



Educação, Pesquisa
e Inovação em Rede

Global P4 Lab – SC23 Results

Marcos Schwarz and Frederic Loui

GNA-G Community VCs Q4 2023

December 06, 2023



Acknowledgements

GNA-G Data Intensive Science WG

GNA-G AutoGOLE / SENSE WG

GEANT RARE Project

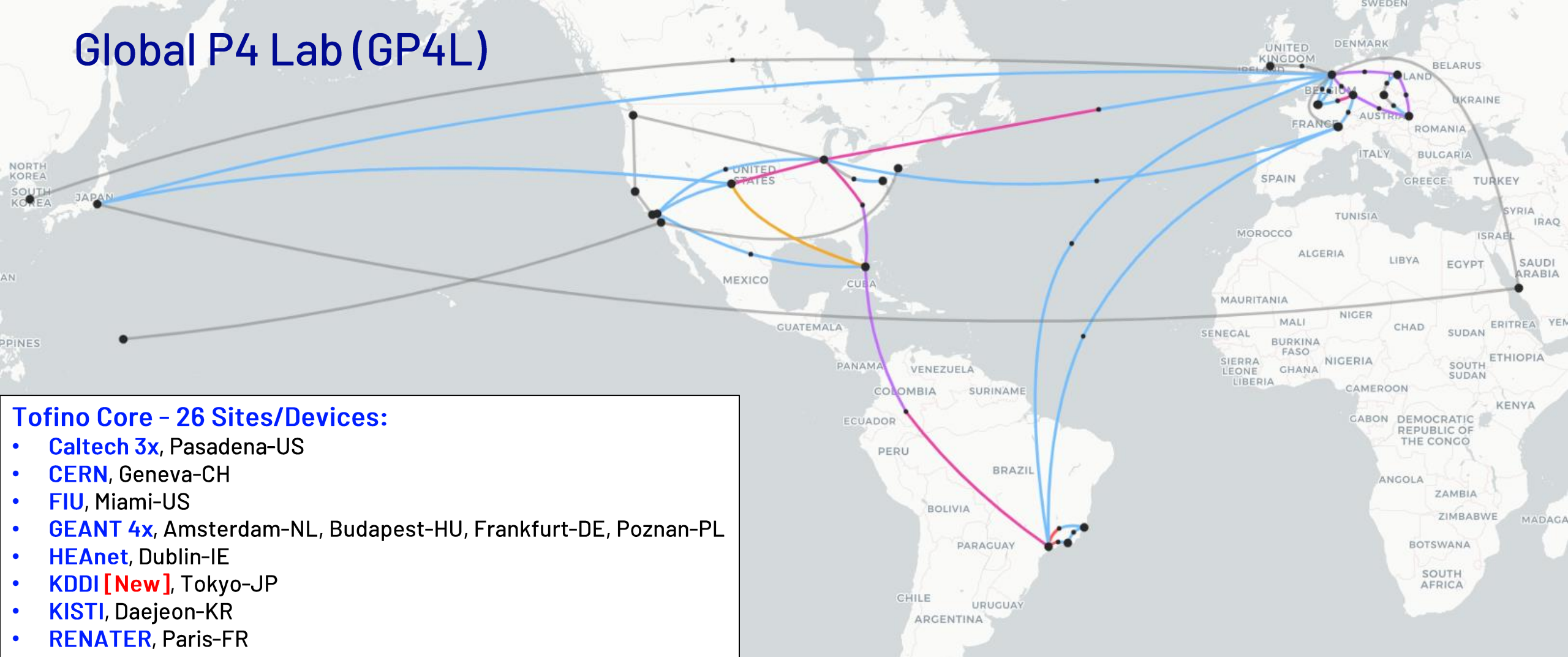
... And all it's collaborating institutions and teams

— Motivation

1. How can we increase the rate of evolution of Research Networks without interfering with production?
2. How to develop and operate end-to-end / multi-domain orchestration services?
 - Resource reservation (guaranteed bandwidth)
 - Resource provisioning (Circuits, VRFs)
 - Underlay observability
 - Dynamic traffic steering/engineering
 - Dynamic creation of L3 VPNs
 - Closed loop multi-domain visibility/intelligence/controllability
3. How can we create/sustain an integration initiative/platform to propose and validate next generation protocol and services?

Proposition: Use programmable devices to experiment on persistent pre-production networks leveraging industry/R&E open ecosystems

Global P4 Lab (GP4L)



Tofino Core - 26 Sites/Devices:

- Caltech 3x, Pasadena-US
- CERN, Geneva-CH
- FIU, Miami-US
- GEANT 4x, Amsterdam-NL, Budapest-HU, Frankfurt-DE, Poznan-PL
- HEAnet, Dublin-IE
- KDDI [New], Tokyo-JP
- KISTI, Daejeon-KR
- RENATER, Paris-FR
- RNP, Rio de Janeiro-BR
- SC23 [New], Denver-US
- SouthernLight, São Paulo-BR
- StarLight, Chicago-US
- SWITCH 6x [New], Geneva-CH
- Tennessee Tech, Cookeville-US
- UFES, Vitória-BR
- UMd/MAX, College Park-US

BlueField-2/DPDK Islands - 7 Sites/Devices [New]:

- Pacific Wave/UCSD, Chicago-US, Guam-GU, Los Angeles-US, New York-US, San Diego-US, Seattle-US, Sunnyvale-US

x86/DPDK Islands - 4 Sites/Devices:

- FABRIC [New], Miami-US
- 2x GEANT, Paris-FR, Prague-CZ
- KAUST [New], Saudi Arabia-SA

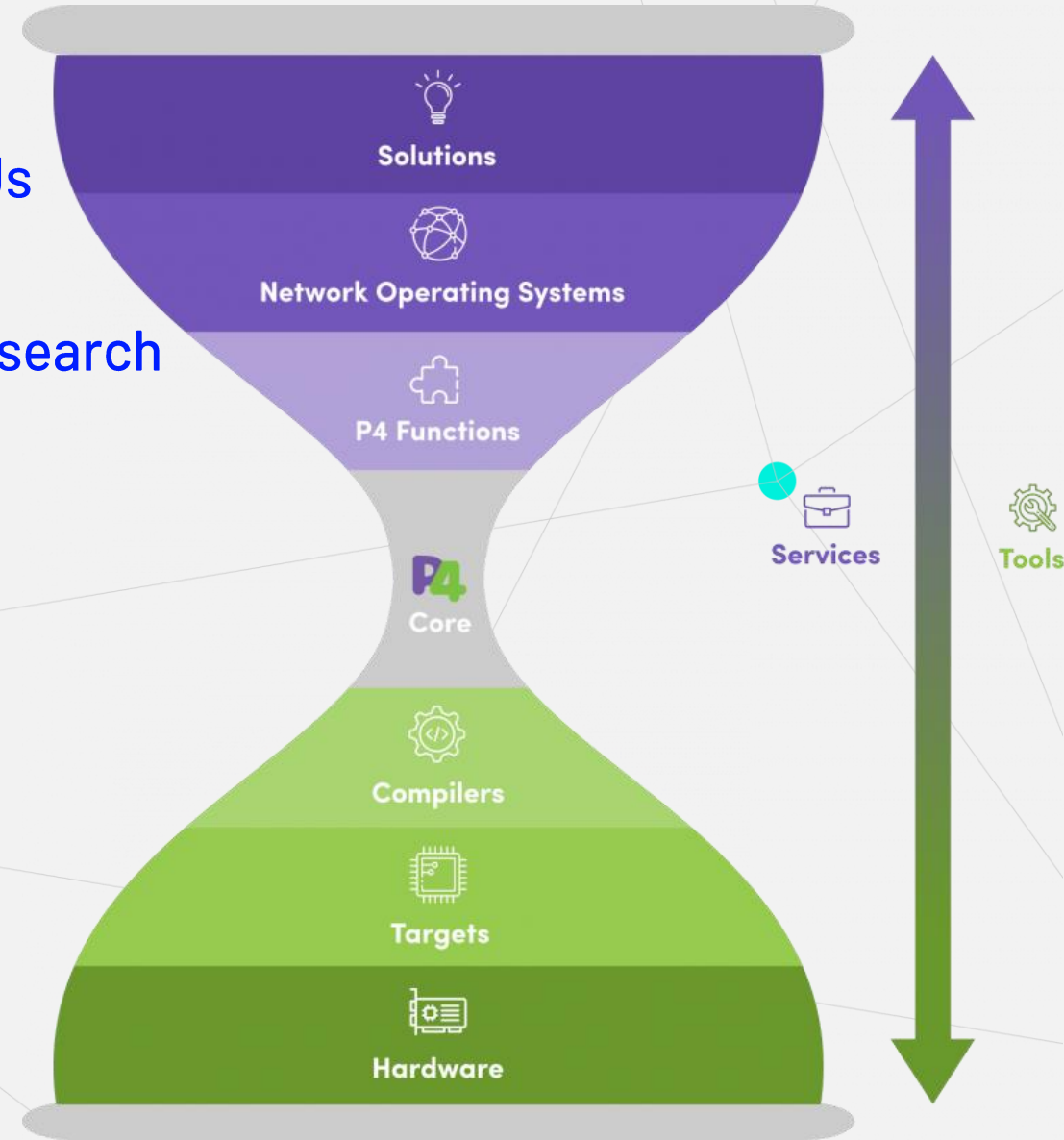
Why P4?

Open Hardware and Software Ecosystem

- Programmable ASICs, FPGAs, SmartNICs and CPUs
- Open implementation of industry standards
- Facilitates development of new features and research

Extensible Production Grade NOSes

- SONiC/PINS (Linux Foundation/ONF)
- RARE/freeRtr (GEANT)



— NOSes Highlights

SONiC/PINS

- Backends (P4, Fixed-function)
- Source Routing (SRv6)
- Telemetry (sFlow, INT, gNMI)
- SDN Controller (P4Runtime)
- Automation (Ansible / Batfish / gNMI)
- Routing (BGP)

New Features

- SRv6 uSID, Path Tracing, TWAMP

RARE/FreeRtr













- Backends (P4, DPDK, eBPF/XDP)
- Source Routing (SR-MPLS, PoIKA)
- Telemetry (Prometheus)
- Telco Protocols (BGP-LS, PCEP, TWAMP)
- Automation (Rest API)
- Routing (OSPF, ISIS, BGP)
- R&E driven development (Packet Marking, PoIKA)

New Features

- BGP-CT, BIER, RIFT

GP4L Management Plane

NMaaS Network Management as a Service

 Grafana ★★★★★5 Open source analytics & monitoring solution for every database	 Prometheus ★★★★★5 Monitoring system & time series database	 Icinga2 ★★★★★0 Scalable and extensible monitoring system	 SPA Inventory ★★★★★0 Resource and Service Inventory with TMF-compliant API
 Routinator ★★★★★0 RPKI Validator	 WebDAV Server ★★★★★5 WebDAV Server with Git versioning	 Uptime Kuma ★★★★★5 Self-hosted monitoring tool like "Uptime Robot"	 NetBox ★★★★★5 Infrastructure resource modeling application
 WiFiMon ★★★★★0 Wireless Crowdsourced Performance Monitoring and...	 Zabbix ★★★★★0 Enterprise-class monitoring solution for networks and applications	 SPA ★★★★★0 SPA for the E-Line service	 Healthchecks ★★★★★0 A cron monitoring tool

Kubernetes
Inside



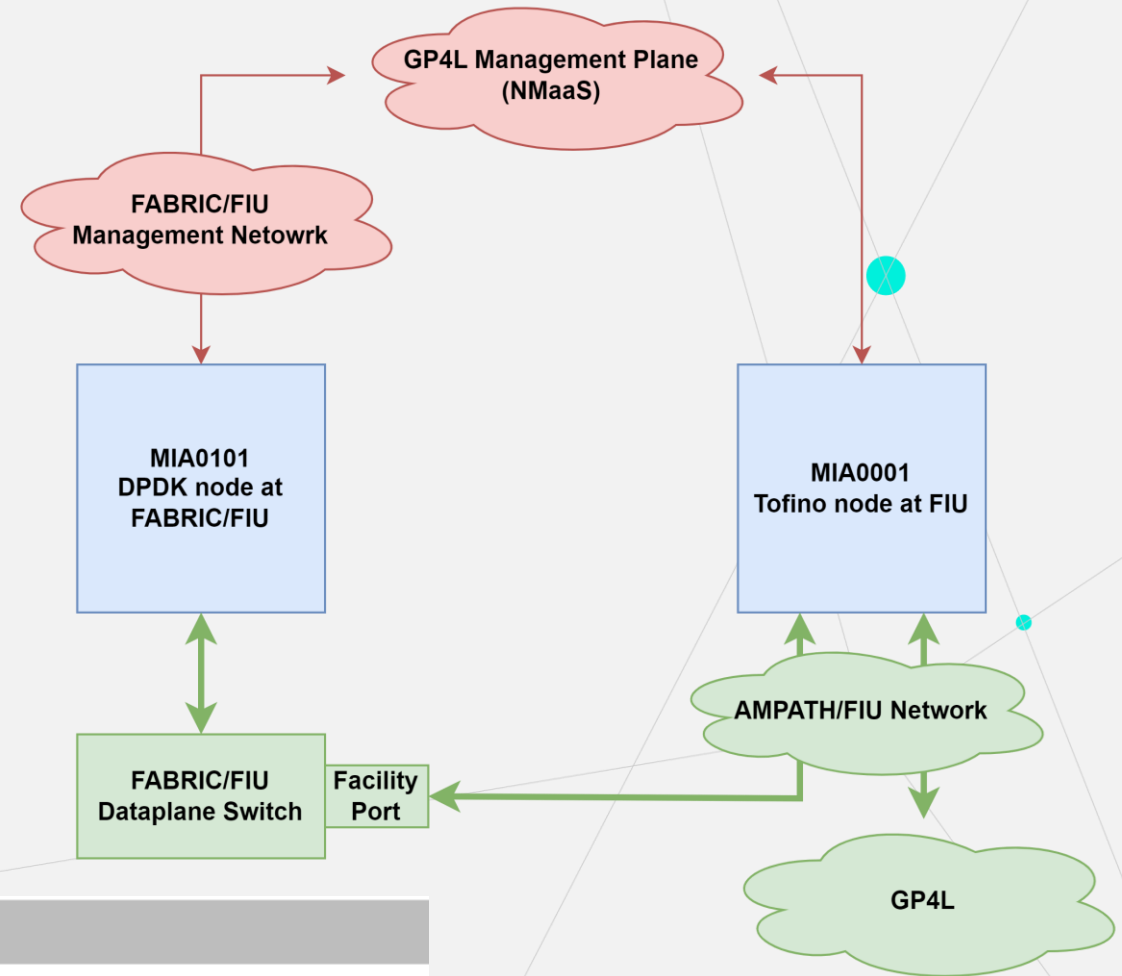


— GP4L Use Cases

- Topology Monitoring with BGP-LS
- Polka - An innovative source routing paradigm, IFES/UFES
- Packet Marking Specification: IPv6 Flow Label, CERN
- SuperComputing 23
 - Integration with FABRIC at FIU
 - Support for Bluefield-2 on Pacific Wave nodes
 - Automated Digital Twin and Topology Visualization

Integration with FABRIC at FIU

- FreeRtr/DPDK container deployed over FABRIC VM with SmartNIC
- Interconnection to FIU P4 Switch through a L2 Facility Port
- Wireguard VPN to GP4L Management Plane
- Extended Slice Lease to SC23 duration

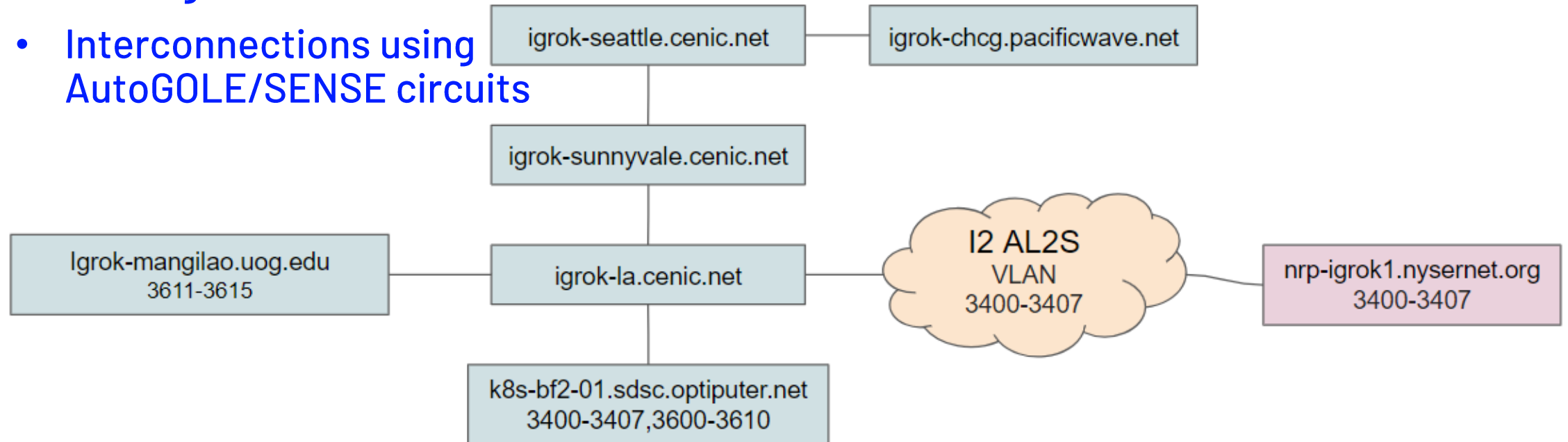


```
facility_port.ipynb x MIA0101 x +
[rocky@Node1 freertr-container]$ docker ps
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS        NAMES
ef5a2f395110   freertr/freertr-dpdk:latest         "/bin/sh -c './scrip..." 22 minutes ago Up 22 minutes        freertr-001
[rocky@Node1 freertr-container]$ telnet localhost 2323
Trying ::1...
telnet: connect to address ::1: Connection refused
Trying 127.0.0.1...
Connected to localhost.
Escape character is '^]'.
welcome
line ready
MIA0101#
```

Support for Bluefield-2

Pacific Wave IRNC
NRP IGROK Bluefield 2
P4 VLAN Topology

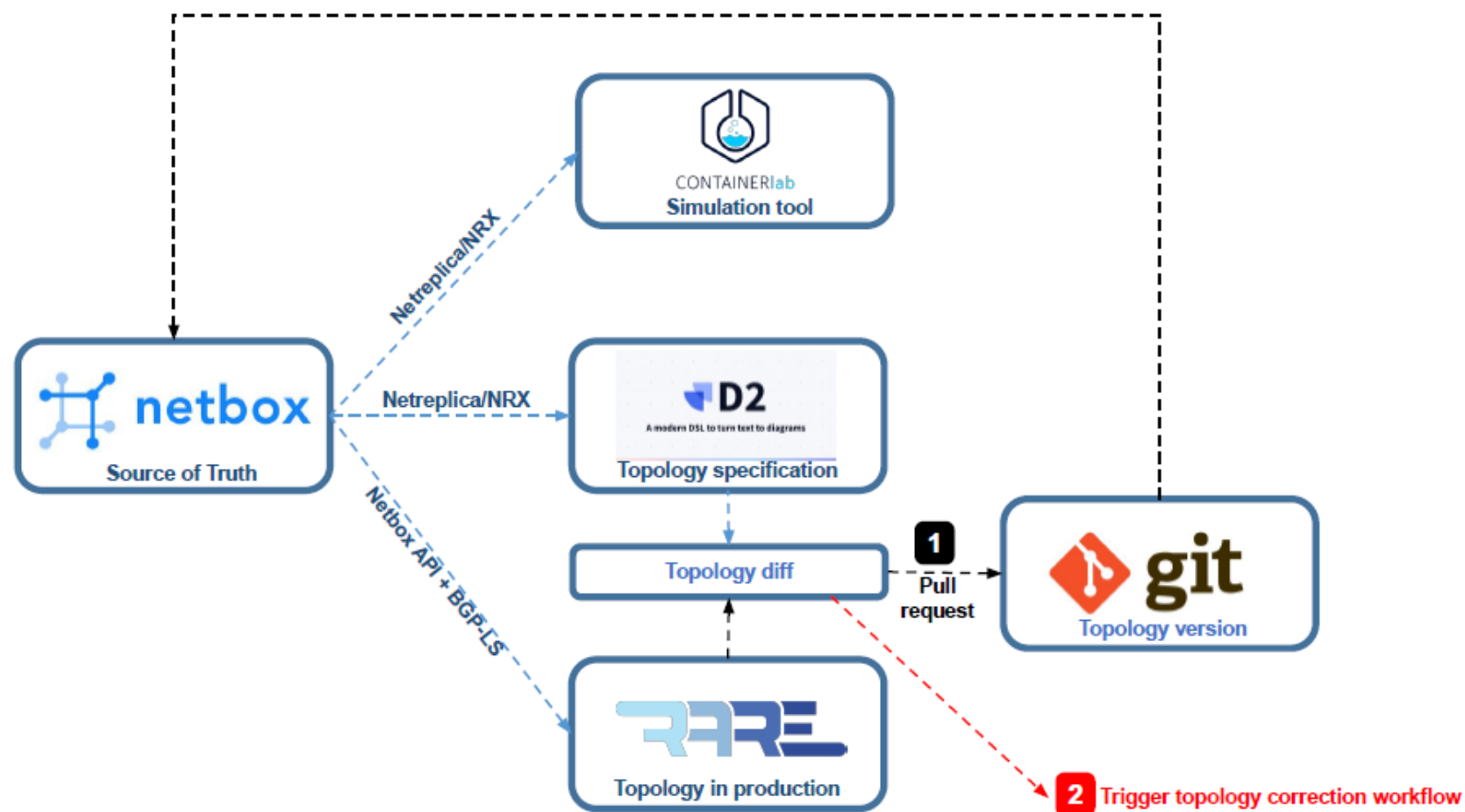
- FreeRtr and VPP ARM64 containers running inside the DPU
- Interconnections using AutoGOLE/SENSE circuits



All nodes are regional sFlow collectors for the IGROK service

We are discussing using these as regional flowd collectors for SCITAGS

GP4L usage: Tools conception “around” the network ecosystem – Digital twin



Via GP4L Automation & Orchestration

GP4L Netbox Instance

https://nbx.rare.nmaas.eu/dcim/devices/?tag=gp4l



Google

Todos os favo



Organization

Devices

DEVICES

Devices

Modules

Device Roles

Platforms

Virtual Chassis

Virtual Device Contexts

DEVICE TYPES

Device Types

Module Types

Manufacturers

DEVICE COMPONENTS

Search



admin

<input type="checkbox"/>	Name	Status	Tenant	Site	Location	Rack	Role	Manufacturer	Type	IP Address	Tags	ID
<input type="checkbox"/>	AMS0001	Active	GÉANT	AMSTERDAM	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	cern-multione GP4L	1
<input type="checkbox"/>	BUD0001	Active	GÉANT	BUDAPEST	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	cern-multione GP4L	4
<input type="checkbox"/>	CHI0041	Active	STARLIGHT	CHICAGO	—	—	GP4L node	Inventec	D5264	—	GP4L	10
<input type="checkbox"/>	FRA0001	Active	GÉANT	FRANKFURT	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	cern-multione GP4L	3
<input type="checkbox"/>	HND0001	Active	KDDI	TOKYO	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	GP4L	24
<input type="checkbox"/>	MIA0001	Active	FIU	MIAMI	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	GP4L	22
<input type="checkbox"/>	PAR0101	Active	GÉANT	PARIS	—	—	GP4L node	Dell	PowerEdge R430/OCN7X8	—	GP4L	13
<input type="checkbox"/>	POZ0001	Active	GÉANT	POZNAN	—	—	GP4L node	Edge-corE Networks	Wedge 100BF-32X	—	cern-multione GP4L	2
<input type="checkbox"/>	PRA0101	Active	GÉANT	PRAGUE	—	—	GP4L node	Dell	PowerEdge R430/OCN7X8	—	GP4L	14

R&D RNP Netbox Instance

https://nbx.rnp.nmaas.eu/dcim/devices/?tag=rnp-core

Google



Organization

Devices

DEVICES

Devices

Modules

Device Roles

Platforms

Virtual Chassis

Virtual Device Contexts

DEVICE TYPES

Device Types

Search

admin

Devices

+ Add

↑ Import

↓ Export

Results 3

Filters 1

× Tags: RNP-CORE (3)

Save

Quick search

Configure Table

<input type="checkbox"/>	Name	Status	Tenant	Site	Location	Rack	Role	Manufacturer	Type	IP Address	Tags	
<input type="checkbox"/>	JPA	Active	Backbone	PoA-PB	—	—	Core	Juniper	MX	—	RNP-CORE	
<input type="checkbox"/>	PB	Active	Backbone	PoP-PB	—	—	Core	Juniper	MX	—	RNP-CORE	
<input type="checkbox"/>	PE	Active	Backbone	PoP-PE	—	—	Core	Juniper	MX	—	RNP-CORE	

```
dtn01:/nrx# ./nrx.py -c rnp.conf
Connecting to NetBox at: https://nbx.rnp.nmaas.eu/
Fetching devices with tags: rnp-core
Created clab topology: rnp/rnp-core.clab.yaml
To deploy this topology, run: sudo -E clab dep -t rnp/rnp-core.clab.yaml
```

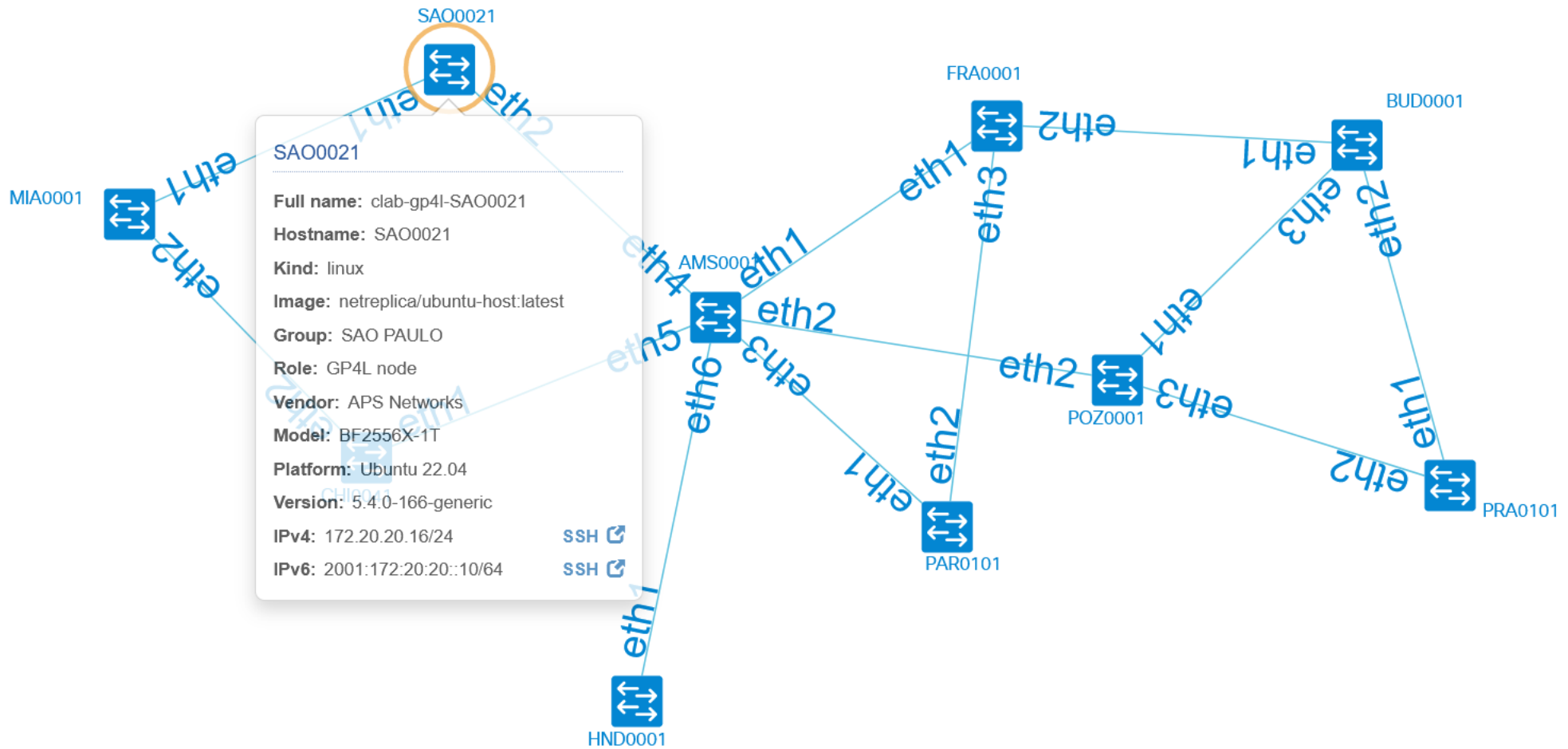
Graphite visualization 🌐 <http://192.168.110.11:8080/graphite>

#	Name	Container ID	Image	Kind
1	clab-rnp-core-JPA	52b3edad0560	crpd:23.1R1.8	crpd
2	clab-rnp-core-PB	20400dc2b2ec	crpd:23.1R1.8	crpd
3	clab-rnp-core-PE	42c8eb125298	crpd:23.1R1.8	crpd
4	clab-rnp-core-graphite	1c8d7bfd6e17	netreplica/graphite:latest	linux

```
dpdadm@dtn01:~/conf/rnp$ docker exec -it clab-rnp-core-PB cli
root@PB> show version
Hostname: PB
Model: cRPD
Junos: 23.1R1.8
cRPD package version : 23.1R1.8 built by builder on 2023-03-19 06:20:36 UTC
```

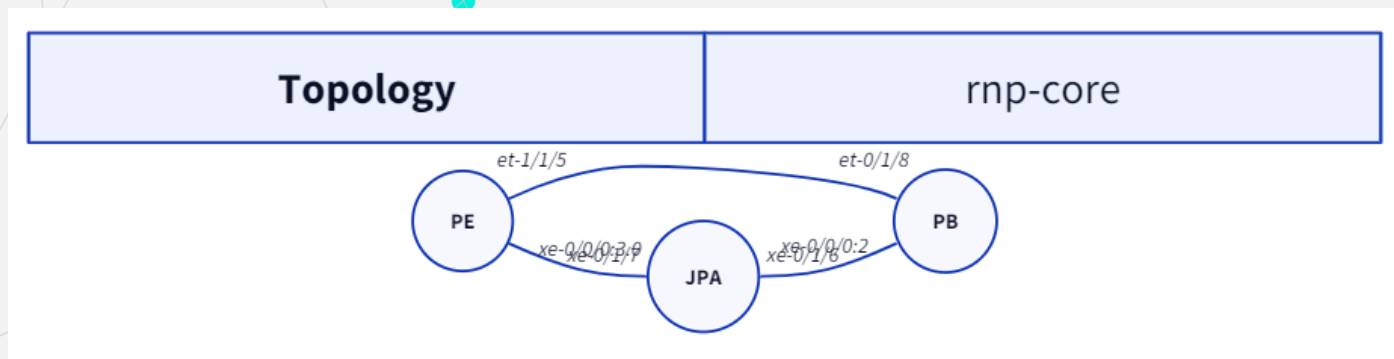
```
dtn01:/nrx# cat rnp/rnp-core.clab.yaml
name: rnp-core
topology:
  nodes:
    graphite:
      kind: linux
      image: netreplica/graphite:latest
    PB:
      kind: crpd
      image: crpd:23.1R1.8
      labels:
        role: Core
        group: PoP-PB
        platform: junos
        vendor: Juniper
        model: MX
    PE:
      kind: crpd
      image: crpd:23.1R1.8
      labels:
        role: Core
        group: PoP-PE
        platform: junos
        vendor: Juniper
        model: MX
    JPA:
      kind: crpd
      image: crpd:23.1R1.8
      labels:
        role: Core
        group: PoA-PB
        platform: junos
        vendor: Juniper
        model: MX
  links:
    - endpoints: ["PE:eth1", "PB:eth1"]
    - endpoints: ["JPA:eth1", "PB:eth2"]
    - endpoints: ["PE:eth2", "JPA:eth2"]
dtn01:/nrx#
```

GP4L Graphite





RNP D2 Diagram



```
1  title: "" {
2    near: top-center
3    grid-columns: 2
4    grid-gap: 0
5    width: 1000
6
7    Topology.width: 500
8    Topology.style.font-size: 32
9    rnp-core.width: 500
10   rnp-core.style.bold: false
11   rnp-core.style.font-size: 32
12 }
13
14 direction: right
15
16 style: {
17   fill: transparent
18 }
19
20 PB: {
21   shape: circle
22 }
23 PE: {
24   shape: circle
25 }
26 JPA: {
27   shape: circle
28 }
29
30 PE -- PB: {
31   source-arrowhead.label: et-1/1/5
32   target-arrowhead.label: et-0/1/8
```

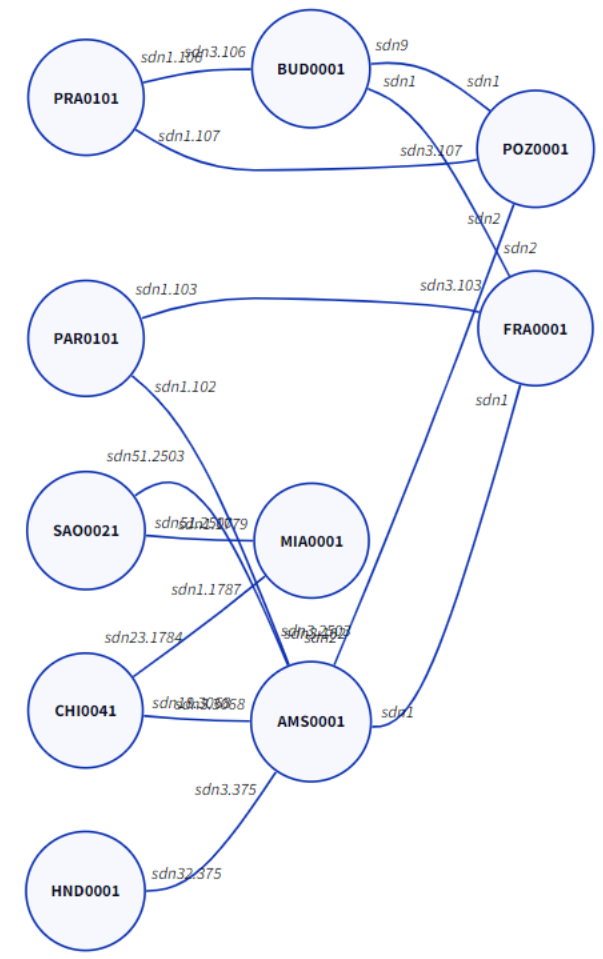

GP4L D2 Diagram

```

51 AMS0001 -- FRA0001: {
52   source-arrowhead.label: sdn1
53   target-arrowhead.label: sdn1
54 }
55 AMS0001 -- POZ0001: {
56   source-arrowhead.label: sdn2
57   target-arrowhead.label: sdn2
58 }
59 BUD0001 -- FRA0001: {
60   source-arrowhead.label: sdn1
61   target-arrowhead.label: sdn2
62 }
63 BUD0001 -- POZ0001: {
64   source-arrowhead.label: sdn9
65   target-arrowhead.label: sdn1
66 }
67 CHI0041 -- AMS0001: {
68   source-arrowhead.label: sdn18.3068
69   target-arrowhead.label: sdn3.3068
70 }
71 CHI0041 -- MIA0001: {
72   source-arrowhead.label: sdn23.1784
73   target-arrowhead.label: sdn1.1787
74 }
75 SAO0021 -- MIA0001: {
76   source-arrowhead.label: sdn51.2500
77   target-arrowhead.label: sdn1.1779
78 }
79 HND0001 -- AMS0001: {
80   source-arrowhead.label: sdn32.375

```

Topology	gp4l
----------	------



— GP4L Progress

Achieved (by SC23)

- Support intercontinental high capacity transfers (100G and over) of next generation protocols, i.e: PolKA
- Creation of on-demand digital twin of GP4L and production networks
- Interconnections with other testbeds (FABRIC) and support for new device types (Bluefield-2)

On Going

- Automated generation of a real time world map dashboards of GP4L
- Capability to support multiple virtual networks, that implement on the same devices different choices of routing stacks, traditional and SDN based
- Integration with initiatives for visibility, controllability and intelligence
- Work as a reference state of the art / next generation R&E network

Thanks! Questions?

Marcos Schwarz
marcos.Schwarz@rnp.br



MINISTÉRIO DO
TURISMO

MINISTÉRIO DA
DEFESA

MINISTÉRIO DA
SAÚDE

MINISTÉRIO DAS
COMUNICAÇÕES

MINISTÉRIO DA
EDUCAÇÃO

MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES

