

The logo for NIC.br features the text 'nic.br' in a bold, lowercase sans-serif font. The '.br' is a vibrant green color, while 'nic' is black. Below the logo, the full name 'Brazilian Network Information Center' is written in a smaller, black, uppercase sans-serif font.

**nic.br**  
Brazilian Network  
Information Center

The logo for EGI.br features the text 'egi.br' in a bold, lowercase sans-serif font. The '.br' is a vibrant green color, while 'egi' is black. Below the logo, the full name 'Brazilian Internet Steering Committee' is written in a smaller, black, uppercase sans-serif font.

**egi.br**  
Brazilian Internet  
Steering Committee

A horizontal list of six .br domain names is displayed in a white, lowercase sans-serif font. The '.br' part of each domain is highlighted in green. The domains are: registro.br, cert.br, cetic.br, ceptro.br, ceweb.br, and ix.br.

registro.br cert.br cetic.br ceptro.br ceweb.br ix.br

# Evoluindo redes layer2 com EVPN e SRv6

Fabio Pessoa Nunes  
<fpnunes@nic.br>  
IX.br Engineering

nic.br

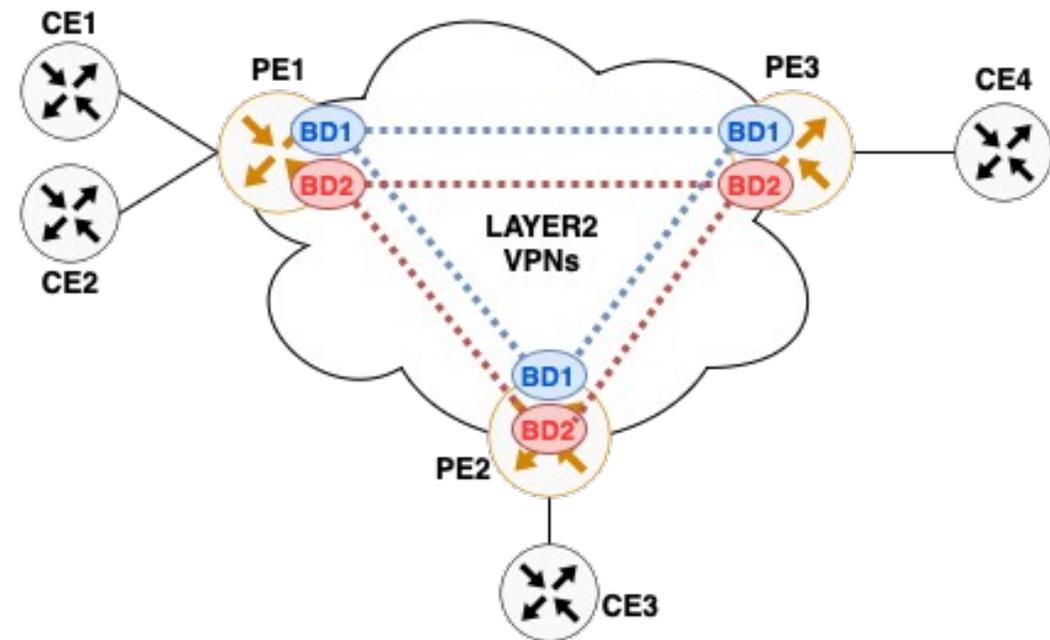
# Agenda

- Importância do L2VPN
- Funcionamento do VPLS
- O que é EVPN?
- Proxy-ARP / Proxy-ND para EVPN
- Migração de VPLS para EVPN
- O que é o Segment Routing?
- IX.br utilizando SRv6 EVPN
- Conclusões



# Importância do L2VPN

- Permite a criação de uma rede virtual privada de camada 2
- Facilidade de expansão da rede layer 2 para pontos remotos
- Melhor convergência de caminhos do que em redes L2 tradicionais
- Facilidade de configuração de novos serviços
- Formas mais comuns de L2VPN:
  - Virtual Private LAN Service (VPLS)
  - Virtual Private Wire Service (VPWS)
  - Ethernet VPN (EVPN) se tornando cada vez mais comum

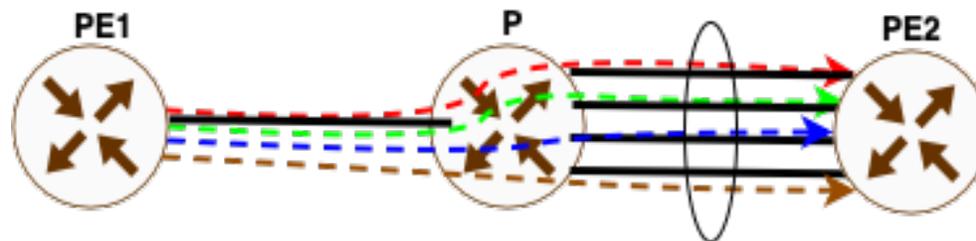


# Funcionamento do VPLS

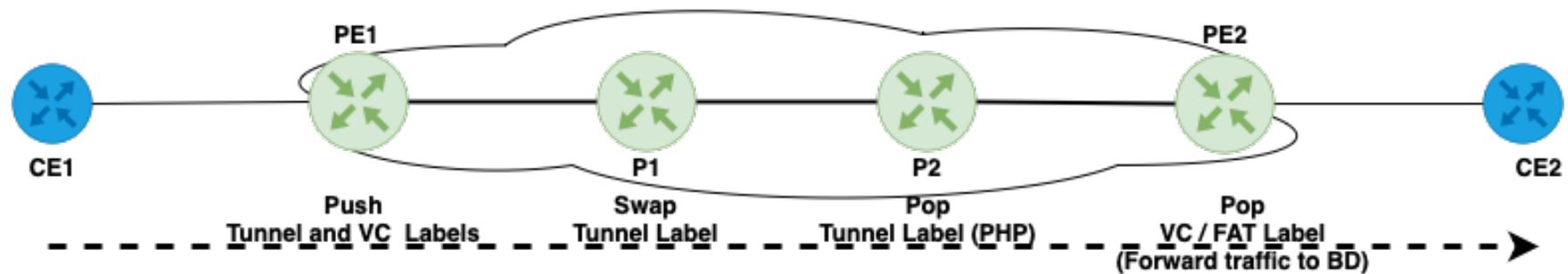
- Emula um switch Ethernet sobre de uma rede MPLS / comutada por labels
- Cria de full-mesh de circuitos virtuais (pseudowires) entre os PEs participantes de um Bridge-domain
- Aprendizado de MAC dinâmico baseado no MAC de origem sendo encaminhado pela rede (plano de encaminhamento)
- Duas formas de sinalização dos pseudowires
  - RFC 4762 (LDP-based VPLS)
  - RFC 4761 (BGP-based VPLS)
- Protocolos envolvidos:
  - IGP: OSPF ou ISIS
  - Comutação de labels de transporte: LDP, RSVP ou Segment Routing MPLS
  - Troca de labels dos Psseudowires / Virtual Circuits (VC): LDP ou BGP

# Funcionamento do VPLS

- Devido ao uso do MPLS, pode exigir técnicas para melhor balanceamento do tráfego nos roteadores do caminho
- **Flow-aware Transport (FAT) label:** Um label é adicionado a pilha pelo ingress PE indicando o flow ao qual aquele frame pertence. Dessa forma, os roteadores no caminho poderão utilizar esse label para balancear o tráfego em um ECMP ou LACP, e tráfego do mesmo flow irá percorrer o mesmo caminho
- **Control-word:** Usado principalmente para indicar que o payload é um frame Ethernet e evitar que payload ethernet com MAC começado em 4 ou 6 seja interpretado como IPv4 ou IPv6 por roteadores P do caminho, causando um balanceamento inadequado do tráfego



# Encaminhamento de tráfego no VPLS



Sem uso de Flow-aware transport (FAT) label e Control Word

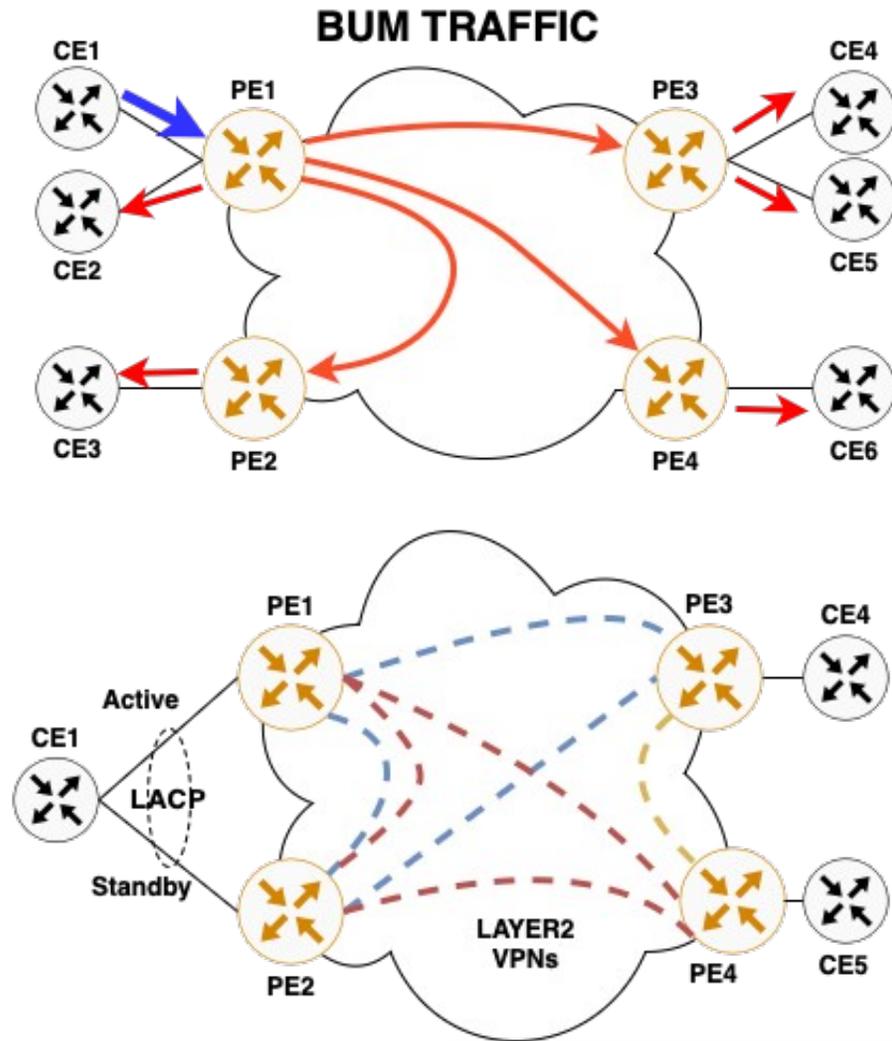
	Tunnel Label: 34	Tunnel Label: 45	Tunnel Label: 40		
	VC Label: 28	VC Label: 28	VC Label: 28	VC Label: 28	
Ethernet	Ethernet	Ethernet	Ethernet	Ethernet	Ethernet
IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP
Data	Data	Data	Data	Data	Data

Com uso de Flow-aware transport (FAT) label e Control Word

	Tunnel Label: 34	Tunnel Label: 45	Tunnel Label: 40		
	VC Label: 28	VC Label: 28	VC Label: 28	VC Label: 28	
	FAT Label: 126	FAT Label: 126	FAT Label: 126	FAT Label: 126	
	Control Word	Control Word	Control Word	Control Word	
Ethernet	Ethernet	Ethernet	Ethernet	Ethernet	Ethernet
IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP	IP/TCP/UDP
Data	Data	Data	Data	Data	Data

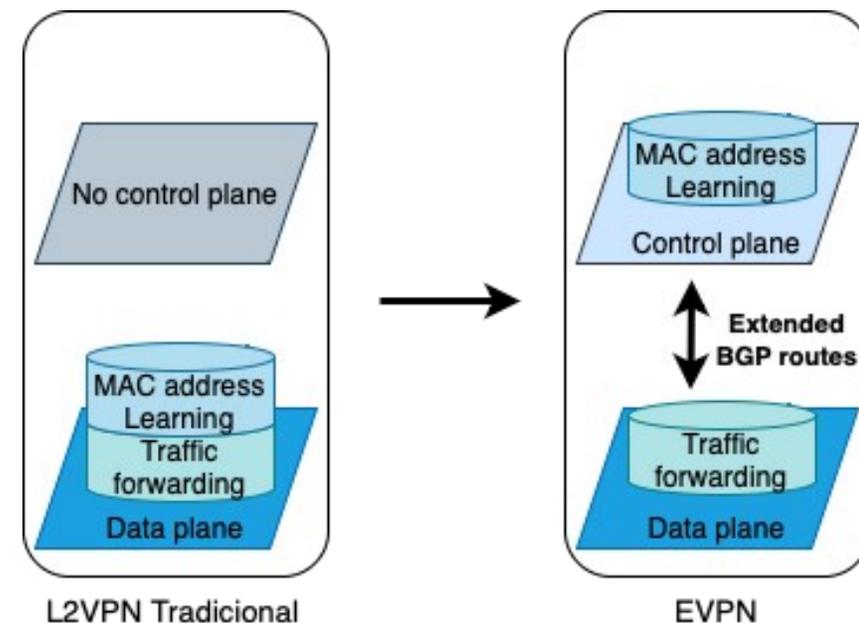
# Limitações do VPLS

- Não é escalável para aplicações que requerem grande quantidade de MAC address e rápida convergência após falhas
- Não há mecanismos para minimizar flood de Broadcast, Unknown-Unicast e Multicast (BUM)
- Não há suporte multi homing all-active no LACP, apenas single-active
- VPLS suporta auto-discovery baseado em BGP, mas ainda requer um grande quantidade de configuração do operador
- Pouca flexibilidade de topologia do L2VPN além do H-VPLS



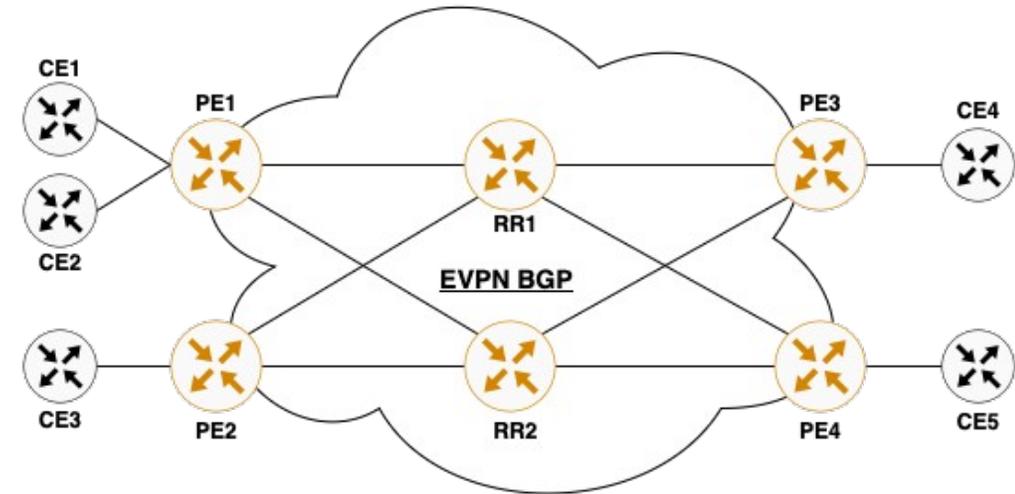
# O que é EVPN?

- Em maio/2014, a RFC7209 definiu os requisitos para o Ethernet VPN (EVPN) endereçar as limitações do VPLS
- Em fevereiro/2015, a RFC7432 (BGP MPLS-Based Ethernet VPN) especificou a implementação da tecnologia se baseando em redes MPLS
- Utiliza extensões no BGP para transmitir informações de serviços de VPN layer 2 e layer 3
- Separação dos planos de controle e encaminhamento:
  - MP-BGP é utilizado como plano de controle, criando os serviços e compartilhando todos os dados necessários entre os PEs
  - Inicialmente o plano de encaminhamento definido foi o MPLS, mas devido a extensibilidade do BGP foi facilmente especificado novas extensões para permitir outros tipos de encaminhamento (VxLAN, SR, SRv6, etc...)



# Vantagens do EVPN

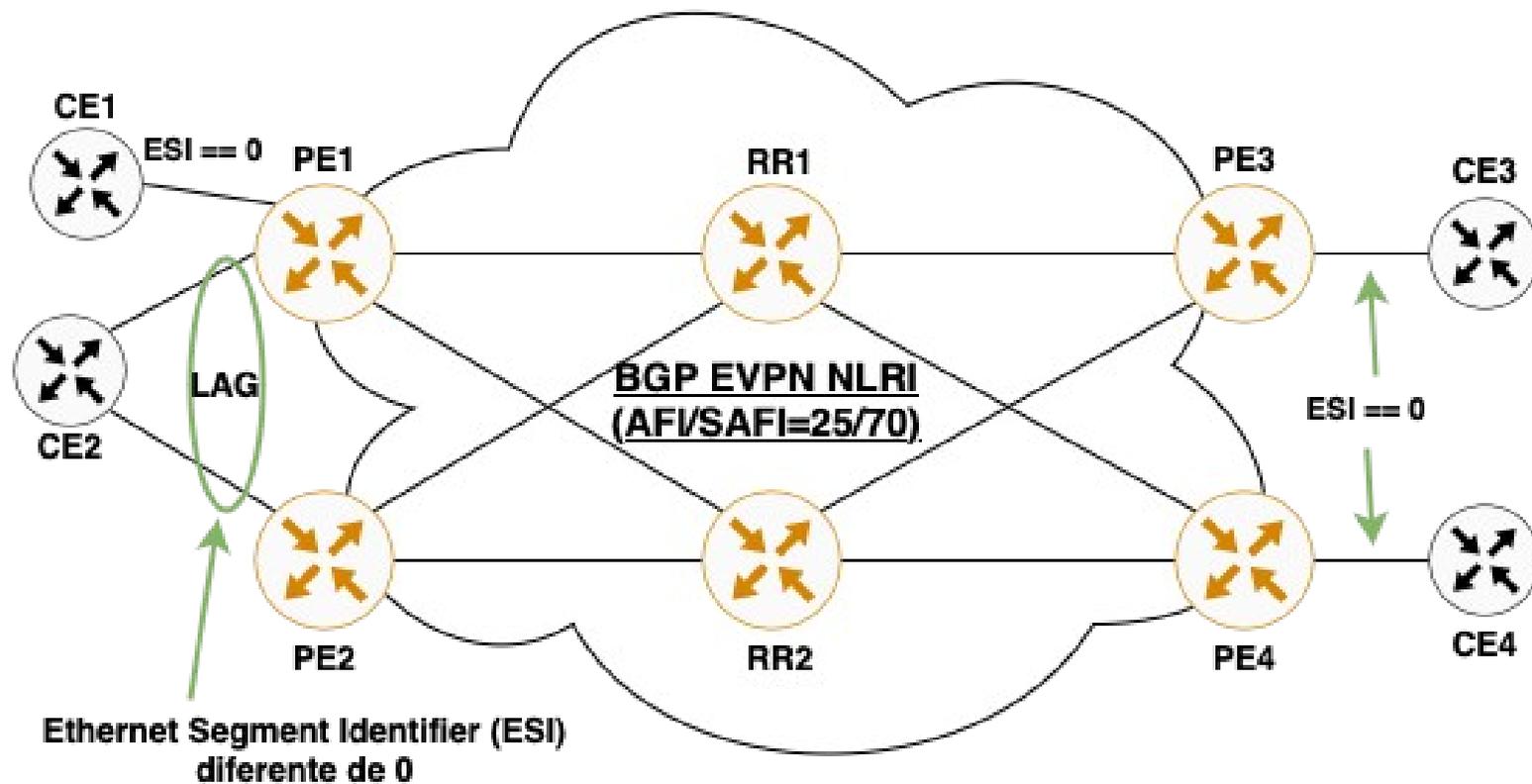
- O uso do BGP permite que o EVPN escale para suportar grande número de serviços / domínios e grande quantidade de MAC / IPs
- Melhora a visibilidade dos serviços da rede, facilitando principalmente automação e troubleshooting
- Rápida convergência dos serviços em casos de falhas ou MAC Move
- Melhor tratamento de tráfego BUM



# EVPN NLRI Route Types

- Utiliza extensões do MP-BGP, tendo o Network Layer Reachability Information (NLRI) com Address Family Identifier (AFI) 25 (L2VPN) e Subsequent Address Family Identifier (SAFI) 70 (EVPN)
- Foram definidos BGP EVPN Route Types que possibilitaram a criação dos serviços e a troca de informações como MAC e IP:
  - **Type 1 – Ethernet auto-discovery**: Anuncia os segmentos ethernet multi homing para permitir rápida convergência em casos de alteração do device conectado e também para permitir balanceamento de carga entre os PEs conectados ao CE multi homed
  - **Type 2 – MAC/IP Advertisement** : Anuncia MAC e IPs aprendidos pelo PE, permitindo que todos os PEs tenham sempre a mesma tabela de encaminhamento. Possibilita a eliminar flood de ARP, ND e Unknown unicast
  - **Type 3 – Inclusive Multicast**: Anuncia como o PE espera receber tráfego Broadcast, Multicast e Unknown unicast (BUM)
  - **Type 4 – Ethernet segment**: Permite que os PEs saibam quando estão conectados ao mesmo CE fazendo multi homing. É usado principalmente para a eleição do designated forwarder (DF), que será o PE responsável por encaminhar o tráfego BUM para o CE
  - **Type 5 – IP Prefix**: Permite um serviço EVPN acessar redes externas, anunciando rotas externas importadas como prefixos IP
- Novos tipos para criação de novos serviços

# Arquitetura do EVPN

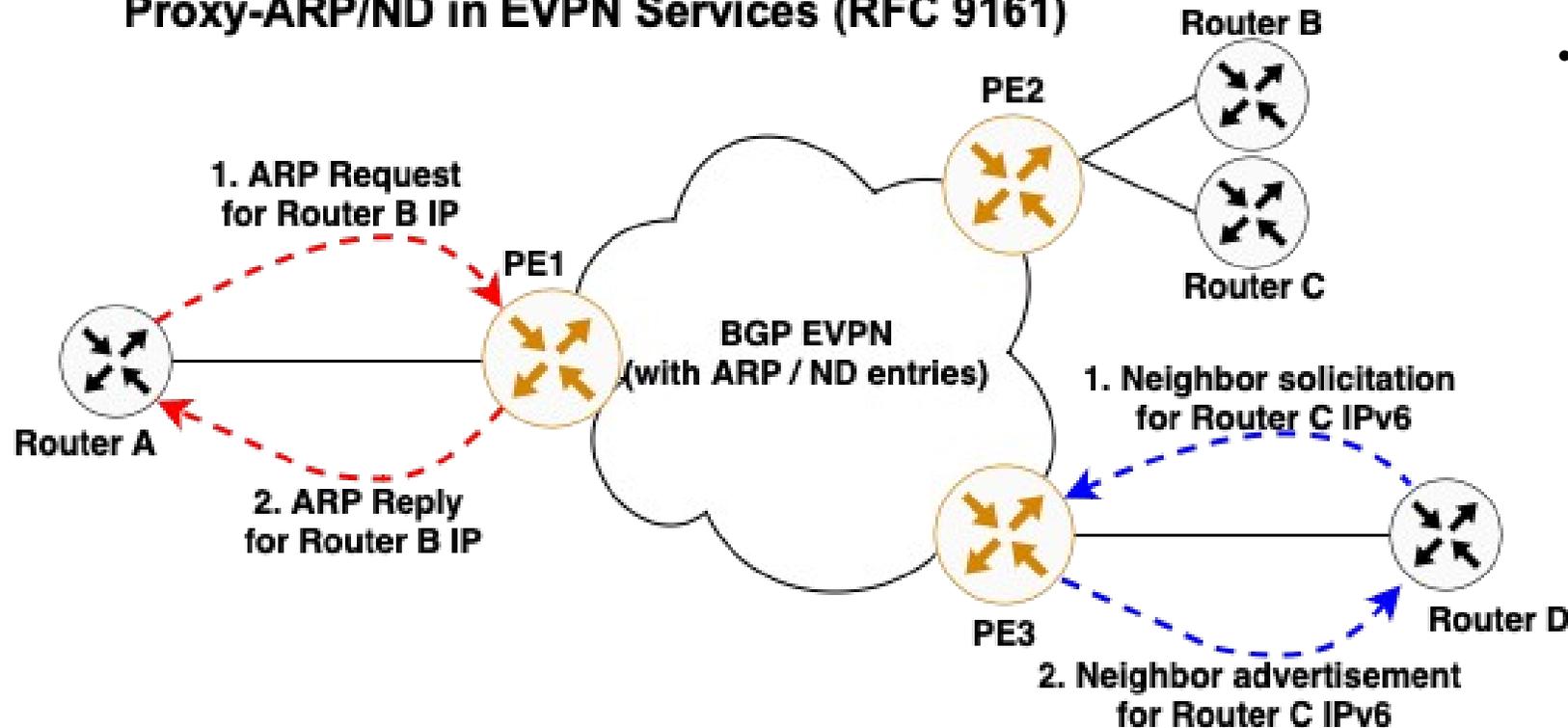


# Optimização do tráfego BUM

- Tráfego BUM pode afetar negativamente os hosts conectados a rede:
  - ARP/Broadcast: É processado diretamente pela CPU dos roteadores e pode afetar performance do equipamento
  - Unknown-unicast: Grande volume pode consumir desnecessariamente banda entre os PEs e os CEs
- EVPN pode carregar as entradas ARP e ND das VPNs, permitindo que os PEs respondam localmente as requisições ARP/ND, evitando que todos CEs recebam desnecessariamente todas as requisições
- Com a separação do plano de controle, existe a garantia que todos os PEs possuem a mesma informação dos MACs presentes na rede. Com isso, é possível desabilitar o flood de unknown-unicast

# EVPN Proxy-ARP / Proxy-ND

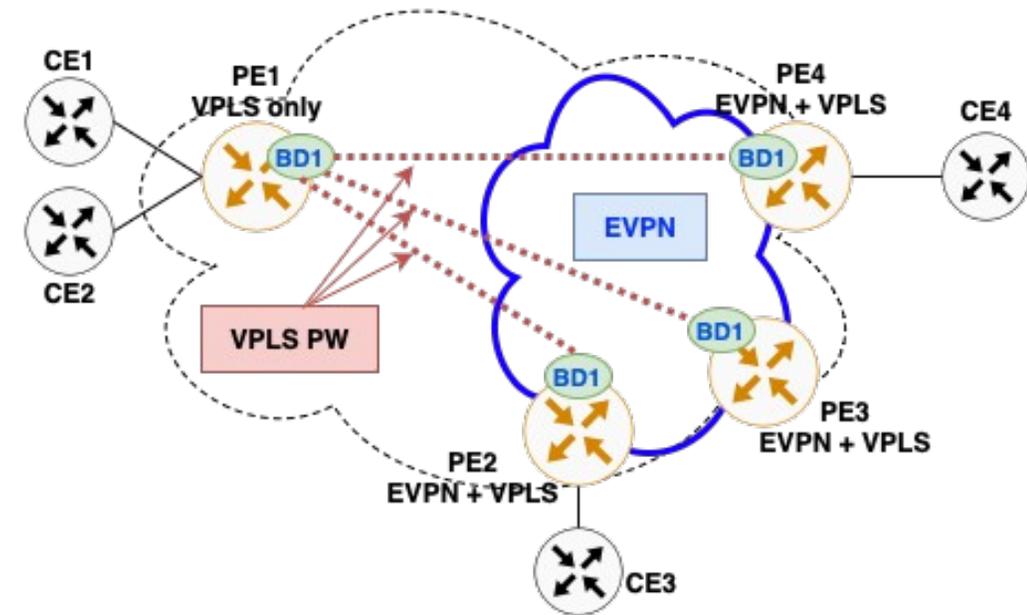
## Proxy-ARP/ND in EVPN Services (RFC 9161)



- Redução / Eliminação de ARP/ND
- ARP/ND estático anunciado com a "ARP/ND Extended Community with Immutable flag (RFC9047)
- Garante que em toda a rede as entradas estáticas de ARP/ND prevaleçam sobre as entradas dinâmicas
- Previne ARP/ND Spoofing
- Garante que o cliente não vai responder ARP/ND para IPs que não lhe pertencem

# Migração do VPLS para EVPN

- Migração transparente com “VPLS and EVPN seamless integration” (RFC 8560)
- Migração incremental, device por device
  - Configuração do VPLS é mantida
  - Configurando o EVPN no mesmo bridge-domain, o roteador automaticamente desabilita o PW VPLS e inicia o encaminhamento utilizando o EVPN
- Se o roteador não suporta EVPN, a comunicação se mantém com todos os peers usando os PWs VPLS já configurados



# EVPN MPLS – Protocolos envolvidos

- Control Plane:
  - BGP
- Data Plane:
  - IGP (ISIS ou OSPF)
  - Comutação de labels de transporte: LDP, RSVP ou SR-MPLS

# O que é Segment Routing?

- Source Routing: Origem define o caminho a ser seguido na rede e representa no pacote como uma lista de segmentos
- **Segmento**: identificador para uma instrução a ser executada na rede (encaminhamento ou serviço)
  - **SR-MPLS**: Lista de segmentos representada como pilha de labels MPLS distribuídos pelo IGP
  - **SRv6**: Lista de segmentos representada no cabeçalho do IPv6
- Benefícios do Segment Routing:
  - Simplificação da rede: Não há necessidade de protocolos adicionais como LDP ou RSVP
  - Fácil implementação de engenharia de tráfego
  - TI-LFA (MPLS tradicional não prove nativamente caminho “loop free-alternate” (LFA) )
  - Evita microloops

# Vantagens do Segment Routing IPv6



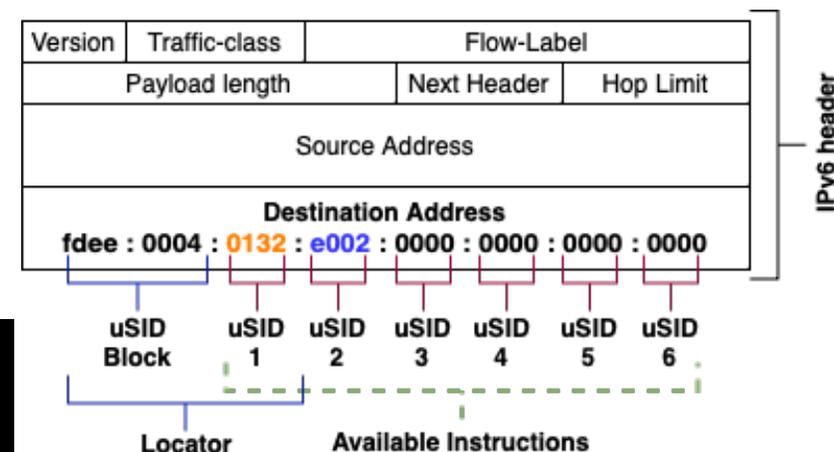
- Combina as vantagens do SR-MPLS com a extensibilidade do puro e simples encaminhamento IPv6
- Equipamentos de core precisam apenas encaminhar IPv6, eles não necessariamente precisam entender SRv6
- O uso de MPLS requer o uso do Flow Aware Transport (FAT) ou inspeção do payload para fazer o load balance
  - SRv6 resolve isso usando o campo Flow-label no cabeçalho IPv6
- Devido a extensibilidade do IPv6, é mais fácil implementar novas features no SRv6
- Alinhado com a transição para o IPv6
- Em resumo, permite uma implementação gradual da tecnologia sem a necessidade de migração de todos os devices no dia zero.

# Como funciona o SRv6 com uSID

As instruções para a rede são representadas por partes do endereço de destino IPv6

- Define-se um prefixo IPv6 de 48 bits chamado “locator,” que representa o device para um grupo de serviços, e se anuncia no IGP indicando a função do SID

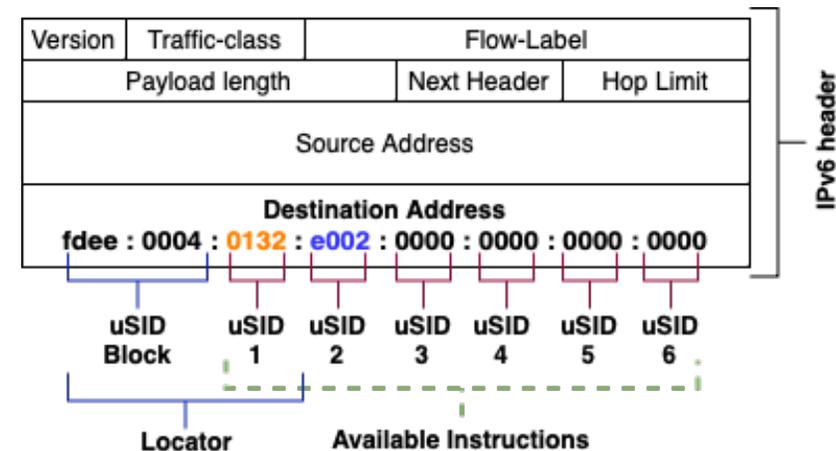
```
SRv6 Locator :  
MT ID : 0  
Metric: ( ) 1 Algo:0  
Prefix : fdee:4:132::/48  
Sub TLV :  
AttrFlags:  
End-SID : fdee:4:132:: flags:0x0, endpoint:End-NXT-CSID-PSP-USD, lbLen:32, lnLen:16, funLen:0, argLen:80
```



- Ao encaminhar um pacote utilizando SRv6 pode-se adicionar uma série de instruções no endereço IPv6, sendo que a instrução ativa sempre será a do uSID1

# Como funciona o SRv6 com uSID

- Serviços e instruções são definidos baseados no Locator e as instruções que requeiram cada um dos serviços são colocadas nos próximos uSID
- No caso do SRv6 EVPN, o BGP anuncia para os peers os uSIDs que referentes a cada um dos serviços e qual ação o PE vai executar ao receber um pacote com aquela instrução
- Se houver necessidade de mais de 6 instruções, usa-se a extensão do cabeçalho IPv6 chamada Segment Routing Header.



# EVPN + SRv6

RFC 9252 - BGP Overlay Services Based on Segment Routing over IPv6 (SRv6)

```
RP/0/RP0/CPU0:lab-pe-132#sh evpn evi vpn-id 10 mac local
Tue Dec 10 17:52:50.205 -03
```

VPN-ID	Encap	MAC address	IP address	Nextthop	Label	SID
10	SRv6	0011.2233.4420 ::		Bundle-Ether69.32	0	fdee:4:132:e002::

```
RP/0/RP0/CPU0:lab-pe-132#sh evpn evi vpn-id 10 encap srv6 detail
Tue Dec 10 17:50:13.748 -03
```

VPN-ID	Encap	Bridge Domain	Type
10	SRv6	sp	EVPN

Stitching: Regular  
Unicast SID: fdee:4:132:e002::  
Multicast SID: fdee:4:132:e003::  
E-Tree: Root  
Forward-class: 0  
Advertise MACs: Yes  
Advertise BVI MACs: No  
Aliasing: Enabled  
UUF: Enabled  
Re-origination: Enabled  
Multicast:  
IGMP-Snooping Proxy: No  
MLD-Snooping Proxy : No  
BGP Implicit Import: Enabled  
VRF Name:  
SRv6 Locator Name: EVPN  
SRv6 SID Function Length: 16 bits  
Preferred Nextthop Mode: Off  
BVI Coupled Mode: No  
BVI Subnet Withheld: ipv4 No, ipv6 No  
L3VRF Label Mode: Per-VRF  
RD Config: none  
RD Auto : (auto) 172.20.0.132:10  
RT Auto : 65000:10

Route Targets in Use	Type
65000:10	Import
65000:10	Export

# EVPN + SRv6 Type 2 Route

## Path attributes

### Path Attribute - MP\_REACH\_NLRI

- Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete

- Type Code: MP\_REACH\_NLRI (14)

- Length: 44

- Address family identifier (AFI): Layer-2 VPN (25)

- Subsequent address family identifier (SAFI): EVPN (70)

- Next hop: 172.20.0.132

- Number of Subnetwork points of attachment (SNPA): 0

### Network Layer Reachability Information (NLRI)

#### EVPN NLRI: MAC Advertisement Route

- Route Type: MAC Advertisement Route (2)

- Length: 33

- Route Distinguisher: 0001ac140084000a (172.20.0.132:10)

- ESI: 00:00:00:00:00:00:00:00:00

- Ethernet Tag ID: 0

- MAC Address Length: 48

- MAC Address: CIMSYS\_33:44:20 (00:11:22:33:44:20)

- IP Address Length: 0

- IP Address: NOT INCLUDED

- 1110 0000 0000 0010 0000 .... = MPLS Label 1: 917536

- Path Attribute - ORIGIN: IGP

- Path Attribute - AS\_PATH: empty

- Path Attribute - LOCAL\_PREF: 100

- Path Attribute - EXTENDED\_COMMUNITIES

- Path Attribute - BGP Prefix-SID

- Frame 1: 515 bytes on wire (4120 bits)
- Ethernet II
- Internet Protocol Version 4
- Transmission Control Protocol
- Border Gateway Protocol - UPDATE Message
  - Marker: ff
  - Length: 152
  - Type: UPDATE Message (2)
  - Withdrawn Routes Length: 0
  - Total Path Attribute Length: 129
  - Path attributes
    - Path Attribute - MP\_REACH\_NLRI
    - Path Attribute - ORIGIN: IGP
    - Path Attribute - AS\_PATH: empty
    - Path Attribute - LOCAL\_PREF: 100
    - Path Attribute - EXTENDED\_COMMUNITIES
    - Path Attribute - BGP Prefix-SID
- Border Gateway Protocol - UPDATE Message
- Border Gateway Protocol - UPDATE Message
- Border Gateway Protocol - UPDATE Message

## Path Attribute - BGP Prefix-SID

- Flags: 0xc0, Optional, Transitive, Complete

- Type Code: BGP Prefix-SID (40)

- Length: 37

### SRv6 L2 Service

- Type: SRv6 L2 Service (6)

- Length: 34

- Reserved: 00

### SRv6 Service Sub-TLVs

#### SRv6 Service Sub-TLV - SRv6 SID Information

- Type: SRv6 SID Information (1)

- Length: 30

- Reserved: 00

- SRv6 SID Value: fdee:4:132::

- SRv6 SID Flags: 0x00

- SRv6 Endpoint Behavior: End.DT2U with NEXT-CSID (0x0043)

- Reserved: 00

### SRv6 Service Data Sub-Sub-TLVs

#### SRv6 Service Data Sub-Sub-TLV - SRv6 SID Structure

- Type: SRv6 SID Structure (1)

- Length: 6

- Locator Block Length: 32

- Locator Node Length: 16

- Function Length: 16

- Argument Length: 0

- Transposition Length: 16

- Transposition Offset: 48

# EVPN + SRv6 Type 2 Route

```
A:admin@IXR-X1-181# show service id 10 fdb detail
```

```
Forwarding Database, Service 10
```

ServId	MAC Transport:Tnl-Id	Source-Identifier	Type	Last Change
10	00:11:22:33:44:20 fdee:4:132:e002::	srv6-1: 172.20.0.132	Evpn	12/09/24 16:54:50
10	00:11:22:33:44:23	sdp:135:10010	LT/360	11/14/24 17:14:05
10	00:11:22:33:44:51	sap:lag-1:81	CStatic: P	11/06/24 14:26:14
10	00:11:22:33:44:87	sdp:135:10010	LT/1140	11/14/24 16:45:42
10	00:11:22:33:44:b7 fdee:3f:183:4001::	srv6-1: 172.20.0.183	Evpn	11/06/24 14:26:36
10	00:31:26:53:f6:cf fdee:3f:183:4001::	srv6-1: 172.20.0.183	EvpnS:P	11/06/24 14:26:36

```
No. of MAC Entries: 6
```

```
BGP EVPN MAC Routes
```

```
RIB In Entries
```

```

Network      : n/a
Nextthop    : 172.20.0.132
Path Id     : None
From        : 172.20.0.137
Res. Nextthop : 10.10.181.0
Local Pref. : 100
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community    : origin:172.20.0.132:10 ext:60e:20 target:65000:10
Cluster      : 172.20.0.137
Originator Id : 172.20.0.132
Origin       : IGP
Flags        : Used Valid Best
Route Source : Internal
AS-Path      : No As-Path
EVPN type    : MAC
ESI          : ESI-0
Tag          : 0
IP Address   : n/a
Route Dist.  : 172.20.0.132:10
Mac Address  : 00:11:22:33:44:20
MPLS Label1 : 57346
MPLS Label2 : n/a
Route Tag    : 0
Neighbor-AS  : n/a
DB Orig Val  : N/A
Add Paths Send : Default
Last Modified : 01d00h39m
SRv6 TLV Type : SRv6 L2 Service TLV (6)
SRv6 SubTLV  : SRv6 SID Information (1)
Sid          : fdee:4:132::
Full Sid     : fdee:4:132:e002::
Behavior     : End.uDT2U (67)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len : 32
Func-Len     : 16
Tpose-Len    : 16
Interface Name : to_SW137
Aggregator    : None
MED           : None
IGP Cost      : 21
Peer Router Id : 172.20.0.137
Final Orig Val : N/A
Loc-Node-Len : 16
Arg-Len      : 0
Tpose-offset  : 48
  
```

# EVPN + SRv6 Type-3 Route

- **Frame 1: 515 bytes on wire (4120 bits)**
- **Ethernet II**
- **Internet Protocol Version 4**
- **Transmission Control Protocol**
- **Border Gateway Protocol - UPDATE Message**
- **Border Gateway Protocol - UPDATE Message**
  - Marker: ffffffffffffffffffffffffffffffff
  - Length: 140
  - Type: UPDATE Message (2)
  - Withdrawn Routes Length: 0
  - Total Path Attribute Length: 117
- **Path attributes**
  - Path Attribute - MP\_REACH\_NLRI
  - Path Attribute - ORIGIN: IGP
  - Path Attribute - AS\_PATH: empty
  - Path Attribute - LOCAL\_PREF: 100
  - Path Attribute - EXTENDED\_COMMUNITIES
  - Path Attribute - PMSI\_TUNNEL\_ATTRIBUTE
  - Path Attribute - BGP Prefix-SID
- **Border Gateway Protocol - UPDATE Message**
- **Border Gateway Protocol - UPDATE Message**

```

Path Attribute - MP_REACH_NLRI
  Flags: 0x90, Optional, Extended-Length, Non-transitive, Complete
  Type Code: MP_REACH_NLRI (14)
  Length: 28
  Address family identifier (AFI): Layer-2 VPN (25)
  Subsequent address family identifier (SAFI): EVPN (70)
  Next hop: 172.20.0.132
  Number of Subnetwork points of attachment (SNPA): 0
  Network Layer Reachability Information (NLRI)
    EVPN NLRI: Inclusive Multicast Route
      Route Type: Inclusive Multicast Route (3)
      Length: 17
      Route Distinguisher: 0001ac140084000a (172.20.0.132:10)
      Ethernet Tag ID: 0
      IP Address Length: 32
      IPv4 address: 172.20.0.132
  
```

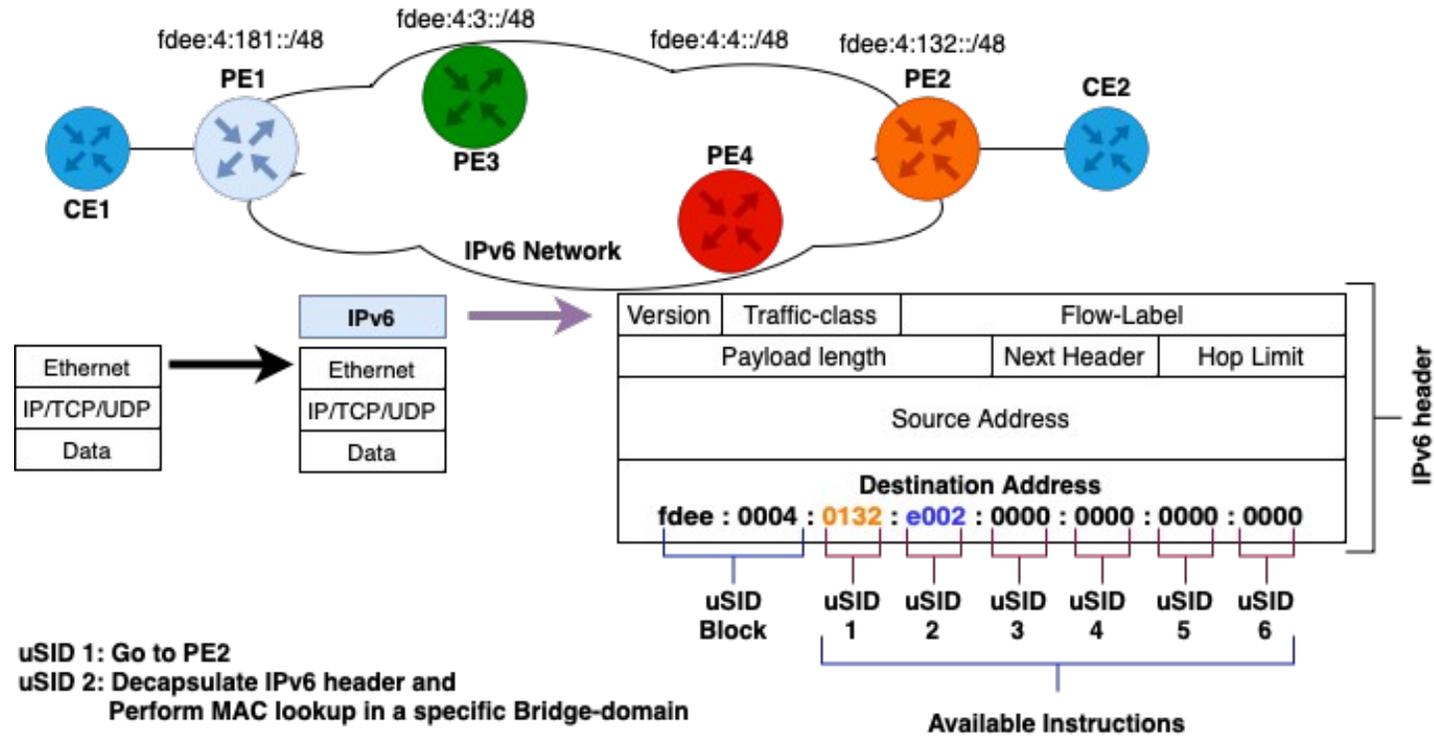
```

Path Attribute - PMSI_TUNNEL_ATTRIBUTE
  Flags: 0xc0, Optional, Transitive, Complete
  Type Code: PMSI_TUNNEL_ATTRIBUTE (22)
  Length: 9
  Flags: 0
  Tunnel Type: Ingress Replication (6)
  1110 0000 0000 0010 0000 .... = MPLS Label: 917536
  Tunnel ID: tunnel end point -> 172.20.0.132
  Path Attribute - BGP Prefix-SID
  Flags: 0xc0, Optional, Transitive, Complete
  Type Code: BGP Prefix-SID (40)
  Length: 37
  SRv6 L2 Service
  Type: SRv6 L2 Service (6)
  Length: 34
  Reserved: 00
  SRv6 Service Sub-TLVs
    SRv6 Service Sub-TLV - SRv6 SID Information
      Type: SRv6 SID Information (1)
      Length: 30
      Reserved: 00
      SRv6 SID Value: fdee:4:132::
      SRv6 SID Flags: 0x00
      SRv6 Endpoint Behavior: End.DT2M with NEXT-CSID (0x0044)
      Reserved: 00
      SRv6 Service Data Sub-Sub-TLVs
  
```

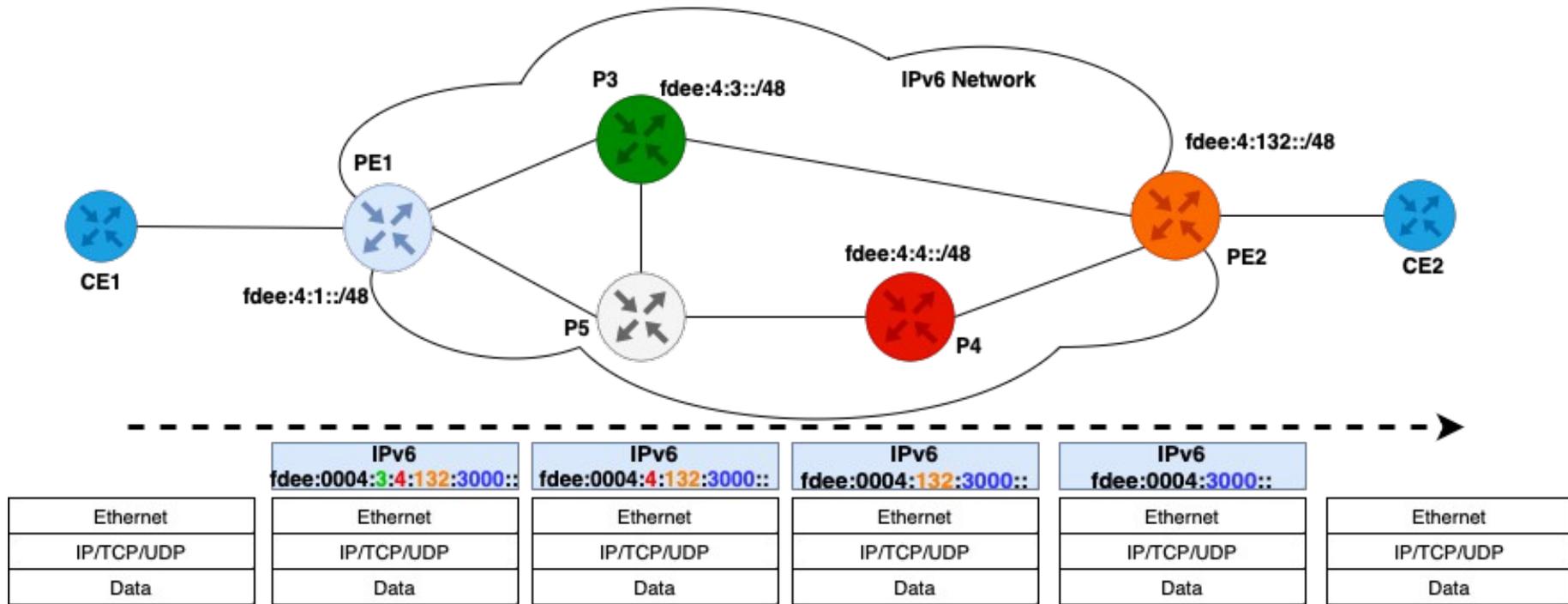
```

BGP EVPN Inclusive-Mcast Routes
-----
Original Attributes
Network      : n/a
Nexthop     : 172.20.0.132
Path Id     : None
From       : 172.20.0.137
Res. Nexthop : 10.10.181.0
Local Pref. : 100
Aggregator AS : None
Atomic Aggr. : Not Atomic
AIGP Metric  : None
Connector    : None
Community   : 12-attribute:MTU: 0 V: 0 M: 0 F: 0 C: 0 P: 0 B: 0
              target:65000:10
Cluster     : 172.20.0.137
Originator Id : 172.20.0.132
Peer Router Id : 172.20.0.137
Origin      : IGP
Flags       : Used Valid Best
Route Source : Internal
AS-Path     : No As-Path
EVPN type  : INCL-MCAST
Tag         : 0
Originator IP : 172.20.0.132
Route Dist. : 172.20.0.132:10
Route Tag   : 0
Neighbor-AS : n/a
DB Orig Val : N/A
Add Paths Send : Default
Last Modified : 01d04h45m
SRv6 TLV Type : SRv6 L2 Service TLV (6)
SRv6 SubTLV  : SRv6 SID Information (1)
Sid          : fdee:4:132::
Full Sid     : fdee:4:132:e003::
Behavior     : End.uDT2M (68)
SRv6 SubSubTLV : SRv6 SID Structure (1)
Loc-Block-Len : 32
Func-Len      : 16
Tpose-Len    : 16
Loc-Node-Len : 16
Arg-Len       : 16
Tpose-offset  : 48
-----
PMSI Tunnel Attributes :
Tunnel-type : Ingress Replication
Flags       : Type: RNVE(0) BM: 0 U: 0 Leaf: not required
MPLS Label  : 14680832
Tunnel-Endpoint: 172.20.0.132
  
```

# SRv6 datapath



# SRv6 Traffic Engineering



## PE1 Instructions:

- 1: Go to P3
- 2: Go to P4
- 3: Go to PE2
- 4: Decapsulate IPv6 header and  
Perform MAC lookup in a specific Bridge-domain

# EPVN SRv6 – Protocolos envolvidos

- Control Plane:
  - BGP
- Data Plane:
  - IGP (ISIS ou OSPF)

# O que o IX.br tem feito?

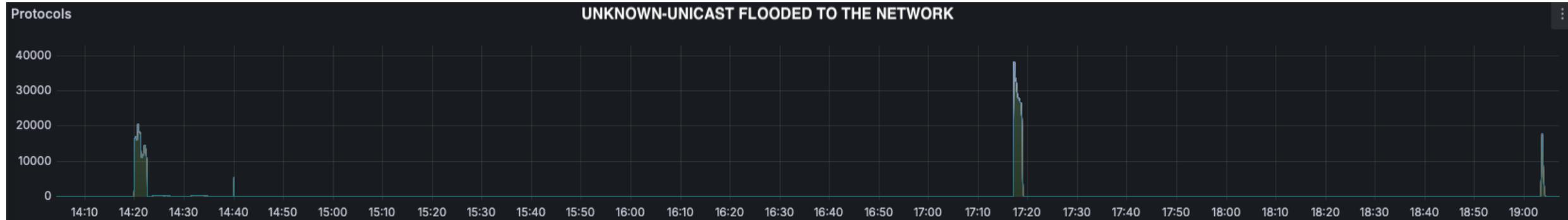
- Desde 2016 o IX.br tem trabalhado com diversos vendedores para ter uma solução EVPN suportando todas necessidades
  - Em 2023 testamos e aprovamos a solução de EVPN da Nokia
  - Em Fevereiro/2024 movemos um IXP (Brasília) completamente para EVPN usando MPLS como plano de encaminhamento
  - Em Junho/2024 o plano de encaminhamento do IX.br DF foi movido de MPLS para SRv6 uSID
  - Em Agosto/2024 foi habilitado o EVPN Proxy-ARP/Proxy-ND
- Solução muito mais simples que layer 2 tradicional ou VPLS
- Protection for Route Servers' IPs
- Prove uma experiência melhor para o participante
  - Menos ARP/ND para processar
  - Tráfego desnecessário não é encaminhado aos participantes
- Rápida recuperação e convergência em casos de falhas
- EVPN ELAN com ARP/ND proxy + SRv6 é requisito para todo novo equipamento adquirido pelo IX.br



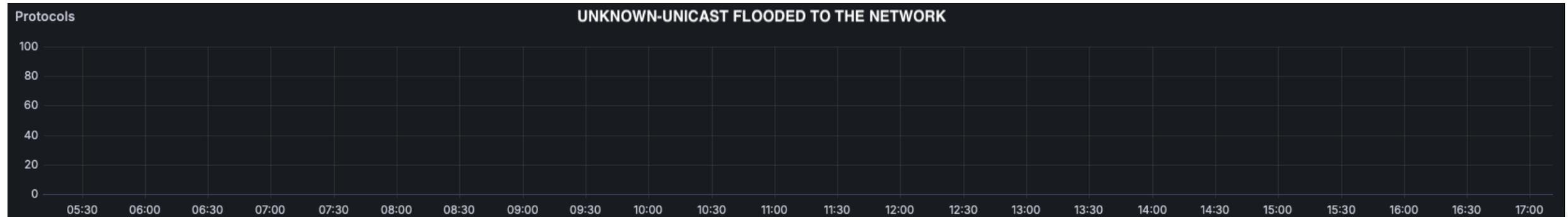
# IX.br Brasília - Unknown-unicast

- Flood de Unknown-unicast foi desabilitado em todos os devices
  - EVPN garante que todos os devices possuem a mesma tabela ARP

- Unknown-unicast flood enabled

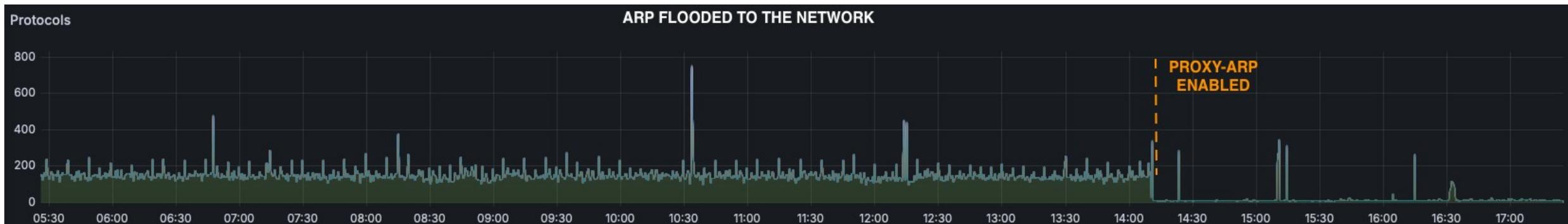


- Unknown-unicast flood disabled



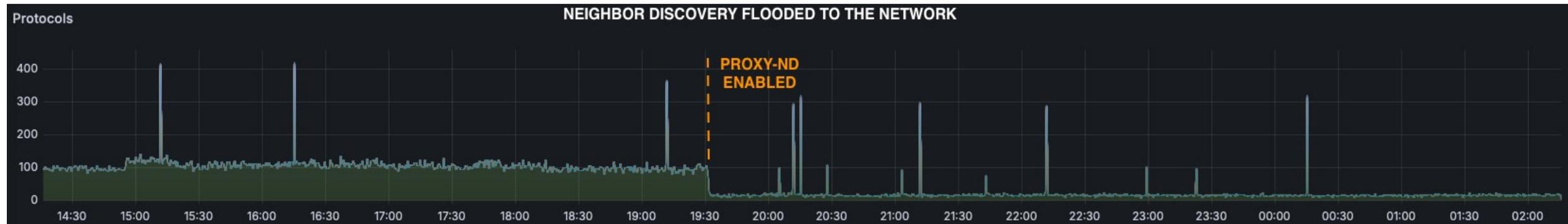
# IX.br Brasília - ARP

- EVPN Proxy-ARP reduziu significativamente a quantidade de ARP Request encaminhados para a rede
  - Entradas ARP desconhecidas pelo EVPN ainda são encaminhadas para a rede
  - Como temos a informação dos MACs/IPs dos participantes, próxima etapa será fornecer esses dados estaticamente para o EVPN e desabilitar por completo o flood de ARP da rede



# IX.br Brasília – IPv6 ND

- EVPN Proxy-ND reduziu significativamente a quantidade de ND Solicitation encaminhada para a rede
  - Entradas ND desconhecidas pelo EVPN ainda são encaminhadas para a rede
  - Como temos a informação dos MACs/IPs dos participantes, próxima etapa será fornecer esses dados estaticamente para o EVPN e desabilitar por completo o flood de ND da rede



# IX.br São Paulo – Principal objetivo

- Mais desafios na implementação:
  - Mais equipamentos e mais fabricantes
  - Infraestrutura crítica
  - Maior IXP em número de participantes e pico de tráfego, nenhum outro IXP possui esse cenário complexo
  - Apesar da extensa lista de testes realizados em laboratório, é impossível simular todos os desafios do IX.br São Paulo
  - Em todas tecnologias encontramos BUGs não triviais para os fabricantes entenderem
- Estamos em um processo de troca dos equipamentos existentes para suporte de SRv6 e EVPN com ARP/ND Proxy
- Já temos devices se comunicando exclusivamente através de SRv6 EVPN ao invés de VPLS (~10% dos MACs do ATM)
- Expectativa que no primeiro semestre/2025 a grande maioria do IX.br São Paulo esteja utilizando SRv6 EVPN

# Conclusões

- Redes layer 2 tradicionais não escalam facilmente
- EVPN é uma das melhores soluções para construir grandes redes layer 2 de forma escalável e facilmente gerenciável
- EVPN Proxy-ARP e Proxy-ND são ferramentas essenciais para grandes redes layer 2, principalmente IXPs
- Uso do SRv6 no plano de encaminhamento pode prover mais resiliência na rede com menos configuração e protocolos
- O SRv6 EVPN ajuda a construir redes preparadas um futuros com novas aplicações
- SRv6 EVPN está permitindo que o IX.br forneça serviços melhores e mais resilientes aos seus participantes

**Thanks**  
[www.ix.br](http://www.ix.br)

© fpnunes@nic.br

December 2024

**nic.br egi.br**  
[www.nic.br](http://www.nic.br) | [www.cgi.br](http://www.cgi.br)