Router for Academia Research Education

RARE/freeRtr in a nutshell

LOUI Frédéric
GÉANT/RENATER – RARE technical leader
MATE Csaba
GÉANT/KIFU – RARE/freeRtr lead core developer

P4 Workshop 2021

May 18th 2021

www.geant.org

RARE project : Group focus

funded project

- Control plane software
- ~~P4~~ (**not only P4**) **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

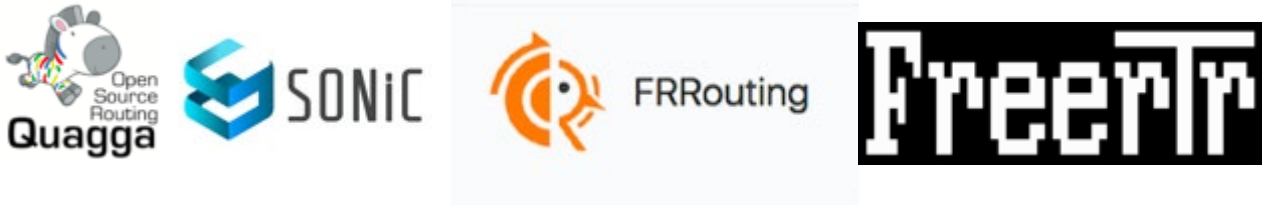
use case

Why RARE now?

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

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

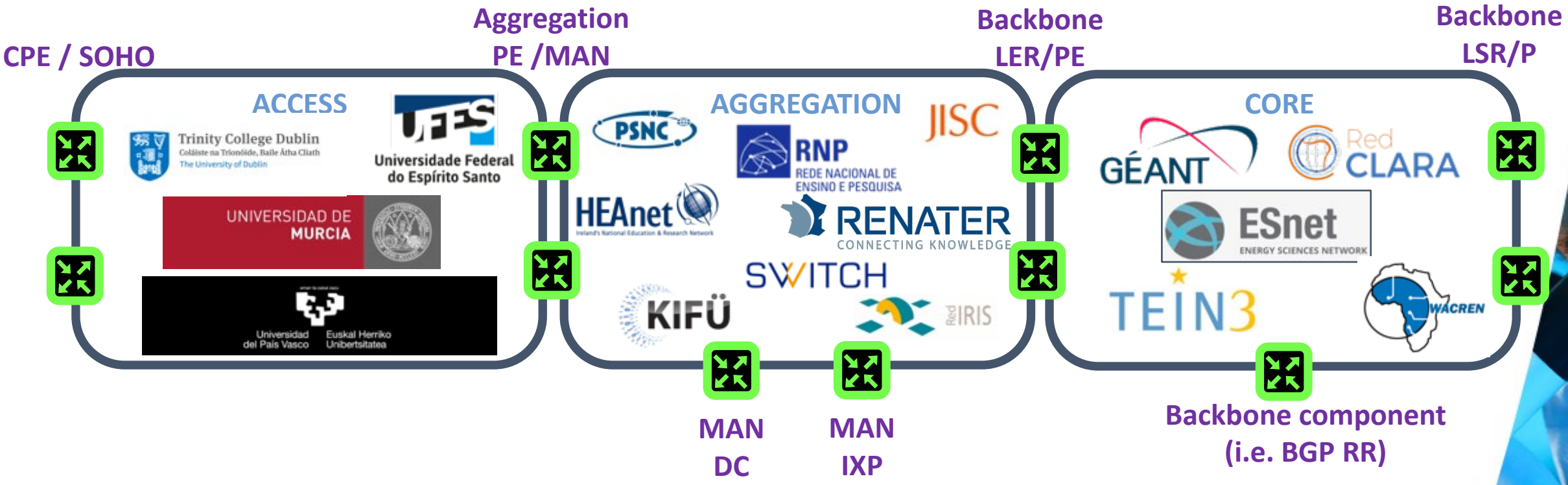
- NOS emergence



- Technology convergence (Hypervisor/VM, K8s/Container, kernel bypass …)

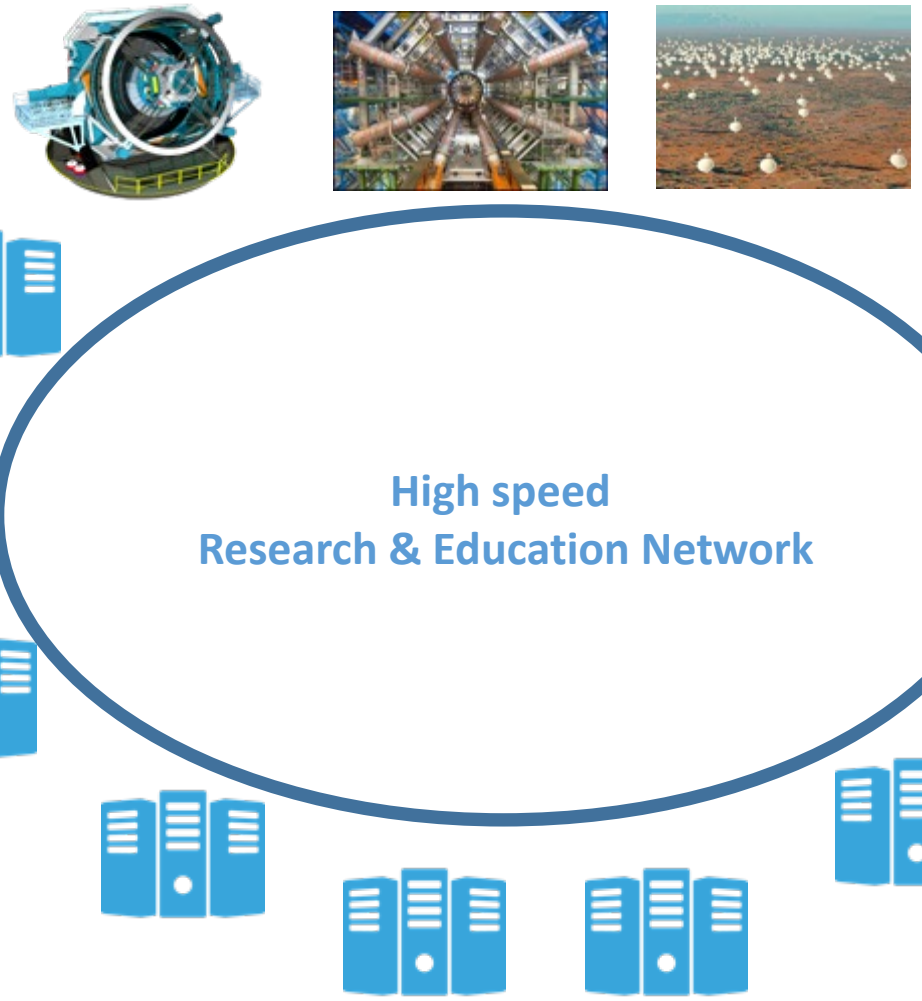
It’s a good time to tie Control Plane and Dataplane!

RARE use cases



IPv4 and IPv6 compliant!

Anatomy of a typical R&E worldwide research project #1



Anatomy of a typical R&E worldwide research project #2

Northwestern University

RENATER
CONNECTING KNOWLEDGE

SWITCH

aarnet
Australia's Academic and Research Network

Caltech

ESnet
ENERGY SCIENCES NETWORK

INTERNET2

GÉANT

TEIN3

NICT

Red CLARA

WACREN

UbuntuNet Alliance

RedCLARA
Ciencia y Educación en Red

RNP
REDE NACIONAL DE ENSINO E PESQUISA

NgREN
Nigerian Research and Education Network

SANReN
South African National Research Network

www.geant.org

GÉANT

RARE is for everyone

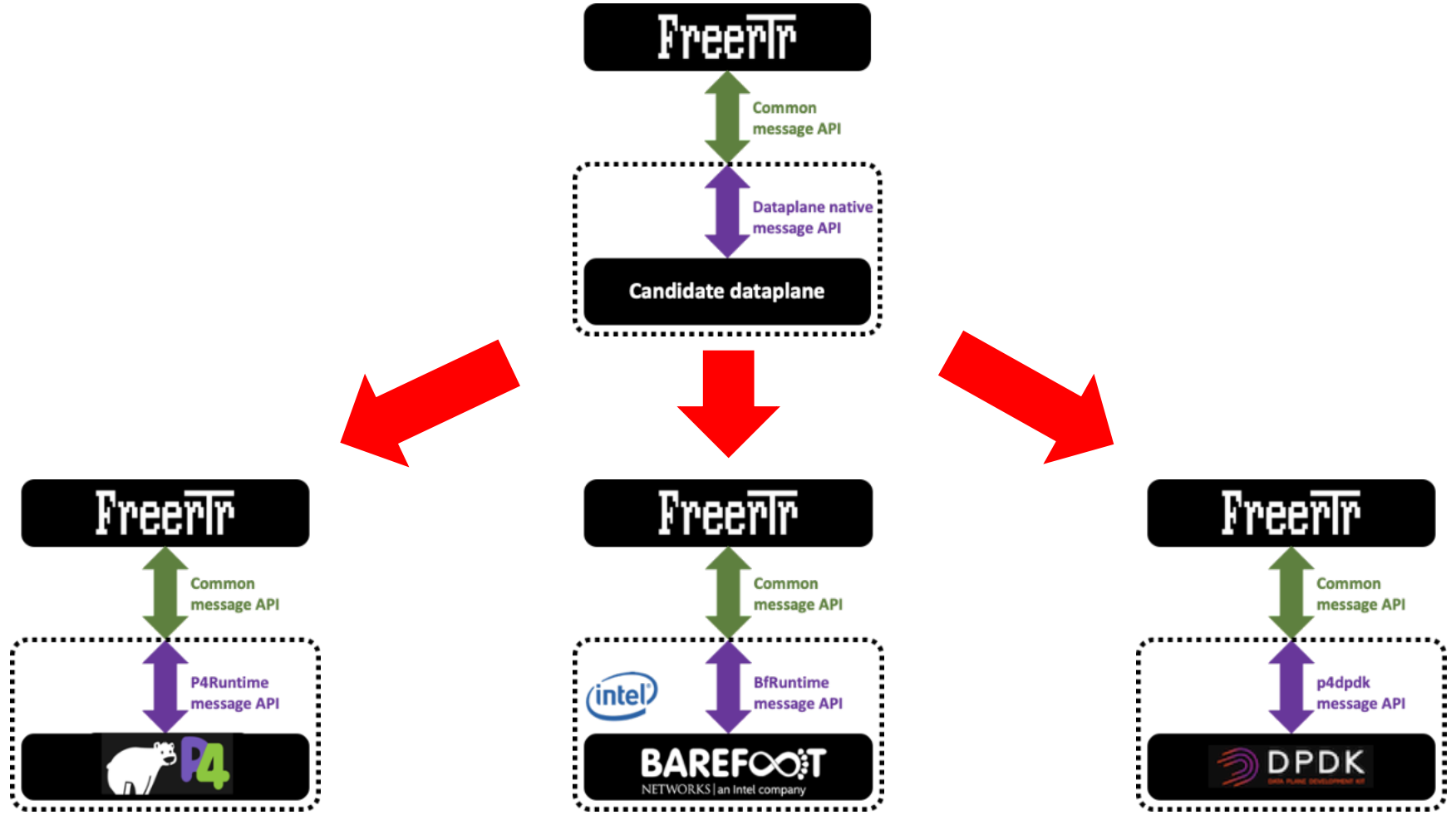
- **Routing (CP+DP) platform solution**
  - Open Platform
  - Programmable
- **RARE for Research and Education connectivity**
  - Emerging NREN
  - Or not …
- **RARE for content provider DCI**
  - IaaS owned by NREN
  - IaaS owned by International Global Research project
- **RARE for end user institution**
  - 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

RARE “target” development



Code / Algorithm validation (Learning reference)

Code port Hardware validation (Core backbone use cases)

DPDK Code port validation (Access layer)

RARE testing framework : ~ 2300 features = 2300 tests

crypt-skip12.tst	success	skip with sha1
crypt-skip13.tst	success	skip with sha256
crypt-skip14.tst	success	skip with sha512
crypt-ssh.tst	success	ssh test
crypt-swipe01.tst	success	swipe over ipv4
crypt-swipe02.tst	success	swipe over ipv6
crypt-swipe03.tst	success	swipe over swipe
crypt-swipe04.tst	success	swipe over loopback
crypt-swipe05.tst	success	swipe with des
crypt-swipe06.tst	success	swipe with blowfish
crypt-swipe07.tst	success	swipe with 3des
crypt-swipe08.tst	success	swipe with aes128
crypt-swipe09.tst	success	swipe with aes192
crypt-swipe10.tst	success	swipe with aes256
crypt-swipe11.tst	success	swipe with md5
crypt-swipe12.tst	success	swipe with sha1
crypt-swipe13.tst	success	swipe with sha256
crypt-swipe14.tst	success	swipe with sha512
crypt-tls.tst	success	tls test
crypt-wireguard01.tst	success	wireguard over ipv4
crypt-wireguard02.tst	success	wireguard over ipv6
crypt-wireguard03.tst	success	wireguard over wireguard
crypt-wireguard04.tst	success	wireguard over loopback
crypt-wireguard05.tst	success	wireguard over asymmetric ports
demo01.tst	success	empty demo network
demo02.tst	success	addressed demo network
intop1-bgp01.tst	success	interop1: ebgp
intop1-bgp02.tst	success	interop1: ibgp
intop1-bgp03.tst	success	interop1: bgp locpref
intop1-bgp04.tst	success	interop1: bgp origin
intop1-bgp05.tst	success	interop1: bgp metric
intop1-bgp06.tst	success	interop1: bgp community
intop1-bgp07.tst	success	interop1: bgp aspath
intop1-bgp08.tst	success	interop1: bgp with labels
intop1-bgp09.tst	success	interop1: bgp addpath
intop1-bgp10.tst	success	interop1: bgp prefix withdraw
intop1-bgp11.tst	success	interop1: bgp vpnv4
intop1-bgp12.tst	success	interop1: bgp authentication
intop1-bgp13.tst	success	interop1: bgp vpnv6

```
description wireguard over ipv6

addrouter r1
int ser1 ser - $1a$ $1b$
!
vrf def v1
 rd 1:1
 exit
int ser1
 vrf for v1
 ipv4 addr 1.1.1.1 255.255.255.0
 ipv6 addr 1234::1 ffff::
 exit
crypto ipsec ips
 key EFw2rJEdqFGDgC80um3fwMmAafwqXno+PsbMHPZ0umM=M6vDV8QdiWDQppVKjKf8xjoKtyGAeRK/Ue48kwKI5Ss=
 exit
int tun1
 tunnel vrf v1
 tunnel prot ips
 tunnel mode wireguard
 tunnel source ser1
 tunnel destination 1234::2
 vrf for v1
 ipv4 addr 2.2.2.1 255.255.255.0
 ipv6 addr 4321::1 ffff::
 exit
!

addrouter r2
int ser1 ser - $1b$ $1a$
!
vrf def v1
 rd 1:1
 exit
int ser1
 vrf for v1
 ipv4 addr 1.1.1.2 255.255.255.0
 ipv6 addr 1234::2 ffff::
 exit
crypto ipsec ips
 key 6JhyvKPutQ9DNLupOPmDnQLRWtUWlUjI6PTJ/IZ9llw=bQMmpCaGVyq9f+v48XGmfH5DMLytkqziID+rBH+qQic=
 exit
int tun1
 tunnel vrf v1
 tunnel prot ips
 tunnel mode wireguard
 tunnel source ser1
 tunnel destination 1234::1
 vrf for v1
 ipv4 addr 2.2.2.2 255.255.255.0
 ipv6 addr 4321::2 ffff::
 exit
!


r1 tping 100 5 2.2.2.2 /vrf v1
r2 tping 100 5 2.2.2.1 /vrf v1
r1 tping 100 5 4321::2 /vrf v1
r2 tping 100 5 4321::1 /vrf v1
```

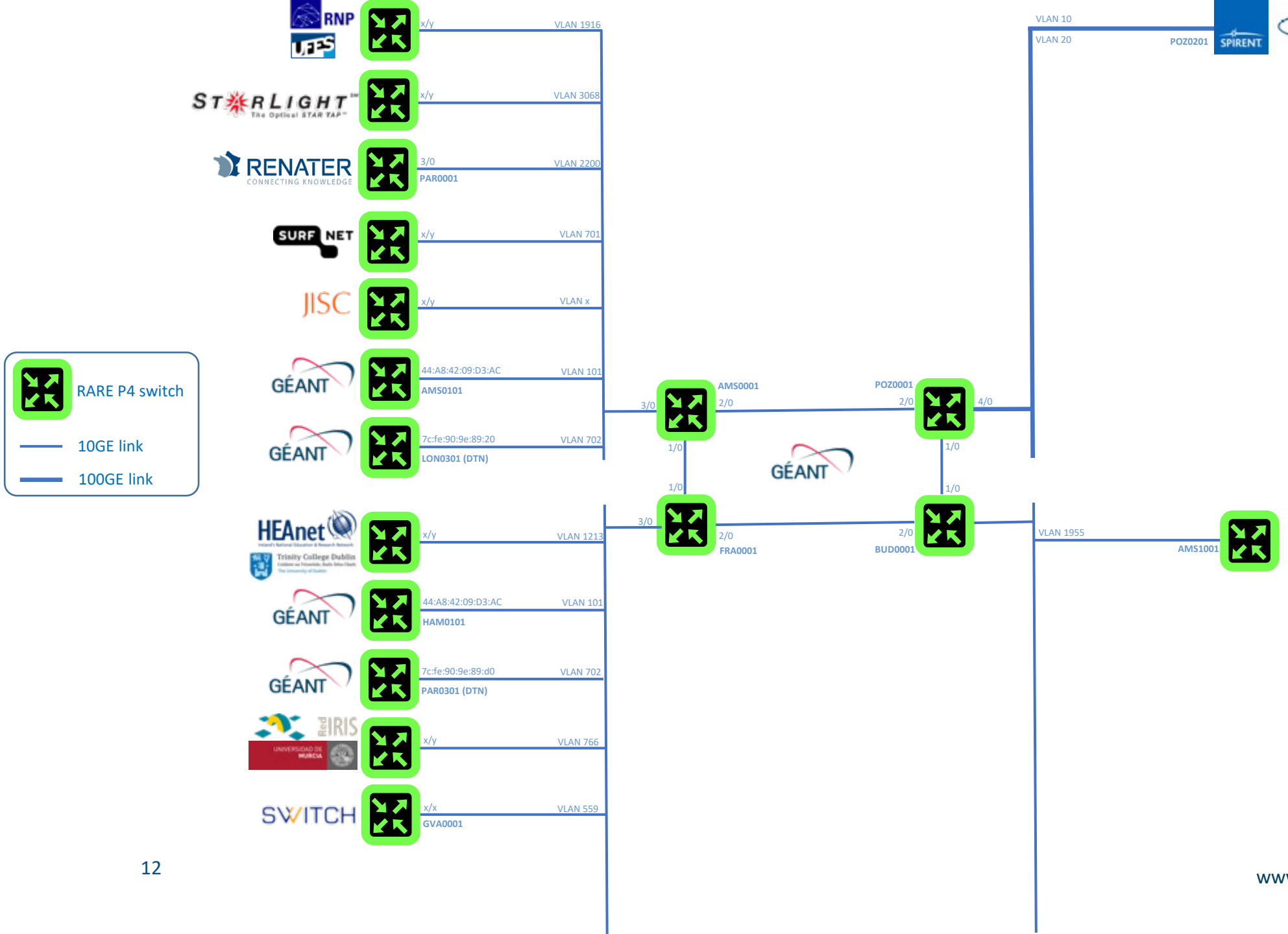
RARE testing framework: Dataplane tests ~300 tests

Complete feature list

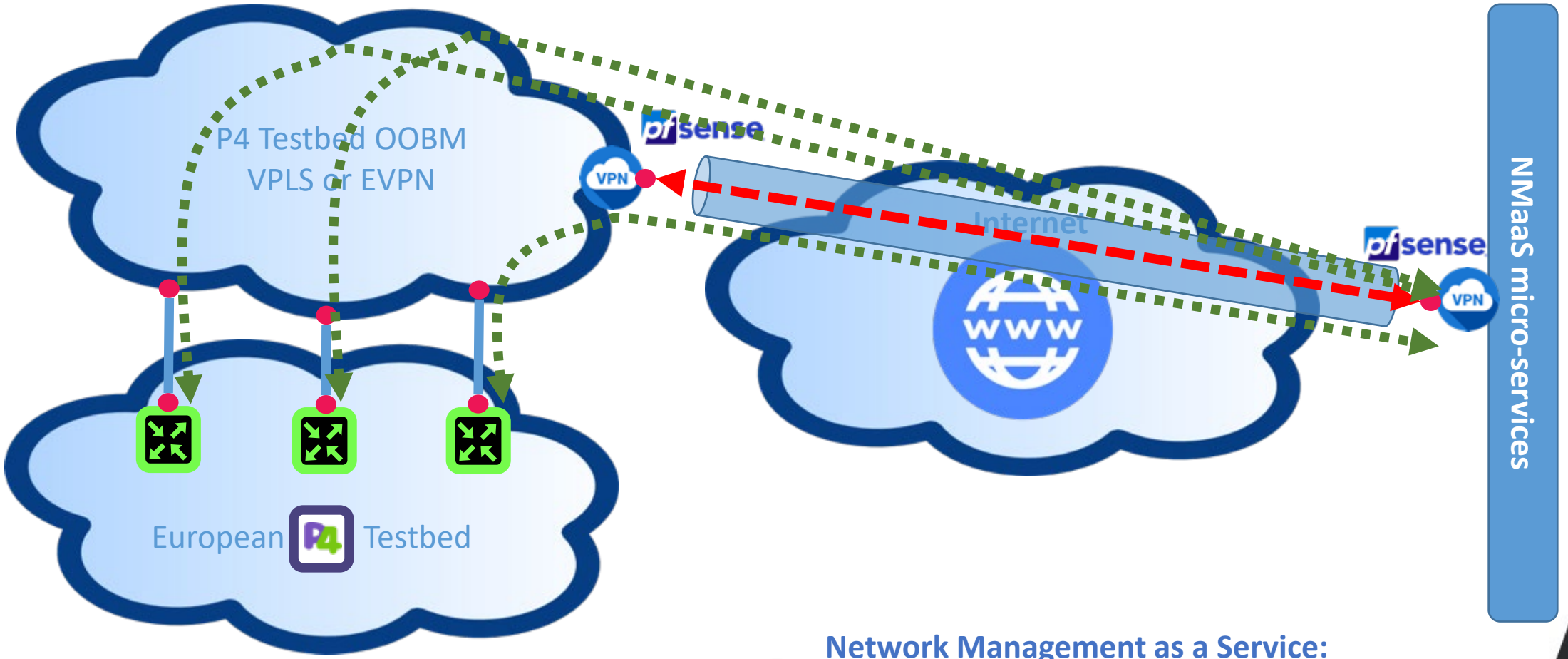
Type	Test #	Name	P4		DPDK
acl	01	copp	✓	✓	✓
acl	02	ingress access list	✓	✓	✓
acl	03	egress access list	✓	✓	✓
acl	04	nat	✓	✓	✓
acl	05	vlan ingress access list	✓	✓	✓
acl	06	vlan egress access list	✓	✓	✓
acl	07	bundle ingress access list	✓	✓	✓
acl	08	bundle egress access list	✓	✓	✓
acl	09	bundle vlan ingress access list	✓	✓	✓
acl	10	bundle vlan egress access list	✓	✓	✓
acl	11	bridge ingress access list	✓	✓	✓

And more features !

Please come @IRC #freertr and submit your idea!

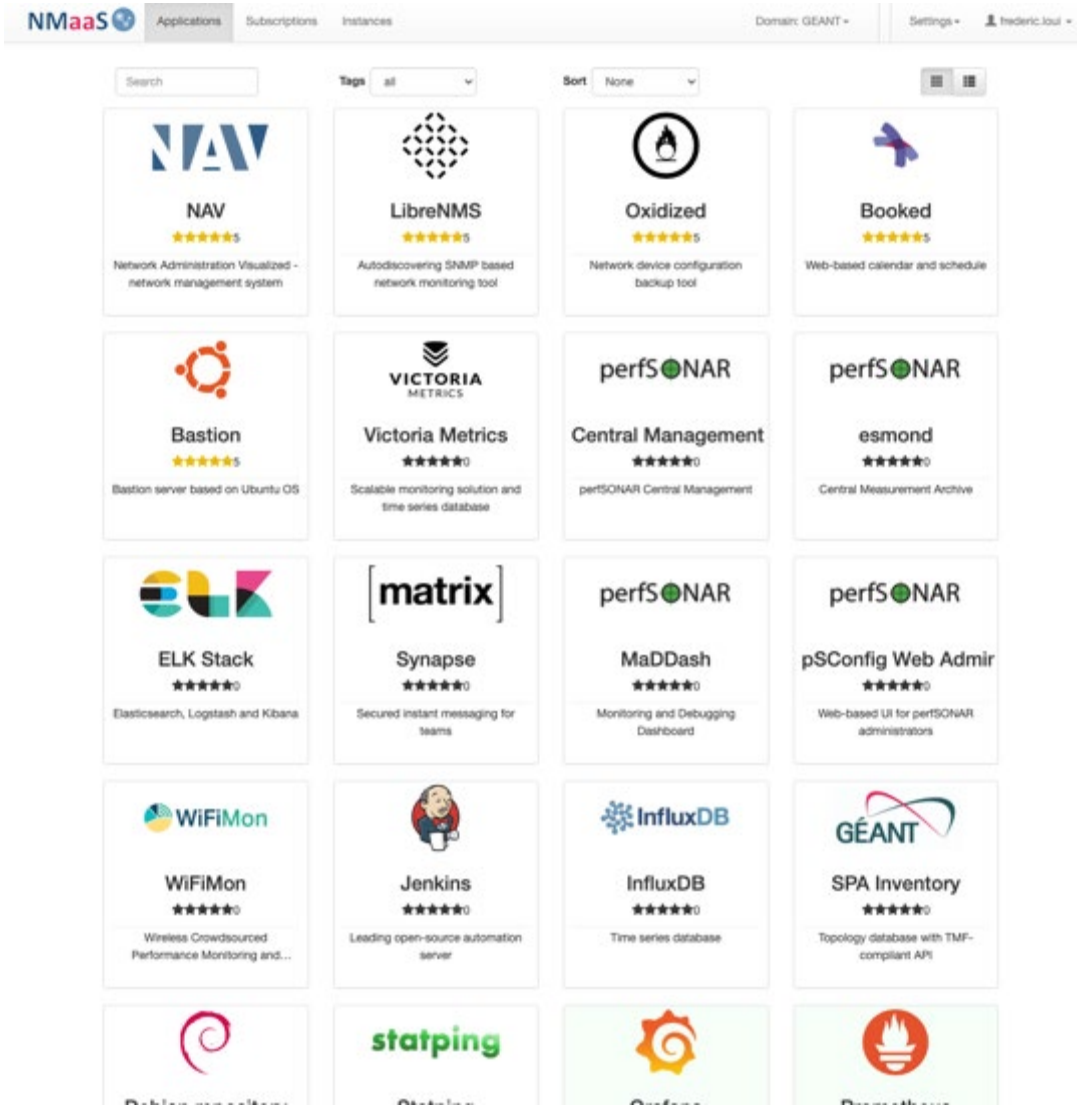

RARE P4 switch
10GE link
100GE link
RNP
UFPS
x/y
VLAN 1916
STARLIGHT
The Optical STAR TAP
x/y
VLAN 3068
RENATER
CONNECTING KNOWLEDGE
3/0
VLAN 2200
PAR0001
SURF NET
x/y
VLAN 701
JISC
x/y
VLAN x
GÉANT
44:A8:42:09:D3:AC
VLAN 101
AMS0101
GÉANT
7c:fe:90:9e:89:20
VLAN 702
LON0301 (DTN)
AMS0001
3/0
2/0
1/0
1/0
POZ0001
2/0
4/0
1/0
GÉANT
VLAN 10
VLAN 20
POZ0201
SPIRENT
PSNC
HEAnet
Trinity College Dublin
x/y
VLAN 1213
3/0
2/0
FRA0001
2/0
BUD0001
VLAN 1955
AMS1001
GÉANT
44:A8:42:09:D3:AC
VLAN 101
HAM0101
GÉANT
7c:fe:90:9e:89:d0
VLAN 702
PAR0301 (DTN)
IRIS
UNIVERSIDAD DE MURCIA
x/y
VLAN 766
SWITCH
x/x
VLAN 559
GVA0001

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

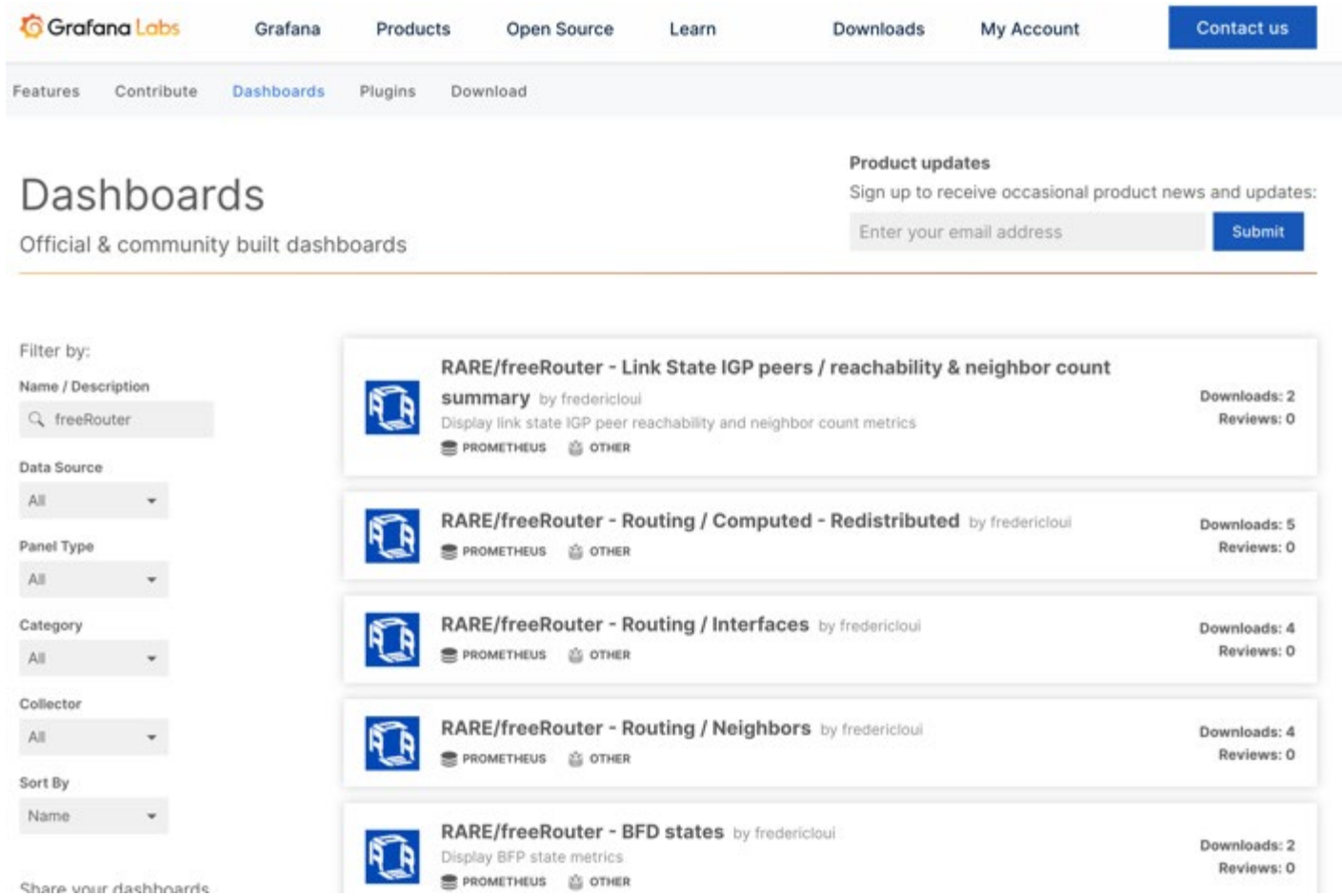
P4 LAB network management via (Network Management as a Service)



Network Management as a Service:
https://nmaas.eu
https://wiki.geant.org/display/NMAAS

Monitoring at node level! (Prometheus agent)

Monitoring at node level! (Grafana dashboard)



https://grafana.com/grafana/dashboards?search=freeRouter

Key take-way – We are ready to roll into production

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

UNIVERSIDAD DE MURCIA

KIFÜ

- Work in progress production deployment

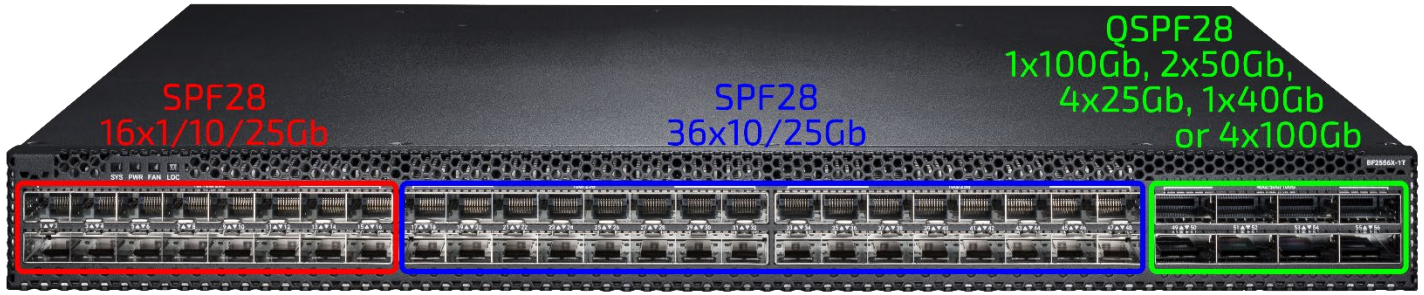
RENATER
CONNECTING KNOWLEDGE

Practicle use case #001 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 #002 BRAS router

- DPDK and P4 dataplane
  - ➔suitable for CAMPUS / EDGE BACKBONE router
- nx1GE, nx10GE, nx100GE



Practical use case #003 LSR router

- P4 dataplane fits perfectly pure **LSR** core router
- NNI: 4 directions with (8x100GE) bundle

Practical use case #004 LER router

- P4 dataplane fits perfectly pure **LER** use case
- NNI: EST/WEST direction @ (8x100GE) bundle
- UNI: 16x100GE left for end user connection!

Practical use case #005 high performance BGP RR

- Recycling new server?
- Ideal for **K8s** cluster using **BGP** as **CNI** network plugin
- Taking advantage of server « huge » amount of RAM
- No need specific high performance dataplane

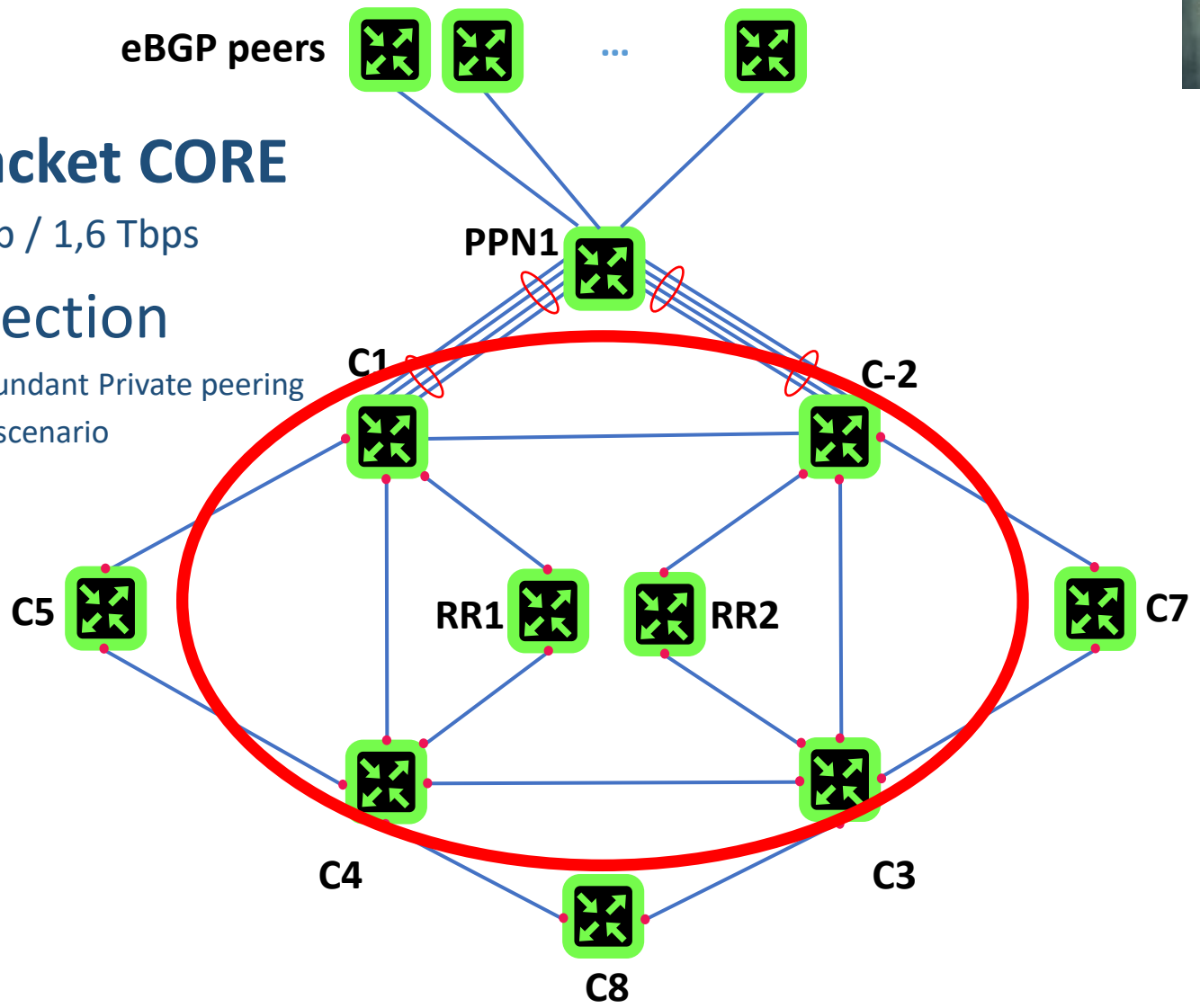
Practical use case #006 « small PE » Practical

Ideal fo aggregation

- 2x10GE or 2x100GE NIC server side
- 2x10g+48x1g or 1x100g+48x1/10g switch

Practical use case #007 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-way – Room for improvement

- **Network Management**
  - Node monitoring
  - Flow Monitoring
- **New Network Management Paradigm**
  - Streaming Telemetry
  - **INT**

  ➔**It is a good opportunity to rethink how Network Management is handled**
- « Closing the dots » with **automation existing project**

Why not joining the effort?

Key take-way – 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 NGN NMS**
  - 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 …

GÉANT

European Commission

FreerTr

P4

DPDK
DATA PLANE DEVELOPMENT KIT

intel

APS Networks

debian

TCPDUMP & LIBPCAP

OpenJDK

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*

Looking ahead: Finalize transition to production

Extend HCL:
new TOFINO based hardware support
new DPDK release

New target:
TOFINO2
NVIDIA DPU
P4 SmartNIC
FPGA

New idea:
Polka
P42VPP
T4P4S ELTE
Leverage Nix paradigm

And more …

Thank you

Any questions?

www.geant.org

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).