RARE routing platform

Potential usage in data-centres network

LOUI Frédéric
GÉANT/RENATER – RARE technical leader

MATE Csaba
GÉANT/KIFU – RARE/freeRtr lead core developer

WLCG Data Center network architecture

7–8th 2021

Public

www.geant.org
RARE project: Group focus

- funded project
  - Control plane software
  - **Programmable** dataplane
  - Interface them and the result is ...

- Feature rich routing platform
  - various hardware line rate
  - Flexible, DIY “hackable/extensible” router
  - Control plane independence

One familiar platform

→ Multiple solutions

→ Each solution addresses

→ R&E

use case
Why RARE now?

• Starting from early 2010:
  • Several valuable Open Source control plane usage besides well know commercial vendor

  BIRD  Exa  VyOS  Quagga

• Starting from 2020:
  • Dataplane solution reached maturity ready to implement production grade use case

  DPDK  Open vSwitch

• NOS emergence

  ARISTA  DELL  CISCO  JUNIPER

  Quagga  SONiC  FRRouting  FreerTr

• Hypervisor Technology convergence

  openstack  kubernetes

It’s a good time to tie Control Plane and Dataplane!
RARE use cases

IPv4 and IPv6 compliant!
Anatomy of a typical R&E worldwide research project #1

High speed Research & Education Network

Data processing computing center

Research project’s Instruments

Eyeballs
Anatomy of a typical R&E worldwide research project #2
RARE is for everyone

- Routing (CP+DP) platform solution
  - Open Platform
  - Programmable

- RARE for Research and Education connectivity
  - Emerging NREN
  - Or not ...

- RARE for content provider DCI
  - IaaS owned by NREN
  - IaaS owned by International Global Research project

- RARE for end user institution
  - Primary/Secondary schools
  - University campus
  - MAN network for Regional network

- RARE for International Global research project connectivity
  - Network research
  - Science research

Positive societal consequences!
RARE latest news (Month 29 of 48)

• RARE p4 targets
  bmv2 software switch
  Programmable Ethernet ASIC on WEDGE-BF100-32X
  under study

• RARE p4 discussion emulation targets
  TCPDUMP & Libpcap
  DPDK
RARE “target” development

- Code / Algorithm validation (Learning reference)
- Code port Hardware validation (Core backbone use cases)
- DPDK Code port validation (Access layer)
FreerTr architecture

Dataplanes:
- p4emu/dpdk, p4emu/libpcap, bmv2, tofino
- maybe npl in the future, anything else? :)

Clockwise:
- ethernet with "cpu_header"
- socat iface:ens3 udp:127.0.0.1
  - packets in udp
  - eth0
- freertr: router processes, mgmt, "server p4lang p4", etc
- grpc
- text<--grpc
- human readable text in tcp

Counters:
- nothing@p4emu
- table-updates

www.geant.org
<table>
<thead>
<tr>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing framework</td>
<td>$\sim 2300$ features = 2300 tests</td>
</tr>
</tbody>
</table>
RARE testing framework: Dataplane tests ~300 tests

<table>
<thead>
<tr>
<th>Type</th>
<th>Test #</th>
<th>Name</th>
<th>P4</th>
<th>∞</th>
<th>DPDK</th>
</tr>
</thead>
<tbody>
<tr>
<td>acl</td>
<td>01^a</td>
<td>copp</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>02^a</td>
<td>ingress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>03^a</td>
<td>egress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>04^a</td>
<td>nat</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>05^a</td>
<td>vlan ingress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>06^a</td>
<td>vlan egress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>07^a</td>
<td>bundle ingress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>08^a</td>
<td>bundle egress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>09^a</td>
<td>bundle vlan ingress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>10^a</td>
<td>bundle vlan egress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>acl</td>
<td>11^a</td>
<td>bridge ingress access list</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

And more features!

Please come @IRC #freertr and submit your idea!
RARE validation designs: P4 LAB network management via NMaaS!* (Network Management as a Service)

Network Management as a Service:
https://nmaas.eu
https://wiki.geant.org/display/NMAAS
P4 LAB network management via (Network Management as a Service)

Network Management as a Service: https://nmaas.eu
https://wiki.geant.org/display/NMAAS
Monitoring at node level! (Prometheus agent)

More API/Agents are available!
Monitoring at node level! (Grafana dashboard)

https://grafana.com/grafana/dashboards?search=freeRouter
Key take-away – We are ready to roll into production

• Automated testing
• 3rd party testing via Spirent usage
  • (thanks PSNC@WB team)
• P4 profile calibration for **only**
• DPDK currently in operation SOHO
• Production deployment

• Work in progress production deployment
Let’s get practical …
and present you
actual real life use cases!
Practical use case #001 100GE DC network

Upstream Service Provider
Providing inter-DCI

Virtual gateway

Spine

Leaf

ToR

www.geant.org
Practical use case #001a ToR

- 2x100GE uplink toward LEAF nodes
- 30x100GE server blade connection
- Each server powered by DPDK NIC
Practical use case #001b BNG / Virtual Gateway DC

- DPDK and P4 dataplane
  - suitable for CAMPUS / EDGE BACKBONE router
- nx1GE, nx10GE, nx100GE
Practical use case #001c LSR or SPINE router

- P4 dataplane fits perfectly pure LSR/SPINE router
- NNI: 4 directions with (8x100GE) bundle
Practical use case #001d LER or LEAF router

• P4 dataplane fits perfectly pure LER/LEAF use case
• NNI: EST/WEST direction @ (8x100GE) bundle
• UNI: 16x100GE left for end user connection!
Practical use case #001e high performance BGP RR

• Recycling old/new server?
• Kubernetes cluster using BGP as CNI network plugin
• Taking advantage of server « huge » amount of RAM
• No need specific high performance dataplane
Practical use case #001f « small PE/L3 ToR »

Ideal for aggregation

- 2x10GE or 2x100GE NIC server side
- 2x10g+48x1g or 1x100g+48x1/10g switch
Practical use case #002 SOHO router

- DPDK flavor ideal for CPE
- nx1GE
- nx10GE small MAN ideal for small campus
- Couple of 100GE (Depending on server generation)
Practical use case #003 100GE Private Peering node

- High resilient **Packet core**
  - 2 direction @ 400Gb / 1,6 Tbps

- User ports connection
  - 24 ports left for 2x12 redundant Private peering
  - 1:3 ratio with redundant scenario
Practical use case #xxx The sky is the limit

- Automation integration
- IXP with MPLS core
- ToR router combined to BGP aware network plugin
- Spine/Leaf DC router
- Global BGP monitoring for your entire BGP domain
- Global IGP guard for your entire IGP domain
- BGP flowspec aware anti DDOS
- AAA servers (TACACS, RADIUS)
- ...

We need **YOUR** creativity!
Key take-away – Final words – RARE vision

• Open Network programming opportunity
  • R&E small institution
  • R&E global project (100GE is real, 400GE just landed)

• Opportunity to define Node/Flow Network Monitoring
  • Scaling new NMS (horizontal scaling with K8s)
  • Streaming Telemetry
  • INT
    ➔ Rethink how Network Management is handled

• Opportunity to integrate existing automation initiatives

**Instantaneous & Flexible**

**Network Services for the users!**
Acknowledgements ...
Useful links

• Project
  freeRtr control plane’s home: freertr.net
  more information on dataplanes: rare.freertr.net
  Project members’ journey: blog.freertr.net
  FreeRtr configuration guide: docs.freertr.net

• Contact
  For daring RARE/freeRtr users: rare-users@lists.geant.org
  For RARE/freeRtr JEDI developer wanabee: rare-dev@lists.geant.org
  For RARE/freeRtr supporters: @rare_freerouter
  IRC@DN42 #freertr
Useful links: Source code!!!!!

freeRtr core: sources.nop.hu/src/

TOFINO ASIC: sources.nop.hu/misc/p4bf/

P4Lang bmv2: sources.nop.hu/misc/p4lang/

p4emu: sources.nop.hu/misc/native/p4*

p4dpk: sources.nop.hu/misc/native/p4*

Breaking news!
A RARE/freeRtr NOS has been developed
And ... We are allowed to distribute it!
Looking ahead: Finalize transition to production

Join the RARE project!

Extend HCL:
- new TOFINO based hardware support
- new DPDK release

New target:
- TOFINO2
- DPU
- SmartNIC
- FPGA

New idea:
- Polka
- P42VPP
- T4P4S ELTE
- Leverage Nix paradigm

And more …
Last word: Worldwide End to end dynamic path?

kubernetes + FreerIn
Workers node
As a CNI plugin

Inter-DCI
End to end LSP!

www.geant.org
Thank you

Any questions?

www.geant.org