



BGP EVPN на Juniper QFX:

більше, ніж VXLAN — MAC-VRF, IP Prefix та DCI

Роман Ткаченко
Technical Lead Network Engineer

ITBIZ telecom

- 
- **Вступ** - Чому класичний Ethernet перестав масштабуватися?
 - **Розділ 1** - EVPN— більше, ніж просто VXLAN
 - **Розділ 2** - Архітектура EVPN у Junos
 - **Розділ 3** – Механізми захисту EVPN у Junos
 - **Розділ 4** - EVPN DCI та мережі Service Provider

Чому класичний Ethernet перестав масштабуватися?

Що дали VLAN

- Ізоляція L2-доменів
- Відкритий стандарт IEEE 802.1Q
- Просте сегментування

Обмеження

- 4094 VLAN
- Прив'язка до фізичної топології
- Складний Stretch L2
- Broadcast-домени ростуть разом із мережею

Рішення

- Використання QinQ: подвійне тегування для Metro Ethernet

Але залишились проблеми:

- Складний Stretch L2
- Broadcast

VLAN чудово працювали в корпоративних мережах, але не масштабувалися для сучасних дата-центрів та мереж провайдерів.

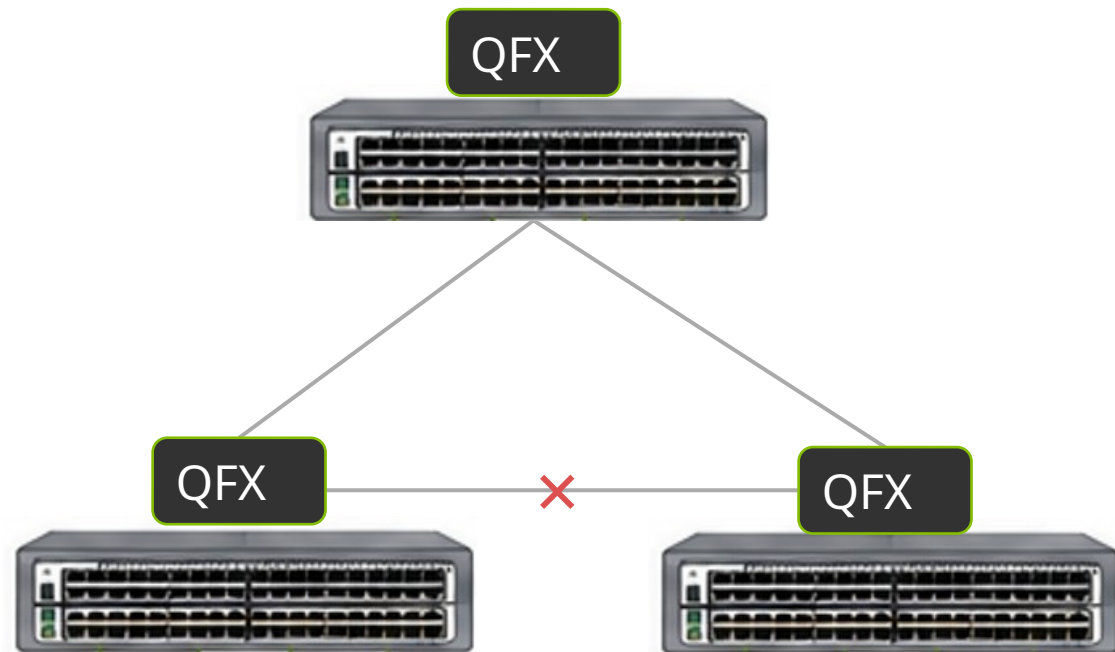
Чому STP став проблемою?

Переваги STP

- Запобігати L2-петлям

Наслідки STP

- Блокування резервних каналів
- Невикористана пропускна здатність
- Повільна конвергенція
- Погано масштабується у великих Fabric



Мережа має декілька шляхів, але STP дозволяє використовувати лише частину з них.

Які проблеми потрібно було вирішити?

Broadcast
Unknown
Unicast

MAC
Learning

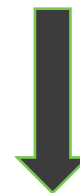
STP

4094 VLAN

Незалежність
логічної
мережі від
фізичної



ERPS



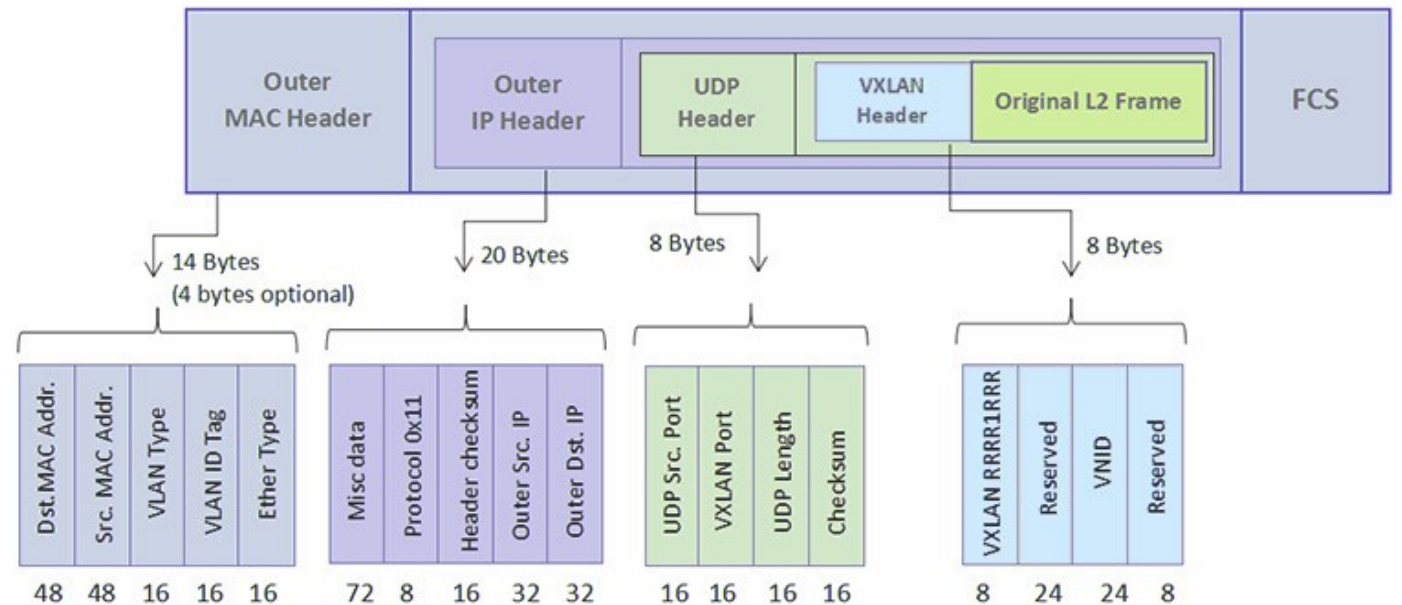
QinQ
(S-Tag + C-
Tag)

Потрібна була нова архітектура масштабування
Layer 2.

Навіщо з'явився VXLAN?

Що вирішує VXLAN

- Overlay поверх IP
- Інкапсуляція MAC-in-UDP
- 24-бітний VNI (~16 млн сегментів)
- Layer2 через Layer3
- Multi-Tenant



VXLAN не замінює VLAN — він забезпечує масштабування Ethernet поверх IP.

Чи достатньо лише VXLAN?

✓ Overlay

- ✓ 24-bit VNI
- ✓ Layer3 Transport
- ✓ Масштабування

✗ Невирішені питання

- MAC Learning
- BUM Traffic
- Flooding
- Control Plane

Наступний крок

BGP EVPN

Control Plane для VXLAN

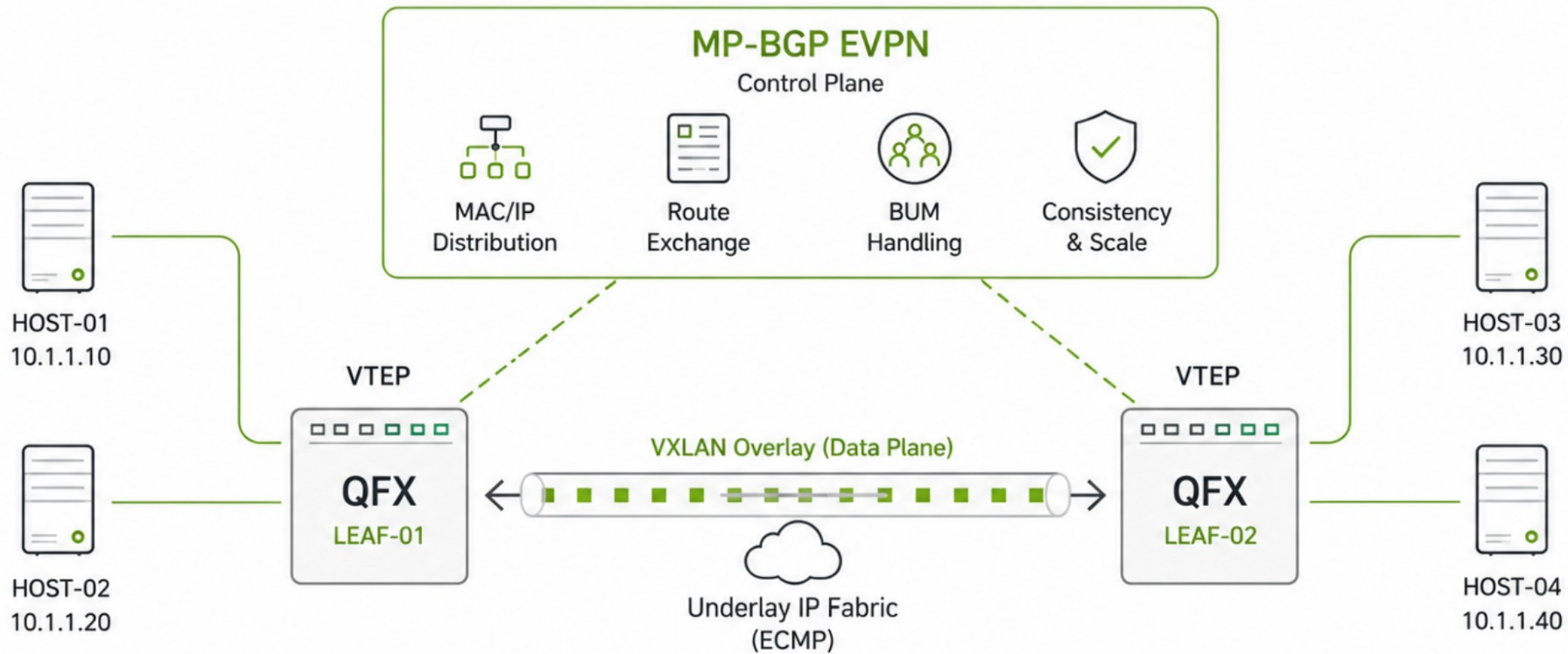
Саме EVPN зробив VXLAN повноцінною технологією побудови сучасних мереж



Розділ 1

EVPN— більше, ніж просто VXLAN

EVPN — керуюча площа сучасної Fabric



- EVPN забезпечує масштабовану та ефективну комунікацію між VTEP



Розповсюдження MAC/IP

Інформація про кінцеві вузли поширюється через MP-BGP EVPN, а не через flooding.



Відсутність Flood & Learn

Не потрібно flood-ити невідомі унікасти, ARP чи інші пакети.



Масштабованість

Централізований контроль стану мережі забезпечує ефективність у великих fabric.

EVPN — це не VXLAN

MP-BGP (Control Plane)

MPLS (Data Plane)

EVPN over MPLS
(MPLS service labels)

Provider Based Bridging
(Data Plane)

EVPN with PBB
(Higher scalability,
deployed over MPLS)

Network Virtualization
Overlay (Data Plane)

EVPN over NVO tunnels
(e.g. VXLAN, MPLSoGRE,
deployed over IP)

EVPN
Control Plane



EVPN-VXLAN
Data Center



Оптимально для сучасних
Data Center мереж

EVPN-MPLS
Service Provider



Оптимально для провайдерських
мереж (L2/L3 VPN, EVPN PBB)



Незалежність
від транспортної
технології



Єдина модель
керування для різних
типів оверлеїв



Гнучкість
вибору найбільш
підходящого
Data Plane

Що дає EVPN?

1 MAC Distribution

Розповсюдження
MAC/IP через MP-BGP



2 ARP Suppression

Зменшення Broadcast
та ARP Flood



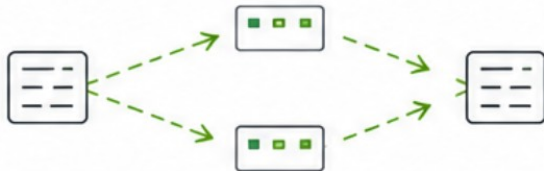
3 Multi-Homing

Активне використання
всіх каналів



4 Aliasing

Балансування до
Multi-Homed вузлів



5 Mass Withdraw

Швидка конвергенція
при відмовах



6 IP Prefix Advertisement

Маршрутизація між VRF
без L2 Flooding



EVPN — це набір мережевих сервісів, а не лише заміна Flood & Learn.
Менше Flooding. Швидша конвергенція. Краща масштабованість.

Що змінює EVPN?

Від класичного Ethernet до сучасної Fabric

Класичний Ethernet



Flood & Learn

Комутатори вивчають MAC через Flooding



Broadcast / ARP

Широкомовні ARP запити по всьому сегменту



STP

Запобігання петлям шляхом блокування портів



Blocked Links

Резервні канали заблоковані STP



Slow Convergence

Повільне відновлення після відмов (секунди)



4094 VLAN ID

Обмежений простір сегментації (12-bit)



EVPN Fabric



MAC/IP Advertisement

Розповсюдження MAC/IP інформації через MP-BGP (Control Plane)



ARP Suppression

Локальна обробка ARP. Менше Broadcast трафіку



ECMP + Multi-Homing

Використання всіх доступних шляхів без блокування



Active/Active

Усі канали активні одночасно. Більше пропускної здатності



Fast Convergence

Швидка конвергенція завдяки BGP EVPN (сотні мілісекунд)



16M VNI (24-bit)

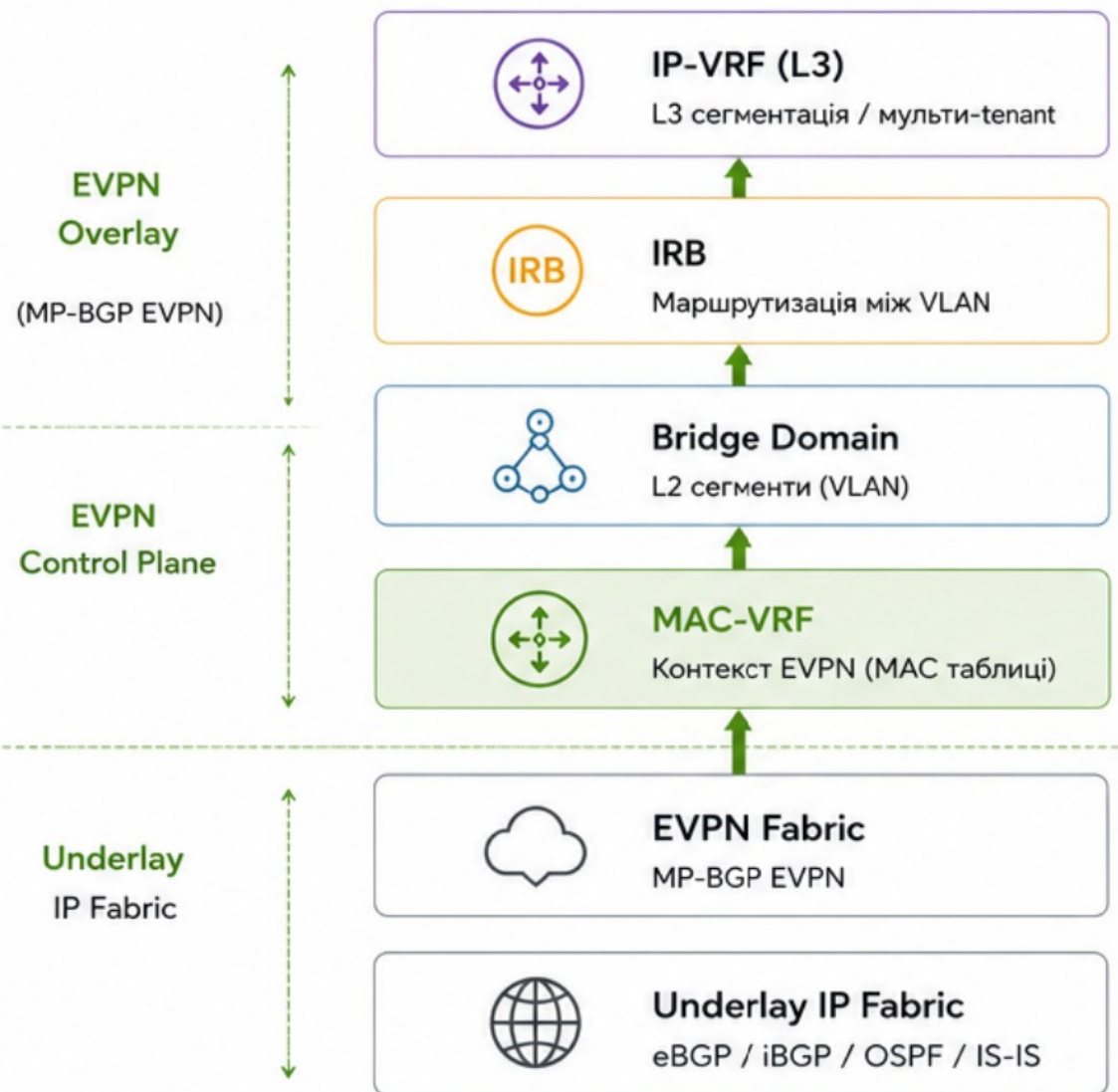
Масштабованість сегментації до ~16 мільйонів



Розділ 2

Архітектура EVPN у Junos

Від Underlay IP Fabric до MAC-VRF та IP-VRF



MAC-VRF — центральний елемент EVPN архітектури в Junos. Усередині нього створюються L2 сервіси (Bridge Domain), додається IRB для маршрутизації та формується L3 сегментація (IP-VRF).

Чому MAC-VRF, а не virtual-switch?

EVPN змінив не лише Control Plane, а й модель побудови L2 сервісів.

	virtual-switch (традиційний)	MAC-VRF (сучасний стандарт EVPN)
 Архітектура	Ієрархія bridge-domains усередині instance	MAC таблиці відокремлені від глобальної таблиці, як VRF для IP
 Control Plane	Локальне навчання MAC (Flood & Learn)	Інформація про MAC/IP розповсюджується через BGP (Route Type 2, 3, 5 тощо)
 Типи сервісів	VLAN-based або VLAN-aware	Підтримує всі типи сервісів: VLAN-based, VLAN-aware, VLAN-bundle в єдиній логіці
 Універсальність	Прив'язаний до конкретного типу сервісу	Єдиний, уніфікований спосіб для всіх EVPN сценаріїв
 Платформи та масштабування	Підтримується практично на всіх платформах (MX, QFX, EX)	Новий стандарт: вимагає свіжих версій Junos, підтримка на сучасних платформах (QFX5k/10k/EX4k, MX тощо)

Що відбувається у BGP?

```
user@leaf01> show route table bgp.evpn.0
...
Route Distinguisher: 10.0.0.1:1
> 2:65000:10100::1:1:1      Type 2
   MAC Advertisement
...

MAC-VRF генерує та анонсує EVPN
Route Type 2 (MAC/IP Advertisement).
```

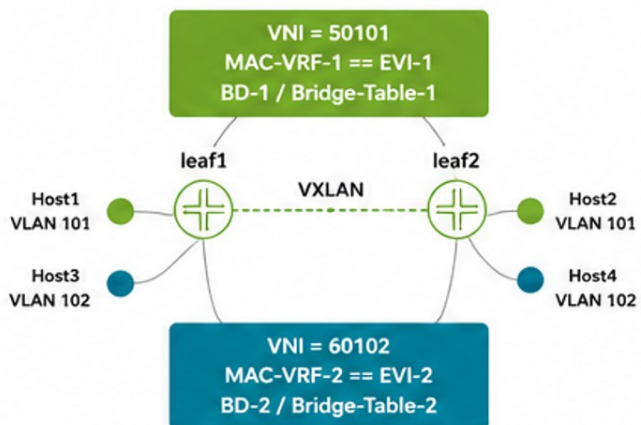
Junos CLI

virtual-switch	MAC-VRF
routing-instances { Tenant-A { instance-type virtual-switch; service-type vlan-aware; } }	routing-instances { Tenant-A { instance-type mac-vrf; service-type vlan-aware; } }

MAC-VRF: Service Types

1. VLAN-based

1:1 VLAN ID → EVI



- 1:1 VLAN ID до EVI
- RT per VLAN-VNI
- Підтримує нормалізацію VLAN (опційно)
- Ефективний flooding (broadcast domain per EVI)
- Перекриття VLAN-id всередині вузла (SP-style)
- Кожен VLAN-VNI має окрему bridge-table
- Type-5 VRF віртуалізація per edge leaf (додаткова опція: ERB та BO дизайн)

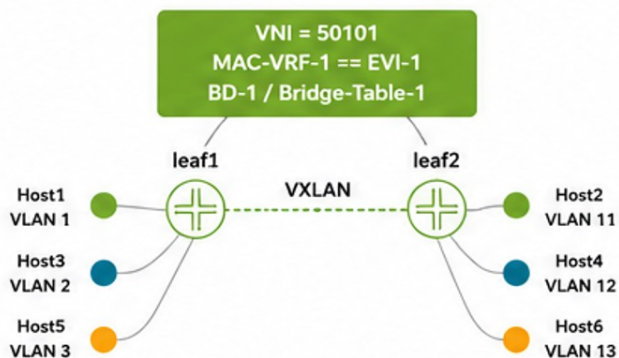


Коли використовувати:

Класичні мережі з унікальними VLAN ID на рівні всього датасентру.

2. VLAN-aware

N:1 (N-VLAN-ID → EVI)



- N:1 (N-VLAN-ID до EVI)
- RT per VRF
- Підтримує нормалізацію VLAN (опційно)
- Ефективний flooding
- Менше EVPN AD routes
- Перекриття VLAN-id всередині вузла (SP-style)
- Легше додавати нові VLAN
- Кожен VLAN-VNI має окрему bridge-table
- Type-5 VRF віртуалізація per edge leaf (додаткова опція: ERB та BO дизайн)

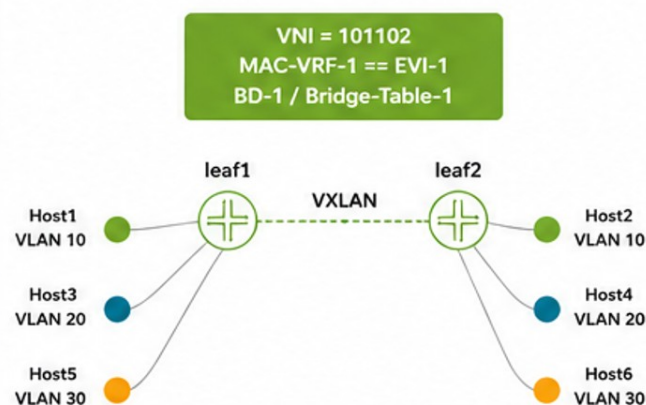


Коли використовувати:

Багато VLAN з можливим перекриттям ID між різними майданчиками.

3. VLAN-bundle

N:1 (N-VLAN-ID → EVI)



- N:1 (N-VLAN-ID до EVI)
- RT per VRF
- Менше EVPN AD routes
- Перекриття VLAN-id всередині вузла (SP-style)
- Немає нормалізації VLAN
- Усі VLAN з даного EVI в одній bridge-table
- Bridged-overlay архітектура



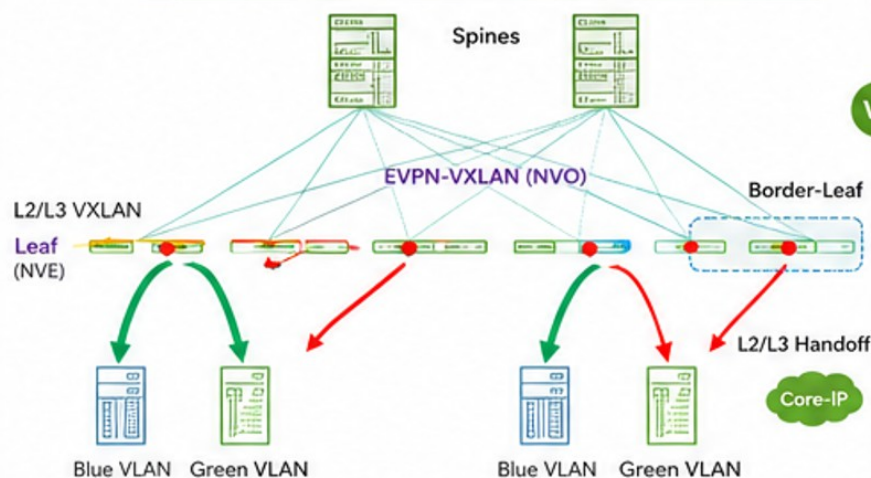
Коли використовувати:

Коли потрібні перекриття VLAN та мінімум EVPN маршрутів (масштабованість).

L3 у EVPN: IRB , VRF та Route T5

CRB vs ERB: де виконується маршрутизація?

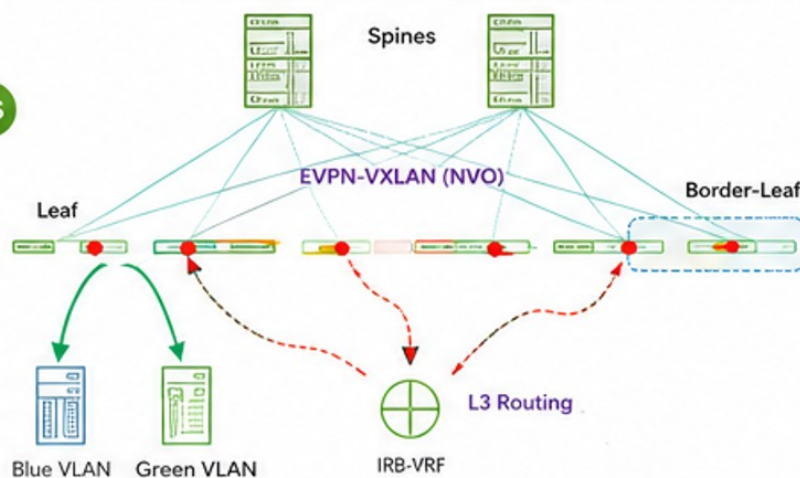
Edge Routed Bridging (ERB)



- Маршрутизація виконується на кожному Leaf
- Локальний шлюз для кожного сегмента L2
- Висока масштабованість та відсутність центрального шлюзу

VS

Central Routed Bridging (CRB)



- Маршрутизація виконується на Border або Central Gateway
- Один шлюз для всіх сегментів
- Простота архітектури, але можливе вузьке місце

Порівняння ERB та CRB

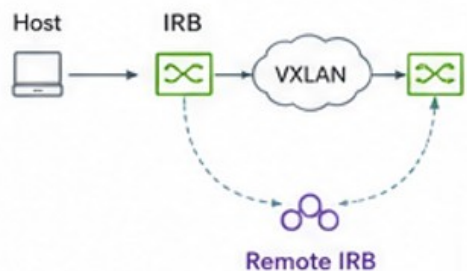
Критерій	ERB (Edge-Routed)	CRB (Central-Routed)
Роль моделі	Розподілена L2/L3 фабрика	Централізована L3 фабрика
Де виконується L3	На кожному Leaf	На Spine / Border
Default Gateway	Anycast IRB на Leaf	Централізований шлюз
IRB інтерфейси	На всіх Leaf	Лише на центральних вузлах
EVPN Route Types	Type-2, Type-5	Type-2 (інколи Type-5)
East-West трафік	Маршрутизується локально	Trombone через центр
Затримка (Latency)	Низька	Середня
Масштабованість	Висока	Середня
Відмовостійкість	Розподілена	Централізована
Типовий use case	Modern DC / Cloud	Міграції / Legacy DC

L3 у EVPN: IRB , VRF та Route T5

Symmetric IRB, Asymmetric IRB та Anycast Gateway

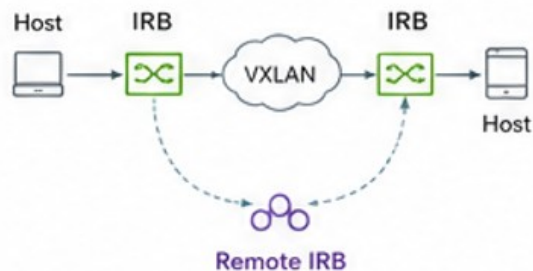
IRB: Symmetric vs Asymmetric

Asymmetric IRB



- Маршрутизація лише на ingress Leaf
- Простіше налаштування
- Менше використання ресурсів

Symmetric IRB



- Маршрутизація на ingress і egress Leaf
- Краща масштабованість
- Підтримка Type-5 та розширених сценаріїв

Anycast Gateway



✓ Однаковий IP та MAC на всіх Leaf

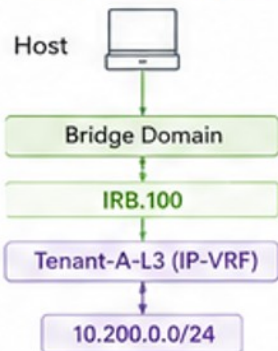
Порівняння

Критерій	Asymmetric	Symmetric
Маршрутизація	Ingress Leaf	Ingress + Egress Leaf
Масштабованість	Середня	Висока
Складність	Низька	Вища
VRF	Менше	Більше
Type 5 Ready	Ні	Так

L3 у EVPN: IRB , VRF та Route T5

IP-VRF та Route Type 5

IP-VRF (Маршрутизація L3)



- Маршрутизація IP в межах VRF
- Ізоляція маршрутів між Tenant
- Використовується Route Type 5

```
routing-instances {
  Tenant-A-L3 {
    instance-type vrf;
  }
}
```

Route Type 5 (EVPN IP Prefix)



Type 5
IP Prefix Advertisement

```
user@leaf01> show route table bgp.evpn.0
5:65000:10100::10.20.0.0/24
```

- Анонсуються IP-префікси без MAC-адрес
- Використовується для L3 VPN
- Не потребує якого MAC в EVPN

CLI: Створення IP-VRF

```
routing-instances {
  Tenant-A-L3 {
    instance-type vrf;
    protocols {
      evpn {
        ip-prefix-routes;
      }
    }
  }
}
```

BGP EVPN: Route Type 5

```
user@leaf01> show route table bgp.evpn.0
5:65000:10100::10.20.0.0/24
Best Active BGP route
```

- Route Type 5 (IP Prefix Advertisement)
- Анонсуються IP-префікси без MAC-адрес
- Не потребує анонсу MAC в EVPN

L3 у EVPN: IRB , VRF та Route T5

Route Type 2 MAC/IP Advertisement



MAC aa:bb:cc:dd:ee:ff
IP 10.100.0.10

Передає інформацію про кінцеві хости

- MAC + IP Binding
- ARP / ND Suppression
- Anycast Gateway
- Host Mobility

```
show route table bgp.evpn.0  
2:65000:10100::aa:bb:cc:dd:ee:ff
```



З'являється після:
створення MAC-VRF та IRB

Route Type 3 Inclusive Multicast (IMET)



Формує VXLAN Overlay та доставляє BUM-трафік

- VTEP Discovery
- BUM Replication (Flood List)
- Broadcast / Unknown Unicast
- Multicast

```
show route table bgp.evpn.0  
3:65000:10100::0
```



З'являється після:
побудови VXLAN Overlay (VTEP)

Route Type 5 IP Prefix Advertisement



IP-VRF

IP Prefix
10.10.0.0/24

Передає IP-префікси між VRF та мережами

- IP Prefix Advertisement
- Inter-VRF Routing
- DCI (Data Center Interconnect)
- External Prefixes

```
show route table bgp.evpn.0  
5:65000:10100::10.10.0.0/24
```



З'являється після:
створення IP-VRF

Type 2



Хости
(MAC/IP)

Type 3



Overlay та BUM
(IMET)

Type 5



IP-префікси
(L3 Routing)

Приклад виводу:

```
user@leaf01> show route table bgp.evpn.0
```

```
2:65000:10100::aa:bb:cc:dd:ee:ff
```

```
3:65000:10100::0
```

```
5:65000:10100::10.10.0.0/24
```

Type 2 - MAC/IP Advertisement

Type 3 - IMET Route

Type 5 - IP Prefix

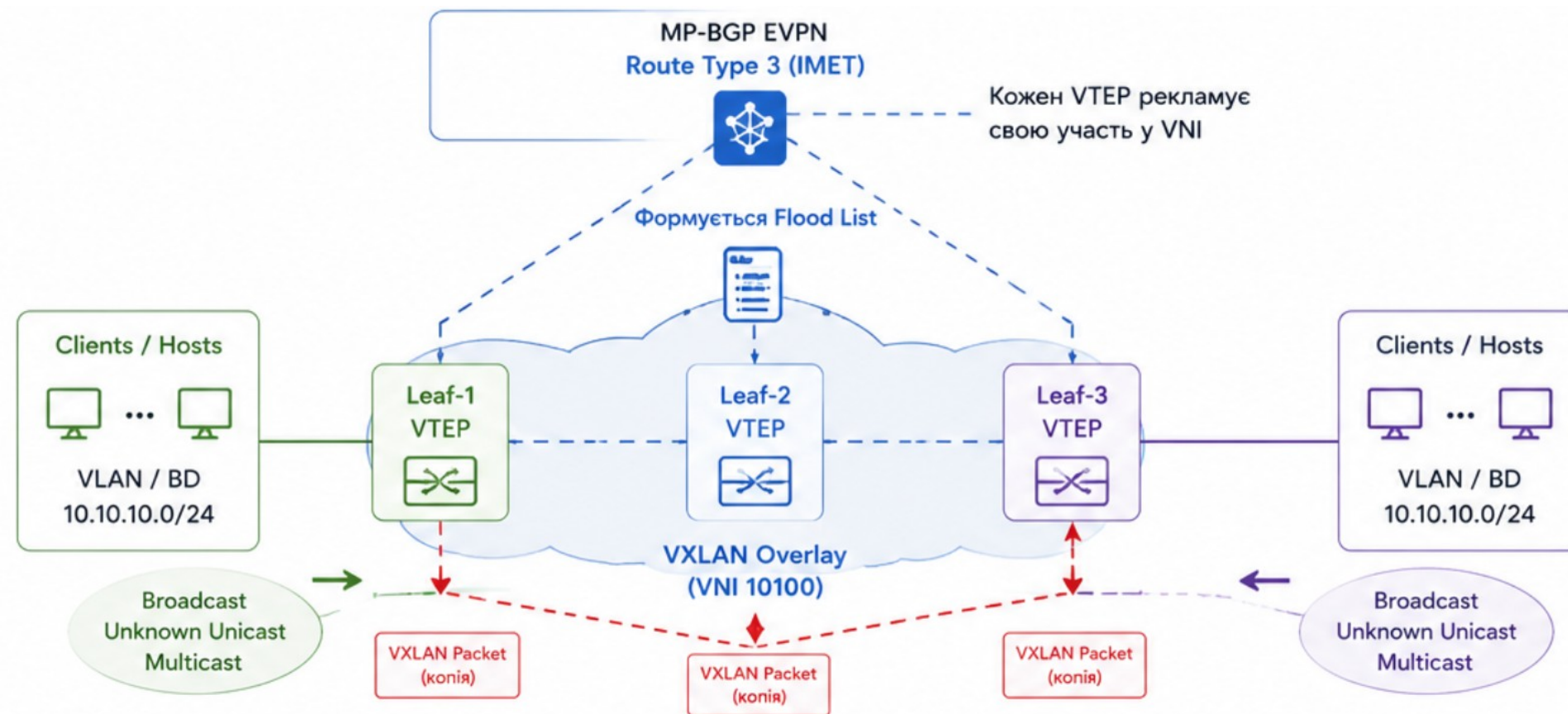


Розділ 3

Механізми захисту EVPN у Junos

Як EVPN доставляє Broadcast, Unknown Unicast та Multicast?

Route Type 3 (IMET) та Ingress Replication

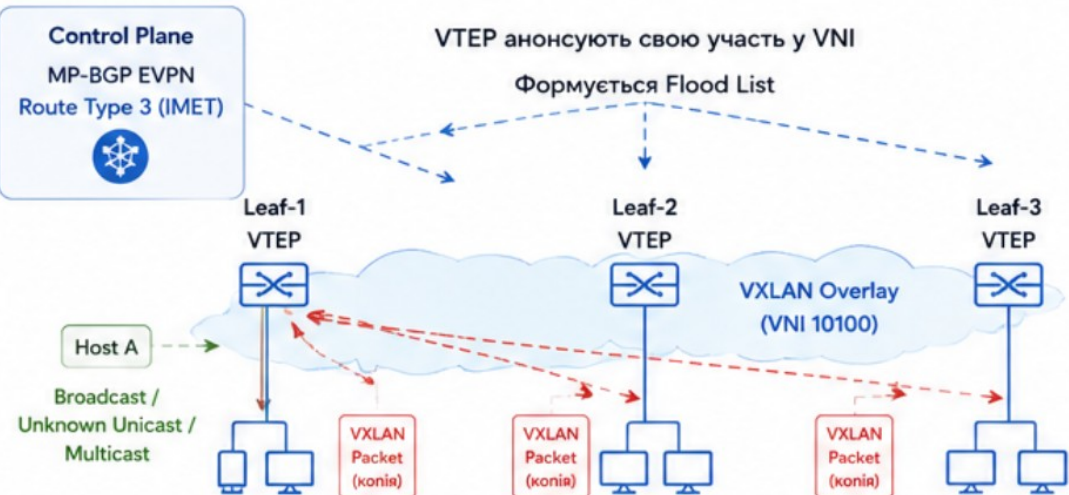


Що відбувається?

1. VTEP анонсує себе через Route Type 3 (IMET) у певному VNI
2. Інші VTEP отримують інформацію та формують Flood List
3. Leaf отримує BUM-трафік від хоста
4. Створюються VXLAN-копії та відправляються кожному VTEP зі списку
5. VTEP отримує та доставляє кадр локальним хостам

Ingress Replication vs Multicast Underlay

1. Ingress Replication (Head-End Replication)



Як це працює

1. Leaf отримує BUM-кадр від хоста.
2. Створює окрему VXLAN-копію для кожного VTEP зі списку Flood List.
3. Кожна копія унікальним чином відправляється відповідному VTEP.

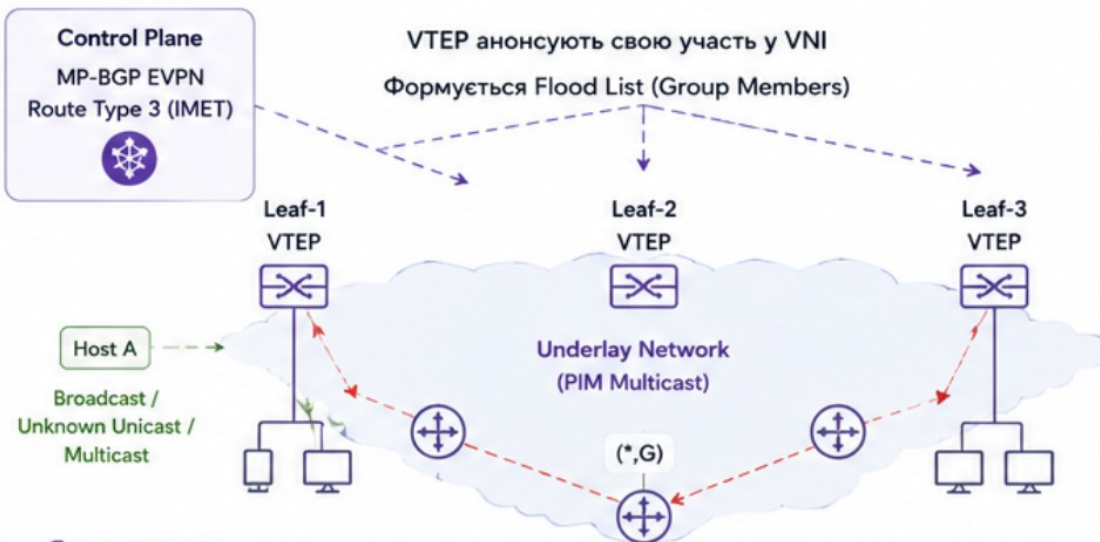
✓ Переваги

- Не потребує Multicast у Underlay
- Простота налаштування
- Працює скрізь (DC, Cloud, Internet)
- Передбачувана поведінка
- Підтримується всіма платформами

✗ Недоліки

- Вищий трафік у Leaf (CPU, пам'ять)
- Більше унікального трафіку в Underlay
- Масштабування обмежене (дуже великі фабрики > 5k+ VTEP)

2. Multicast Underlay (PIM/IGMP)



Як це працює

1. Leaf відправляє BUM-кадр у Multicast групу (*,G).
2. Underlay (PIM/IGMP) будує дерево розповсюдження.
3. Кадр реплікується в мережі лише на гілках, де є VTEP-отримувачі.

✓ Переваги

- Нижче навантаження на Leaf (менше копій)
- Менше трафіку в Underlay
- Краща масштабованість (10k+ VTEP)
- Ефективно для дуже великих фабрик

✗ Недоліки

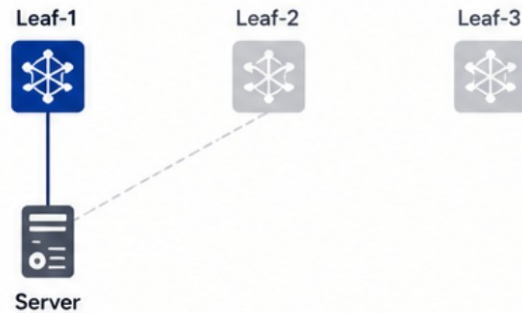
- Потребує Multicast у Underlay (PIM)
- Складніше налаштування та експлуатація
- Потрібна підтримка Multicast на всіх пристроях Underlay

EVPN Single-Homing vs Multi-Homing

Від одного підключення до Active-Active відмовостійкості

Single-Homing

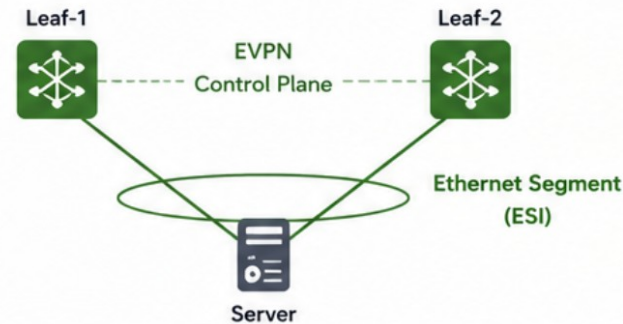
Сервер підключений до одного Leaf-комутатора



- ✓ Просте підключення
- ✓ Мінімальна конфігурація
- ✗ Single Point of Failure
- ✗ Один активний uplink

EVPN Multi-Homing

Сервер одночасно підключений до двох Leaf-комутаторів



- ✓ Active / Active
- ✓ Немає STP
- ✓ Немає MC-LAG / ICCP
- ✓ Швидке перемикання
- ✓ Балансування трафіку
- ✓ Стандартизований EVPN

Характеристика	Single-Homing		EVPN Multi-Homing
Підключення сервера	До одного Leaf	→	До двох Leaf
Доступність	Single Point of Failure	→	Відмовостійкість
Активні uplink'и	Один	→	Два (Active-Active)
Залежність від STP	Може знадобитися	→	Не потрібен
Технологія	Простий LACP	→	EVPN Мульти-хомінг (ESI, DF Election)

Як це працює?



Сервер підключений до двох Leaf через один **Ethernet Segment Identifier (ESI)**.



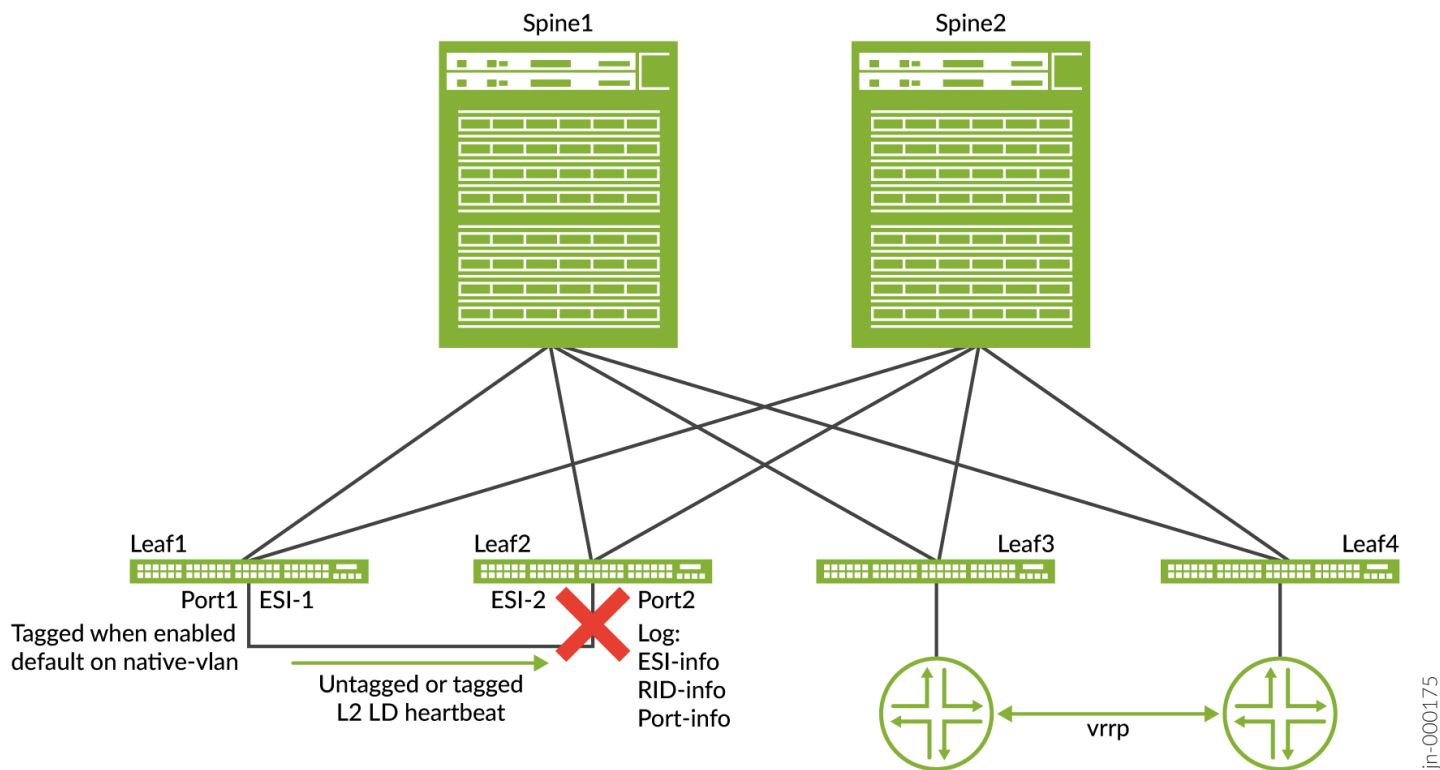
Координація та синхронізація між Leaf виконується через **BGP EVPN**.



EVPN використовує **Route Type 1** та **Type 4** для забезпечення узгодженості та швидкого відновлення після відмов.

EVPN Loop Detection

Захист від L2-петель на Leaf-Server з'єднаннях



- EVPN Control Plane не може запобігти всім фізичним L2-петлям
- в **Junos** використовується механізм Lightweight Loop Detection
- Loop Detection працює незалежно від EVPN Control Plane

```
protocols {  
  loop-detect {  
    enhanced {  
      interface ae0.100 {  
        vlan-id 100;  
        loop-detect-action interface-  
down;  
      }  
    }  
  }  
}
```



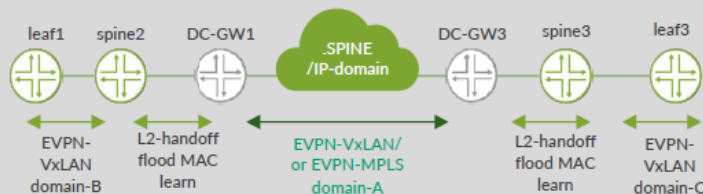
Розділ 3

EVPN DCI та мережі Service Provider

EVPN DCI: варіанти побудови між сайтами

Option 1

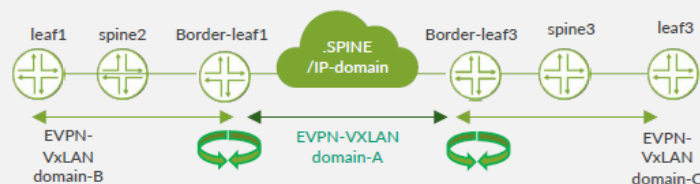
Decoupled Interconnect model using L2 VLAN handoff – L2 DCI



- Harder to manage - more devices/equipment
- Traditional demarcation points
- Flood based MAC learning

Option 3

Seamless EVPN-VxLAN to EVPN-VXLAN seamless stitching - L2 DCI

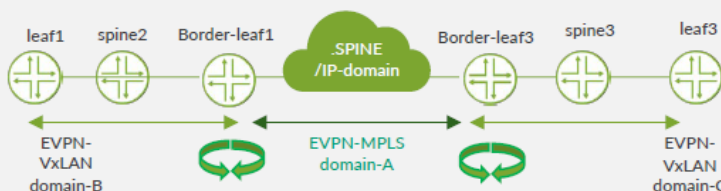


- Unified EVPN layer 2 solution
- Pure overlay with controlled scaling
- Higher scale—more DC sites and pods



Option 4

Seamless EVPN-VxLAN to EVPN-MPLS seamless stitching – L2 DCI

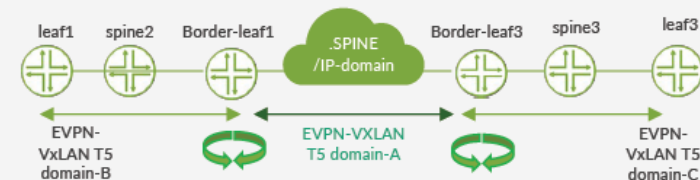


- Unified EVPN layer 2 solution
- Pure overlay with controlled scaling
- Higher scale—more DC sites and pods
- MPLS direct connect from border-leaf



Option 5

EVPN Type-5 to EVPN Type-5 VxLAN to VxLAN stitching - L3 DCI

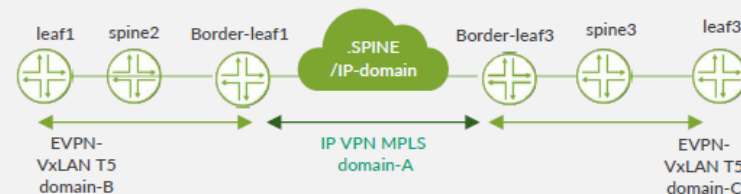


- Unified EVPN layer 3 solution
- Pure overlay with controlled scaling
- Higher scale—more DC sites and pods
- VXLAN for IP only DCI



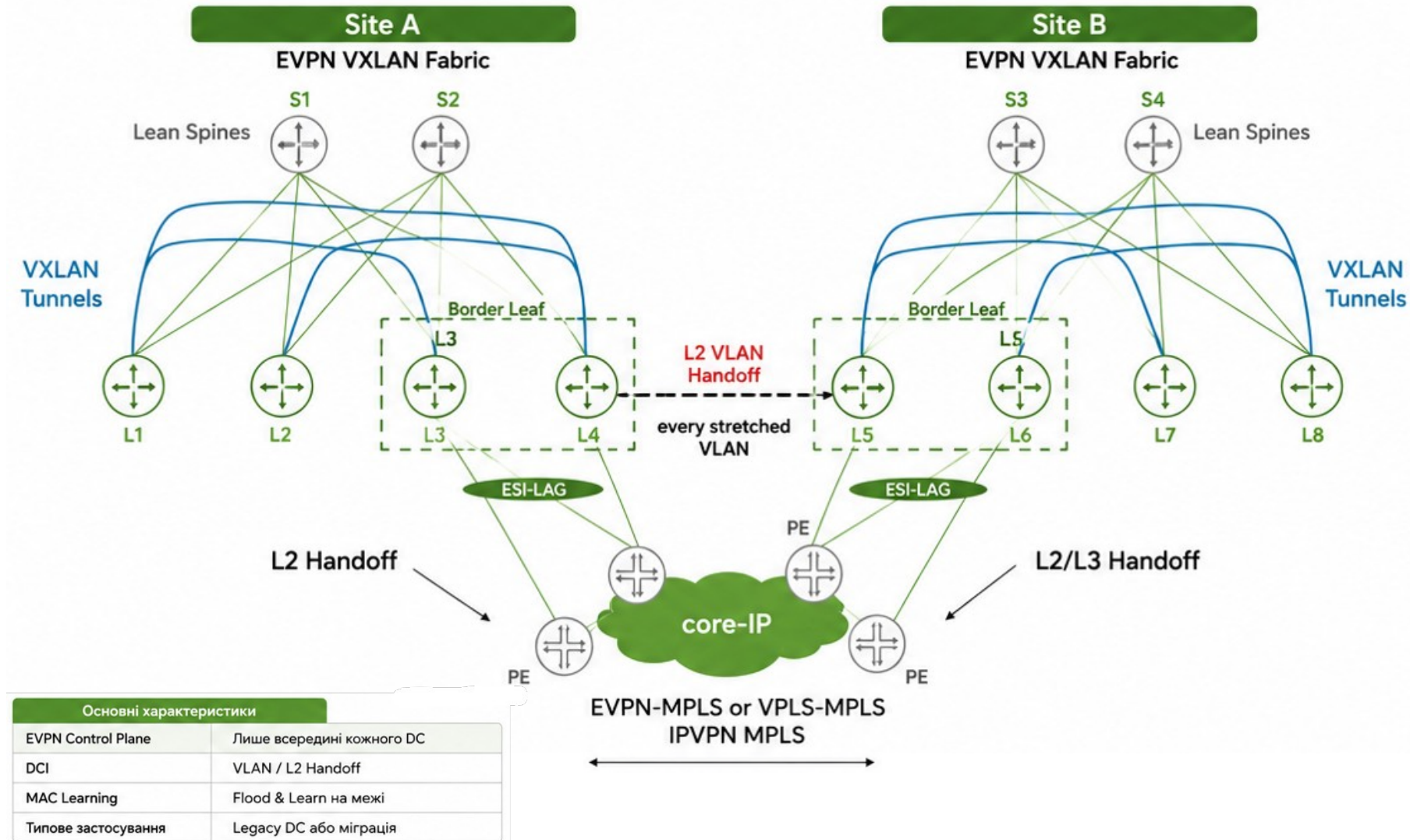
Option 6

EVPN-VXLAN Type-5 to IPVPN MPLS – L3 DCI



- Unified EVPN layer 3 solution
- Pure overlay with controlled scaling
- Higher scale—more DC sites and pods
- MPLS IPVPN using IP only DCI

DCI Option 1: VLAN Handoff (Legacy)



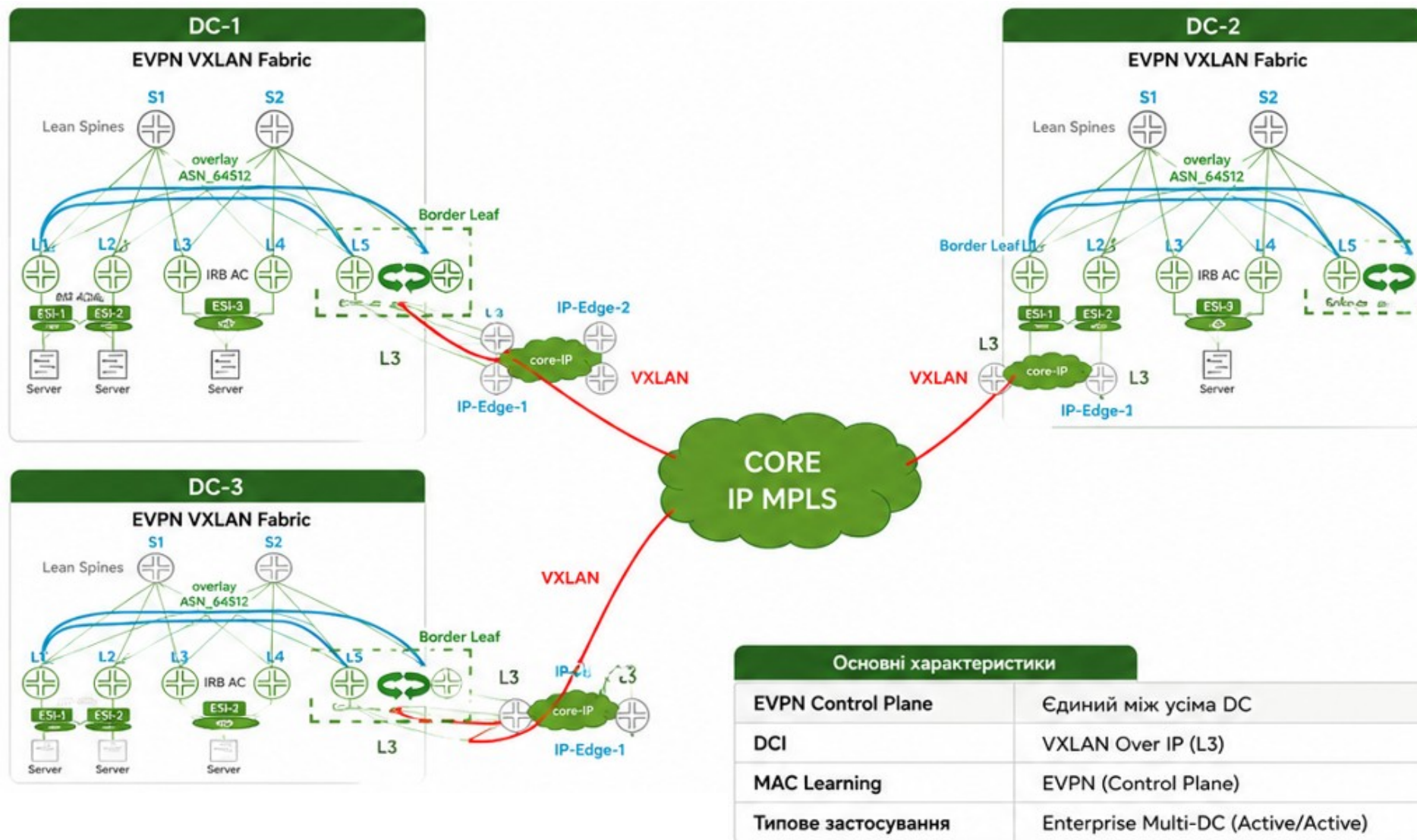
- Option 1 – найпростіший спосіб об'єднання і повертає класичні механізми Layer 2
- Не потребує перебудови старої мережі
- Можливе поєднання з EVPN-PBB на MPLS Core



Обмеження

- ✗ EVPN не проходить між DC
- ✗ Flood & Learn на стику доменів
- ✗ Окремі Border Gateway
- ✗ Складніше масштабування
- ✗ Більше операційних витрат

DCI Option 3: EVPN VXLAN ↔ VXLAN

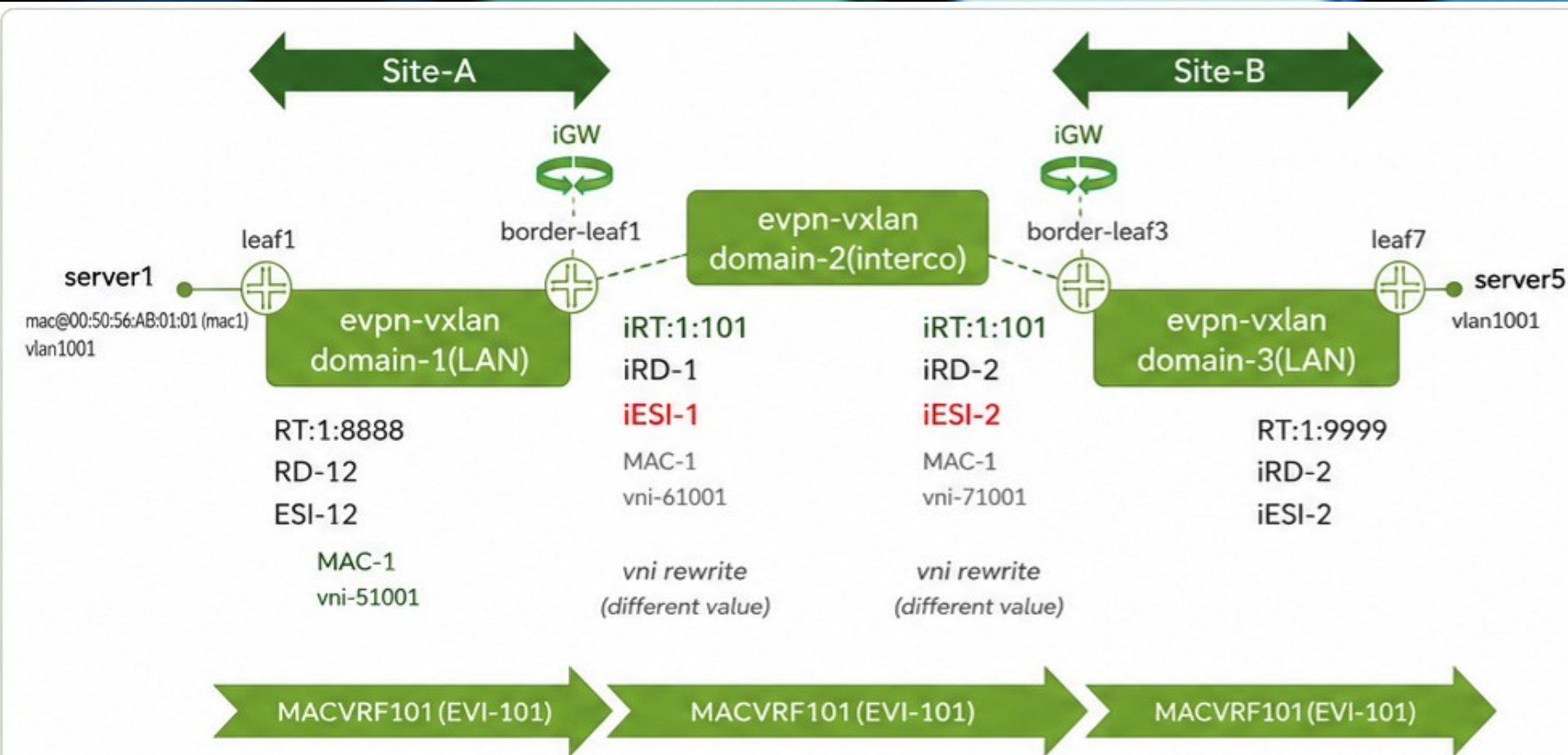


- Option 3 – єдиний EVPN CP
- MAC вивчаються глобально
- Активні лінки між майданчиками
- Краща масштабованість

⚠ Обмеження

- ✗ Потрібна IP/MPLS інфраструктура
- ✗ Необхідна уніфікація конфігурацій
- ✗ Планування IP та VTEP адресації
- ✗ Залежить від якості WAN

DCI Option 3: EVPN VXLAN ↔ VXLAN



Як це працює

- 1 Обмін EVPN маршрутами**
Border Leaf аноншують EVPN маршрути (Type 2, Type 3, Type 5) в єдиний BGP EVPN домен (AS).
- 2 Оголошення MAC (Type 2)**
Коли хост відправляє трафік, Leaf1 вивчає MAC-адресу mac1 у VLAN 1001 і анонсує Type 2 маршрут.
- 3 Поширення MAC**
Маршрут Type 2 з VNI та MAC поширюється між усіма DC. Border Leaf3 отримує інформацію про MAC.
- 4 Встановлення VXLAN тунелю**
Border Leaf автоматично створює VXLAN тунель до віддаленого VTEP (Border Leaf1).
- 5 Перепис VNI**
На міжсайтовому стику виконується VNI rewrite (різне значення VNI в кожному домені).
- 6 Доставка трафіку**
Трафік доставляється напряму між VTEP через VXLAN Over IP без розтягування традиційного L2.

EVPN Route Types у DCI

- Type 2 – MAC/IP Advertisement
- Type 3 – Inclusive Multicast
- Type 5 – IP Prefix (за потреби L3 DCI)

Що важливо

- ✓ Єдиний AS для всіх DC
- ✓ Унікальний Loopback для кожного VTEI
- ✓ Reachability між VTEP по Underlay
- ✓ Однакова політика EVPN (RT/Import/Export)
- ✓ Active/Active між DC

DCI Option 3: EVPN VXLAN ↔ VXLAN

```
macvrfAB-vlan-aware {  
  instance-type mac-vrf;  
  protocols {  
    evpn {  
      encapsulation vxlan;  
      default-gateway no-gateway-community;  
      extended-vni-list [ 5055 5056 ];  
      interconnect {  
        vrf-target target:1:55;  
        route-distinguisher 172.16.7.11:55;  
        esi {  
          00:00:99:99:99:99:99:55:55;  
          all-active;  
        }  
        interconnected-vni-list [ 6055 ];  
      }  
    }  
  }  
}
```

1

New tunnel
stitching
configuration
block

2 VNI translation
specific to
interconnect
function

```
vtep-source-interface lo0.0;  
service-type vlan-aware;  
route-distinguisher 172.16.7.11:155;  
vrf-target target:1:155;  
vlangs {  
  vlang55 {  
    vlang-id 55;  
    vxlan {  
      vni 5055;  
      translation-vni 6055;  
    }  
  }  
  vlang56 {  
    vlang-id 56;  
    vxlan {  
      vni 5056;  
      translation-vni 6056;  
    }  
  }  
}
```

Що робить ця конфігурація

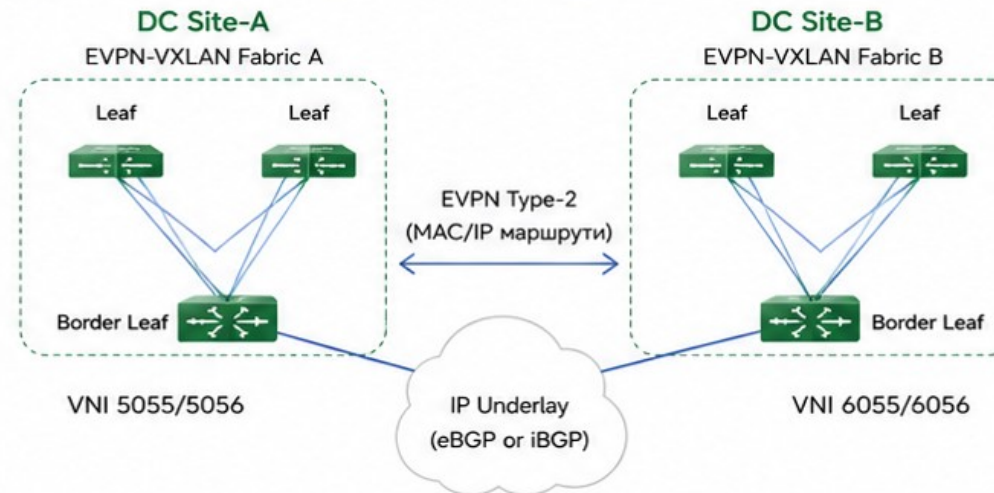
1 Interconnect block – Tunnel Stitching

- Визначає параметри Tunnel Stitching між двома VXLAN фабриками
- Задає RT, RD та ESI для ECMP multi-homing (all-active)
- interconnected-vni-list вказує VNI іншої фабрики, з якою виконується stitching

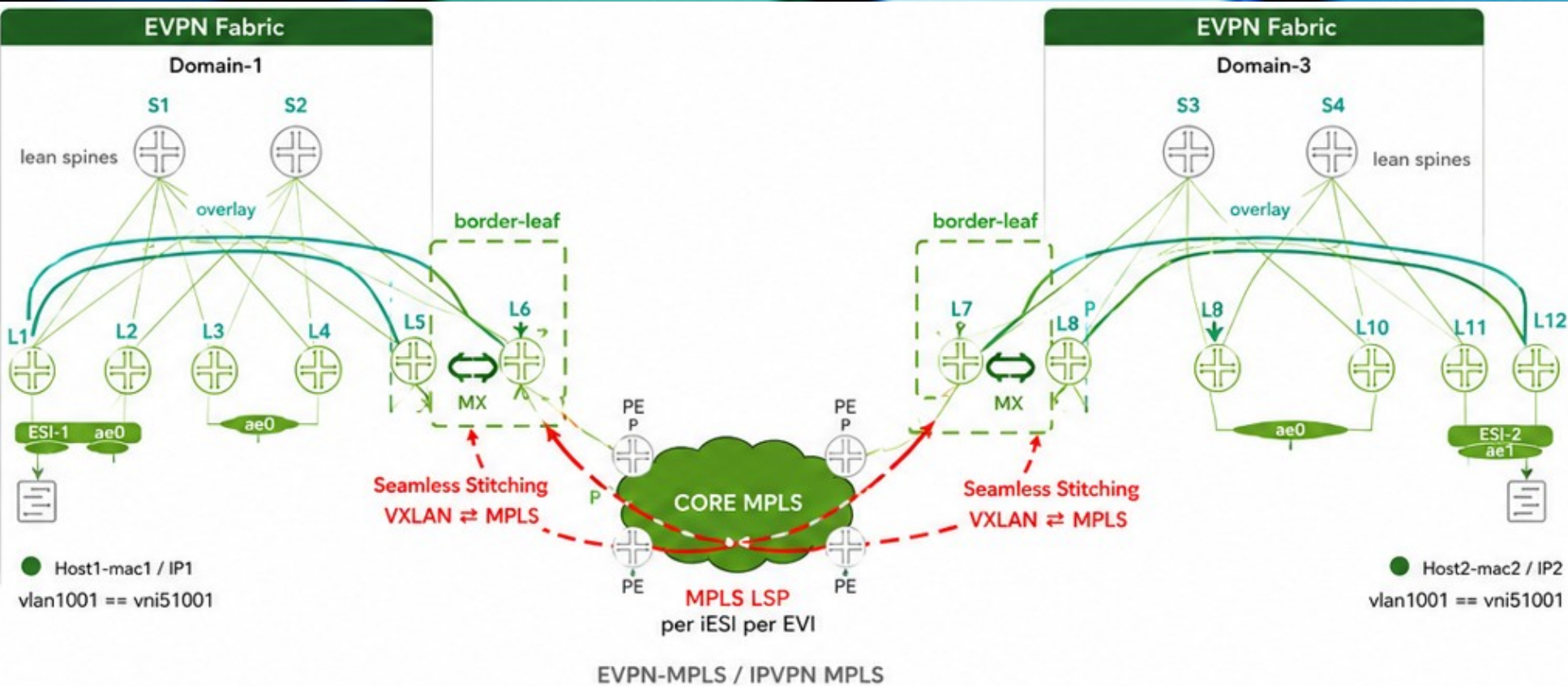
2 VNI translation

- translation-vni визначає VNI віддаленої фабрики (інший VXLAN домен)
- Локальний VNI (5055/5056) ↔ Віддалений VNI (6055/6056)
- Забезпечує обмін Type-2 EVPN маршрутами між фабриками

Схема: Type-2 to Type-2 VXLAN Stitching



DCI Option 4: EVPN VXLAN ↔ MPLS



Переваги

- ✓ EVPN між Fabric, але трафік через MPLS Core
- ✓ Підтримка Multi-Domain MPLS
- ✓ Масштабування через Service Provider
- ✓ Підтримка Active/Active між DC

Обмеження

- ✗ Потрібна MPLS інфраструктура провайдера
- ✗ Складніше налаштування і експлуатація
- ✗ Залежить від якості MPLS Core
- ✗ Додаткові витрати на MPLS сервіси

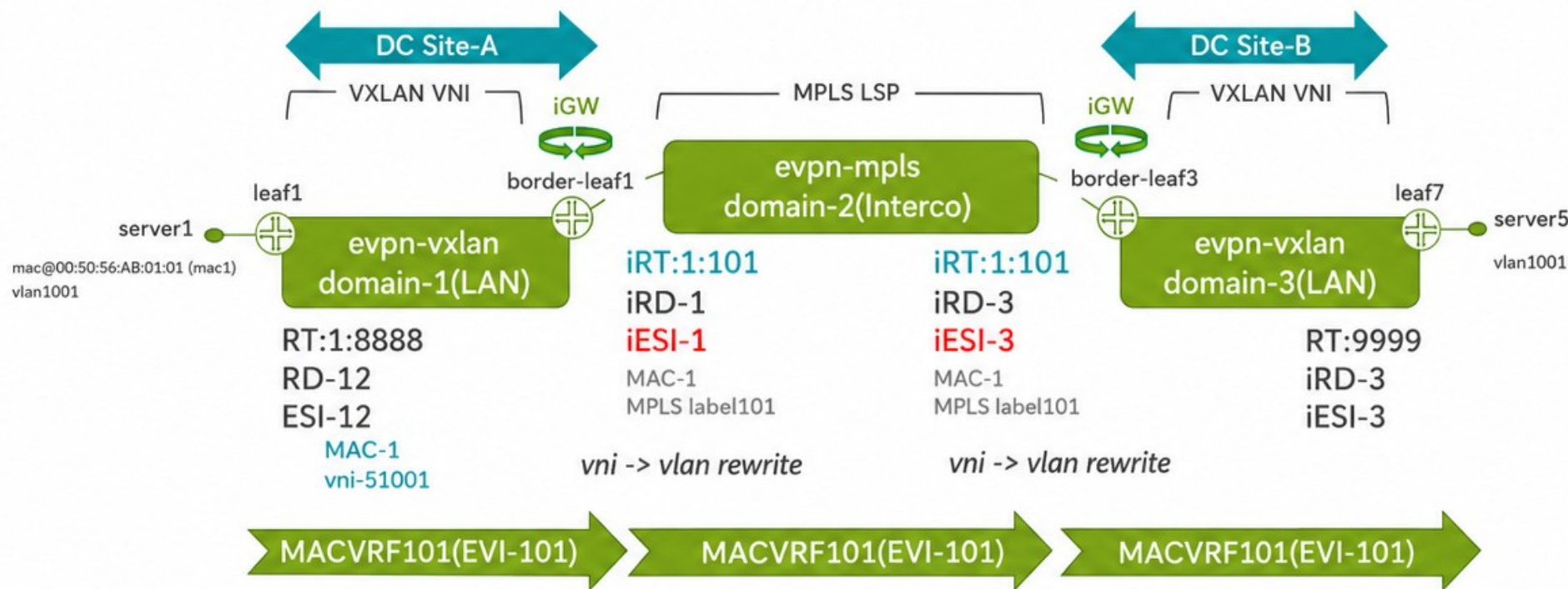
Основні характеристики

EVPN Control Plane	Єдиний між усіма DC (через MPLS)
DCI	EVPN over MPLS (MP-BGP EVPN між PE)
Transport	MPLS LSP (per ESI per EVI)
Stitching	Seamless Stitching VXLAN ↔ MPLS на Border Leaf
Типове застосування	Service Provider / Multi-DC між майданчиками

Що важливо

- ✓ Один EVI = один iESI між DC
- ✓ MPLS LSP формується на рівні PE
- ✓ MP-BGP EVPN між PE для обміну маршрутами
- ✓ VNI та iESI можуть бути однакові в різних DC

DCI Option 4: EVPN VXLAN ↔ MPLS



Маршрути EVPN

- Type 2 - MAC/IP Advertisement
- Type 3 - Inclusive Multicast
- Type 5 - IP Prefix (за потреби L3 DCI)

Що важливо

- ✓ Один EVI = один iESI між DC
- ✓ MPLS LSP формується на рівні PE
- ✓ MP-BGP EVPN між PE для обміну маршрутами
- ✓ VNI та ESI можуть бути однакові в різних DC

Що відбувається

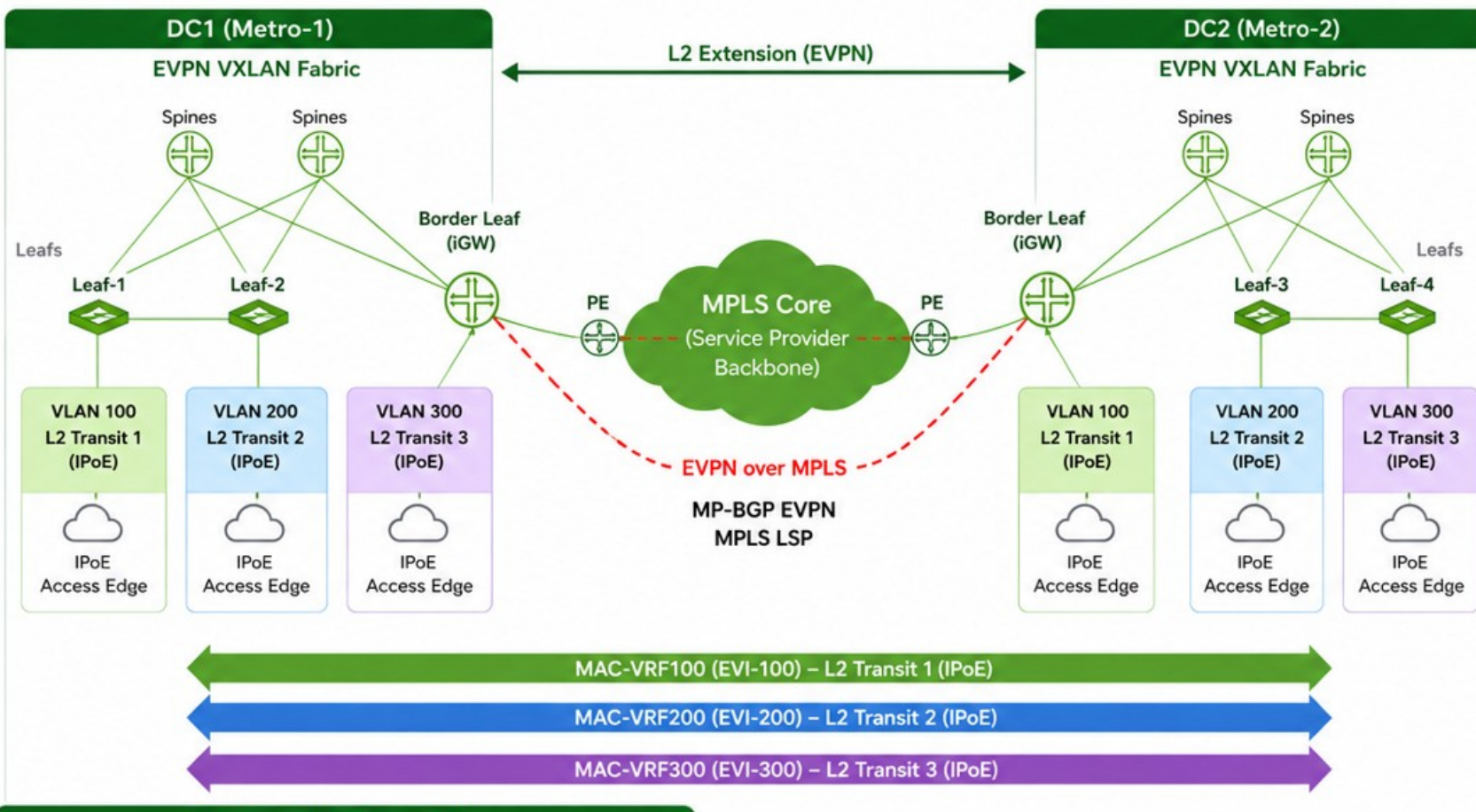
- 1 Leaf1 (DC Site-A) видає MAC-адресу mac1 у VLAN1001 та анонсує Type 2 маршрут із VNI 51001 через EVPN.
- 2 Border-Leaf1 (iGW) імпортує маршрут, виконує VNI -> MPLS label mapping (label 101) та передає маршрут у EVPN-MPLS домен.
- 3 Маршрут поширюється по EVPN-MPLS (iRT/iRD/iESI, label 101) до Border-Leaf3.
- 4 Border-Leaf3 (iGW) виконує MPLS label -> VNI mapping та VNI -> VLAN rewrite (за потреби).
- 5 Leaf7 (DC Site-B) імпортує Type 2 маршрут та навчає MAC mac1 у VLAN1001.
- 6 Трафік між server1 та server5 передається через MPLS LSP (label 101) між PE.

Позначення:

- iGW iGW - Interconnect Gateway
(EVPN Stitching)
- Border Leaf - VTEP + PE

- PE - Provider Edge (MPLS PE)
LSP - Label Switched Path
EVI - Ethernet Virtual Instance
ESI - Ethernet Segment Identifier

DCI Option 4: EVPN VXLAN ↔ MPLS



- Як це працює**
- 1 Leaf у DC1 анонсує MAC через EVPN (Type 2) для VLAN 100/200/300.
 - 2 Border Leaf (iGW) виконує VNI → MPLS label mapping.
 - 3 Маршрут передається через MP-BGP EVPN між PE.
 - 4 На стороні DC2 виконується MPLS → VNI mapping.
 - 5 Трафік у тому ж L2 сегменті доставляється до Leaf у DC2.

Позначення	
	iGW – Interconnect Gateway (EVPN Stitching)
	Border Leaf – VTEP + PE
	PE – Provider Edge (MPLS PE)
	MP-BGP EVPN
	VXLAN (Overlay)
	MPLS LSP

L2 сегменти (IPoE приклад)			
VLAN	Сегмент (L2 Transit)	MAC-VRF (EVI)	VNI / MPLS Label
100	L2 Transit 1 (IPoE)	EVI-100	VNI 5100 / Label 100
200	L2 Transit 2 (IPoE)	EVI-200	VNI 5200 / Label 200
300	L2 Transit 3 (IPoE)	EVI-300	VNI 5300 / Label 300

DCI Option 4: EVPN VXLAN ↔ MPLS

MX Configuration Example (JSON)

```
{
  "routing-instances": {
    "MACVRF101": {
      "instance-type": "virtual-switch",
      "protocols": {
        "evpn": {
          "encapsulation": "vlan",
          "extended-vni-list": "all",
          "interconnect": {
            "vrf-target": "target:1:101",
            "route-distinguisher": "172.16.7.115:101",
            "esi": "00:00:11:11:11:11:11:11:11:11",
            "all-active": true,
            "interconnected-vlan-list": ["1001"],
            "encapsulation": "mpls"
          }
        }
      },
      "vrf-target": ["target:1:8888"]
    },
    "vtep-source-interface": "lo0.0",
    "bridge-domains": {
      "bd1001": {
        "domain-type": "bridge",
        "vlan-id": 1001,
        "routing-interface": "irb.1001",
        "vlan": {
          "vni": 51001
        },
        "ingress-node-replication": true
      }
    },
    "route-distinguisher": "172.16.7.115:1"
  },
  "interfaces": {
    "lo0.1": {
      "route-distinguisher": "172.16.7.115:100",
      "vrf-target": ["target:1100:1100"],
      "vrf-table-label": true
    }
  }
}
```

EVPN interconnect для зшивання з EVPN-MPLS (Type-2 / Type-5) з ESI all-active

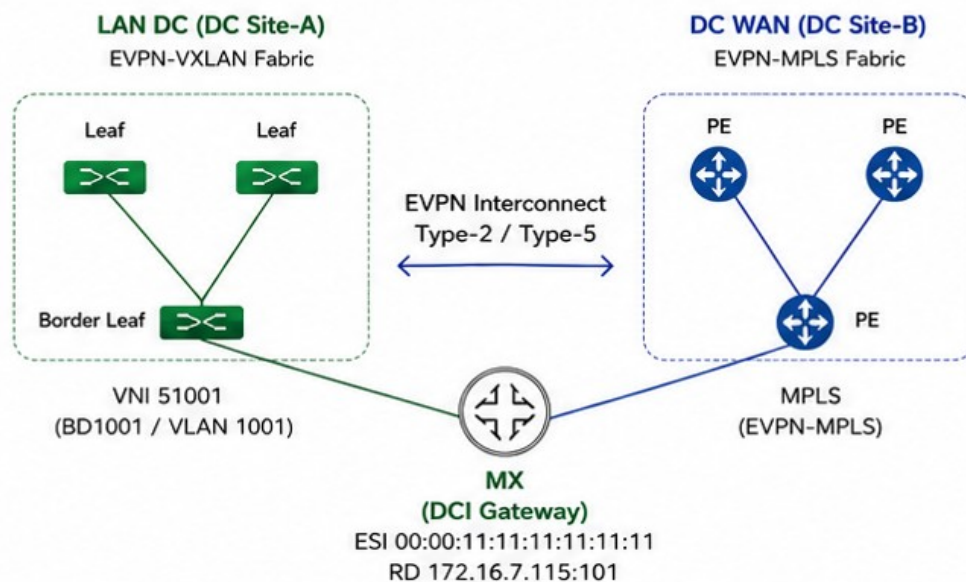
VRF-Target для обміну EVPN маршрутами (Type-2 / Type-5 з префіксами)

Bridge Domain для VLAN 1001. VNI 51001 використовується у VLAN фабриці (VXLAN)

Унікальний RD для EVPN (LAN EVPN-VXLAN) per node

Interface RD для EVPN. RT 1100:1100 – для LAN EVPN intra-fabric маршрутизації

Схема: EVPN-VXLAN to EVPN-MPLS Stitching

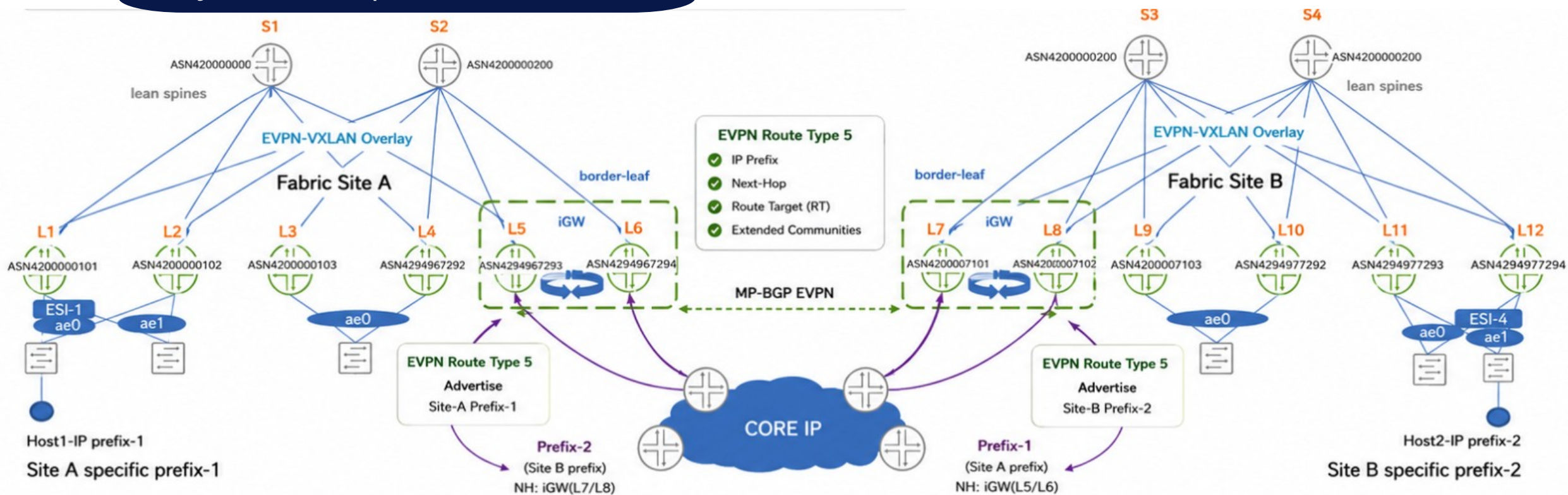


Ключові параметри

- ✓ VXLAN encapsulation для LAN фабрики
- ✓ MPLS encapsulation для DCI в EVPN-MPLS
- ✓ Type-2 (MAC) + Type-5 (IP) маршрути
- ✓ Extended-VNI-List all – для підтримки всіх VNI
- ✓ iESI + all-active – для multi-homing
- ✓ Interconnected-VLAN-List 1001 – VLANи для stitching
- ✓ VNI 51001 – локальний VXLAN VNI в LAN фабриці
- ✓ Унікальні RD та RT для DC та LAN сторін

DCI Option 5: EVPN Type-5 ↔ EVPN Type-5

Layer 3 DCI через IP-VRF



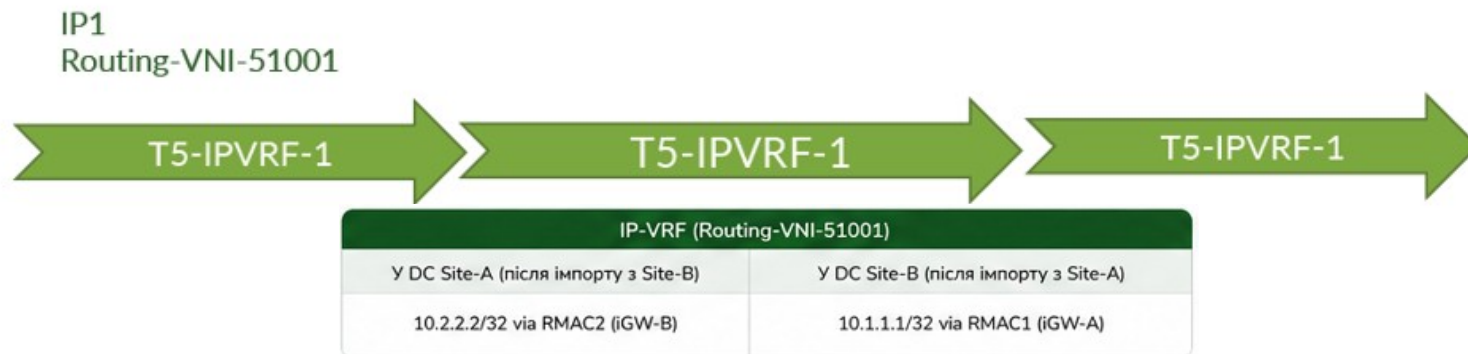
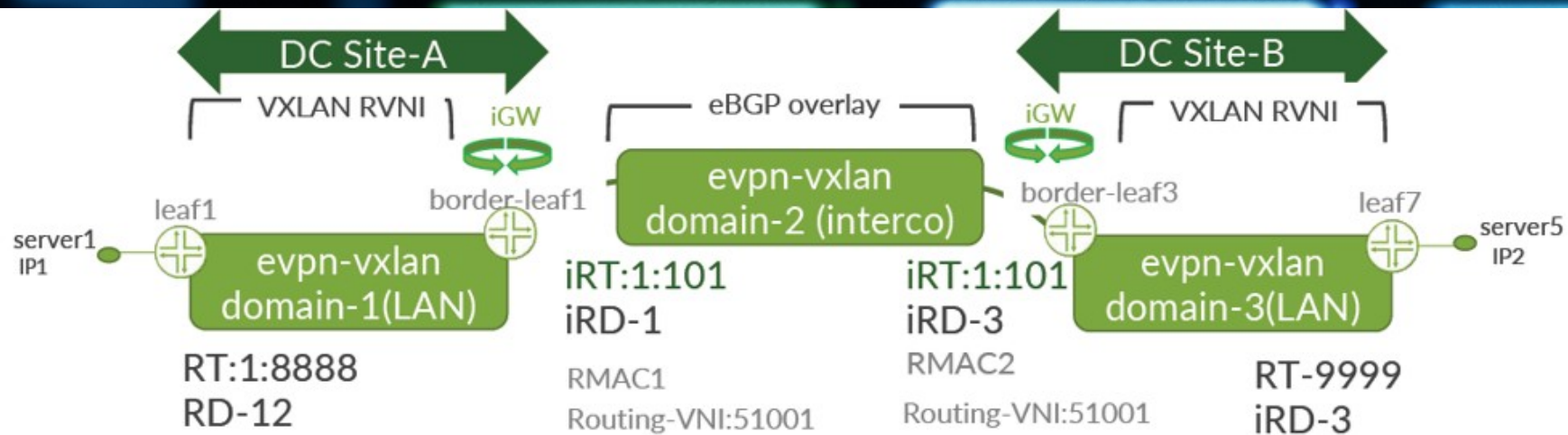
Що передається між DC

- ✓ EVPN Route Type 5 (IP Prefixes)
- ✓ IP маршрути для IP-VRF / L3 сервісів
- ✓ Route Target (RT) для імпорту маршрутів

Що НЕ передається між DC

- ✗ MAC адреси (Route Type 2)
- ✗ Broadcast, Unknown Unicast
- ✗ VLAN та ESI

DCI Option 5: EVPN Type-5 ↔ EVPN Type-5



Позначення	
	iGW – Interconnect Gateway (EVPN Stitching)
	Border Leaf – VTEP + IP Gateway
	Leaf – VTEP
	eBGP MP-EVPN
	EVPN Route Type 5 (IP Prefix Exchange)

Що анонсує Site-A (Route Type 5)

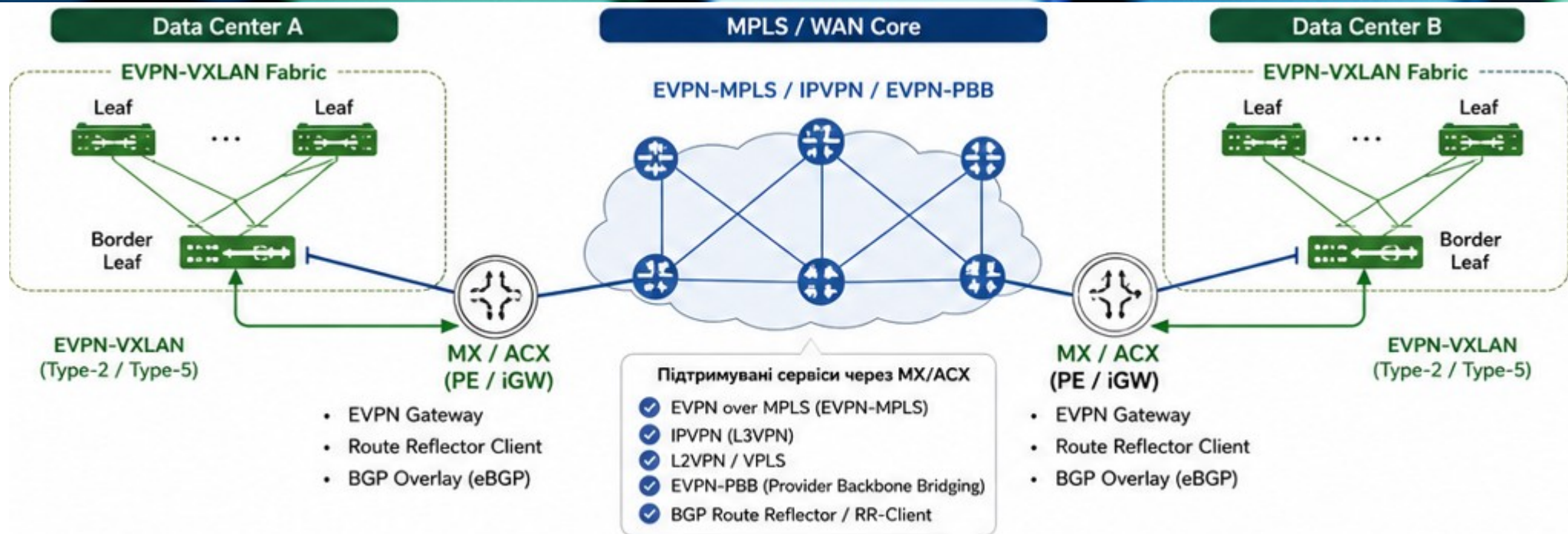
IP Prefix: IP1 (server1)
 Next-Hop: iGW Site-A (RMAC1)
 RT: 1:101
 VN: Routing-VNI-51001

Route Type 5
 (IP Prefix Exchange)

Що анонсує Site-B (Route Type 5)

IP Prefix: IP2 (server5)
 Next-Hop: iGW Site-B (RMAC2)
 RT: 1:101
 VN: Routing-VNI-51001

MX / ACX як DCI Gateway для EVPN Fabrics



Порівняння платформ	
QFX Border Leaf	MX / ACX
✓ EVPN-VXLAN	✓ EVPN-VXLAN
✓ VXLAN ↔ VXLAN DCI	✓ EVPN-MPLS PE
✓ VXLAN ↔ MPLS Stitching	✓ MPLS L2VPN / L3VPN (IPVPN)
✓ Route Type 2 / 5 (EVPN Type-2 / Type-5)	✓ Route Type 2 / 5 (EVPN Type-2 / Type-5)
✓ Data Center Focus	✓ Carrier Ethernet / Service Provider
⚠ Обмежене масштабування WAN	✓ Metro / WAN / DCI
⚠ Типові Enterprise сценарії	✓ Service Provider сценарії

Коли використовувати MX / ACX як DCI Gateway	
	MPLS Core між дата-центрами (EVPN-MPLS)
	Інтеграція з IPVPN (L3VPN) або L2VPN / VPLS
	Metro Ethernet сервіси
	Великі Enterprise / Service Provider мережі
	EVPN-PBB для максимальної масштабованості
	Потреба у BGP Route Reflector та великій кількості EVPN маршрутів

DC Gateway Portfolio

Border Leaf 1 (DC Site-A)

```
root@border-leaf1> show configuration routing-instances T5-VRF1
```

```
instance-type vrf;
```

```
routing-options {
```

```
  static {
```

```
    route 10.10.100.113/32 discard;
```

```
  }
```

```
  multipath;
```

```
}
```

```
protocols {
```

```
  evpn {
```

```
    interconnect {
```

```
      vrf-target target:102:102;
```

```
      route-distinguisher 172.16.7.113:102;
```

```
    }
```

```
  ip-prefix-routes {
```

```
    advertise direct-nexthop;
```

```
    encapsulation vxlan;
```

```
    vni 1100;
```

```
    export my-t5-export-vrf1;
```

```
  }
```

```
}
```

```
interface lo0.1;
```

```
  route-distinguisher 172.16.7.113:100;
```

```
  vrf-target target:1100:1100;
```

```
  vrf-table-label;
```

Border Leaf 2 (DC Site-B)

```
root@border-leaf2> show configuration routing-instances T5-VRF1
```

```
instance-type vrf;
```

```
routing-options {
```

```
  static {
```

```
    route 10.10.100.114/32 discard;
```

```
  }
```

```
  multipath;
```

```
}
```

```
protocols {
```

```
  evpn {
```

```
    interconnect {
```

```
      vrf-target target:102:102;
```

```
      route-distinguisher 172.16.7.114:102;
```

```
    }
```

```
  ip-prefix-routes {
```

```
    advertise direct-nexthop;
```

```
    encapsulation vxlan;
```

```
    vni 1100;
```

```
    export my-t5-export-vrf1;
```

```
  }
```

```
}
```

```
interface lo0.1;
```

```
  route-distinguisher 172.16.7.114:100;
```

```
  vrf-target target:1100:1100;
```

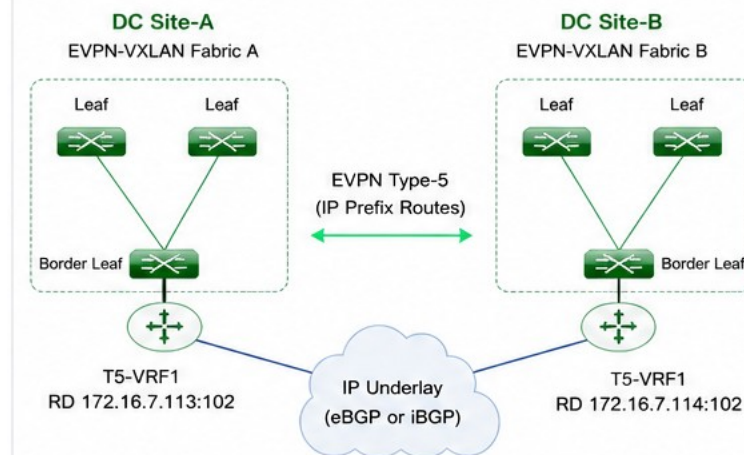
```
  vrf-table-label;
```

Що робить ця конфігурація

1 Interconnect block – Type-5 DCI Stitching

- Визначає параметри для обміну IP-маршрутами (Type-5) між двома VXLAN фабриками
- Однаковий VRF-Target (target:102:102) дозволяє обмінюватися Type-5 маршрутами
- Унікальний RD на кожному сайті забезпечує коректну ідентифікацію префіксів

Схема: Type-5 to Type-5 VXLAN Stitching

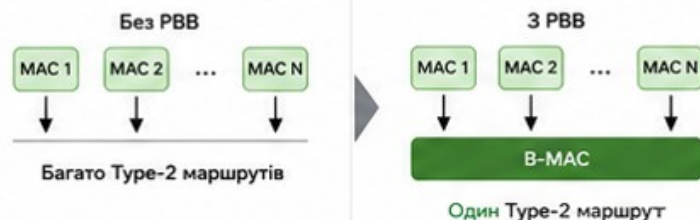


EVPN PBB (Provider Backbone Bridge)

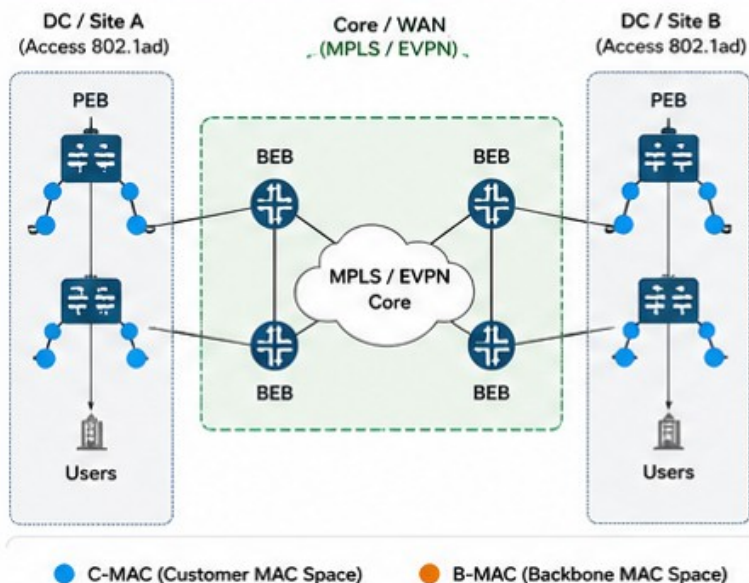
1. Що таке EVPN-PBB?

- Розширення EVPN для масштабованого перенесення L2 через MPLS/EVPN мережу.
- Використовує концепцію Backbone MAC (B-MAC) для агрегації великої кількості Customer MAC (C-MAC).
- Значно зменшує кількість EVPN Type-2 маршрутів у мережі.
- Підходить для великих DCI, Metro Ethernet та Service Provider мереж.

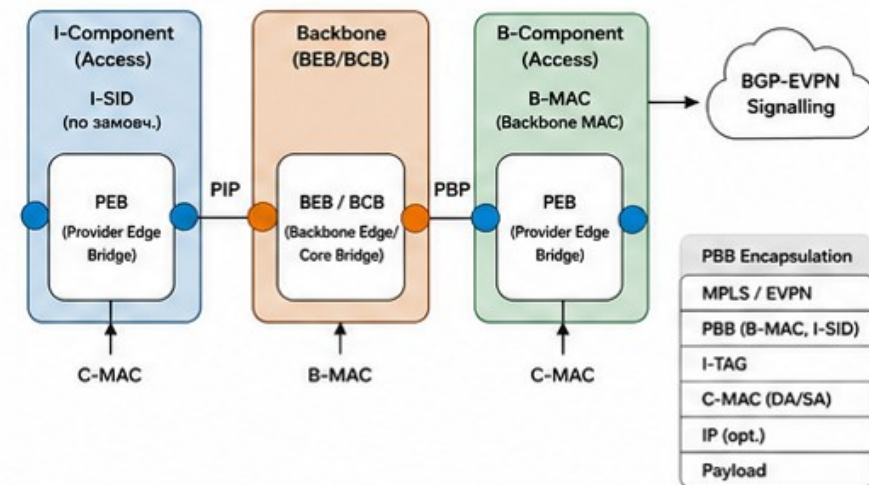
Ідея EVPN-PBB



2. Архітектура EVPN-PBB

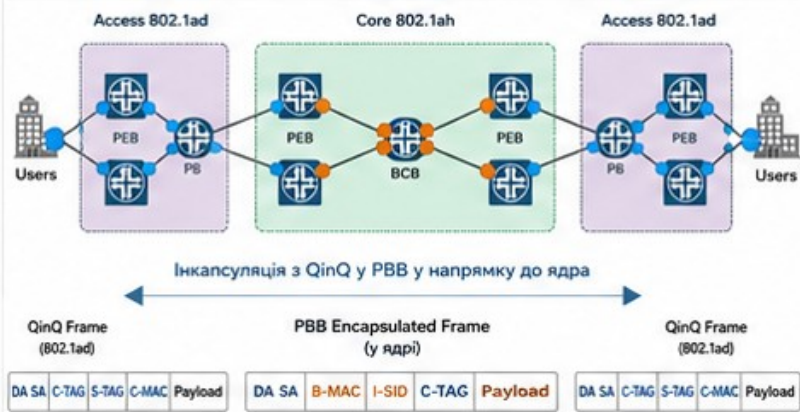


3. EVPN-PBB компоненти

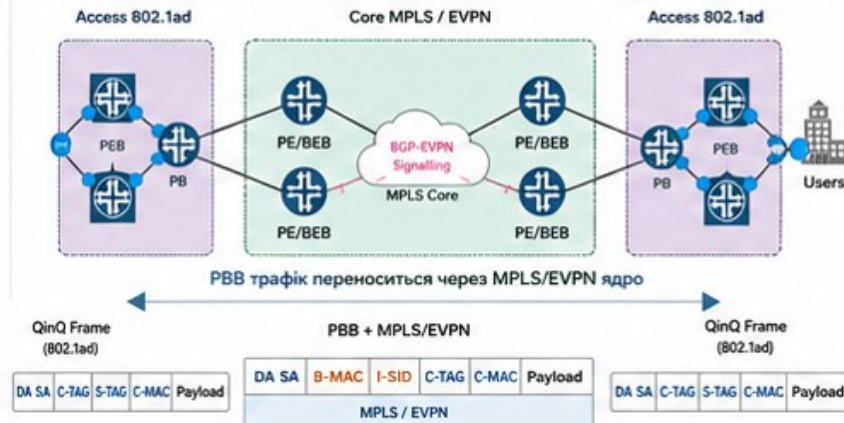


EVPN PBB (Provider Backbone Bridge)

4. Frame Flow – QinQ (802.1ad) до PBB



5. Frame Flow – PBB через MPLS / EVPN



6. Переваги EVPN-PBB

- ✓ Масштабування: один B-MAC замість тисяч/мільйонів C-MAC у CP
- ✓ Менше EVPN Type-2 маршрутів
- ✓ Краща продуктивність Control Plane
- ✓ Швидша конвергенція
- ✓ Підтримка великих Layer-2 доменів через WAN / MPLS
- ✓ Сумісність з EVPN, MPLS, L3VPN сервісами



Коли використовувати

DCI між дата-центрами через MPLS / EVPN

Metro Ethernet та Wholesale сервіси

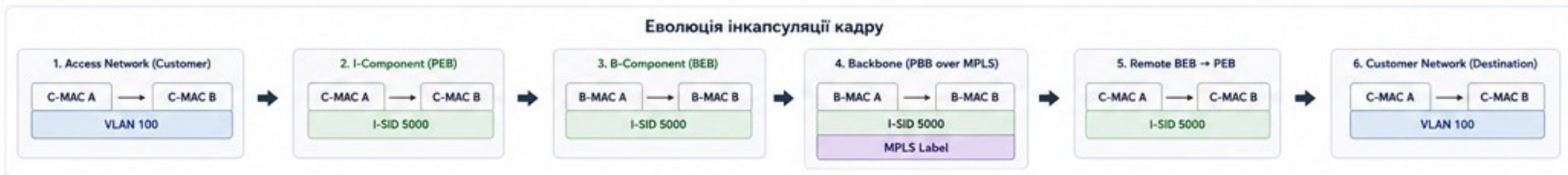
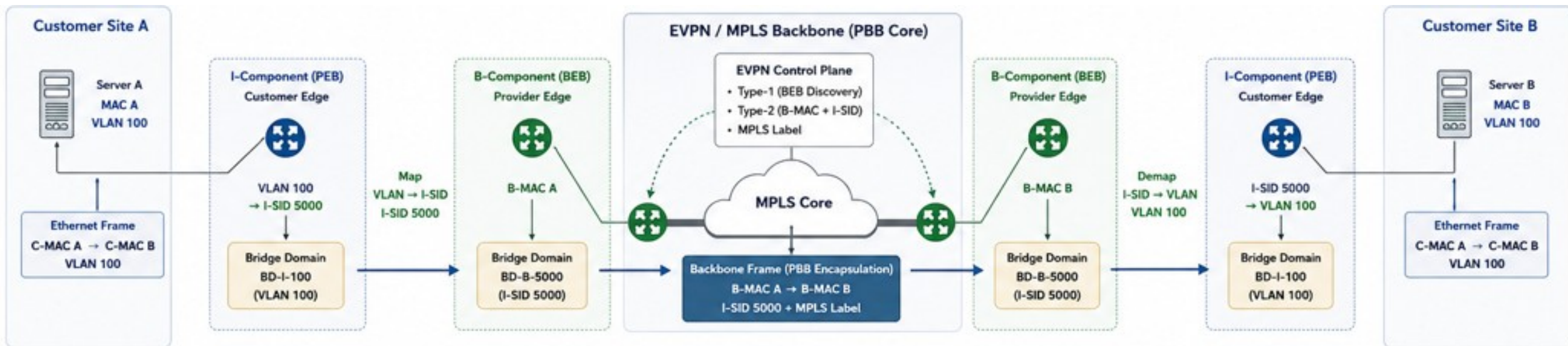
Коли потрібно масштабувати велику кількість MAC-адрес

Коли важлива швидка конвергенція та стабільність Control Plane

Ключові терміни

- PEB (Provider Edge Bridge) – на межі клієнтської мережі
- BEB (Backbone Edge Bridge) – на краю провайдерського ядра
- BCB (Backbone Core Bridge) – у середині ядра
- B-MAC – Backbone MAC (агрегує C-MAC)
- I-SID – Identifier для I-Component (домен клієнта)
- PIP / PBP – Provider Instance Ports (службові порти PBB)

EVPN PBB (Provider Backbone Bridge)



EVPN PBB (Provider Backbone Bridge)

```
routing-instances {
  DC2-PBB-B-COMP {
    instance-type virtual-switch;
    interface cbp0.999;

    route-distinguisher 62.0.0.2:999;
    vrf-target target:65500:999;

    protocols {
      evpn {
        control-word;
        pbb-evpn-core;
        extended-isid-list 999;
      }
    }

    bridge-domains {
      BD-B-COMP-999 {
        vlan-id 999;
        isid-list 999;
        vlan-id-scope-local;
      }
    }
  }

  DC2-PBB-I-COMP {
    instance-type virtual-switch;
    interface pip0.999;

    bridge-domains {
      BD-I-COMP-999 {
        vlan-id 999;
        interface ge-0/0/1.999;
      }
    }

    pbb-options {
      peer-instance DC2-PBB-B-COMP;
    }

    service-groups {
      CUSTOMER2 {
        service-type elan;
        pbb-service-options {
          isid 999;
          vlan-id-list 999;
        }
      }
    }
  }
}

interfaces {
  cbp0 {
    unit 999 {
      family bridge {
        interface-mode trunk;
        bridge-domain-type bvlan;
        isid-list all;
      }
    }
  }

  pip0 {
    unit 999 {
      family bridge {
        interface-mode trunk;
        bridge-domain-type svlan;
        isid-list all-service-groups;
      }
    }
  }

  ge-0/0/1 {
    unit 999 {
      encapsulation vlan-bridge;
      vlan-id 999;
      family bridge;
    }
  }
}
```

Links & Literature

Design DC

- <https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/concept/evpn-vxlan-implementing.html>
- <https://www.juniper.net/documentation/us/en/software/nce/sg-005-data-center-fabric/topics/concept/solution-cloud-data-center-components.html>
- <https://www.juniper.net/documentation/us/en/software/nce/sg-005-data-center-fabric/topics/concept/solution-cloud-data-center-topology-overview.html#hardware-summary>

DCI

- https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/concept/data-center-interconnect-evpn-vxlan-evpn-mpls-wan-overview.html#evpn-vxlan-data-center-interconnect-through-evpn-mpls-wan-overview_d1897e255
- <https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/concept/pbb-evpn-integration-for-dci-overview.html>

QinQ

- <https://www.juniper.net/documentation/us/en/software/junos/multicast-l2/topics/topic-map/q-in-q.html#id-understanding-qinq-tunneling-and-vlan-translation>
- <https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/topic-map/evpn-vxlan-flexible-vlan-tag.html#id-overview-and-topology>
- https://www.juniper.net/documentation/us/en/software/nce/sg-005-data-center-fabric/topics/topic-map/static-vxlan-q-in-q.html#task_ctq_rwh_qbc

L3 over VXLAN

- <https://www.juniper.net/documentation/us/en/software/nce/sg-005-data-center-fabric/topics/task/edge-routed-overlay-cloud-dc-configuring.html#id-configuring-an-edgerouted-bridging-overlay-on-a-leaf-device>
- <https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/topic-map/evpn-sym-t2-routing-ov.html>
- <https://www.juniper.net/documentation/us/en/software/junos/evpn/topics/example/evpn-vxlan-qfx5110-l3-vxlan-gateway-collapsed.html#basic-customer-profile-configuration683>



Q&A

ДОЦІЛЬНА ІНТЕГРАЦІЯ МОЖЛИВОСТЕЙ

м. Київ, вул. Бучанська, буд. 25/28

+38 044 597 10 90

itbiz.ua



ITBIZ telecom