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

Ivana Golub (PSNC)
Frédéric Loui (RENATER)

27 STF meeting
October 19-20 2022
Zurich, Switzerland / SWITCH

www.geant.org
GÉANT RARE P4 testbed

- GÉANT project environment
- RARE
- GÉANT P4 Lab
- Global P4 Lab
- GP4L Use Cases
- Looking Ahead
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
The GÉANT Project Structure

WP Leaders: Tim Chown (Jisc), Ivana Golub (PSNC)
WP6 budget: > 6.2 mil EUR
33 R&E organisations from 23 countries
88 team members
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/deparser, 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
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](https://wiki.geant.org/display/rare)

For more features or details, contact:

- [rare-users@lists.geant.org](mailto:rare-users@lists.geant.org)
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 Going Global
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)

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 in GÉANT: https://nmaas.geant.org/
• NMaaS production instance 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

MULTICAST
BIER
AMT relay/gateway
Unicast2multicast translator

The GÉANT P4 lab multicast
BIER domain

AMT-relay
Public
unicast2multicast

Private
unicast2multicast

[VLC 4 needed][https://nightlies.videolan.org/] -- vlc amt://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 a 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 policies 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
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**

---

**P4 + SONIC Programmable Global Persistent Testbed**

- **Active GNA-G/RARE P4, SONIC Testbed Sites**
  - Caltech, Pasadena-US: FreeB/E P4 + SONIC
  - CLRC, Daresbury-UK: FreeB/E P4
  - CERN, Geneva-CH: FreeB/E P4
  - FIU, Miami-FL: FreeB/E P4
  - GÉANT, Amsterdam-NL: FreeB/E P4
  - GEANT, Budapest-HU: FreeB/E P4
  - GEANT, Frankfurt-DE: FreeB/E P4
  - GEANT, Paris-FR: FreeB/E P4
  - GEANT, Geneva-CH: FreeB/E P4
  - GEANT, Prague-CZ: FreeB/E P4

- **Expected sites (by SC22):**
  - HEAdl, Dublin-IE: FreeB/E P4
  - JIRC, London-UK: FreeB/E P4
  - KAUST, Saudi Arabia: FreeB/E P4
  - KISTI, South Korea: SONIC/P4
  - RNP, Rio de Janeiro-BR: FreeB/E P4
  - SONIC/P4
  - SC22 Caltech Booth, Dallas-US: FreeB/E P4
  - UCSD, San Diego-US: SONIC/P4
  - UNC, Chapel-Hill-NC: FreeB/E P4
  - UMD, College Park-MD: FreeB/E P4
  - University of Texas, Austin-TX: FreeB/E P4
  - Tennessee Tech: FreeB/E P4

---

**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/AutoGOL, P4/PINS, segment routing with SRv6 and/or PolKA, BGP/PCEP
  - **Network OSEs and Tools:** GÉANT RARE/freeRtr, SONIC, Calico VPP, Bstruct-Mininet environment, ... 
  - **Orchestration:** SENSE, Kubernetes (+K8s namespace), dedicated code and APIs for interoperation and progressive integration

---

**Slides:** by courtesy of Harvey Newman, Caltech

www.geant.org
RARE/freeRtr and GP4L at 3rd GRP (Salt Lake City)
(Thanks Ivana !)
Looking ahead

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 ...

Validate your use case with GP4L!
Useful Links

Documentation:
- **GP4L project:** [https://wiki.geant.org/display/GP4L/](https://wiki.geant.org/display/GP4L/)
- **RARE/FreeRtr:** [https://wiki.geant.org/display/RARE](https://wiki.geant.org/display/RARE)
  [https://blog.freertr.org](https://blog.freertr.org)
  [https://docs.freertr.org](https://docs.freertr.org)
  [https://blog.freertr.org](https://blog.freertr.org)
- **GÉANT NETDEV:** [https://wiki.geant.org/display/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