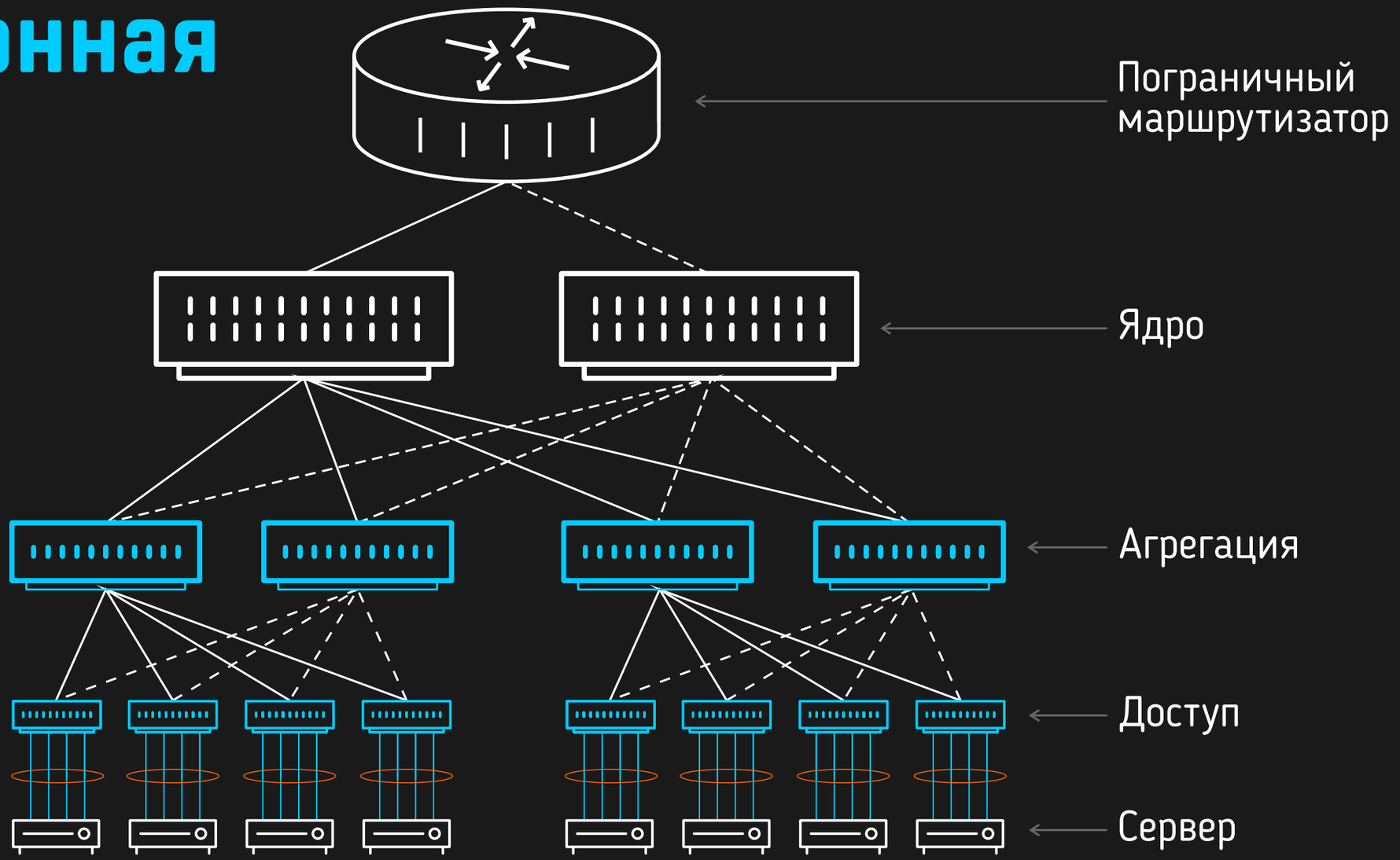
Эволюция сетей центров обработки данных.

Оверлейные сети и SDN

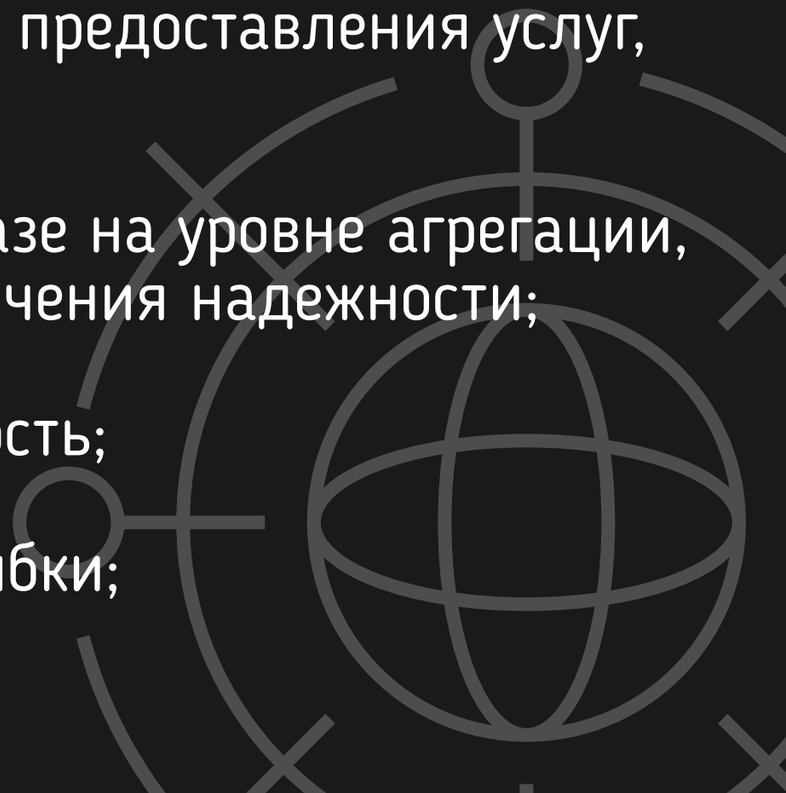
Рассказывает
Станислав Титов

Традиционная сетевая модель ЦОДа



Традиционная сетевая модель ЦОДа и её недостатки

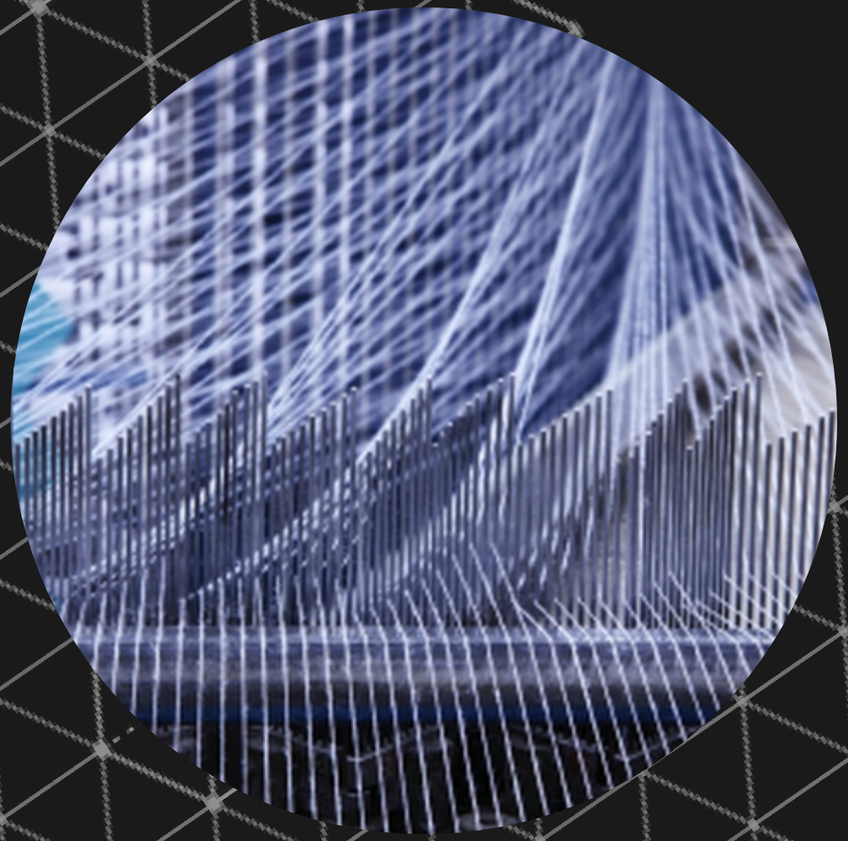
- Отсутствие гибкости взаимодействия, скорость предоставления услуг, внесение изменений;
- Резкое снижение производительности при отказе на уровне агрегации, и возрастание сложности/стоимости для увеличения надежности;
- Недостаточная стабильность и масштабируемость;
- VLAN'ы конфигурирование и ограничения, ошибки;



Традиционная сетевая модель ЦОДа и её недостатки

- MAC лимиты и BUM траффик;
- ARP устаревание -> неактуальность;
- Перегруженность аплинка горизонтальным трафиком;
- Неэффективное использование ресурсов для избыточных соединений STP. Не панацея от возникновения петель;
- Соединение нескольких ЦОДов.

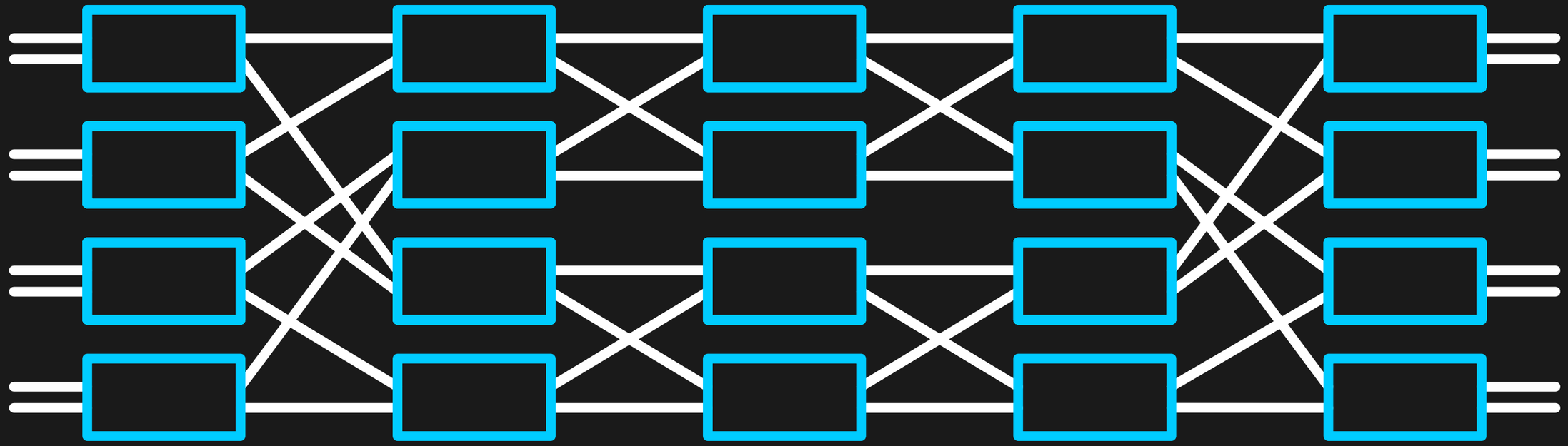
Фабрика



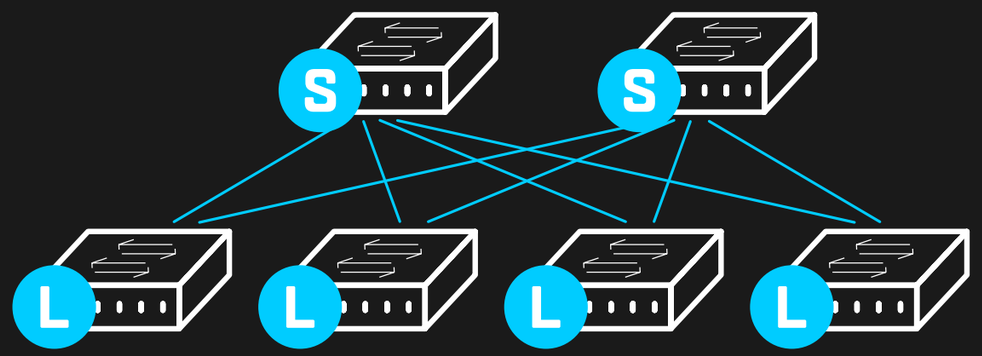
калька с используемого в англоязычной практике слова fabric, основное значение которого – «ткань».

Некая матрица которая может быть представлена как решетка из вертикальных и горизонтальных линий, соответствующих её входным и выходным каналам.

Топология для фабрики



Трёхступенчатая сеть Клоза

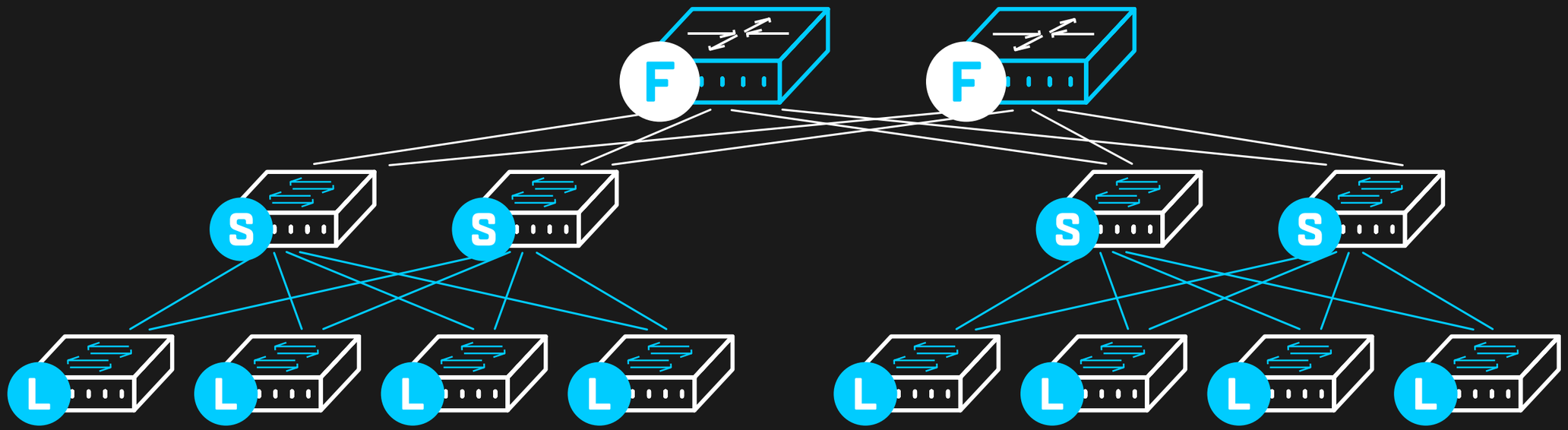


L = лист S = хребет

Мы можем использовать те же принципы сети Clos и применять их для создания фабрик.

Многие современные сети уже разработаны таким образом и часто называются сетями Spine/Leaf (хребта и листьев). Листовые устройства обеспечивают подключение к доступу, в то время как устройства позвоночника действуют как точки агрегации. Обычные многоступенчатые модели ткани IP в ключают в себя 3-ступенчатые Clos (три уровня: этапы 1 и 3 (лист), ступень 2 (позвоночник)

Пятиступенчатая сеть Клоза

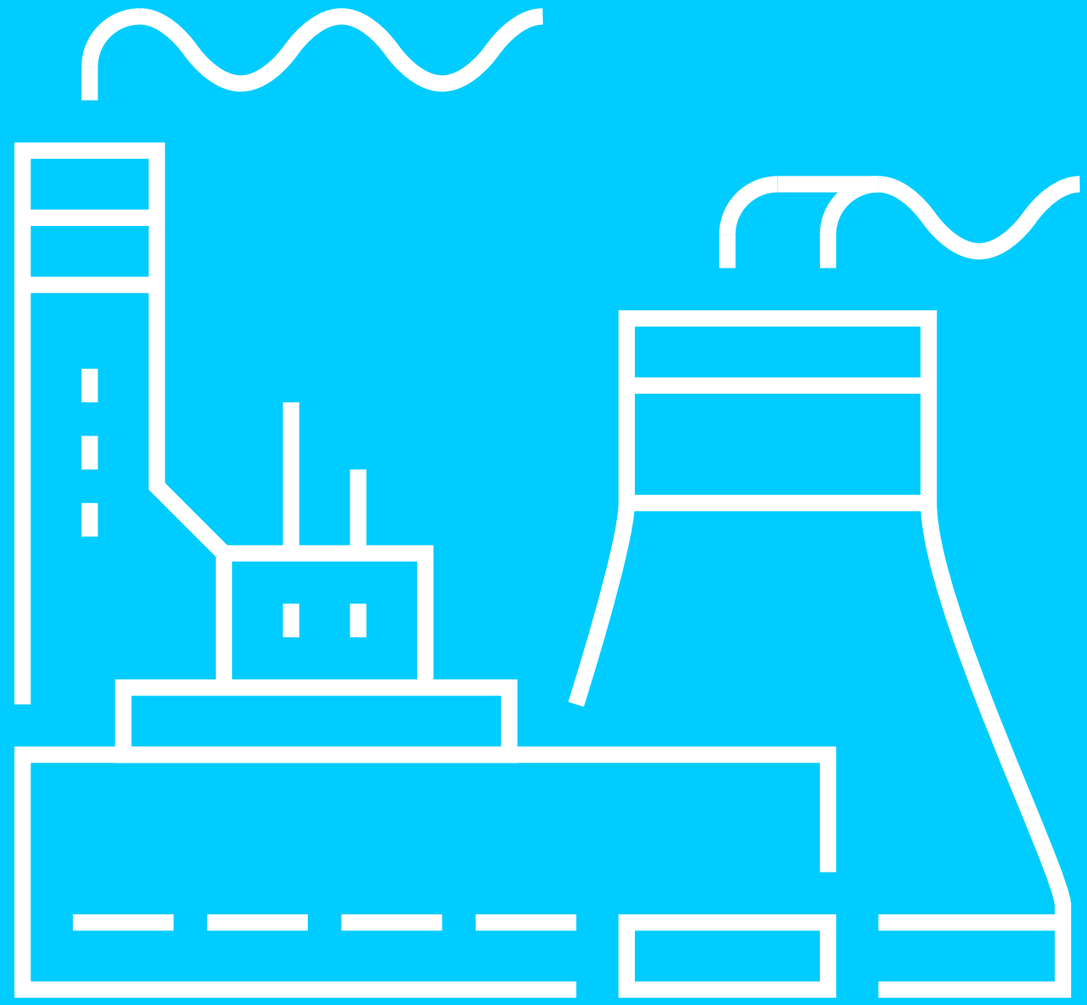


L = лист **S = хребет** **F = суперхребет/фабрика**

5-ступенчатые Clos (три уровня: этапы 1 и 5 (лист), этап 2 и 4 (позвоночник), и этап 3 (ткань / супер позвоночник для обеспечения связи между хребтами)

Виды фабрик

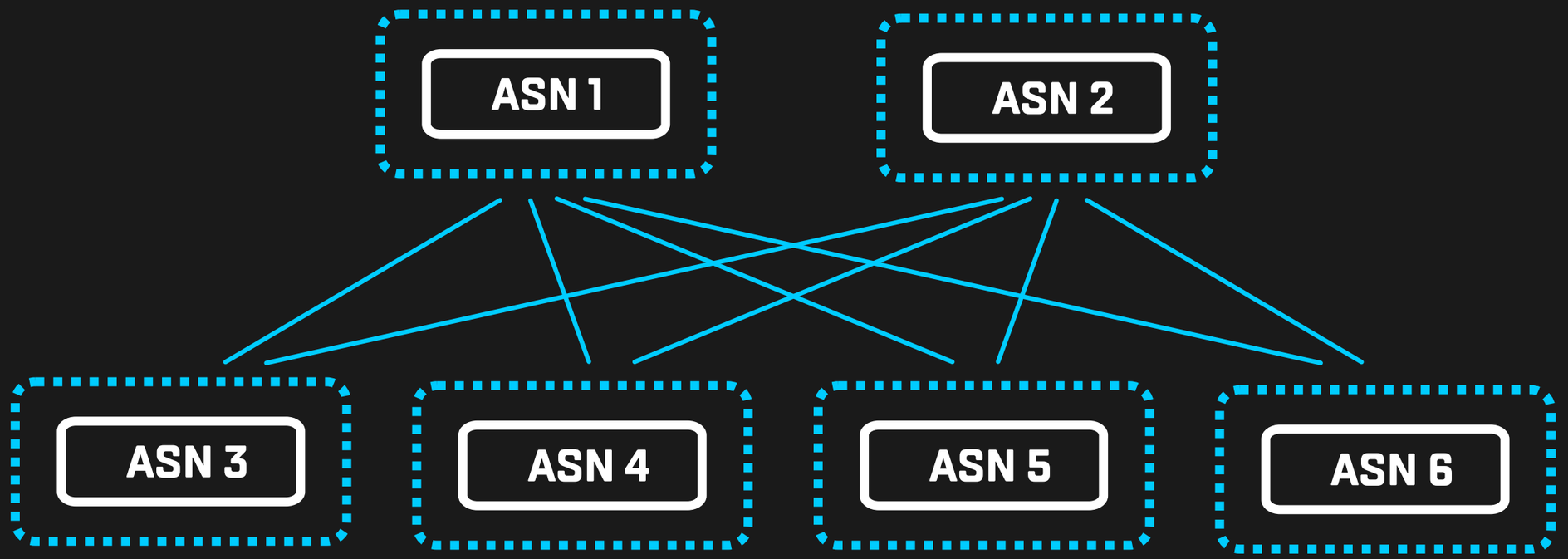
- MPLS
- Ethernet
- IP



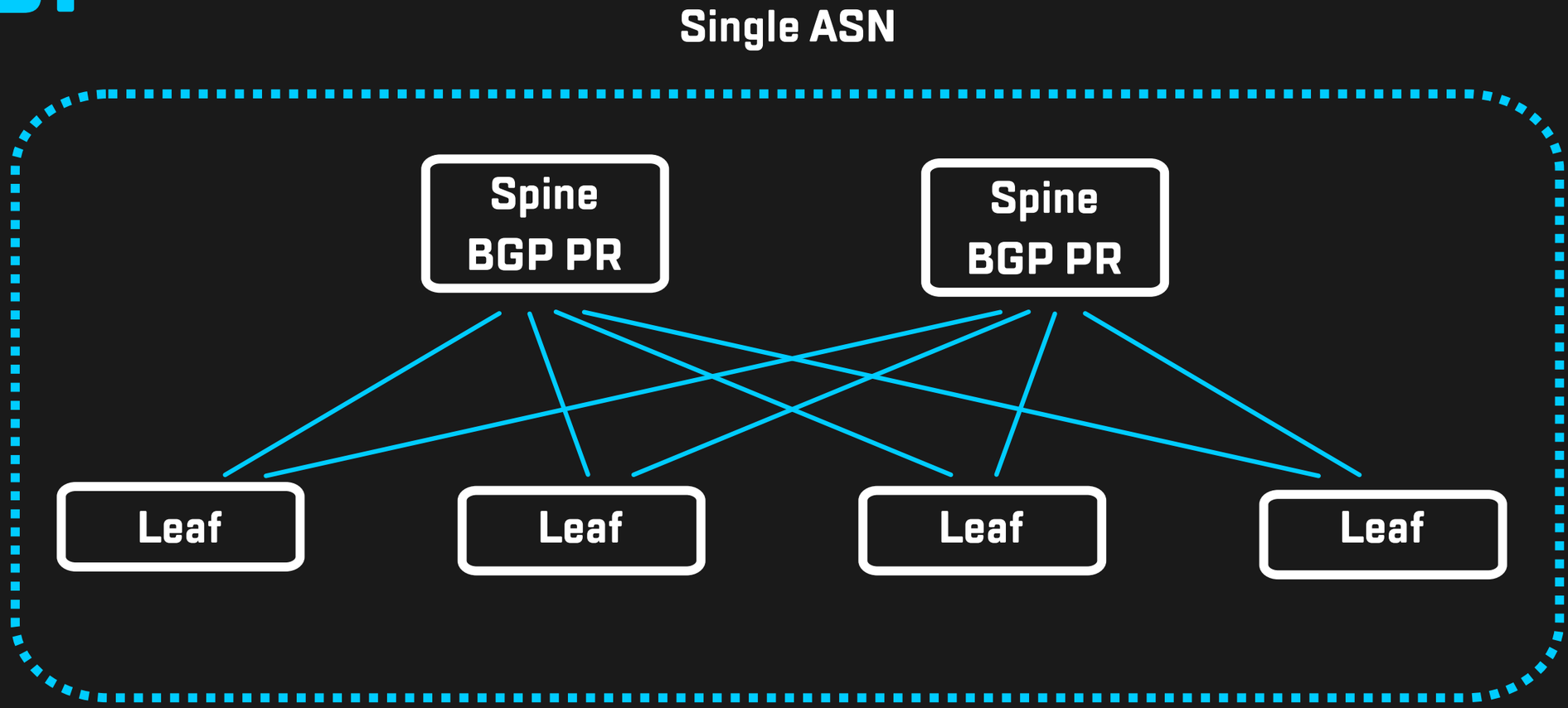
Виды протоколов плоскости управления для управления IP-фабрики

	OSPF	IS-IS	BGP
Advertise Prefixes	YES	YES	YES
Scale	Limited	Limited	Extensive
Traffic Engineering	Limited	Limited	Extensive
Traffic Tagging	Limited	Limited	Extensive
Multivendor Stability	YES	YES	Extensive

eBGP



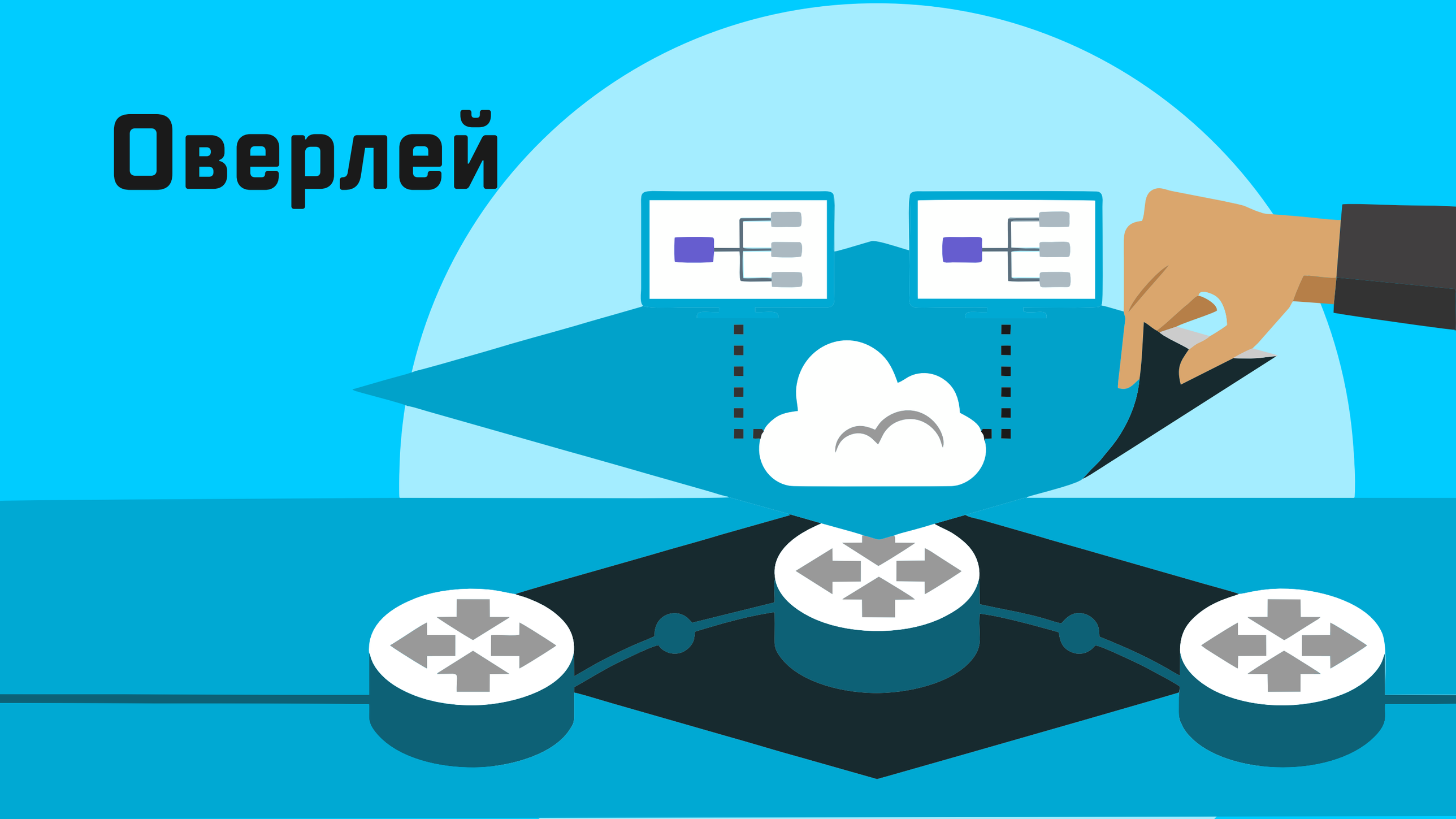
iBGP

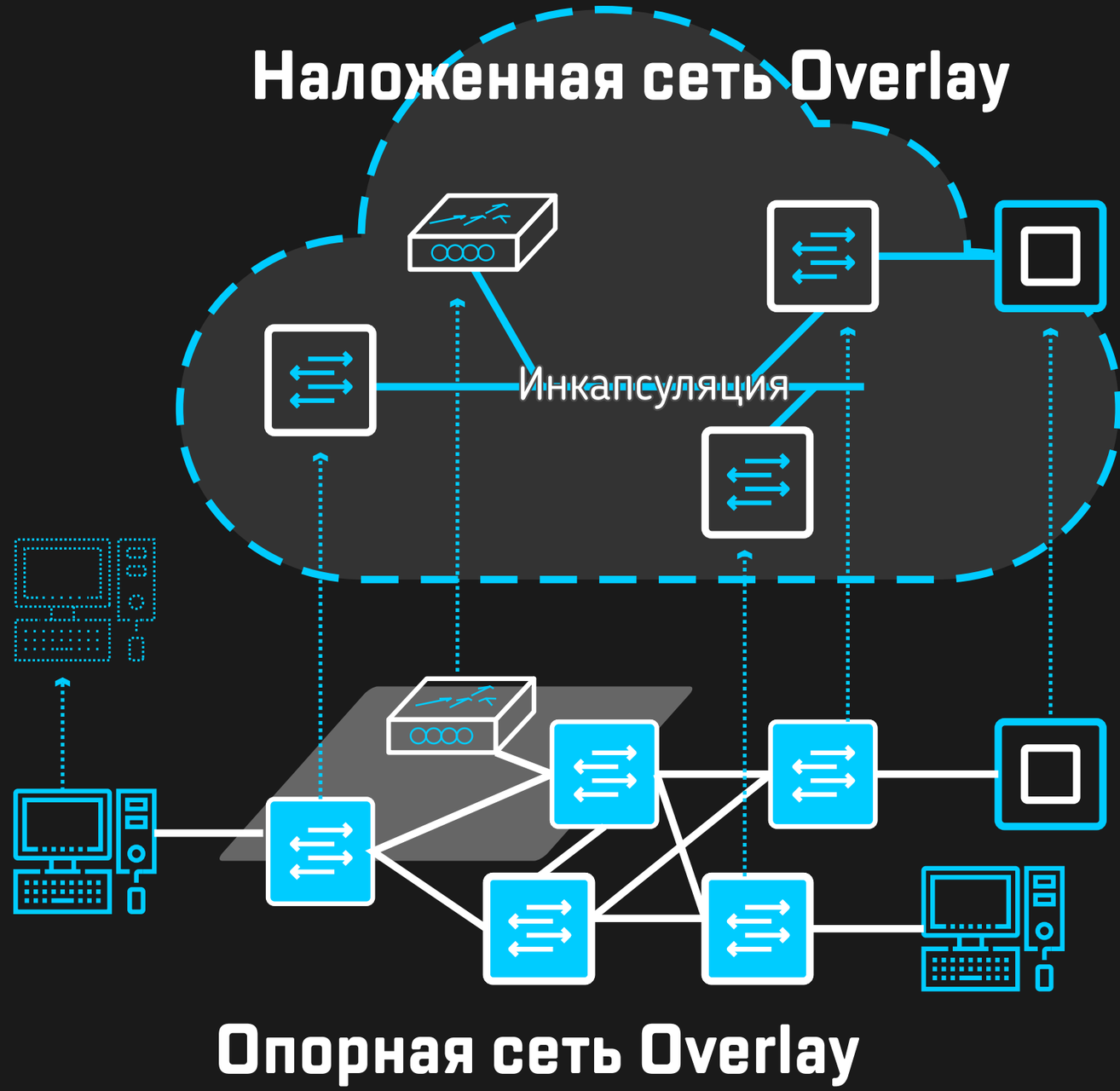


Требования к внедрению:

- Базовый IP prefix
- Point-to-point network mask
- Loopback addressing
- BGP autonomous system numbers
- BGP export policy
- BGP import policy
- Equal cost multipath routing
- BFD

Оверлей




Наложенная сеть Overlay
Инкапсуляция
Опорная сеть Overlay

Underlay

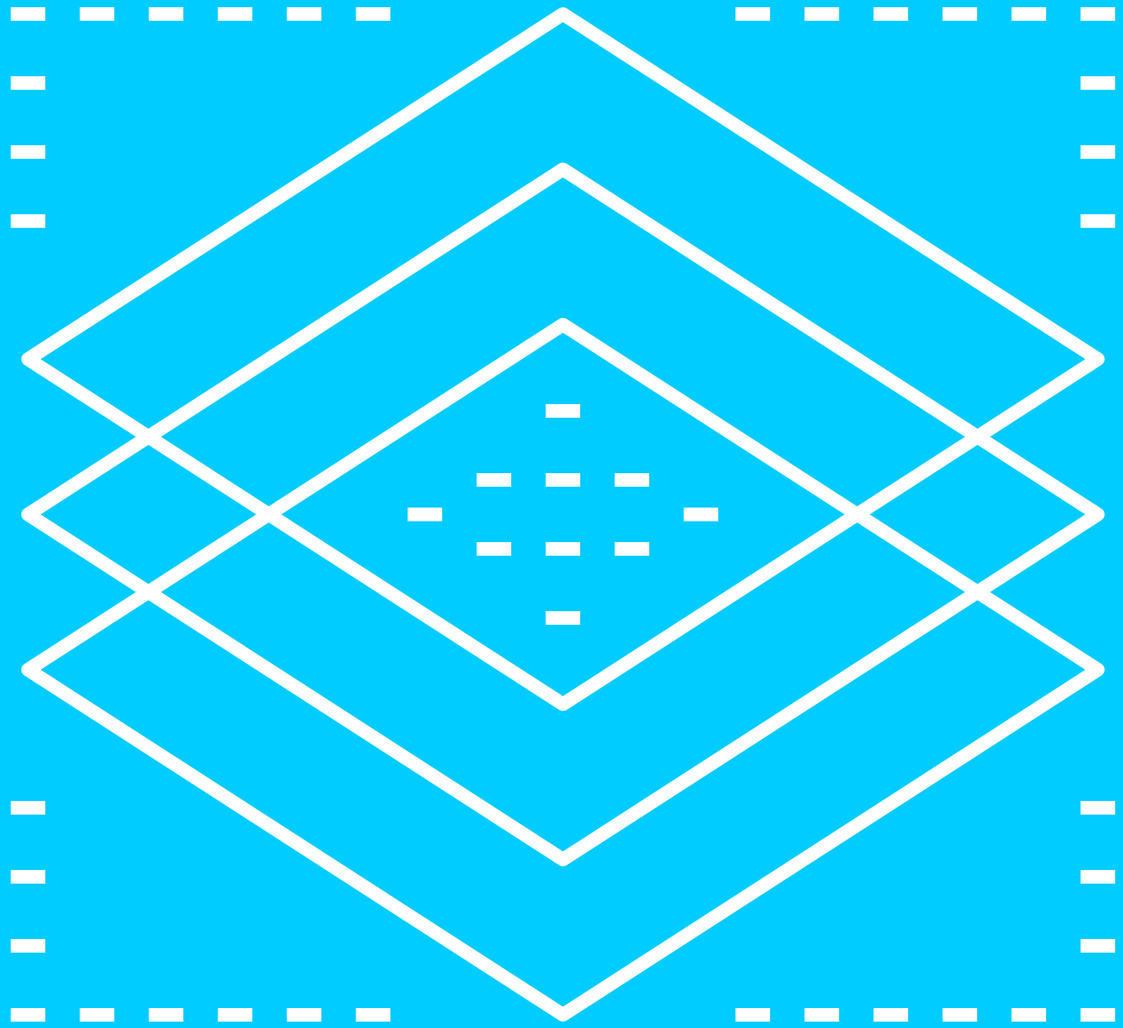
- Эффективный транспорт;
- Резервирование и производительность;
- Эффективное управление трафиком;
- Управление, диагностика, программируемость.

Overlay

- Гибкая наложенная сеть;
- Мобильность подключенных адресов;
- Изолированные частные сети;
- Масштабирование – снижение сложности в ядре;
- Гибкость, программируемость.

Типы Оверлеев:

- MPLSoMPLS
- MPLSoGRE
- MPLSoUDP
- VXLAN/EVPN



Типы наложенных сетей: вид сервиса

Layer 2

- Эмуляция сегмента LAN,
- Передача кадров Ethernet,
- Мобильность в подсети (L2 домене),
- Риск L2 фладинга,
- Имитация физической топологии

Layer 3

- Абстракция связности на основе IP,
- Передача IP пакетов,
- Мобильность адреса без растягивания L2,
- Ограничение доменов сбоя

EVPN/VXLAN

Обеспечивает сеть с сегментацией, мобильностью и масштабированием.

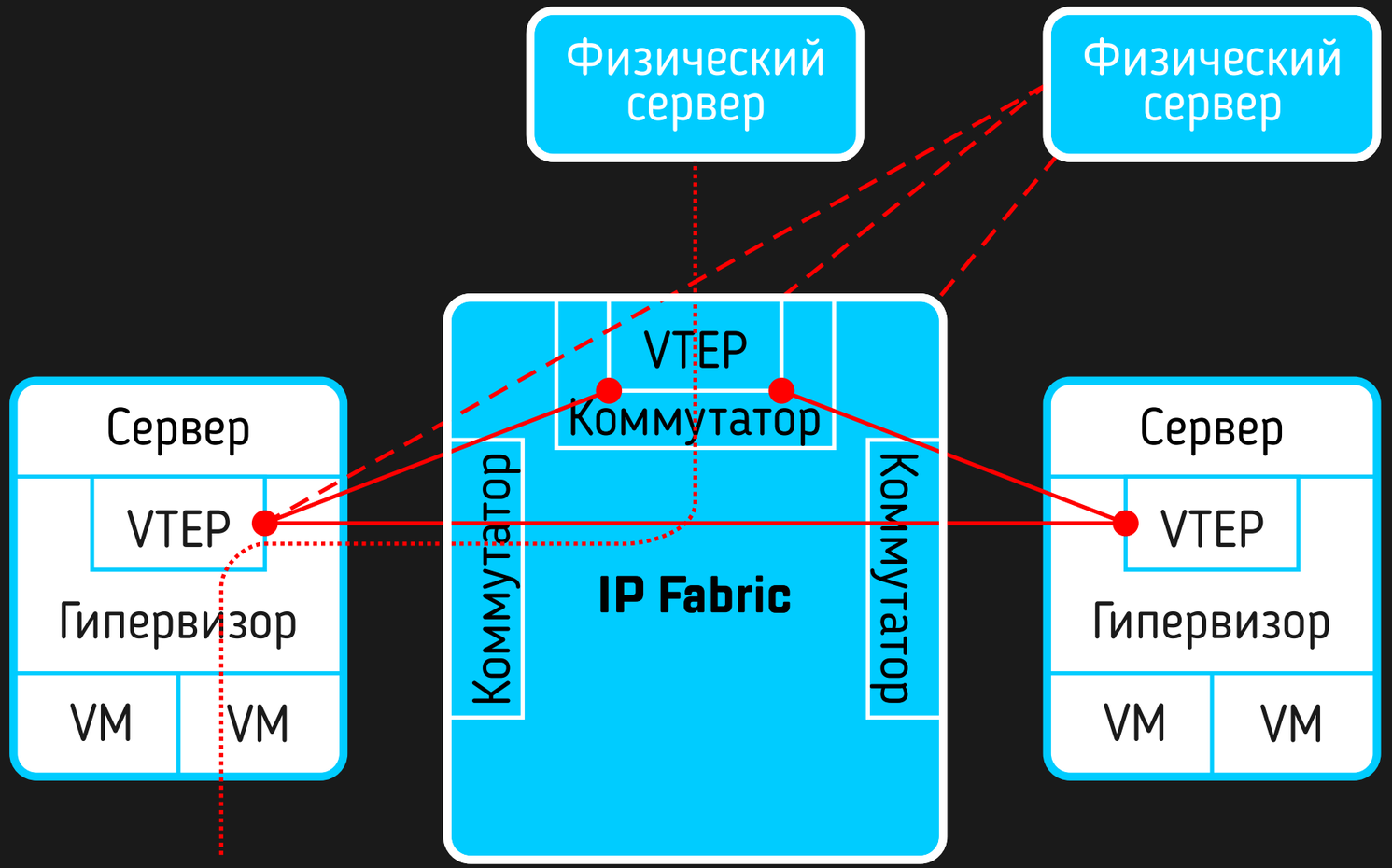
VXLAN и EVPN являются стандартами и вы можете надеяться

на согласованную работу в мультивендорной сети

EVPN/VXLAN

- Используется IP фабрика с Layer-3 ECMP,
- «Пространство имен» сегментов до 16 миллионов,
- Развёртывание сегментов без изменений на промежуточных устройствах,
- Поддержка физическими и виртуальными коммутаторами,
- Поддержка разными производителями

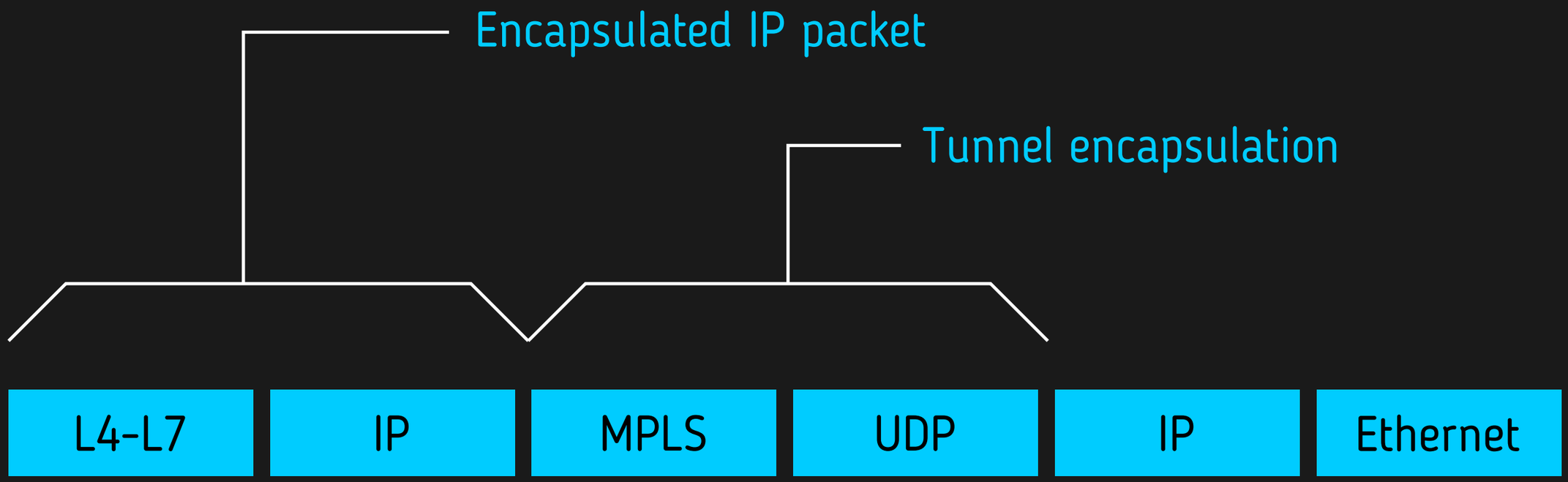
EVPN/VXLAN



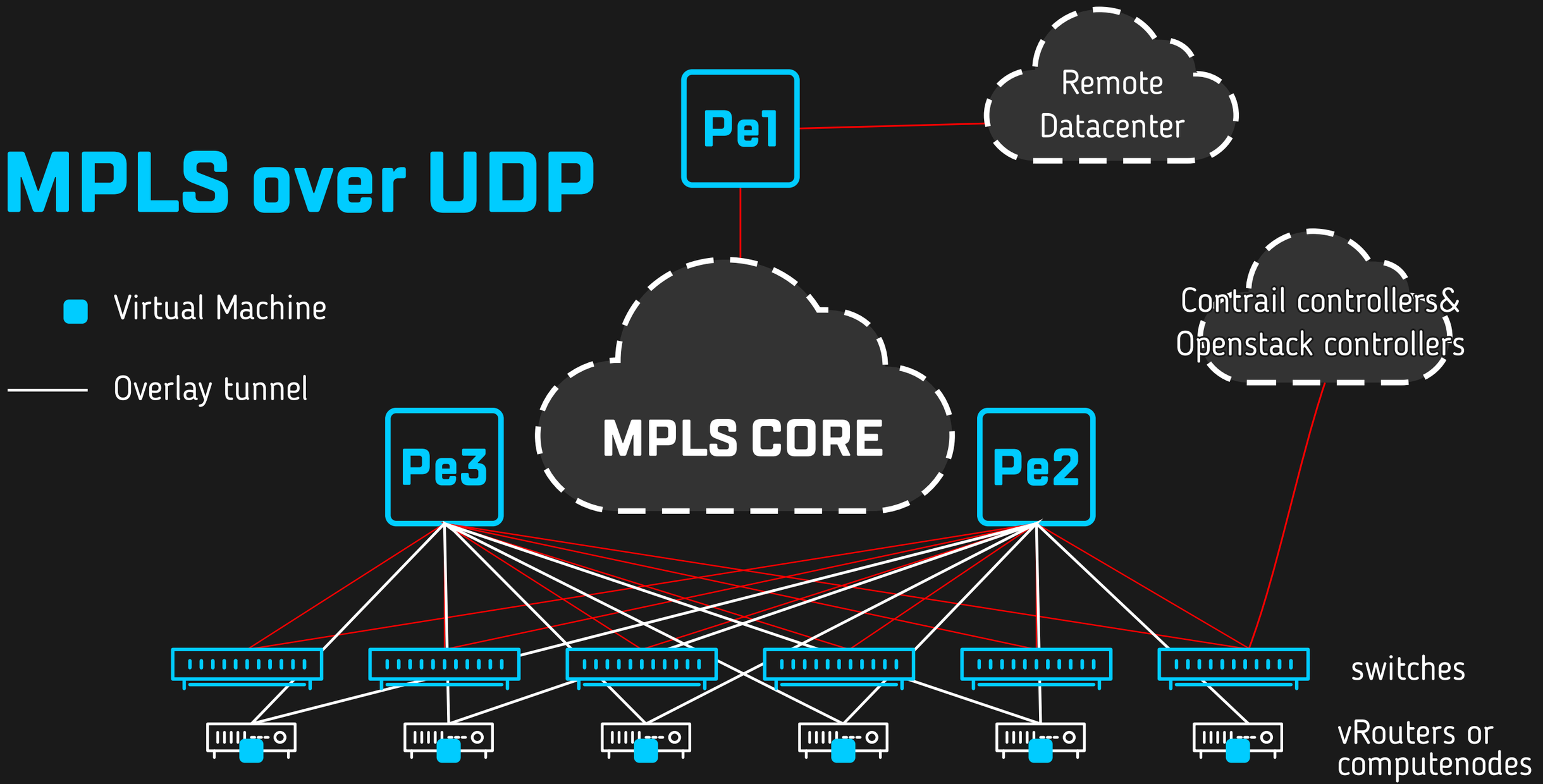
Помимо подключения виртуальных рабочих нагрузок (как показано на рисунке), шлюзы VTEP также могут обрабатывать мост VLAN на VXLAN для подключения физических и виртуальных рабочих нагрузок.

VTEP концы виртуального туннеля
DataFlow VM и физического сервера
Плоскость Управления SDN

MPLS over UDP



MPLS over UDP



```
root@ixc-mx100003# run show route receive-protocol bgp 82.202.175.131

Demo-VRF.inet.0: 4 destinations, 5 routes (4 active, 0 holddown, 0 hidden)
  Prefix  Nexthop       MED     Lclpref    AS path
* 195.140.146.3/32       82.202.175.129       100     200       ?

bgp.l3vpn.0: 1 destinations, 2 routes (1 active, 0 holddown, 0 hidden)
  Prefix  Nexthop       MED     Lclpref    AS path
  82.202.175.129:3:195.140.146.3/32
*                       82.202.175.129       100     200       ?
root@ixc-mx100003# run show dynamic-tunnels database
*- Signal Tunnels #- PFE-down
Table: inet.3
```

```
Destination-network: 82.202.175.0/24
Tunnel to: 82.202.175.129/32
  Reference count: 4
  Next-hop type: UDP
    Source address: 172.25.11.1
    Next hop: tunnel-composite, 0x71133dc, nhid 651
      VPN Label: Push 23 Reference count: 3
      Ingress Route: with 0.0.0.0/0
      Traffic Statistics: Packets 0, Bytes 0
      State: Up
```

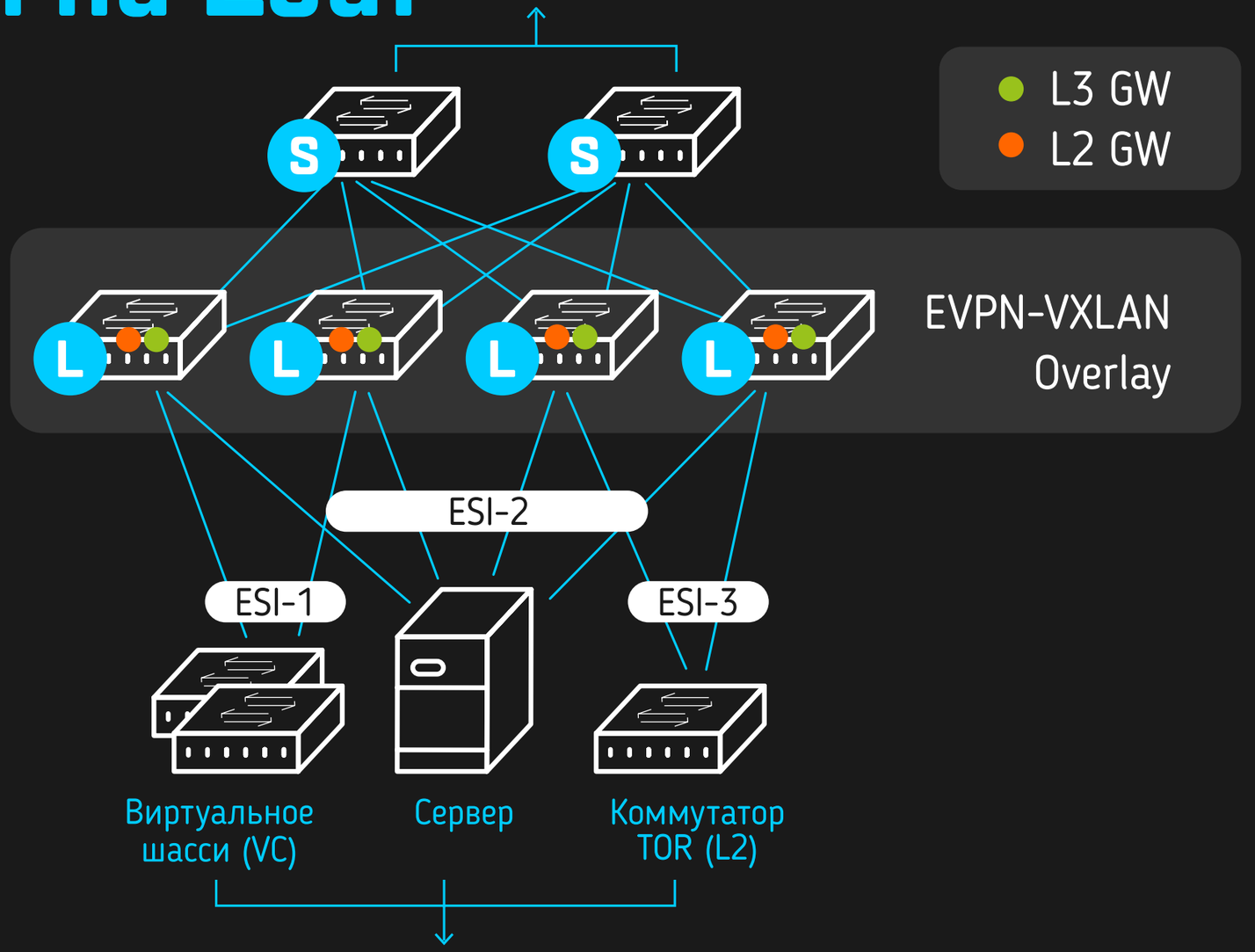
L2/L3 GW

Применительно к симметричной топологии
ЦОД маршрутизацию обычно реализуют в двух вариантах:

- **на всех Spine коммутаторах,**
- **на всех Leaf коммутаторах**

Маршрутизация на Leaf

- Маршрутизация или бриджинг на входе в оверлей
- Распределенный шлюз по умолчанию (Anycast)
- Распределяет задачу
- Лучше масштабируется



Маршрутизация на Spine

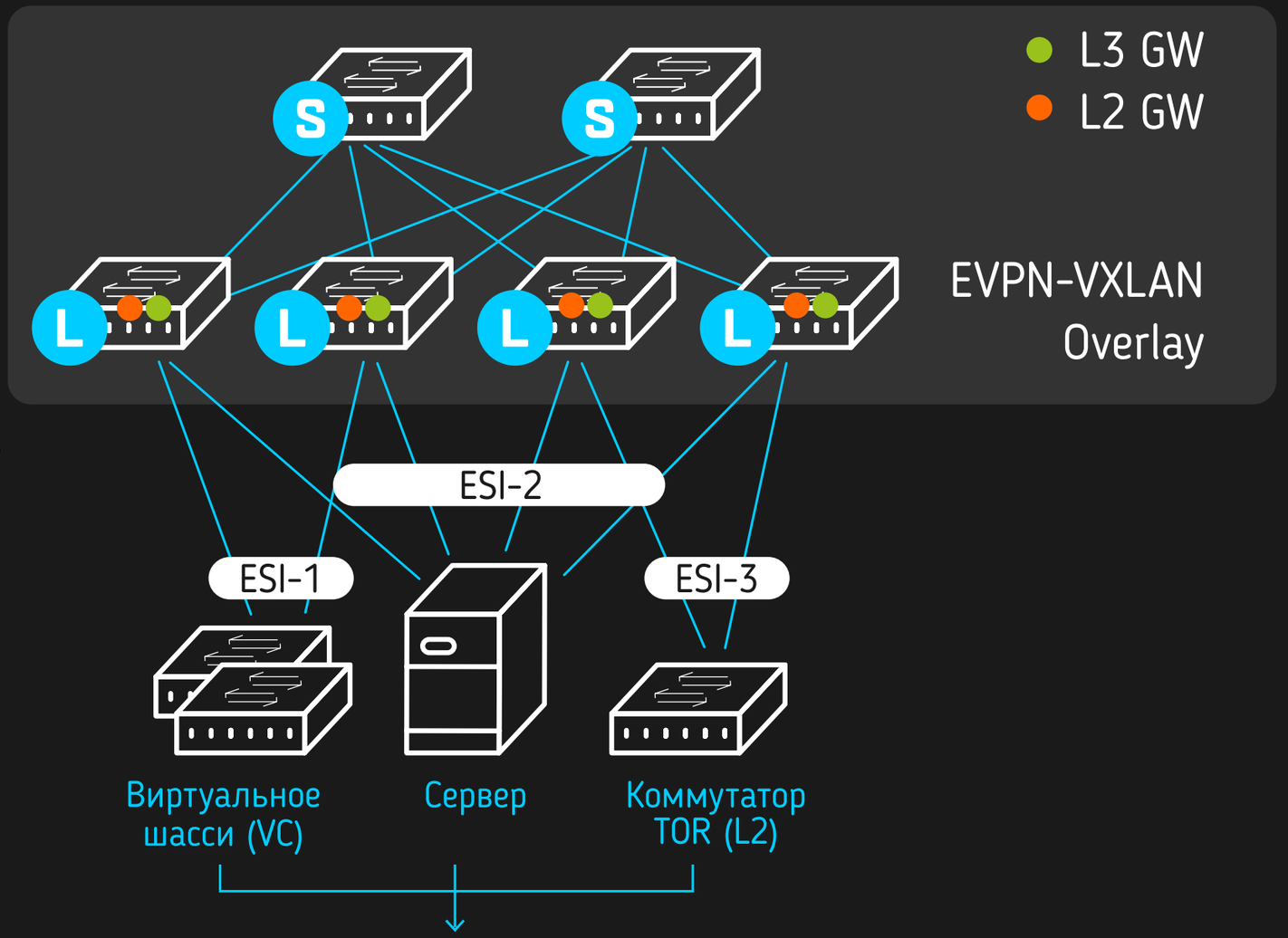
- Централизованный шлюз,

  Бриджинг, потом роутинг,

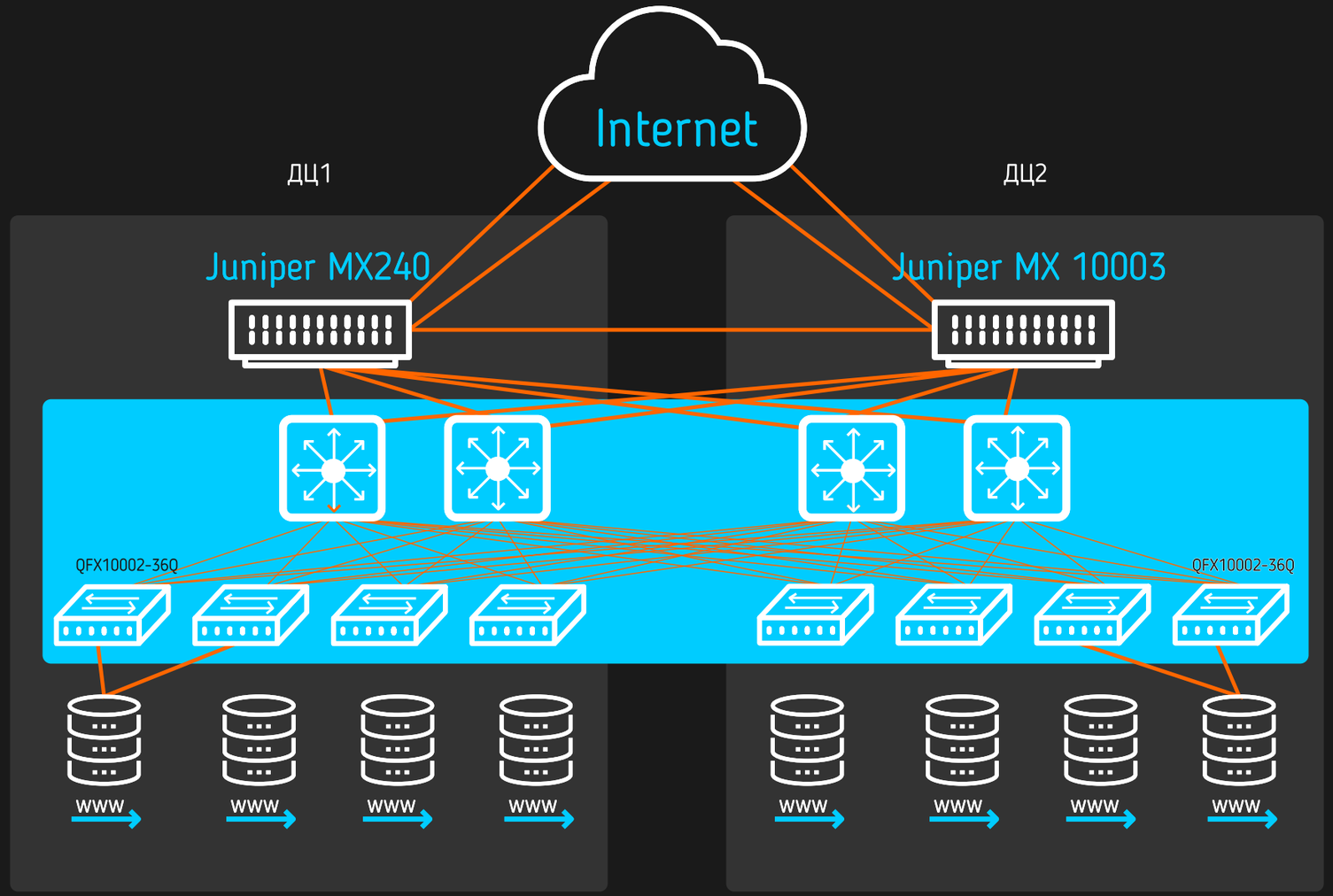
- Центральная точка маршрутизации (агрегация),

- Большие потребности в масштабируемости,

- Трудно внедрять в больших L2 доменах.



Data Center Interconnect



SDN

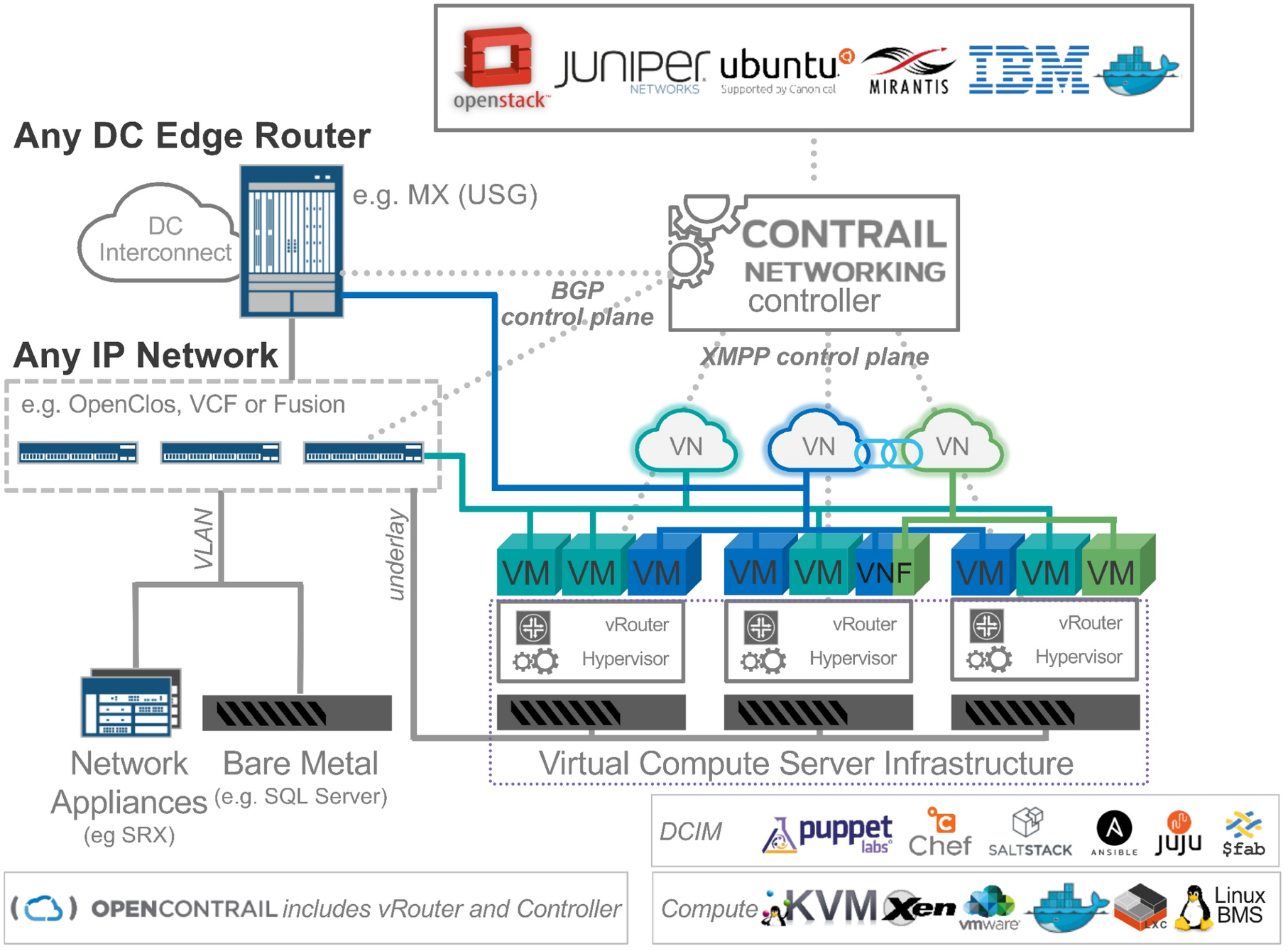
- Software Defined Networking (SDN) - это новая архитектура, которая разделяет управление устройством (network control plane) от передачи «данных (forwarding hardware), позволяя централизованно определять политику прохождения трафика в сети на контроллерах;
- Управление сетью - реализуется на основе использования стандартных интерфейсов и протоколов;
- Абстракция от железа, и, как следствие, абстракция всех политик обслуживания, возможностей, уход от проблем совместимости и масштабирования.

Contrail

Contrail – оверлейный SDN
для облачных систем оркестрации,
например OpenStack

Contrail

- Масштабируемая отказоустойчивая платформа для построения облачной среды с виртуализацией сетевых сервисов
- Поддерживает полный набор сетевых сервисов
- Решение с открытым исходным кодом
- Интеграция с сетевыми сервисами сторонних производителей
- Аналитика и визуализация наложенной и транспортной сетевой инфраструктуры


openstack
JUNIPER NETWORKS
ubuntu Supported by Canonical
MIRANTIS
IBM
Any DC Edge Router
DC Interconnect
e.g. MX (USG)
CONTRAIL NETWORKING controller
BGP control plane
XMPP control plane
Any IP Network
e.g. OpenClos, VCF or Fusion
VN
VN
VN
VLAN
underlay
VM
VNF
vRouter Hypervisor
Network Appliances (eg SRX)
Bare Metal (e.g. SQL Server)
Virtual Compute Server Infrastructure
DCIM
puppet labs
Chef
SALTSTACK
ANSIBLE
juju
$fab
OPENCONTRAIL includes vRouter and Controller
Compute
KVM
Xen
vmware
LXC
Linux BMS

Преимущества решения Contrail

- В качестве транспортной инфраструктуры может использоватьсялюбая IP сеть;
- Широкие возможности по автоматизации;
- Поддержка открытых стандартов для построения наложенной сети:
  - BGP, EVPN, OVSDB control plane
  - MPLS over GRE/UDP, VXLAN data plane overlays

SDN-контроллер состоит из четырех основных типов нод:

Конфигурационные ноды —

отвечают за предоставление REST API оркестратору и другим приложениям. «Компилирует» инструкции, приходящие «сверху», в конфигурации, применимые в конкретной сети на низком уровне

Контрольные ноды —

принимают конфигурацию от конфигурационных нодов и программируют vRouter-ы (см. далее) и физические роутеры.

SDN-контроллер состоит из четырех основных типов нод:

Аналитика —

ноды, собирающие статистику потоков, логи, и прочее.

База данных —

БД Cassandra, в которой хранится конфигурация и собранная аналитикой информация.

Контроллер Contrail

логически централизованный, но физически распределенный контроллер Software Defined Networking (SDN), отвечающий за предоставление управления, контроля и аналитики виртуальной сети.

Контроллер Contrail обеспечивает логически централизованную плоскость управления и плоскость управления системой и оркестрирует все vRouters.

Contrail vRouter

- плоскость пересылки (распределенного маршрутизатора), который работает совместно с гипервизором виртуализации. Он расширяет сеть с физических маршрутизаторов и коммутаторов в центре обработки данных в виртуальную оверлейную сеть, размещенную на серверах предоставляющую виртуализацию.

Обзор Contrail: вычислительный узел

- vRouter заменяет модуль Linux Bridge или OVS в ядре гипервизора
- vRouter выполняет бриджинг (E-VPN) и маршрутизацию (L3VPN)
- vRouter выполняет сетевые службы, такие как политики безопасности, NAT, многоадресную рассылку, зеркалирование и балансировку нагрузки
- Нет необходимости в сервисных узлах или шлюзах L2 / L3 для маршрутизации, широковещательной/многоадресной передачи, NAT

Обзор Contrail: вычислительный узел

- Маршруты автоматически просачиваются в VRF на основе политик
- Поддержка нескольких интерфейсов на виртуальных машинах
- Поддержка нескольких интерфейсов от вычислительного узла до IP Fabric

Обзор Contrail: контрольная плоскость

- Все узлы контрольных плоскостей активны
- Каждый vRouter использует XMPP для подключения с несколькими узлами контрольной плоскости для избыточности
- Каждый узел контрольной плоскости подключается к нескольким узламконфигурации для избыточности
- BGP используется для подключения к физическому маршрутизатору или коммутаторы шлюза
- Узлы управляющих плоскостей объединяются с использованием BGP

На этом и закончим