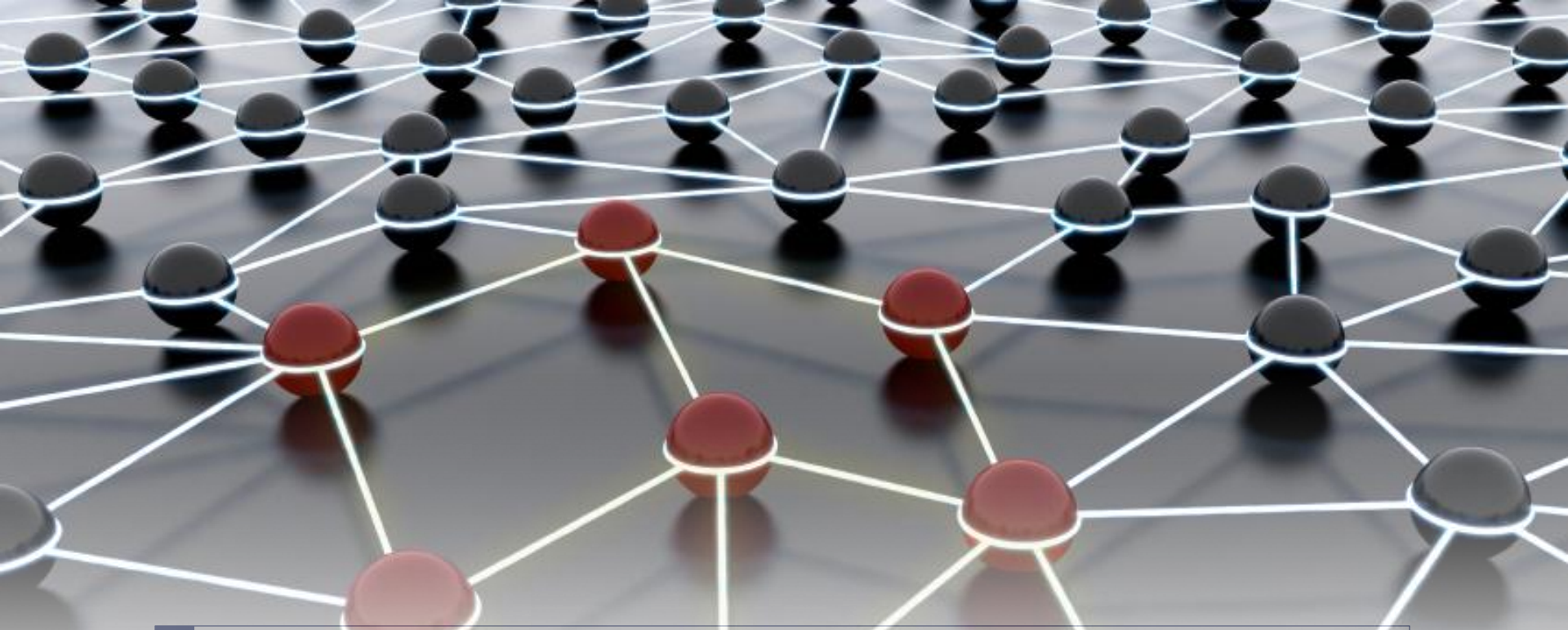


Software-Defined Networking

Robert Keahey

IEEE-Consultants' Network of Silicon Valley – 2012/10/16

SummaLogic

CPLANE

Agenda

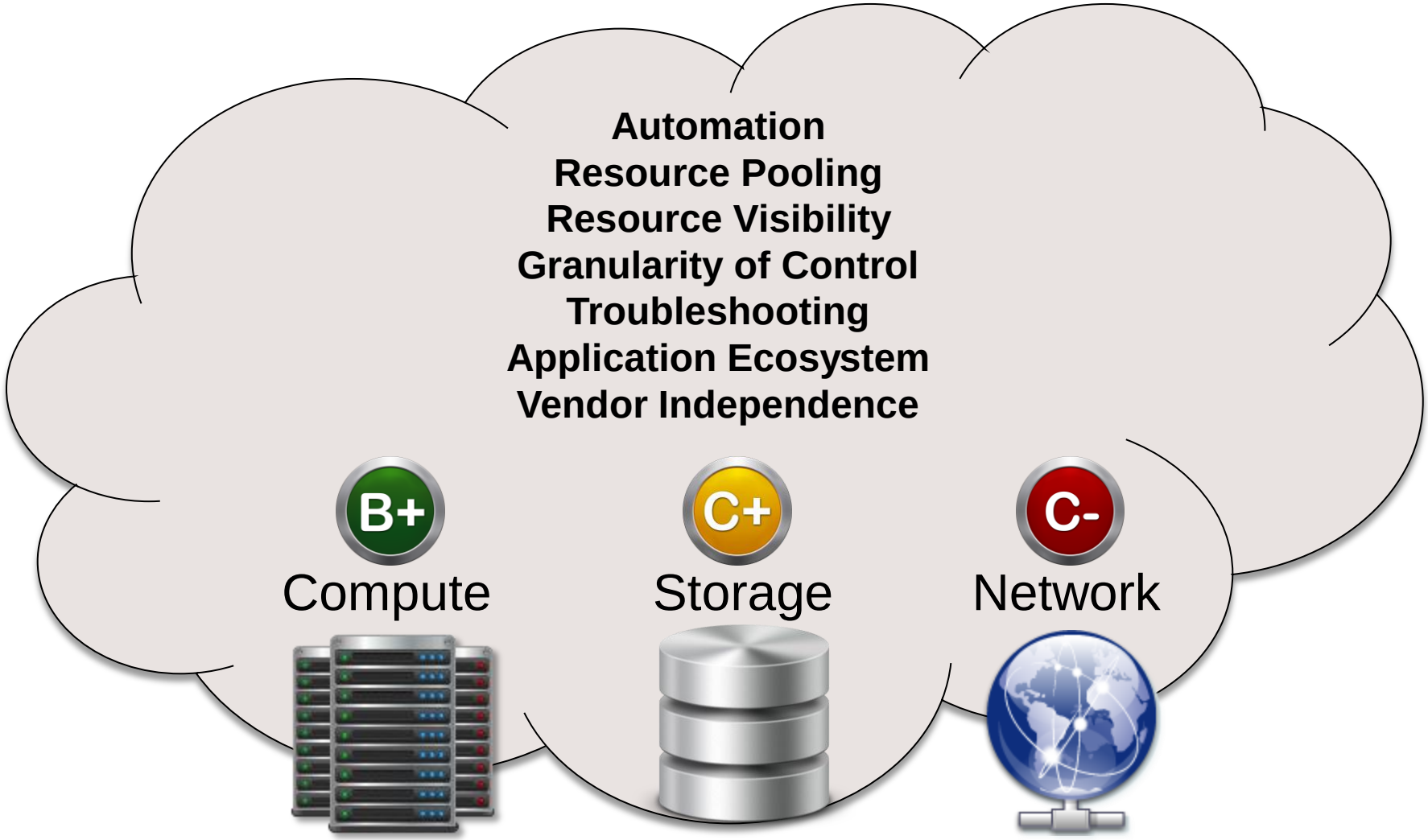
- SDN
  - Trends, Drivers, Models, Use Cases
- OpenFlow
- Network Virtualization
- Essential things to know

The Next (R)evolution?

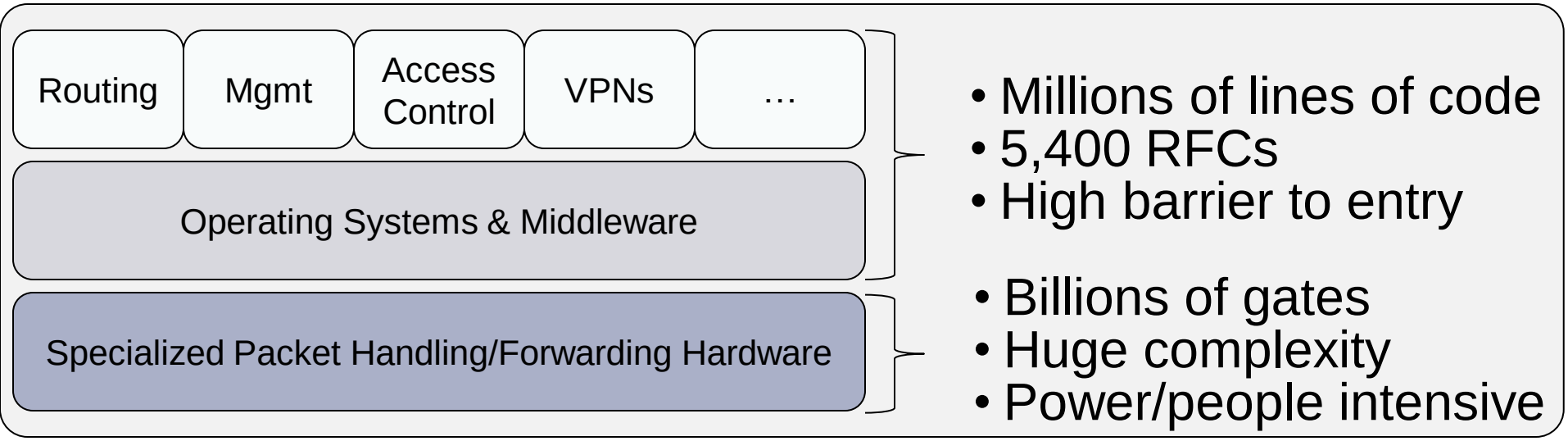
- What is it?
- When will it be here?
- Who will use it?
- Why will they use it?
- Where will they use it?

How big <u>is</u> this thing?

Cloud Enablement Report Card



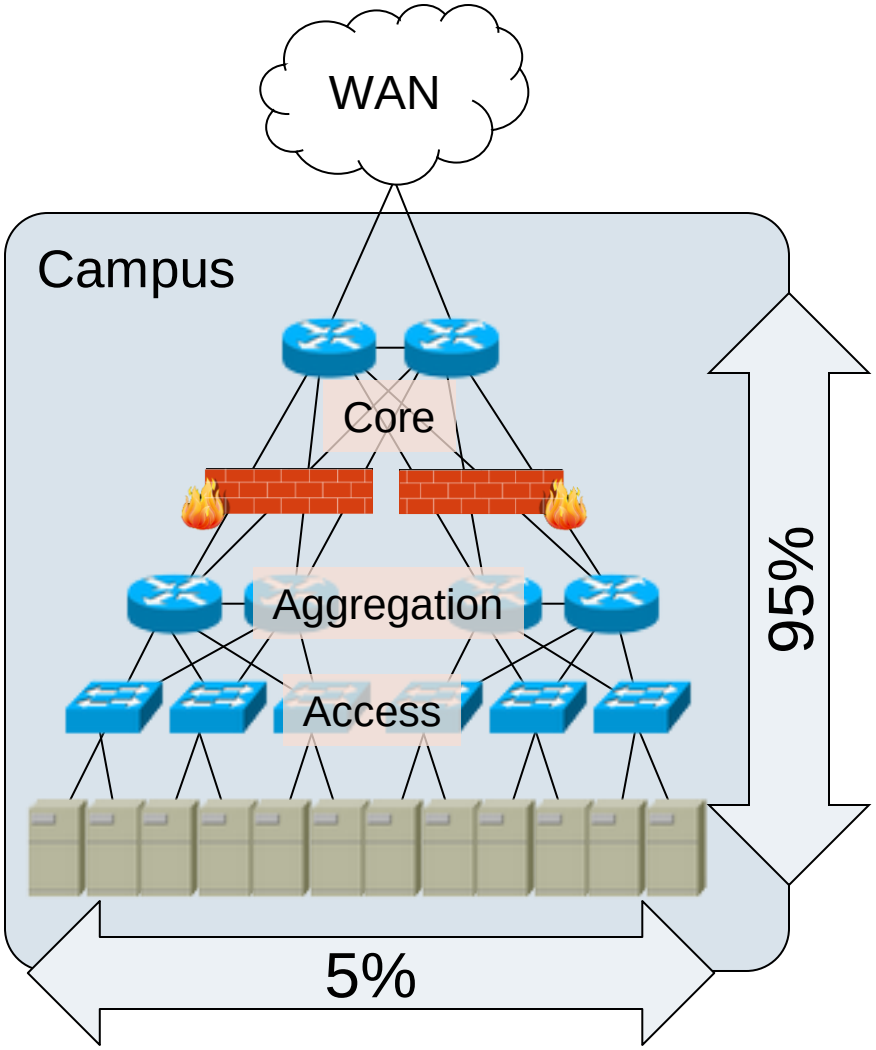
Looking Back – Networking 2007



- Many complex functions embedded into the infrastructure
  - OSPF, BGP, Multicast, NAT, TE, MPLS, Firewalls, …
  - Redundant layers/services
  - Unique “differentiation”
- Mainframe mentality industry
- Functionality → standards → hardware → nodes

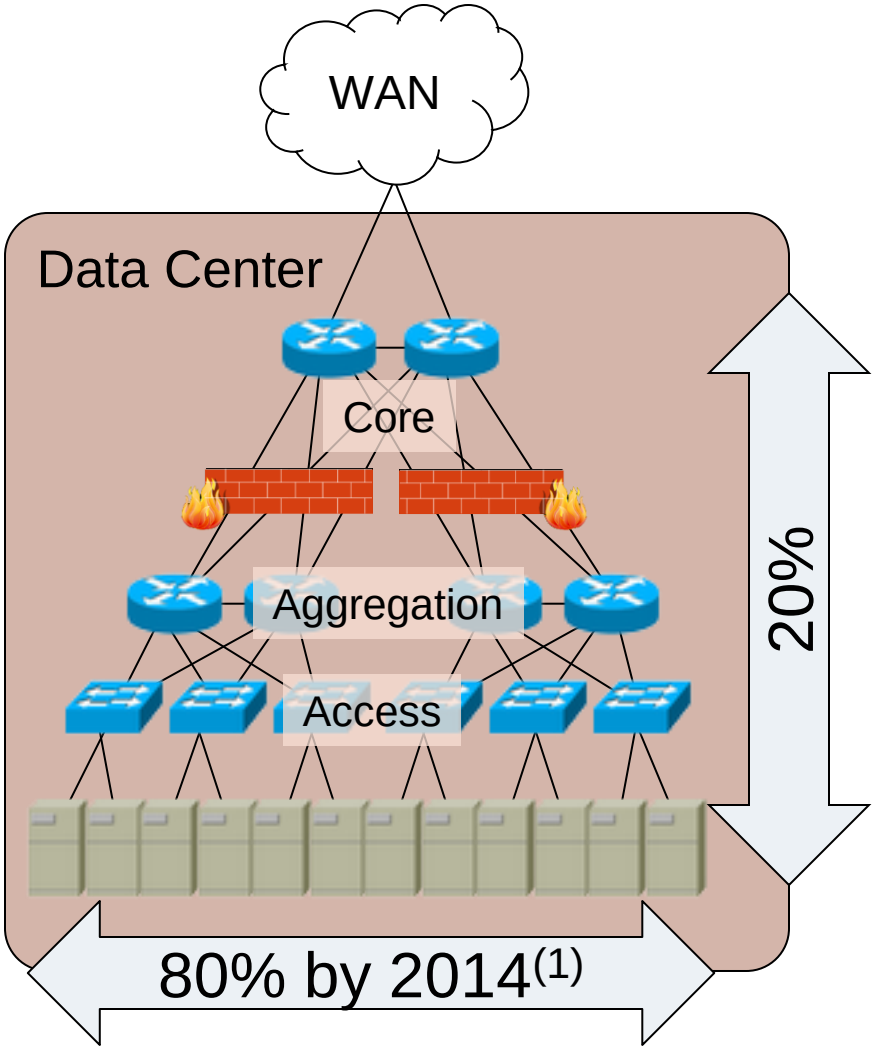
Source: Adapted from ONS12 presentation by Brandon Heller, et al

Compounding the problem…



- Evolved campus Ethernet model into tree structure
  - Core
  - Aggregation
  - Access
- Most (95%) of traffic is “north-south”
- Segregated campus networks at Access to avoid spanning tree problems

Compounding it further…



- Applied same model to the data center
- Different traffic patterns
  - Majority “east-west”
- Different performance needs
  - Lossless storage traffic
  - Low latency, high bandwidth
- Different service needs to support virtual compute model
  - Static to dynamic
  - Multi-tenancy
  - Workload management

Source: (1) Gartner Synergy Report

And even further…

- Significant gap exists
  - Architectural
  - Operational
  - Organizational
- Assumption that physical fabric is “up and running”
  - *YoYo* mindset
- Extends to intra- and inter-data center deployments
  - Metro
  - WAN
  - Carrier

Additional market drivers

Video and Mobility are transforming business communications

Up to

10X

Increase in network capacity to support new wave of business video applications

INCREASE IN BANDWIDTH REQUIREMENTS

At least

50 Billion

Devices will connect to wireless networks by the year 2020

UNIFIED WIRED AND WIRELESS CAMPUS NETWORKS, IT CONSUMERIZATION

More than

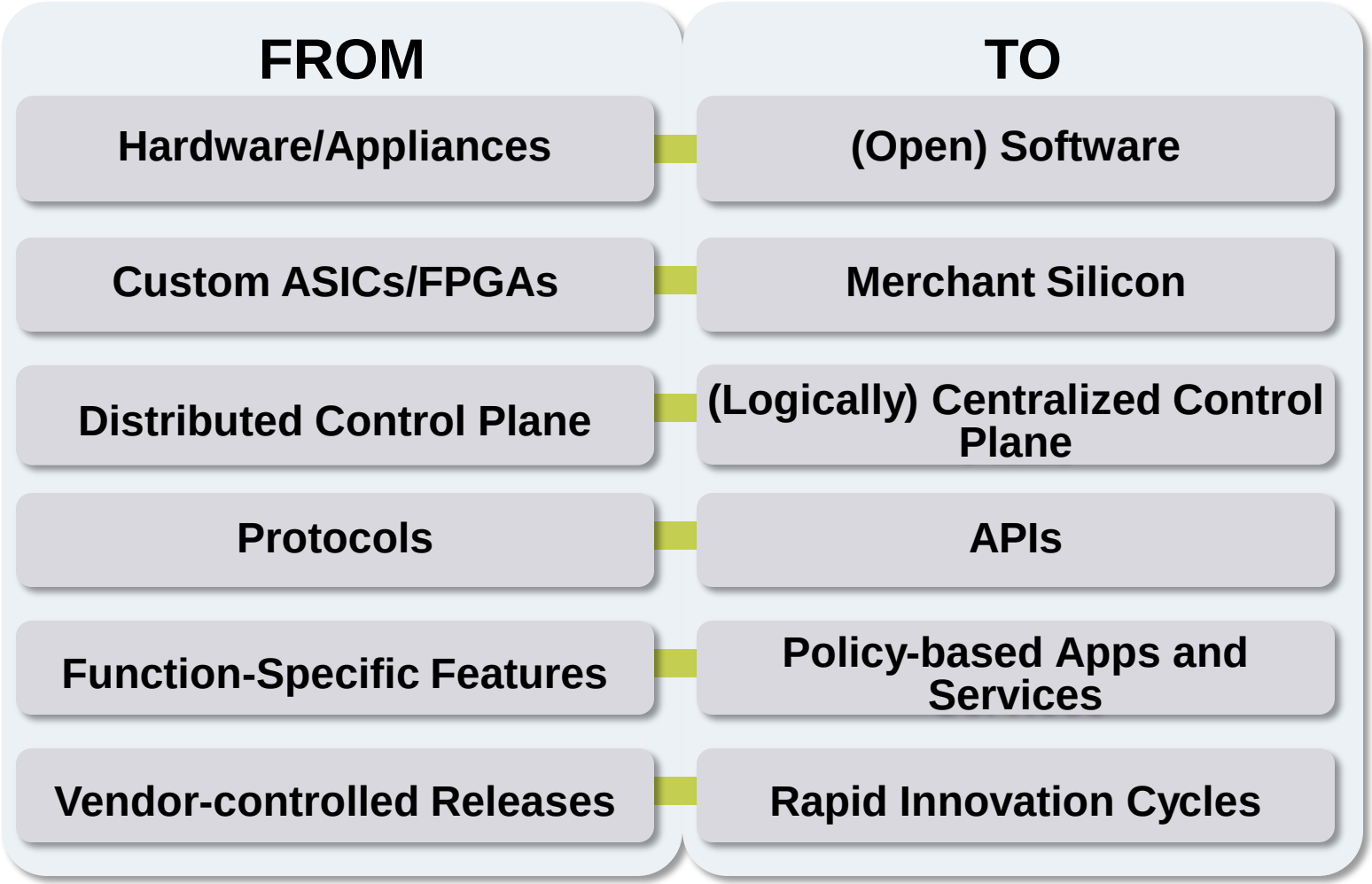
25%

Of all daily business communications will be video or multi-media communications by 2013

COLLABORATION, TRAINING, PRODUCTIVITY

Source: Gartner – G00207476 Key Technology Analysis
Gartner – G00175764 Key Issues For Communications Strategies, 2010
ONS12

General shift in networking



Source: Adapted from ONS12 Presentation by Dan Pitt

Which leads us to SDN

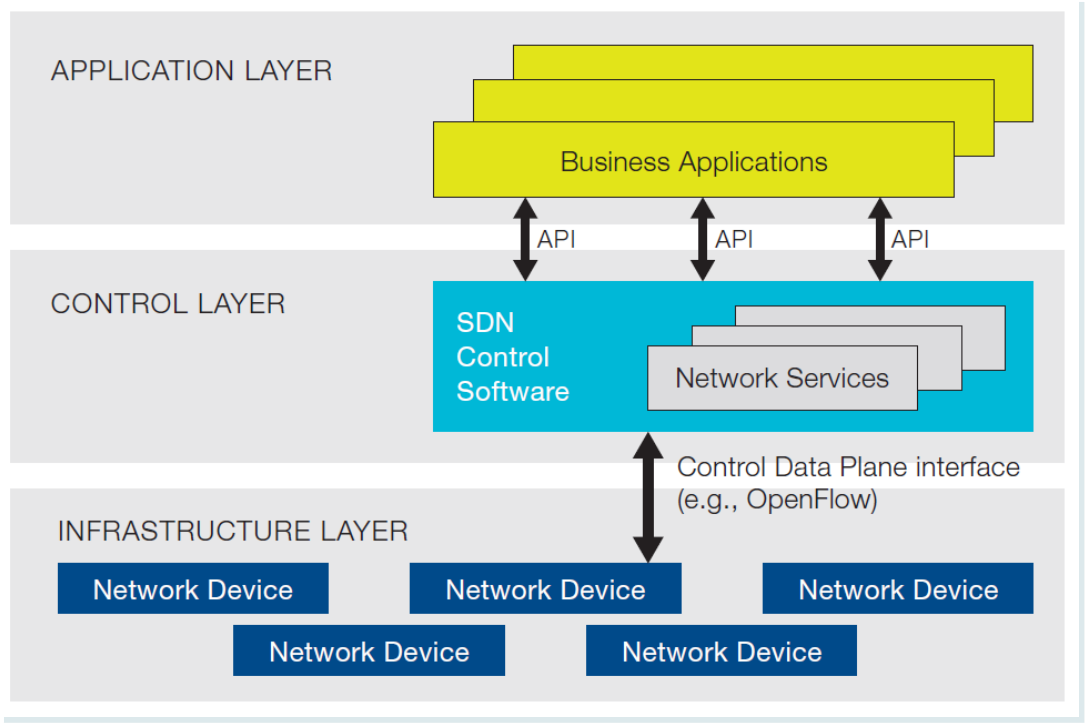
ONF

In the SDN architecture, the control and data planes are decoupled, network intelligence and state are logically centralized, and the underlying network infrastructure is abstracted from the applications.

Wikipedia

A network architecture in which the network control plane is decoupled from the physical topology.

SDN Model



Source: ONF White Paper *Software-Defined Networking: The New Norm for Networks* – April 13, 2012

SDN theory and practice

The Premise…

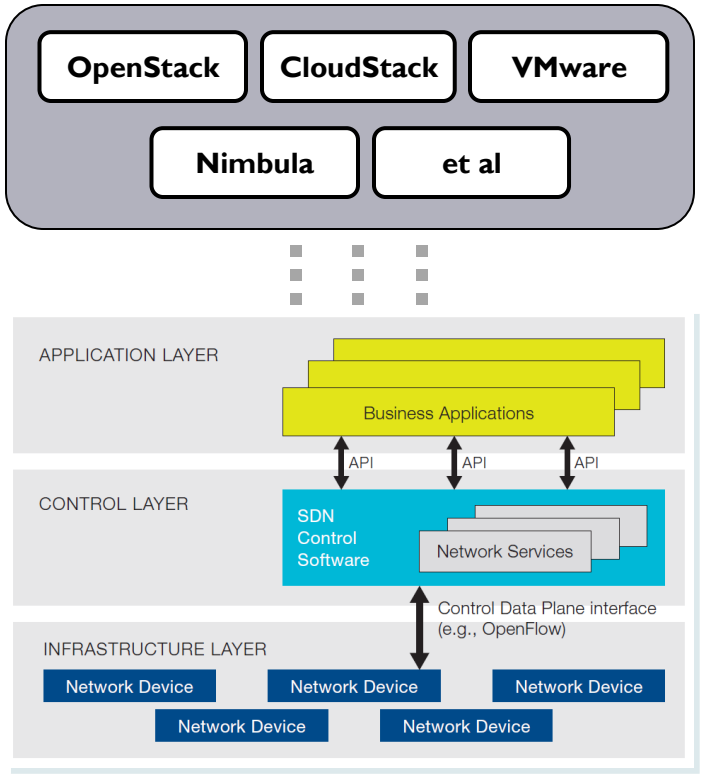
- Commodity (merchant silicon) solutions can be exploited
- Control plane can be distributed
- State can be externalized
- Acceptable performance can be maintained
- Standards will evolve
- Networking manufacturers will adopt SDN-enabling protocols and features

The Promise…

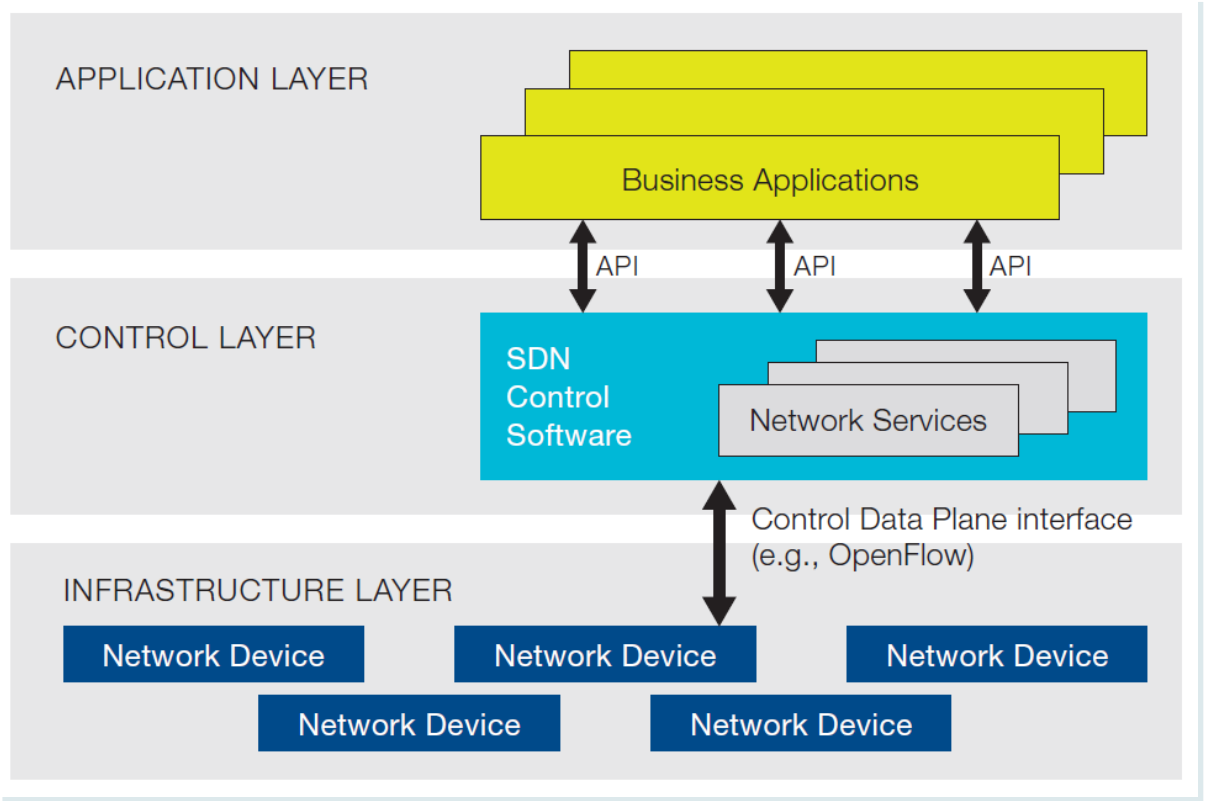
- Centralized management and control
- More granular network control
- Improved automation and management
- Rapid innovation
- Programmability
- Increased network reliability and security
- Better end-user experience

Essential elements of SDN

- Abstraction
- Pooling
- Orchestration
- Automation
- Service Insertion
- Apps
- Programmability (APIs)



Inside the layers



- **Virtual network overlays**
- **Slicing**
- **Tenant-aware broadcast**
- **Application-aware path computation**
- **Traffic engineering**
- **Network services (FW, LB, Security)**

- **Data plane resource management**
- **Common services and libraries**
- **Topology**
- **Metadata**
- **State abstraction**

- **Packet forwarding**
- **Packet manipulation**
- **Statistics gathering**

Source: Adapted from Dan Pitts, ONF

Typical use cases to date

- Multi-Tenancy
- Network Access Control
- Load Balancing
- Network Taps
- Cut-Through Applications
- Network Virtualization (overlays)

OpenFlow

OpenFlow ≠ **SDN**

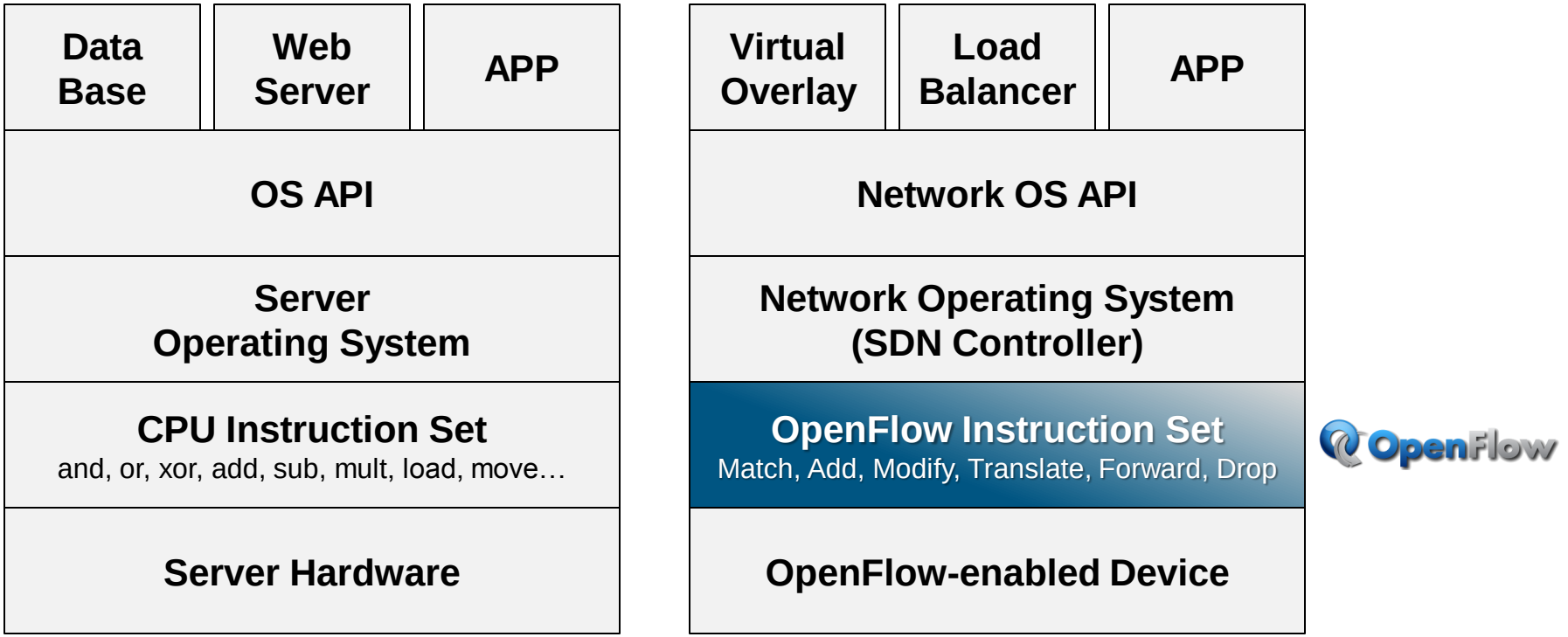
(by itself)

OpenFlow

- A protocol specification
  - Open Networking Foundation

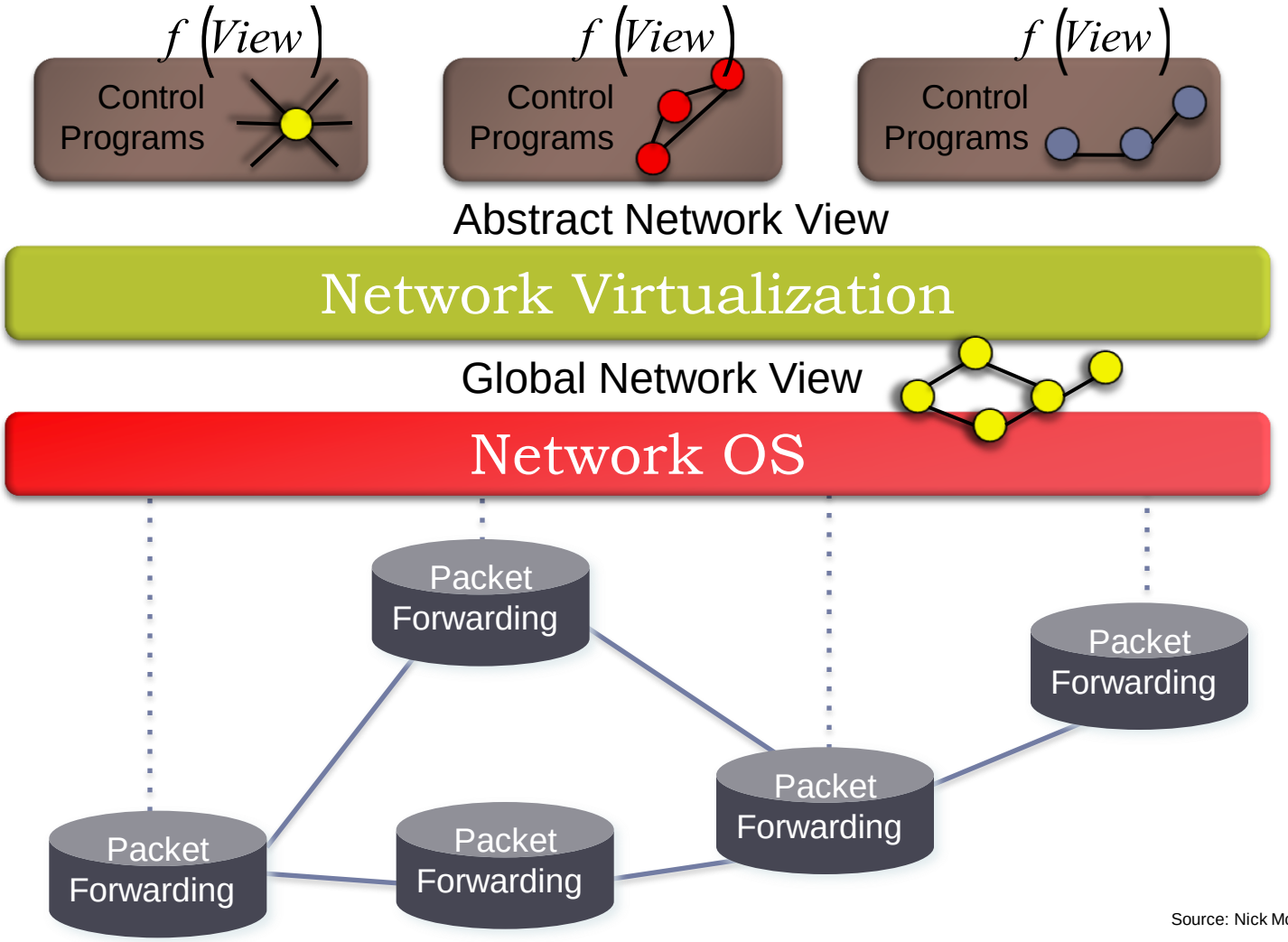
- Requires OpenFlow-enabled devices
  - Switches
- Defines controller messages
  - PACKET_IN, PACKET_OUT, FLOW_REMOVED, etc.
- Enables construction of Flow Tables
  - Match/Action

Analogy



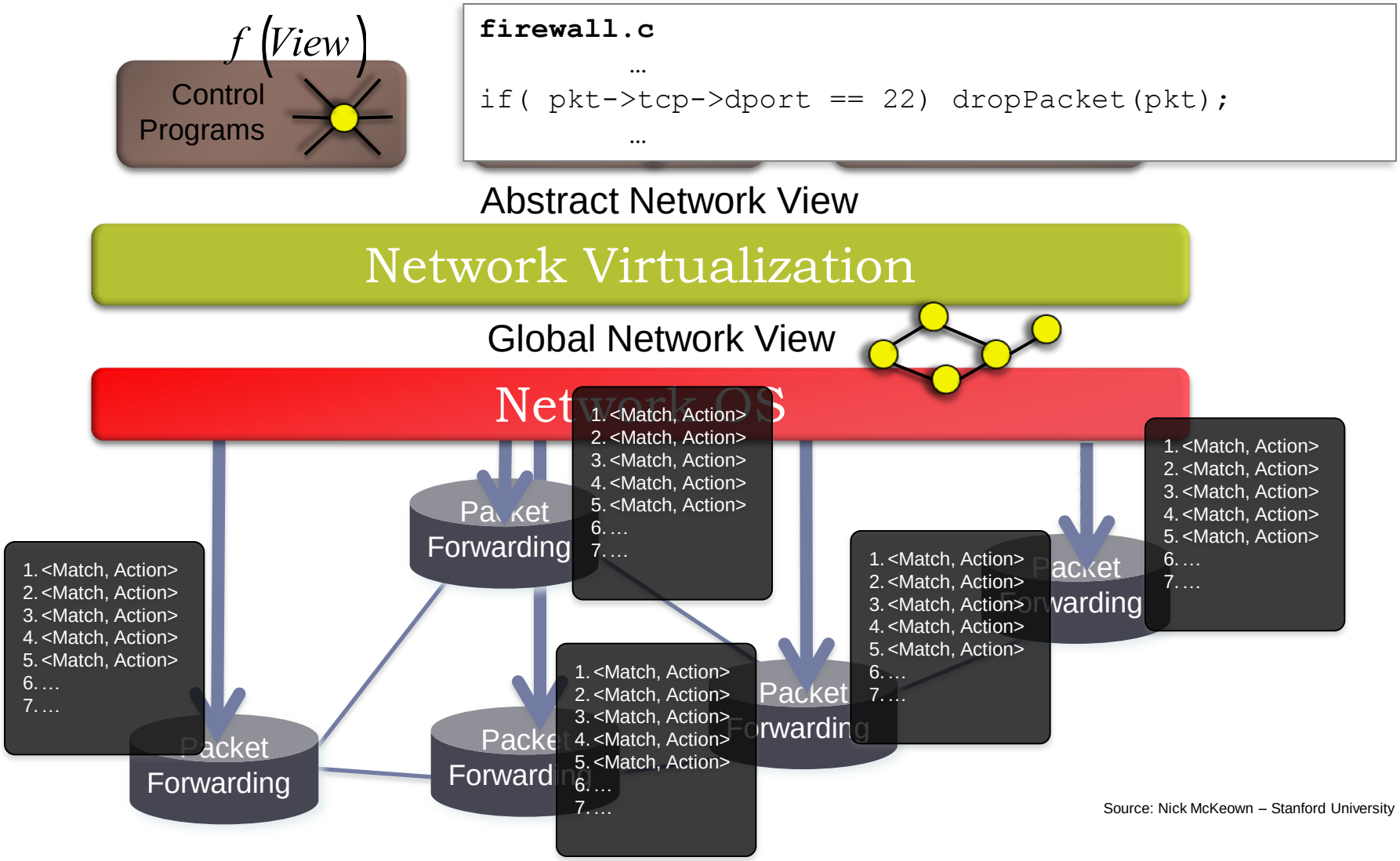
Source: Adapted from IBM ONS12 presentation by Rakesh Saha, IBM & Amit Agarwal, Google

Simple OpenFlow-enabled Example



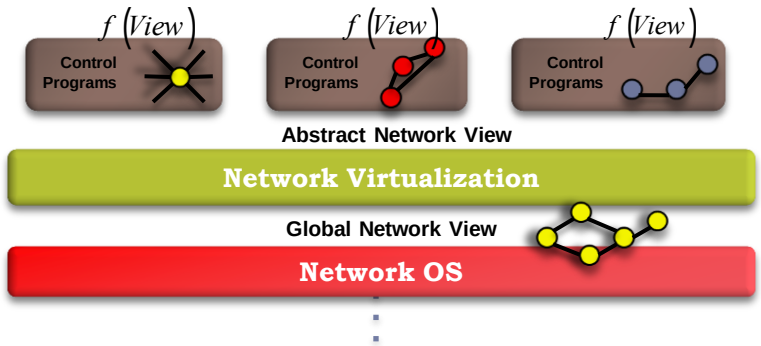
Source: Nick McKeown – Stanford University

Simple OpenFlow-enabled Example



Source: Nick McKeown – Stanford University

Flow Table Example



OpenFlow-enabled switch

Flow Table

MAC src	MAC dst	IP Src	IP Dst	TCP dport	...	Action	Count
*	10:20:.	*	*	*	*	port 1	250
*	*	*	5.6.7.8	*	*	port 2	300
*	*	*	*	25	*	drop	892
*	*	*	192.*	*	*	local	120
*	*	*	*	*	*	controller	11

Source: Adapted from ONS12 Presentation by Dan Pitt

Flow Table

Generic primitive that sits on top of (virtual) switch TCAM, designed to match well with common ASICs

Example actions:

1. Switching and routing (port)
2. Firewall (drop)
3. Use with switch's non-OpenFlow logic (local)
4. Send to controller for processing (controller)

Foundation network functions are split between switch and high-level decisions at the controller

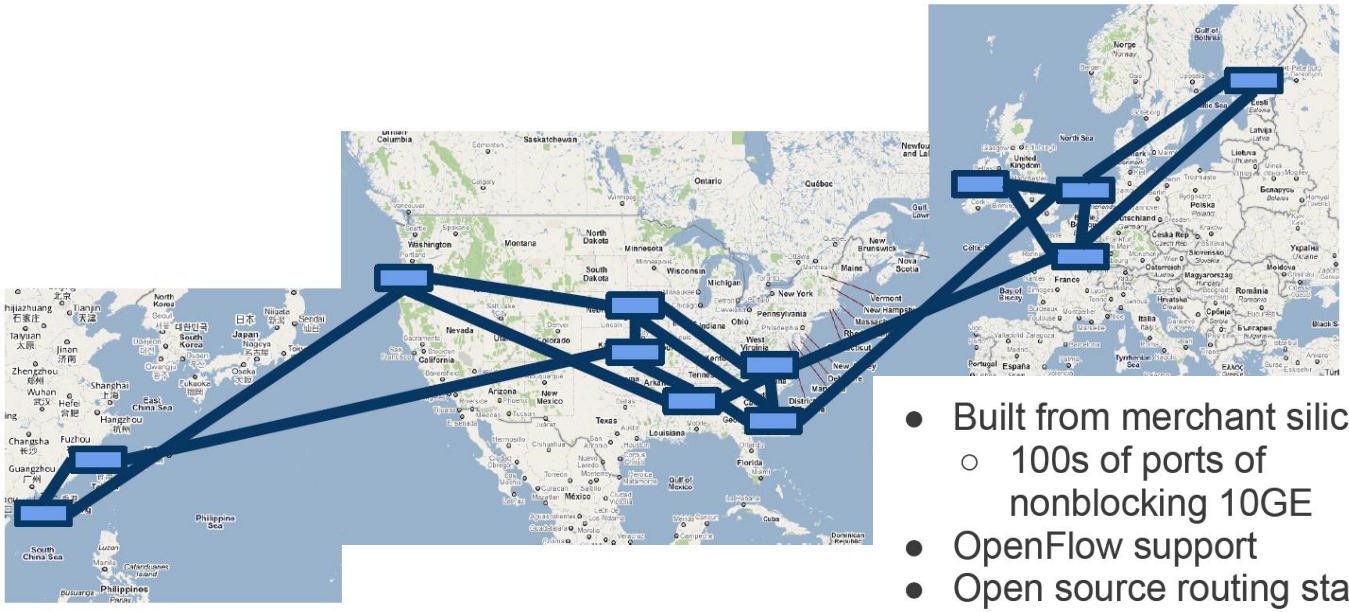
Major OpenFlow (SDN) Controllers

Controller	Language	Platform(s)	License	Originator
Beacon	**Java**	**Linux**	**GPL**	**Stanford University**
Floodlight	**Java**	**Linux**	**GPL**	**Big Switch (based on Beacon)**
Maestro	**Java**	**Windows, Mac, Linux, Android**	**GPL (core), FOSS Licenses for your code**	**Rice University**
NOX	**Python, C++**	**Linux**	**OpenFlow**	**Stanford University**
OpenTransit	**Java, Python**	**Linux**	**CPlane License**	**CPlane (LAYERZngn)**
ProgrammableFlow Controller	**Ruby, C**	**Linux (RHEL 6.1)**	**GPL/NEC**	**NEC (based on Trema)**
Programmable Network Controller	**Ruby, C**	**Linux (RHEL 6.1)**	**GPL/IBM**	**IBM (NEC OEM)**
Open Network Environment (ONE)				***Cisco***
Virtual Application Networks SDN Controller				***HP – Available 2H13 (Rumored to be Big Switch)***
NetScaler SDX				***Citrix – Early 2013***
?				***Juniper Networks***

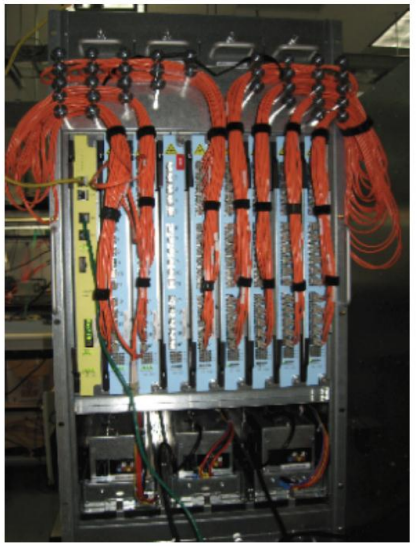
OpenFlow-Enabled Switches

- **Arista**
  - 7050
- **Brocade**
  - MLX Series
  - NetIron CER 2000 Series
- **Cisco**
  - Nexus 7000 Series
- **Dell**
  - Force10 MXL 10/40GbE
- **Extreme Networks**
  - Black Diamond X Series
- **HP**
  - 3500 Series
  - 3800 Series
  - 5400 Series
  - 8200 Series
- **IBM**
  - IBM RackSwitch G8264
- **Juniper Networks**
  - MX Series
- **Marvell**
  - Prestera
- **NEC**
  - PF5240
  - PF5280
- **NETGEAR**
  - ProSafe Plus Series
- **Pica8**
  - 3290, 3295, 3780, 3920
- **Pluribus Networks**
  - F64 Series

Real-world “G-scale” OF example



- Built from merchant silicon
  - 100s of ports of nonblocking 10GE
- OpenFlow support
- Open source routing stacks for BGP, ISIS
- Does not have all features
  - No support for AppleTalk...
- Multiple chassis per site
  - Fault tolerance
  - Scale to multiple Tbps

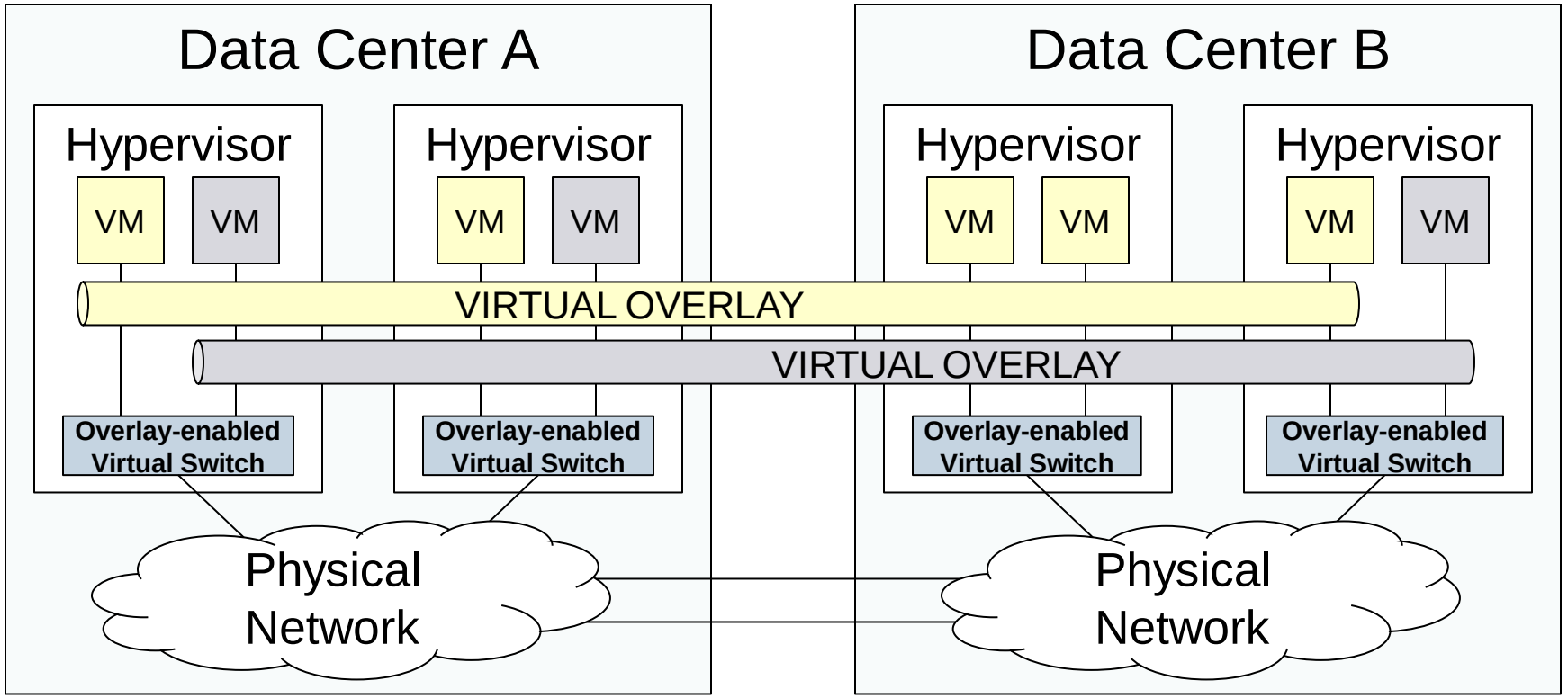


Source: Google ONS12 presentation

Data Center/Cloud Networking Issues

- VLAN limits (4,094)
- Spanning Tree Protocol disabled links
- Reconfiguration to extend VLANs
- MAC address contention
- MAC address table size in ToR switches
- Layer 3 address contention
- Security "choke points"
- …

Virtual Overlays Using IP-encapsulation



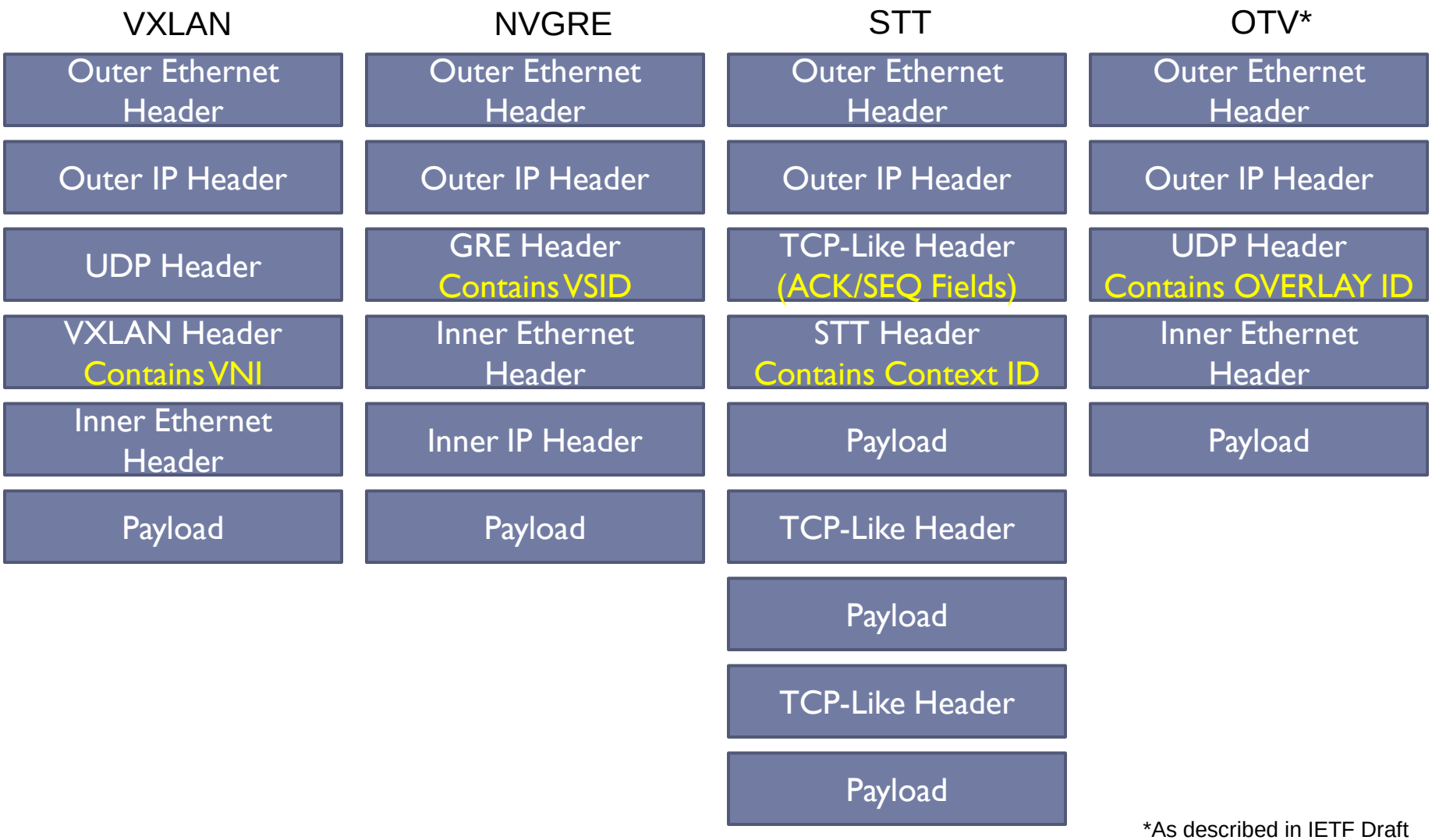
- "Similar" to other tunneling methods (L2TPV3, AToM, VPLS, LISP)
- Encapsulation via tunnel "endpoints"
- Not dependent on specific transports
- Layer 2 over Layer 3 (e.g., ECMP/OSPF)
- VMs see only Layer 2
- "Customer-edge" → easier to set up

Common IP-based Encapsulation Methods

Method	Full Name	Sponsors	Approach
DOVE	Distributed Overlay Virtual Ethernet	IBM	Leverages OTV and VXLAN
NVGRE	Network Virtualization using Generic Routing Encapsulation	Arista Networks, Broadcom, Dell, Emulex, HP, Intel, Microsoft	24-bit Virtual Subnet Identifier (VSI) in GRE header
OTV	Overlay Transport Virtualization	Cisco	~~24-bit Overlay ID in OTV header inside UDP~~ VLAN extension via GRE/MPLS (Nexus 7000)
STT	Stateless Transport Tunneling	Nicira (VMware)	64-bit Context ID in STT header, "TCP-like" header, leverages TSO/LRO
VXLAN	Virtual Extensible Local Area Network	Arista Networks, Broadcom, Cisco, Citrix, Red Hat, VMware	24-bit VXLAN Network Identifier (VNI) in VXLAN header inside UDP packet

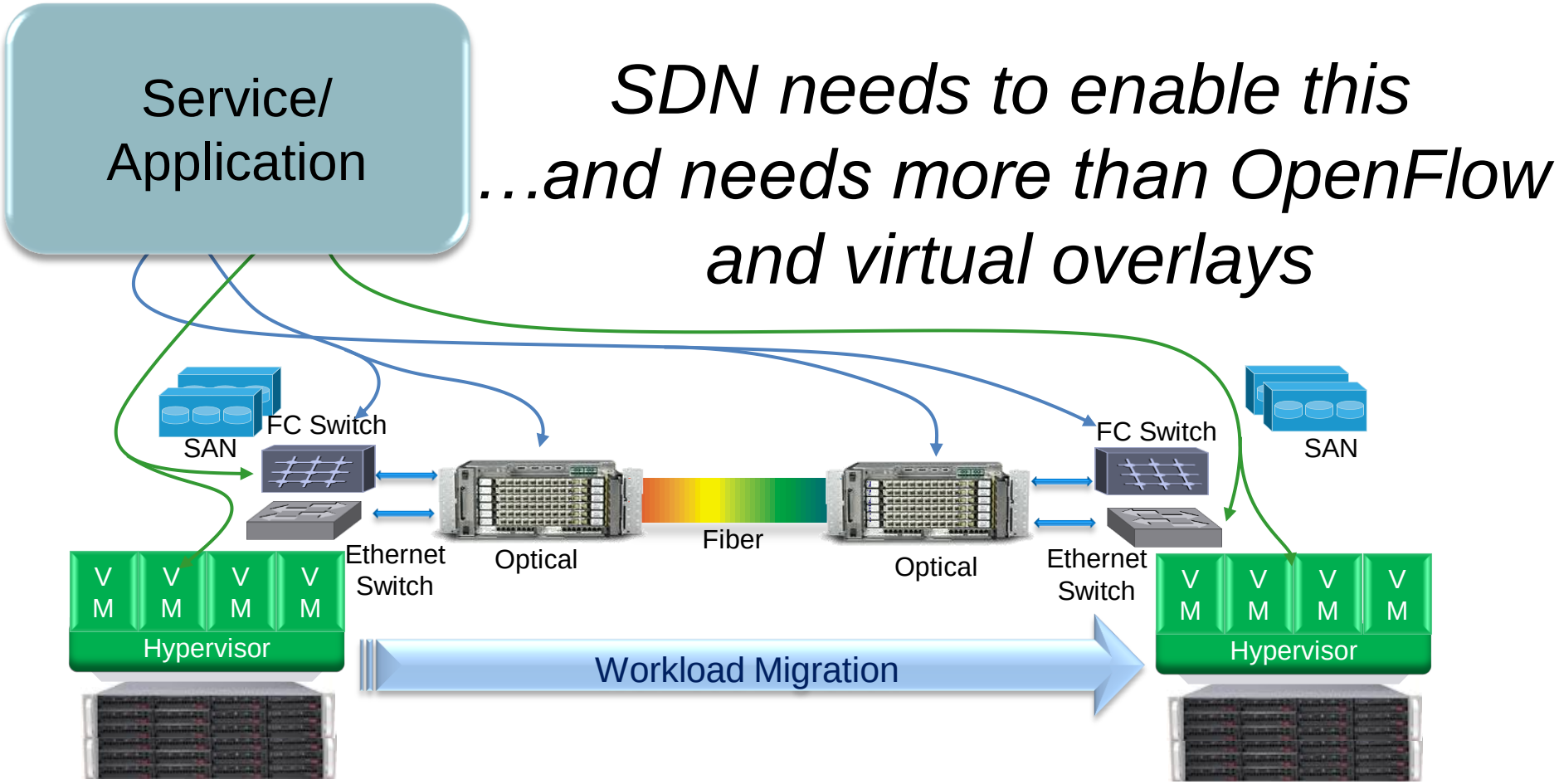
- Different approaches to destination endpoint identification
- Different approaches to load balancing for efficiency
- Can be negatively impacted by "middle boxes" (firewalls, intrusion protection, etc.)
- Some increased exposure to MAC-over-IP security threats

Encapsulation Headers



*As described in IETF Draft

But wait…


Service/
Application
SDN needs to enable this
…and needs more than OpenFlow
and virtual overlays
SAN
FC Switch
Ethernet
Switch
Optical
Fiber
Optical
FC Switch
Ethernet
Switch
SAN
V
M
Hypervisor
Workload Migration

Still work to be done…

Data Plane

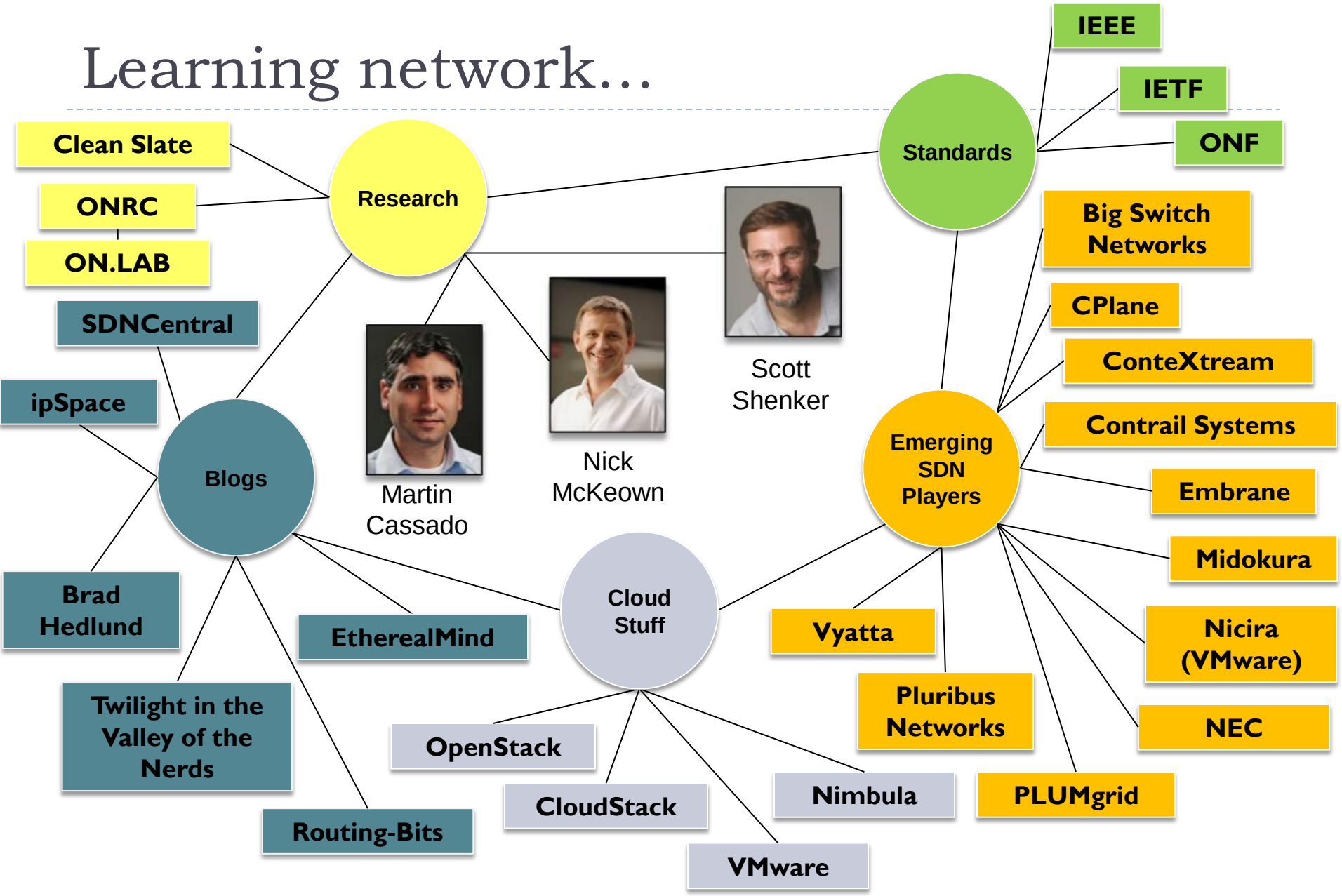
- State of specifications
  - Maturity
  - Changes across releases
- Silicon Concerns
  - Specifications outpace silicon development
  - Merchant silicon not optimized for OF
- Performance
  - Scalability of Flow-Matches (limited by TCAM size)
  - Cost driver excludes rich multi-core xPU ecosystem

Control Plane

- Scalability
  - Centralized vs distributed
  - State coherence between control and data plane
- Interoperability
  - SDN to non-SDN
  - Inter-Controller
  - Multi-orchestrator conflicts
  - Virtual overlays

Source: Adapted from ONS12 Presentation by Geng Lin - Dell

Learning network…



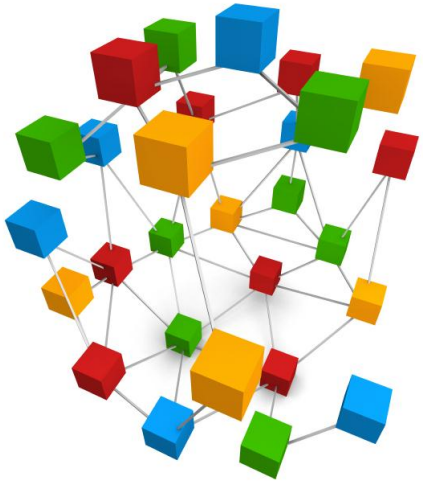
Resource Links

- **Big Switch Networks:** www.bigswitch.com
- **Brad Hedlund:** www.bradhedlund.com
- **CloudStack:** www.cloudstack.org
- **ConteXtream:** www.contextream.com
- **Contrail Systems** www.contrailsystems.com
- **CPlane:** www.cplane.net
- **Embrane:** www.embrane.com
- **Ethereal Mind:** www.etherealmind.com
- **IEEE:** www.ieee.org
- **IETF:** www.ietf.org
- **ipSpace:** www.ipspace.net
- **Midokura:** www.midokura.com
- **NEC:** www.necam.com/pflow
- **Nicira:** www.nicira.com
- **Nimbula:** www.nimbula.com
- **ONRC:** onrc.stanford.edu
- **ON.LAB:** onlab.us
- **ONF:** www.opennetworking.org
- **OpenStack:** www.openstack.org
- **ONF:** www.opennetworking.org
- **PLUMgrid:** www.plumgrid.com
- **Pluribus Networks:** www.pluribusnetworks.com
- **Routing-Bits:** www.routing-bits.com
- **SDNCentral:** www.sdncentral.com
- **Stanford Clean Slate:** cleanslate.stanford.edu
- **Twilight in the Valley of the Nerds:** nerdtwilight.wordpress.com
- **VMware:** www.vmware.com
- **Vyatta:** www.vyatta.com

Martin Cassado

Nick McKeown

Scott Shenker



Thank you!

Robert Keahey

robert.keahey@summalogic.com

www.summalogic.com

robert@cplane.net

www.cplane.net