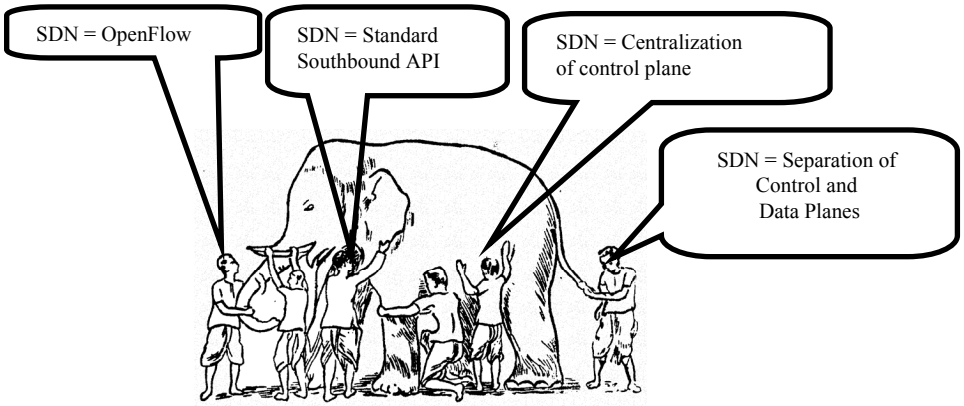
Virtualization and Software Defined Networking (SDN) for *Multi-Cloud Computing*



RAJ JAIN
Washington University in Saint Louis
Saint Louis, MO 63130, Jain@cse.wustl.edu

Invited Talk at PES University, Bangaluru, Sept 17, 2014

These slides and video recording of this presentation are at:

http://www.cse.wustl.edu/~jain/talks/apf_pes.htm

Overview

1. Five concepts/events that have changed the networking world: Virtualization, Cloud, Smart Phones, SDN, NFV
2. What really is SDN?: SDN 1.0 vs. SDN 2.0
3. Network Function Virtualization
4. Mobile Apps ⇒ Global Cloud of Clouds

1. Virtualization

- Internet ⇒ Virtualization

- No need to get out for
  - Office
  - Shopping
  - Education
  - Entertainment
- Virtual Workplace
- Virtual Shopping
- Virtual Education
- Virtual Sex

Virtualization

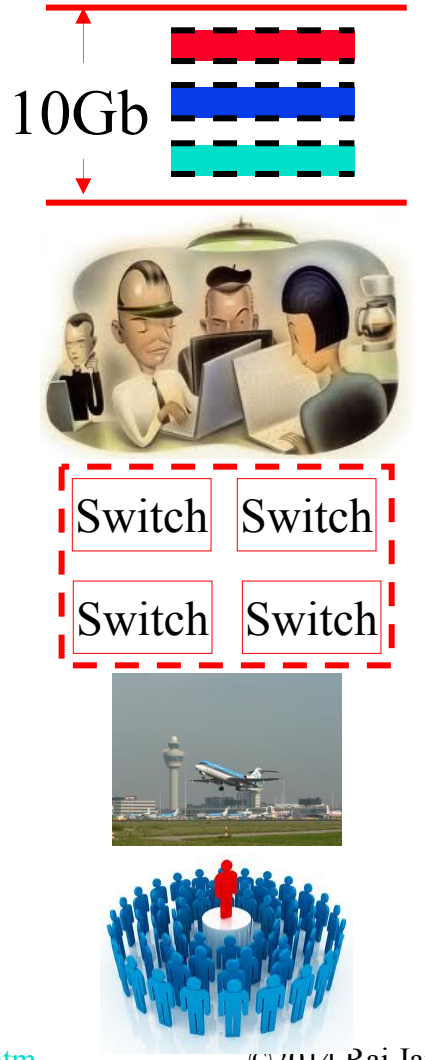
"*Virtualization means that Applications can use a resource without any concern for where it resides, what the technical interface is, how it has been implemented, which platform it uses, and how much of it is available.*"

-Rick F. Van der Lans

in Data Virtualization for Business Intelligence Systems

5 Reasons to Virtualize

1. Sharing: Break up a large resource
   Large Capacity or high-speed
   ⇒ Multi-Tenant
2. Isolation: Protection from other tenants
3. Aggregating: Combine many resources in to one
4. Dynamics: Fast allocation, Change/Mobility, Follow the sun (active users) or follow the moon (cheap power)
5. Ease of Management
   ⇒ Cost Savings. fault tolerance



2. Cloud Computing

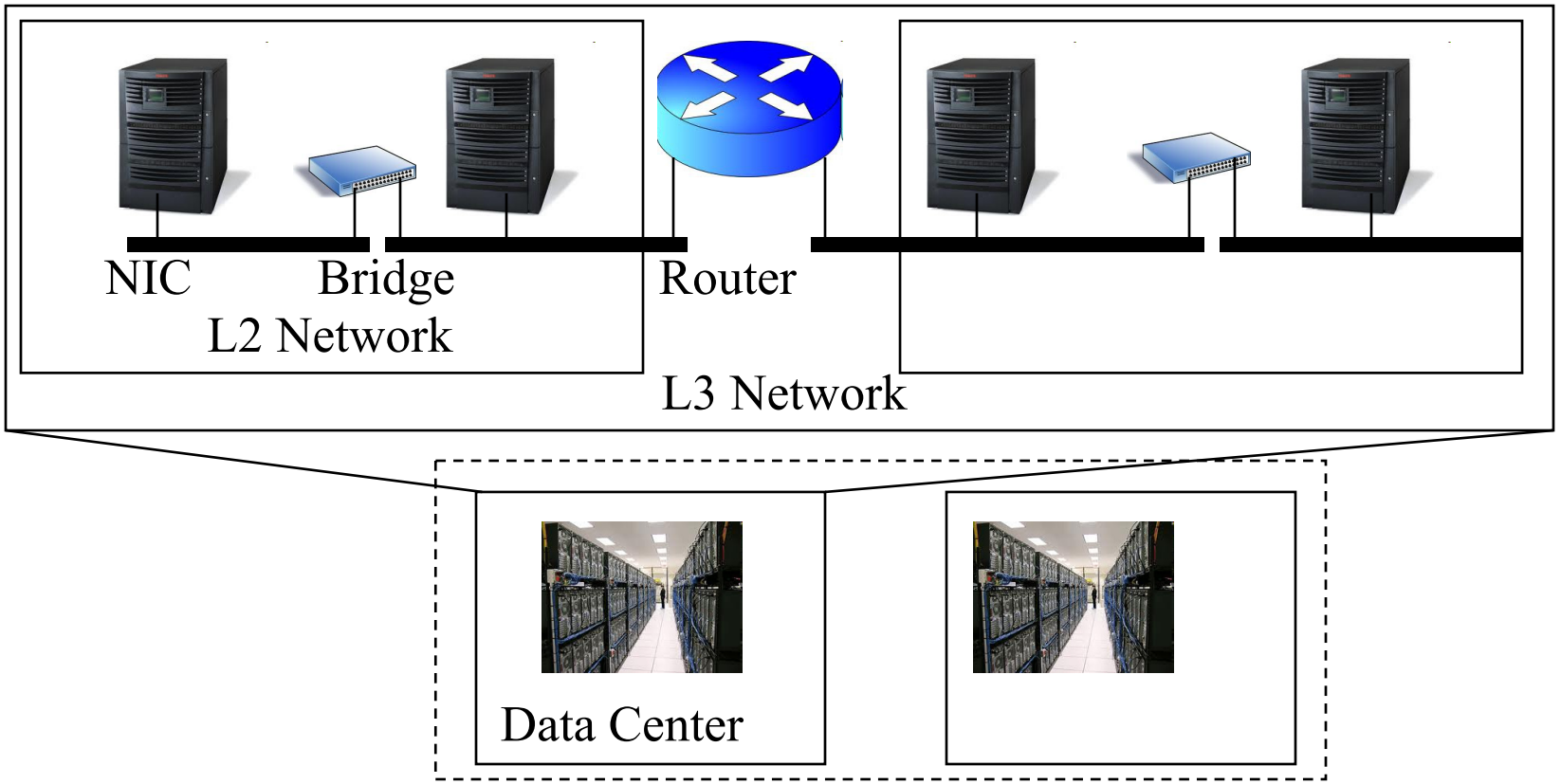
- August 25, 2006: Amazon announced EC2 ⇒ Birth of Cloud Computing in reality (Prior theoretical concepts of computing as a utility)
- *Web Services To Drive Future Growth For Amazon* ($2B in 2012, $7B in 2019) - Forbes, Aug 12, 2012

- Cloud computing was made possible by computing virtualization
- **Networking**: Plumbing of computing
  - IEEE: Virtual Bridging, …
  - IETF: Virtual Routers, …
  - ITU: Mobile Virtual Operators, …

Why Virtualize a Network?

1. Network virtualization allows tenants to form an overlay network in a multi-tenant network such that tenant can control:
   1. Connectivity layer: Tenant network can be L2 while the provider is L3 and vice versa
   2. Addresses: MAC addresses and IP addresses
   3. Network Partitions: VLANs and Subnets
   4. Node Location: Move nodes freely
2. Network virtualization allows providers to serve a large number of tenants without worrying about:
   1. Internal addresses used in client networks
   2. Number of client nodes
   3. Location of individual client nodes
   4. Number and values of client partitions (VLANs and Subnets)
3. Network could be a single physical interface, a single physical machine, a data center, a metro, … or the global Internet.
4. Provider could be a system owner, an enterprise, a cloud provider, or a carrier.

Levels of Network Virtualization



- Networks consist of: **Network Interface Card (NIC)** – **L2 Links** - **L2 Bridges** - **L2 Networks** - L3 Links - L3 Routers - L3 Networks – **Data Centers** – **Global Internet.**
- Each of these needs to be virtualized

Network Virtualization Techniques

Entity	Partitioning	Aggregation/Extension/Interconnection**
NIC	SR-IOV	MR-IOV
Switch	VEB, VEPA	VSS, VBE, DVS, FEX
L2 Link	VLANs	LACP, Virtual PortChannels
L2 Network using L2	VLAN	PB (Q-in-Q), PBB (MAC-in-MAC), PBB-TE, Access-EPL, EVPL, EVP-Tree, EVPLAN
L2 Network using L3	NVO3, VXLAN, NVGRE, STT	MPLS, VPLS, A-VPLS, H-VPLS, PWoMPLS, PWoGRE, OTV, TRILL, LISP, L2TPv3, EVPN, PBB-EVPN
Router	VDCs, VRF	VRRP, HSRP
L3 Network using L1		GMPLS, SONET
L3 Network using L3*	MPLS, GRE, PW, IPSec	MPLS, T-MPLS, MPLS-TP, GRE, PW, IPSec
Application	ADCs	Load Balancers

*All L2/L3 technologies for L2 Network partitioning and aggregation can also be used for L3 network partitioning and aggregation, respectively, by simply putting L3 packets in L2 payloads.

**The aggregation technologies can also be seen as partitioning technologies from the provider point of view.

Names, IDs, Locators

Name: John Smith

ID: 012-34-5678

Locator:
1234 Main Street
Big City, MO 12345
USA

- Locator changes as you move, ID and Names remain the same.
- **Examples**:
  - Names: Company names, DNS names (Microsoft.com)
  - IDs: Cell phone numbers, 800-numbers, Ethernet addresses, Skype ID, VOIP Phone number
  - Locators: Wired phone numbers, IP addresses

Fallacies Taught in Networking Classes

1. Ethernet is a local area network (Local $\leq$ 2km)
2. Token ring, Token Bus, and CSMA/CD are the three most common LAN access methods.
3. Ethernet uses CSMA/CD.
   No CSMA/CD in 10G and up
   No CSMA/CD in practice now even at home or at 10 Mbps
4. Ethernet bridges use spanning tree for packet forwarding.
5. Ethernet frames are limited to 1518 bytes.
6. Ethernet does not provide any delay guarantees.
7. Ethernet has no congestion control.
8. Ethernet has strict priorities.

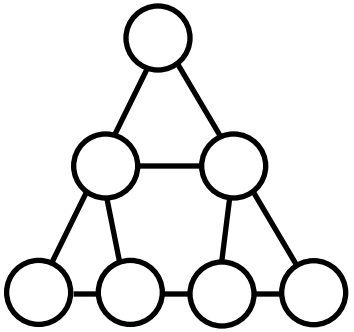
Ethernet has changed.
All of these are now false or are becoming false.

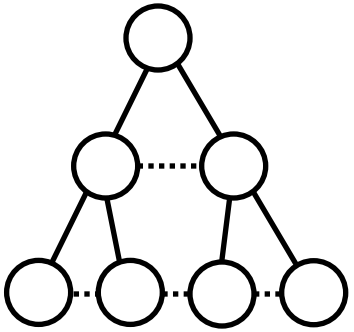
Residential vs. Data Center Ethernet

Residential	Data Center/Cloud
❑ Distance: up to 200m	❑ No limit
❑ Scale: ➢ Few MAC addresses ➢ 4096 VLANs	❑ Millions of MAC Addresses ❑ Millions of VLANs Q-in-Q
❑ Protection: Spanning tree	❑ Rapid spanning tree, … (Gives 1s, need 50ms)
❑ Path determined by spanning tree	❑ Traffic engineered path
❑ Simple service	❑ Service Level Agreement. Rate Control.
❑ Priority ⇒ Aggregate QoS	❑ Need per-flow/per-class QoS
❑ No performance/Error monitoring (OAM)	❑ Need performance/BER

Spanning Tree and its Enhancements

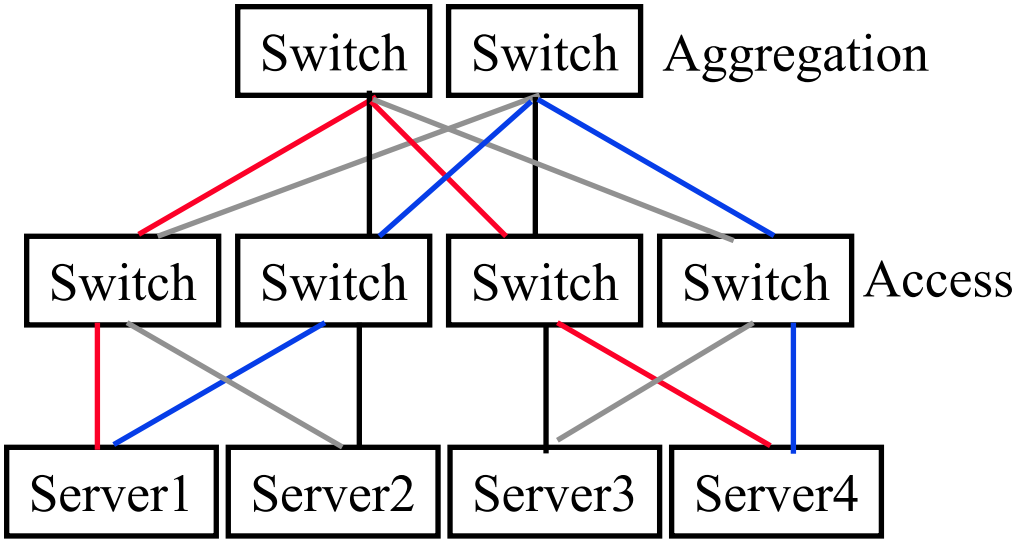
- Helps form a tree out of a mesh topology
- A topology change can result in 1 minute of traffic loss with STP ⇒ All TCP connections break
- Rapid Spanning Tree Protocol (RSTP)
  IEEE 802.1w-2001 incorporated in IEEE 802.1D-2004
- One tree for all VLANs
  ⇒ Common spanning tree
- Many trees
  ⇒ Multiple spanning tree (MST) protocol
  IEEE 802.1s-2002 incorporated in IEEE 802.1Q-2005
- One or more VLANs per tree.





Shortest Path Bridging

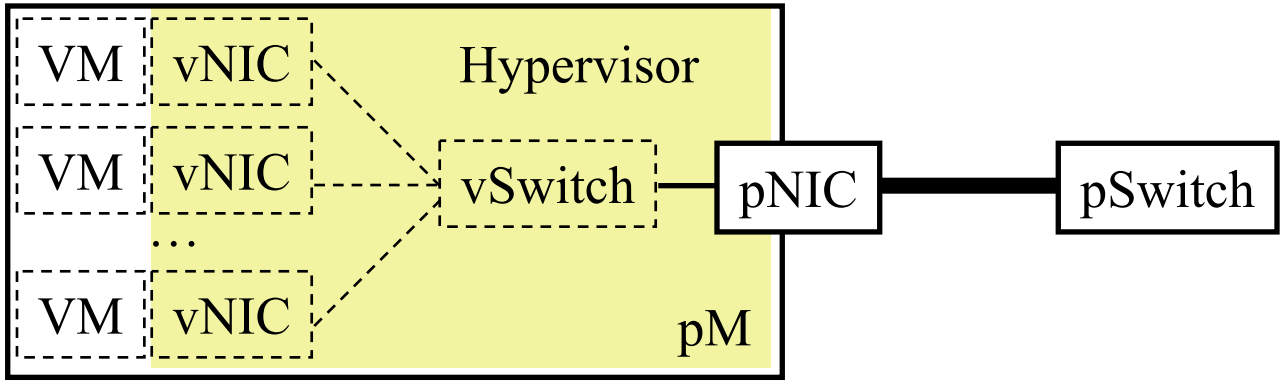
- IEEE 802.1aq-2012
- Allows all links to be used ⇒ Better CapEx
- IS-IS link state protocol (similar to OSPF) is used to build shortest path trees for each node to every other node within the SPB domain
- Equal-cost multi-path (ECMP) used to distribute load



Ref: http://en.wikipedia.org/wiki/Shortest_Path_Bridging

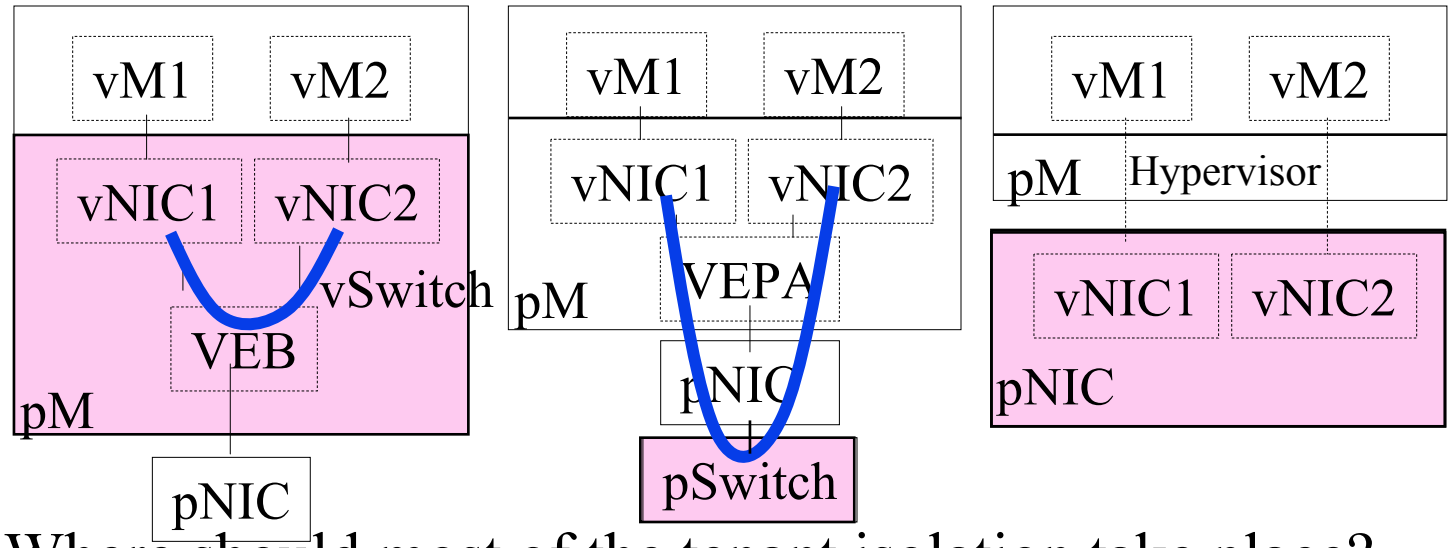
vSwitch

- **Problem**: Multiple VMs on a server need to use one physical network interface card (pNIC)
- **Solution**: Hypervisor creates multiple vNICs connected via a virtual switch (vSwitch)
- pNIC is controlled by hypervisor and not by any individual VM
- **Notation**: From now on prefixes p and v refer to physical and virtual, respectively. For VMs only, we use upper case V.



Ref: G. Santana, "Datacenter Virtualization Fundamentals," Cisco Press, 2014, ISBN: 1587143240

Virtual Bridging



Where should most of the tenant isolation take place?

1. VM vendors: S/W NICs in Hypervisor w Virtual Edge Bridge (**VEB**)(overhead, not ext manageable, not all features)
2. Switch Vendors: Switch provides virtual channels for inter-VM Communications using virtual Ethernet port aggregator (**VEPA**): **802.1Qbg** (s/w upgrade)
3. NIC Vendors: NIC provides virtual ports using Single-Route I/O virtualization (**SR-IOV**) on PCI bus

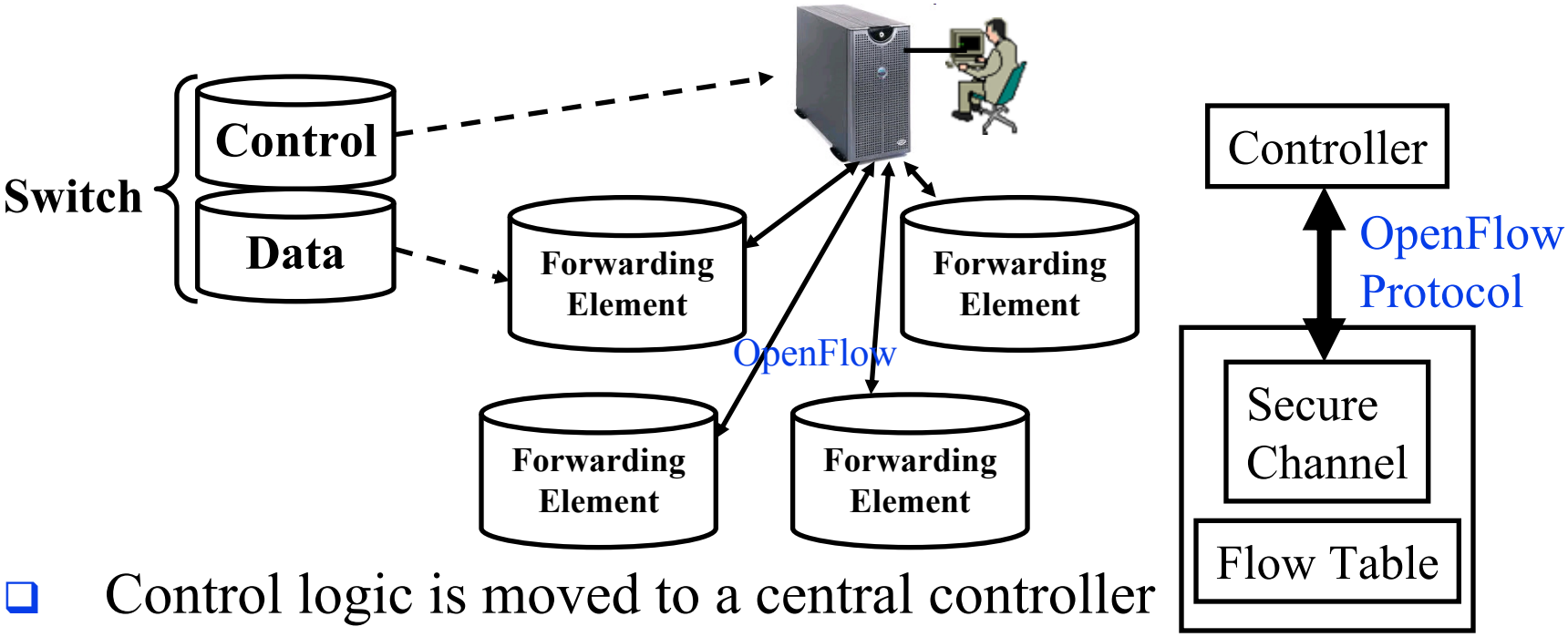
Planes of Networking

- **Data Plane**: All activities involving as well as resulting from data packets sent by the end user, e.g.,
  - Forwarding
  - Fragmentation and reassembly
  - Replication for multicasting
- **Control Plane**: All activities that are necessary to perform data plane activities but do not involve end-user data packets
  - Making routing tables
  - Setting packet handling policies (e.g., security)

Dest.	Output Port	Next Hop

Ref: Open Data Center Alliance Usage Model: Software Defined Networking Rev 1.0,"
http://www.opendatacenteralliance.org/docs/Software_Defined_Networking_Master_Usage_Model_Rev1.0.pdf

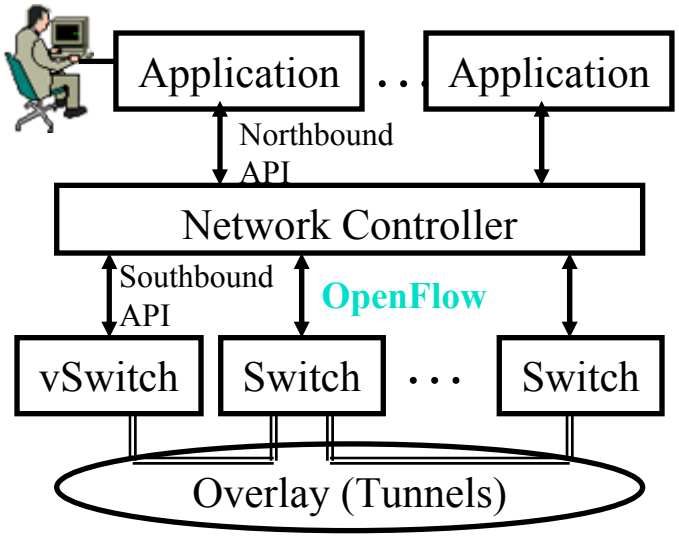
Separation of Control and Data Plane



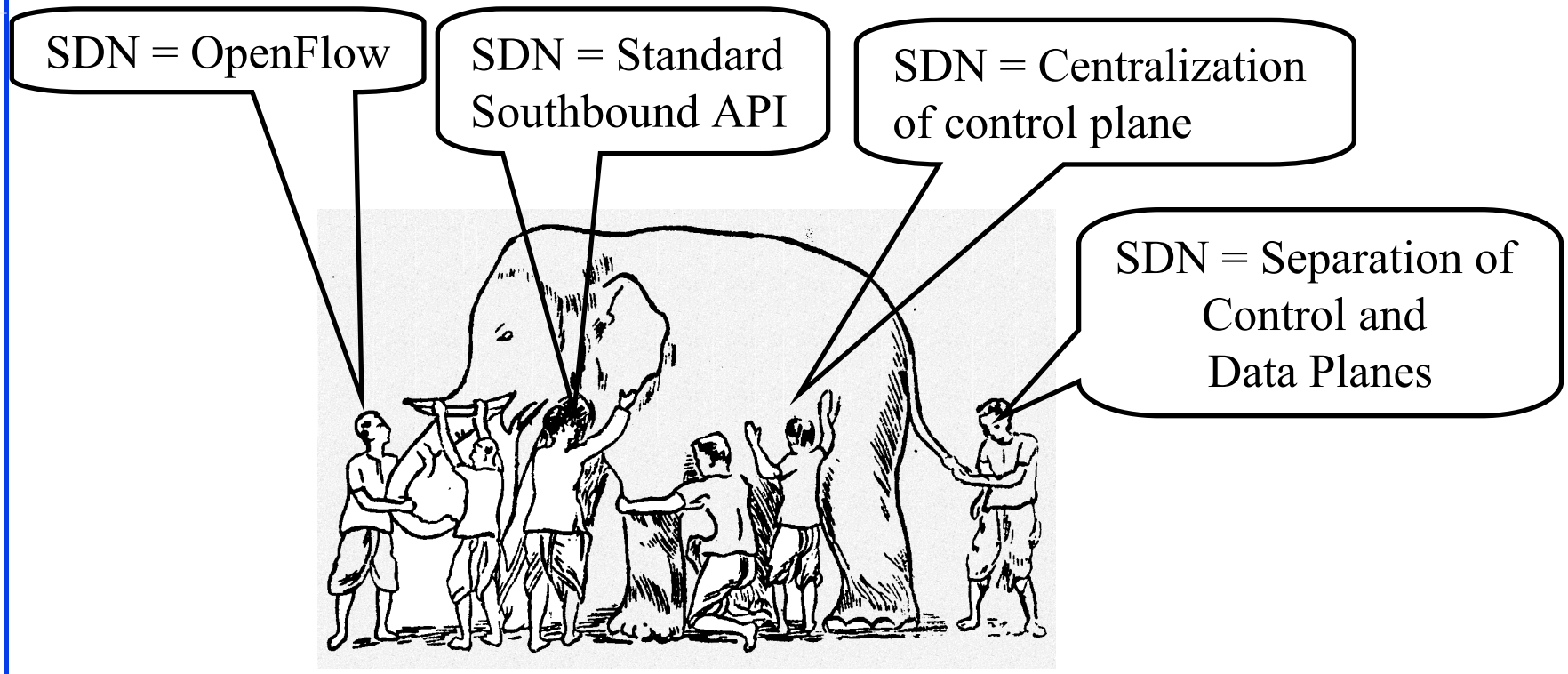
- Control logic is moved to a central controller
- Switches only have forwarding elements
- One expensive controller with a lot of cheap switches
- OpenFlow is the protocol to send/receive forwarding rules from controller to switches

SDN 1.0: SDN Based on OpenFlow

- SDN originated from OpenFlow
- Centralized Controller
  ⇒ Easy to program
  ⇒ Change routing policies on the fly
  ⇒ Software Defined Network (SDN)
- Initially, SDN = OpenFlow



What is SDN?

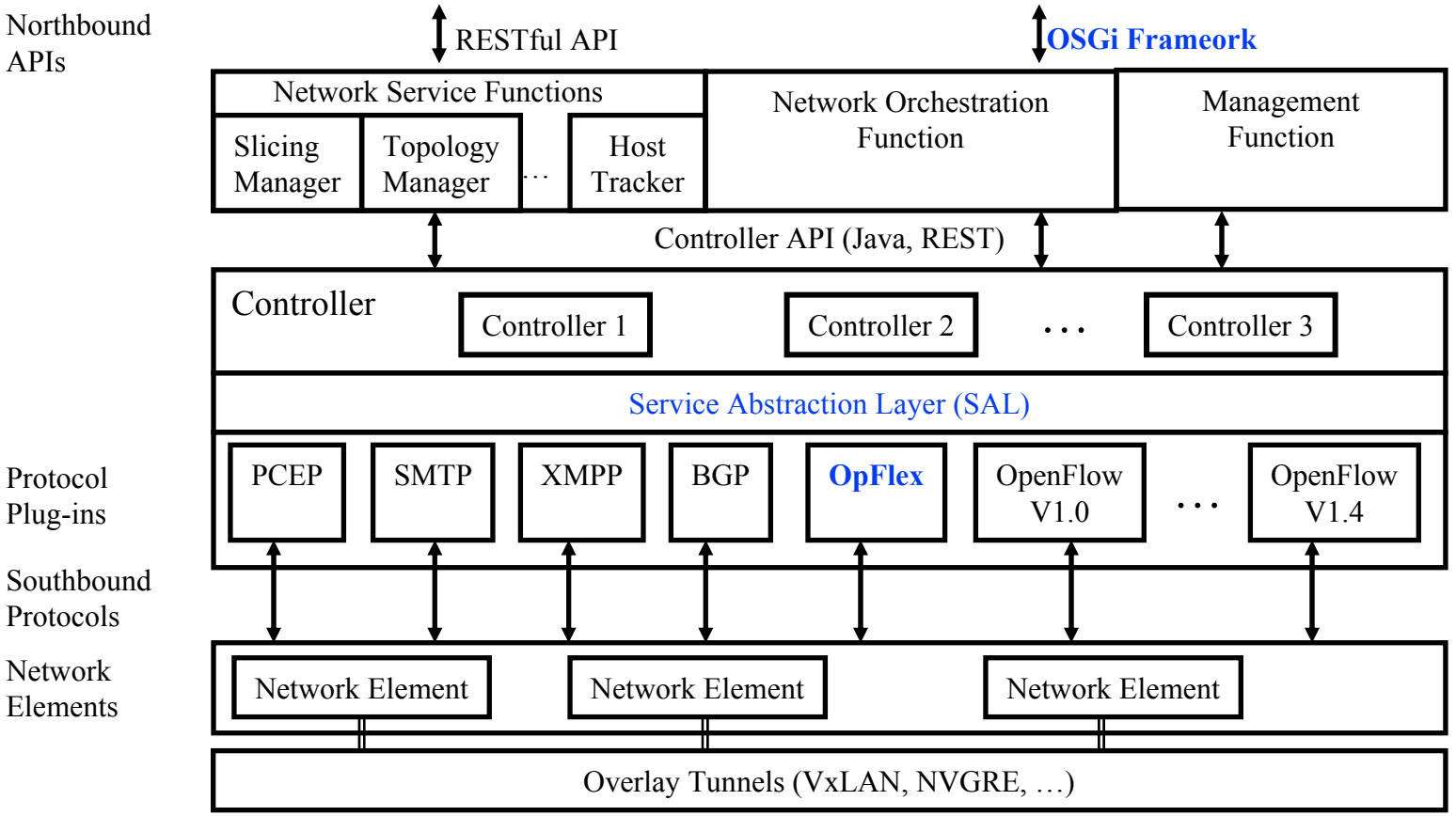


- All of these are mechanisms.
- SDN is *not* about a mechanism.
- It is a framework to solve a set of problems ⇒ Many solutions

What do We need SDN for?

1. **Virtualization**: Use network resource without worrying about where it is physically located, how much it is, how it is organized, etc.
2. **Orchestration**: Manage thousands of devices
3. **Programmable**: Should be able to change behavior on the fly.
4. **Dynamic Scaling**: Should be able to change size, quantity
5. **Automation**: Lower OpEx
6. **Visibility**: Monitor resources, connectivity
7. **Performance**: Optimize network device utilization
8. **Multi-tenancy**: Sharing expensive infrastructure
9. **Service Integration**
10. **Openness**: Full choice of Modular plug-ins
11. **Unified management** of computing, networking, and storage

SDN 2.0: OpenDaylight Style SDN



- **NO-OpenFlow** (**Not Only** OpenFlow) Multi-Protocol
- New work in **IETF** XMPP, ALTO, I2RS, PCEP, ….
- Linux Foundation

Open Everything

- Open Networking Foundation
- OpenFlow
- OpenStack
- OpenDaylight
- Open Access
- Open Source

Current SDN Debate: What vs. How?

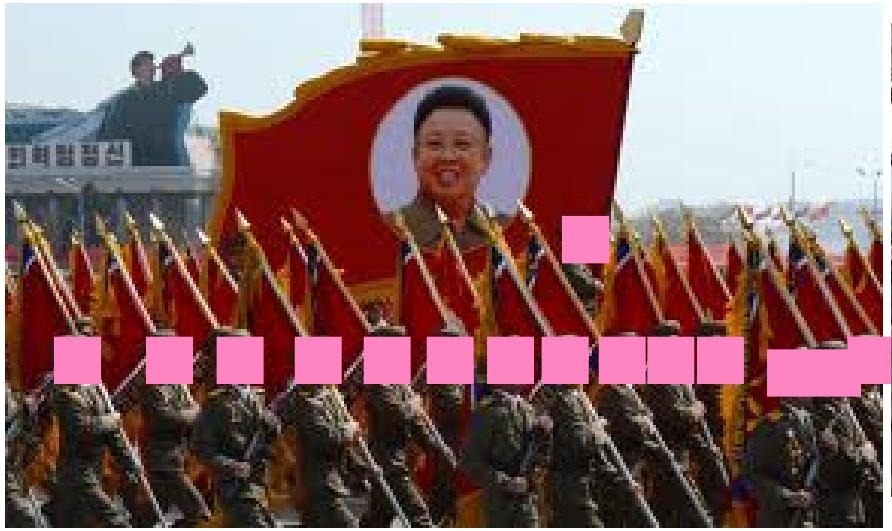
- SDN is easy if control plane is centralized but not necessary. Distributed solutions may be required for legacy equipment and for fail-safe operation.
- Complete removal of control plane may be harmful. Exact division of control plane between centralized controller and distributed forwarders is yet to be worked out
- SDN is easy with a standard southbound protocol like OpenFlow but one protocol may not work/scale in all cases
  - Diversity of protocols is a fact of life.
  - There are no standard operating systems, processors, routers, or Ethernet switches.
- If industry finds an easier way to solve the same problems by another method, that method may win. E.g., ATM vs. MPLS.

How to SDN?

Separation vs. Centralization

Separation of Control Plane

Centralization of Control ~~Plane~~

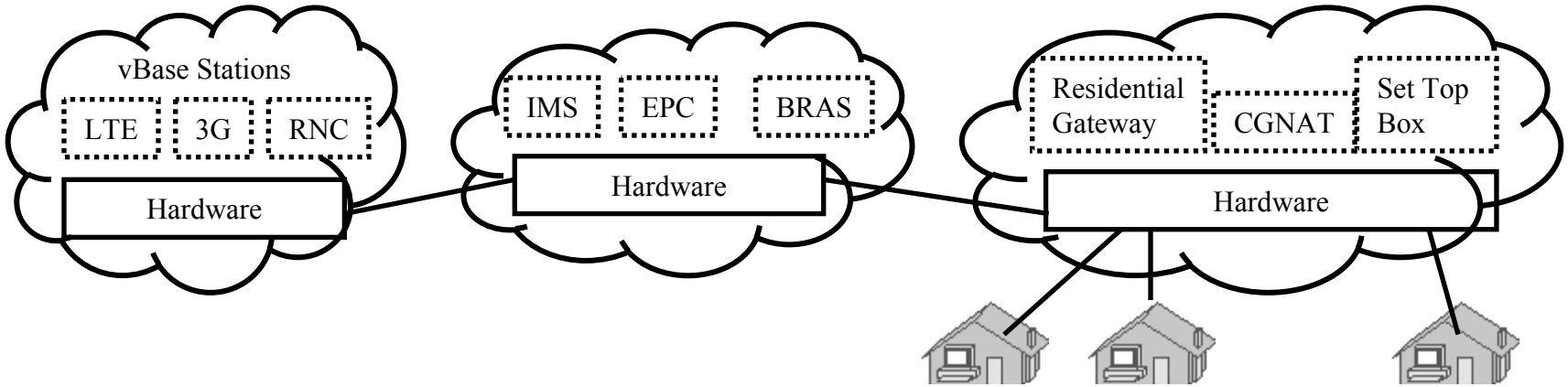


Micromanagement is not scalable

5. Network Function Virtualization (NFV)

1. Fast standard hardware ⇒ **Software based Devices**
   Routers, Firewalls, Broadband Remote Access Server (BRAS) ⇒ A.k.a. *white box* implementation

2. **Virtual Machine implementation**
   ⇒ Virtual appliances
   ⇒ All advantages of virtualization (quick provisioning, scalability, mobility, Reduced CapEx, Reduced OpEx, …)

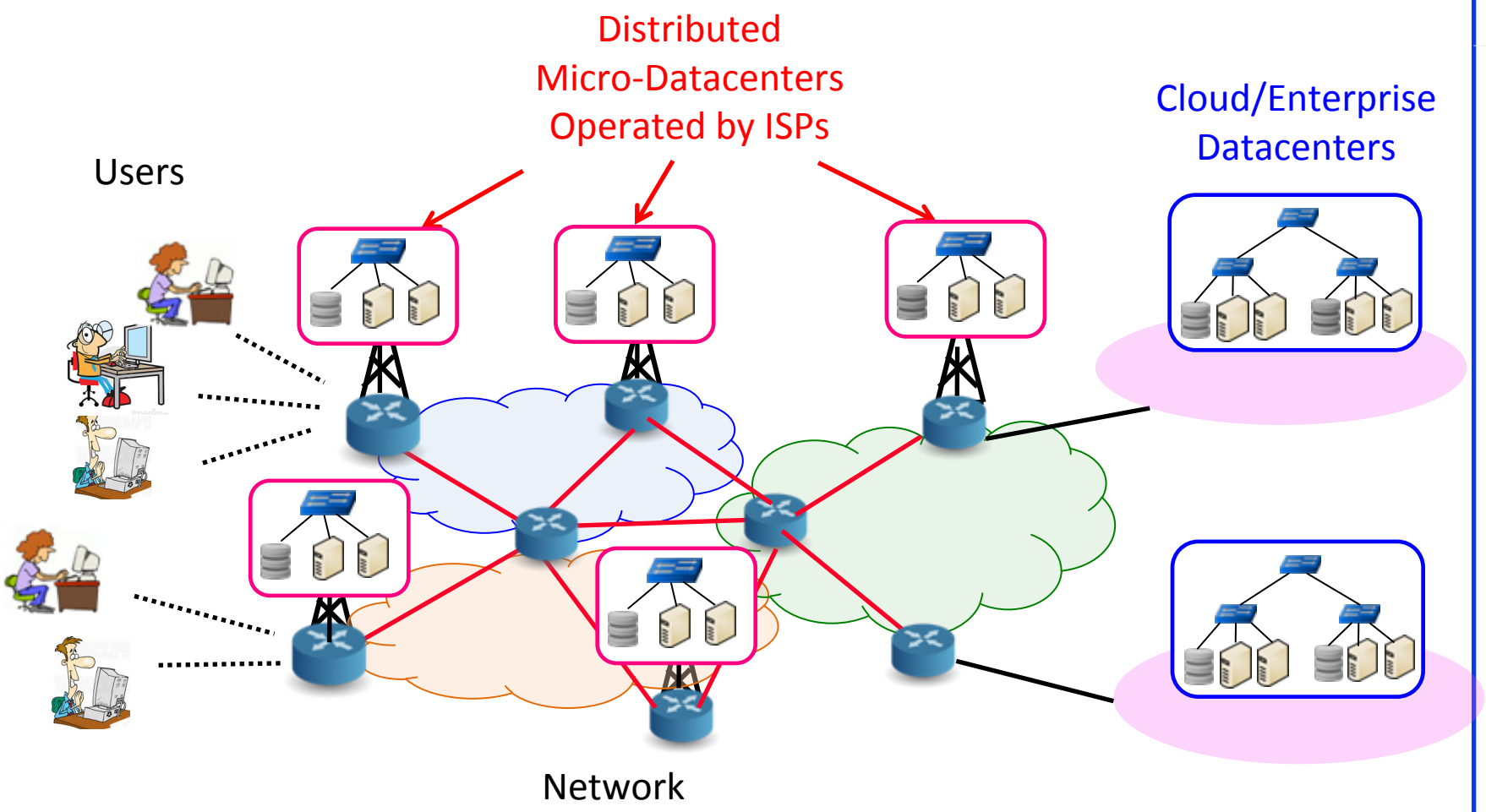


Ref: ETSI, "NFV – Update White Paper," Oct 2013, http://www.tid.es/es/Documents/NFV_White_PaperV2.pdf (Must read)

Service-Infrastructure Separation

- With cloud computing, anyone can super-compute on demand.
  - Physical infrastructure is owned by Cloud Service Provider (CSP). Tenants get virtual infrastructure
  - **Win-Win** combination
- With virtualization, an ISP can set up all virtual resources on demand
  - Physical Infrastructure owned by NFV infrastructure service provider (NSP) and tenant ISPs get virtual NFVI services
  - **Win-Win** combination

Micro-Clouds on Cell-Towers



New Business Opportunities: Domain 2.0, Datacenters on Towers, IoT, NFV, FV, Elastic Networks

Any Function Virtualization (FV)

- Network function virtualization of interest to Network service providers
- But the same concept can be used by any other industry, e.g., financial industry, banks, stock brokers, retailers, mobile games, …
- Everyone can benefit from:
  - Functional decomposition of there industry
  - Virtualization of those functions
  - Service chaining those virtual functions (VFs)
    ⇒ A service provided by the next gen ISPs

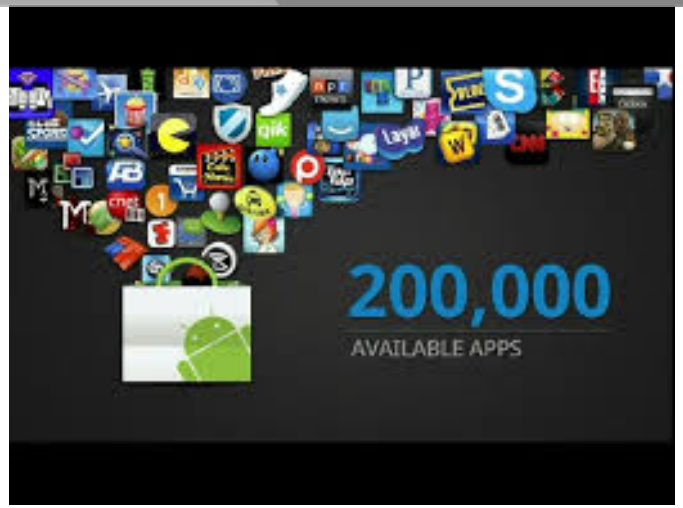
Carrier App Market: Lower CapEx

Virtual IP Multimedia System

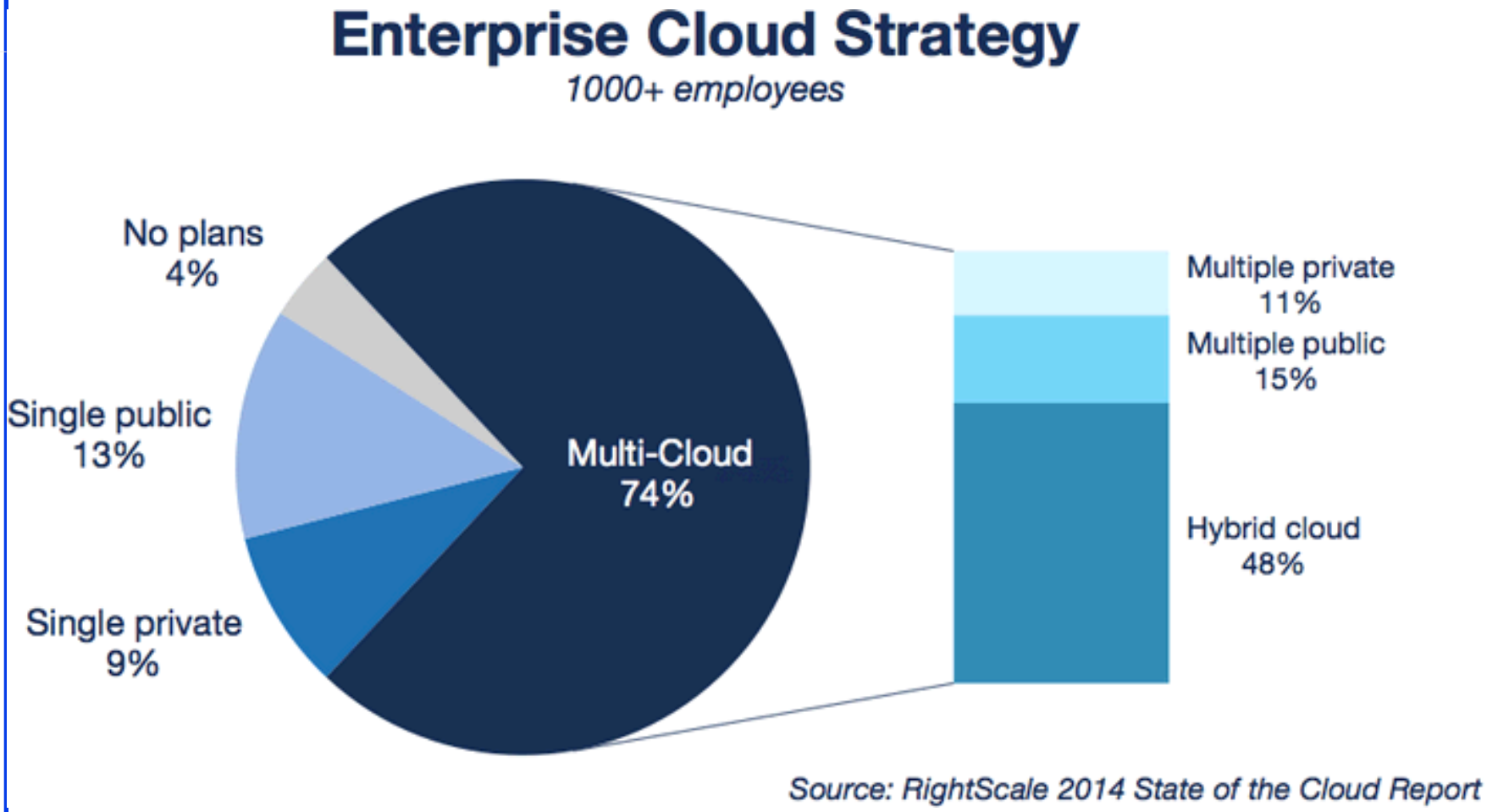
Available on the App Store

amazon.com® and you're done.™ App Store

ANDROID



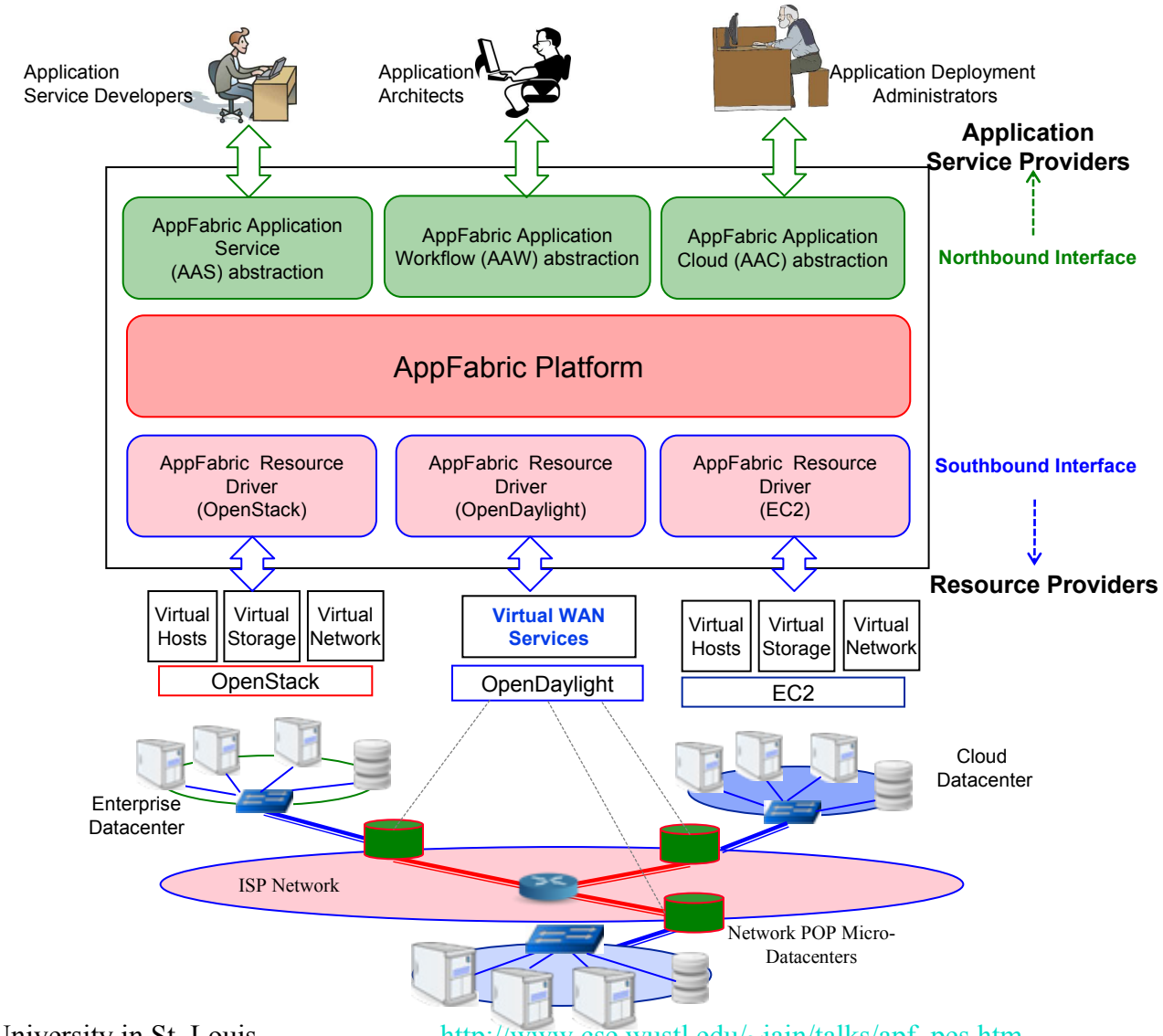
Trend: Multi-Clouds



- Most companies use more than one cloud.

Ref: http://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2014-state-cloud-survey

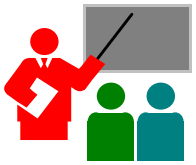
Services in a Cloud of Clouds



10 SDN Research Issues

1. Centralization ⇒ Reliability ⇒ Distributed Controllers, Controller Synchronization
2. Performance of Controllers: Scalability, Caching
3. Multi-controller Load balancing, Latency Minimization
4. Security in the Control Plane: Confidentiality, Integrity, Authentication, Monitoring, Detection, Recovery, Trust
5. SDN in a Multi-Domain Environment: Hierarchical Organization of Policy Control
6. SDN in Specific Applications: High-Performance Computing, Network Virtualization, Big Data, IoT
7. Live traffic monitoring and fault detection in the Data Plane
8. Rules consistency checking
9. Live network reconfiguration and optimization
10. Security in data plane

Note: This is not a complete list.



Summary

1. Virtualization is revolutionizing networking. NFV allows virtual mobile services using virtual modules in a shared cloud environment ⇒ Key to CapEx OpEx reduction.
2. SDN is about centralized policy control. Separation of control plane is not necessary.
3. Virtual functions useful not only for networking but also for **all other global enterprises** and games
   ⇒ New business opportunity for FV Infrastructure service
4. **AppFabric** allows customers to select **multiple clouds** from different providers and **share wide area network** infrastructure and specify their policies

References

- Raj Jain and Subharthi Paul, "**Network Virtualization and Software Defined Networking for Cloud Computing - A Survey**," IEEE Communications Magazine, Nov 2013, pp. 24-31, http://www.cse.wustl.edu/~jain/papers/net_virt.htm