



Educação, Pesquisa
e Inovação em Rede

Use of SONIC and FreeRTR on programmable switches

Marcos Schwarz

LHCOPN-LHCONE meeting #49

October 25 2022, CERN, Geneva CH



—● Agenda

Extensible Production-Grade P4 NOS

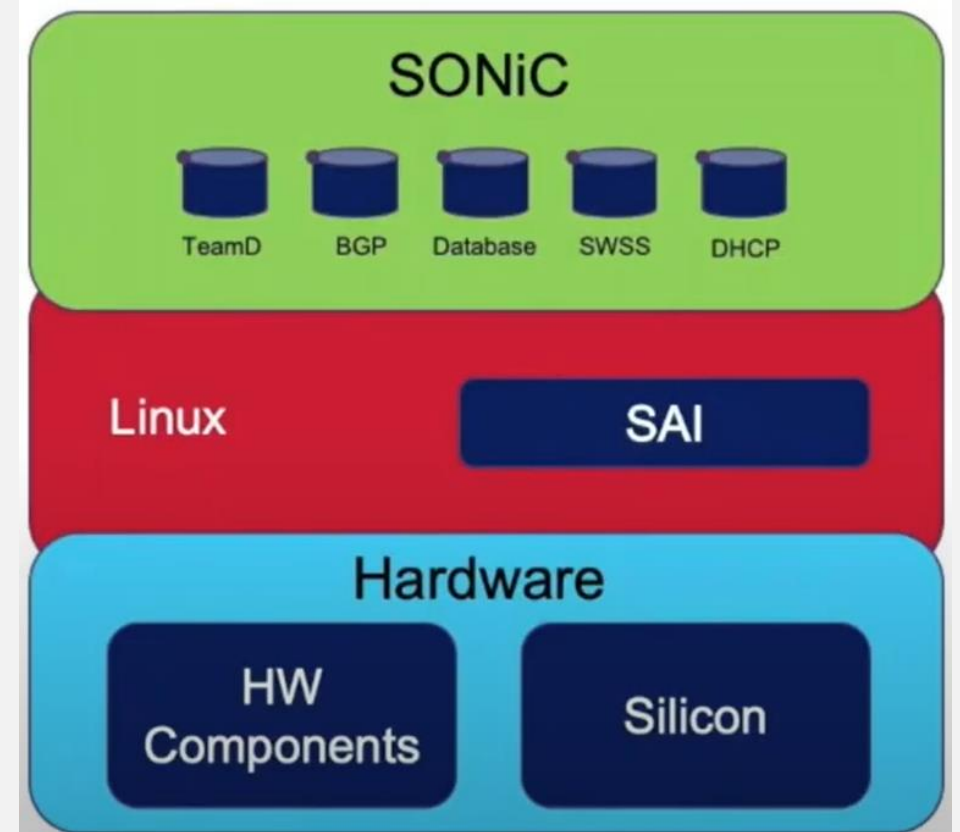
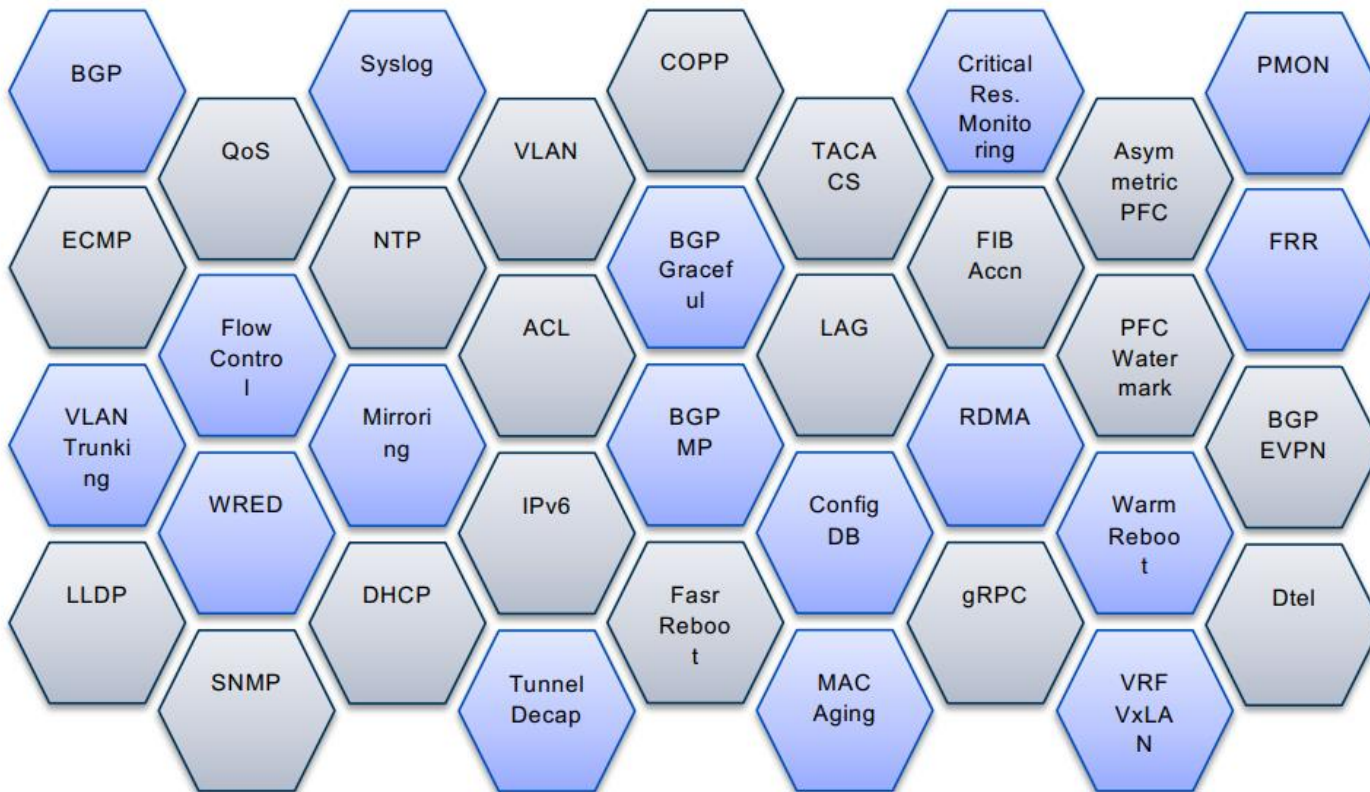
- SONiC/PINS, RARE/FreeRtr (PoIKA)
- Highlights
- Pipeline Constraints

Global P4 Lab

- Topology / Sites
- Goals
- Integration Path

SONiC (Linux Foundation)

- Open and multi-vendor Network Operating System focused on datacenter requirements/features
- Support for traditional (fixed-function) and programmable switches



SONiC Ecosystem

Merchant Silicon



Switch Platform



Adoption



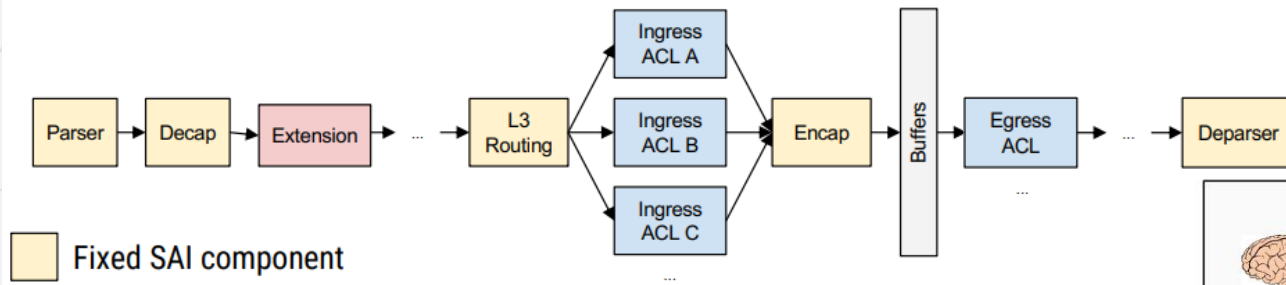
System/Service



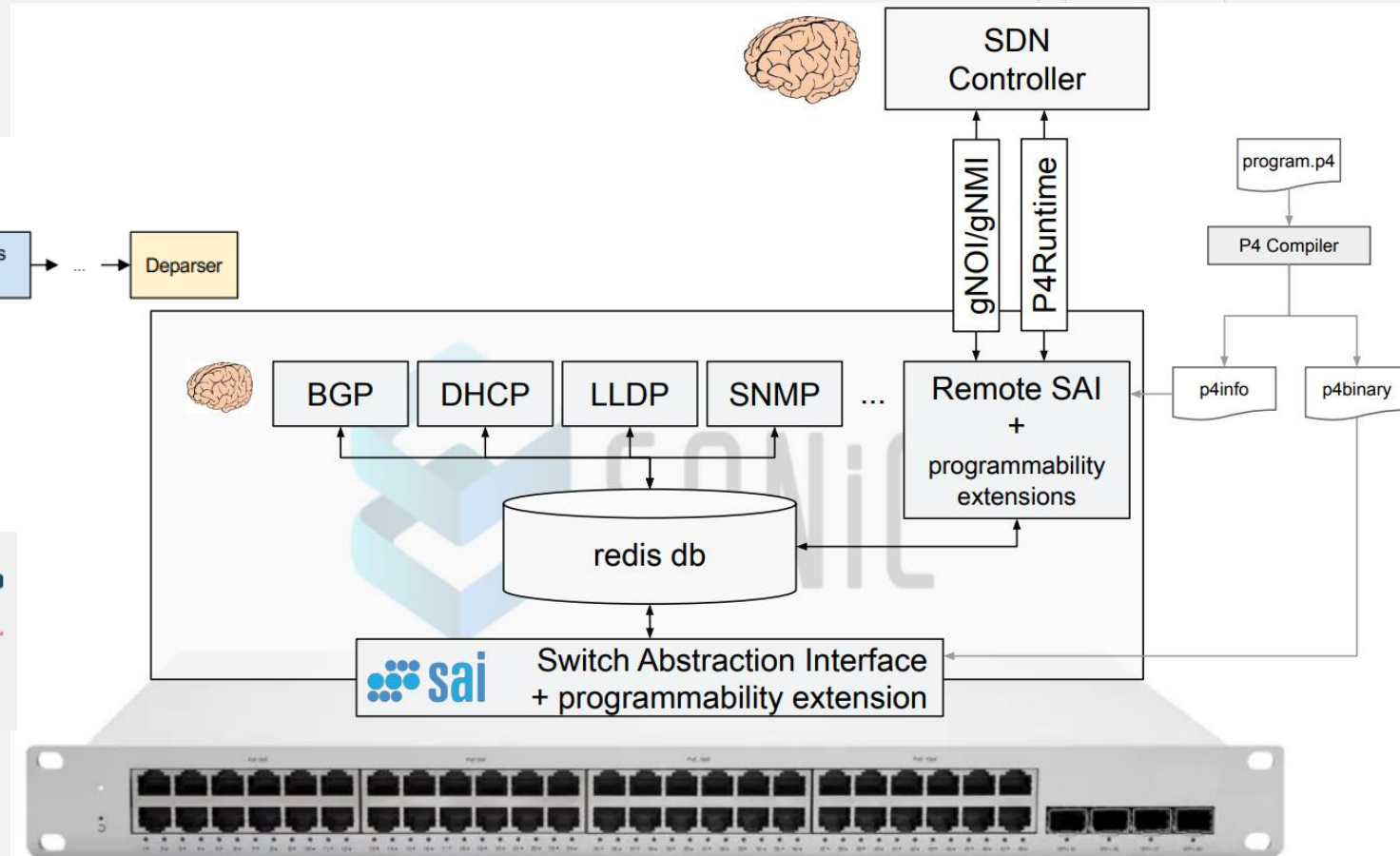
PINS – P4 Integrated Networking Stack (ONF)

Hybrid control plane with path to SAI extensibility (SONiC + Stratum)

- Great for traditional network users that want some SDN capabilities
- Opt-in SDN



- Fixed SAI component
- Configurable SAI component (ACL Tables)
- P4 extension (for parts not exposed by SAI)



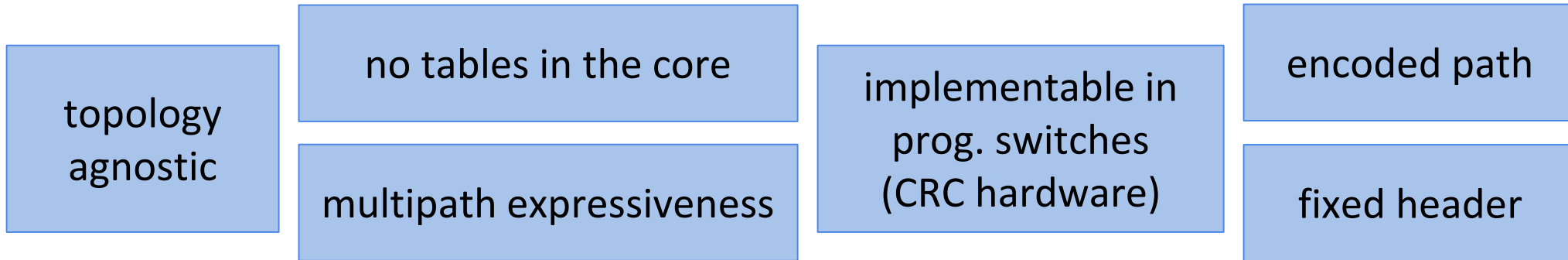
— RARE / FreeRtr (GEANT)

R&E driven router implementation with support for telco protocols

- Supports multiple backends: Tofino/P4, DPDK, eBPF/XDP
- Modular design and rich feature set:
 - **forwarding**: ipv4, ipv6, ipx, mpls, nsh, layer2, irb, atom, eompls, vpls, **evpn**
 - **routing protocols**: ospf, isis, bgp, rip, eigrp, babel, olsr, pim, msdp
 - **isp support**: p2p, p2mp, mp2mp built by bgp, ldp, rsvp-te, **sr**, **sr-te**, bier, **polka**
 - **crypto**: macsec, ipsec, ikev1, ikev2, tls, dtls, ssh, openvpn, wireguard, sgt
 - **tunnel**: gre, ipip, l2tp, pptp, lisp, geneve, nvgre, vxlan, etherip, amt
 - **encapsulation**: ethernet, vlan, ppp, framerelay, pwether, virtppp, hairpin
 - **misc**: acl, **qos**, nat, pbr, **srv6**, vrrp, hsrp, inspect, 6to4, rpl, tunnel, vpdn, **pcep**
- Simple P4 pipeline for integration/extension of new features (PolKA, Packet Marking)



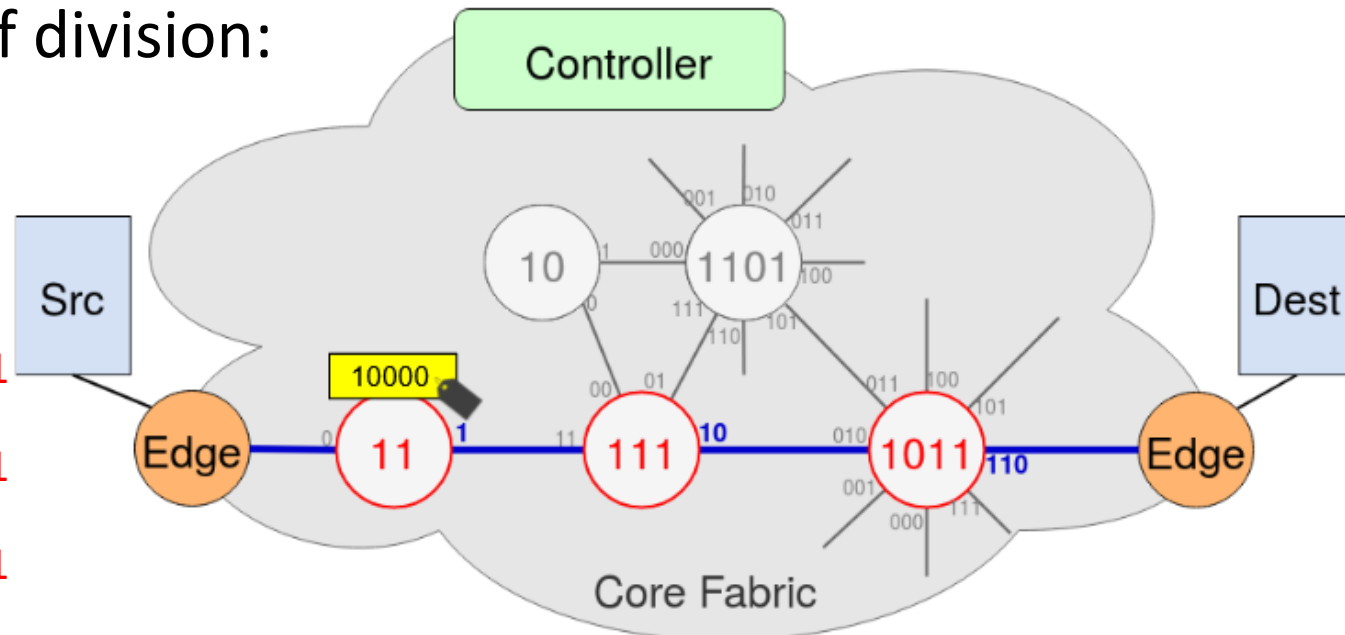
- A **Source Routing** approach that simultaneously meets the requirements:



- Polynomial Residue Number System (RNS)
- Forwarding based on remainder of division:

$$\text{portID} = \langle \text{routeID} \rangle_{\text{nodeID}}$$

1	=	$\langle 10000 \rangle_{0011}$
10	=	$\langle 10000 \rangle_{0111}$
110	=	$\langle 10000 \rangle_{1011}$



Timeline

 PolKA received the 2021 Google Research Scholar Award

 M-PolKA received the Intel Connectivity Research Grant (Fast Forward Initiative)

2020

2021/1

2021/2

2022

**PolKA paper
IEEE NetSoft**

Novel Polynomial RNS-based SR and reuse of CRC hardware

Emulated prototype in Mininet

**ONDM paper
Deploy @RARE**



Hardware prototype in Tofino

**Integration with
RARE+FreeRtr**

PolKA data & control plane implementation + integration

Emulated prototype in FreeRtr &
Hardware prototype in Tofino w/ FreeRtr control plane

**M-PolKA paper
IEEE TNSM**

Extension to multipath SR for reliable communications

Innovative apps: inband network telemetry, and load balance

**PolKA@pangr
IETF 113**

Lightning Talk Path Aware Networking RG

**PolKA@Global
P4 Lab**

PolKA deployment @Caltech SDN Lab
PolKA Talk at LHC-ONE

PolKA Demo at SC-22

Motivation

Proposal

Design

Prototype

Applications

Conclusions

— Highlights

SONiC/PINS

- Backends (P4, Fixed-function)
- Source Routing (SRv6)
- Telemetry (sFlow, INT, gNMI)
- SDN Controller (P4Runtime)
- Automation (Ansible / Batfish)

Known Limitations

- No pipeline with all features
- Routing (BGP only)

Roadmap

- SRv6 uSID
- PINS SAI Generic Extensions
- SONiC PINS P4DPDK SoftSwitch

RARE/FreeRtr

- Backends (P4, DPDK, eBPF/XDP)
- Source Routing (SR-MPLS, PoIKA)
- Telemetry (Prometheus)
- Telco Protocols (BGP-LS, PCEP, TWAMP)
- Routing (OSPF, ISIS, BGP)
- R&E driven development (Packet Marking)

Known Limitations

- No pipeline with all features
- SR-MPLS (label depth 1)
- SRv6 (no traffic steering)

— Pipeline Constraints

Due to limitations of current gen programmable devices

It's necessary to implement pipelines specialized to specific use cases

i.e. Intel P4 pipeline for SAI/SONiC on Tofino1 (x1, x2..) and Tofino2 (y1, y2...)

Model	DTEL (INT)	sFlow	NAT	VXLAN	L2 VXLAN	MPLS	SRv6
Tofino1 (x1)	Yes	Yes	No	No	No	No	No
Tofino1 (x2)	No	No	Yes	Yes	No	No	No
Tofino1 (x6)	No	No	No	No	No	Yes	Yes
Tofino2 (y1)	Yes	Yes	No	Yes	Yes	No	No
Tofino2 (y2)	No	Yes	Yes	Yes	Yes	No	No
Tofino2 (y3)	No	No	No	No	No	Yes	Yes

Global P4 Lab Goals

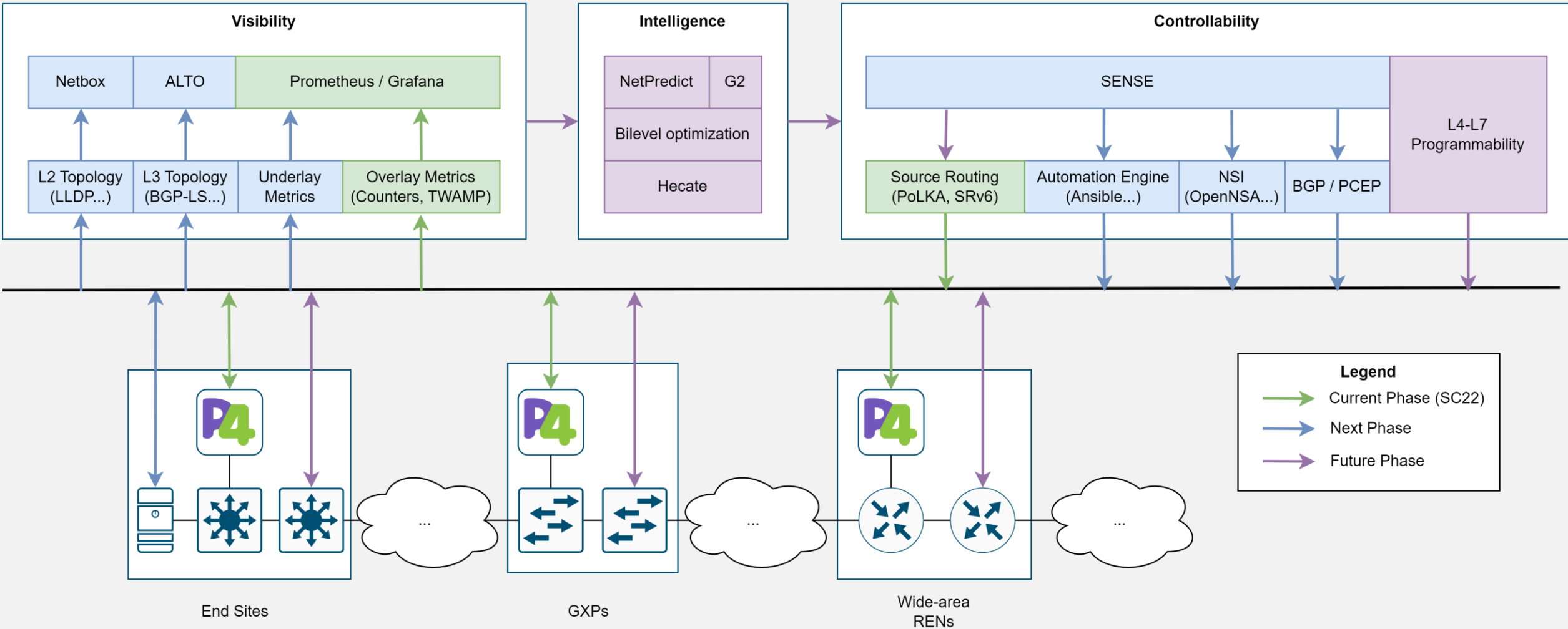
Current (SC22)

- Persistent global L3 overlay network based on P4 switches
- Intercontinental high capacity transfers (100G and over) exploring multiple source routing solutions (Segment Routing and PoIKA)
- Management infrastructure and tools used to operate this global network
- Core network based on RARE/FreeRtr and edge networks based also on SONiC

Future (2023)

- Capability to support multiple virtual networks, that implement on the same devices different choices of routing stacks, traditional and SDN based
- Integration with initiatives for visibility, controllability and intelligence
- Work as a reference state of the art / next generation R&E network

Integration Path with other Initiatives



—● Challenges / Conclusions

Can we increase the rate of evolution without interfering with production networks?

To develop and operate end-to-end / multi-domain orchestration services:

- Resource reservation (guaranteed bandwidth)
- Resource provisioning (Circuits, VRFs)
- Underlay observability
- Dynamic traffic steering/engineering
- Dynamic creation of L3 VPNs
- Closed loop multi-domain visibility/intelligence/controllability

How can we create/sustain an integration initiative/platform to propose and validate next generation protocol and services?

- Proposition: Use programmable P4 devices to experiment on pre-production networks leveraging industry/R&E open ecosystems



— Acknowledgements

GNA-G Data Intensive Science WG

GNA-G AutoGOLE / SENSE WG

GEANT RARE Project

IETF ALTO WG

... And all it's collaborating institutions and teams

Thanks! Questions?

Marcos Schwarz
marcos.Schwarz@rnp.br



MINISTÉRIO DO
TURISMO

MINISTÉRIO DA
DEFESA

MINISTÉRIO DA
SAÚDE

MINISTÉRIO DAS
COMUNICAÇÕES

MINISTÉRIO DA
EDUCAÇÃO

MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÕES

