



nic.br
Brazilian Network
Information Center

egi.br
Brazilian Internet
Steering Committee

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

nic.br cgi.br



APIX#31

Petaling Jaya, Malaysia | **02/23/2025**

IX.br Evolution to SRv6 EVPN

Julio Sirota (jsirota@nic.br)

Fábio Pessoa Nunes (fpnunes@nic.br)

ix.br nic.br cgi.br

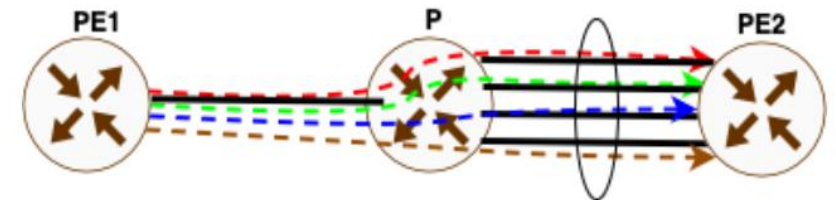
Agenda

- IXP challenges
- EVPN in na IXP network
- Main goal with EVPN
- IX.br São Paulo puzzle
- Segment Routing and SRv6
- What IX.br have already done
- Conclusions



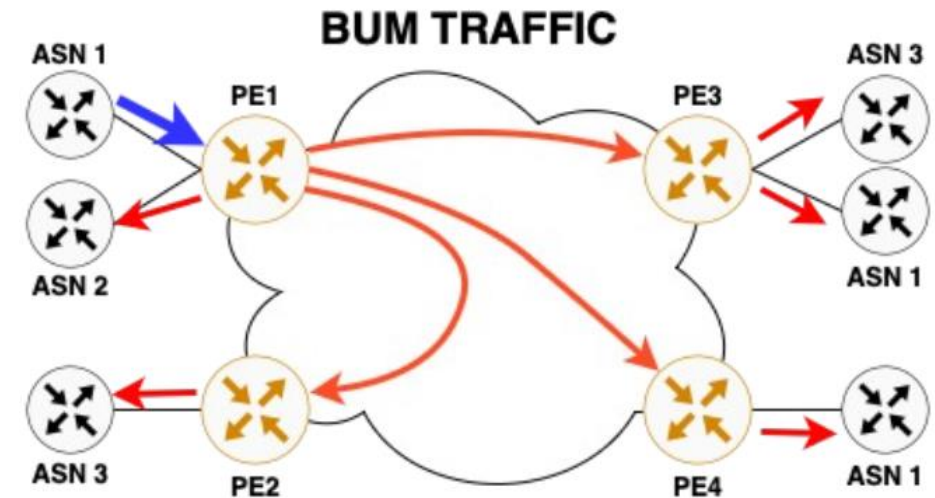
IXPs challenges

- IXP are mainly layer 2 Ethernet networks providing connectivity between routers
- Traditional layer 2 networks learn MAC in the data plane
 - Can't assure all devices have all same MACs in TCAM
- VPLS provides more resilience and intelligence however still relies on flood to learn MACs
- Need to make sure overlay traffic is properly balanced in underlay links and paths
 - Using overlay network based on MPLS, network devices must inspect the MPLS payload or rely on Entropy label / FAT to balance traffic
- IXP requires fast network convergence (paths and MAC addresses)
 - MPLS doesn't provide a native loop free-alternate (LFA) path
 - MPLS TE FRR requires na additional protocol and is complex to maintain



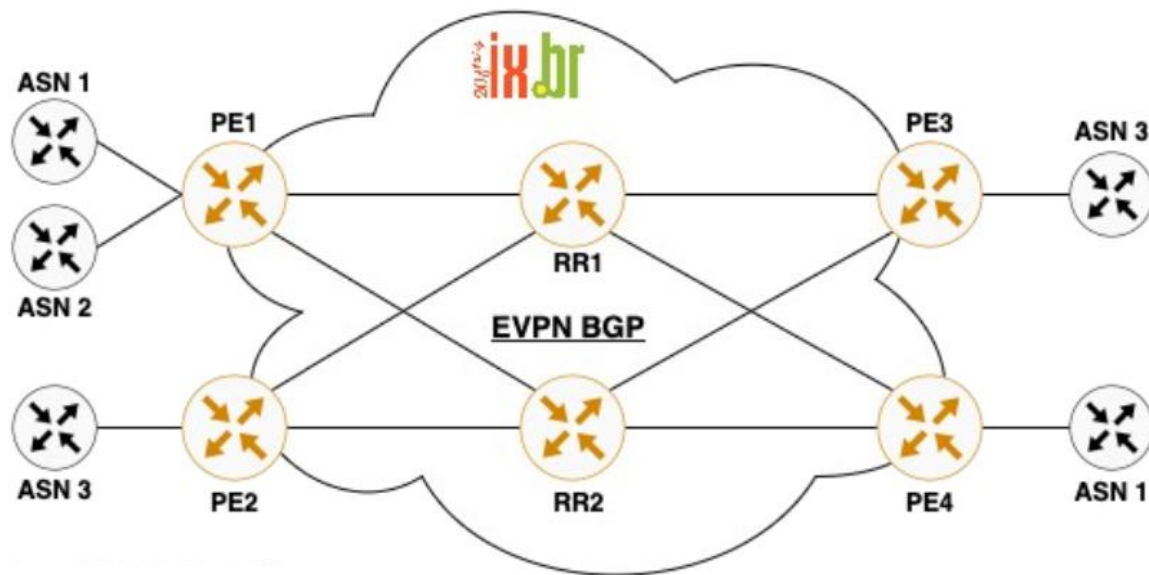
IXPs challenges

- BUM – **B**roadcast, **U**nknown-Unicast and **M**ulticast
 - ARP (ALL ARP requests are processed by ALL routers)
 - More participants, more routers making legit ARP Requests
 - Low ARP timeout configured in routers increases ARP rate
 - Sometimes one router leaves the network and the others keeps trying to reach it Forever
 - Any router can reply na ARP Request for Route Servers' IPs
 - Neighbour Discovery (similar to ARP, but DST MAC is multicast and only routers in the group will process the ND)
 - Unknown-Unicast
 - Common in Ethernet networks, but high volumes can consume most of the participant's link to the IXP
 - Possible causes:
 - Participant leaving unexpectedly the network
 - IXP device leaving the network (device fault or link fault)



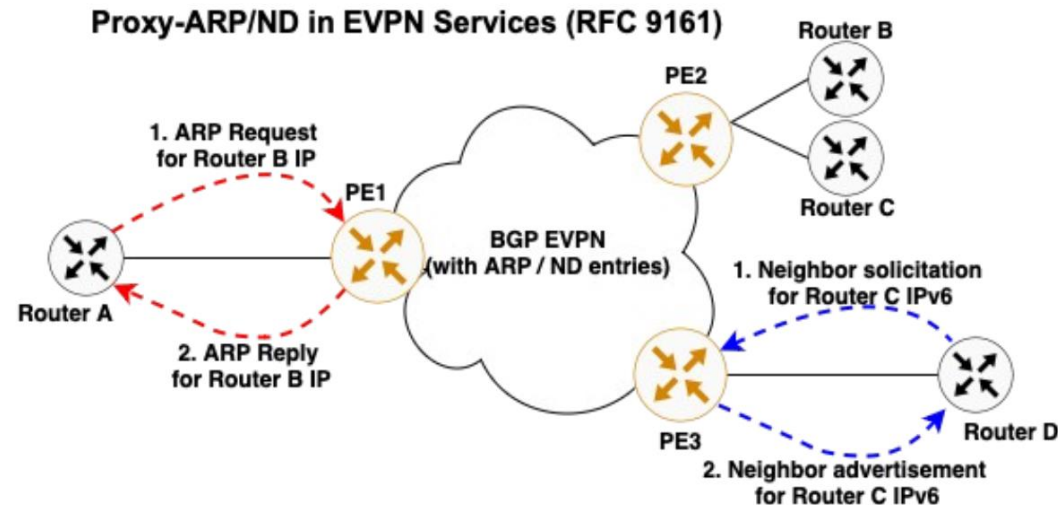
EVPN in a IXP network

- Creates Virtualized Ethernet Networks over IP or MPLS
- Uses BGP as its control plane to create services between PEs and advertise MAC to all devices
- Scalability: easily scale to support large number of domains and MAC addresses
- Enhanced visibility of the network services
- Fast service convergence
- Unknown-unicast flood can be disabled
- Smooth migration using VPLS and EVPN seamless integration (RFC 8560)



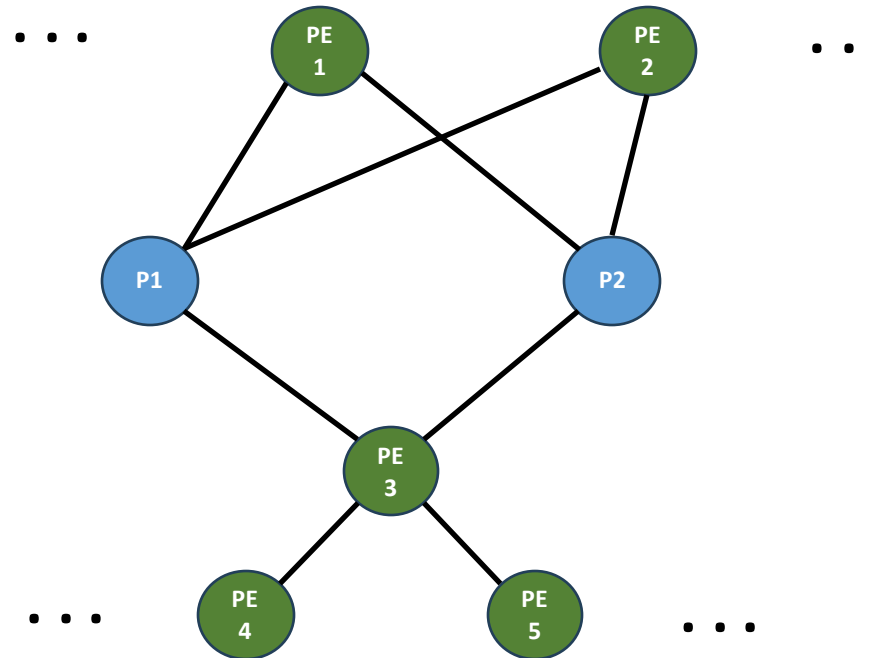
EVPN in a IXP network

- Proxy-ARP/ND in EVPN Services (RFC 9161)
 - PE routers learn ARP and ND from the services and advertise it to all PEs using BGP
 - If a ARP Request / ND reaches the PE and it has the entry, the request is not flooded to the network and PE answers it locally
 - Reduces unnecessary ARP Requests and ND flooded to the network
- Static ARP / ND entries are advertised with ARP/ND Extended Community with Immutable flag (RFC 9047)
 - Immutable binding entry overrides an existing non-immutable binding for the same [IP, MAC]
 - Prevents that some participant's router answer ARP / ND Requests for IPs out of their control
 - Mainly protects Route Servers's IPs



Main goal: use EVPN with Proxy-ARP/ND in São Paulo

- São Paulo is the biggest IX.br infrastructure
- 140 network equipment from four vendors, 2.000 participants ports
- Spine and leaf topology, with two P switches in the spine and 30 datacenters as leaves



Puzzle pieces

- **Core: 2 twins P switches**
 - Forward IPv4 packets: ✓
 - Forward IPv6 packets: ✓
 - Forward MPLS: ✓
 - Support OSPFv2: ✓
 - Support IS-IS: ✓
 - Load balance:
 - IPv4 MPLS/VPLS without FAT / Entropy label: ✗
 - IPv4 MPLS/VPLS with FAT / Entropy label: ✓
 - IPv6 native (using header Flow label): ✓

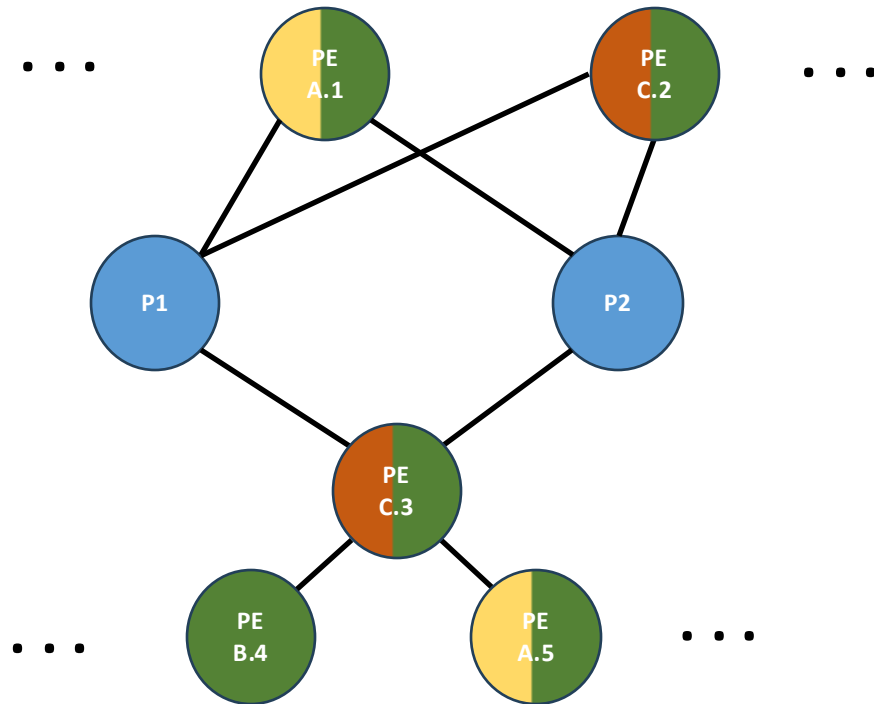
Puzzle pieces

- PEs – three vendors: A, B and C

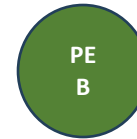
Feature / Vendor	A	B	C
VPLS/MPLS	✓	✓	✓
SRv6	✓	✗	✓
EVPN & MPLS	✓	✗	✓
EVPN & MPLS & FAT label	✗✓	✗	✗
EVPN & ARP/ND Proxy & MPLS	✓	✗	✗
EVPN & SRv6	✓	✗	✗✓
EVPN & ARP/ND Proxy & SRv6	✓	✗	🔨
OSPFv2 & IS-IS	✓	✓	✓
VPLS and EVPN SRv6 seamless integration	✓	✗	✓

✗✓ - Not all models with the feature implemented

Puzzle pieces together



- ✓ OSPFv2 (IPv4)
- ✓ IS-IS (SRv6)
- ✓ **MPLS/VPLS (green)**
- ✓ **EVPN + SRv6 + ARP/ND Proxy (yellow)**
- ✓ VPLS and EVPN SRv6 seamless integration



- ✓ OSPFv2 (IPv4)
- ✓ **MPLS/VPLS (green)**



- ✓ OSPFv2 (IPv4)
- ✓ IS-IS (SRv6)
- ✓ **MPLS/VPLS (green)**
- ✓ **EVPN + SRv6 ~~ARP/ND Proxy~~ (orange)**
- ✓ VPLS and EVPN SRv6 seamless integration

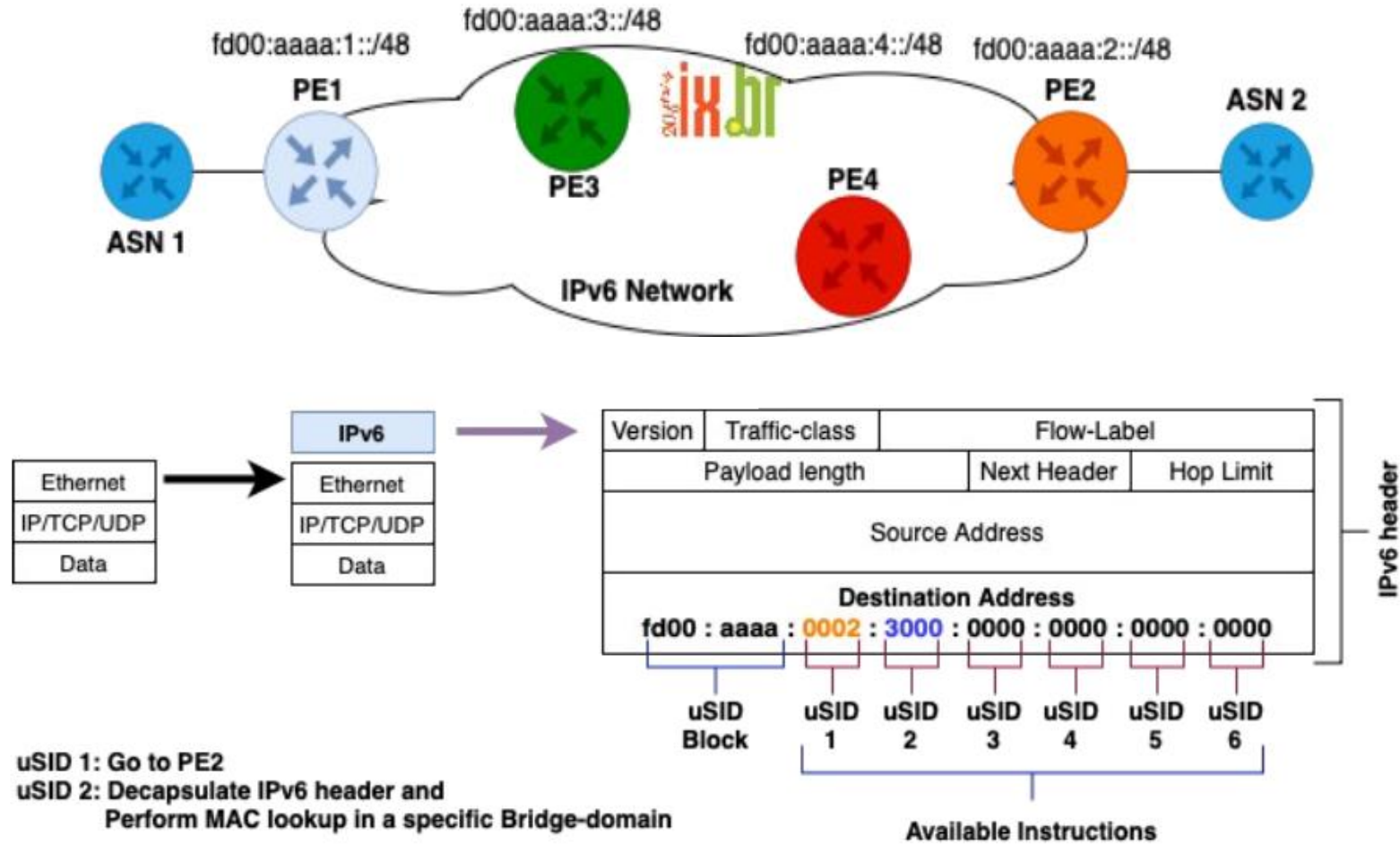
Segment Routing

- Source Routing: source chooses a path and encodes it in the packet as list of segments
- Segment: an identifier for an instruction (forwarding or service)
 - SR-MPLS: list of segments is represented as a Stack of MPLS labels
 - SRv6: list of segments is encoded in the IPv6 header
- Benefits of Segment Routing:
 - Simplify the network: don't need additional protocols like LDP or RSVP
 - All forwarding information is exchanged within the IGP
 - Easy implementation of Traffic Engineering
 - TI-LFA
 - Microloop avoidance

SRv6

- Combines the advantages of SR-MPLS with extensibility, simple and neat IPv6 forwarding
- Core devices just need to forward properly IPv6, it doesn't need SRv6 on it
- The use of MPLS requires Flow Aware Transport (FAT) or look inside the payload to properly load balance traffic
- SRv6 solves it using Flow-label at IPv6 header
- Due to extensibility of IPv6, it's easier to implement new features to SRv6
- Aligned with the transition to IPv6
- In summary, it permits us to start deployment without segment routing in all devices

SRv6 datapath

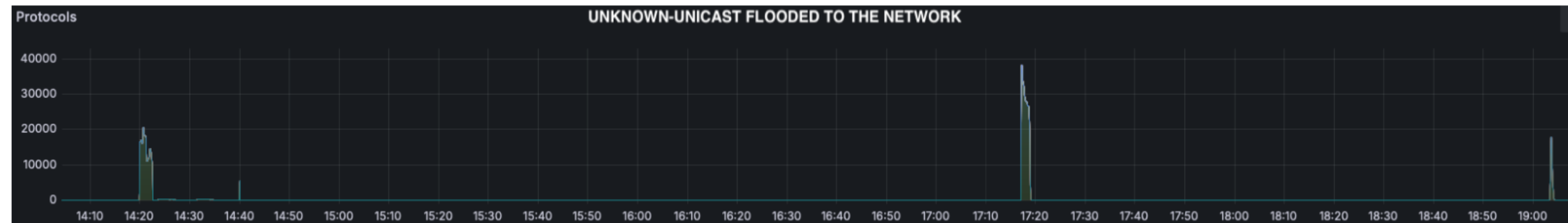


What IX.br have already done

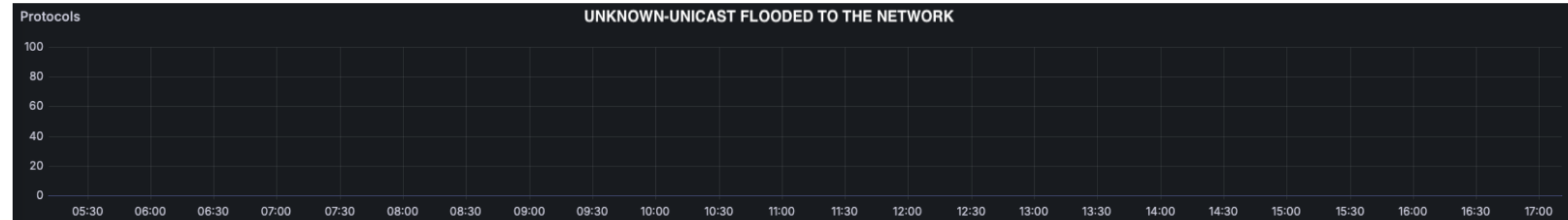
- Since 2016 IX.br is working with vendors to have na EVPN solution supporting all our needs
- In 2023 we tested and approved a complete EVPN solution (A)
- In February, 2024 we moved one IXP (Brasilia/DF) completely to EVPN using MPLS as underlay (A)
- In June, 2024 moved from MPLS to SRv6 (A)
- In August, 2024 Proxy-ARP/Proxy-ND enabled (A)
- In September, 2024 we moved one IXP (Curitiba/PR) completely to EVPN using SRv6 as underlay (B)
- H2 2024, start replacing IX.br São Paulo network equipment (B / C) by (A) or new equipment (A)
- Start using SRv6 and EVPN in part of the new equipment
- In November, 2024 we moved IX.br Manaus/AM completely to EVPN using SRv6 as underlay (B)
- Now in IX.br São Paulo:
 - ✓ 46% of the network equipments are capable to do EVPN & Proxy-ARP/Proxy-ND
 - ✓ The EVPN control plane has 50% of the MAC addresses
 - ✓ The biggest datacenter is in migration
 - ✓ And after this migration Proxy-ARP/Proxy-ND will be activated (two to three weeks)

IX.br Brasília/DF

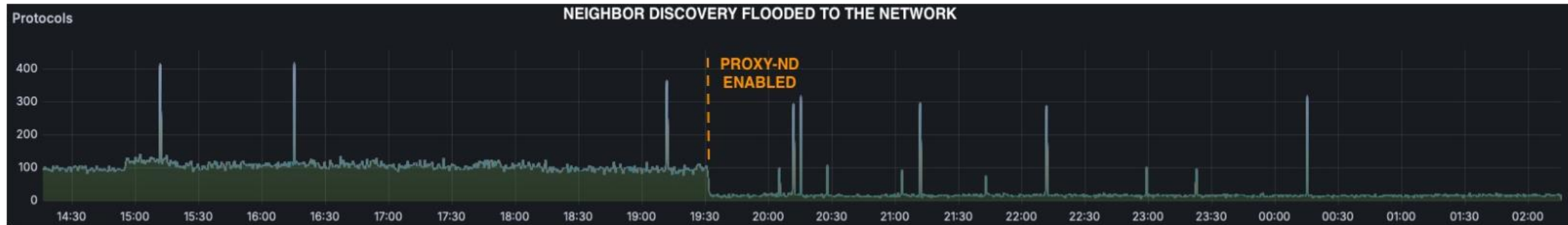
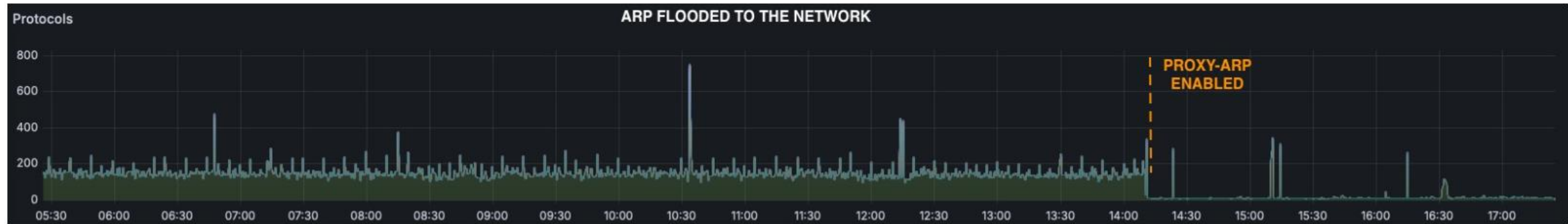
- Unknown-unicast flood enabled



- Unknown-unicast flood disabled



IX.br Brasília/DF



Conclusions

- IXP's have some specific challenges that are not so evident in traditional networks
- EVPN is one of the best solution to provide a layer 2 network for IXP's
- Proxy-ARP, Proxy-ND and no Unknown-unicast flood are essential for medium and large IXP's
- SRv6 can provide a resilient network with simple configuration and less protocols
- SRv6 EVPN empowers IX.br IXP's to provide a better, more resilient and scalable service to IXP members

Thanks !

www.ix.br

© jsirota@nic.br

February, 2025

nic.br **cgi.br**
www.nic.br | www.cgi.br