



Networking
For everyone

Уровень доступа

Основные задачи

- Предоставление отказоустойчивого доступа для серверов
- Коммутация фреймов в сторону уровня агрегации
 - маршрутизация в современных Leaf/Spine дизайнах с наложенной сетью
- Обеспечение мобильности



Что такое мобильность?

- Такие технологии как VMware vMotion позволяют ВМ прозрачно переезжать с одного ESXi хоста на другой
 - чаще всего IP настройки не должны изменяться
- Результатом является необходимость обеспечения непрерывного широковещательного L2 сегмента



Как обеспечить мобильность?

- Классическое растягивание L2 сегмента
 - spanning-tree
- Построение L3 сети и растягивание L2 с помощью наложенных технологий
 - OTV, EVPN
 - Fabricpath
 - VPLS



Что такое отказоустойчивость?

- Отказоустойчивость – способность сети безболезненно переживать отказ одного или более компонентов
- Конвергенция (возвращение в работоспособное состояние) должно занимать минимальный период времени
 - какой – каждый сам решает для себя



Обеспечение отказоустойчивости

- Отказоустойчивость может достигаться на нескольких уровнях
- Уровень ПО
- Сетевой уровень
 - уровень сетевых протоколов
 - уровень железа



Отказоустойчивость на уровне железа

- Современные коммутаторы ЦОД имеют внутри себя дублирующие компоненты
 - 2 блока питания
 - 2 управляющих модуля
 - в т.ч. container-based SUP 😊
 - несколько линейных карт
 - и пр.



Отказоустойчивость сетевых протоколов

- L2 отказоустойчивость
 - чаще всего STP
- L3 отказоустойчивость
 - обычная маршрутизация
 - ECMP (чаще) или Active/Standby



Отказоустойчивость с помощью STP

- NX-OS поддерживает RSTP и MSTP
- STP работает как Active/Standby с т.з. утилизации полосы пропускания на интерфейсах
- Нет таких понятий как Graceful Restart
- Долгая конвергенция
- Результат – STP в ЦОД должен быть сведён к минимуму



Как обеспечить подключение серверов?

- Сервер – не всегда Bare Metal, он может быть со встроенным коммутатором
- Для увеличения полосы используем агрегационные протоколы (LACP)
- Но что делать с отказоустойчивость между коммутаторами доступа?
 - можно сделать резервирование на уровне серверов
 - можно сделать Multi-Chassis агрегационный канал



Агрегация интерфейсов Ethernet

- Объединение нескольких физических портов (агрегируемых, bundled) Ethernet в один агрегатный логический порт (bundle)
- Технология известна как
 - Port Channel
 - Link Aggregation (LAG)
 - NIC Teaming



Как работает

- Агрегированный канал состоит из 2-ух частей
 - Интерфейс Port-Channel
 - Логический интерфейс, который представляет собой бандл интерфейсов
 - Member интерфейс
 - Физический интерфейс, являющийся частью бандла
- Основная цель – скрыть Member интерфейсы от выше-стоящих протоколов
 - Напр, STP должен видеть не 2x1 Гб интерфейса, а 1x2 Гб
 - В результате – Active/Active поведение для STP

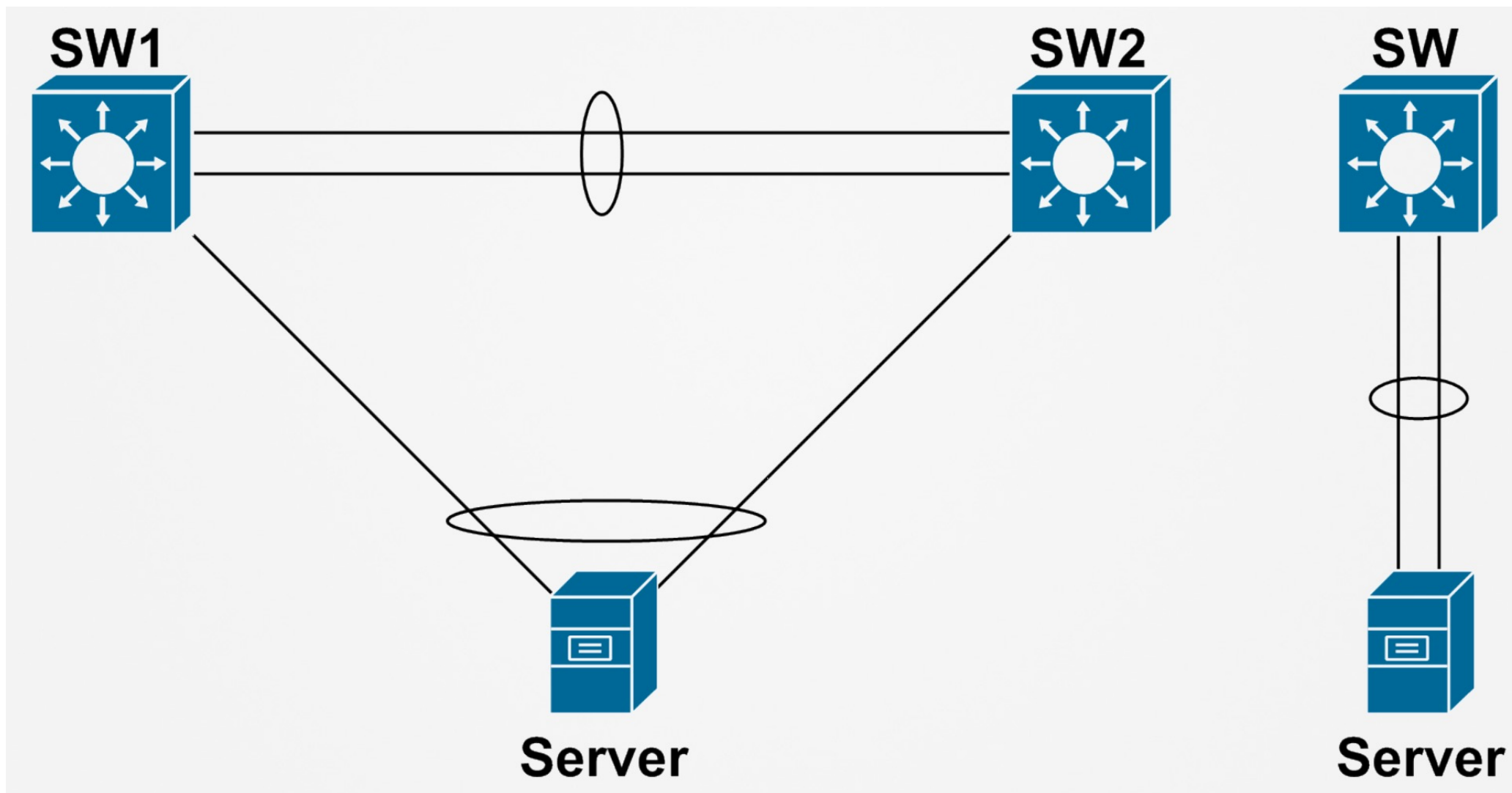


Плюсы и минусы

- Плюсы
 - Проще и дешевле провести апгрейд сети
 - Добавляется отказоустойчивость линка
- Минусы
 - Неравномерная загрузка физических интерфейсов
 - Максимальная пропускная способность не увеличивается



Агрегация между разными шасси



Агрегация между разными шасси

- StackWise
 - Применяется на платформах доступа Cisco Catalyst 3750/3850
 - Синхронизация Control Plane через специальный кабель
 - Логически - одно устройство
- Virtual Switching System (VSS)
 - Применяется на платформах агрегации Cisco Catalyst 4500/6500/6800
 - Синхронизация Control Plane через Virtual Switch Link (VSL)
 - Логически – одно устройство
- Virtual Port-Channel (vPC)
 - Применяется на платформах Cisco Nexus
 - Логически – независимые устройства
 - Частичная синхронизация



Создание агрегационного канала

- Статическая
- Link Aggregation Control Protocol (LACP)
 - стандарт
- Port Aggregation Group Protocol (PAGP)
 - Cisco proprietary
 - Отличается от LACP примерно также, как ISL от 802.1Q



Варианты согласования бандла

- On
 - Нет согласования
- Desirable
 - Активное согласование PAGP
- Auto
 - Пассивное согласование PAGP
- Active
 - Активное согласование LACP
- Passive
 - Пассивное согласование LACP



Бандл поднимется в случаях

- On – On
- Desirable – Desirable
- Desirable – Auto
- Active – Active
- Active - Passive



Балансировка внутри бандла

- Балансировка настраивается только для исходящих кадров
- Варианты балансировки (зависит от платформы):
 - *src-mac*: все кадры от одного MAC-адреса отправляются с одного порта
 - *dst-mac*: все кадры на один и тот же MAC-адрес отправляются с одного порта
 - *src-ip*: все кадры (пакеты) от одного IP отправляются с одного и того же порта
 - *dst-ip*: все кадры на один и тот же IP отправляются с одного и того же порта
 - *src-dst-mac* (-ip): для определения выходного порта берется результат функции XOR от MAC-адресов (IP-адресов) отправителя и получателя



L2 и L3 бандлы

- LAG не зависит от типа порта
 - Т.е. может быть access, trunk, tunnel, L3 и т.д.
- Перевод из L2 в L3 (и наоборот) не может быть сделан «на горячую»
 - Требуется изменение типа физического интерфейса *switchport* -> *no switchport*
 - Использование LAG на L3 – тема дискуссионная





Networking
For everyone

Настройка L2 коммутации



Networking
For everyone