MIGRATING AN IXP FROM VPLS TO EVPN: A JOURNEY FROM VPLS TO EVPN + SR-MPLS AT NYIIX

Akio Sugeno, Telehouse America – presented by Greg Hankins, Nokia

RIPE 90 Connect Working Group

May 15, 2025

WHO WE ARE

- Telehouse America was established in 1987 and provides data center services in the US
  - Telehouse group has 48 data centers in 16 major global cities
- Telehouse America runs and operates an IXP, NYIIX, New York International Internet eXchange, since 1996
- The parent company is KDDI in Japan
  - The 2nd largest cell phone operator (au) / carrier in Japan
  - The global consolidated of KDDI in 2024 was $32B

NYIIX OVERVIEW

- POP locations (NY)
  - Telehouse (7 Teleport, 85 10th Ave.)
  - DRT (111 8th Ave., 60 Hudson St.)
  - Coresite (32 Ave. of Americas)
  - 165 Halsey meet-me-room
  - Equinix (275 Hartz Way)
- Number of members: 245 (as of April 4, 2025)
- Traffic peak: 2.3 Tbps (as of April 4, 2025)
- v4/v6 routes in the route servers
  - 260,000 IPv4 and 110,000 IPv6
- Route server RPKI support
- IX address range coverage by RPKI/ROA

ISSUES WITH THE PREVIOUS INFRASTRUCTURE

- Outdated VPLS + MPLS/LDP design
- Supported only up to 100G peering ports
  - Could not meet the growing need for 400G
- Sudden spikes in BUM (broadcast, unknown unicast, multicast) traffic were difficult to control
  - We needed more proactive traffic management
- ECMP did not quite work
  - Traffic was not evenly distributed to N x 100G trunks

PLANS TO RESOLVE THE ISSUES

- Support 400G peering ports
- Flexible port configuration 100G/400G if possible
- Well-baked reliable code
- Strong technical support by the vendor
- Short lead time (semiconductor shortage issues)
- Adapt EVPN to suppress and control BUM
- Adapt SR to simplify the configuration and operations
- We chose Nokia

NYIIX ASTRON PLATFORM

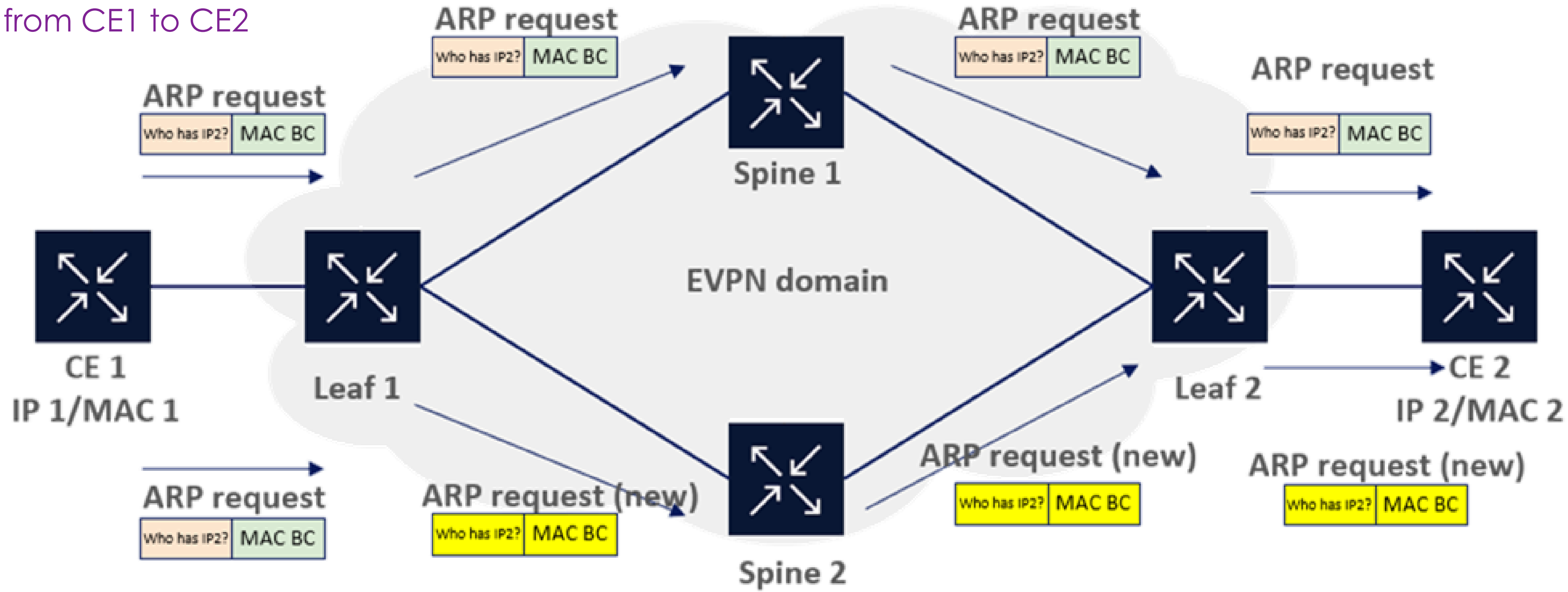
- New features and technical benefits
  - 400G ports for future-proof bandwidth capacity
  - EVPN for more efficient MAC learning and BUM traffic management
  - SR-MPLS for simpler routing configurations and native ECMP support for trunks
  - SR (underlay) and EVPN (overlay) in an integrated manner
    - Systematic configuration of layers with clear organization
  - MD-CLI offers a structured, tree-based configuration model
    - Supports advanced operations, efficient template-based management, and more consistent deployment across multiple devices
- Member migration completed on March 31, 2025
  - All existing NYIIX members were successfully migrated to the Astron platform
  - Thorough testing and phased transition ensured minimal downtime

EVPN OVERVIEW

- EVPN (Ethernet VPN) is a Layer 2 VPN technology defined in RFC 7432
- Provides MAC/IP advertisement over BGP, improving MAC learning and reducing flooding
  - Uses Route Type 2 (MAC/IP advertisement), sharing locally learned MAC/IP addresses via IBGP and enhancing overall visibility
- Supports active-active multihoming, split-horizon, and integrated routing
- Enables fine-grained BUM traffic control via ingress replication
  - Unicast instead of flooding

BUM REDUCTION WITH PROXY ARP/ND OVER EVPN

Proxy ARP/ND disabled: ARP is flooded from CE1 to CE2



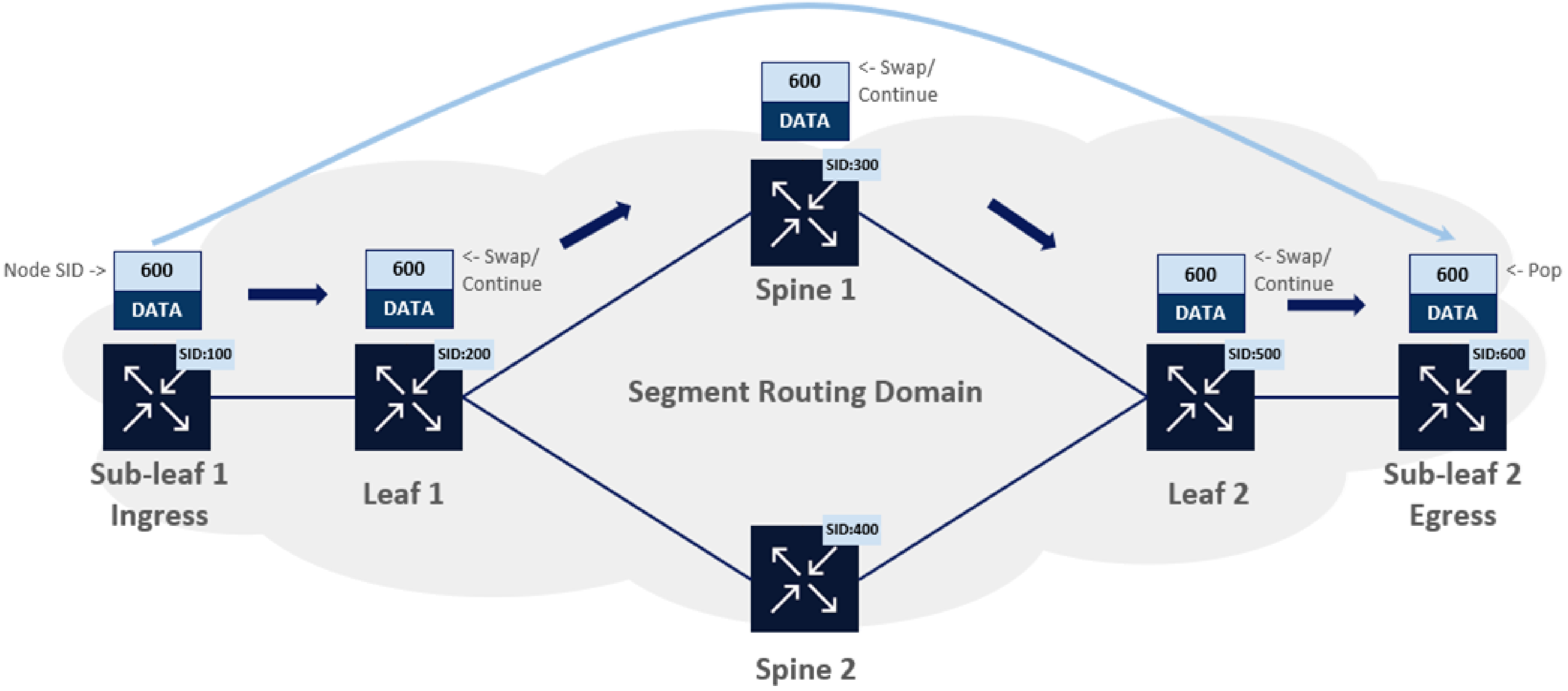
Proxy ARP/ND enabled: ARP request intercepted by Leaf 1 and forwarded to Leaf 2

- ARP request from CE1 is intercepted by Leaf 1
- Leaf 1 generates a new ARP request and sends it to Leaf 2
- Leaf 2 updates its ARP table and sends the ARP request to CE2

SR-MPLS OVERVIEW

- SR-MPLS (segment routing over MPLS) is defined in RFC 8402 as an architecture for source-based MPLS forwarding
- Encodes paths as sequences of segment identifiers (SIDs)
  - Segment IDs (SIDs) are assigned based on the SR Global Block (SRGB)
  - Distributed with the IGP, in this case OSPF
  - Eliminates the need for LDP/RSVP, simplifying the control plane
- ECMP is supported natively in the backbone, distributing traffic nearly evenly across multiple paths
- Anycast SIDs allow high availability for services and control-plane endpoints
  - Configured by assigning the same SID to two spine routers, making them appear as a single endpoint and balancing traffic evenly between them.

SEGMENT ROUTING OVER MPLS



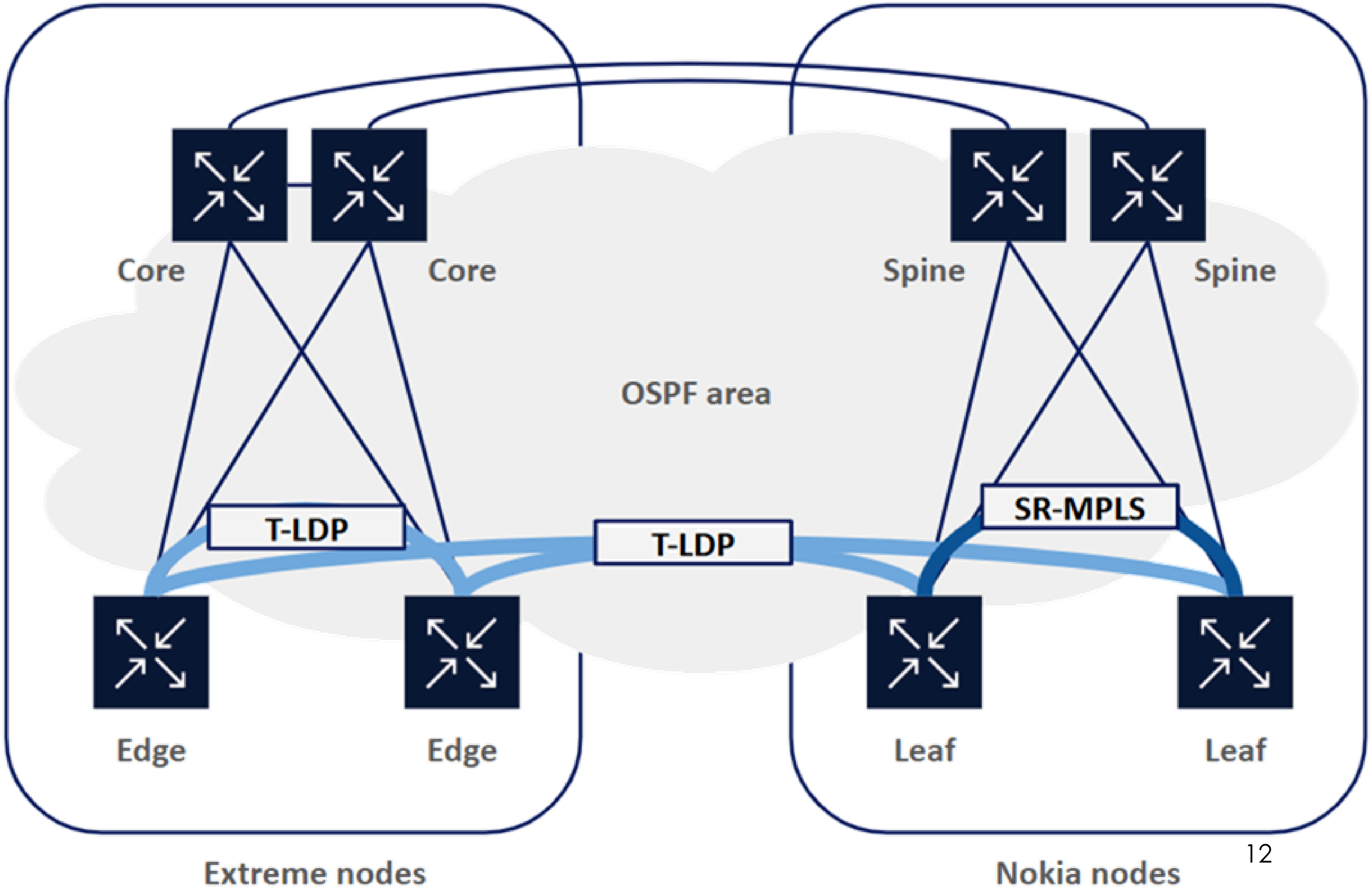
- This is an MPLS network, packets are forwarded based on MPLS labels
- Sender sends a packet with MPLS labels to the destination
- Node SID is distributed to every router with OSPF
- MPLS control plane protocols like LDP and RSVP are not used

TESTING

- Used Containerlab (cloud) extensively for lab testing
- Ran Nokia SR OS as containers with vrnetlab in Containerlab
- Extreme did not offer a virtual environment, so we imported existing configurations into SR OS for testing
- Created an environment as close as possible to production for reliable pre-deployment testing
- Initially considered MPLS-TE, but ECMP did not work properly with multiple LSPs (on IXR only)
- Tried SR-MPLS and found ECMP worked fine on both SR and IXR, so we chose SR-MPLS

PARALLEL OPERATION DURING MIGRATION

- Interconnect the old platform’s VPLS backbone with the Astron EVPN backbone to form a single bridge domain
- Run both systems in parallel and gradually switch traffic, minimizing service impact during the migration



PARALLEL OPERATION DURING MIGRATION

- Pre-migration
  - Added new devices to the OSPF backbone (area 0) with legacy devices, forming a single, unified area 0
  - Created a full mesh of T-LDP tunnels between the old and new devices
  - Under this setup, customer traffic on legacy devices is carried over T-LDP tunnels, while traffic on the new device is forwarded via SR-MPLS
  - This configuration enabled a parallel operation environment in preparation for the migration
- Migration day steps
  - Disconnect the cable from the old device
  - Connect the cable to the new device
  - Clear the MAC address entry if necessary
- Downtime is limited to the duration of the cable swap

MIGRATION CHALLENGES

- We encountered two main issues

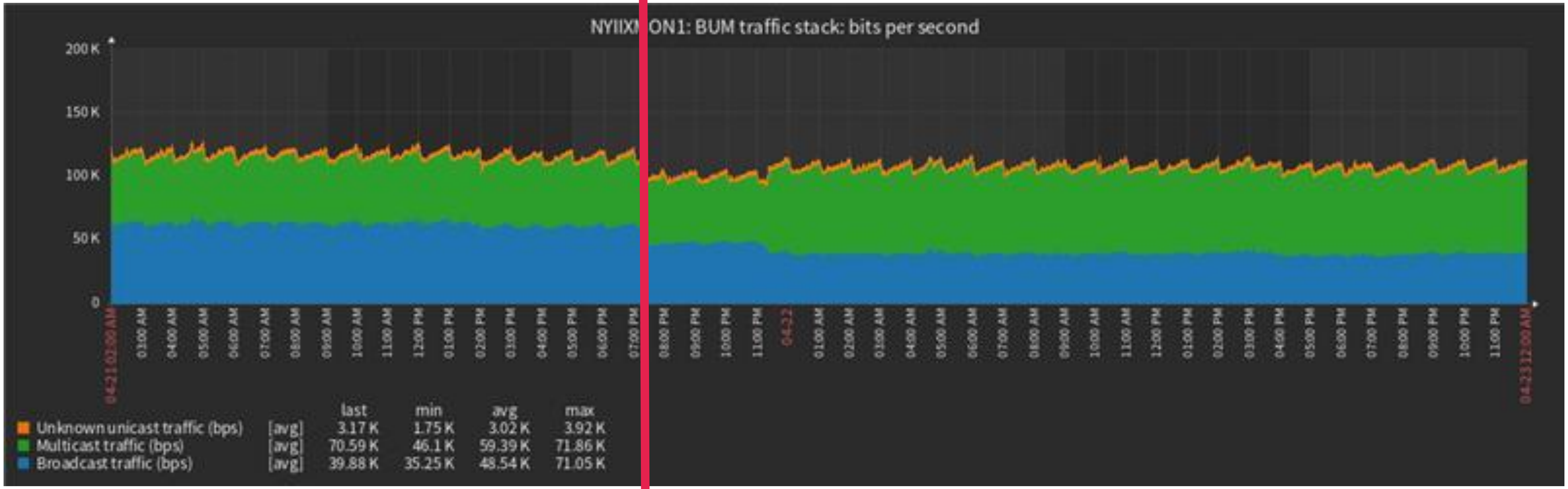
1. Burst BUM traffic
    - Caused a loop between VPLS and EVPN on Nokia
    - Resolved by placing the two domains into the same split-horizon group
2. MAC address not updating
    - When we moved cables from Extreme to Nokia, the MAC address remained as an entry learned from the Extreme on Nokia
    - Manually cleared the MAC entry on Nokia to fix the issue

BENEFITS GAINED FROM THE NEW NYIIX ASTRON PLATFORM

- Stability
  - Balanced traffic across multiple paths to avoid congestion and ensure overall network stability, plus faster convergence
- Scalability
  - By using 400G-capable hardware and a leaf-spine architecture, we can flexibly handle higher traffic and plan for future services
- Visibility
  - Centralized MAC learning with EVPN and simpler SR-MPLS configuration (no LDP or RSVP) make troubleshooting easier and give a clearer view of the whole network
  - BUM traffic before/after migration: 48k → 37k pps (or 23% reduction)

BUM TRAFFIC UNDER EVPN (BEFORE AND AFTER)

Before | After



Legend: **Green** Multicast, **Blue** Broadcast

FUTURE PLANS

- Planning to use weighted ECMP and PCEP for more flexible path and load management
- Exploring ways to take advantage of FlexAlgo for traffic engineering

QUESTIONS?