

RARE routing platform

Potential usage in data-centres network

LOUI Frédéric

GÉANT/RENATER – RARE technical leader

MATE Csaba

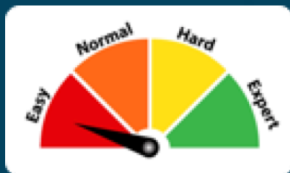
GÉANT/KIFU – RARE/freeRtr lead core developer

WLCG Data Center network architecture

7-8th 2021

Public

www.geant.org



GÉANT 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



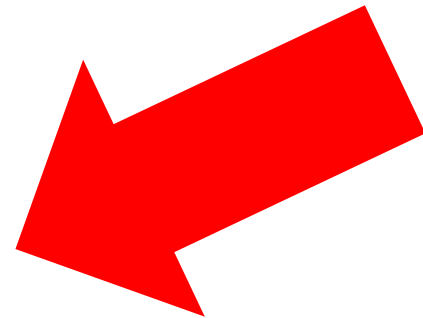
- Starting from 2020:
 - Dataplane solution reached maturity ready to implement production grade use case



- NOS emergence

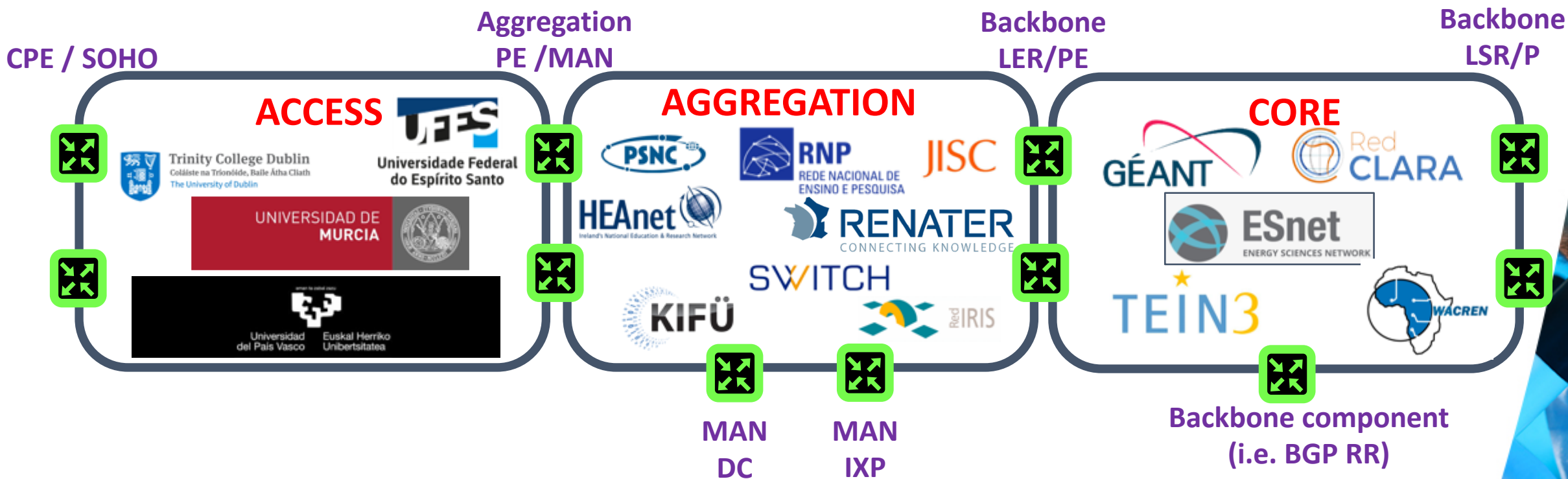


- Hypervisor Technology convergence
 - openstack.
 - kubernetes



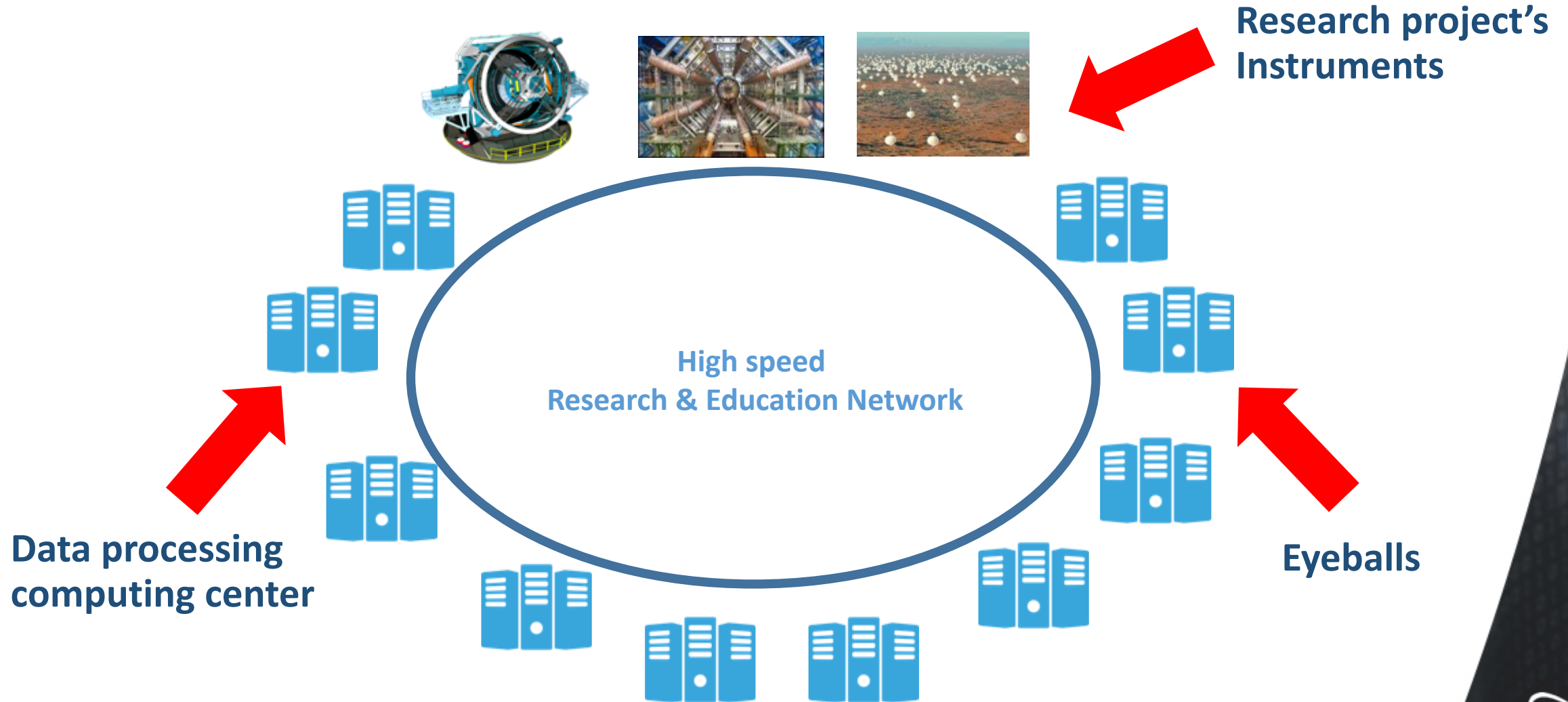
It's a good time to tie Control Plane and Dataplane!

RARE use cases

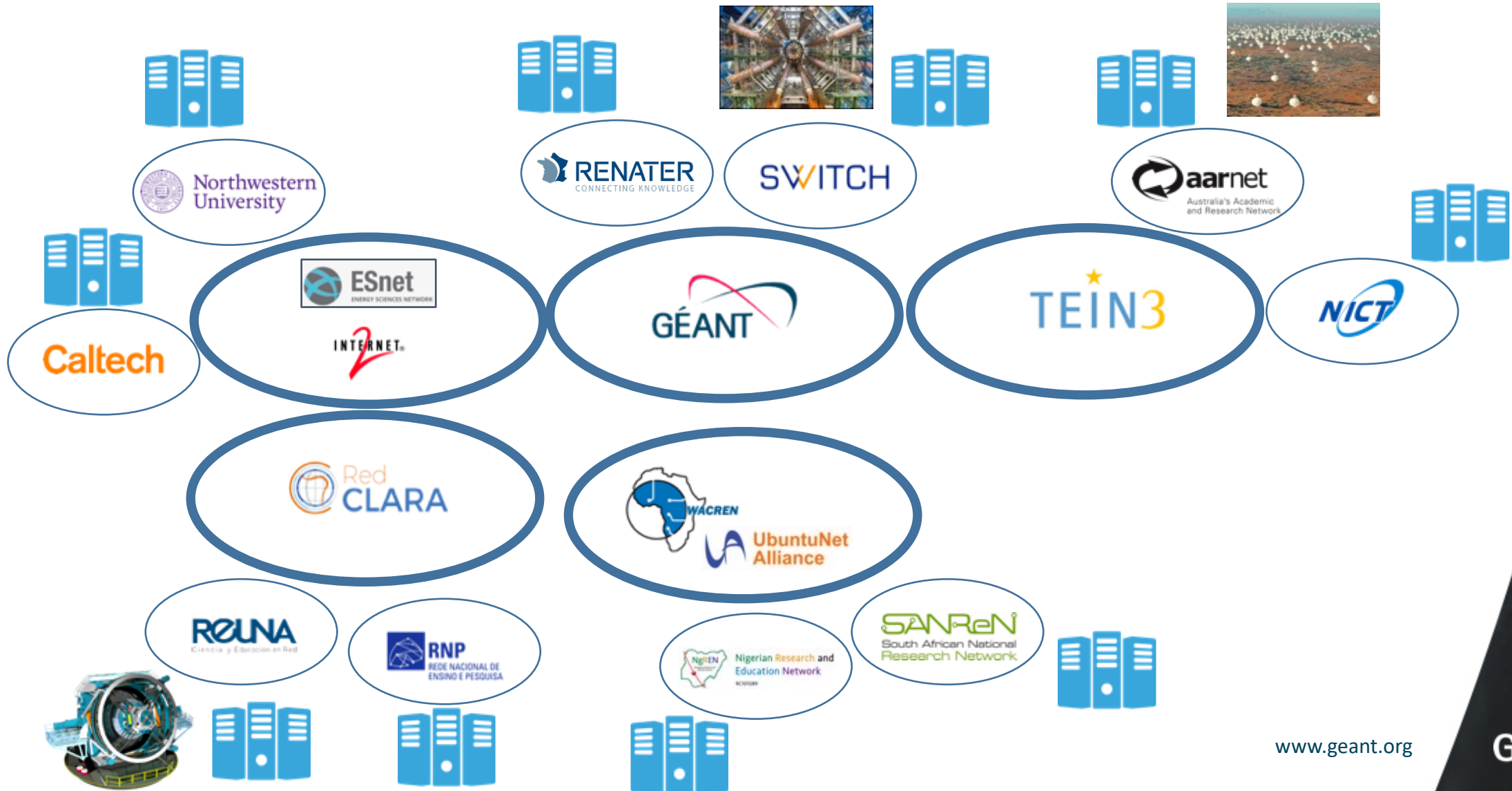


IPv4 and IPv6 compliant!

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



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

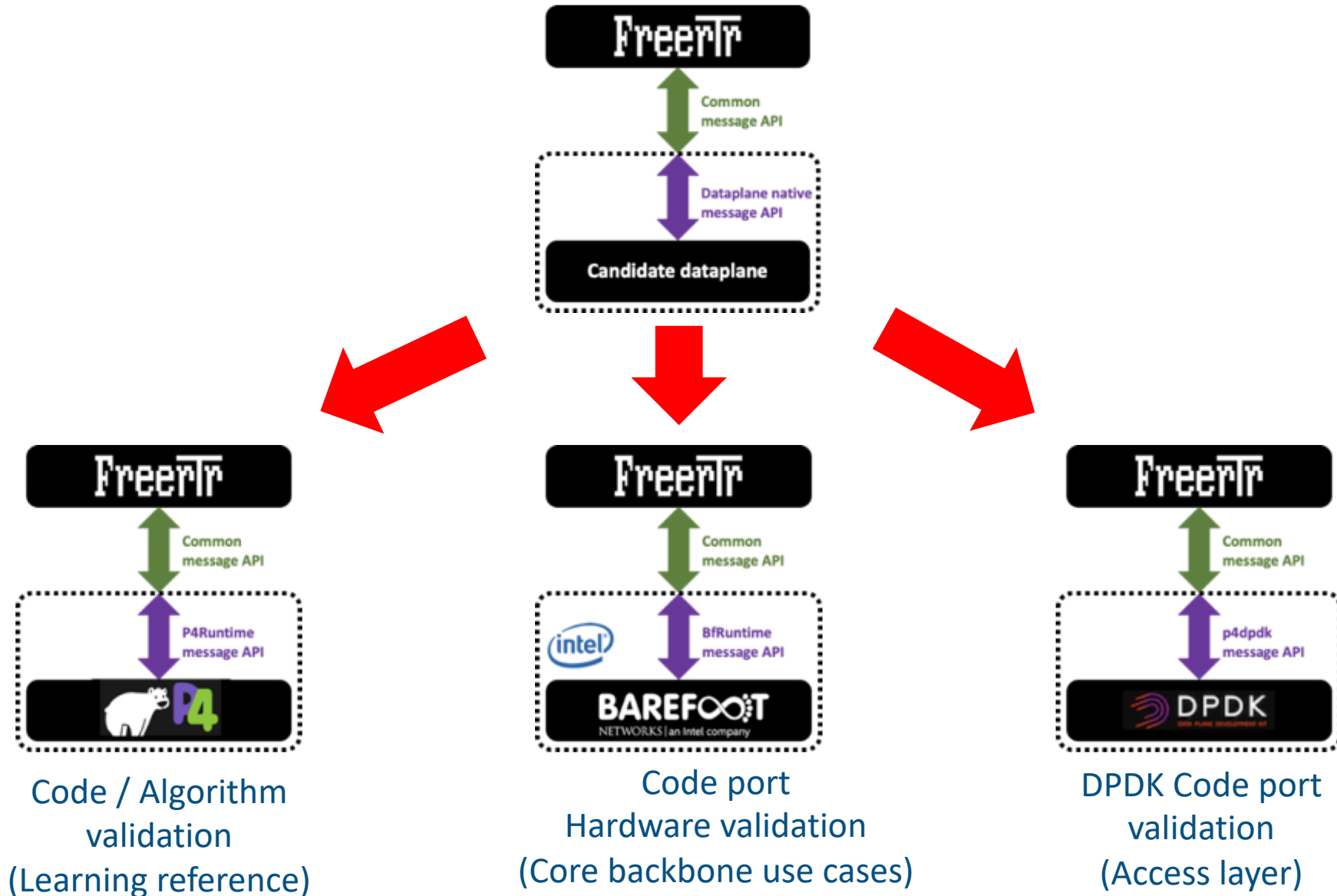


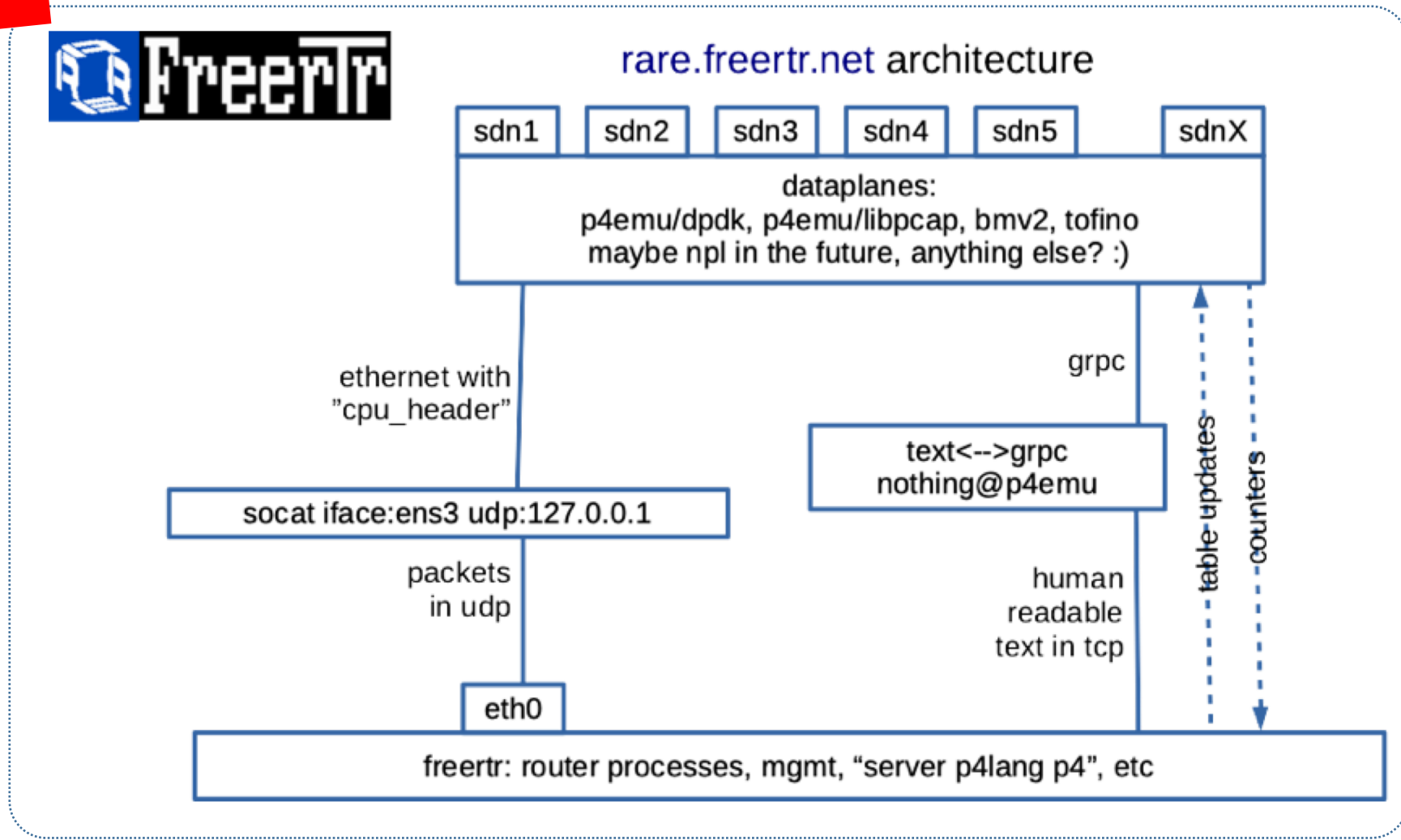
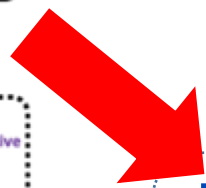
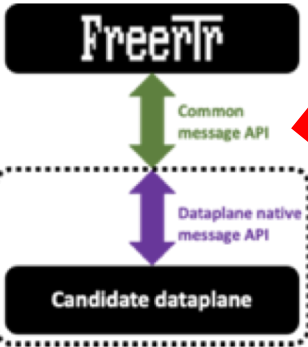
under study

- RARE p4 discussion emulation targets



RARE “target” development





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+PsbMHP20umM=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 6JhyvKPatQ9DNLuPOpMdnQLRWtUWlUjI6PTJ/I291lw=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
!

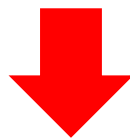
r1 tping 100 5 2.2.2.2 /vrf v1
r2 tping 100 5 2.2.2.1 /vrf v1
r1 tping 100 5 4321::2 /vrf v1
r2 tping 100 5 4321::1 /vrf v1
    
```



RARE testing framework: Dataplane tests ~300 tests

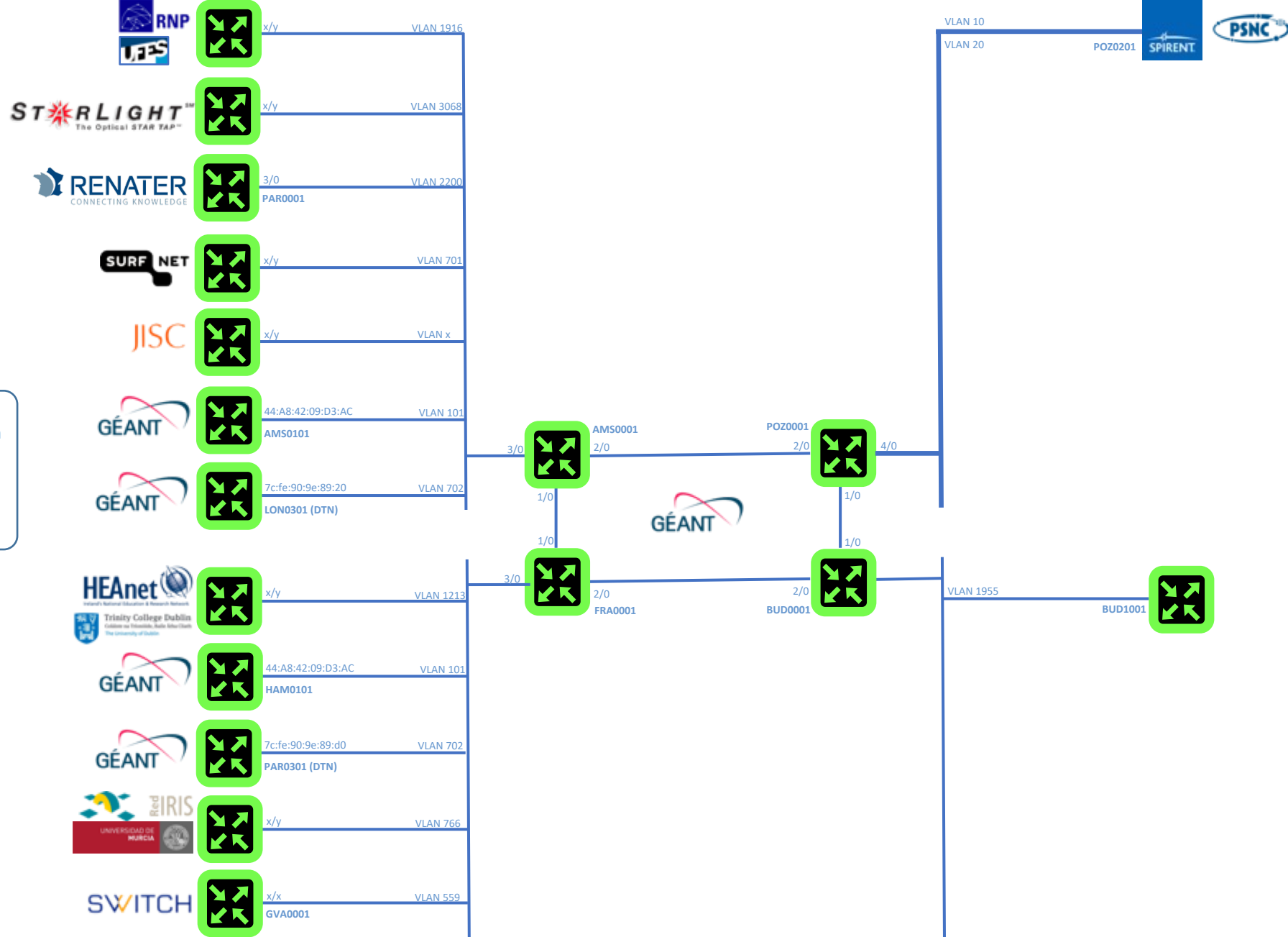
Complete feature list


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





And more features !

Please come @IRC #freertr and submit your idea!

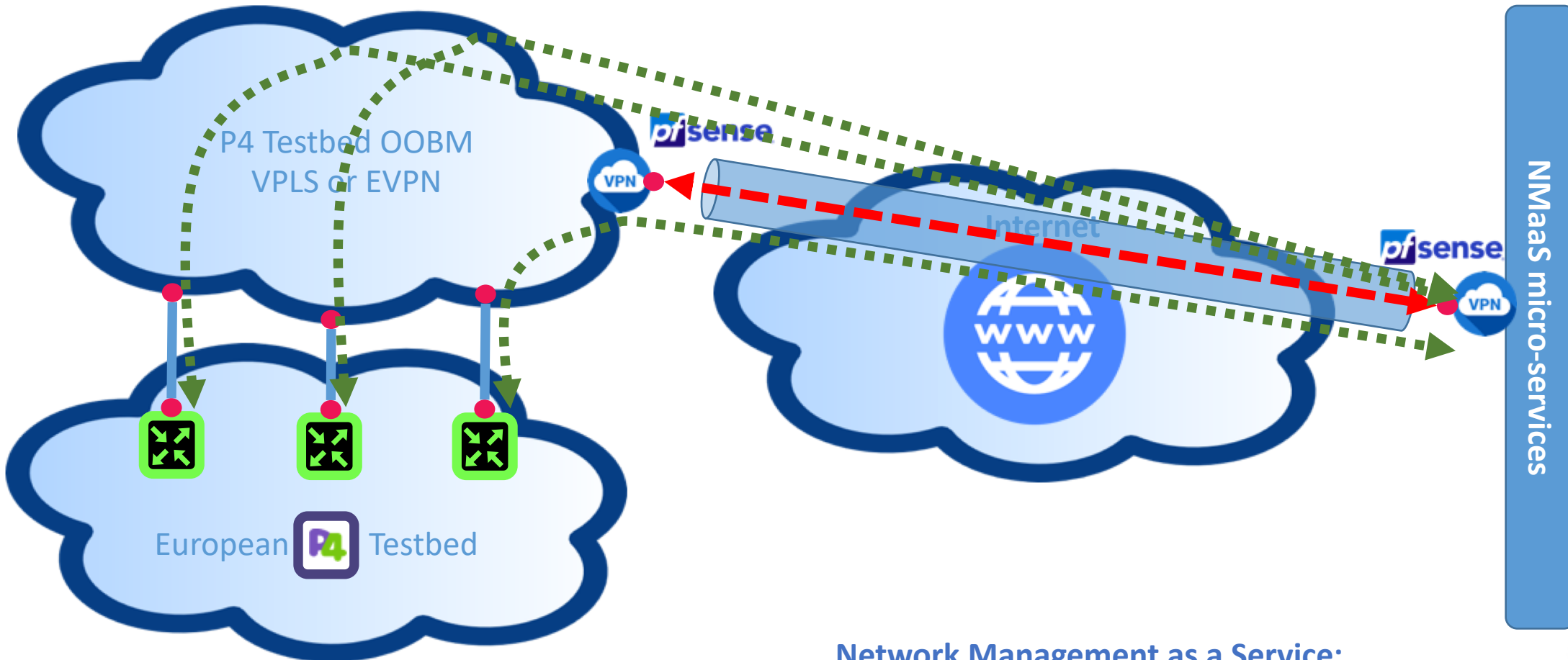


 RARE P4 switch

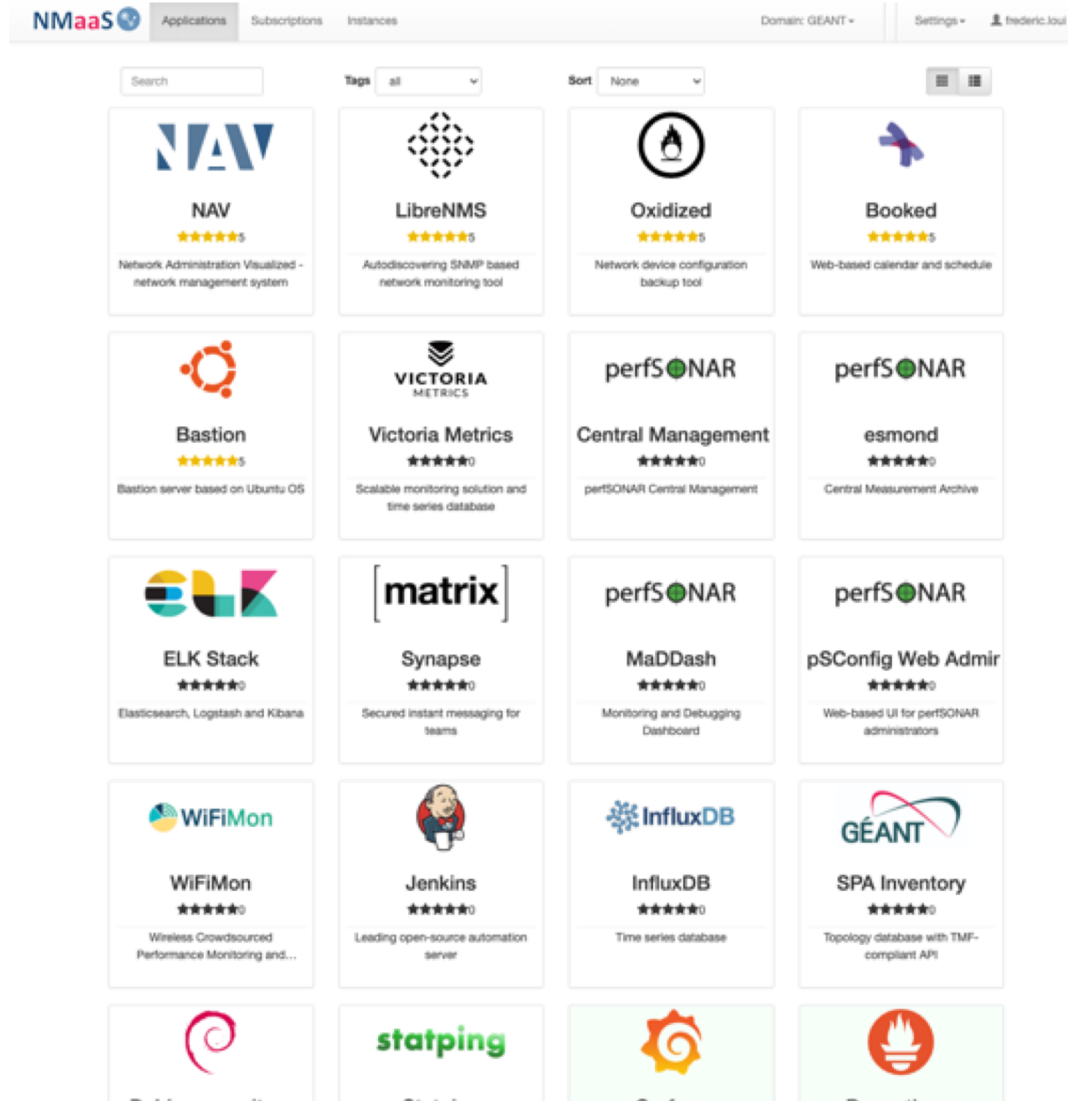
 10GE link

 100GE link

RARE validation designs: P4 LAB network management via NMaas*! (Network Management as a Service)



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)

A screenshot of the Grafana Labs website. The top navigation bar includes 'Grafana Labs', 'Grafana', 'Products', 'Open Source', 'Learn', 'Downloads', 'My Account', and 'Contact us'. Below this is a secondary navigation bar with 'Features', 'Contribute', 'Dashboards', 'Plugins', and 'Download'. The main heading is 'Dashboards' with the subtitle 'Official & community built dashboards'. On the right, there is a 'Product updates' section with a sign-up form. On the left, there is a filter sidebar with options for 'Name / Description', 'Data Source', 'Panel Type', 'Category', 'Collector', and 'Sort By'. The search term 'freeRouter' is entered in the search box. The main content area displays five dashboard cards, each with a 'freeRouter' icon, a title, a description, and download/review statistics.



Name / Description	Downloads	Reviews
RARE/freeRouter - Link State IGP peers / reachability & neighbor count summary by fredericloui Display link state IGP peer reachability and neighbor count metrics	2	0
RARE/freeRouter - Routing / Computed - Redistributed by fredericloui	5	0
RARE/freeRouter - Routing / Interfaces by fredericloui	4	0
RARE/freeRouter - Routing / Neighbors by fredericloui	4	0
RARE/freeRouter - BFD states by fredericloui Display BFP state metrics	2	0

<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
-  currently in operation SOHO
- Production deployment



- Work in progress production deployment

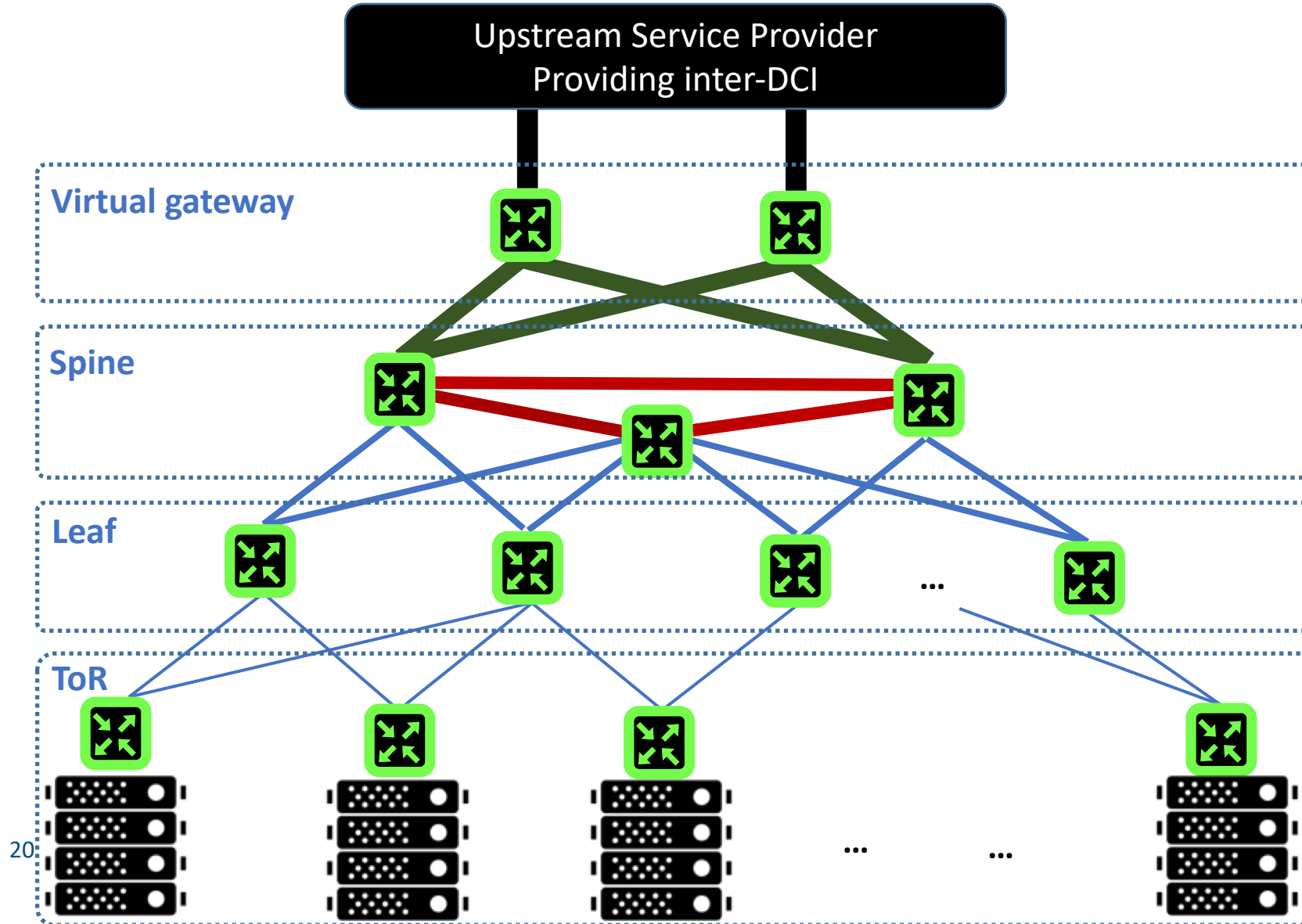




**Let's get practical ...
and present you
actual real life use cases !**



Practical use case #001 100GE DC network



20

Practical use case #001a ToR

- 2x100GE uplink toward LEAF nodes
- 30x100GE server blade connection
- Each server powered by DPDK NIC



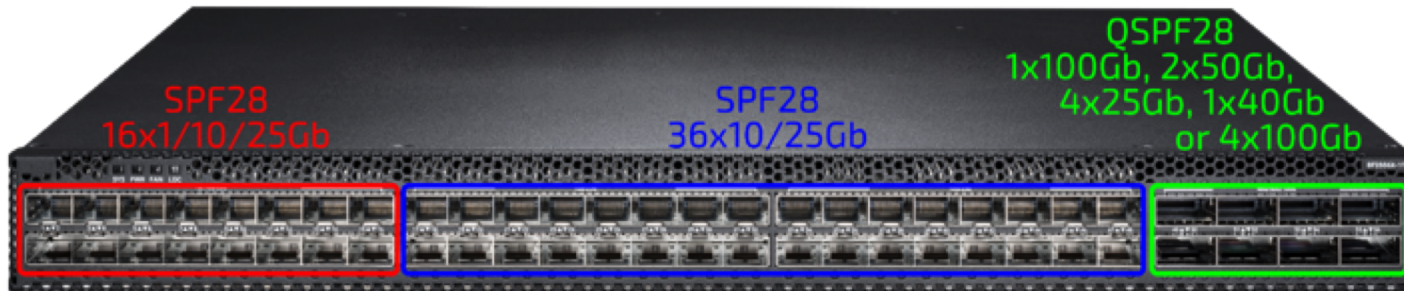
Edge-core
NETWORKS

WEDGE-100BF-32X



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



Edge-core
NETWORKS

WEDGE-100BF-32X



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!




Edge-core
NETWORKS

WEDGE-100BF-32X



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)



DPDK
DATA PLANE DEVELOPMENT KIT

TCPDUMP & LIBPCAP



Practical use case #003 100GE Private Peering node

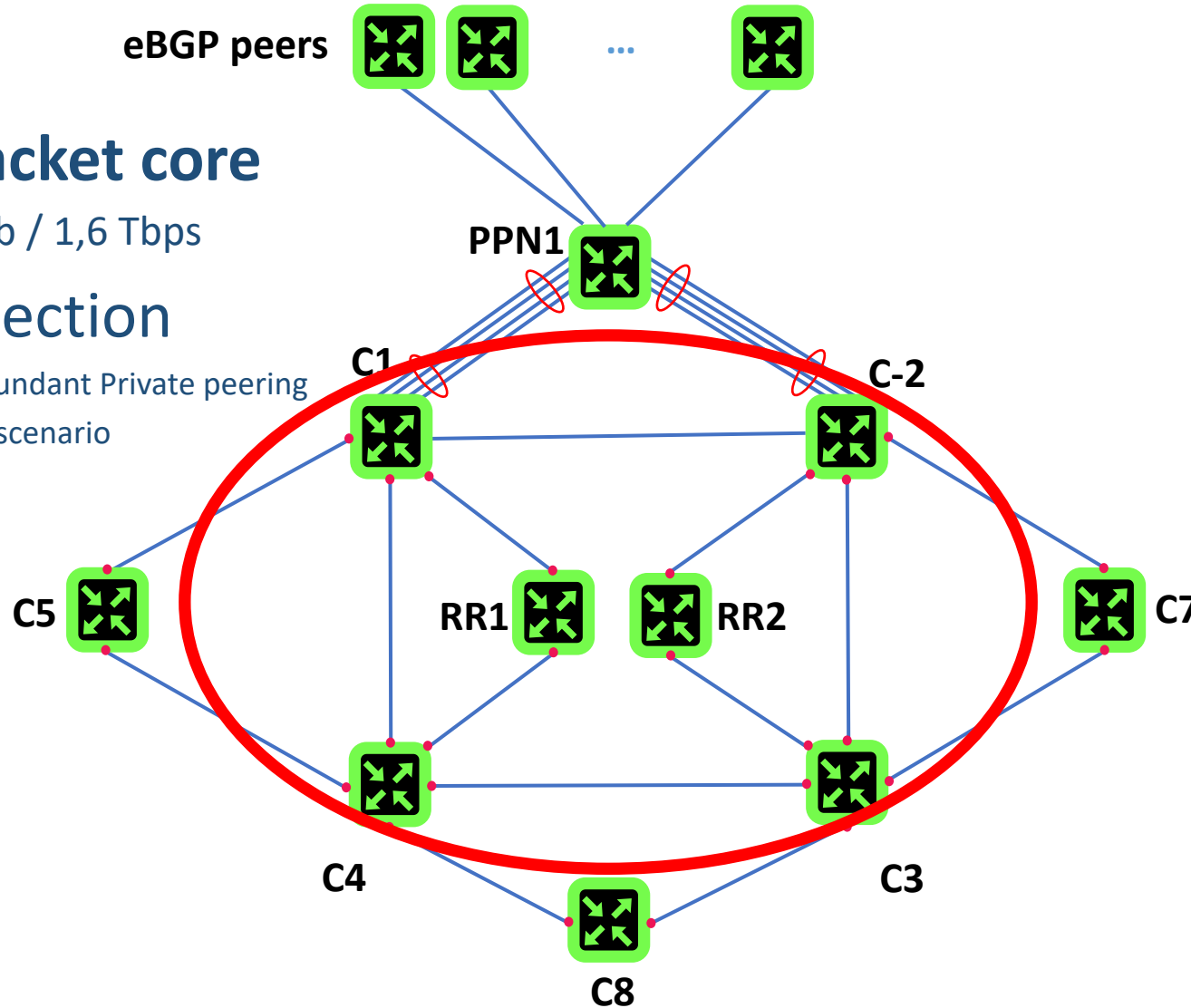


- High resilient **Packet core**

- 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 ...



APS Networks



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](https://twitter.com/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 ?



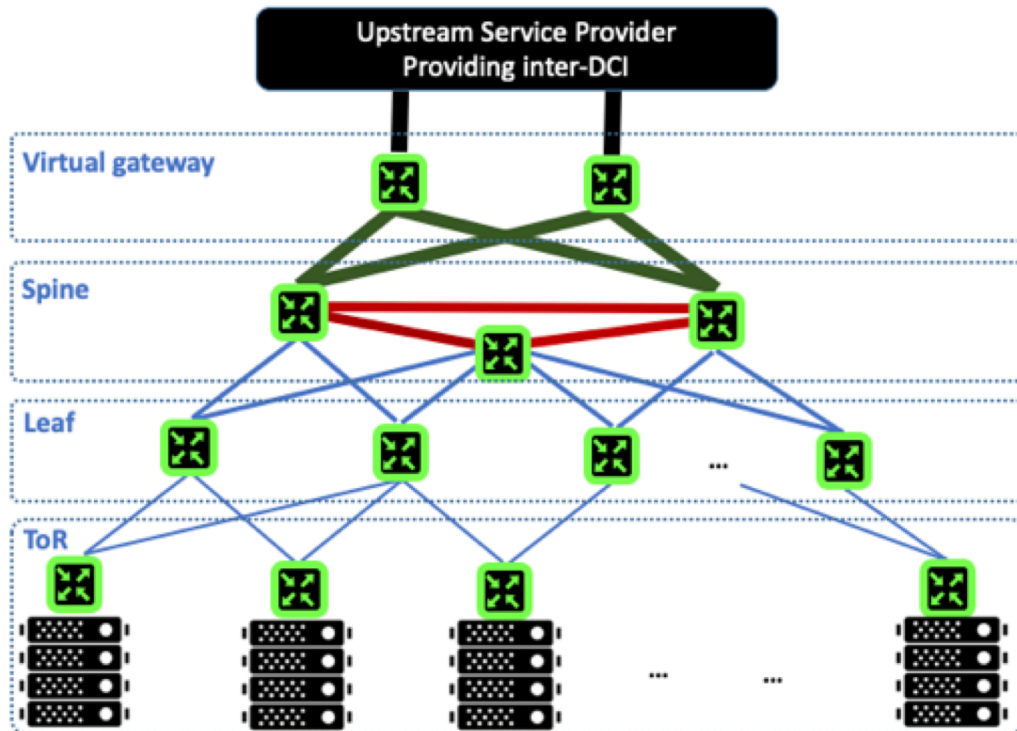
kubernetes

Workers node

+



As a CNI plugin



Inter-DCI
End to end LSP !

Thank you

Any questions?

www.geant.org

