OPEN POSSIBILITIES.

Migration to SONiC from Three-tiered Legacy Network - EPFL Cast Study

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Migration to SONiC from 3-tiered Legacy Network – EPFL Case Study

Eric Krejci, Infrastructure Architect, EPFL
Kamran Naqvi, Principal Architect, Broadcom
Mehdi Abdelouhab, Product Manager, Juniper

EPFL Today

NETWORKING

Campus

11,449 students, of whom 2,199 PhD students

344 faculty

6,134 staff (incl. PhD)

Structure

5 Schools (13 study prog. leading to an MSc)

2 Colleges

20 Institutes

44 research centers

371 laboratories

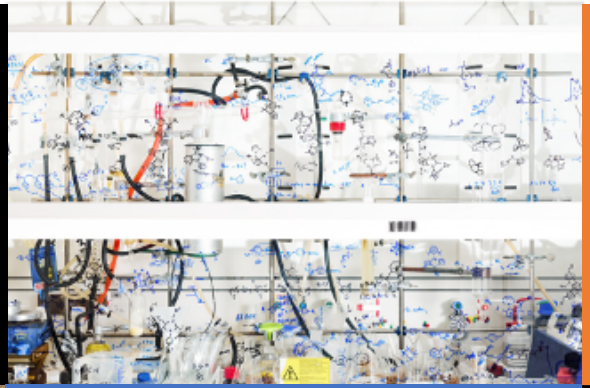
OCP GLOBAL SUMMIT

NOVEMBER 9-10, 2021

EPFL’s three missions according to the Federal Act

NETWORKING

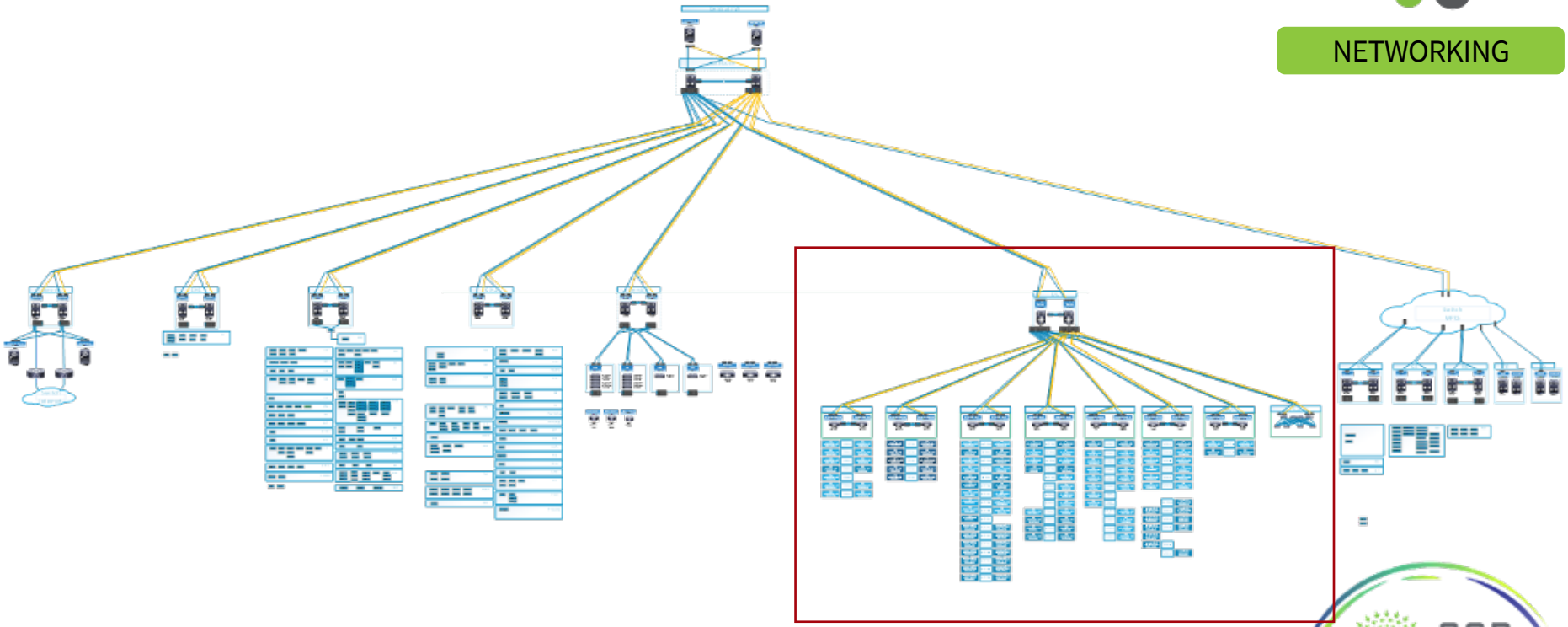
Education



Innovation

Research

Current Network Architecture



NETWORKING

OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

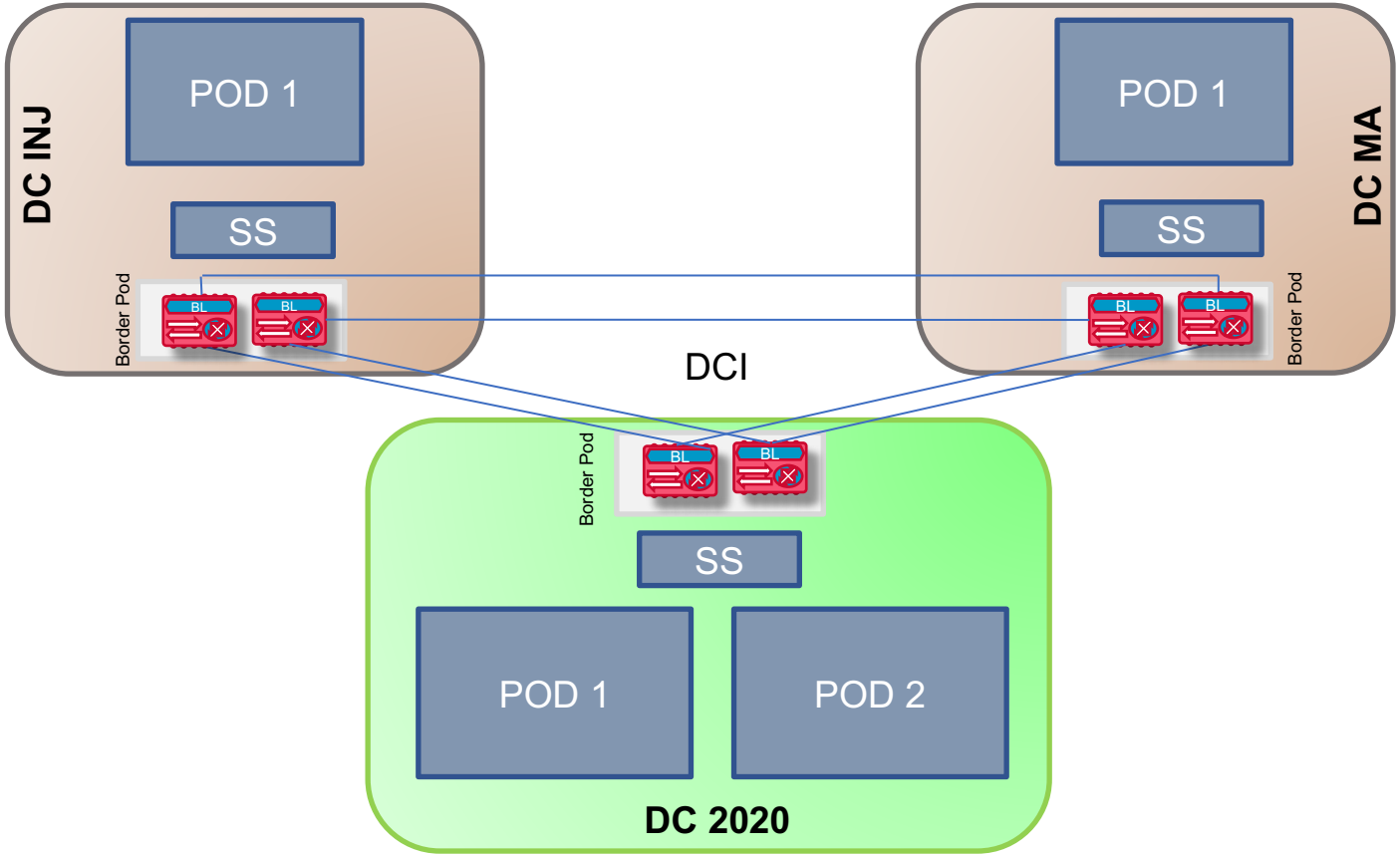
Goals for the Data Center Network

- No more blocking links for redundancy
  - No spanning tree
- Place a server everywhere within a POD
  - Optimize placement without compromising bandwidth or latency
- Ease to add links, hence bandwidth
- Flexibility for host connectivity
  - From 10Gb to 100Gb
- Provide HaaS capabilities with dynamic network and security assignments through automation

Why SONiC

- The same NOS, thus functionalities and reliability, across different Hardware Vendors to avoid locking and the freedom to choose the best hardware that suits our needs
- Being able to separate the hardware and the software
- Advanced telemetry. Deep visibility leads to rapid troubleshooting
- A modern NOS able to answer our needs in dissimilar workloads and capable to evolve accordingly with the best standards

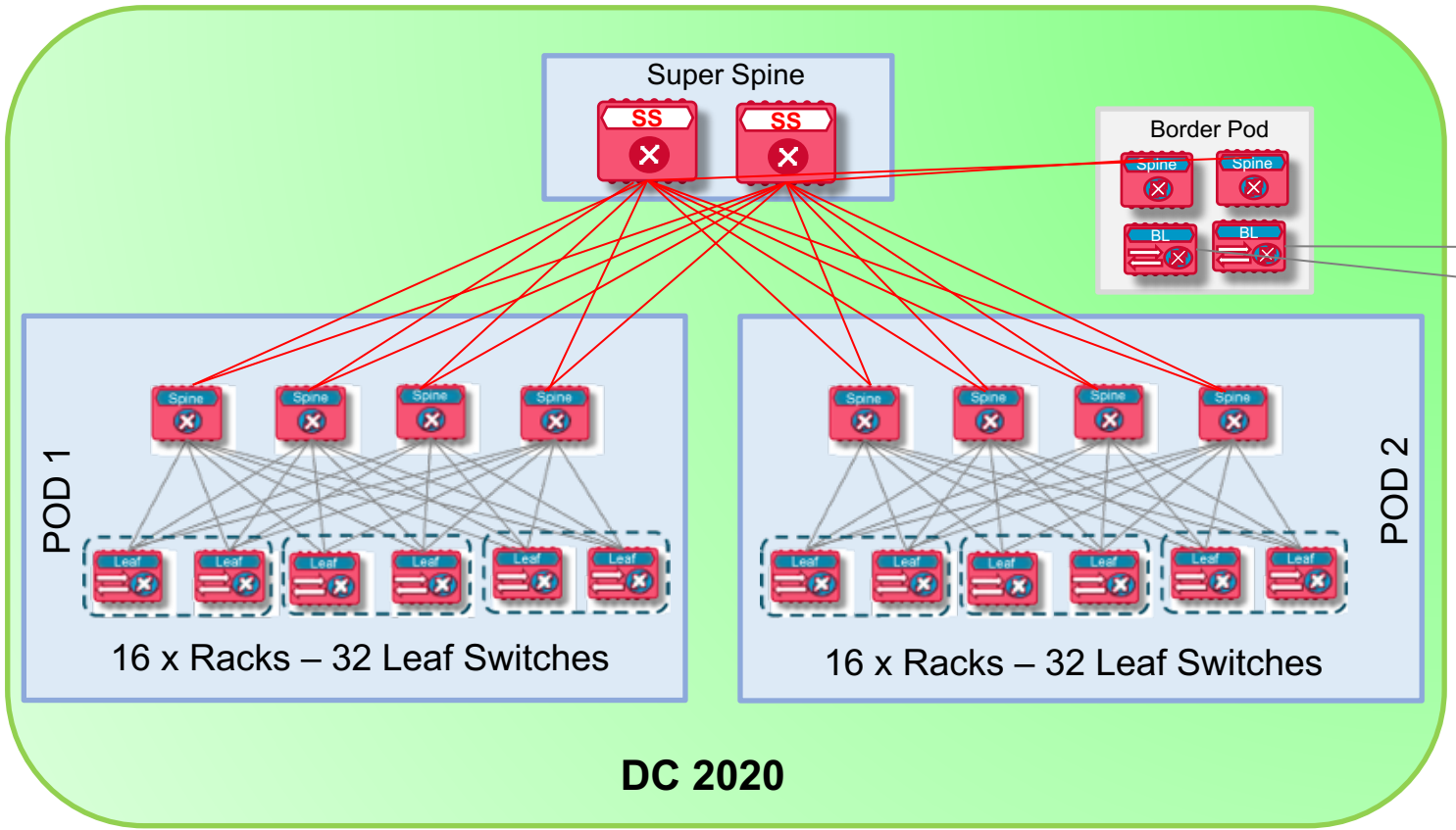
New Fabric Architecture



NETWORKING

200G

DC 2020



NETWORKING

Legacy

ESXi

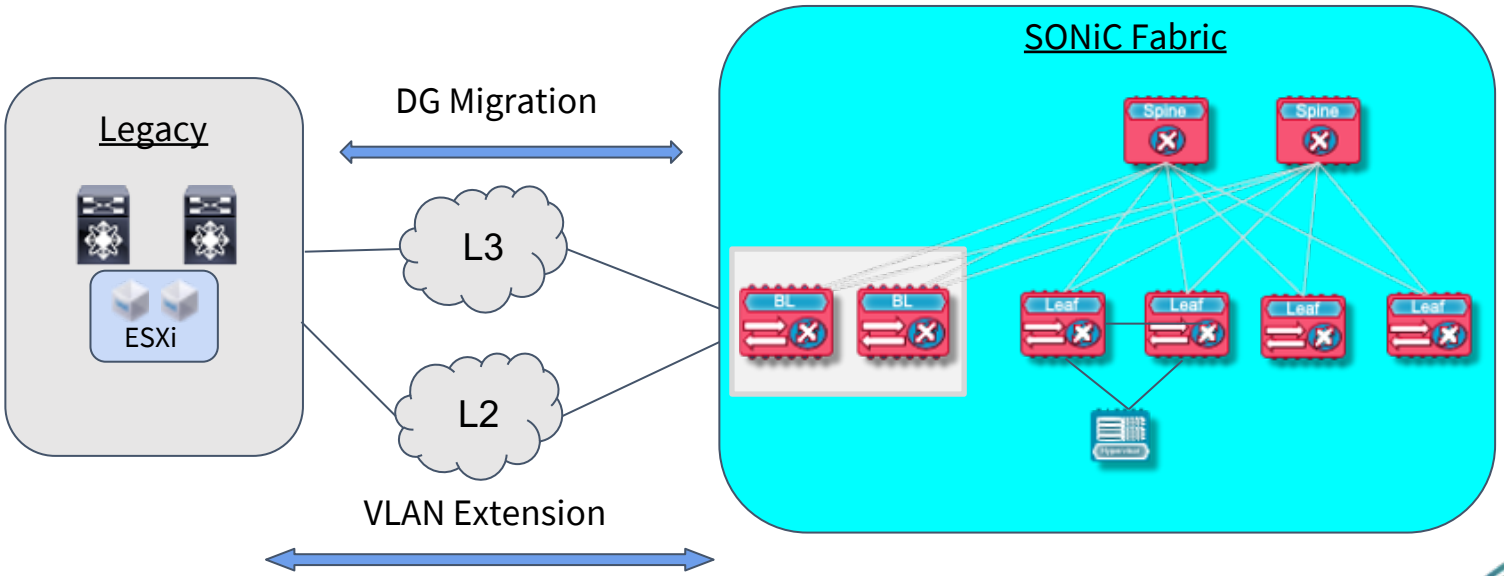
400G

100G

Switch Silicon

Sup Spine: TD4
Spine: TD4
Leaf : TD3

Migration Strategy



NETWORKING

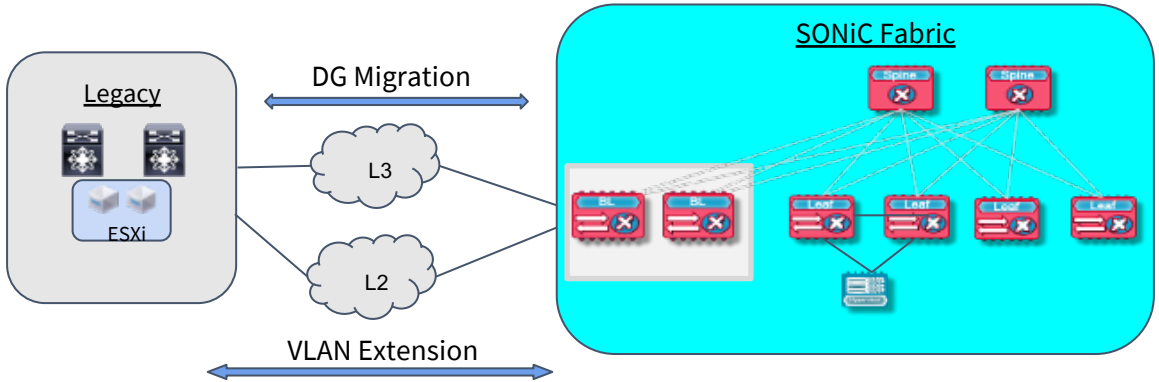
OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Migration Strategy

NETWORKING

- Extend VLAN to SONiC Fabric
- vMotion VMs to New Fabric
  - DG is still in Legacy Network
  - Tromboning traffic
- Once 50% or more workloads are migrated, it's time to migrate DG



- a. Configure the Default Gateway address of the VLAN as anycast gateway in the Fabric.
- b. Shutdown sub-interface of the corresponding VLAN on Legacy L3 Switch. This will also stop BGP advertisement from N7700-L3 for this subnet (VLAN).
- c. Advertise the subnet (VLAN) from Border Leaf to Legacy L3 Switch.

How to Scale!

- Initial Validation was done with CLI
- Default gateway migration is an intrusive step:
  - **6 sec outage** was observed for hosts in legacy network
  - **1 sec outage** was observed for hosts in new SONiC Fabric

Scale Requirements

- Migration of 1000s of VLANs requires automation
- EPFL expectations from Automation Tool:
  - **Ability to stage** all changes before the change window
  - **Feedback Loop** after committing the changes
  - **Ability to Roll-back** changes from 100s of devices with one push

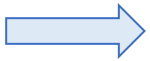
Approach to automating network operations

- "**Apstra IBN**" is about three distinct automation aspects:
  - Validation of intent correctness — eliminates operator error
  - Configuration generation — stateless automation focuses only on this
  - Operational expectations validation — most important aspect of automation.

- The first two are prerequisites, but the end goal is the correct outcome.

Step1) Preprovision all EVPN Tenants and VLANs

NETWORKING



VRF Name	Type	VLAN ID	Route Target	VNI
default	L3 Fabric	N/A	N/A	N/A
tenant_1	EVPN	2	10000:1	10000
tenant_2	EVPN	3	10001:1	10001

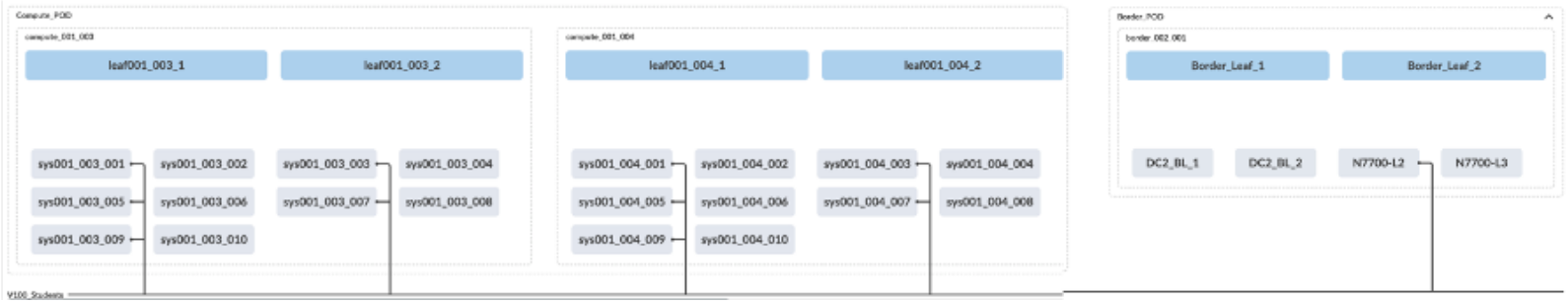
Name	Routing Zone	Type	VN ID	Assigned to	IPv4 Connectivity
V100_Students	tenant_1	VXLAN	10004	3 nodes	Disabled
V200_PhD	tenant_1	VXLAN	10003	3 nodes	Disabled
V300_Research	tenant_1	VXLAN	10002	3 nodes	Disabled

OPEN POSSIBILITIES.

Step2) Migrate VLANs

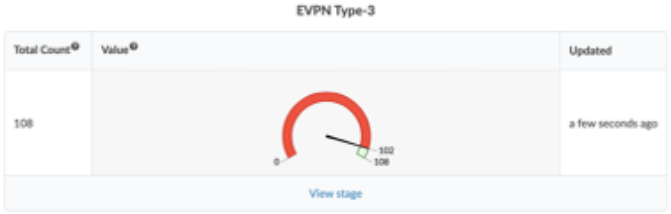
NETWORKING

- Use server labeling (unique names) or tags (meta-data) for batch assignment of interfaces
- Example: VLAN stretched over two compute racks, 10 servers per rack



Step2) Validate network state

NETWORKING



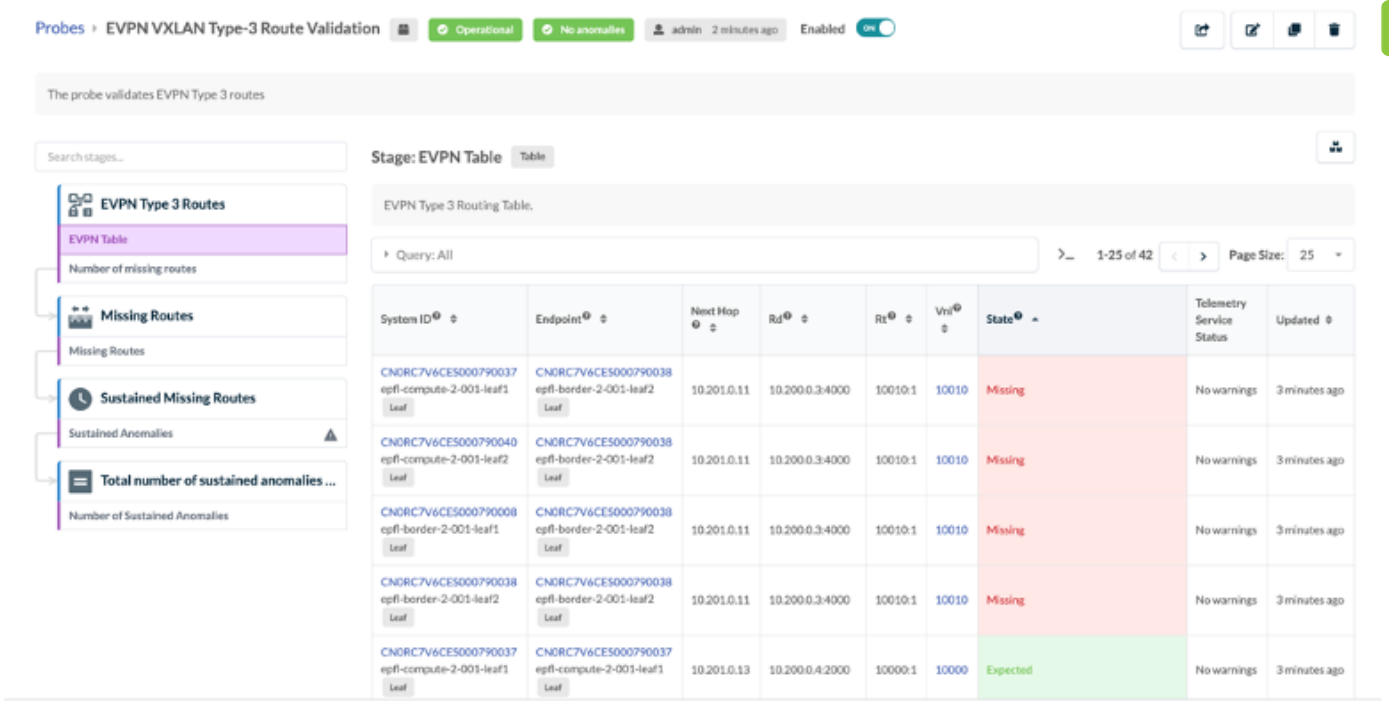
Validate EVPN control plane:

- Look at BGP routing table (RIB) for EVPN routes, check EVPN Route-Type 3 generated routes on a per device and VN basis and validates them against routes expectations derived from intent.

Validate EVPN data plane:

- Check device's forwarding plane (FIB) and validates them against Flood List expectation derived from intent.

Step2) Validate network state (cont'd)



NETWORKING

OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

Step3) Move L3 GW to EVPN domain, and validate

NETWORKING

- Augment VNs with L3 intent.

Incremental configuration (SONiC API)

State validation (EVPN Route-Type 5)

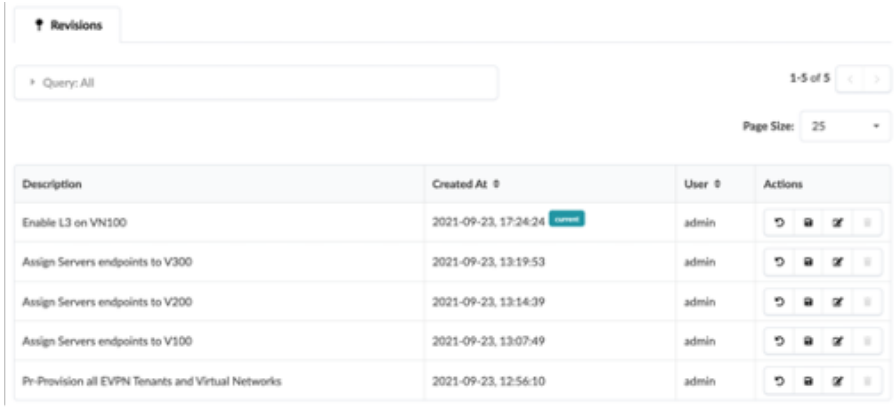
OPEN POSSIBILITIES.

OCP GLOBAL SUMMIT
NOVEMBER 9-10, 2021

In case of issues, rollback to previous fabric state

NETWORKING

- Store system-wide definitions of your intent, to roll-back to.
- Storing individual switch configurations is not scalable.
- The automation tool must render device's configurations at run time by selecting a blueprint revision.



Call to Action

NETWORKING

- SONiC is ready for Enterprise Deployment.
- There is no suitable Enterprise grade Orchestration option in Open Source community
- Apstra integration with SONiC provides a turnkey Day 0- Day 2 operations solution
- Trial of SONiC + Apstra

OPEN POSSIBILITIES.

Thank you!

NOVEMBER 9-10, 2021