TURNING
INTO REALITY

Explorando Soluções de Roteamento IP em Redes OpenFlow

Christian Esteve Rothenberg, Ph.D.
Converged Networks Division (DRC)

Agenda

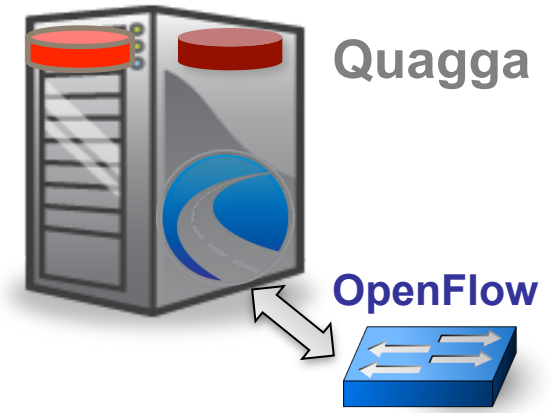
Brief Introduction

Basics of OpenFlow/SDN

RouteFlow

Project
Architecture
Prototype
Demos
Use Cases

Open discussion!



Brief Introduction

CPqD

- Major telecom R&D center in LATAM with expertise in various areas:
  - Optical (WDM, PON), Wireless (WiMax, LTE), IP (IMS/NGN, SDN/OpenFlow), OSS/BSS, Digital TV, SmartGrid/SmartCities
  - Today with ~1200 highly-skilled employees
  - Created in 1976 as R&D branch of Telebras - Brazilian telecom monopoly
  - Not-for-profit private foundation since 1998 after Telebras was privatized

- Purpose is to foster innovation to help (mainly) Brazilian companies and society
  - Focus on technology R&D
  - Bridge the gap between universities and the industry

- Near highly-ranked universities in Brazil
  - History of collaborations

CPqD – Network division

- ~200 researchers and engineers doing R&D
- Various product technologies transferred to the industry (mostly Brazilian companies) since privatization
- 37 patents filled in 2011
- 52 papers published in 2011
- Most successful spin-offs (and consumers of CPqD technology)
  - Tropico: created in 1999 with focus on NGN/IMS; US$ 120M revenue in 2011, with growing presence in South America
  - Padtec: created in 2001 with focus on WDM; US$ 150M revenue in 2011; WDM market leader in Brazil, with growing presence in South America and Europe

Trópico TELECOMUNICAÇÕES AVANÇADAS
Promon
CISCO
Padtec
ideiasnet

OpenFlow data plane developments at CPqD

OpenFlow 1.0 Switch (done 2010!)

- 24 x 10/100/1000
- 2 x 10Gb
- L2/L3
- ~2000 flow entries
- No protocol stack

OpenFlow 1.3 ROADM (ongoing)

- 5-degree WSS for mesh networks
- Contentionless-Directionless-Colorless
- Flexgrid
- Virtualization-capable
- Multicast-capable

OpenFlow 1.3 – 1a. no mundo

ERICSSON

Search Mailing List Archives

Search openflow-discuss!

Limit search to: Subject & Body Subject Author
Sort by: Date Reverse Sort
Limit to: All This Week Last Week This Month Last Month
Select Date Range Apr 13 2012 through Apr 13 2012

[openflow-discuss] Announcing OpenFlow version 1.3 software switch

Christian Esteve Rothenberg esteve at cpqd.com.br
Fri Nov 9 11:41:49 PST 2012

- Previous message: [openflow-discuss] POX and cbench
- Next message: [openflow-discuss] Announcing OpenFlow version 1.3 software switch and cont prototyp
- **Messages sorted by:** [date] [thread] [subject] [author]

```
Dear OpenFlow fellows,

giving continuity to the OpenFlow 1.X developments by Ericsson and
CPqD (cf. v1.2 Tool-Kit [1]) we are glad to announce the release of a
version 1.3 software switch [2] (based on Stanford reference design
extended by Ericsson for v1.1) and a companion NOX OpenFlow 1.3
controller [3] (based on NOX Zaku and
using oflib 1.3 of the software switch).
```

Forte candidata a se tornar a implementação de referência da ONF

Propósito da Apresentação

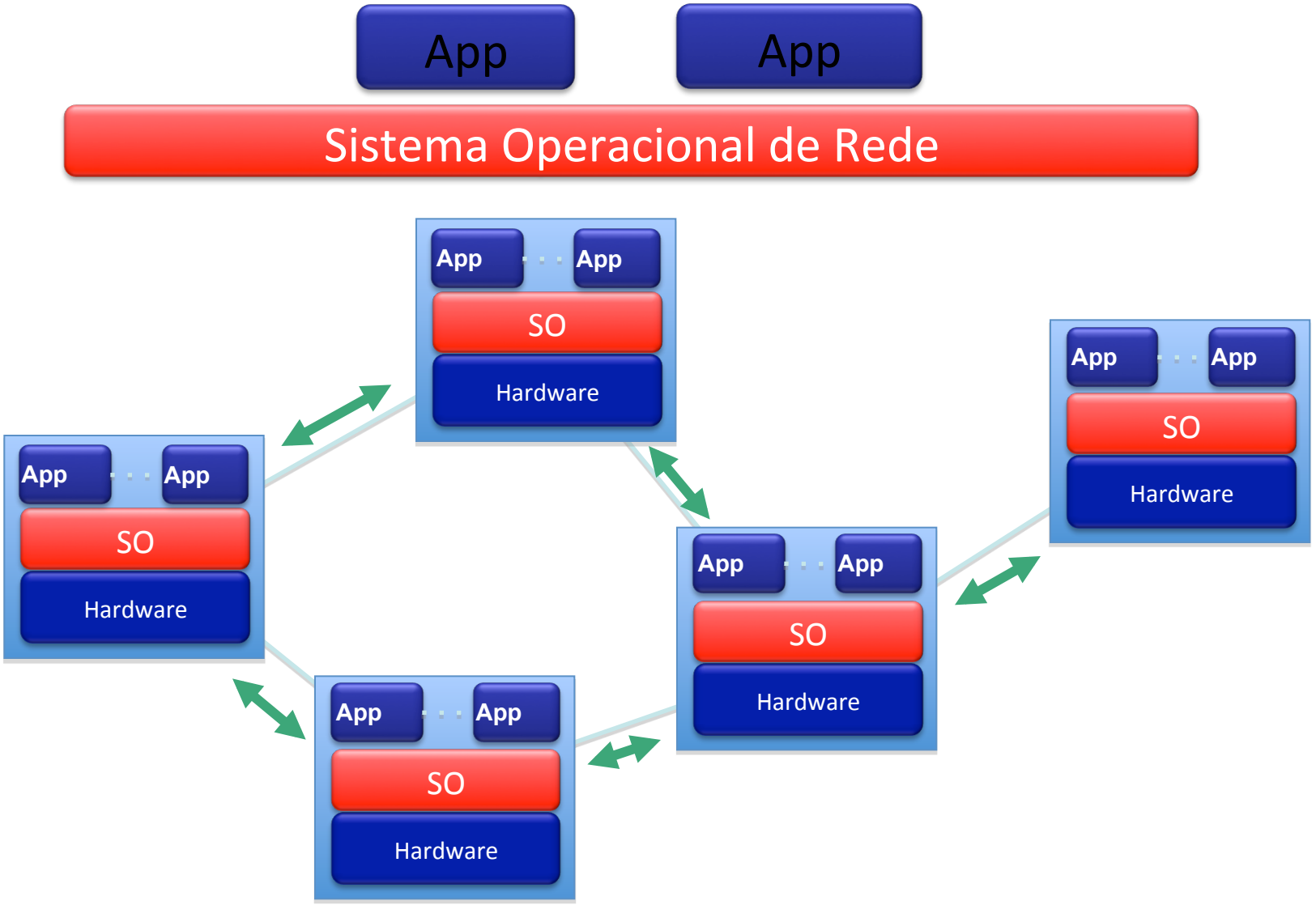
- Oferecer uma solução de software gratuita que pode
  - resolver seus problemas prontamente
  - ser usada como base para resolver seus problemas
- Ouvir de vocês, que projetam e operam redes diariamente
  - que problemas precisam resolver
  - se RouteFlow tem potencial para resolver esses problemas
  - se gostariam de usar RouteFlow
- Trabalhar com vocês para resolver esses problemas usando RouteFlow

SDN/OpenFlow

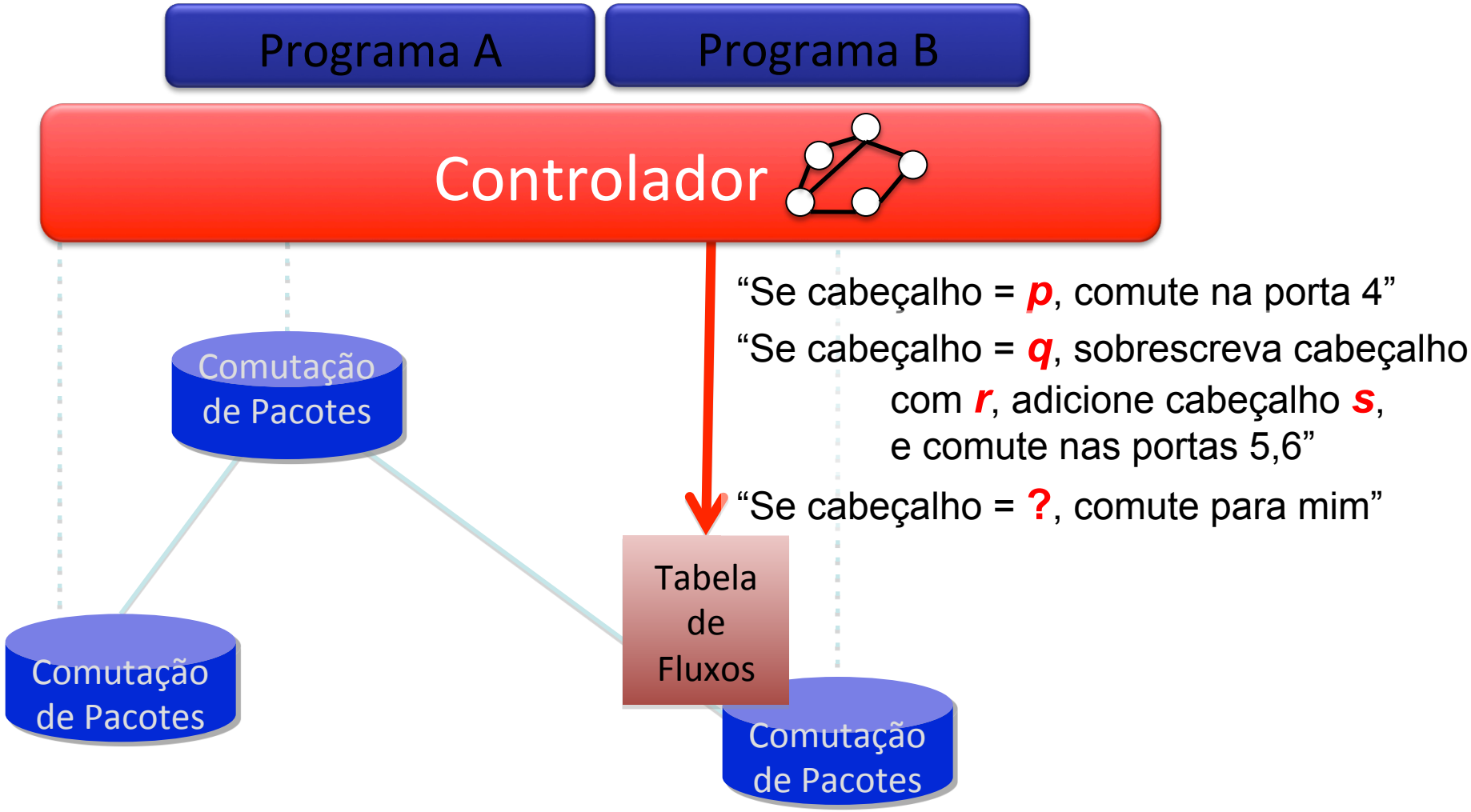
Brief Introduction

(slides from Nick Mckeown, U. Stanford, USA)

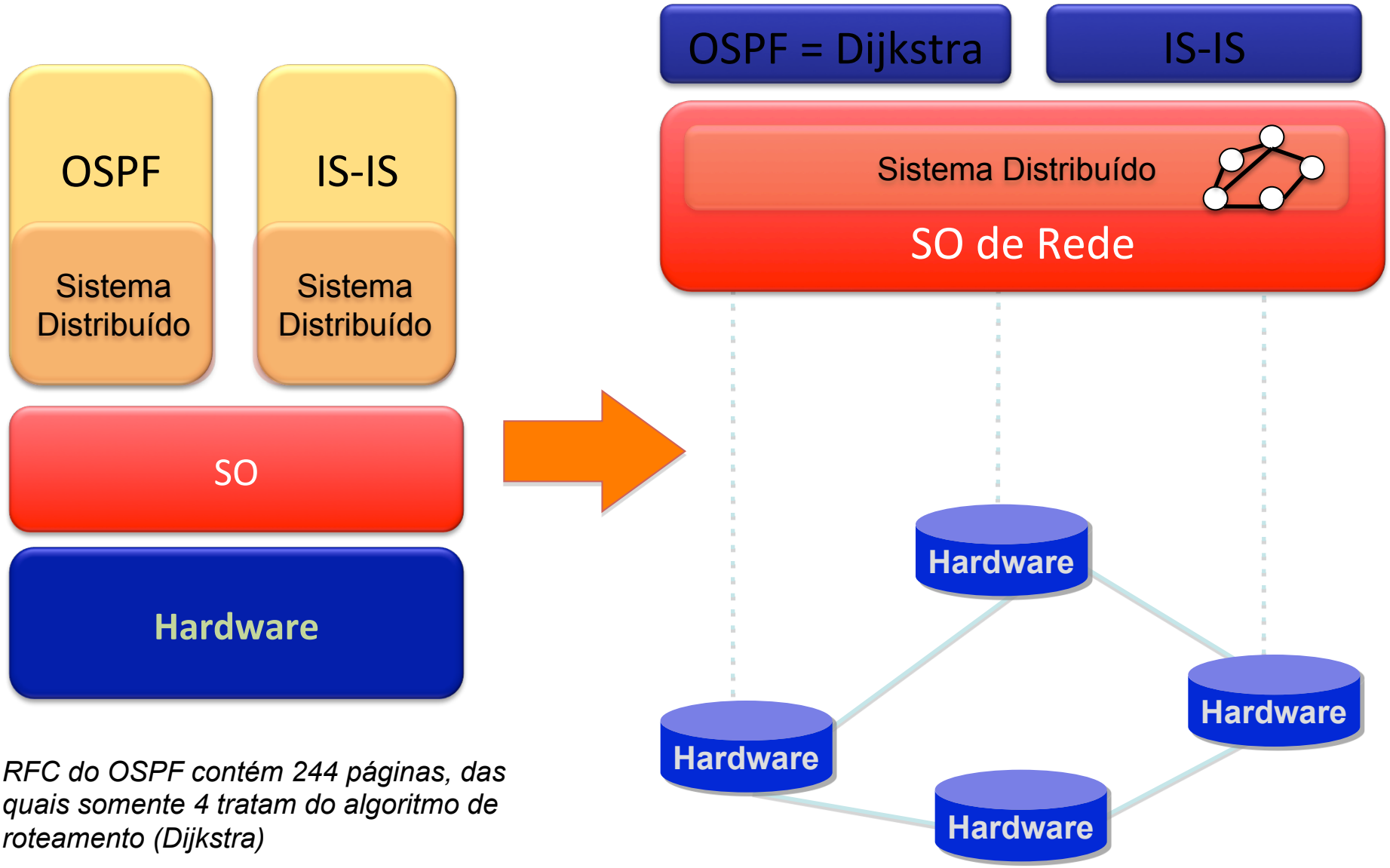
Rede Definida por Software (SDN)



OpenFlow: Exemplo básico

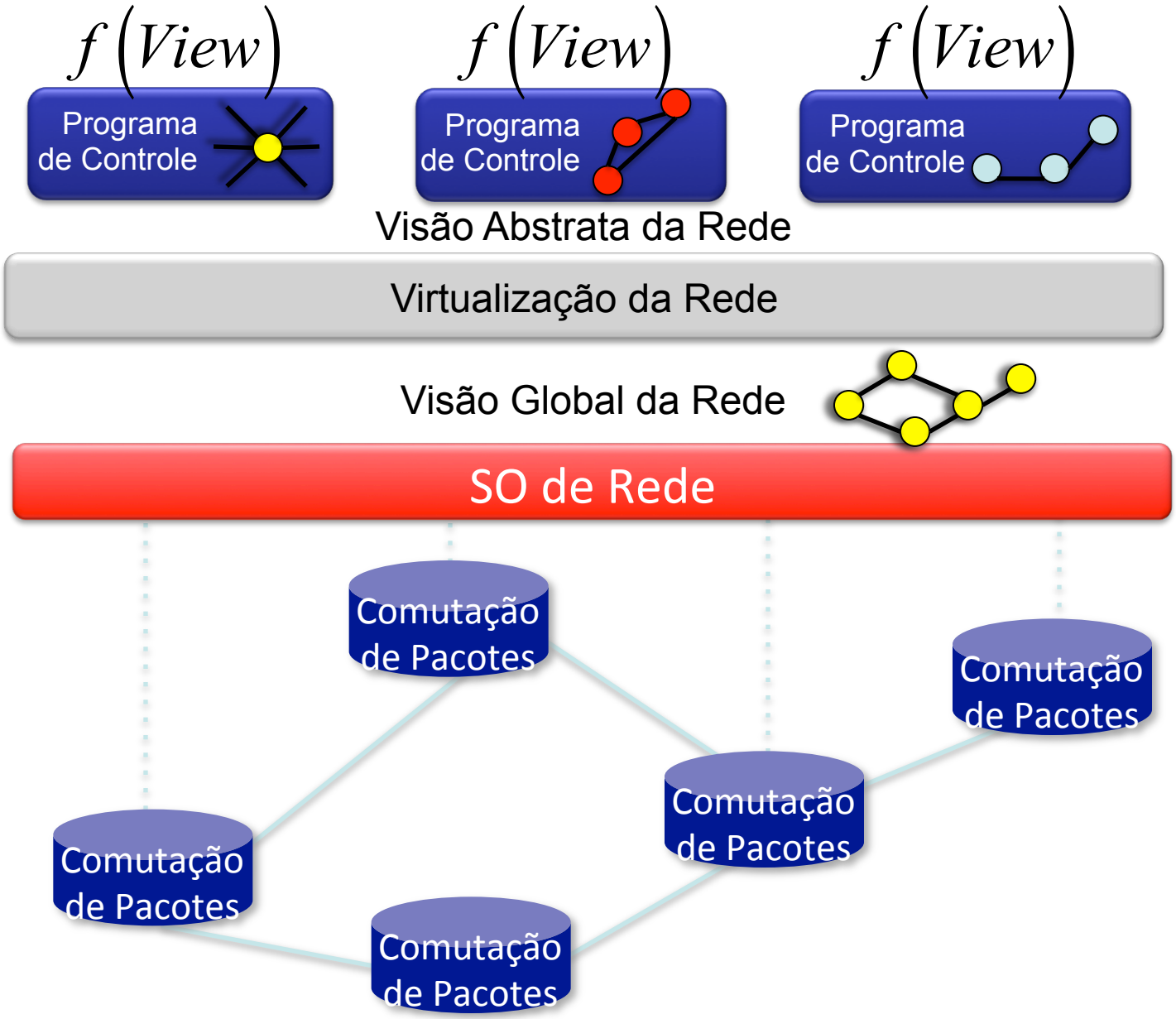


Exemplo: Roteamento Traditional x SDN



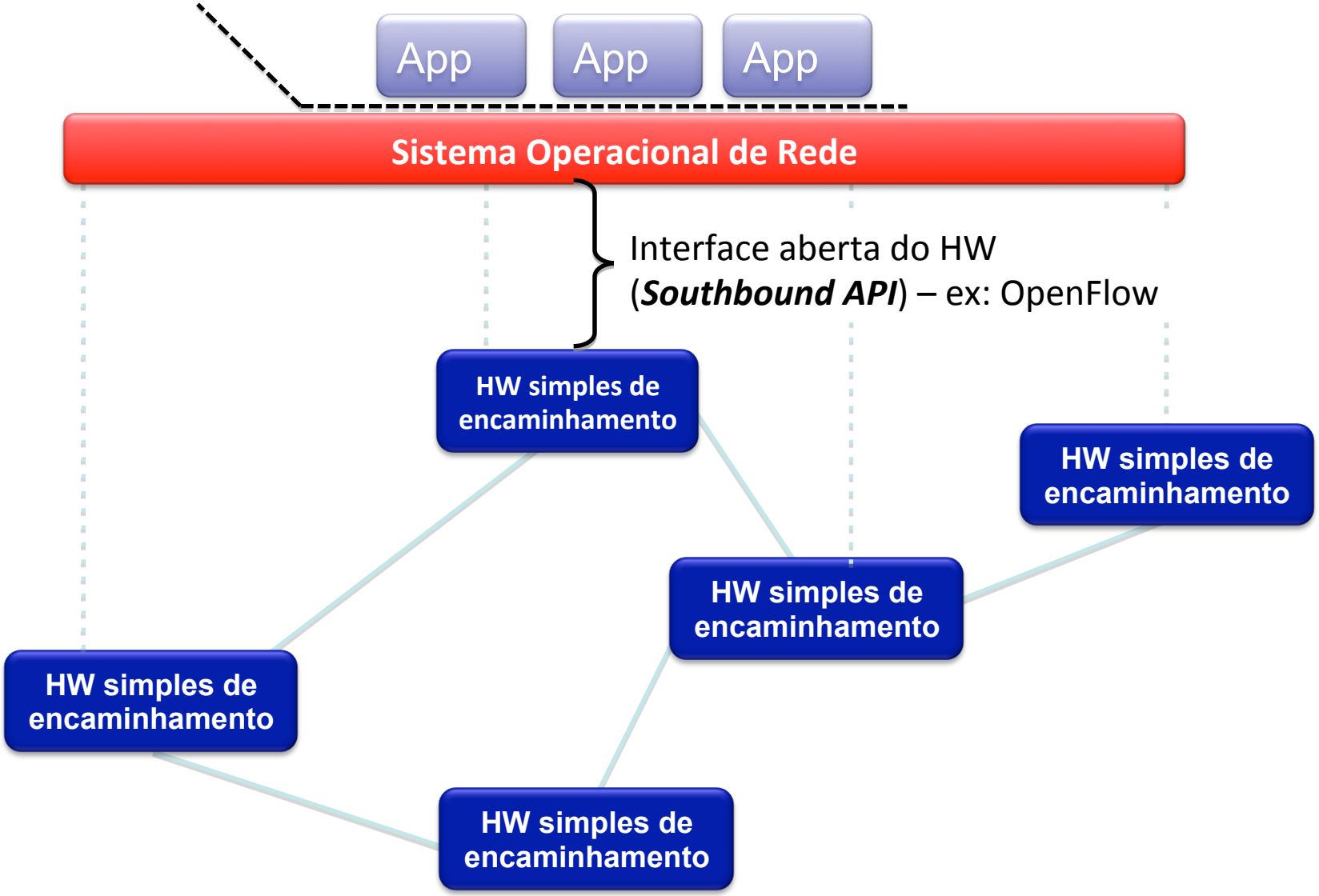
RFC do OSPF contém 244 páginas, das quais somente 4 tratam do algoritmo de roteamento (Dijkstra)

Visão atual de SDN: Programabilidade e Virtualização



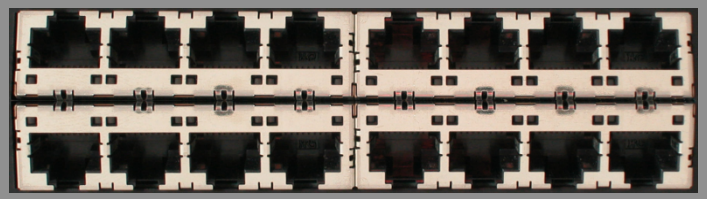
Arquitetura SDN

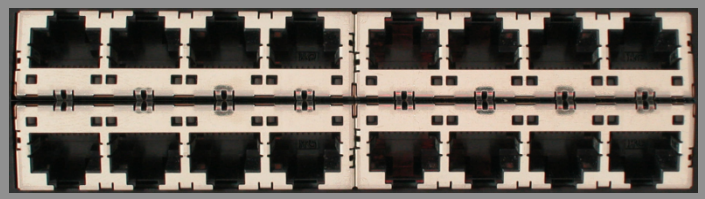
Interface aberta do SO de Rede (***Northbound API***)

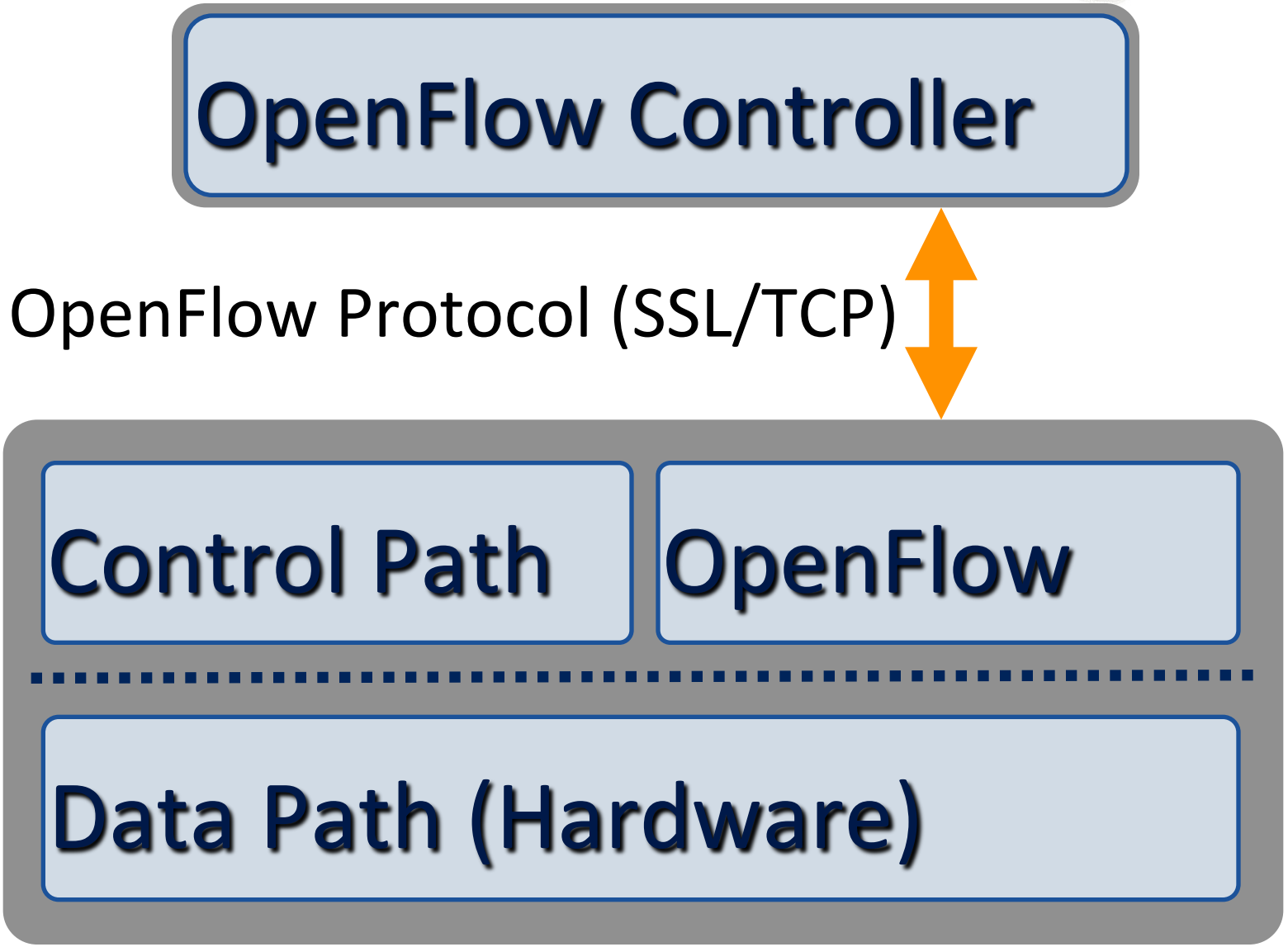


How does OpenFlow work?

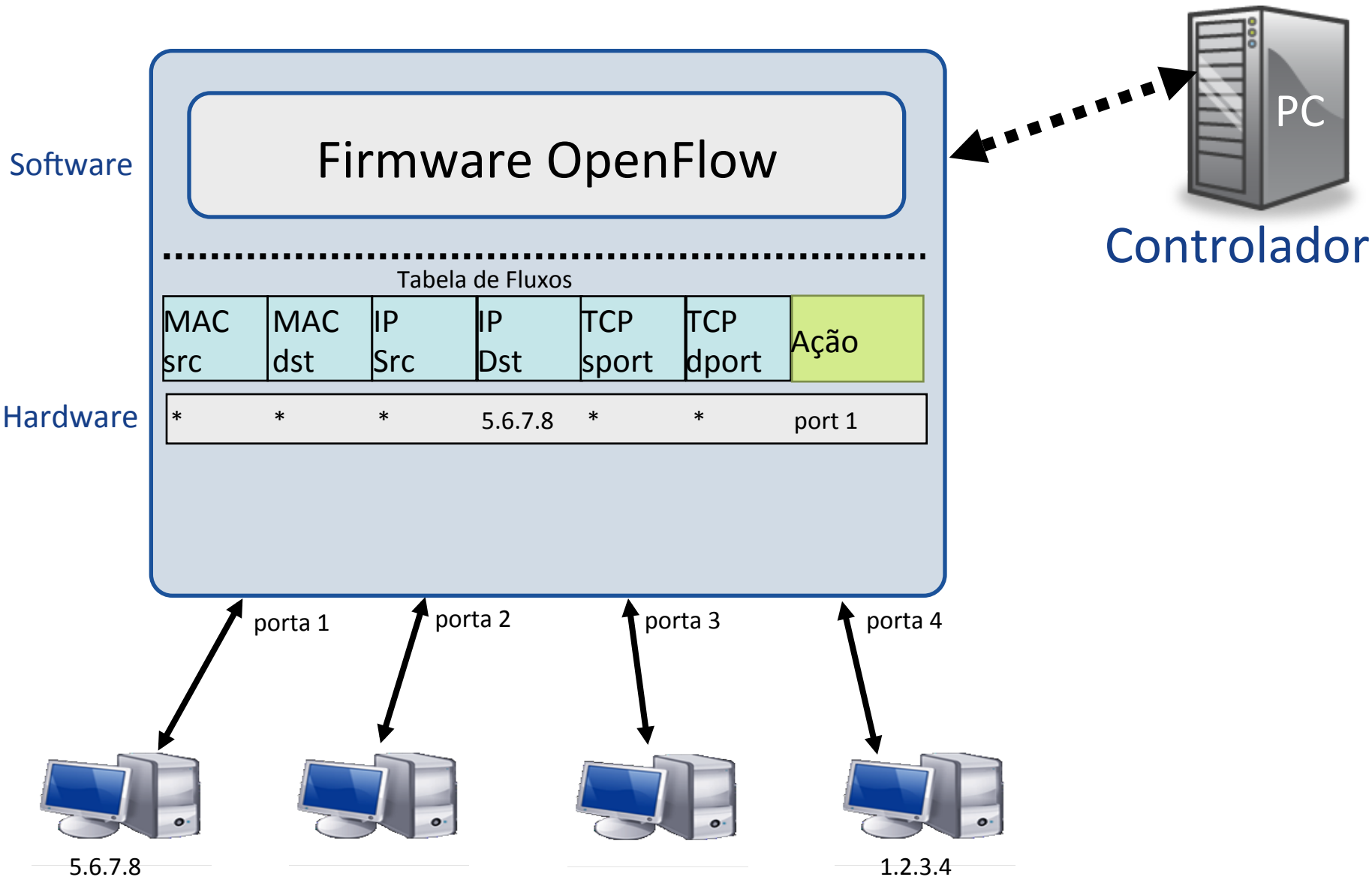
Ethernet Switch





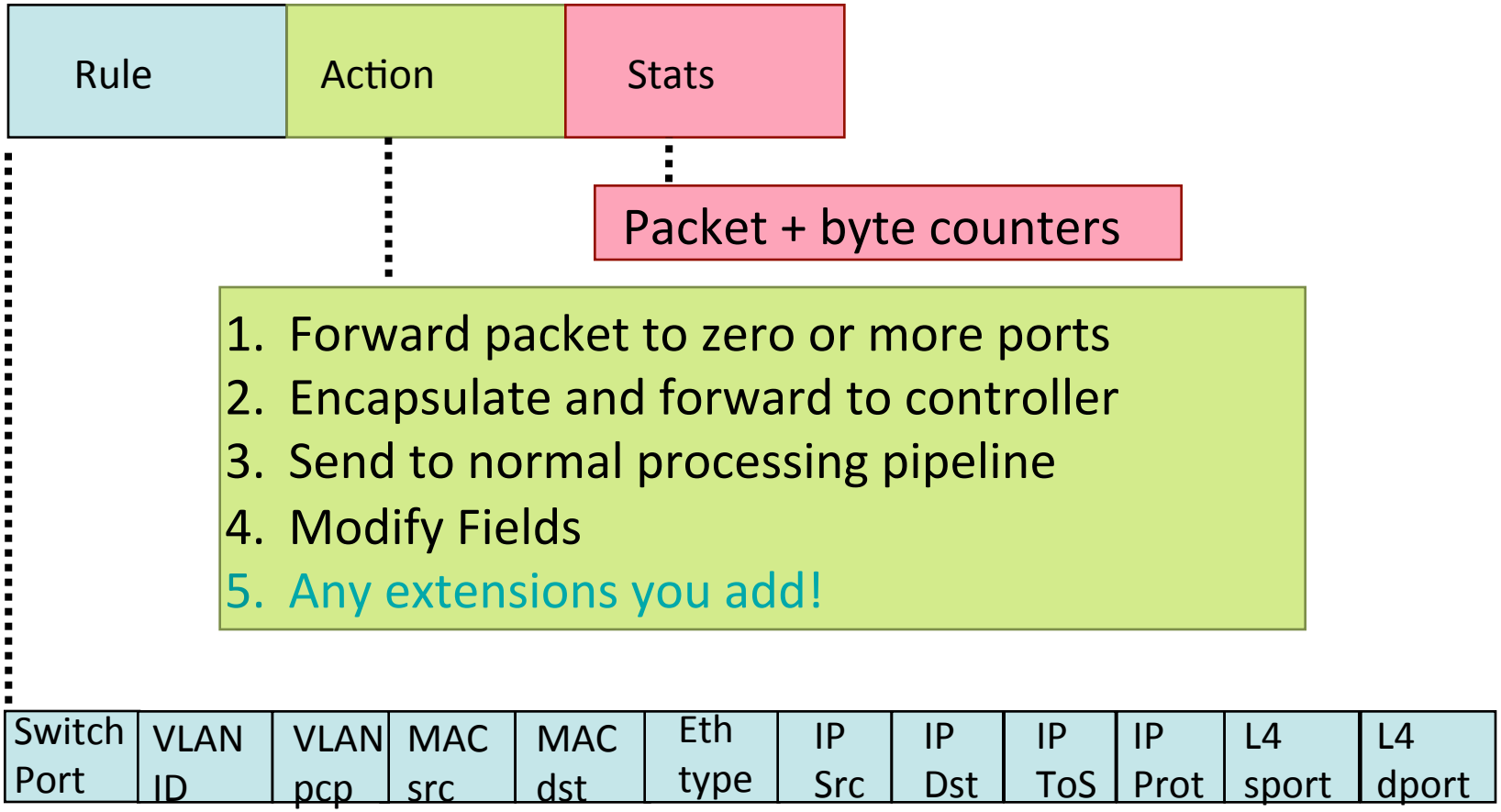

OpenFlow Controller
OpenFlow Protocol (SSL/TCP)
Control Path
OpenFlow
Data Path (Hardware)

OpenFlow: Interface de Programção do Hardware



OpenFlow Basics

Flow Table Entries



+ mask what fields to match

OpenFlow: Flexibilidade da abstração de fluxo

Comutação de quadros

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
*	*	00:1f:..	*	*	*	*	*	*	*	Porta 6

Comutação de fluxos

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
port3	00:20..	00:1f..	0800	vlan1	1.2.3.4	5.6.7.8	4	17264	80	Porta 6

Firewall

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
*	*	*	*	*	*	*	*	*	22	descartar

OpenFlow: Flexibilidade da abstração de fluxo

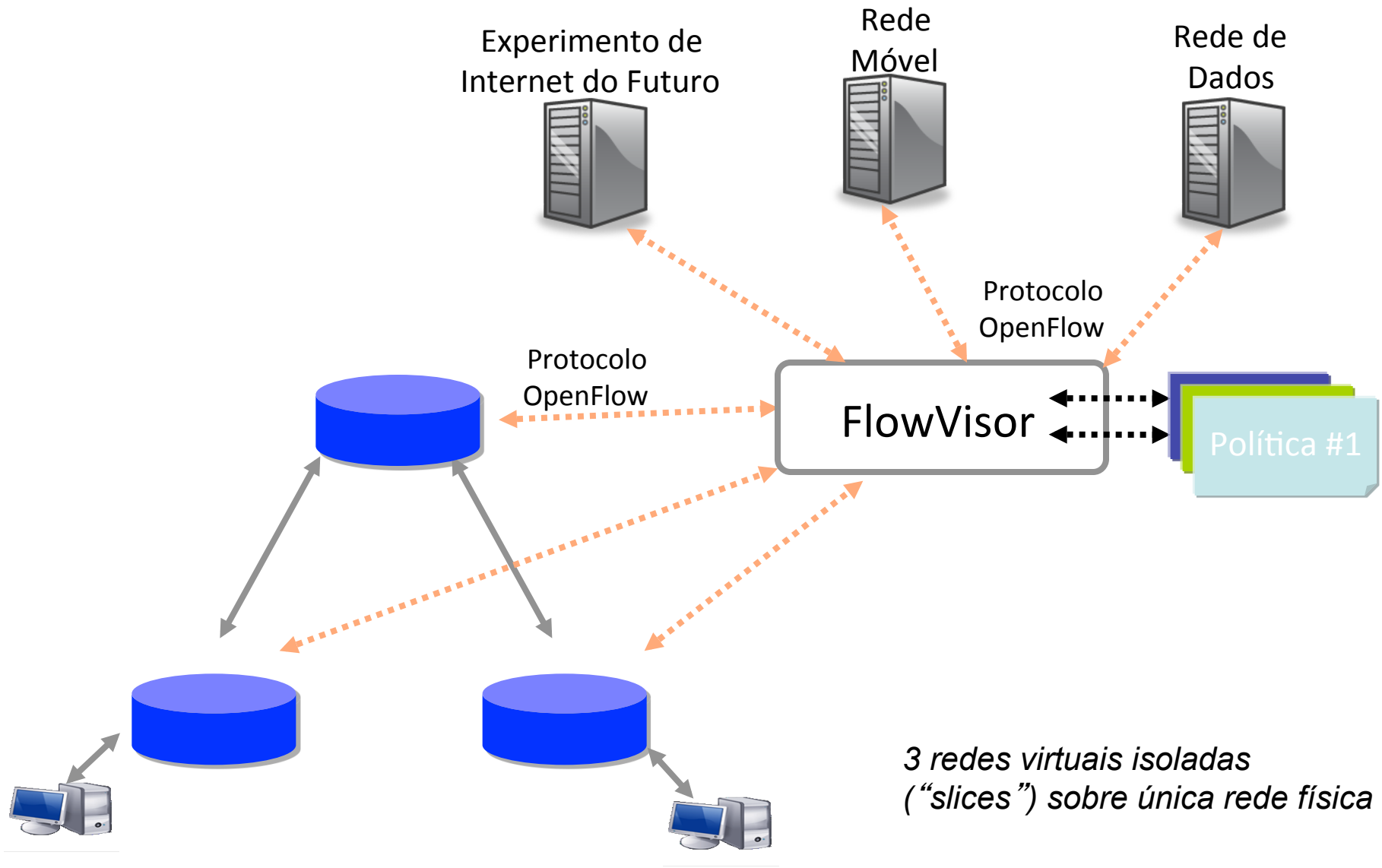
Roteamento

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
*	*	*	*	*	*	5.6.7.8	*	*	*	Porta 6

Comutação de VLAN

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
*	*	00:1f..	*	vlan1	*	*	*	*	*	Porta 6, Porta 7, Porta 9

FlowVisor: Virtualização OpenFlow

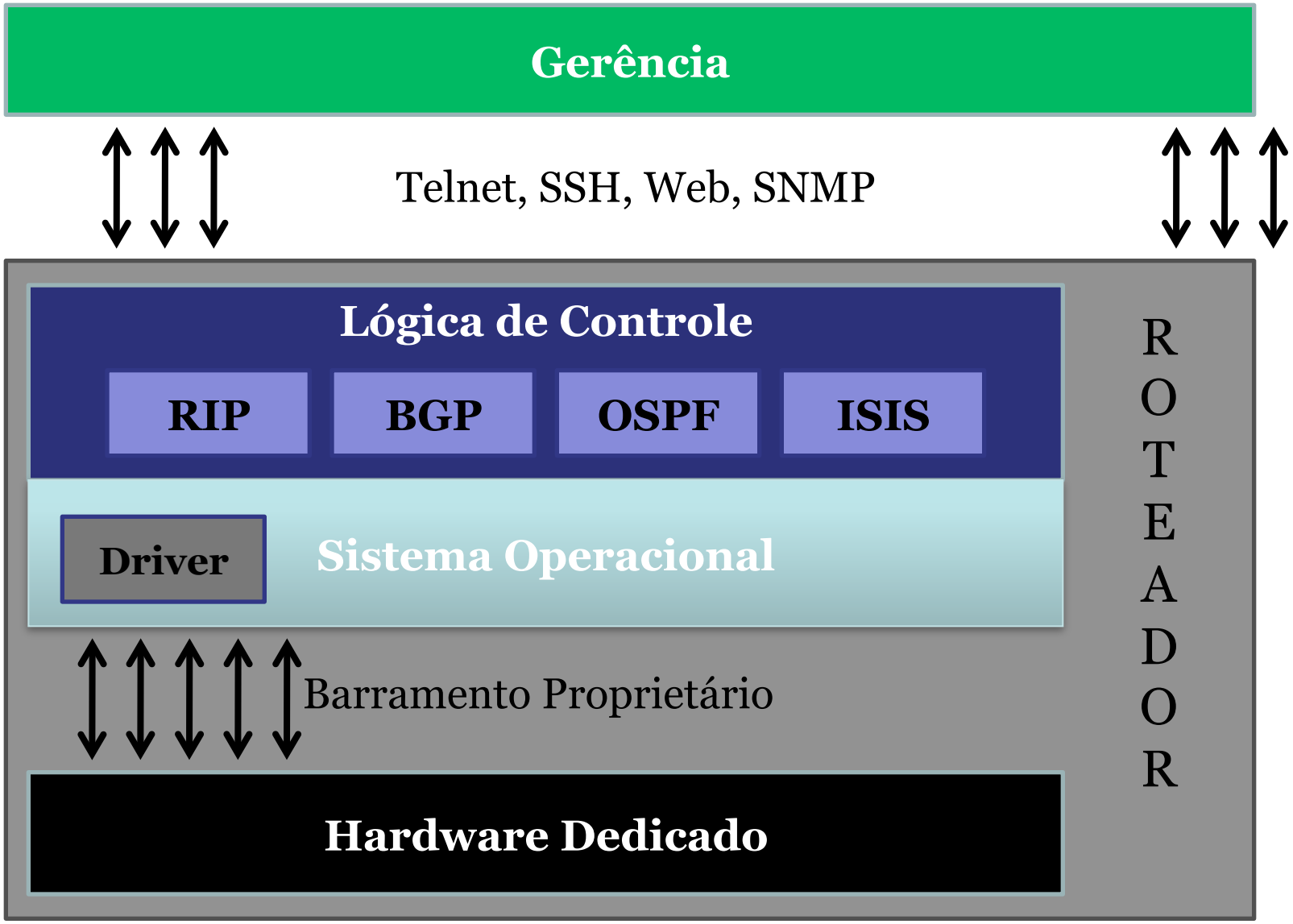


3 redes virtuais isoladas ("slices") sobre única rede física

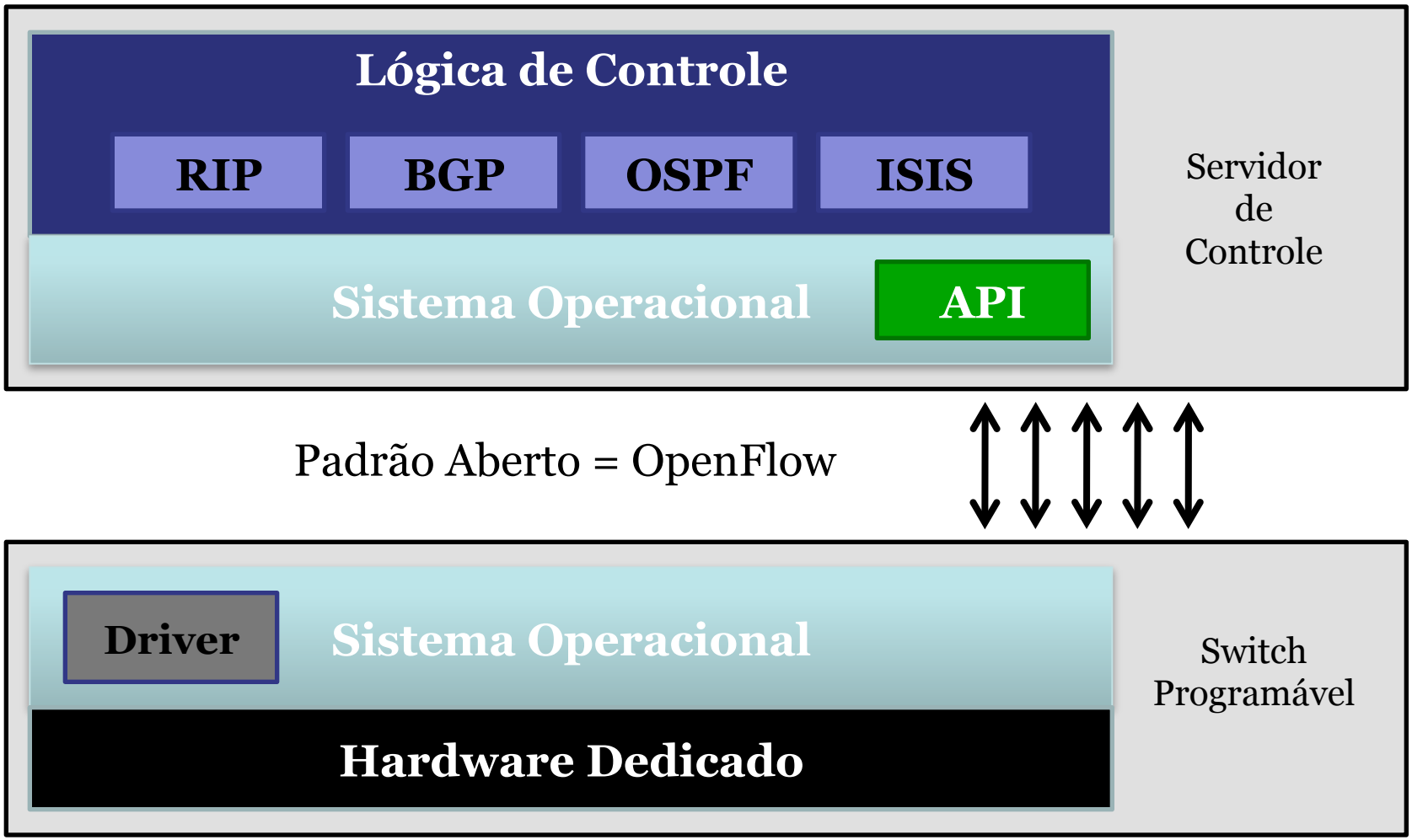
RouteFlow

(Transição de roteamento tradicional para roteamento ideal em SDN)

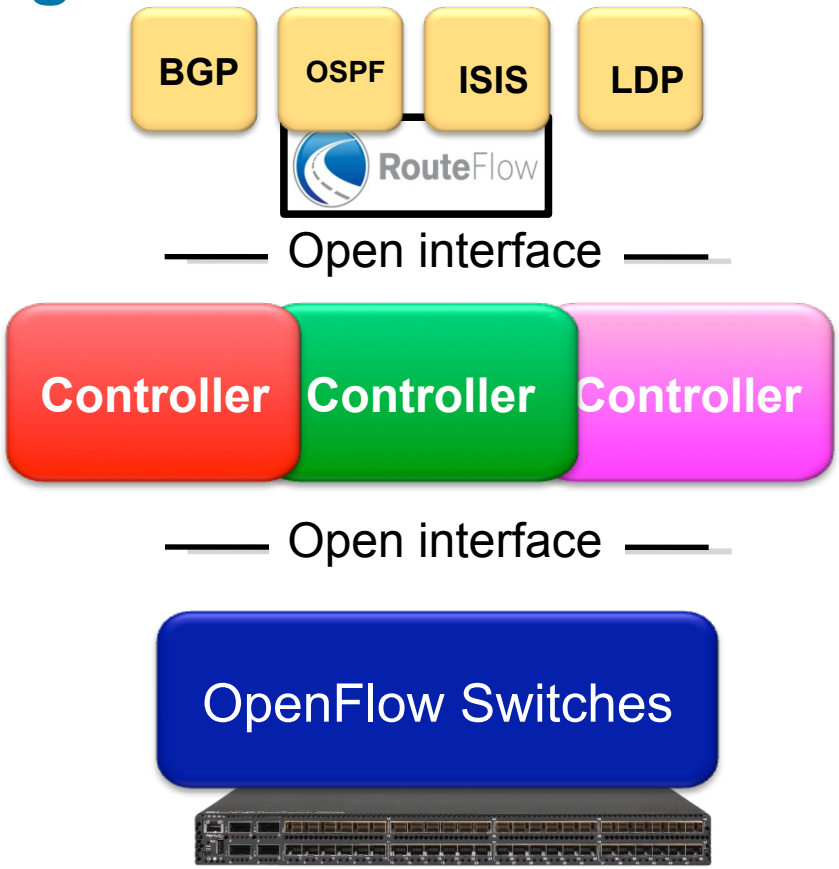
Modelo Atual: planos de controle e encaminhamento fortemente acoplados



Modelo novo: planos de controle e encaminhamento separados por meio do protocolo OpenFlow



Software Defined IP Routing

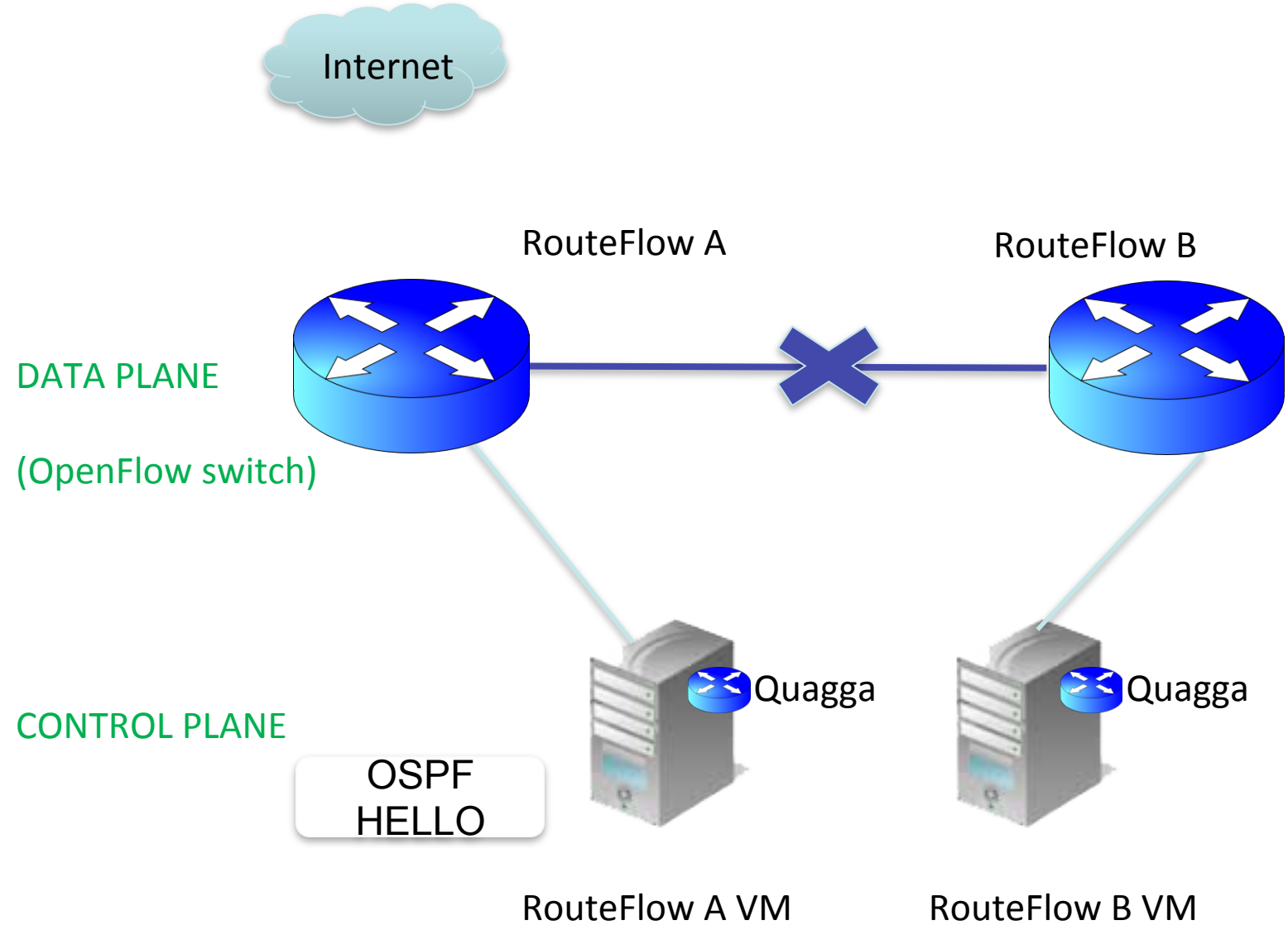


High cost
Specialized config.
Closed source
Slow innovation

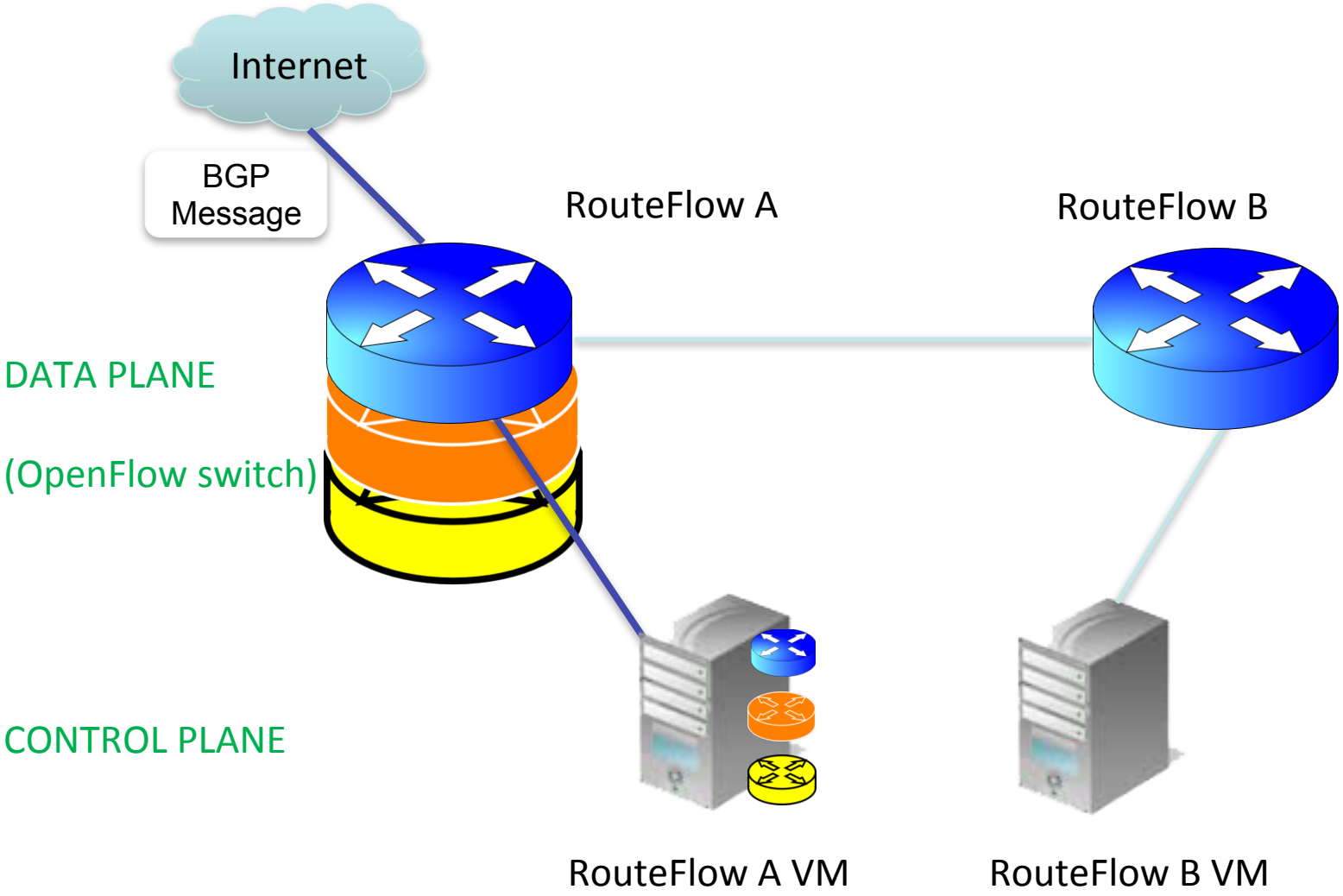
Source: McKeown

Low cost (commodity)
Multi-vendor modularity
Open source
Rapid innovation

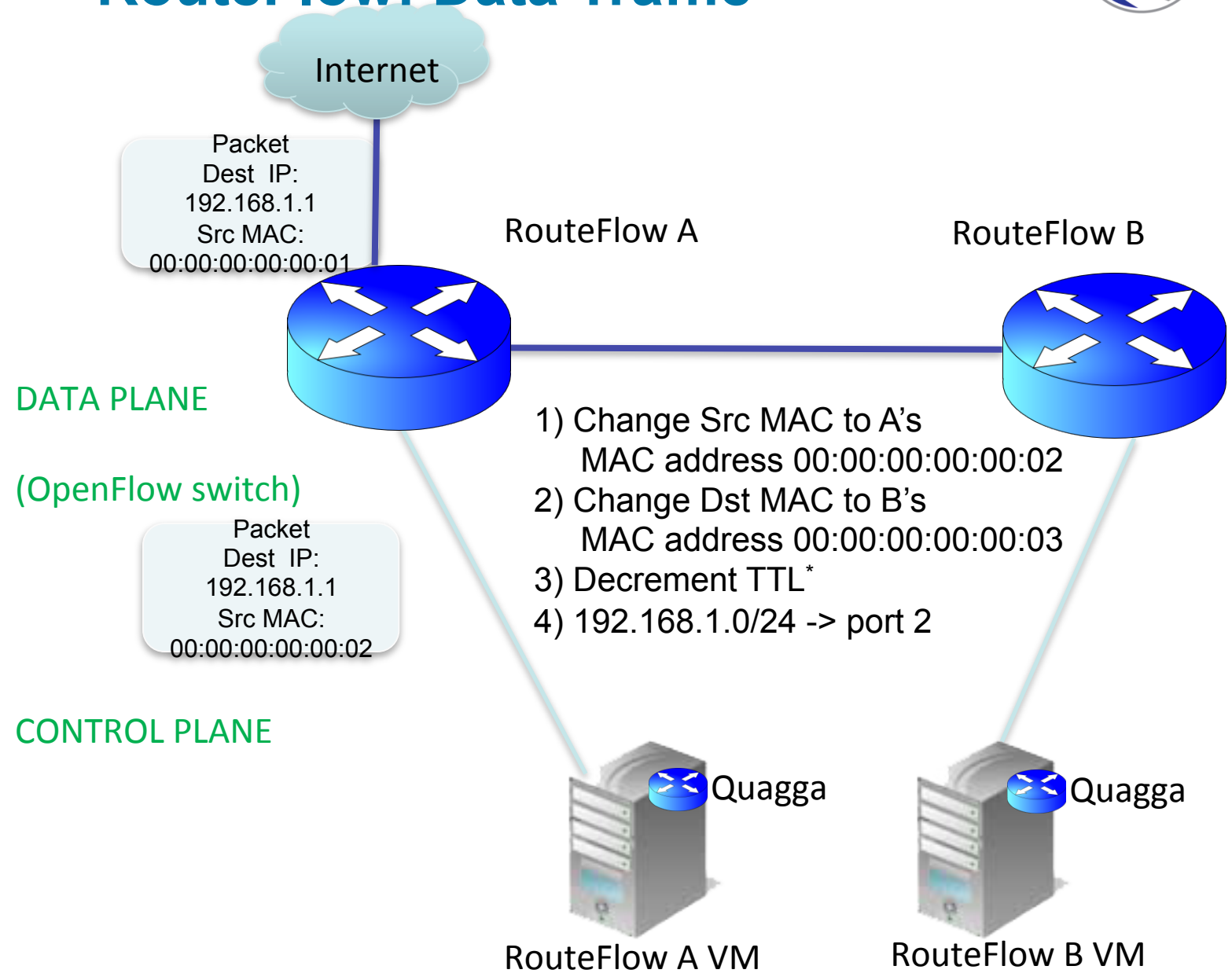
RouteFlow: Control Traffic (IGP)



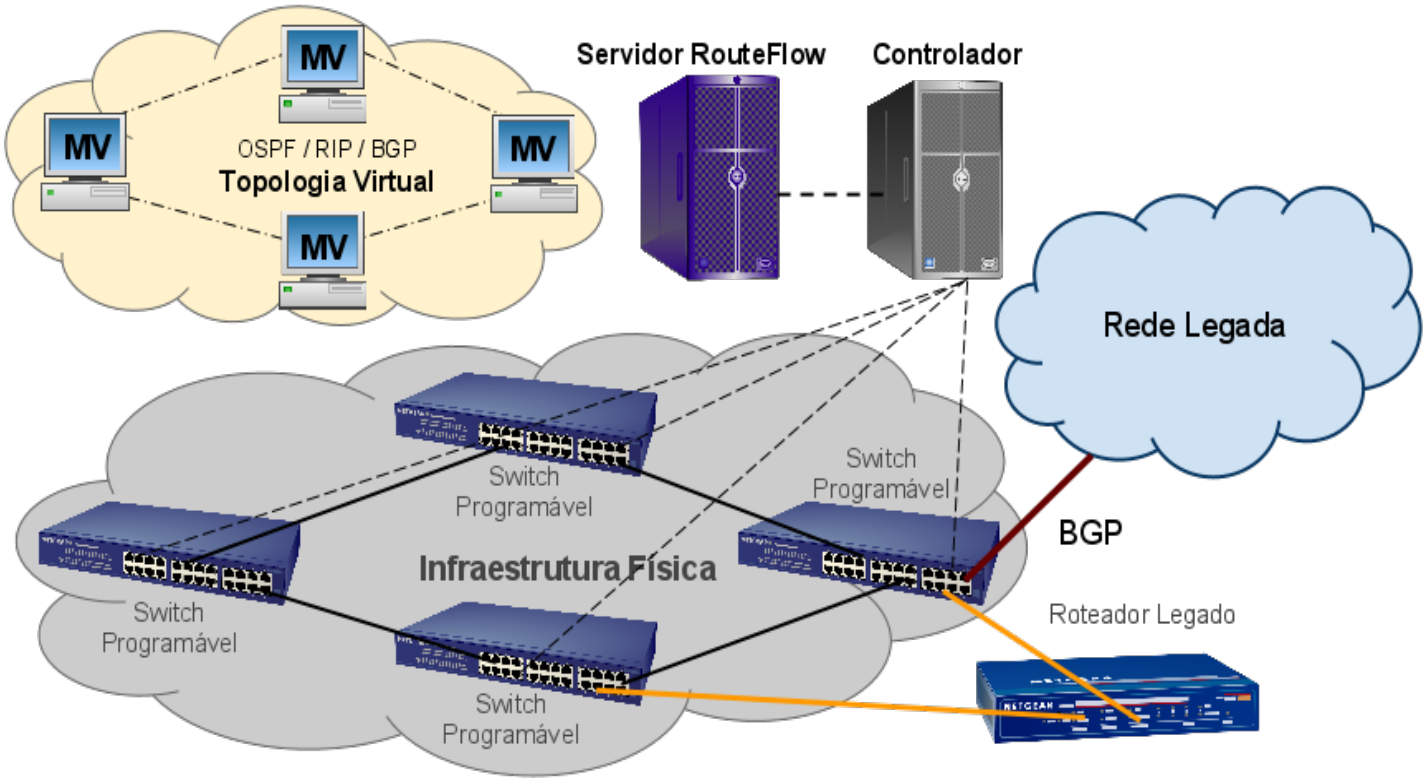
RouteFlow: Control Traffic (BGP)



RouteFlow: Data Traffic

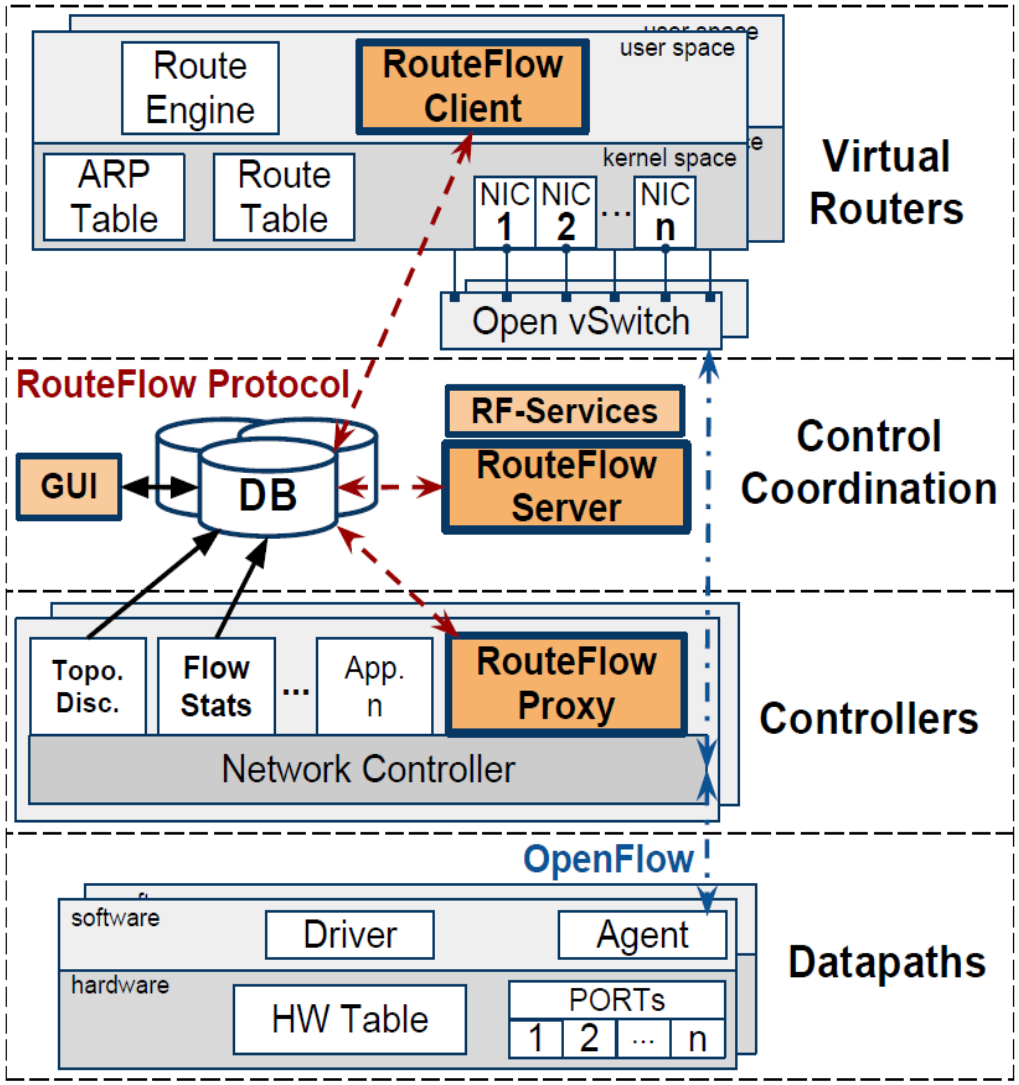


RouteFlow



1. Quagga on steroids
2. Low cost routing solutions
3. Innovative routing solutions
4. Transition path to OpenFlow/SDN

Arquitetura

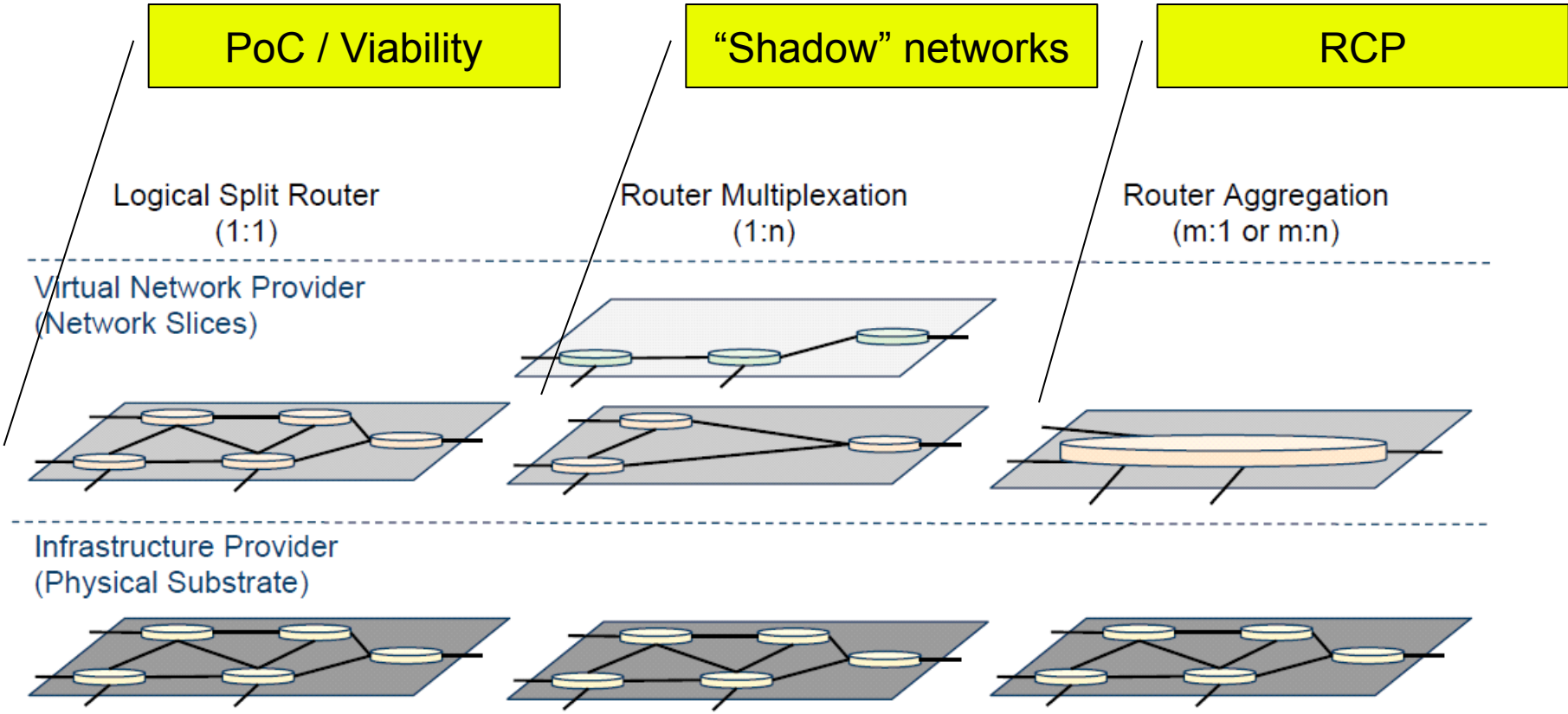


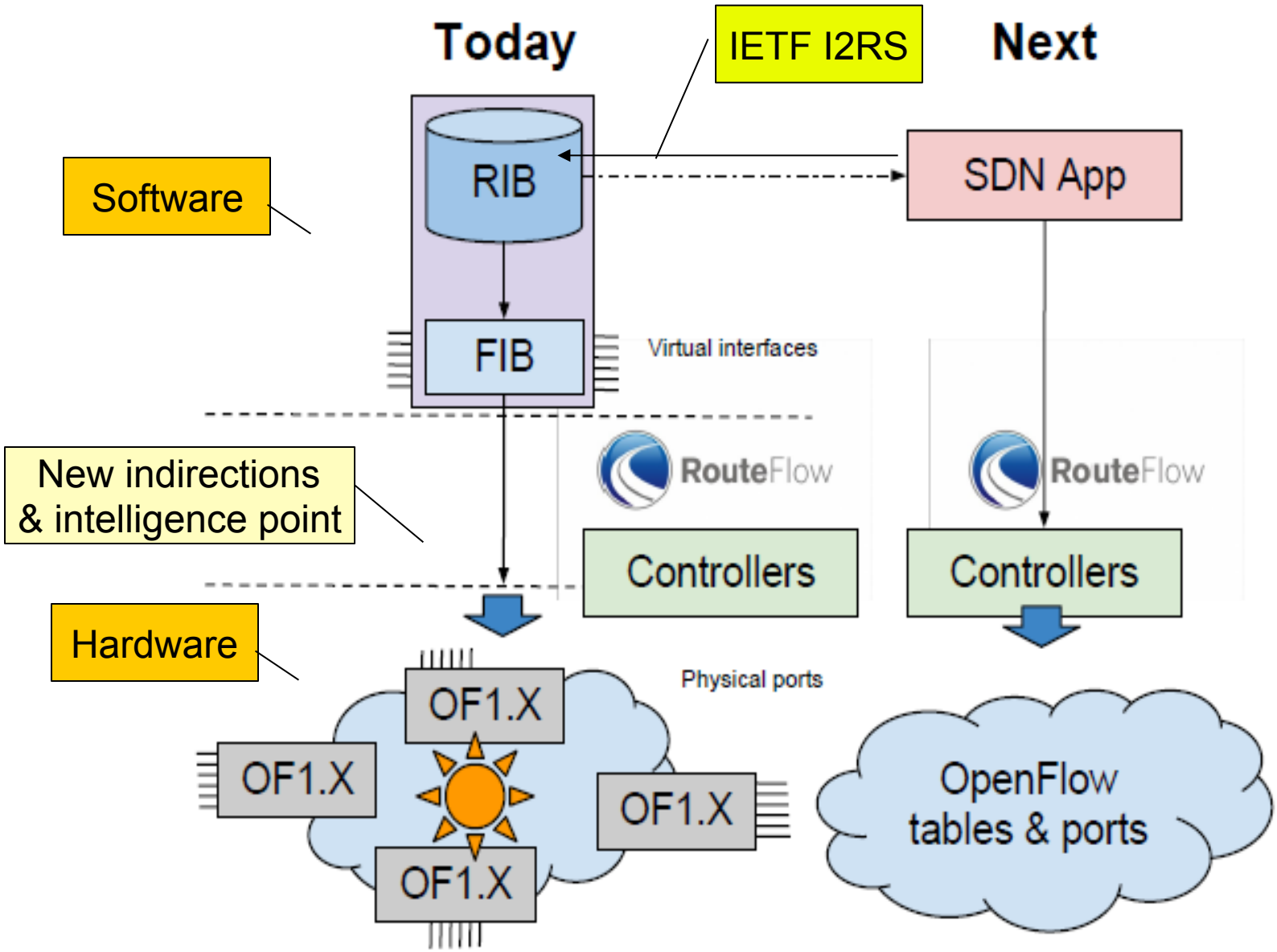
Key Features

- Modular architecture
  - RF-Proxy
  - RF-Server
  - RF-Client
- Database layer
  - JSON-based IPC
  - Resillient core state
  - Programmer-friendly
- Multi-Controller support
  - NOX, POX, Ryu, Floodlight,
  - Trema (ongoing)

Modes of operation

- From logical routers (akin VRFs) to single node abstractions over flexible virtual networks.
- New design choices on the distribution of the control nodes.

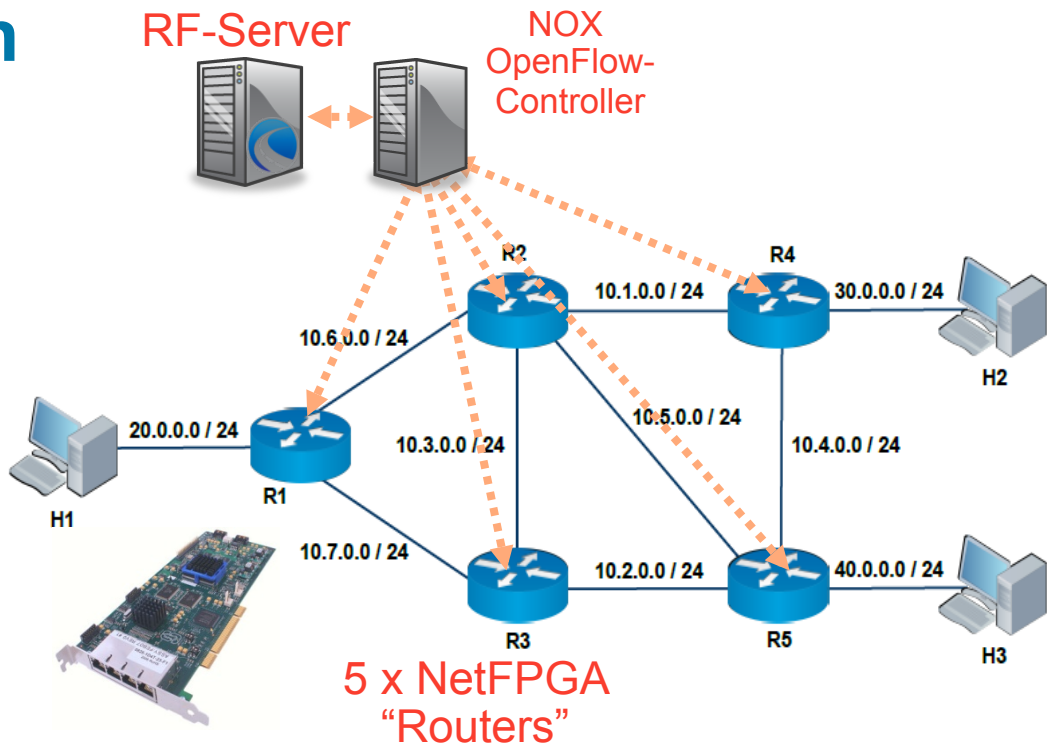



Today
IETF I2RS
Next
Software
RIB
SDN App
FIB
Virtual interfaces
New indirections
& intelligence point
RouteFlow
RouteFlow
Controllers
Controllers
Hardware
Physical ports
OF1.X
OF1.X
OF1.X
OF1.X
OpenFlow
tables & ports

Prototype evaluation

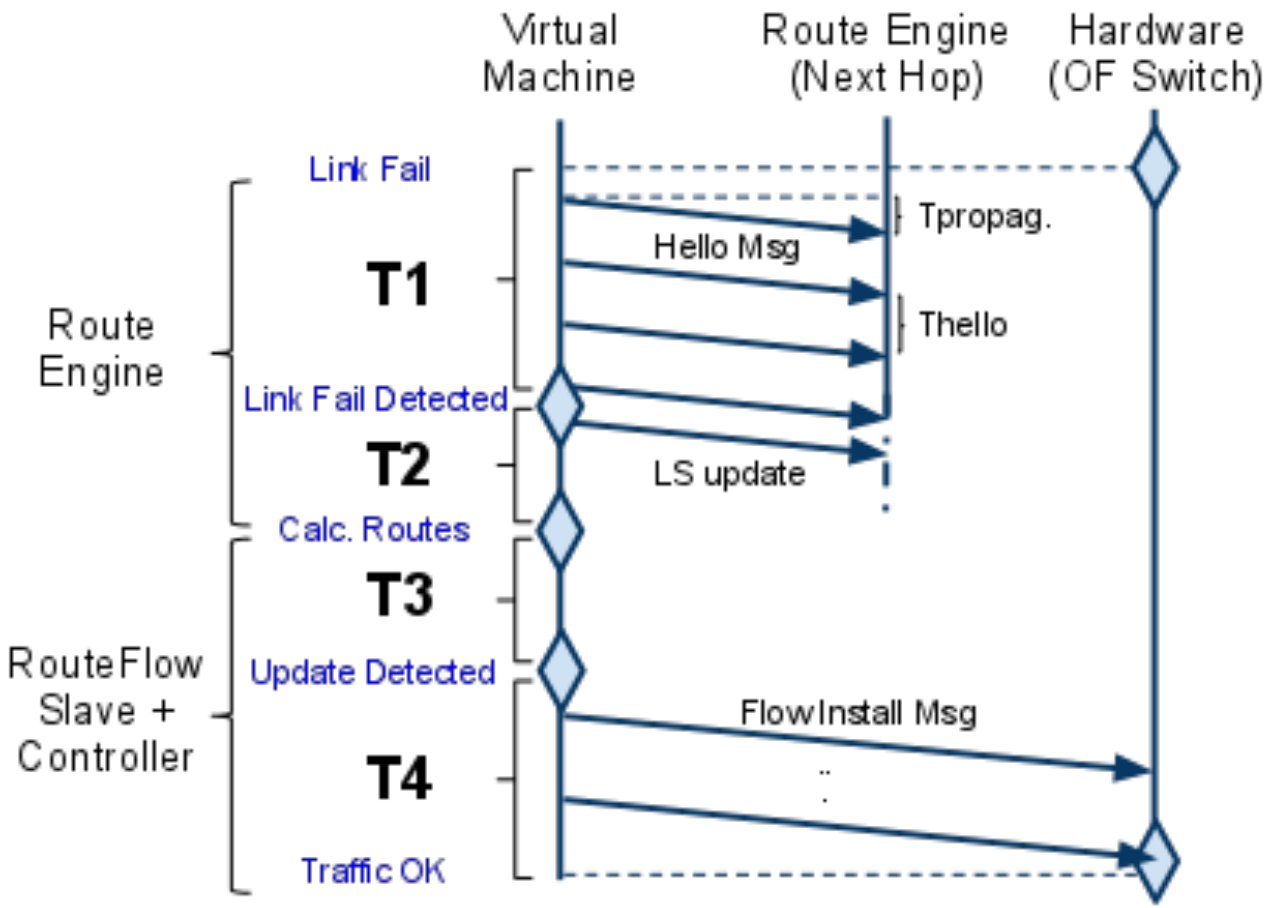
- Back in 2010…

- Lab Setup
  - NOX controller
  - Quagga routing engine
  - 5 x NetFPGAs switches

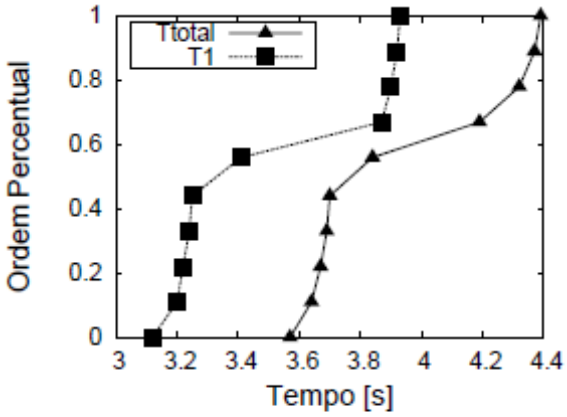


- Results
  - Interoperability with traditional networking gear
  - Route convergence time is dominated by the protocol time-out configuration (e.g., 4 x HELLO in OSPF) not by slow-path operations
  - Compared to commercial routers (Cisco, Extreme, Broadcom-based), larger latency only for those packets that need to go to the slow-path:
    - Miss FIB entry, need processing by the OS networking / routing stack e.g., ARP, PING, routing protocol messages

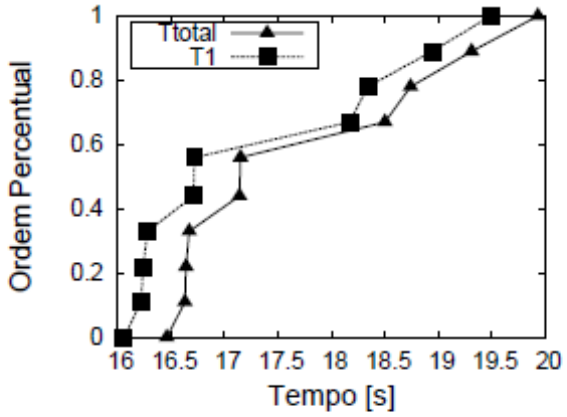
Teste de Convergência com OSPF



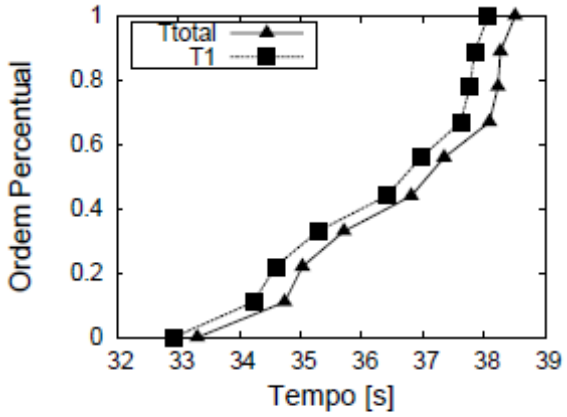
Teste de Convergência com OSPF



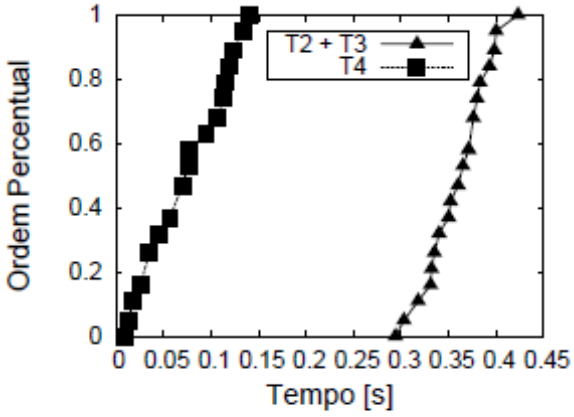
(a) T_{total} e T_1 para T_{Hello} de 1s.



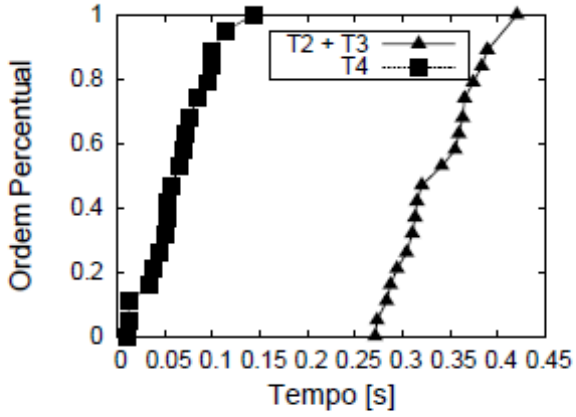
(b) T_{total} e T_1 para T_{Hello} de 5s.



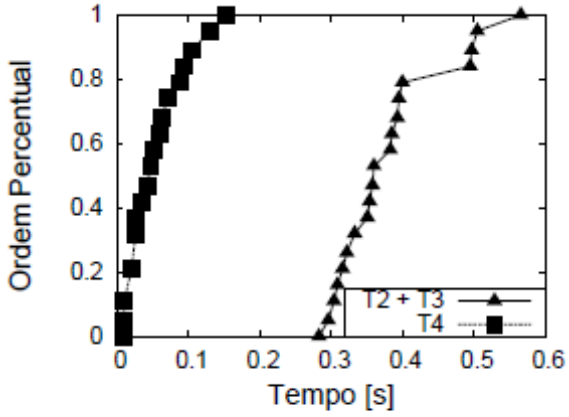
(c) T_{total} e T_1 para T_{Hello} de 10s.



(d) T_{2+3} e T_4 para T_{Hello} de 1s.



(e) T_{2+3} e T_4 para T_{Hello} de 5s.



(f) T_{2+3} e T_4 para T_{Hello} de 10s.

Teste de SlowPath e FastPath

	Slow Path [ms]		Fast Path [ms]	
Equipamento	Tmed.	T90%	Tmed.	T90%
CISCO 3560-e Catalyst	5,46	7,75	0,100	0,130
Extreme x450-e	11,30	14,00	0,106	0,141
CPqD Enterprise	14,20	17,30	0,101	0,147
RouteFlow*	22,00	28,00	0,082	0,119

- Cálculo do RTT com pacotes ICMP;
- Avaliação do overhead das mensagens direcionadas a uma pilha de controle remoto;
- Fast Path: encaminhamento em line rate;

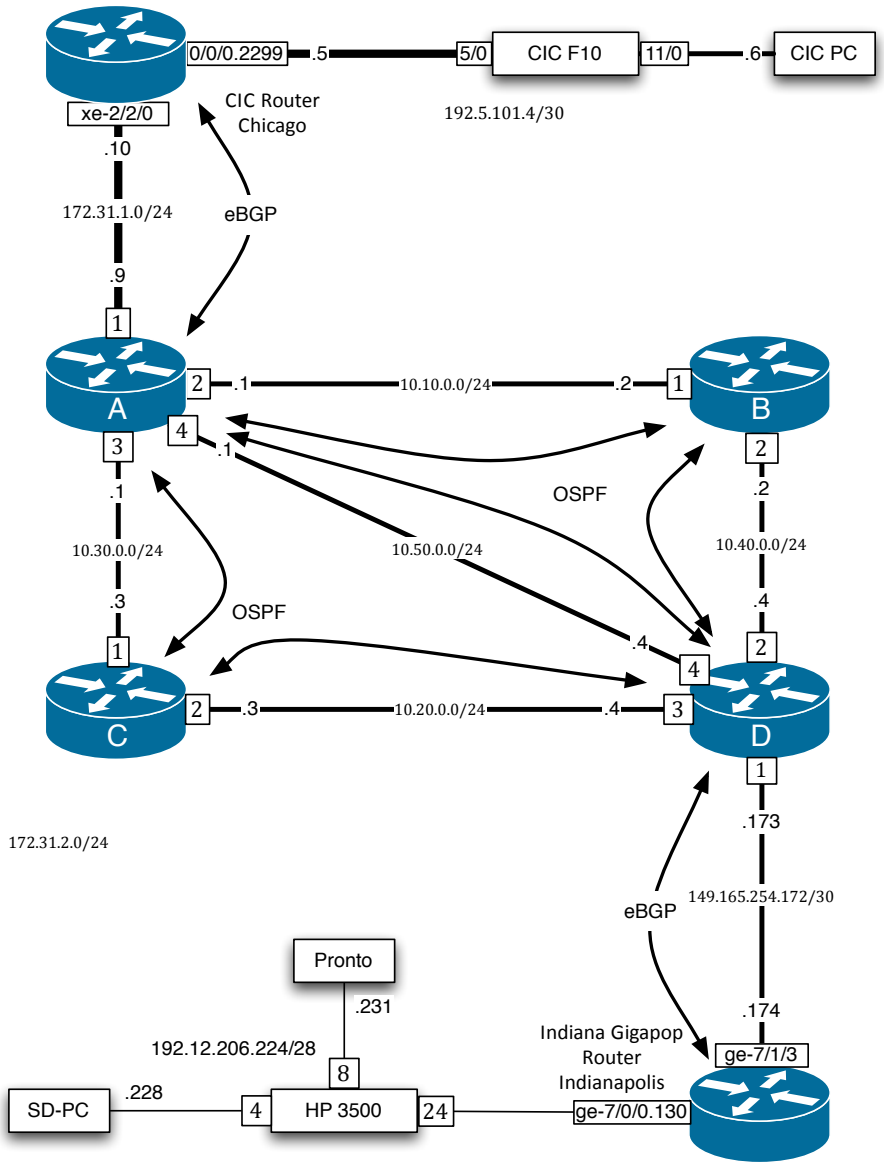
*Newest RouteFlow design and NetFPGA

Demonstration at I Open Networking Summit (Stanford)

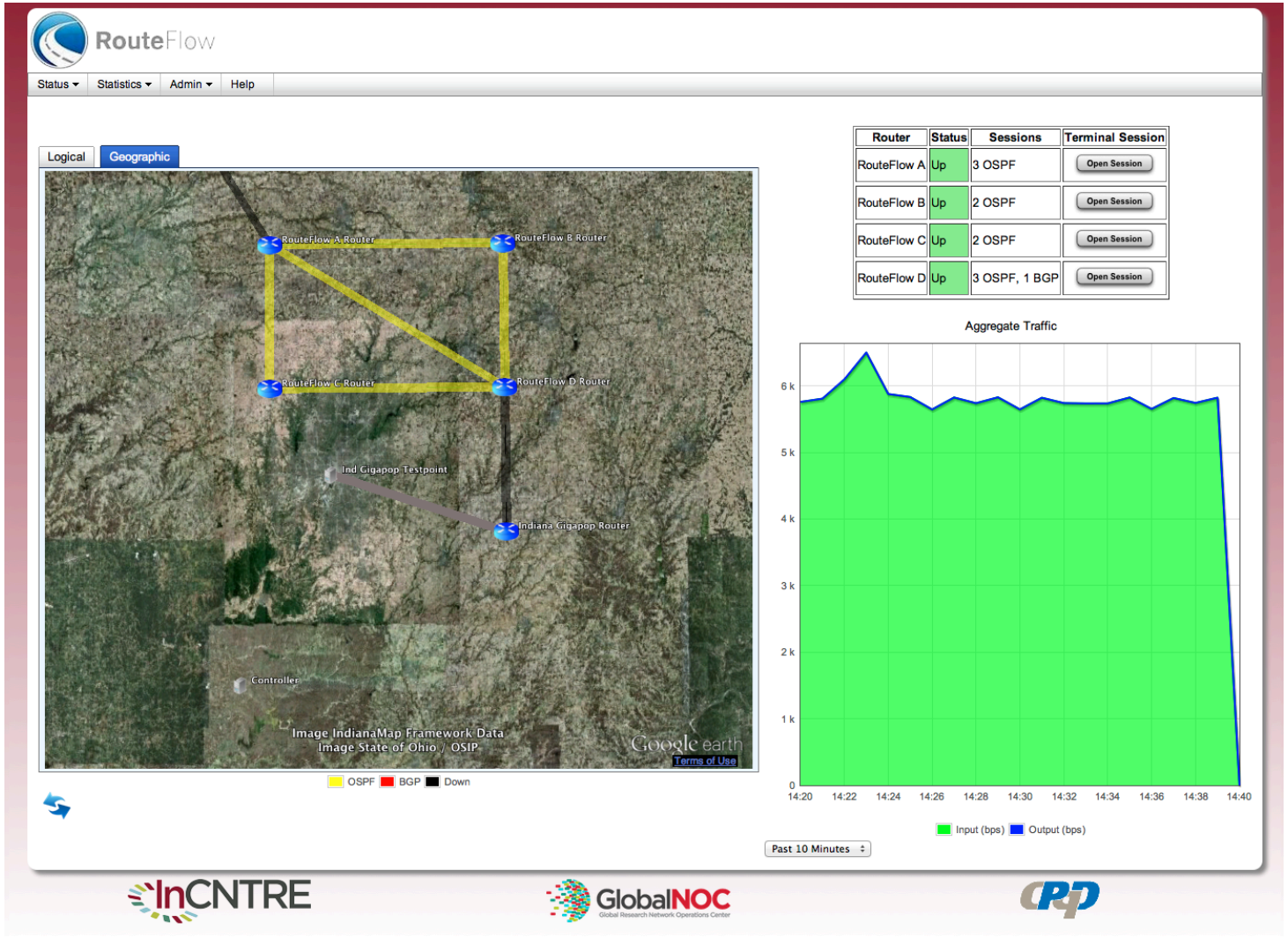
Field Trial at the University of Indiana

Network setup

- 1 physical OpenFlow switch
  - Pronto 3290
- 4 Virtual routers out of the physical OpenFlow switch
- 10 Gig and 1 Gig connections
- 2 BGP connections to external networks
  - Juniper routers in Chicago and Indianapolis
- Remote Controller
- New User Interface

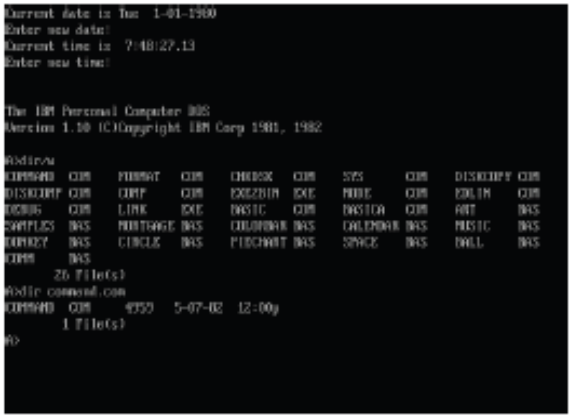


Field Trial at the University of Indiana
User Interface

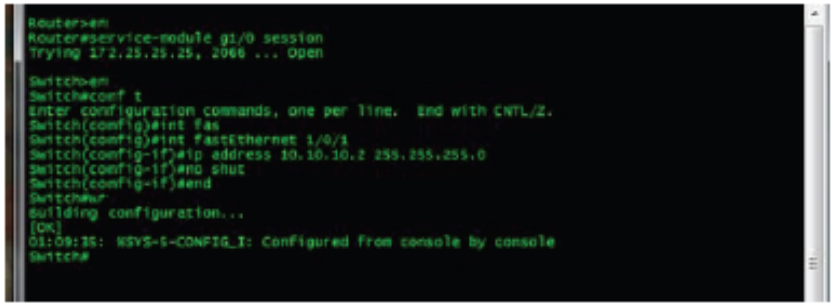


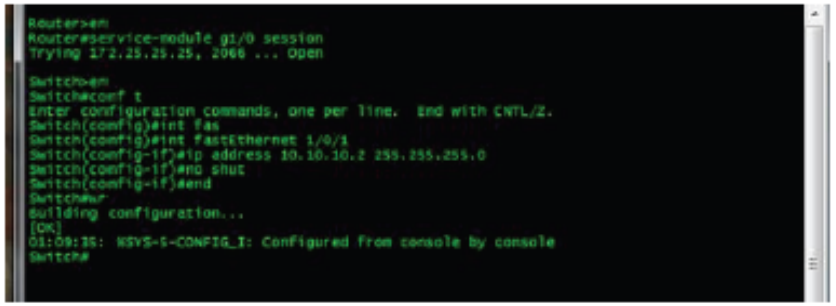
Compare interfaces over the last 30 years

"PC" user interfaces



Network user interfaces

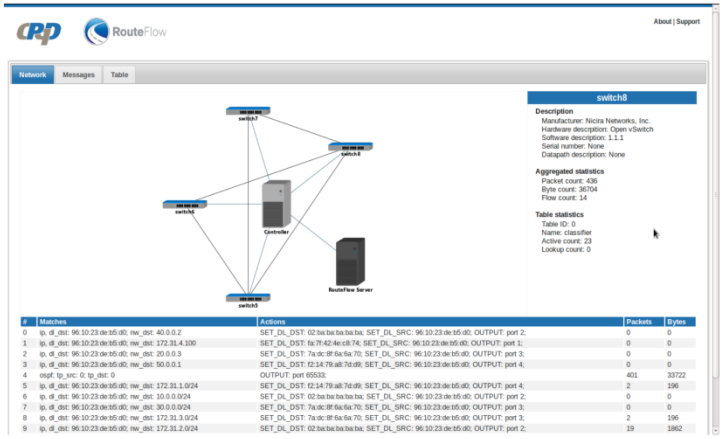




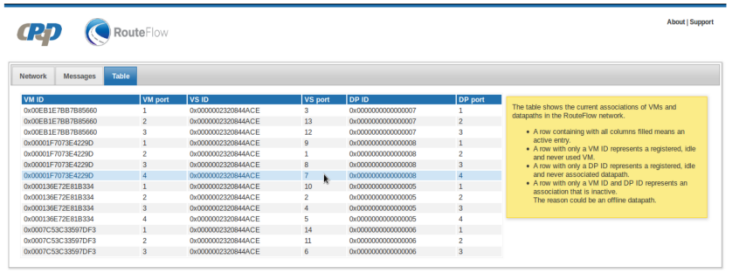
Source: Chris Small (Indiana)

Demystifying Configuation Challenges and Tradoffs in Network Based ISP Services (Benson, Akella,Shaikh SIGCOMM 2011)

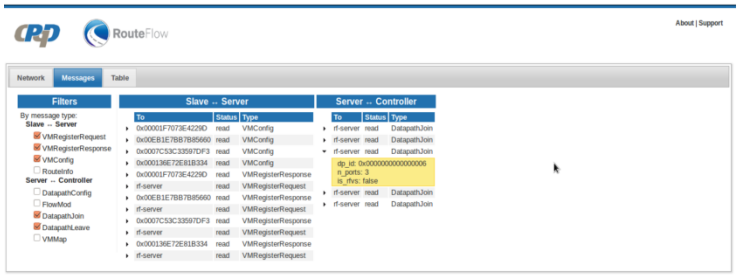
Advancing the RouteFlow GUI...



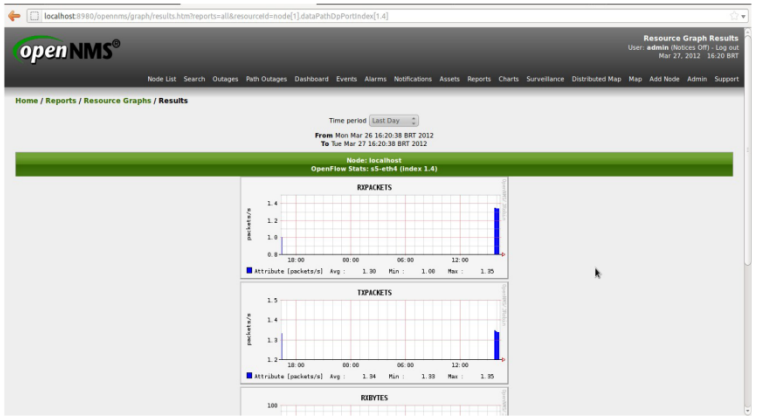
Topology and Statistics



Resource Status and Mapping



RouteFlow Protocol



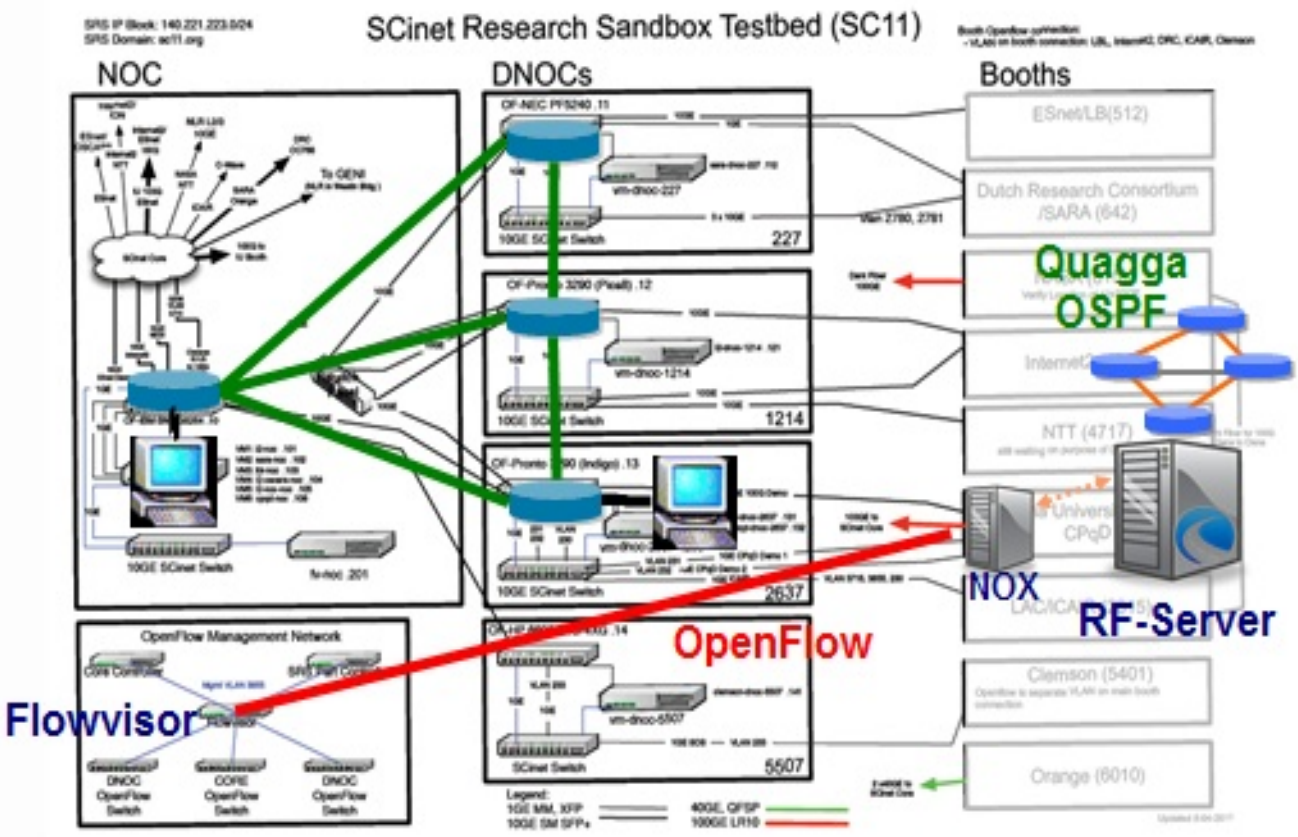
OpenNMS SNMP

Demonstration at Supercomputing 11

RouteFlow

SRS Demo @ SC11

SC11 Seattle, WA Connecting communities through HPC

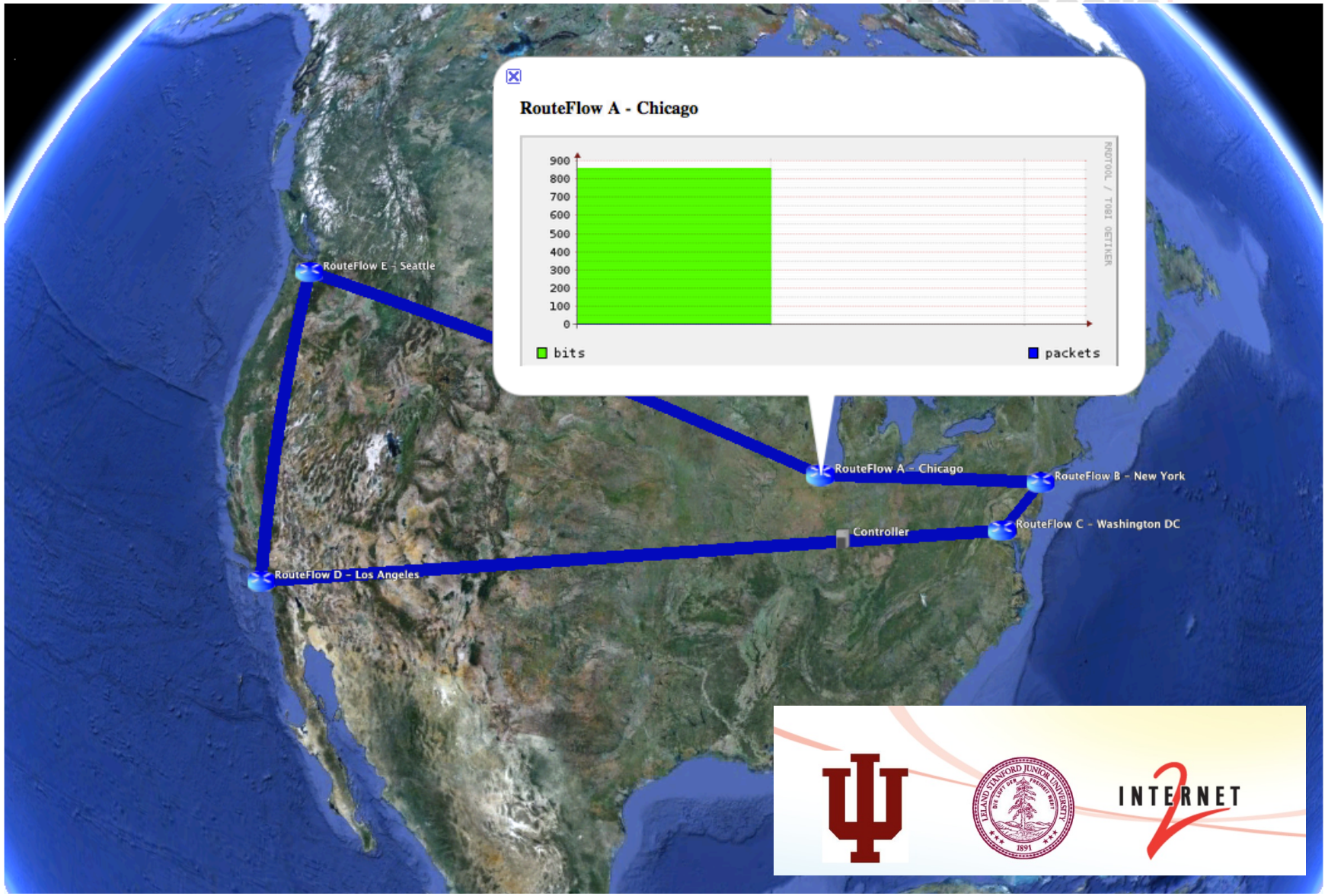


Routing configuration at your fingertips

II Open Networking Summit (Santa Clara)

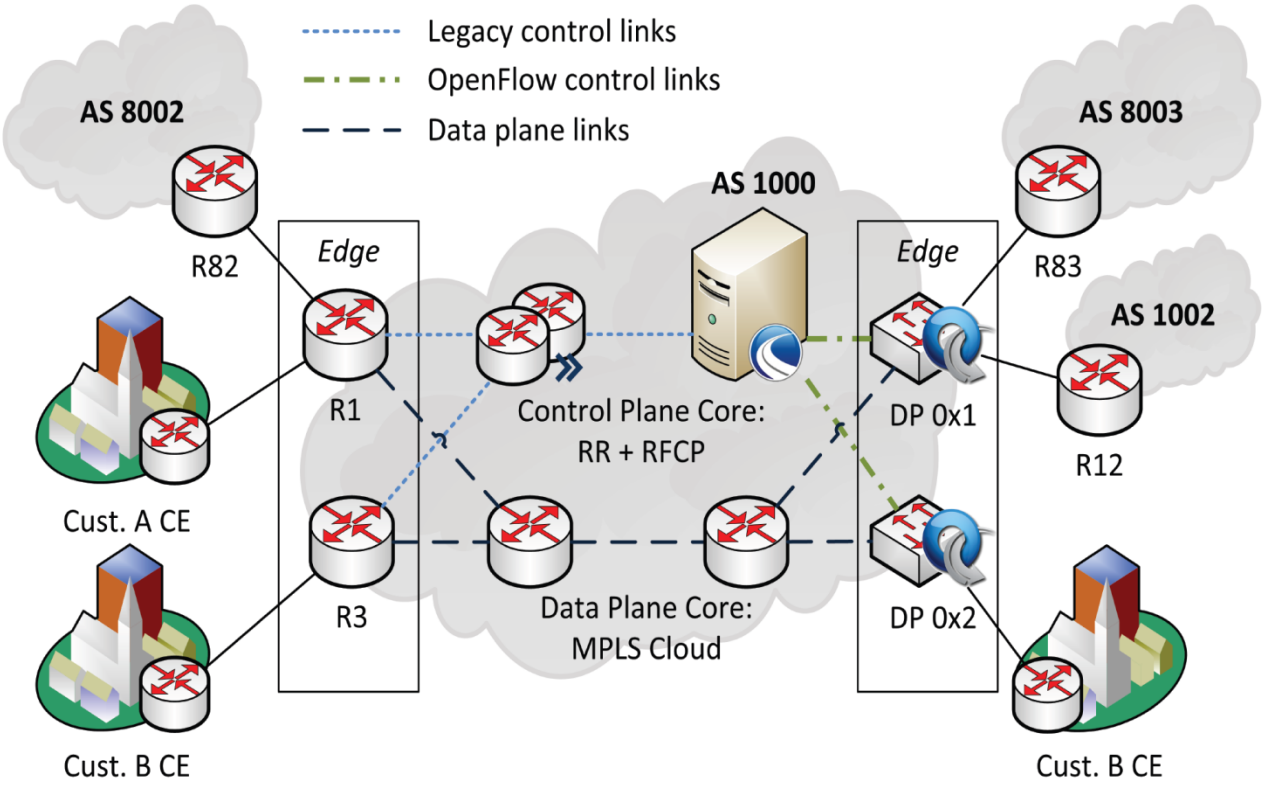
RouteFlow NDDI Deployment

[Joint Techs]



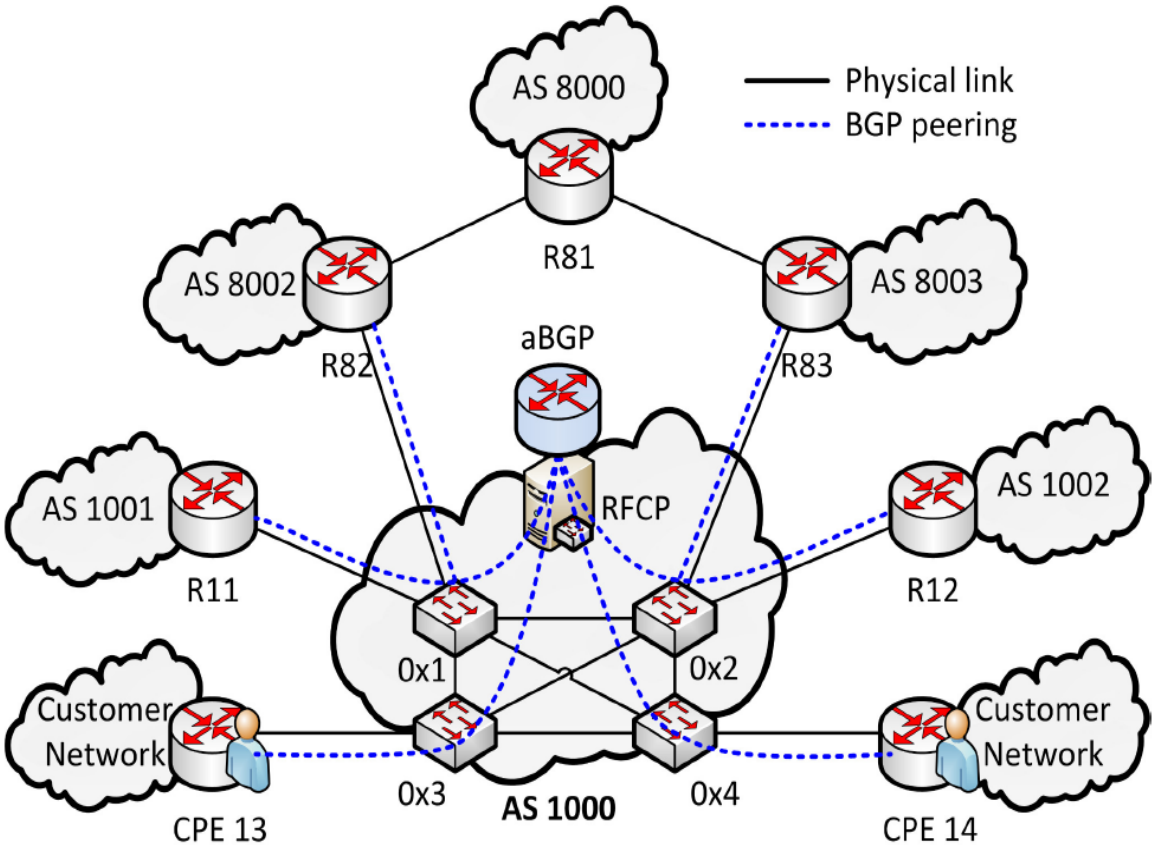
Controller-Centric Hybrid Networking

- A migration path to roll out OpenFlow technology
- Not a revolution, but an evolution of current iBGP RRs to essentially eBGP Route Controllers
  - "BGP-free edge": A cost-effective simplified edge for SW-driven innovations

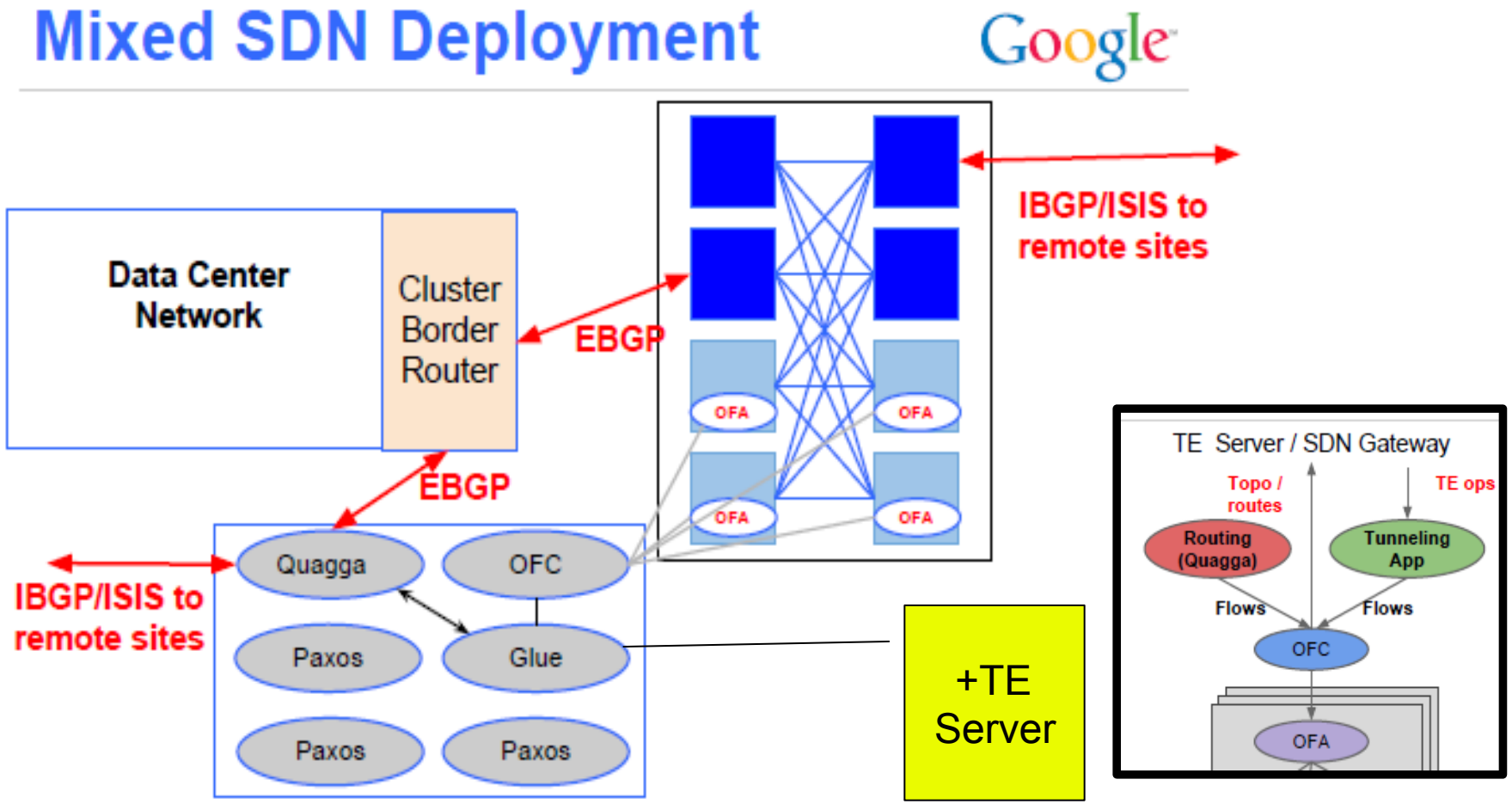


Aggregated BGP routing service → RCP

- Single node abstraction of a domain-wide eBGP router
  - Think modern multi-chasis routing architectures with external route processors and OpenFlow switches acting as line cards
- Aggregation logic defined in the RF-Server

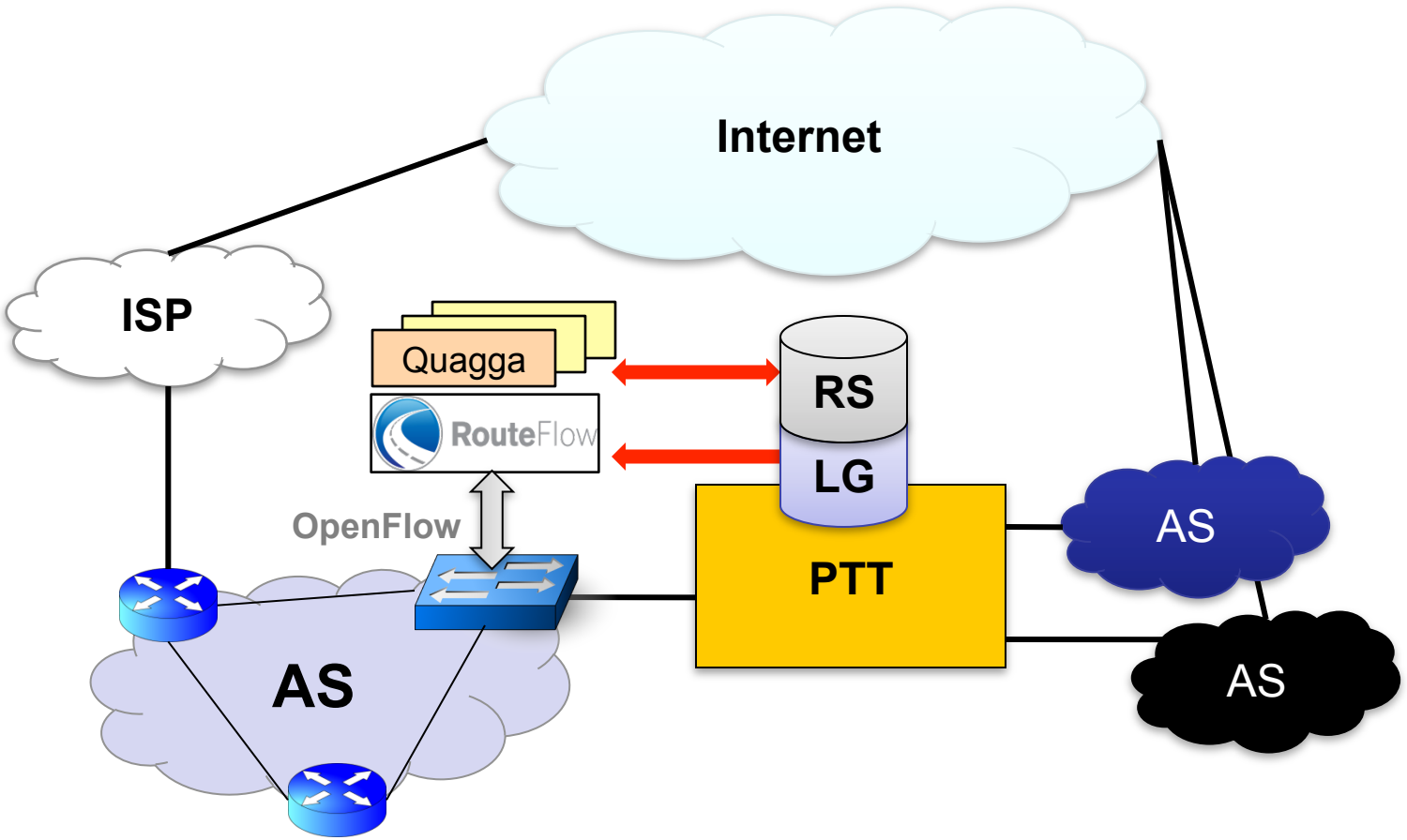


Google Software Defined WAN Architecture



Source: vahdat-wed-sdnstack.pdf @ONS12

RouteFlow em rede cliente do PTT

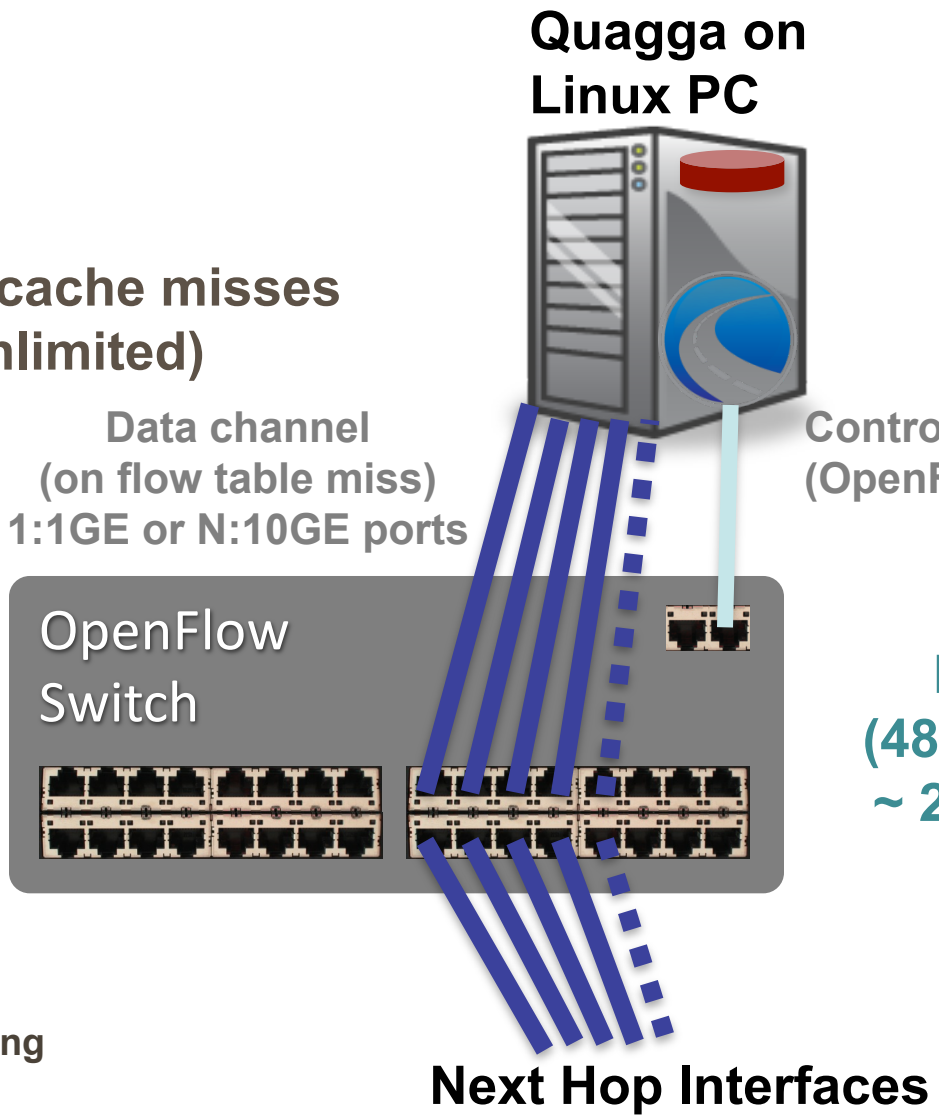


Quagga on Steroids

SW Forwarding
- Only for OpenFlow cache misses
- Full FIB (virtually unlimited)

HW Forwarding
- FIB cache*

(2 – 4 K entries)



* use Simple VA + smart caching

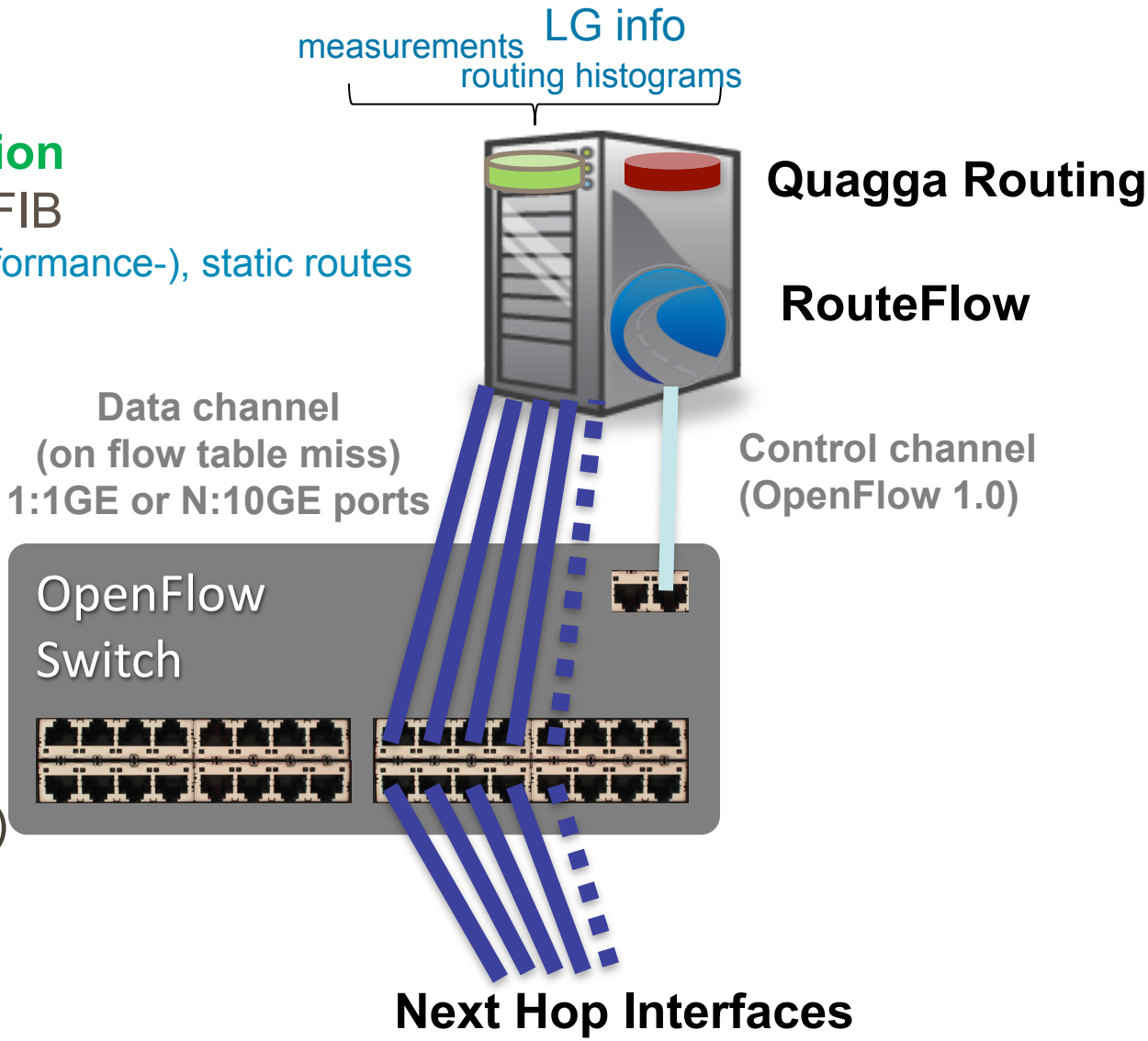
Quagga on Steroids + Customized Control

New control application

- Overwrites dynamic FIB
  - e.g., PBR (cost- / performance-), static routes
- Multi-homing/pathing
- Back-up routes
- Security rules

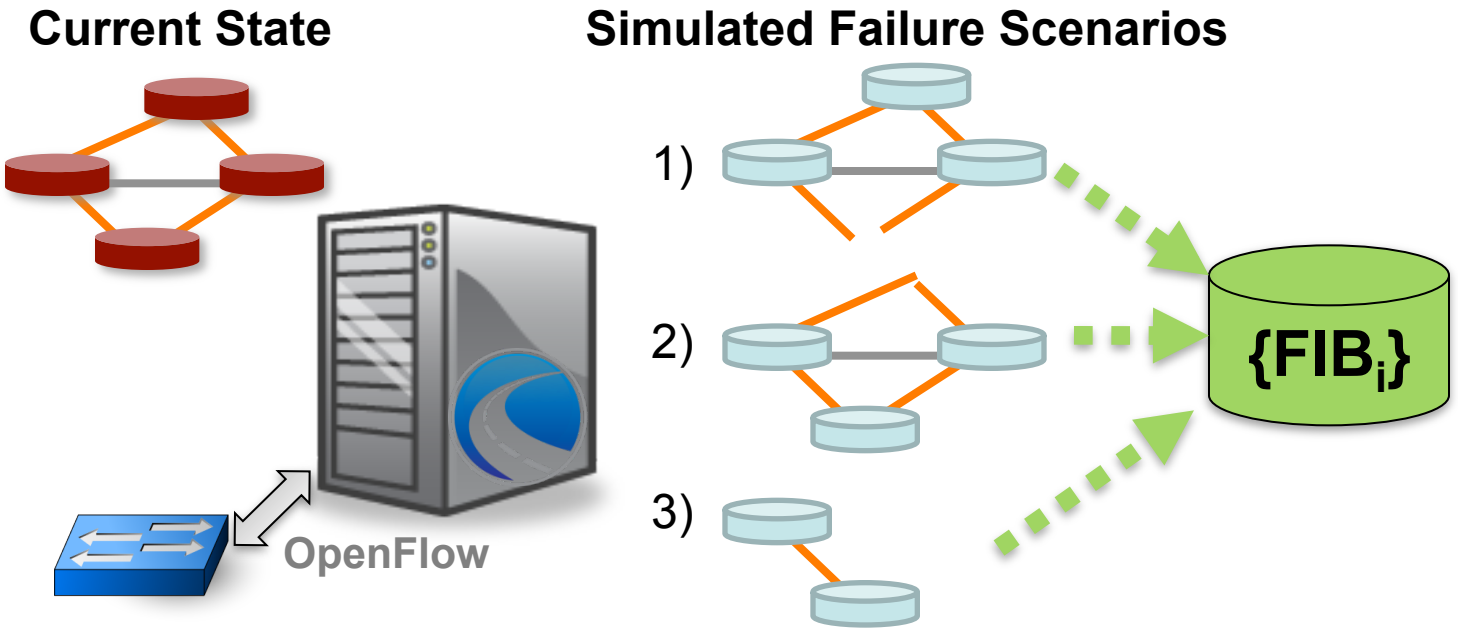
HW Forwarding

- FIB cache
- + new prioritized entries (e.g. ACL, PBR)



Fast convergence and availability

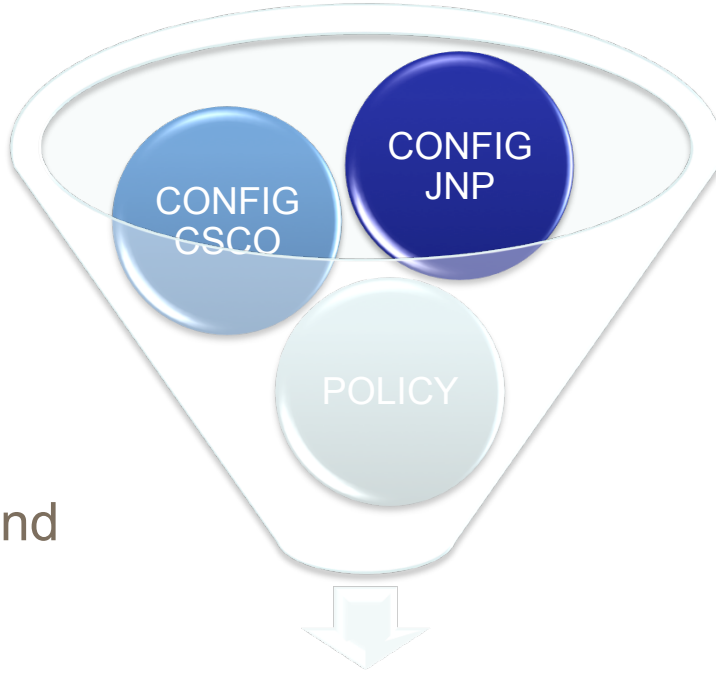
- Non-stop forwarding ready!
  - OpenFlow controller and IP control plane can be unavailable but FIB state is kept stale in the switches to forward packets
- Shadow control plane to test failure events and store next valid consistent states.
  - Upon failure, delta flow entries can be pushed down



Security

Towards a network-wide firewall

- Import existing device-level configurations
- Consolidate and review
- Import OpenFlow topology
- Transform into OpenFlow-rules
- Provide single point of configuration, roll-back, annotated histogram (who and why and for how long inserted a rule?), etc.
- Develop APIs for other systems, e.g., IDS



Network-wide OpenFlow rules

Secure Inter-domain routing (S*-BGP, etc..)

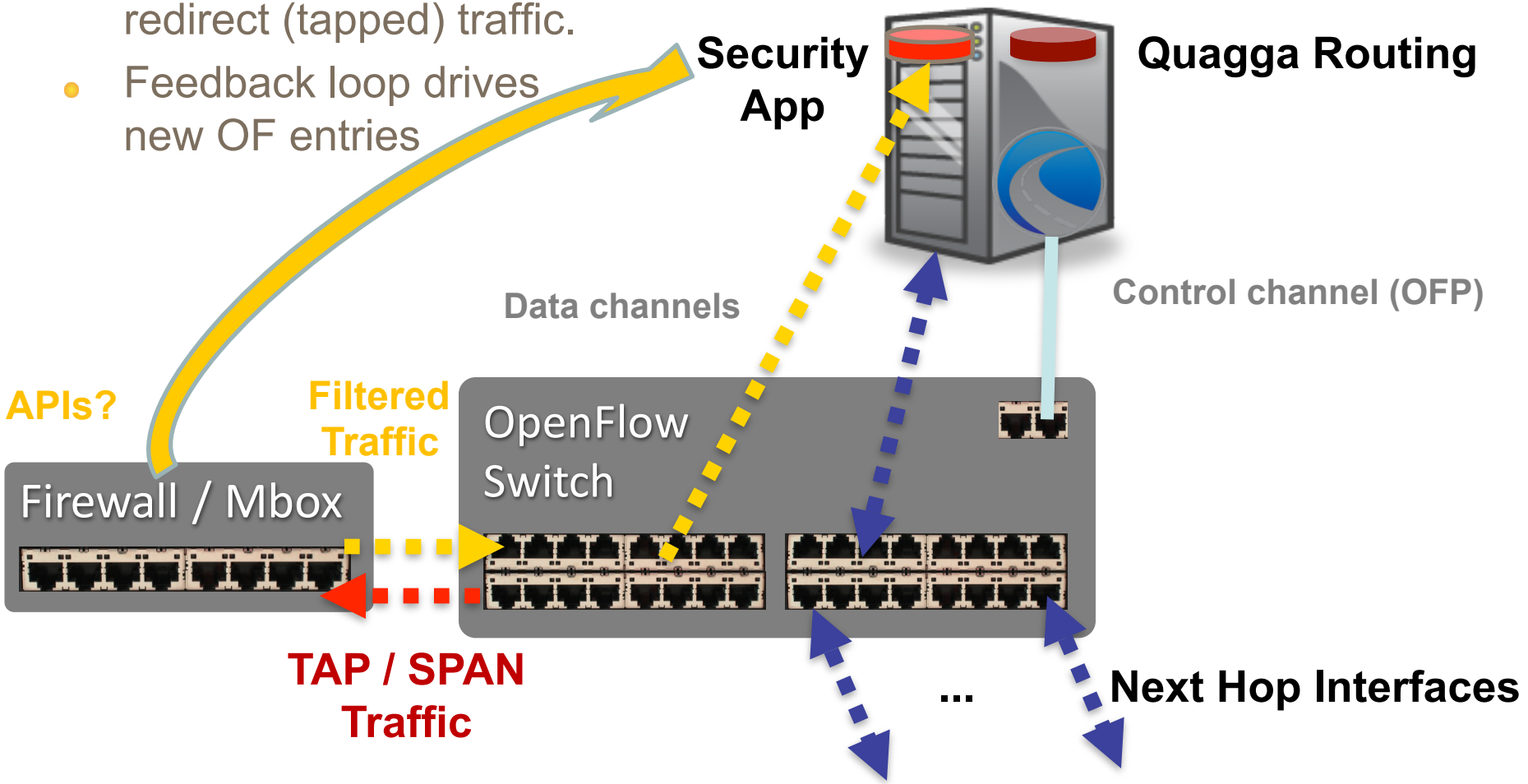
Data plane blackholes

Middlebox injections

Security

Hybrid firewalling

- Re-use existing firewalls / middleboxes (IDS, Net/SDlow) and redirect (tapped) traffic.
- Feedback loop drives new OF entries



Routing-centric use cases under research

- Engineered path selection
  - Think Google WAN, performance-based routing, etc.
- Optimal best path reflection
  - Per ingress/customer [draft-ietf-idr-bgp-optimal-route-reflection-01]
- Path protection with prefix independent convergence
  - Hierarchical FIBs w/ OF 1.X Tables + LFA route-precomputation
- Security
  - Data plane blackholes and middlebox injections,
  - Secure Inter-domain routing ideas (crypto intense S*-BGP, etc..)
- Simplifying customer multi-homing
  - Easy to set and control cost/performance/policy-based routing
- IPv6 migration
  - Flow matching for service termination in v4-v6 migration solutions

Challenges

- Centralized BGP
  - Shown to scale well in modern CPU architectures
  - Centralized does not mean not distributed (but removal from edge)
- Small OpenFlow table sizes
  - Transient limitation?
  - Expose existing FIB data structures as an IP lookup OF table?
  - Smart RIB&FIB reduction (e.g., simple [draft-ietf-grow-simple-va-04]
  - HW/SW flow offloading
- Limited OpenFlow processing in datapath
  - Transient / Un-optimized implementations
- High availability
  - Previous ideas from disitributed RCPs
  - Database-centric designs
  - Development in-progress of “BGP SHIM” for transparent eBGP redundancy

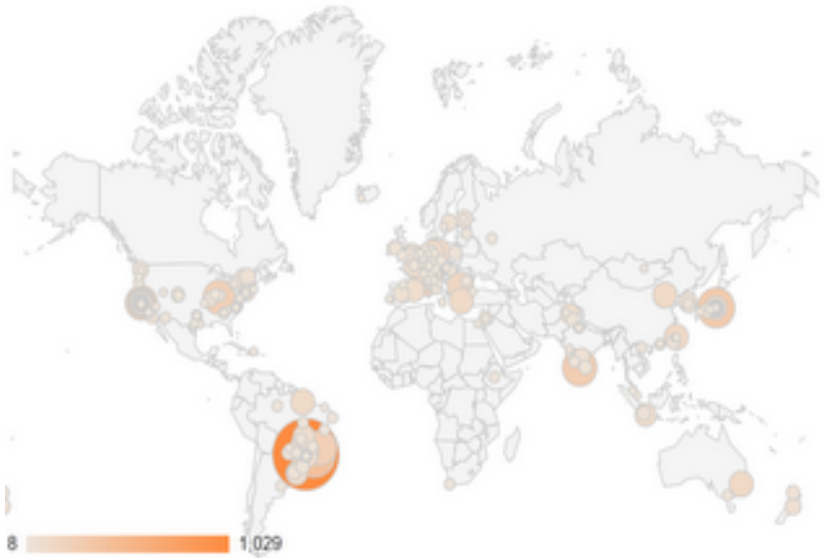
RouteFlow Project History

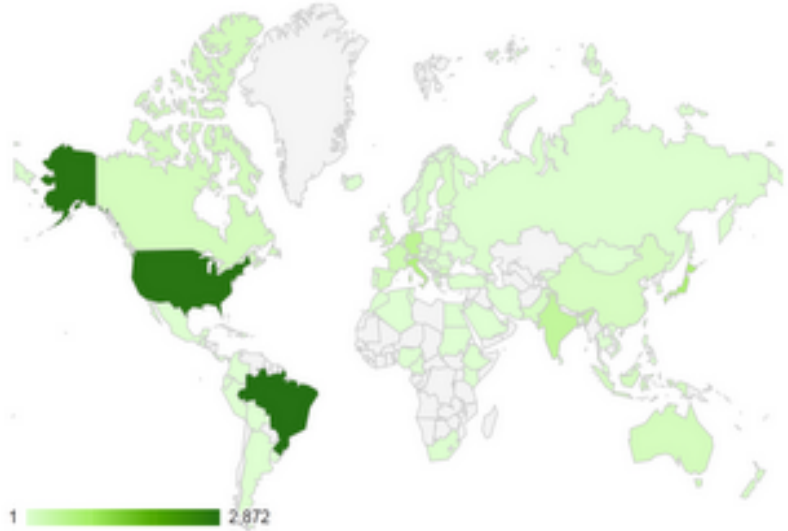
- Start Msc. Thesis work by Marcelo N.
- First Prototype
- First Short-Paper @ WPEIF
- Evalaluation on NetFPGA testbed
- QuagFlow Poster @ SIGCOMM
- Open-Source Release
- Demos @ ONS11
- Tutorial & Demo @ OFELIA/CHANGE SS
- Indiana University - Pronto OF switches + BGP peering with Juniper MX
- Demo @ SuperComputing 11
- Demos @ ONS12
- HotSDN Paper
- Running on FIBRE / OFELIA testbed
- Collaboraion with NTT

… building a community

Visits: 24,000+ (11,000+ Unique)

From over 2000 cities of 110+ countries all over the globe!





http://go.cpqd.com.br/routeflow/

RouteFlow

582

days since
Project Launch

GIGA
REDE EXPERIMENTAL DE ALTA VELOCIDADE

FUNTTEL FINEP

1000s downloads!

github SOCIAL CODING

Apache

Collaborations and community developments

- Web-based UI & Internet 2 HW pilot [C. Small, Indiana] ✓
- Aggregated BGP Routing Service [C. Corrêa, Unirio] ✓
- SNMP plugin [J. Stringer, Google] ✓

- Optimal BGP best path reflection [R. Raszuk, NTT-MCL] ◷
- Open Label Switched Router [OSRF; Google] ◶
- OpenFlow v1.2 and v1.3 [w/ Ericsson] ◵

Evolving the IP routing landscape with OpenFlow/SDN

Comentário finais

RouteFlow como “*Quagga on steroids*”

- Open-source routing SW + line-rate forwarding on HW

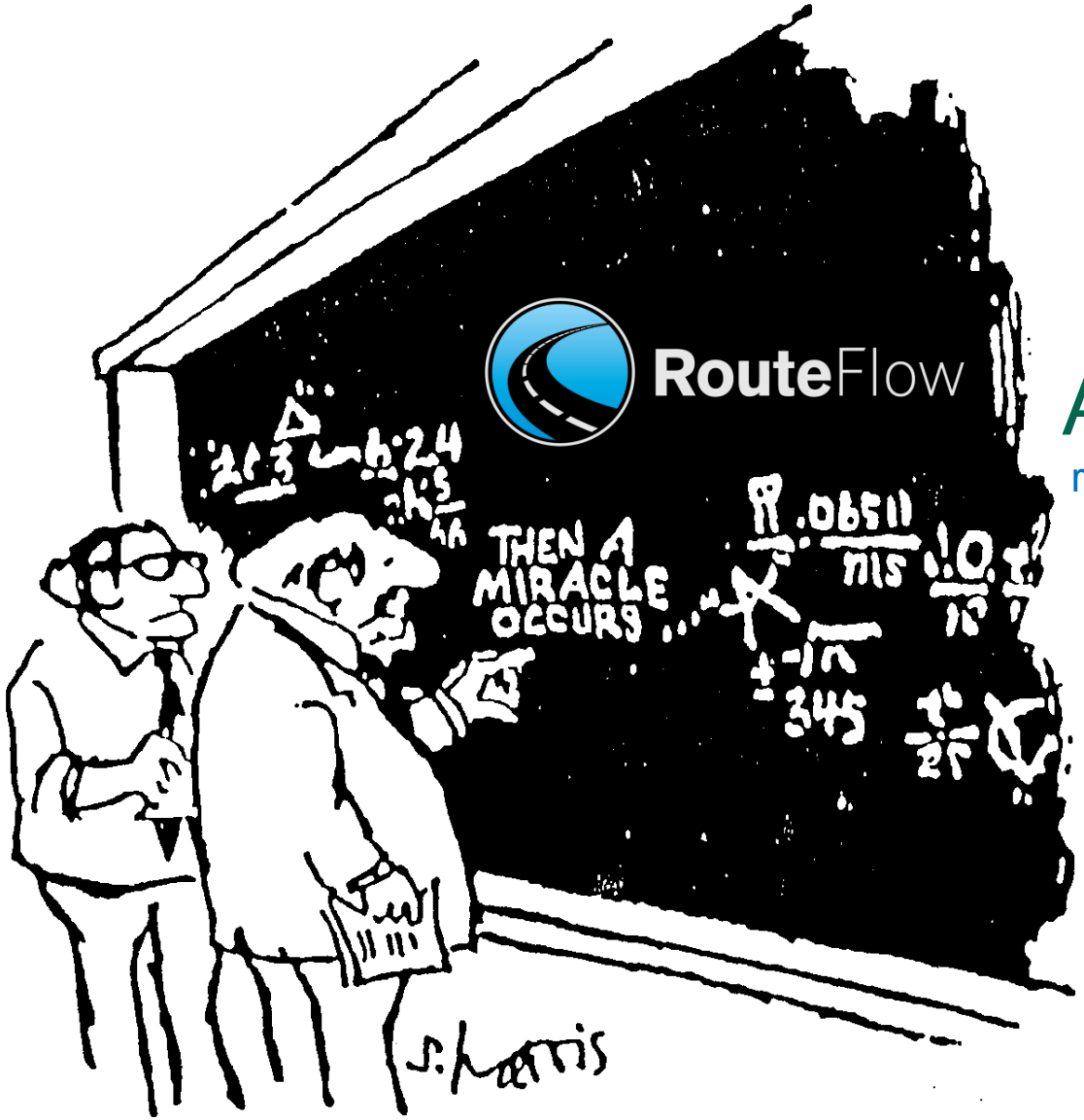
RouteFlow como exemplo e oportunidade para inovar no roteamento IP

- Parceria com usuários “beta” nacionais
  - com requisitos e dores de cabeça reais,
  - com orçamentos restritos,
  - sem limites na vontade de fazer acontecer
- Ex: Policy-based routing (cost-, performance), convergência rapida, segurança, migração para IPv6, etc.

OpenFlow/SDN como tecnologia nos PTT

- Automation, self-service provisioning, validation, security, etc.

Questions?



Thank you!

Ask and contribute!

routeflow-discuss@googlegroups.com

Learn more!

https://go.cpqd.com.br/routeflow

Get the Code!

https://github.com/CPqD/RouteFlow

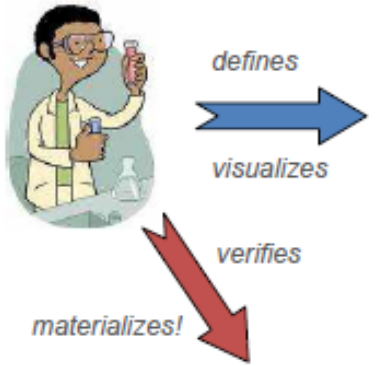
BACKUP

fibre

FUTURE INTERNET TESTBEDS EXPERIMENTATION BETWEEN BRAZIL AND EUROPE

CPqD

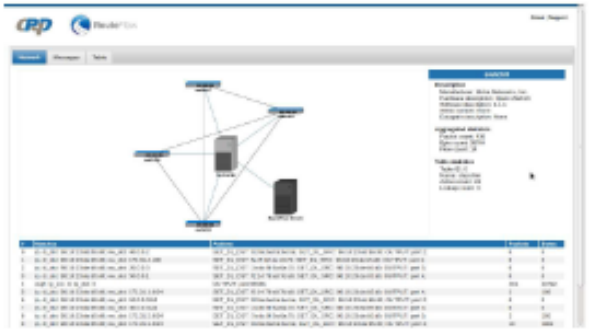
IP Network as a Service (NaaS)

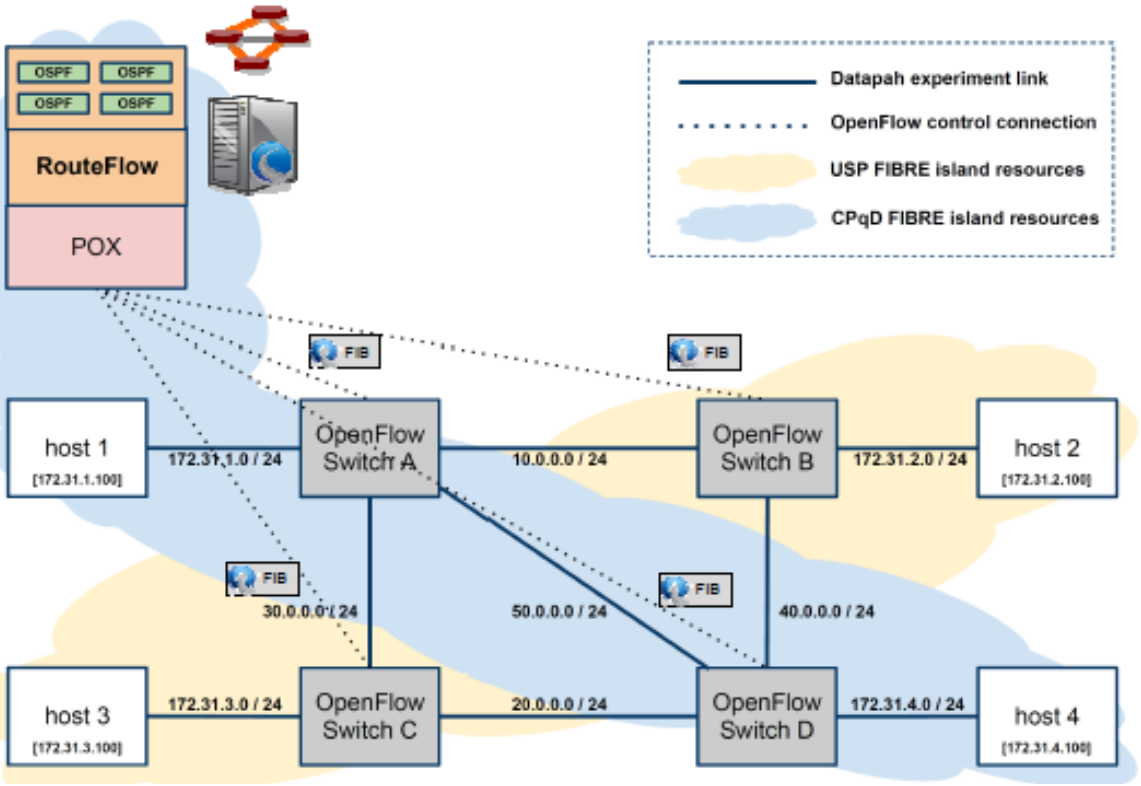


FIBRE Aggregate

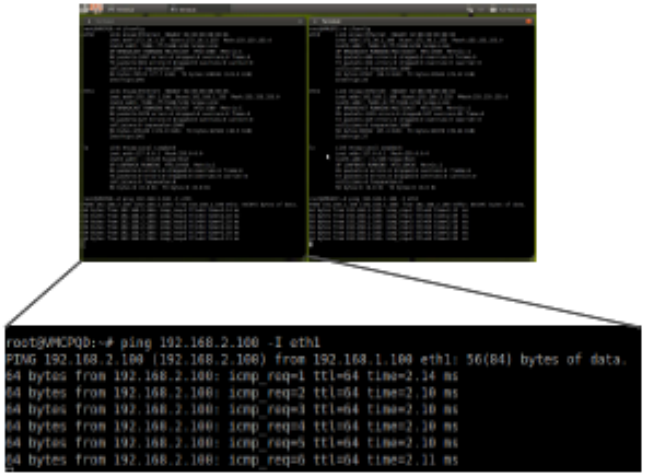
FIBRE Slice

RouteFlow GUI





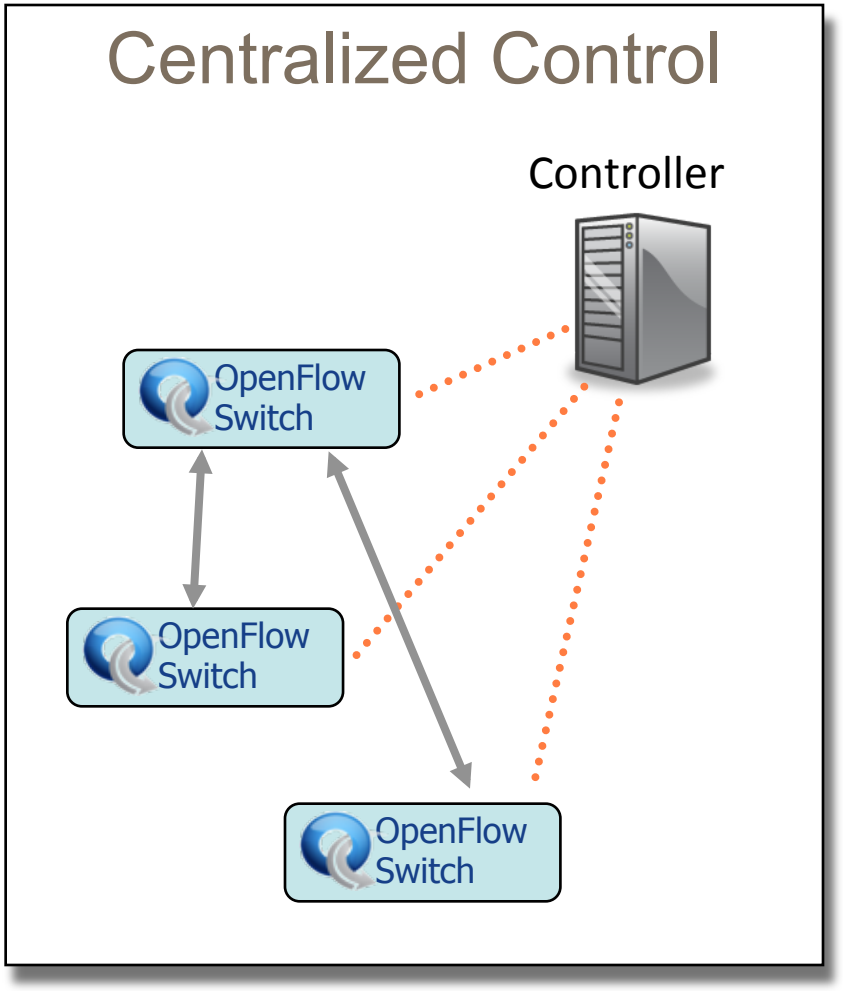
VM end-to-end PING

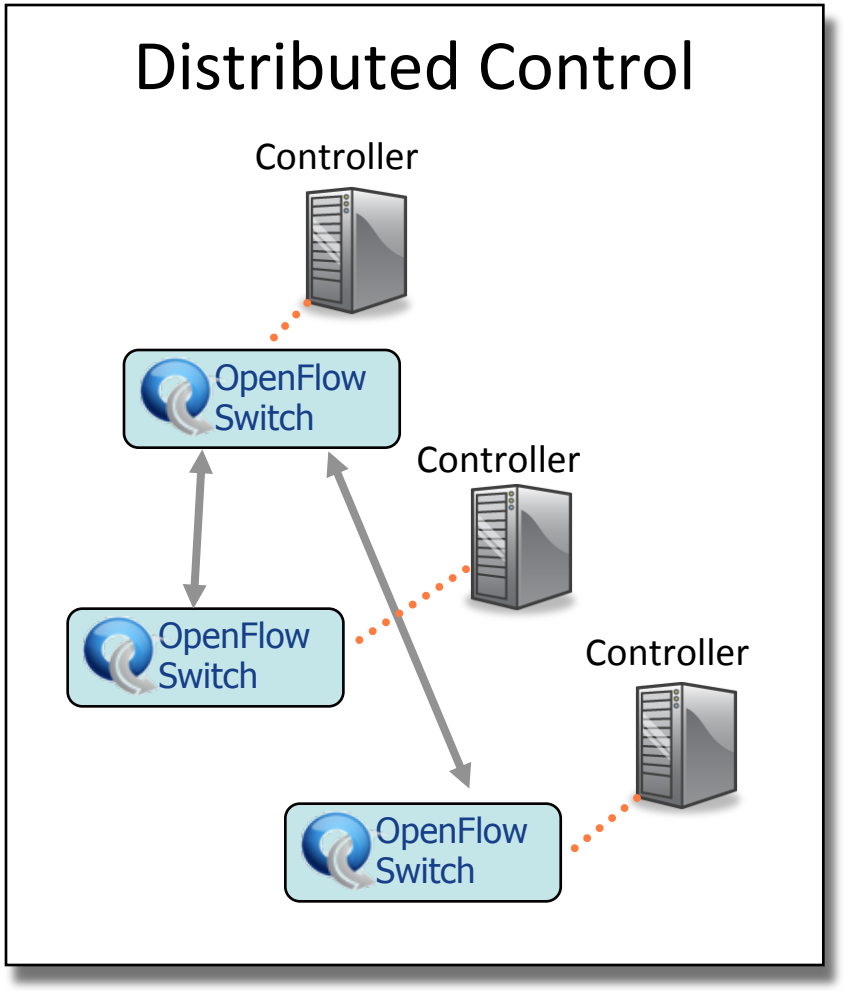


RouteFlow

Centralized vs Distributed Control

Both models are possible with OpenFlow





Flow Routing vs. Aggregation

Both models are possible with OpenFlow

Flow-Based

- Every flow is individually set up by controller
- Exact-match flow entries
- Flow table contains one entry per flow
- Good for fine grain control, e.g. campus networks

Aggregated

- One flow entry covers large groups of flows
- Wildcard flow entries
- Flow table contains one entry per category of flows
- Good for large number of flows, e.g. backbone

Reactive vs. Proactive (pre-populated)

Both models are possible with OpenFlow

Reactive

- First packet of flow triggers controller to insert flow entries
- Efficient use of flow table
- Every flow incurs small additional flow setup time
- If control connection lost, switch has limited utility

Proactive

- Controller pre-populates flow table in switch
- Zero additional flow setup time
- Loss of control connection does not disrupt traffic
- Essentially requires aggregated (wildcard) rules