Novas arquiteturas de redes: Um futuro definido por software

Seminário de Pesquisa , UFABC

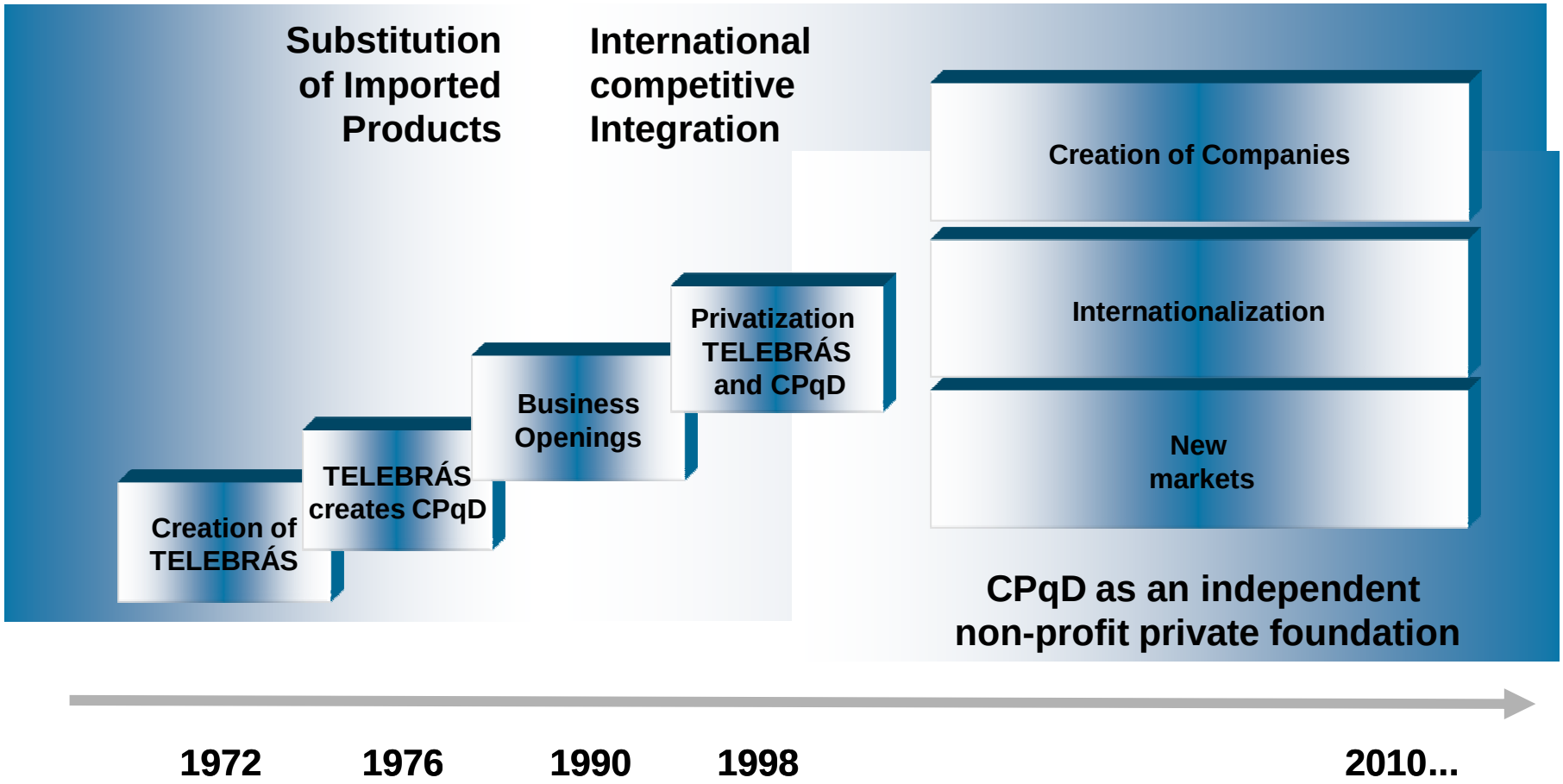
Março 2013

Christian Esteve Rothenberg

About CPqD

- Major telecom R&D center in LATAM with expertise in various areas:
  - Optical (WDM, PON), Wireless (WiMax, LTE), IP (IMS/NGN, OpenFlow), OSS/BSS, Digital TV
  - Today with ~1200 highly-skilled employees
- Created in 1976 as R&D branch of Telebras - Brazilian telecom monopoly
- Private foundation since 1998 after Telebras was privatized
- Purpose 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

Company History and Evolution



TELEBRÁS – Brazil telephony system monopoly

(Typical) Model of Operation

Government funds

Own funds

Government funds
Own funds

Licensing of technology to partners
Direct provider of services in the market

** Also work as independent contractor for national and international companies*

CPqD – Network division

- Various product technologies transferred to the Brazilian industry (since privatization)
  - 18 patents filled in 2011
  - 24 patents filled in 2012 (until Sep/12)

- Most successful spin-offs (and consumers of CPqD technology)
  - Tropico: created in 1999 with focus on NGN/IMS; US$ 120M revenue in 2008, with growing presence in South America
  - Padtec: created in 2001 with focus on WDM; US$ 150M revenue expected in 2011; WDM market leader in Brazil, with growing presence in South America and Europe

- Universo CPqD
  - formado por organizações com características distintas (Atlântico, Já!, ClearTech, SPAT, Civcom, CPqD USA, Trópico, Padtec, ZELOX, WxBr) criadas a partir da iniciativa direta ou indireta do CPqD
  - R$ 740 milhões em 2011

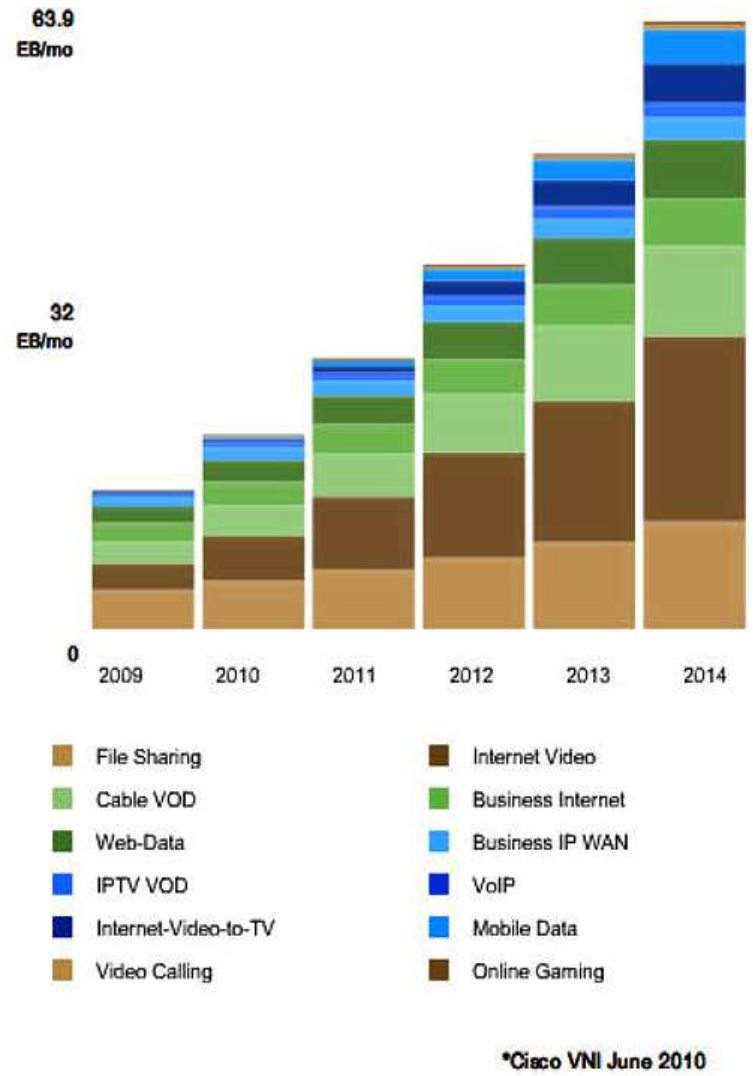
Agenda

- Tendências em Redes / TICs
- Networking for the Cloud
  - Inside the data center
  - Between data centers
  - Evolving Internet exchange landscape
- Software-Defined Networking
  - What? Why? How?
- Projetos de P&D em OpenFlow/SDN no CPqD
  - RouteFlow

Tendências em Redes (TICs): Drivers

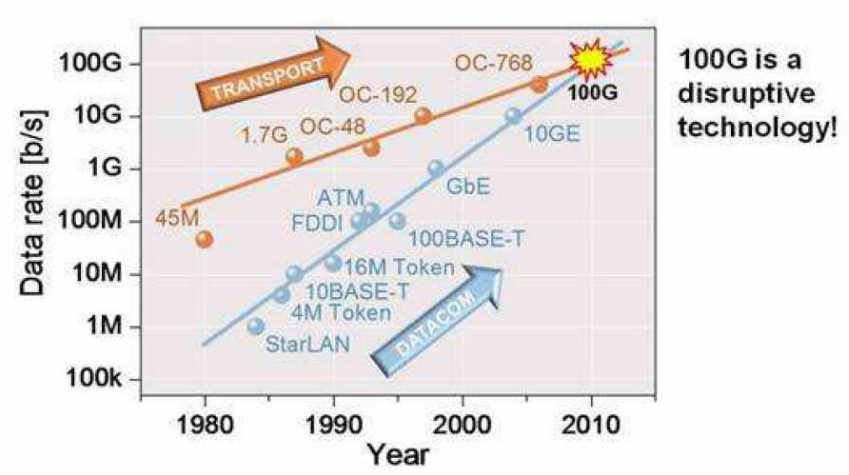
- **Drivers de Usuários:**
  - Crescimento do Tráfego
  - Aumento do # Terminais / Usuário
  - Qualidade de Experiência (QoE)
  - Mobilidade
  - Identidade digital

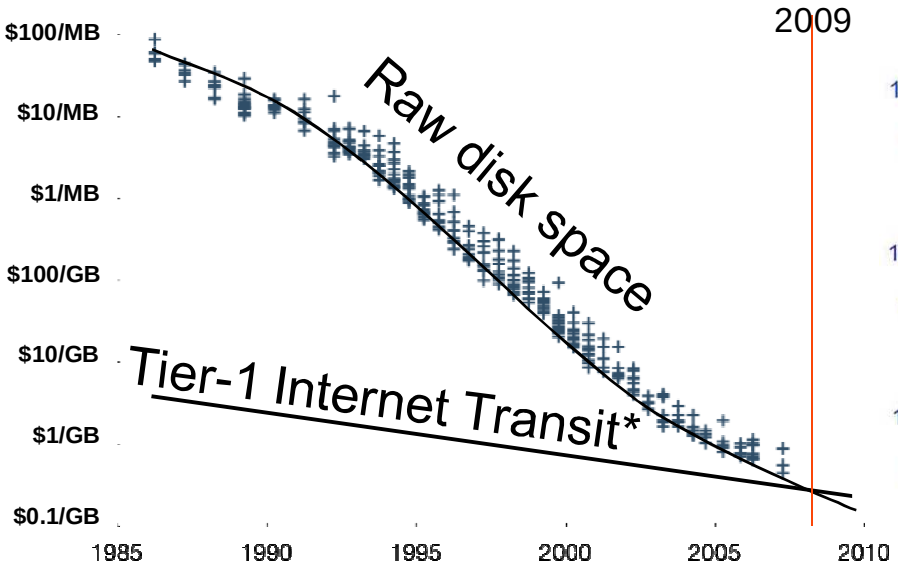
- **Drivers Operacionais (Service Provider)**
  - Redução de Custo
    (TCO = CAPEX + OPEX)
  - Segurança, QoS
  - Convergência
    - IP/MPLS/Ethernet/**OTN**/WDM, FMC, Virtualização,
      Computing + Storage + Network
  - Demandas de Novos Serviços
    - P.ex. Clouds, CDNs, VPN+,
  - Novos modelos de negócios (XaaS)

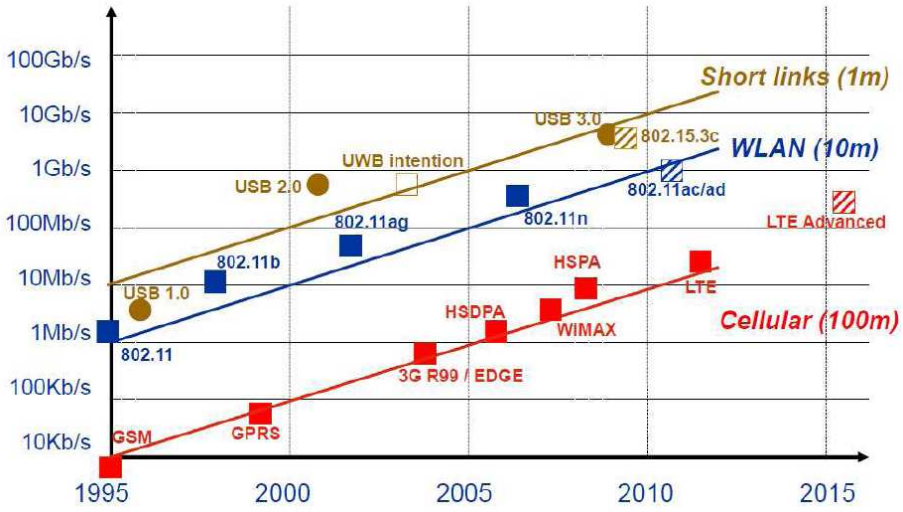


Tendências em Redes (TICs): Drivers

- **Drivers Tecnológicos**
  - Altas taxas de dados (óptica, sem fio)
  - Infraestruturas energeticamente eficientes (“*green networks*”)
  - Esgotamento IPv4: Migração IPv6
  - Comoditização de HW e modularização da industria
  - *Open-Source meets networking*







[1] Preliminary data [Nikander’09]

Tendências?

IPv6
(agora sim...)

Banda larga

Convergência

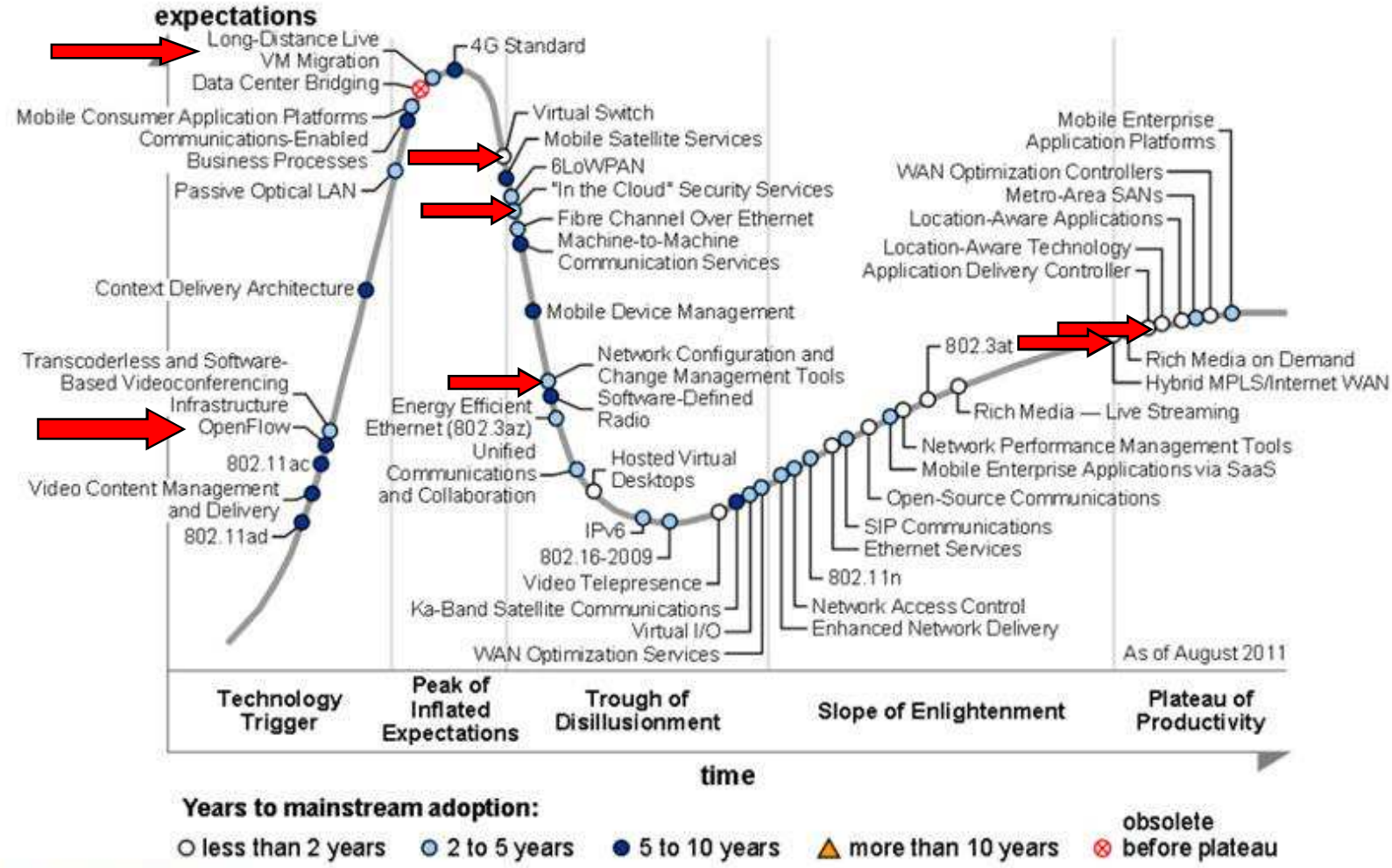
Mobilidade

Cloud

Virtualização

Tendências? Ask Gartner...

Figure 1. Hype Cycle for Networking and Communications, 2011

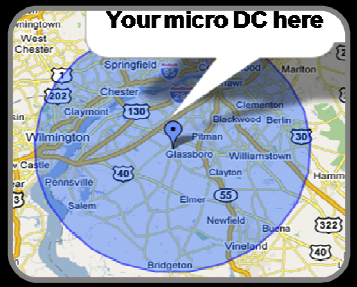


Source: Gartner (August 2011)

Types of cloud service data centers

Macro Data Center

- Specially dedicated facilities
- 100.000 or more servers and 10s of Mega-Watts of power at peak
- Computation in the cloud
  (e.g., Amazon EC2, Windows Azure, Google AppEngine)



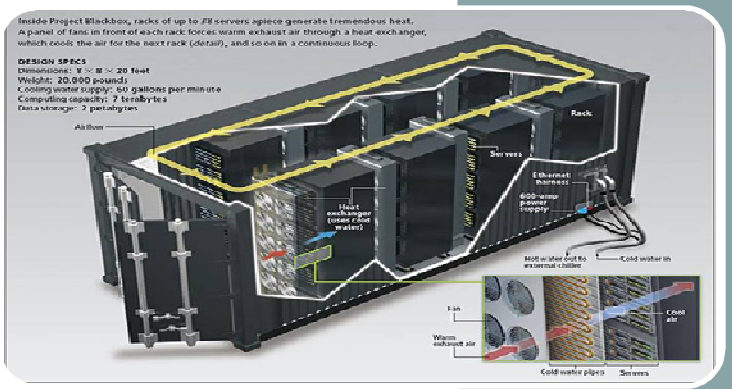
Micro Data Center

- Geo-diverse placed close to major population centers (e.g. CDN)
- 1000s of servers and 100s of kilowatts
- Higher degree of independence between physical DC outages
- Opportunity to economically reach data center customers with low latency (e.g., front-end cloud apps)

Nano Data Center

- Located in the customer premises equipment (e.g., set-top-box)
- "Why don't we try to take the functionality in the data center, and distribute it across hundreds of thousands of set top boxes so that we have these 'Nano Data Centers" [EU FP7 NADA]
- P2P-like resource management. Low latency. Low cost.

Data center in a box



Container-based modular DC

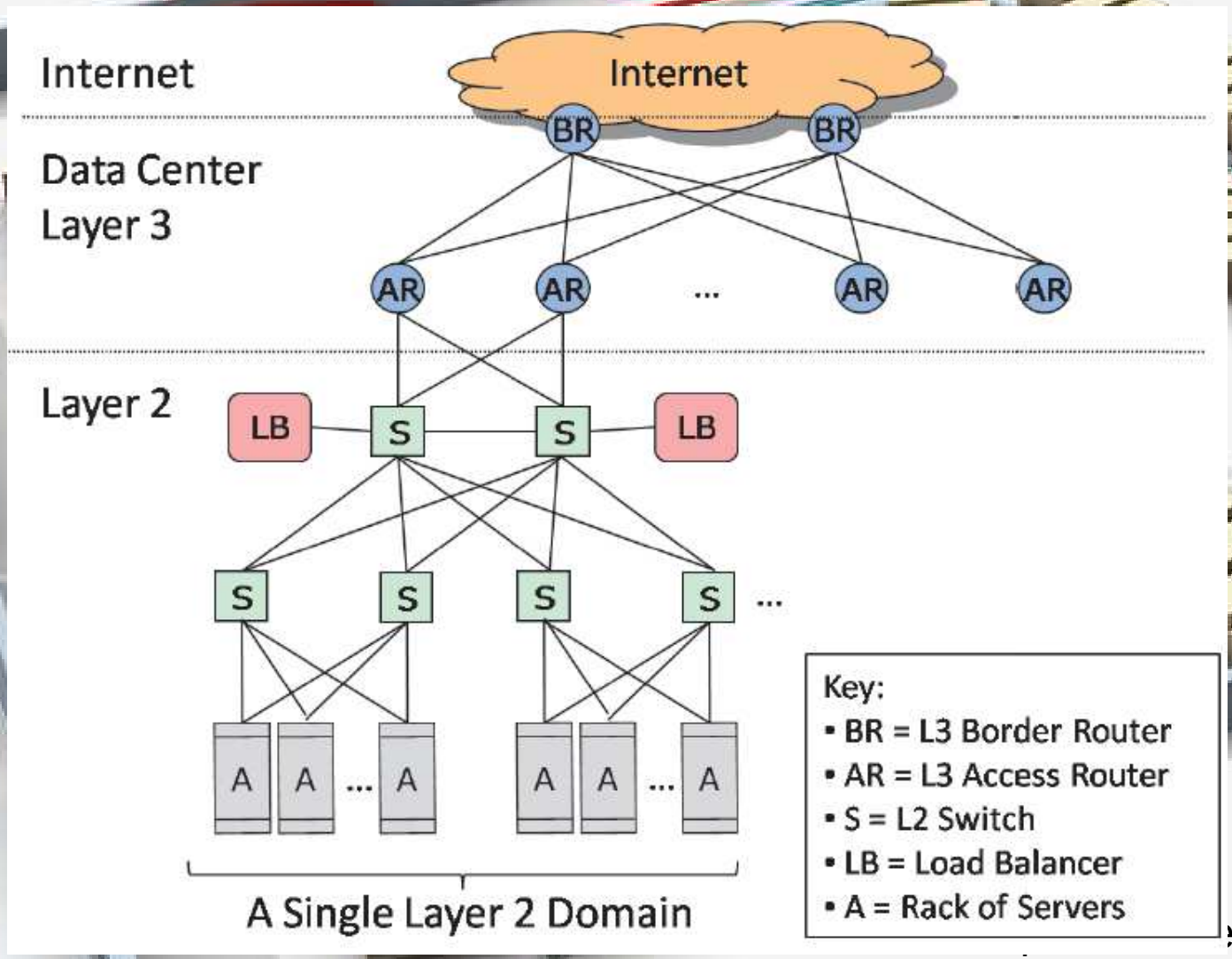
- Efficient way to deliver computing and storage services
- 1000-2000 servers in a single container
- Sun Project Black Box (242 systems in 20’)

Rackable Systems Container
2800 servers in 40’

Core benefits:

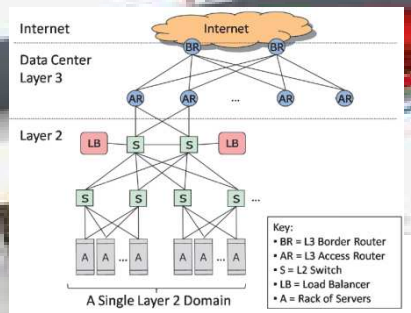
- Easy deployment
  - High mobility
  - Just plug in power, network, & chilled water
- Increased cooling efficiency
- Manufacturing & H/W Admin. Savings
- Push modularity throughout the DC

Current DC network architectures



eb. ´10

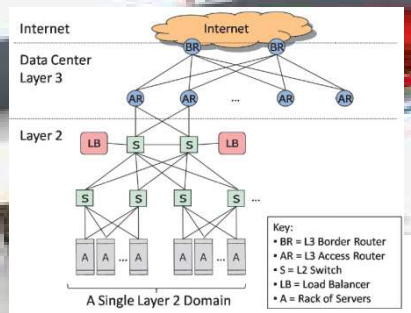
Some issues with conventional DCN designs



Networking constraints of traditional L2/L3 hierarchical organization:

- Fragmentation of resources VLANs / subnets ("any VM to any server")
- Limited server-to-server capacity
- Ethernet scalability
- Low performance under cloud application traffic patterns (e.g., "east-west" traffic)

New designs: DC Fabrics



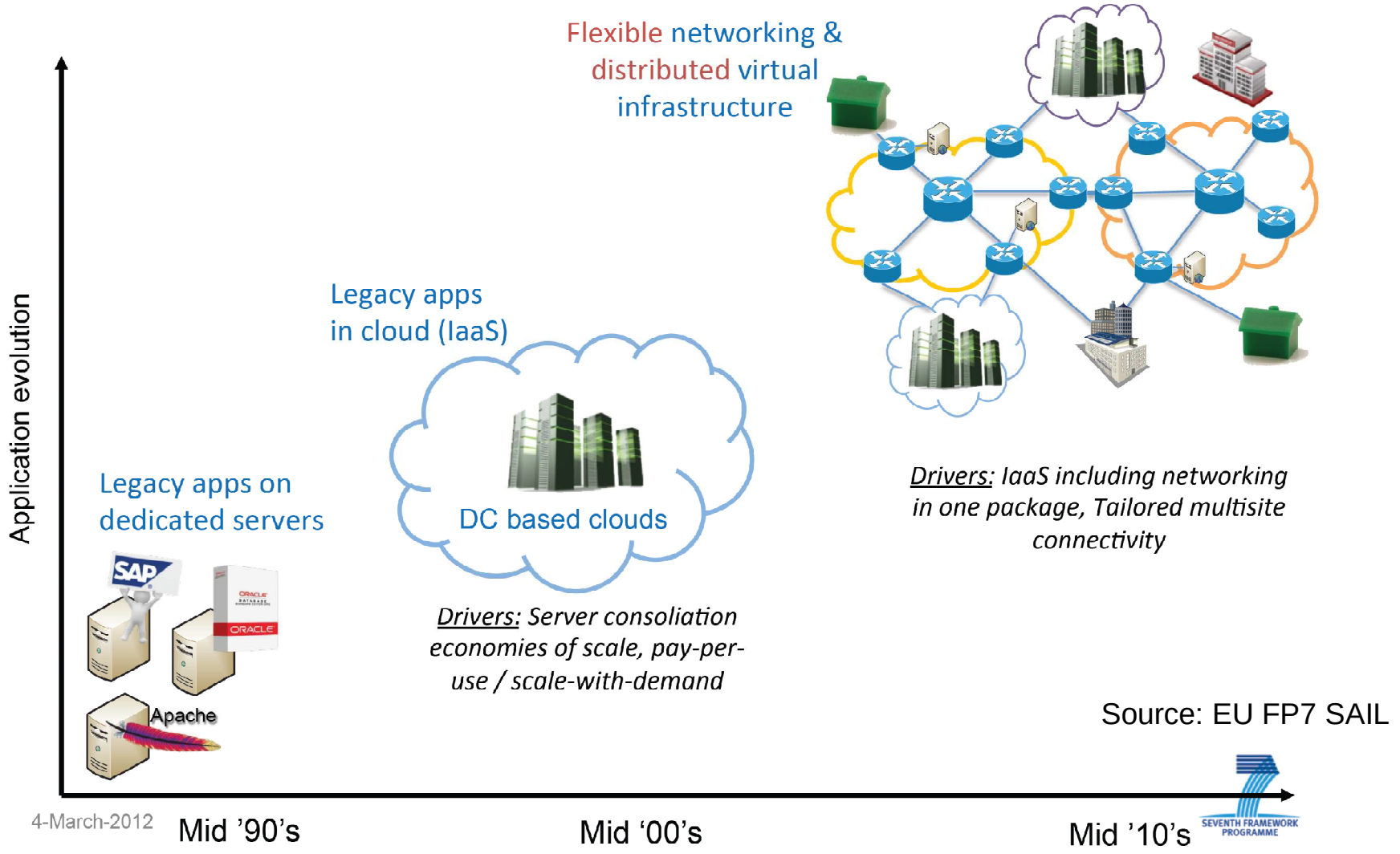
Active R&D space:

- New Ethernet protocols, e.g., TRILL
- Fabric Computing
- Virtualization 3.0 (Computer + Storage + Network)
- Hypervisor vSwitch is the first networking hop
- Virtual L2 networks (Overlays: L2-over-IP, tunnels)
- OpenFlow & Software Defined Networking
- …

Agenda

- Networking for the Cloud
  - Inside the data center
  - Between data centers
  - Evolving Internet exchange landscape
- Software-Defined Networking
  - What? Why? How?
- Projetos de P&D em OpenFlow/SDN no CPqD
  - RouteFlow

Cloud & WAN

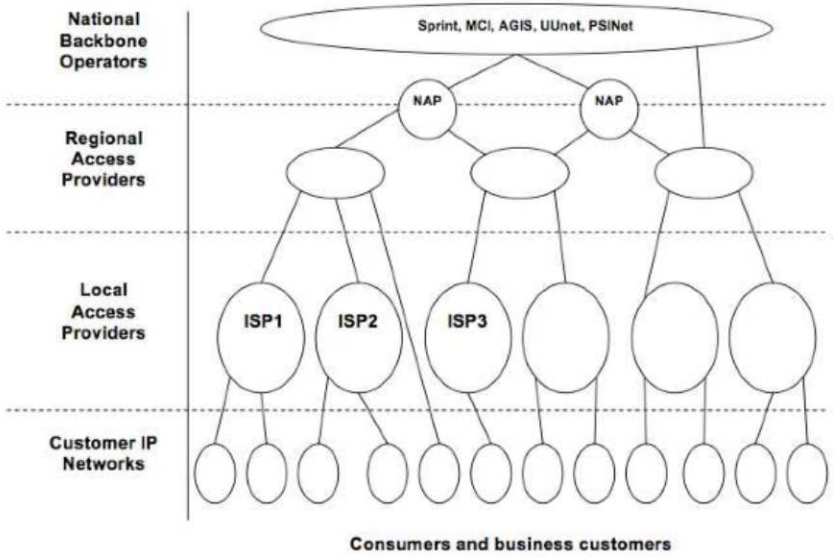

Flexible networking & distributed virtual infrastructure
Legacy apps in cloud (IaaS)
DC based clouds
Legacy apps on dedicated servers
SAP
ORACLE
Apache
Application evolution
Drivers: IaaS including networking in one package, Tailored multisite connectivity
Drivers: Server consoliation economies of scale, pay-per-use / scale-with-demand
Source: EU FP7 SAIL
4-March-2012
Mid '90's
Mid '00's
Mid '10's
SEVENTH FRAMEWORK PROGRAMME

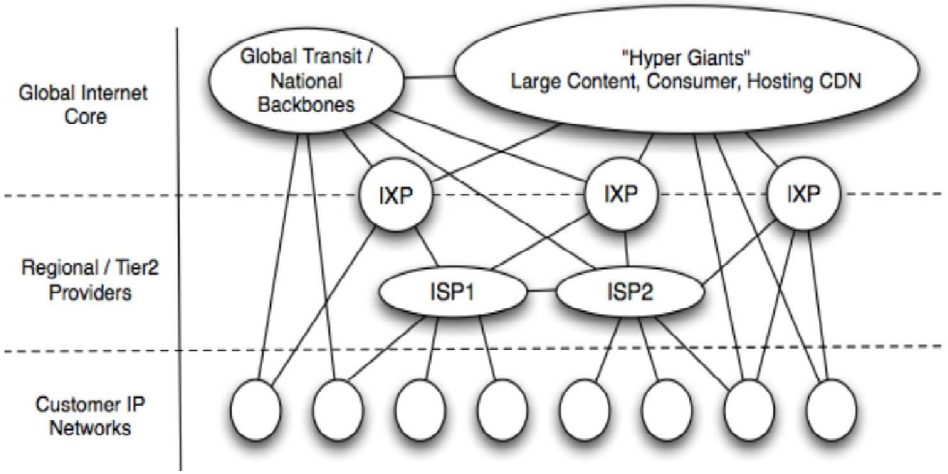
World Clouds

http://www.datacentermap.com/

The Internet Exchange Point Ecosystem

- **Internet Exchange Point (IXP) – PTT (Ponto de Troca de Tráfego)**
  - L2 service to facilitate the interconnection between ISPs / SPs / CDNs
  - PTTMetro project by CGI.br / NIC.br
- **“Flattening Internet”:**
  - Traffic increasingly exchanged directly between CDN / Clouds & regional ISPs





Further reading

B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger. ***Anatomy of a Large European IXP***. Proc. of ACM SIGCOMM 2012.
William B. Norton. The Internet Peering Playbook : ***Connecting to the Core of the Internet***. DrPeering Press, 2012.
C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. ***Internet Interdomain Traffic.*** Proc. of ACM SIGCOMM 2010

INTER-CLOUD

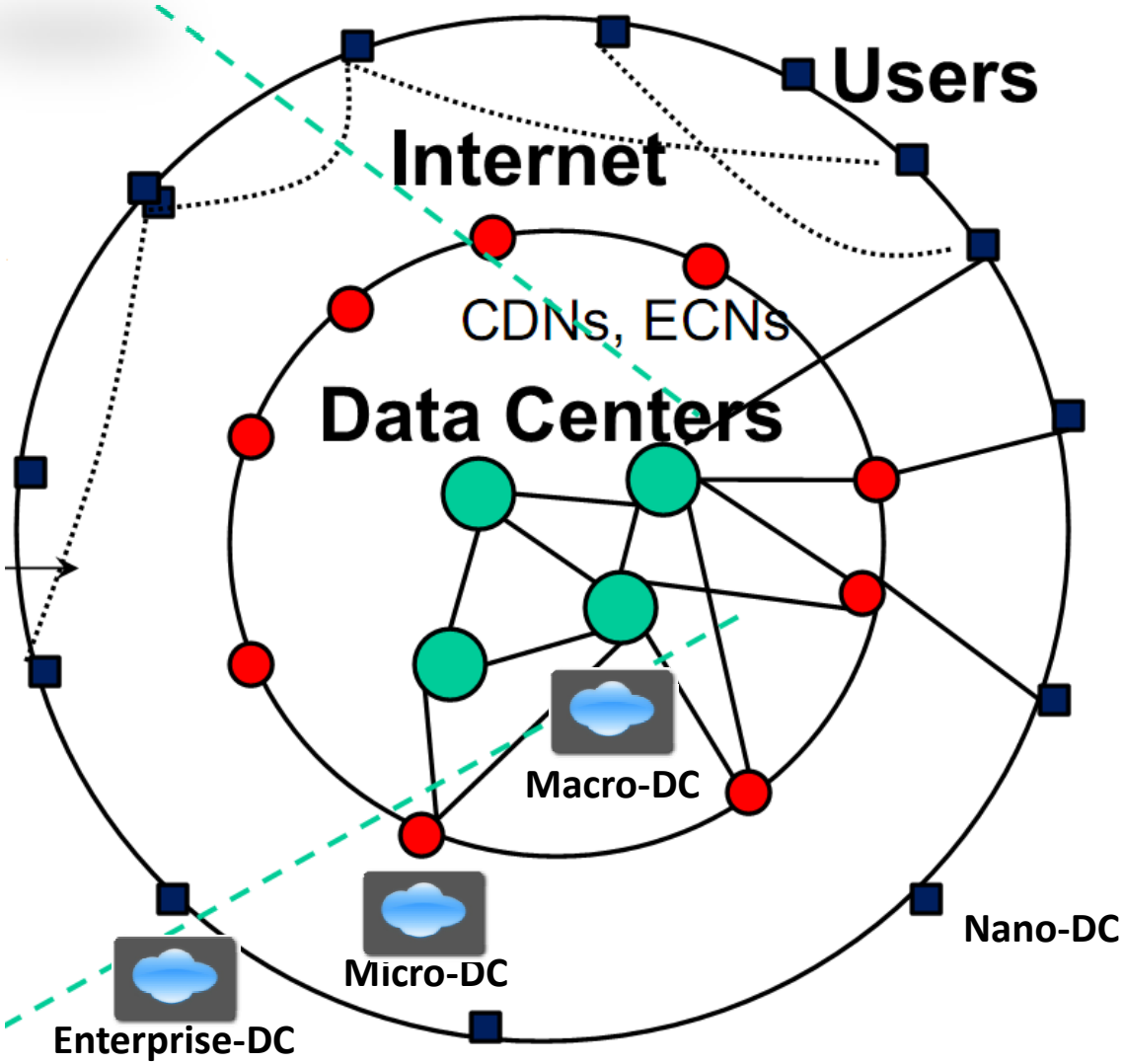
Interconnecting the Clouds

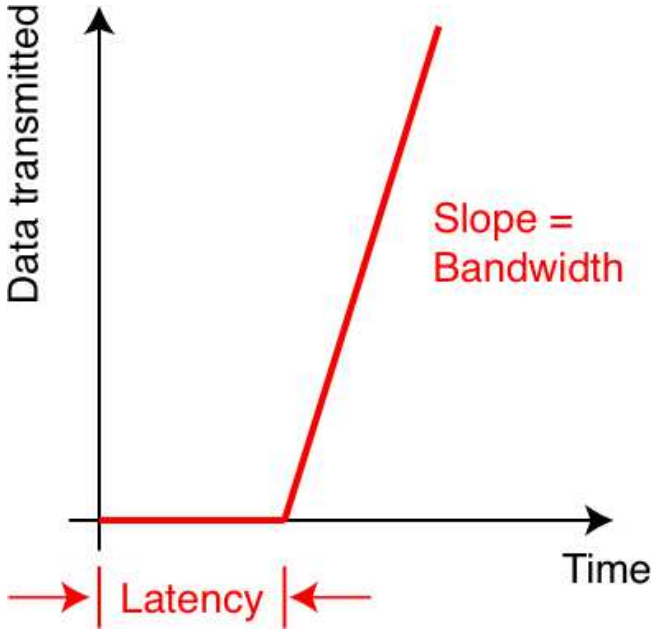
Connecting to the Cloud

Enabling hybrid Clouds

Networking the Cloud

Address high availability and low latency

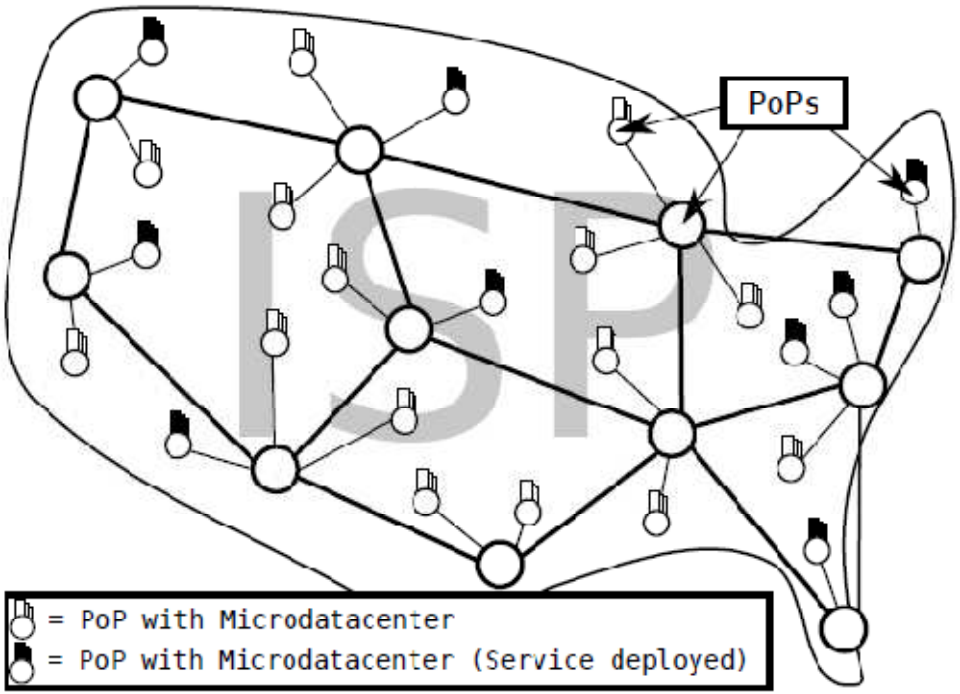




[Core image adapted from A. Greenberg SIGMETRICS 09 Tutorial]

Content Delivery Networks 3.0

- Hybrid infrastructures: Akamai, PPTV
- Meta-CDNs, e.g., Conviva
- Virtual CDNs through ISP micro-datacenters



Source: S. Uhlig
http://www.informatics.sussex.ac.uk/research/projects/ngn/slides/msn12talks/uhlig-forces.pdf

Van Jacobson's waves of networking

"If a Clean Slate is the solution, what was the problem?"

99%

Named chunks of data (Web, P2P, Video, etc.)

New problem!

- Dissemination of named pieces of data

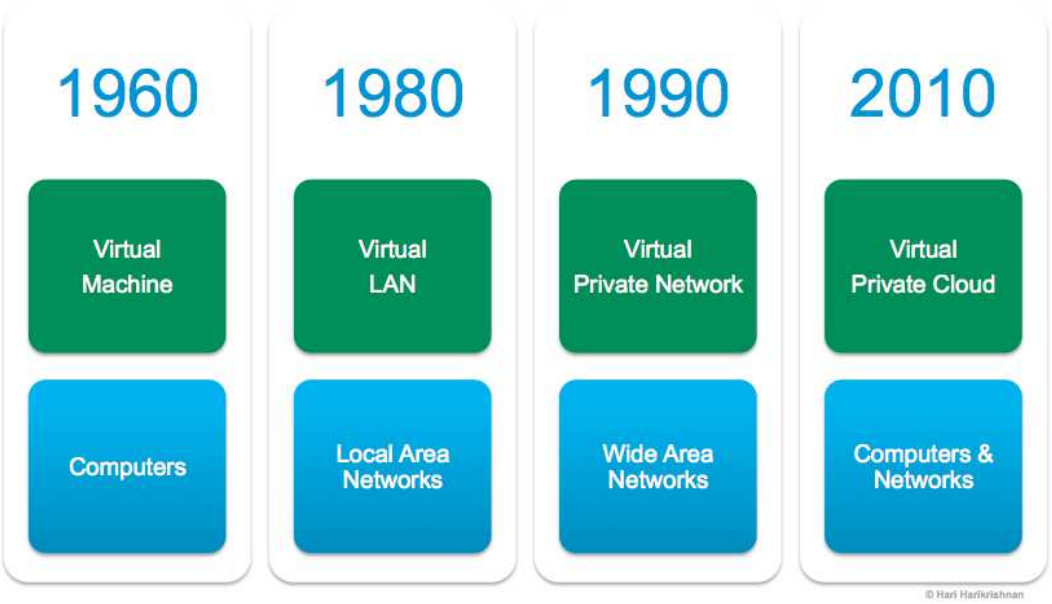
Answer:

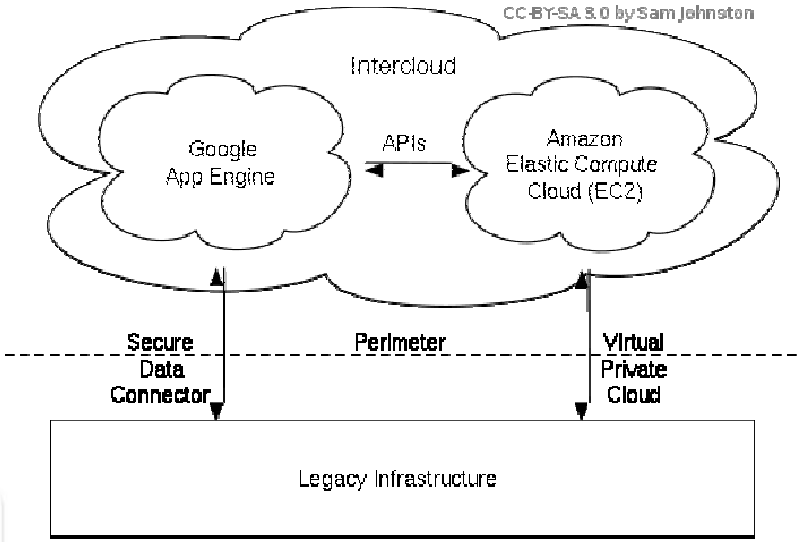
- Content-Centric Networking / Named Data Networking

http://video.google.com/videoplay?docid=-6972678839686672840

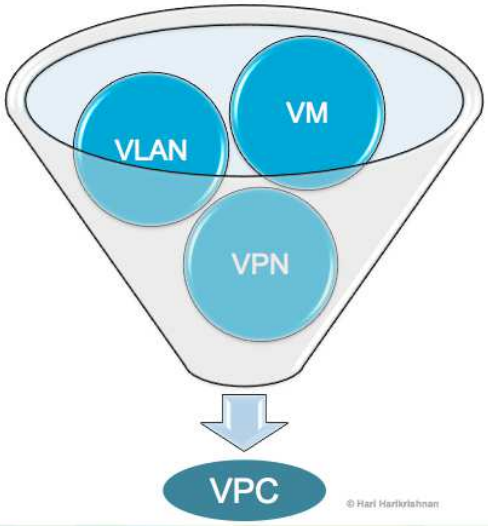
Cloud inter-networking

- Demand for Virtual Private Clouds
  - QoS, privacy, security, availability, etc.
- Inter-Cloud Connectivity
  - Identity of information, security, agility, cost, etc.



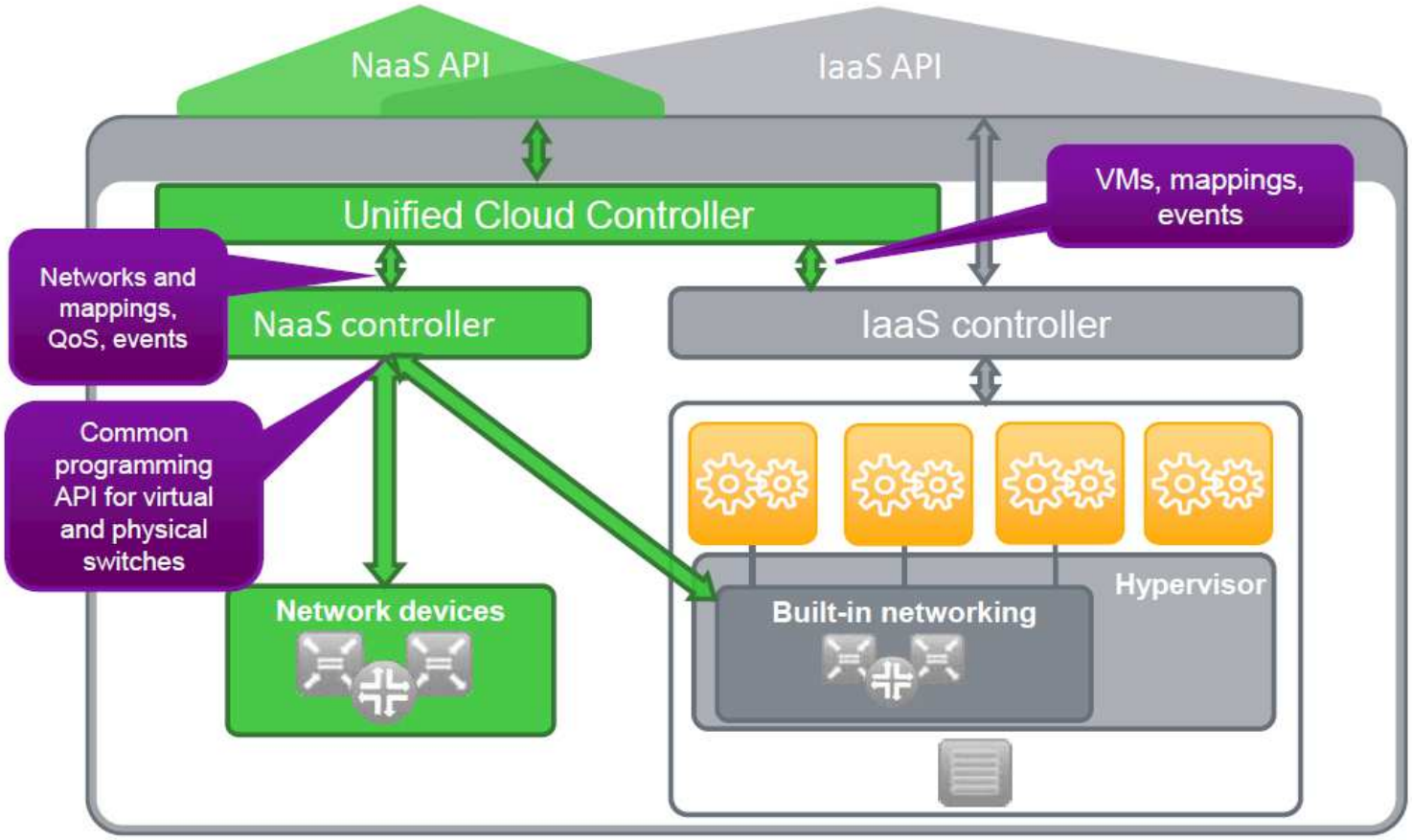


Virtual Private Cloud (VPC)



Network-as-a-Service (NaaS)

- Telco Grade Clouds?
- Cloud Operator? Virtual Private Cloud Operator / Broker?

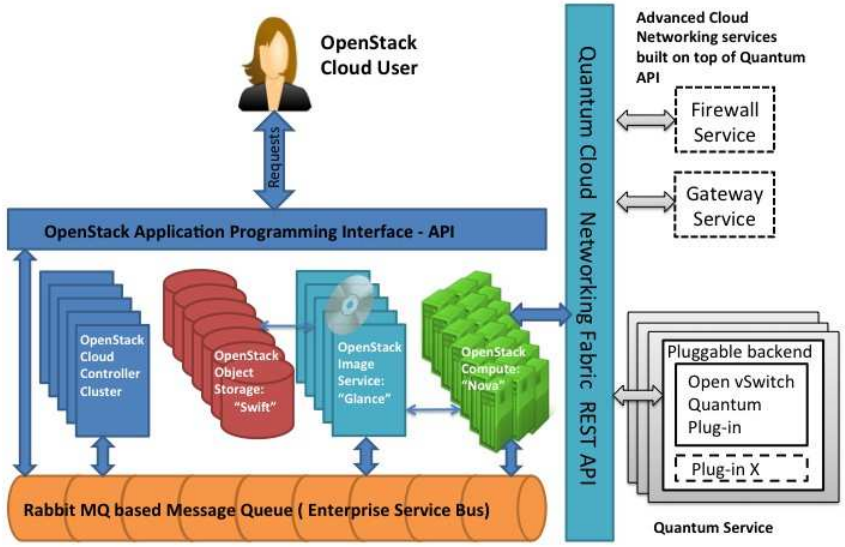


Source: EU FP7 4CaaSt

Open-Source Cloud & Networking

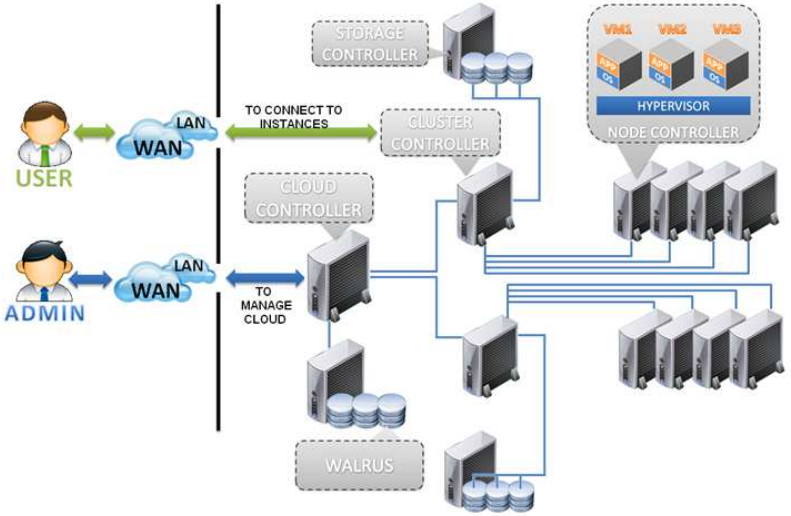
openstack™
CLOUD SOFTWARE

EUCALYPTUS



OpenStack + Quantum Integration Architecture

Source: somik@nicira



OpenFlow

Source: http://yoyoclouds.files.wordpress.com/2011/12/eucalyptus-cloud-setup.png

Agenda

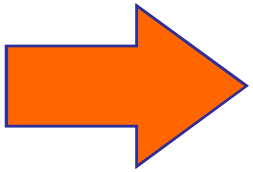
- Networking for the Cloud
  - Inside the data center
  - Between data centers
  - Evolving Internet exchange landscape
- Software-Defined Networking
  - What? Why? How?
- Projetos de P&D em OpenFlow/SDN no CPqD
  - RouteFlow

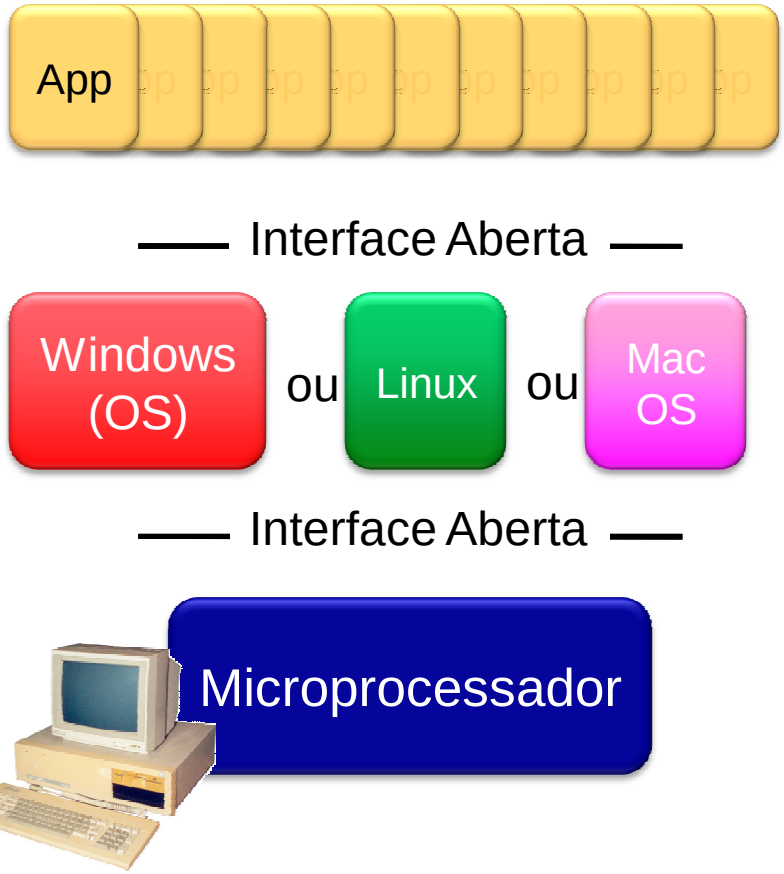
Industria de Redes de Pacotes = Switches e Roteadores

- Industria verticalizada
  - Interfaces fechadas/proprietárias
  - (Pouca) Inovação pelos fabricantes
  - Pouca competição = industria pequena = preços altos
- Produto segmentado
  - Preço alto por fração útil pequena
  - Necessidades específicas não atendidas
- Valor percebido no HW
  - 60-80% SW Embarcado

A Solução da Industria da Computação

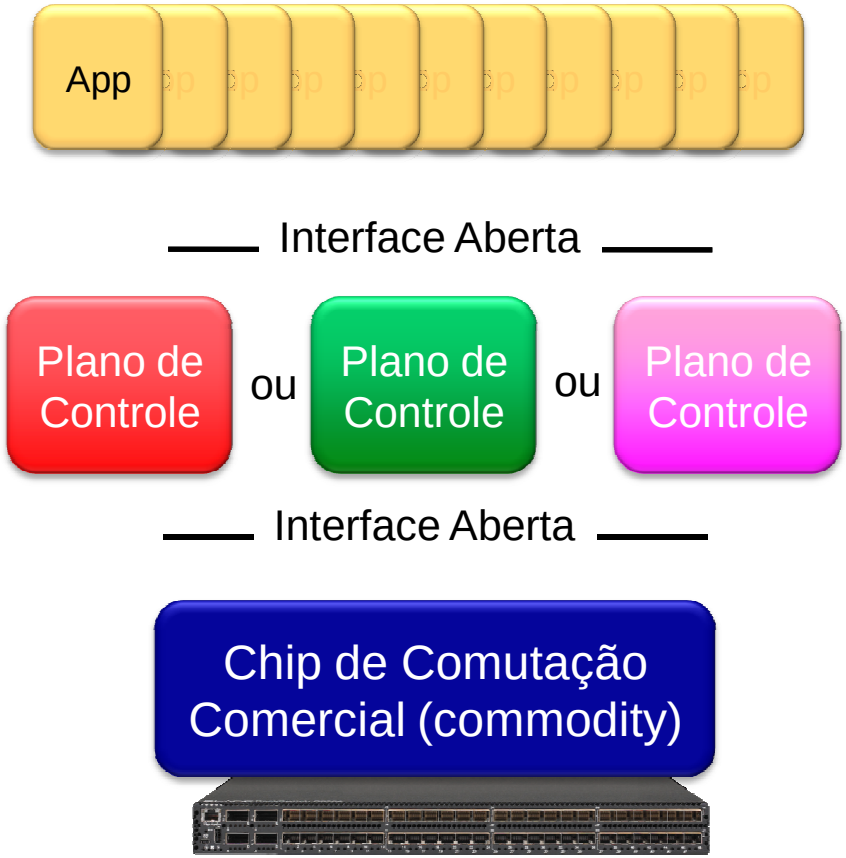
Verticalmente integrado
Fechado, proprietário
Inovação lenta
Industria pequena





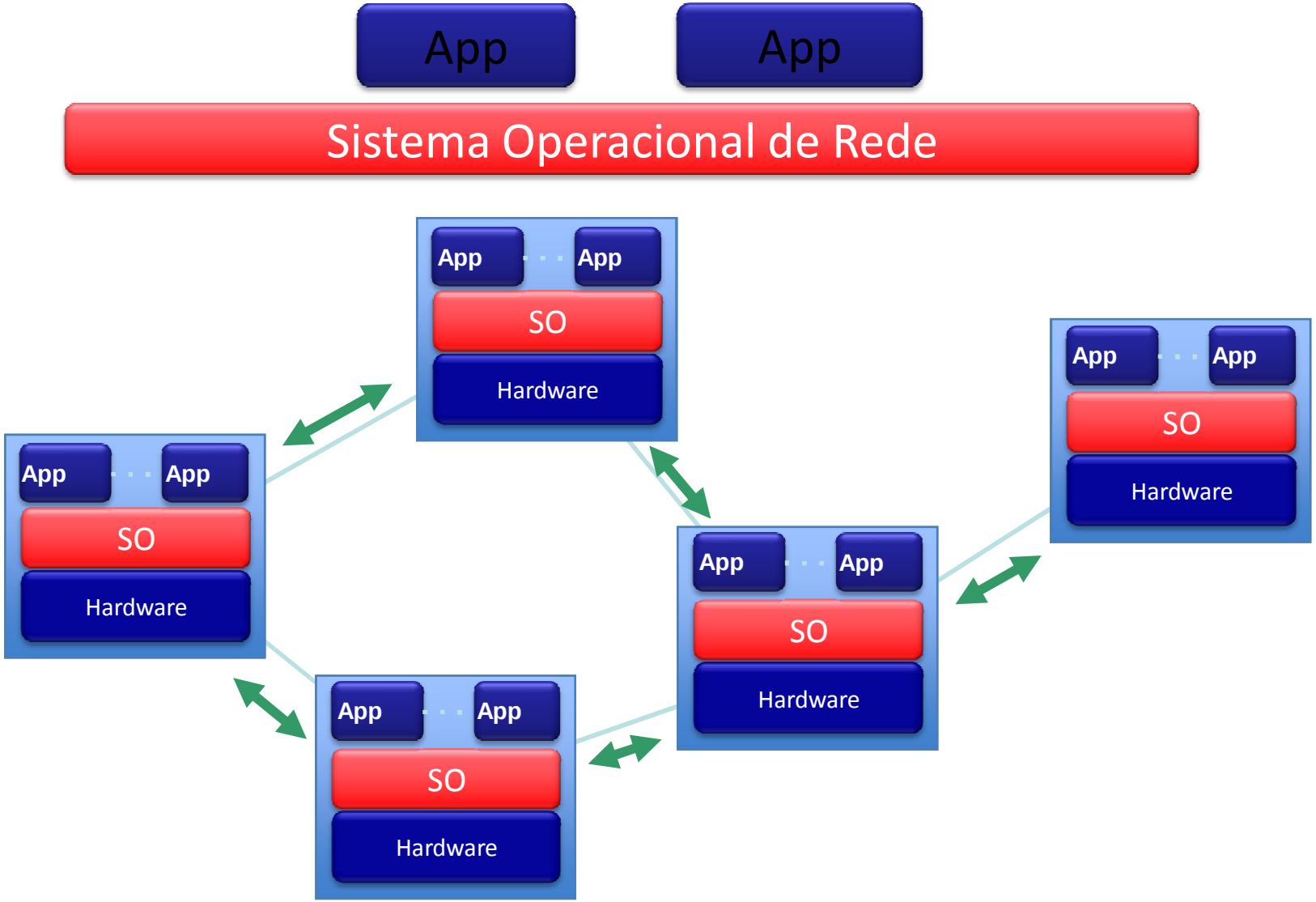
Horizontal
Interfaces abertas
Inovação acelerada
Industria enorme

Industria de Redes de Pacotes Definidas por Software

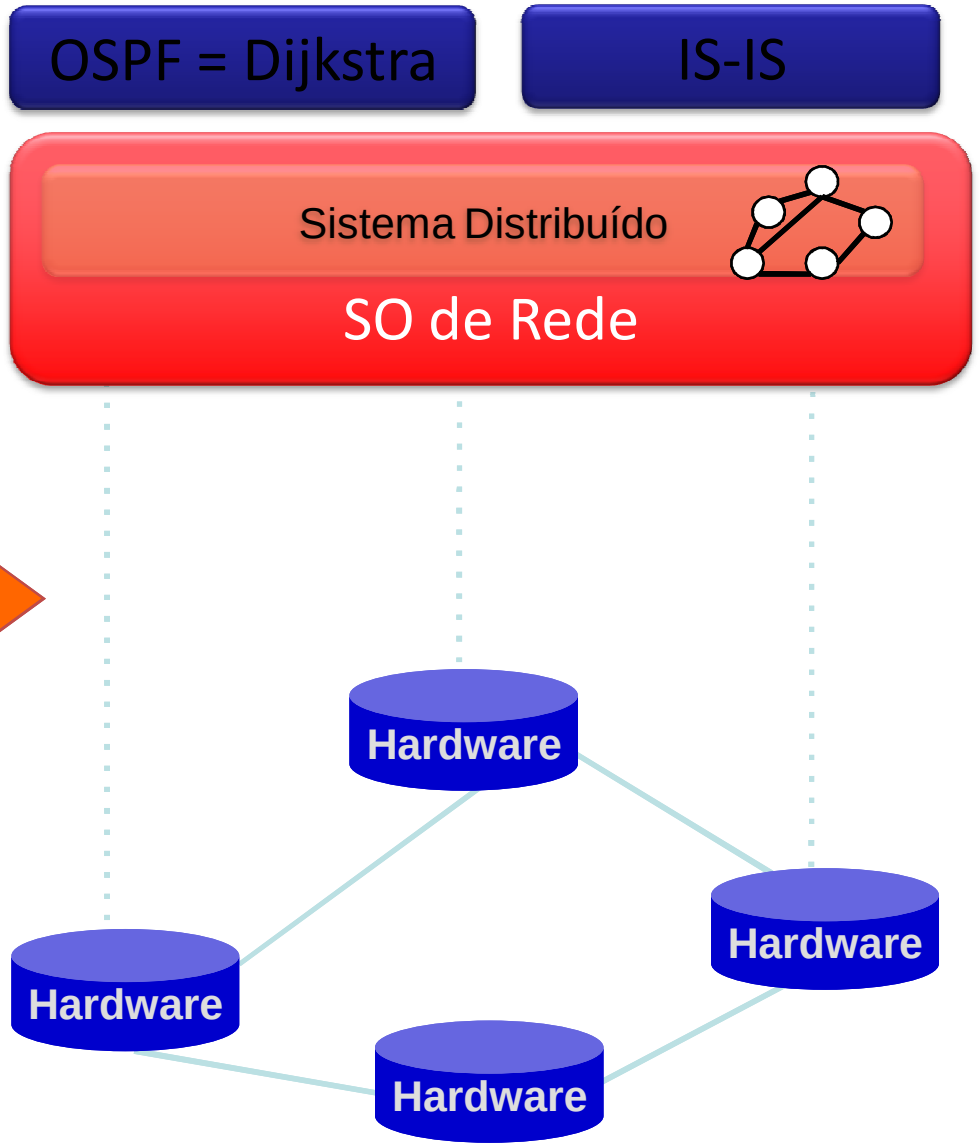


- Industria horizontalizada
  - Interface abertas = Inovação ilimitada
  - Competição em (quase) todas as camadas = industria grande = preços mais baixos
- Customização
  - Atendimento de necessidades específicas
  - Preço justo por utilidade e flexibilidade
- Valor percebido no SW
  - No HW também se for produto high-end

Rede Definida por Software (SDN)

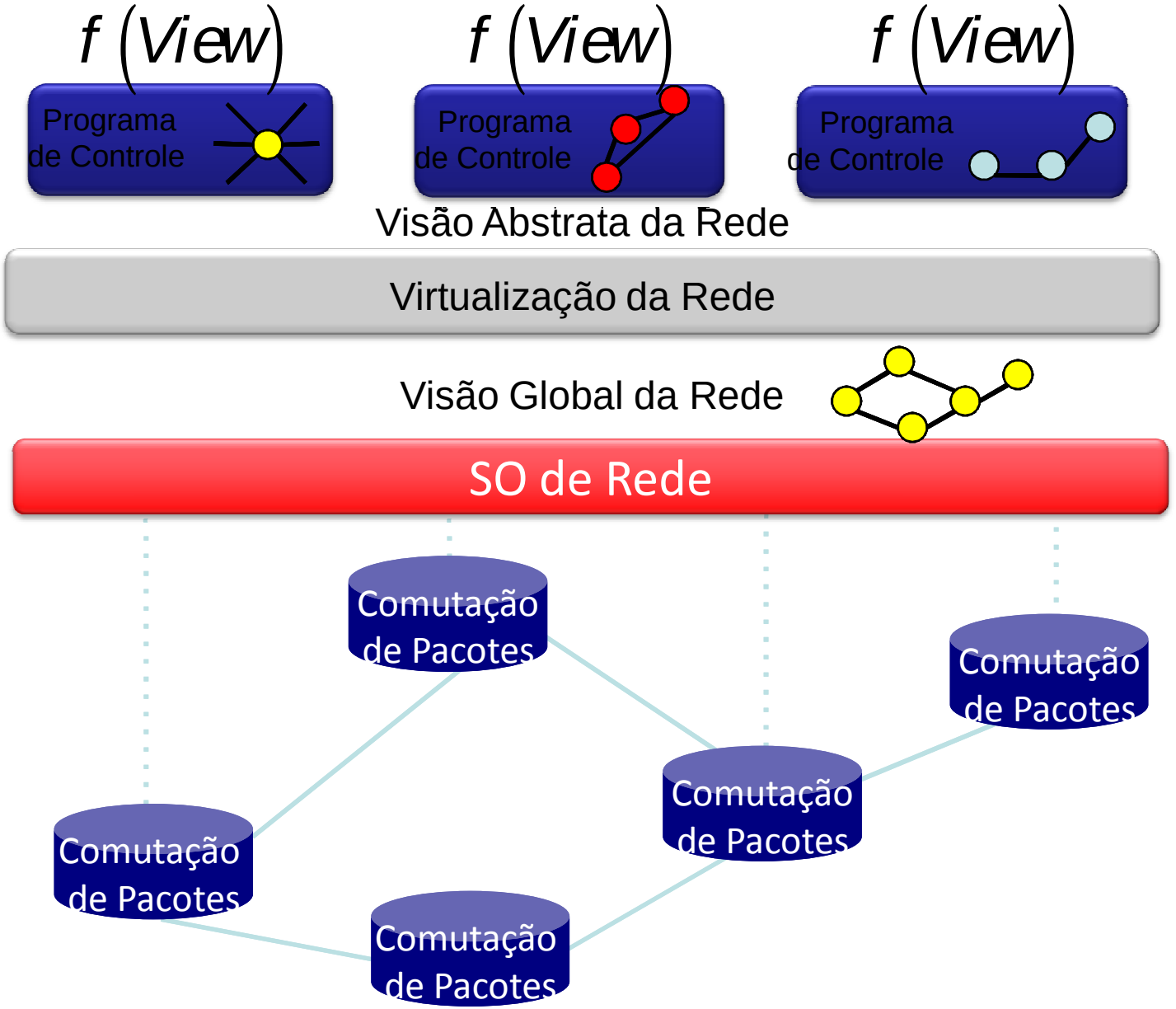


Exemplo: Roteamento Traditional x SDN



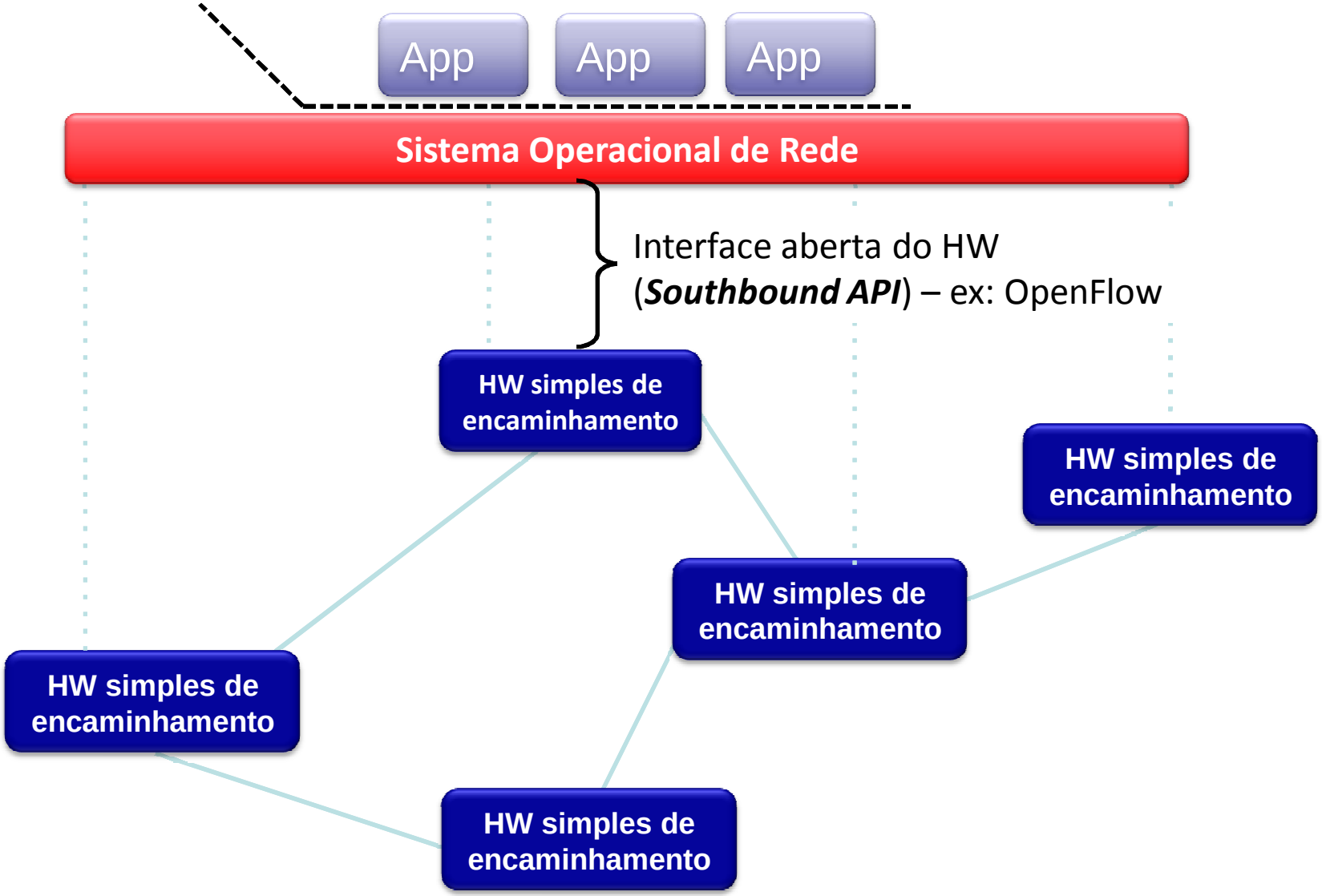
RFC do OSPF contém 244 páginas, das quais somente 4 tratam do algoritmo de roteamento (Dijkstra)

Visão atual de SDN: Programabilidade e Virtualização



Arquitetura SDN

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



Software-Defined Networking (SDN)

- Not just an idle academic daydream
  - Tapped into some strong market need
- One of those rare cases where we "*know the future*"
  - Still in development, but consensus on inevitability
- Much more to story than "OpenFlow" trade rag hype
  - A revolutionary paradigm shift in the network control plane

See Scott Shenker talk!

Layers are Main Network Abstractions

- Layers provide nice **data plane** service abstractions
  - IP's best effort delivery
  - TCP's reliable byte-stream

- *Aside*: good abstractions, **terrible interfaces**
  - Don’t sufficiently hide implementation details

- *Main Point*: No **control plane** abstractions
  - No sophisticated management/control building blocks

Source: Scott Shenker

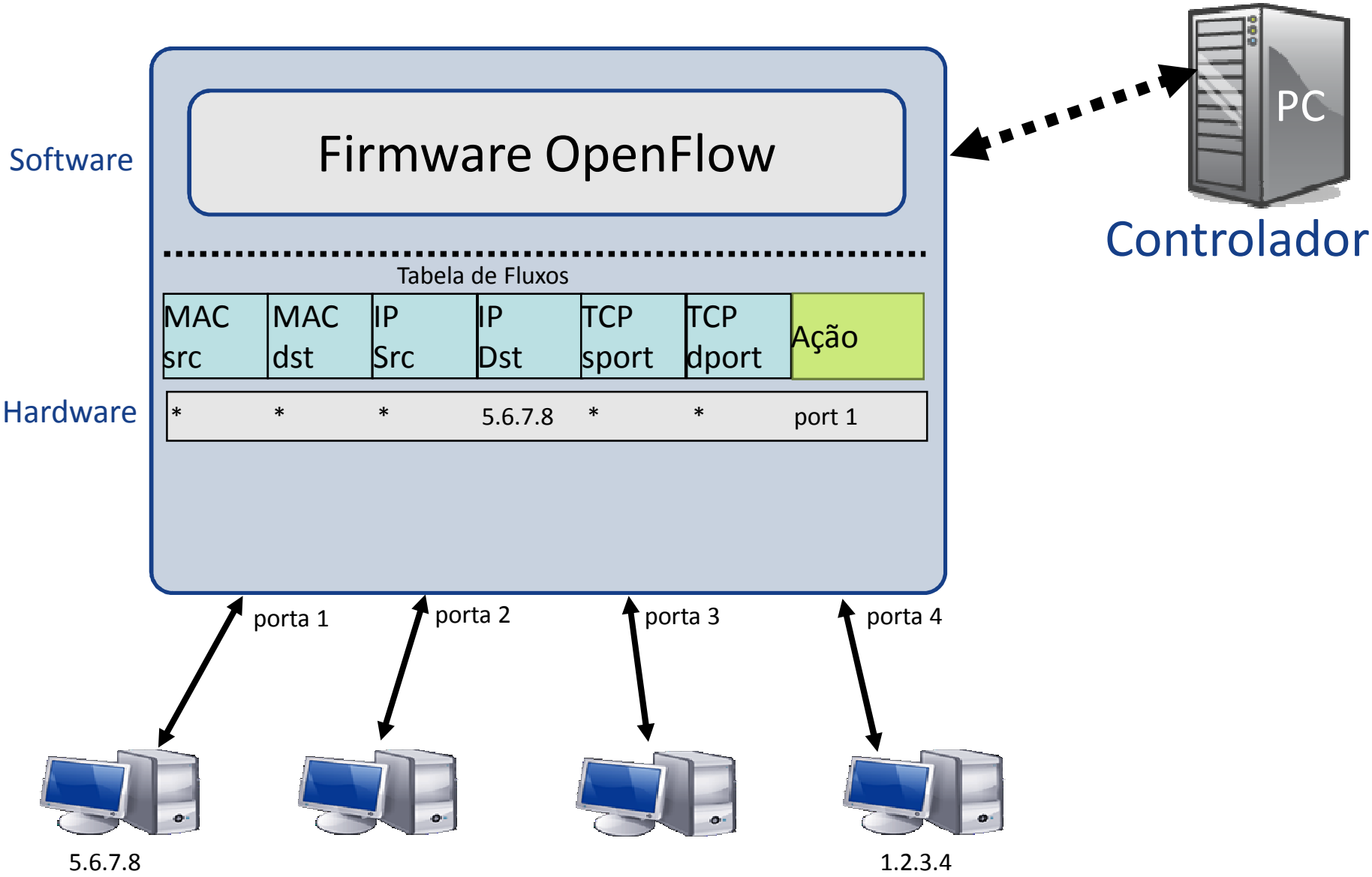
No Abstractions = Increased Complexity

- Each control requirement leads to new mechanism
  - TRILL, LISP, etc.
- We are really good at designing mechanisms
  - So we never tried to make life easier for ourselves
  - And so networks continue to grow more complex
- But this is an unwise course:
  - Mastering complexity cannot be our only focus
  - Because it helps in short term, but harms in long term
  - **We must shift our attention from mastering complexity to extracting simplicity….**

Source: Scott Shenker


OpenFlow

OpenFlow: Interface de Programção do Hardware



OpenFlow: Abstração de Fluxo

Comparação

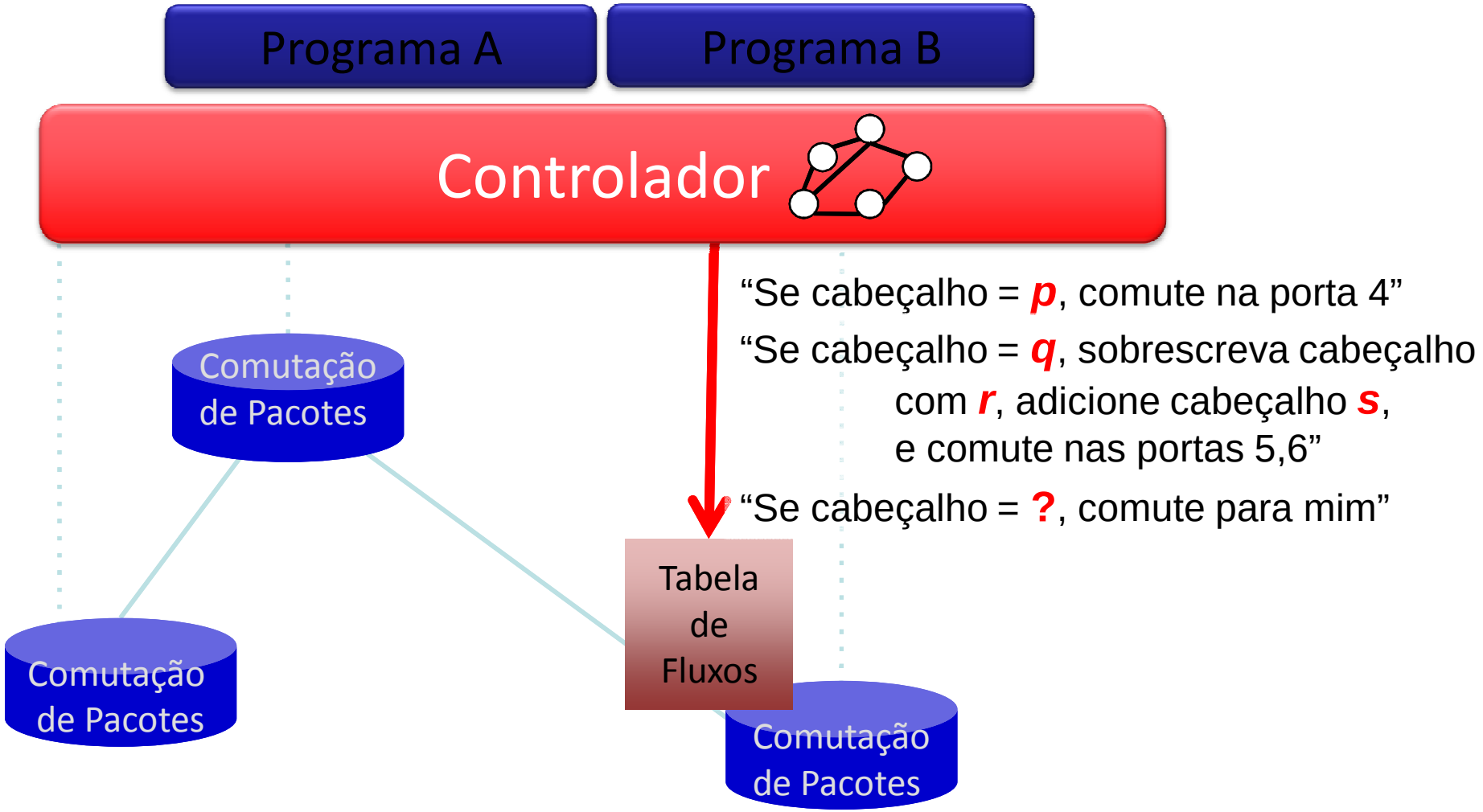
Comparação: 1000x01xx0101001x

- Compara com qualquer cabeçalho (existente ou novo)
- Permite qualquer granularidade de fluxo

Ação

- Encaminha para a porta(s), descarta, envia ao controlador
- Sobrescreve cabeçalhos com mascara, coloca cabeçalho, remove cabeçalho
- Encaminha na taxa de transmissão específica

OpenFlow: Exemplo básico



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

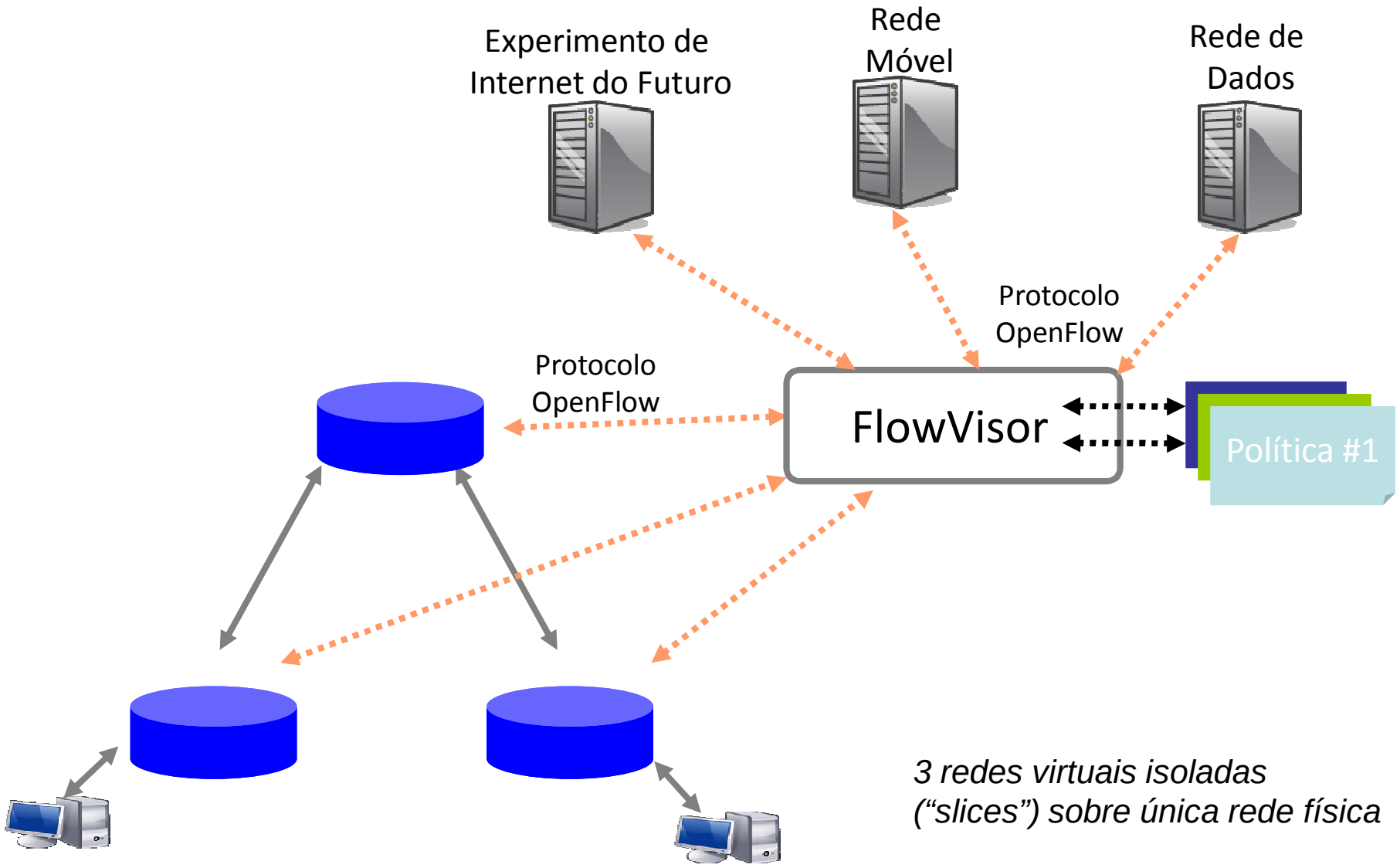
Roteamento

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

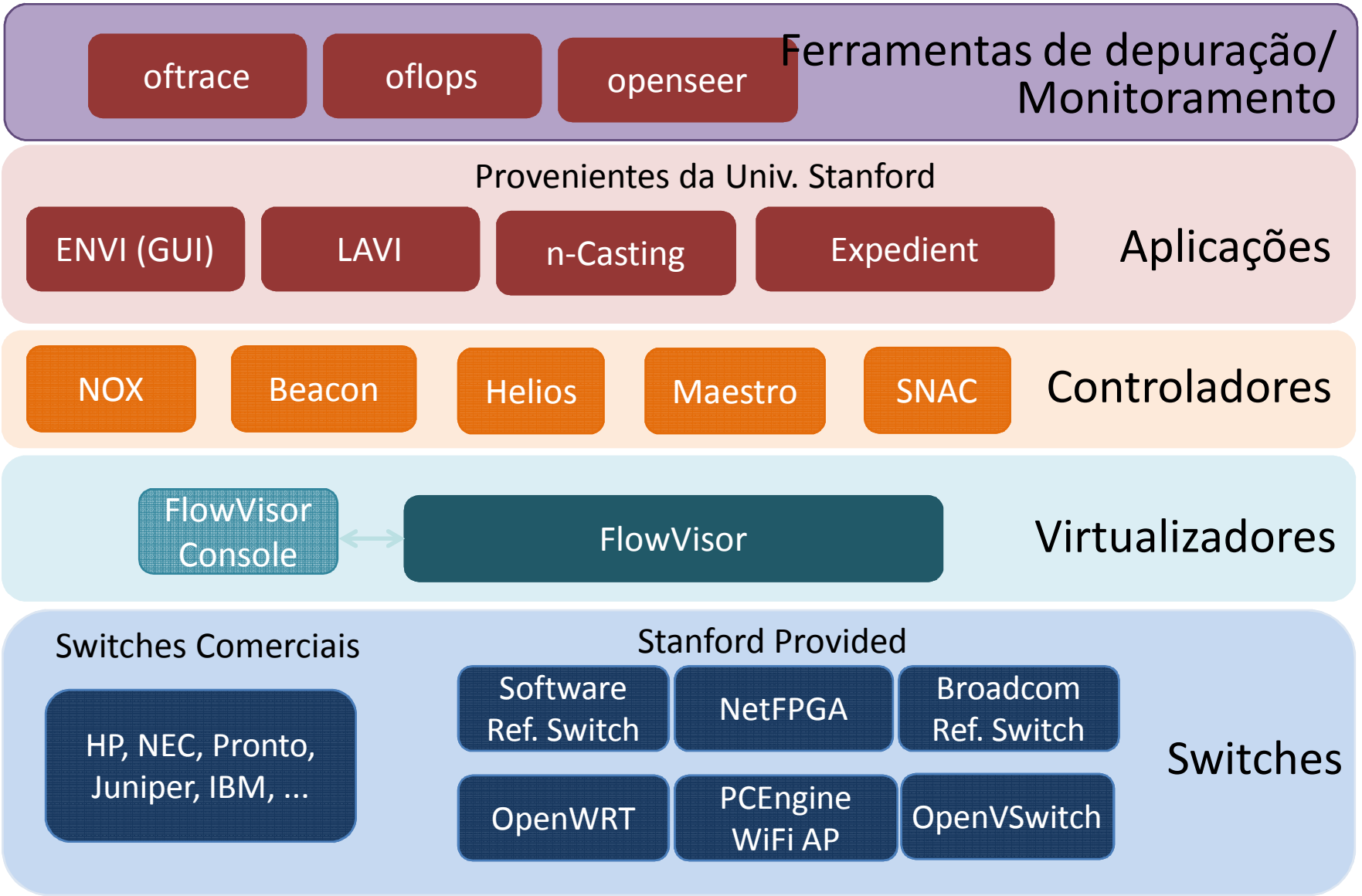
Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Ação
*	*	00:1f..	*	vlan1	*	*	*	*	*	Porta 6, Porta 7, Porta 9

FlowVisor: Virtualização OpenFlow



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

Ecossistema OpenFlow (visão 2010)



Ecossistema OpenFlow (visão 2010)

Provedores de Datacenter e Operadoras de Telecomunicações

Google Microsoft Deutsche Telekom Level(3) BT amazon.com facebook NTT docomo verizon

Redes Experimentais

Fabricantes

NEC hp JUNIPER NETWORKS ARISTA CISCO extreme networks ciena HUAWEI FUJITSU DELL NETGEAR big switch networks nicira BROADCOM MARVELL

Padronização de SDN e do OpenFlow

Board of Directors

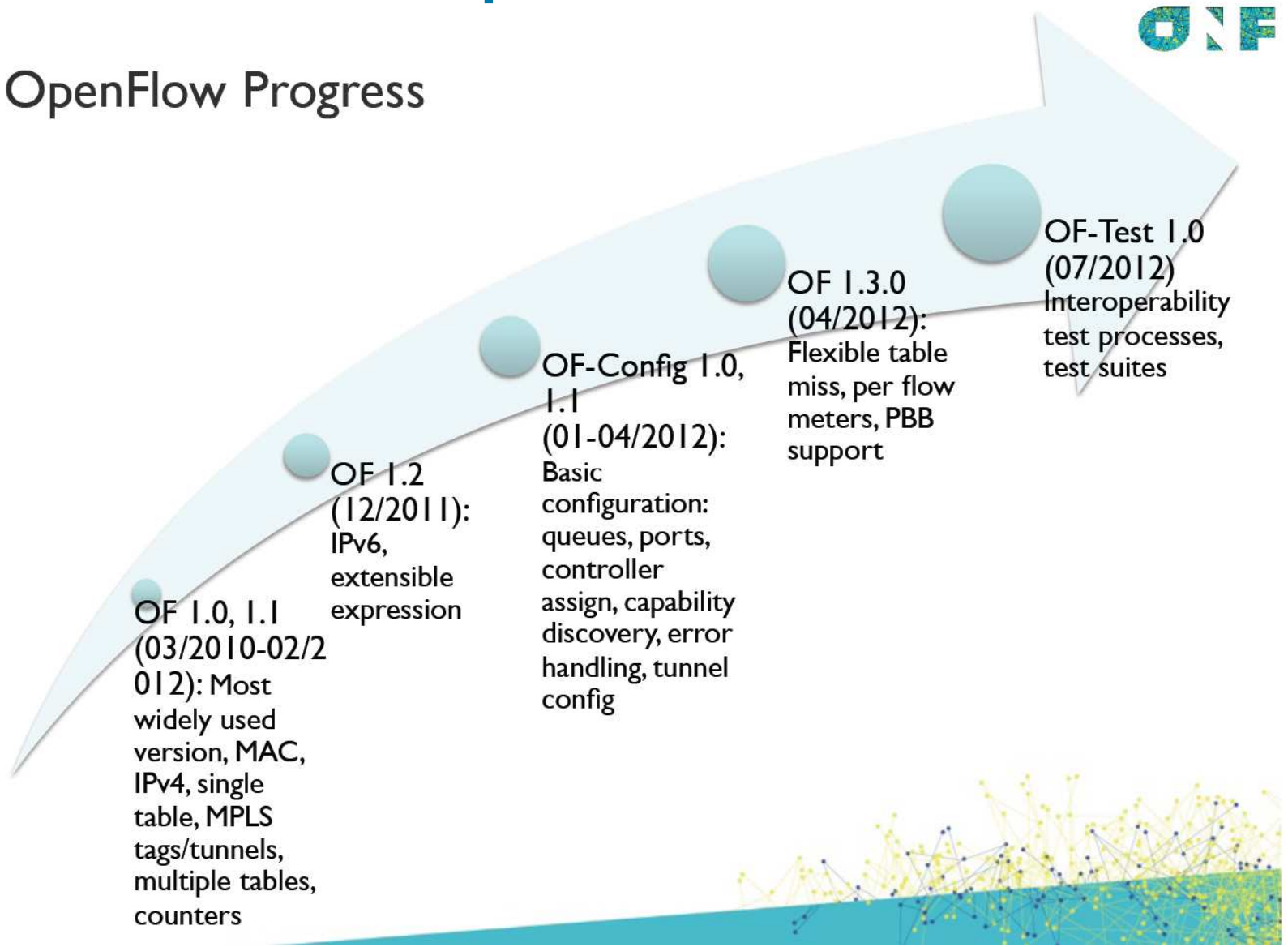
- Deutsche Telekom
- Facebook
- Google
- Microsoft
- Verizon
- Yahoo!

Members

- Broadcom
- Brocade
- Ciena
- Cisco
- Citrix
- Dell
- Ericsson
- Force10
- HP
- IBM
- Juniper Networks
- Marvell
- NEC
- Netgear
- NTT
- Riverbed Technology
- VMware

http://www.opennetworkingfoundation.org/

OpenFlow Roadmap



Apelo de Mercado

Dar aos proprietários e operadores de rede o poder de

- Customizar redes às suas necessidades
- Eliminar (e deixar de pagar por) funcionalidades desnecessárias
- Simplificar (e reduzir custos de) operação
- Aumentar vida útil da infraestrutura de redes
- Criar redes virtuais isoladas e especializadas

Acelerar o ritmo da inovação em toda cadeia de valor do setor

- Inovação na velocidade do software (e não do hardware)
- Reuso de softwares disponíveis na comunidade
- Efeito multiplicador da Inovação Aberta

Oportunidades para o País

- Ter um papel relevante na definição da nova arquitetura e da nova tecnologia da Internet
  - Impactos científico e principalmente tecnológico
- Criar uma industria nacional de rede competitiva internacionalmente, em particular no software
  - Fabricantes existentes
  - *Startups*
- Reduzir a dependência tecnológica e as importações
  - Melhorar balança comercial
- Gerar e explorar propriedade intelectual
- Capacitar os cientistas e engenheiros de amanhã
  - Saber fazer # saber como se faz

Oportunidades para a Industria

- Barreira de entrada mais baixa facilita entrada de novas (e pequenas) empresas no mercado nacional e global de telecomunicações, em particular focadas em software de rede
  - *Startups* e *spin-offs* de sucesso (pequenas notáveis)

- Percepção de valor em software representa uma mudança de paradigma que requer novos modelos de negócio e poderá levar a uma nova estrutura produtiva e de poder na industria
  - Fragmentação da cadeia de valor com empresas fornecendo soluções de software diretamente para os clientes finais ou indiretamente aos fabricantes e integradores

Agenda

- Networking for the Cloud
  - Inside the data center
  - Between data centers
  - Evolving Internet exchange landscape
- Software-Defined Networking
  - What? Why? How?
- Projetos de P&D em OpenFlow/SDN no CPqD
  - RouteFlow

Exemplos de atividades de P&D em andamento

- RouteFlow+
  - Low-cost routing, migration to IPv6, BGP Sec.
- Software-based OpenFlow switch v1.2 and v1.3
  - Collaboration with Ericsson
  - First OpenFlow v1.2 and 1.3 open-source toolkit worldwide!
- OpenFlow and Optical
  - OpenFlow-enabled ROADM (EU/Brazil FIBRE Project)
  - Autonomic Optical Reconfiguration
- OpenFlow/SDN and Wireless
  - Home networking use cases. OF 1.3 TP-Link wireless router.
- Networking for the Cloud
  - Integration of OpenFlow w/ OpenStack and transport networks

• HOW REAL IS OPENFLOW 1. X? EXPERIENCES AND ROADMAP ON SOFTWARE DEFINED CONVERGED NETWORKING FROM AN R&D CENTER FOR TELECOMMUNICATIONS , Open Networking Summit (ONS´13), Research Track Poster, April, 2013.
• "Software Defined Home Networking: Research Challenges and Innovation Opportunities", In International Workshop on Telecommunications (IWT'13), May, 2013.

OpenFlow/SDN R&D: Forwarding plane

OpenFlow Switch – first in South America

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

OpenFlow ROADM – first in South America

- WSS for mesh networks
- 3 / 5 degree
- Directioned / Coloured
- Virtualization

RouteFlow

RouteFlow is an open-source project to provide IP routing & forwarding services in OpenFlow networks

CPqD

Marcelo Nascimento

Christian E. Rothenberg

Marcos Salvador

Eder Leao Fernandes

Rodrigo Denicol

Alisson Soares

Tomas Benedotti

UniRio

Carlos Corrêa

Sidney Lucena

Unicamp

Mauricio Magalhães

Indiana University

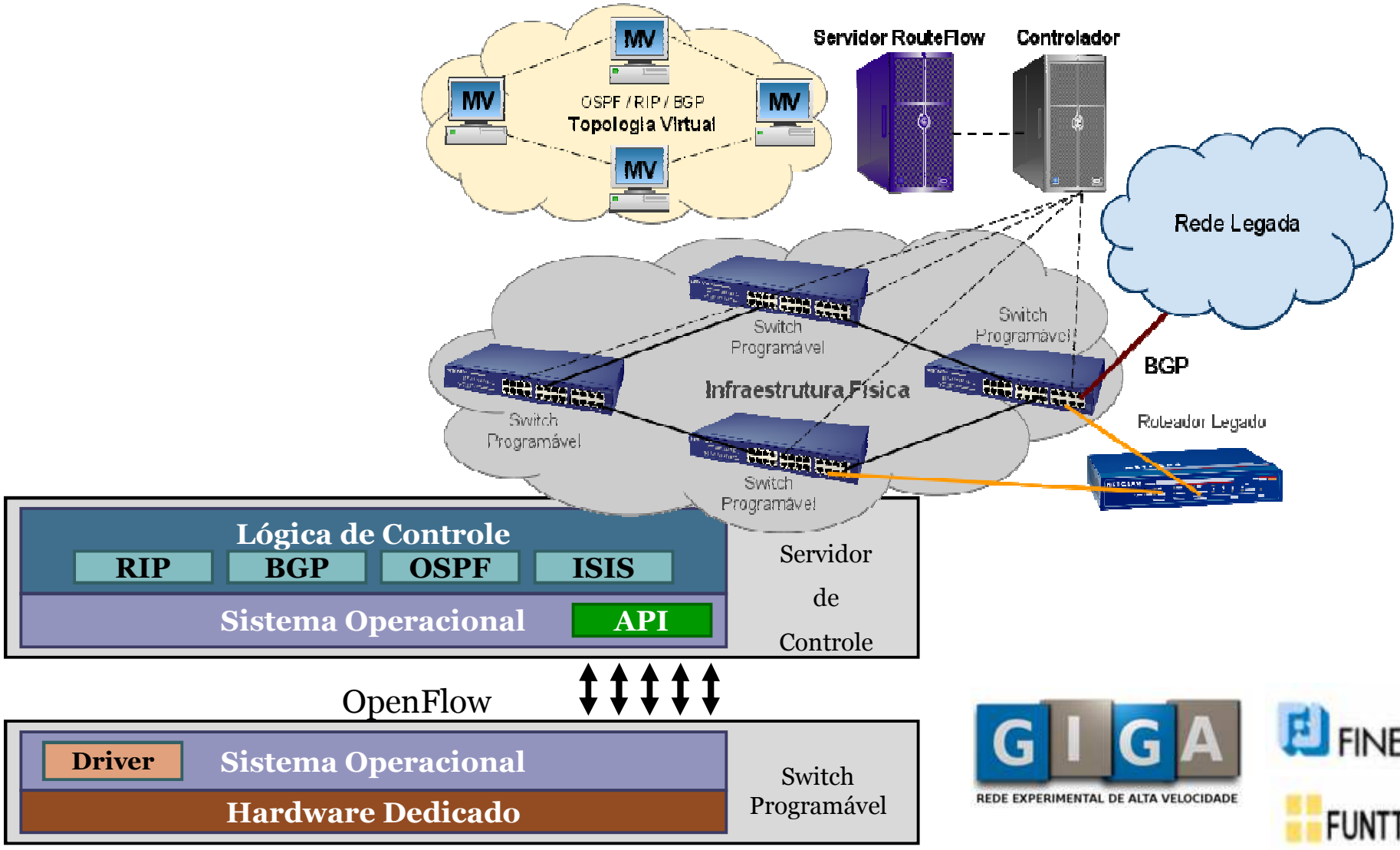
Stanford University

UFSCAR

UFPA

...

RouteFlow



GIGA
REDE EXPERIMENTAL DE ALTA VELOCIDADE

FINEP

FUNTTEL

RouteFlow Project History

- Start Msc. Thesis work by Marcelo N.
- First Prototype
- First Short-Paper @ WPEIF
- Evaluation 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: 25,000+ (12,000+ Unique)

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

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

RouteFlow

688

days since
Project Launch

GIGA
REDE EXPERIMENTAL DE ALTA VELOCIDADE

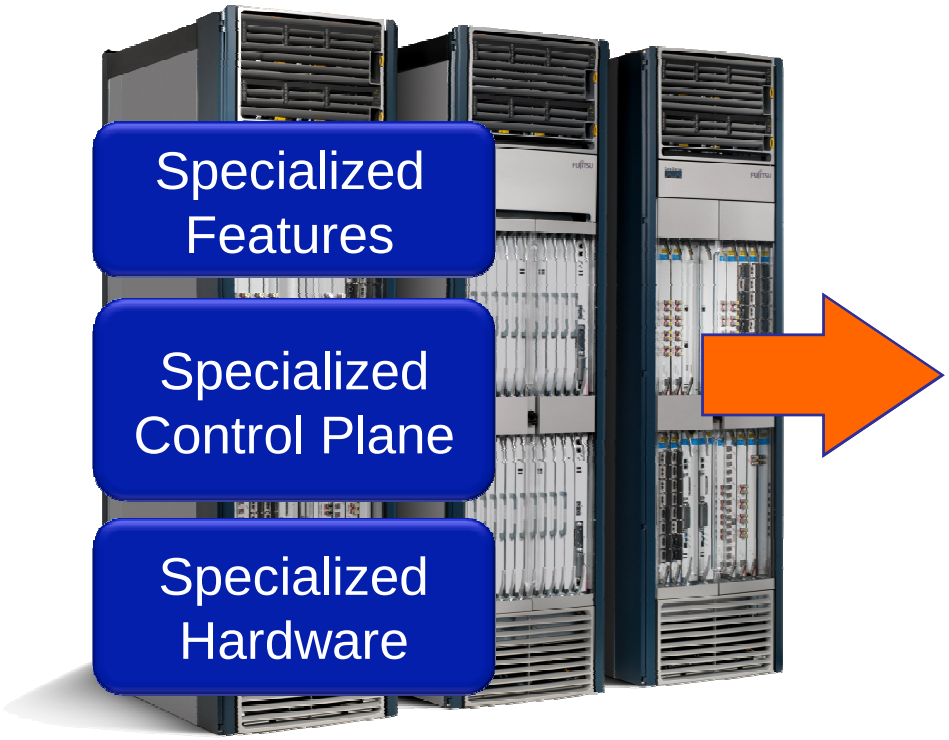
FUNTTEL FINEP

1000s downloads!

github SOCIAL CODING

Apache

Software Defined IP Routing



BGP OSPF ISIS LDP

RouteFlow

Open interface

Controller Controller Controller

Open interface

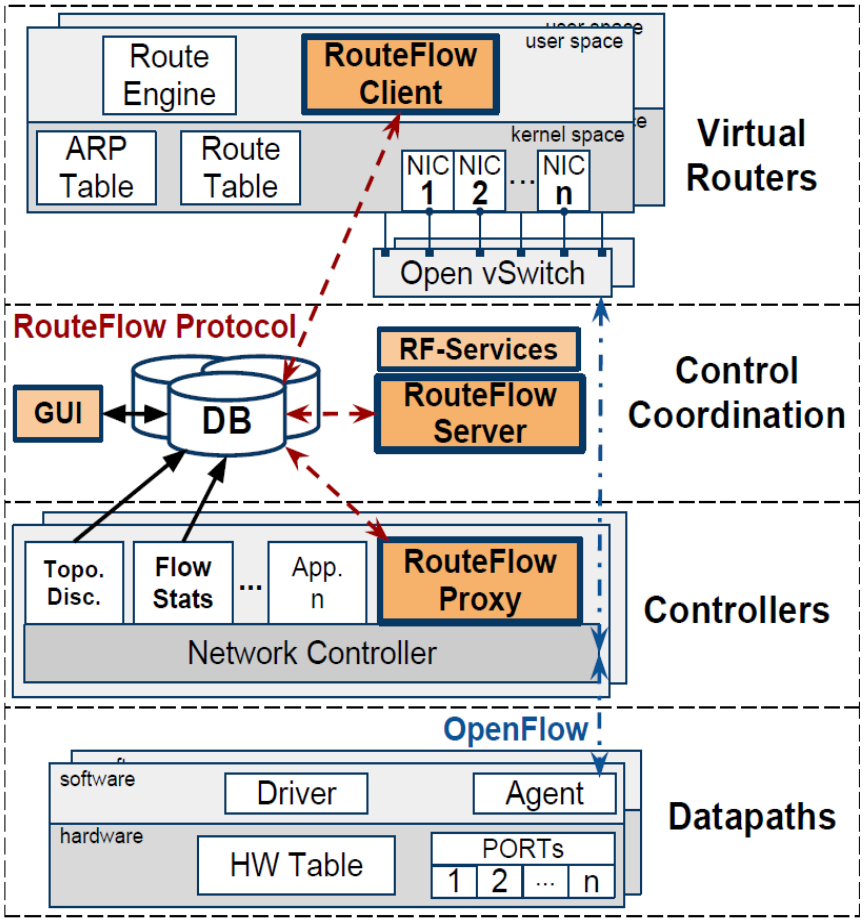
OpenFlow Switches

High cost
Specialized config.
Closed source
Slow innovation

Source: McKeown

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

Design

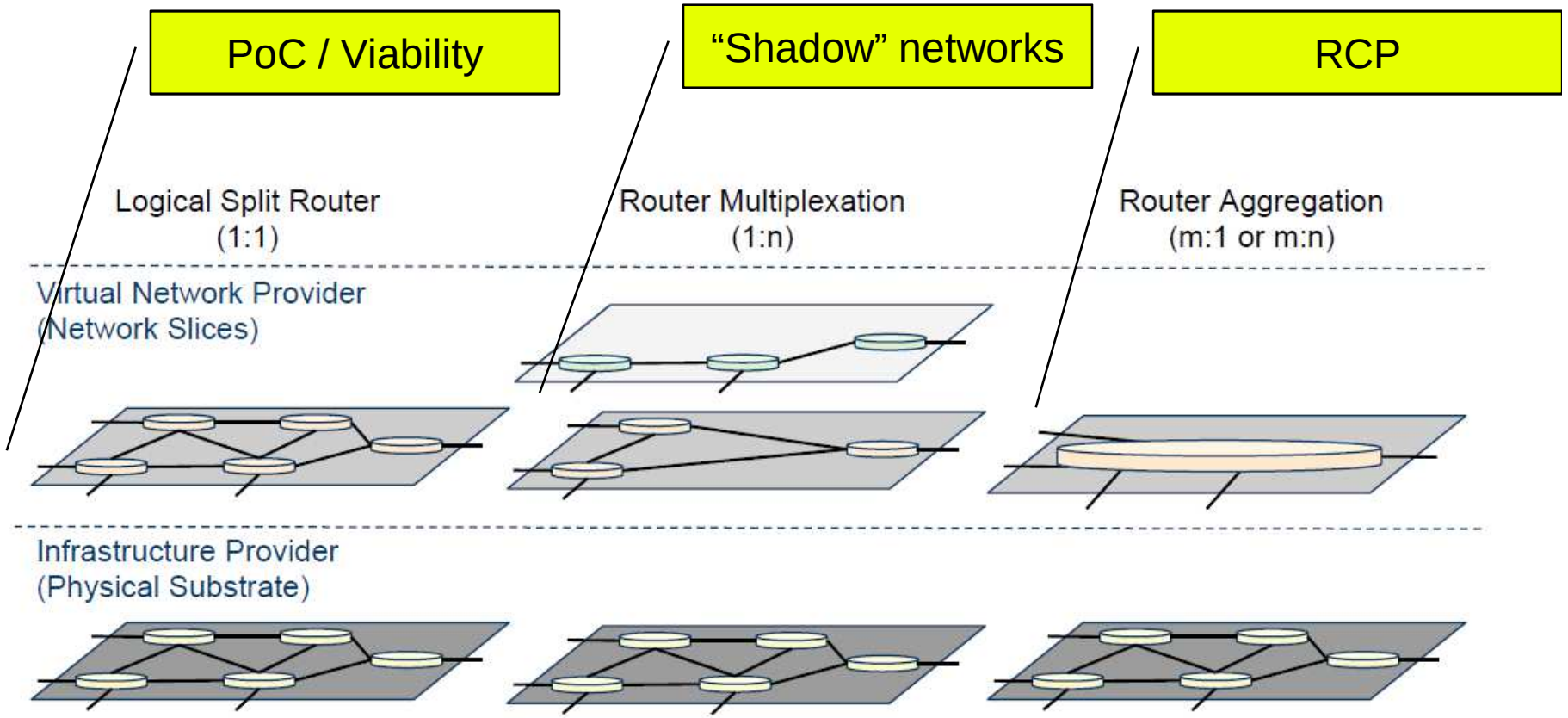


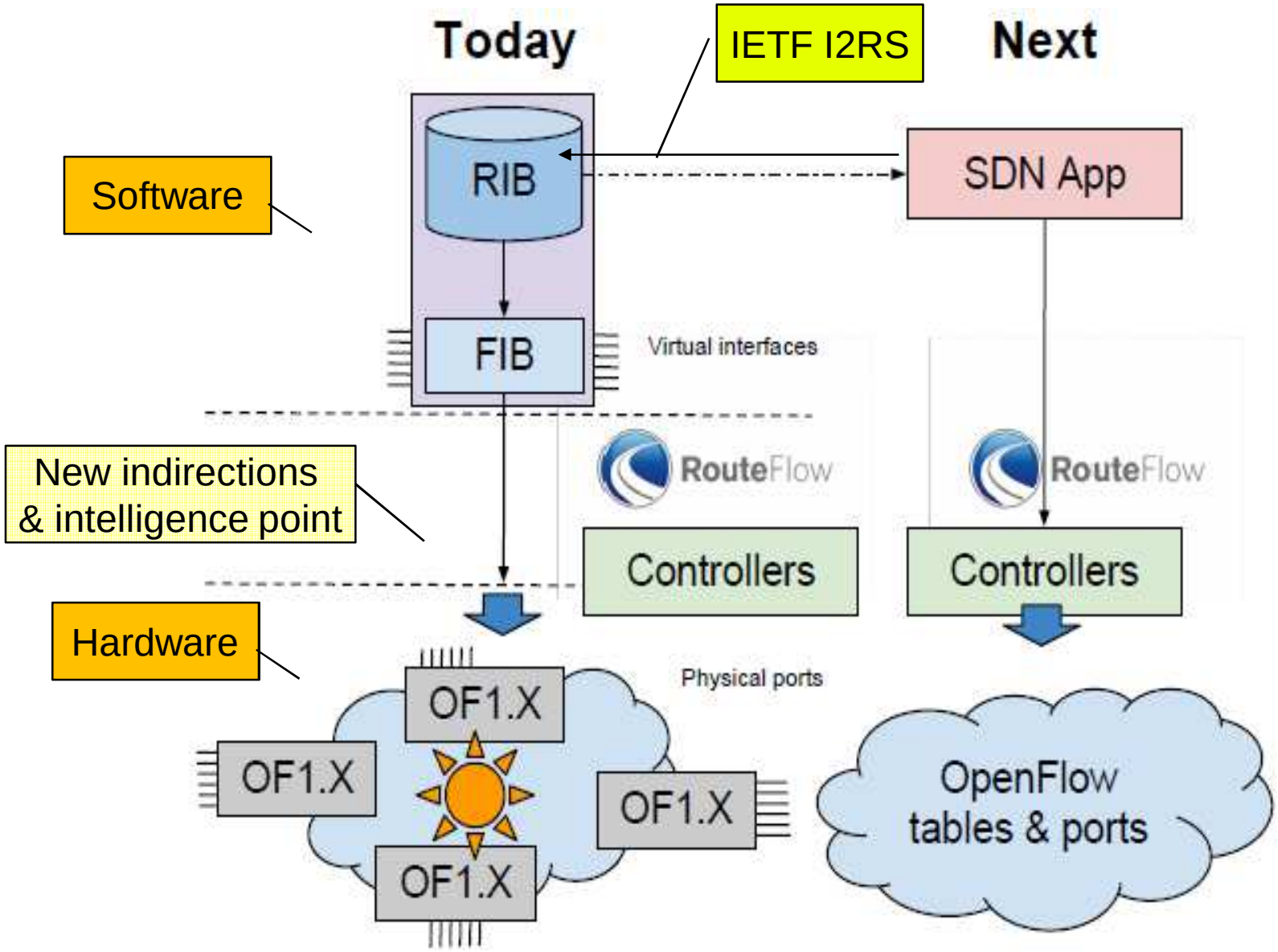
Key Features

- 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

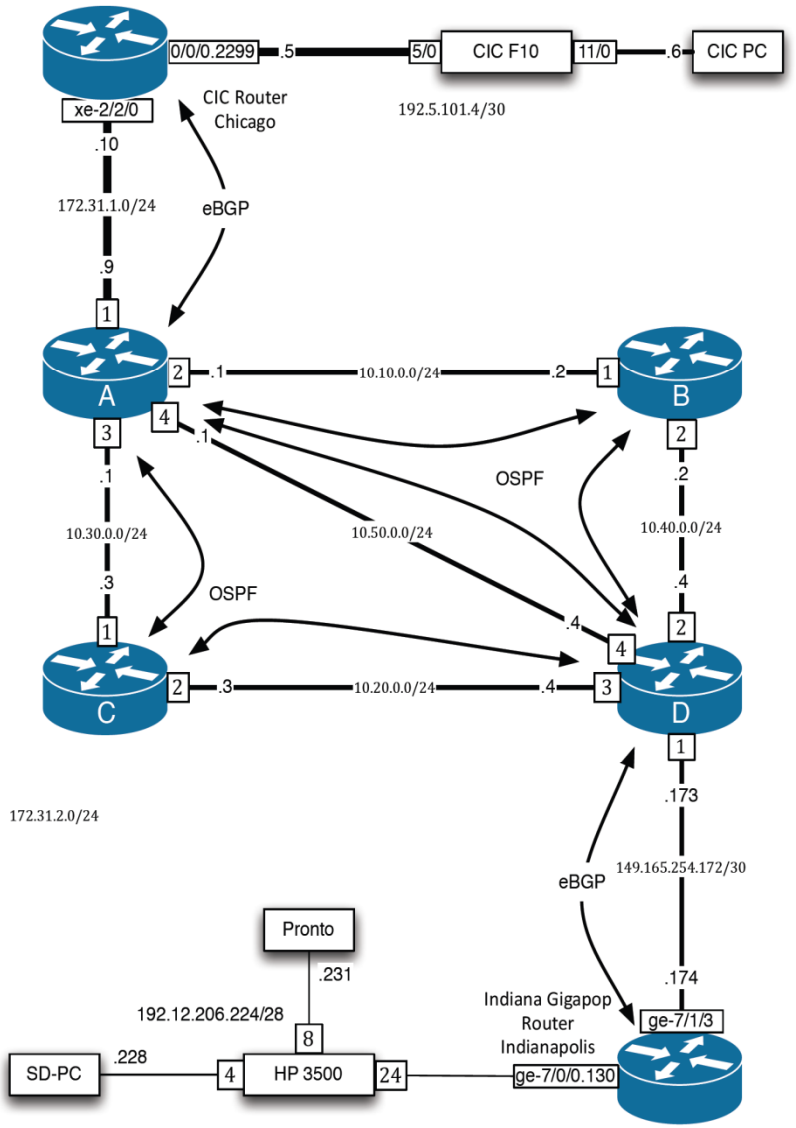
- From logical routers (akin VRFs) to single node abstractions over flexible virtual networks.
- New design choices on the distribution of the control nodes.



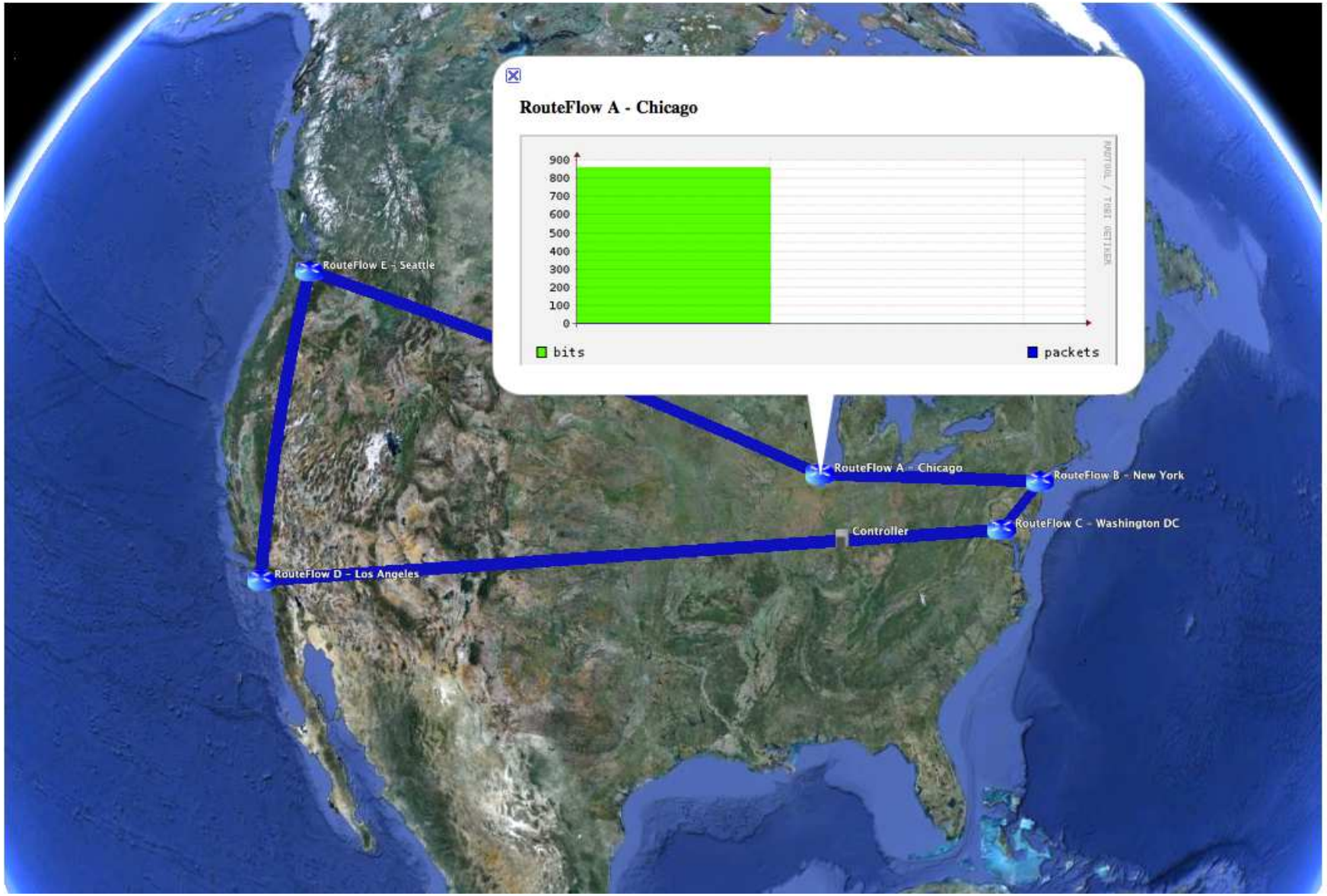

Today
IETF I2RS
Next
Software
RIB
SDN App
FIB
Virtual interfaces
RouteFlow
RouteFlow
New indirections
& intelligence point
Controllers
Controllers
Hardware
Physical ports
OF1.X
OF1.X
OF1.X
OF1.X
OpenFlow
tables & ports

InCNTRE Deployment

- 4 Virtual routers
- 10 Gig and 1 Gig connections
- 2 BGP connections to external networks
- Remote Controller
- New User Interface



RouteFlow NDDI Deployment



Demos @ ONS 2011, 2012, SC´11

Indiana University

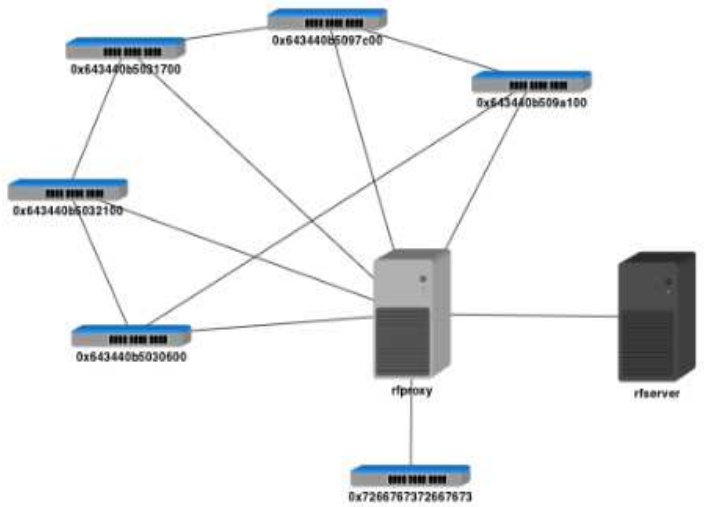
Pronto 3240/3290

+ Commercial switches from IBM, NEC, Pronto

RouteFlow

About | Support

Network | Messages | Table | Map



0x643440b5032100

Description

Manufacturer: Blade Network Technologies
Hardware descrpition: Hardware Revision: 0
Hardware PartNumber: BAC-00017-00
Manufacturing date: 11/43
Software description: Software Version 1.0.2.0 (FLASH image1), active configuration.
Serial number: MY213903NR
Datapath description: None

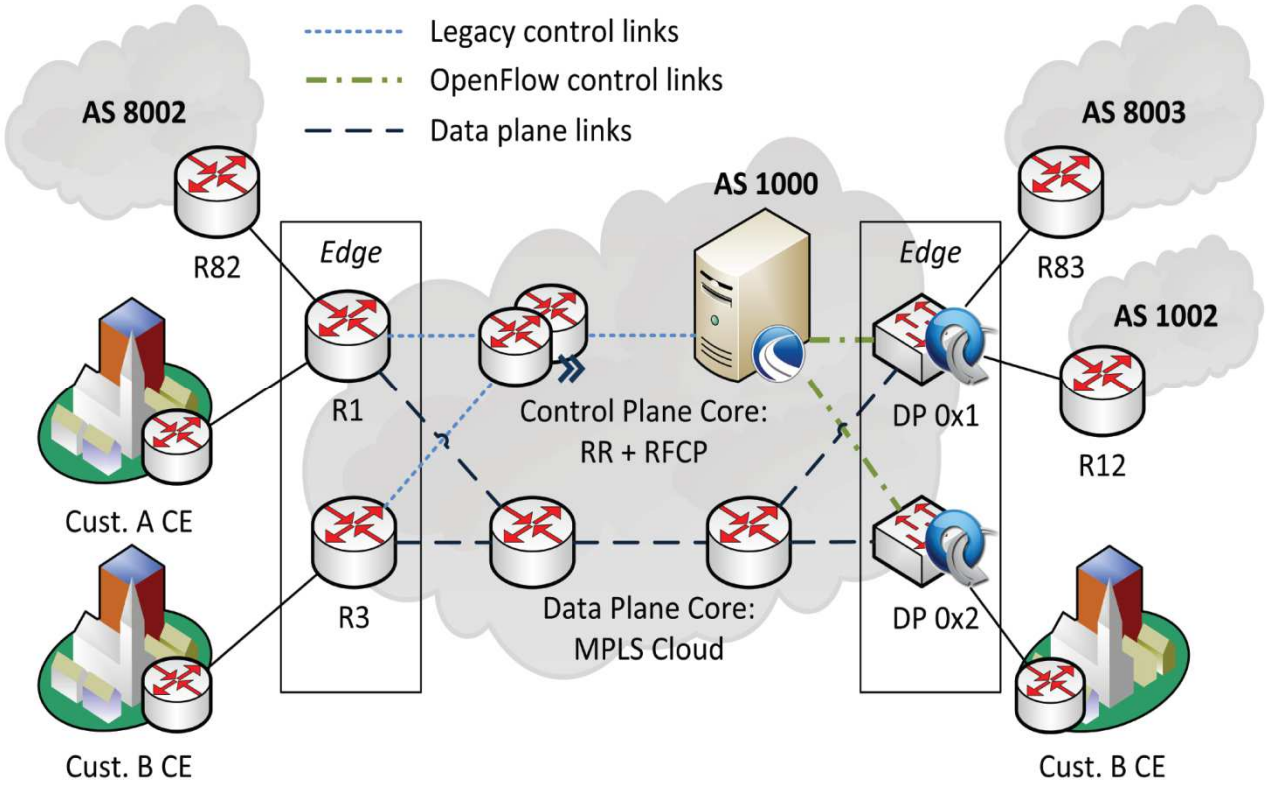
Aggregated statistics

Packet count: 603
Byte count: 48400
Flow count: 9

#	Match	Actions	Packets	Bytes
0	dl_type: 0x800, nw_proto: 89, tp_dst: 0, tp_src: 0	OUTPUT(port: 65533)	454	38228
1	dl_type: 0x800, nw_proto: 6, tp_dst: 179	OUTPUT(port: 65533)	0	0
2	dl_type: 0x800, nw_proto: 17, nw_dst: 224.0.0.9	OUTPUT(port: 65533)	0	0
3	dl_type: 0x806	OUTPUT(port: 65533)	137	9188
4	dl_type: 0x800, nw_proto: 1	OUTPUT(port: 65533)	0	0
5	dl_type: 0x800, dl_dst: 02:c1:c1:c1:c1:c1, nw_dst: 10.0.0.0	SET_DL_SRC(dl_addr: 02:c1:c1:c1:c1:c1), SET_DL_DST(dl_addr: 02:b1:b1:b1:b1:b1), OUTPUT(port: 17)	0	0
6	dl_type: 0x800, dl_dst: 02:c1:c1:c1:c1:c1, nw_dst: 20.0.0.1	SET_DL_SRC(dl_addr: 02:c1:c1:c1:c1:c1), SET_DL_DST(dl_addr: 02:b1:b1:b1:b1:b1), OUTPUT(port: 17)	0	0
7	dl_dst: 78:2b:cb:5e:cb:cd, dl_vlan: 702, dl_src: 78:2b:cb:48:ff:44, in_port: 63	OUTPUT(port: 18)	12	984
8	dl_dst: 78:2b:cb:48:ff:44, dl_vlan: 702, dl_src: 78:2b:cb:5e:cb:cd, in_port: 18	OUTPUT(port: 63)	0	0

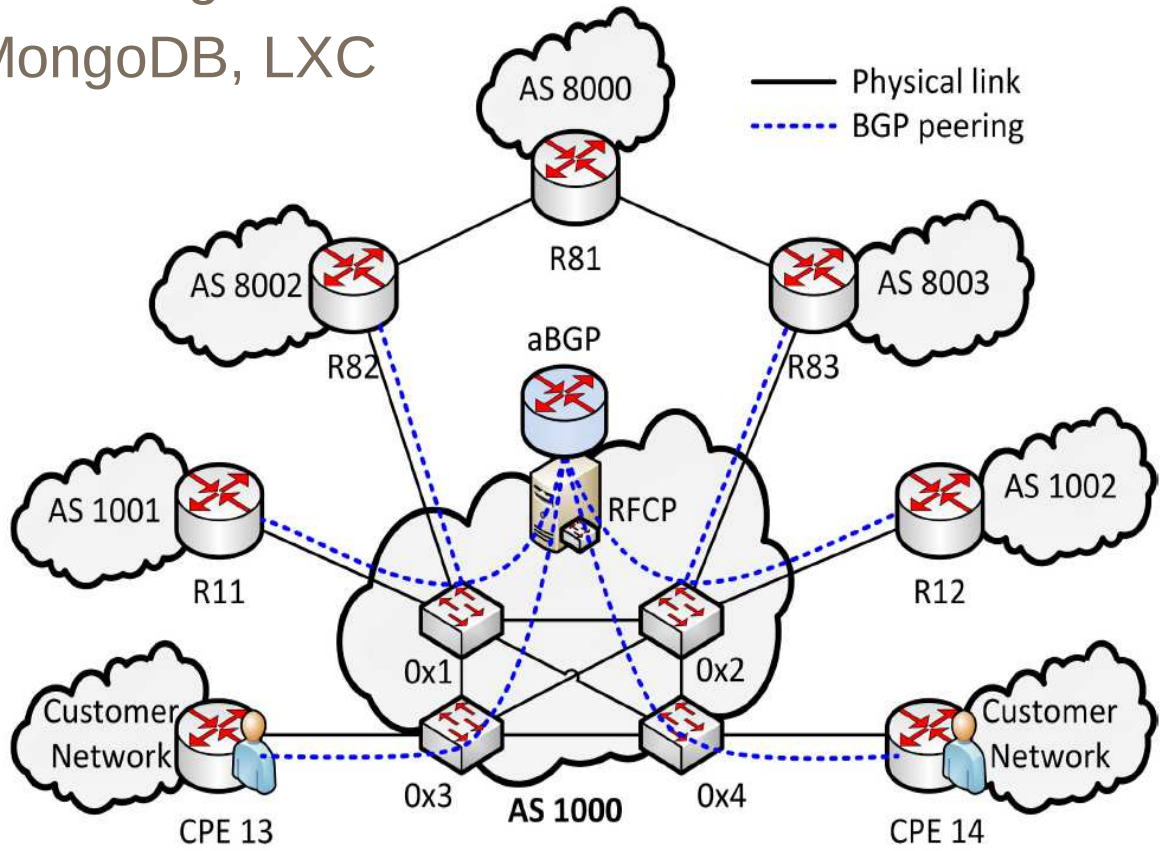
Controller-Centric Hybrid Networking

- 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



Prototyped: Aggregated BGP routing service

- 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
- NOX, MongoDB, LXC



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]
- OpenFlow v1.2 and v1.3 [w/ Ericsson]
- Open Label Switched Router [OSRF; Google]
- Multi-path, Fast-ReRoute, BGP-Sec, IPv6, ... [YOU?]

RouteFlow

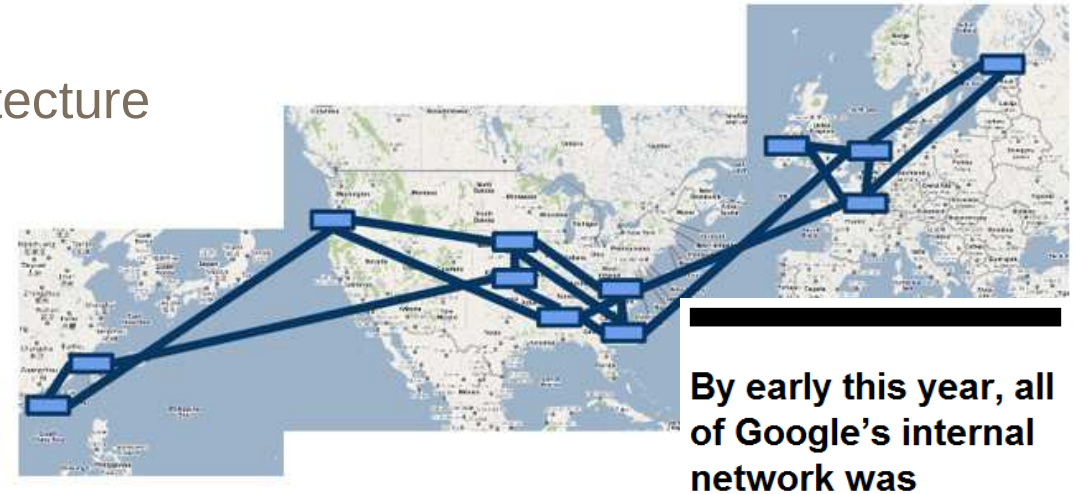
CPqD UNICAMP UNIRIO InCNTRE ufscar NTT Communications Google Open Source Routing

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

Case: Software-Defined WAN @ Google

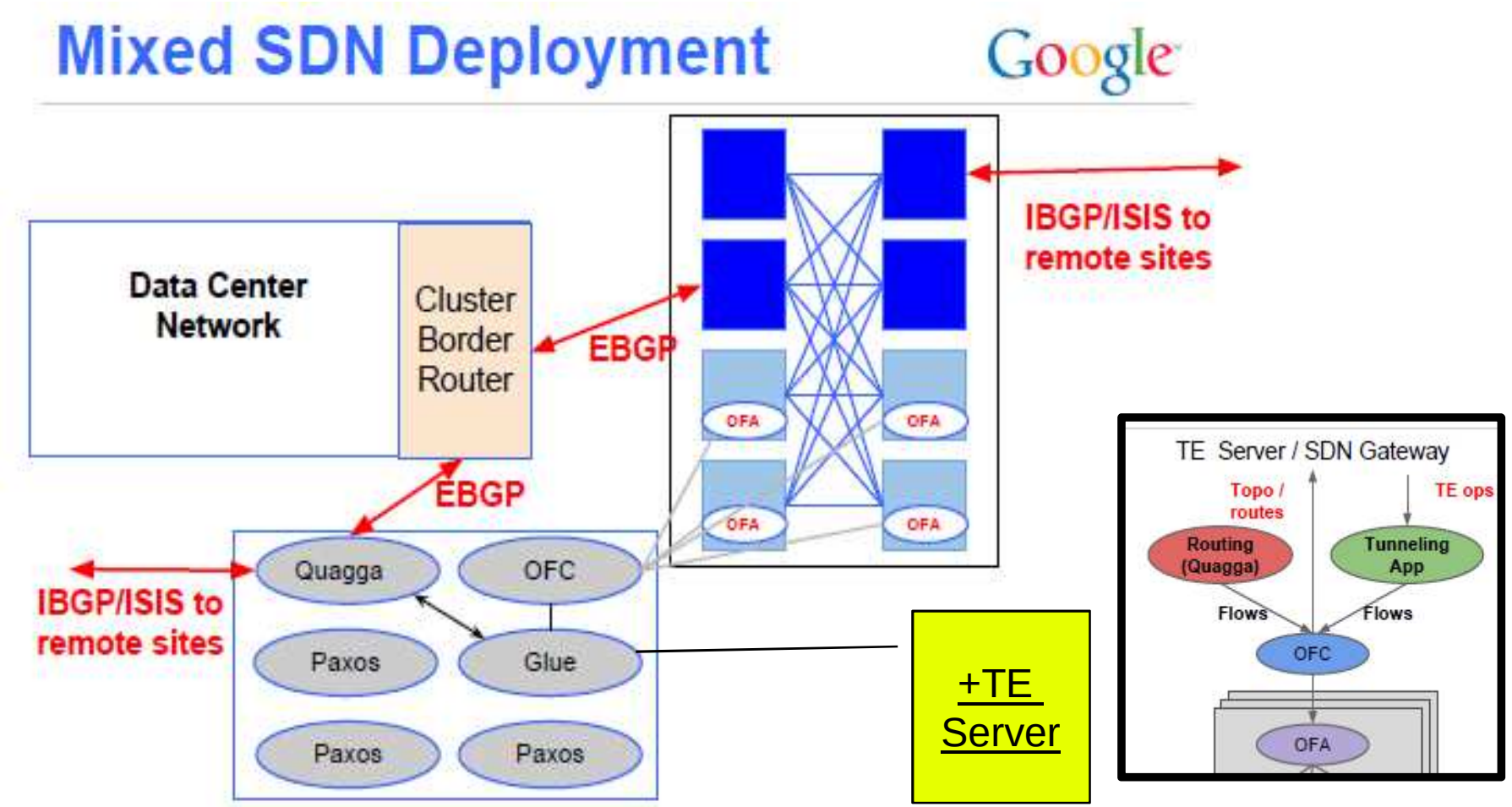
- Inter-DC WAN architecture
  - OpenFlow-based



By early this year, all of Google's internal network was running on OpenFlow. 'Soon we will able to get very close to 100 percent utilization of our network,' Hölzle says.

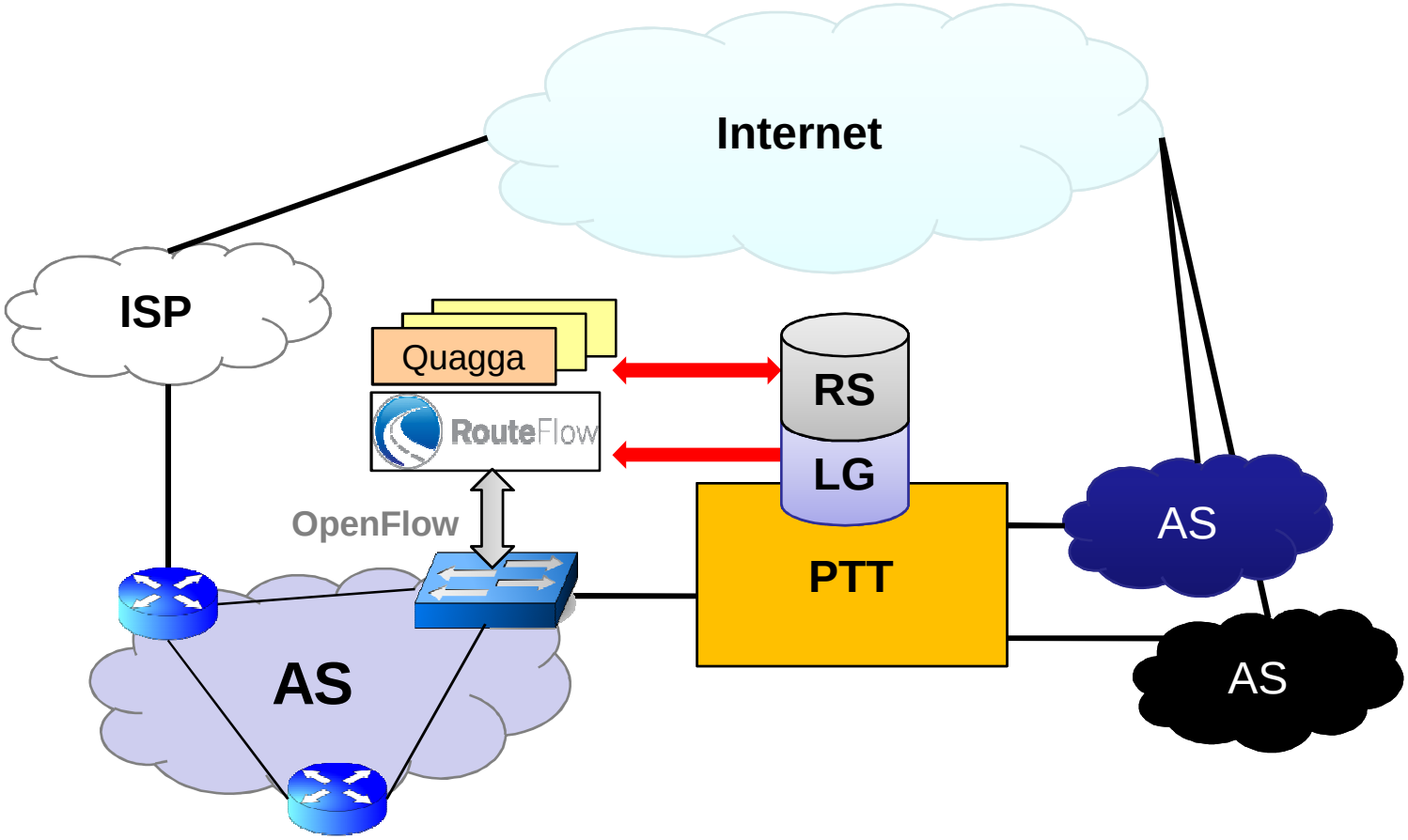
- Benefits:
  - $$$
  - Efficiency (100% utilization)
  - Planning big data moves
  - Simulate everything offline
  - Products can be rolled out more quickly
  - Moving the control plane out of the box:
    - Networking equipment is simpler and enduring, requiring less labor
  - Makes network management much easier.

Google Software Defined WAN Architecture



Source: vahdat-wed-sdnstack.pdf @ONS12

RouteFlow em rede cliente do PTT



Preview of ONS´13 Demos (1/3): Cardigan

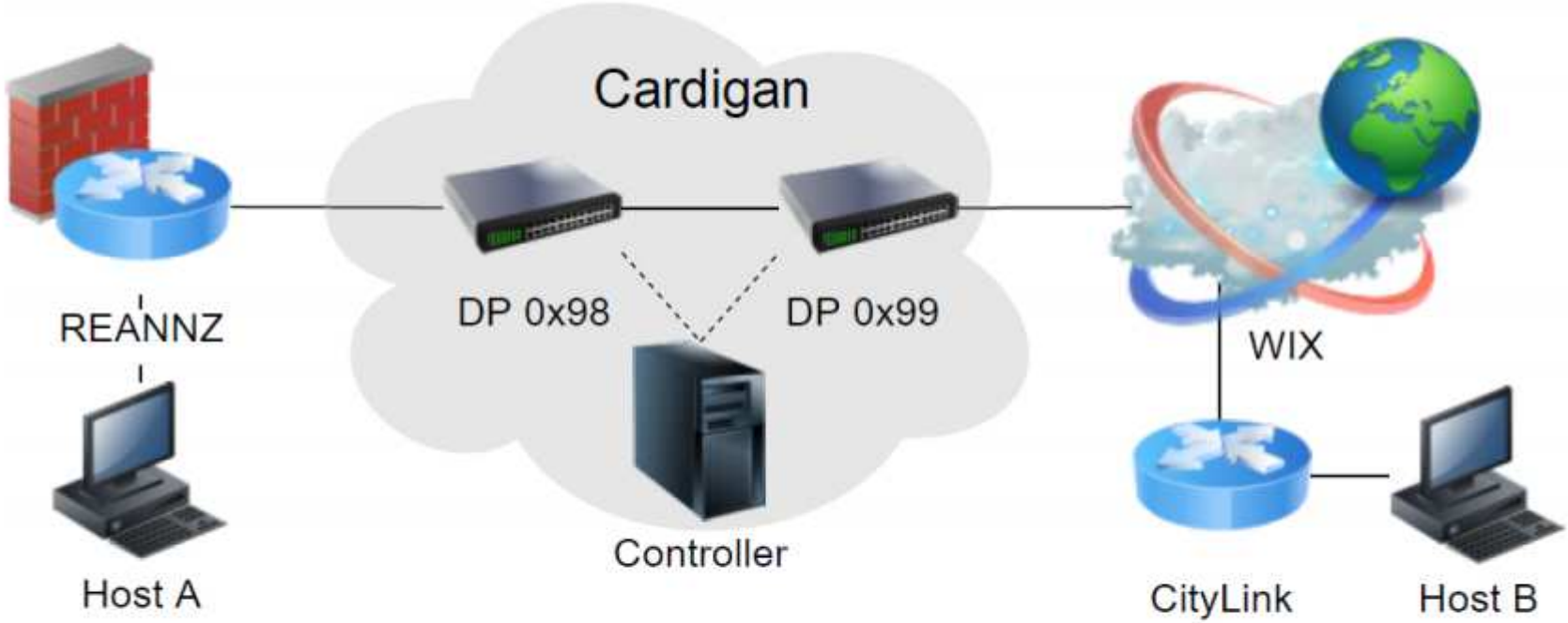


Figure 3 - Distributed IX Router

Production traffic!

Preview of ONS´13 Demos (2/3): FIBRE

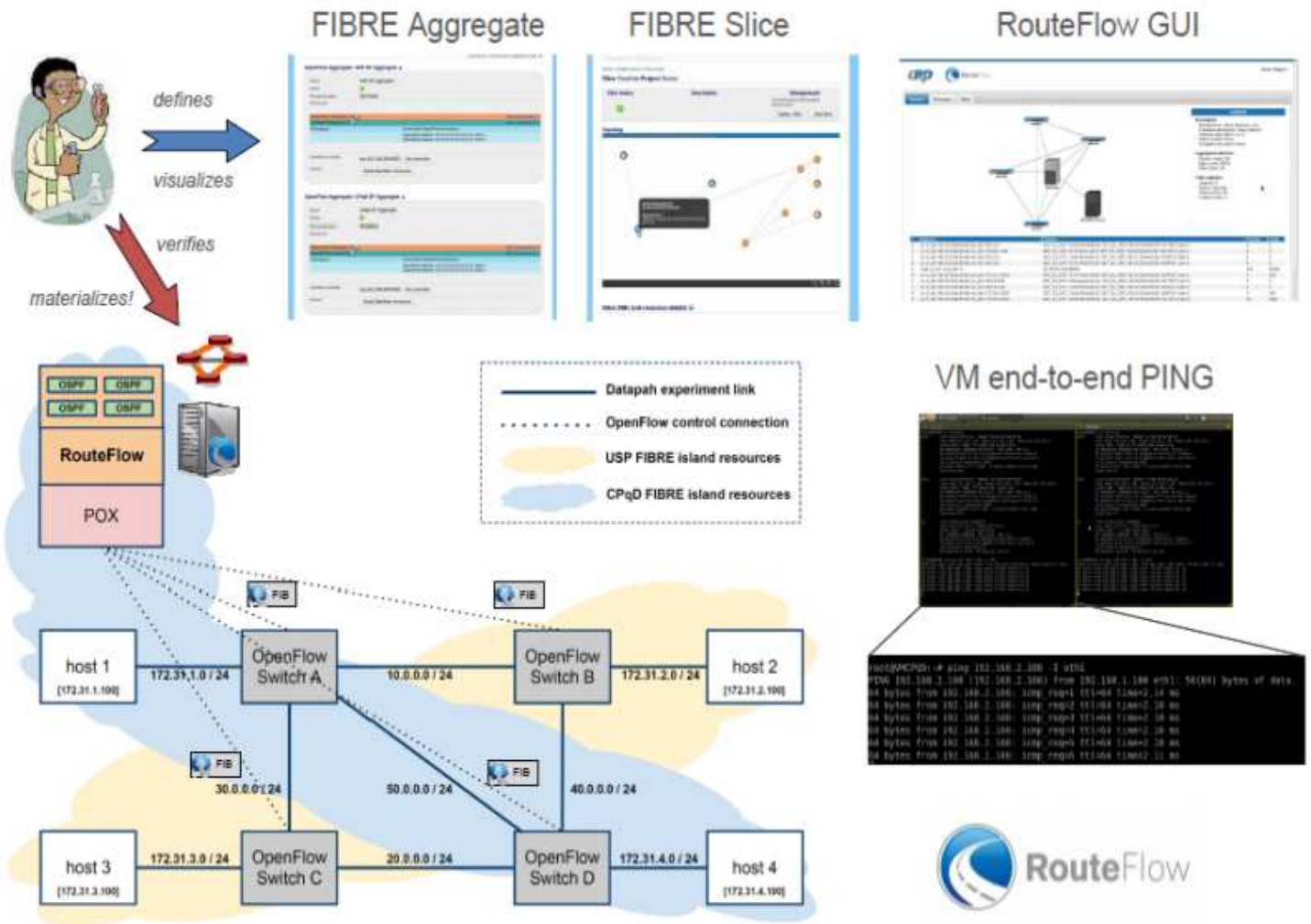


Figure 1 - IP-Routed Network on Demand

Preview of ONS´13 Demos (3/3): IVR

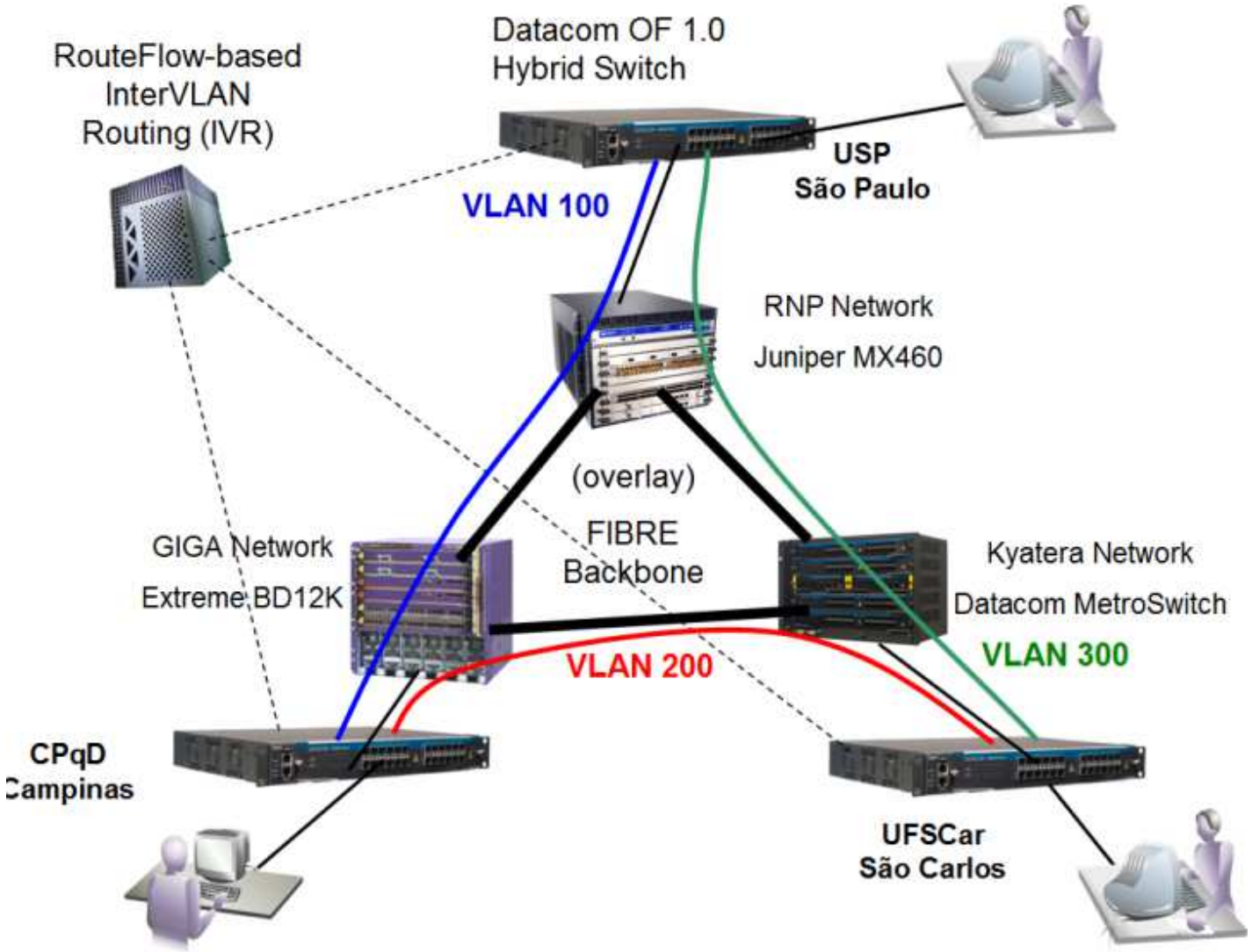
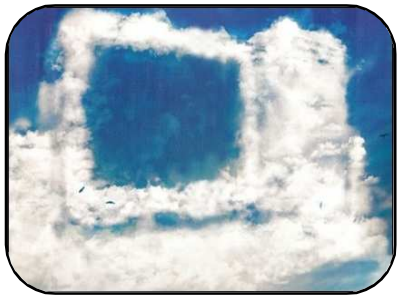
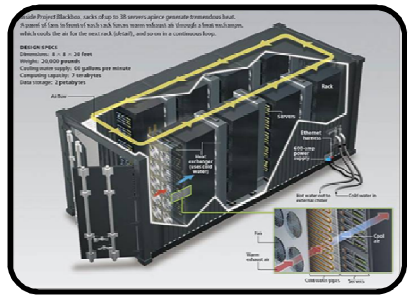


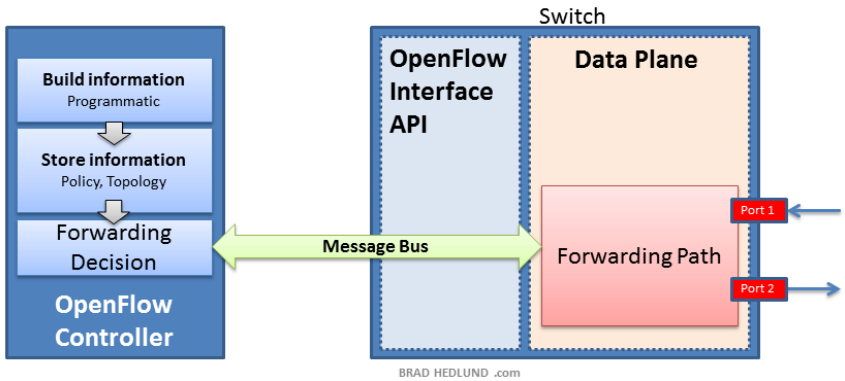
Figure 2 - InterVLAN Routing

Conclusões

- Networking for the Cloud
- Software-Defined Networking
  - Foco da inovação e da percepção de valor migra:
    - do **Hardware**: proprietário, box com SW embarcado
    - para **Software**: do estado-da-arte, que utiliza plataformas HW padrão de mercado
  - Software: Atuação de novos players nas camadas de Controle e de Serviços (industria nacional de SW)
    - permitindo a formação de um eco-sistema multi-vendor (flexibilidade, agilidade) e open source (?)
    - acelerando time-to-market para customizações de novas funções (serviços)

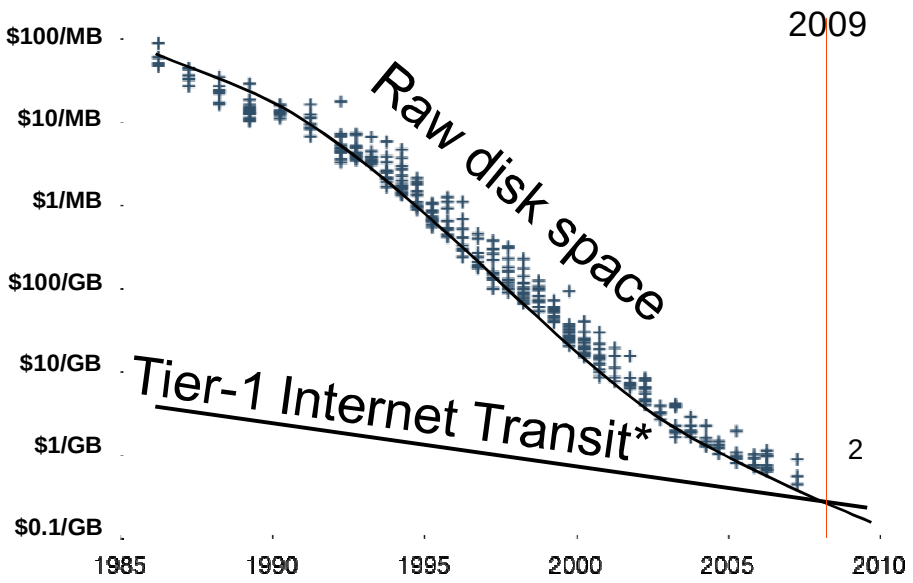


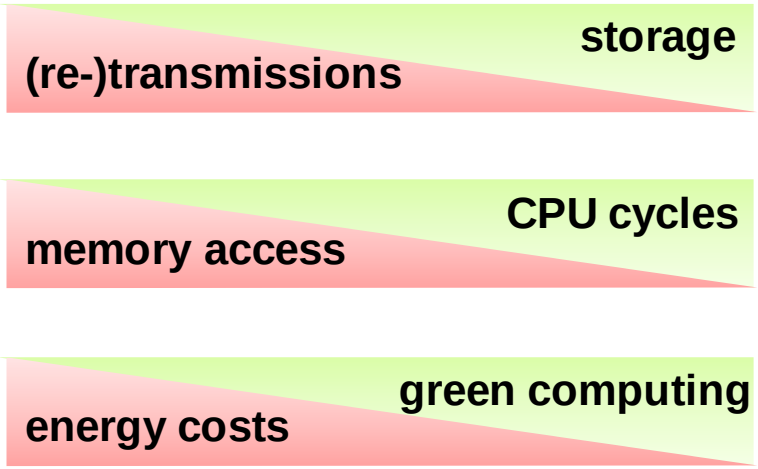




Network economics & Future Internet

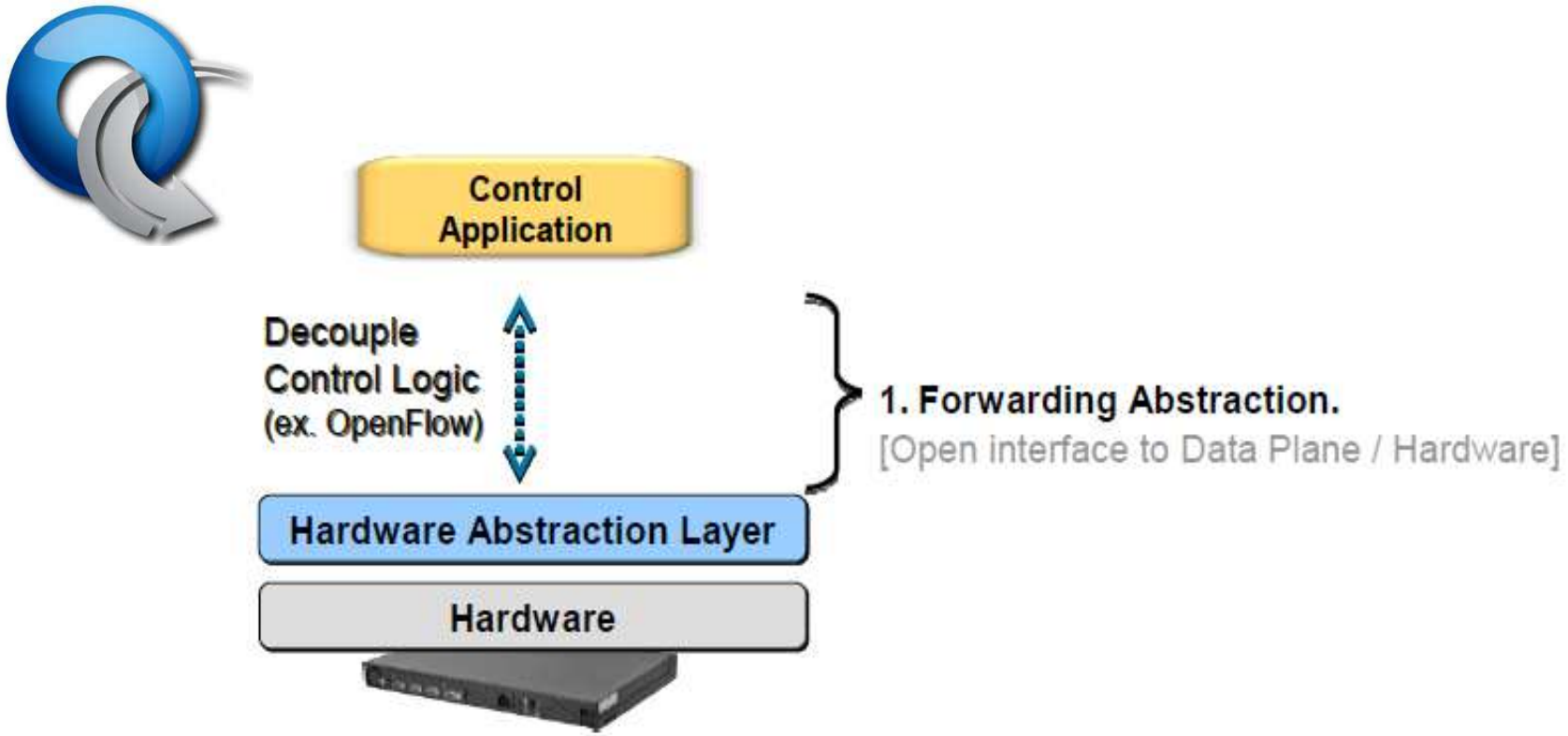
- **Data Centers are like Factories[1]**
  - Number 1 Goal: Maximize useful work per dollar spent
- **And the future network of networks?**
  - Incentives for re-architeturing the Internet? DC-driven incentives???
- **Think like an economist/industrial engineer as well as a computer scientist**
  - Understand where the dollar costs come from
  - Use computer science to reduce/eliminate the costs / complexity





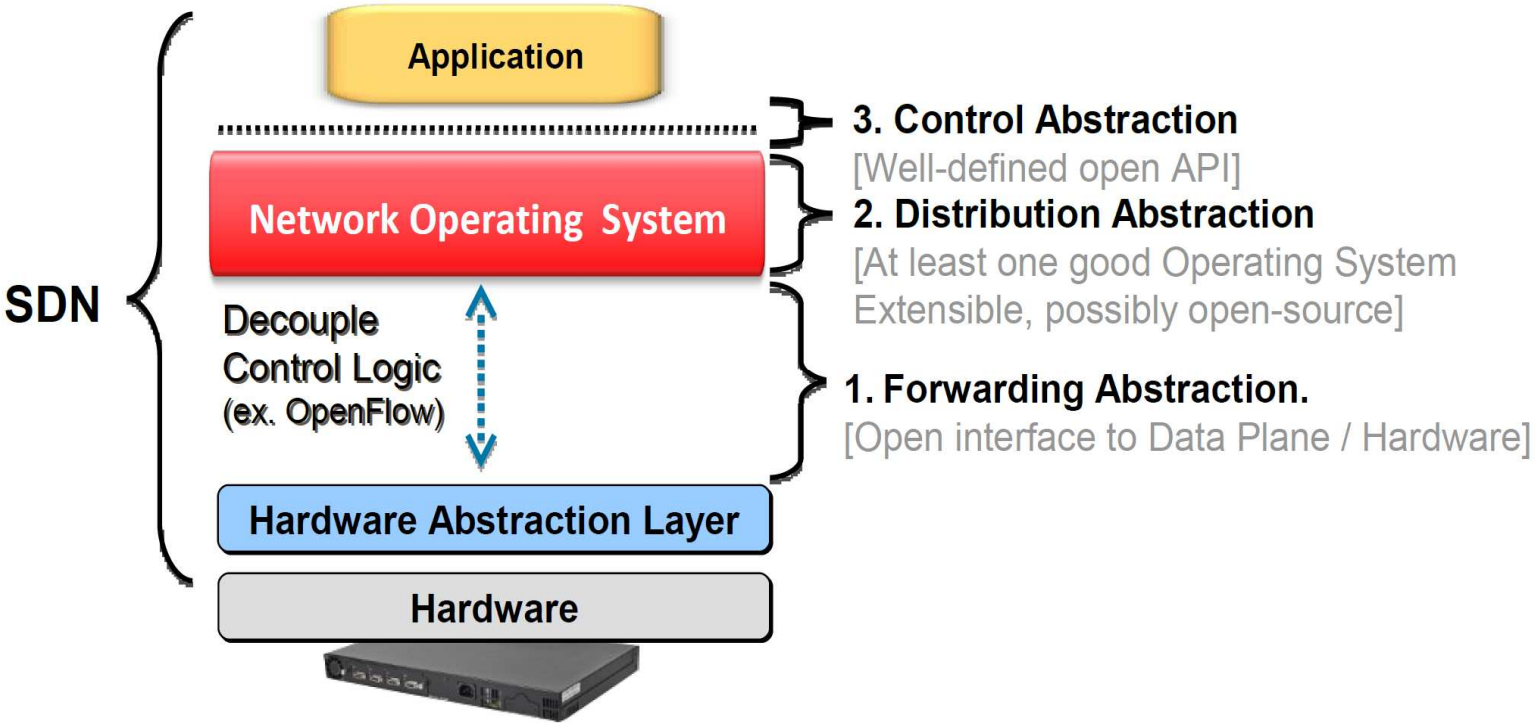
[1] cf. Greenberg SIGMETRICS tutorial
[2] Preliminary data [Nikander'09]

OpenFlow é um protocolo, SDN é uma arquitetura



- Software Defined Networking (SDN) refactors the relationship between network devices and the software that controls them.
- Open interfaces to network switches (e.g. OpenFlow) enable more flexible and predictable network control, and they make it easier to extend network function. [HotSDN´12 CFP]

OpenFlow é um protocolo, **SDN é uma arquitetura**



- Software Defined Networking (SDN) refactors the relationship between network devices and the software that controls them.
- Open interfaces to network switches (e.g. OpenFlow) enable more flexible and predictable network control, and they make it easier to extend network function. [HotSDN´12 CFP]

Christian Esteve Rothenberg, PhD
Diretoria de Redes Convergentes (DRC)
esteve@cpqd.com.br

Obrigado!

Perguntas?

www.cpqd.com.br

REFERENCES

- A. Greenberg and et al., “The cost of a cloud: research problems in data center networks.” *SIGCOMM CCR.*, 2009.
- A. Greenberg and et al.,“Monsoon: Towards a Next Generation Data Center Architecture: Scalability and Commoditization”
- A. Greenberg and et al.,“VL2: A Scalable and Flexible Data Center Network” , SIGCOMM 09
- R.Niranjan et al., “PortLand: A Scalable Fault-Tolerant Layer 2 Data Center Network Fabric” , SIGCOMM 09
- “BCube: A High Performance, Server-centric Network Architecture for Modular Data Centers”, SIGCOMM 09
- Wu et al. “MDCube: A High Performance Network Structure for Modular Data Center Interconnection”, CoNext09.
- Benson et al., “Understanding Data Center Traffic Characteristics”, WREN 09
- Vaquero et al., “Break in the Clouds: Towards a Cloud Definition”
- Valancius et al. “Transit Portal: Bringing Connectivity to the Cloud”

REFERENCES

- Guo et al., “DCell: A Scalable and Fault-Tolerant Network Structure for Data Centers”
- Joseph et al. “A Policy-aware Switching Layer for Data Centers”
- Al-Fares et al. “A Scalable, Commodity Data Center Network Architecture”
- Nano Data Center, EU FP7 NADA, www.nanodatacenters.eu
- Laoutaris et al., “ECHOS: Edge Capacity Hosting Overlays of Nano Data Centers”
- Prachi Patel-Predd et al., “Cutting the Power in Data Centers”,
- Qureshi “Cutting the Electric Bill for Internet-Scale Systems” SIGCOMM 09
- Christian Esteve Rothenberg, "Re-architected Cloud Data Center Networks and Their Impact on the Future Internet". In New Network Architectures, Studies in Computational Intelligence, 2010, Volume 297/2010, 179-187,
- Christian Esteve Rothenberg et al.,"Revisiting Routing Control Platforms with the Eyes and Muscles of Software-Defined Networking." In Hot Topics in Software Defined Networking (HotSDN), Helsinki, Finland, Aug 2012.
- Marcelo R. Nascimento et al., . "RouteFlow: Roteamento Commodity Sobre Redes Programáveis." In RB-RESD - Revista Brasileira de Redes de Computadores e Sistemas Distribuídos, vol. 4, no. 2, Dez 2011 - 1983-4217

Material

Credits

- Sudipta Sengupta, Slides on “issues with conventional DC designs”, from “Oblivious Routing and Applications”, Tutorial at IEEE ICC 2009.
- Guo et al, Slide on “Container-based modular DC”
- A. Greenberg and D.A. Maltz, „What Goes into a Data Center”, SIGMETRICS 2009 Tutorial, Image on slide on “Impacts of the cloud on the FI?”.

Images

- Switch slide 10, http://lcg.web.cern.ch/LCG/lhcgridfest/partners.htm
- Interconnection hw, slide 4, http://www.microsoft.com/presspass/events/msrtechfest/images/LowPowerProcessors_print.jpg
- OpenFlow Switch, Brad Hedlund, http://bradhedlund.com/2011/04/21/data-center-scale-openflow-sdn/

Examples of Optical R&D activities at CPqD

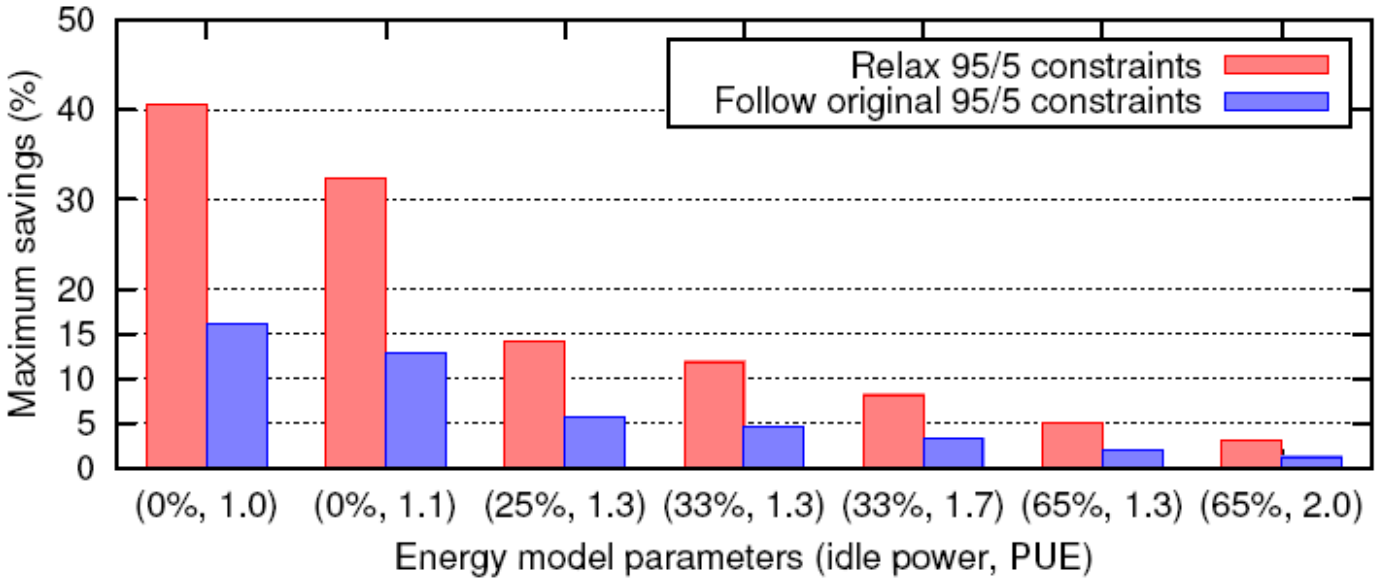
- High-speed transmission
  - 100Gb WDM transmission (DP-QPSK with coherent detection)
  - 1Tb transmission (Optical OFDM, xxQAM)
- Elastic optical networks
  - Subsystems (e.g., hybrid EDFA/Raman adaptive amplification, dynamic optical spectrum equalizer, low-cost multiparameter optical monitor, rate-adaptive tunable transponder)
  - Optimization algorithms (e.g., distance, power consumption)
  - Intelligent control plane
- Access
  - XGPON, WDM-PON
  - RoF (GPON as WiMax/WiFi backhaul)
  - Programmability and virtualization of the optical layer

5000km 32x100Gbps DP-QPSK setup

Cost-Aware Internet Routing

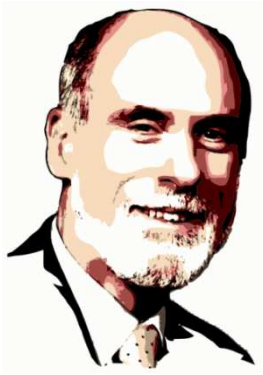
40% savings of a cloud computing installation's power usage by dynamically re-routing service requests to wherever electricity prices are lowest on a particular day, or perhaps even where the data center is cooler.

From "Follow the energy price! " to "Follow the wind, the sun or the moon!"



[Qureshi et al, "Cutting the Electric Bill for Internet-Scale Systems", SIGCOMM´09]

Cloud



"The Cloud represents a new layer in the Internet architecture and, like the many layers that have been invented before, it is an open opportunity to add functionality to an increasingly global network" - Vint Cerf, 2009 [1]

"History doesn't repeat itself, but it does rhyme." - Mark Twain

Cloud Initiatives that have an analogue in the Internet's past [2]:

- The rising importance of *academia*.
- Increasing interest in *interoperability* among cloud vendors.
  - Today's clouds like network islands before IP
- *Carrier* interest in new service opportunities.

[1] http://googleresearch.blogspot.com/2009/04/cloud-computing-and-internet.html

[2]

Switch 10Gb OpenFlow 1.0 spec

1o. Switch OpenFlow desenvolvido na América Latina

OpenFlow

Search

News | Research | About

« Pantou : OpenFlow 1.0 for OpenWrt now available! (alpha relea ... OpenFlow at GEC9 »

CPqD Ports OpenFlow to New Platform

October 13th, 2010, dtalayco in OpenFlow Blog

CPqD is a private non-profit Brazilian R&D foundation. Recently they announced the first switch in South America to support the OpenFlow 1.0 specification. The switch uses Broadcom L2/L3 silicon with 24 x 1Gb ports and 2 x 10Gb ports. It has a high performance CPU running the Indigo-beta-4 release from Stanford.

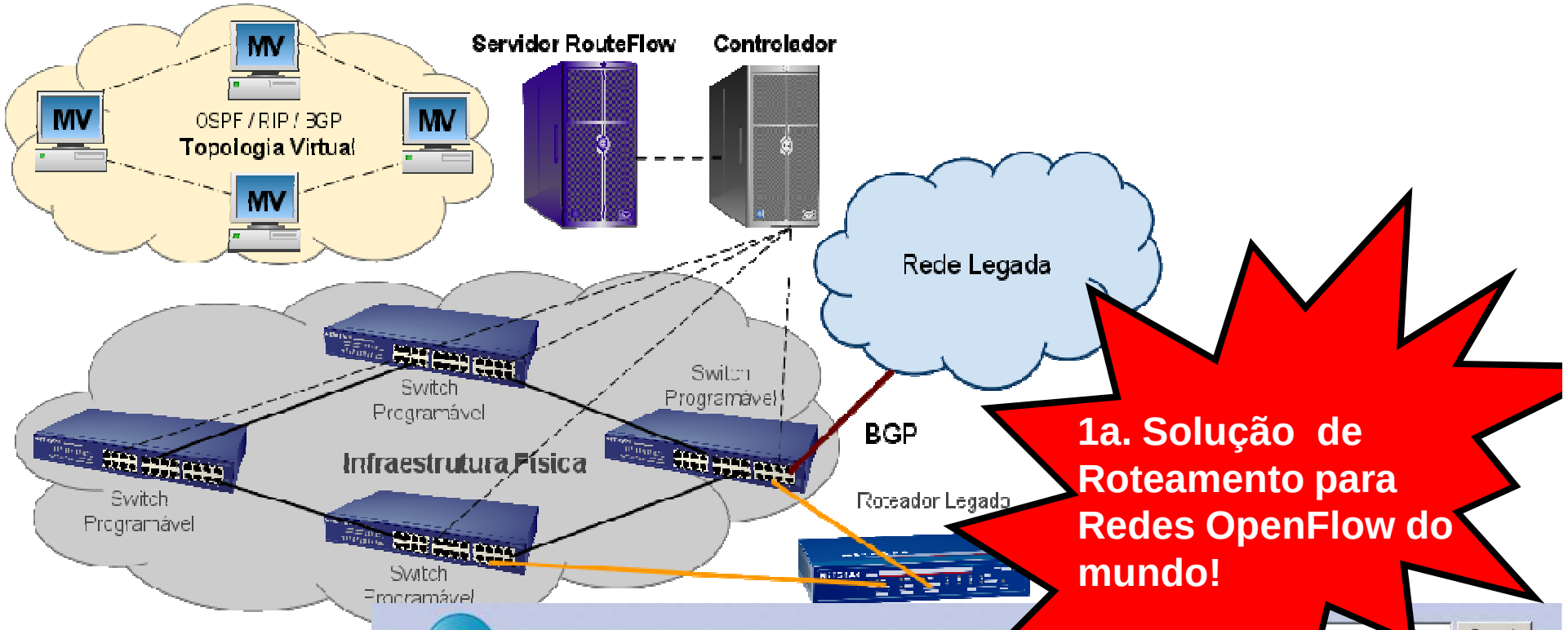
Tens of these OpenFlow switches will be deployed in Project GIGA´s High-speed Experimental Network, an IP/Ethernet/WDM network testbed run by CPqD and RNP (Brazilian NREN). Today the GIGA network connects 66 research labs at multigigabit per second rates in the southeast region of Brazil, but will soon be upgraded to support 100Gb/s bit rates, using technology developed in project GIGA, and expanded to all the other regions of the country, using RNP´s network resources.

This large-scale OpenFlow infrastructure will be fundamental to the national initiative on Future Internet that CPqD and RNP, amongst others, are leading, as well as to support collaborative experiments related to projects GIGA and GENI.

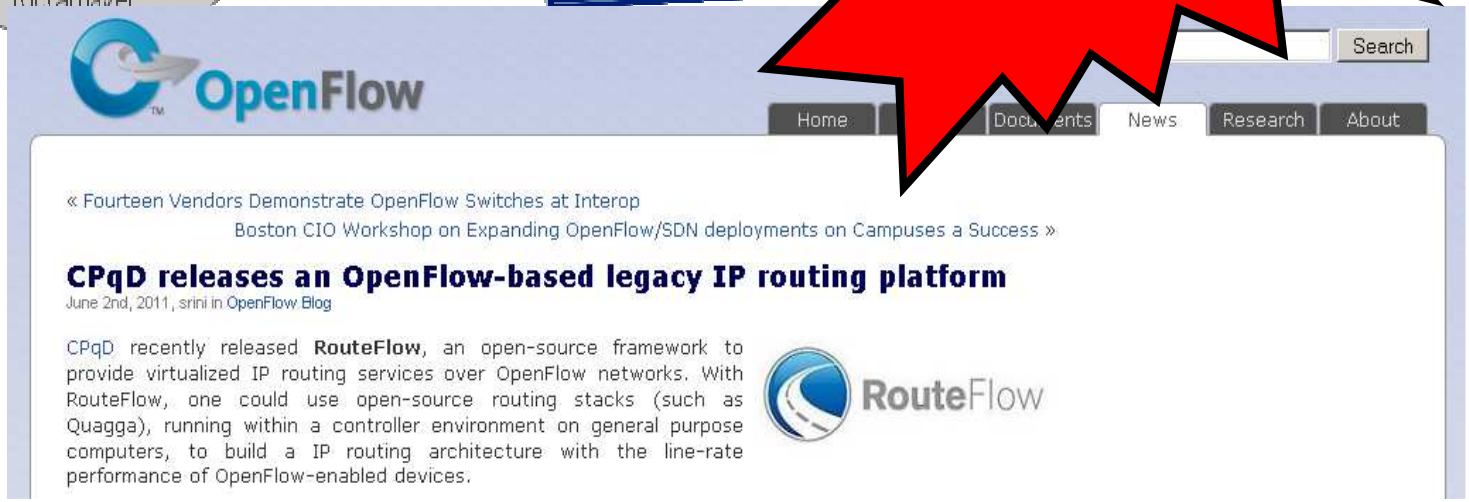
The development of the OpenFlow switch and the development of an IP routing stack solution that runs outside the switches on top of NOX (stay tuned!) are under the Future Internet umbrella of CPqD´s current R&D program, which includes a number of projects, such as Project GIGA.

CPqD was the R&D branch of Brazil´s telephony monopoly system until 1997, when the whole system was privatized. Since then CPqD is a private foundation with the goal of bridging the gap between university research and product development, helping (mainly) local companies to innovate and compete in the market. Today CPqD has more than 30 years of existence and 1200 employees carrying out R&D activities on various ICT sectors.

RouteFlow



1a. Solução de Roteamento para Redes OpenFlow do mundo!



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