Five Trends Leading to Opportunities in Multi-Cloud Global Application Delivery

RAJ JAIN

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

Research Seminar at Cisco, San Jose, CA
March 18, 2016

These slides are available on-line at:

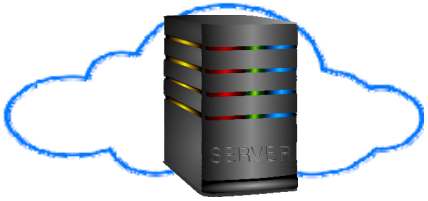
http://www.cse.wustl.edu/~jain/talks/cisco16.htm

Overview

- Why Multi-Cloud?
  - 1. Cloud Computing
  - 2. Software Defined Networking (SDN)
  - 3. Mobile Traffic Explosion
  - 4. Smart Everything
  - 5. Multi-Cloud Hierarchy
- OpenADN Multi-Cloud Management
- Service Function Placement Problem

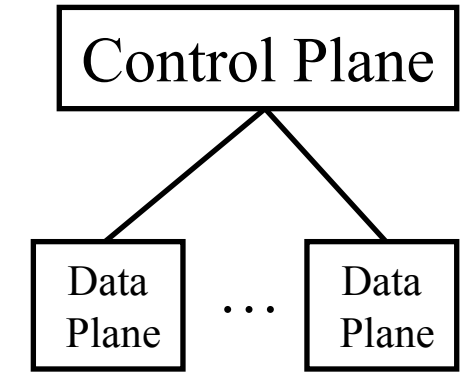
1. Cloud Computing

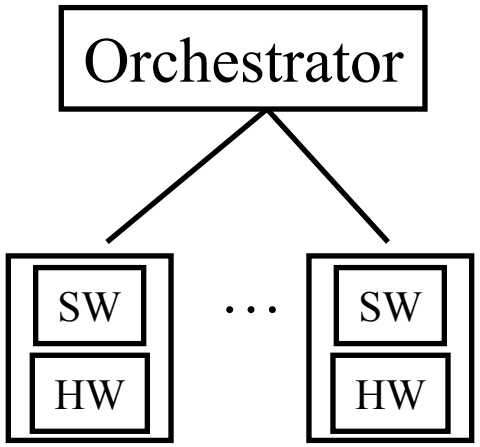
- Cloud computing was invented in 2006
- Then: Cloud = Large Data Center Multiple VMs managed by a cloud management system (OpenStack)
- Today: Cloud = Computing using virtual resources
  - µCloud = Cloud in a server with multiple VMs.
  - Each VM with Multiple Containers ⇒ Multiple Services



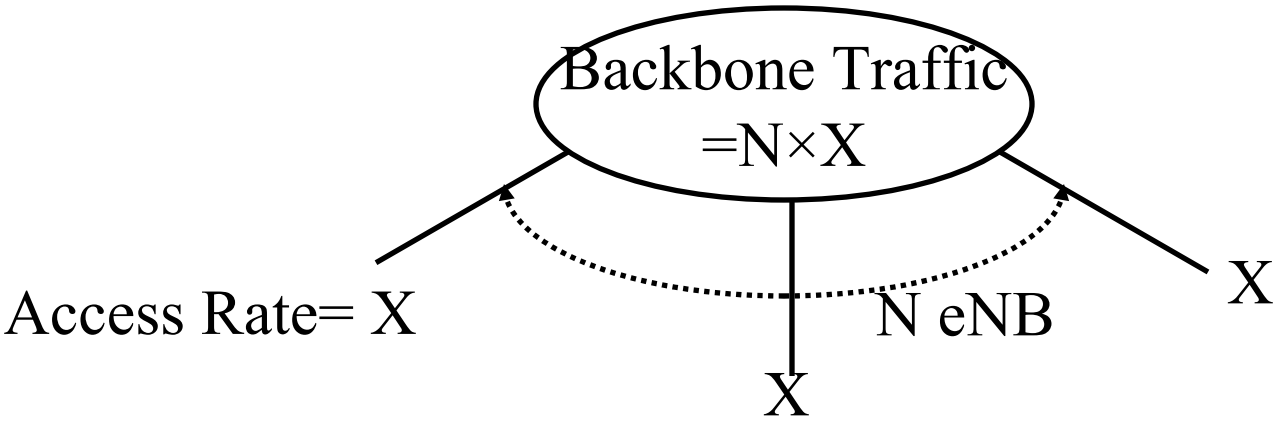
2. Software Defined Networking (SDN)

- SDN was invented in 2009
- Then: SDN:
  - OpenFlow Southbound
  - Separation of control and data planes
  - Centralization of Control
- Now: SDN = **Disaggregation** of hw/sw
  - Commodity hardware
  - Software that runs on commodity hw
  - Open Source Software
    ⇒ Service industry
  - Controller replaced by Orchestrator



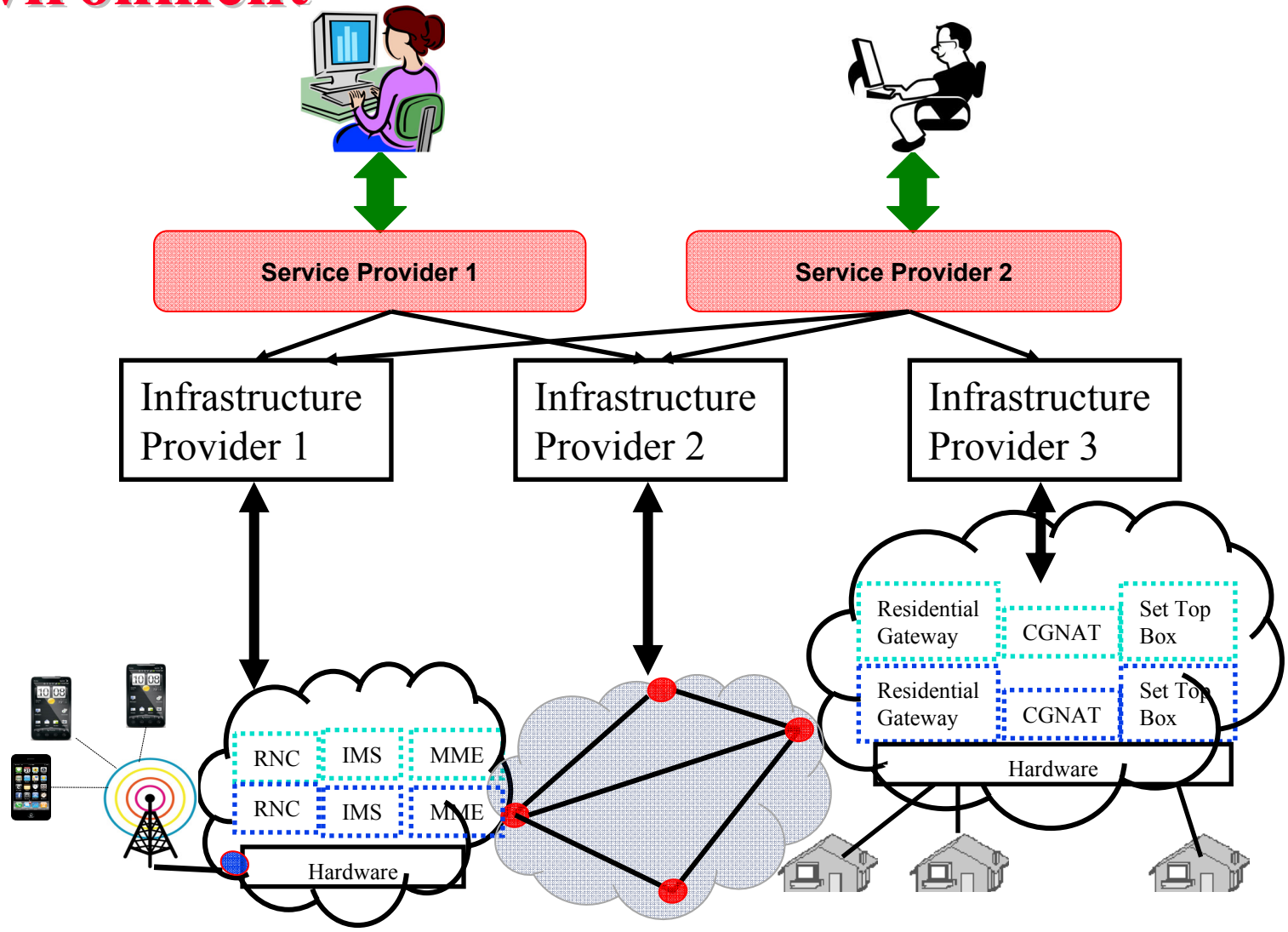


3. Mobile Traffic Explosion



- With **small cells**, N=# of edge points can be large
- While 5G radio access technologies may see only 1000× increase, backhaul may see much larger increase in capacity, highly dynamic, local
- Need dynamic capacity management
  ⇒ Network Function Virtualization (NFV)
- Need aggregation ⇒ Computing in the Edge

NFV in a Multi-Cloud Multi-Tenant Environment



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) or **Apps**

Networking App Market: Lower CapEx

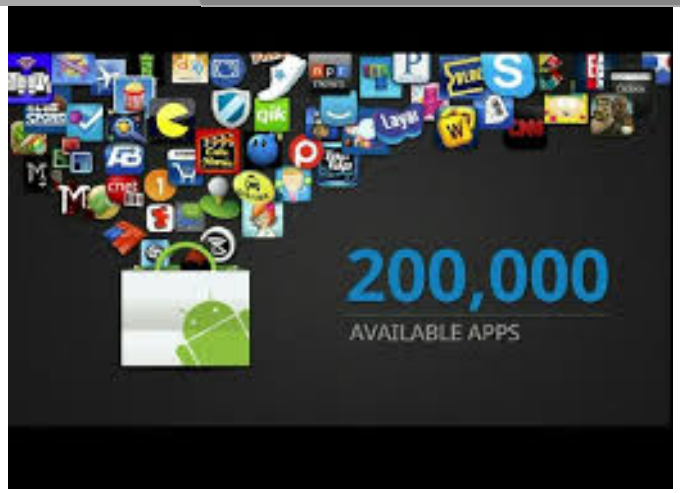
Virtual IP Multimedia System

Available on the App Store

CISCO

App Store

vRouter vRADIUS vVPN

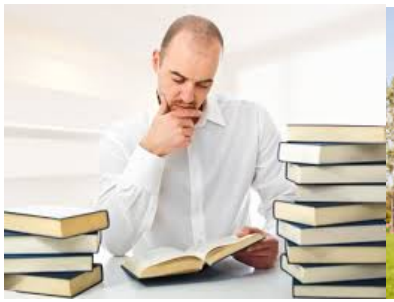


4. Smart Everything

- What's Smart?
  - Old: Smart = Can think fast ⇒ High compute power
  - Then: Smart = Can remember everything ⇒ High storage
  - Now: Smart = Can communicate ⇒ Good Networking
- Smart Grid, Smart Meters, Smart Cars, Smart homes, Smart Cities, Smart Factories, Smart Smoke Detectors, …

Think

Communicate

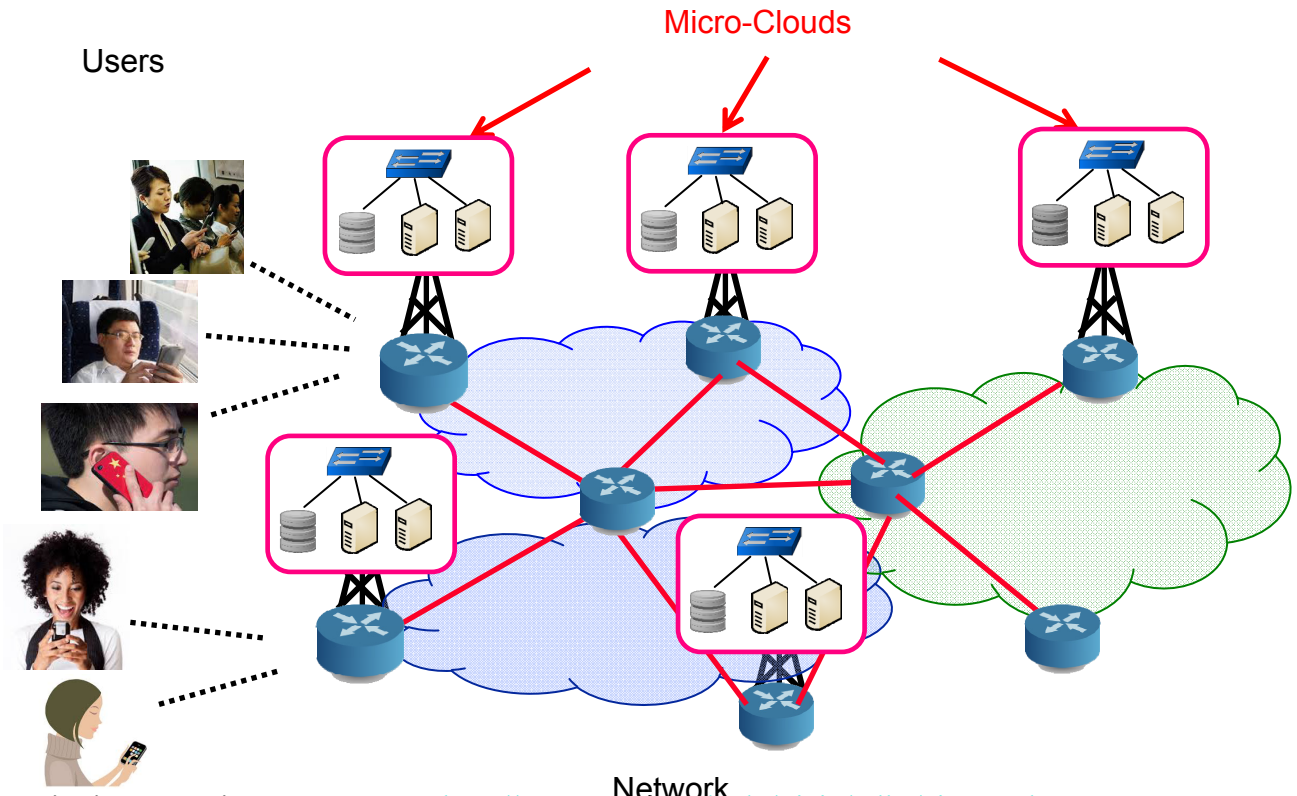


Not-Smart

Smart

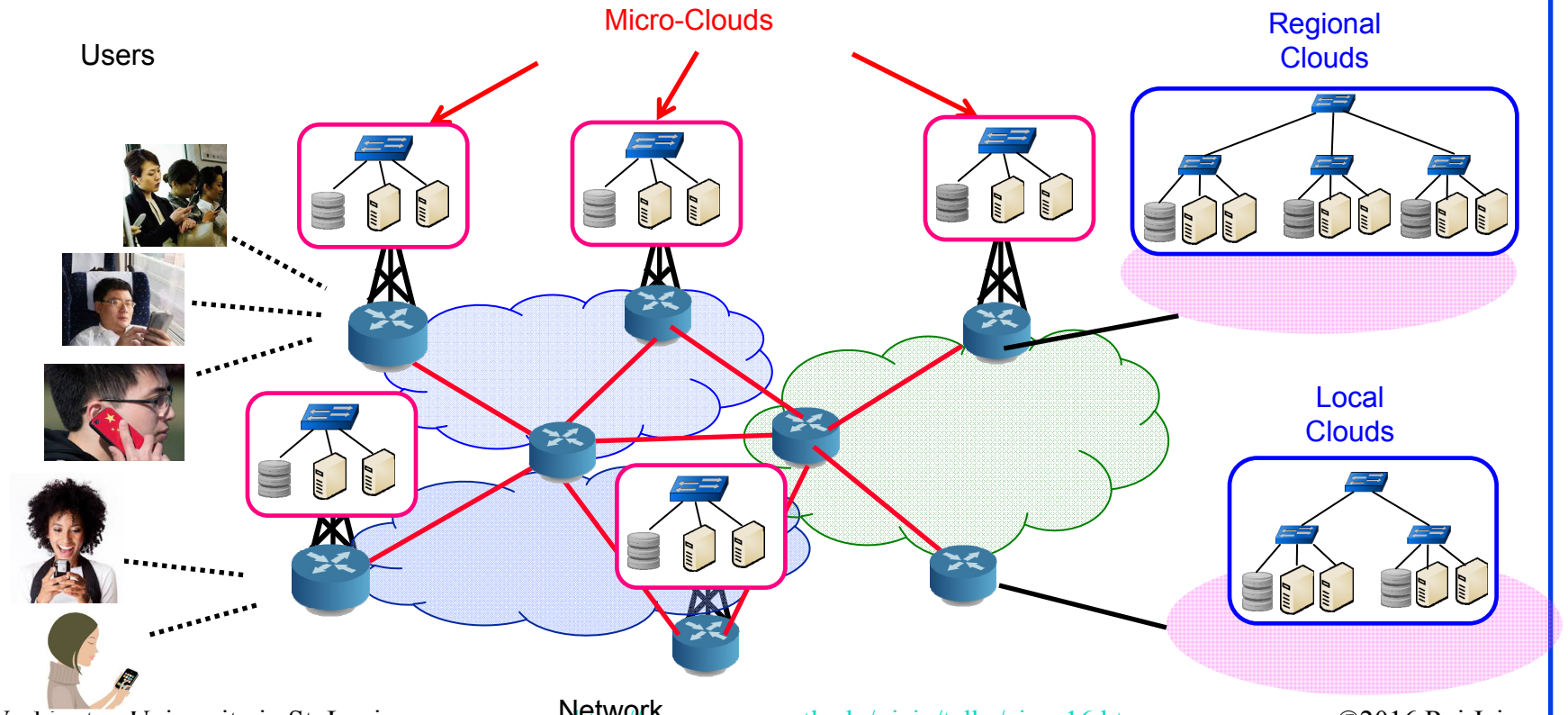
Mobile Edge Computing

- To service mobile users/IoT, the computation needs to come to edge ⇒ Mobile Edge Computing

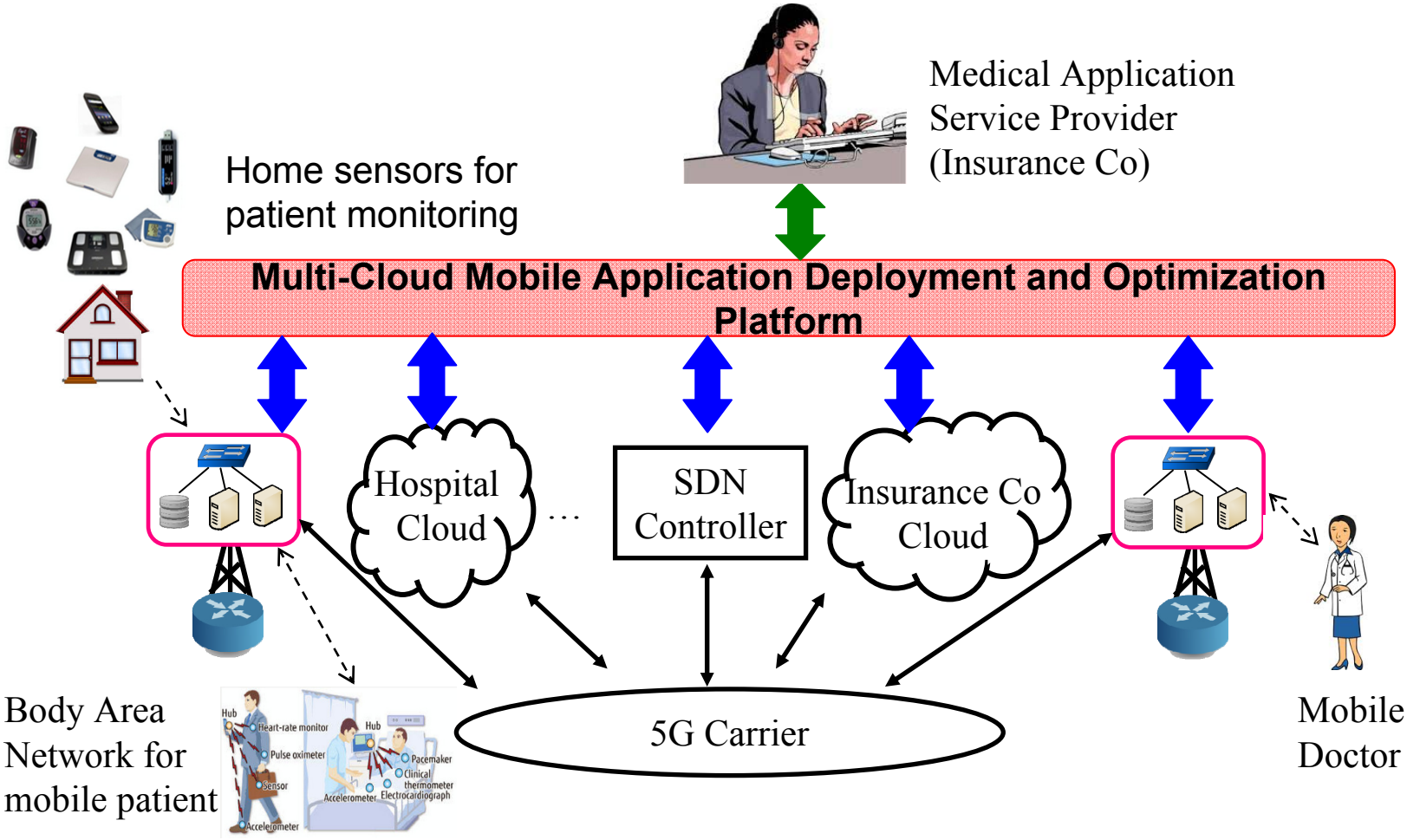


5. Multi-Cloud Hierarchy

- Wide area clouds, local area clouds (home routers with cloud features), Personal area clouds (cars), body area clouds (smart phone) ⇒ Fog Computing

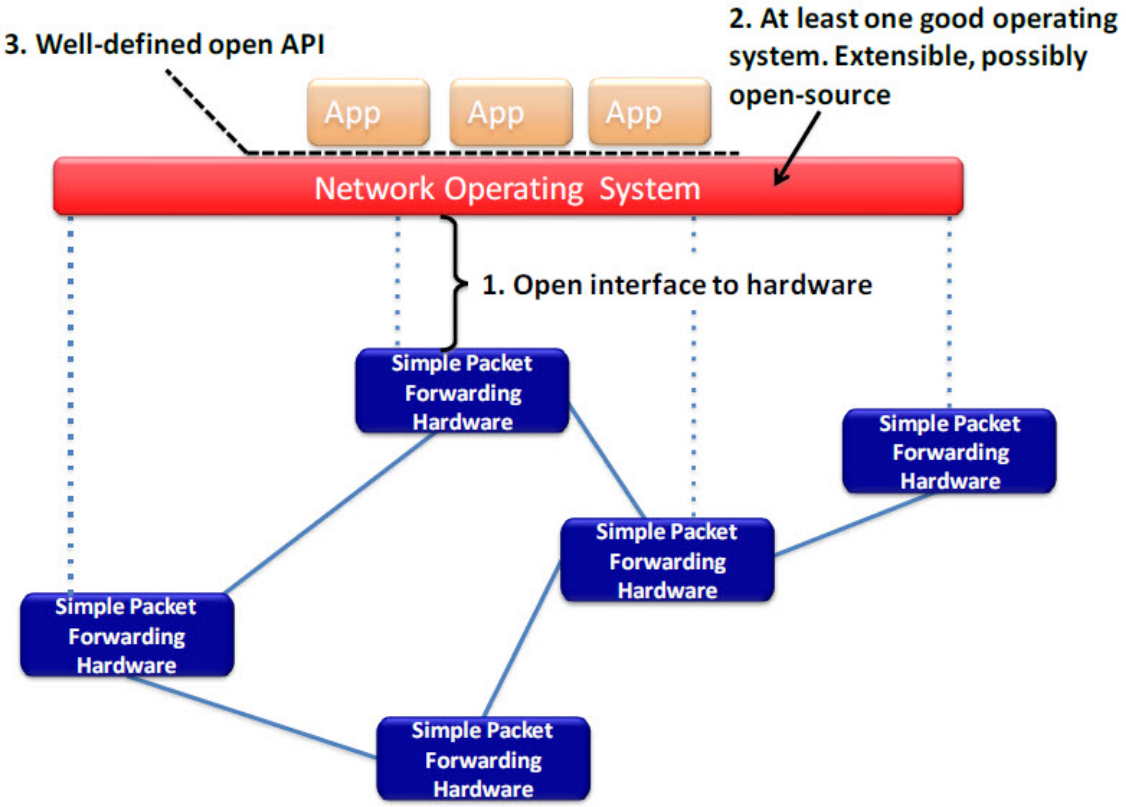


Mobile Healthcare Use Case



Domain 2.0

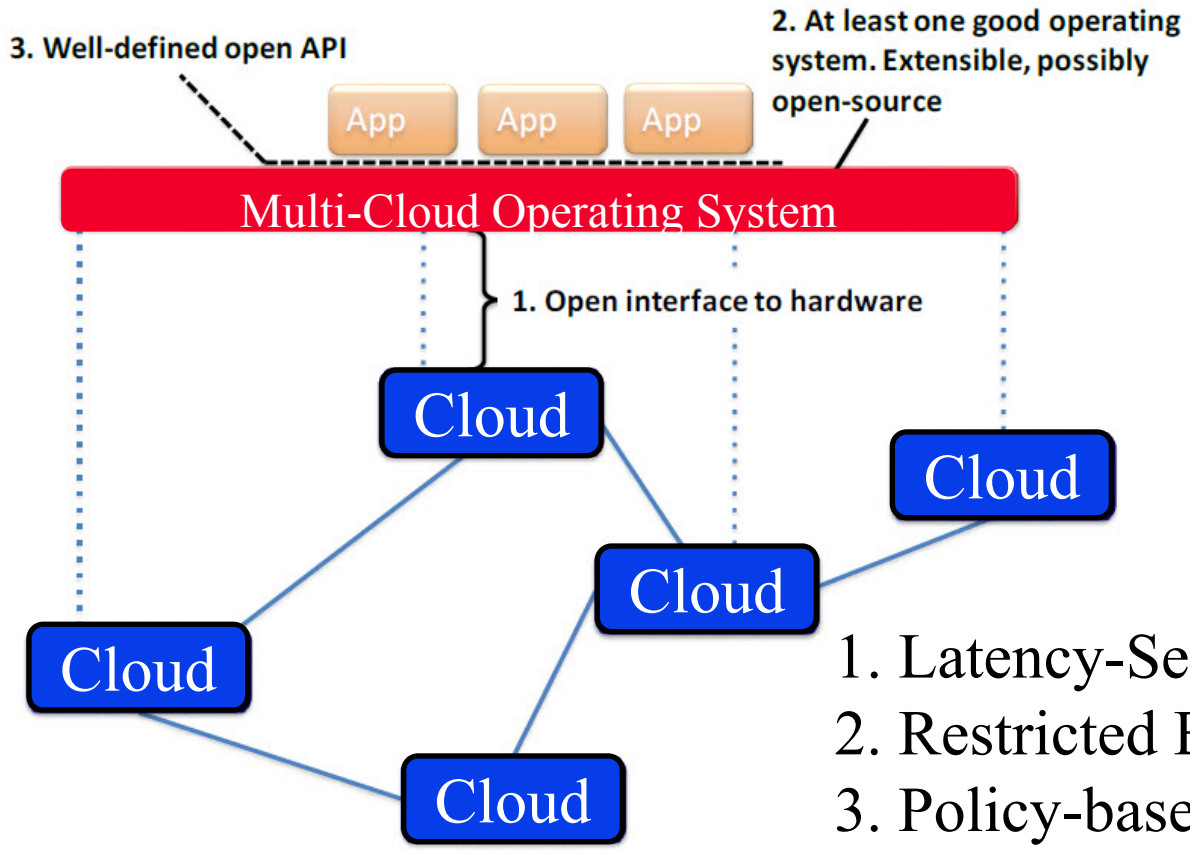
- Today: Virtualization of switches



Ref: AT&T, "Domain 2.0 White paper,"
https://www.att.com/Common/about_us/pdf/AT&T%20Domain%202.0%20Vision%20White%20Paper.pdf

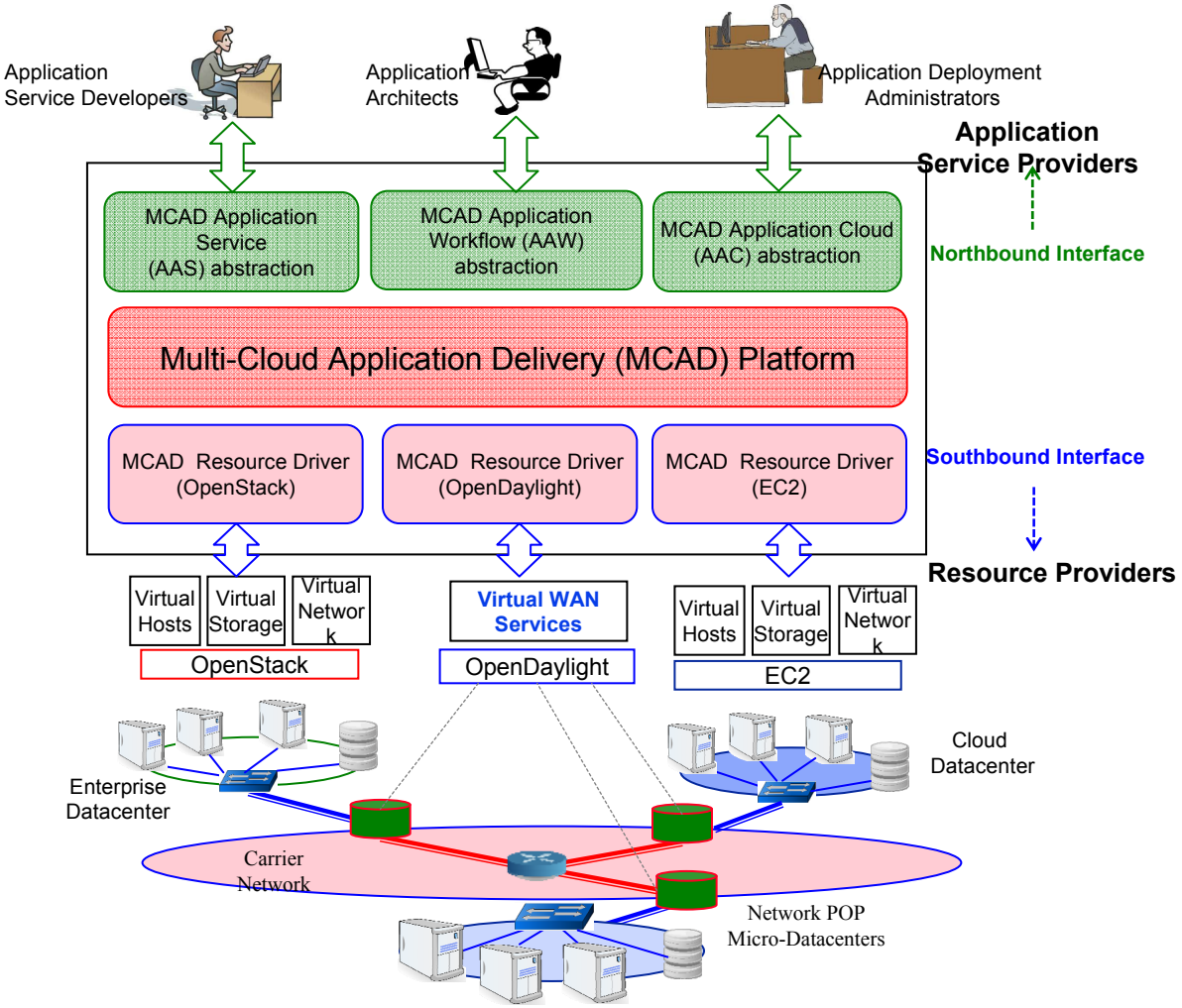
Domain 2.1

- Tomorrow: Virtualization of Clouds



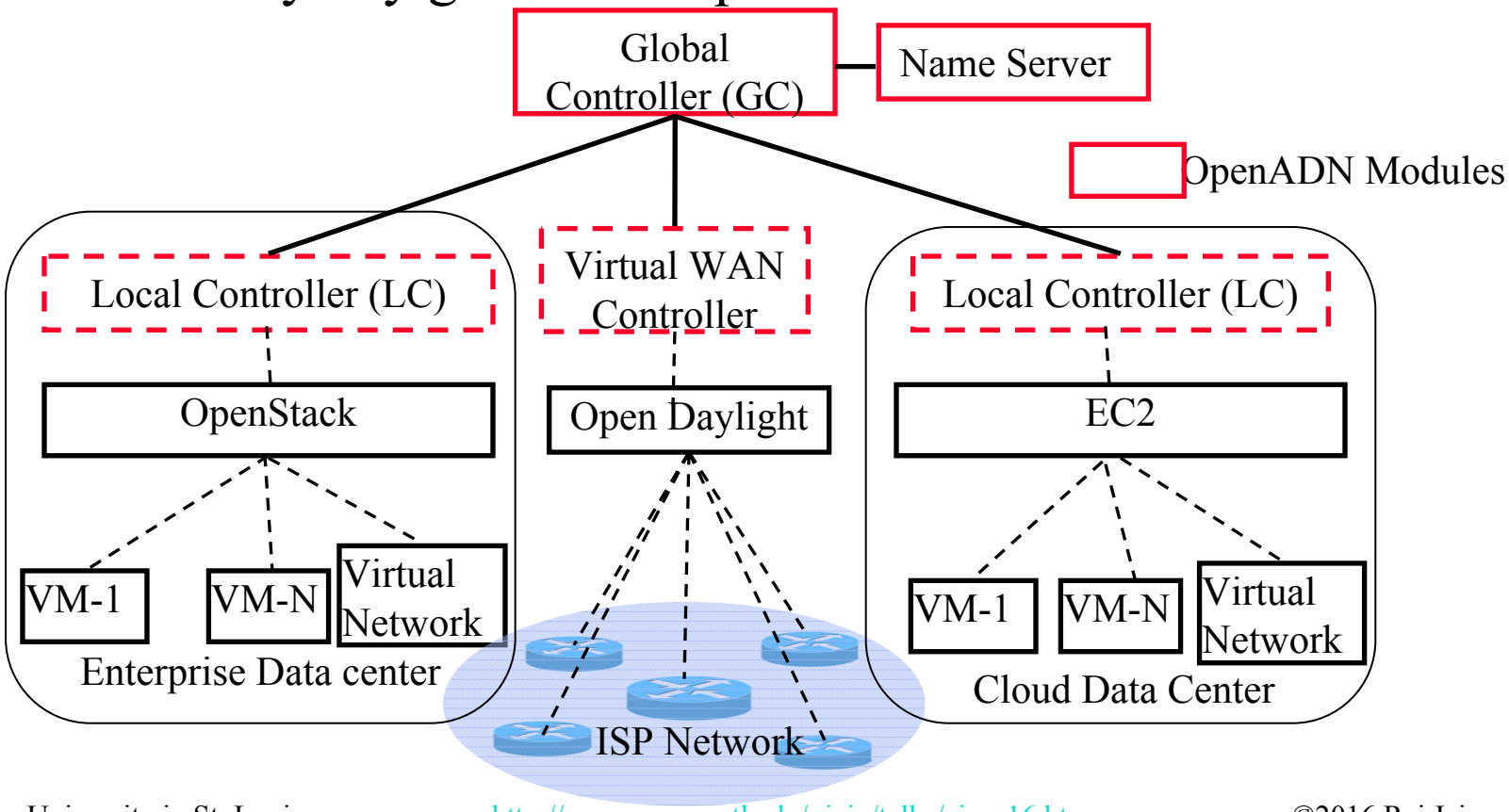
1. Latency-Sensitive
2. Restricted Bandwidth
3. Policy-based automation

OpenADN Multi-Cloud Management



Key aspects of OpenADN Architecture

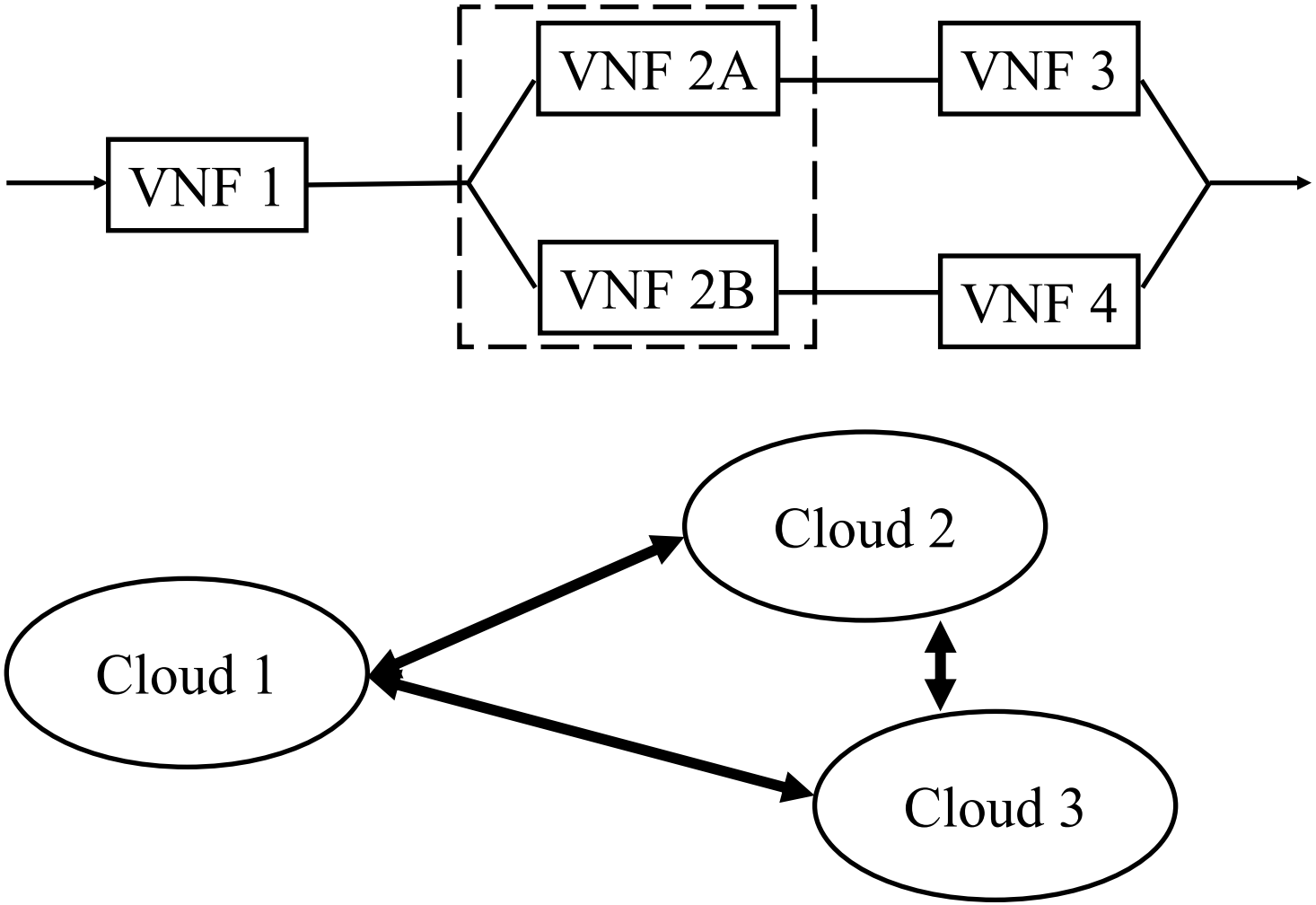
- Global and local controllers
- All services, servers, instances accessed by name
- Can be used by any global enterprise or carrier



Challenges in Multi-Cloud Deployment

- **Dynamic**: Forwarding changes with state of the servers, links
- **Heterogeneous**: Different cloud providers, different services, different policies
- **Distributed Control**:
  - Equipment belongs to infrastructure provider
  - Data belongs to Tenants
- **Massive Scale**: Millions of enterprise applications sharing networks provided by many ISPs using cloud services from many CSPs

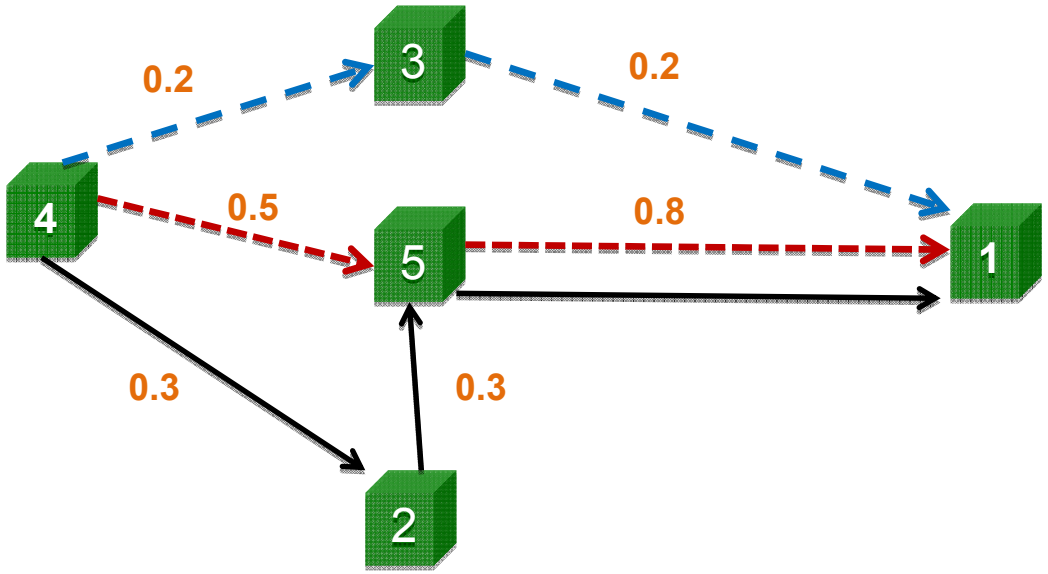
Service Function Placement Problem



Challenges in Service Placement

- **Delay constraints**
- **WAN links bottleneck:** Need to model link queues
- **Complexity**: NP-complete ⇒ Need efficient heuristics
- **Affinity:** VNF1 and VNF2 should be co-located
  - Significant communication exchanges
  - Duplicate memory pages in VMs (same OS and Libraries)
- **Anti-Affinity**: VNF1 and VNF2 should not be placed on the same physical server.
  - CPU-intensive applications
  - VMs belonging to different users in a cloud may cause security risk such as cross-VM attacks
  - Duplicate VMs used to improve fault tolerance and availability

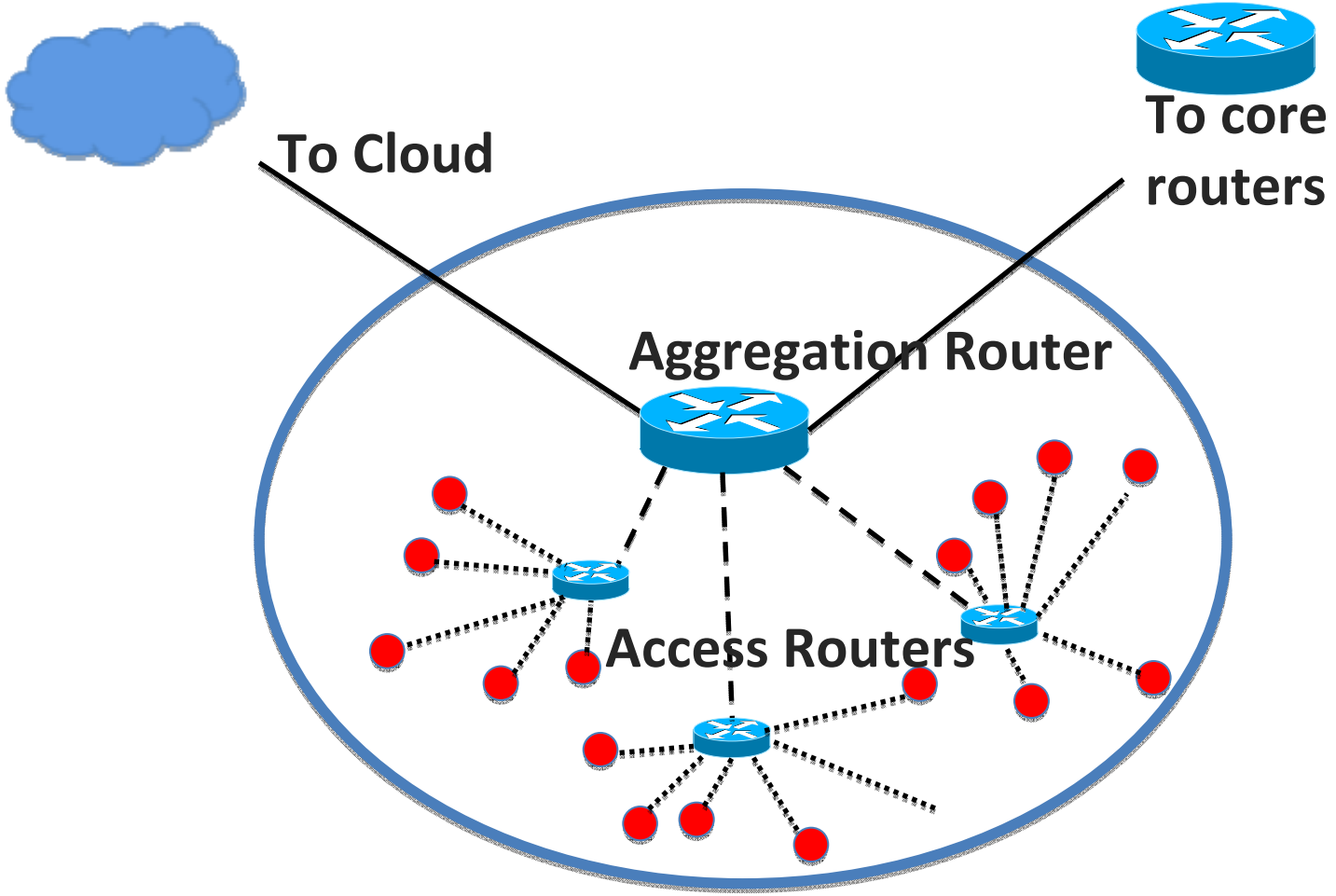
Sample Service Chains



- 5 VFs and 3 possible service chains based on business logic

User Clusters

- A cluster of end users are modeled as one

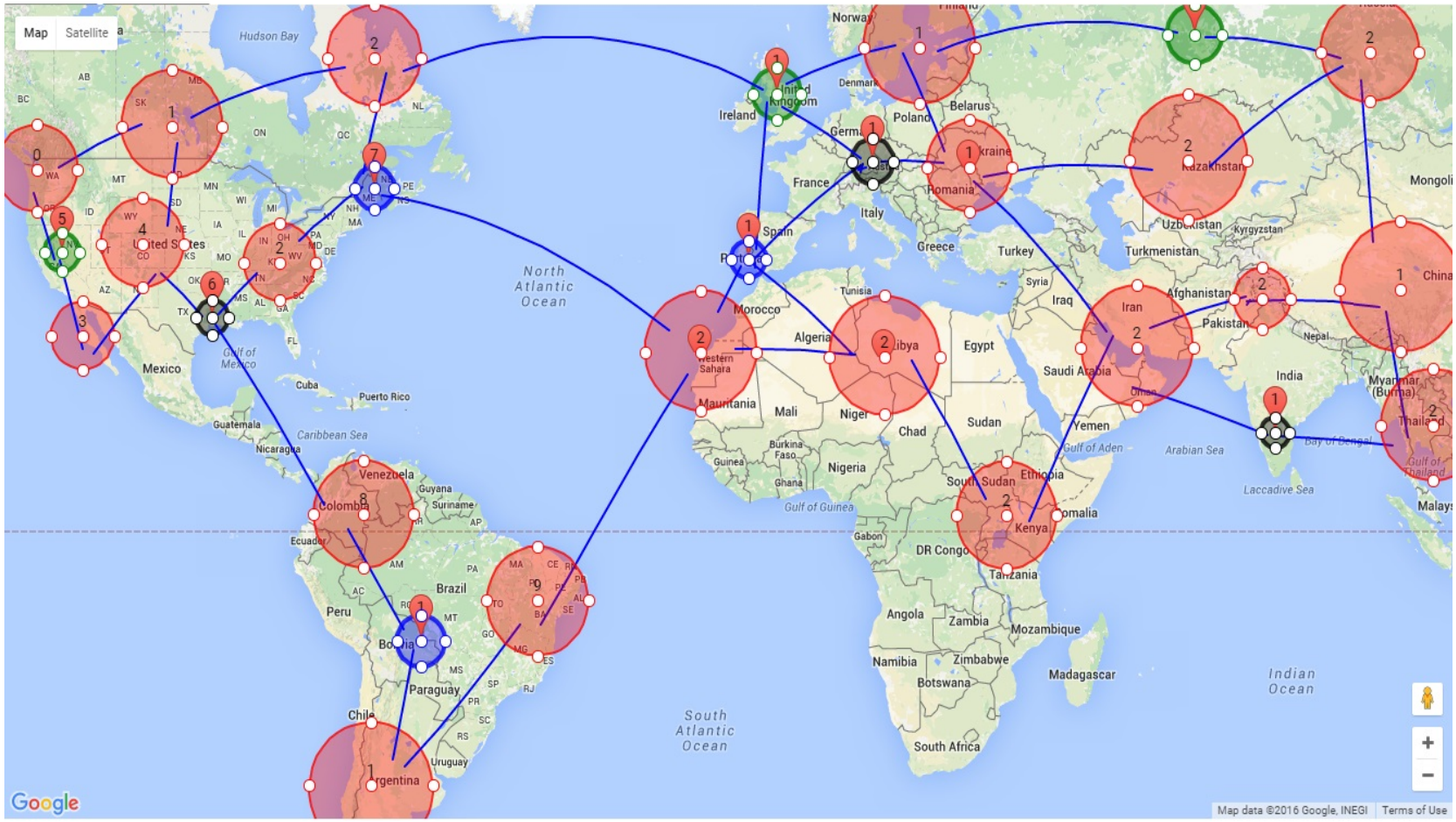


VM configurations

- **Resource configurations from Amazon EC2**

Configuration	Memory (GB)	Compute Unit	Disk (GB)	Platform (bit)	Cost ($/h)
m1.samll	1.7	1	160	32	0.1
m1.large	7.5	4	850	64	0.4
m1.xlarge	15	8	1690	64	0.8
c1.medium	1.7	5	350	32	0.2
c1.xlarge	7	20	1690	64	0.8

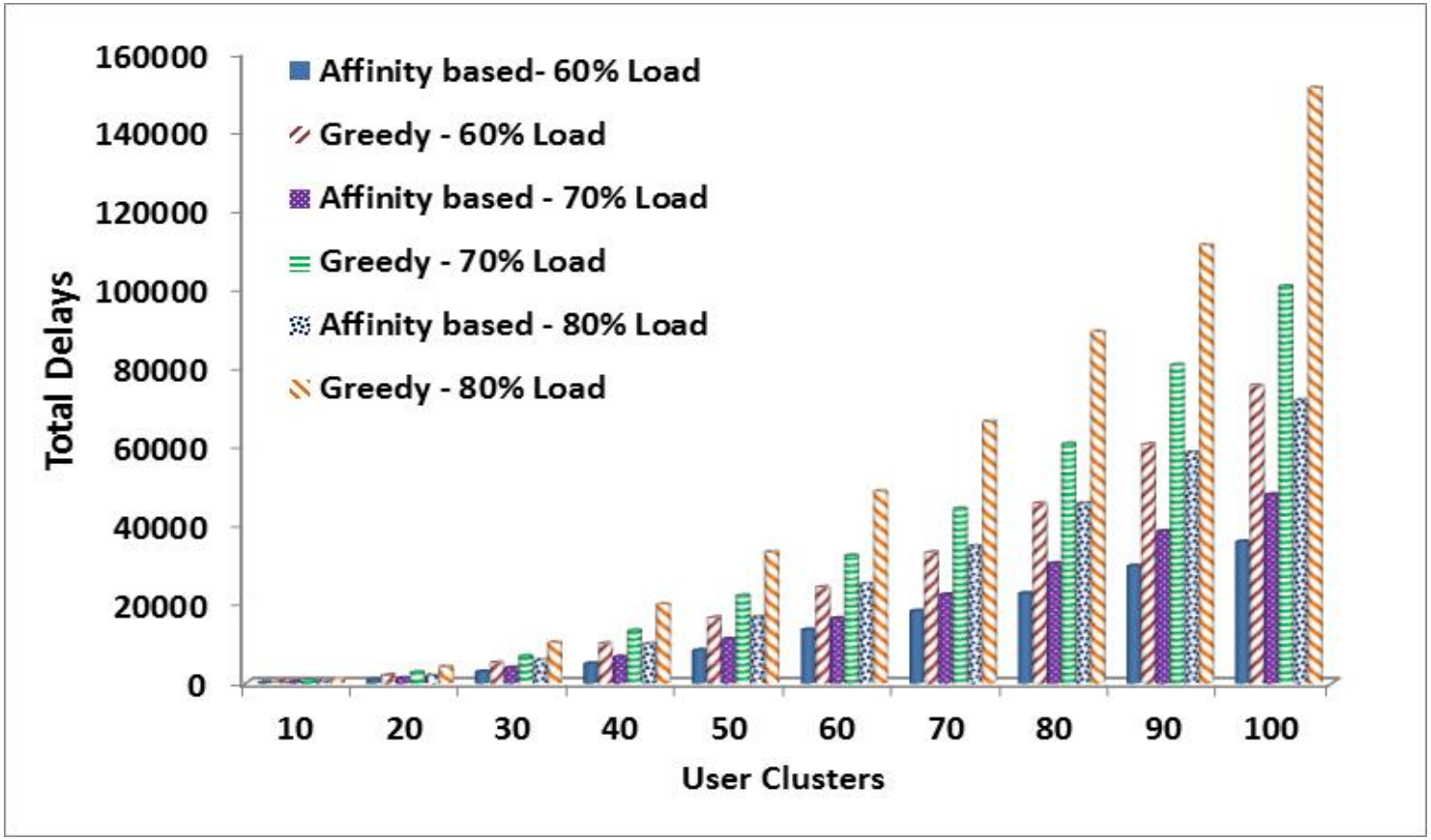
Topology UI


Map
Satellite
Hudson Bay
North Atlantic Ocean
South Atlantic Ocean
Indian Ocean
Caribbean Sea
Gulf of Mexico
Gulf of Guinea
Gulf of Aden
Arabian Sea
Laccadive Sea
Mexico
Cuba
Puerto Rico
Guatemala
Nicaragua
Venezuela
Colombia
Ecuador
Peru
Brazil
Paraguay
Chile
Uruguay
Argentina
Guyana
Suriname
Ireland
France
Spain
Italy
Greece
Turkey
Poland
Belarus
Romania
Norway
Denmark
Morocco
Algeria
Tunisia
Libya
Egypt
Mauritania
Mali
Niger
Chad
Sudan
Nigeria
Ghana
Guinea
Burkina Faso
Western Sahara
South Sudan
Ethiopia
Kenya
Somalia
Gabon
DR Congo
Tanzania
Angola
Zambia
Mozambique
Namibia
Zimbabwe
Botswana
Madagascar
South Africa
Saudi Arabia
Yemen
Iraq
Iran
Syria
Afghanistan
Pakistan
India
Nepal
Kazakhstan
Uzbekistan
Turkmenistan
Kyrgyzstan
Mongolia
China
Myanmar (Burma)
Thailand
Malaysia
Google
Map data ©2016 Google, INEGI
Terms of Use

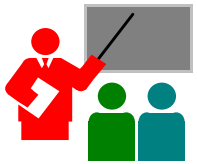
Analysis Results

- Integer Linear Programming (ILP) to find the optimal solution
  - Exponential complexity (NP-Complete)
  - Takes days ⇒ Not usable in real-time deployment scenarios
- Used two heuristics:
  - Greedy (First Fit)
  - Affinity-based approach

Results: Greedy vs. Affinity-based



- For cluster size 100 at 60% traffic load, total delays are 486.75 s for Greedy vs. 209.25s for Affinity



Summary

1. Clouds getting smaller, SDN definition changing to disaggregation, Carriers and enterprises moving to clouds, Internet of things are leading to clouds everywhere => multi-cloud applications.
2. Our multi-cloud application management system (MCAD) allows policy-based deployment and management of multi-cloud application. Handles heterogeneous clouds and respects resource ownerships
3. Service function placement problem is NP complete. Challenges included delay constraints, WAN Link bottlenecks, and affinity
4. We have developed an efficient heuristic for placement that is 2X faster than greedy based heuristics.

References

- Deval Bhamare, Raj Jain, Mohammed Samaka, Gabor Vaszkun, Aiman Erbad, "Multi-Cloud Distribution of Virtual Functions and Dynamic Service Deployment: OpenADN Perspective," Proceedings of 2nd IEEE International Workshop on Software Defined Systems (SDS 2015), Tempe, AZ, March 9-13, 2015, 6 pp. http://www.cse.wustl.edu/~jain/papers/vm_dist.htm
- Subharthi Paul, Raj Jain, Mohammed Samaka, Aiman Erbaud, "Service Chaining for NFV and Delivery of other Applications in a Global Multi-Cloud Environment," ADCOM 2015, Chennai, India, September 19, 2015, http://www.cse.wustl.edu/~jain/papers/adn_in15.htm
- Raj Jain, Mohammed Samaka, "Application Deployment in Future Global Multi-Cloud Environment," The 16th Annual Global Information Technology Management Association (GITMA) World Conference, Saint Louis, MO, June 23, 2015, http://www.cse.wustl.edu/~jain/papers/apf_gitp.htm

References (Cont)

- Subharthi Paul, Raj Jain, Mohammed Samaka, Jianli Pan, "Application Delivery in Multi-Cloud Environments using Software Defined Networking," Computer Networks Special Issue on cloud networking and communications, Available online 22 Feb 2014, http://www.cse.wustl.edu/~jain/papers/comnet14.htm
- 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

Acronyms

- ATM Asynchronous Transfer Mode
- ECN Explicit congestion notification
- EFCI Explicit Forward Congestion Indication
- FECN Forward Explicit Congestion Notification
- GB Gigabyte
- IEEE Institution of Electrical and Electronic Engineering
- IETF Internet Engineering Task Force
- IoT Internet of Things
- IP Internet Protocol
- IRTF Internet Research Task Force
- ITU International Telecommunications Union
- LAN Local Area Network
- LTE Long Term Evolution
- MHz Mega Hertz
- OpenADN Open Application Delivery Networking
- SDN Software Defined Networking

Acronyms (Cont)

- TCP Transmission Control Protocol
- TV Television
- VM Virtual Machine
- WAN Wide Area Network
- WiFi Wireless Fidelity
- WiMAX Worldwide Interoperability for Microwave Access