструктуре заказчиков. Компания Техносерв	Развитие семейства коммутаторов Nexus: Nexus 7000, Nexus 5000/2000
	<p>Дизайн небольшого ЦОД (2 часа)</p> <p>Сессия неформального общения по тематике ЦОД</p>	

## Дополнительные потоки по тематике ЦОД

<b>Облачные вычисления. День 2</b>	<b>Демо-поток по ЦОД. День 3</b>
<b>Стратегия реализации "облачных вычислений" предприятия. На что нужно обратить внимание CIO?</b>	<b>Демо-сессия Citrix - OCS</b>
<b>Облачные технологии от оператора связи Компания Orange</b>	<b>Технологии распределенного уровня доступа, Cisco vPC+ и Cisco FabricPath для построения масштабируемых L2-сетей в ЦОД.</b>
<b>Реализация подхода десктоп как услуга на базе совместного решения Citrix XenDesktop и Cisco UCS. Компания Citrix</b>	<b>Связь распределенных ЦОД</b>
<b>"Unified Communications as a Service". Предложение Cisco</b>	<b>Технологии коммутации трафика физических и виртуальных машин: Cisco Adapter-FEX и VM-FEX.</b>
<b>Customer Case Study. Tieto</b>	<b>Vblock demo Сергей Елкин, iTesco</b>
<b>"Security as a Service". Предложение Cisco</b>	<b>Система управления Cisco Unified Computing Systems Manager</b>



# Эволюция идеологии ЦОД

Консолидация и управляемость

Фаза 1

Мейнфреймы



Централизация

Фаза 2

Распределенная  
обработка



Децентрализация

Фаза 3

Сервис-ориентированный ЦОД



Виртуализация  
Облачные вычисления

Эволюция архитектуры приложений

# Эволюция ЦОД: сервера

- Традиционный ЦОД:
  - Одно приложение – один сервер
  - Тысячи и десятки тысяч серверов
  - Средняя утилизация < 10-15%
  - Типичный сервер: 1-4 процессорных ядра, 4-8 ГБ памяти
- ЦОД нового поколения
  - Виртуализация серверов и рабочих мест
  - Сокращение числа серверов, площадей и мощности на тот же объём задач
  - Средняя утилизация 50-60% и более
  - Типичный сервер: 8-20+ процессорных ядер, 48-192+ ГБ памяти
  - Миграция RISC систем на стандартные архитектуры



48 блейдов  
Cisco UCS B230

960 процессорных  
ядер Westmere EX

До 24 ТБ DRAM

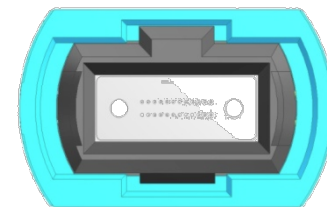
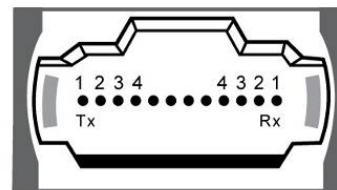
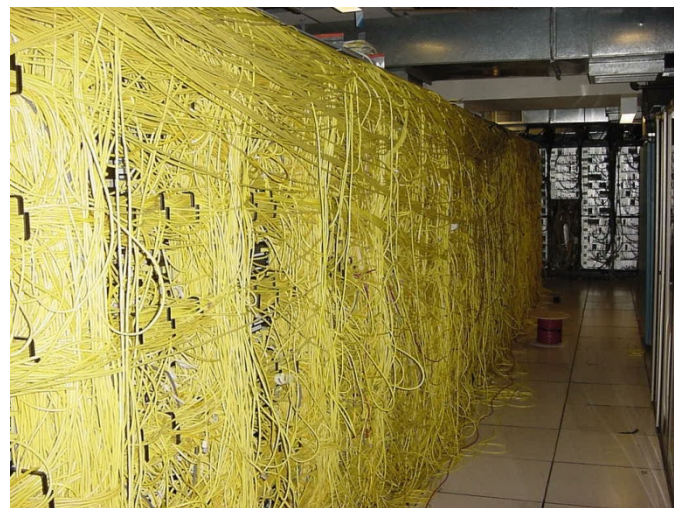
# Эволюция ЦОД: питание и охлаждение

- Традиционный ЦОД:
  - 2-5 кВт мощности на стойку
  - Охлаждение «зала в целом»
- ЦОД нового поколения
  - До 20-30 и более кВт на стойку
  - Управление потреблением
  - Эффективное разделение потоков воздуха: горячие/холодные коридоры, «дымоходные трубы», жидкостное охлаждение
  - Новые технологии гарантированного питания: роторные ИБП
  - Новые технологии охлаждения: «естественное охлаждение» (free cooling)



# Эволюция ЦОД: кабельная система

- Традиционный ЦОД:
  - «Витая пара» до серверов – до 1 Гбит/с
  - Многомодовый магистральный кабель (1-10Гбит/с)
  - Отдельная СКС для SAN
- ЦОД нового поколения
  - 10 Гбит/с до серверов: твинаксиальные кабели/Cat 6A/Cat 7
  - Магистральные соединения готовые к 40/100 Гбит/с: «групповые» оптические кабели
  - Единый конвергентный транспорт для SAN
  - Влияние на топологию СКС



# Требования к сети ЦОД нового поколения И их реализация в решениях Cisco

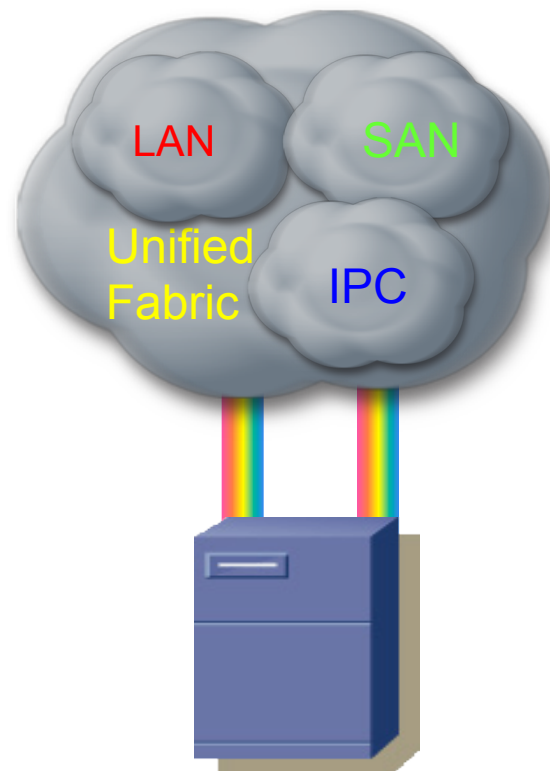
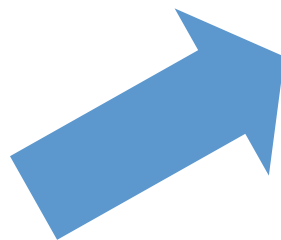
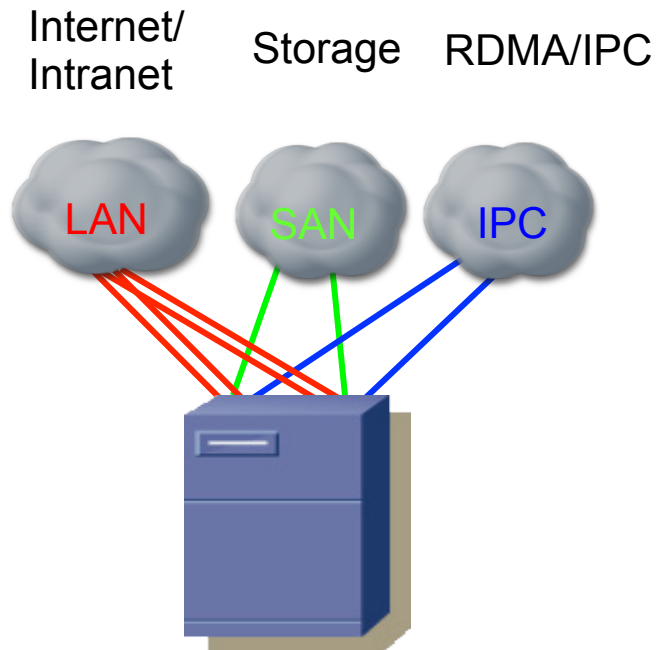
- Повышение производительности подключения серверов и магистралей
  - 10 Gigabit Ethernet, 40/100GE, TRILL/FabricPath
- Консолидация ввода-вывода серверов
  - Fibre Channel over Ethernet (FCoE/DCB), Adapter-FEX
- Более простая и «плоская» архитектура сети ЦОД без опоры на STP
  - Virtual Portchannel, FEX, TRILL/FabricPath
- Сетевая поддержка виртуализации
  - Nexus 1000V, VM-FEX
- Надёжная и производительная связь ЦОД
  - OTV, LISP

# Ключевые технологии Cisco для ЦОД

- Консолидация транспорта
  - FCoE и DCB
- Эволюция дизайна сети ЦОД
  - VPC
  - FEX
  - FabricPath/TRILL
- Сетевая поддержка виртуализации
  - Nexus 1000V
  - VM-FEX/Adapter-FEX
- Надёжная и производительная связь ЦОД
  - OTV
  - LISP



# Консолидация ввода-вывода: Объединенный транспорт FCoE/IEEE DCB



## Сегодня

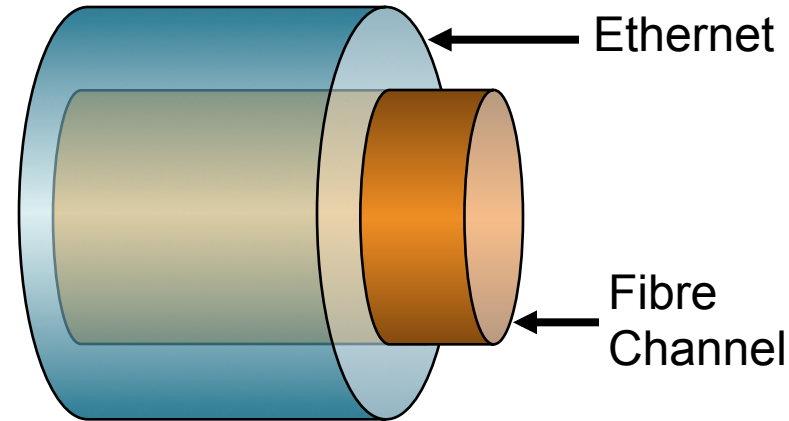
- Много портов ввода-вывода
- Высокие расходы на оборудование и эксплуатацию

## Используя FCoE/DCB

- Общий транспорт
- Обеспечение совместимости

# Fibre Channel over Ethernet (FCoE)

- Метод передачи фреймов FC по Ethernet
  - Выглядит как FC для серверов и сети
  - Сохраняет текущую инфраструктуру и управление FC
  - Фрейм FC остается неизменным
- Может работать на стандартных коммутаторах (с jumbo фреймами)
- Priority Flow Control обеспечивает отсутствие потерь
  - Имитирует систему буферных кредитов FC
- Стандарт утвержден 3 июня 2009 года (ANSI T11 FC-BB-5)

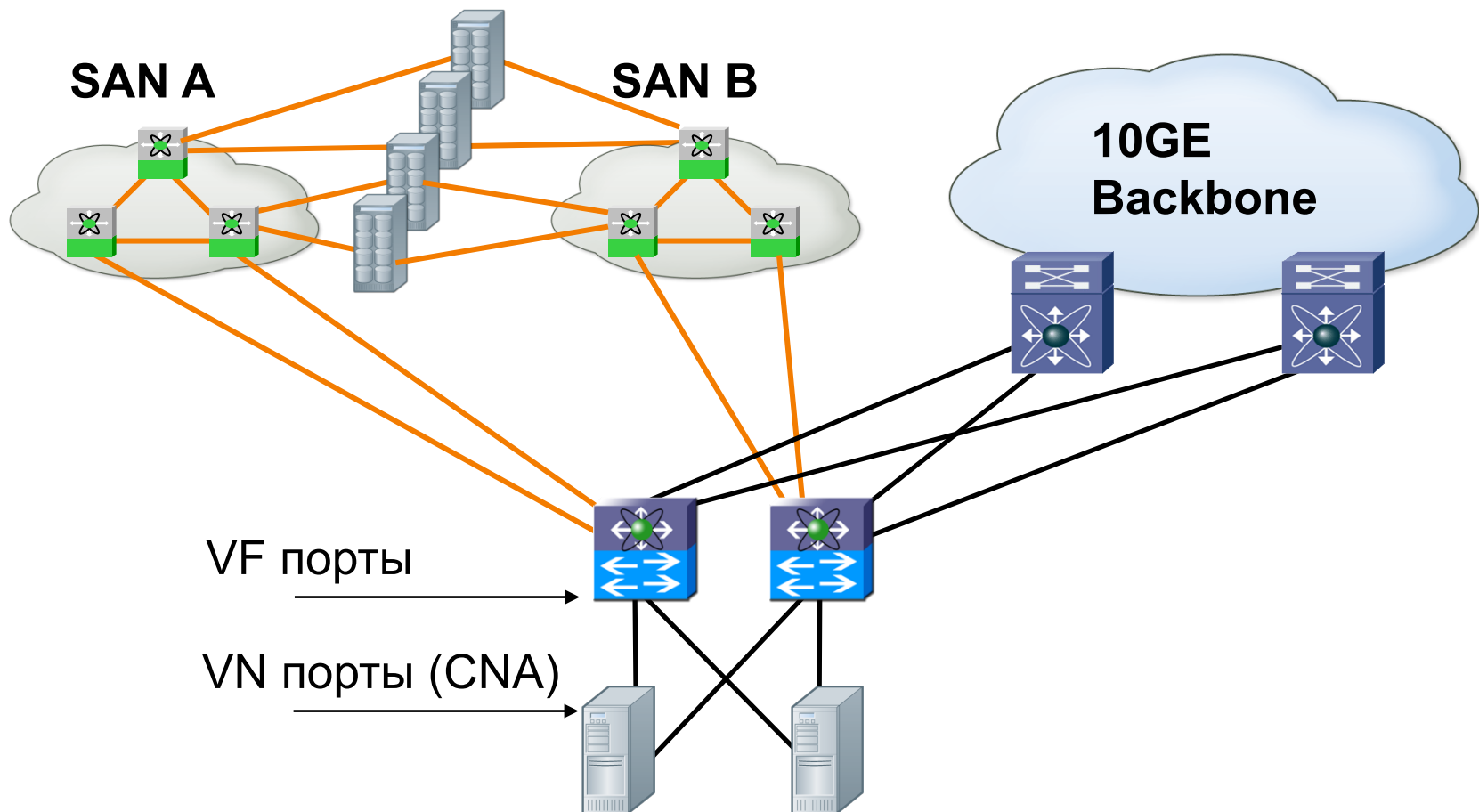


Cisco первой представила основанный на стандартах коммутатор FCoE Cisco Nexus 5000

Весна 2008

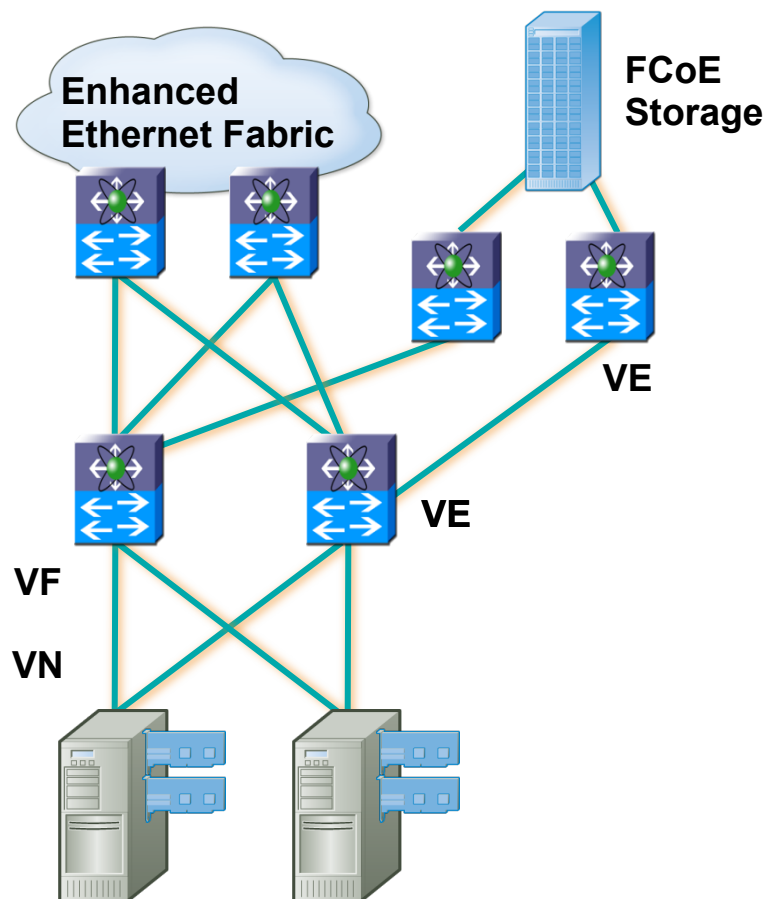
# FCoE: консолидация на уровне доступа

- Первый шаг – «консолидация доступа» («Unified Wire»)
- Существенная экономия при сохранении существующего ядра сетей Ethernet и Fibre Channel



# FCoE: консолидация в масштабах сети

- Расширение консолидации ввода-вывода на магистраль
- Поддержка разделения SAN фабрик для отказоустойчивости
- Поддержка систем хранения с подключением по DCB/FCoE



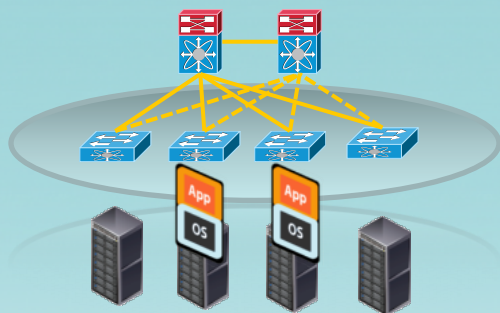
# Ключевые технологии Cisco для ЦОД

- Консолидация транспорта
  - FCoE и DCB
- Эволюция дизайна сети ЦОД
  - VPC
  - FEX
  - FabricPath/TRILL
- Сетевая поддержка виртуализации
  - Nexus 1000V
  - VM-FEX/Adapter-FEX
- Надёжная и производительная связь ЦОД
  - OTV
  - LISP

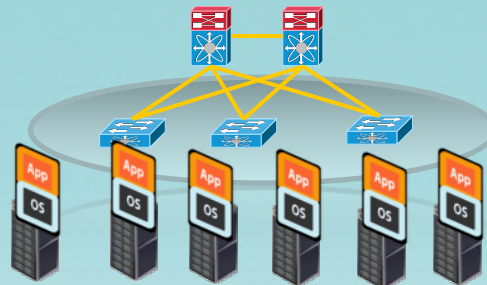


# Развитие архитектуры сети с Cisco Nexus

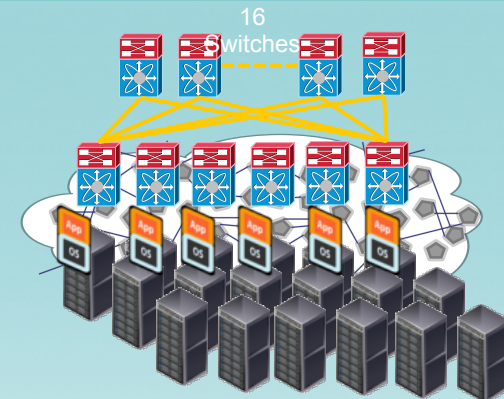
## Spanning-Tree



## vPC



## FabricPath



Активных путей

Один

Два

16

Производительность блока

До 15 Тбит/с

До 30 Тбит/с

До 240 Тбит/с

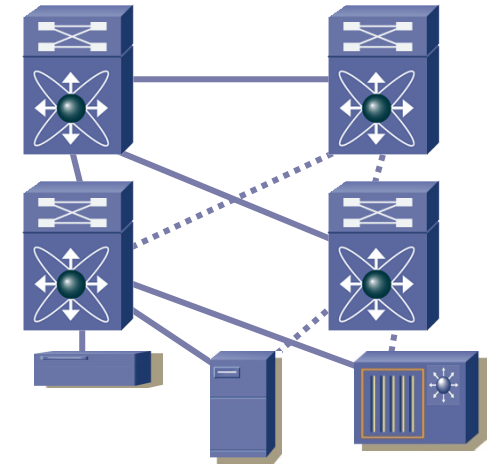
Масштабируемость Layer 2

Виртуализация инфраструктуры и производительность

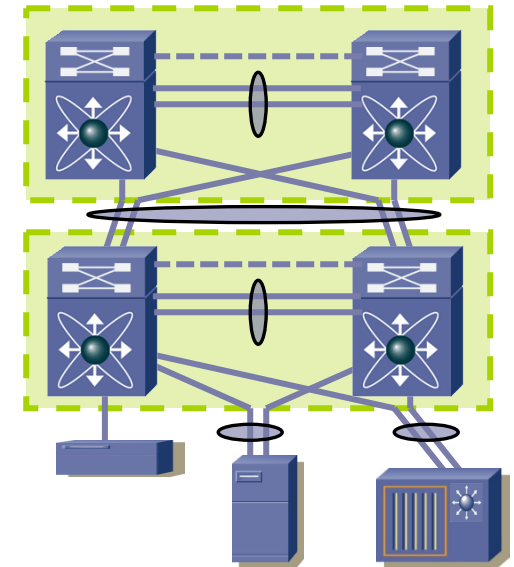
# Virtual Port Channel (VPC) на Cisco Nexus

## Обзор функции

- Возможность организации агрегированного канала (port channel) приходящего на два разных коммутатора
- Избавляет от опоры на STP
- Использование полосы всех имеющихся соединений
- Быстрая сходимость при отказе устройства или канала
- Обеспечение отказоустойчивости и масштабируемости при подключении серверов
- Сокращение CAPEX и OPEX
- Совместимость со всеми функциями



Без vPC

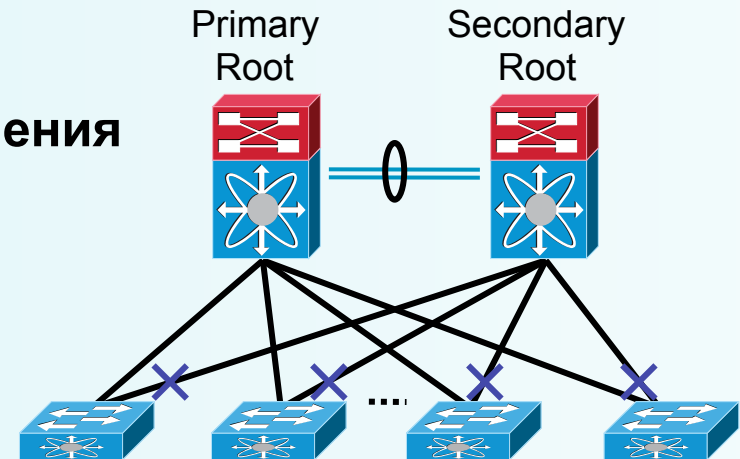


С использованием vPC

# Миграция на VPC с традиционного дизайна

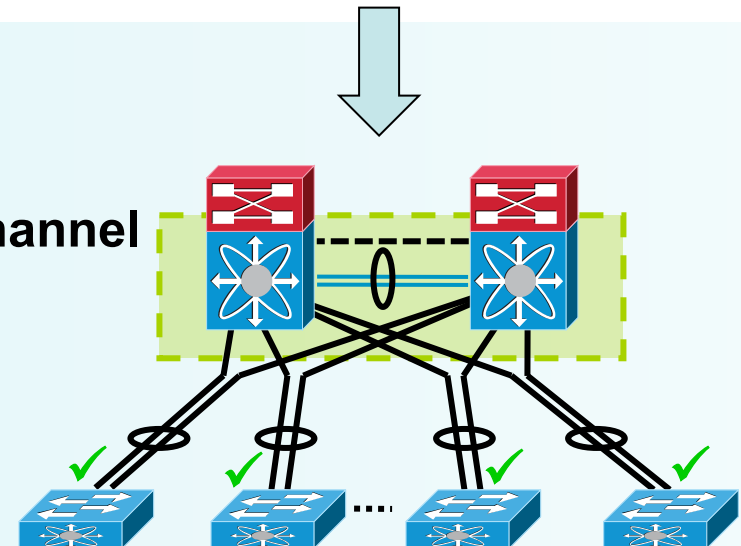
- **Без vPC**

- STP блокирует резервные соединения
- Балансировка загрузки по VLAN
- «Петли» заблокированные STP
- Сбой протокола → 💣\*



- **С vPC**

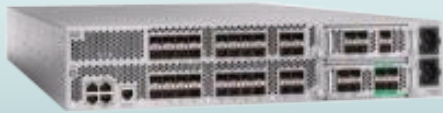
- Соединения не заблокированы
- Балансировка с помощью EtherChannel
- Выше производительность
- Отсутствие «петель»



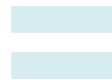
# Виртуальное модульное шасси Nexus 5000 + Nexus 2000

Зима 2009

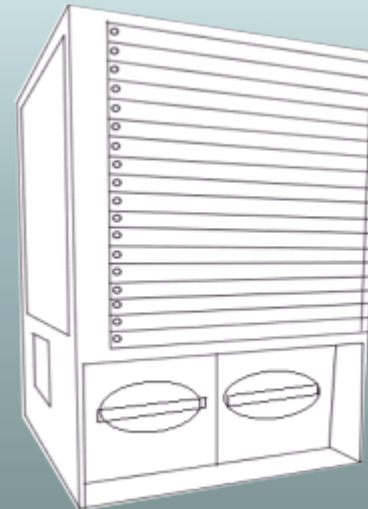
Nexus 5500



Nexus 2000



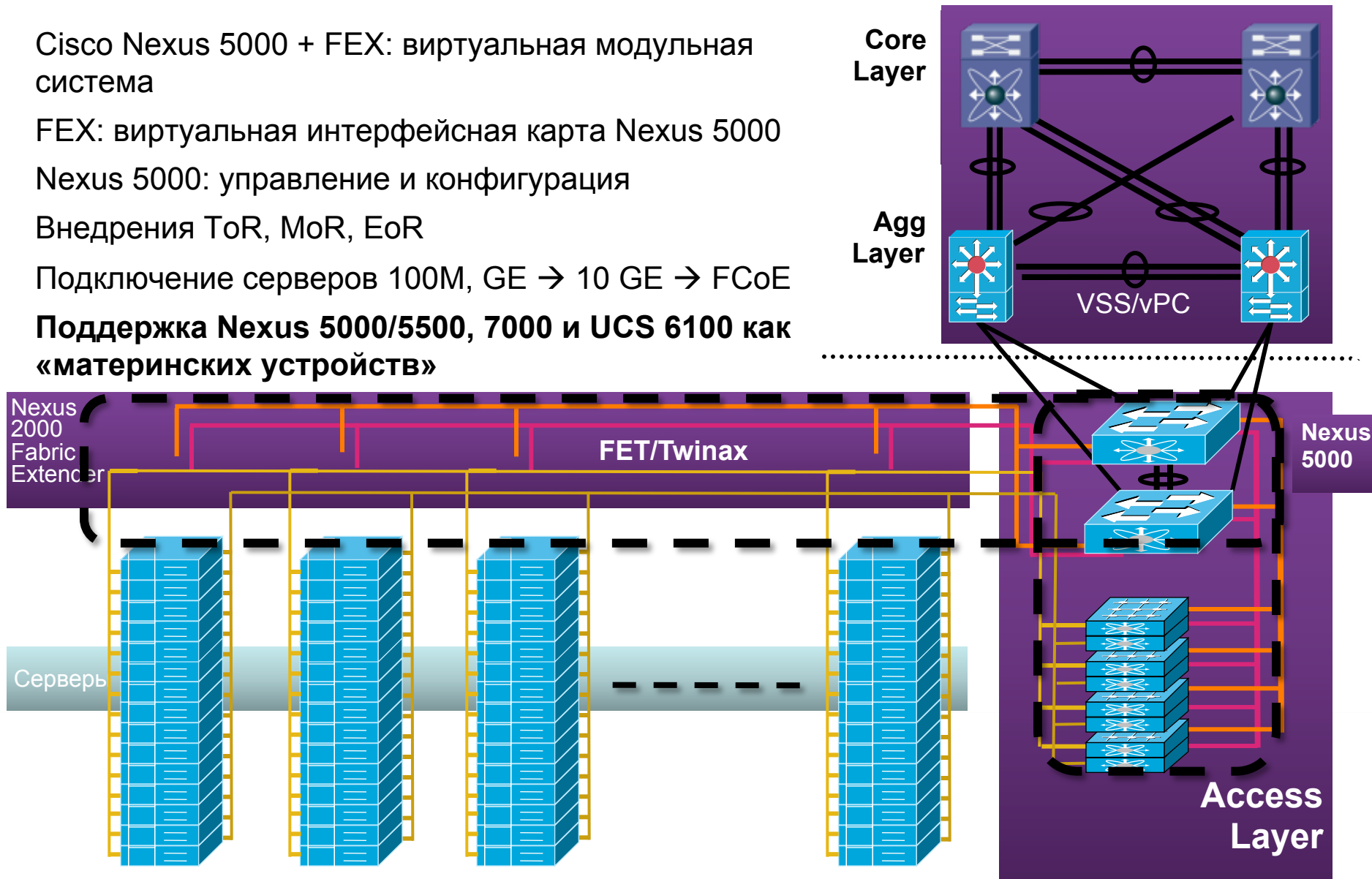
Виртуальное модульное  
шасси



- Nexus 2000 FEX выполняет роль виртуальной карты для Nexus 5000
- Единый конфигурационный файл на Nexus 5000
- Между FEX и Nexus 5000 не используется STP

# Fabric Extender (FEX): унификация архитектуры уровня доступа

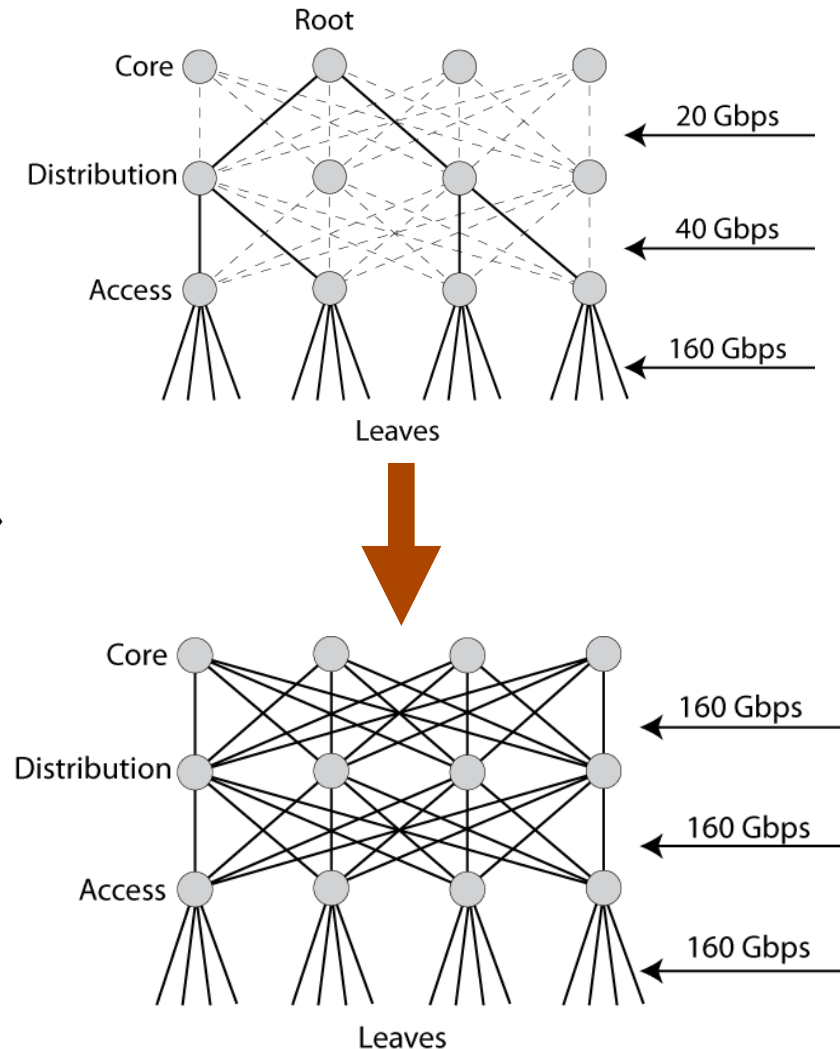
- Cisco Nexus 5000 + FEX: виртуальная модульная система
- FEX: виртуальная интерфейсная карта Nexus 5000
- Nexus 5000: управление и конфигурация
- Внедрения ToR, MoR, EoR
- Подключение серверов 100M, GE → 10 GE → FCoE
- **Поддержка Nexus 5000/5500, 7000 и UCS 6100 как «материнских устройств»**



# Потребность в L2MP

## Spanning Tree превращает многосвязную сеть в дерево

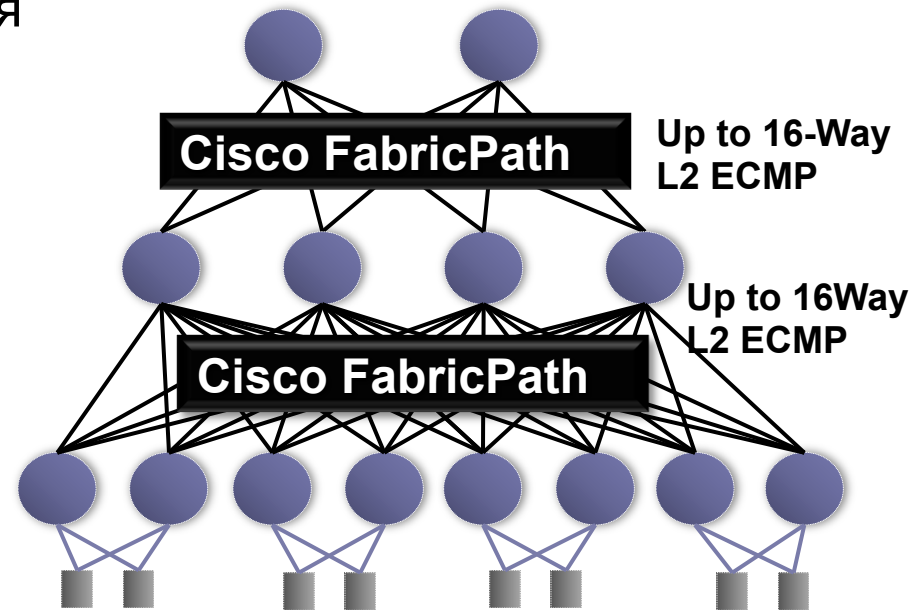
- Нужно альтернативное решение, позволяющее:
  - Задействовать все соединения
  - Нарращивать производительность путем увеличения числа связей
  - Принципиально исключить возможность бесконечных «петель»
  - Обеспечить быструю и надежную сходимость
- ...и всё это – для коммутации на втором уровне!



# Cisco FabricPath: ключевые возможности

Лето 2010

- Продолжение возможностей TRILL
- Маршрутизация на втором уровне с использованием до 16 альтернативных путей (ECMP)
- Заголовок FabricPath: иерархическая адресация со встроенным предотвращением «защипывания»
- Выучивание MAC «по диалогам»: эффективное использование аппаратных ресурсов
- Совместимость с «классическим» Ethernet
  - VPC+ обеспечивает VPC в L2MP сеть
  - STP Boundary Termination
- Поддержка множественных топологий – возможность Traffic Engineering



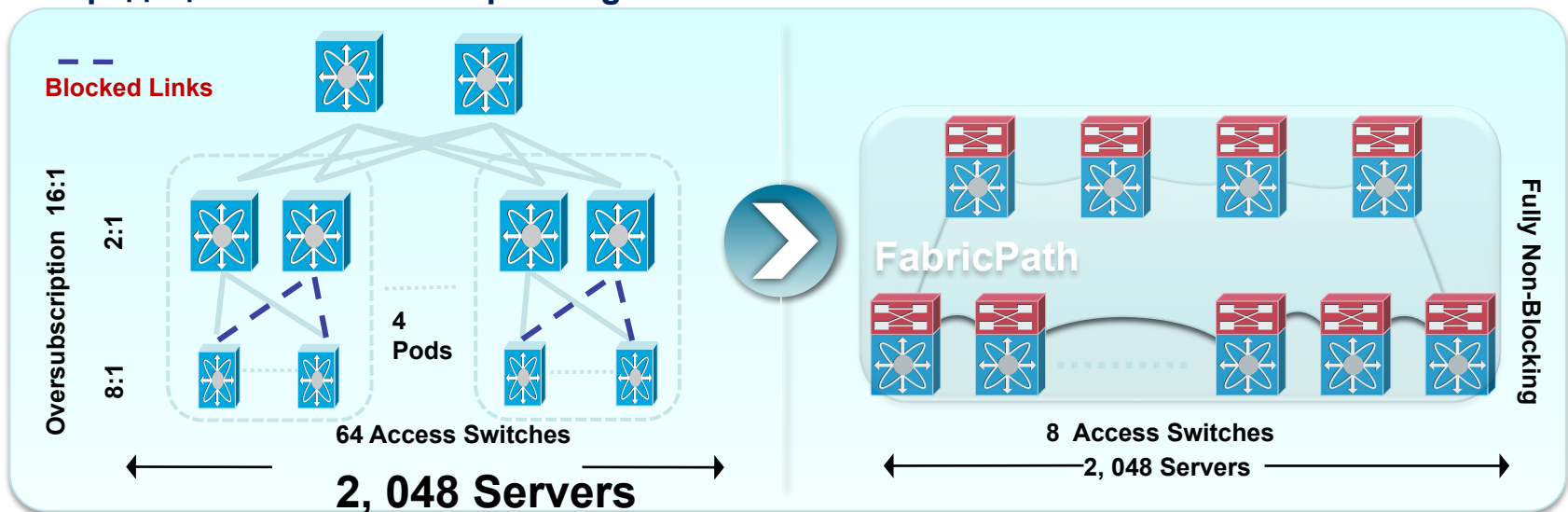
# Использование FabricPath в корпоративном ЦОД

## Альтернатива традиционной архитектуре с STP

- Существенное повышение производительности
- Повышение надёжности
- Упрощение эксплуатации
  - меньше устройств
  - проще настройка

Традиционная сеть со Spanning Tree

Сеть на базе FabricPath

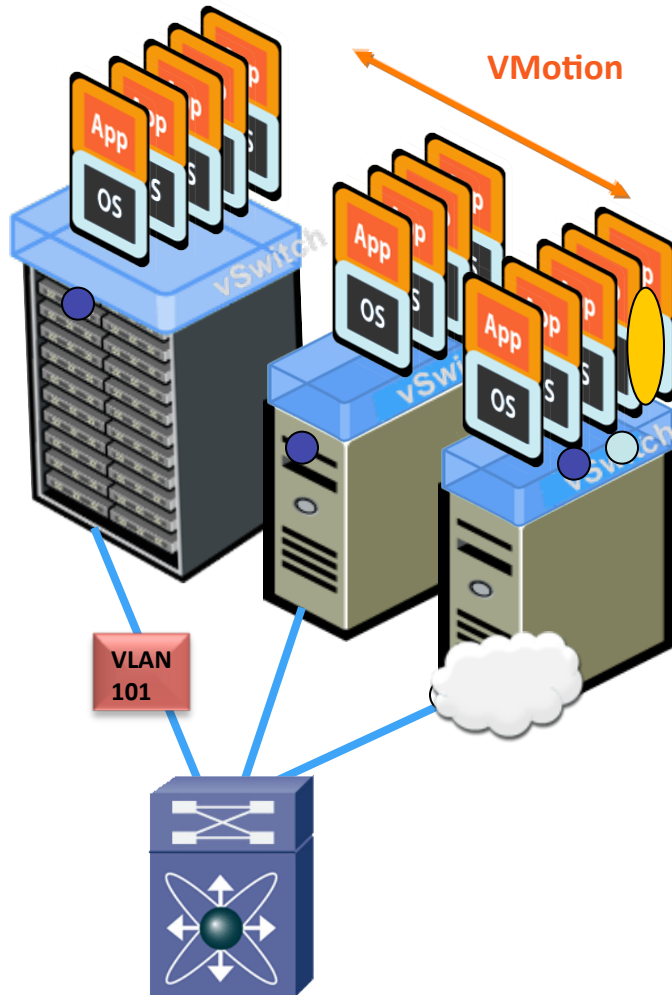


# Ключевые технологии Cisco для ЦОД

- Консолидация транспорта
  - FCoE и DCB
- Эволюция дизайна сети ЦОД
  - VPC
  - FEX
  - FabricPath/TRILL
- Сетевая поддержка виртуализации
  - Nexus 1000V
  - VM-FEX/Adapter-FEX
- Надёжная и производительная связь ЦОД
  - OTV
  - LISP



# Распознать в сети трафик виртуальной машины



## Проблемы:

- VMotion может переместить VM на другой сервер. Политика должна следовать вслед за VM
- Сеть не «видит» локально коммутируемый трафик и не может применить к нему политику
- Сеть не может выделить на порту трафик конкретной VM
- Изменение модели эксплуатации

## Сетевая поддержка виртуализации:

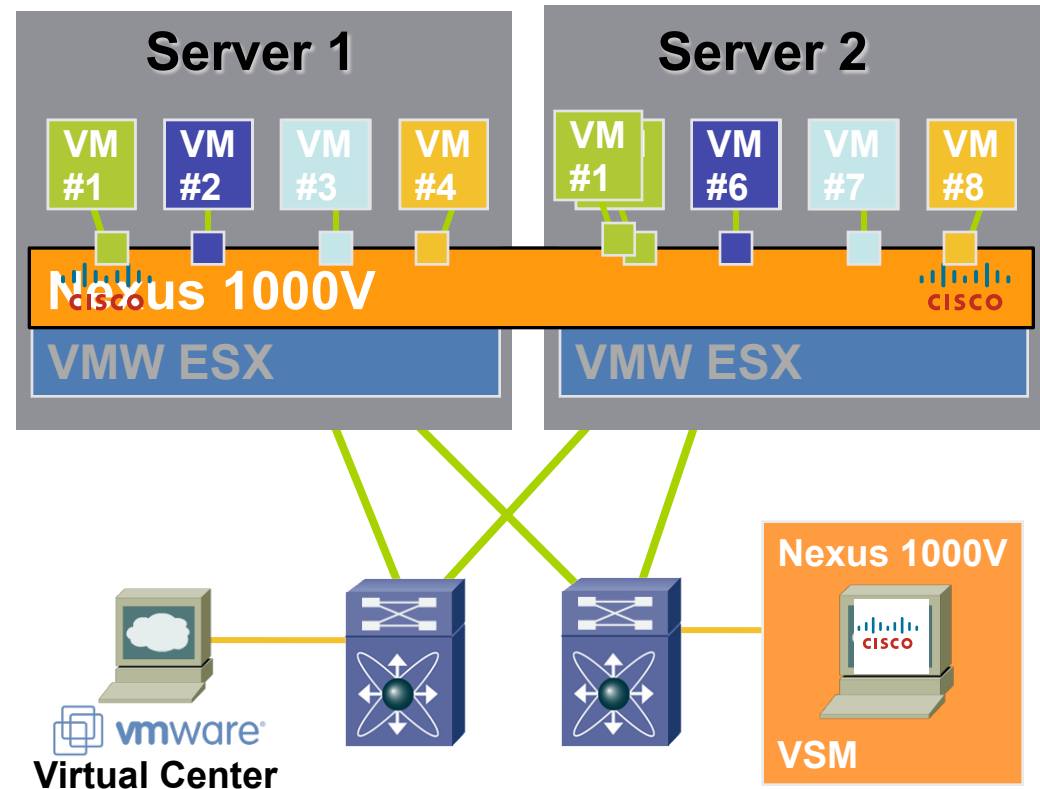
- Расширяет сеть до VM
- Общие с физическими коммутаторами функции сетевых сервисов
- Скоординированное с VM управление

# Cisco Nexus 1000V

## Виртуальный распределенный программный коммутатор

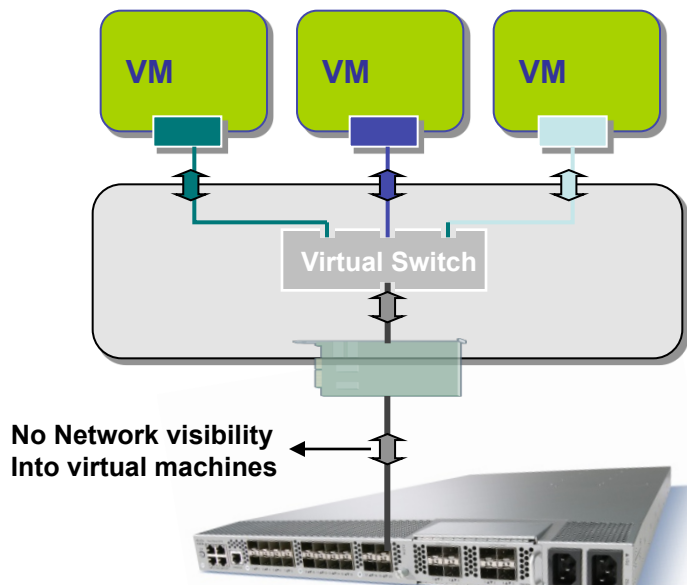
- Nexus 1000V обеспечивает полнофункциональную коммутацию для VMWare ESX
- Ключевые возможности:
  - Управление VM по политикам
  - Функции безопасности, поддержка Netflow, ERSPAN, мультикаста, etherchannel
  - Мобильность настроек сети, безопасности и мониторинга
  - Сохраняет эксплуатационную модель
- Сохранение политик и связи с сетью при использовании VMotion

BEST OF  
vmworld 2008

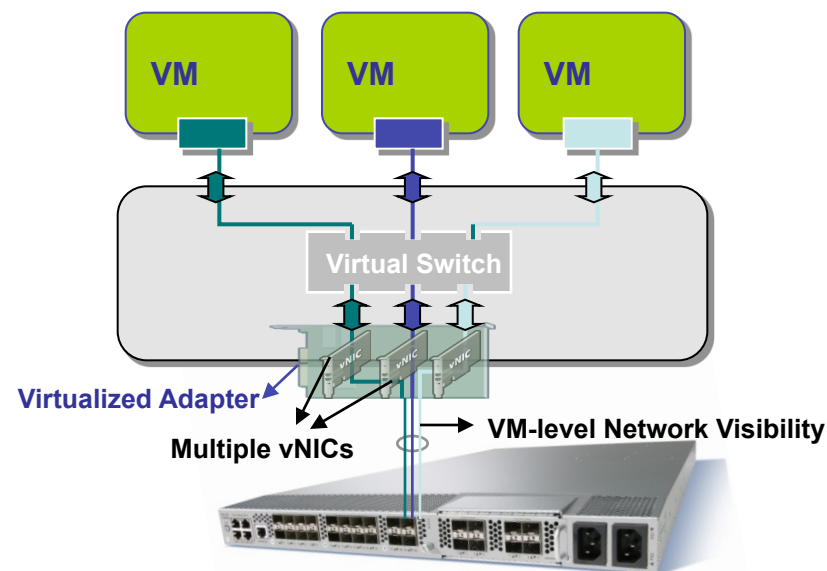


# VM-FEX с использованием виртуализированного адаптера Cisco VIC

Сейчас



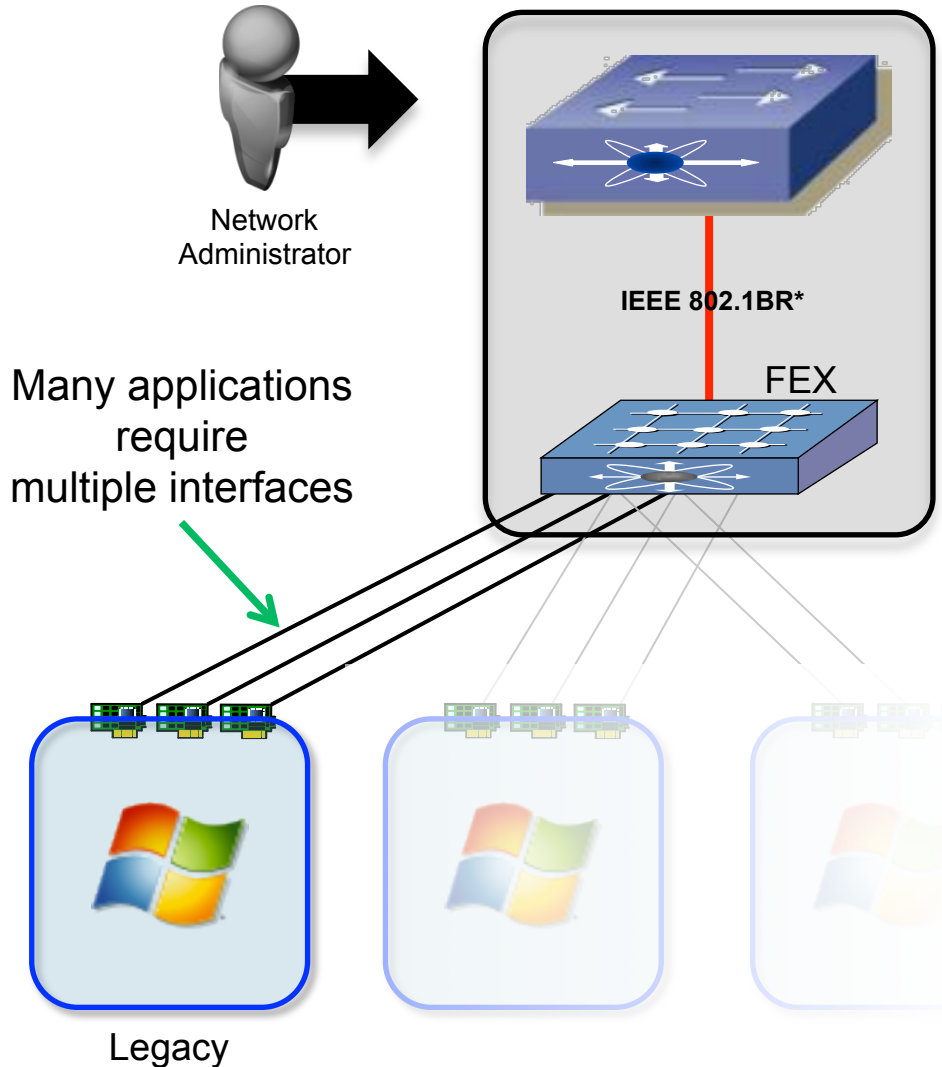
Cisco VM-FEX



- **Изоляция подключений VM – защищенная среда виртуализации**
  - Применение сетевых политик к конкретным VM
- **«Видимость» для сети отдельных VM – упрощение диагностики**
  - Однозначная идентификация трафика VM

# Эволюция технологии Fabric Extender

Распределённый коммутатор до уровня стойки, интерфейсов сервера и VM



## Единое устройство

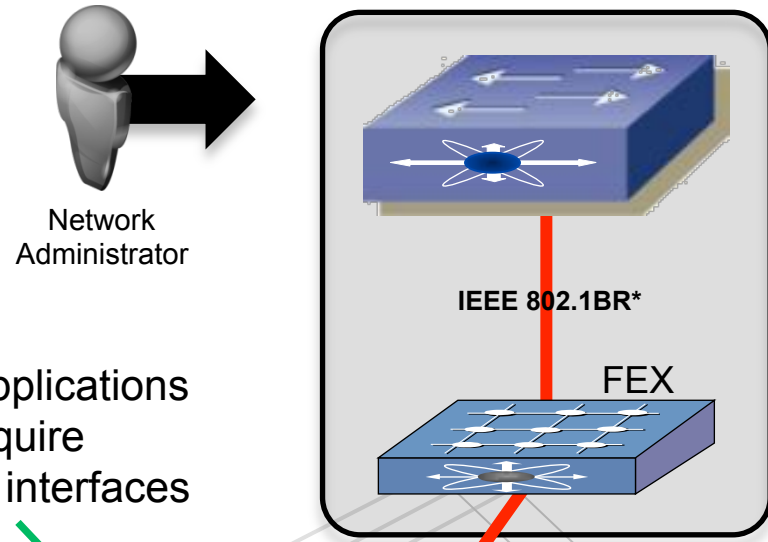
### Fabric Extender

- Консолидация управления сетью
- FEX является частью «родительского коммутатора»
- Использует пре-стандартную реализацию IEEE 802.1BR

\*IEEE 802.1BR pre-standard

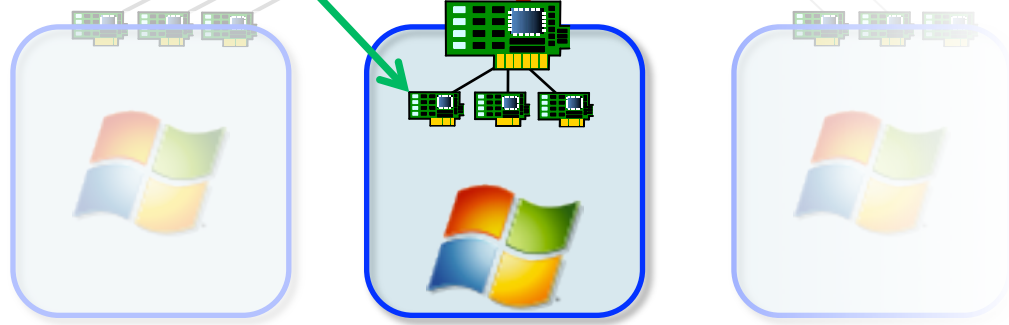
# Эволюция технологии Fabric Extender

Распределённый коммутатор до уровня стойки, интерфейсов сервера и VM



Network Administrator

Many applications require multiple interfaces



Legacy

Adapter FEX

## Единое устройство

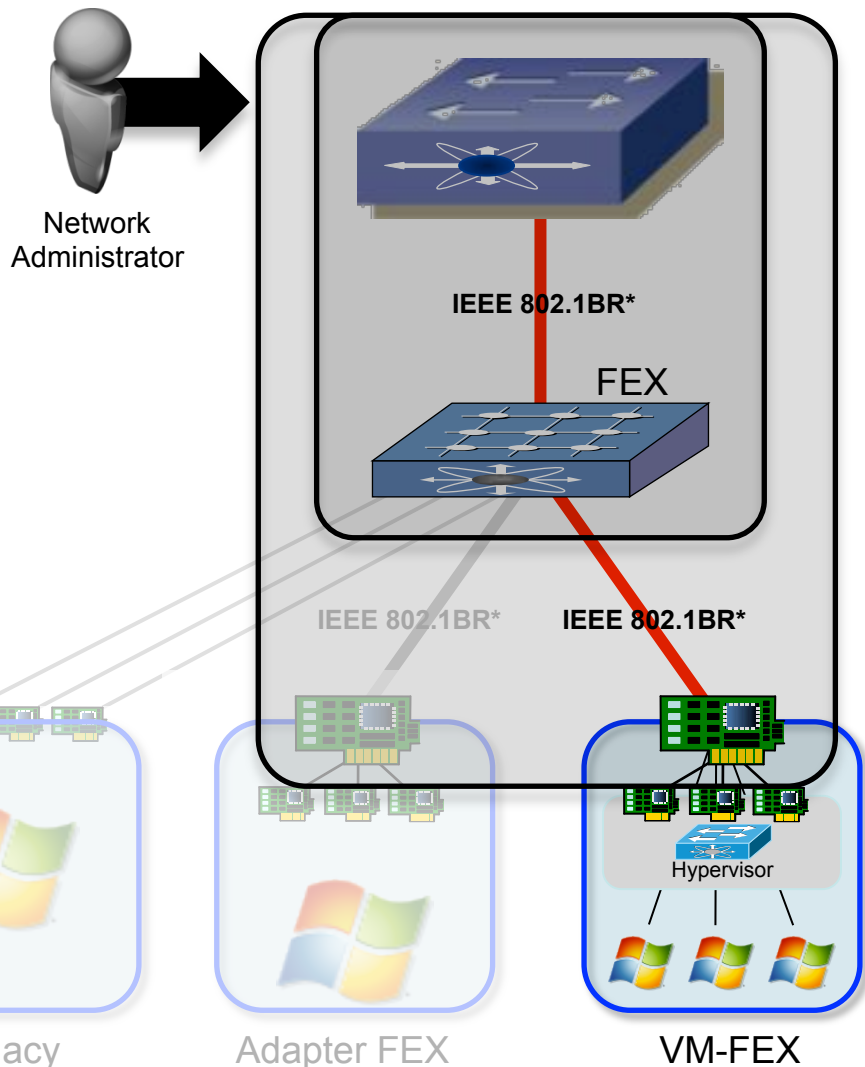
### Adapter FEX

- Консолидация многих интерфейсов в единое 10GE подключение
- Расширение сети внутрь сервера
- Использует пре-стандартную реализацию IEEE 802.1BR

\*IEEE 802.1BR pre-standard

# Эволюция технологии Fabric Extender

Распределённый коммутатор до уровня стойки, интерфейсов сервера и VM



## Единое устройство

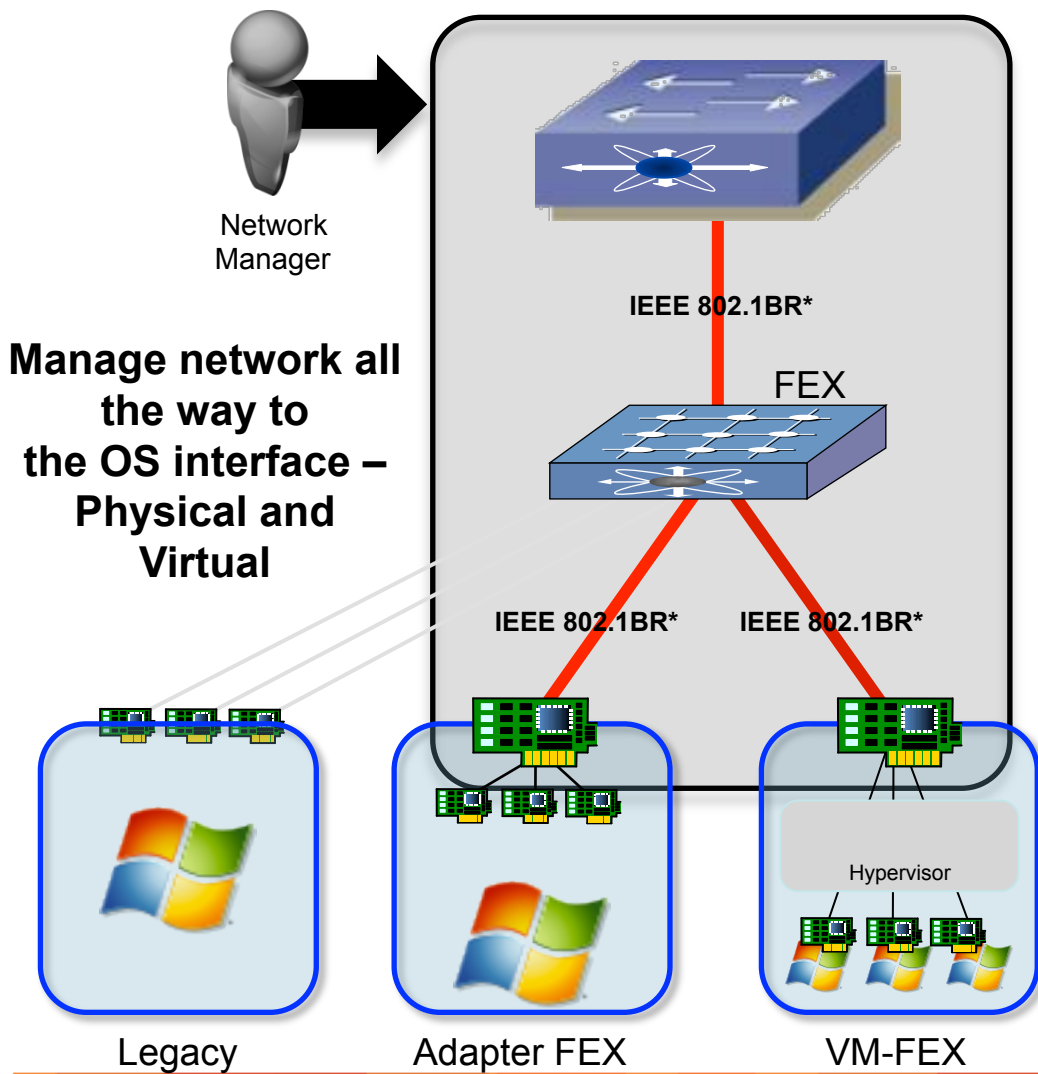
### VM-FEX

- Консолидация физической и виртуальной сети
- Каждая VM получает порт на распределённом коммутаторе
- Использует пре-стандартную реализацию IEEE 802.1BR

\*IEEE 802.1BR pre-standard

# Эволюция технологии Fabric Extender

Распределённый коммутатор до уровня стойки, интерфейсов сервера и VM



## Единое устройство

- Порты коммутатора
- Порты FEX
- Виртуальные адаптеры
- Виртуальные машины

## Fabric Extender

- Консолидация управления
- FEX выглядит линейной картой головного коммутатора

## Adapter FEX

- Консолидация многих интерфейсов в единое 10GE подключение
- Расширение сети внутрь сервера

## VM-FEX

- Консолидация физической и виртуальной сети
- Каждая VM получает порт на распределённом коммутаторе

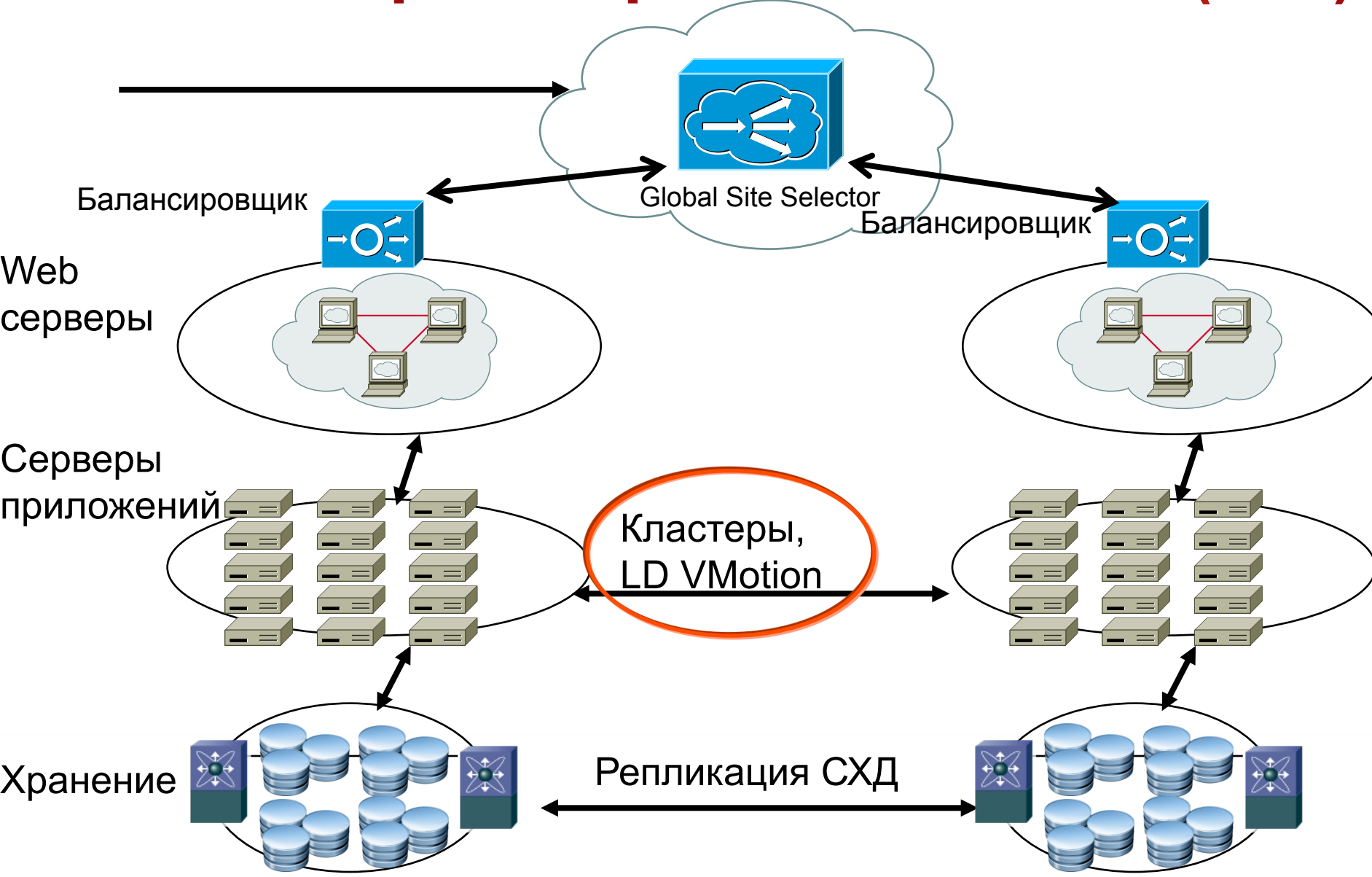
\*IEEE 802.1BR pre-standard

# Ключевые технологии Cisco для ЦОД

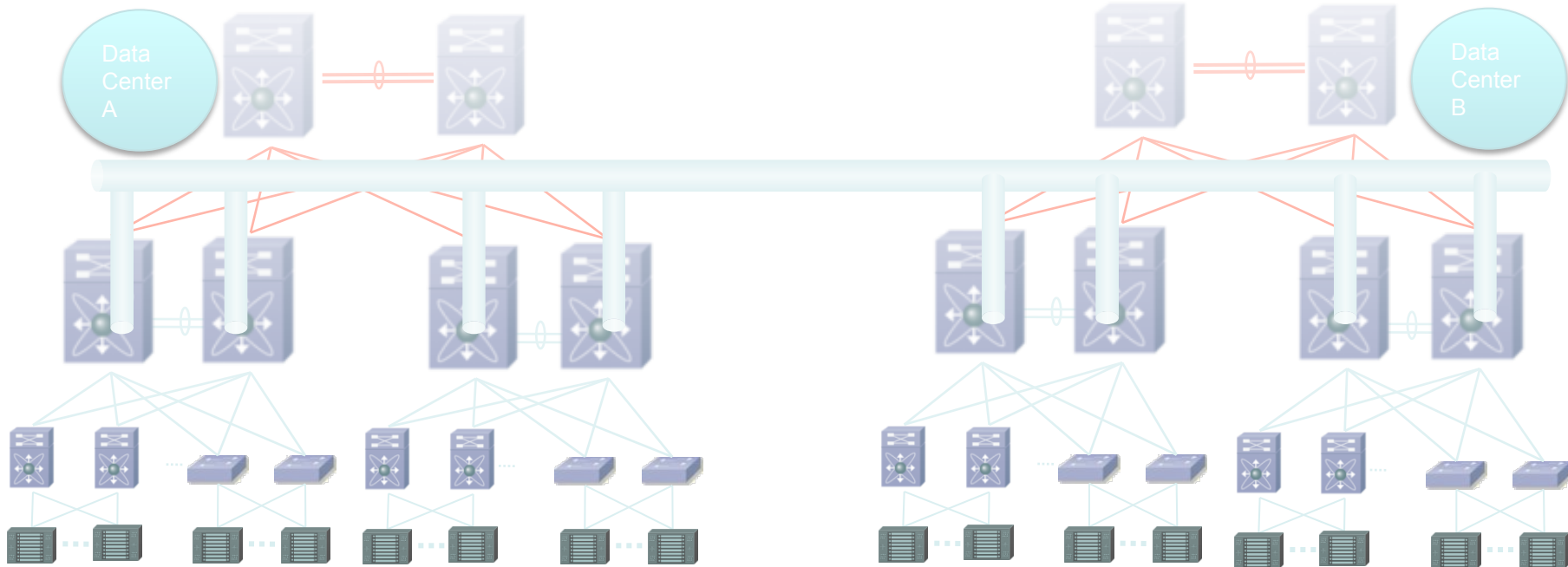
- Консолидация транспорта
  - FCoE и DCB
- Эволюция дизайна сети ЦОД
  - VPC
  - FEX
  - FabricPath/TRILL
- Сетевая поддержка виртуализации
  - Nexus 1000V
  - VM-FEX/Adapter-FEX
- Надёжная и производительная связь ЦОД
  - OTV
  - LISP



# Связь центров обработки данных(DCI)



# «Растягивание» LAN: внутри и между ЦОД



- Ряд приложений требуют смежности на 2 уровне
  - Кластеры (Veritas, MSFT)
  - vMotion
  - «Доморощенные» приложения

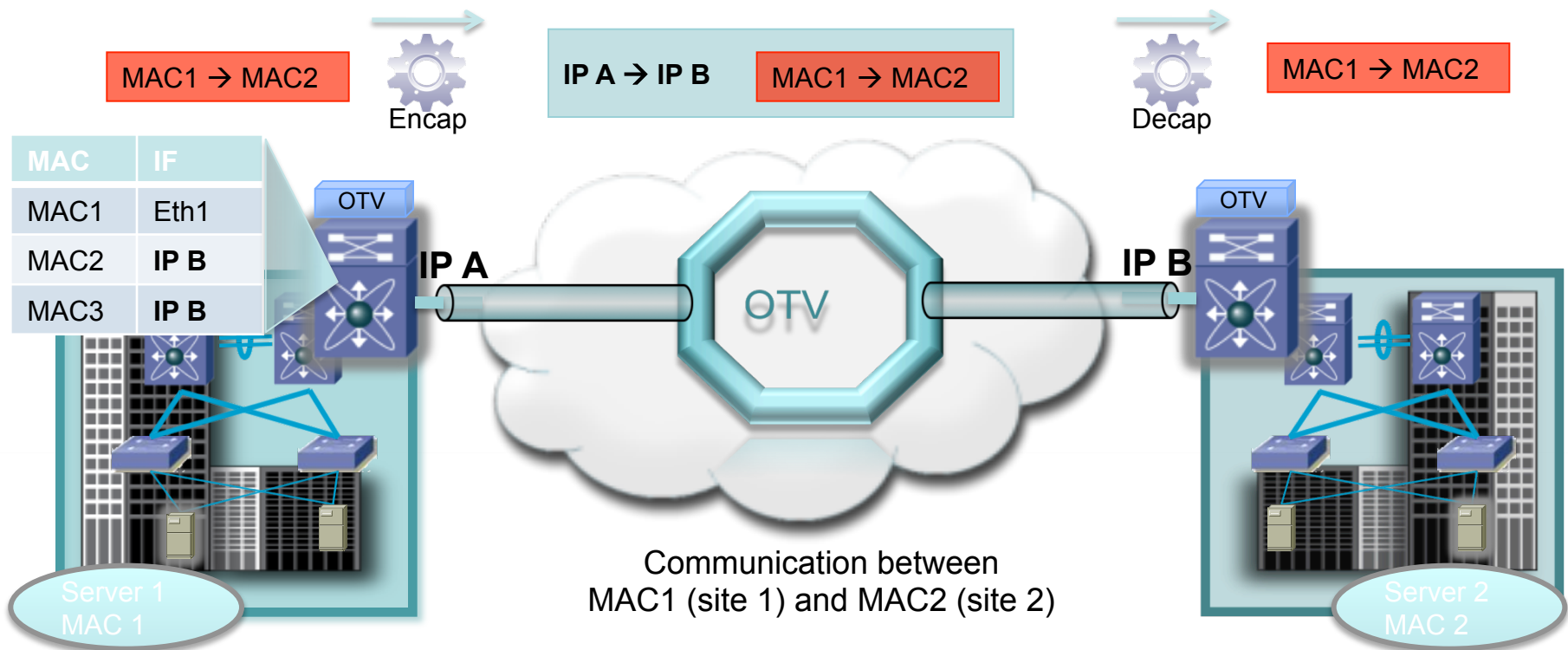
- Миграция серверов
- Высокая доступность
- Распределенные служебные и прикладные сервисы

# Overlay Transport Virtualization

## Принципы работы протокола



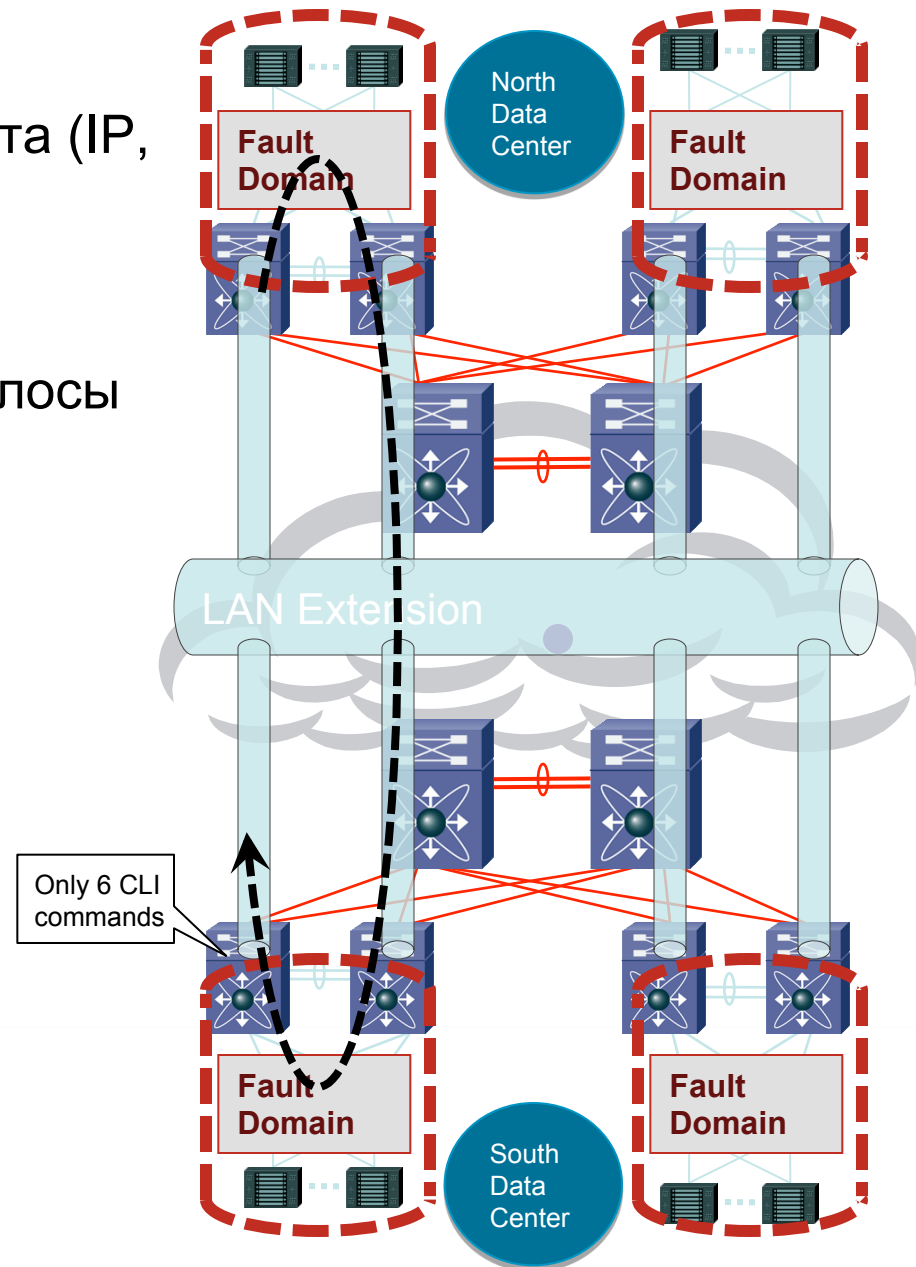
- Ethernet трафик инкапсулируется в IP: “MAC in IP”
- Динамическая инкапсуляция с использованием таблицы маршрутизации MAC
- Не строится Pseudo-Wire или туннель



# Проблемы «растягивания» LAN

## Решаемые OTV

- Работа поверх любого транспорта (IP, MPLS)
- Изоляция доменов сбоев
- Независимость сайтов
- Оптимальное использование полосы
- Встроенная отказоустойчивость
- Встроенная защита от «петель»
- Связь многих сайтов
- Масштабируемость
  - VLANs, сайты, MACs
  - ARP, broadcasts/floods
- Простота настройки
- Легкость добавления сайтов



# Оптимальный путь

## В чём именно проблема?

10.1.1.0/25 & 10.1.1.128/25 advertised into L3  
DC A is the primary entry point



10.1.1.0/24 advertised into L3  
Backup should main site go down

Layer 3 Core

DCI IP  
Cloud

Agg

Access

Agg

Access

Node A

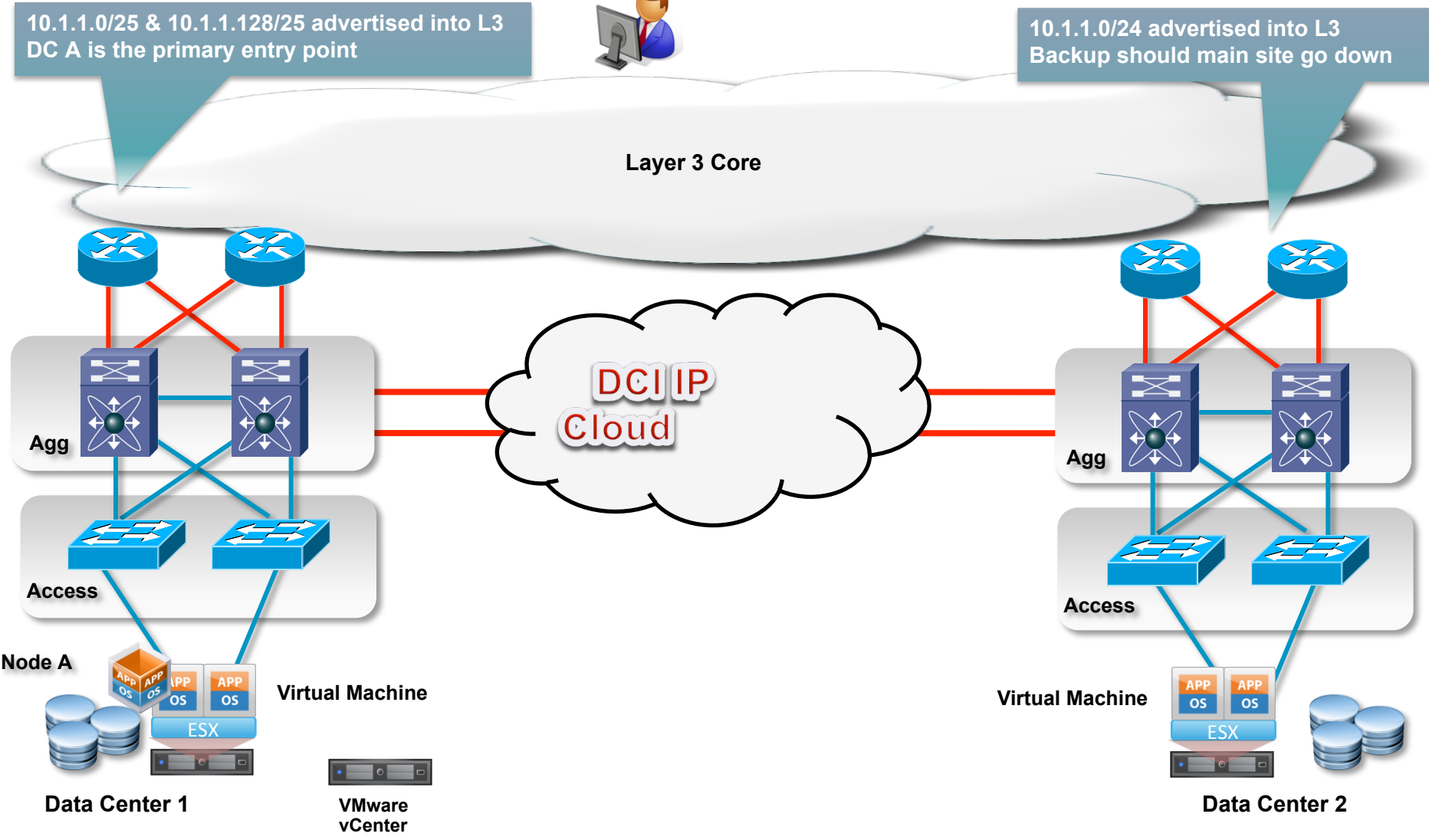
Virtual Machine

Virtual Machine

Data Center 1

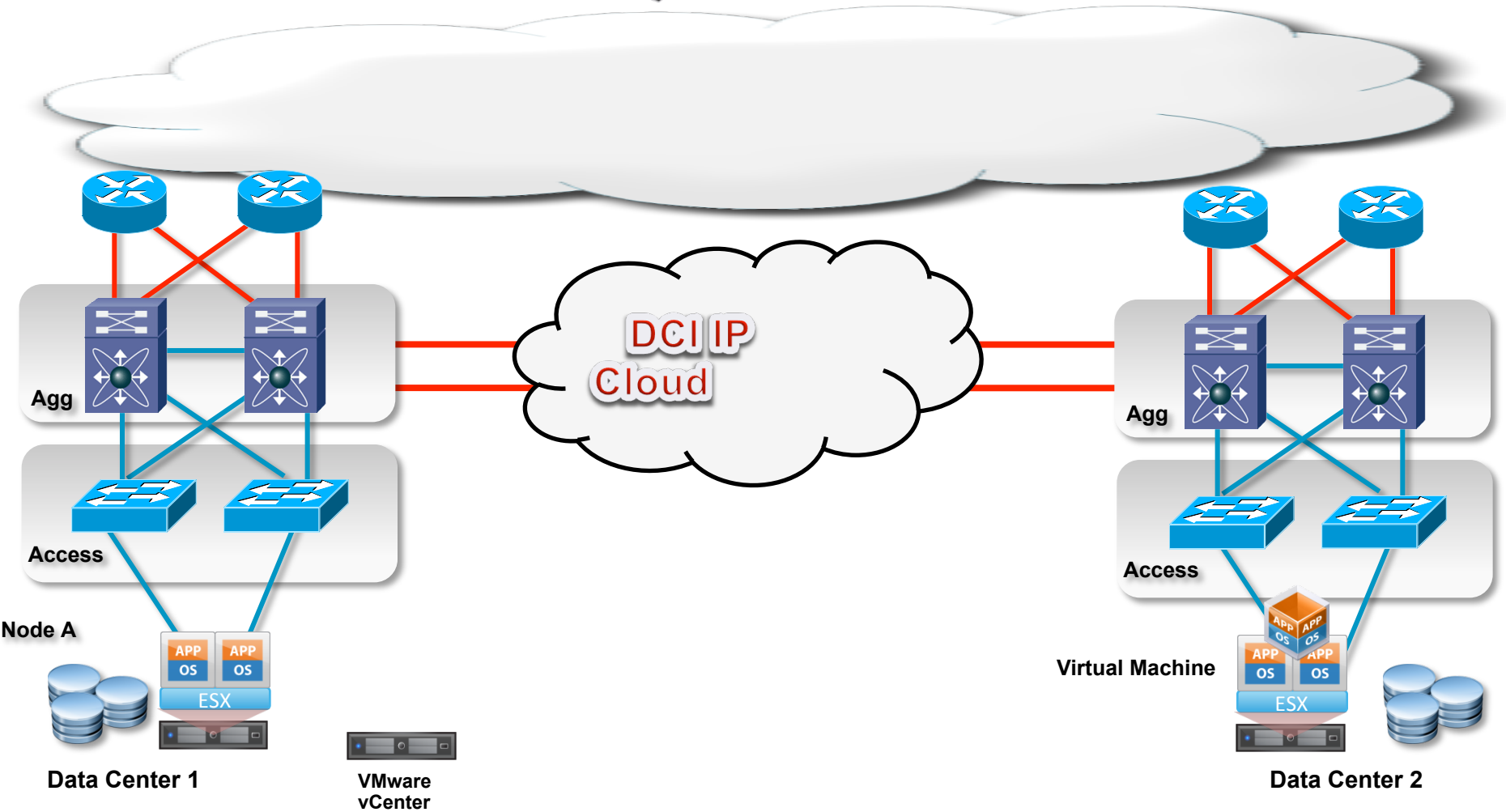
VMware  
vCenter

Data Center 2



# Оптимальный путь

Хотелось бы так...



Data Center 1

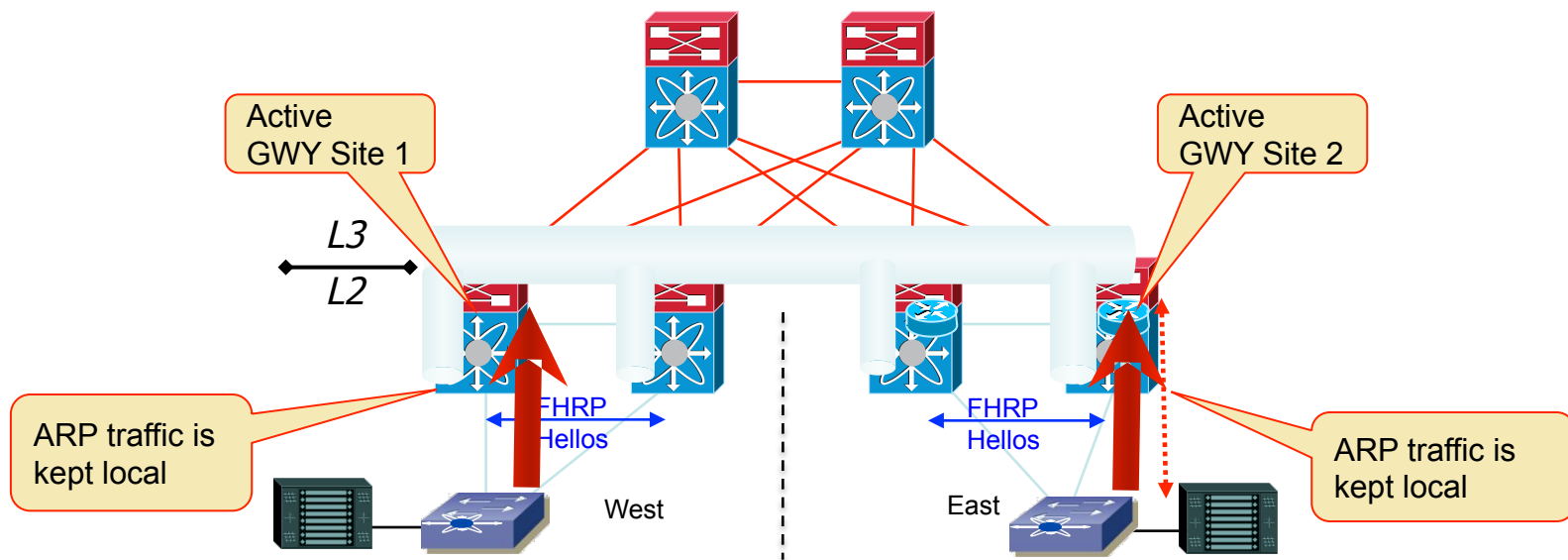
VMware  
vCenter

Data Center 2

# Оптимизация пути «на выход»

## Локализация FHRP с помощью OTV

- Одна и та же HSRP группа на всех сайтах с тем же виртуальным MAC адресом
- Каждый сайт обеспечивает исходящую маршрутизацию
- **OTV локализует исходящий трафик за счёт фильтрации HSRP hello сообщений между сайтами**
- ARP запросы перехватываются на OTV edge устройстве чтобы обеспечить ответы именно от локального шлюза



# Оптимизация пути «на вход»

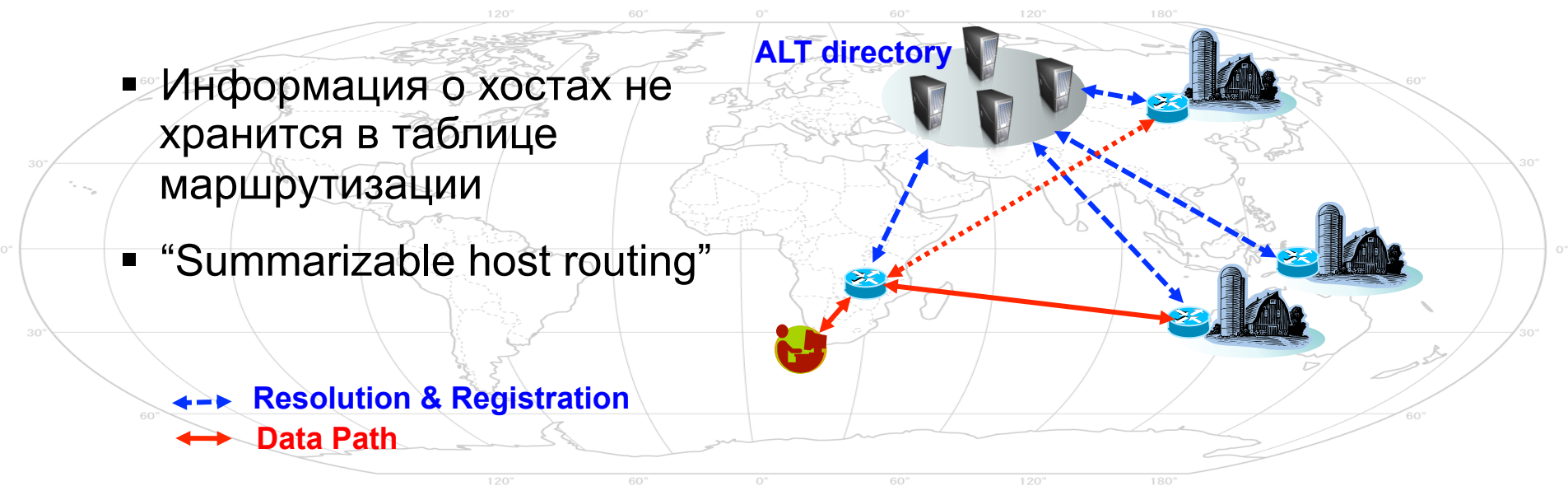
## Locator-ID Separation Protocol (LISP)

- Отделяет идентификатор сервиса (IP адрес) от его местоположения
- Маршрутизация исходя из местоположения, а не адреса хоста
- Соотношение адреса и его местоположение хранятся в директории
- Поиск метоположения IP адреса по информации из директории
- Инкапсуляция трафика (IP in IP) и передача по месту нахождения хоста
- Директория – распределенная база данных

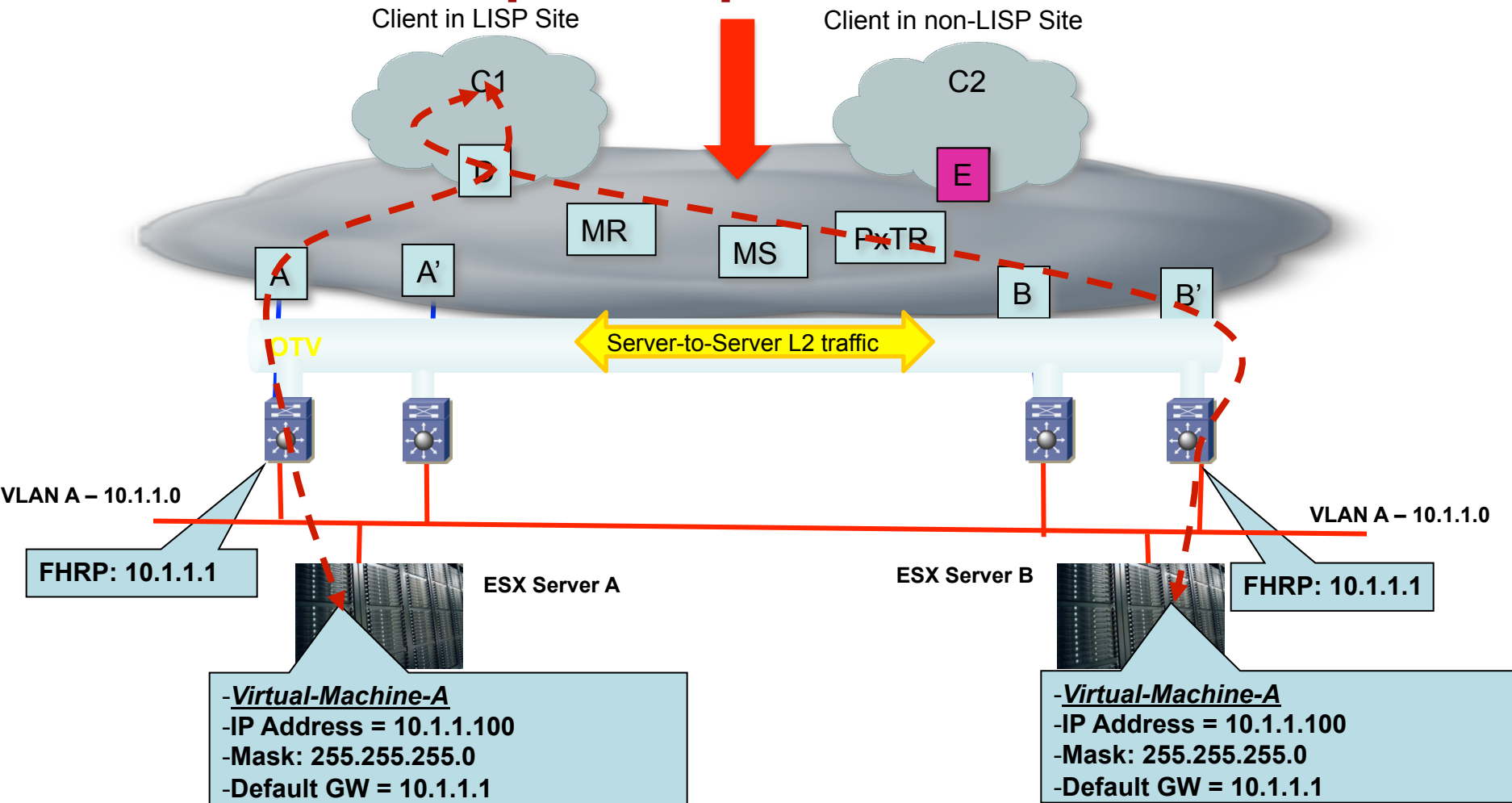
- Информация о хостах не хранится в таблице маршрутизации
- “Summarizable host routing”

↔ Resolution & Registration  
↔ Data Path

ALT directory



# Оптимальный транспорт с помощью LISP и OTV



## LISP: L3 Client-to-Server

- Оптимизация маршрутизации с детальной информацией о местоположении
- Оптимизация мобильности внутри или между подсетями
- Масштабирование прикладных сервисов

## OTV: L2 Server-to-Server

- Оптимизация расширения LAN
- Распределение прикладных систем
- Надежная связь на втором уровне для мобильности виртуальных сервисов и кластерных систем

Cisco Expo 2011



# Спасибо!

Просим Вас заполнить анкеты.  
Ваше мнение очень важно для нас!

innovate *together*