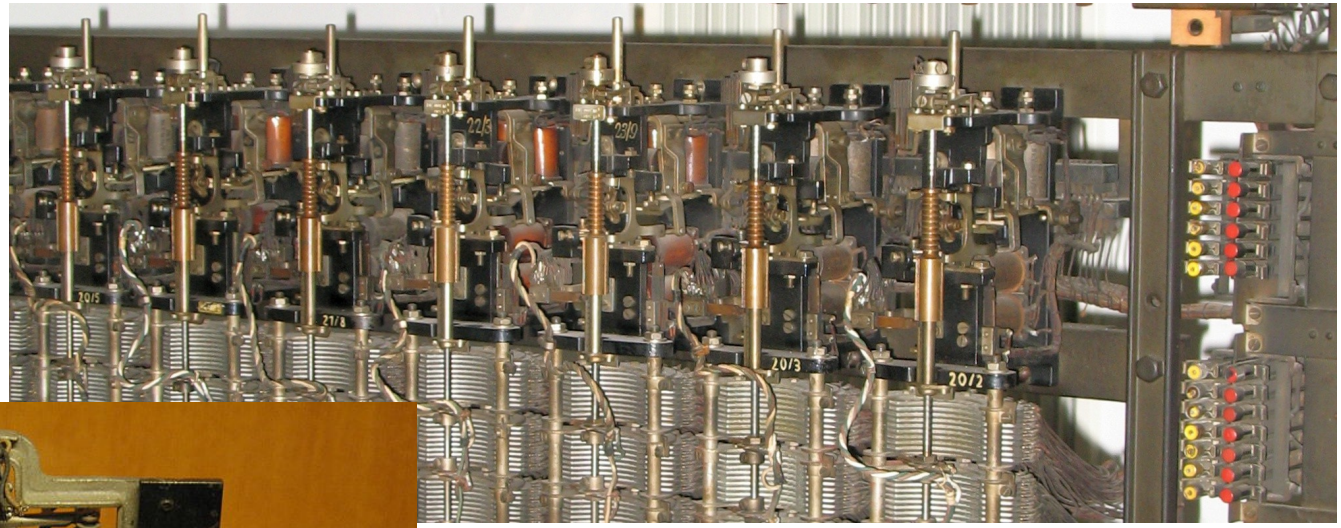
Networking in the cloud: An SDN primer

Ben Cherian
Chief Strategy Officer
@bencherian
Midokura


The current state of networking is too manual

Telecom has this problem before

Almon Strowger – mortician, inventor, and possibly paranoid



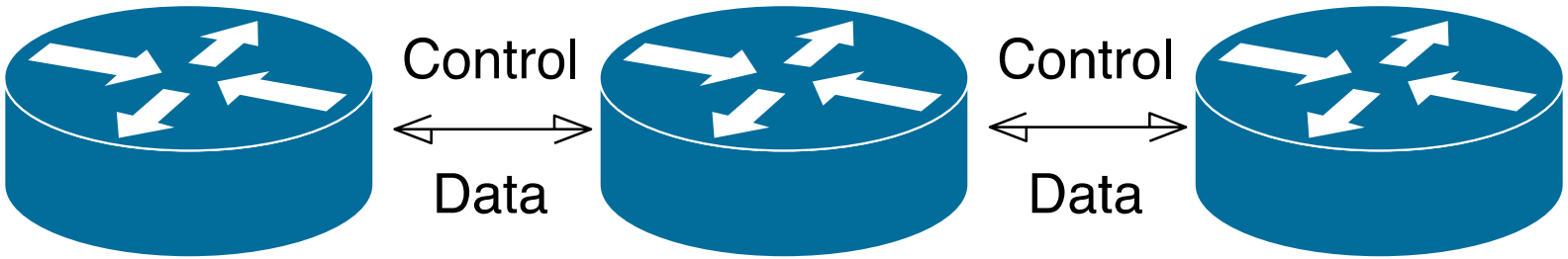
Strowger’s switches

Wanted to solve:
Privacy
Intended human errors

Solved these too:
Unintended human errors
Speed of connections
Lowering operational costs of running a local exchange

What is SDN?

Traditional networks

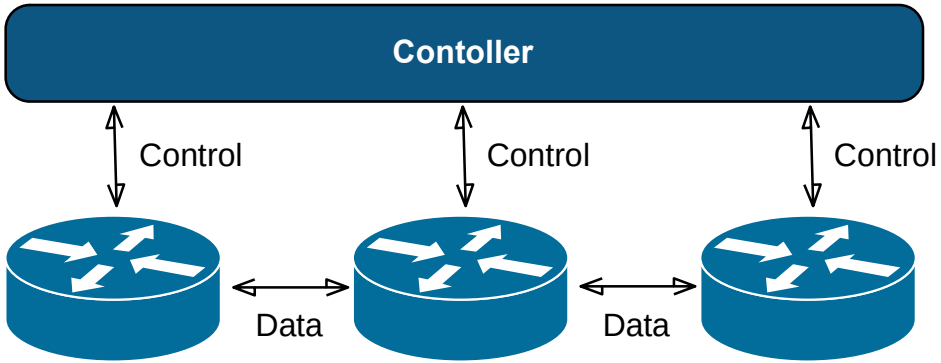


Control Plane: Responsible for making decision on where the traffic is sent

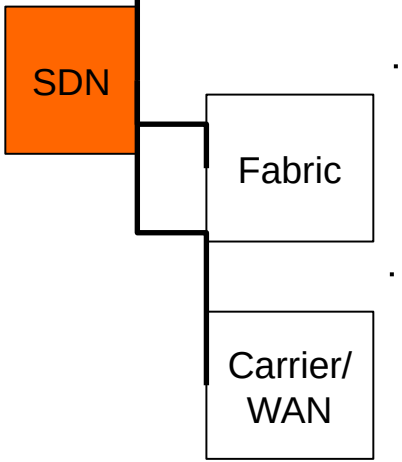
Data Plane:
Responsible for forwarding traffic to the selected destination

The network needs better abstraction

A basic example of SDN



Categories of SDN



	Definition	Example
Cloud Service Providers (CSP / CAP)	· Distributed control plane · Software solution for IaaS Cloud use	· Midokura · VMWare/Nicira · Nuage
Enterprises		
Fabric	· Centralized control plane · Hardware solution for DC use	· Juniper Qfabric · NEC Programmable Flow · Big Switch

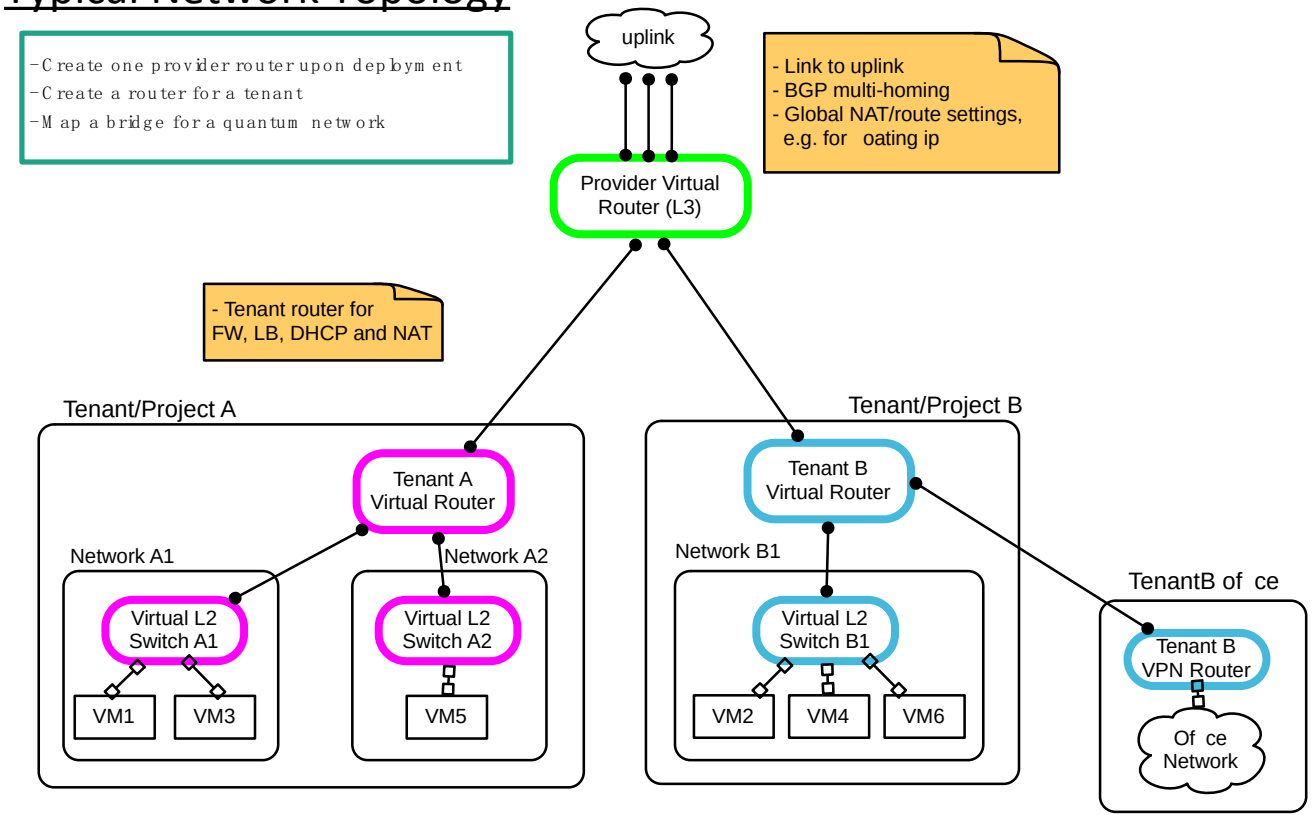
Carrier/ WAN	· Hybrid control plane (CP)
	  · Distributed CP with BGP
	  · Centralized CP using OpenFlow
		· Google

Iaas Cloud Networking Requirements

- Multi-tenancy
- L2 isolation
- L3 routing isolation
  - VPC
  - Like VRF (virtual routing and forwarding)
- Scalable control plane
- ACLs
- Stateful (L4) Firewall
  - Security Groups
- VPN
  - IPSec
- BGP gateway
- REST API
- Integration with CMS

Iaas Cloud Networking Requirements

Typical Network Topology



Candidate models

- Traditional network
- Centrally controlled OpenFlow based hop-by-hop switching fabric
- Edge to edge overlays

Traditional Network

- Ethernet VLANs for L2 isolation
  - 4096 limit
  - VLANs will have large spanning trees terminating on many hosts
  - High churn in switch control planes doing MAC learning non-stop
  - Need MLAG for L2 multi-path
    - Vendor specific
- VRFs for L3 isolation
  - Not scalable to cloud scale

OpenFlow Fabric

- State in switches
  - Proportional to virtual network state
  - Need to update all switches in path when provisioning
  - Not scalable, not fast enough to update, no atomicity of updates
- Not good for IaaS cloud virtual networking

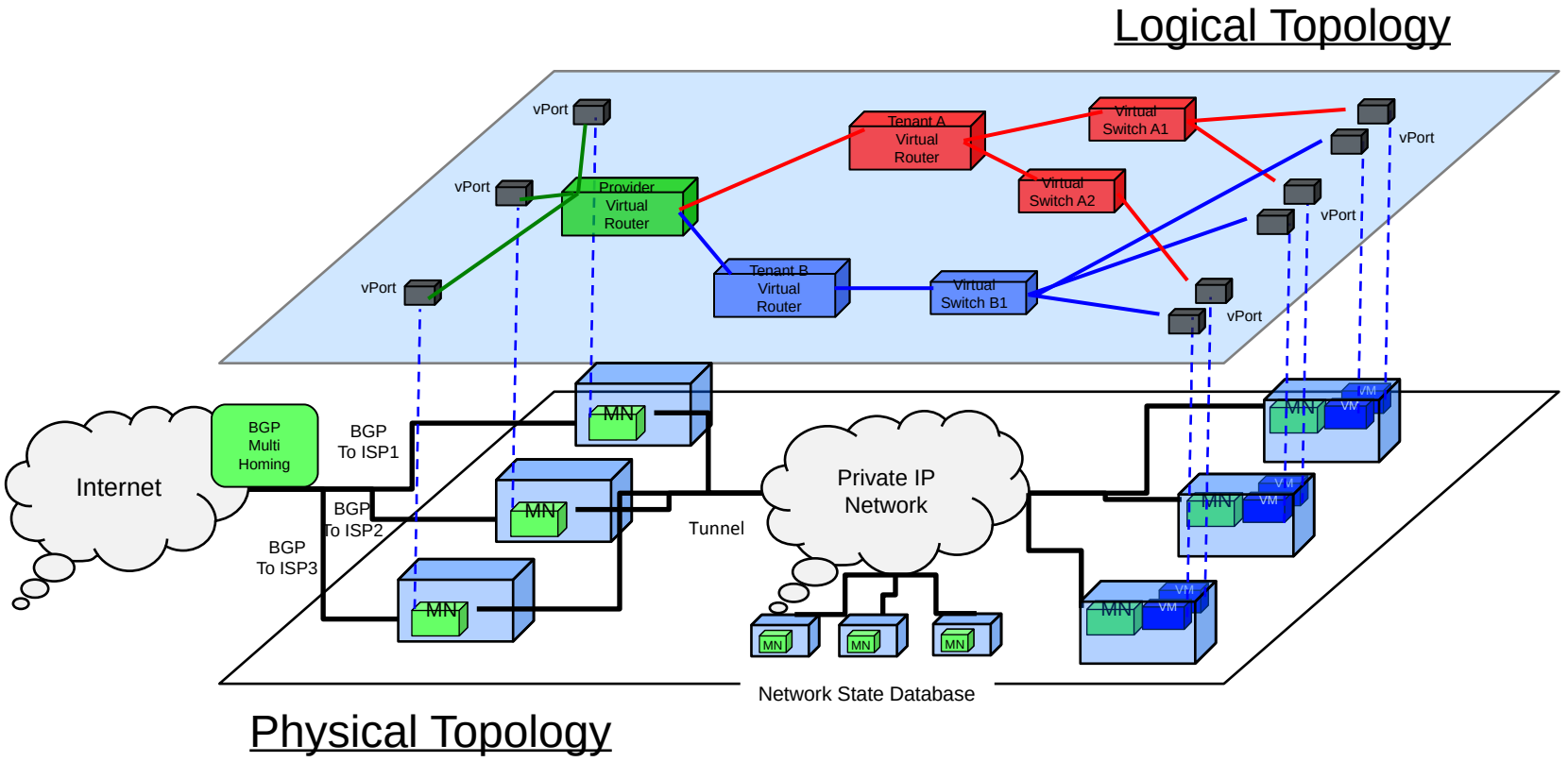
Edge to Edge IP Overlays

- Isolation not using VLANs
  - IP encapsulation
- Decouple from physical network
- Provisioning VM doesn’t change underlay state
- Underlay delivers to destination host IP
  - Forwarding equivalence class (FEC)
- Use scalable IGP (iBGP, OSPF) to build multi-path underlay

Market trends supporting overlay solutions

- Packet processing on x86 CPUs (at edge)
  - Intel DPDK facilitates packet processing
  - Number of cores in servers increasing fast
- Clos Networks (for underlay)
  - Spine and Leaf architecture with IP
  - Economical and high E-W bandwidth
- Merchant silicon (cheap IP switches)
  - Broadcom, Intel (Fulcrum Micro), Marvell

Example of an overlay solution



Overlays are the right approach!

But not sufficient.
We still need a scalable control plane.

Questions?