SDN in Wide-Area Networks

May 8-10, 2017, Valencia, Spain

Oliver Michel

University of Colorado
Boulder
Global IP Traffic Growth

[Cisco® Visual Networking Index (VNI) 2016]
Global IP Traffic Growth

Global IP Traffic [Exabytes/Month]

- 2015: 72.5
- 2016: 88.7
- 2017: 108.5
- 2018: 132.1
- 2019: 160.6
- 2020: 194.4

52% → 66%

[Cisco® Visual Networking Index (VNI) 2016]
Global IP Traffic Growth

- Global IP Traffic [Exabytes/Month]

- 2015: 72.5
- 2016: 88.7
- 2017: 108.5
- 2018: 132.1
- 2019: 160.6
- 2020: 194.4

[52% → 66%  70% → 82%]

[Cisco® Visual Networking Index (VNI) 2016]
SDN in one Slide

Data Plane

Control Plane

centralized, consolidated control
SDN Evolution
1. Active Networks
SDN Evolution

1. Active Networks

2. Data Plane / Control Plane Separation
   [Casado. Ethane. SIGCOMM ’07, Greenberg. 4D. SIGCOMM CCR ’05, Caesar. RCP. ‘05]
SDN Evolution

1. Active Networks

2. Data Plane / Control Plane Separation
   [Casado. Ethane. SIGCOMM ’07, Greenberg. 4D. SIGCOMM CCR ’05, Caesar. RCP. ‘05]

3. Control Protocols
   [McKeown. OpenFlow. SIGCOMM CCR ’08]
1. Network Virtualization

2. Resource Management
WAN Challenges
WAN Challenges

- controlled environment
- dedicated control networks
- fewer external factors
WAN Challenges

- controlled environment
- dedicated control networks
- fewer external factors

- fibers in ducts along highways/pipelines
- in-band control
WAN Challenges

• Legacy Equipment, Protocols and Practices
• Different Domains, Stakeholders
• Interoperability Requirements

Facebook Wedge Platform

Juniper Networks PTX3000 Core Router
WAN Opportunities

- tree-like networks with high degree of parallel links
- less expensive copper cabling

- mesh network with fewer parallel links
- expensive wide-area fibers and optics
WAN Opportunities

Any link or node in the server layer that is shared by multiple abstract links can be the basis for a separate SRLG, and an abstract link will typically be associated with a string of SRLGs.

Multiple abstract links can share the same server-layer links, in which case they are part of the same SRLG. This can be achieved by summarizing the detailed design of the transport layer topology to the minimum set of information required to address relevant multilayer TE use cases. The IP/MPLS layer only requires network topology and reachability on the server layer, as well as the metrics of these links such as bandwidth, latency, SRLGs, and so on.

This allows for higher scalability in engineering accuracy. Detailed information on network element connectivity and optical transmission impairments between both network layers.

Figure 3 shows the abstraction of the transport layer into a set of abstract nodes and links. A link that connects a transport-layer node and a node in the IP/MPLS layer is referred to as an access link. Every node on the transport layer (client layer). This abstracted topology model consists of a set of abstract links that represent the end-to-end information is shared between both network layers.

Abstract link

SDN in Wide-Area Networks | O. Michel, University of Colorado Boulder

IP/MPLS

WDM OTN
Challenges for SDN in Wide-Area Networks
Challenges for SDN in Wide-Area Networks

1. Distributing SDN Controller State
Challenges for SDN in Wide-Area Networks

1. Distributing SDN Controller State

2. Placing Controller Instances
   [Heller. Controller Placement. HotSDN ‘12]
Challenges for SDN in Wide-Area Networks

1. Distributing SDN Controller State

2. Placing Controller Instances
   [Heller. Controller Placement. HotSDN ‘12]

3. Updating SDN Switches in a consistent Manner
   [Reitblatt. Consistent Updates. SIGCOMM ’12, Jin. Dionysus. SIGCOMM ‘14]
Distributing SDN Controller State

- **Kandoo** [Yeganeh HotSDN ’12]
  - hierarchical model, reduces controller traffic
Distributing SDN Controller State

- **Kandoo** [Yeganeh HotSDN ’12]
  - hierarchical model, reduces controller traffic
- **ONOS** [Berde. ONOS. HotSDN ’14]
  - distributed, eventually consistent network graph through Cassandra backend
Placing SDN Controller Instances

The Controller Placement Problem

[Heller ’12]

- 3 fundamental underlying problems
  1. average-case latency
     minimum k-median
  2. worst-case latency
     minimum k-center
  3. nodes within latency bound
     maximum cover

- cost/benefit analysis: single or pair of controllers often enough

Figure 1: Optimal placements for 1 and 5 controllers

Figure 2: Latency CDFs for all possible controller placements

Figure 3: Cost-benefit ratio for optimized average latency

Figure 4: Number of controllers required for different placements
Consistent Data Plane Updates

Problem 1

current state
Consistent Data Plane Updates

Problem 1

current state

target state
Consistent Data Plane Updates

Problem 1

Current state

Target state

Possible intermediate state
Consistent Data Plane Updates

Problem 2

current state
Consistent Data Plane Updates

Problem 2

current state

target state
Consistent Data Plane Updates

Problem 2

current state

intermediate state

target state
Consistent Data Plane Updates

Consistent Network Updates
[Reitblatt ’12]

• abstract update operation where a set of packets is guaranteed to receive consistent treatment

• per-packet or per-flow consistency

• implementation on top of NOX
Consistent Data Plane Updates

Dynamic Scheduling of Network Updates
[Jin ’14]

- schedule update order dynamically at runtime accounting for runtime variations
- critical path scheduling through dependency graph
SD-WAN Deployments and Benefits
SD-WAN Deployments and Benefits

Intra-Domain
- Distributed Applications
- Inter-DC Networks
- Synchronization
- Backup
SD-WAN Deployments and Benefits

Intra-Domain
- Distributed Applications
- Inter-DC Networks
- Synchronization
- Backup

Inter-Domain
- Content Delivery
- Peering
- BGP inflexibilities
Achieving High Utilization with Software-Driven WAN

[Hong ’13]

- central control of
  - bandwidth allocation for different services
  - centrally computing globally-optimal paths
- frequent data plane updates to maintain high utilization
- congestion-free updates through scratch capacity
SD-WAN Deployments and Benefits

Achieving High Utilization with Software-Driven WAN

[Hong ’13]

- central control of
  - bandwidth allocation for different services
  - centrally computing globally-optimal paths
- frequent data plane updates to maintain high utilization
- congestion-free updates through scratch capacity

B4: Experience with a Globally-Deployed Software Defined WAN

[Jain ’13]

- integration with legacy routing protocols
- evaluation in production network over three years
Expanding beyond a single Domain

SDX: A Software Defined Internet Exchange

[Gupta ’14]

- BGP inflexibilities: indirect control over forwarding
- new use-cases: e.g. application specific peering
- SDN advantages: direct, fine-grained control
- IXPs: natural starting point
Traffic Engineering, Data Plane Fault Tolerance, and Low-Latency Routing

- high uncertainty and randomness in path quality
- active probing and SDN control can help to dynamically change paths
- can in part be done in the data plane (e.g., P4 technologies)
Packet-Optical Convergence

- routing over a more complex topology
- IP layer routing can use transport layer properties for CSPF routing
Internet-Scale Attacks

• use centralized logic for analysis and mitigation of Internet-scale attacks across domains
• fine-grained filtering with programmability
Internet-Scale Attacks

- use centralized logic for analysis and mitigation of Internet-scale attacks across domains
- fine-grained filtering with programmability
Internet-Scale Attacks

- use centralized logic for analysis and mitigation of Internet-scale attacks across domains
- fine-grained filtering with programmability
Conclusion

- WANs gaining important with mobile traffic rising
- some deployments, typically within domains
- still space for extensive research
Q&A / Discussion

Oliver Michel

oliver.michel@colorado.edu
http://nsr.colorado.edu/oliver
Backup Slides
Active Networks

- rapid traffic growth in mid-‘90s, slow standardization through IETF
Active Networks

• rapid traffic growth in mid-‘90s, slow standardization through IETF

• programmability
Active Networks

- rapid traffic growth in mid-‘90s, slow standardization through IETF
- programmability
- code embedded in packets
Active Networks

- rapid traffic growth in mid-‘90s, slow standardization through IETF
- programmability
- code embedded in packets
- no clear use-cases or applications
Control- and Data Plane Separation

- networks rapidly increasing in size and complexity
Control- and Data Plane Separation

- networks rapidly increasing in size and complexity
- scalability issues
Control- and Data Plane Separation

- networks rapidly increasing in size and complexity
- scalability issues
- manageability issues
Scalability
Scalability

Routing Control Platform (RCP) [Caesar ’05]

- routers peer with RCP
- mimics full iBGP mesh
- single best route advertised via standard iBGP
- intrinsic correctness of full mesh with scalability of RR
- no route oscillations or forwarding loops
Manageability

**Set Interface Configuration**
- `set interfaces ge-1/2/0 unit 0 family inet address 172.16.1.1/24`
- `interface gigabitethernet 2 ip address 172.16.1.1 255.255.255.0`

**Protocols and Features**
- OSPF
- BGP
- RIP
- IS-IS
- VLAN
- SNMP
- RSVP
- LDP
Manageability

• novel architectural pattern for networks based on layers [Greenberg '05]

  • four different layers
  • control/data separation

• high-level network policies through centralized controller [Casado '07]

  • simple switch architecture
  • evaluated in real-world deployment
Control Protocols

- need for standardized control between control and data plane
- generalization of networking equipment
Control Protocols

• OpenFlow [McKeown ’08]

• open protocol that gives applications control over a switches data plane

• designed around a set of header match fields and forwarding actions

• forwarding abstraction balancing…

  1. general match/action (TCAM model)

  2. fixed-function switch ASICs

• not the only protocol
fundamental scaling problems in SDX architecture

- composition of rules requires large state
- more rules than policies defined needed due to BGP congruence checking
Open Network Operating System [Berde ’14]

- global network view shared across all instances
- scale-out and failure resiliency
- each switch connected to primary OF Manager
- new primary selected at failure through consensus protocol by Zookeeper
- distributed, eventually consistent network graph through Cassandra backend
NorthStar Controller

- Multi-Layer WAN Traffic Engineering Solution
- Controller-Controller Interface
Expanding beyond a single Domain

SDX-enabled Route Server (RS)

AS A router

SDX-Fabric

AS B router

AS C router

BGP session

AS A's outbound policy:
application-specific peering

(match(dstport=80) >> fwd(B)) +
(match(dstport=443) >> fwd(C))

AS B's inbound policy:
traffic engineering

(match(srcip={0/1}) >> fwd(B1)) +
(match(srcip={128/1}) >> fwd(B2))

virtual switch

virtual switch

virtual switch

physical port

virtual port

virtual switch

virtual switch

virtual switch

AS A

B

A

C

AS B

AS C

AS A has a virtual switch connecting to AS B

AS A has an outbound policy. Inbound
AS B has an inbound policy that
directs traffic with source IP addresses
starting with 0 to B's
Virtual Switch. To do so, the SDX allows
each AS to write routing policies for its
virtual switches. Each AS can define
inbound and outbound policies as if
it is the only participant at the SDX.

Yet, the virtual switches of ASes B and C,
where each AS can write policies, interact
with each other. For example, AS A
has an outbound policy that
directories traffic with source IP addresses
starting with 0 to AS B's virtual switch.
AS B has an inbound policy that
directs traffic with source IP addresses
starting with 0 to AS B's virtual switch.

The SDX must combine the
policies of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.

The SDX must combine the
courses of ASes B and C, where each AS can
write policies, to determine
how to forward packets on
their virtual switches.