

EVPN для провайдеров

Кирилл Малеванов
Selectel



Классический дата-центр –

Facebook, Amazon, Yandex, ВКонтакте

- Все серверы принадлежат одному (внутреннему) клиенту
- Доля сетевых аномалий мала
- Каждый сервер имеет предсказуемые IP-адреса, профиль трафика, политику доступа
- Выход из строя одного сервера, как правило, не приводит к падению всего сервиса

Дата-центр Selectel

- много разных услуг, от legacy до облачных

- Каждый сервер (физический или виртуальный)

 - может принадлежать отдельному клиенту

- Любое количество сетевых аномалий – как в ШПД

- Каждый сервер может произвольно менять IP-адреса,

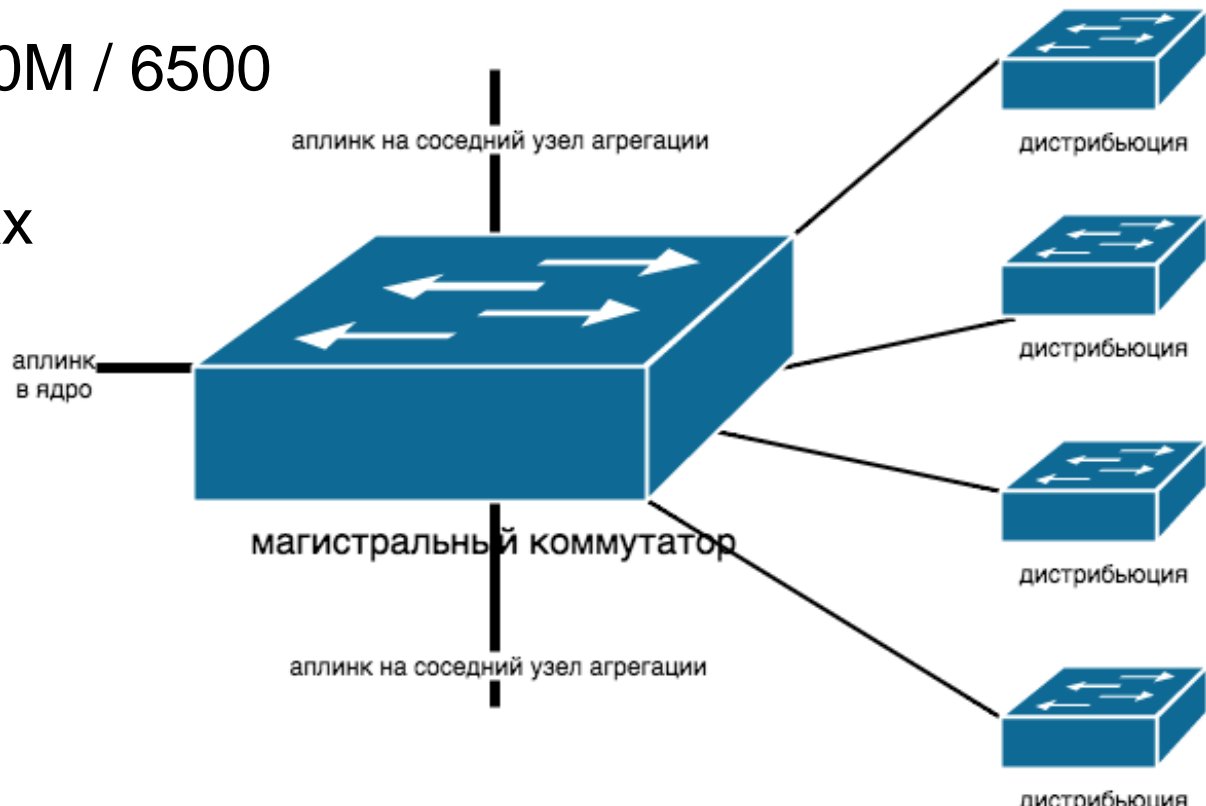
 - профиль трафика,

- Вероятность DDoS выше на порядки

- Выход из строя одного сервера у клиента зачастую критичен

Магистральный коммутатор

- Extreme x460
- Cisco ME3600 / 4900M / 6500
- D-Link DGS-36xx
- Huawei S53xx, S58xx
- Juniper EX4xxx
- ?

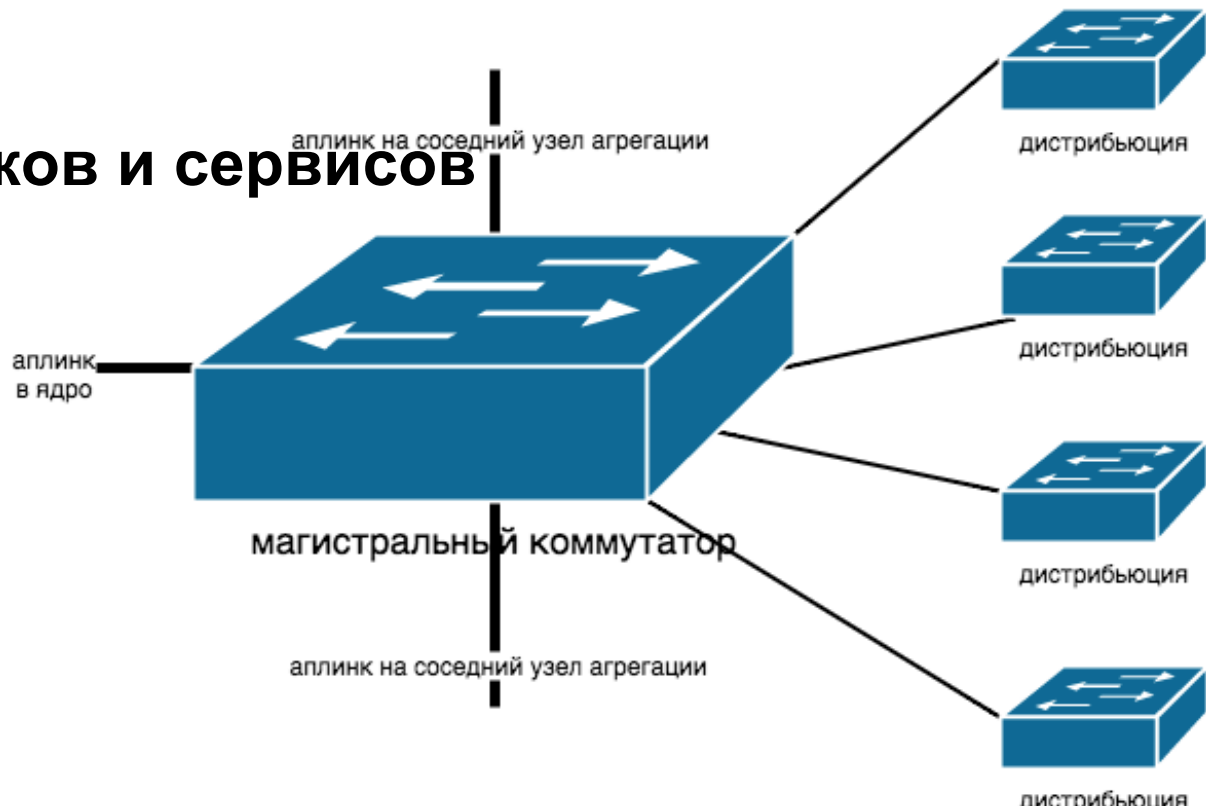


Магистральный коммутатор

- какой?

Резервирование линков и сервисов

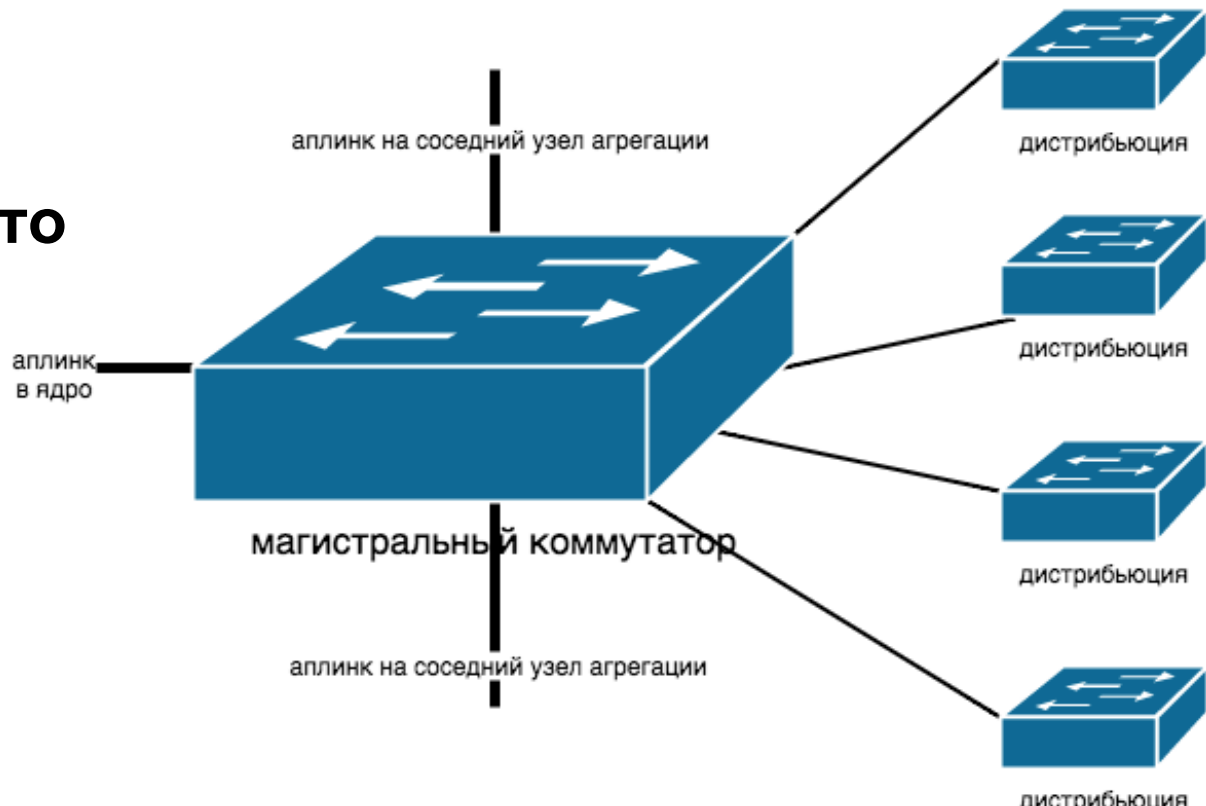
- как?



MPLS

- Juniper MX
- Cisco ASR900

Практика показала, что
MPLS pizza box нет



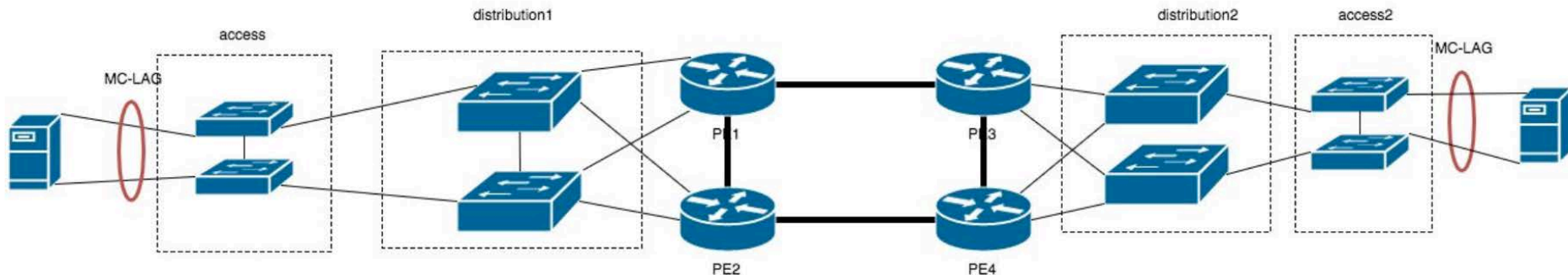
Резервирование VLAN

VPLS multihoming

- Наличие LDP/RSVP между PE-устройствами

Использование туннелей

- MPLS туннель

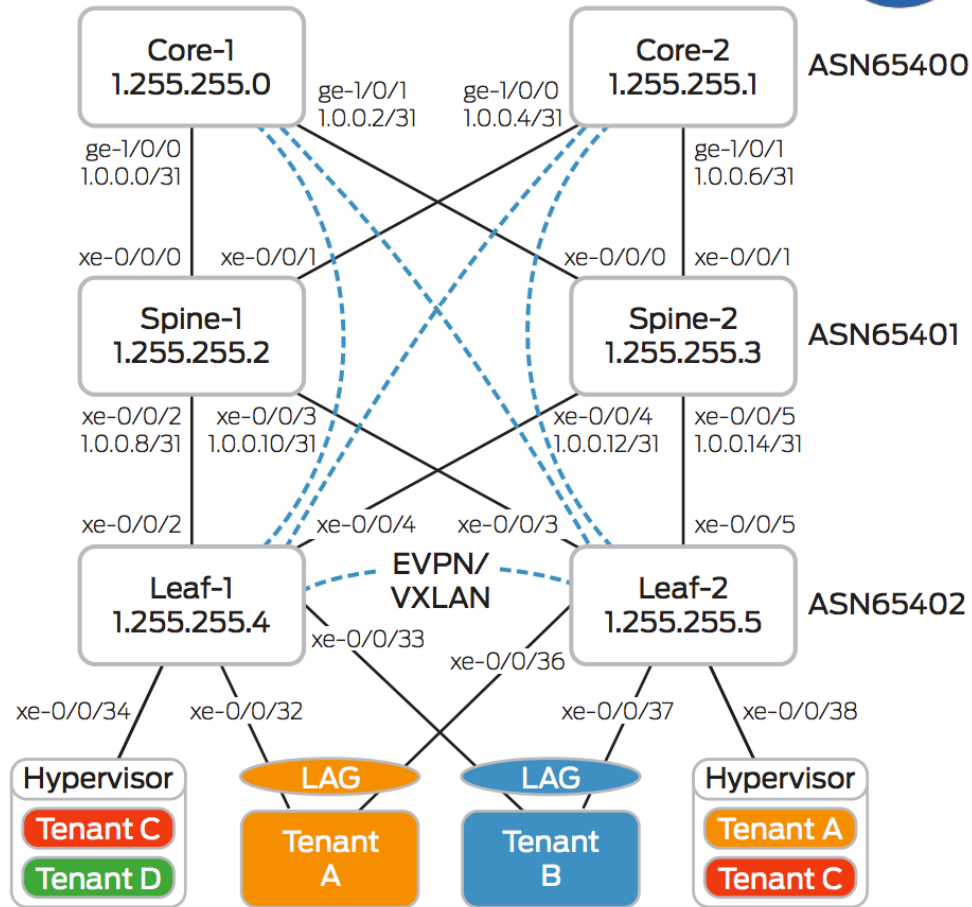


Underlay

- Пересылка пакетов «в железе»
- Пересчет контрольных сумм фреймов

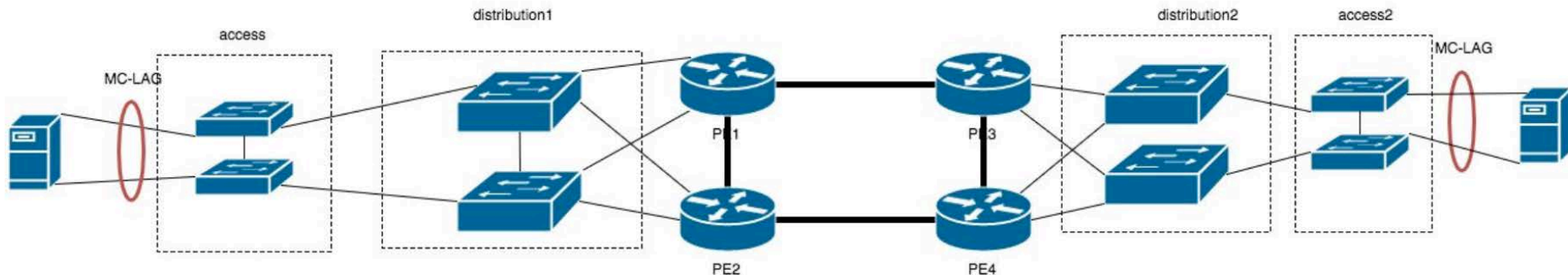
Overlay

- Сигнализация
- Резервирование при обрывах и неисправностях



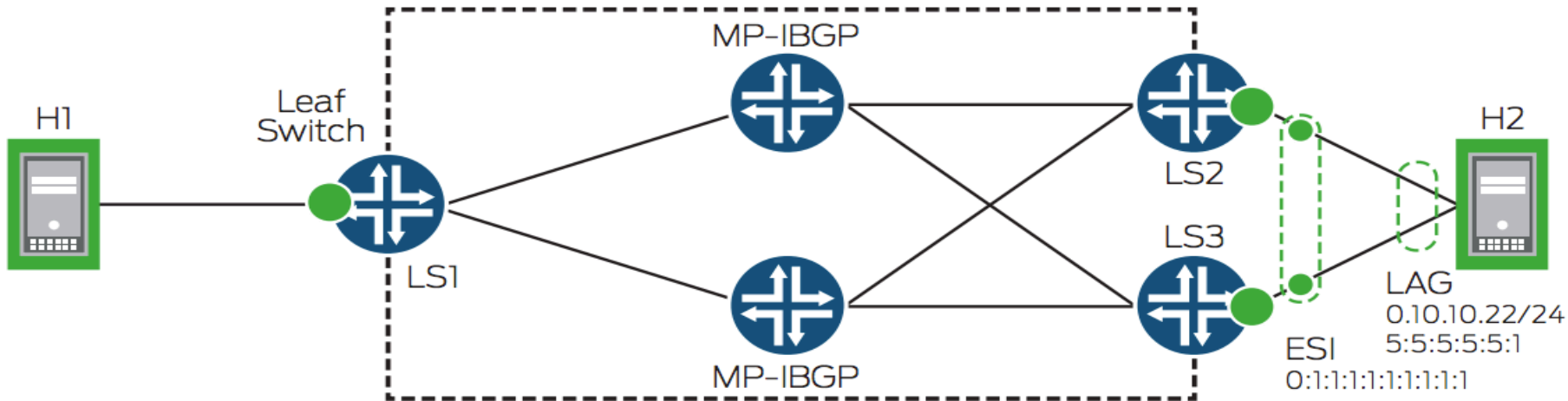
1U коммутатор умеет VXLAN + BGP

- Juniper QFX51x0 / QFX10k
- Cisco Nexus 5672, 9318
- Brocade VDX6740, SLX
- Huawei CE6800



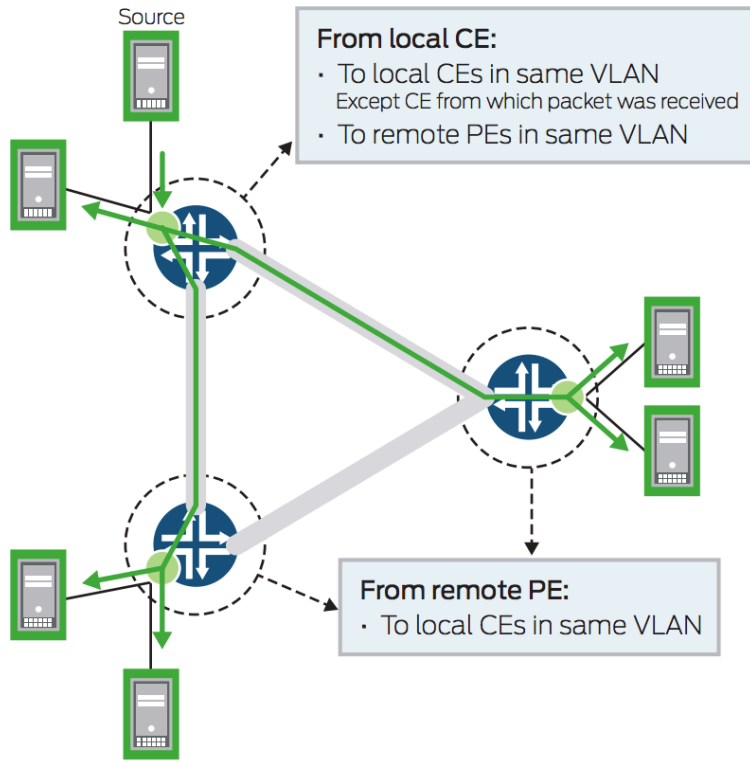
EVPN

- Между конечными точками требуется только L3 связность
- Резервирование сети между точками по общей L3 сети
- Active-standby и active-active подключения
- Штатная возможность организации LACP с независимых устройств (ESI – Ethernet Segment Identifier)

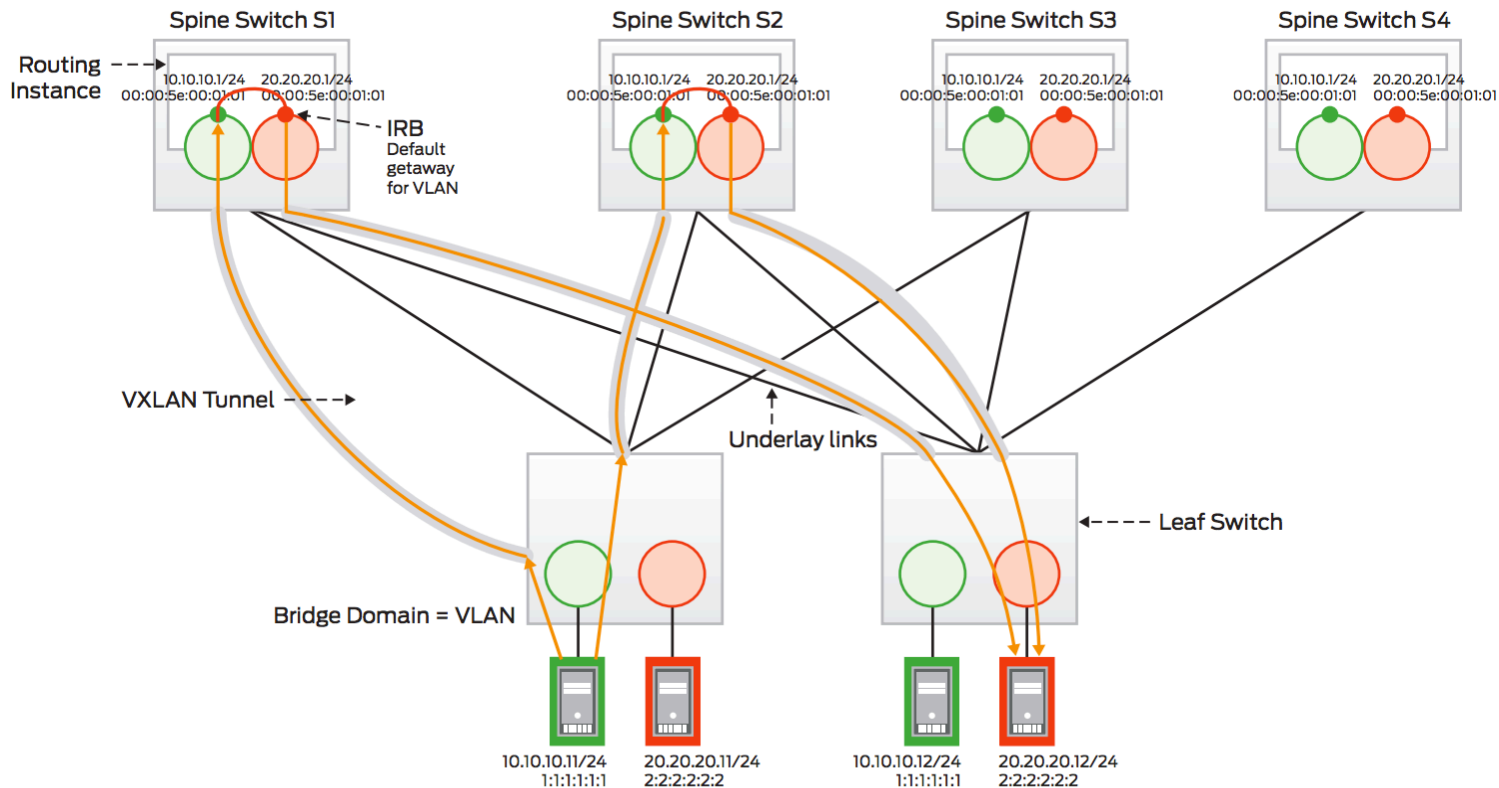


EVPN

- Чтобы трафик не дублировался, применяется правило split horizon



Layer 3 – самое интересное

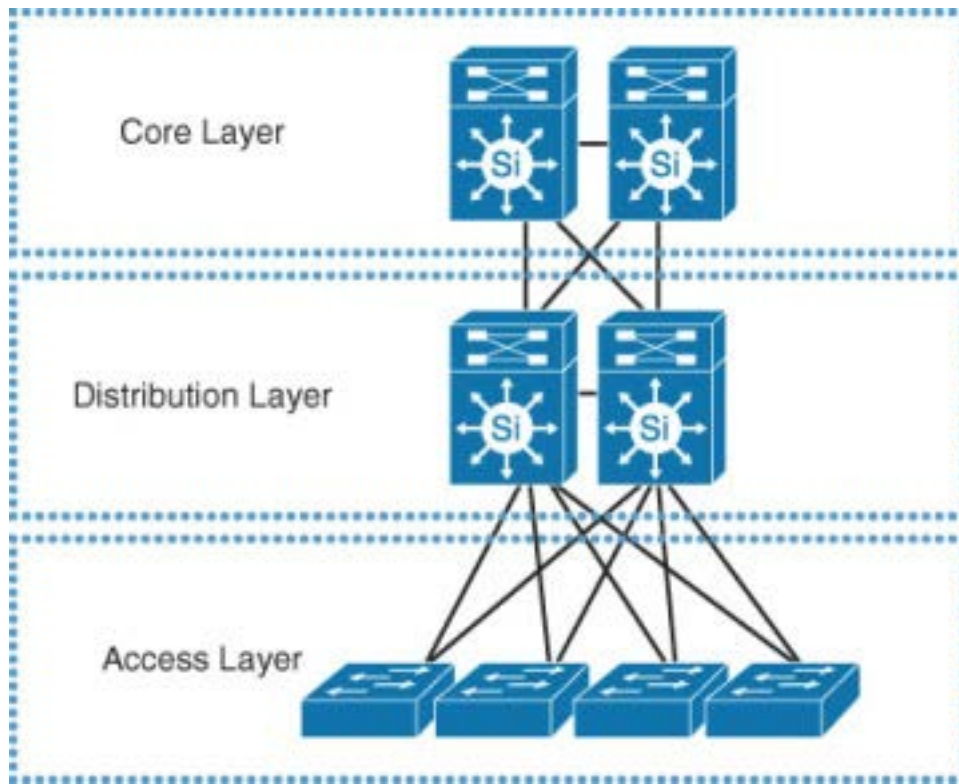


	Starta XGS				Dune , StartaDNX		
Chip Series	Trident 2	Trident 2	Trident 2+	Tomahawk	Qumran	jericho	
Chip Number	BCM56854	BCM56850	BCM56864	BCM56960	BCM88375	BCM88670	BCM88750
Switching Capacity (half Duplex)	720Gbps	1.28 Tbps	720Gbps	3.3Tbps	800Gbps	720Gbps	NO INFO
Forwarding Capacity	1 Bpps	1.44Bpps	1 Bpps	4.7Bpps	Variable, can be very large , Utilizes an external TCAM Chip such as BCM52311		
IPv4 Table	64k	64k	64k	72k			
IPv6 Table	20k	20k	20k	36k			
Buffer	12Mb	12Mb	12Mb	24Mb	Variable, supports external DDR4 memory for buffer. More than 8GB		
Features	Very common chip in Leaf switches	Very common chip in Spine switches	First chip to support VXLAN Routing	100G Switch with VXLAN Routing	Expandable buffer and TCAM tables. Gigabytes of Buffer and Internet BGP fit tables.		
Ports	48 x 10G 6 x 40G	32 x 40G	48 x 10G 6 x QSFP	32 x 100G	48 x 10G 4 x 100G	NO INFO	NO INFO

Поддержка VXLAN routing

- Juniper Q5 (QFX10k)
- Juniper Trio (MX)
- Huawei ?
- Brocade VDX6740/6940 (Brocade Cobra)
- Brocade SLX (Cavium)
- Cisco Nexus 9318
- Cisco ASR1K, ASR9K

Резервирование шлюза по умолчанию



Кирилл Малеванов

Технический директор

malevanov@selectel.ru

<https://selectel.ru>

