Student Online Teaching Advice Notice

The materials and content presented within this session are intended solely for use in a context of teaching and learning at Trinity.

Any session recorded for subsequent review is made available solely for the purpose of enhancing student learning.

Students should not edit or modify the recording in any way, nor disseminate it for use outside of a context of teaching and learning at Trinity.

Please be mindful of your physical environment and conscious of what may be captured by the device camera and microphone during videoconferencing calls.

Recorded materials will be handled in compliance with Trinity's statutory duties under the Universities Act, 1997 and in accordance with the University's policies and procedures.

Further information on data protection and best practice when using videoconferencing software is available at https://www.tcd.ie/info_compliance/data-protection/.

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

CSU33031 Computer Networks

Rerun: Routing

Stefan Weber

email: sweber@tcd.ie

Office: Lloyd 1.41

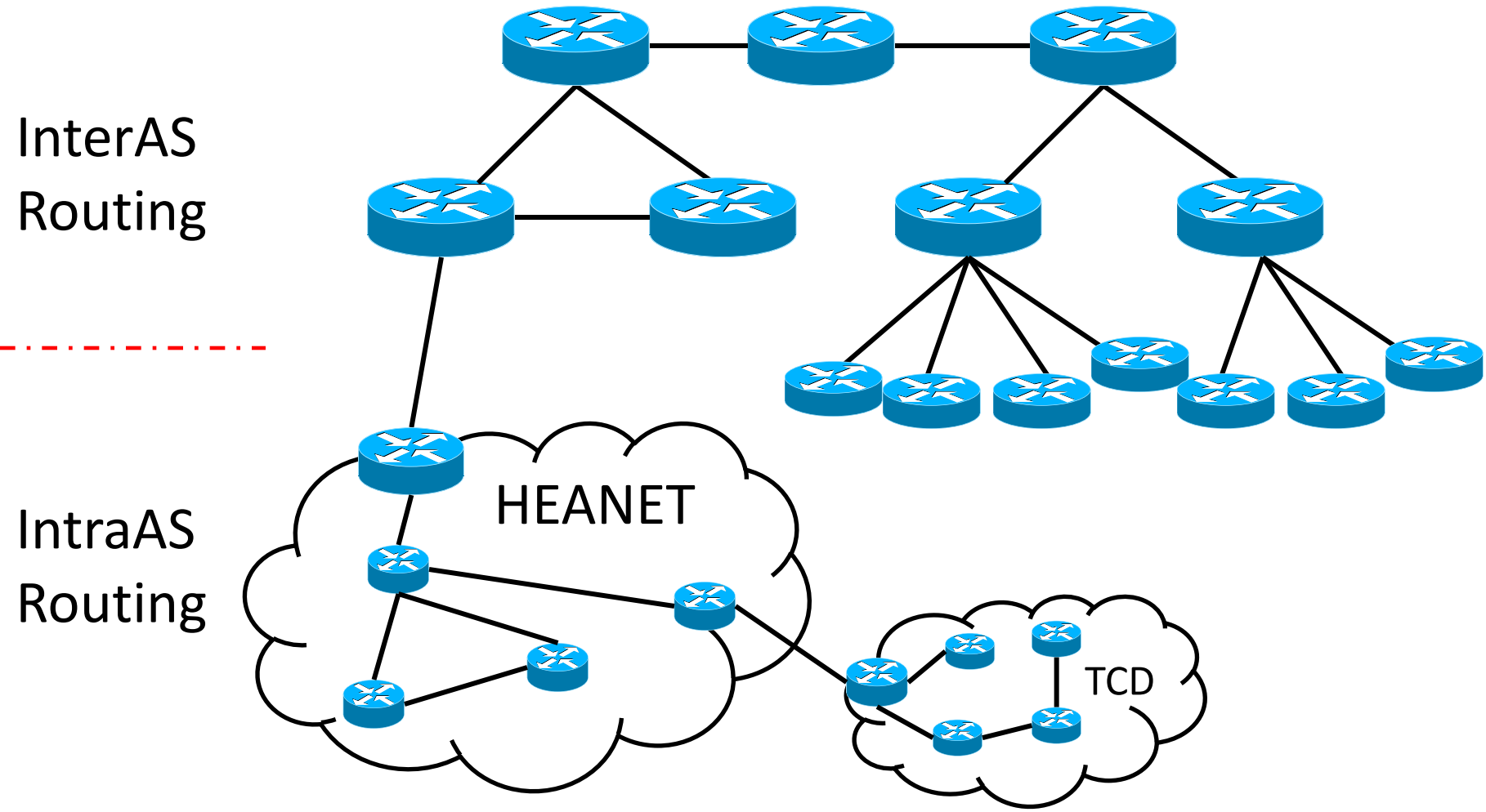
Plan for today

- Rerun of Routing
- OpenFlow
- SD-WAN
- Assignment 2

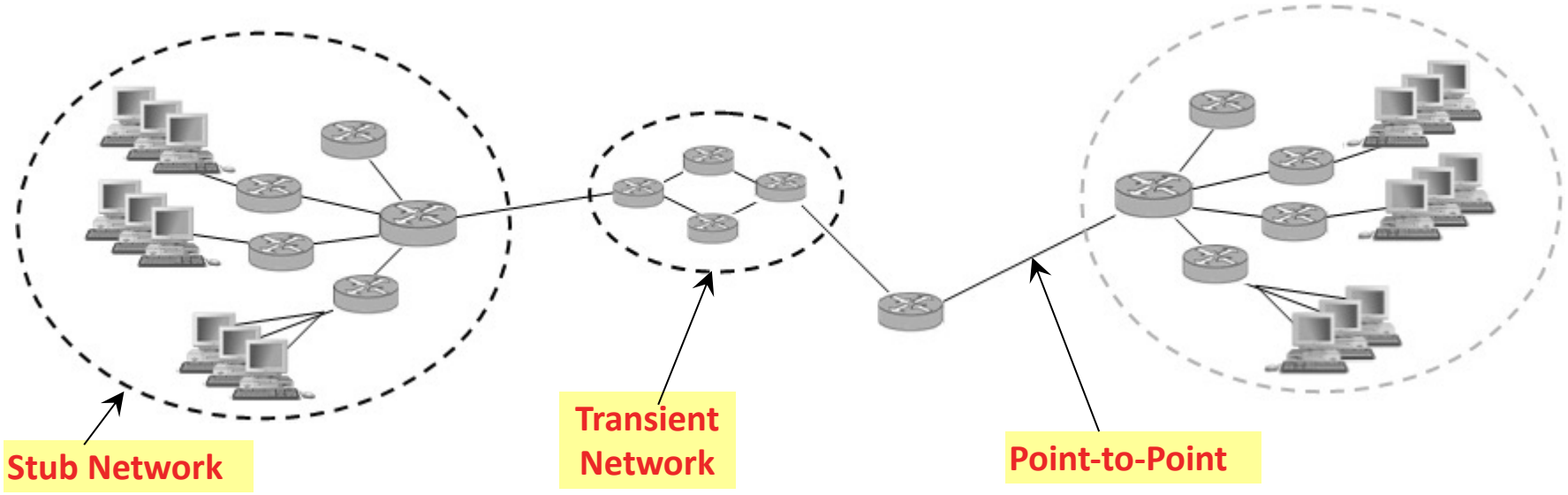
InterAS vs IntraAS Routing

InterAS Routing

IntraAS Routing

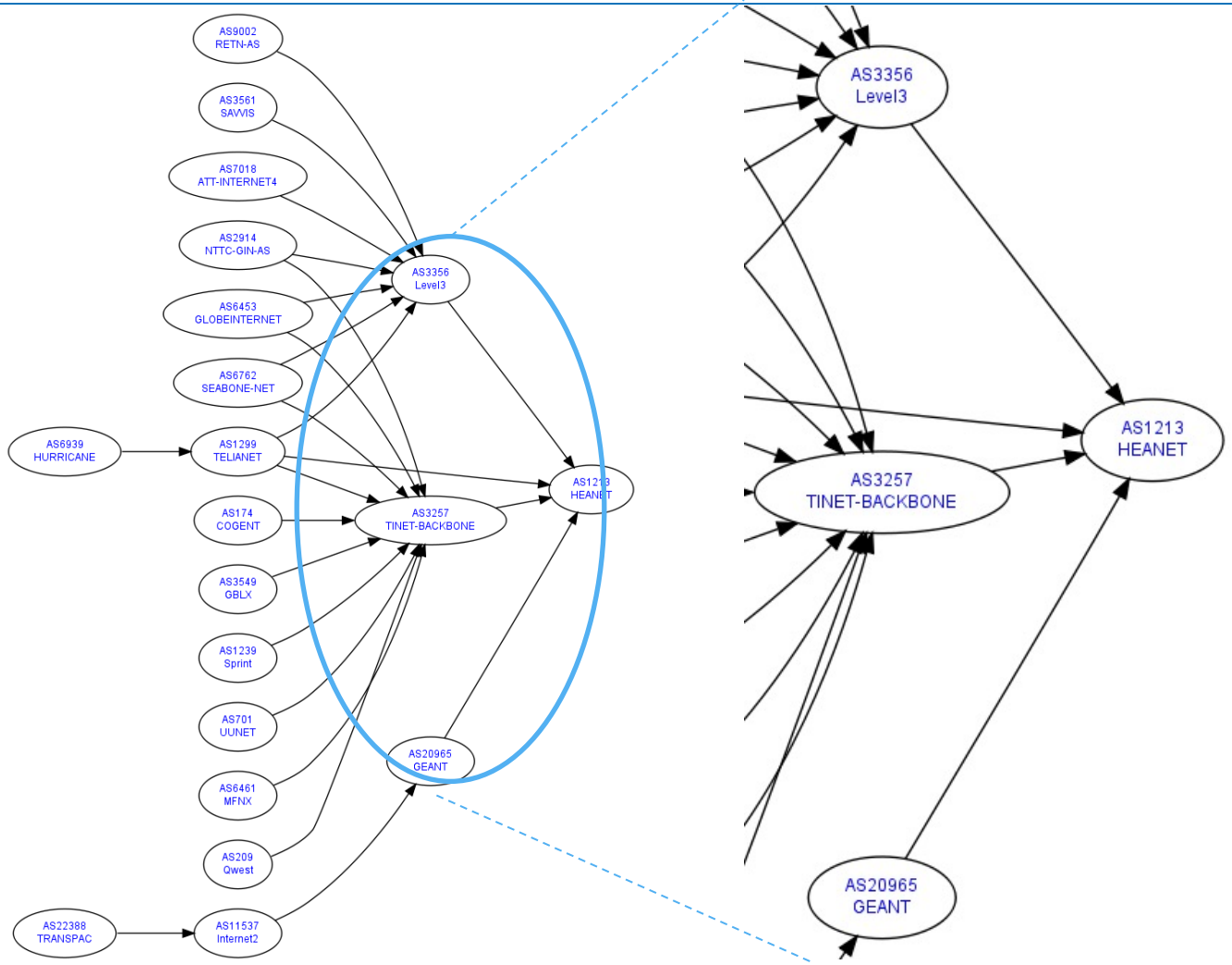


Autonomous Systems

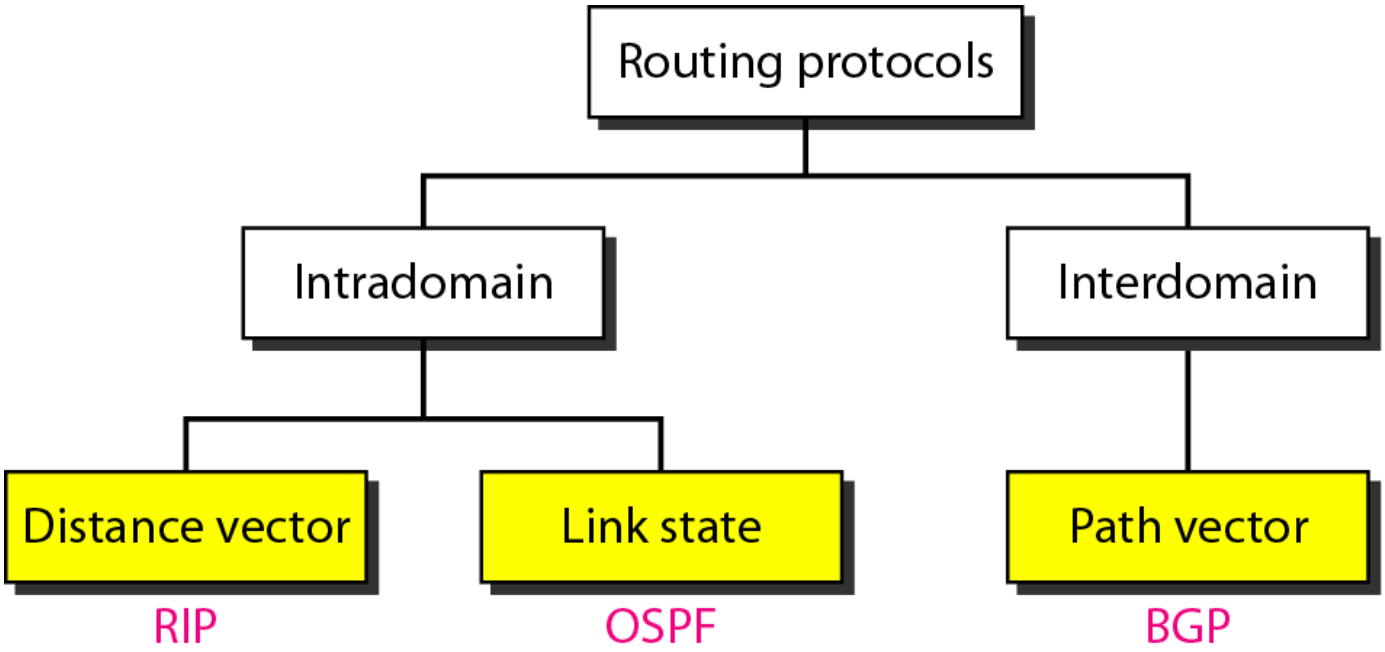


- Stub network
  - Network that does not forward to other network
- Transient network
  - Network that forwards traffic between other networks
- Point-to-point link

AS1213 - HEANET



Routing Protocols



- Interior Routing Protocols
  - Routing within ASs
- Exterior Routing Protocols
  - Routing between ASs

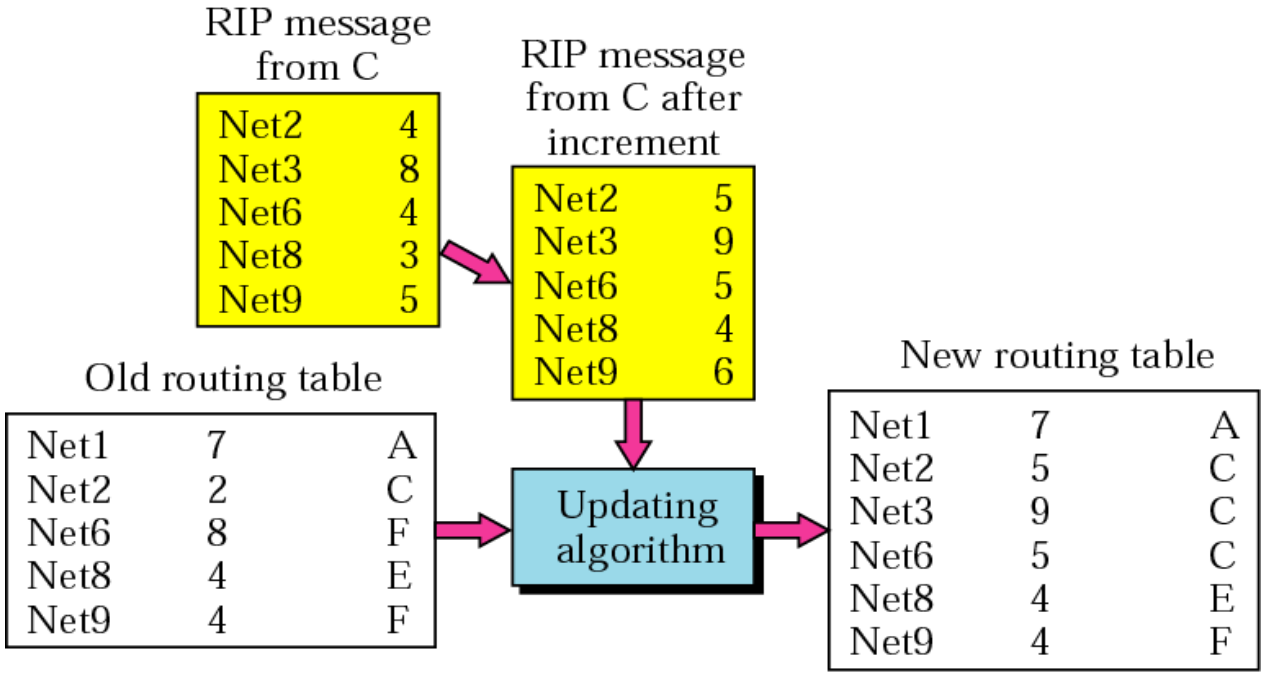
* Figure is courtesy of B. Forouzan

Routing Approaches

- Distance Vector routing
  - Routes propagate through exchange of routing tables
  - Based on communication with neighbours

- Link State routing
  - Establishing view of complete topology
  - Makes use of Dijkstra's Shortest-Path Algorithm

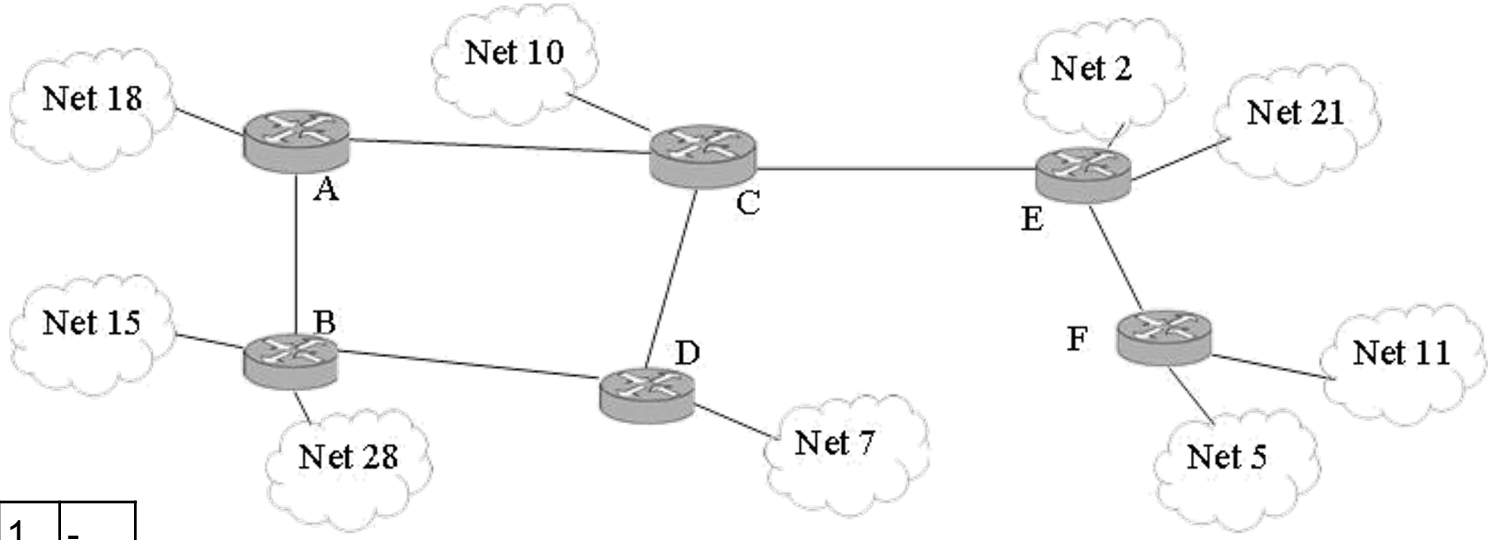
Distance Vector Routing



Net1: No news, do not change
Net2: Same next hop, replace
Net3: A new router, add
Net6: Different next hop, new hop count smaller, replace
Net8: Different next hop, new hop count the same, do not change
Net9: Different next hop, new hop count larger, do not change

* Figure is courtesy of B. Forouzan

Distance Vector: Example IV

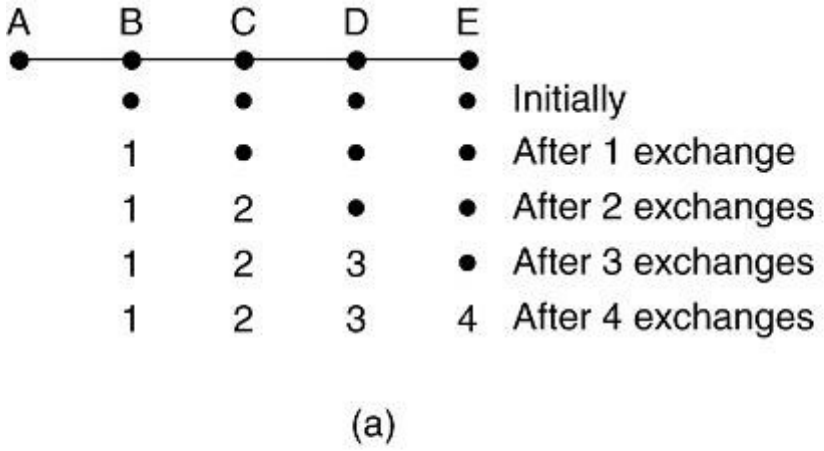


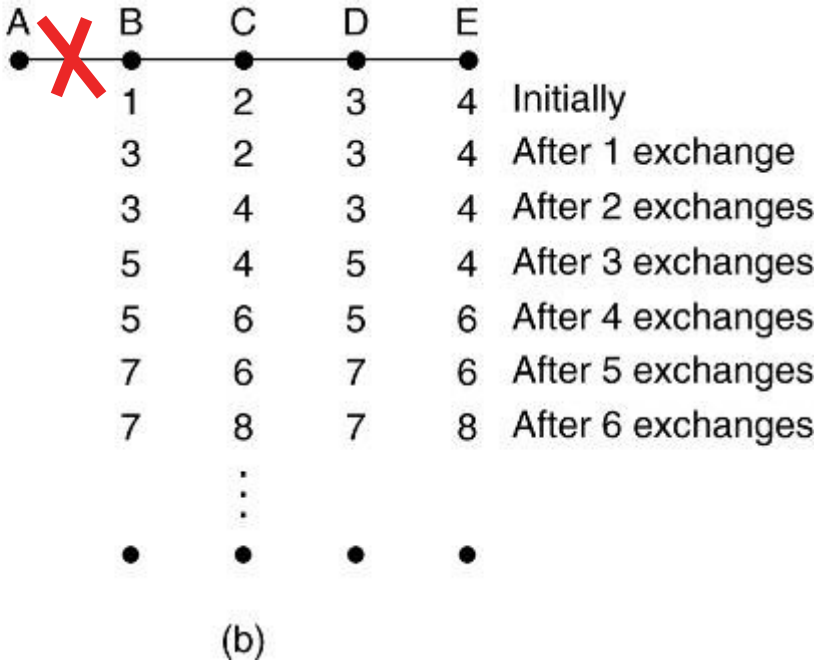
A

Net18	1	-
Net15	2	B
Net28	2	B
Net10	2	C
Net7	3	B
…	…	…
Net11	5	C
Net5	5	C

- Time required to build complete routing tables for all nodes
  - this is known as convergence

Count-to-Infinity Problem





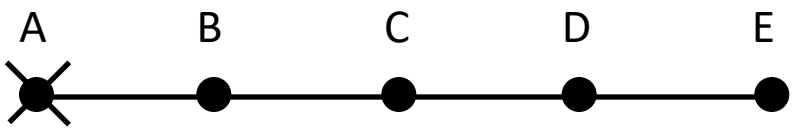
- Routers exchange updates
- After 4 steps the network converges
- Every router knows how to get to router A

- Link between A and B breaks
- B receives update from C advertising route to A

* Figure is courtesy of A. Tanenbaum

Split Horizon with Poisoned Reverse

Report “split-horizon” routes as infinity to break loops on the first routing exchange.



inf.	2	3	4	B learns A is dead
inf.	inf. → 2	3	4	B reports to C that A’s metric is inf.
inf.	inf.	3	4	After 1 exchange
inf.	inf.	inf.	4	After 2 exchanges
inf.	inf.	inf.	inf.	After 3 exchanges

* Taken from X. Li, Rutgers

Link State Routing

- Sharing knowledge about neighbourhood
- Sharing with every other router

- Uses Dijkstra's Shortest-Path Algorithm to calculate the routing table

- Open Shortest Path First (OSPF)
  - Divides autonomous system into areas
  - Border routers summarize information for an area

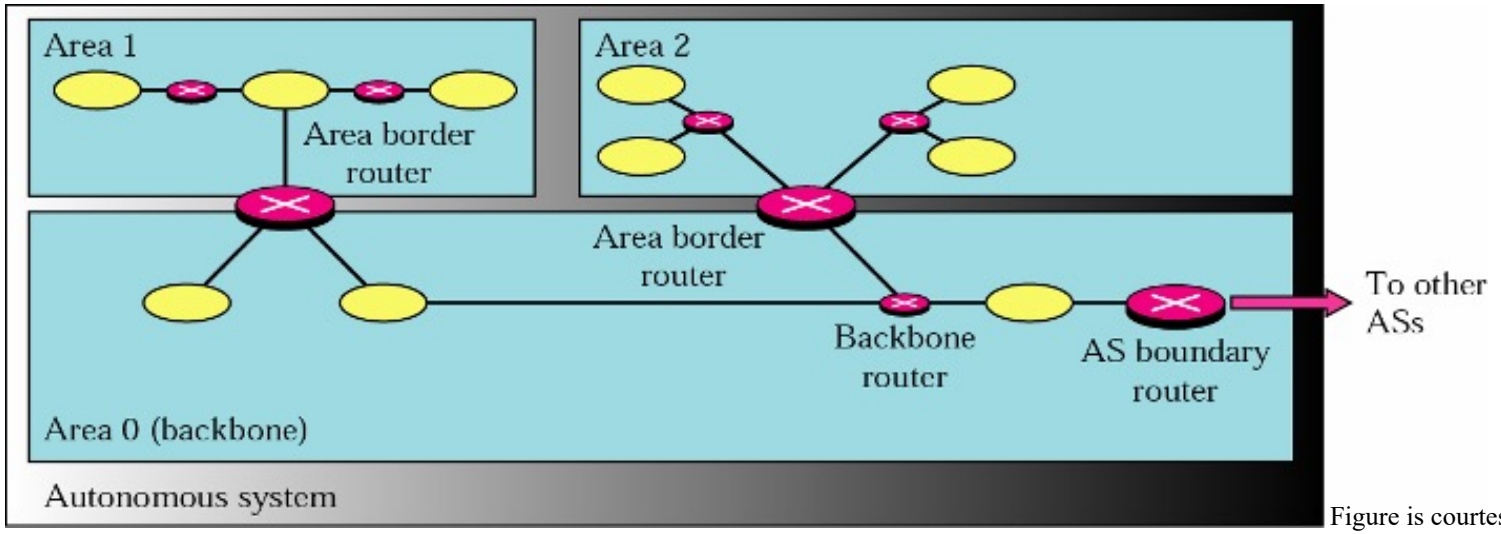
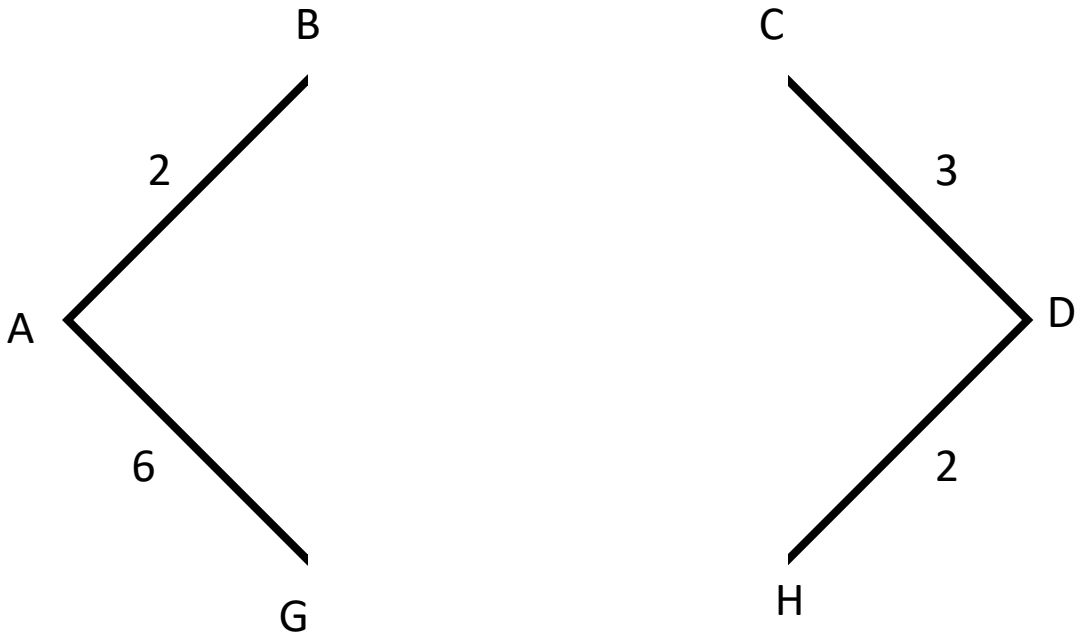


Figure is courtesy of B. Forouzan

Example: Building a network graph

A receives an update from D



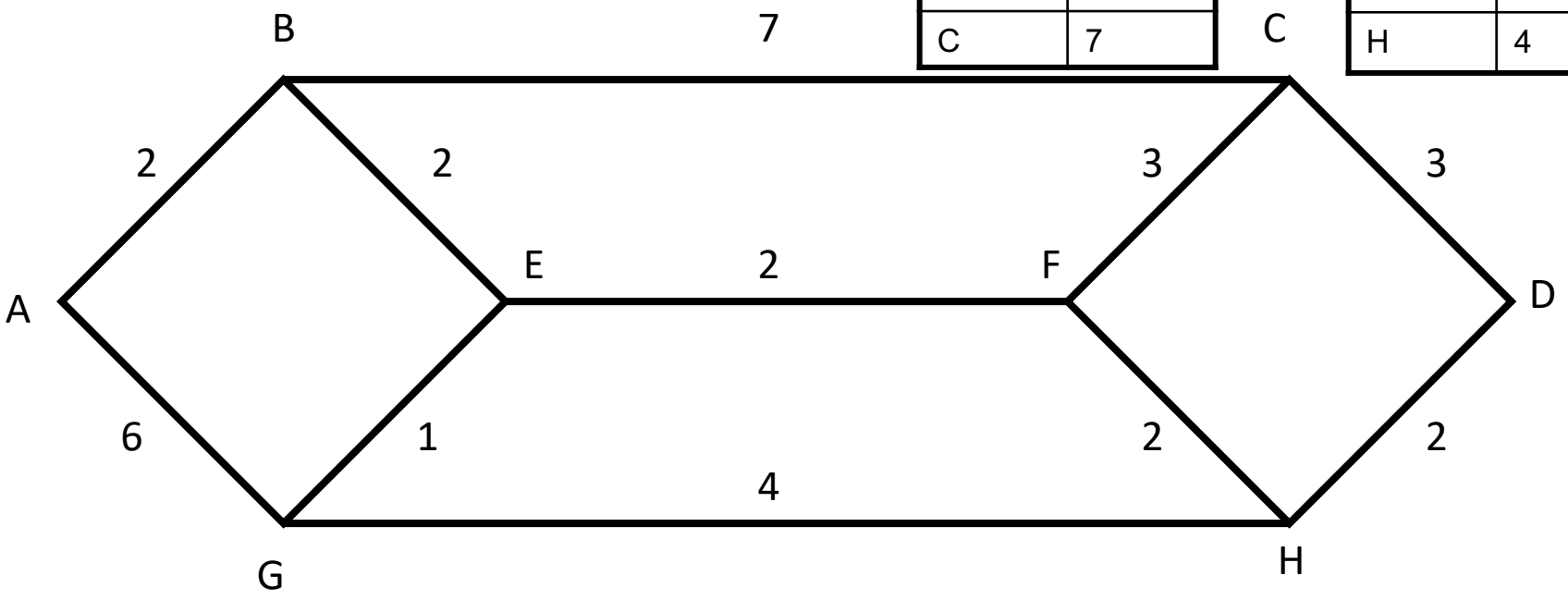
ID	D
SEQNO	0
TTL	21
C	3
H	2

Example: Building a network graph

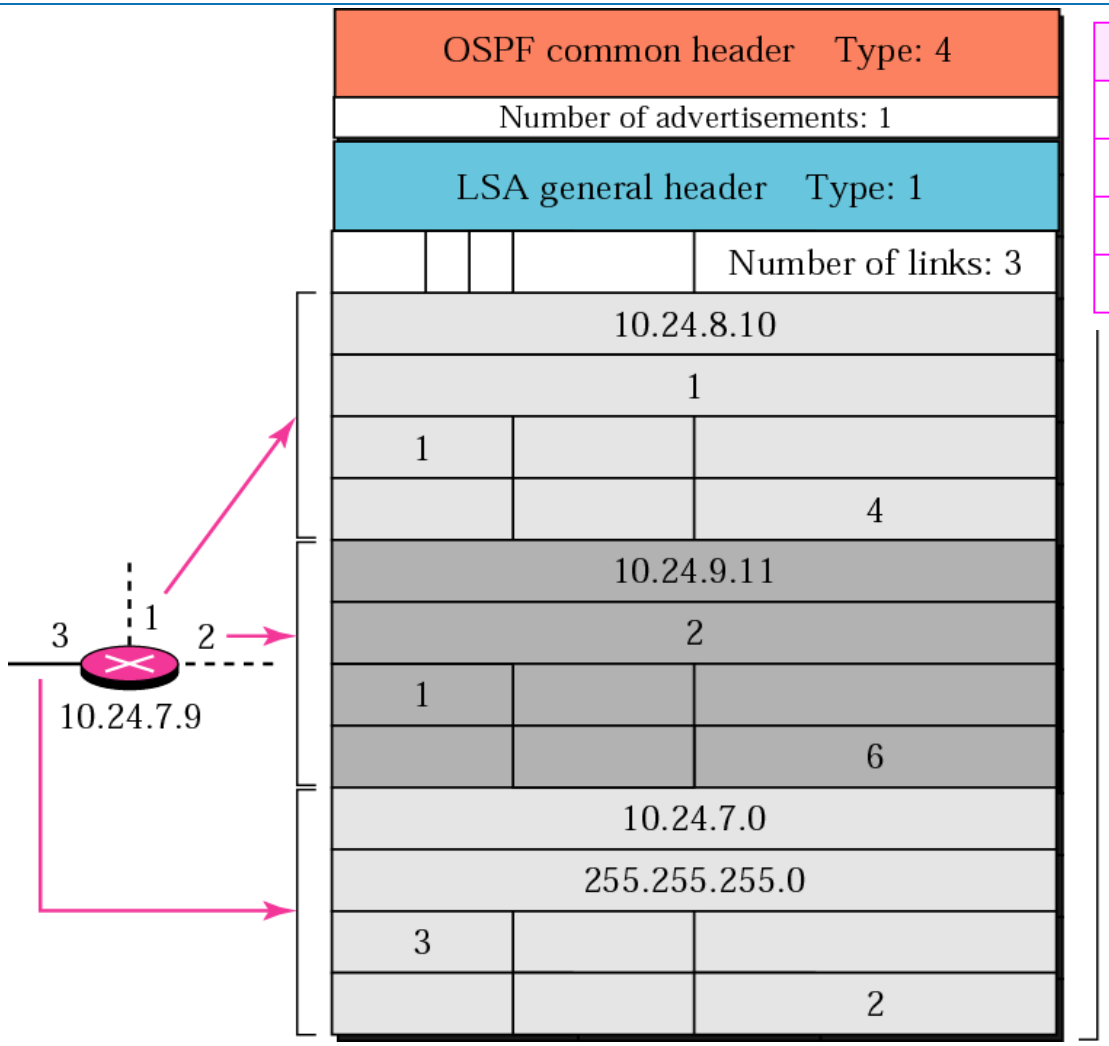
A receives updates from B and G

ID	B
SEQNO	2
TTL	23
A	2
E	2
C	7

ID	G
SEQNO	3
TTL	21
A	6
E	1
H	4



OSPF LSA Example



Link Type	*Link Identification*
Type 1: Point-to-point	Address of neighbor router
Type 2: Transient	Address of designated router
Type 3: Stub	Network address
Type 4: Virtual	Address of neighbor router

Link Type	*Link Data*
Type 1: Point-to-point	Interface number
Type 2: Transient	Router address
Type 3: Stub	Network mask
Type 4: Virtual	Router address

* Figure is courtesy of B. Forouzan

Dijkstra's Algorithm

1. Start with the local node (router): the root of the tree.
2. Assign a cost of 0 to this node and make it the first permanent node.
3. Examine each neighbour node of the node that was the last permanent node.
4. Assign a cumulative cost to each node and make it tentative.
5. Among the list of tentative nodes
    a) Find the node with the smallest cumulative cost and make it permanent.
    b) If a node can be reached from more than one direction
        I. Select the direction with the shortest cumulative cost.
6. Repeat steps 3 to 5 until every node becomes permanent.

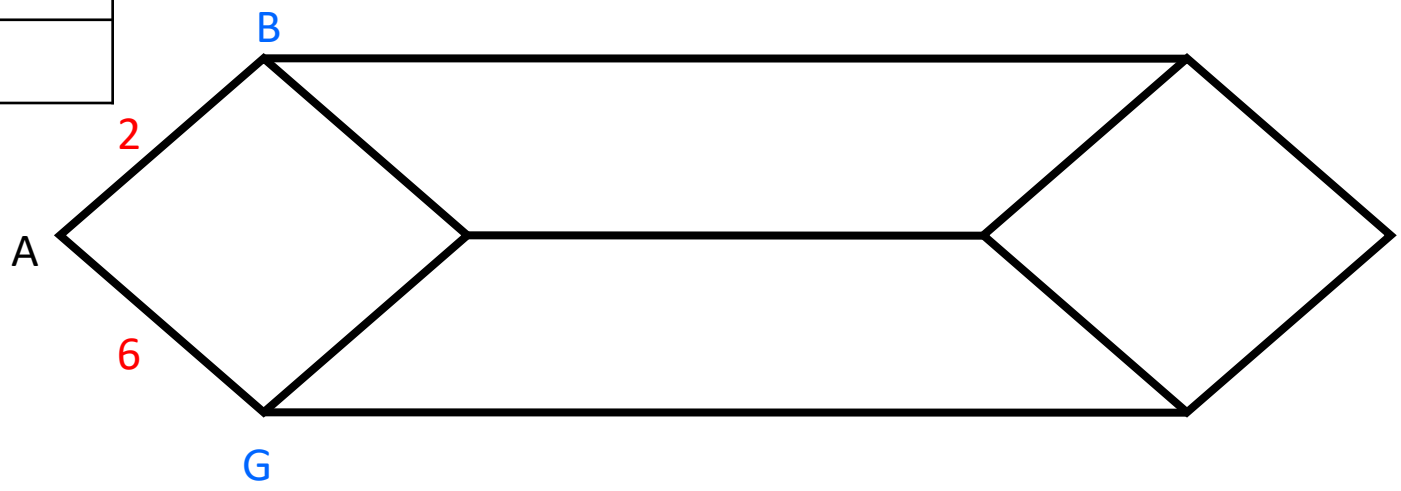
Example: Dijkstra's Algorithm

Routing Table

A	0	-

Tentative Nodes

B	2	A
G	6	A



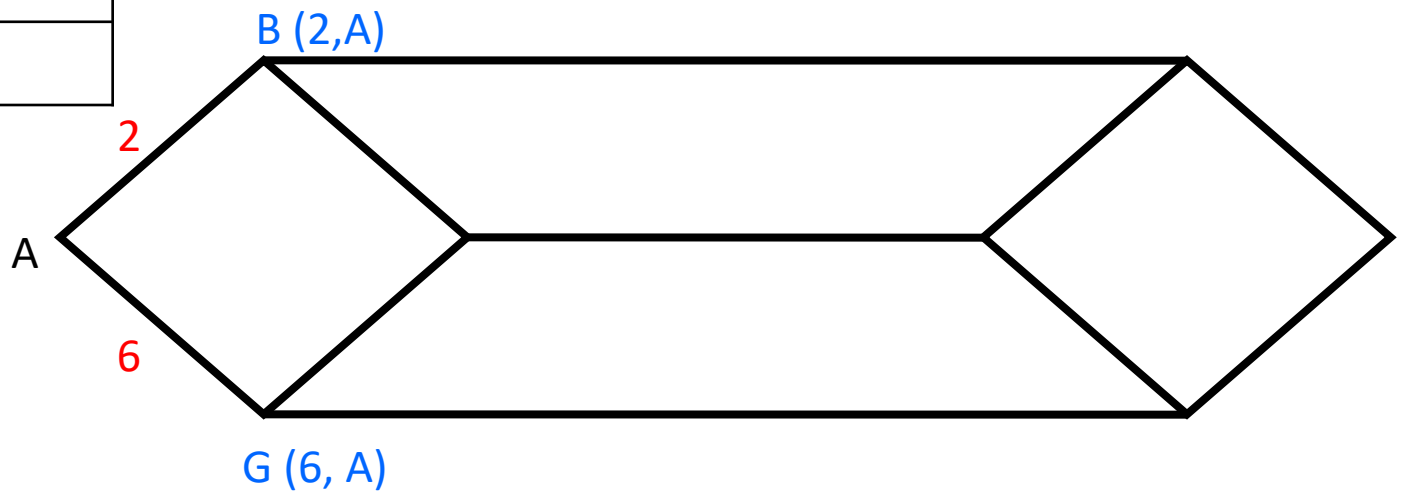
Example: Dijkstra's Algorithm

Routing Table

A	0	-
B	2	A

Tentative Nodes

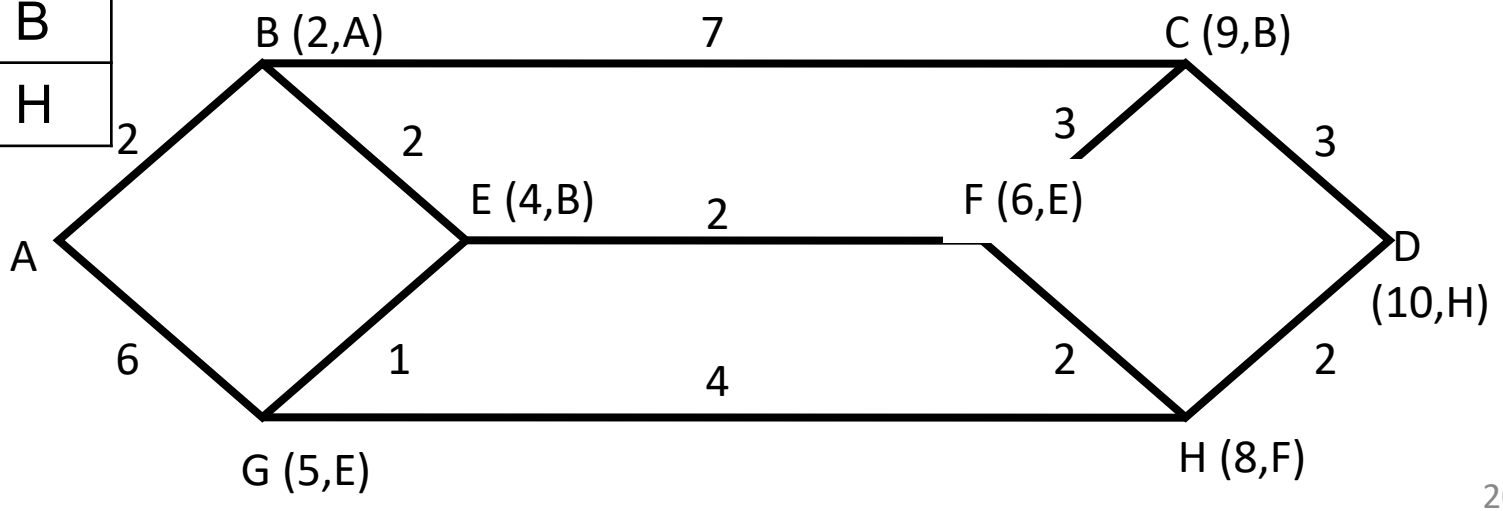
G	6	A



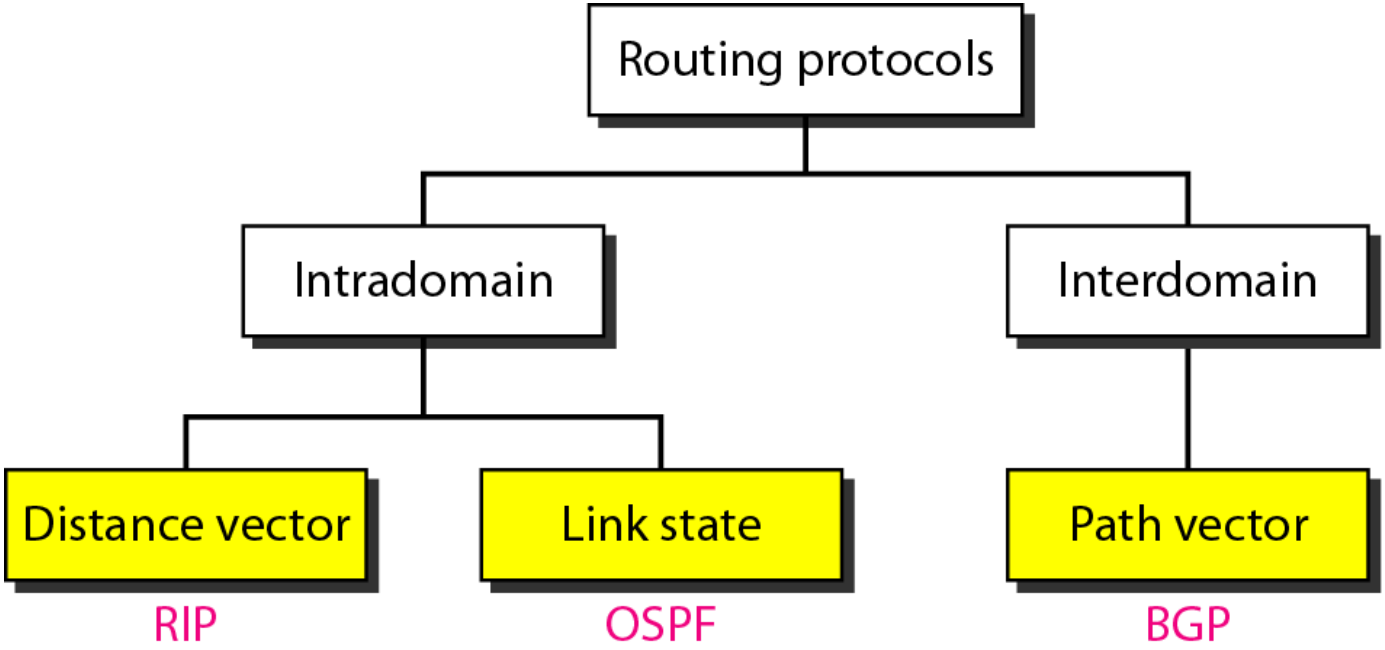
Example: Dijkstra's Algorithm

A	0	-
B	2	A
E	4	B
G	5	E
F	6	E
H	8	F
C	9	B
D	10	H

- Shortest connection from A to any other node
- Greedy algorithm! Follows minima!



Routing Protocols



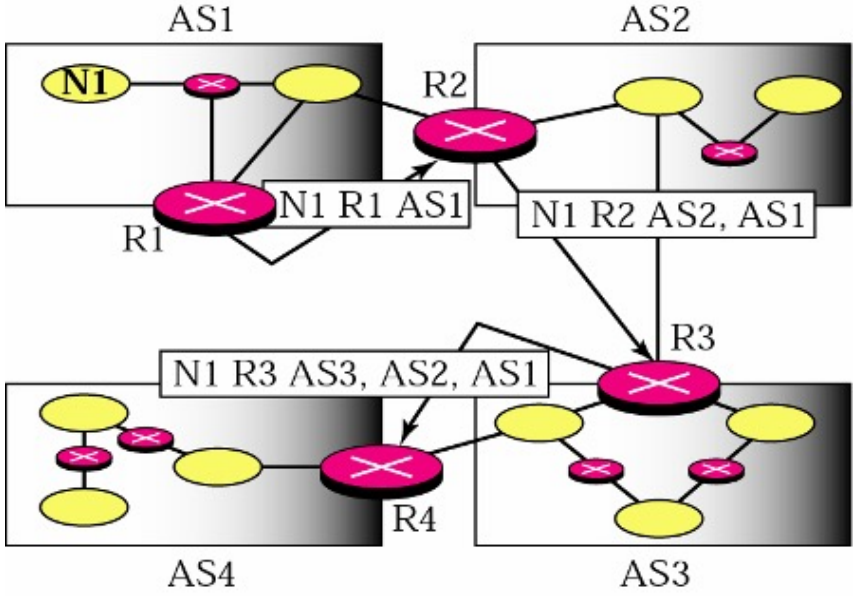
- Interior Routing Protocols
  - Routing within ASs
- Exterior Routing Protocols
  - Routing between ASs

* Figure is courtesy of B. Forouzan

Border Gateway Protocol (BGP)

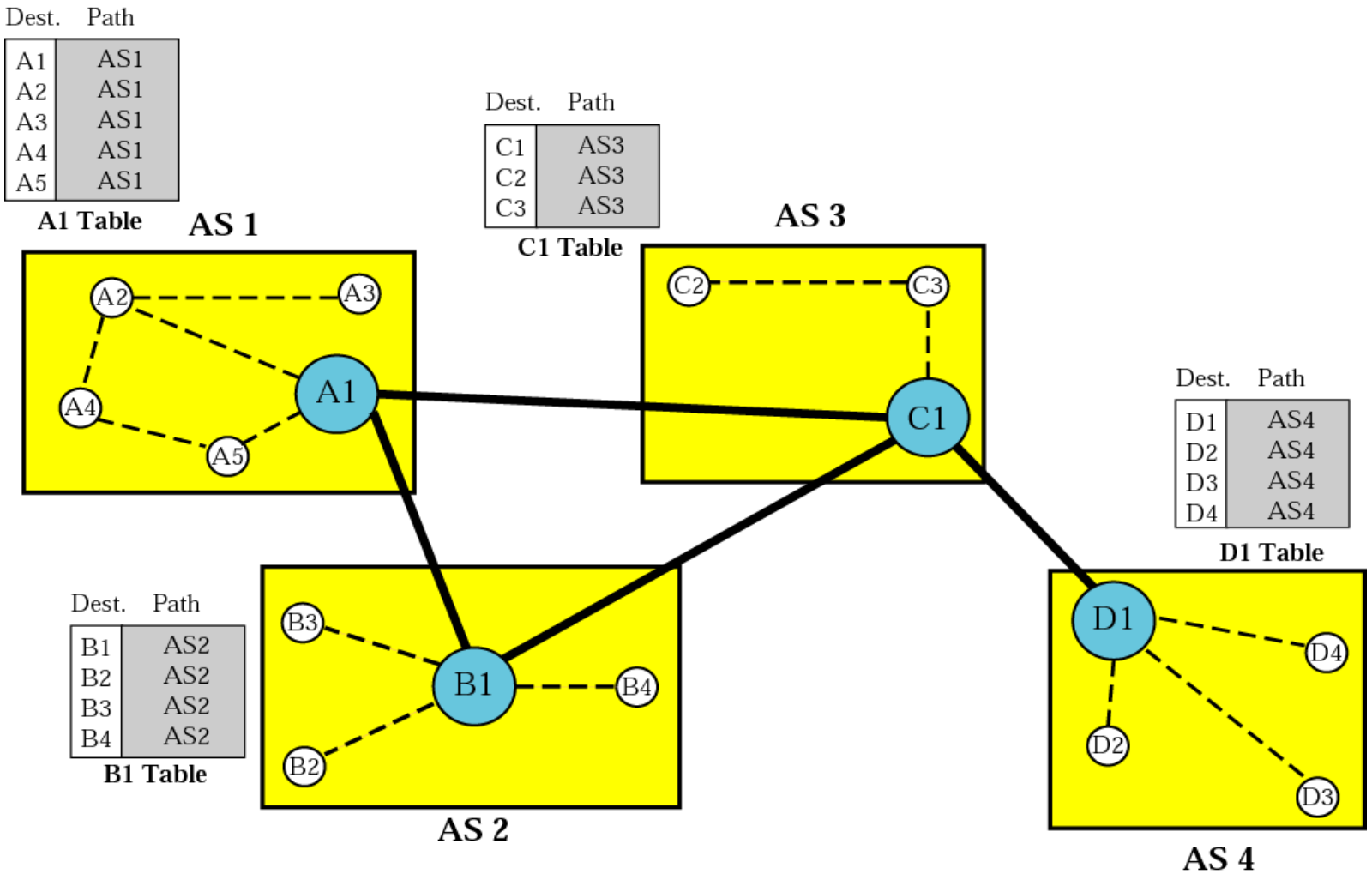
- Uses Path Vector Routing
- Advertisements include complete path to destination
- Router that forwards advertisement adds itself to the list
- Path can be checked for loops
- Policies are applied when incorporating new routes

Network	Next Router	Path
N01	R01	AS14, AS23, AS67
N02	R05	AS22, AS67, AS05, AS89
N03	R06	AS67, AS89, AS09, AS34
N04	R12	AS62, AS02, AS09



* Figure is courtesy of B. Forouzan

Tables at Autonomous Systems



* Figure is courtesy of B. Forouzan

BGP-4: Border Gateway Protocol

- AS Types
  - Stub AS: has a single connection to one other AS
    - Carries local traffic only
  - Multihomed AS: has connections to more than one AS
    - Refuses to carry transit traffic
  - Transit AS: has connections to more than one AS
    - Carries both transit and local traffic

- Each AS has:
  - One or more border routers
  - One BGP *speaker* that advertises:
    - Local networks
    - Other reachable networks (transit AS only)
    - Gives *path* information

* Taken from Vivek Pai, Princeton - L Peterson & B Davie

Summary: Routing

- Autonomous Systems
  - Stub network
  - Transient network
  - Point-to-point link

- Distance Vector routing
  - Share complete information with neighbours
  - Count-to-Infinity problem
  - Example: Routing Information Protocol (RIP)

- Link State routing
  - Share information about neighbours with everyone
  - Dijkstra's Shortest-Path Algorithm
  - Example: Open Shortest Path First (OSPF)

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

CSU33031 Computer Networks

OpenFlow

Stefan Weber

email: sweber@tcd.ie

Office: Lloyd 1.41

Original OpenFlow Paper

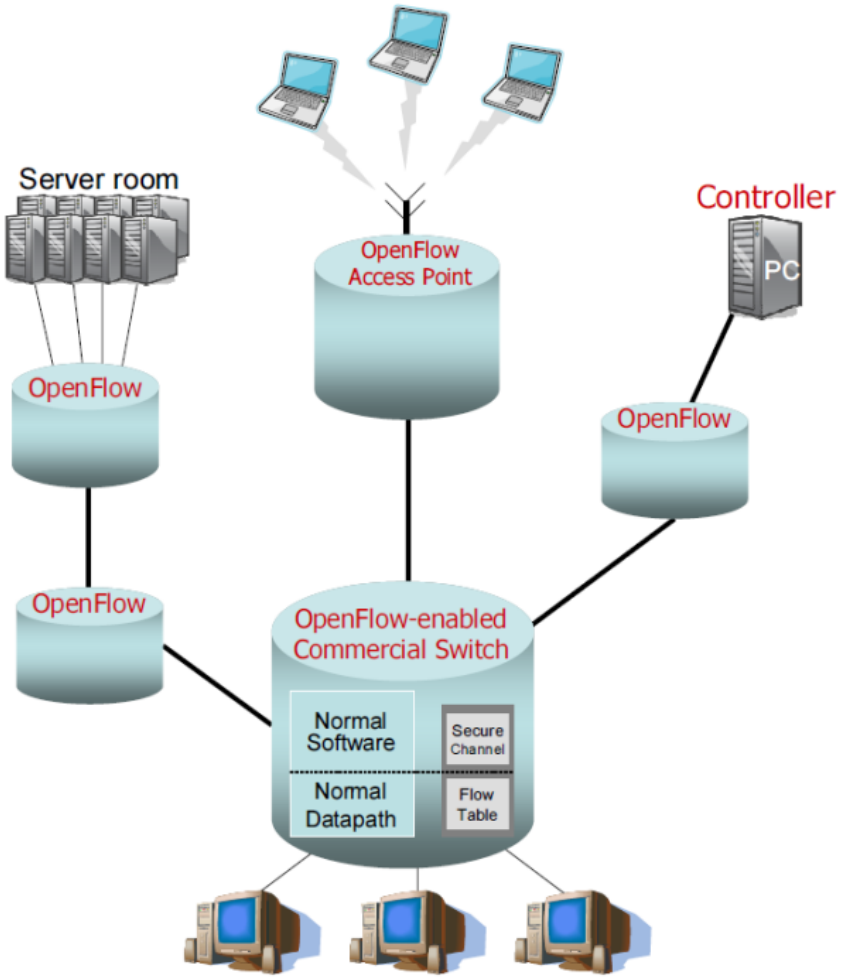
OpenFlow: Enabling Innovation in Campus Networks

March 14, 2008

Nick McKeown, Stanford University
Tom Anderson, University of Washington
Hari Balakrishnan, MIT
Guru Parulkar, Stanford University
Larry Peterson, Princeton University
Jennifer Rexford, Princeton University
Scott Shenker, University of California, Berkeley
Jonathan Turner, Washington University in St. Louis

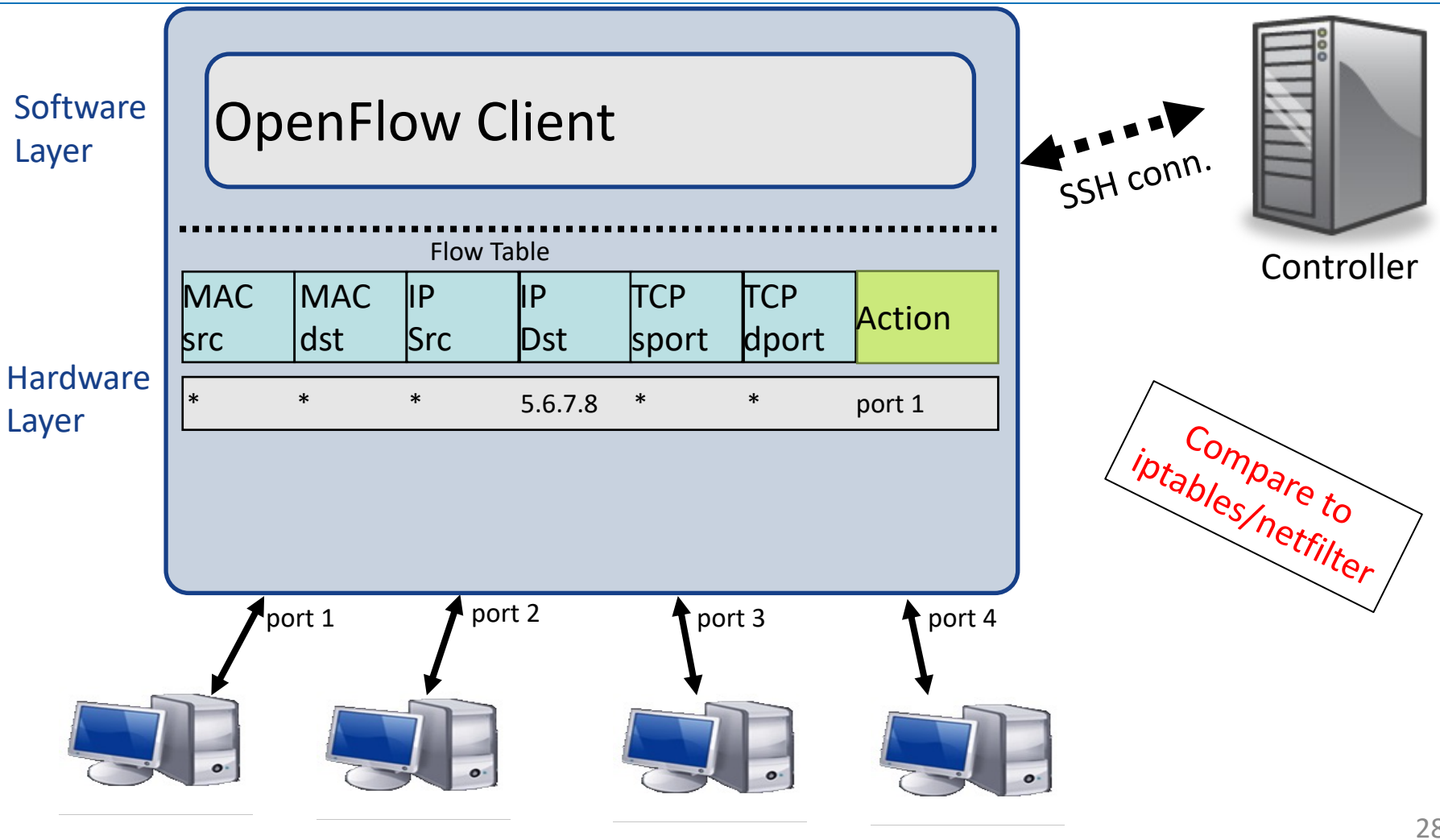
Header of 2008 Paper

- OpenFlow switches
- Controller w/ secure connection
- Configurable flow tables



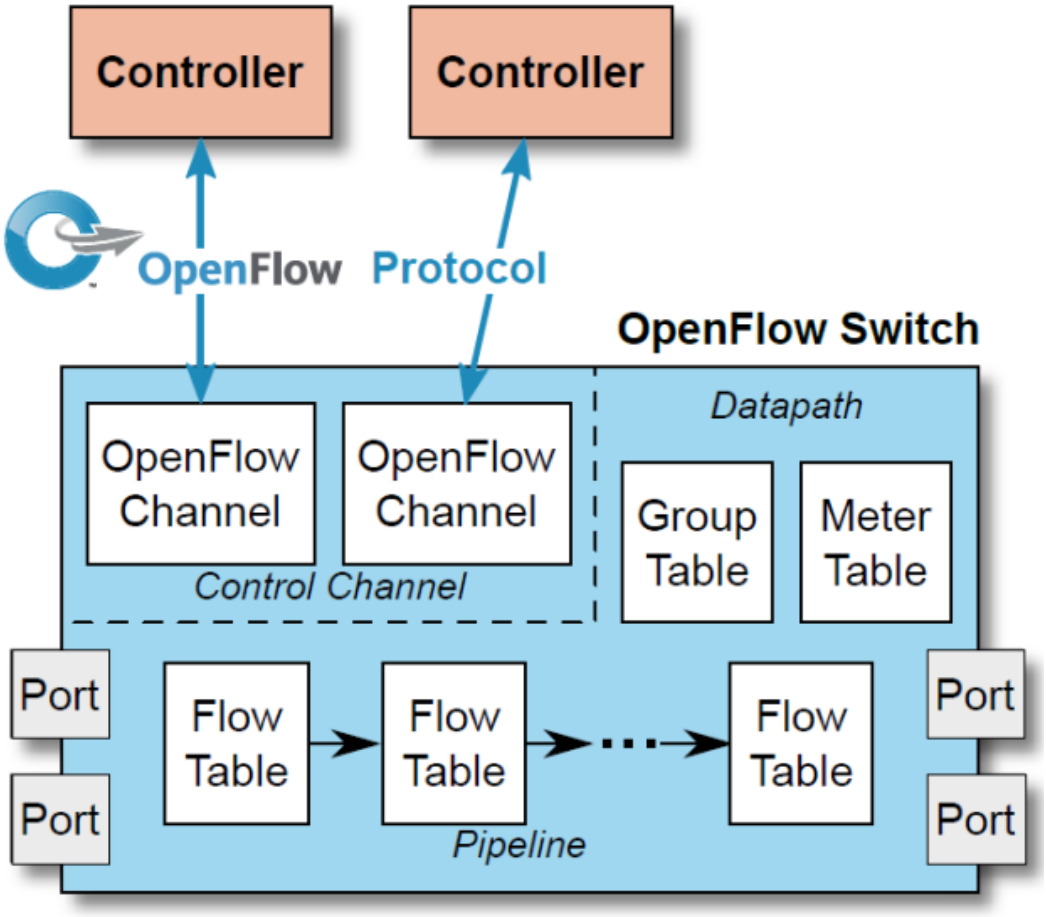
•Nick McKeown et al, SIGCOMM 2008

OpenFlow Switch



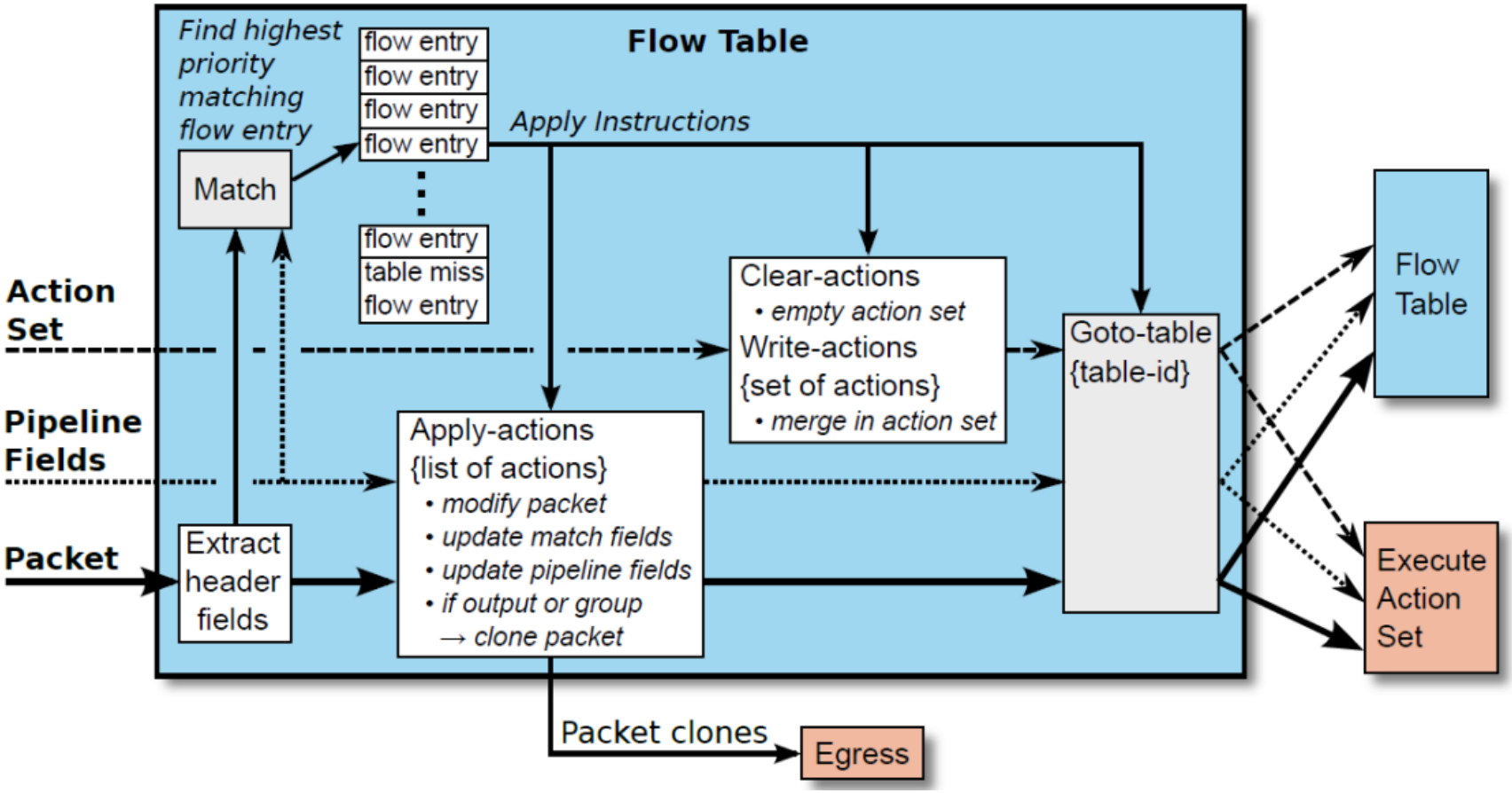
•Yiannis Yiakoumis, et al, HOT-Interconnects 2010

Components of OpenFlow Switches



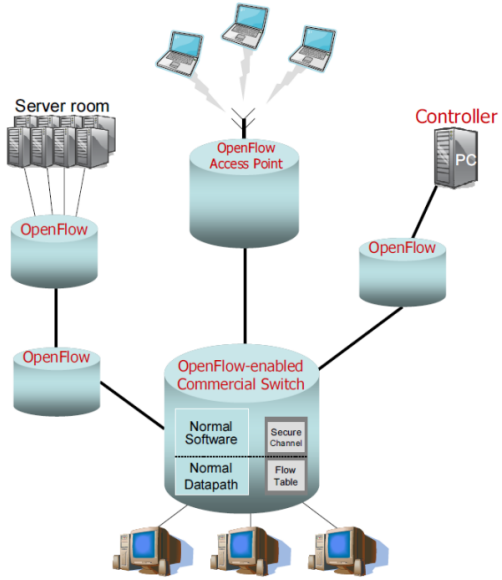
OpenFlow Switch Specification,v1.5.1, Open Net. Found.

Flow Table Processing



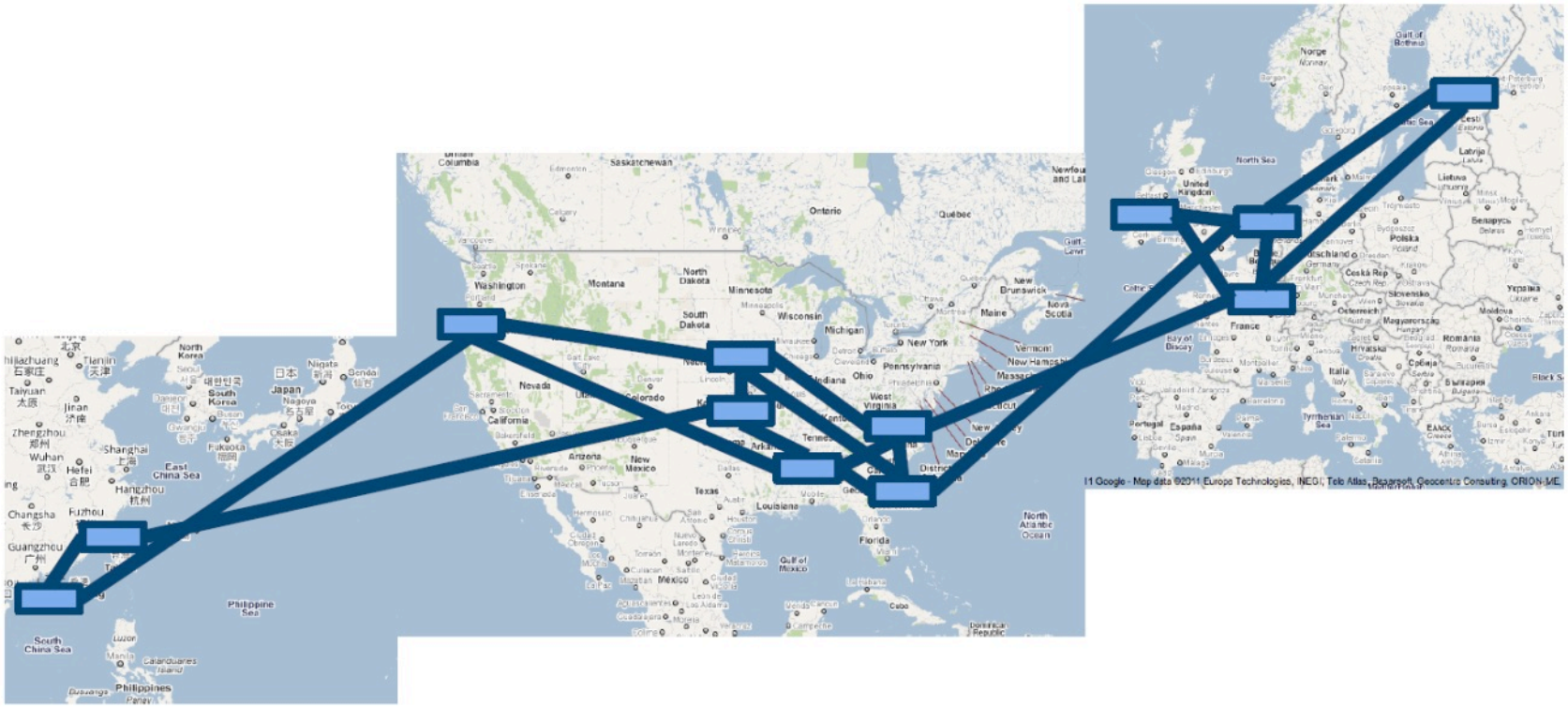
OpenFlow Switch Specification,v1.5.1, Open Net. Found.

Why????



As in: Why OpenFlow? ☺

Google's WAN

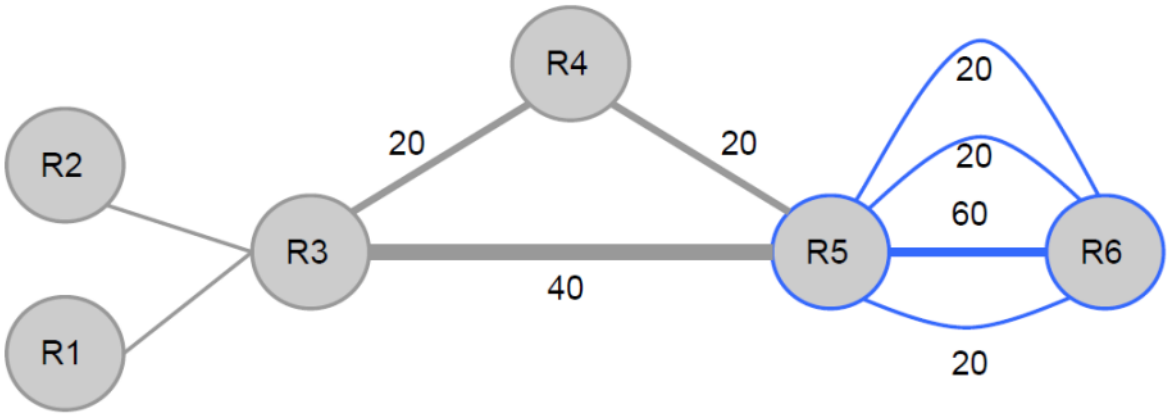


32

•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

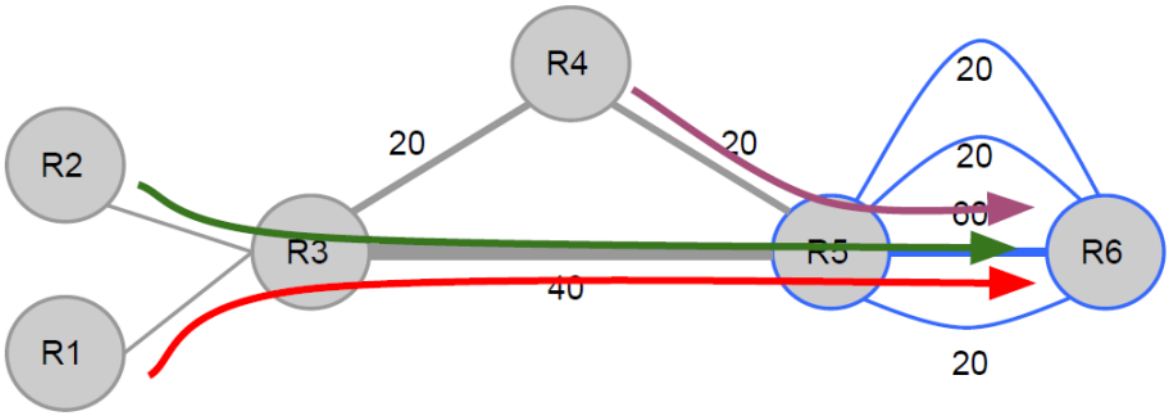
- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20



•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

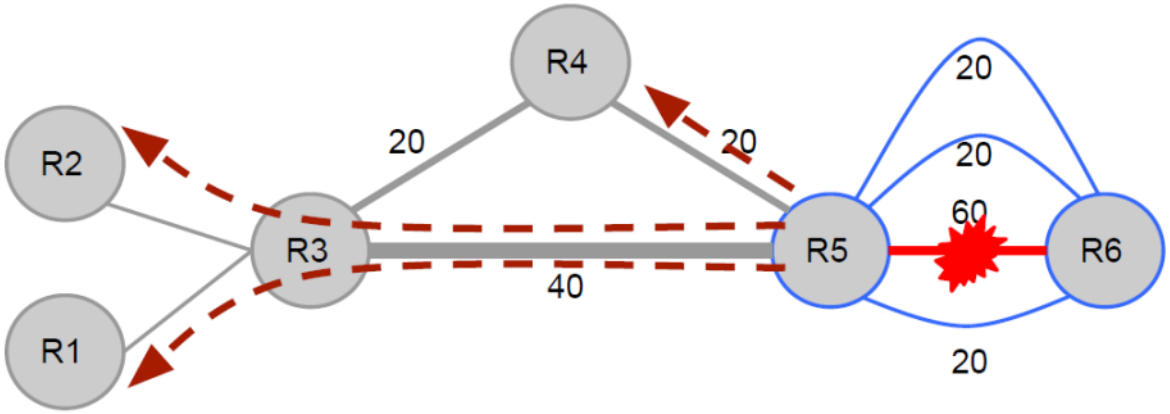
- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20



•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

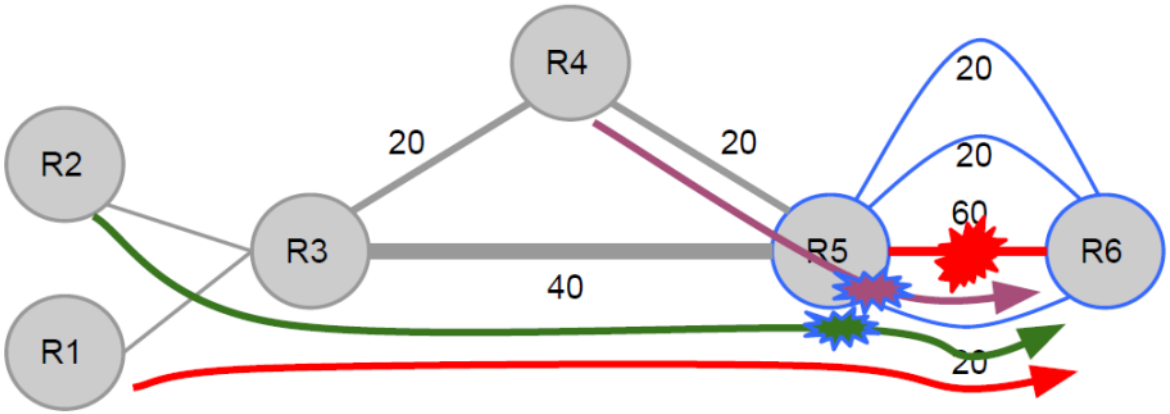


- R5-R6 link fails
  - R1, R2, R4 *autonomously* try for next best path

•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

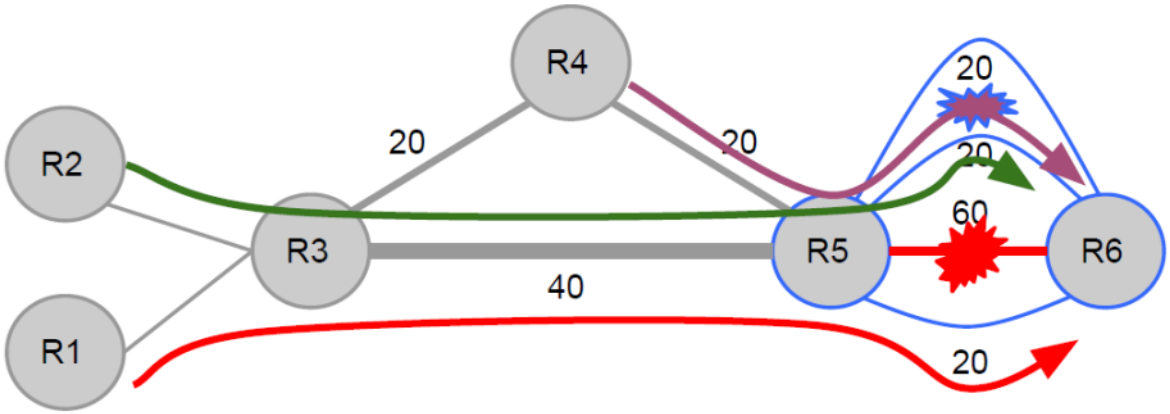


- R5-R6 link fails
  - R1, R2, R4 *autonomously* try for next best path
  - R1 wins, R2, R4 retry for next best path

•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

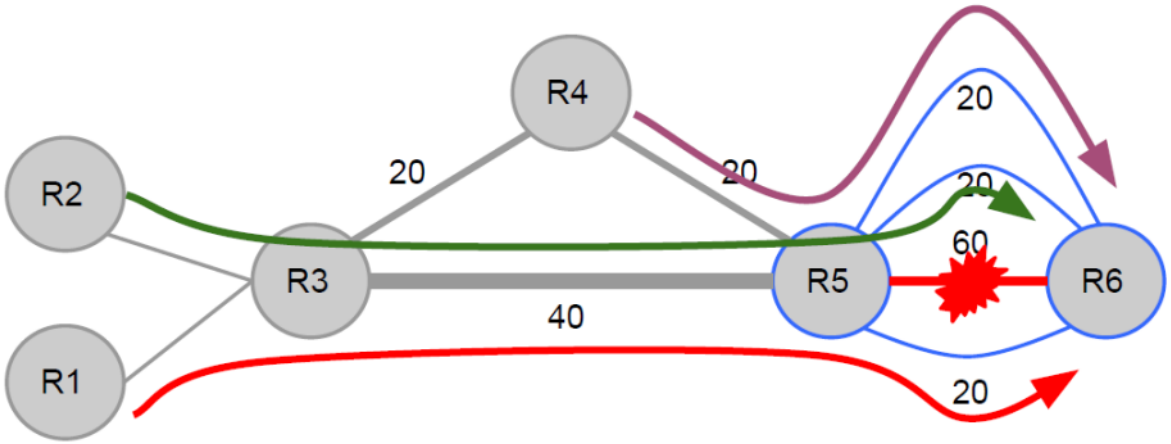


- R5-R6 link fails
  - R1, R2, R4 *autonomously* try for next best path
  - R1 wins, R2, R4 retry for next best path
  - R2 wins this round, R4 retries again

•Urs Hoelzle, Open Network Summit 2012

Traditional Net. Example

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

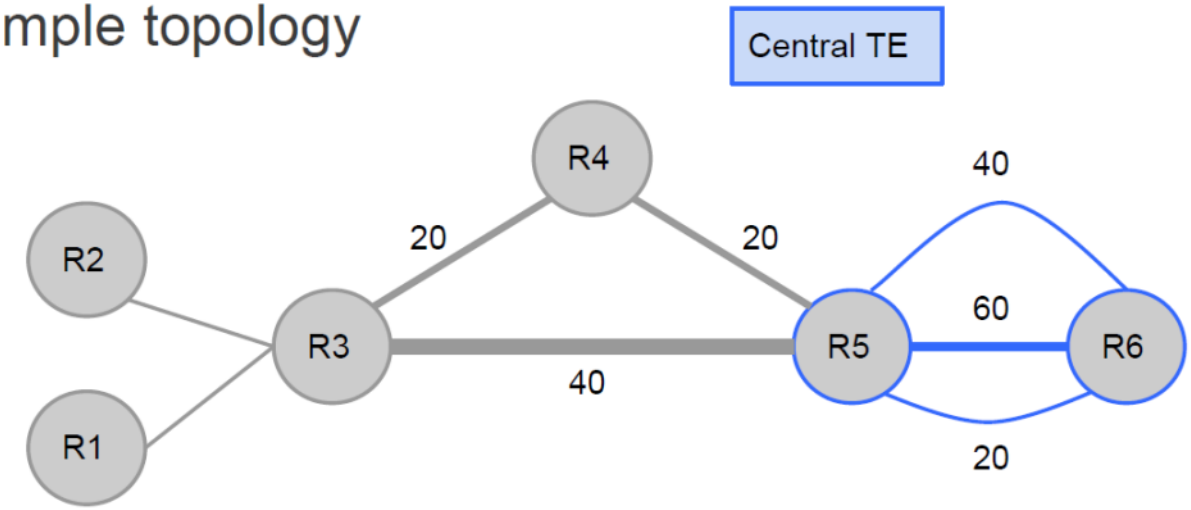


- R5-R6 link fails
  - R1, R2, R4 *autonomously* try for next best path
  - R1 wins, R2, R4 retry for next best path
  - R2 wins this round, R4 retries again
  - R4 finally gets third best path

•Urs Hoelzle, Open Network Summit 2012

Topology with Central TE

- Simple topology



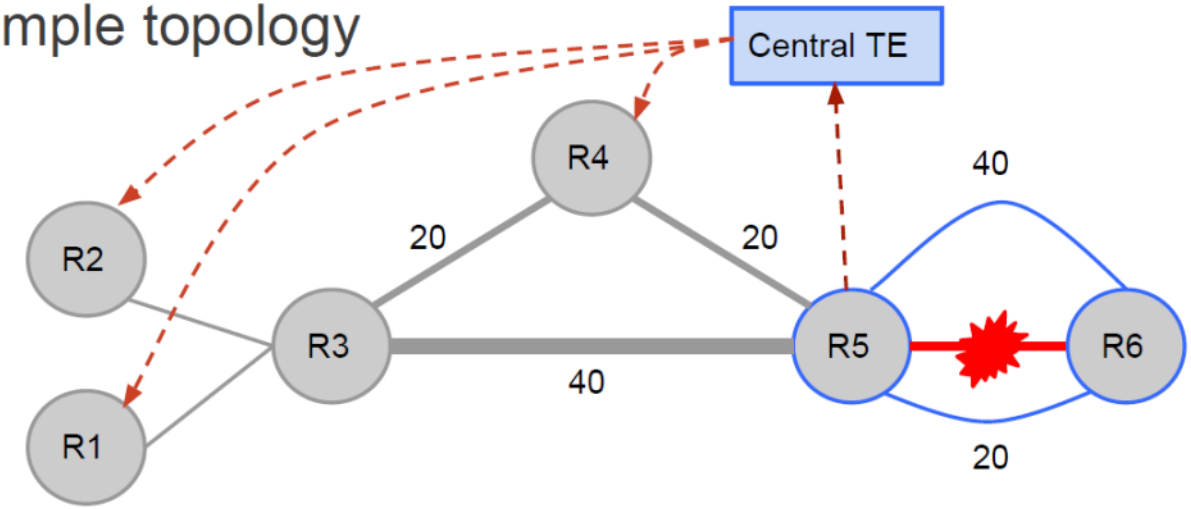
- Flows:
  - R1->R6: 20; R2->R6: 20; R4->R6: 20

TE = Traffic Engineering

•Urs Hoelzle, Open Network Summit 2012

Topology with Central TE

- Simple topology

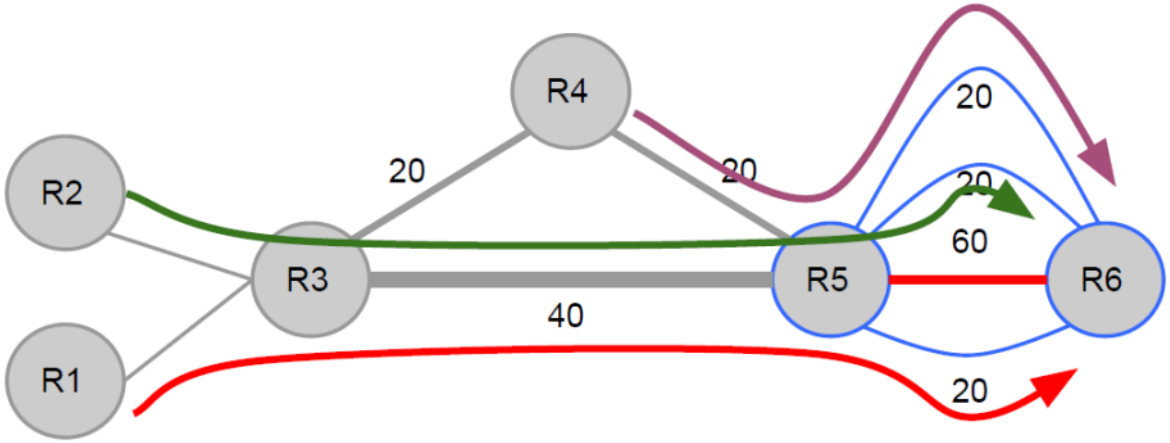


- Flows:
  - R1->R6: 20; R2->R6: 20; R4->R6: 20
- R5-R6 fails
  - R5 informs TE, which programs routers in one shot

•Urs Hoelzle, Open Network Summit 2012

Topology with Central TE

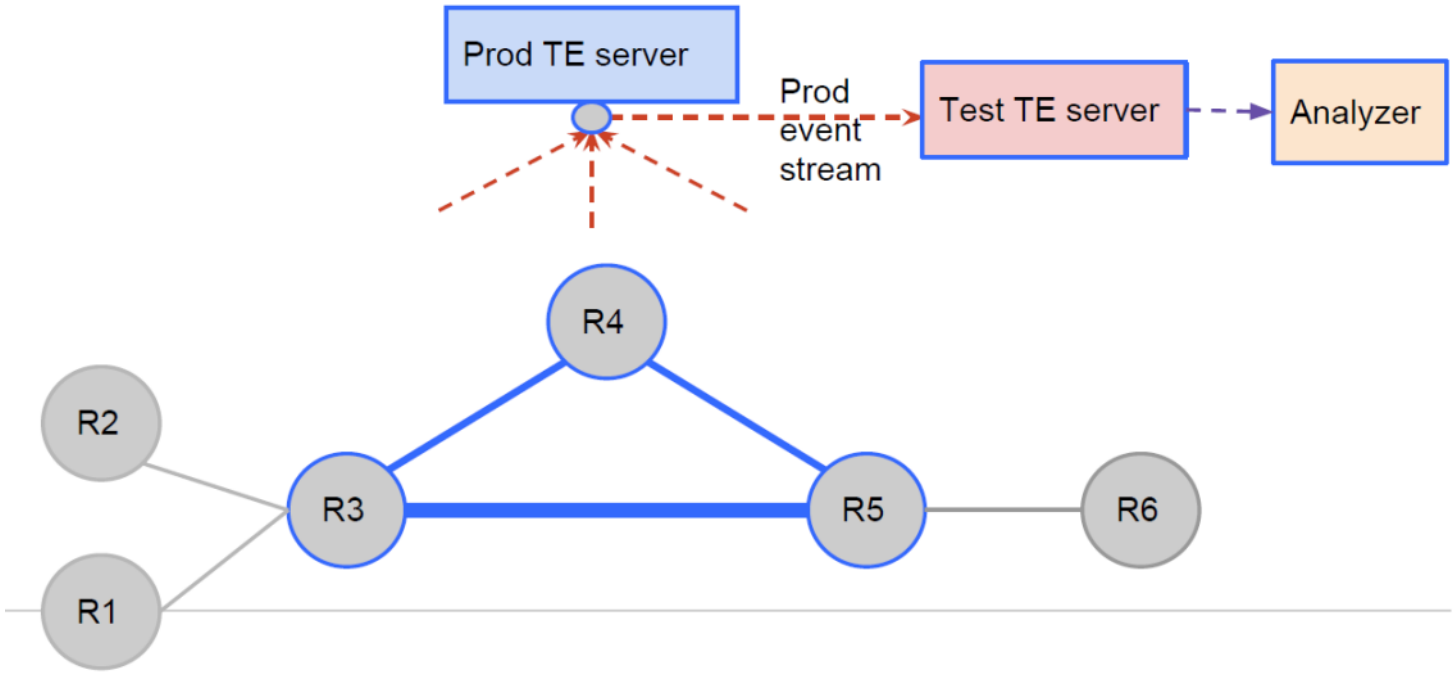
- Simple topology



- Flows:
  - R1->R6: 20; R2->R6: 20; R4->R6: 20
- R5-R6 link fails
  - R5 informs TE, which programs routers in one shot
  - Leads to faster realization of target optimum

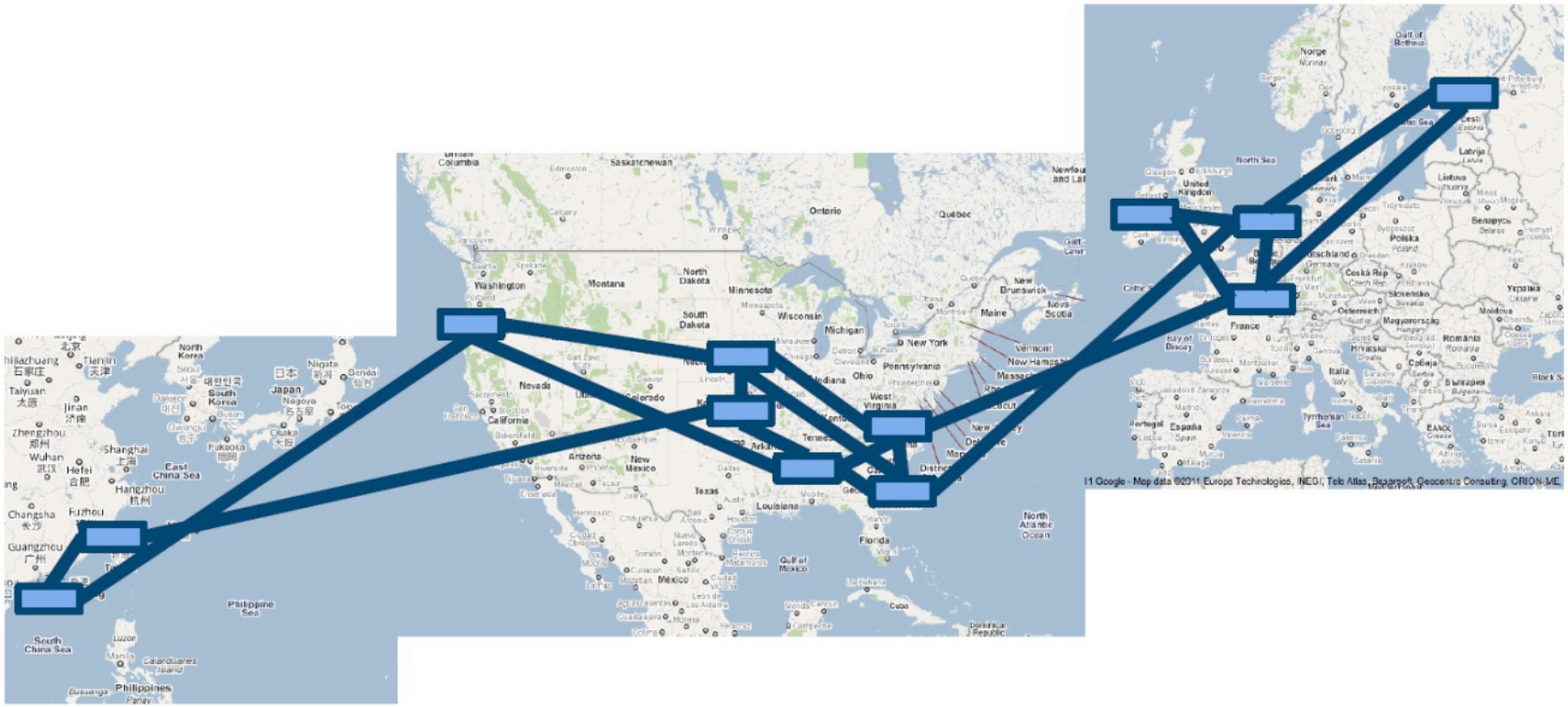
•Urs Hoelzle, Open Network Summit 2012

Traffic Engineering with Analyzer



•Urs Hoelzle, Open Network Summit 2012

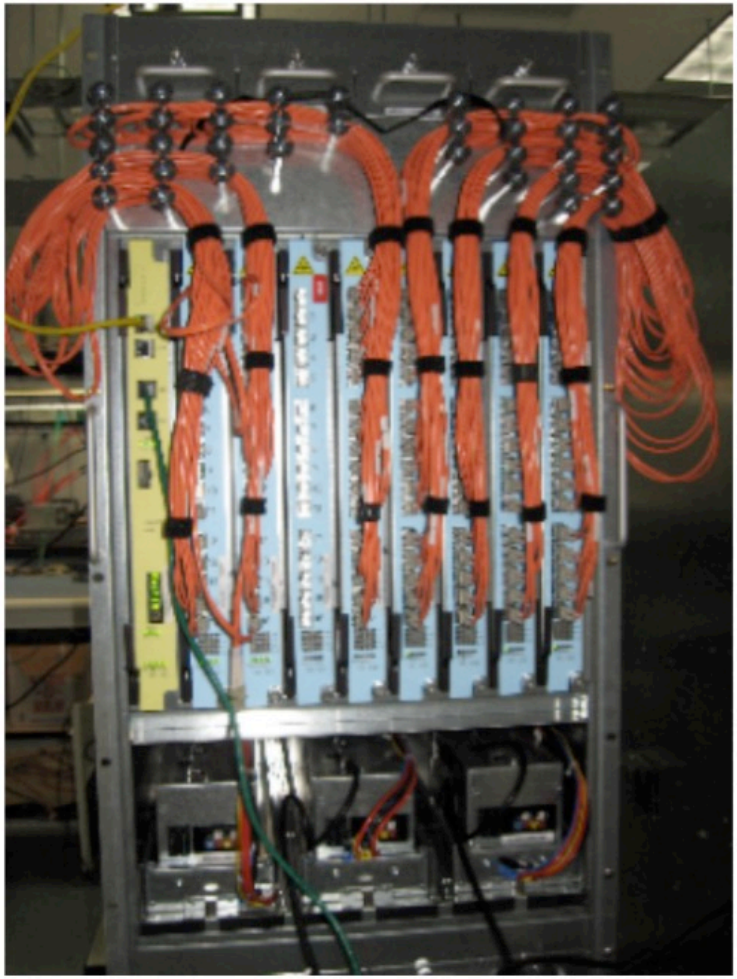
Google's WAN



•Urs Hoelzle, Open Network Summit 2012

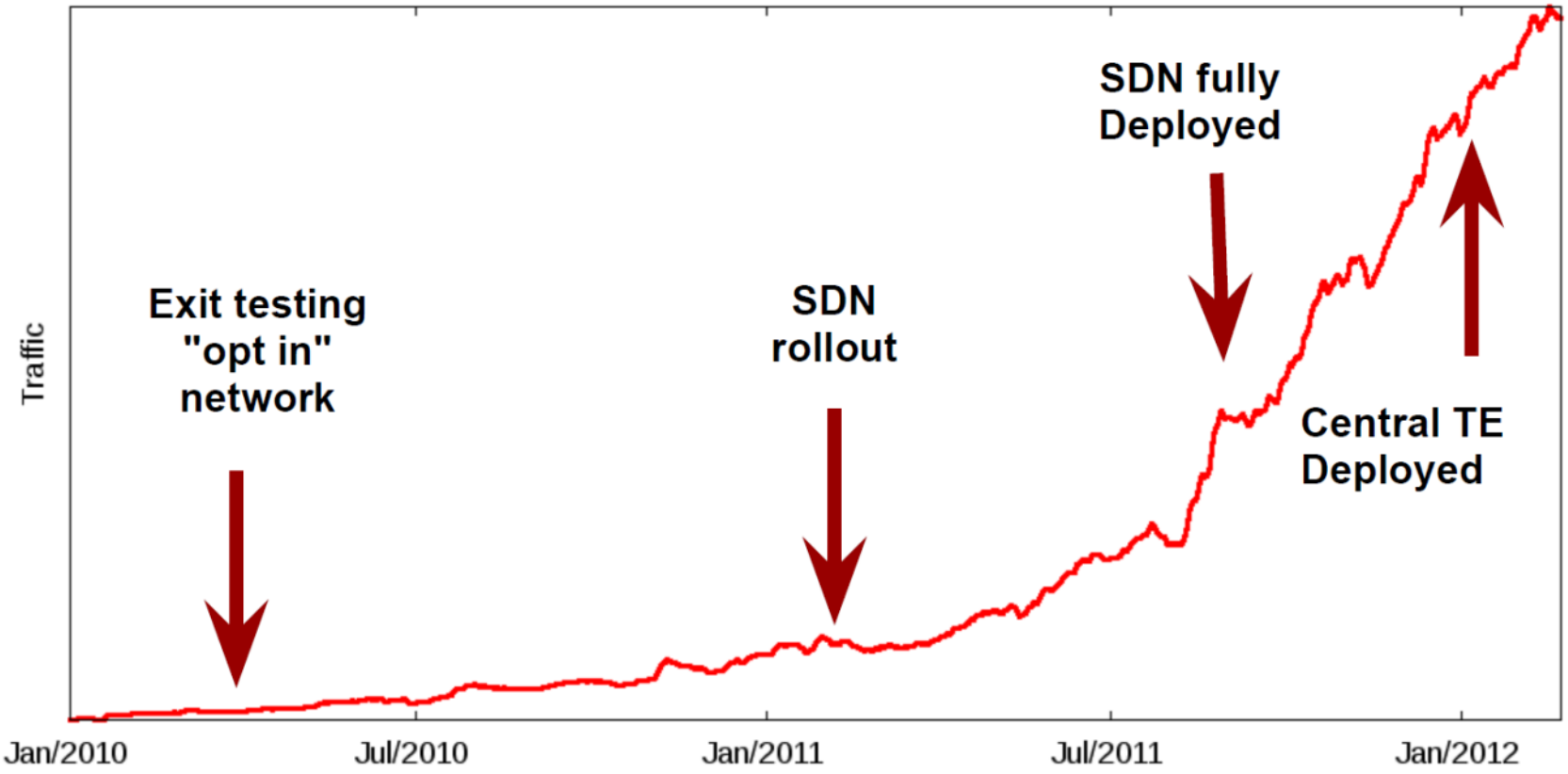
Google's Hardware

- Built from merchant silicon
  - 100s of ports of nonblocking 10GE
- OpenFlow support
- Open source routing stacks for BGP, ISIS
- Does not have all features
  - No support for AppleTalk...
- Multiple chassis per site
  - Fault tolerance
  - Scale to multiple Tbps



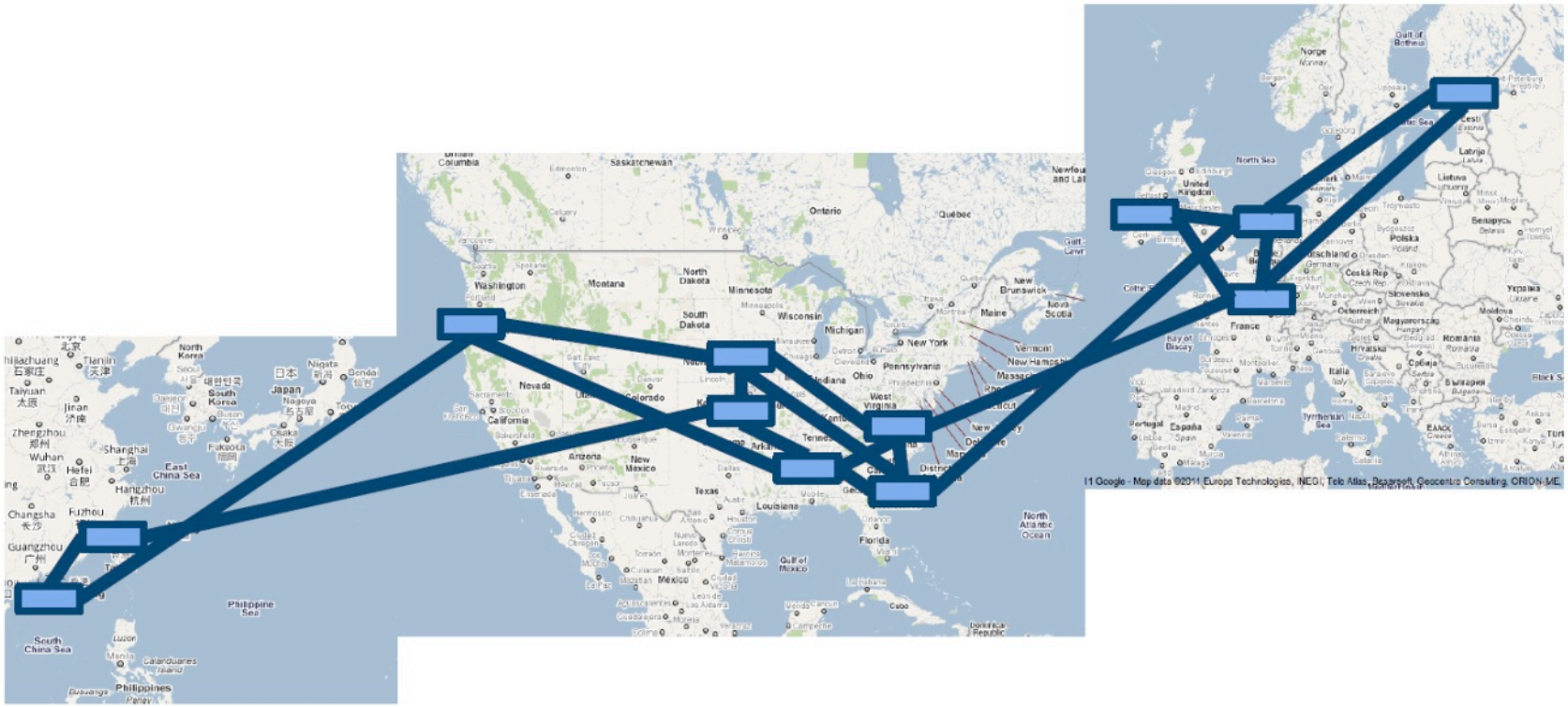
•Urs Hoelzle, Open Network Summit 2012

Timeline of Deployment



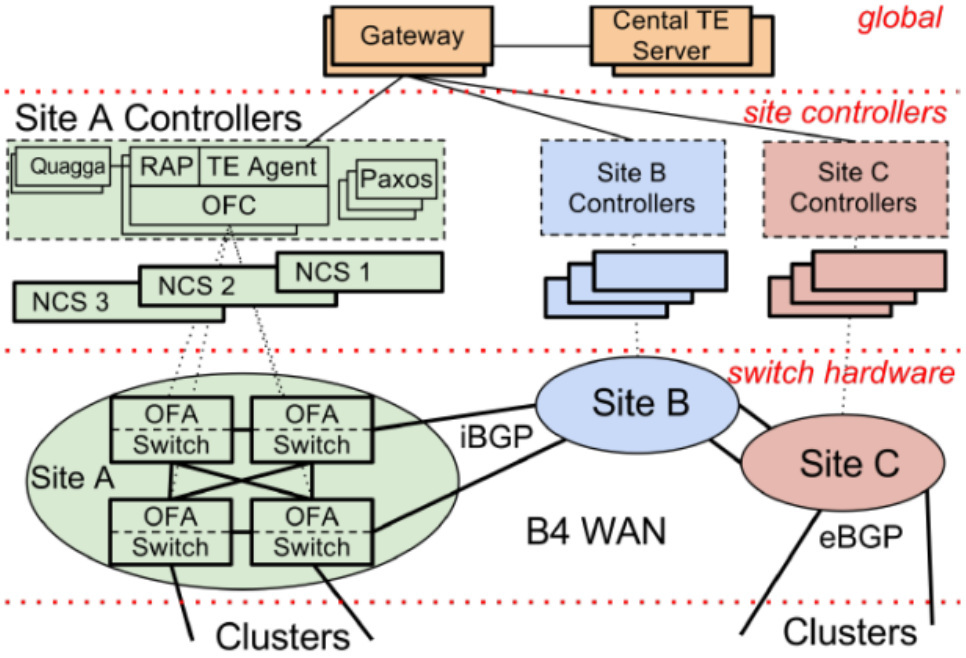
•Urs Hoelzle, Open Network Summit 2012

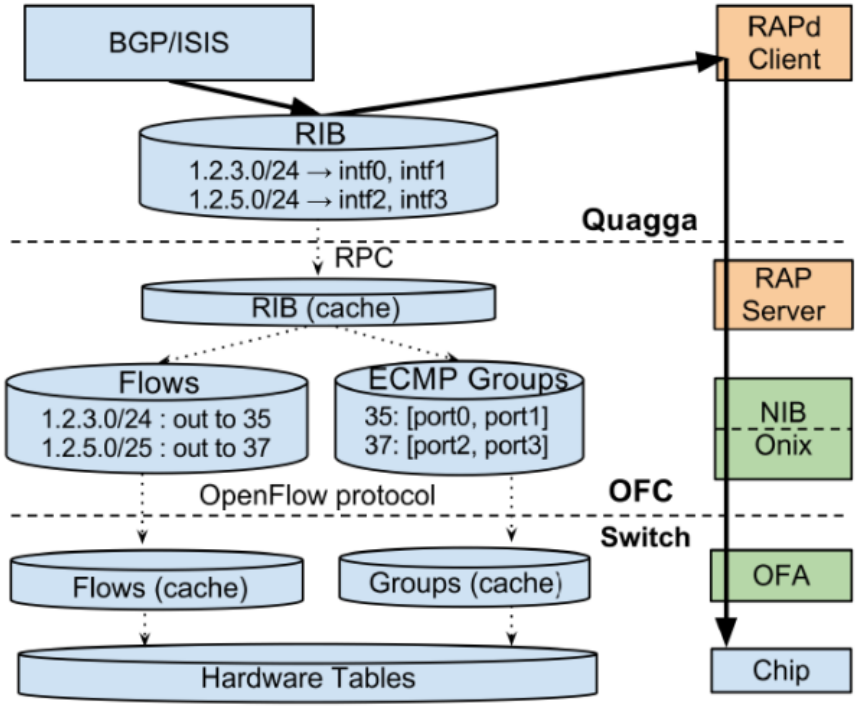
Google's WAN



•Urs Hoelzle, Open Network Summit 2012

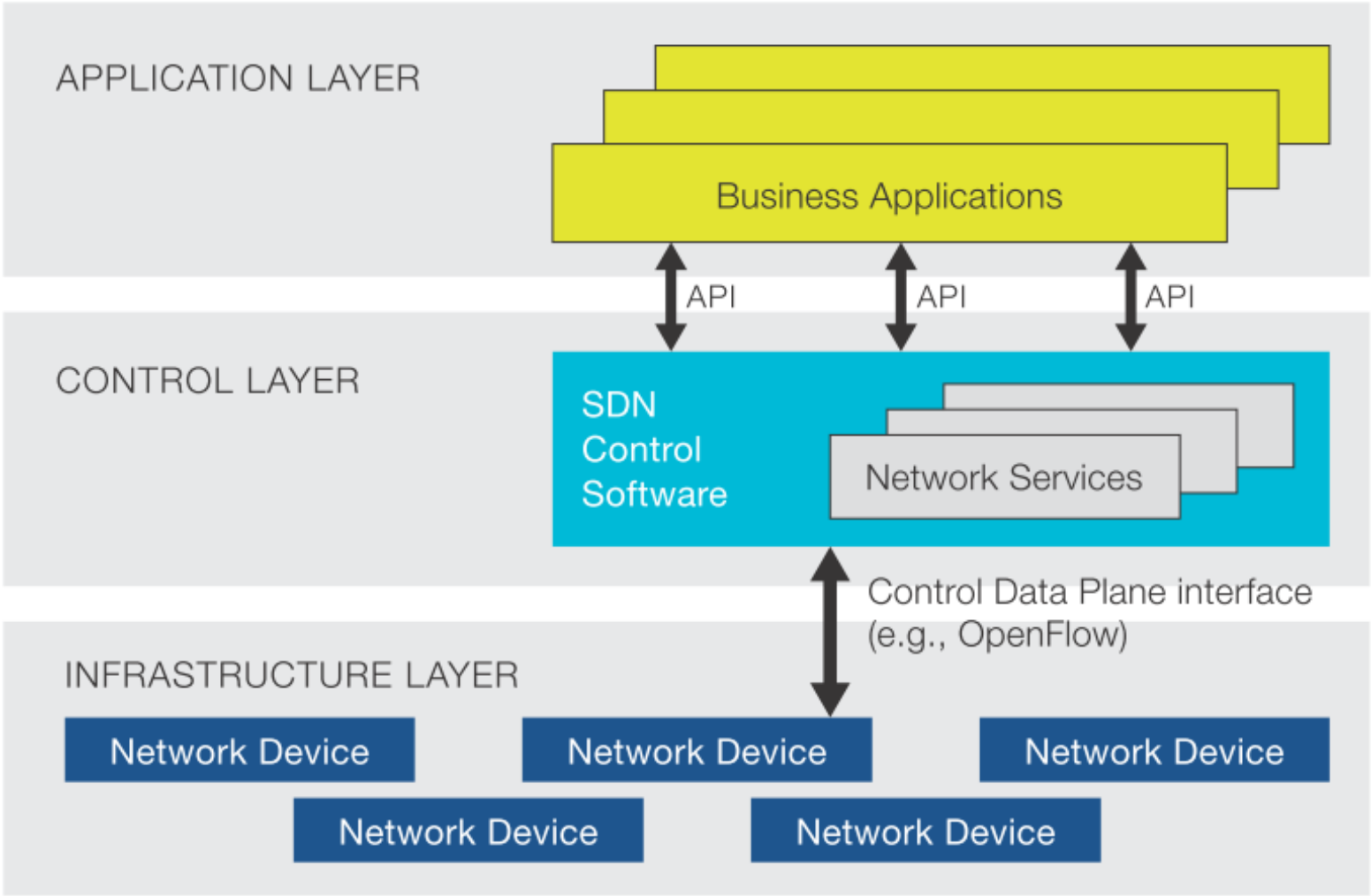
OpenFlow in B4





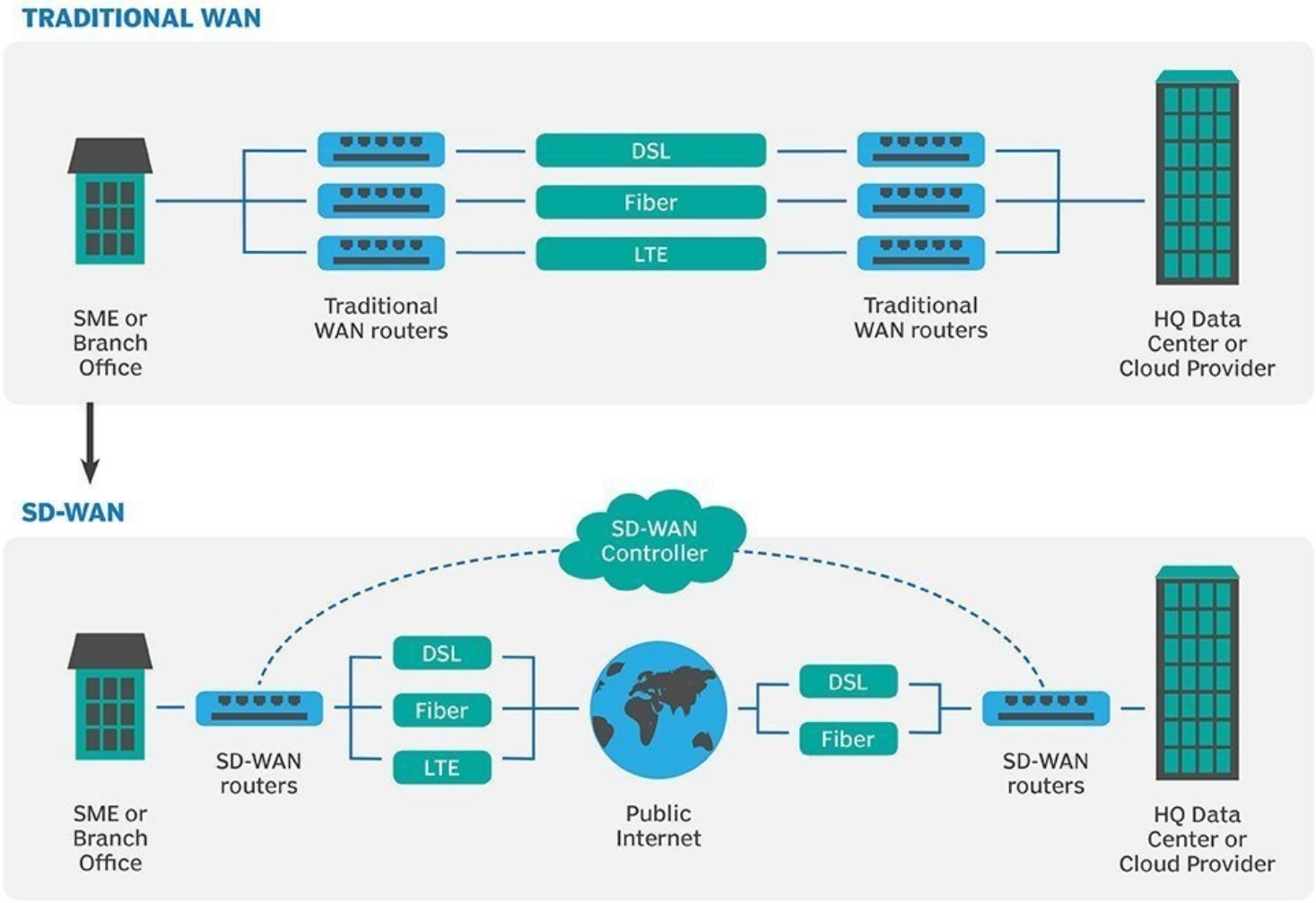
B4: Experience with a Globally Deployed Software Defined WAN, Sushant Jain, et al, 2013

SDN Architecture incl. OpenFlow



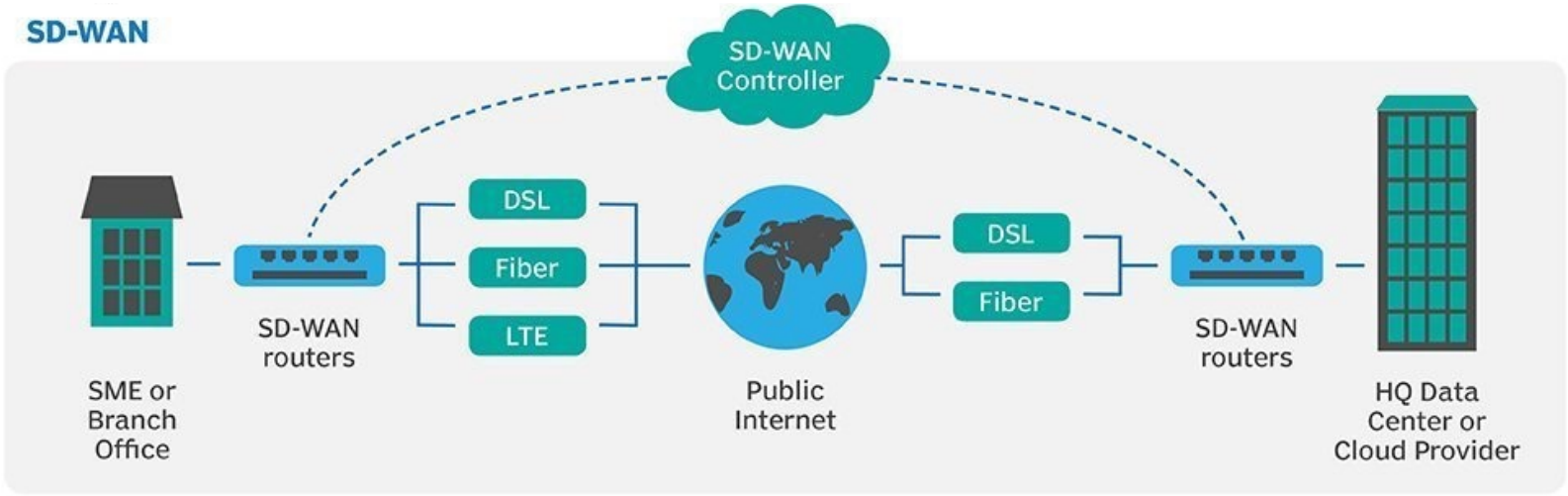
OpenFlow Whitepaper, Open Net. Found.

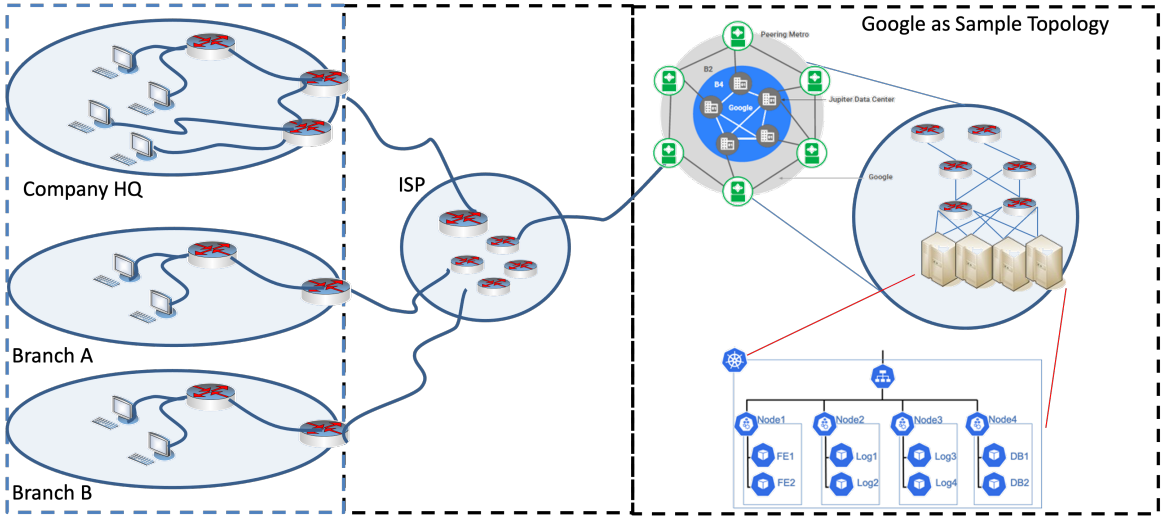
Traditional WAN vs SD-WAN



https://searchsdn.techtarget.com/definition/SD-WAN-software-defined-WAN

Traditional WAN vs SD-WAN





https://searchsdn.techtarget.com/definition/SD-WAN-software-defined-WAN

Example from a Local Provider

[redacted] BUSINESS

Order Online | Why Virgin | Search

Small (1-20 people) | Medium Business (20-250 people) | Enterprise Business (250+ people) | Public Sector | Wholesale | Help & Advice

SD-WAN

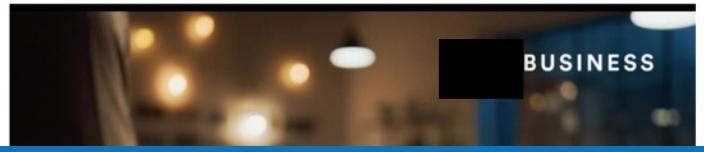
Embrace a cloud-based network and pave the way for digital transformation

If you know what you are looking for and want to get in touch, call us on 0808 231 6758 or fill in our web form and we will get back to you as soon as possible.

How SD-WAN works | Advice | SD-WAN deployment | Managed Service

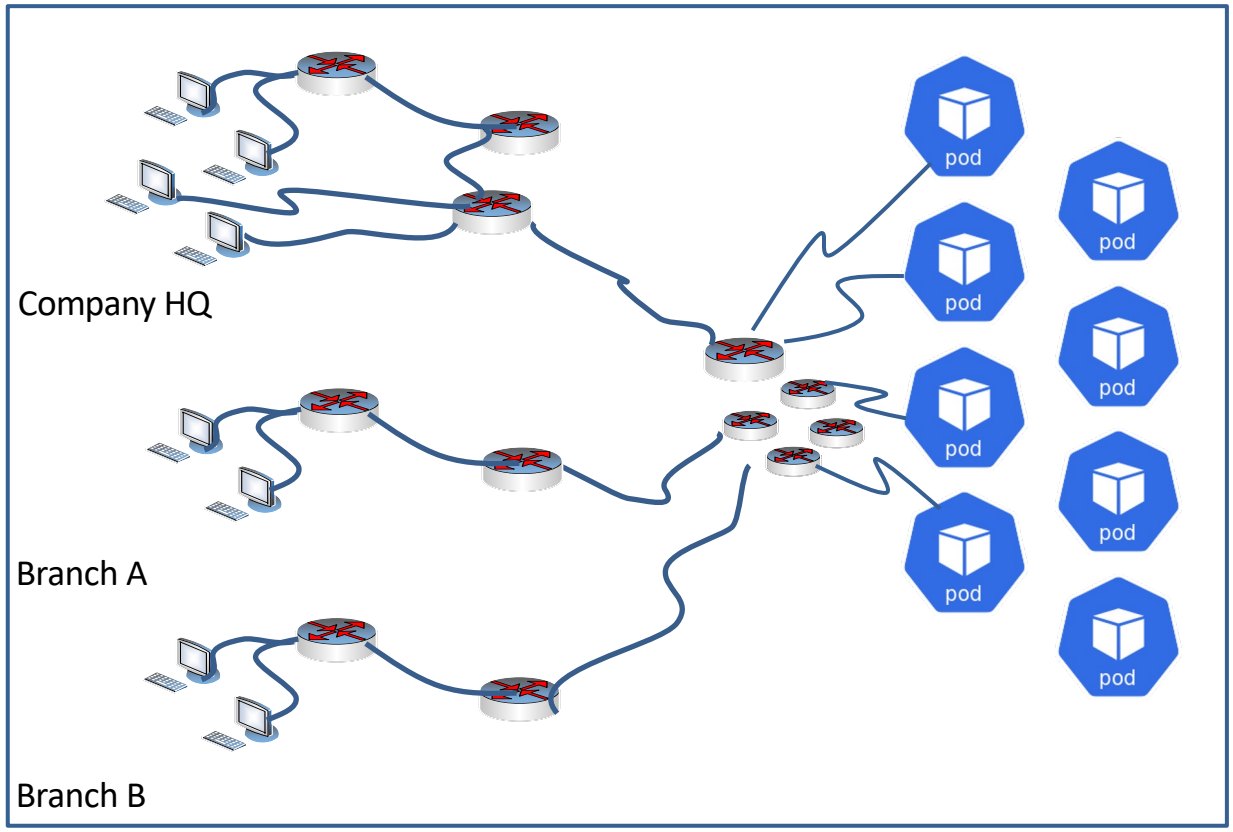
WHAT IS SD-WAN?

SD-WAN is a wide area network – or WAN – architecture that works through a centralised platform, detaching the management of the network from the hardware itself.

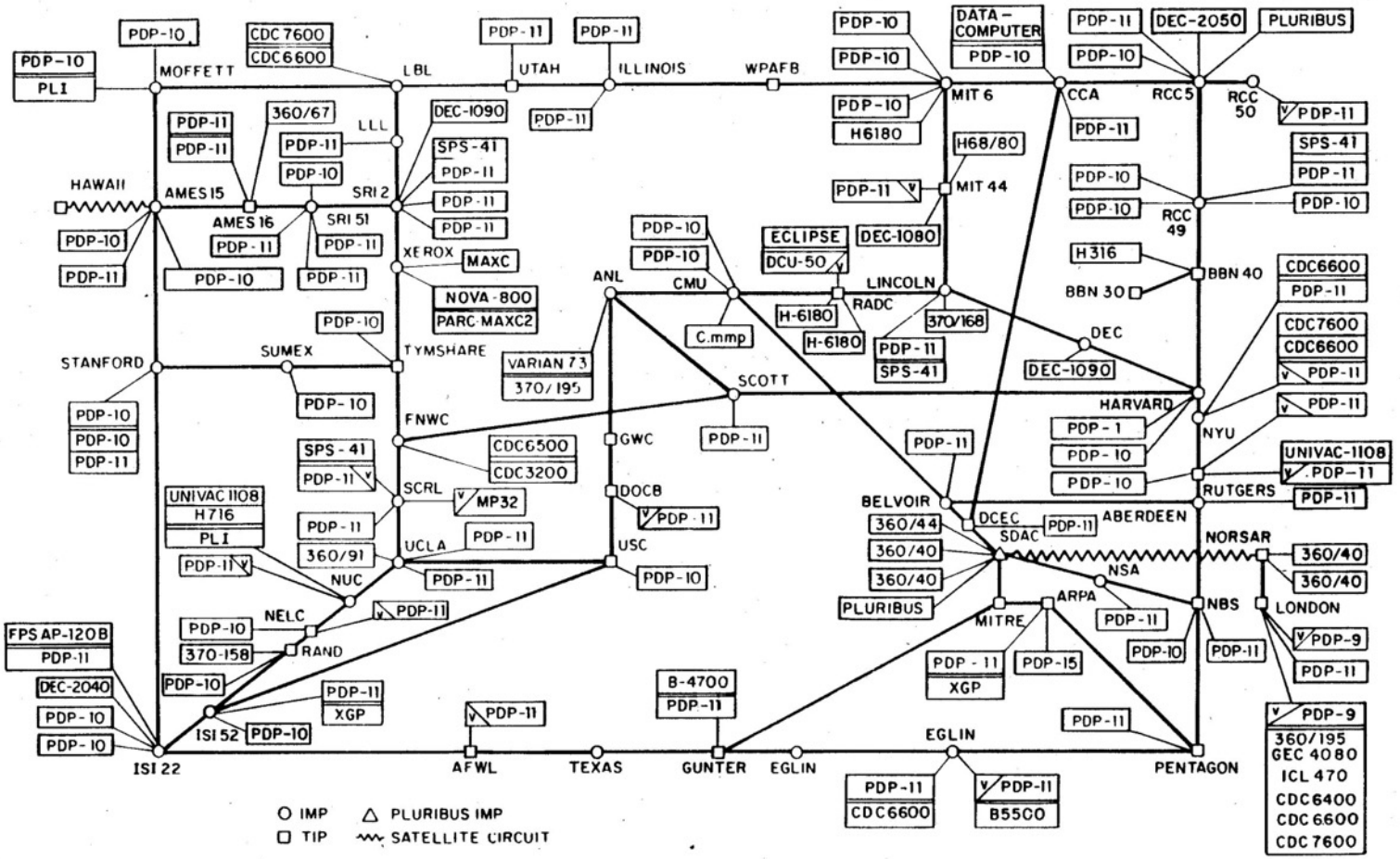


Software-Defined Wide-Area Networks

(Private Networks)

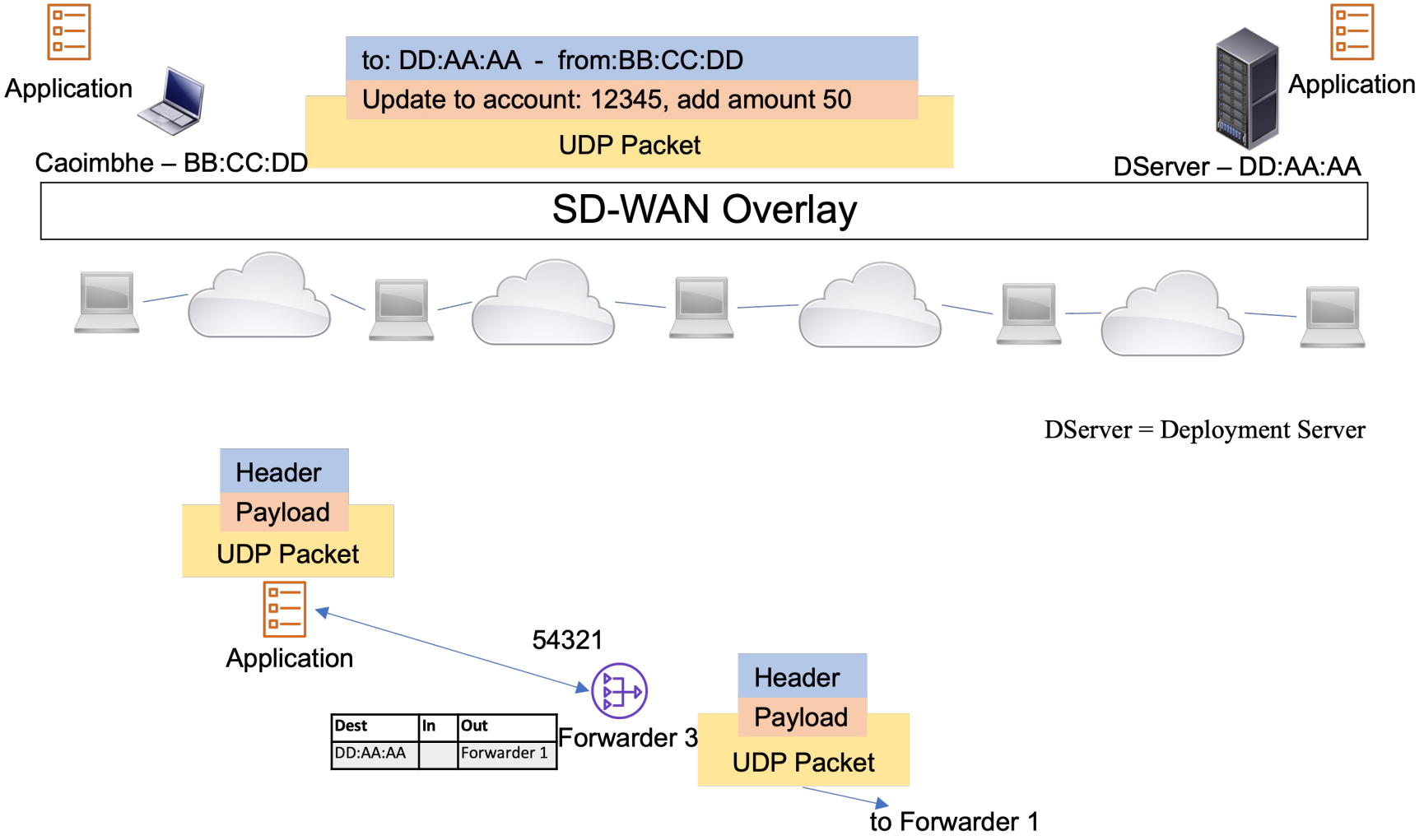


ARPANET 1977

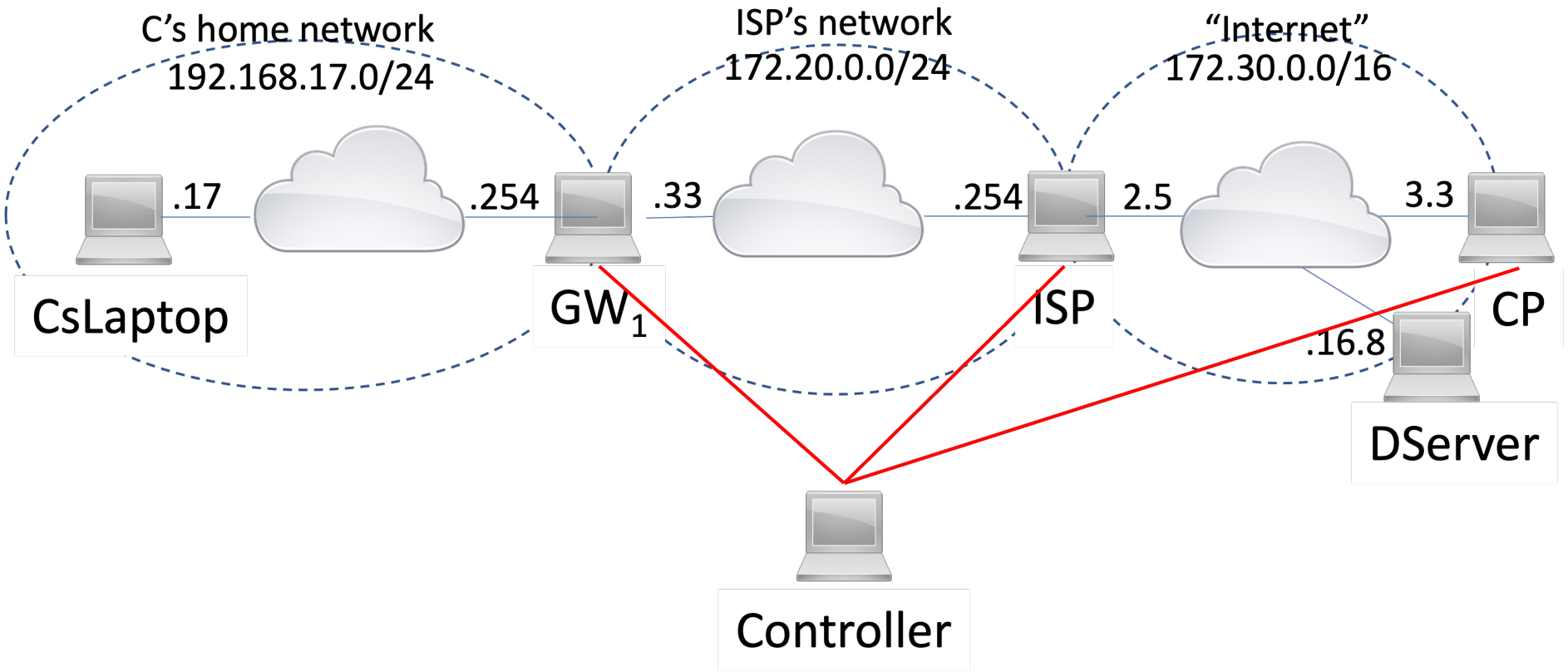


Assignment 2

Conceptual Idea of the Overlay

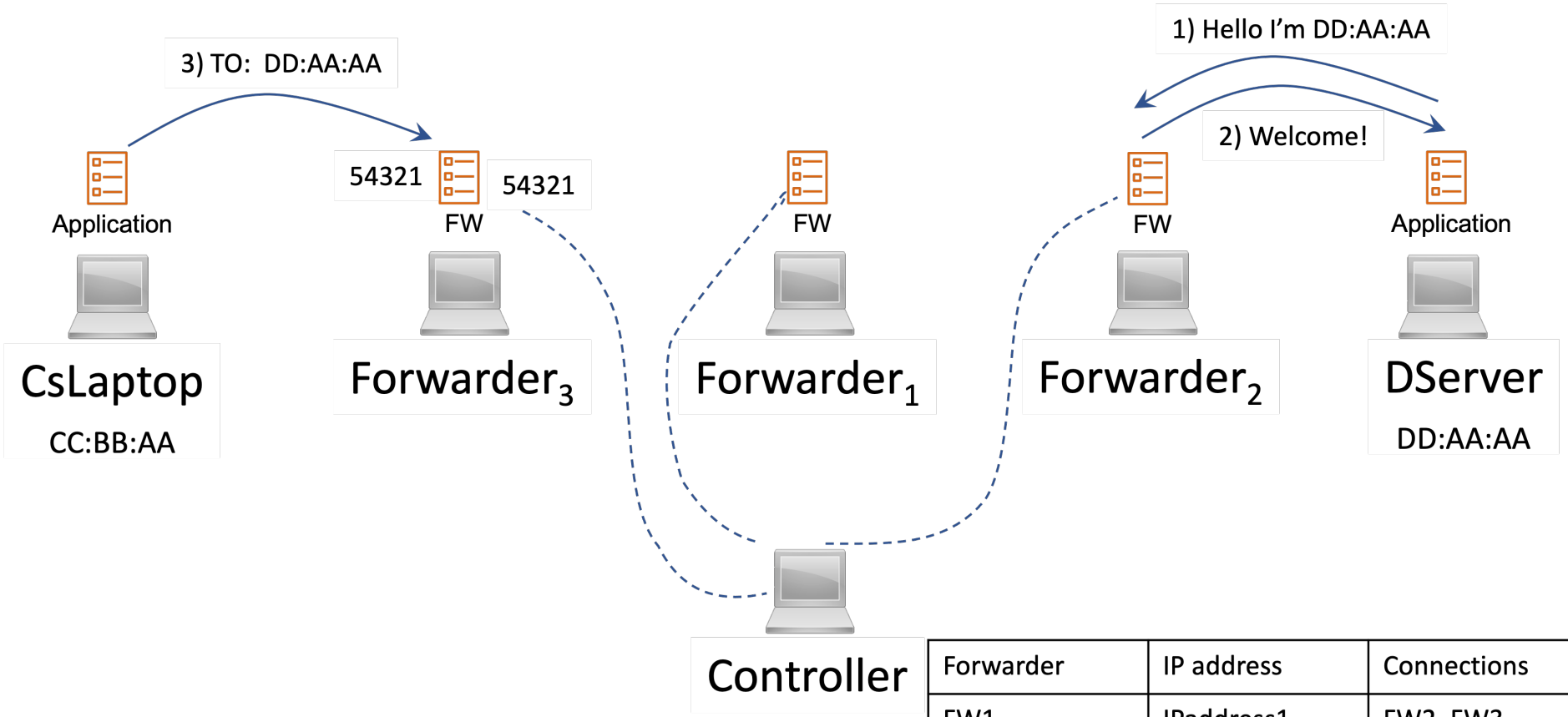


Assignment 2: Possible Topology



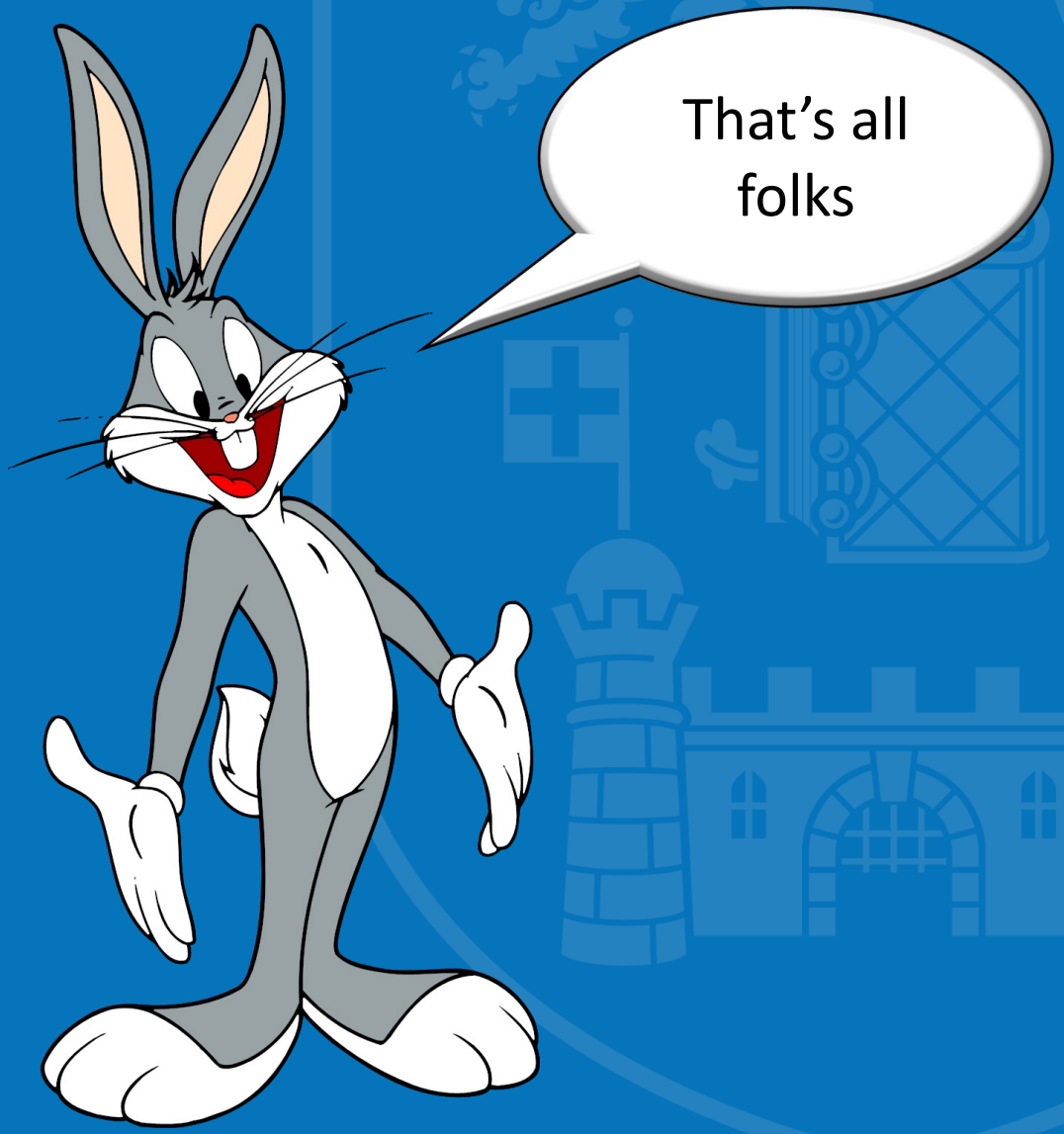
Controller fills Tables of Forwarders

e.g. implementations for Forwarder, Controller, and Application



Forwarder	IP address	Connections
FW1	IPaddress1	FW2, FW3
FW2	IPaddress2	DD:AA:AA, FW1
FW3	IPaddress3	CC:BB:AA, FW1

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin



Recent Publication/Apr 2021

Revitalizing the Public Internet By Making it Extensible

Hari Balakrishnan[1], Sujata Banerjee[2], Israel Cidon[2], David Culler[3], Deborah Estrin[4], Ethan Katz-Bassett[5], Arvind Krishnamurthy[6], James McCauley[7], Nick McKeown[8], Aurojit Panda[9], Sylvia Ratnasamy[10], Jennifer Rexford[11], Michael Schapira[12], Scott Shenker[10,13], Ion Stoica[10], David Tennenhouse[2], Amin Vahdat[3], Ellen Zegura[14*]

[1] MIT, [2] VMware, [3] Google, [4] Cornell Tech, [5] Columbia, [6] UWashington, [7] Mount Holyoke, [8] Stanford, [9] NYU, [10] UC Berkeley, [11] Princeton, [12] HUJI, [13] ICSI, [14] Georgia Tech

Contact:shenker@icsi.berkeley.edu

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.

The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ACM SIGCOMM Computer Communication Review Volume 51 Issue 2, April 2021