

Проблемы доступности в сети и методы их решения (дублирование агрегации)

Малеванов Кирилл

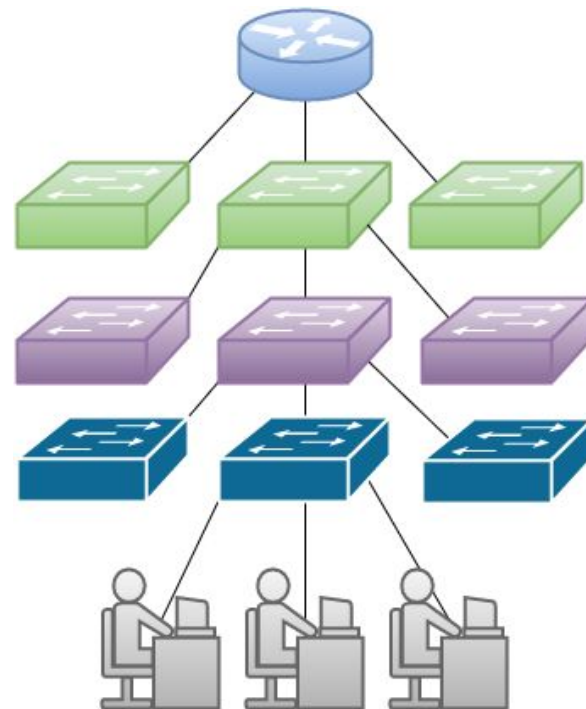
Содержание

- Ядро-агрегация-дистрибьюция-доступ
 - L2 и L3 схемы подключения абонентов
 - L2 как отдельная услуга
 - MPLS на агрегации
- Схема узла агрегации
 - Агрегация как PE
 - Дистрибьюция для уменьшения стоимости порта
- Проблемы при дублировании
 - Стек
 - MC-LAG
 - L3
 - MPLS
- Тестирование оборудования
- Выводы

Схема сети

Классическая схема сети предусматривает несколько уровней:

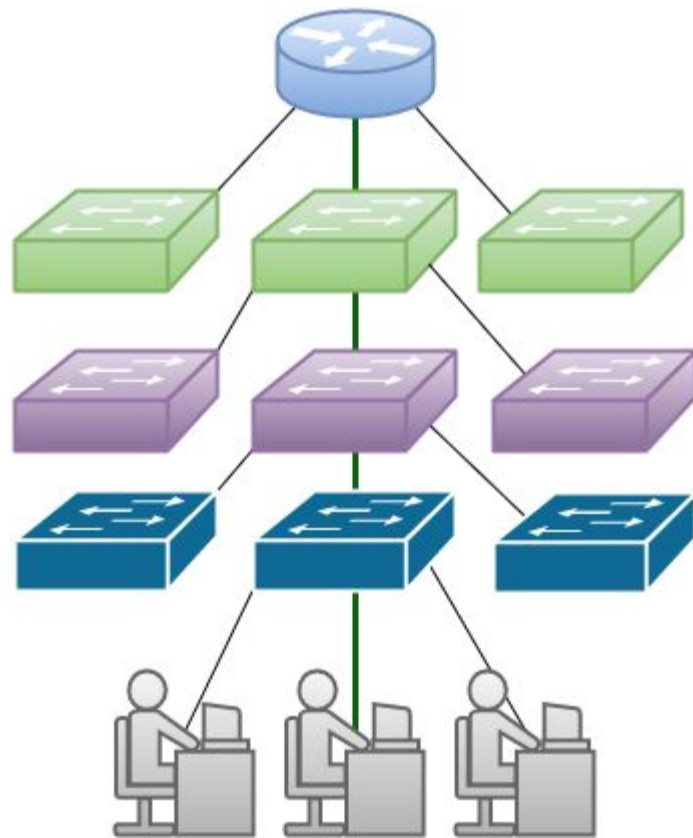
- ядро
- агрегация
- дистрибьюция
- доступ
- абоненты



L2 и L3 абоненты

В зависимости от места терминции, абоненты могут быть L2 или L3

В случае L2 схемы, абоненты терминируются в ядре, как правило, на BRAS (BNG)

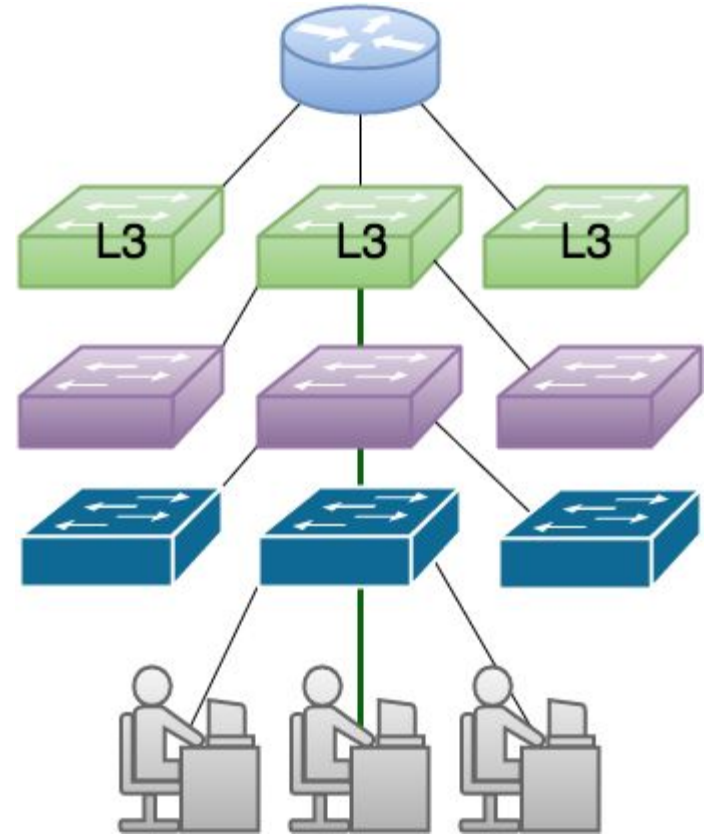


L2 и L3 абоненты

В зависимости от места терминции, абоненты могут быть L2 или L3

В случае L3 схемы, абоненты терминируются на уровне агрегации

Как правило, установлен L3 коммутатор или же отдельный маршрутизатор

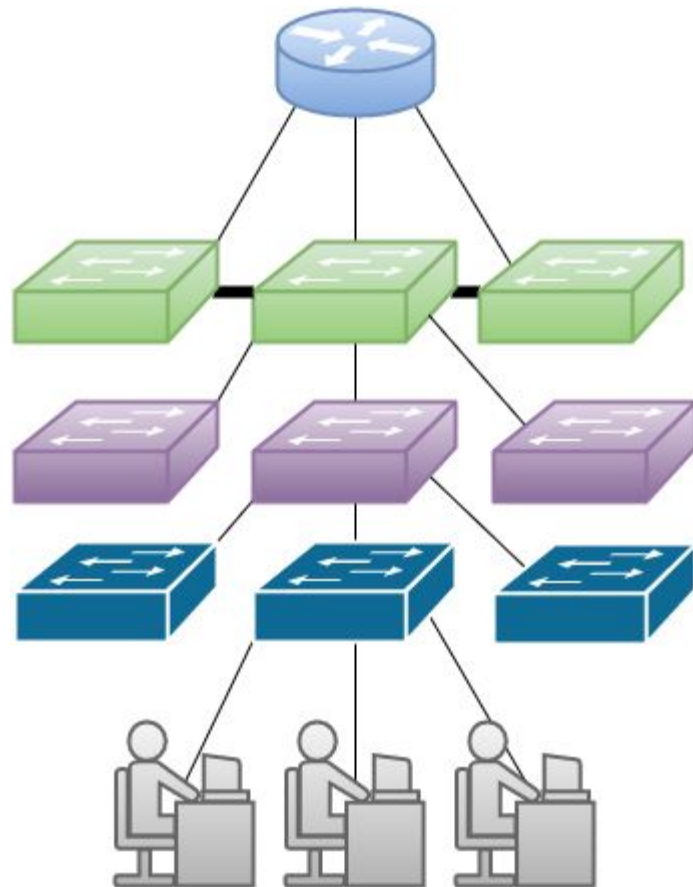


Схемы подключения

Для чего это все?

Для резервирования на уровне агрегации

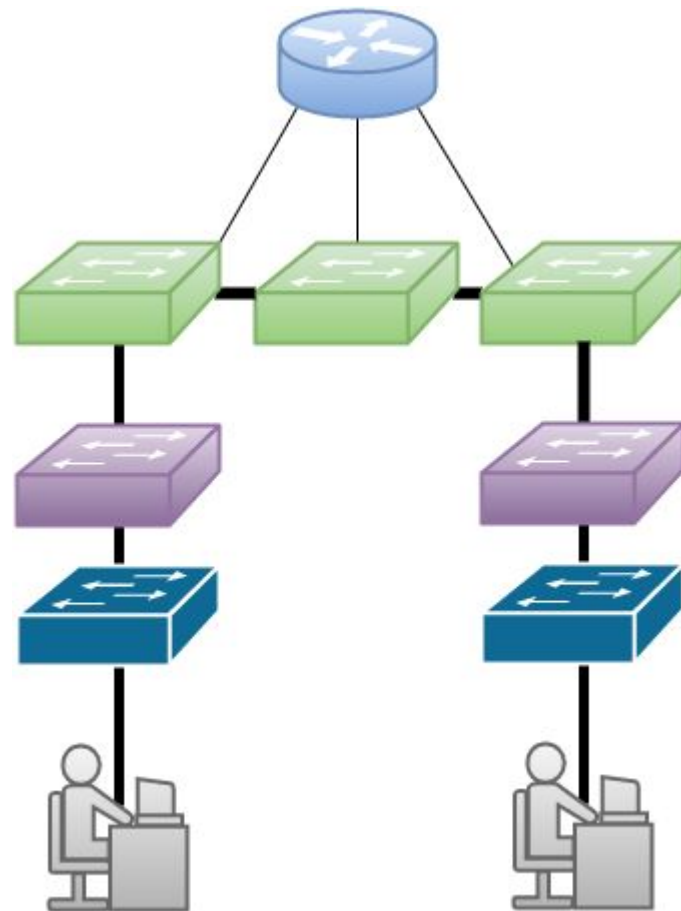
- ядро, как правило, далеко от агрегации
- потеря одного узла агрегации влечет за собой проблемы у тысяч абонентов



L2 как сервис

Предоставление L2-услуг для абонентов
(L2VPN, VPLS, L3VPN...)

- VLAN
- MPLS



L2 как сервис

Предоставление L2-услуг для абонентов
(L2VPN, VPLS, L3VPN...)

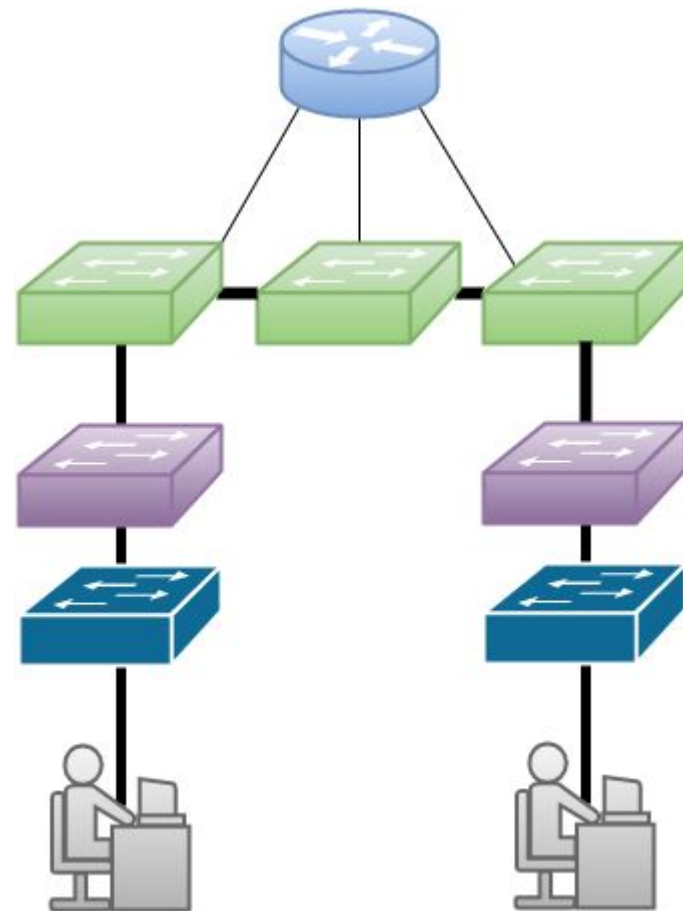
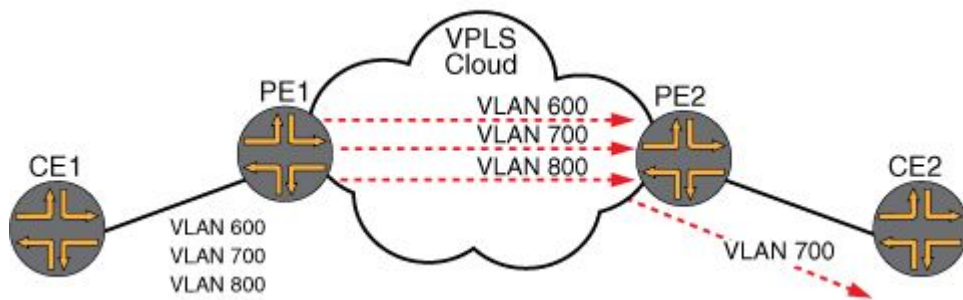


Схема узла агрегации

- аплинки, как правило, 10GE или N*10GE
- линк на дистрибьюцию 1/10GE
- MPLS PE или L3 терминация

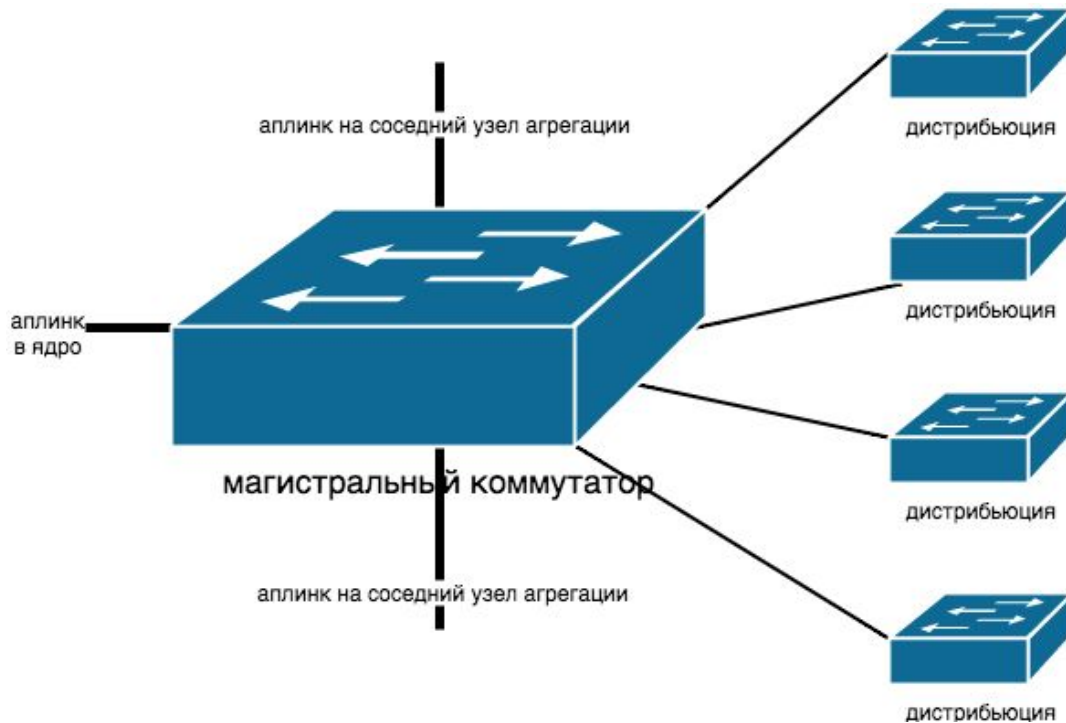


Схема узла агрегации

- дистрибьюция используется для увеличения количества портов (2-3 ВОК 24-48)
- выход из строя магистрального аплинка не ведет к простою
- выход из строя линка на дистрибьюцию = простой
- отсюда вывод - дублируем линки на дистрибьюцию

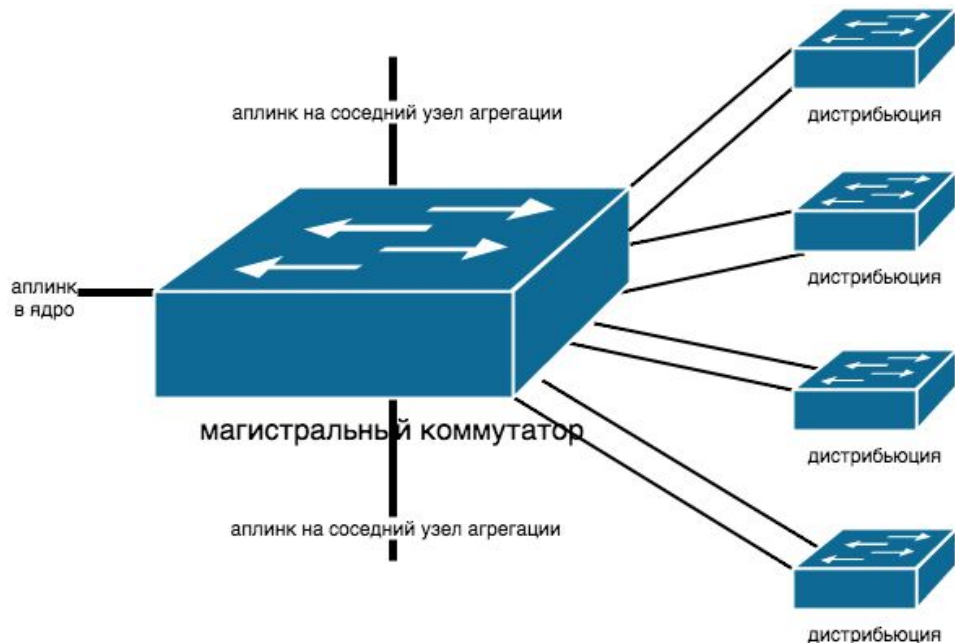


Схема узла агрегации

- увеличение количества трафика ведет к появлению агрегированных линков в ядро и на соседние узлы агрегации
- с ростом количества трафика возрастает требование по увеличению коэффициента доступности

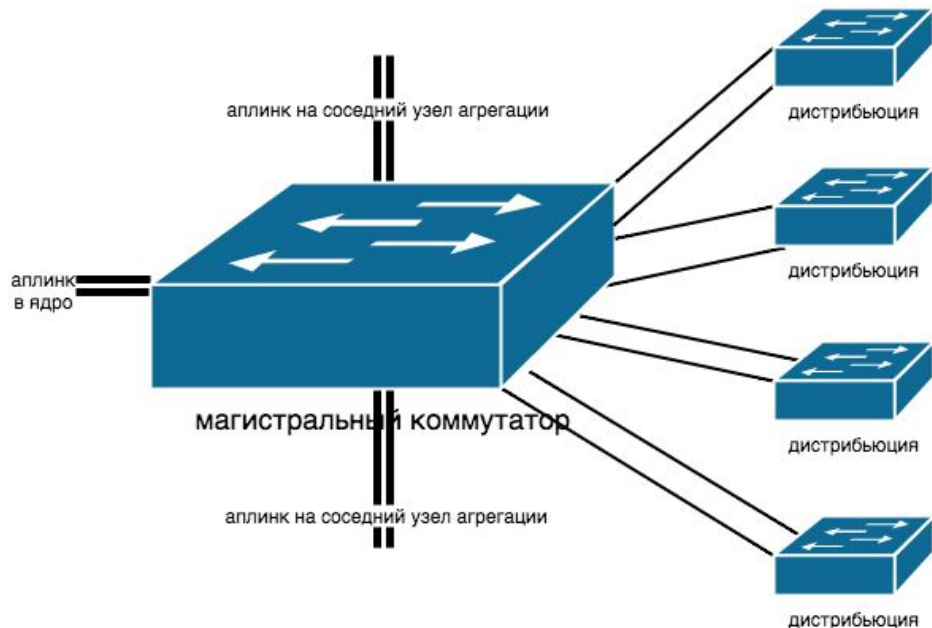


Схема узла агрегации

- увеличение количества трафика ведет к появлению агрегированных линков в ядро и на соседние узлы агрегации
- с ростом количества трафика возрастает требование по увеличению коэффициента доступности
- **SLA 99.9% - 8 часов в год**

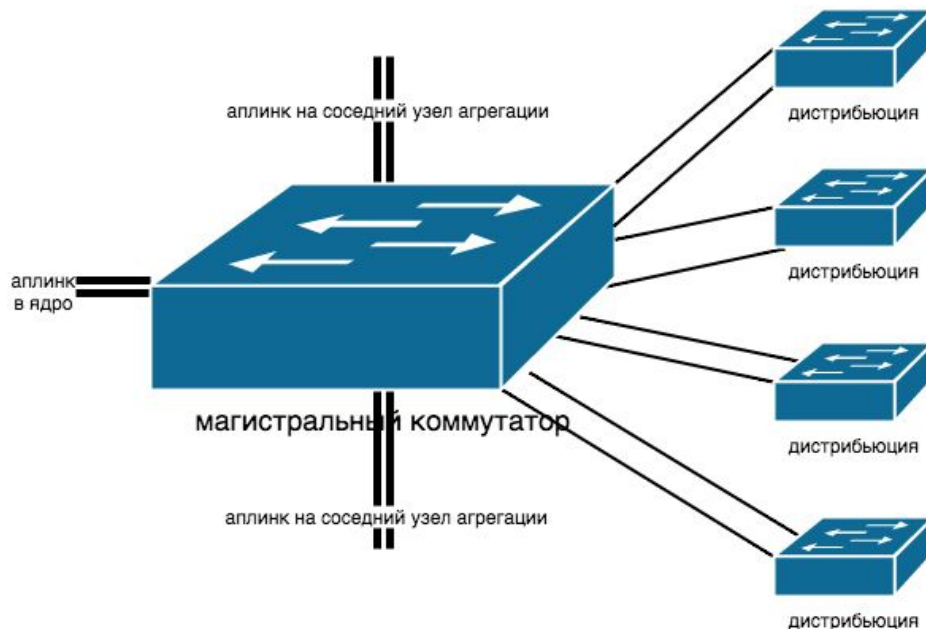


Схема узла агрегации

- отсюда выход - а давайте дублировать агрегацию

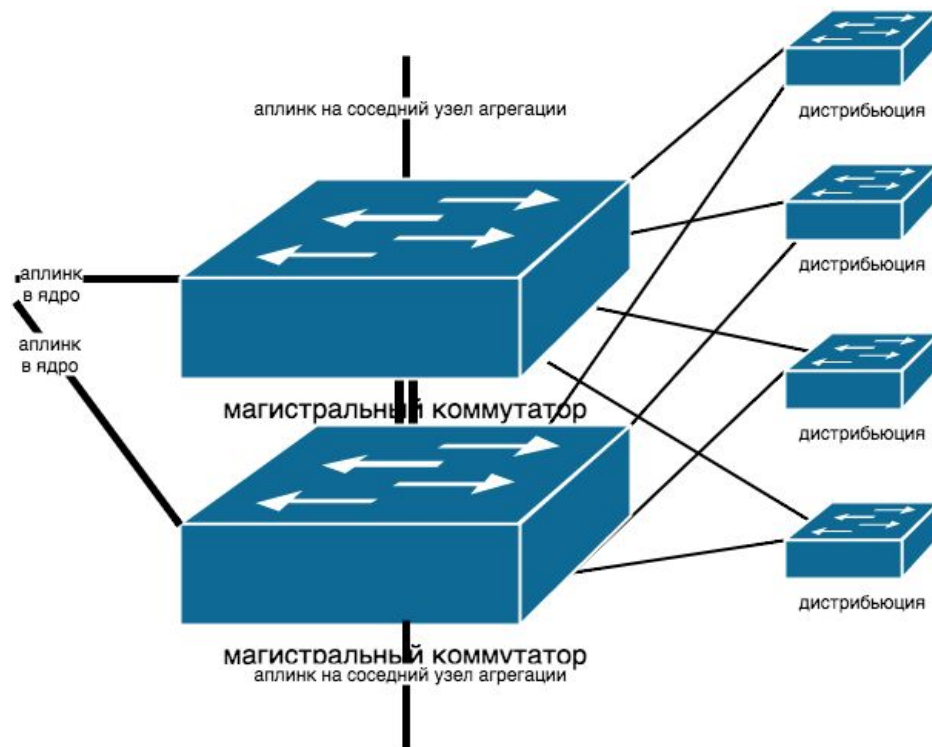


Схема узла агрегации

- независимые коммутаторы
- стек/фабрика

Подключение уровней ниже:

- MC-LAG
- STP / ERPS etc

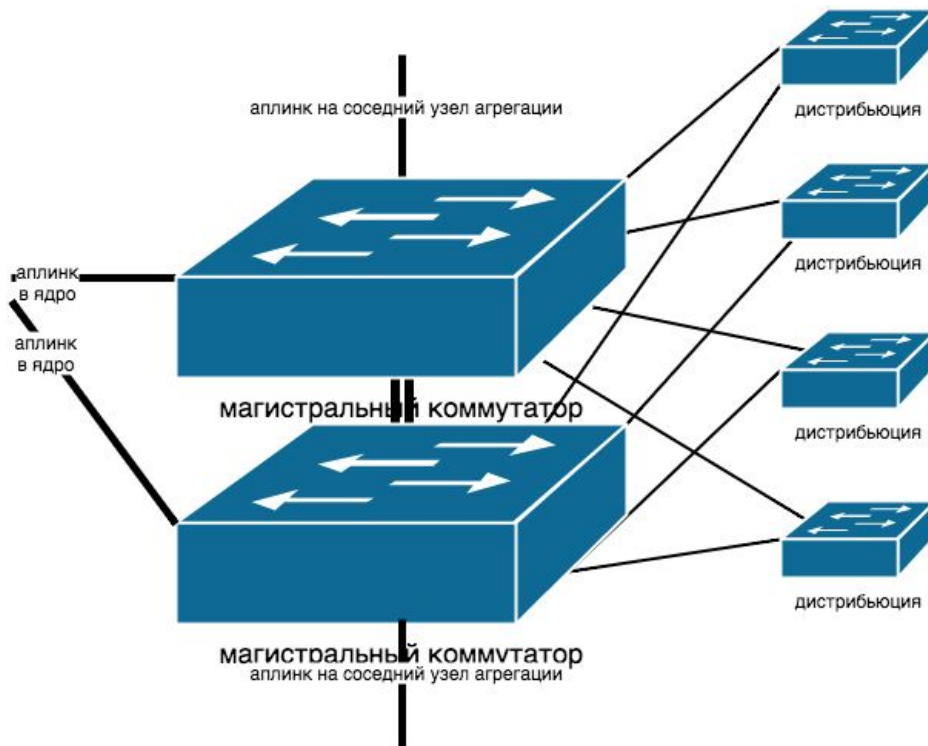


Схема узла агрегации

Агрегация

- DLink DGS-3620
- Huawei 5300
- Cisco ME3600
- Extreme x460

Дистрибьюция

- DLink DGS-3120 / 3420
- SNR
- EdgeCore

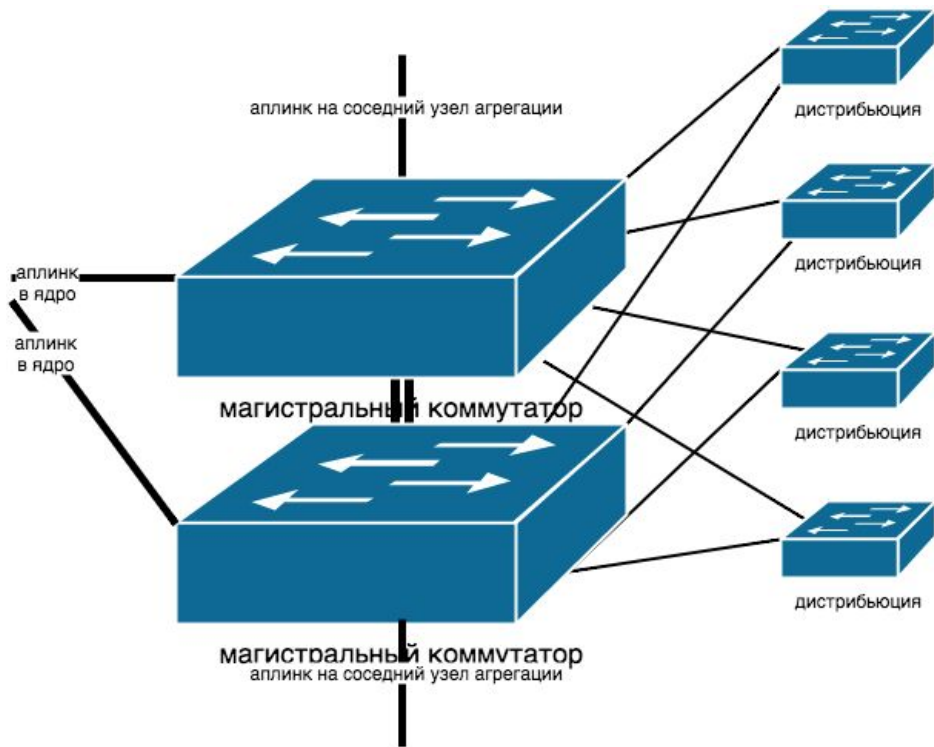


Схема узла агрегации

независимые коммутаторы

- L2 резервирование на уровень дистрибуции
- VRRP при L3 терминации
- VPLS multihoming
- L3VPN + VRRP + ...
- Невозможность работы независимых коммутаторов как одной сущности

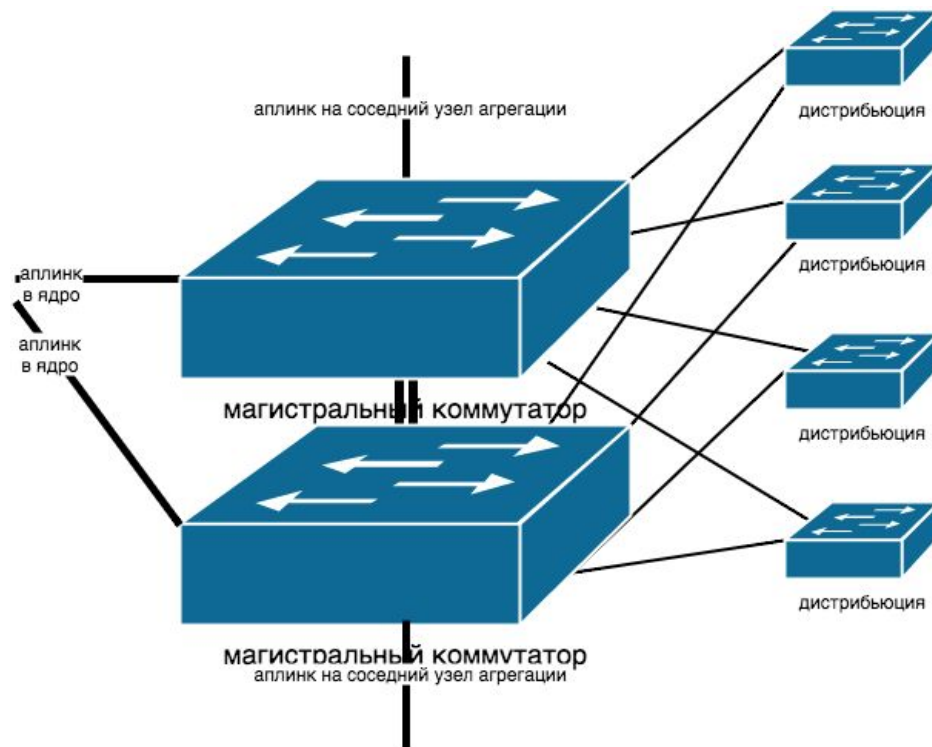


Схема узла агрегации

независимые коммутаторы

- L2 резервирование на уровень дистрибуции
- VRRP при L3 терминации
- VPLS multihoming
- L3VPN + VRRP + ...
- Невозможность работы независимых коммутаторов как одной сущности
- **усложнение обслуживания**

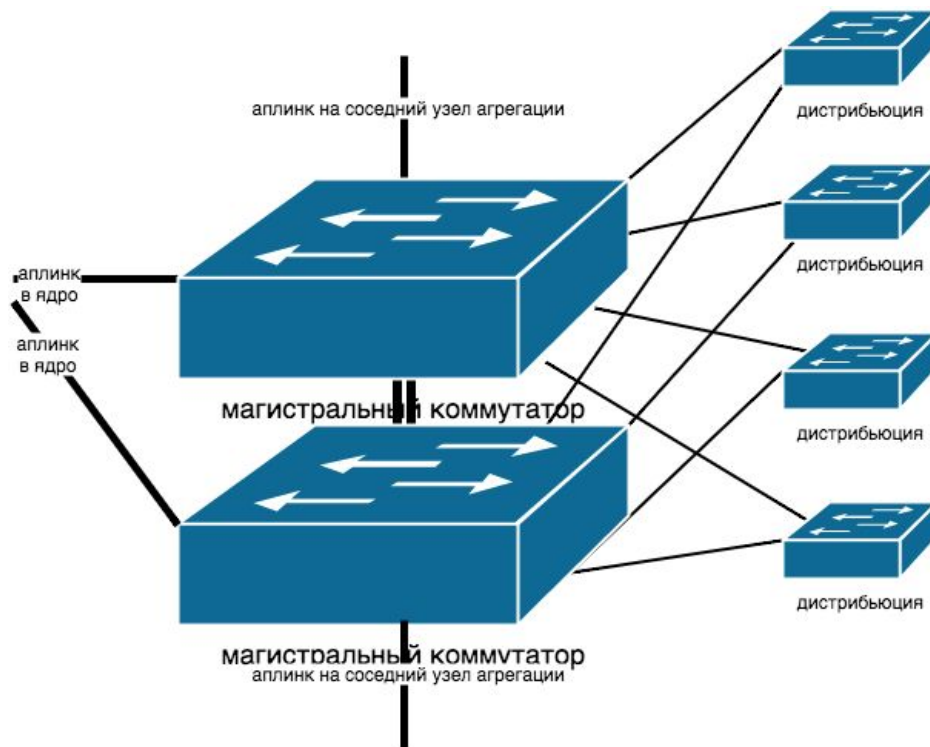


Схема узла агрегации

стек/фабрика

- это логически 1 коммутатор
- нет вопросов с MPLS PE, L2VPN, L3, L3VPN, не нужен VRRP
- **есть шанс потери всего стека**

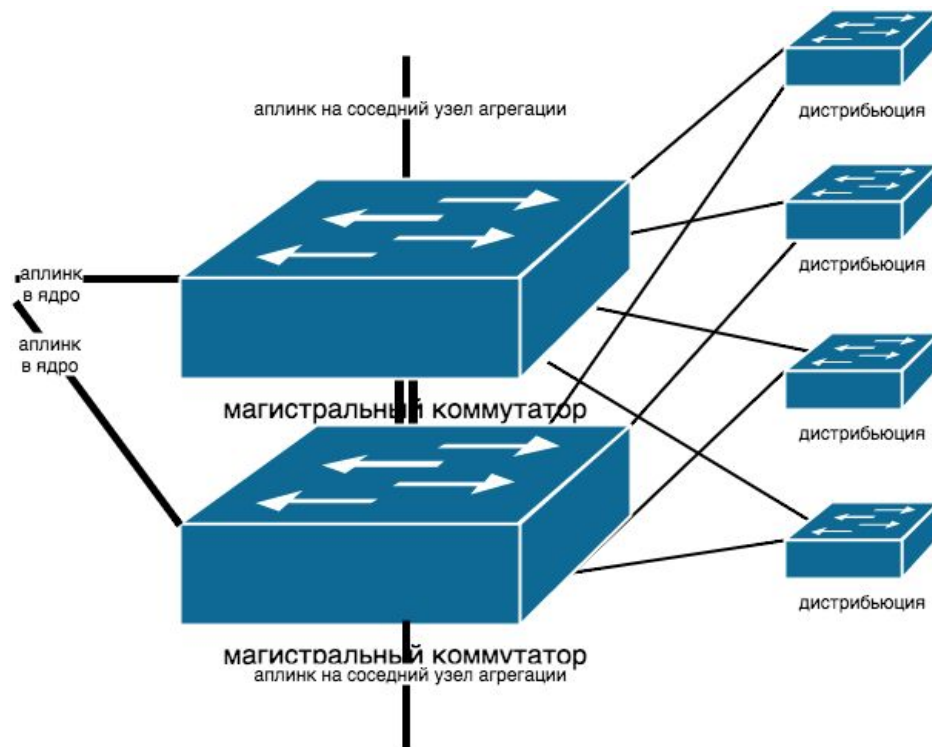
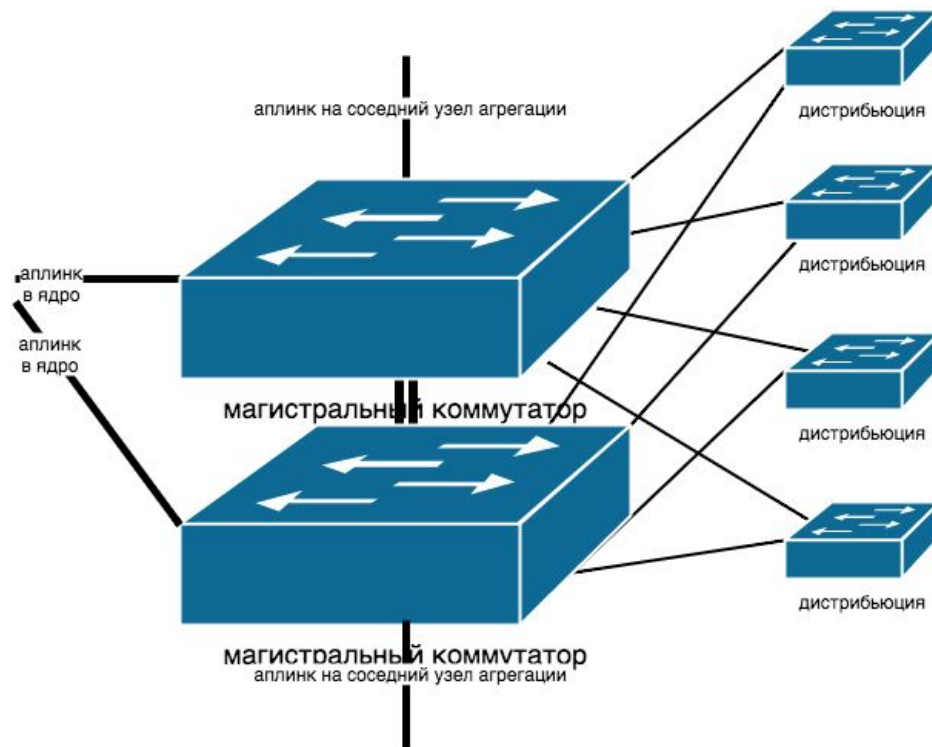


Схема узла агрегации

стек/фабрика

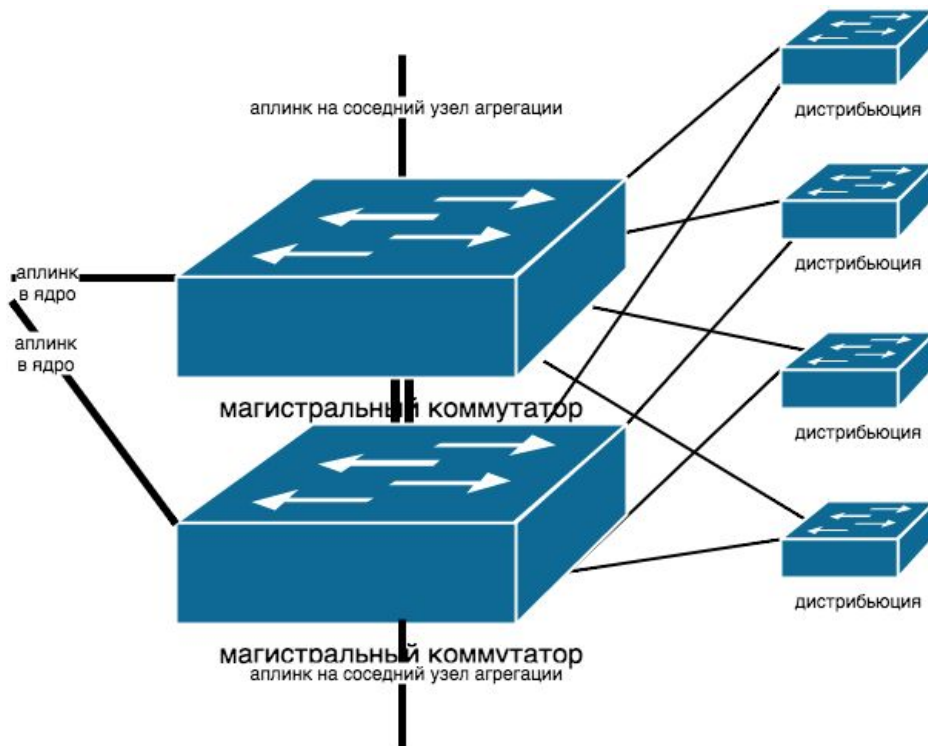
- это логически 1 коммутатор
- нет вопросов с MPLS PE, L2VPN, L3, L3VPN, не нужен VRRP
- **есть шанс потери всего стека**
- **упрощение обслуживания**



Независимые коммутаторы

MC-LAG

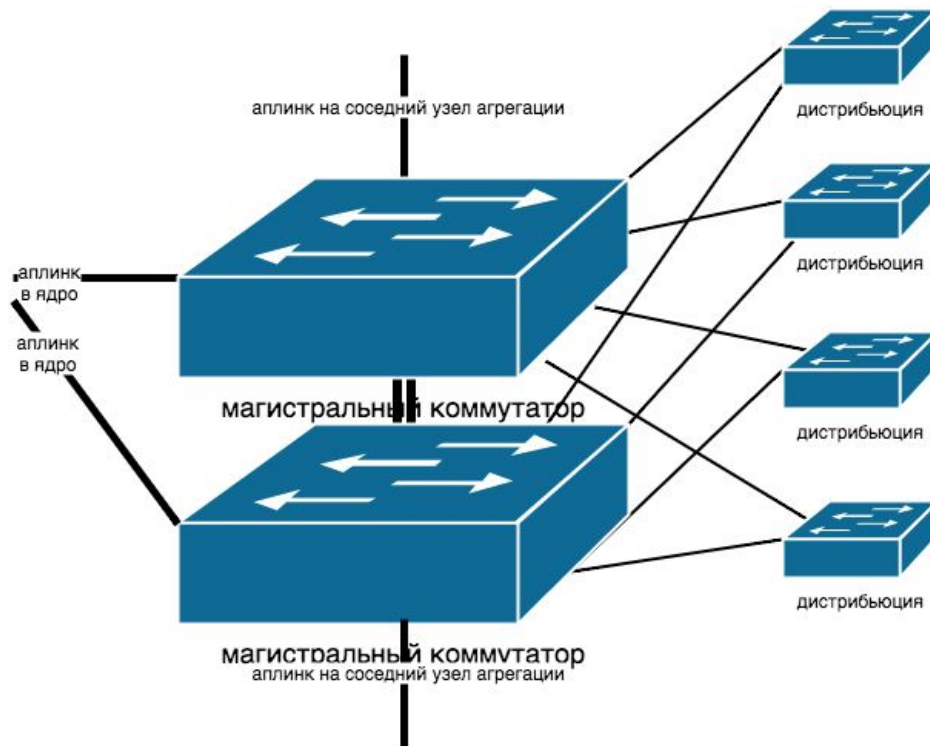
- все недостатки независимых коммутаторов
- есть шанс потери двух RE, если железо “сходит с ума”
- у некоторых вендоров есть требование одинаковых версий ПО на соседях



Независимые коммутаторы

VPLS multihoming

- все недостатки независимых коммутаторов
- Немногие pizza-box умеют VPLS multihoming
Juniper MX80+
Cisco ASR900+
- **L3VPN + dynamic routing = СЛОЖНО**



Требования к узлу агрегации

- непрерывность работы в нормальных условиях
- плановые работы не должны вызывать простой
- минимизация последствий аварии

Требования к узлу агрегации

- непрерывность работы в нормальных условиях
 - качественный софт
 - работа заявленных фич
 - интероперабельность
- плановые работы не должны вызывать простой
- минимизация последствий аварии

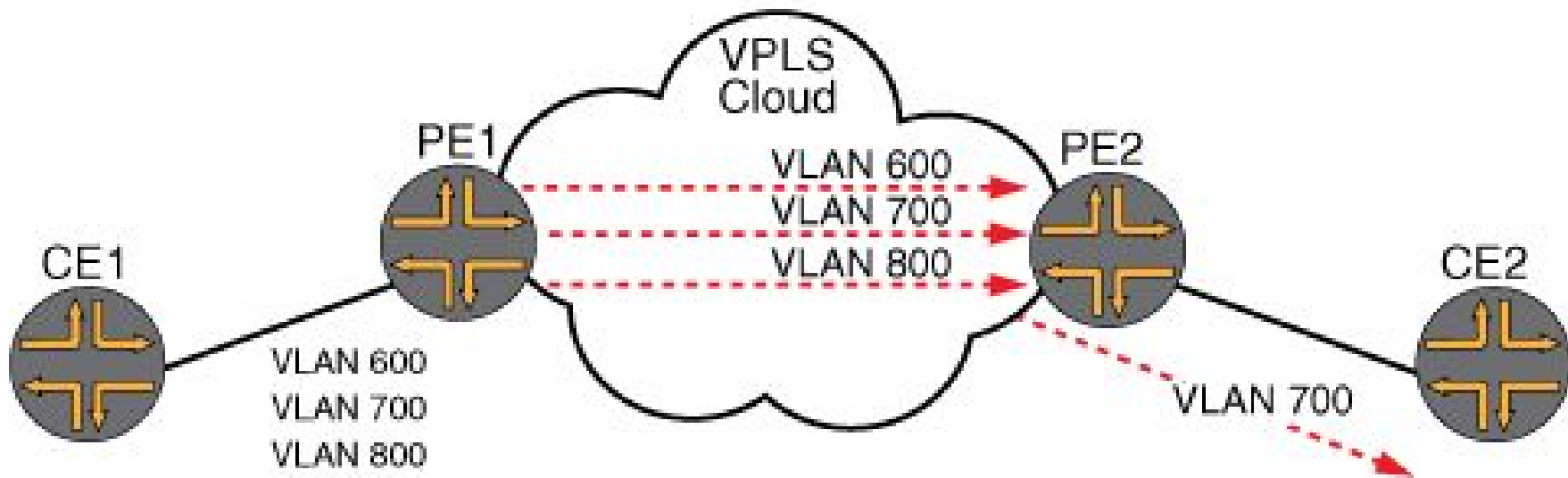
Требования к узлу агрегации

- непрерывность работы в нормальных условиях
- плановые работы не должны вызывать простой
 - обновление ПО по очереди, без простоя стека целиком
 - NSSU / ISSU
 - резервирование всех линков
 - бесшовное переключение master/backup
- минимизация последствий аварии

Требования к узлу агрегации

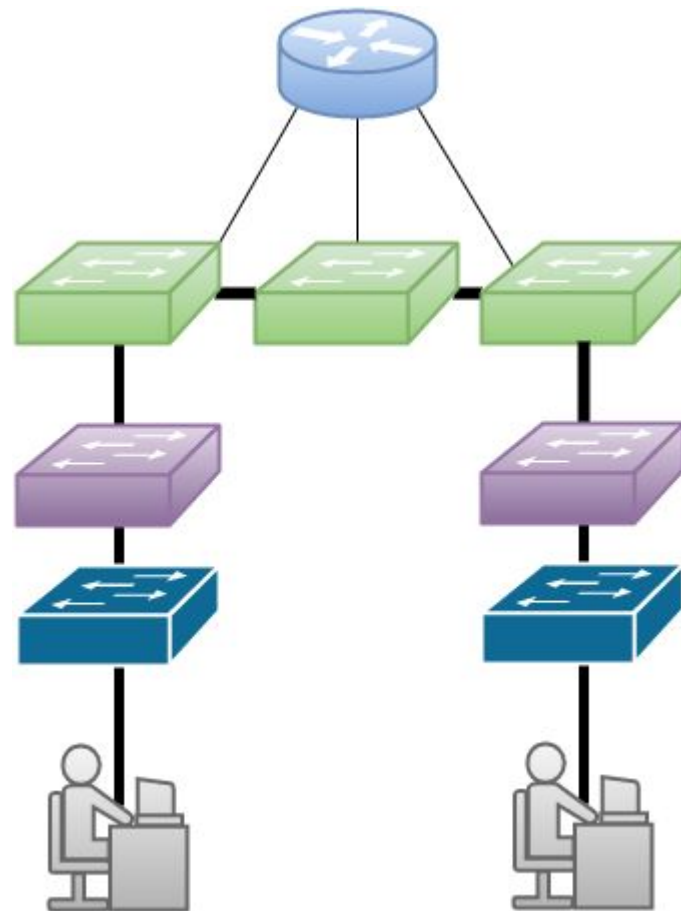
- непрерывность работы в нормальных условиях
- плановые работы не должны вызывать простой
- минимизация последствий аварии
 - non-stop bridging
 - non-stop routing
 - бесшовное переключение master/backup
 - бесшовное добавление и удаление членов стека

VXLAN - предпосылки



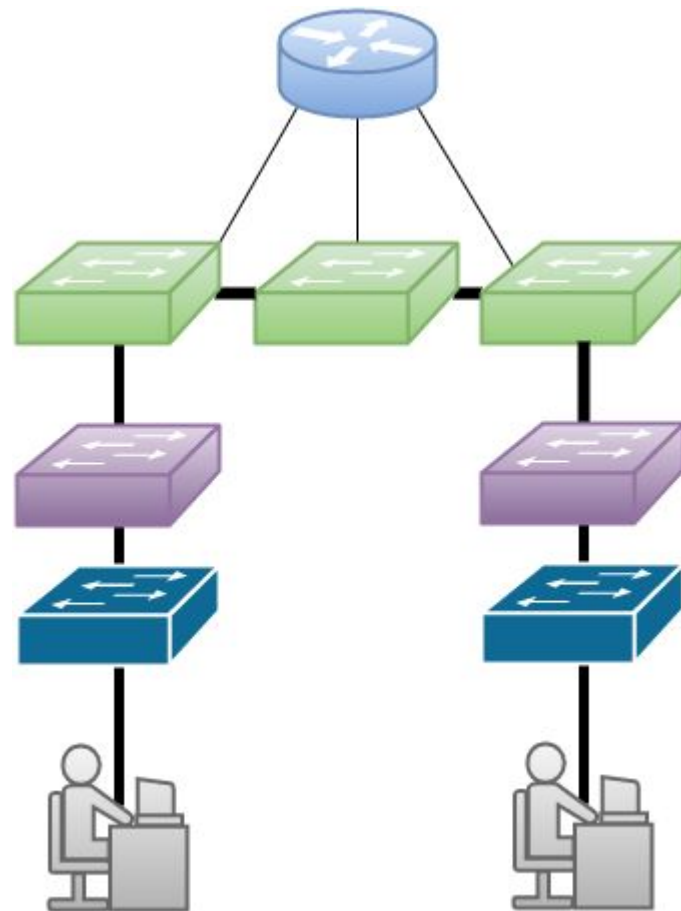
VXLAN - предпосылки

- L3-резервирование проще (у STP/ERPS/etc есть куча ограничений)
- VXLAN, в отличие от MPLS, требует на промежуточных точках только увеличенного L3MTU (не нужен label switching)
- MPLS есть на малом количестве pizza-box (Extreme x460 и x670, Brocade CER, Cisco ME3600/3800, Huawei S5300)



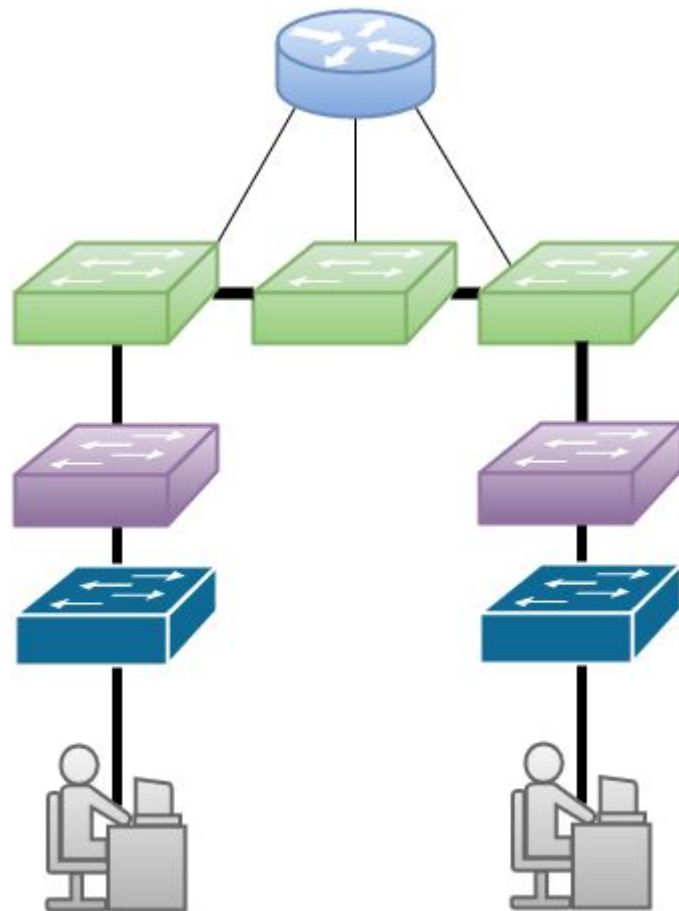
VXLAN - предпосылки

- VXLAN есть на современных чипсетах (Trident II, Intel Fulcrum, ...) и на многих коммутаторах (Extreme, Brocade VDX, Cisco Nexus 5k, Mellanox, Arista, Juniper)
- VXLAN уже есть на OpenVSwitch
- VXLAN есть на BareMetal



VXLAN - предпосылки

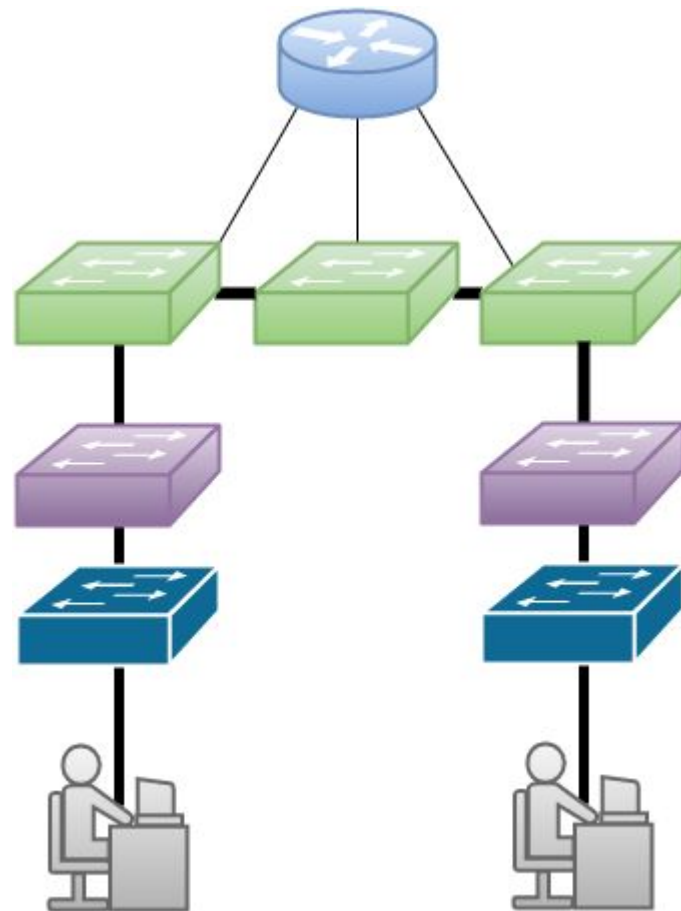
- замена оборудования нужна только в тех местах, где планируется использование VXLAN
- VXLAN не потребляет ресурсы транзитных коммутаторов (метки, LSP)
- Использование существующей схемы резервирования



VXLAN - модели коммутаторов

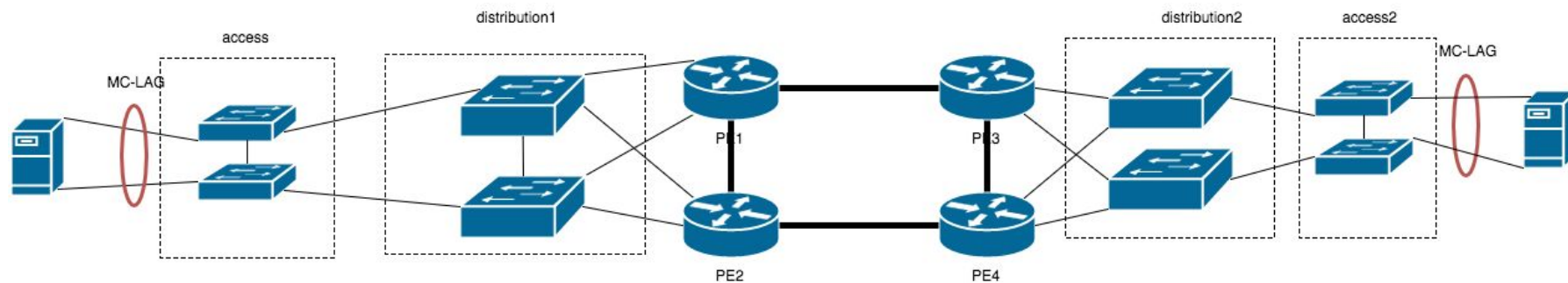
- Juniper QFX5100, QFX10002
- Brocade VDX 6740
- Cisco Nexus 5K
- Extreme x460, x670
- Huawei CE6850
- ALU OS6900

Большинство коммутаторов 48xSFP+ / 6xQFSP



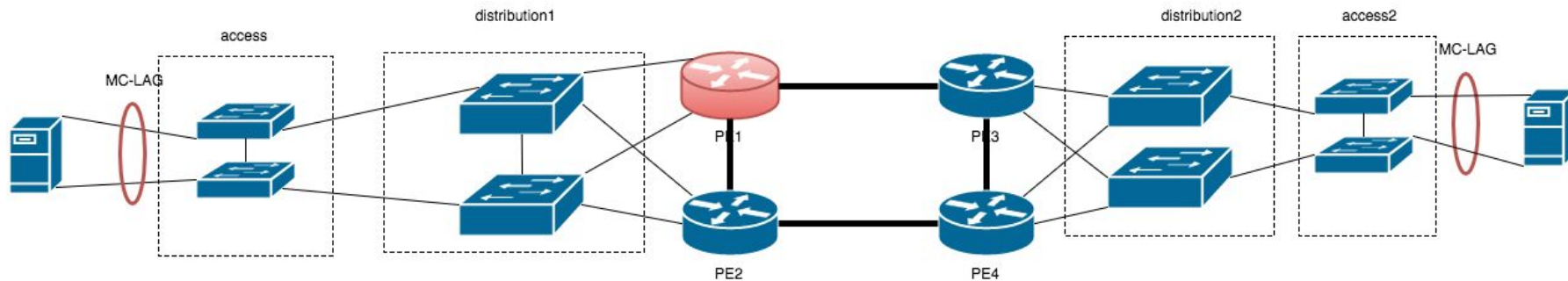
Тестирование оборудования

Начальная конфигурация схемы



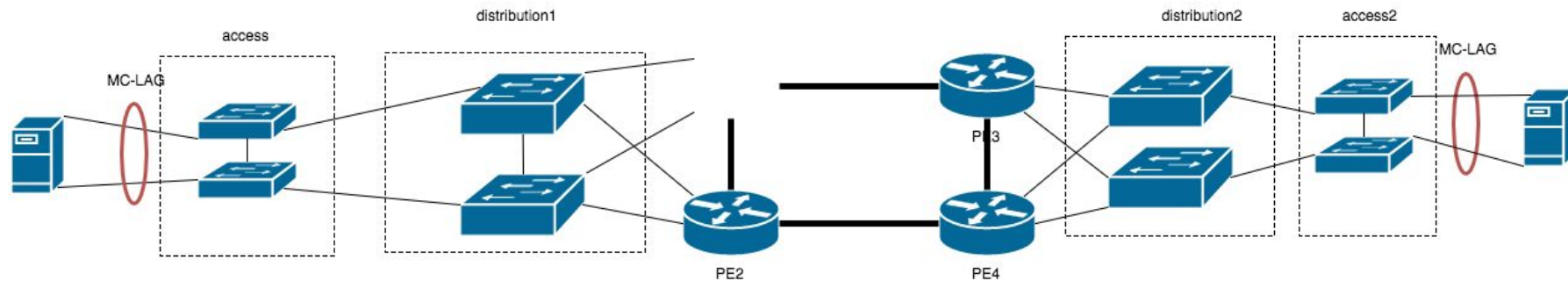
Тестирование оборудования

Падения и перезагрузки коммутаторов



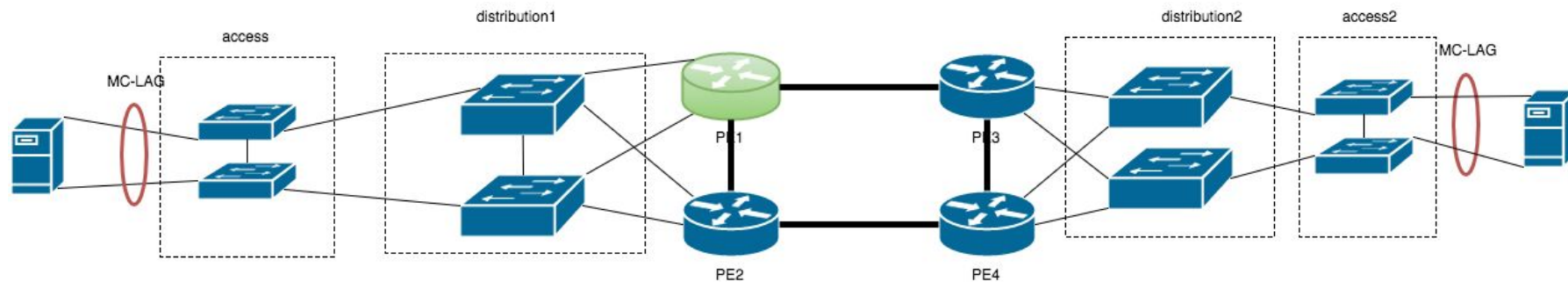
Тестирование оборудования

Удаление коммутаторов для обслуживания



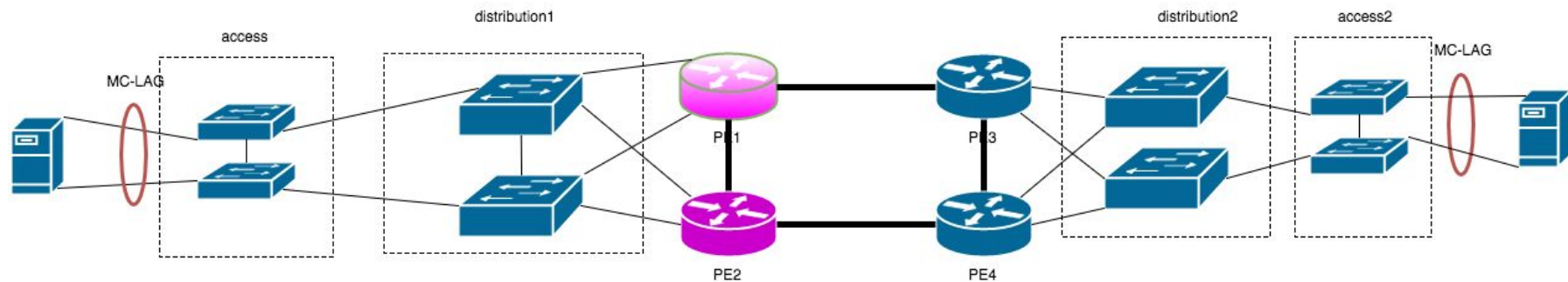
Тестирование оборудования

Замена коммутатора



Тестирование оборудования

Обновление ПО



Тестирование оборудования

	начальная конфигурация	перезагрузка/ поломка коммутатора	удаление коммутатора	добавление коммутатора	обновление ПО
Juniper QFX	версия ПО	OK	OK	OK	стек по очереди
Cisco Nexus 5k	---	---	---	---	---
Brocade VDX	OK	OK	OK	OK	ISSU
Extreme x670	перезагрузка при стекировании	OK	OK	OK	весь стек сразу
Huawei CE6850	OK	OK	OK	OK	ISSU, перерыв 20 сек
ALU OS6900	OK	OK	OK	OK	ISSU, перерыв 5 сек

Выводы

- Стоимость одного коммутатора ~ \$30-40K GPL
- Агрегация “на чердаке” превращается в дистрибьюцию
- Новые требования к помещениям под размещение агрегации

Выводы

- Стоимость одного коммутатора ~ \$30-40K GPL
- Агрегация “на чердаке” превращается в дистрибьюцию
- Новые требования к помещениям под размещение агрегации
- MPLS остался на больших маршрутизаторах
- VXLAN как наследник L3 резервирования
- Решения для дата-центров интересны и для SP

Вопросы и обсуждение