

OPEN POSSIBILITIES.

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



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

OPEN POSSIBILITIES.



OPEN
COMMUNITY®



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



EPFL's three missions according to the Federal Act



NETWORKING

Education

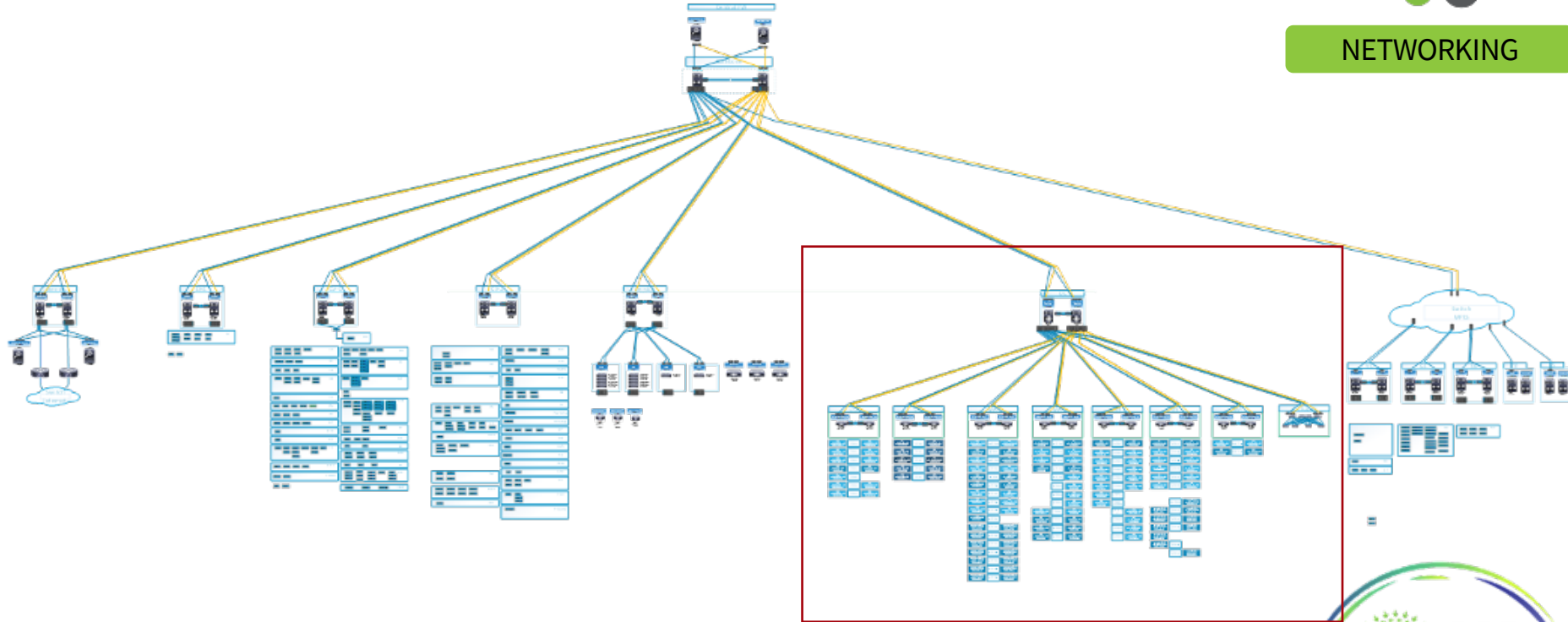
Innovation

Research

Current Network Architecture



NETWORKING



OPEN POSSIBILITIES.



Goals for the Data Center Network



NETWORKING

- 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

OPEN POSSIBILITIES.



Why SONiC



NETWORKING

- 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

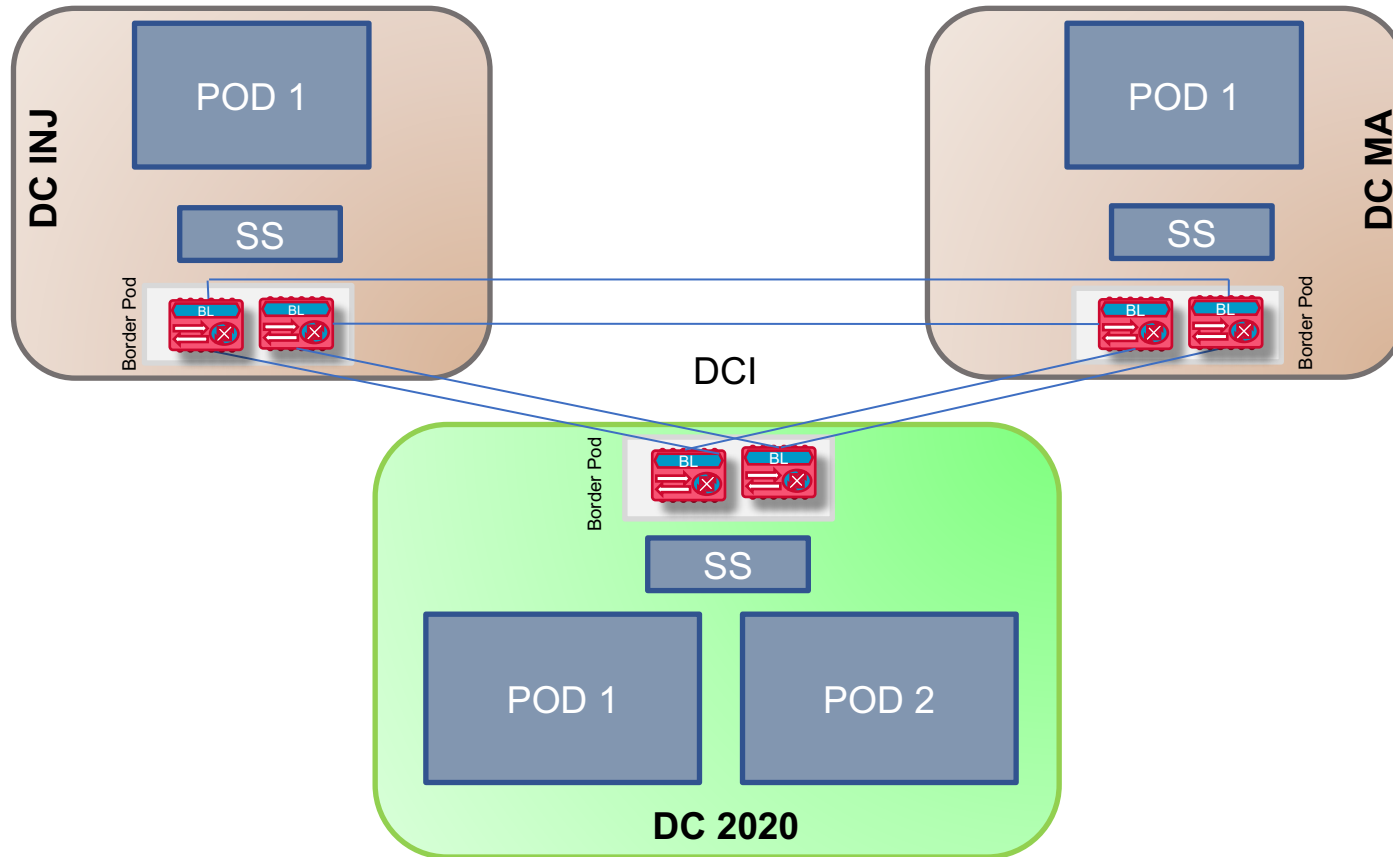
OPEN POSSIBILITIES.



New Fabric Architecture



NETWORKING



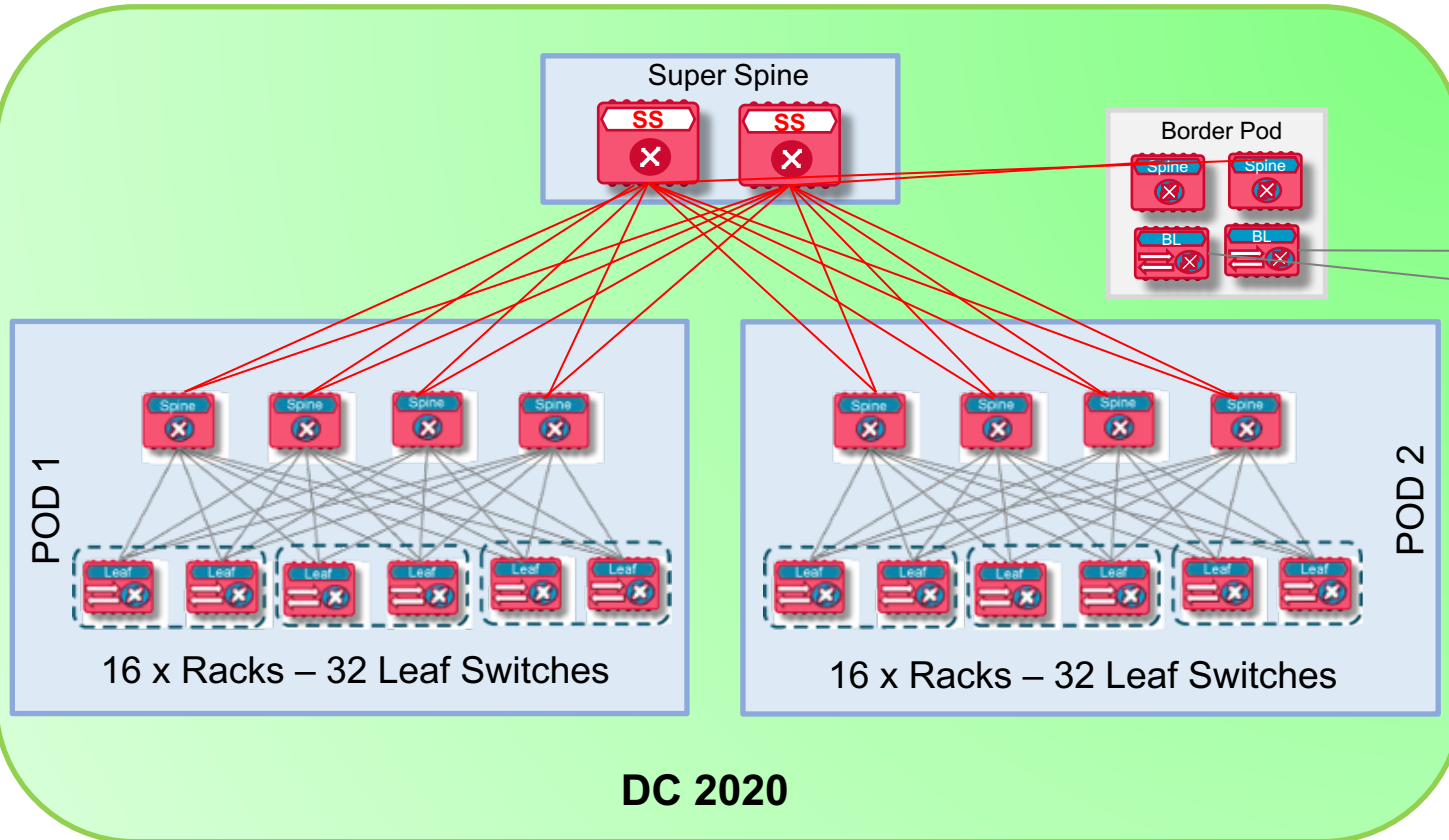
200G



DC 2020



NETWORKING



Legacy



400G

100G

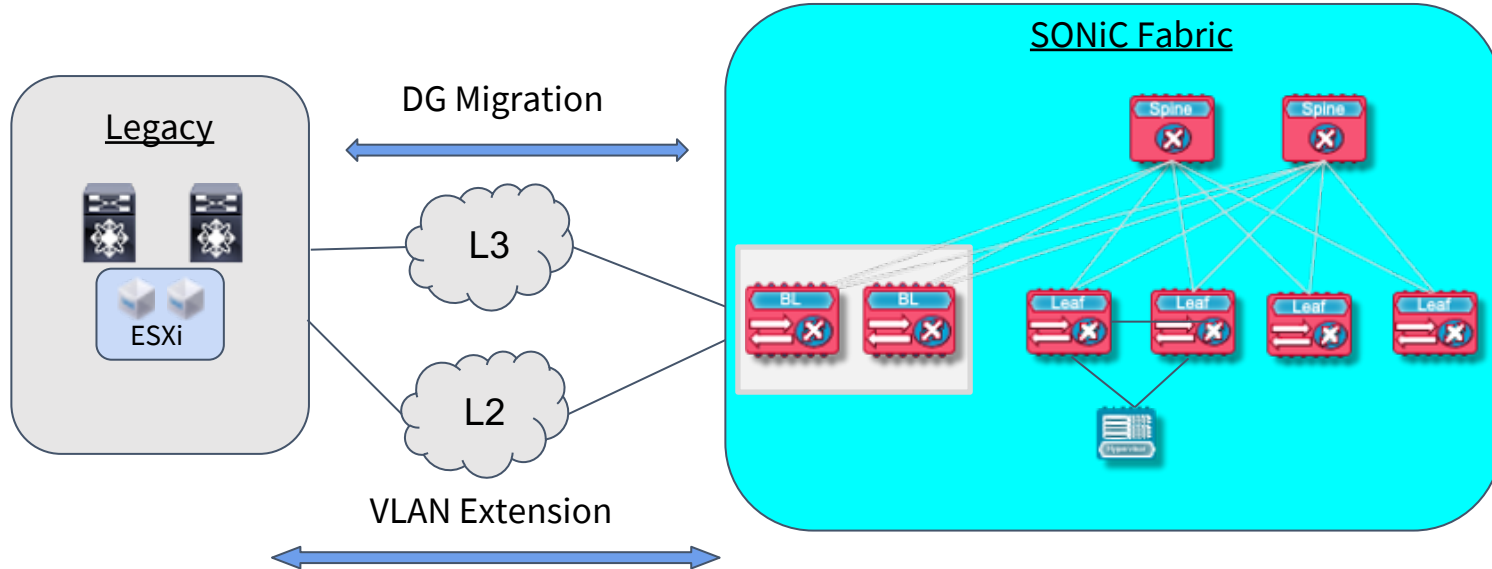
Switch Silicon

Sup Spine: TD4
Spine: TD4
Leaf : TD3

Migration Strategy



NETWORKING



OPEN POSSIBILITIES.

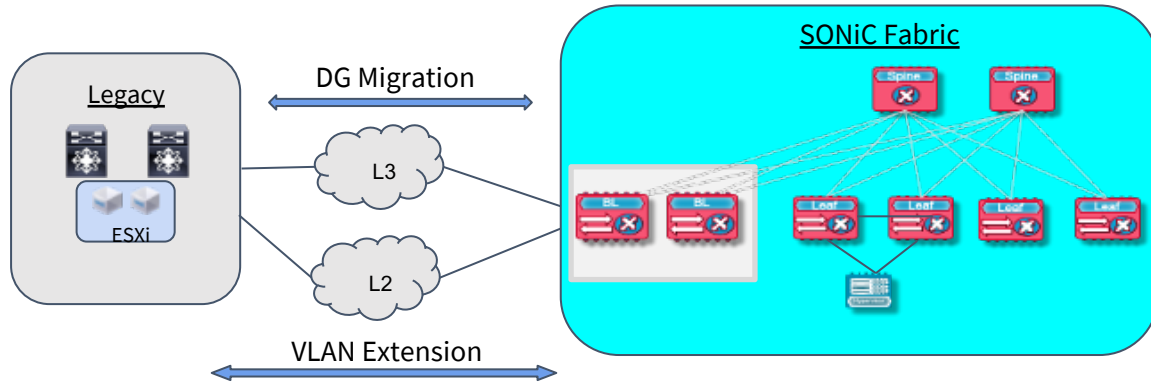


Migration Strategy



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!



NETWORKING

- 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

OPEN POSSIBILITIES.



Approach to automating network operations



NETWORKING

- “**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.

OPEN POSSIBILITIES.



Step1) Preprovision all EVPN Tenants and VLANs



NETWORKING

Create Routing Zone

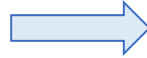
VRF Name *

VLAN ID *

VNI

Routing Policies

Select...



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

Create Virtual Network

Name *

Routing Zone

x

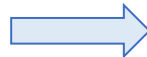
☒ tenant_1

☐ tenant_2

Set same VLAN ID on all leafs?

☒

VLAN ID (on leaf)



<input type="radio"/>	Name	Routing Zone	Type	VN ID	Assigned to	IPv4 Connectivity
<input type="radio"/>	V100_Students	tenant_1	VXLAN	10004	3 nodes	Disabled
<input type="radio"/>	V200_PhD	tenant_1	VXLAN	10008	3 nodes	Disabled
<input type="radio"/>	V300_Research	tenant_1	VXLAN	10002	3 nodes	Disabled

OPEN POSSIBILITIES.

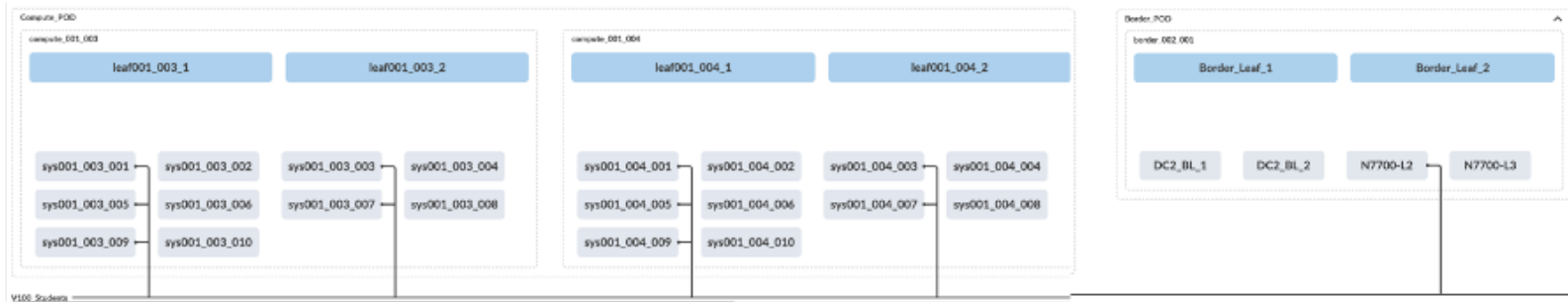


Step2) Migrate VLANs



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



OPEN POSSIBILITIES.



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.

OPEN POSSIBILITIES.



Step2) Validate network state (cont'd)



NETWORKING

Probes > EVPN VXLAN Type-3 Route Validation Operational No anomalies admin 2 minutes ago Enabled on

The probe validates EVPN Type 3 routes

Search stages...

- EVPN Type 3 Routes
- EVPN Table**
- Number of missing routes
- Missing Routes
- Sustained Missing Routes
- Sustained Anomalies
- Total number of sustained anomalies ...
- Number of Sustained Anomalies

Stage: EVPN Table Table

EVPN Type 3 Routing Table.

Query: All 1-25 of 42 Page Size: 25

System ID	Endpoint	Next Hop	Rd	Rt	Vni	State	Telemetry Service Status	Updated
CN0RC7V6CES000790037 epfl-compute-2-001-leaf1 <small>Leaf</small>	CN0RC7V6CES000790038 epfl-border-2-001-leaf2 <small>Leaf</small>	10.201.0.11	10.200.0.3:4000	10010:1	10010	Missing	No warnings	3 minutes ago
CN0RC7V6CES000790040 epfl-compute-2-001-leaf2 <small>Leaf</small>	CN0RC7V6CES000790038 epfl-border-2-001-leaf2 <small>Leaf</small>	10.201.0.11	10.200.0.3:4000	10010:1	10010	Missing	No warnings	3 minutes ago
CN0RC7V6CES000790008 epfl-border-2-001-leaf1 <small>Leaf</small>	CN0RC7V6CES000790038 epfl-border-2-001-leaf2 <small>Leaf</small>	10.201.0.11	10.200.0.3:4000	10010:1	10010	Missing	No warnings	3 minutes ago
CN0RC7V6CES000790038 epfl-border-2-001-leaf2 <small>Leaf</small>	CN0RC7V6CES000790038 epfl-border-2-001-leaf2 <small>Leaf</small>	10.201.0.11	10.200.0.3:4000	10010:1	10010	Missing	No warnings	3 minutes ago
CN0RC7V6CES000790037 epfl-compute-2-001-leaf1 <small>Leaf</small>	CN0RC7V6CES000790037 epfl-compute-2-001-leaf1 <small>Leaf</small>	10.201.0.13	10.200.0.4:2000	10000:1	10000	Expected	No warnings	3 minutes ago

OPEN POSSIBILITIES.



Step3) Move L3 GW to EVPN domain, and validate



NETWORKING

Incremental configuration (SONiC API)

- Augment VNs with L3 intent.

Edit Virtual Network

Virtual Network Parameters

Type
VLAN

Name
/s100.Students

Routing Zone
tenant_1

VNI
30004

Set same VLAN ID on all leafs?
☐

Route Target

10004:1

DHCP Service
☒ Disabled
☐ Enabled

IPv4 Connectivity
☐ Disabled
☒ Enabled

IPv4 Subnet
192.168.10/24

Virtual Gateway (IPv4)
Enabled?
☒

Virtual Gateway IPv4
192.168.10.254

leaf001_004_1 Incremental Config Preview

```
1 -----RESTCONF-----
2 {
3   "url": "/data/openconfig-interfaces:interfaces/interface=Vlan100/openconfig-vlan:routed-vlan/openconfig-lf-ip:ip=4/openconfig-interfaces:interfaces-external-static-anycast-gateway": {
4     "data": {
5       "openconfig-interfaces-external-static-anycast-gateway": {
6         "192.168.10.254/24": {
7           }
8         }
9       },
10      "method": "PATCH",
11      "config_gen_mode": "incremental",
12      "expect": 204
13    }
```

State validation (EVPN Route-Type 5)

EVPN Type5 IPV4

Total Count	Value	Updated
50		a few seconds ago
View stage		

OPEN POSSIBILITIES.



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.

Revisions			
Query: All			
1-5 of 5			
Page Size: 25			
Description	Created At	User	Actions
Enable L3 on VN100	2021-09-23, 17:24:24 Cancel	admin	↶ 🗑 🔍 ⌵
Assign Servers endpoints to V300	2021-09-23, 13:19:53	admin	↶ 🗑 🔍 ⌵
Assign Servers endpoints to V200	2021-09-23, 13:14:39	admin	↶ 🗑 🔍 ⌵
Assign Servers endpoints to V100	2021-09-23, 13:07:49	admin	↶ 🗑 🔍 ⌵
Pre-Provision all EVPN Tenants and Virtual Networks	2021-09-23, 12:56:10	admin	↶ 🗑 🔍 ⌵

OPEN POSSIBILITIES.



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