

OpenFlow and Open Networking

An Introduction and Overview

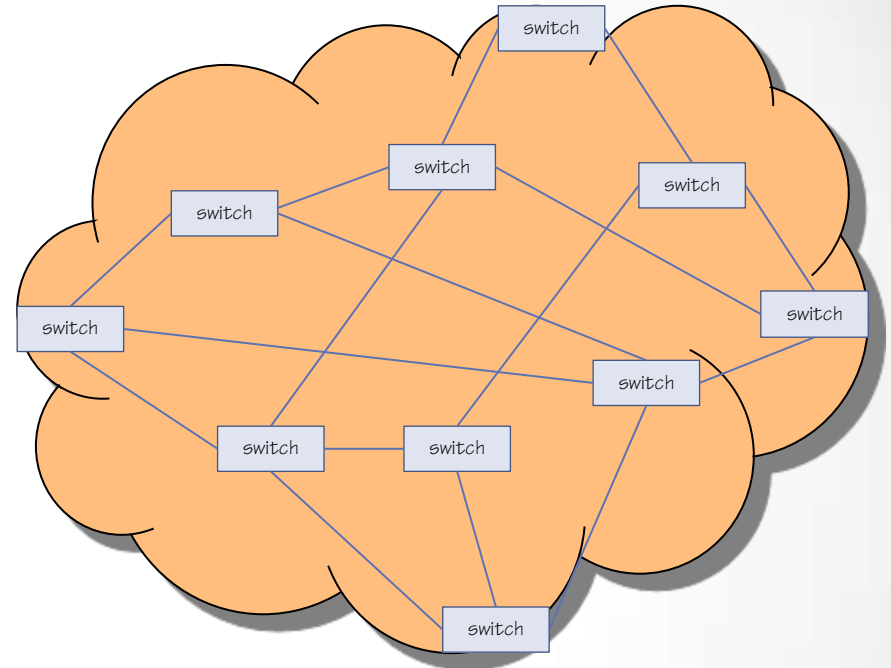
Peter Christy, IRG, pchristy@irg-intl.com

February 2012

What?:

Modernizing the Control Plane

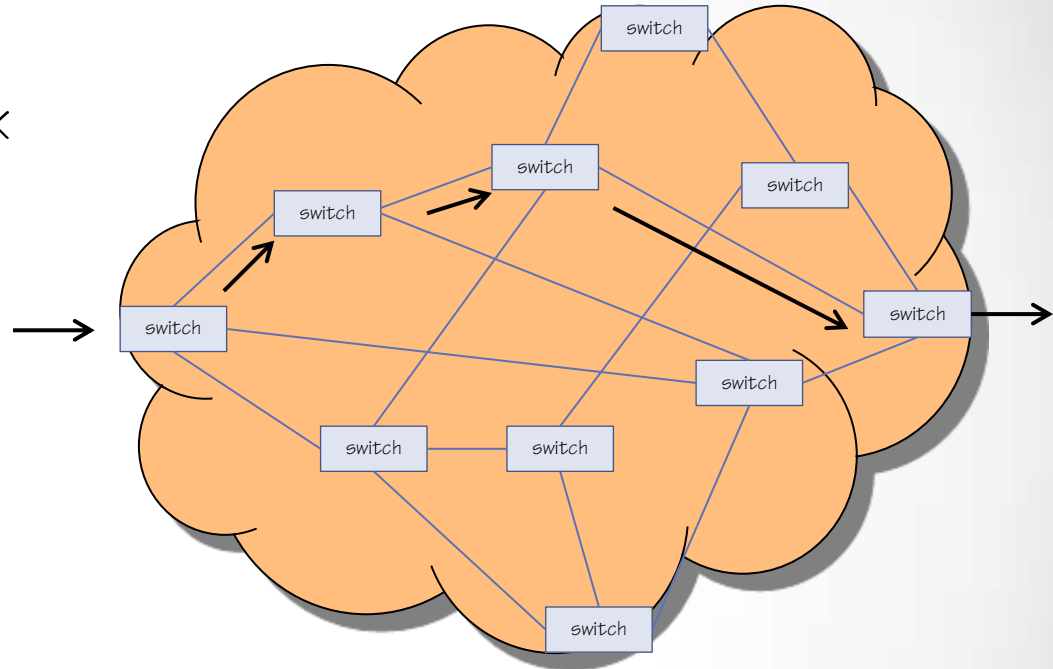
- This is all about **modernizing the control plane** of a network



What?:

Modernizing the Control Plane

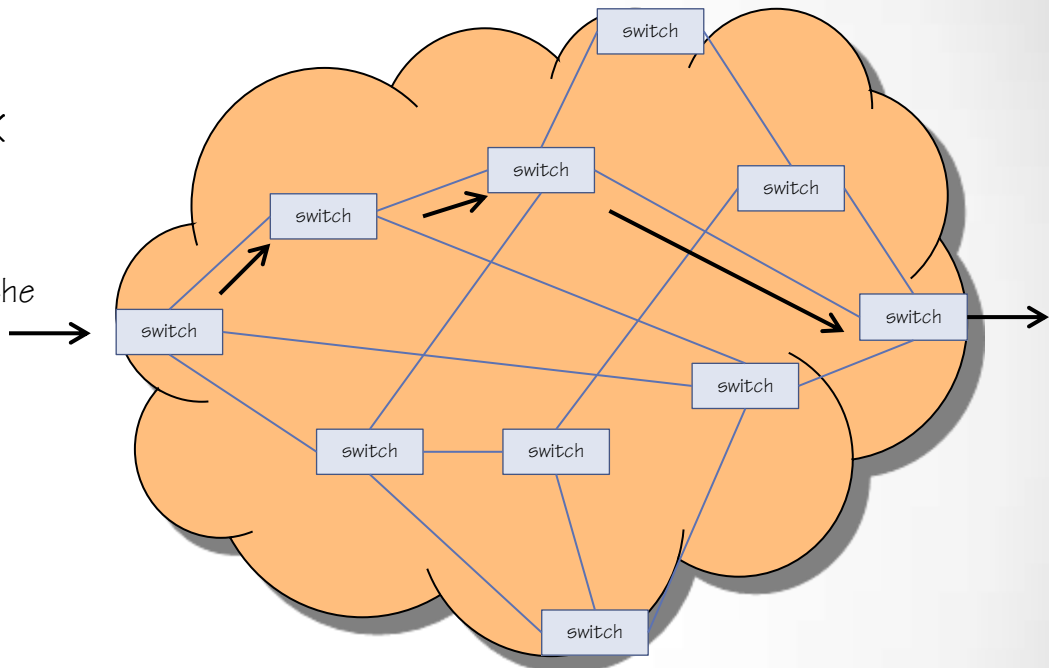
- This is all about **modernizing the control plane** of a network
 - The **data plane** does the work of moving packets through the network



What?:

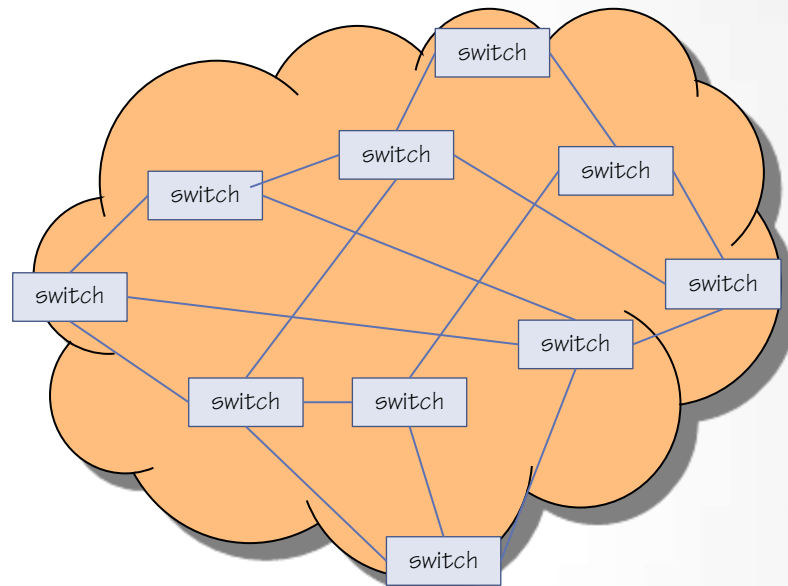
Modernizing the Control Plane

- This is all about **modernizing the control plane** of a network
 - The **data plane** does the work of moving packets through the network
 - The **control plane** is the means by which the data plane is configured, and how exceptions are handled



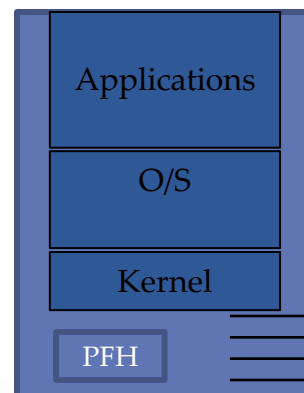
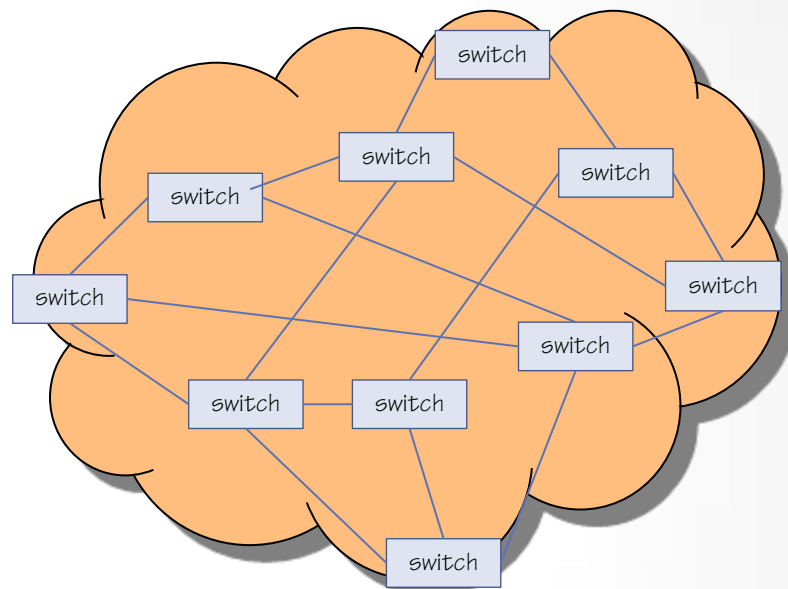
The Data Plane

- A network consists of a set of interconnected switches.



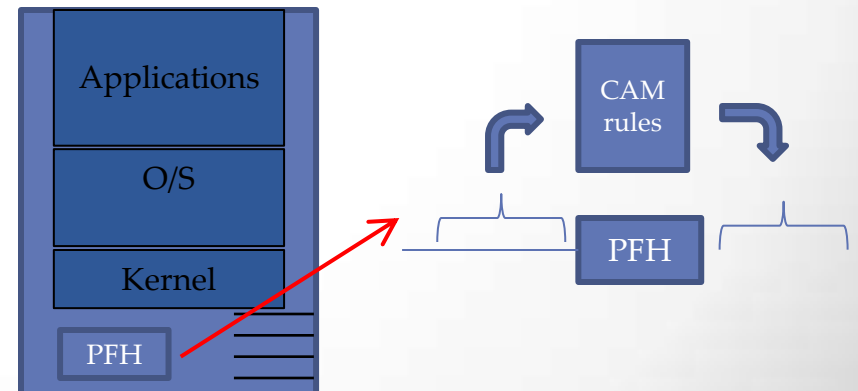
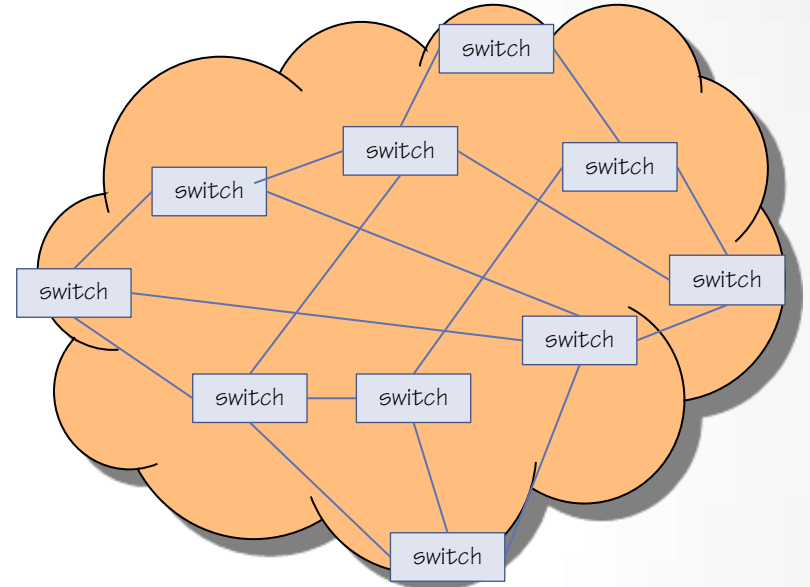
The Data Plane

- A network consists of a set of interconnected switches.
- Each switch is a “server” system with a purpose built operating system and application software, as well as specialized hardware in the form of line cards and packet forwarding hardware (PFH).



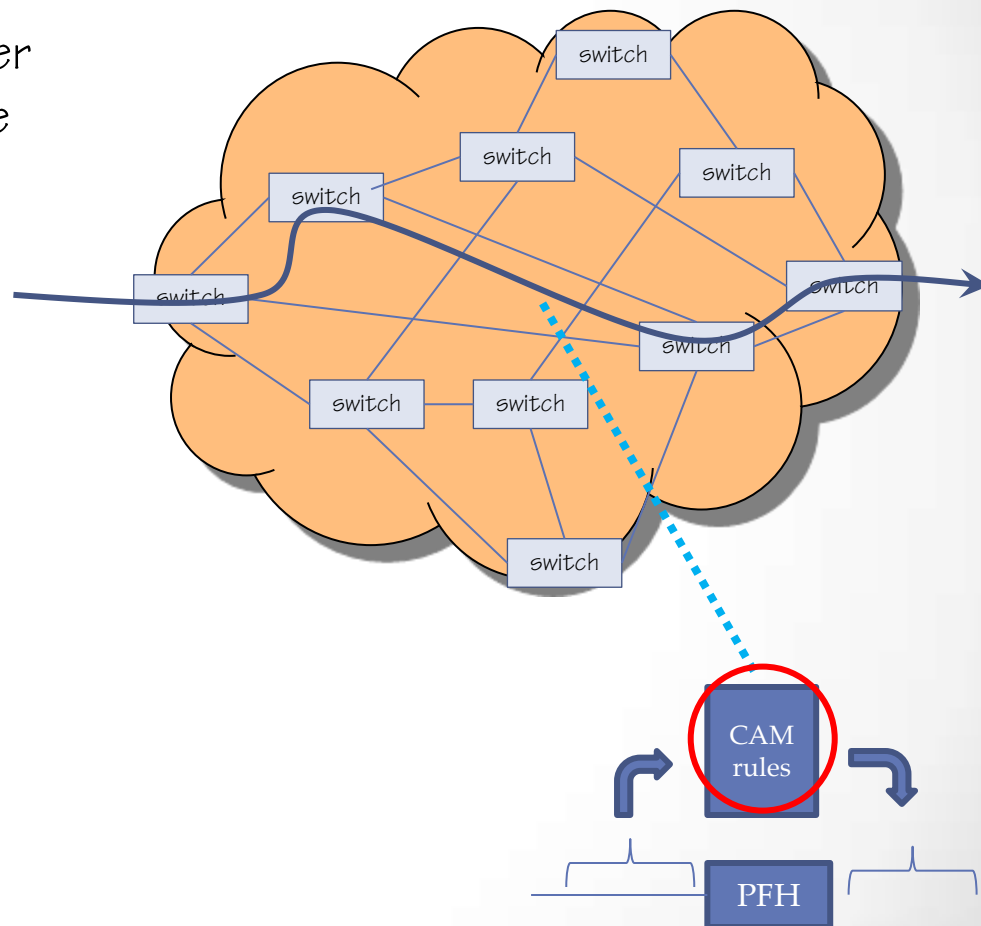
The Data Plane

- A network consists of a set of interconnected switches.
- Each switch is a “server” system with a purpose built operating system and application software, as well as specialized hardware in the form of line cards and packet forwarding hardware (PFH).
- The packet forwarding hardware operates by examining the various header bits on the packet, consulting a set of rules stored in high speed associative memory (CAM) and if an applicable rule is found, using the rule data to select the output link, assign a queuing priority, and possibly rewrite some of the header bits.



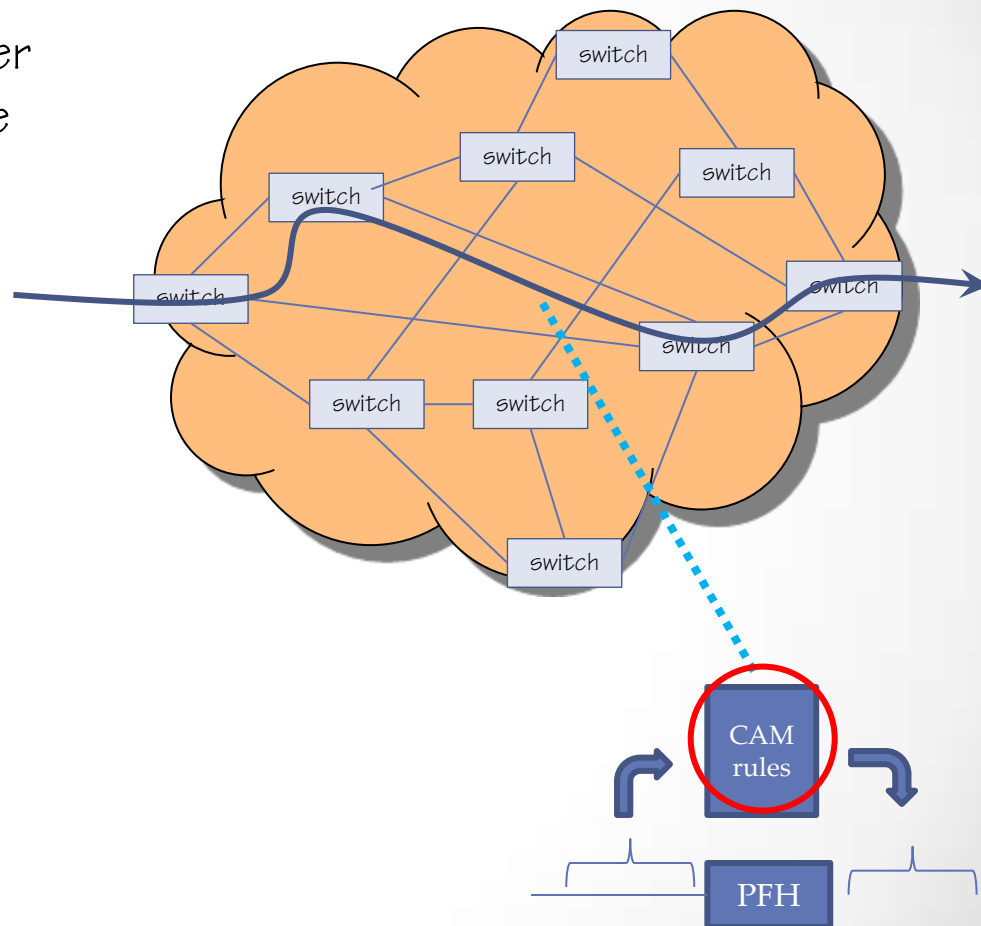
The Control Plane

- For most packets in a data center or SP network, the actions of the network are dictated by stored rules, executed by the PFH



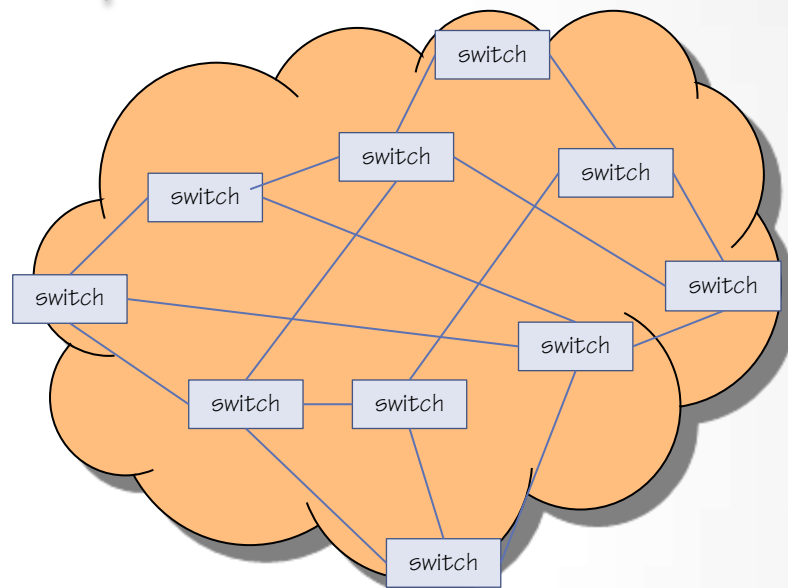
The Control Plane

- For most packets in a data center or SP network, the actions of the network are dictated by stored rules, executed by the PFH
- The “control plane” is the means by which the PFH rules are managed.



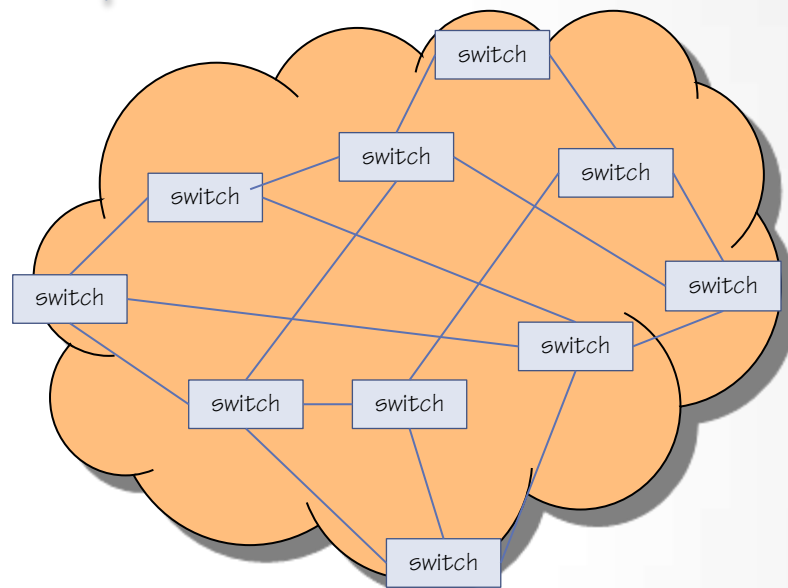
History of OpenFlow

- “OpenFlow” began as an idea about how research networks could be built.



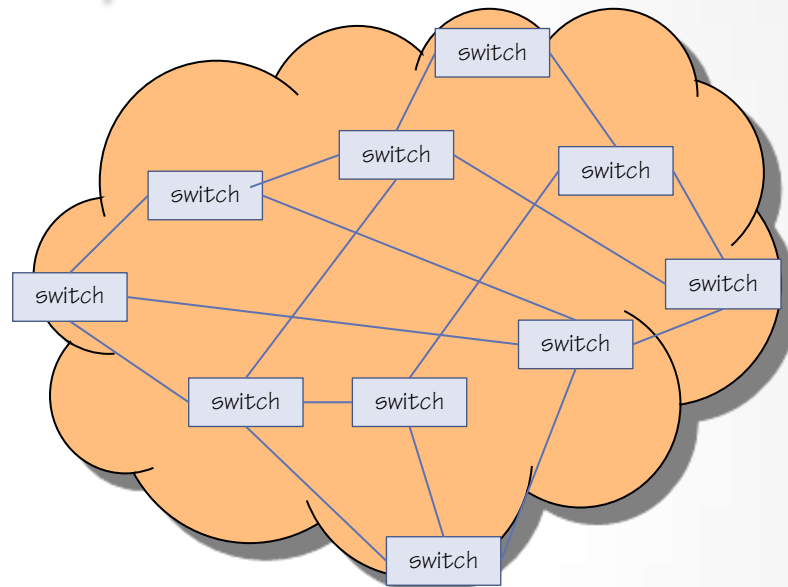
History of OpenFlow

- “OpenFlow” began as an idea about how research networks could be built.
- Why not just build an experimental network?
 - Researchers can't generally afford to buy new devices and build a new network.



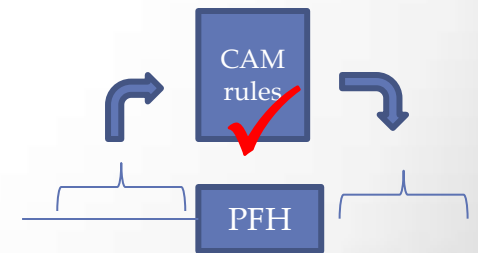
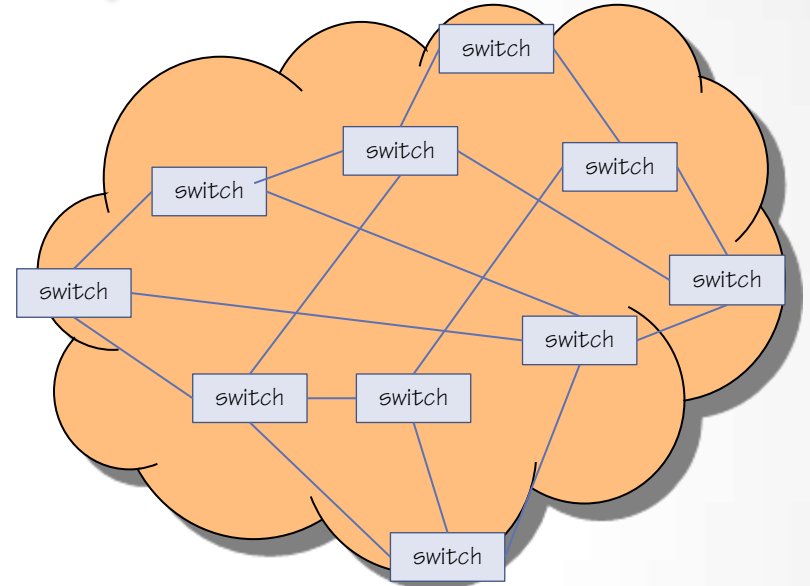
History of OpenFlow

- “OpenFlow” began as an idea about how research networks could be built.
- Why not just build an experimental network?
 - Researchers can’t generally afford to buy new devices and build a new network.
- Why not add experimental software to existing switches?
 - Modern switches are all different because they each have proprietary and different software architecture internally, so writing additional code for them is difficult or impractical.



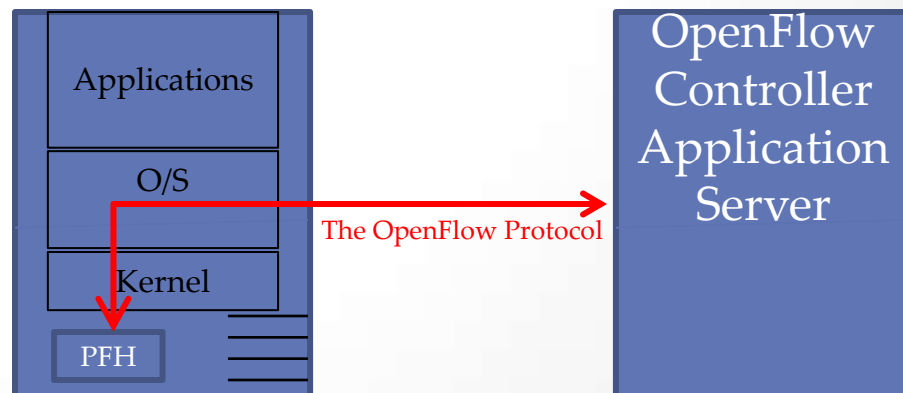
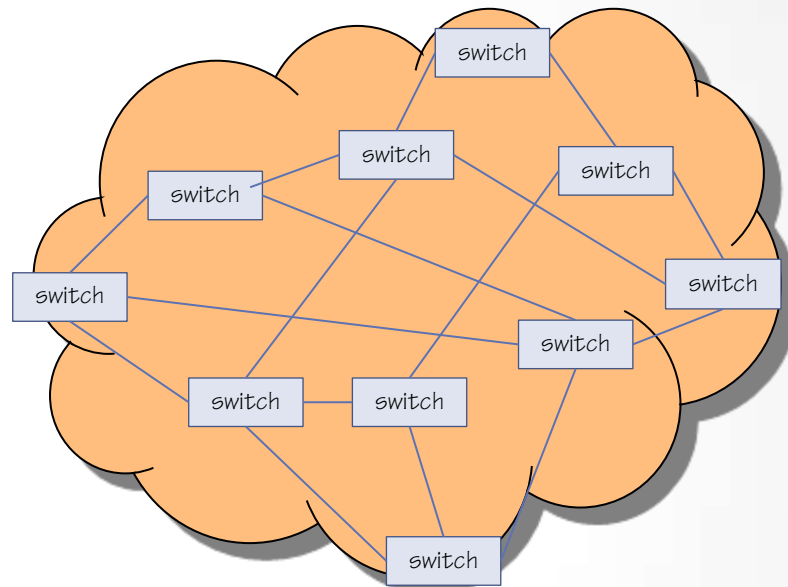
History of OpenFlow

- “OpenFlow” began as an idea about how research networks could be built.
- Why not just build an experimental network?
 - Researchers can’t generally afford to buy new devices and build a new network.
- Why not add experimental software to existing switches?
 - Modern switches are all different because they each have proprietary and different software architecture internally, so writing additional code for them is difficult or impractical.
- What about programming the PFH?
 - In contrast to the software designs, PFH hardware is all similar. What about directly programming the Packet Forwarding Hardware?



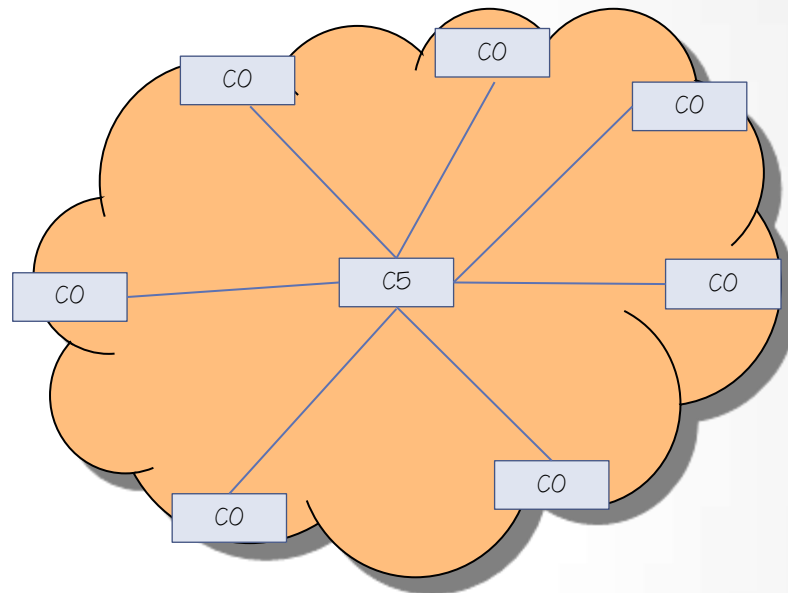
OpenFlow: a PFH Control Protocol

- OpenFlow is a protocol by which the PFH hardware in a switch can be managed by software executing in a separate server, external to the switch, in a standardized way.



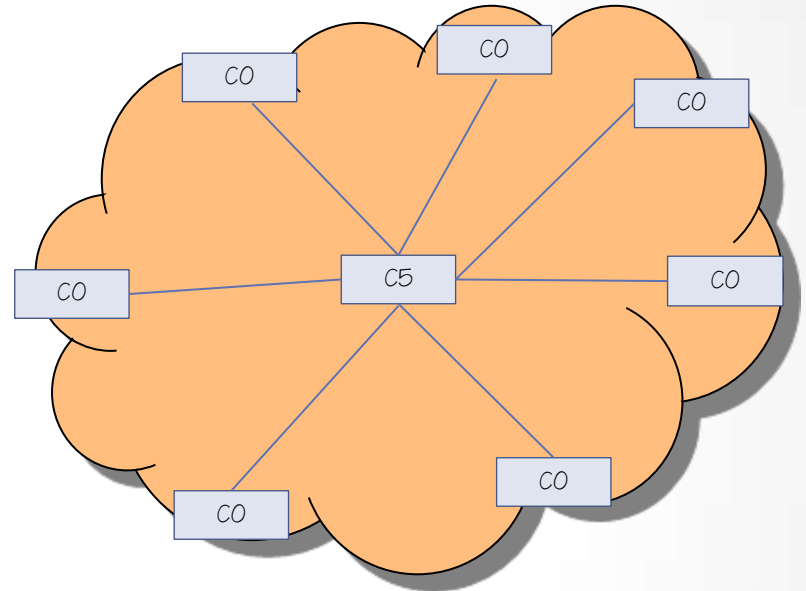
Why Modernize?

- Paul Baran invented the Internet at Rand in the early 1960's.



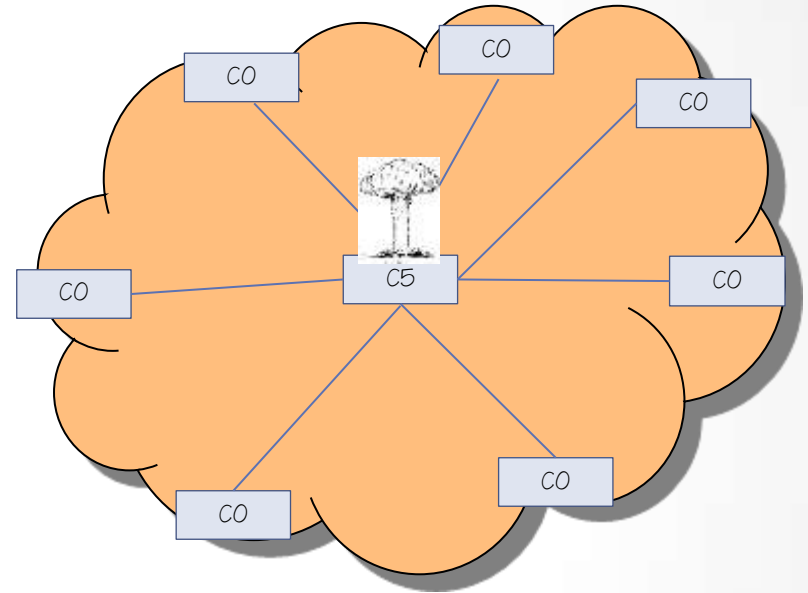
Why Modernize?

- Paul Baran invented the Internet at Rand in the early 1960's.
- WWII and the ensuing Cold War demonstrated
 1. the necessity of reliable communications for defense or in a war and
 2. the lethal destruction possible with missiles, long-range bombers, and nuclear weapons.



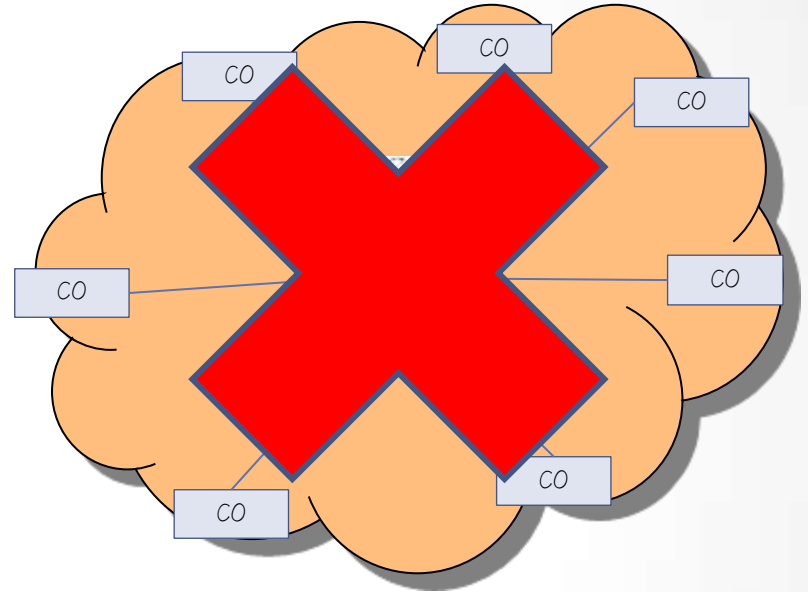
Why Modernize?

- Paul Baran invented the Internet at Rand in the early 1960's.
- WWII and the ensuing Cold War demonstrated
 1. the necessity of reliable communications for defense or in a war and
 2. the lethal destruction possible with missiles, long-range bombers, and nuclear weapons.
- The phone system in the US at the time was a centralized circuit-switching system that would be very easy to disrupt or destroy.



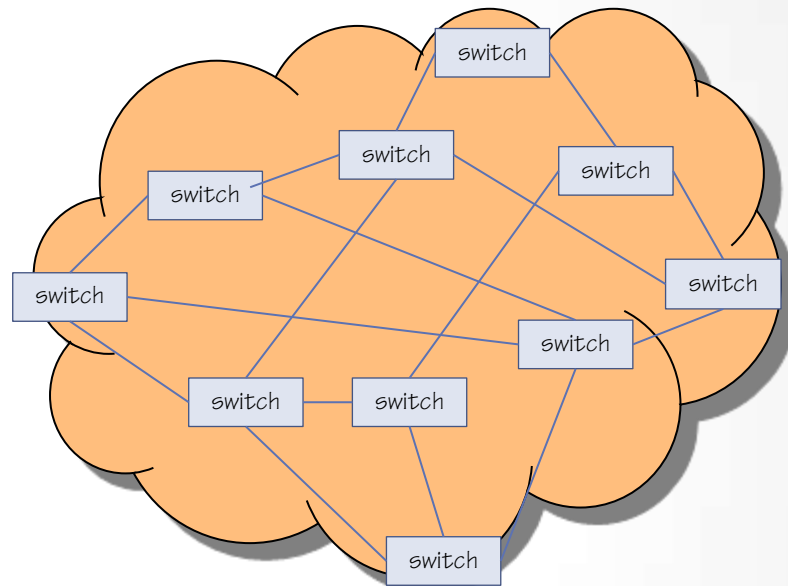
Why Modernize?

- Paul Baran invented the Internet at Rand in the early 1960's.
- WWII and the ensuing Cold War demonstrated
 1. the necessity of reliable communications for defense or in a war and
 2. the lethal destruction possible with missiles, long-range bombers, and nuclear weapons.
- The phone system in the US at the time was a centralized circuit-switching system that would be very easy to disrupt or destroy.



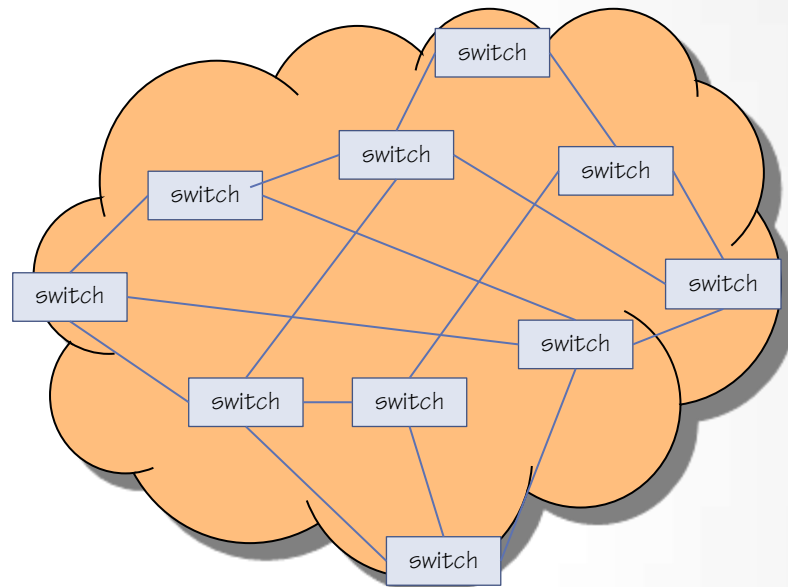
Baran's Clever Solution

- Packetized Voice Communications
 - Transform voice communications connections into sequences of packets of voice data.



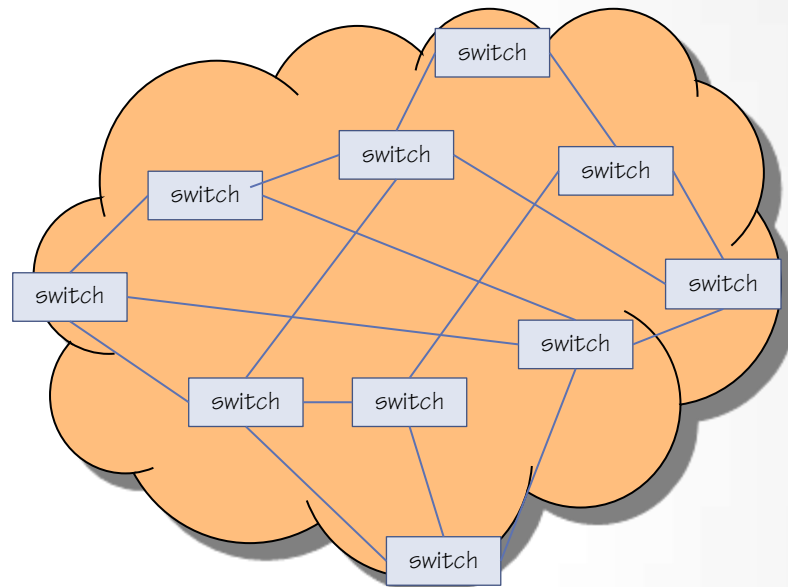
Baran's Clever Solution

- Packetized Voice Communications
 - Transform voice communications connections into sequences of packets of voice data.
 - Transmit each packet independently.



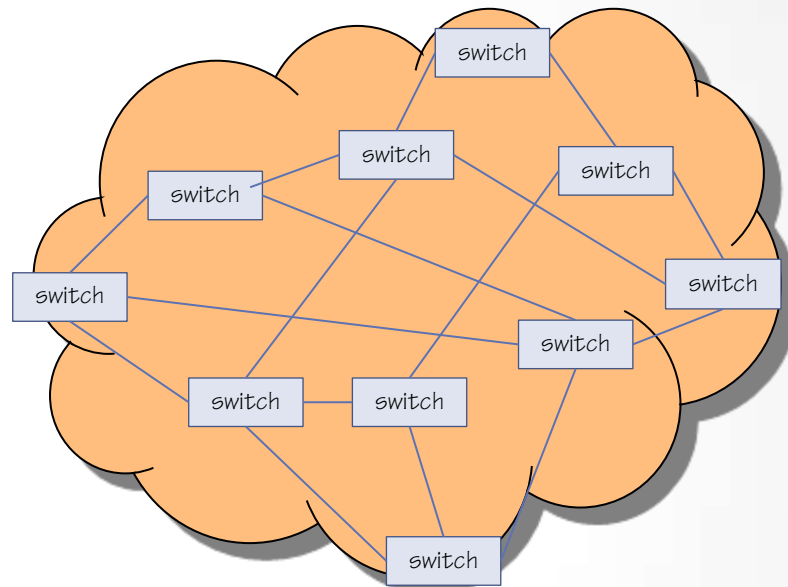
Baran's Clever Solution

- Packetized Voice Communications
 - Transform voice communications connections into sequences of packets of voice data.
 - Transmit each packet independently.
 - Have the packet forwarding devices strictly autonomous systems so that the destruction or failure of one causes no additional failures.



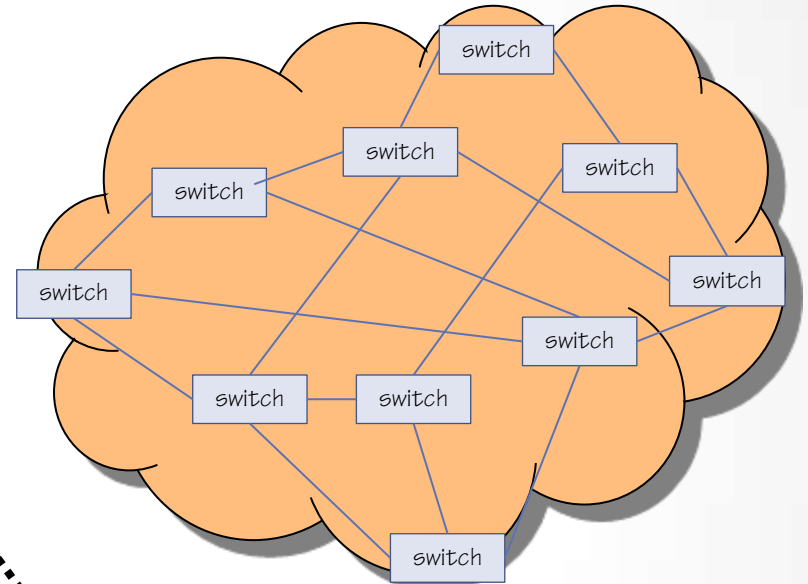
Baran's Clever Solution

- Packetized Voice Communications
 - Transform voice communications connections into sequences of packets of voice data.
 - Transmit each packet independently.
 - Have the packet forwarding devices strictly autonomous systems so that the destruction or failure of one causes no additional failures.
 - Baran's analysis demonstrated that the resulting voice communications system could still function if 50% of the packet forwarding devices were destroyed!



Baran's Clever Solution

- Packetized Voice Communications
 - Transform voice communications connections into sequences of packets of voice data.
 - Transmit each packet independently.
 - Have the packet forwarding devices strictly autonomous systems so that the destruction or failure of one causes no additional failures.
 - Baran's analysis demonstrated that the resulting voice communications system could still function if 50% of the packet forwarding devices were destroyed!



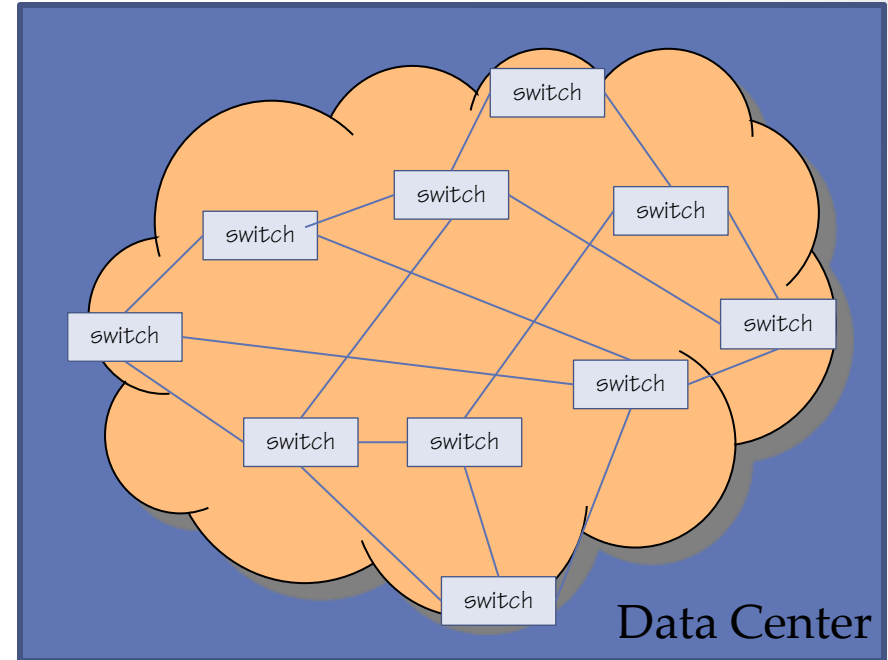
the legacy Baran "survivability" control plane architecture

Now fast-forward 50 years

- Networks are used much more for data communications than voice communications

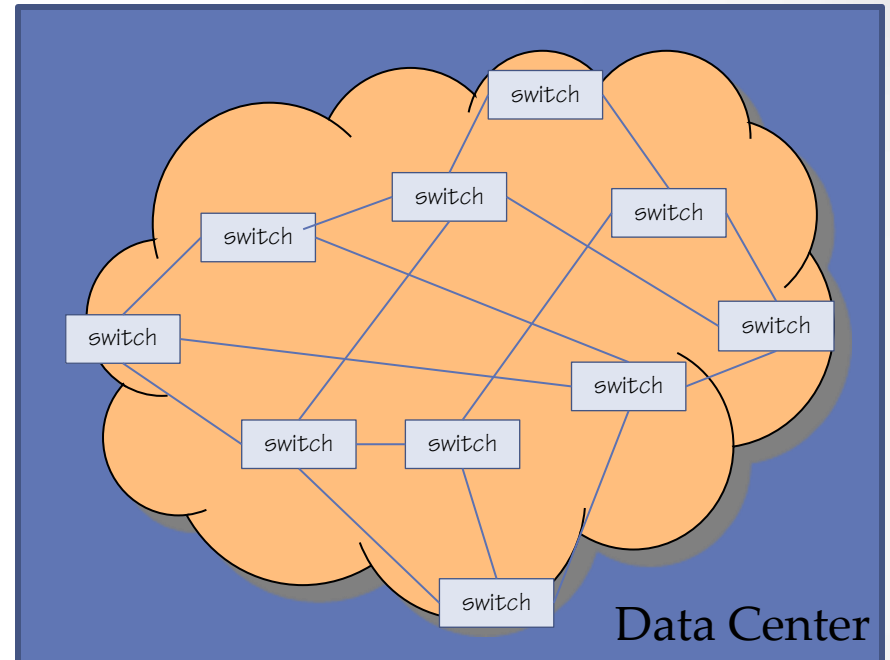
Now fast-forward 50 years

- Networks are used much more for *data communications* than *voice communications*
- A large data center network or SP network will have thousands of switches.



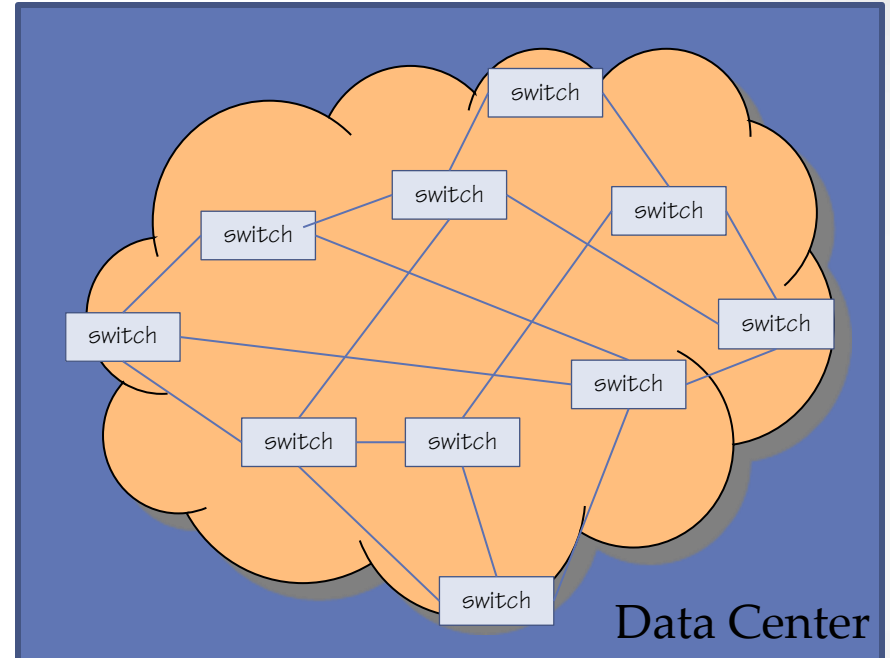
Now fast-forward 50 years

- Networks are used much more for *data communications* than *voice communications*
- A large data center network or SP network will have thousands of switches.
- Survivability is no longer a worry.



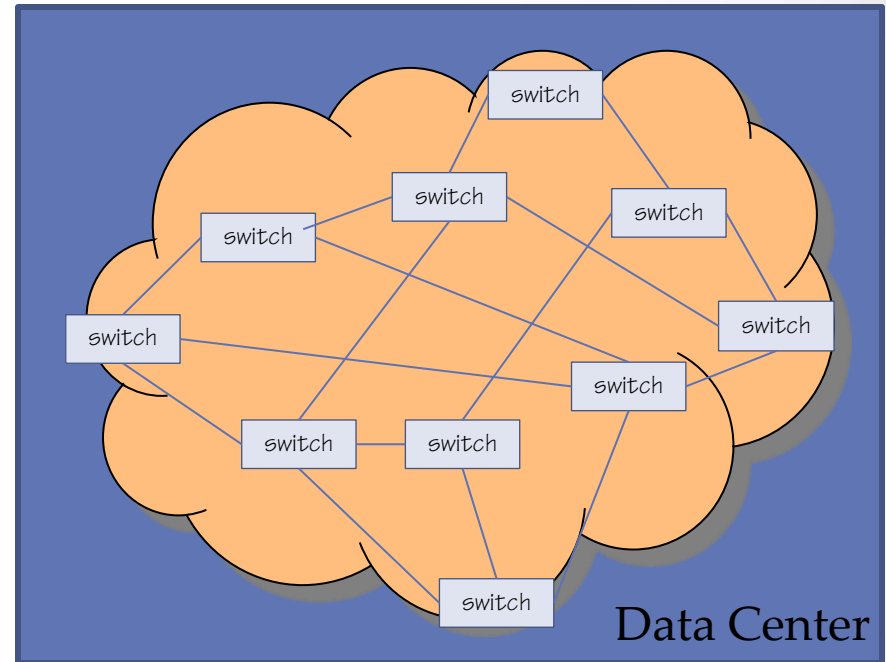
Now fast-forward 50 years

- Networks are used much more for *data communications* than *voice communications*
- A large data center network or SP network will have thousands of switches.
- Survivability is no longer a worry.
- We want the network to operate as a coherent whole.



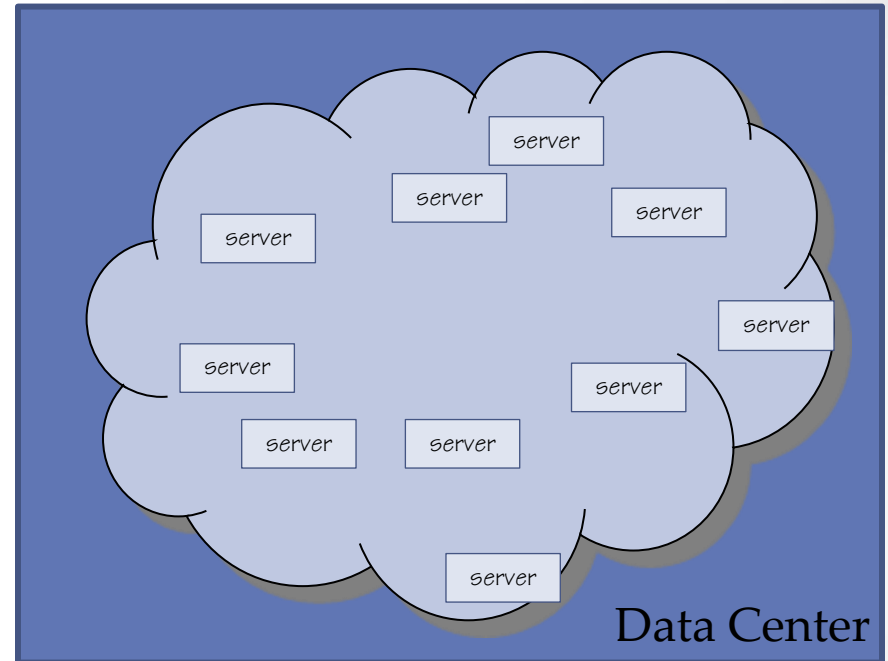
Now fast-forward 50 years

- Networks are used much more for data communications than voice communications
- A large data center network or SP network will have thousands of switches.
- Survivability is no longer a worry.
- We want the network to operate as a coherent whole.
- Managing the network with the legacy “Baran” control plane architecture is really hard



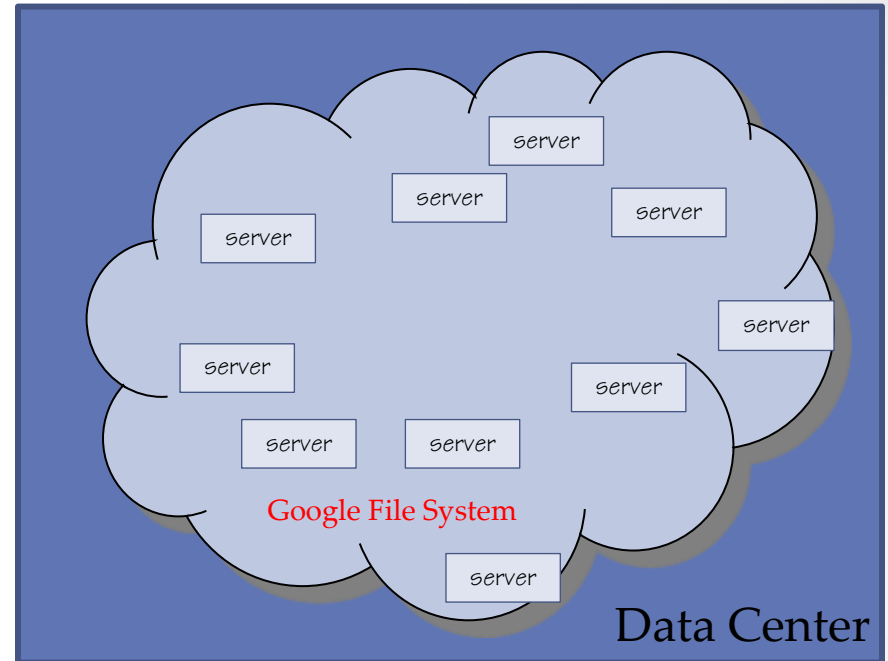
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).



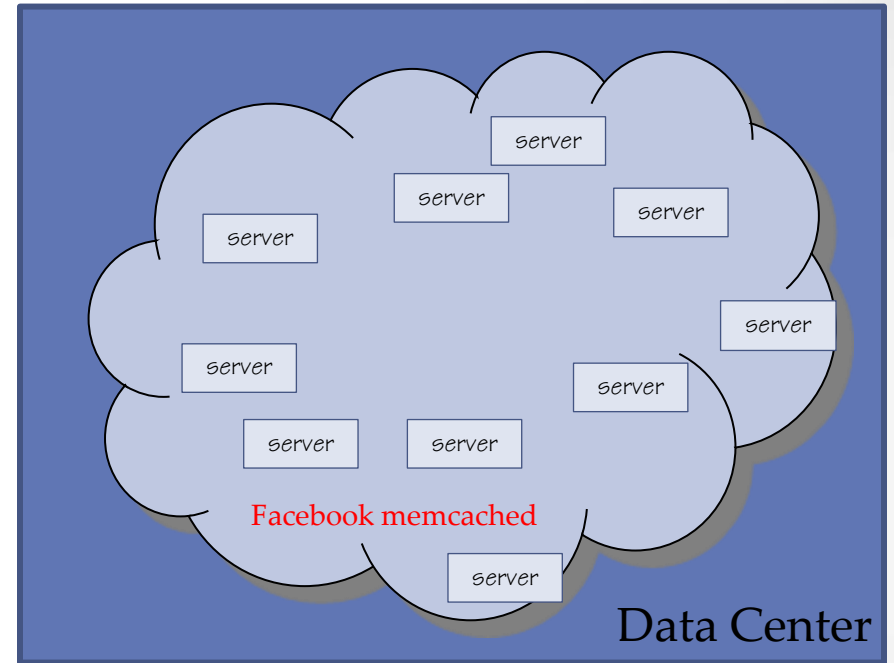
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).
- Consider how servers are coordinated in data centers today
 - A shared file system like GFS (Google)



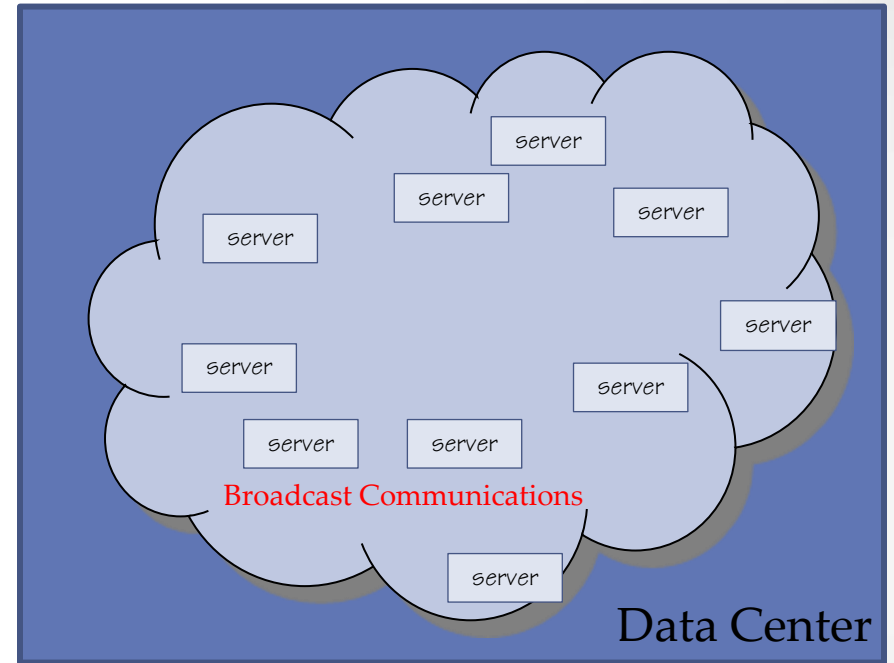
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).
- Consider how servers are coordinated in data centers today
 - A shared file system like GFS (Google)
 - A shared caching system like memcached (Facebook).



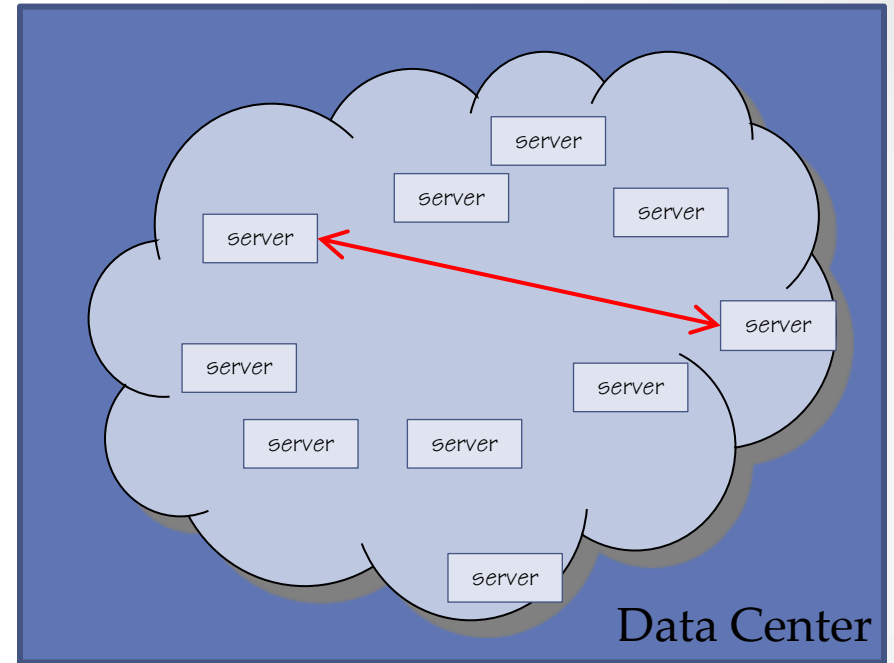
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).
- Consider how servers are coordinated in data centers today
 - A shared file system like GFS (Google)
 - A shared caching system like memcached (Facebook).
 - Broadcast network communications to all servers.



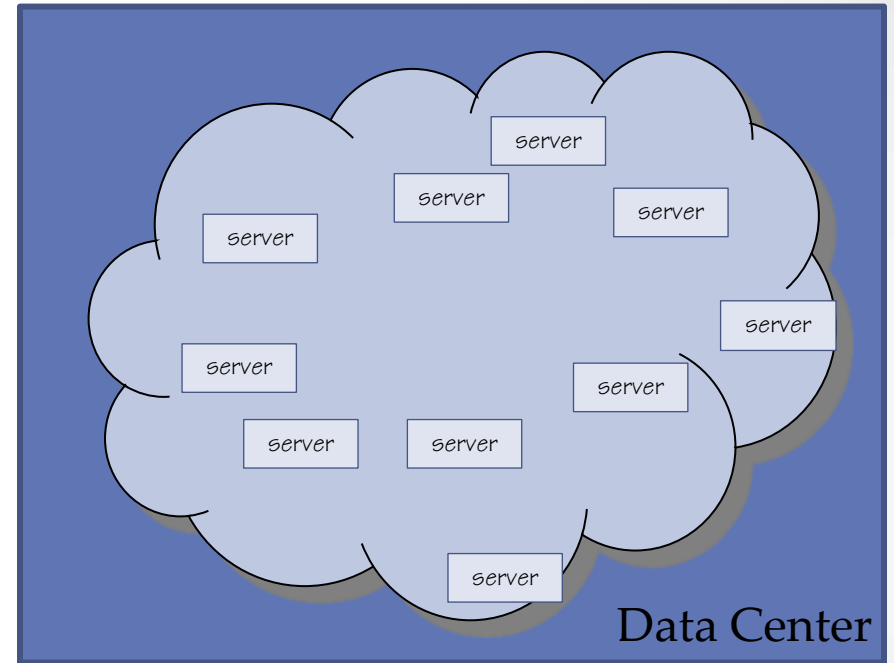
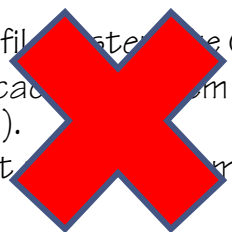
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).
- Consider how servers are coordinated in data centers today
 - A shared file system like GFS (Google)
 - A shared caching system like memcached (Facebook).
 - Broadcast network communications to all servers.
 - Direct network communication between any two servers.



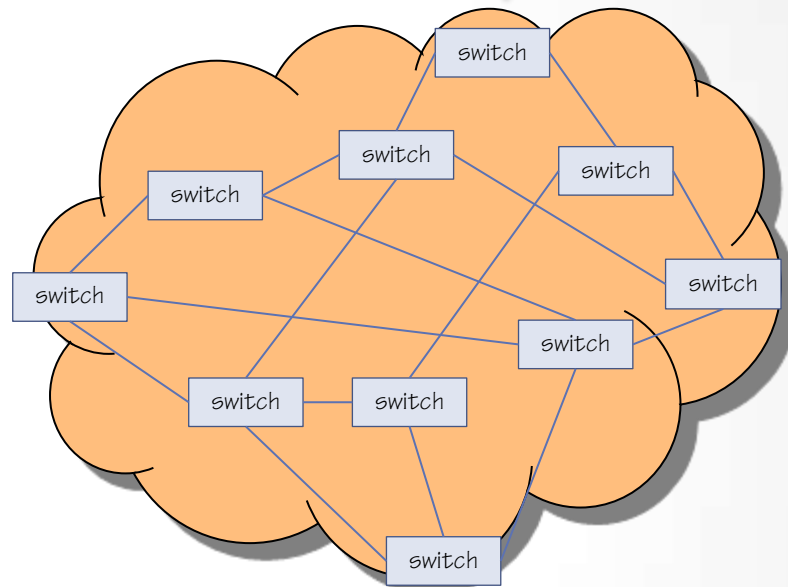
What does modernization mean?

- Think of a data center network as just another distributed server application (switches are servers with specialized PFH included).
- Consider how servers are coordinated in data centers today
 - A shared file system like GFS (Google)
 - A shared cache system like memcached (Facebook).
 - Broadcast network communications to all servers.
 - Direct network communication between any two servers.
- All of these are prohibited by a legacy control plane architecture that was created to maximize survivability.



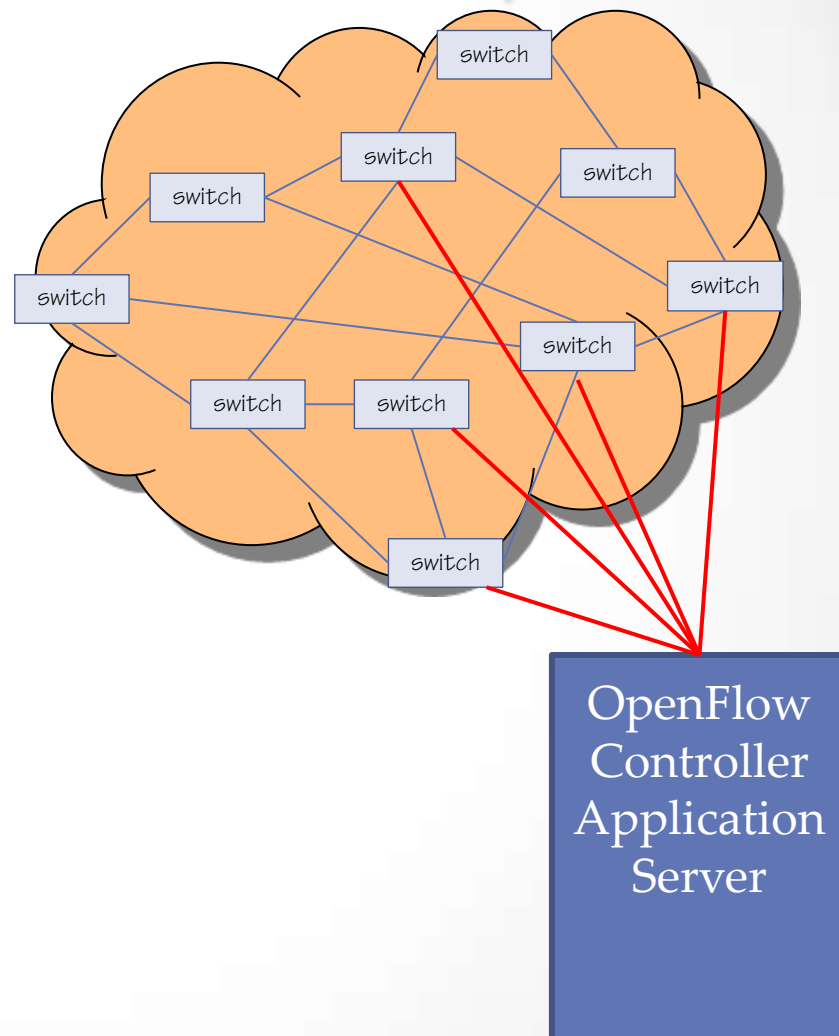
An OF implemented control plane

- We use existing switches modified to add an OpenFlow control port



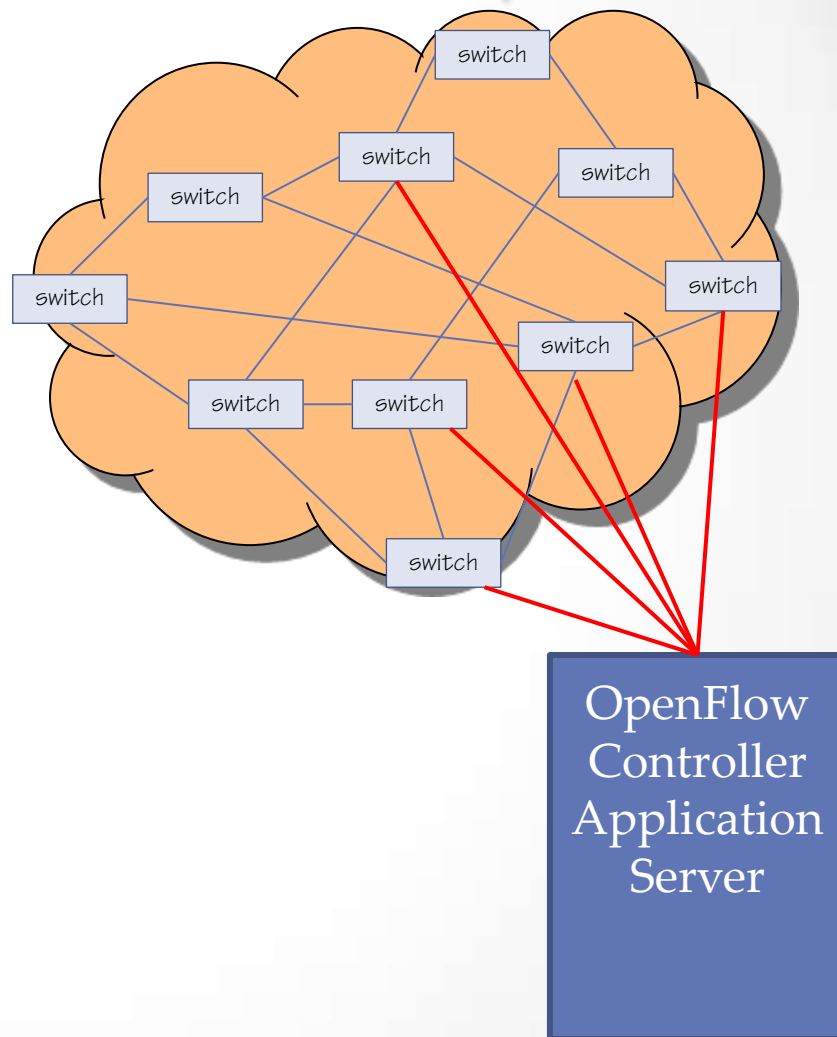
An OF implemented control plane

- We use existing switches modified to add an OpenFlow control port
- We implement the control plane in the OpenFlow controller (remember, it's just a server application)



An OF implemented control plane

- We use existing switches modified to add an OpenFlow control port
- We implement the control plane in the OpenFlow controller (remember, it's just a server application)
- We do anything we want, no longer constrained by the survivability design architecture or by the internal software design of the switches



What is a “Flow” Anyway?

- Ethernet networking is between the MAC addresses of the network interfaces on systems.
- IP networking assigns each system an IP number and then switches traffic based on the IP number throughout the internet until the final switch that knows the associated MAC address of the destination
- The Internet was initially designed to switch primarily on IP number.

What is a “Flow” Anyway?

- Ethernet networking is between the MAC addresses of the network interfaces on systems.
- IP networking assigns each system an IP number and then switches traffic based on the IP number throughout the internet until the final switch that knows the associated MAC address of the destination
- The Internet was initially designed to switch primarily on IP number.
- But networking is really done between logical ports on the systems, not with the system generally. A browser makes a connection to Port 80 on the destination system -- the Web server on a system by convention.
- A flow represents the packets between a specific port on the transmitting system to a specific port on the destination system.

What is a “Flow” Anyway?

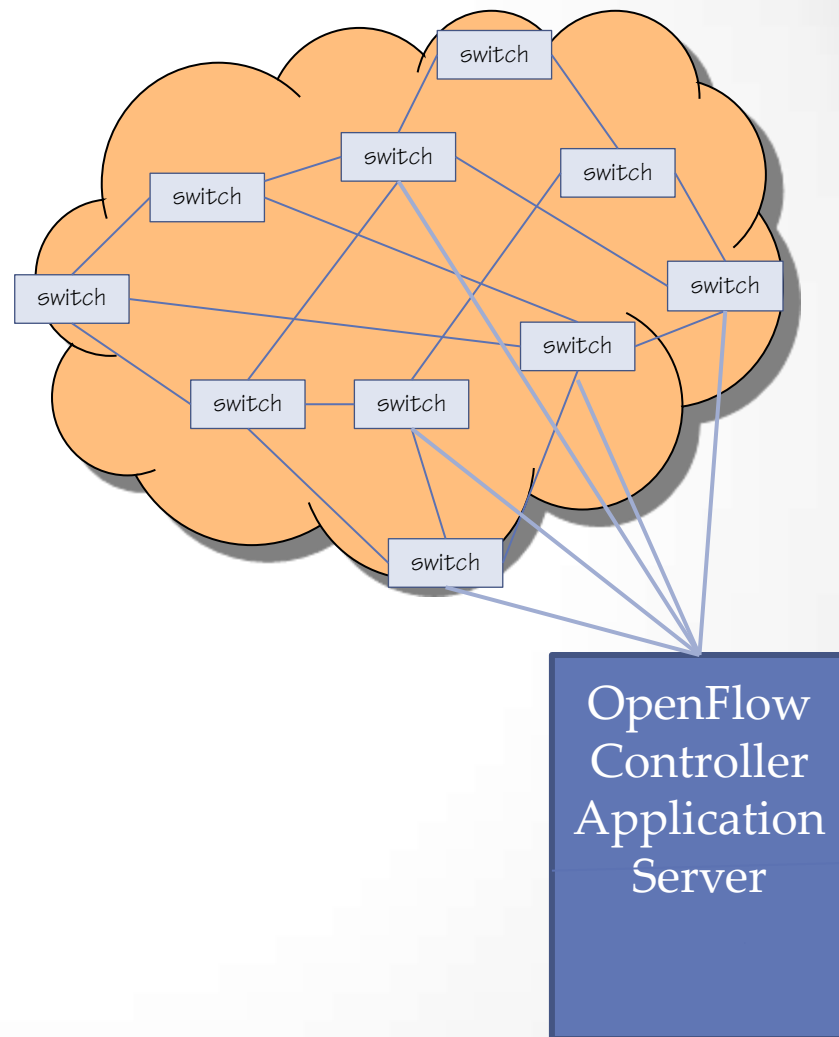
- Ethernet networking is between the MAC addresses of the network interfaces on systems.
- IP networking assigns each system an IP number and then switches traffic based on the IP number throughout the internet until the final switch that knows the associated MAC address of the destination
- The Internet was initially designed to switch primarily on IP number.
- But networking is really done between logical ports on the systems, not with the system generally. A browser makes a connection to Port 80 on the destination system -- the Web server on a system by convention.
- A flow represents the packets between a specific port on the transmitting system to a specific port on the destination system.
- Larry Roberts, the individual that funded the original ARPAnet research, tried to build two companies based on the premise that routing across the entire Internet should be done on a flow basis (Caspian Networks and the Anagran networks each built flow-based routers).
- But to do what Roberts wanted to do required that flow-based routing be accepted throughout the Internet. That wasn't going to happen.

What is a “Flow” Anyway?

- Ethernet networking is between the MAC addresses of the network interfaces on systems.
- IP networking assigns each system an IP number and then switches traffic based on the IP number throughout the internet until the final switch that knows the associated MAC address of the destination
- The Internet was initially designed to switch primarily on IP number.
- But networking is really done between logical ports on the systems, not with the system generally. A browser makes a connection to Port 80 on the destination system -- the Web server on a system by convention.
- A flow represents the packets between a specific port on the transmitting system to a specific port on the destination system.
- Larry Roberts, the individual that funded the original ARPAnet research, tried to build two companies based on the premise that routing across the entire Internet should be done on a flow basis (Caspian Networks and the Anagran networks each built flow-based routers).
- But to do what Roberts wanted to do required that flow-based routing be accepted throughout the Internet. That wasn't going to happen.
- OpenFlow adopts flow-based processing but only within a single network. To the rest of the network it is still a conventional TCP/IP network.
- Being able to switch on an individual flow basis enable innovative ways of doing security (for example) but requires larger (more expensive) CAM's in the PFH, and requires fast controller response to a new flow.

OpenFlow for Research

- Modify an existing network (e.g., the Stanford campus network) so that the switches support OpenFlow.
- Create a research control plane
- Gets better if “Hybrid” switch operation is possible
 - Let the existing flows be managed by the legacy network
 - Let the OpenFlow controller manage the new “research” flows



OpenFlow Beyond Research

- Network management requirements for virtualized, cloud data centers
 - Fast provisioning of virtual networks
- Impact of the edge “switch” now being in the hypervisor
 - A modern control plane can be implemented within the Virtual Switch that is part of the hypervisor
- Desire for more flexible asset use in network service providers
 - Move as much “network” processing as possible (e.g., firewall, VPN termination) to execute as virtual machine on conventional servers (less expensive CPU, use of shared resources).
 - Be able to provision these now virtualized services on-demand without having to speculatively provision features in a network device
- Desire to speed up innovation in networking broadly
- Cost minimization and use of common processor resources in very large web properties

A Radical Thought: SDN without OpenFlow!

- The primary SDN goal is to modernize the control plane in order to solve real, vexing network management issues (like creating virtual networks).
- Juniper QFabric:
 - Clearly SDN (a modernized control plane)
 - “Controller” distributed to each Top of Rack Switch; communications between controller and PFH would not use OF.
- Nicira:
 - Designed for building virtual networks in Cloud systems based on a hypervisor.
 - The switch is a vSwitch (see Nicira’s contribution to the Open vSwitch effort)
 - The controller is (I think) distributed to each virtual server cluster.
 - I don’t think OF is used to talk to the vSwitches in the cluster.
- Contexstream
 - Building SDN’s for Verizon and Comcast
 - The switching elements are VM’s that use an interesting distributed-hash algorithm
 - Clearly a form of SDN, but no OF
- SDN Controllers
 - An interesting and challenging distributed system problem with tradeoffs
 - OF is great for research and for limited adaptation to existing networks (e.g., NEC).
 - Direct controller / PFH interface is much higher performance (seems important for doing pure flow-based processing)

Current status

- Openflow.org -> ONF (upcoming 2nd ONF Summit at Stanford)
- Continuing work
 - on OF in conjunction with controller implementation
 - on the best way to create the data plane fabric
 - in virtual switch technology and “overlay” networks
- Some interesting issues
 - No enterprise “killer apps” yet
 - Market insertion in general is a challenge
 - Transfer of value from hardware to software

Follow up

- pchristy@irg-intl.com
- Thanks!