Software-Defined Networking Architecture

Brighten Godfrey

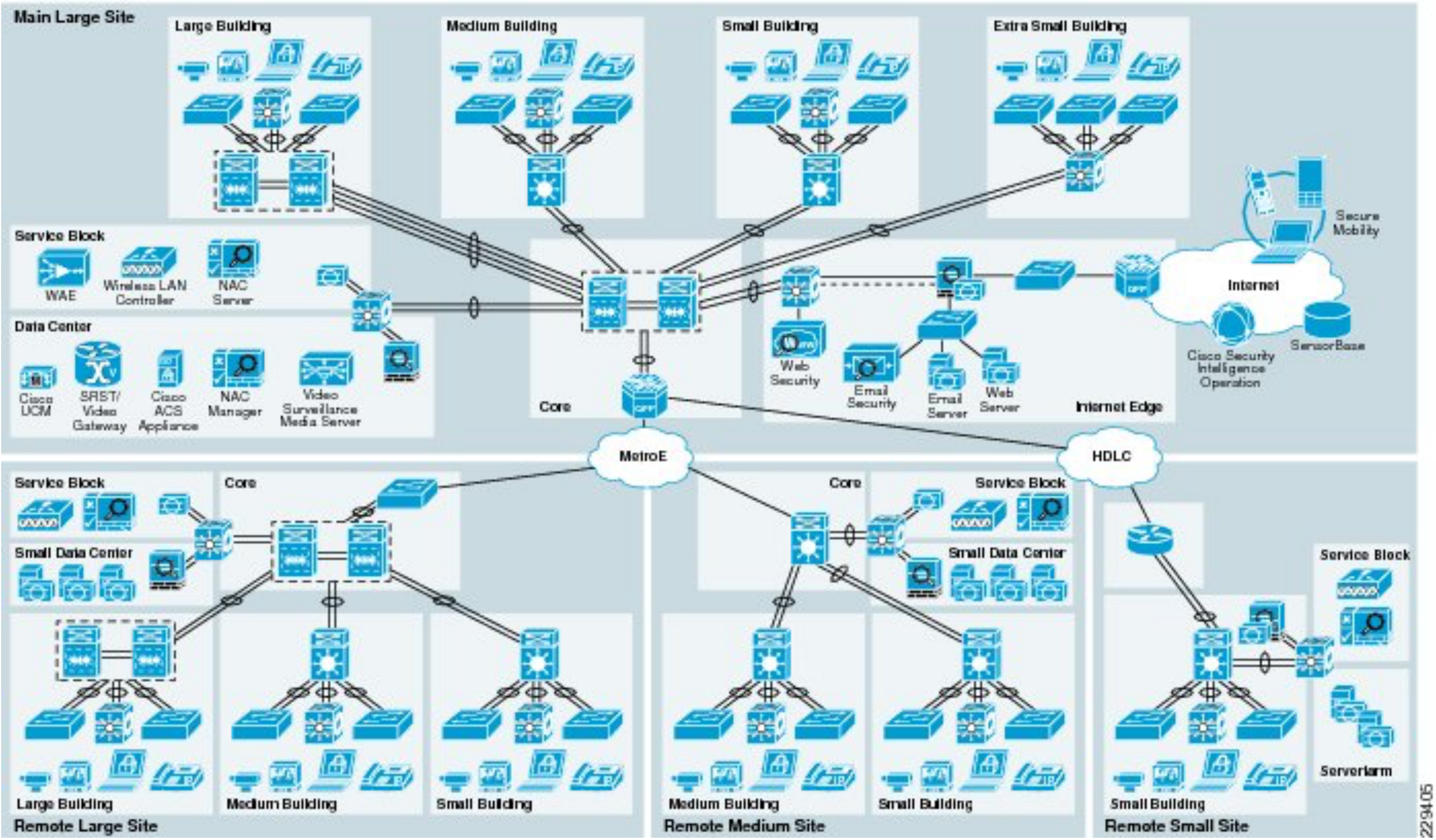
CS 538 February 21 2018

The Problem

Networks are complicated

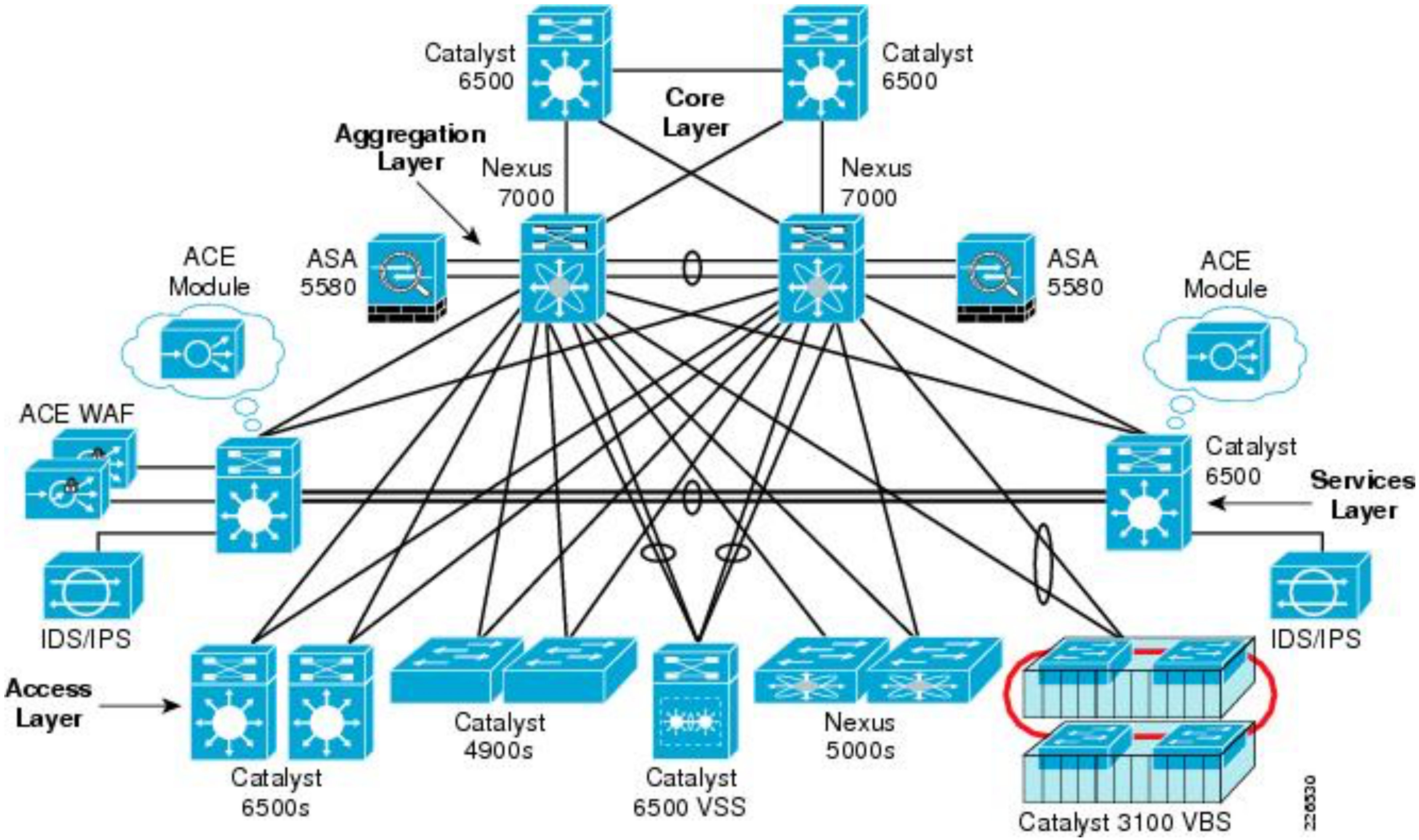
- Just like any computer system
- Worse: it's distributed
- Even worse: no clean programming APIs, only "knobs and dials"

Inside a typical enterprise network



Source: http://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Medium_Enterprise_Design_Profile/MEDP/chap5.html

Inside a typical enterprise data center



Source: http://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Security/SAFE_RG/SAFE_rg/chap4.html

Many protocols and features used

Layer 1 protocols (physical layer)

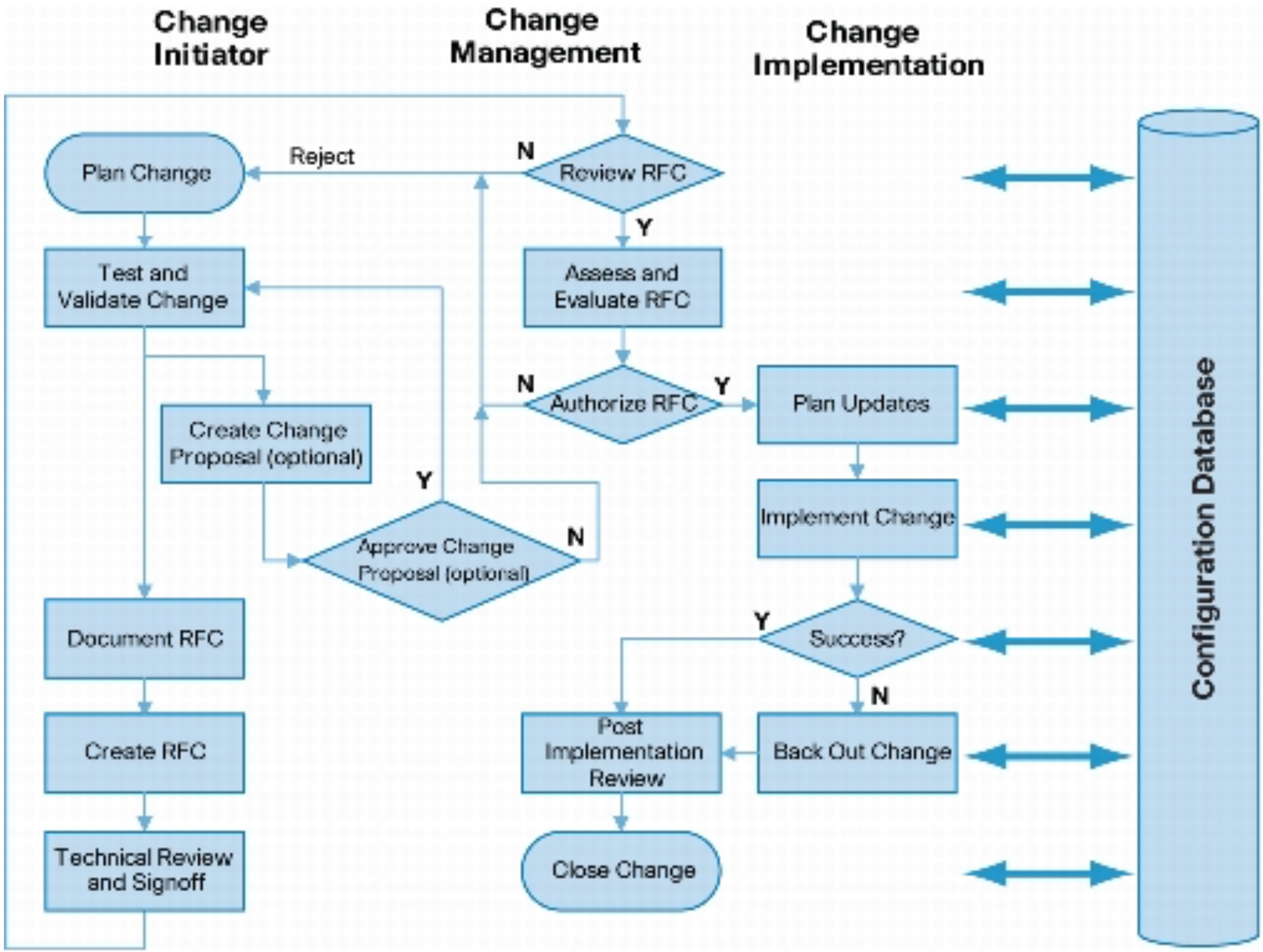
USB Physical layer
Ethernet physical layer including 10 BASE T, 100 BASE T,100 BASE TX,100 BASE FX, 1000 BASE T and other variants
varieties of 802.11 Wi-Fi physical layers
DSL
ISDN
T1 and other T-carrier links
E1 and other E-carrier links
Bluetooth physical layer

List of protocols commonly encountered by CCNAs
https://learningnetwork.cisco.com/docs/DOC-25649

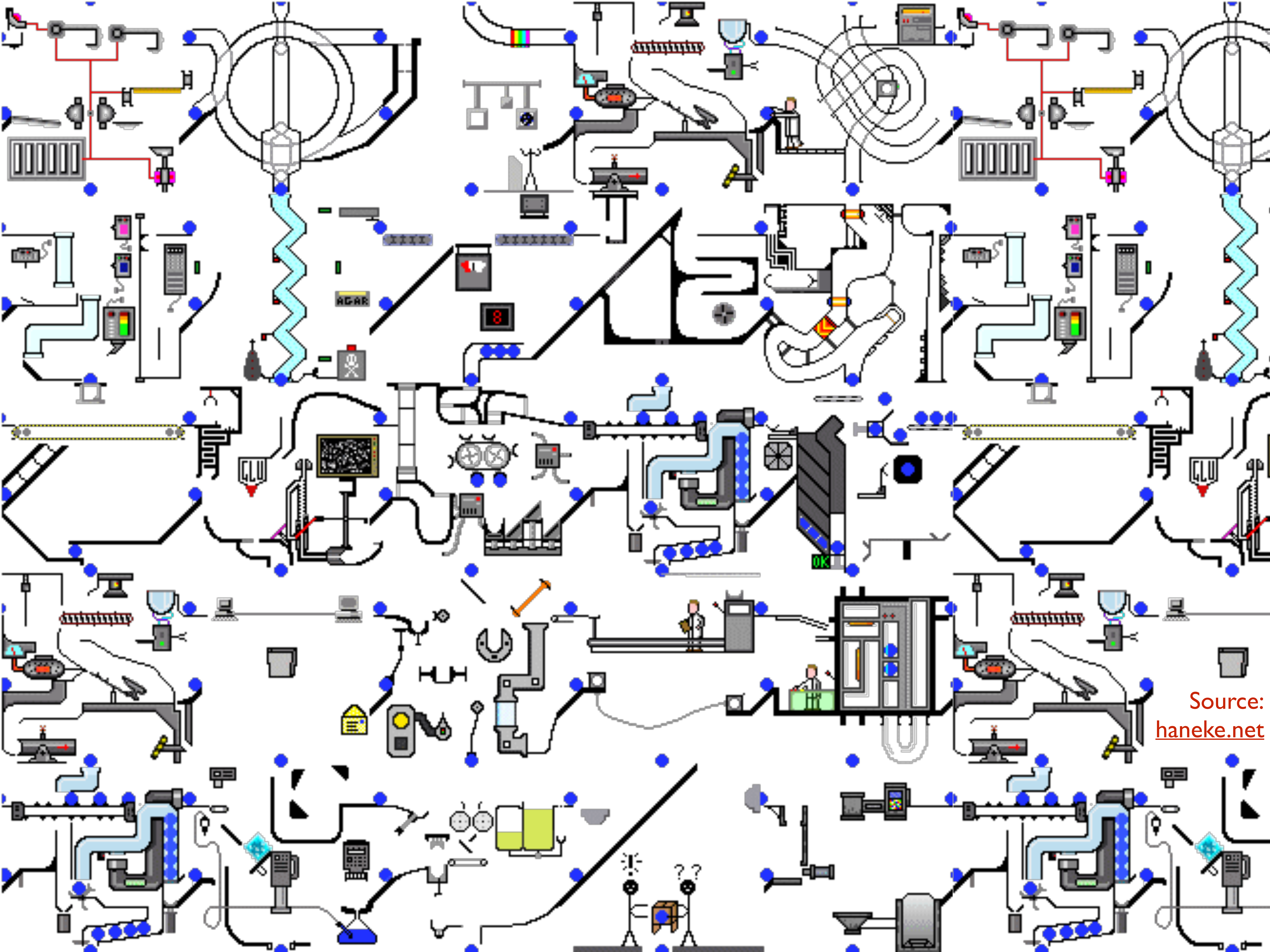
Many protocols and features used

```
version 12.4
service timestamps debug datetime msec
service timestamps log datetime msec
no service password-encryption
!
hostname PrimaryR1
!
boot-start-marker
boot-end-marker
!
!
no aaa new-model
!
!
ip cef
!
interface Loopback100
no ip address
!
interface GigabitEthernet0/1
description LAN port
ip address 64.X.X.1 255.255.255.224
ip nat inside
ip virtual-reassembly
duplex auto
speed auto
media-type rj45
no negotiation auto
standby 1 ip 64.X.X.5
standby 1 priority 105
standby 1 preempt delay minimum 60
standby 1 track Serial3/0
!
interface GigabitEthernet0/2
description conn to Backup Lightpath
ip address 65.X.X.66 255.255.255.240
ip nat outside
ip virtual-reassembly
duplex full
speed 100
media-type rj45
no negotiation auto
!
interface GigabitEthernet0/3
description LAN handoff from P2P to Denver
ip address 10.30.0.1 255.254.0.0
duplex auto
speed auto
media-type rj45
no negotiation auto
!
interface Serial1/0
description p-2-p to Denver DC
ip address 10.10.10.1 255.255.255.252
dsu bandwidth 44210
framing c-bit
cablelength 10
clock source internal
serial restart-delay 0
!
interface Serial3/0
description DS3 XO WAN interface
ip address 65.X.X.254 255.255.255.252
ip access-group 150 in
encapsulation ppp
dsu bandwidth 44210
framing c-bit
cablelength 10
serial restart-delay 0
!
router bgp 16XX
no synchronization
bgp log-neighbor-changes
network 64.X.X.0 mask 255.255.255.224
network 64.X.X.2
aggregate-address 64.X.X.0 255.255.255.0 summary-only
neighbor 64.X.X.2 remote-as 16XX
neighbor 64.X.X.2 next-hop-self
neighbor 65.X.1X.253 remote-as 2828
neighbor 65.X.X.253 route-map setLocalpref in
neighbor 65.X.X.253 route-map localonly out
no auto-summary
!
no ip http server
!
ip as-path access-list 10 permit ^$
ip nat inside source list 101 interface GigabitEthernet0/2 overload
!
access-list 101 permit ip any any
access-list 150 permit ip any any
!
route-map setLocalpref permit 10
set local-preference 200
!
route-map localonly permit 10
match as-path 10
!
control-plane
!
gatekeeper
shutdown
!
!
end
```

Example basic BGP+HSRP config from https://www.myriadsupply.com/blog/?p=259



http://www.cisco.com/c/en/us/products/collateral/services/high-availability/white_paper_c11-458050.html



Source: haneke.net

The Problem

Networks are complicated

- Just like any computer system
- Worse: it's distributed
- Even worse: no clean programming APIs, only "knobs and dials"

Network equipment is proprietary

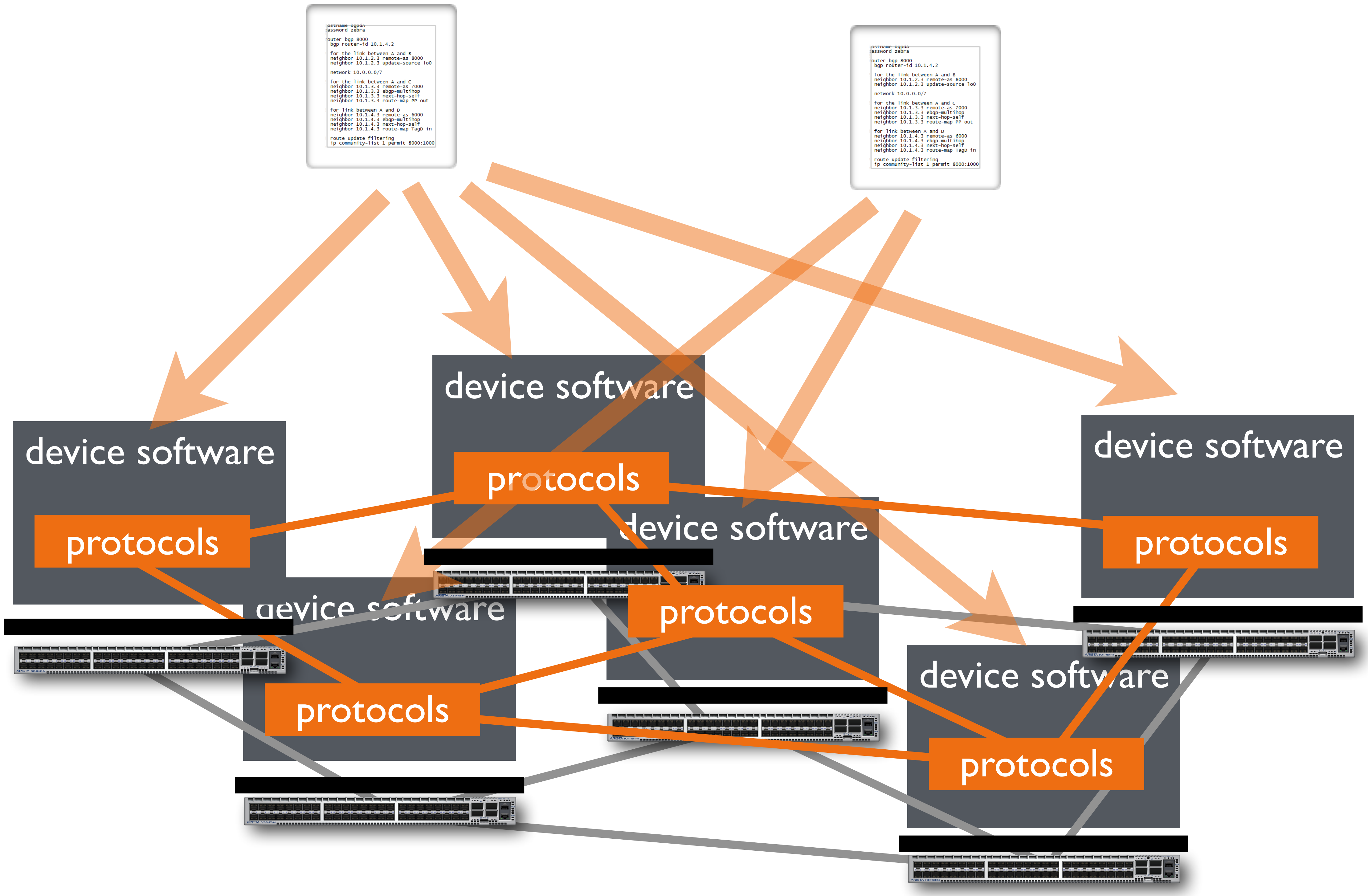
- Integrated solutions (software, configuration, protocol implementations, hardware) from major vendors

Result: Hard to innovate and modify networks

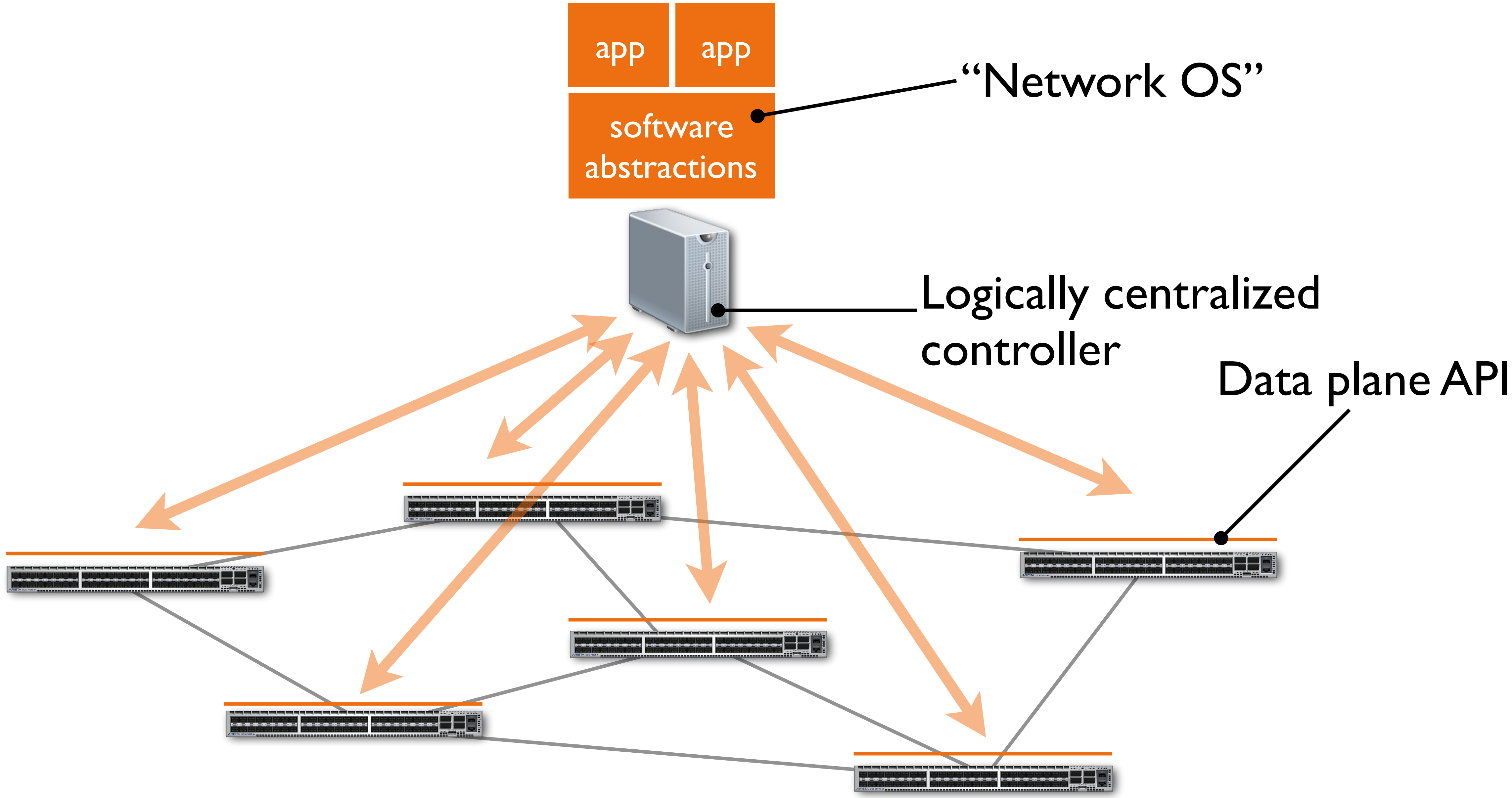
Traditional network

```
hostname bgpdA
password zebra
!
router bgp 8000
  bgp router-id 10.1.4.2

! for the link between A and B
  neighbor 10.1.2.3 remote-as 8000
  neighbor 10.1.2.3 update-source lo0

  network 10.0.0.0/7

! for the link between A and C
  neighbor 10.1.3.3 remote-as 7000
  neighbor 10.1.3.3 ebgp-multihop
  neighbor 10.1.3.3 next-hop-self
  neighbor 10.1.3.3 route-map PP out

! for link between A and D
  neighbor 10.1.4.3 remote-as 6000
  neighbor 10.1.4.3 ebgp-multihop
  neighbor 10.1.4.3 next-hop-self
  neighbor 10.1.4.3 route-map TagD in

! route update filtering
  ip community-list 1 permit 8000:1000
!
```

Traditional network



Software-defined network



Example

From NOX [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]

```
# On user authentication, statically setup VLAN tagging
# rules at the user's first hop switch
def setup_user_vlan(dp, user, port, host):
    vlanid = user_to_vlan_function(user)

    # For packets from the user, add a VLAN tag
    attr_out[IN_PORT] = port
    attr_out[DL_SRC] = nox.reverse_resolve(host).mac
    action_out = [(nox.OUTPUT, (0, nox.FLOOD)),
        (nox.ADD_VLAN, (vlanid))]
    install_datapath_flow(dp, attr_out, action_out)

    # For packets to the user with the VLAN tag, remove it
    attr_in[DL_DST] = nox.reverse_resolve(host).mac
    attr_in[DL_VLAN] = vlanid
    action_in = [(nox.OUTPUT, (0, nox.FLOOD)), (nox.DEL_VLAN)]
    install_datapath_flow(dp, attr_in, action_in)

nox.register_for_user_authentication(setup_user_vlan)
```

Example

From NOX [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]

```
# On user authentication, statically setup VLAN tagging
# rules at the user's first hop switch
def setup_user_vlan(dp, user, port, host):
    vlanid = user_to_vlan_function(user)

    # For packets from the user, add a VLAN tag
    attr_out[IN_PORT] = port
    attr_out[DL_SRC] = nox.reverse_resolve(host).mac
    action_out = [(nox.OUTPUT, (0, nox.FLOOD)),
        (nox.ADD_VLAN, (vlanid))]
    install_datapath_flow(dp, attr_out, action_out)

    # For packets to the user with the VLAN tag, remove it
    attr_in[DL_DST] = nox.reverse_resolve(host).mac
    attr_in[DL_VLAN] = vlanid
    action_in = [(nox.OUTPUT, (0, nox.FLOOD)), (nox.DEL_VLAN)]
    install_datapath_flow(dp, attr_in, action_in)

nox.register_for_user_authentication(setup_user_vlan)
```

Match specific set of packets

Example

From NOX [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]

```
# On user authentication, statically setup VLAN tagging
# rules at the user's first hop switch
def setup_user_vlan(dp, user, port, host):
    vlanid = user_to_vlan_function(user)

    # For packets from the user, add a VLAN tag
    attr_out[IN_PORT] = port
    attr_out[DL_SRC] = nox.reverse_resolve(host).mac
    action_out = [(nox.OUTPUT, (0, nox.FLOOD)),
        (nox.ADD_VLAN, (vlanid))]
    install_datapath_flow(dp, attr_out, action_out)

    # For packets to the user with the VLAN tag, remove it
    attr_in[DL_DST] = nox.reverse_resolve(host).mac
    attr_in[DL_VLAN] = vlanid
    action_in = [(nox.OUTPUT, (0, nox.FLOOD)), (nox.DEL_VLAN)]
    install_datapath_flow(dp, attr_in, action_in)

nox.register_for_user_authentication(setup_user_vlan)
```

Match specific set of packets

Construct action

Example

From NOX [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]

```
# On user authentication, statically setup VLAN tagging
# rules at the user's first hop switch
def setup_user_vlan(dp, user, port, host):
    vlanid = user_to_vlan_function(user)

    # For packets from the user, add a VLAN tag
    attr_out[IN_PORT] = port
    attr_out[DL_SRC] = nox.reverse_resolve(host).mac
    action_out = [(nox.OUTPUT, (0, nox.FLOOD)),
        (nox.ADD_VLAN, (vlanid))]
    install_datapath_flow(dp, attr_out, action_out)

    # For packets to the user with the VLAN tag, remove it
    attr_in[DL_DST] = nox.reverse_resolve(host).mac
    attr_in[DL_VLAN] = vlanid
    action_in = [(nox.OUTPUT, (0, nox.FLOOD)), (nox.DEL_VLAN)]
    install_datapath_flow(dp, attr_in, action_in)

nox.register_for_user_authentication(setup_user_vlan)
```

Match specific set of packets

Construct action

Install (match, action) in a specific switch

Example

From NOX [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]

```
# On user authentication, statically setup VLAN tagging
# rules at the user's first hop switch
def setup_user_vlan(dp, user, port, host):
    vlanid = user_to_vlan_function(user)

    # For packets from the user, add a VLAN tag
    attr_out[IN_PORT] = port
    attr_out[DL_SRC] = nox.reverse_resolve(host).mac
    action_out = [(nox.OUTPUT, (0, nox.FLOOD)),
        (nox.ADD_VLAN, (vlanid))]
    install_datapath_flow(dp, attr_out, action_out)

    # For packets to the user with the VLAN
    attr_in[DL_DST] = nox.reverse_resolve(h
    attr_in[DL_VLAN] = vlanid
    action_in = [(nox.OUTPUT, (0, nox.FLOOD
    install_datapath_flow(dp, attr_in, acti

nox.register_for_user_authentication(setup_
```

Match specific set of packets

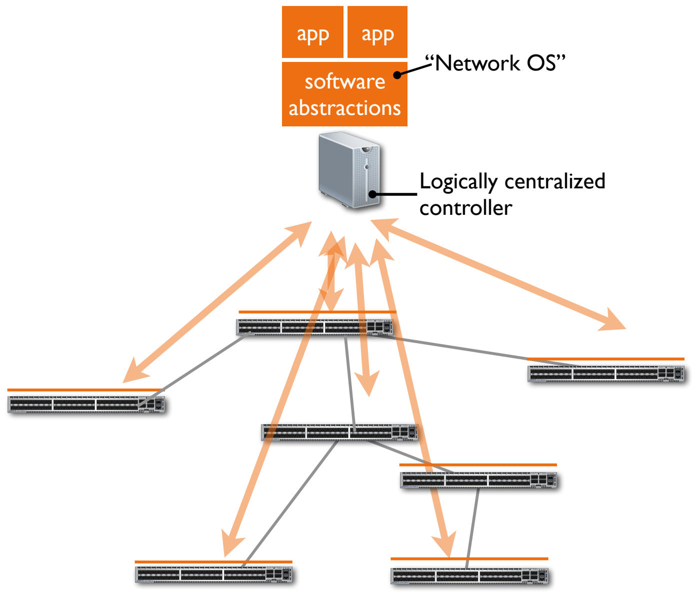
Construct action

Install (match, action) in a specific switch

Common primitives:

- Match packets, execute actions (rewrite, forward packet)
- Topology discovery
- Monitoring

Evolution of SDN



[Graphic: José-Manuel Benitos]

Flexible Data Planes

Label switching / MPLS (1997)

- "Tag Switching Architecture Overview", [Rekhter, Davie, Rose, Swallow, Farinacci, Katz, Proc. IEEE, 1997]
- Set up explicit paths for classes of traffic

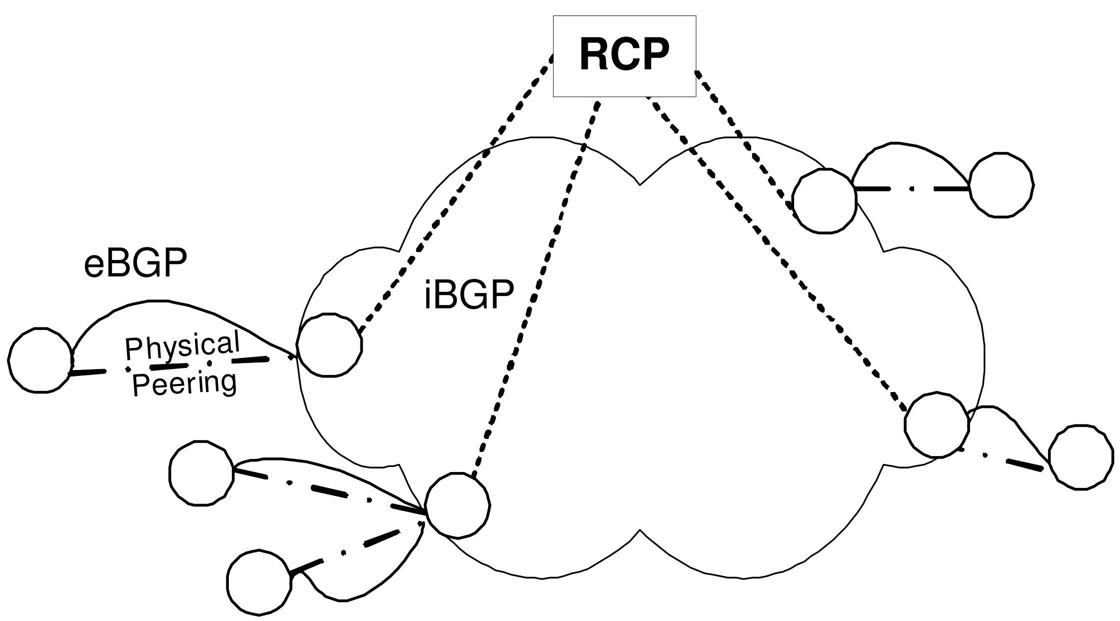
Active Networks (1999)

- Packet header carries (pointer to) program code

Logically Centralized Control

Routing Control Platform (2005)

- [Caesar, Caldwell, Feamster, Rexford, Shaikh, van der Merwe, NSDI 2005]
- Centralized computation of BGP routes, pushed to border routers via iBGP

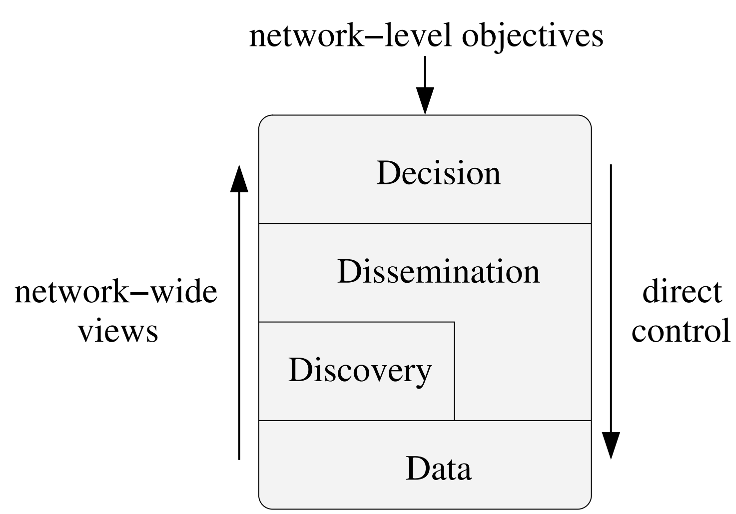


Logically Centralized Control

Routing Control Platform (2005)

4D architecture (2005)

- A Clean Slate 4D Approach to Network Control and Management [Greenberg, Hjalmtysson, Maltz, Myers, Rexford, Xie, Yan, Zhan, Zhang, CCR Oct 2005]
- Logically centralized "decision plane" separated from data plane



Logically Centralized Control

Routing Control Platform (2005)

4D architecture (2005)

Ethane (2007)

- [Casado, Freedman, Pettit, Luo, McKeown, Shenker, SIGCOMM 2007]
- Centralized controller enforces enterprise network Ethernet forwarding policy using existing hardware

Logically Centralized Control

Routing Control Platform (2005)

4D architecture (2005)

Ethane (2007)

- [Casado, Freedman, Pettit, L
  SIGCOMM 2007]
- Centralized controller enfo
  Ethernet forwarding policy

```
# Groups —
desktops = ["griffin","roo"];
laptops = ["glaptop","rlaptop"];
phones = ["gphone","rphone"];
server = ["http_server","nfs_server"];
private = ["desktops","laptops"];
computers = ["private","server"];
students = ["bob","bill","pete"];
profs = ["plum"];
group = ["students","profs"];
waps = ["wap1","wap2"];
%%
# Rules —
[(hsrc=in("server")∧(hdst=in("private"))] : deny;
# Do not allow phones and private computers to communicate
[(hsrc=in("phones")∧(hdst=in("computers"))] : deny;
[(hsrc=in("computers")∧(hdst=in("phones"))] : deny;
# NAT-like protection for laptops
[(hsrc=in("laptops")] : outbound-only;
# No restrictions on desktops communicating with each other
[(hsrc=in("desktops")∧(hdst=in("desktops"))] : allow;
# For wireless, non-group members can use http through
# a proxy. Group members have unrestricted access.
[(apsrc=in("waps"))∧(user=in("group"))] :allow;
[(apsrc=in("waps"))∧(protocol="http)] : waypoints("http-proxy");
[(apsrc=in("waps"))] : deny;
[]: allow; # Default-on: by default allow flows
```

Figure 4: A sample policy file using *Pol-Eth*

Logically Centralized Control

Routing Control Platform (2005)

4D architecture (2005)

Ethane (2007)

OpenFlow (2008)

- [McKeown, Anderson, Balakrishnan, Parulkar, Peterson, Rexford, Shenker, Turner, CCR 2008]
- Thin, standardized interface to data plane
- General-purpose programmability at controller

Evolution of SDN:

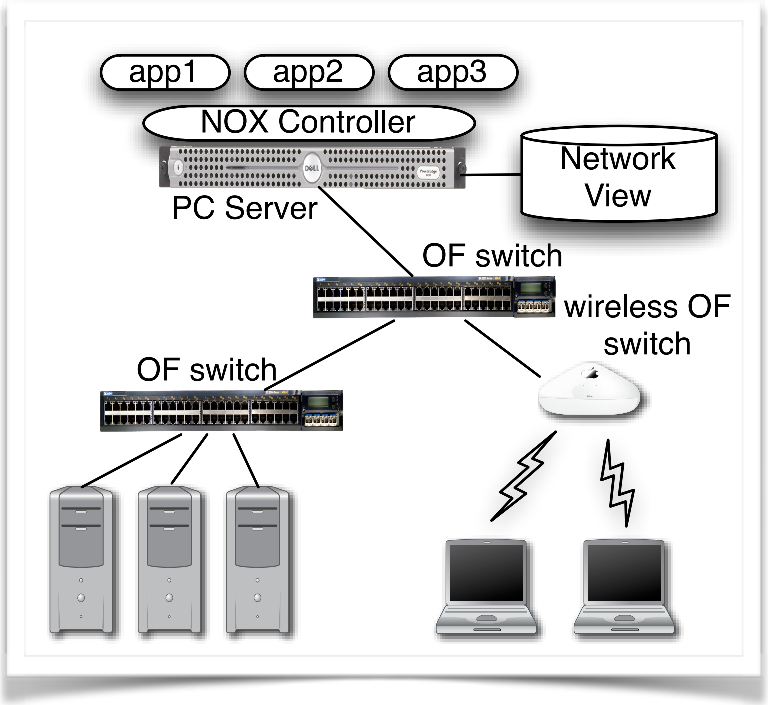
Routing Control Platform (2005)

4D architecture (2005)

Ethane (2007)

OpenFlow (2008)

NOX (2008)



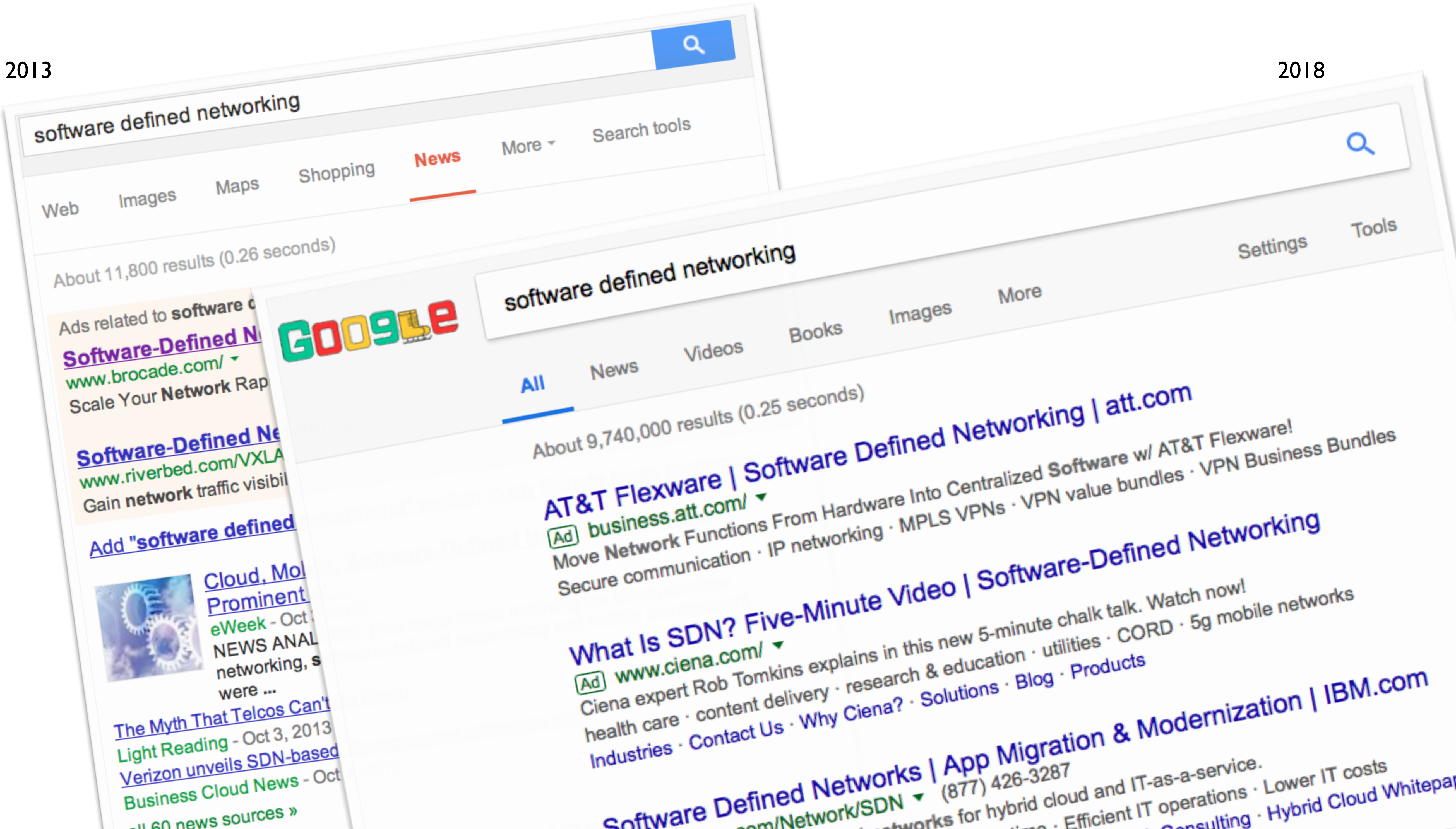
- [Gude, Koponen, Pettit, Pfaff, Casado, McKeown, Shenker, CCR 2008]
- First OF controller: centralized network view provided to multiple control apps as a database
- Behind the scenes, handles state collection & distribution

Evolution of SDN

Industry explosion (~2010+)

2013

2018



Opportunities

Open data plane interface

- Hardware: Easier for operators to change hardware, and for vendors to enter market
- Software: Can more directly access device behavior

Centralized controller

- Direct programmatic control of network

Software abstractions on the controller

- Solve dist. sys. problems once, then just write algorithms
- Libraries/languages to help programmers write net apps
- Systems to write high level policy instead of programming

Opportunities

Open data plane interface

- Hardware: Easier for operators to change hardware, and for vendors to enter market
- Software: Can more directly access device behavior

Centralized controller

- Direct programmatic control of network

Software abstractions on the cont

- Solve dist. sys. problems once, th
- Libraries/languages to help progr
- Systems to write high level policy
  programming

All active areas of current research!

Challenges for SDN

Performance and scalability

Distributed system challenges still present

- Resilience of "logically centralized" controller
- Imperfect knowledge of network state
- Consistency issues between controllers

Challenges for SDN

Reaching agreement on data plane protocol

- OpenFlow? NFV functions? Whitebox switching? Programmable data planes?

Devising the right control abstractions

- Programming OpenFlow: far too low level
- But what are the right high-level abstractions to cover important use cases?

Q: When do you control the net?

When does the SDN controller send instructions to switches?

- ...in the OpenFlow paper?
- ...other options?

Q: When do you control the net?

When does the SDN controller send instructions to switches?

- ...in the OpenFlow paper? Reactive (when packet arrives needing forwarding rule)
- ...other options? Proactive (in advance of need)

Q: How does SDN affect reliability?

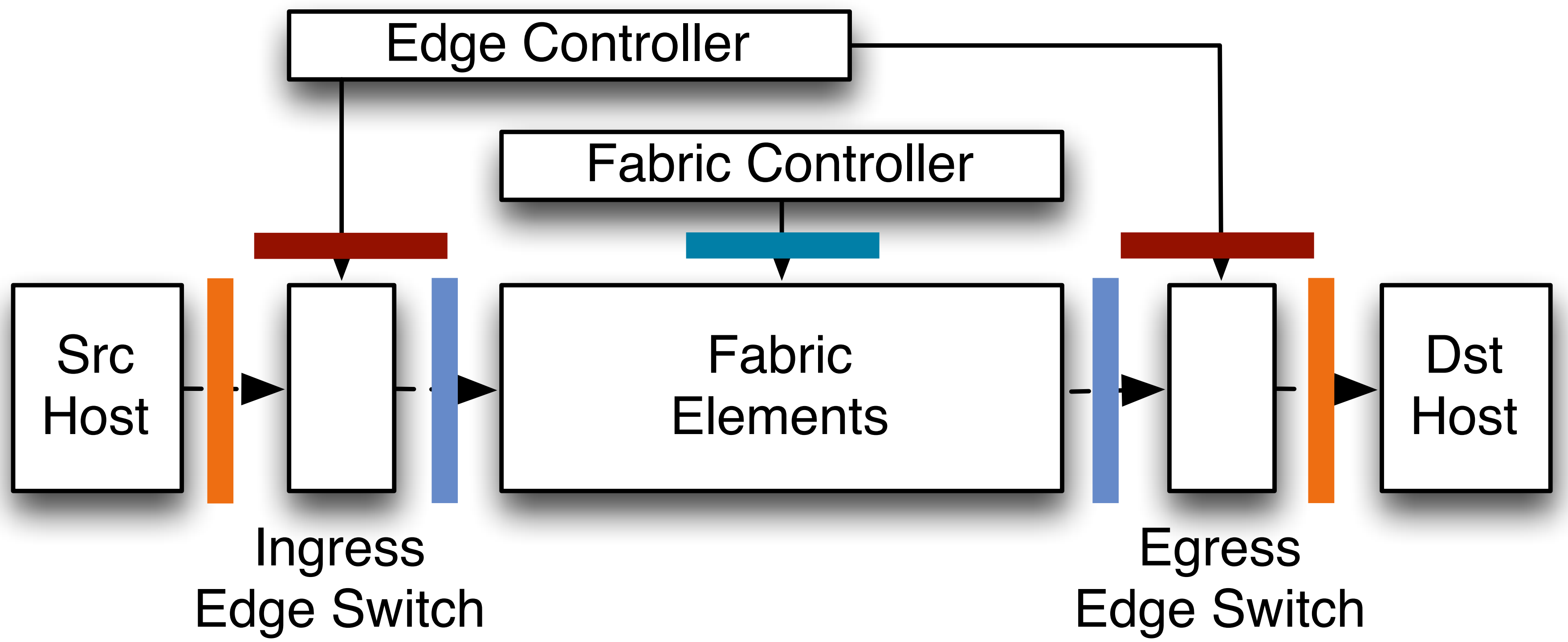
More bugs in the network, or fewer?

From SDN to Fabric

[Casado, Koponen, Shenker, Tootoonchian, HotSDN'12]

Separate interfaces:

- Host-network (external-to-internal data plane)
- Operator-network
- Packet-switch (internal data plane)



Fabric discussion

Q: "Host-Network and Packet-Switch interfaces were identical" in the Internet. How is this a simplification?

Q: Does OF meet the "ideal network" goals the Fabric paper lays out?:

- Simplified hardware
- Vendor-neutral hardware
- "Future-proof" hardware
- Flexible software

Q: Drivers of early deployment?

What drove early deployment of OpenFlow & SDN?

Access control in enterprises? Net research?

- Good ideas, are already valuable
- But not the "killer apps" for initial large-scale deployment

The first "Killer Apps" for SDN

Inter-datacenter traffic engineering

- Drive utilization to near 100% when possible
- Protect critical traffic from congestion

Cloud virtualization

- Create separate virtual networks for tenants
- Allow flexible placement and movement of VMs

Key characteristics of the above use cases

- Special-purpose deployments with less diverse hardware
- Existing solutions aren't just inconvenient, *they don't work!*

SDN today

Software Defined WAN (SD-WAN)

- Overlay network connecting enterprise sites across the Internet instead of traditional MPLS service
- Note: Not the same as Google's B4

SDN in service provider networks

- Central control of virtualized network functions (VNFs)

Controllers that use traditional configs instead of OF

- e.g., "API" into the device is a BGP config
- Automate configuring a data center or cluster in an enterprise

Next up

Monday: SDN in the WAN

Brighten out on jury duty next week