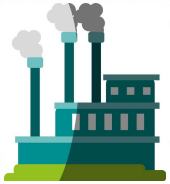
Workshop Agile Data Center

Москва, 11 апреля 2019 г.

Что Вас ждет?



И снова о фабрике… о-ох
- Откуда, зачем и как
- Как проектируем?
- Какие протоколы используем?

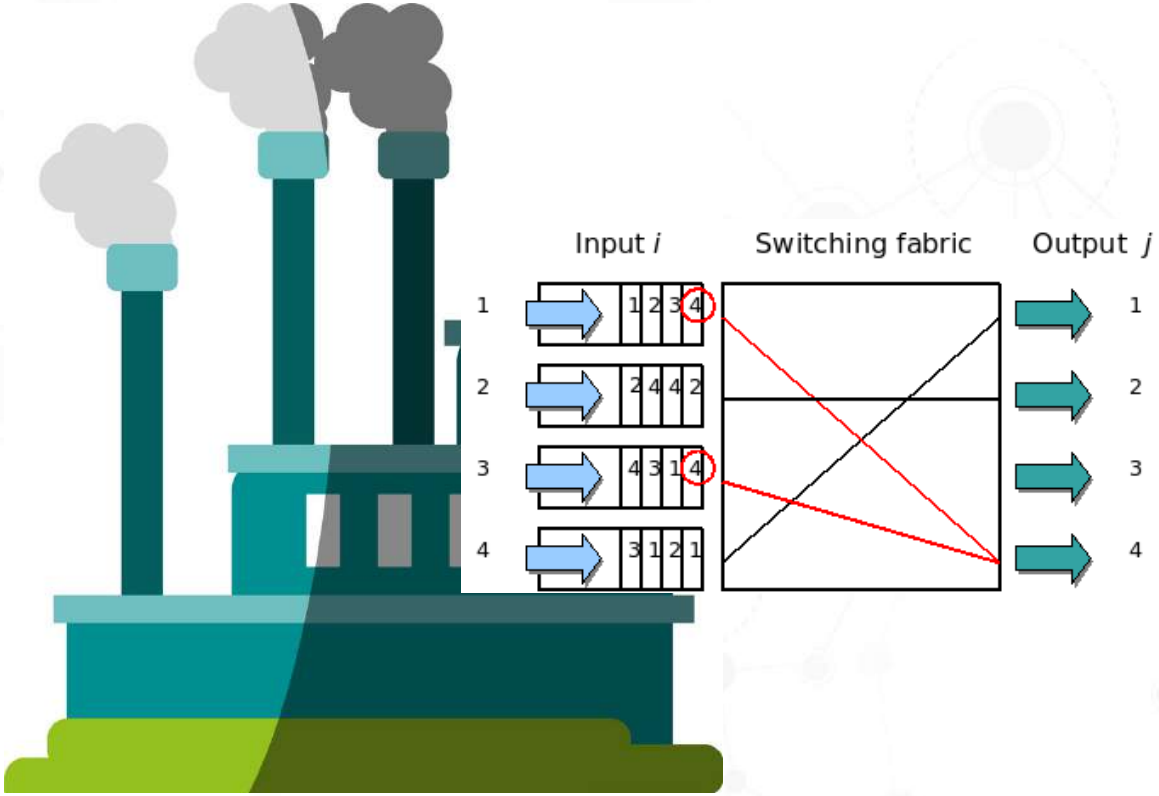
Настройка
- Пример конфигурации фабрики

Автоматизация
- Embeded fabric automation

Словарь

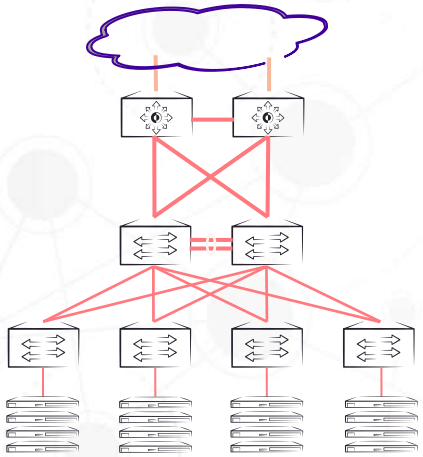
agile [ˈædʒaɪl] прил

1. проворный, поворотливый, ловкий, маневренный, верткий, юркий (nimble, maneuverable, dexterous)
2. подвижный, подвижной, гибкий (flexible, mobile)
3. быстрый, расторопный, прыткий (quick)
4. живой (alive)
5. динамичный (dynamic)
6. шустрый (nimble)
7. сообразительный (smart)


Input i
Switching fabric
Output j
1
2
3
4
1 2 3 4
2 4 4 2
4 3 1 4
3 1 2 1
1
2
3
4

Эволюция сетевых архитектур

Традиционная

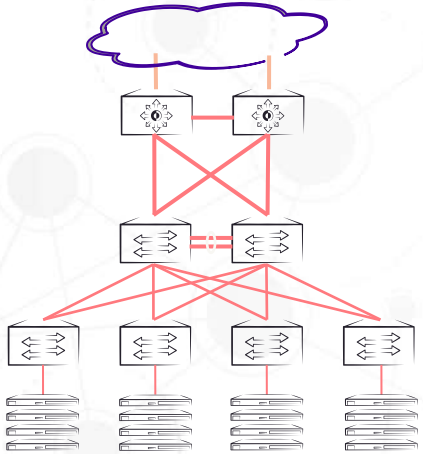


3-уровневая
Fat Tree

- Layer 2: только 1 путь
  - STP disables other paths
- 40~50% линков не используется
- MSTP повысит использование (STP per VLAN)
  - Увеличенная сложность
  - Создается множество single-path сетей
  - Ограниченная зона мобильности
- Отказ линка
  - Сходимость STP – сеть не работает
  - Broadcast шторма – сеть может не работать
- Layer 3 как альтернатива
  - Сложнее?
  - Выше стоимость?
  - Ограниченная мобильность VM (в пределах стойки)

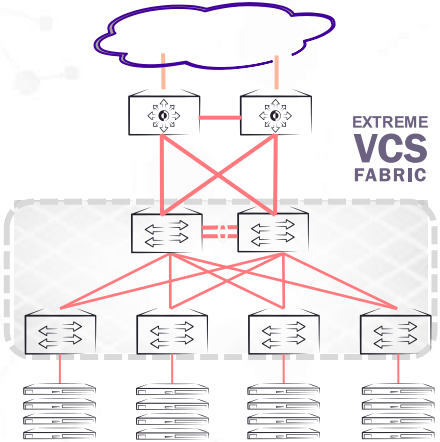
Эволюция сетевых архитектур

Традиционная



3-уровневая
Fat Tree

Влияние виртуализации



L2 Scale-out
TRILL, SPB[v,m], FP, VCS

- IETF → TRILL
- IEEE → SPB

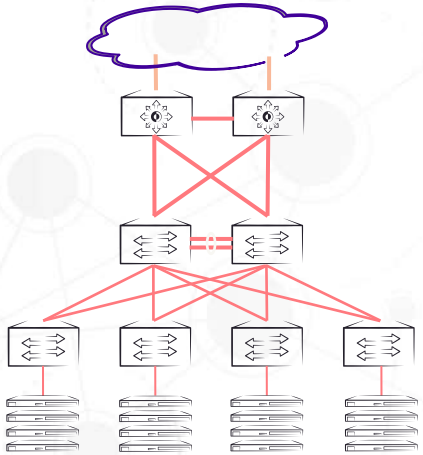
- Маршрутизация (условно)
- Поддержка IP storage (iSCSI, NFS)
- Миграция VM (AMPP)
- Любые топологии

- Особенности реализации у каждого вендора
- Все еще единый L2 домен
- Масштабируемость ограничена

У сложных систем – сложные сбои

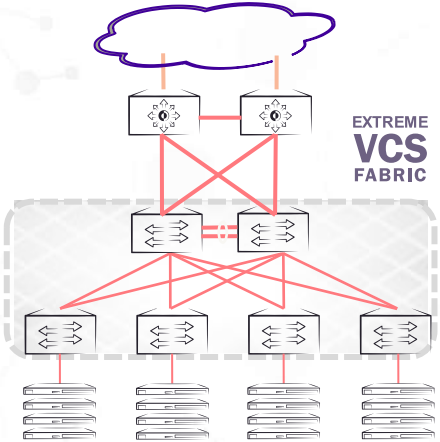
Эволюция сетевых архитектур

Традиционная



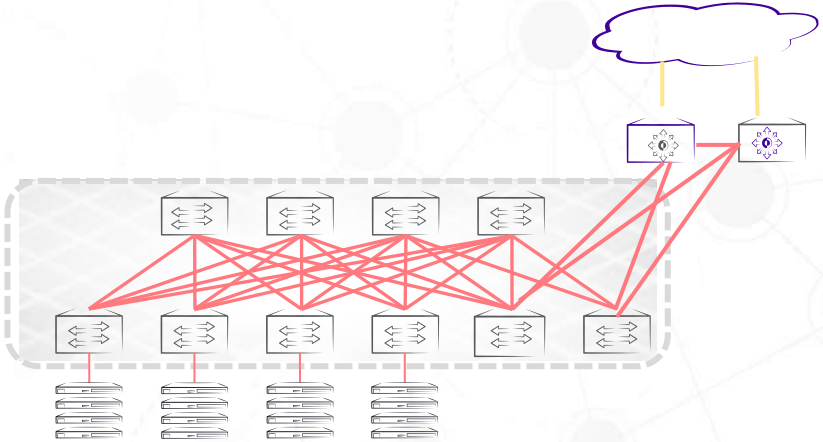
3-уровневая
Fat Tree

Влияние виртуализации



L2 Scale-out
TRILL, SPB[v,m], FP, VCS

Влияние облачных сервисов



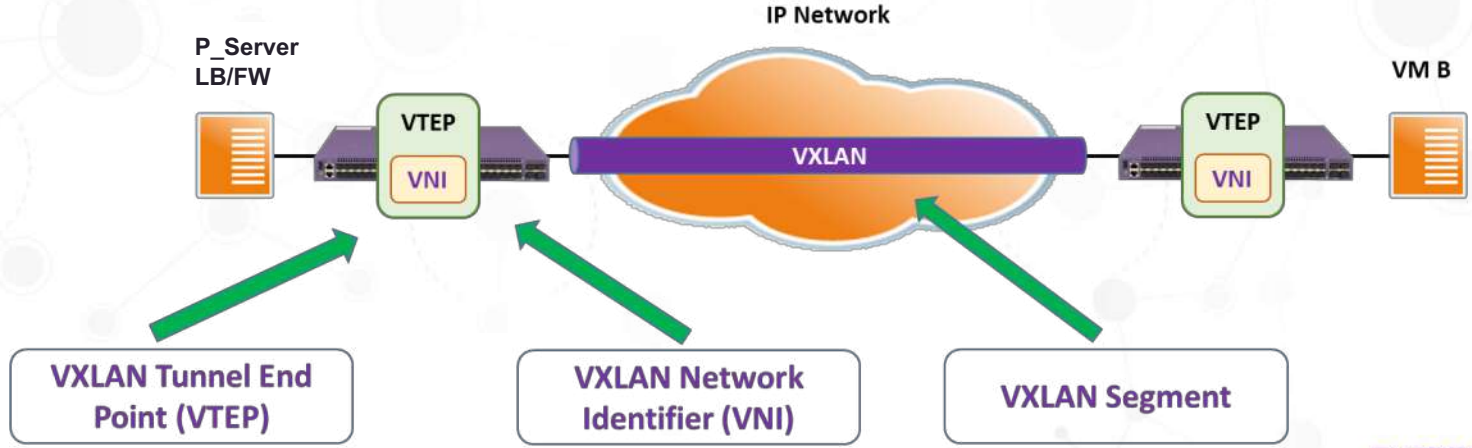
L3 Scale-out
VxLAN, BGP EVPN, Vmware NSX, OpenStack

У сложных систем – сложные сбои

HW VTEP vs vSW VTEP

VXLAN: терминология

- Работа VXLAN encap/decap происходит на VTEP (Virtual Tunnel Endpoint)
  - Local VTEP, Remote VTEP
  - IP/UDP инкапсуляция
  - S_port UDP обеспечивают отличную энтропию для балансировки
- VNI (VXLAN Network Identifier) - это тэг сегмента VXLAN
  - Длина 24bit ➔ ~16,7M уникальных сегментов



VXLAN: control plane

- VXLAN – просто инкапсуляция, не протокол!
- Multicast
  - Рекомендован RFC 7348
  - Требуется PIM-BiDir
  - Можно но все-таки сложно (dDP и sCP)

VXLAN: control plane

- VXLAN – просто инкапсуляция, не протокол!
- Multicast – кто хочет этим заниматься?
  - Рекомендован RFC 7348
  - Требуется PIM-BiDir
  - Можно но все-таки сложно (dDP и sCP)
- Controller with OVSDB
  - VMware NSX и другие
  - Дополнительные компоненты, цена, сложность?

VXLAN: control plane

- VXLAN – просто инкапсуляция, не протокол!
- Multicast – кто хочет этим заниматься?
  - Рекомендован RFC 7348
  - Требуется PIM-BiDir
  - Можно но все-таки сложно (dDP и sCP)
- Controller with OVSDB – есть необходимость?
  - VMware NSX и другие
  - Дополнительные компоненты, цена, сложность?
- IGP
  - OSPF – расширения LSA
- EVPN с MP-BGP
  - CP для overlay (чаще VXLAN)
  - MAC learning между VTEP
  - Аналог MPLS L3VPN
  - Multitenancy (RD и RT)

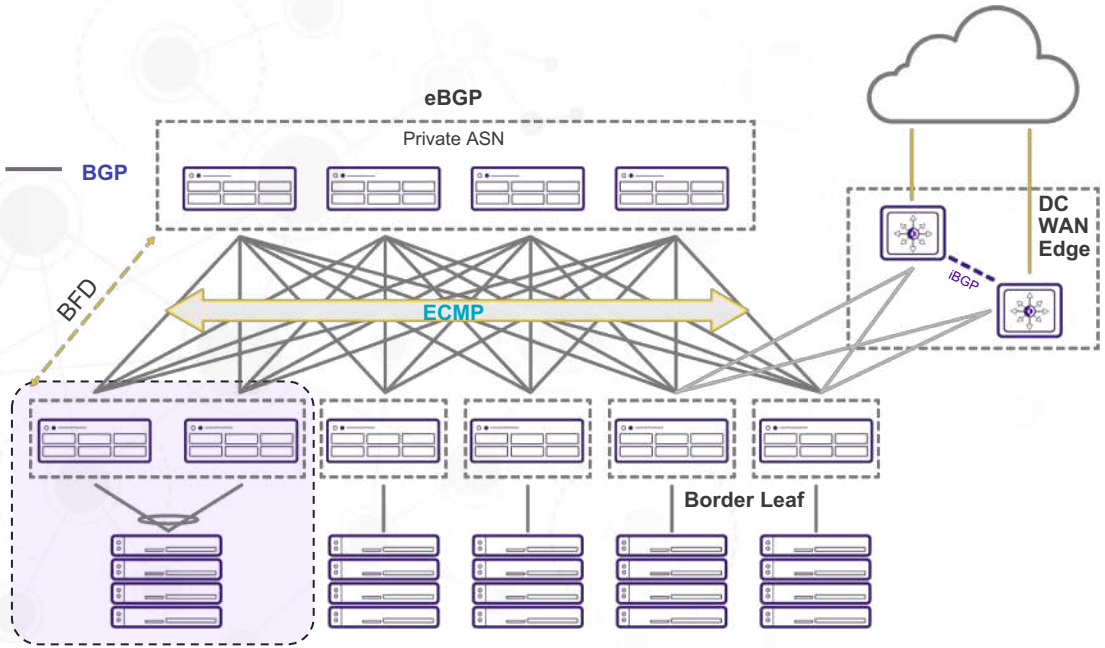
VXLAN: важно иметь ввиду

- Maximum transmission unit (MTU)
  - IP оверхед
  - Внутренние заголовки коммутаторов
  - Включаем Jumbo Frames! или уменьшаем на серверах.
- Для underlay можем использовать любой IGP и iBGP
  - Смотрим на возможности оборудования, но все умеют eBGP ☺
- IP адресация
  - Point-2-Point (/30 или .31) или IP unnumbered

```
!---Globally
device# configure terminal
device(config)# mtu 2000
!
!---per interface
device# configure terminal
device(config)# interface ethernet 0/1
device(conf-if-eth-0/1)# mtu 2000
```

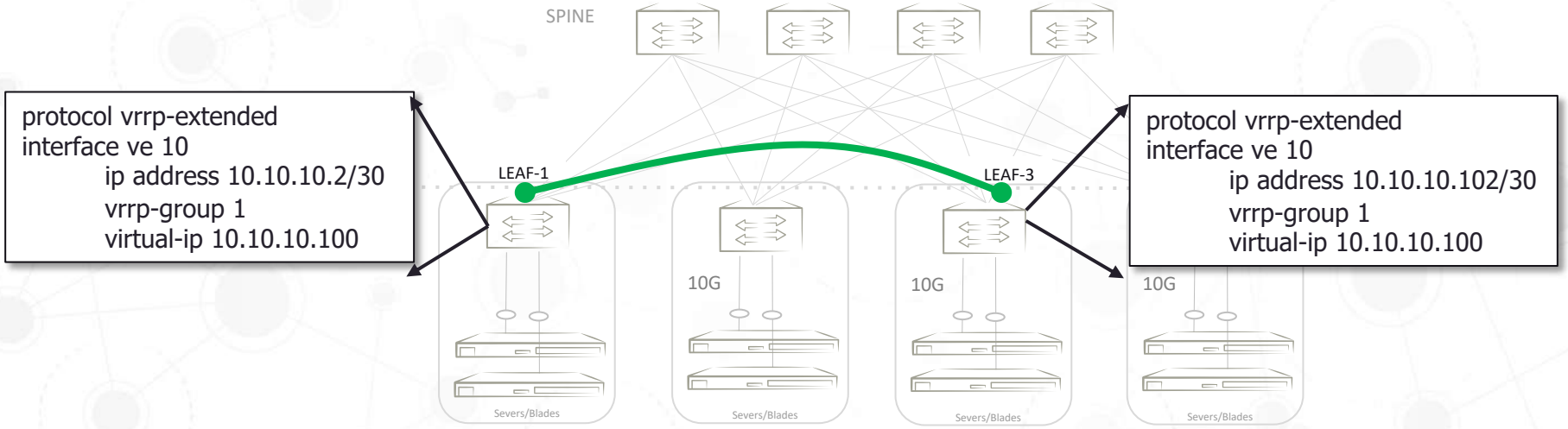
```
!---per port-channel
device# configure terminal
device(config)# interface po 10
device(conf-if-eth-0/1)# mtu 2000
!
device#sh run inter po 10
interface Port-channel 10
mtu 2000
switchport
switchport mode trunk
switchport trunk allowed vlan all
```

Extreme IP Fabric - Multihoming



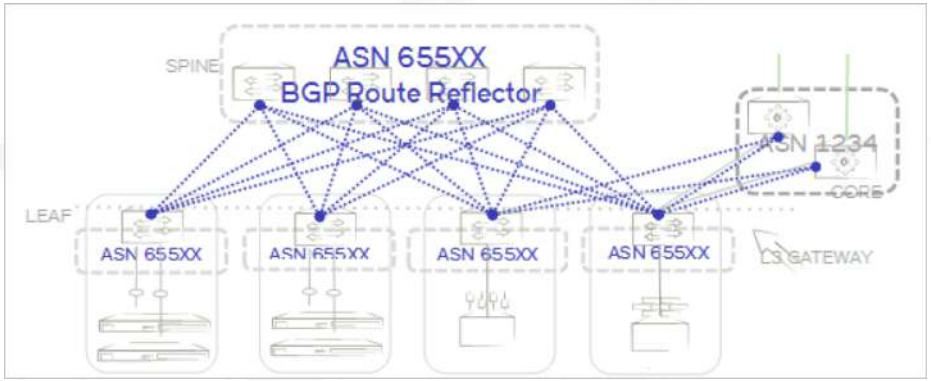
- Стандарт и проприетарные
  - EVPN
  - mLAG
- mLAG
  - vIP для VTEP
  - Только 2 узла
  - Поддержка LACP на клиенте
  - Требуется соединение двух Leaf
  - Отказоустойчивость (2S и L3 линк)
- EVPN
  - Стандарт
  - Более 2 узлов
  - A/S, 1A, mA
  - Не требуется линк между Leaf

L3 gateway – VRRP-E

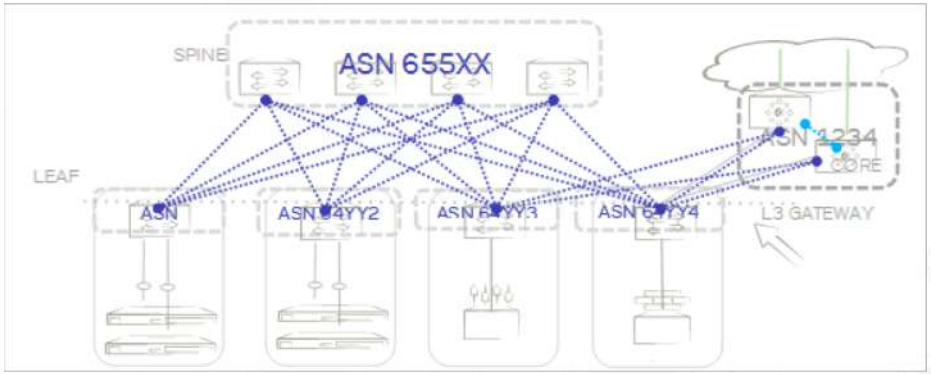


- Одна подсеть для всех Leaf
- Один vIP и MAC на всех Leaf для одного VXLAN туннеля
- VRRP-E = Short Path forwarding at Leaf.

Underlay routing



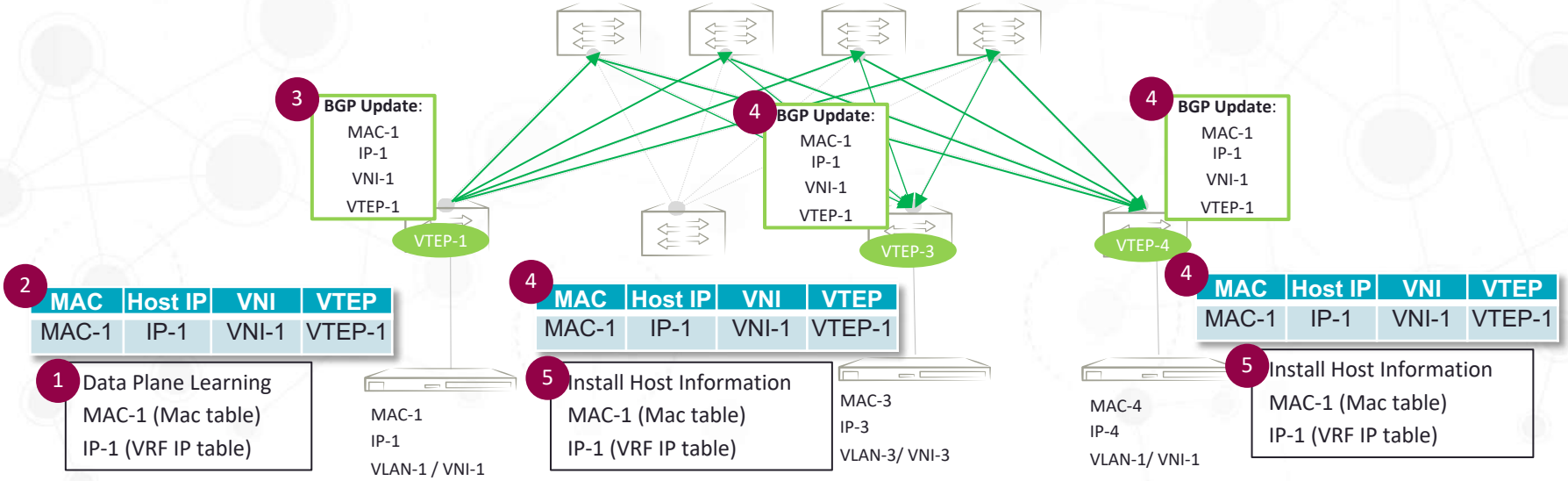
- iBGP
  - Spine как RR
  - Невозможно управлять (фильтровать) маршрутную информацию



- eBGP
  - Spine и Leaf в разных AS
  - Гибкое управление маршрутной информацией

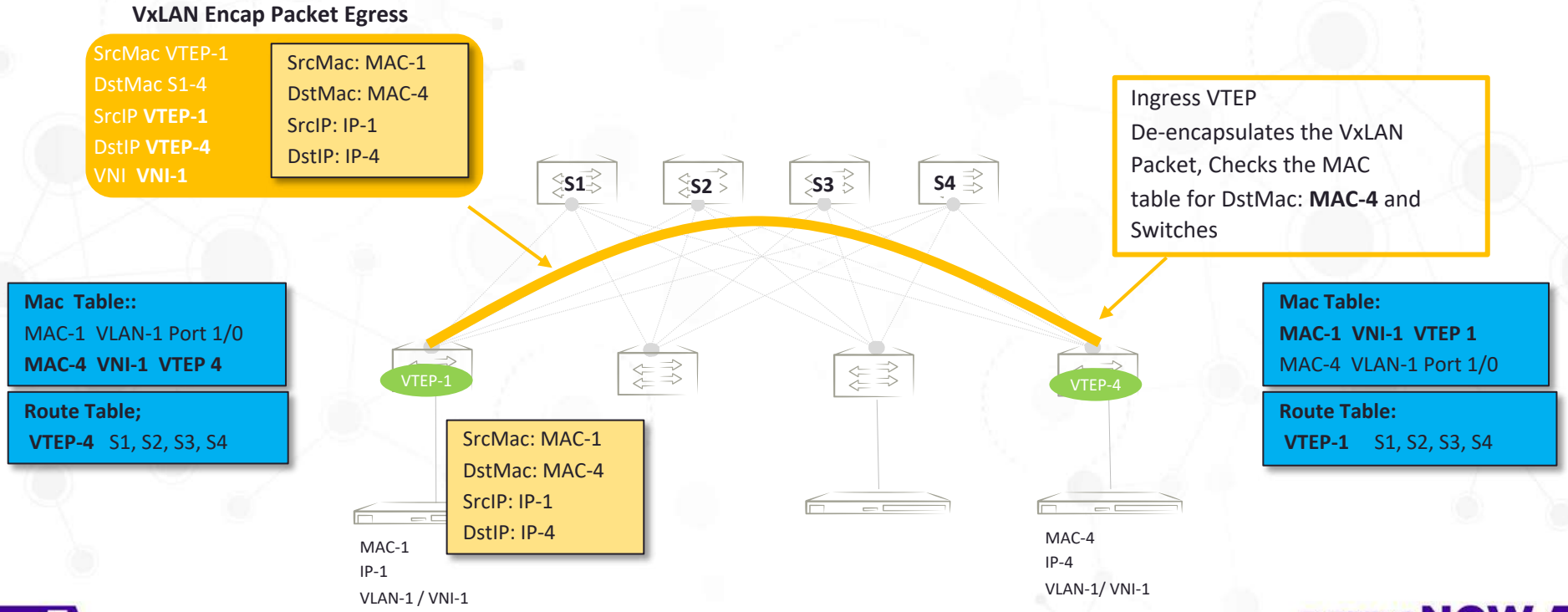
EVPN: Host NLRI Learning and Distribution

Control Plane Signaling.

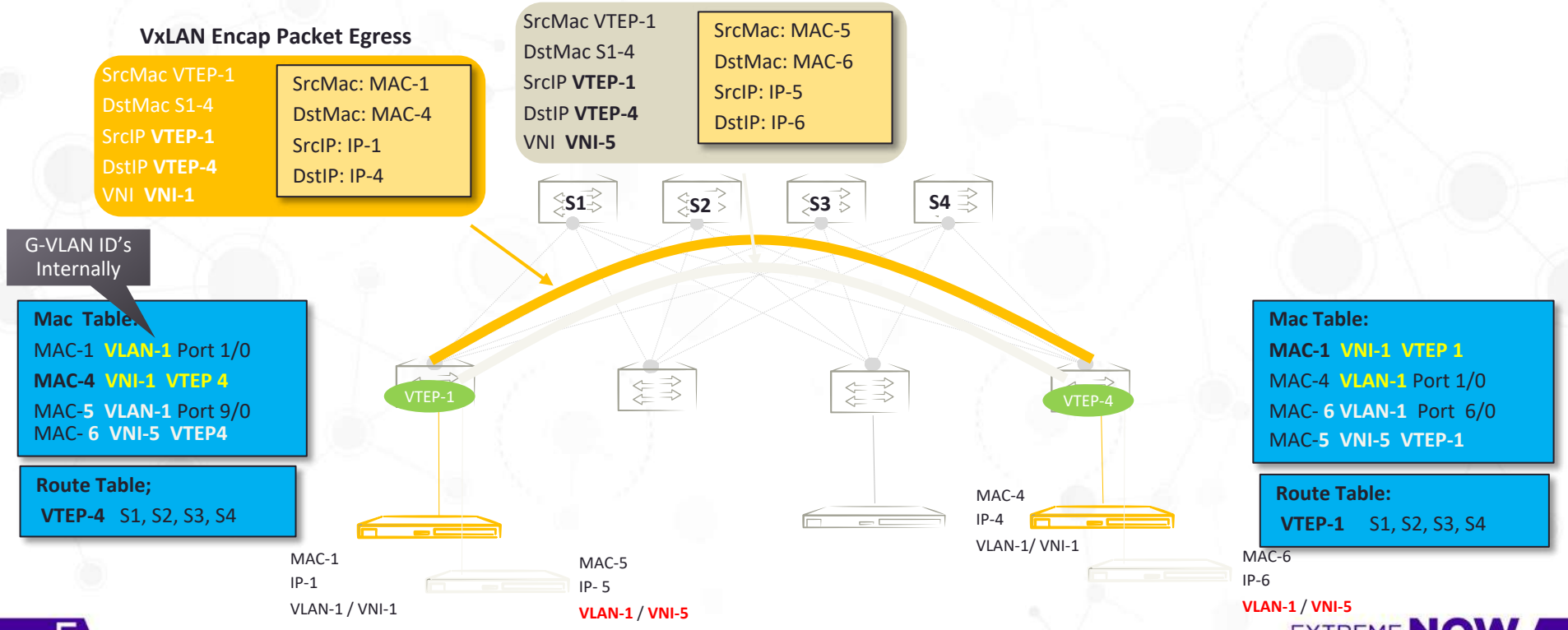


Use Case: Intra-Subnet Bridging

L2 Bridging

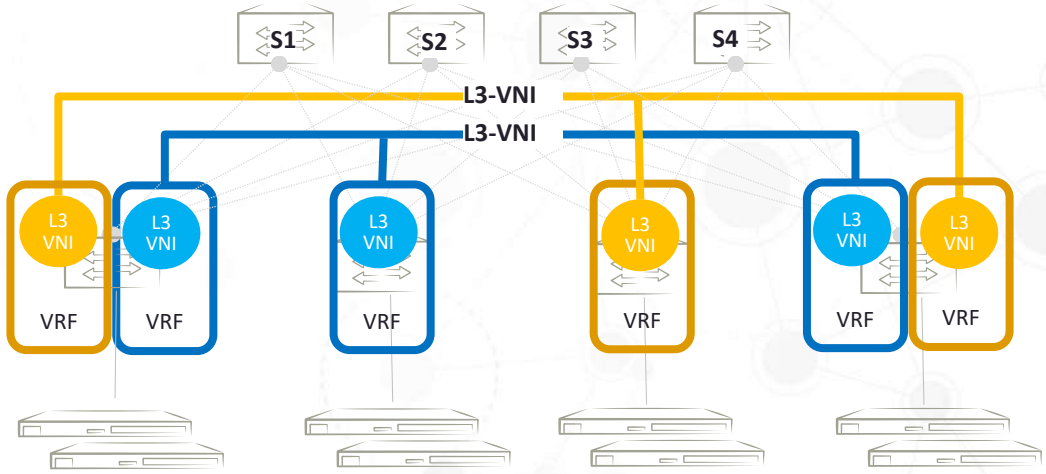


Use Case: L2 Bridging & Mult-Tenancy



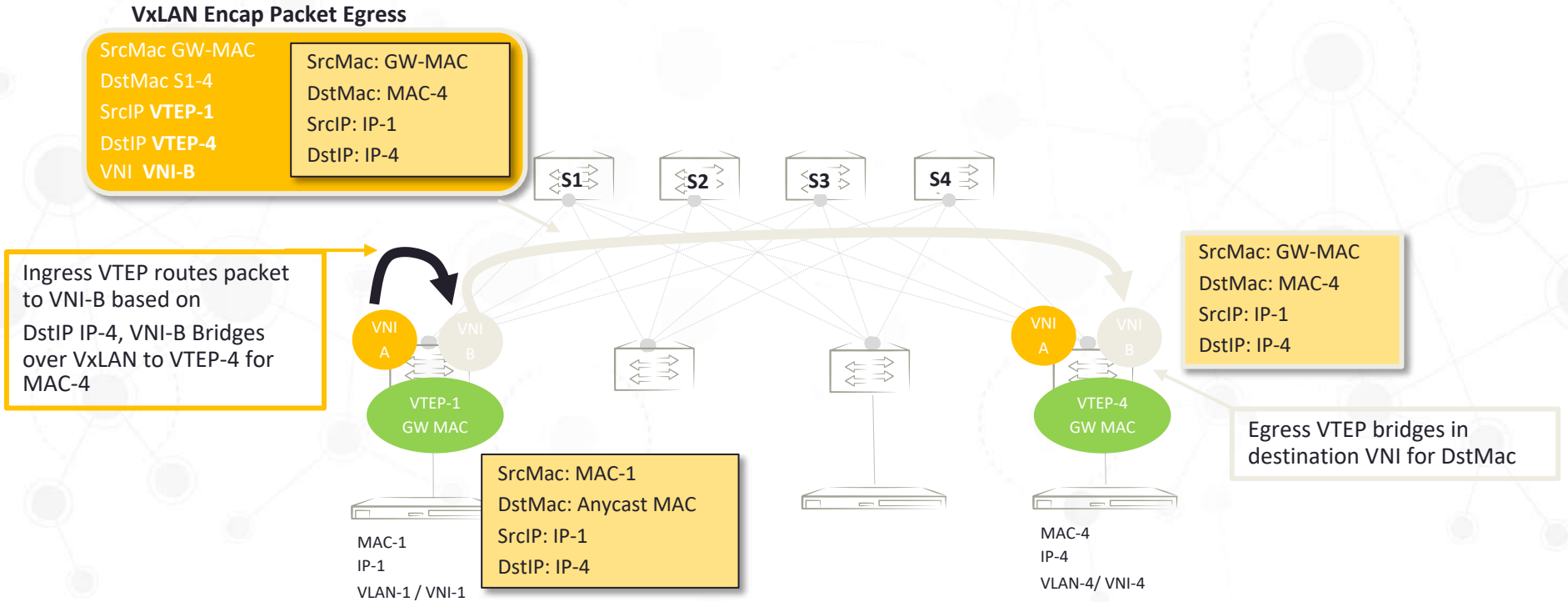
L3 Multi-Tenancy w/ VxLAN

- VxLAN Based L3 Multi-Tenancy
- VRF + L3 VNI
- Standards based Interop
- Single Protocol Instance for multiple VRF.
- No MPLS complexity
- RT/RD Import Export Policies supported
- Scale TBD Tenants/TOR



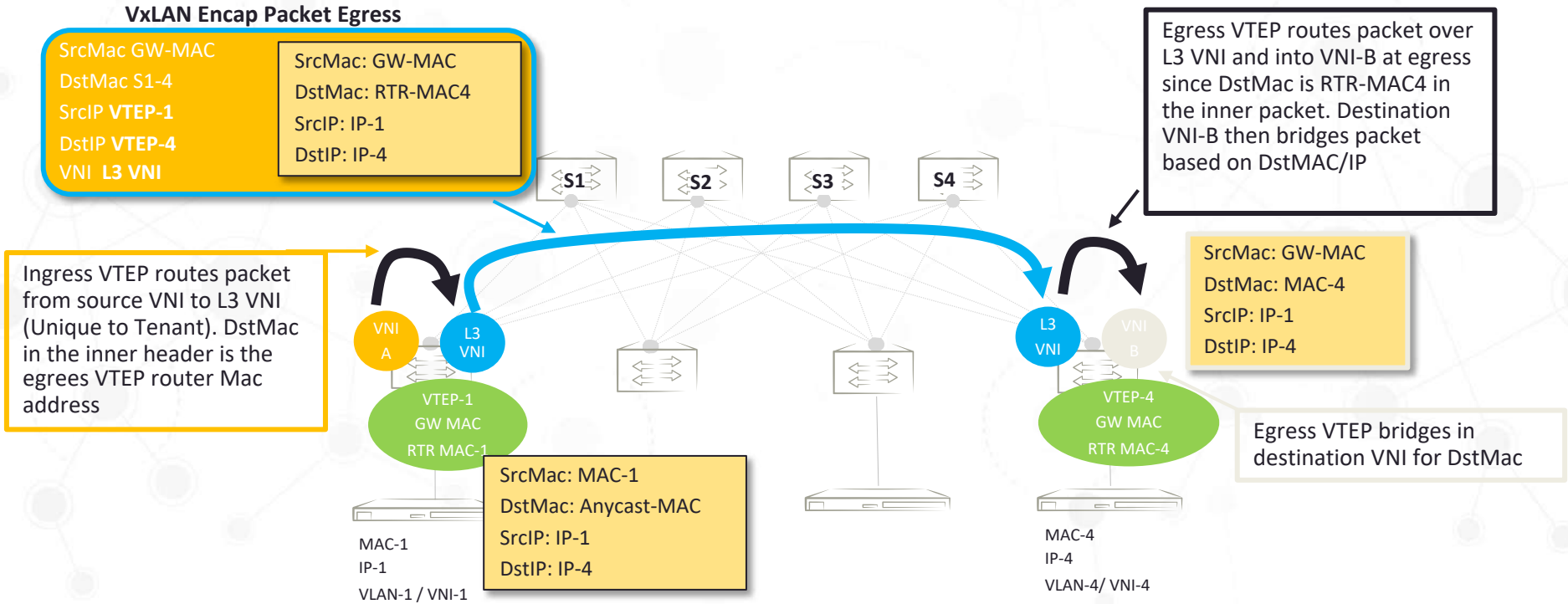
Use Case: Inter-Subnet Routing

Asymmetric Integrated Routing and Bridging

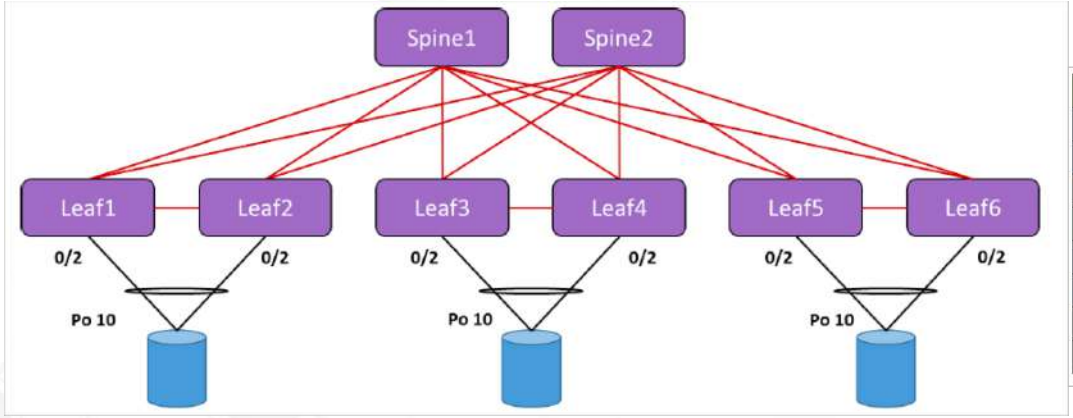


Use Case: Inter-Subnet Routing

Symmetric Integrated Routing and Bridging



Настройка IP фабрики



Switch	ASN	Loopback1 (router-id)	Loopback2 (VTEP IP)
Leaf1	64100	10.1.1.1	10.10.10.1
Leaf2	64100	10.1.1.2	10.10.10.1
Leaf3	64101	10.1.1.3	10.10.10.2
Leaf4	64101	10.1.1.5	10.10.10.2
Leaf5	64102	10.1.1.6	10.10.10.3
Leaf6	64102	10.1.1.7	10.10.10.3
Spine1	65000	10.1.1.8	N/A
Spine2	65000	10.1.1.9	N/A

Настройка IP фабрики - ручная

- Настраиваем MTU
- *Опционально* - если есть MCT, то начинаем с него
- Настраиваем eBGP между Spine и Leaf
- Настраиваем VXLAN

Настройка IP фабрики - ручная

- Настраиваем MTU
- *Опционально* - если есть MCT, то начинаем с него
- Настраиваем eBGP между Spine и Leaf
- Настраиваем VXLAN

- Всё! Кто сказал, что сложно!?

Настройка IP фабрики – Настраиваем MTU

```
!--- Globally
Leaf3(config)# mtu 9216
Leaf3(config)# ip mtu 9168
```

Настройка IP фабрики – Настраиваем MCT

```
!--- Globally
Leaf3(config)# mtu 9216
Leaf3(config)# ip mtu 9168
```

SLX-OS Layer 2 MCT is based on RFC 7432 (BGP Ethernet VPN).

☺ ☺ ☺ ☺ ☺ ☺ ☺ ☺ ☺ ☺

Настройка IP фабрики – Настраиваем MCT

```
Leaf3(config)# router ospf
Leaf3(config-router-ospf-vrf-default-vrf)# area 0
Leaf3(config-router-ospf-vrf-default-vrf)# exit
Leaf3(config)# vlan 4090
Leaf3(config-vlan-4000)# router-interface ve 4090
Leaf3(config-vlan-4000)# exit
Leaf3(config)# interface ve 4090
Leaf3(config-if-Ve-4000)# ip ospf area 0
Leaf3(config-if-Ve-4000)# ip address 10.10.1.5/30
Leaf3(config-if-Ve-4000)# no shut
Leaf3(config-if-Ve-4000)# exit
Leaf3(config)# interface Loopback 1
Leaf3(config-Loopback-1)# ip ospf area 0
Leaf3(config-Loopback-1)# ip address 10.1.1.3/32
Leaf3(config-Loopback-1)# no shut
Leaf3(config-Loopback-1)# exit
Leaf3(config)# interface Ethernet 0/1
Leaf3(conf-if-eth-0/1)# description MCT
Leaf3(conf-if-eth-0/1)# switchport
Leaf3(conf-if-eth-0/1)# switchport mode trunk
Leaf3(conf-if-eth-0/1)# switchport trunk allowed vlan add 4090
Leaf3(conf-if-eth-0/1)# no shut
Leaf3(conf-if-eth-0/1)# router bgp
Leaf3(config-bgp-router)# local-as 64102
Leaf3(config-bgp-router)# neighbor 10.1.1.4 remote-as 64102
Leaf3(config-bgp-router)# neighbor 10.1.1.4 update-source loopback 1
Leaf3(config-bgp-router)# address-family l2vpn evpn
Leaf3(config-bgp-evpn)# neighbor 10.1.1.4 encapsulation [mpls | nsh |
vxlan]
Leaf3(config-bgp-evpn)# neighbor 10.1.1.4 activate
```

- Настраиваем OSPF – IGP для iBGP
- Настраиваем VLAN 4090
  - Контрольный VLAN
  - Рекомендуемый номер для MCT
  - Создаем VRI
- Создаем Lo для iBGP
- Добавляем интерфейс в MCT VLAN 4090
- Запускаем iBGP, используя MPLS инкапсуляцию
- Всё?

Настройка IP фабрики – Настраиваем MCT

```
Leaf3(config)# router ospf
Leaf3(config-router-ospf-vrf-default-vrf)# area 0
Leaf3(config-router-ospf-vrf-default-vrf)# exit
Leaf3(config)# vlan 4090
Leaf3(config-vlan-4000)# router-interface ve 4090
Leaf3(config-vlan-4000)# exit
Leaf3(config)# interface ve 4090
Leaf3(config-if-Ve-4000)# ip ospf area 0
Leaf3(config-if-Ve-4000)# ip address 10.10.1.5/30
Leaf3(config-if-Ve-4000)# no shut
Leaf3(config-if-Ve-4000)# exit
Leaf3(config)# interface Loopback 1
Leaf3(config-Loopback-1)# ip ospf area 0
Leaf3(config-Loopback-1)# ip address 10.1.1.3/32
Leaf3(config-Loopback-1)# no shut
Leaf3(config-Loopback-1)# exit
Leaf3(config)# interface Ethernet 0/1
Leaf3(conf-if-eth-0/1)# description MCT
Leaf3(conf-if-eth-0/1)# switchport
Leaf3(conf-if-eth-0/1)# switchport mode trunk
Leaf3(conf-if-eth-0/1)# switchport trunk allowed vlan add 4090
Leaf3(conf-if-eth-0/1)# no shut
Leaf3(conf-if-eth-0/1)# router bgp
Leaf3(config-bgp-router)# local-as 64102
Leaf3(config-bgp-router)# neighbor 10.1.1.4 remote-as 64102
Leaf3(config-bgp-router)# neighbor 10.1.1.4 update-source loopback 1
Leaf3(config-bgp-router)# address-family l2vpn evpn
Leaf3(config-bgp-evpn)# neighbor 10.1.1.4 encapsulation [mpls | nsh |
vxlan]
Leaf3(config-bgp-evpn)# neighbor 10.1.1.4 activate
```

- Настраиваем OSPF – IGP для iBGP
- Настраиваем VLAN 4090
  - Контрольный VLAN
  - Рекомендуемый номер для MCT
  - Создаем VRI
- Создаем Lo для iBGP
- Добавляем интерфейс в MCT VLAN 4090
- Запускаем iBGP, используя MPLS инкапсуляцию

- НЕТ!

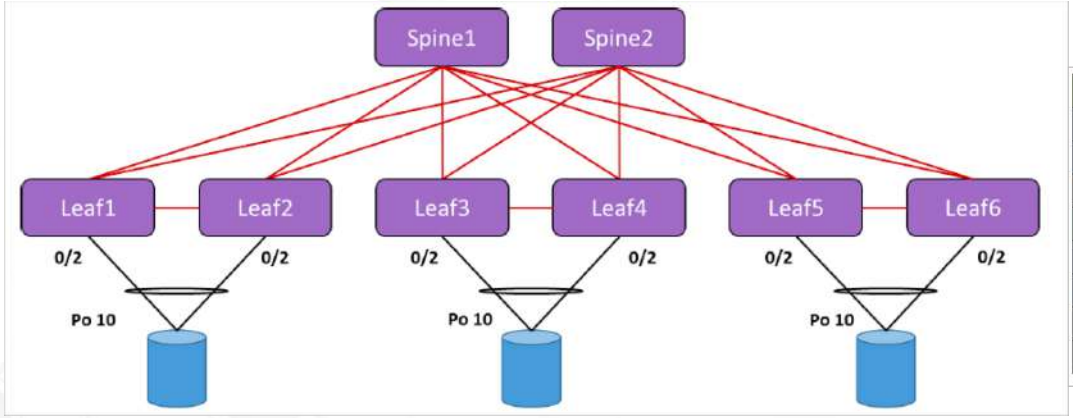
Настройка IP фабрики – Настраиваем MCT

```
Leaf3# cluster management node-id 1
!
Leaf3(config)# cluster MCT1 1
Leaf3(config-cluster-1)# peer-interface ve 4090
Leaf3(config-cluster-1)# peer 10.1.1.4
Leaf3(config-cluster-1)# deploy
!
Leaf3(config)# interface Ethernet 0/2
Leaf3(conf-if-eth-0/2)# channel-group 10 mode active type standard
Leaf3(conf-if-eth-0/2)# no shut
Leaf3(conf-if-eth-0/2)# exit
Leaf3(config)#
Leaf3(config)# interface Port-channel 10
Leaf3(config-Port-channel-10)# switchport
Leaf3(config-Port-channel-10)# switchport mode trunk
Leaf3(config-Port-channel-10)# switchport trunk allowed vlan add
100,300
Leaf3(config-Port-channel-10)# no shutdown
Leaf3(config-Port-channel-10)# exit
Leaf3(config)#
Leaf3(config)# cluster MCT1 1
Leaf3(config-cluster-1)# client mlag1 1
Leaf3(config-cluster-client-1)# client-interface port-channel 10
Leaf3(config-cluster-client-1)# esi auto lacp
Leaf3(config-cluster-client-1)# deploy
```

- НЕТ! Теперь настраиваем кластер))
- Настраиваем одинаковый cluster ID
- Добавляем клиентские порты в MCT
- Создаем Po10
- И привязываем его к кластеру

- Вот теперь всё!

Настройка IP фабрики – Настраиваем BGP



Switch	ASN	Loopback1 (router-id)	Loopback2 (VTEP IP)
Leaf1	64100	10.1.1.1	10.10.10.1
Leaf2	64100	10.1.1.2	10.10.10.1
Leaf3	64101	10.1.1.3	10.10.10.2
Leaf4	64101	10.1.1.5	10.10.10.2
Leaf5	64102	10.1.1.6	10.10.10.3
Leaf6	64102	10.1.1.7	10.10.10.3
Spine1	65000	10.1.1.8	N/A
Spine2	65000	10.1.1.9	N/A

Настройка IP фабрики – Настраиваем BGP

- Настраиваем базовую связность

```
Leaf1(config)# interface Ethernet 0/3
Leaf1(conf-if-eth-0/3)# ip address 10.20.10.0/31
Leaf1(conf-if-eth-0/3)# description To-Spine1
Leaf1(conf-if-eth-0/3)# no shut
Leaf1(conf-if-eth-0/3)# exit
Leaf1(config)#
Leaf1(config)# interface Ethernet 0/4
Leaf1(conf-if-eth-0/4)# ip address 10.30.10.0/31
Leaf1(conf-if-eth-0/4)# description To-Spine2
Leaf1(conf-if-eth-0/4)# no shut
Leaf1(conf-if-eth-0/4)# exit
```

- Настраиваем underlay BGP
- Leaf to Spine1: 10.20.x0.0/31 with x the leaf number.
- Leaf to Spine2: 10.30.x0.0/31 with x the leaf number.
- The Spine is always .1.
- Настраиваем только **один** раз в Day0

```
Leaf1(config)# router bgp
Leaf1(config-bgp-router)# local-as 64100
Leaf1(config-bgp-router)# fast-external-fallover
Leaf1(config-bgp-router)# bfd interval 300 min-rx 300 multiplier 3
Leaf1(config-bgp-router)# neighbor 10.20.10.1 remote-as 65000
Leaf1(config-bgp-router)# neighbor 10.20.10.1 bfd
Leaf1(config-bgp-router)# neighbor 10.30.10.1 remote-as 65000
Leaf1(config-bgp-router)# neighbor 10.30.10.1 bfd
Leaf1(config-bgp-router)# address-family ipv4 unicast
Leaf1(config-bgp-ipv4u)# maximum-paths 8
Leaf1(config-bgp-ipv4u)# exit
Leaf1(config-bgp-router)# exit
```

Настройка IP фабрики – Настраиваем BGP

- Настраиваем базовую связность

```
Leaf1(config)# interface Ethernet 0/3
Leaf1(conf-if-eth-0/3)# ip address 10.20.10.0/31
Leaf1(conf-if-eth-0/3)# description To-Spine1
Leaf1(conf-if-eth-0/3)# no shut
Leaf1(conf-if-eth-0/3)# exit
Leaf1(config)#
Leaf1(config)# interface Ethernet 0/4
Leaf1(conf-if-eth-0/4)# ip address 10.30.10.0/31
Leaf1(conf-if-eth-0/4)# description To-Spine2
Leaf1(conf-if-eth-0/4)# no shut
Leaf1(conf-if-eth-0/4)# exit
```

- Настраиваем underlay BGP
- Leaf to Spine1: 10.20.x0.0/31 with x the leaf number.
- Leaf to Spine2: 10.30.x0.0/31 with x the leaf number.
- The Spine is always .1.
- Настраиваем только **один** раз в Day0

```
Leaf1(config)# router bgp
Leaf1(config-bgp-router)# local-as 64100
Leaf1(config-bgp-router)# fast-external-fallover
Leaf1(config-bgp-router)# bfd interval 300 min-rx 300 multiplier 3
Leaf1(config-bgp-router)# neighbor 10.20.10.1 remote-as 65000
Leaf1(config-bgp-router)# neighbor 10.20.10.1 bfd
Leaf1(config-bgp-router)# neighbor 10.30.10.1 remote-as 65000
Leaf1(config-bgp-router)# neighbor 10.30.10.1 bfd
Leaf1(config-bgp-router)# address-family ipv4 unicast
Leaf1(config-bgp-ipv4u)# maximum-paths 8
Leaf1(config-bgp-ipv4u)# exit
Leaf1(config-bgp-router)# exit
```

Настройка IP фабрики – Настраиваем BGP

- Настраиваем Spine

```
Spine2(config)# interface Loopback 1
Spine2(config-Loopback-1)# ip address 10.1.1.9/32
Spine2(config-Loopback-1)# no shut
Spine2(config-Loopback-1)# exit
Spine2(config)#
Spine2(config)# ip router-id 10.1.1.9
Spine2(config)#
Spine2(config)# router bgp
Spine2(config-bgp-router)# local-as 65000
Spine2(config-bgp-router)# fast-external-fallover
Spine2(config-bgp-router)# bfd interval 300 min-rx 300 multiplier 3
Spine2(config-bgp-router)# neighbor leaf-group peer-group
Spine2(config-bgp-router)# neighbor leaf-group bfd
Spine2(config-bgp-router)# neighbor 10.30.10.0 remote-as 64100
Spine2(config-bgp-router)# neighbor 10.30.10.0 peer-group leaf-group
Spine2(config-bgp-router)# neighbor 10.30.20.0 remote-as 64100
Spine2(config-bgp-router)# neighbor 10.30.20.0 peer-group leaf-group
Spine2(config-bgp-router)# neighbor 10.30.30.0 remote-as 64101
Spine2(config-bgp-router)# neighbor 10.30.30.0 peer-group leaf-group
Spine2(config-bgp-router)# neighbor 10.30.40.0 remote-as 64101
Spine2(config-bgp-router)# neighbor 10.30.40.0 peer-group leaf-group
Spine2(config-bgp-router)# neighbor 10.30.50.0 remote-as 64102
Spine2(config-bgp-router)# neighbor 10.30.50.0 peer-group leaf-group
Spine2(config-bgp-router)# neighbor 10.30.60.0 remote-as 64102
Spine2(config-bgp-router)# neighbor 10.30.60.0 peer-group leaf-group
Spine2(config-bgp-router)# exit
```

Настройка IP фабрики – Настраиваем BGP

- Проверяем

```
Leaf1# show ip bgp summary
  BGP4 Summary
  Router ID: 10.1.1.1   Local AS Number: 64100
  Confederation Identifier: not configured
  Confederation Peers:
  Maximum Number of IP ECMP Paths Supported for Load Sharing: 8
  Number of Neighbors Configured: 3, UP: 3
  Number of Routes Installed: 0
  Number of Routes Advertising to All Neighbors: 0 (0 entries)
  Number of Attribute Entries Installed: 0
  '+': Data in InQueue '>': Data in OutQueue '-': Clearing
  '*': Update Policy 'c': Group change 'p': Group change Pending
  'r': Restarting 's': Stale '^': Up before Restart '<': EOR waiting
  Neighbor Address  AS#         State     Time      Rt:Accepted Filtered Sent     ToSend
  10.1.1.2          64100       ESTAB     13h41m38s     0          0        0          0
  10.20.10.1        65000       ESTAB     13h57m32s     0          0        0          0
  10.30.10.1        65000       ESTAB     0h9m39s       0          0        0          0
```

Настройка IP фабрики – Настраиваем BGP

- Проверяем

```
Leaf1# show ip bgp summary
  BGP4 Summary
  Router ID: 10.1.1.1   Local AS Number: 64100
  Confederation Identifier: not configured
  Confederation Peers:
  Maximum Number of IP ECMP Paths Supported for Load Sharing: 8
  Number of Neighbors Configured: 3, UP: 3
  Number of Routes Installed: 0
  Number of Routes Advertising to All Neighbors: 0 (0 entries)
  Number of Attribute Entries Installed: 0
  '+': Data in InQueue '>': Data in OutQueue '-': Clearing
  '*': Update Policy 'c': Group change 'p': Group change Pending
  'r': Restarting 's': Stale '^': Up before Restart '<': EOR waiting
  Neighbor Address  AS#         State     Time      Rt:Accepted Filtered Sent     ToSend
  10.1.1.2          64100       ESTAB     13h41m38s     0          0        0          0
  10.20.10.1        65000       ESTAB     13h57m32s     0          0        0          0
  10.30.10.1        65000       ESTAB     0h9m39s       0          0        0          0
```

- В фабрике пока нет связности между серверами, т.к. еще не настроили overlay

Настройка IP фабрики – Настраиваем overlay BGP

- Нужно настроить CP и VXLAN

```
Leaf1(config)# interface Loopback 2
Leaf1(config-Loopback-2)# ip address 10.10.10.1/32
Leaf1(config-Loopback-2)# no shut
Leaf1(config-Loopback-2)# exit
Leaf1(config)#
Leaf1(config)# router bgp
Leaf1(config-bgp-router)# address-family l2vpn evpn
Leaf1(config-bgp-evpn)# neighbor 10.20.10.1 encapsulation vxlan
Leaf1(config-bgp-evpn)# neighbor 10.20.10.1 allowas-in 1
Leaf1(config-bgp-evpn)# neighbor 10.20.10.1 enable-peer-as-check
Leaf1(config-bgp-evpn)# neighbor 10.20.10.1 activate
Leaf1(config-bgp-evpn)# neighbor 10.30.10.1 encapsulation vxlan
Leaf1(config-bgp-evpn)# neighbor 10.30.10.1 allowas-in 1
Leaf1(config-bgp-evpn)# neighbor 10.30.10.1 enable-peer-as-check
Leaf1(config-bgp-evpn)# neighbor 10.30.10.1 activate
Leaf1(config-bgp-evpn)# exit
Leaf1(config-bgp-router)# address-family ipv4 unicast
Leaf1(config-bgp-ipv4u)# network 10.10.10.1/32
Leaf1(config-bgp-ipv4u)# exit
Leaf1(config-bgp-router)# exit
```

Настройка IP фабрики – Настраиваем overlay BGP

- Нужно настроить CP и VXLAN

```
Spine1(config)# router bgp
Spine1(config-bgp-router)# address-family l2vpn evpn
Spine1(config-bgp-evpn)# retain route-target all
Spine1(config-bgp-evpn)# neighbor leaf-group encapsulation vxlan
Spine1(config-bgp-evpn)# neighbor leaf-group next-hop-unchanged
Spine1(config-bgp-evpn)# neighbor leaf-group activate
Spine1(config-bgp-evpn)# exit
```

Настройка IP фабрики – Настраиваем VXLAN

- Нужно настроить CP и VXLAN

```
Leaf1(config)# overlay-gateway PoD1
Leaf1(config-overlay-gw-PoD1)# type layer2-extension
Leaf1(config-overlay-gw-PoD1)# ip interface Loopback 2
Leaf1(config-overlay-gw-PoD1)# map vni auto
Leaf1(config-overlay-gw-PoD1)# activate
Leaf1(config-overlay-gw-PoD1)# exit
```

Настройка IP фабрики – Настраиваем VXLAN

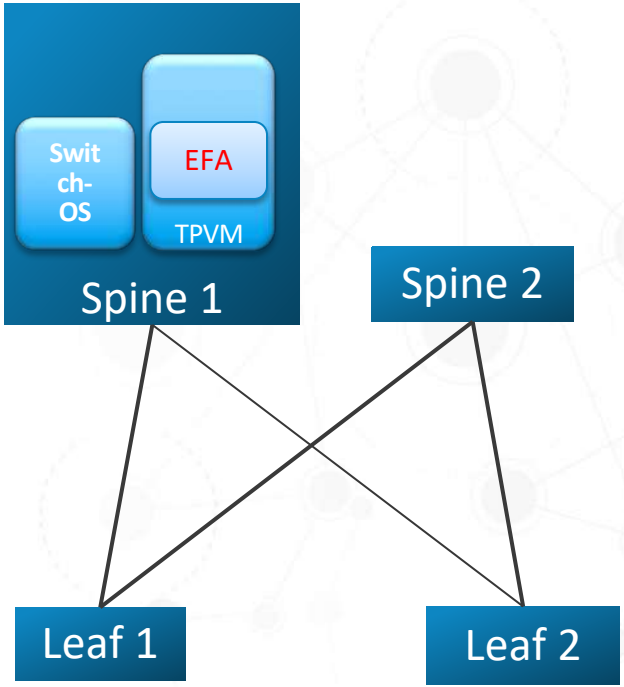
- Нужно настроить CP и VXLAN

```
Leaf1(config)# overlay-gateway PoD1
Leaf1(config-overlay-gw-PoD1)# type layer2-extension
Leaf1(config-overlay-gw-PoD1)# ip interface Loopback 2
Leaf1(config-overlay-gw-PoD1)# map vni auto
Leaf1(config-overlay-gw-PoD1)# activate
Leaf1(config-overlay-gw-PoD1)# exit
```

- Настройка завершена! Фабрика в рабочем состоянии, но сервисы не подняты, т.к. нет хостов, подключенных к Leaf.
- Дальше начинается операционная деятельность:
  - Настройка “MLAG”, VLAN, VRF, добавление vlan в EVPN.
  - Настройка маршрутизации между VXLAN

Embedded Fabric Automation

- Приложение для авто настройки фабрики
- Встроено в ПО коммутатора (bin pack)
  Не требуется сторонний софт
- Устанавливается в TPVM
- Одна команда для инсталляции



Процесс запуска EFA и настройки фабрики

Step 1: Run the command:
“*efa deploy*” on a Spine switch

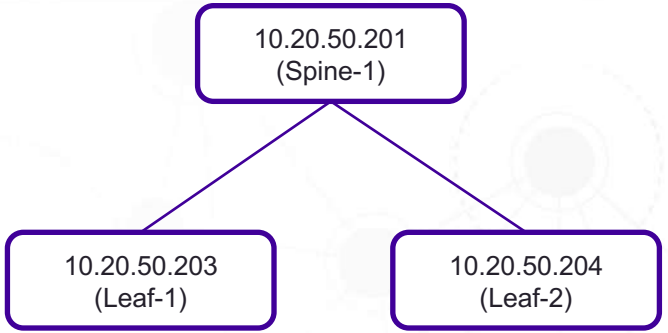
Step 2: run the command :
“*show tpvm ip-address*”,

Step 3: SSH to the 3rd party virtual machine (tpvm) (to the ip in Step 2 above)

Step 4: Run the command
“*efa fabric deploy --spine x.x.x.x --leaf x.x.x.x*”
(x.x.x.x – ip address of Spine and Leaf switches respectively)

Configure IP Fabric

- Configure IP Fabric having one Spine(10.20.50.201) and two leaves (10.20.50.203,10.20.50.204)
- Configurations pushed to the switches
  - Interfaces between Leaf and Spine switches
  - Router BGP Configurations
  - MCT (in case of two node leaves)
  - Overlay Configurations



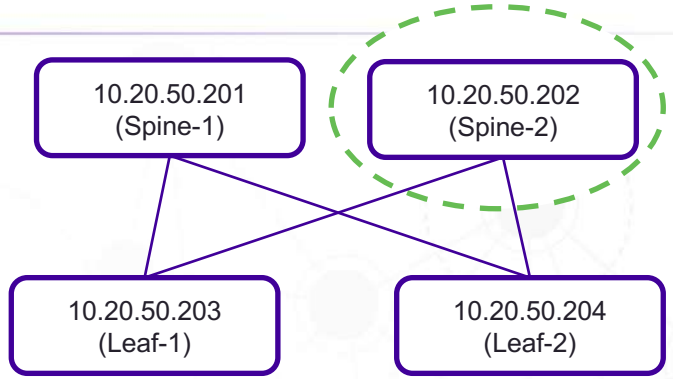

```
root@TPVM:~# efa fabric configure --spine 10.20.50.201 --leaf 10.20.50.203,10.20.50.204
Add Device(s) [Success]
        Addition of Spine device with ip-address = 10.20.50.201 [Succeeded]
        Addition of Leaf device with ip-address = 10.20.50.204 [Succeeded]
        Addition of Leaf device with ip-address = 10.20.50.203 [Succeeded]

Validate Fabric [Success]

Configure Fabric [Success]
--- Time Elapsed: 41.599848748s ---
```

Adding more Switches to the fabric

- Adding one Spine (10.20.50.202)
  - This will add the Spine (10.20.50.202) to the existing IP Fabric.
- Similarly one or more Spine or Leaf devices can be added to the IP Fabric



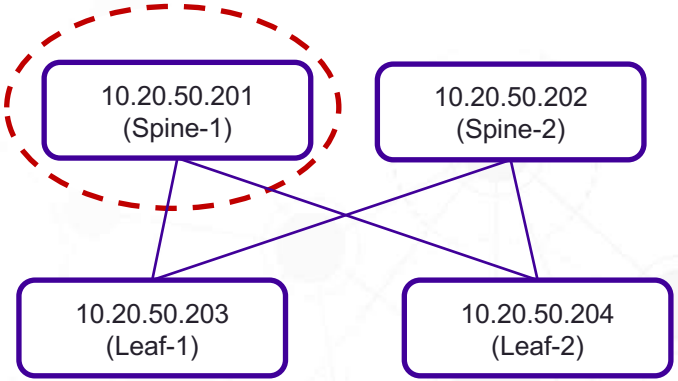

```
root@TPVM:~# efa fabric configure --spine 10.20.50.202
Add Device(s) [Success]
        Addition of Spine device with ip-address = 10.20.50.201 [Succeeded]
        Addition of Leaf device with ip-address = 10.20.50.203 [Succeeded]
        Addition of Spine device with ip-address = 10.20.50.202 [Succeeded]
        Addition of Leaf device with ip-address = 10.20.50.204 [Succeeded]

Validate Fabric [Success]

Configure Fabric [Success]
--- Time Elapsed: 9.98110426s ---
```

Deconfigure – Remove Switches from the IP Fabric

- Removing Spine Device (10.20.50.201) from the IP Fabric
  - This would remove the device and delete the configurations from the switch.
- Similarly one or more Spine or Leaf devices can be deleted from the IP Fabric




```
root@TPVM:~# efa fabric deconfigure --device 10.20.50.201
Delete Device(s) [Success]
        Deletion of  device with ip-address = 10.20.50.201 [Succeeded]
--- Time Elapsed: 3.536408776s ---
```

Fabric Setting Show

```
root@TPVM:~# efa fabric setting show
+---------------------------+-----------------+
|           NAME            |      VALUE      |
+---------------------------+-----------------+
| Fabric Name               | default         |
| Link IP Range             | 10.10.10.0/23   |
| Loopback IP Range         | 172.32.254.0/24 |
| Loopback Port Number      | 1               |
| VTEP Loopback Port Number | 2               |
| Spine ASN Block           | 64512           |
| Leaf ASN Block            | 65000-65534     |
| P2P IP Type               | numbered        |
+---------------------------+-----------------+
```

- IP Fabric gets created using the default settings
- Fabric Settings can be displayed using “efa fabric setting show”
  - “—advanced” flag can be used to display all the setting attributes

```
root@TPVM:~# efa fabric setting show --advanced
+---------------------------------+-----------------+
|              NAME               |      VALUE      |
+---------------------------------+-----------------+
| Fabric Name                     | default         |
| Link IP Range                   | 10.10.10.0/23   |
| Loopback IP Range               | 172.32.254.0/24 |
| Loopback Port Number            | 1               |
| VTEP Loopback Port Number       | 2               |
| Spine ASN Block                 | 64512           |
| Leaf ASN Block                  | 65000-65534     |
| P2P IP Type                     | numbered        |
| Any cast MAC                    | 0201.0101.0101  |
| IPV6 Any cast MAC               | 0201.0101.0102  |
| ARP Aging Timeout               | 300             |
| MAC Aging Timeout               | 1800            |
| MAC Aging Conversational        | 300             |
| Timeout                         |                 |
| MAC Move Limit                  | 20              |
| Duplicate MAC Timer             | 5               |
| Duplicate MAC Timer MAX Count   | 3               |
| Configure Overlay Gateway       | Yes             |
| BFD Tx                          | 300             |
| BFD Rx                          | 300             |
| BFD Multiplier                  | 3               |
| BGP MultiHop                    | 2               |
| MaxPaths                        | 8               |
| AllowAsIn                       | 1               |
| MTU                             | 9216            |
| IPMTU                           | 9100            |
| Leaf PeerGroup                  | spine-group     |
| Spine PeerGroup                 | leaf-group      |
| MCT Link IP Range               | 20.20.20.0/24   |
| MCT PortChannel                 | 1024            |
| Control Vlan                    | 4090            |
| Control VE                      | 4090            |
| VNI Auto Map                    | Yes             |
+---------------------------------+-----------------+
-- Time Elapsed: 53.744005ms ---
```

Fabric Setting Update

```
root@TPVM:~# efa fabric setting update --help
Update fabric settings.

Usage:
  efa fabric setting update [flags]

Flags:
      --p2p-link-range string                 Range Of IP Address.
      --loopback-ip-range string              Range Of IP Address
      --loopback-port-number string           Loopback Port Number <NUMBER: 1-255>
      --vtep-loopback-port-number string      VTEP Loopback Port Number <NUMBER: 1-255>
      --spine-asn-block string                Spine ASN Range Separated -;Or Single AS
      --leaf-asn-block string                 Leaf ASN Range Separated -
      --p2p-ip-type string                    IP Type numbered/unnumbered
      --anycast-mac-address string            IPV4 ANY CAST MAC address.mac address HHHH.HHHH.HHHH
      --ipv6-anycast-mac-address string       IPV6 ANY CAST MAC address.mac address HHHH.HHHH.HHHH
      --arp-aging-timeout string              Determines how long an ARP entry stays in cache <NUMBER: 60-100000
      --mac-aging-timeout string              MAC Aging Timeout <NUMBER: 0|60-100000>
      --mac-aging-conversation-timeout string MAC Conversational Aging time in seconds<NUMBER: 0|60-100000>
      --mac-move-limit string                 MAC move detect limit <NUMBER: 5-500>
      --duplicate-mac-timer string            Duplicate Mac Timer
      --duplicate-mac-timer-max-count string  Duplicate Mac Timer Max Count
      --configure-overlay-gateway string      ConfigureOverlayGateway Enabled Yes/No
      --bfd-tx string                         BFD desired min transmit interval in milliseconds <NUMBER: 50-3000
      --bfd-rx string                         BFD desired min receive interval in milliseconds <NUMBER: 50-30000
      --bfd-multiplier string                 BFD detection time multiplier <NUMBER: 3-50>
      --bgp-multihop string                   Allow EBGP neighbors not on directly connected networks <Number:1-
      --max-paths string                      Forward packets over multiple paths<Number:1-64>
      --allow-as-in string                    Disables the AS_PATH check of the routes learned from the AS<Numbe
      --mtu string                            The MTU size in bytes <Number:1548-9216>
      --ip-mtu string                         For SLX IPV4/IPV6 MTU size in bytes <Number:1300-9194>
      --leaf-peer-group string                Leaf Peer Group Name <WORD: 1-63>
      --spine-peer-group string               Spine Peer Group Name <WORD: 1-63>
      --mctlink-ip-range string               Range Of IP Address
      --mct-port-channel string               Portchannel interface number <NUMBER: 1-1024>
      --control-vlan string                   vlan number <NUMBER: 1-4090>
      --control-ve string                     vlan number <NUMBER: 1-4090>
      --vni-auto-map string                   VNI Auto Map <STRING Yes/No>
  -h, --help                                  help for update
```

Fabric settings can be updated using “efa fabric setting update” command

Fabric Setting Update Example

```
root@TPVM:~# efa fabric setting update --p2p-link-range 30.30.30.0/24 --loopback-port-number 100
default Fabric Update Successful
FabricId: 1
--- Time Elapsed: 139.949425ms ---
root@TPVM:~# efa fabric setting show
+--------------------------+-----------------+
|          NAME            |      VALUE      |
+--------------------------+-----------------+
| Fabric Name              | default         |
| Link IP Range            | 30.30.30.0/24   |
| Loopback IP Range        | 172.32.254.0/24 |
| Loopback Port Number     | 100             |
| VTEP Loopback Port Number | 2               |
| Spine ASN Block          | 64512           |
| Leaf ASN Block           | 65000-65534     |
| P2P IP Type              | numbered        |
+--------------------------+-----------------+
--- Time Elapsed: 52.249752ms ---
```

Updating multiple fabric setting using a single command

Почему Extreme?

Scale And Performance
- Ethernet и IP фабрика
- BGP-EVPN Network Virtualization
- VxLAN Single Pass RIOT (!)
- BGP-EVPN L2 and L3 Multi-tenancy (VNI)
- VMware NSX L2 and L3 Fabric Integration
- VMware vCenter Port Profiles

ZTP and Programmability

YANG model based REST & RESTCONF, NETCONF APIs

On box scripting – Python scripts

vSLX

Insight Architecture

Extreme Validated Designs

Полный спектр решений для инфраструктуры ЦОД

Коммутаторы и маршрутизаторы

Фиксированные платформы


SLX 9030


SLX 9140


SLX 9240


SLX 9540


SLX 9640


VDX 6740


VDX 6740-T


VDX 6940


VDX 6940-144S

Модульные платформы

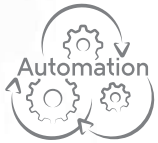

SLX 9850


VDX 8770


MLXe Routers

Программное обеспечение

Автоматизация



Workflow Composer
Powered by StackStorm

Workflow Composer
Automation Suites

- Network Essentials
- Data Center Fabrics
- IXP

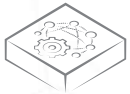
vmware NSX

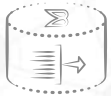
openstack

OPEN DAYLIGHT

Аналитика


Virtual Packet Broker


Session Director

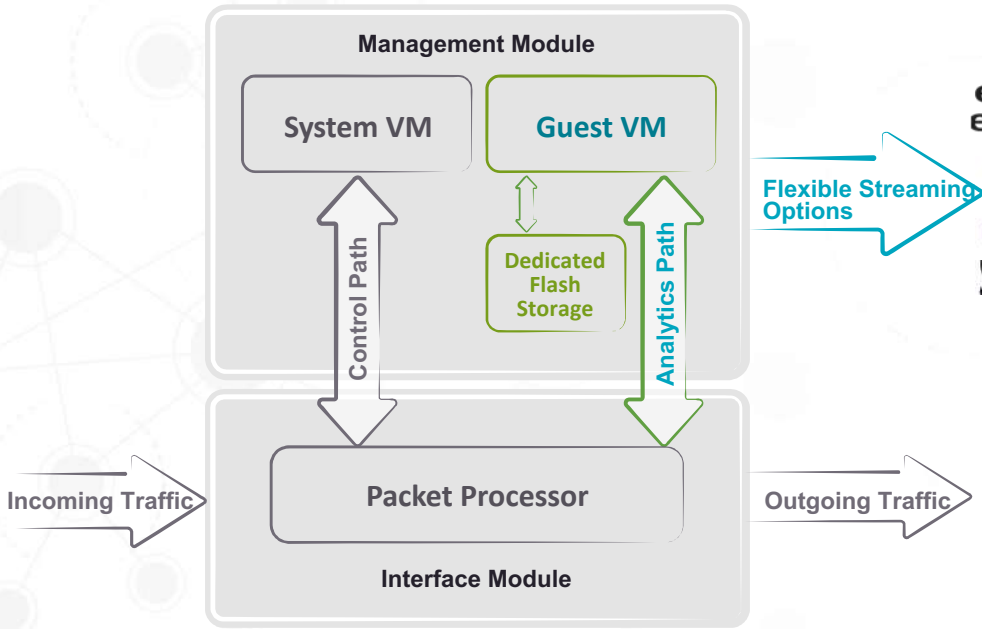

Flow Optimizer


Virtual TAP


Visibility Manager

Extreme SLX Insight Architecture

Упрощаем операционную деятельность



Доступно на всей линейке SLX

(кроме SLX 9030)

- Основное назначение monitoring and troubleshooting приложения и клиенты для автоматизации RESTful
- TPVM НЕ предназначена для NFV, LB, VR, FW и т.д.
- Поддерживаемые приложения:
  - TCPdump, Tshark, Wireshark
  - cURL – command line RESTful access
  - Chrome browser – GUI RESTful access
  - PerfSonar
  - Puppet
- Поддержка Docker и LXC 1.0
- Выделенный CPU/Cores
- Выделенный DIMM
- Выделенный SSD

Extreme SLX Insight Architecture
Простота развертывания

Guest-VM
---

1. tpvm install
2. tpvm start

 # Enable Analytic port
insight enable -- *Inside Port-Channel)*

 # Опционально
3. tpvm disk add
4. tpvm auto-boot
5. tpvm show ipadd

Extreme SLX Insight Architecture
Простота развертывания

Guest-VM
---

1. tpvm install
2. tpvm start

 # Enable Analytic port
insight enable -- *Inside Port-Channel)*

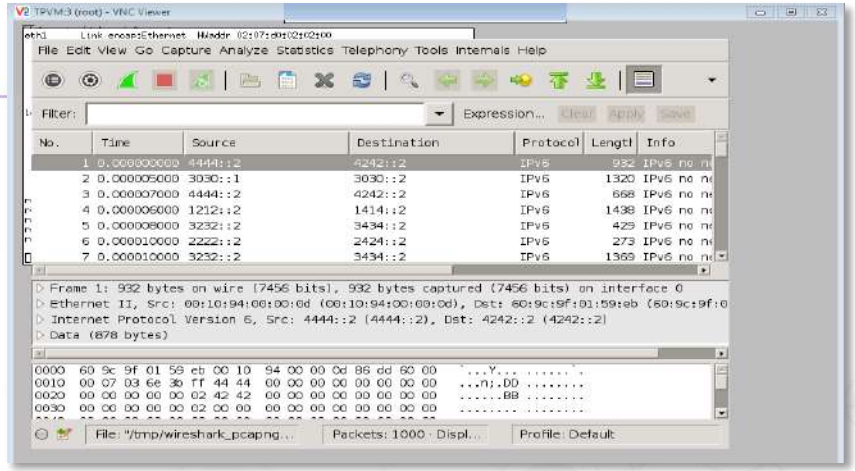
 # Опционально
3. tpvm disk add
4. tpvm auto-boot
5. tpvm show ipadd

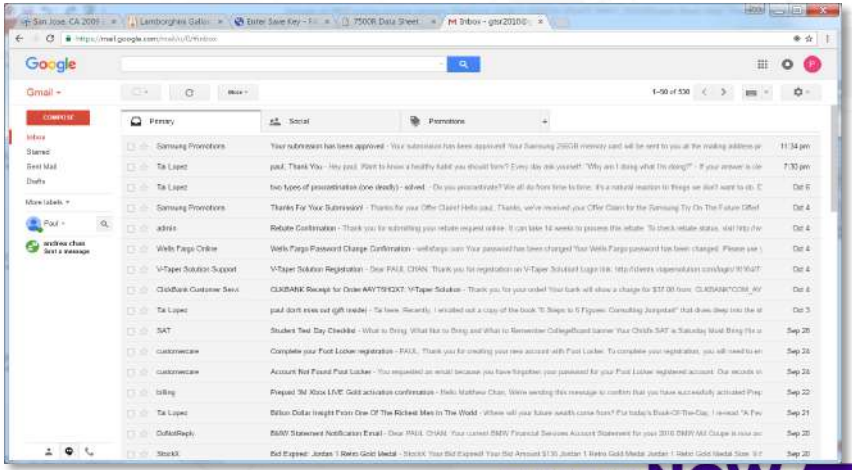
Wireshark on Guest-VM
---

1. sudo apt-get install wireshark
2. apt-get install vnc4server
3. wireshark (-ni eth1 -c 1000 ñw /mnt/disk2/PCAP4)

Example: On-Box Troubleshooting

1. I'm having a network issue!
2. Let's SPAN traffic to the Wireshark App running on the Guest VM giving us direct visibility.
3. Wireshark allows NetOps to analyze captured data directly on the box.
4. Captured PCAP files are stored on the Guest VM.
5. PCAP files can easily be emailed out of the Guest VM using the Chrome Browser via any email client of choice.





Eliminates costly network probes/sniffers & shortens time to resolution

Application Possibilities for Guest VM

Networking tools on every router to inform IT decisions

Monitor

- Packet sniffers, analyzers
- Health of the device, CPU, memory
- SLA monitoring for network parameters

TCPDUMP

WIRESHARK

Analyze Data

- Traffic patterns on links for load balancing and capacity planning
- Analysis of data flows in real time and forensics
- Faster DDoS attack detection, debugging

splunk>

DATADOG

New Relic.

Network Validation

- Test latency from different points in the network
- Packet generation and playback captured packets
- Measure end to end performance

perfSONAR

SPIRENT

Troubleshoot

- TAC can use customer tools, and apps for troubleshooting
- Report congestion, latency, and performance issues
- Faster time-to-recovery

Remedy

Extend on-box Capabilities

- Wrapper functions using native API for simplified integration
- Scripting agents for provisioning, monitoring and reporting
- LXC support for application hosting agility and flexibility

docker

ANSIBLE

puppet labs

Обновление серии SLX

Leaf

Spine

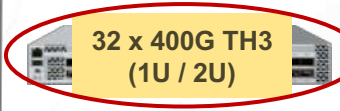
Super Spine

Border Router

Border Leaf

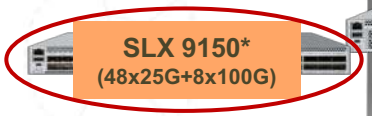
Extreme SLX 9850
4 Slot

Extreme SLX 9850
8 Slot



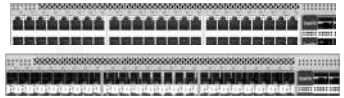
Extreme SLX 9640

Extreme SLX 9540



Extreme SLX 9240

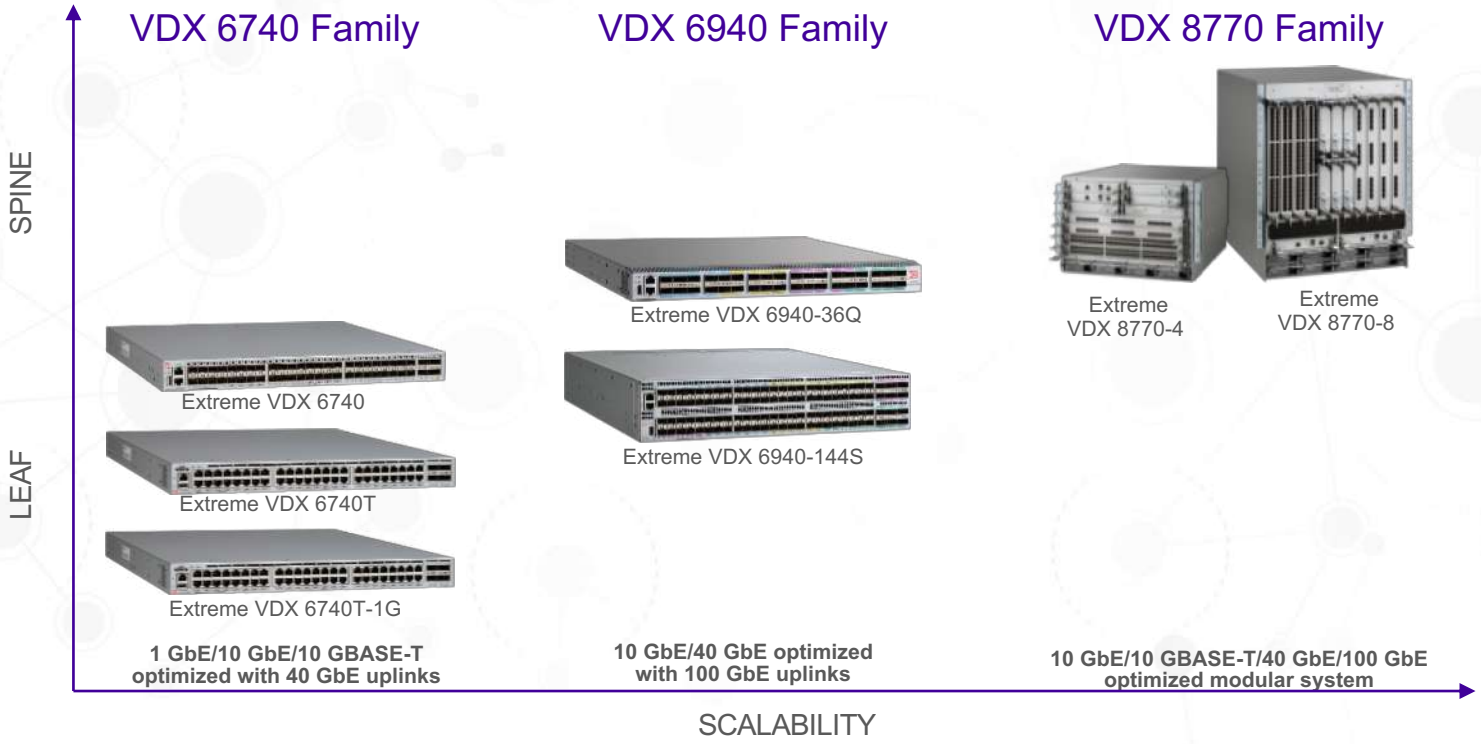
Extreme SLX 9140



Extreme SLX 9030

EXTREME NOW

VDX Switch Portfolio



СПАСИБО!

SLX PoD / CoD & Software Licensing

Extreme SLX 9540

Base SKU	Optional PoD/CoD
24 ports 1/10G enabled 24 ports 1G enabled 6 ports 100/40G not enabled	24 ports 10G upgrade
	2 ports 100/40G
48 ports 1/10G enabled 6 ports 100/40G enabled	None

Advanced SW license
- BGP EVPN
- Guest VM
- gRPC Streaming
- Precision Time Protocol (1588 BC)
- Packet Timestamping
- OpenFlow
- MPLS
- OptiScale Routing

Extreme SLX 9850

-M Module Base SKUs	Optional PoD
72 port 1/10G	None
36 port 100G / 60 port 40G	None
12 port 100G / 60 port 40G	6 ports 40/100G

Advanced SW license
- No SW licenses required

Extreme SLX 9640

Base SKU	Optional PoD
24 ports 1/10G enabled 4 ports 100/40G enabled 8 ports 100/40G not enabled	4 ports 100/40G
24 ports 1/10G enabled 12 ports 100/40G enabled	None

Advanced SW license
- BGP EVPN
- Guest VM
- gRPC Streaming
- Precision Time Protocol (1588 BC)
- Packet Timestamping
- OpenFlow
- MPLS
- OptiScale Routing (future)

SLX PoD / CoD & Software Licensing

Extreme SLX 9240

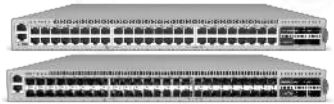
Base SKU	Optional PoD/CoD
32 ports 100/40G enabled	None

Advanced SW license
- BGP EVPN
- Guest VM
- gRPC Streaming
- Precision Time Protocol (1588 BC)
- Packet Timestamping

Extreme SLX 9140

Base SKU	Optional PoD/CoD
48 ports 1/10/25G enabled 6 ports 40/100G enabled	None

Advanced SW license
- BGP EVPN
- Guest VM
- gRPC Streaming
- Precision Time Protocol (1588 BC)
- Packet Timestamping



Extreme SLX 9030

Base SKU	Optional PoD/CoD
48 ports 1/10G TX enabled 4 ports 40/100G enabled	None
48 ports 1/10G enabled 4 ports 40/100G enabled	None

Advanced SW license
- BGP EVPN
- gRPC Streaming