



RARE and ~~GÉANT~~ (GLOBAL) P4 Lab (aka GP4L)

Ivana Golub (PSNC)

Frédéric Loui (RENATER)

17th SIG-NOC meeting

November 16-17 2022

Paris, France / RENATER

www.geant.org

GÉANT RARE P4 testbed

- GÉANT project environment
- RARE
- GÉANT P4 Lab
- Global P4 Lab
- GP4L Use Cases
- Looking Ahead
- IGP topology monitoring 2.0



The GÉANT Project



GÉANT's vision is to ensure **equal network access for all scientists across Europe** to the research **infrastructures** and the **e-infrastructure resources** available to them.



A part of the European Union's Horizon 2020 research and innovation programme - GÉANT 2020 Framework Partnership Agreement (FPA)



500 contributors from 40 partners - European R&E Institutions



50 M users



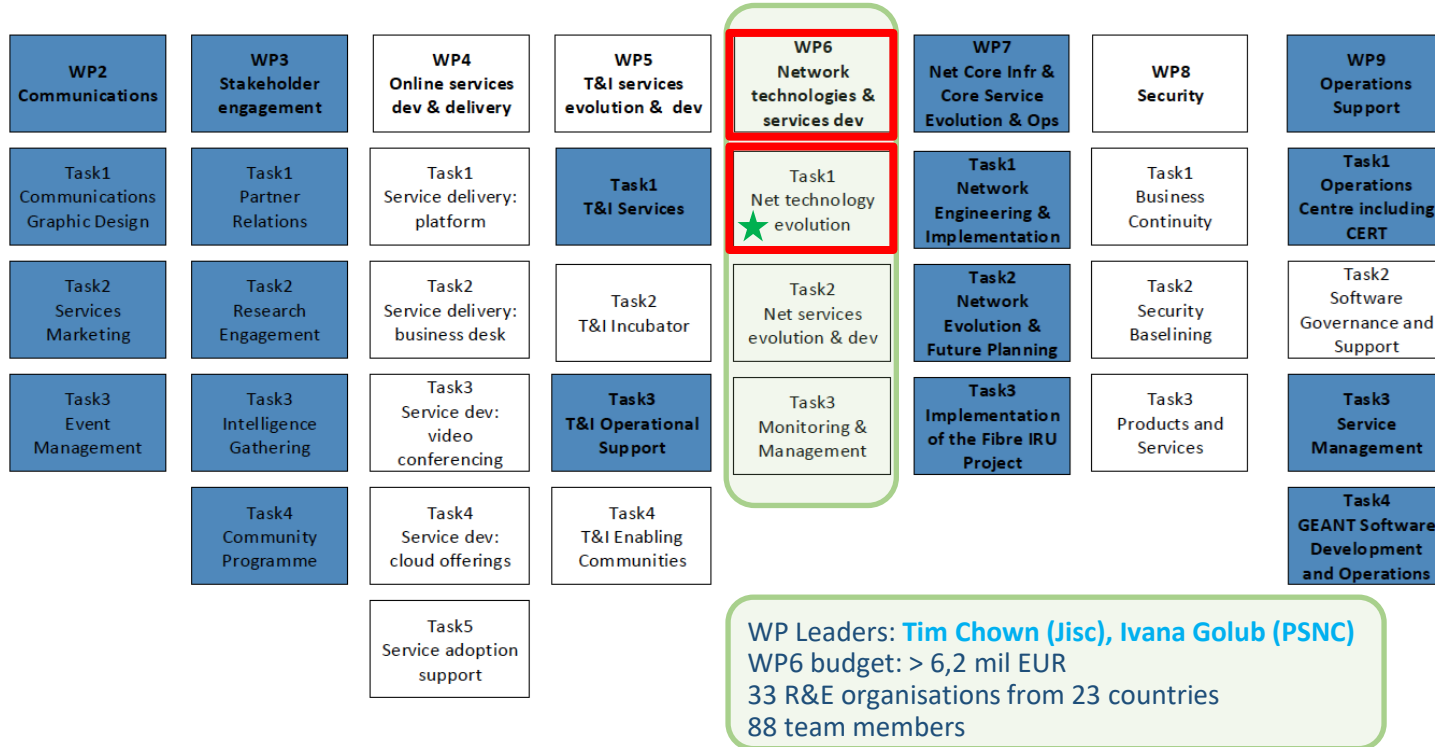
GN4-3 duration: 1 Jan 2019 – 31 December 2022

The GÉANT Project Structure

WP2 Communications	WP3 Stakeholder engagement	WP4 Online services dev & delivery	WP5 T&I services evolution & dev	WP6 Network technologies & services dev	WP7 Net Core Infr & Core Service Evolution & Ops	WP8 Security	WP9 Operations Support
Task1 Communications Graphic Design	Task1 Partner Relations	Task1 Service delivery: platform	Task1 T&I Services	Task1 Net technology evolution	Task1 Network Engineering & Implementation	Task1 Business Continuity	Task1 Operations Centre including CERT
Task2 Services Marketing	Task2 Research Engagement	Task2 Service delivery: business desk	Task2 T&I Incubator	Task2 Net services evolution & dev	Task2 Network Evolution & Future Planning	Task2 Security Baselining	Task2 Software Governance and Support
Task3 Event Management	Task3 Intelligence Gathering	Task3 Service dev: video conferencing	Task3 T&I Operational Support	Task3 Monitoring & Management	Task3 Implementation of the Fibre IRU Project	Task3 Products and Services	Task3 Service Management
	Task4 Community Programme	Task4 Service dev: cloud offerings	Task4 T&I Enabling Communities				Task4 GÉANT Software Development and Operations
		Task5 Service adoption support					



The GÉANT Project Structure



Router for Academia, Research and Education (RARE)

RARE is an open source routing platform, used to create a network operating system (NOS) on commodity hardware (a white box switch).



RARE uses FreeRtr as a control plane software and is thus often referred to as RARE/FreeRtr

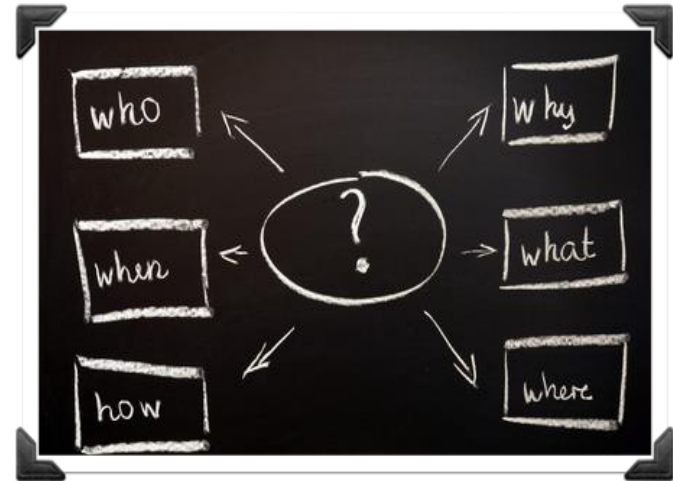


More information:

<https://wiki.geant.org/display/rare>

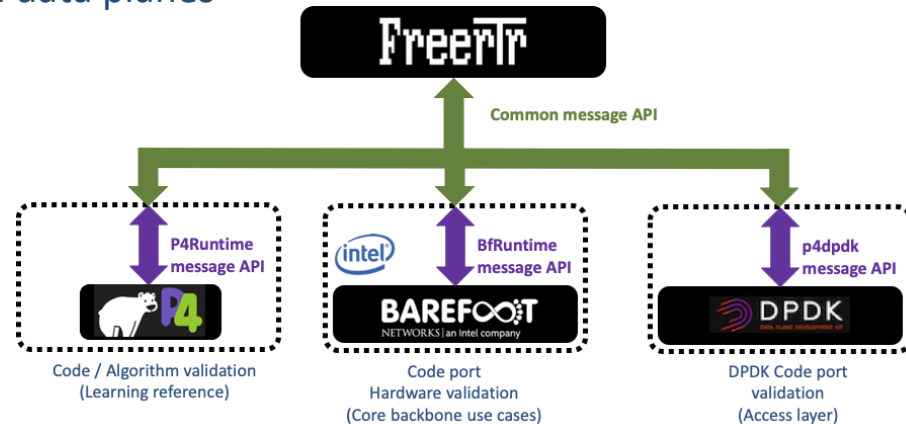
Why RARE?

- Needs of network-aware applications and application-aware network
- Reduce vendor lock-in
- Ability to implement ad-hoc features
- Ability for use-case based solutions
- Reduce digital divide with affordable network solution without functionality trade-off



RARE/FreeRtr Basics

- Free and open source routing platform
- Controls the data plane by managing entries in Match Action Unit (MAU) tables
- Every routed interface must be in a virtual routing table, every layer interface in a bridge table
- One control, several data planes
- Exports control plane computation results to DPDK or hardware switches
- Uses Data Plane Programming (DPP) Language such as **Programming Protocol-independent Packet Processors: P4** language



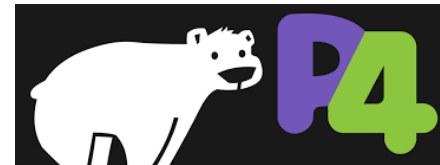
Programming Protocol-independent Packet Processors: P4 language

Language for **programming the data plane** of network devices

- Define how packets are processed
- P4 program structure: header types, parser/deparsers, match-action tables, user-defined metadata and intrinsic metadata

Domain-specific language designed to be implementable on a large variety of targets

- Programmable network interface cards, FPGAs, software switches and hardware ASICs



P4 Programmable Switches

EdgeCore Wedge100BF-32QS:

100GbE Data Center Switch

- Bare-Metal Hardware
- L2/L3 Switching
- 32xQSFP28 Ports

Data-Plane Programmability

- Intel Tofino Switch Silicon
- Barefoot Networks

Quad-Pipe Programmable Packet Processing Pipeline

- 6.4 Tbps Total Bandwidth
- CPU: Intelx86 Xeon 2.0GHz
- 8-core/48GB/2TB SSD



TOFINO 1™
6.4 Tbps



TOFINO 2™
12.8 Tbps
32x400 GE ports



TOFINO 3™
25.6 Tbps
64x400 GE ports

RARE IPv4/IPv6 Features

Include, but not limited to:

- Interior Routing Protocol
- Dataplane forwarding
- External Routing Protocol
- Link local protocol
- Network management

Supported platforms:

- BMv2, TOFINO, DPDK, XDP

List updated regularly:

- <https://wiki.geant.org/display/rare>

For more features or details, contact:

- rare-users@lists.geant.org

Complete feature list

Type	Test #	Name				
acl	01*	copp	✓	✓	✓	✗
acl	02*	ingress access list	✓	✓	✓	✗
acl	03*	egress access list	✓	✓	✓	✗
acl	04*	nat	✓	✓	✓	✗
acl	05*	vlan ingress access list	✓	✓	✓	✗
acl	06*	vlan egress access list	✓	✓	✓	✗
acl	07*	bundle ingress access list	✓	✓	✓	✗
acl	08*	bundle egress access list	✓	✓	✓	✗
acl	09*	bundle vlan ingress access list	✓	✓	✓	✗
acl	10*	bundle vlan egress access list	✓	✓	✓	✗
acl	11*	bridge ingress access list	✓	✓	✓	✗
acl	12*	bridge egress access list	✓	✓	✓	✗
acl	13*	vlan bridge ingress access list	✓	✓	✓	✗
acl	14*	vlan bridge egress access list	✓	✓	✓	✗
acl	15*	ingress pppoe access list	✓	✓	✓	✗
acl	16*	egress pppoe access list	✓	✓	✓	✗
acl	17*	ingress vlan pppoe access list	✓	✓	✓	✗
acl	18*	egress vlan pppoe access list	✓	✓	✓	✗
acl	19*	hairpin ingress access list	✓	✓	✓	✗
acl	20*	hairpin egress access list	✓	✓	✓	✗
acl	21*	hairpin vlan ingress access list	✓	✓	✓	✗
acl	22*	hairpin vlan egress access list	✓	✓	✓	✗
acl	23*	hairpin pppoe ingress access list	✓	✓	✓	✗
acl	24*	hairpin pppoe egress access list	✓	✓	✓	✗
acl	25*	hairpin vlan pppoe ingress access list	✓	✓	✓	✗
acl	26*	hairpin vlan pppoe egress access list	✓	✓	✓	✗
acl	27*	ingress gre access list	✓	✓	✓	✗
acl	28*	egress gre access list	✓	✓	✓	✗
acl	29*	ingress vlan gre access list	✓	✓	✓	✗

GÉANT P4 Lab – GP4L

Initially aimed to **validate the RARE/FreeRtr** open source routing stack software

- 4 switches in Europe: AMS, POZ, FRA, BUD

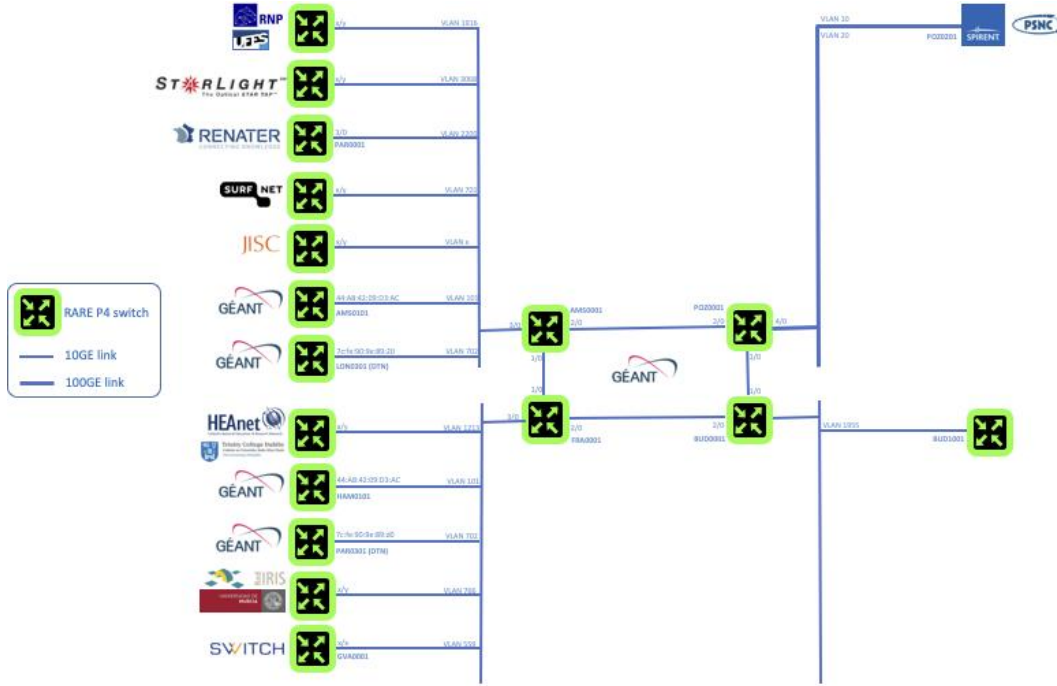
With growing interest, offering **experimental dataplane programming facilities to researchers** to perform geographically distributed network experiments:

- With the usage of RARE/FreeRtr NOS
- Using a clean slate environment (i.e use exclusively GP4L without RARE/FreeRtr dataplane & control plane)

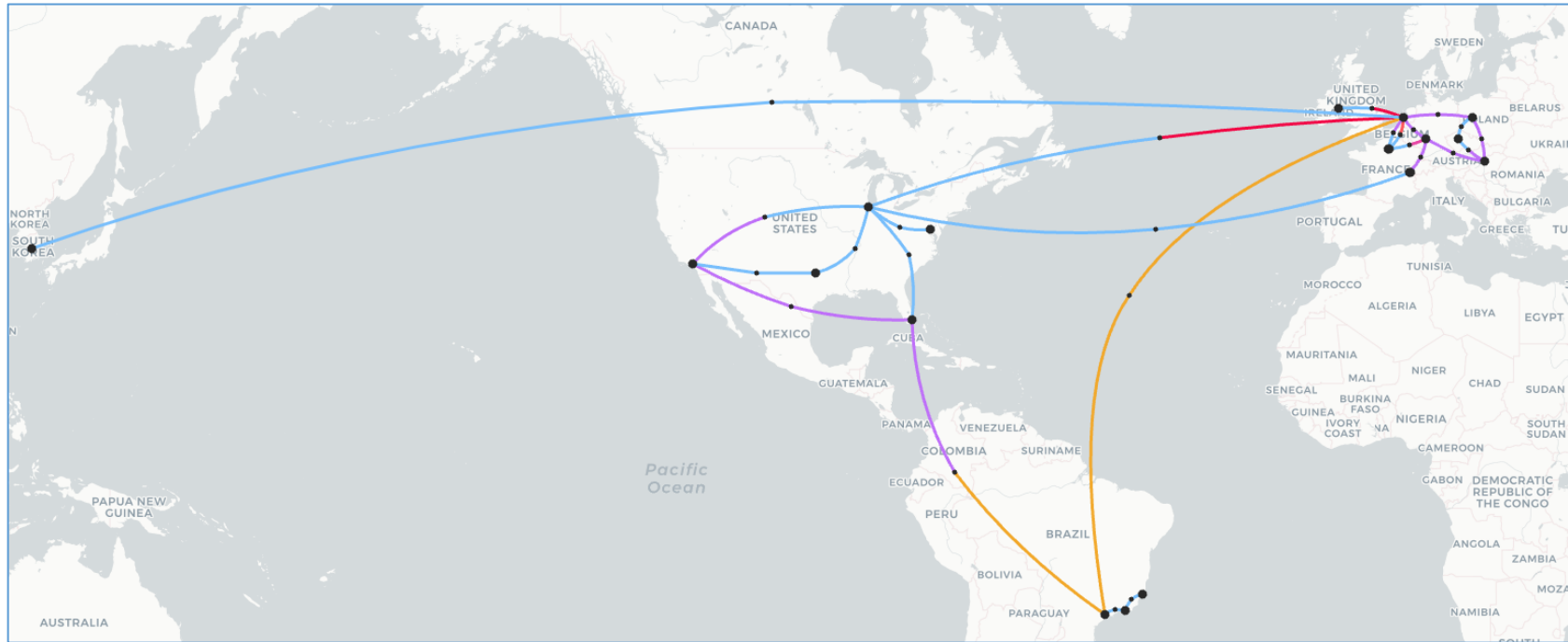


GP4L
GÉANT P4 LAB

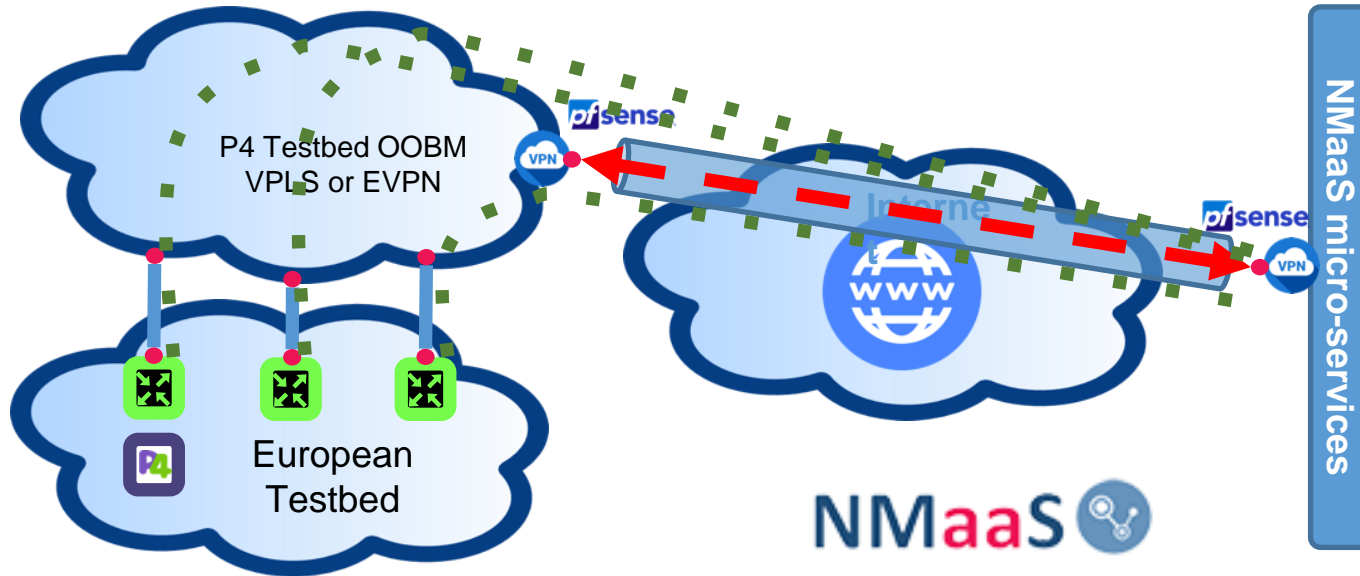
GP4L Going Global



GP4L November 2022



GP4L Monitoring and Management Using GÉANT NMaaS Service



Network Management as a Service:

<https://nmaas.eu>

<https://wiki.geant.org/display/NMaas>

Network Management as a Service (NMaaS)



Network Management as a Service (NMaaS) provides a portfolio of network management applications run as dedicated per-user instances in the cloud.

GÉANT's NMaaS service includes three aspects: providing, managing and maintaining the infrastructure of the NMaaS service portal, platform and selected tools, supporting users in using the system, and the selected tools for monitoring their networks via NMaaS, as well as supporting users that contribute their software to NMaaS system.



Target users

NMaaS users are organisations that do not want to own NMS infrastructure themselves and/or want to outsource network management, as well as organisations and/or individuals that are searching for quality network management software or who want to share their software within the community.

NMaaS Marketplace

NMaaS Marketplace is a catalogue of available open source tools, supported by community, distributed free, chosen by administration. There is also place for your application choice - you can propose new applications.



NMaaS is a platform for network management providing

- A portfolio of network management and monitoring applications
- Per-user, secured network monitoring infrastructure
- Dockerised images implemented through a Kubernetes cluster

NMaaS Usage

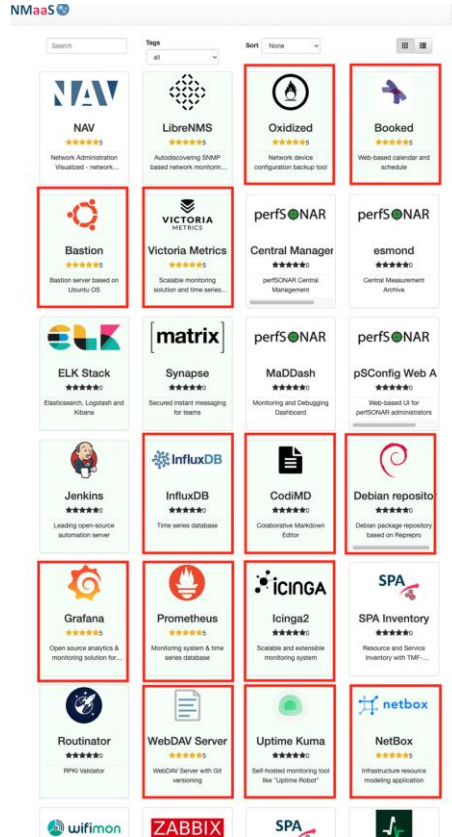
- On GÉANT instances or deployed locally
- NMaaS [sandbox instance](https://nmaas.geant.org/) in GÉANT: <https://nmaas.geant.org/>
- NMaaS [production instance](https://nmaas.eu/) in PSNC: <https://nmaas.eu/>

NMaaS Update

- Version 1.5.1 released
- The work on providing IPv6 support is ongoing
- [NMaaS OAV Architecture Analysis](#) was published



NMaaS Tools Portfolio for GP4L Monitoring and Management



Network Management as a Service:

<https://nmaas.eu>

<https://wiki.geant.org/display/NMaaS>

GP4L Use cases

- Topology Monitoring with BGP-LS
- Next Generation Multicast with AMT relay/gateway and Unicast to Multicast translator, Juniper and Akamai
- Polka - an innovative source routing paradigm, IFES/UFES
- Packet Marking Specification: IPv6 Flow Label, CERN
- SuperComputing22 Demo, GNA-G DIS

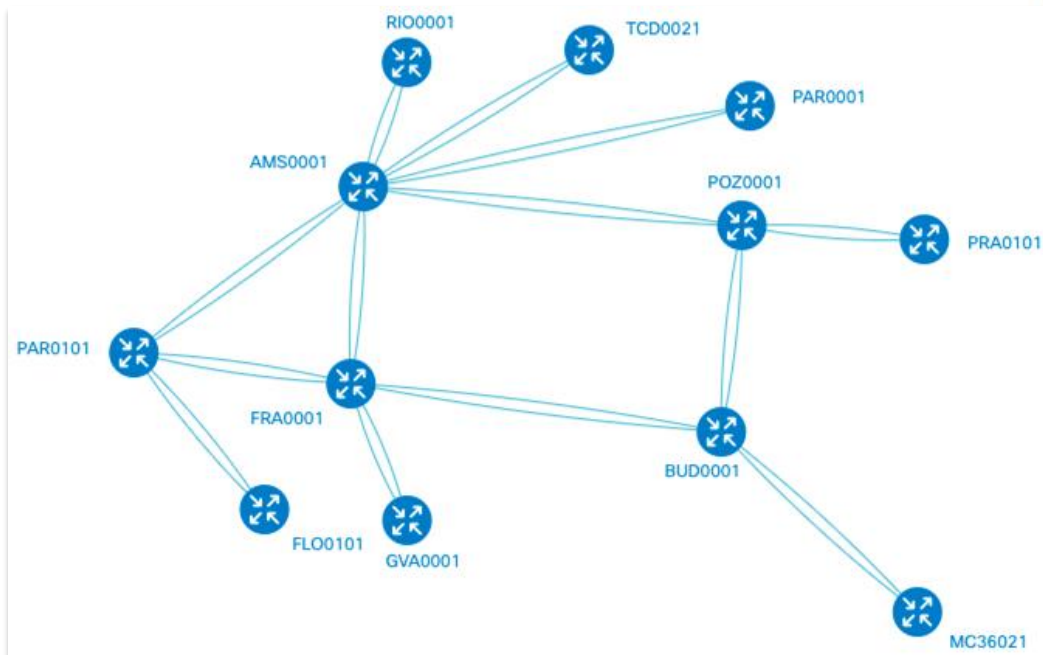
Topology monitoring with BGP-LS

Network topology rendering using BGP-LS

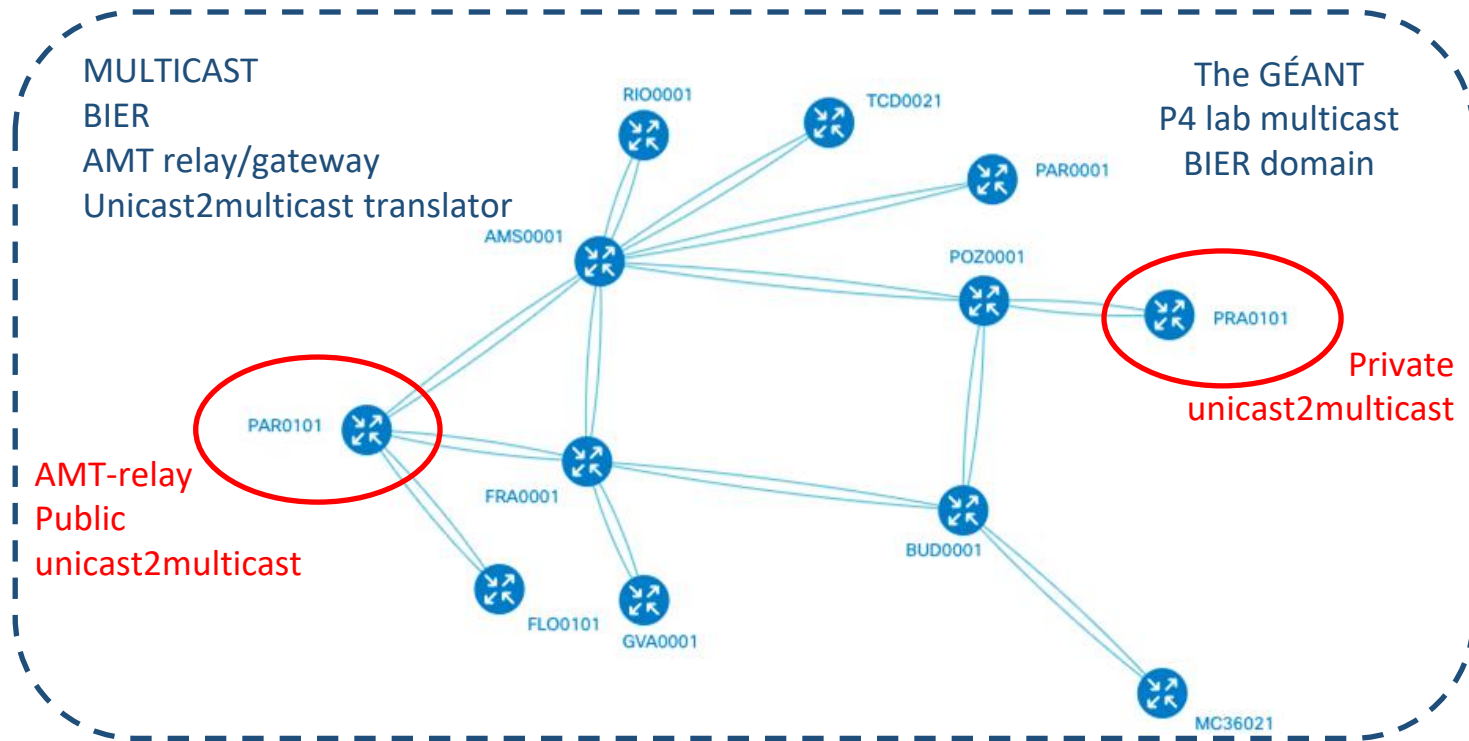
- BGP-LS feed translated to a JSON model
- The model then used to visualise as a map
- Per-minute updates

Available at:

<http://gp4l.geant.org>



GP4L AMT relay / AMT gateway / Unicast --> Multicast



[VLC 4 needed][<https://nightlies.videolan.org/>] -->
vlc <vlc://62.40.109.31@232.123.86.28:1234> --amt-relay amt-relay.geant.org

PolKA - Polynomial Key-based Architecture for Source Routing in Network Fabrics

- GP4L has been used to validate a [Research Paper](#) describing an innovative source routing paradigm: [Polka](#)
- After successful publication of Polka paper, it has been decided to implement this routing paradigm to RARE/FreeRtr routing stack

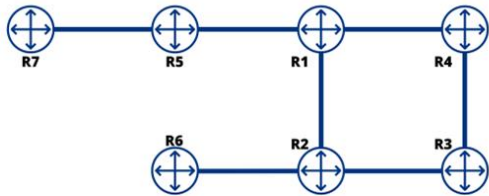


Figure 3. Edge-Core Experiment

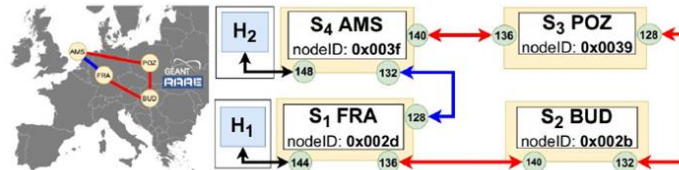
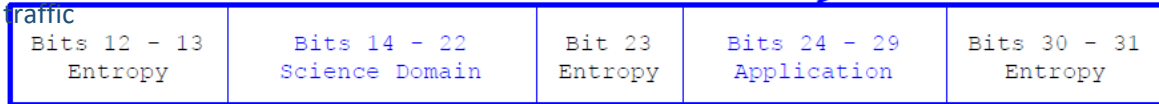


Figure 4. RARE/GEANT testbed

Figure source: <https://sol.sbc.org.br/index.php/wpeif/article/download/21490/21314/>
by Federal Institute of Education Science and Technology of Espírito Santo, and
Federal University of Espírito Santo, Espírito Santo, Brazil

Packet Marking Specification: IPv6 Flow Label

- A packet marking technique proposed by the Research Network Technology WG
- Identifying the LHC experiment and the application that has generated a transmission packet
- The Experiment-Application tag inserted in the IPv6 packet header flow label field
- Primary goal: traffic count, but special routing polices could be applied
- Flow label field of IPv6 header: 20 bits
 - 5 entropy bits to match RFC 6436
 - 9 bits to define the science domain
 - bits to define the application/type of traffic



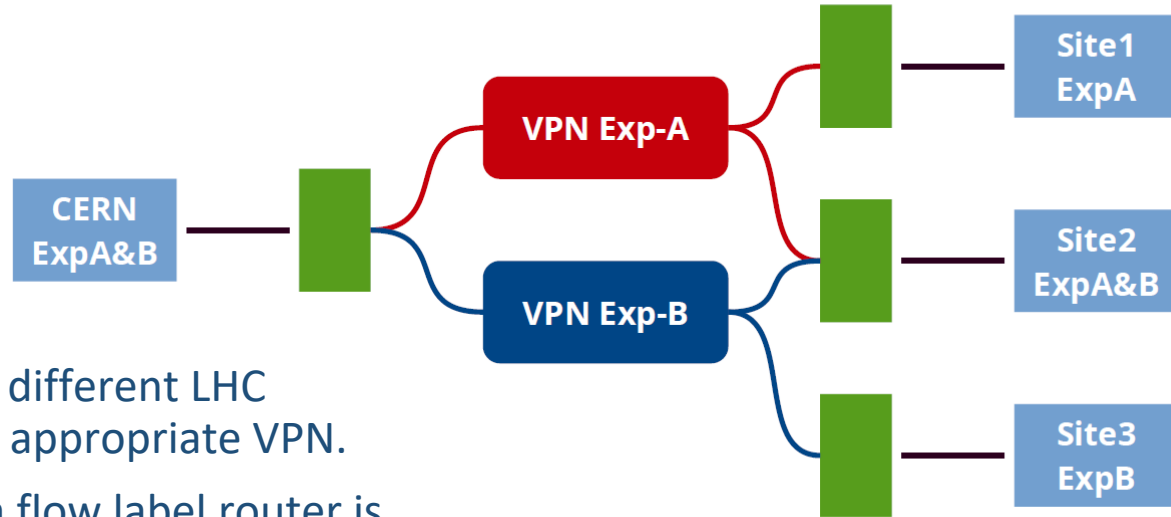
Astro/HEP Science Domains:

- Reserved - 0
- Default - 65536
- ATLAS - 32768
- CMS - 98304
- LHCb - 16384
- ALICE - 81920
- BelleII - 49152
- SKA - 114688
- LSST - 73728
- DUNE - 8192

Application:

- Reserved - 0
- Default - 4
- perfsONAR - 8
- Cache - 12
- DataChallege - 16

MultiONE multiple “LHCONES”: Traffic separation with IPv6 flow labels

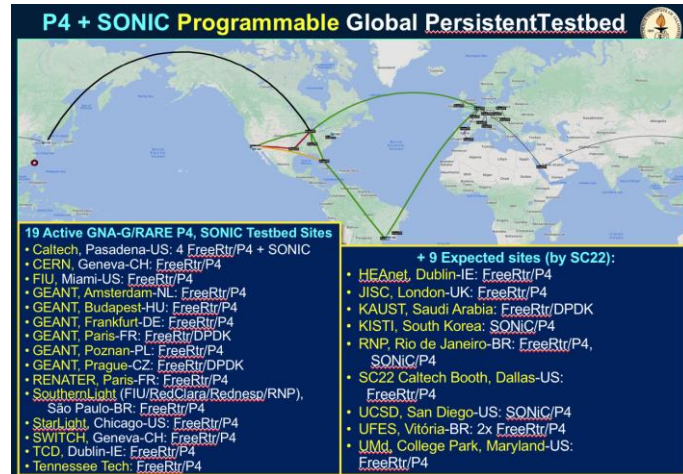


Routing traffic of the different LHC experiments into the appropriate VPN.

- A prototype of a flow label router is being developed using a P4 programmable switch (EdgeCore Wedge100BF-32QS with an Intel Tofino processor)

SuperComputing22 Demo

Over **20** locations expected in the SC22 **Global P4Lab**, including the GÉANT P4 Lab
Several areas in scope: Visibility, Intelligence, Controllability, NOS and tools, Orchestration
In collaboration with the GNA-G Data Intensive Sciences Working group – **GNA-G DIS WG**



SC22: Global Petascale to Exascale Workflows for Data Intensive Sciences

- **Advances Embedded and Interoperate** within a 'composable' architecture of subsystems, components and interfaces, organized into several areas:
- **Visibility:** Monitoring and information tracking and management including IETF ALTO/OpenALTO, BGP-LS, sFlow/NetFlow, Perfsonar, Traceroute, Qualcomm Gradient Graph congestion information, Kubernetes statistics, LibreNMS, P4/Inband telemetry
- **Intelligence:** Stateful decisions using composable metrics (policy, priority, network- and site-state, SLA constraints, responses to 'events' at sites and in the networks, ...), using NetPredict, Hecate, RL-G2, Yale Bilevel optimization, Coral, Elastiflow/Elastic Stack
- **Controllability:** SENSE/OpenNSA/AutoGOLE, P4/PINS, segment routing with SRv6 and/or PolKA, BGP/PCEP
- **Network Oses and Tools:** GEANT RARE/freeRtr, SONIC, Calico VPP, Bstruct-Mininet environment, ...
- **Orchestration:** SENSE, Kubernetes (+k8s namespace), dedicated code and APIs for interoperation and progressive integration

RARE/freeRtr and GP4L at 3rd GRP (Salt Lake City)

(Thanks Ivana !)

GEANT (Now Global) P4 Lab
and Dataplane of P4 Programmable Switches

Frederic Loui,
Casaba Mate,
Marcos Schwarz
et al.

A new worldwide platform for

- New agile and flexible feature development
- New use case development corresponding to research programs' requirements
- Multiple research network overlay slices

A global playing field for

- Next generation network monitoring tools and systems
- Development of fully automated network deployment
- New network operations paradigm development

VISION: Federate and integrate multiple testbeds and toolsets such as AutoGOLE / SENSE



Looking ahead



**Validate your use case
with GP4L!**

Orchestrate and automate GP4L:

Lab reservation

Persistent testbed interaction at global scale

New hardware:

TOFINO2, NVIDIA DPU, P4 SmartNIC, TOFINO/FPGA

Global worldwide footprint:

Interconnection with other persistent testbed

🔍 New idea:

Validate new use cases

Focus on use case scalability

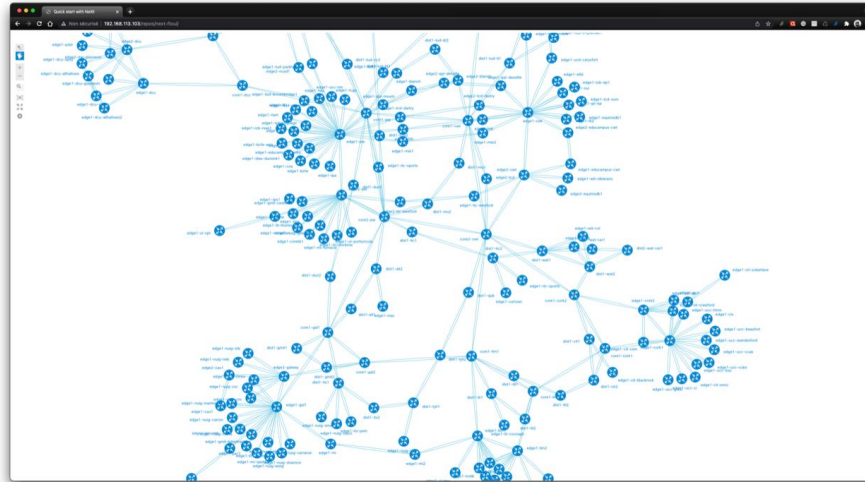
100/400 GE DTN automation

Control plane scalability

And more ...

IGP topology monitoring 2.0

- **IGP topology monitoring**
 - Next generation alarm notification system
 - Map physical circuit when it maps IGP topology
 - Applicable to 1 or ... thousands of nodes



IGP topology monitoring 2.0 (Recipe in 3 steps)

- **Deploy RARE/freeRtr in your network**
 - 1 server or VM only
 - Insert it into your IGP
- **On RARE/freeRtr CLI:**
 - `server nrpe <my_server>`
 - Define event you are interested globally:
 - `check <my_check>`
- **On your NRPE system (ICINGA2)**
 - Configure the notification event matching `<my_check>`
 - Bind to your favorite communication client available from your NRPE system
 - Slack, Matrix client, email. SMS, phone etc.

IGP topology monitoring 2.0

- **MUST** have:
 - if you don't have existing alarms notification system
 - **Add** an additional non intrusive lightweight alarm notification system
 - **Replace** your existing alarm system by a lightweight and cost effective system
 - You want to redirect alarm to specific groups ...

Cost effective (almost 0 CAPEX) !!!!!

Useful Links

Documentation:

GP4L project: <https://wiki.geant.org/display/GP4L/>

RARE/FreeRtr: <https://wiki.geant.org/display/RARE>

<https://blog.freertr.org>

<https://docs.freertr.org>

<https://blog.freertr.org>

GÉANT NETDEV: <https://wiki.geant.org/display/NETDEV>

Contact:

Users: gp4l-users@lists.geant.org, rare-users@lists.geant.org

Developers: gp4l-dev@lists.geant.org, rare-dev@lists.geant.org

Project: gp4l@lists.geant.org, rare@lists.geant.org

Thank you

Any questions?

Email: *netdev@lists.geant.org*

www.geant.org

