Improving the Internet

From Fragility to Resilience

Laurent Vanbever

D-ITET

ETH Zürich

December, 1st 2015

000 08

80 000

estimated # of Internet hosts

in 1990

2.8 billion

estimated* # of Internet hosts

in 2014, 6 months before I joined ETH

* Cisco Visual Networking Index 2015

3.9 billion

estimated # of Internet hosts

in 2019

~2 exabytes

estimated global IP traffic in 2015

per day

If 

= 1 Gigabyte


volume(Great Wall of China) = 1 exabyte

~2 exabytes

estimated global IP traffic in 2015

per day

~6 exabytes

estimated global IP traffic in 2019

per day

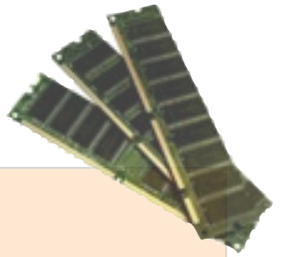
Looks like everything is going well?

Looks like everything is going well?

Well, not exactly…

Here are some “amuse bouches”
of the challenges the Internet faces

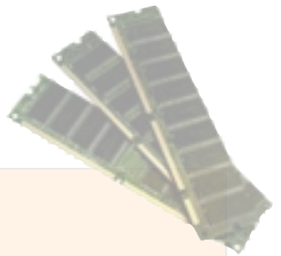
manageability &
reliability



scalability

security

manageability & reliability



scalability

security

JUL 8, 2015 @ 03:36 PM 11,261 VIEWS

United Airlines Blames Router for Grounded Flights

Alexandra Talty, CONTRIBUTOR

I cover personal finance and travel.

FOLLOW ON FORBES (110)

Opinions expressed by Forbes Contributors are their own.

FULL BIO

After a computer problem caused nearly two hours of grounded flights for United Airlines this morning and ongoing delays throughout the day, the airline announced the culprit: a faulty router.

Spokeswoman Jennifer Dohm said that the router problem caused "degraded network connectivity," which affected various applications.

A computer glitch in the airline's reservations system caused the Federal Aviation Administration to impose a groundstop at 8:26 a.m. E.T. Planes that were in the air continued to operate, but all planes on the ground were held. There were reports of agents writing tickets by hand. The ground stop was lifted around 9:47 a.m. ET.

The outage was due to

one faulty Internet device

Facebook, Tinder, Instagram suffer widespread issues

IMAGE: GETTY IMAGES

BY JENNI RYALL
AUSTRALIA

JAN 27, 2015

***UPDATED**: Tuesday, Jan. 27 / 4:32 a.m. EST* — A Facebook spokeswoman told *Mashable* that the outage was due to a change to the site's configuration systems, and not a hacker attack. "Earlier this evening many people had trouble accessing Facebook and Instagram. This was not the result of a third party attack but instead occurred after we introduced a change that affected our configuration systems. We moved quickly to fix the problem, and both services are back to 100% for everyone.", she said.

***UPDATED**: Tuesday, Jan. 27 / 2:14 a.m. EST* — Facebook, Tinder and Twitter appear to be back to normal after a 40 minute outage and mass freak out.

The outage was due to a change to the site's configuration systems

“Human factors are responsible for 50% to 80% of network outages”

Juniper Networks, *What’s Behind Network Downtime?*, 2008

The Internet Under Crisis Conditions

Learning from September 11

Committee on the Internet Under Crisis Conditions:
Learning from September 11

Computer Science and Telecommunications Board
Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

National Research Council. The Internet Under Crisis Conditions: Learning from September 11

The Internet Under Crisis Conditions

Learning from September 11

Committee on the Internet Under Crisis Conditions:
Learning from September 11

Computer Science and Telecommunications Board
Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Internet advertisements rates

suggest that

The Internet was more stable

than normal on Sept 11

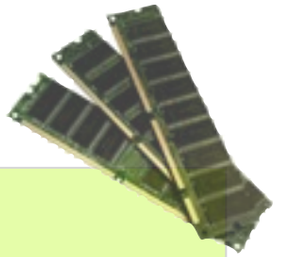
Internet advertisements rates suggest that The Internet was more stable than normal on Sept 11

Information suggests that operators were watching the news instead of making changes to their infrastucture

“Cost per network outage
can be as high as 750 000$”

Smart Management for Robust Carrier Network Health
and Reduced TCO!, NANOG54, 2012

manageability & reliability



scalability

security

Scalable routing systems maintain

- detailed information about nearby destination
- coarse-grained information about far-away destination

The internet maintains

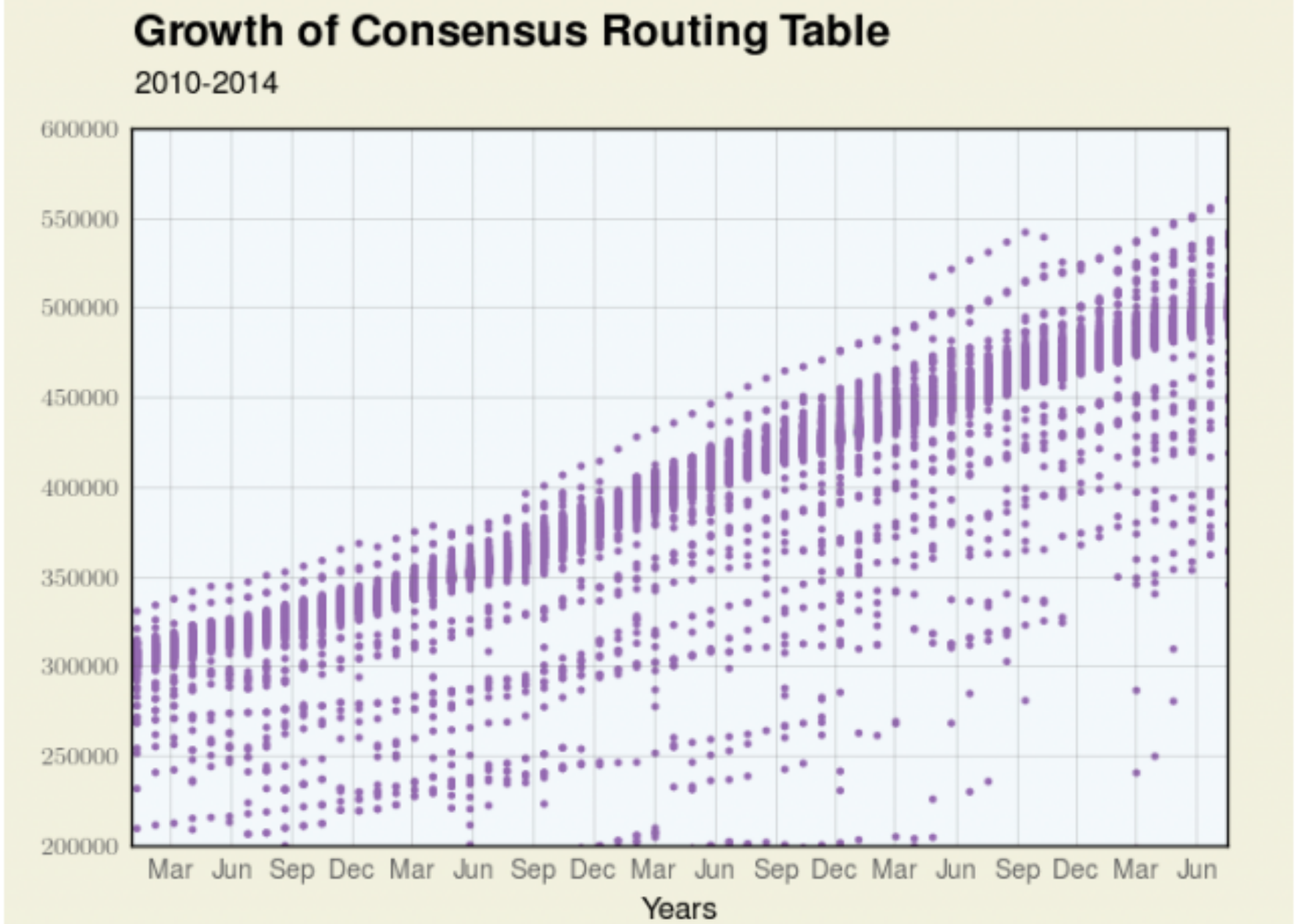
detailed information about every destination

Sign Post Forest, Watson Lake, Yukon

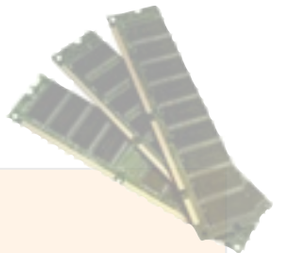
Internet routers hitting 512K limit, some become unreliable

Table limit in some routers causes minor Internet outages; more to come.

by **Robert Lemos** - Aug 13, 2014 9:03pm CEST



manageability & reliability



scalability

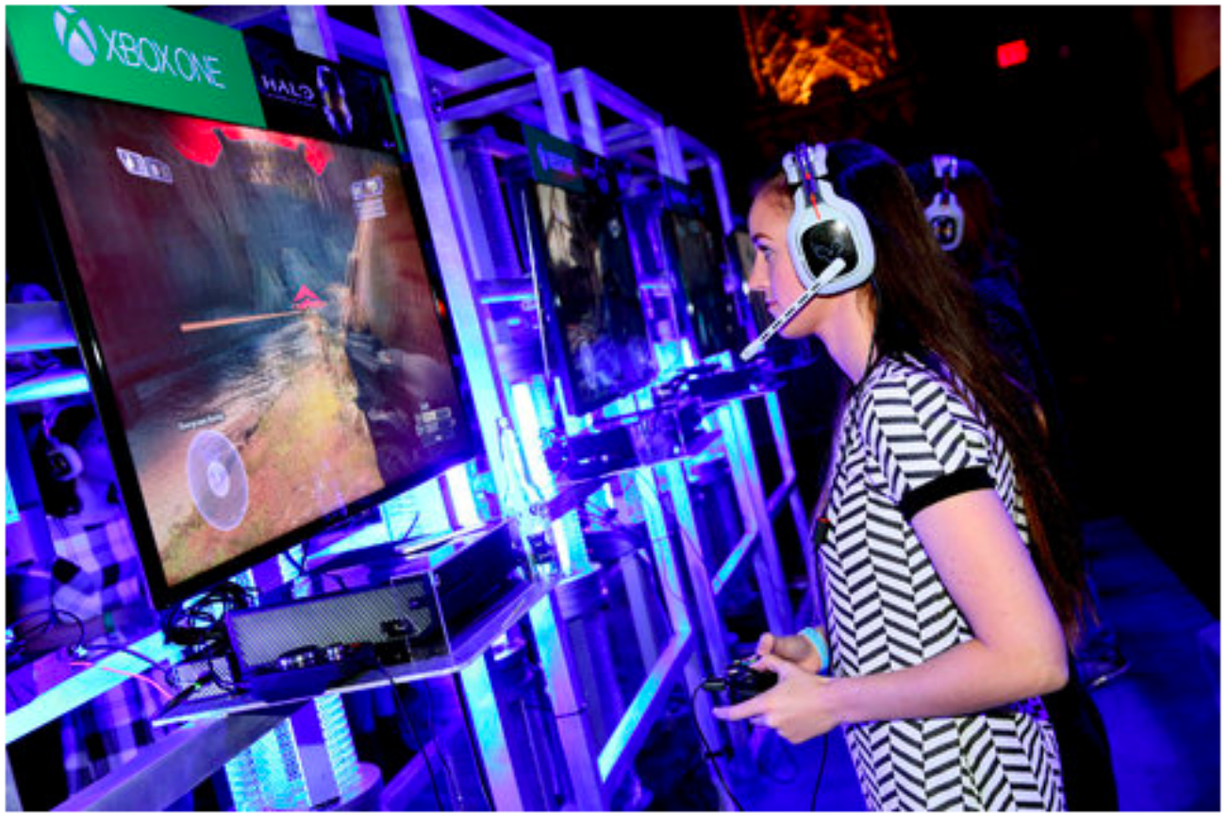
security

Bits

GAMING

Sony PlayStation and Microsoft Xbox Live Networks Attacked by Hackers

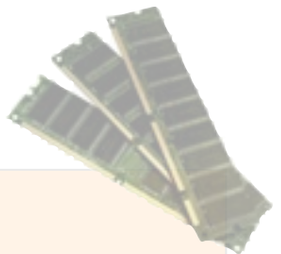
By NICOLE PERLROTH and BRIAN X. CHEN DECEMBER 26, 2014 4:11 PM 31 Comments



Xbox fans playing games from the Halo series last month at the HaloFest event in Los Angeles. Matt Sayles/Invision for Microsoft, via Associated Press

The computer networks for Microsoft's Xbox and Sony's PlayStation 4 video game consoles were off line for most of Christmas Day, possibly because of an attack by a group of hackers with a history of targeting video games.

manageability &
reliability



scalability

security

Solving these problems was hard because network devices are completely locked down

closed software

closed hardware

Cisco™ device

There is hope for change though:

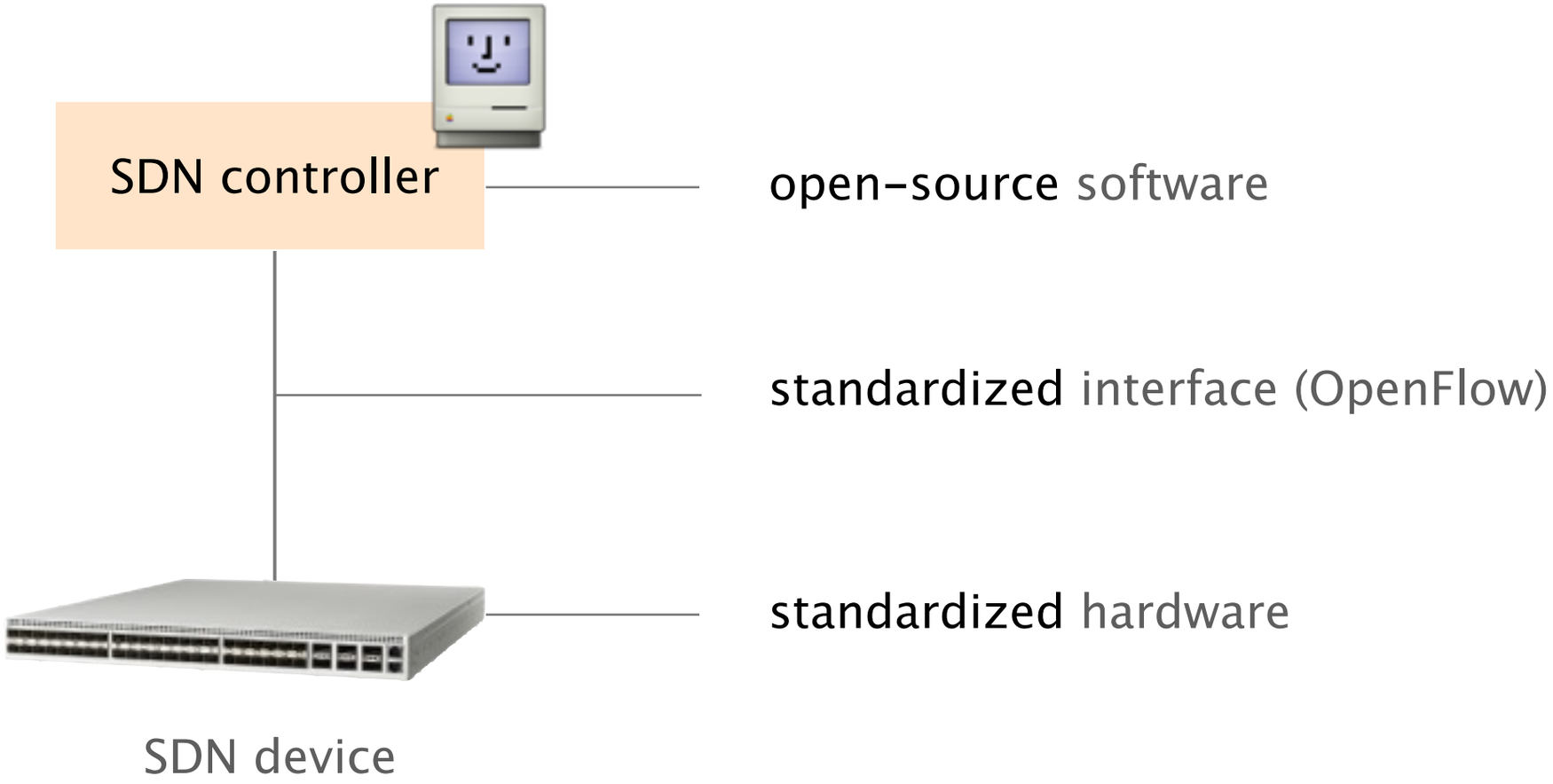
networks are on the verge of a paradigm shift

Software-Defined Network

Software-Defined Network

enable network programmability

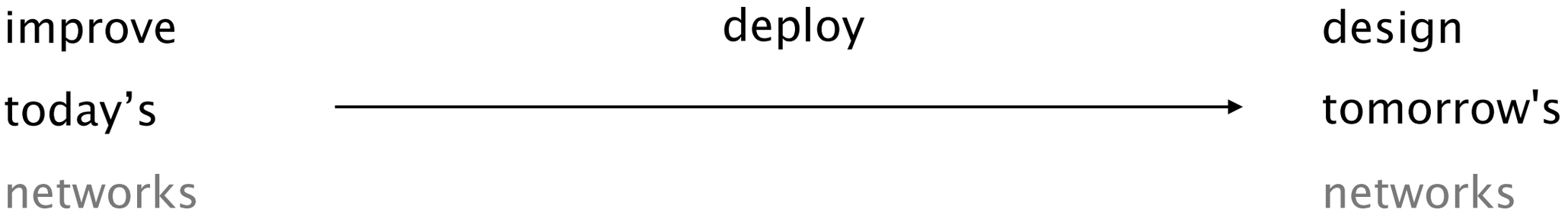
SDN also enables us, researchers,
to innovate, at a much faster pace



Our focus

these days

Leverage network programability to…



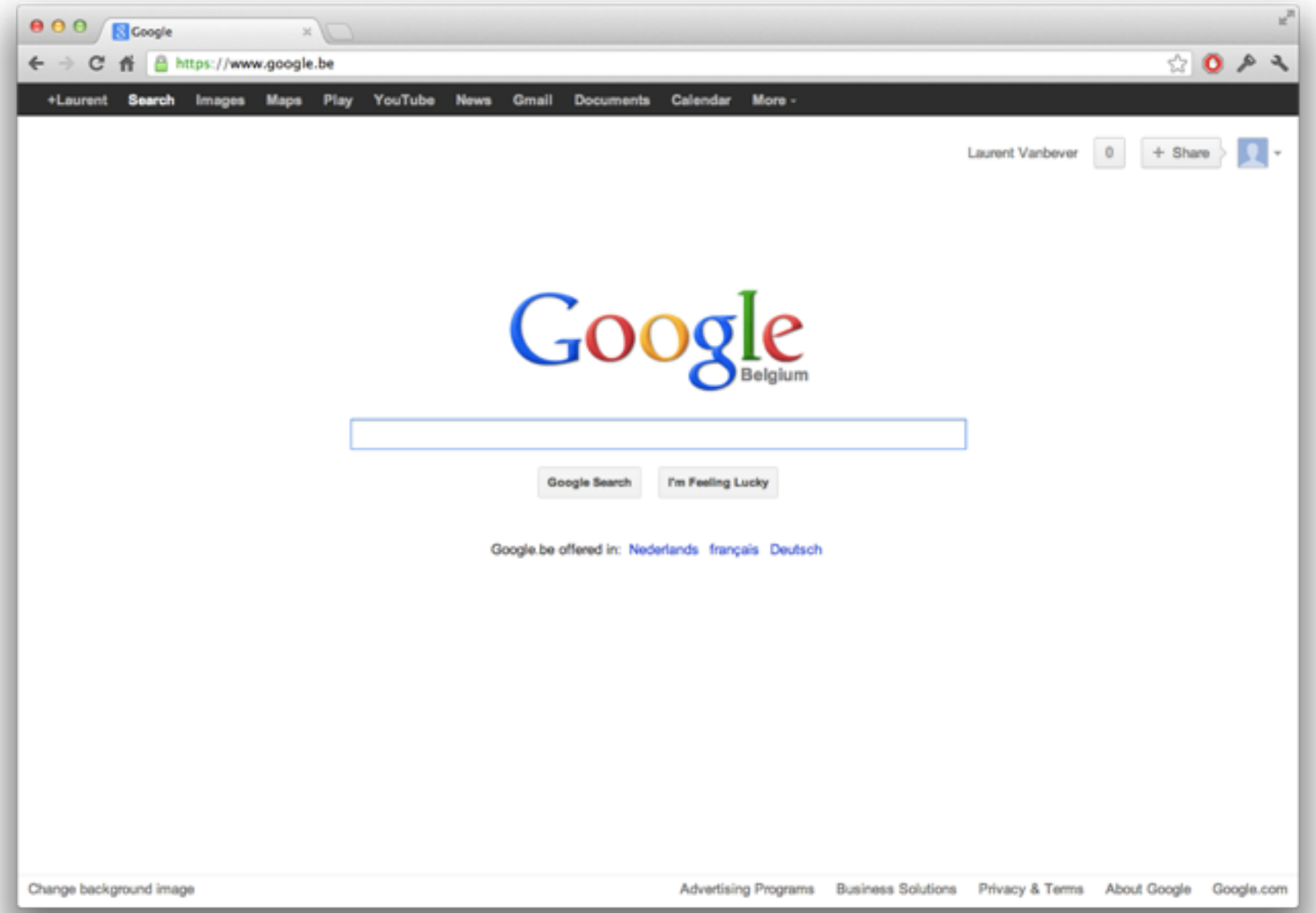
What is a network?

For a lot of people,

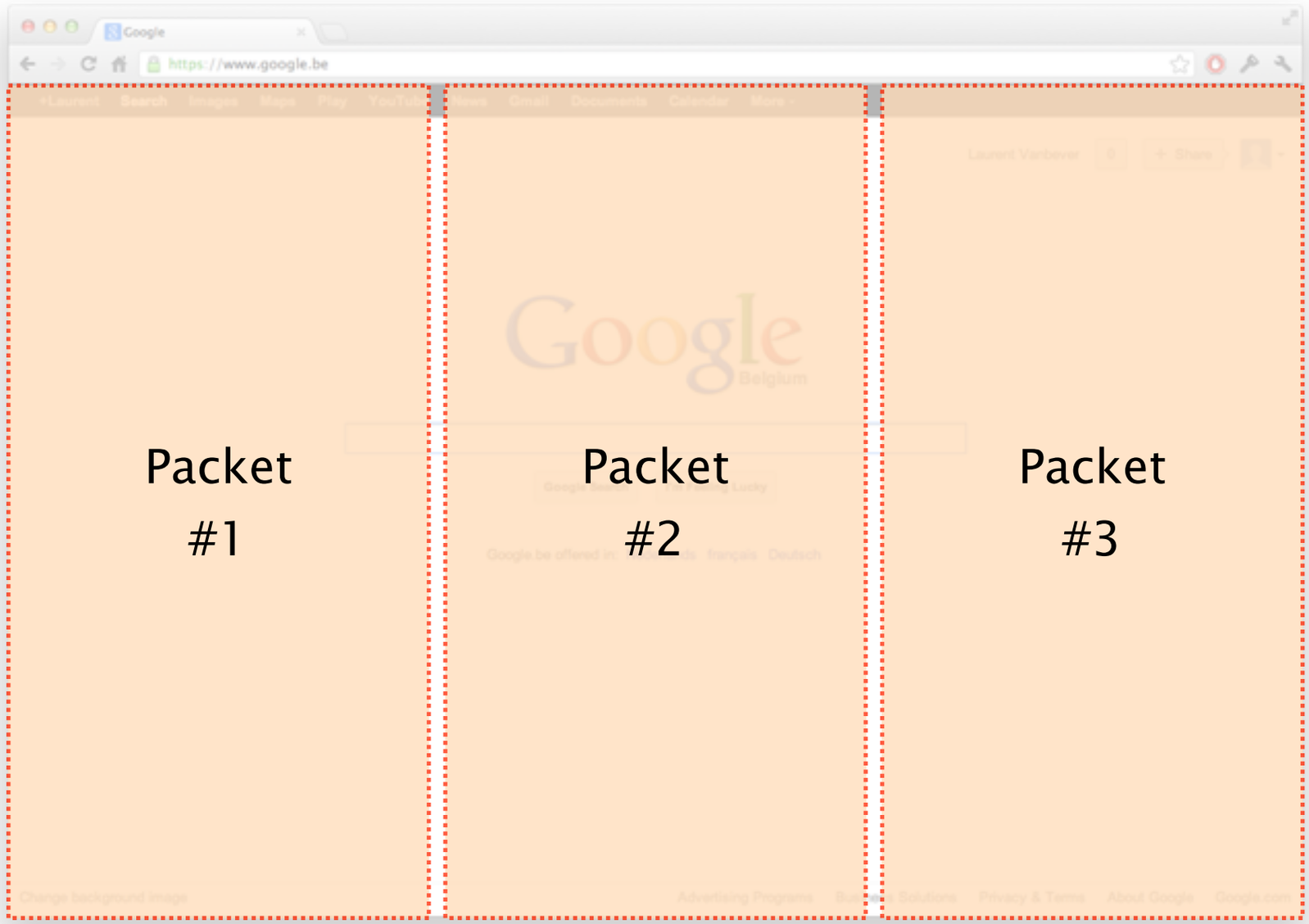
this is what the Internet looks like

Network connection

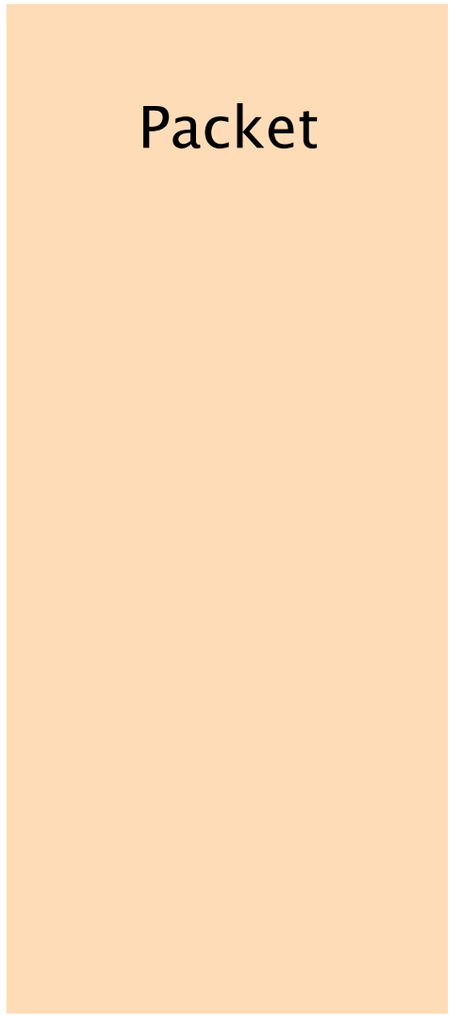
Data exchanged on the Internet is fragmented into small chunks, called IP packets



Data exchanged on the Internet is fragmented into small chunks, called IP packets

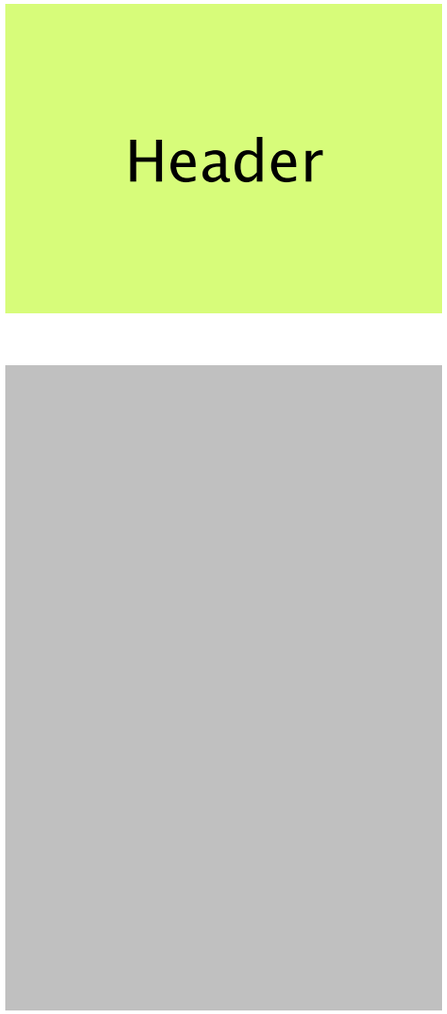


You can think of packets as envelopes



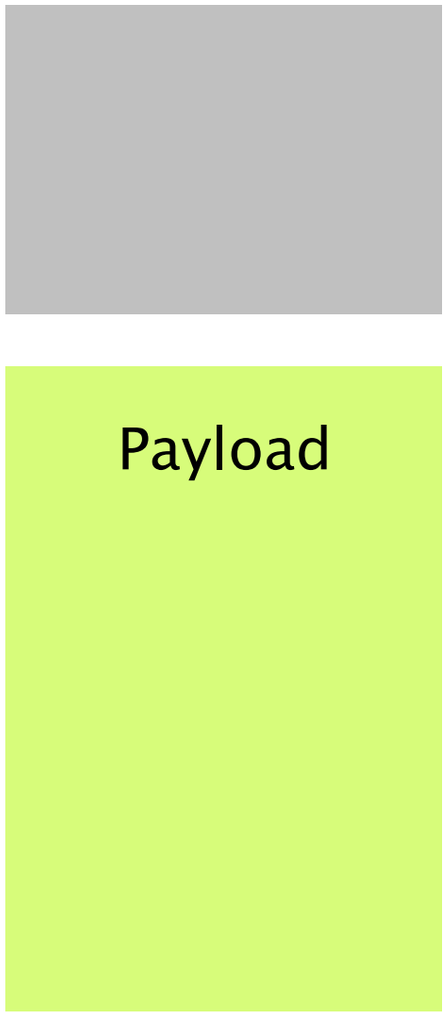
Like an envelope,

packets have a header



Like an envelope,

packets have a payload

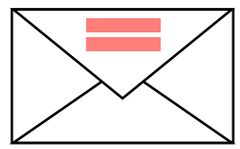


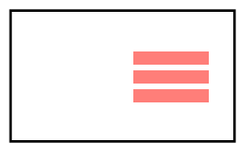
The header contains the metadata needed to forward the packet

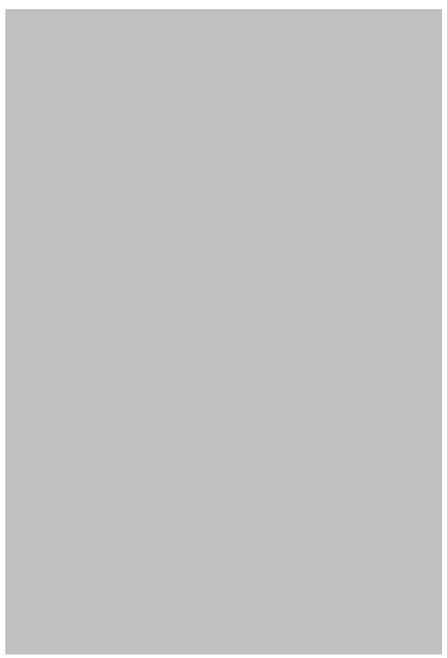
Identify the

src address — source

dst address — destination





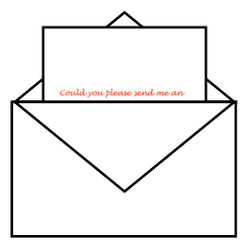


of the communication

The payload contains
the data to be delivered

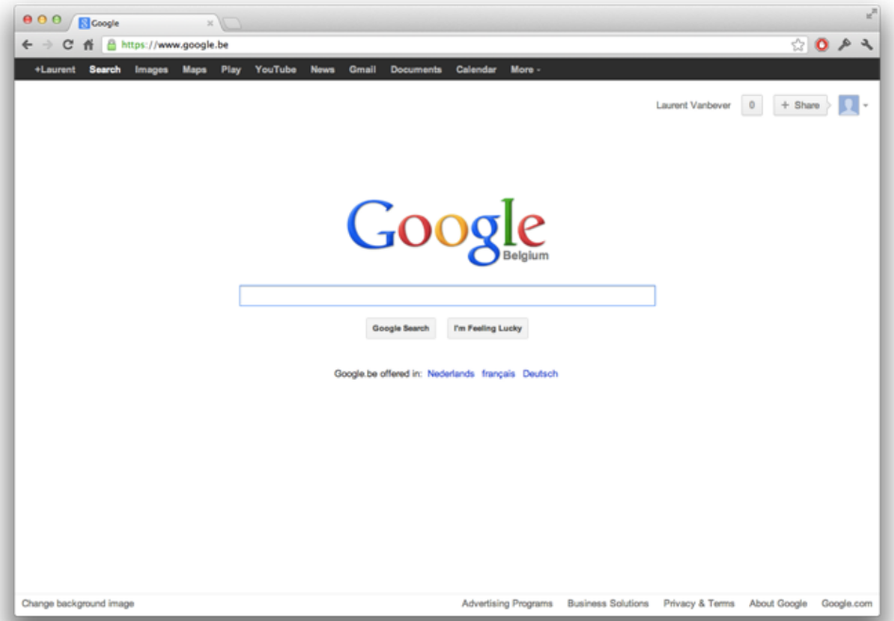
Payload

```
<html><head>
<meta http-equiv="content-type" content="text/html; charset=UTF-8">
<title>Google</title>
</head><body>
  <img alt="Google" height=110 src="images/logo.gif" width=276>
  <form action="/search" name=f>
   <input name=hl type=hidden value=en>
   <input name=q size=55 title="Google Search" value="">
   <input name=btnG type=submit value="Google Search">
   <input name=btnI type=submit value="I'm Feeling Lucky">
 </form>
</body></html>
```



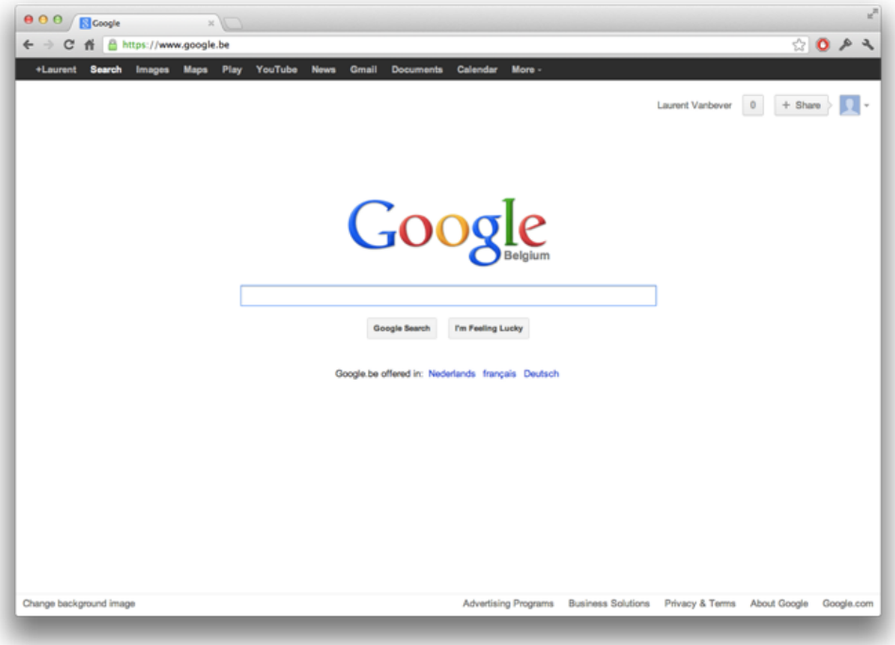
Google

Network connection



src	Laurent
dst	Google
Payload	

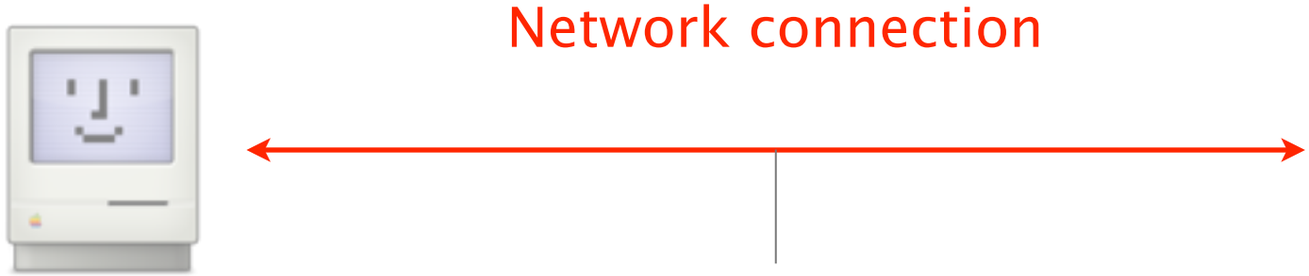
Network connection



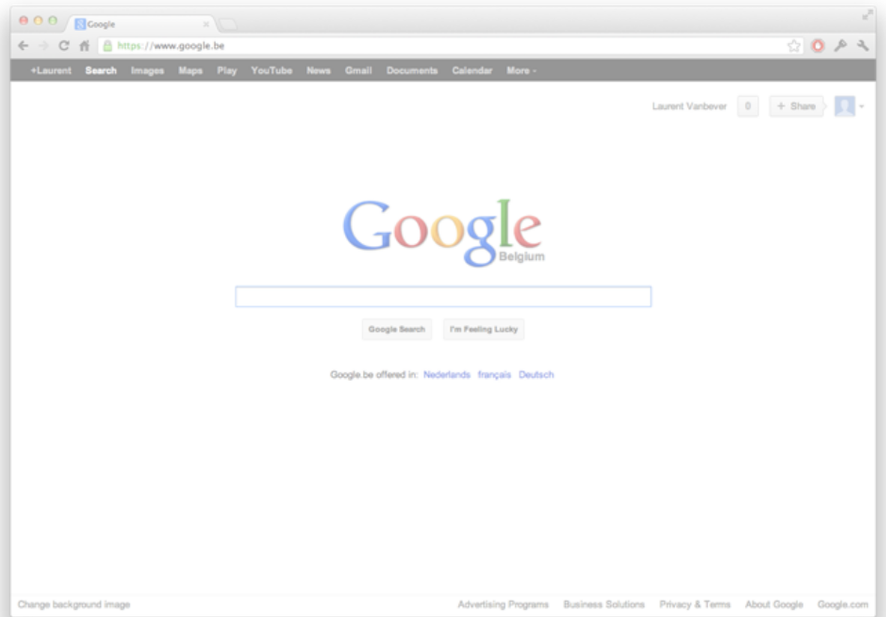
src Google

dst Laurent

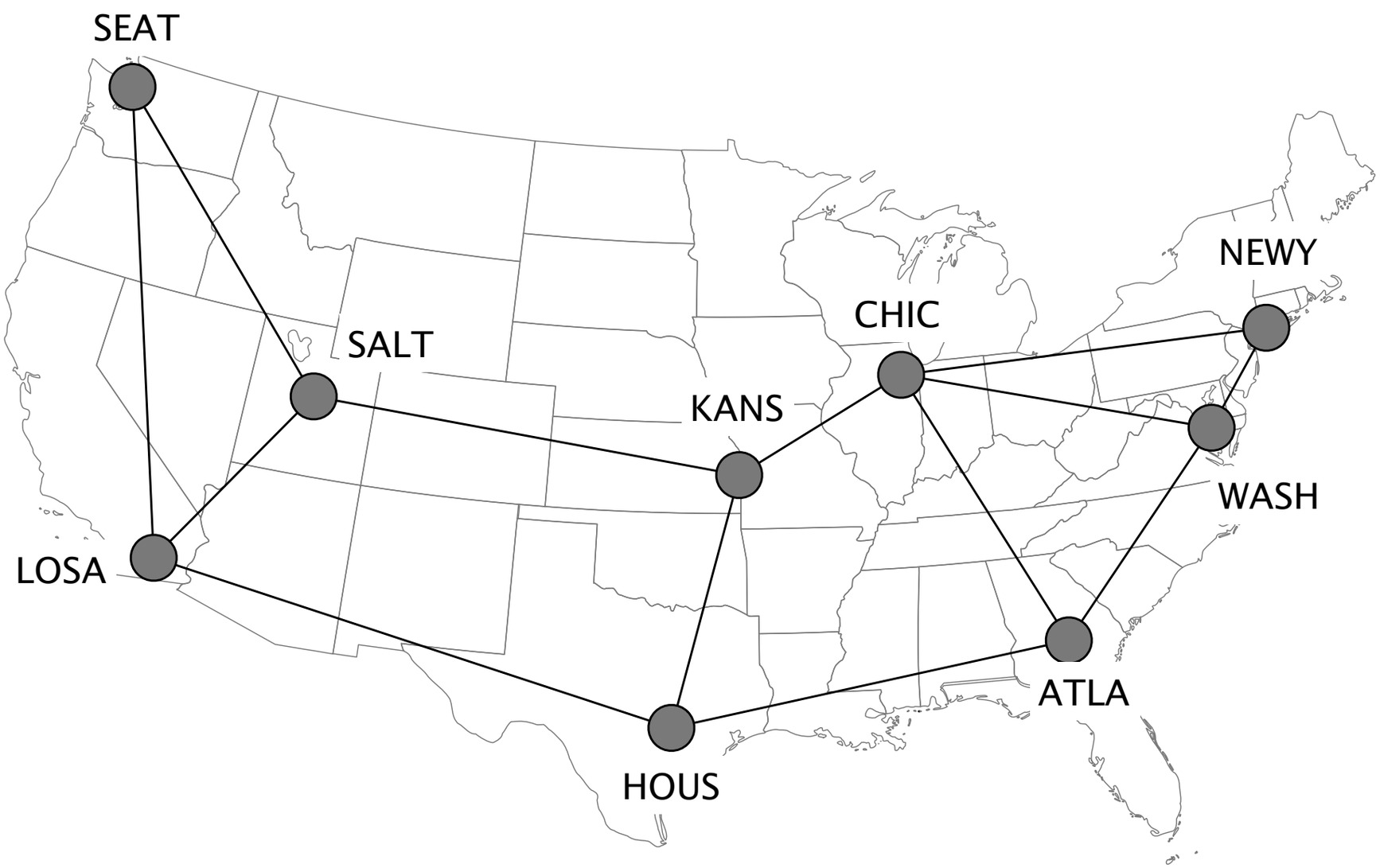
Payload


Network connection

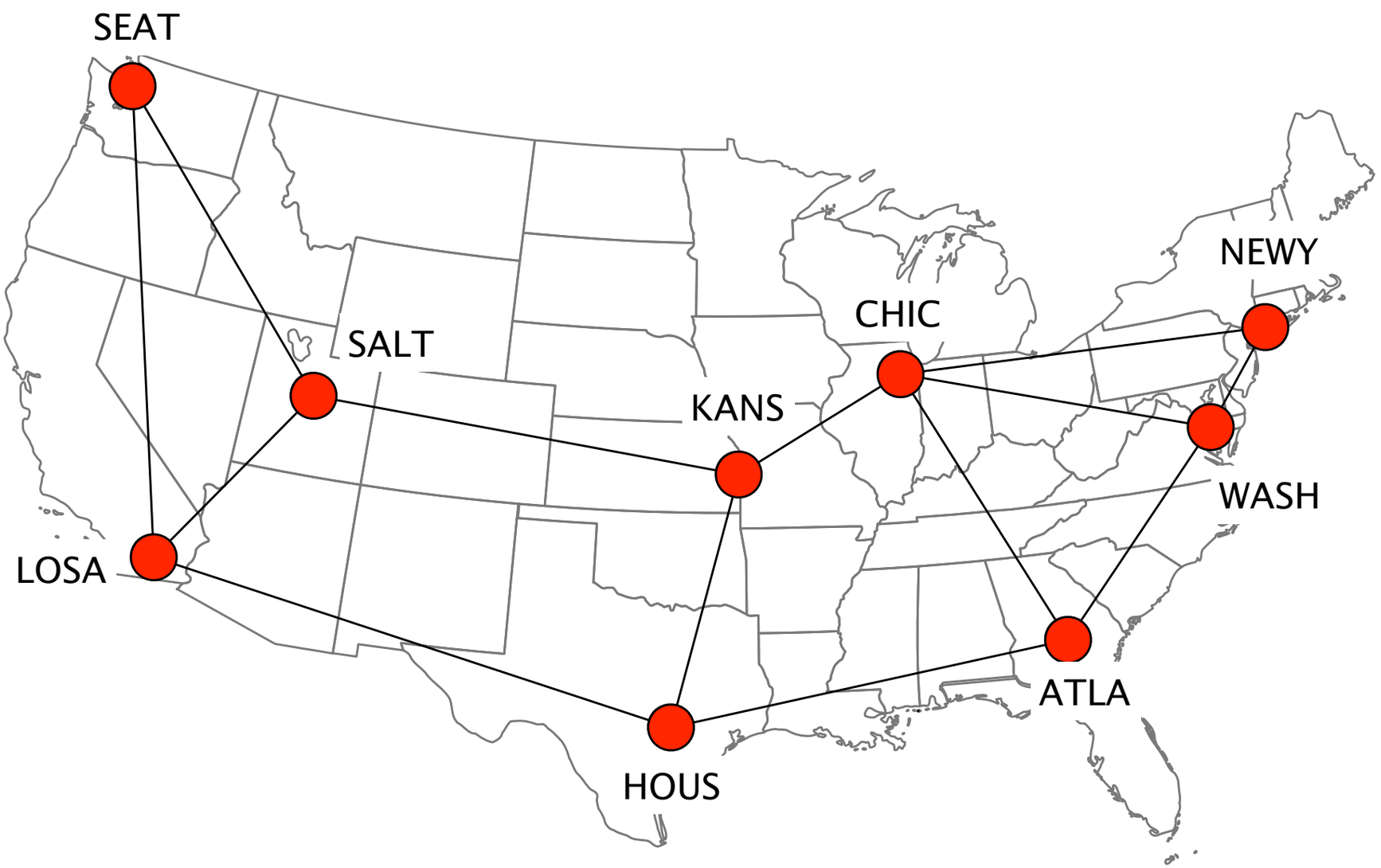
What is this?


Google
https://www.google.be
+Laurent
Search
Images
Maps
Play
YouTube
News
Gmail
Documents
Calendar
More
Laurent Vanbever
Share
Google
Belgium
Google Search
I'm Feeling Lucky
Google.be offered in: Nederlands français Deutsch
Change background image
Advertising Programs
Business Solutions
Privacy & Terms
About Google
Google.com

A network is a distributed system



A network is a distributed system composed of routers



Routers vary in size and usage

Home router

~20 cm

0,5 kg

1 Gbps

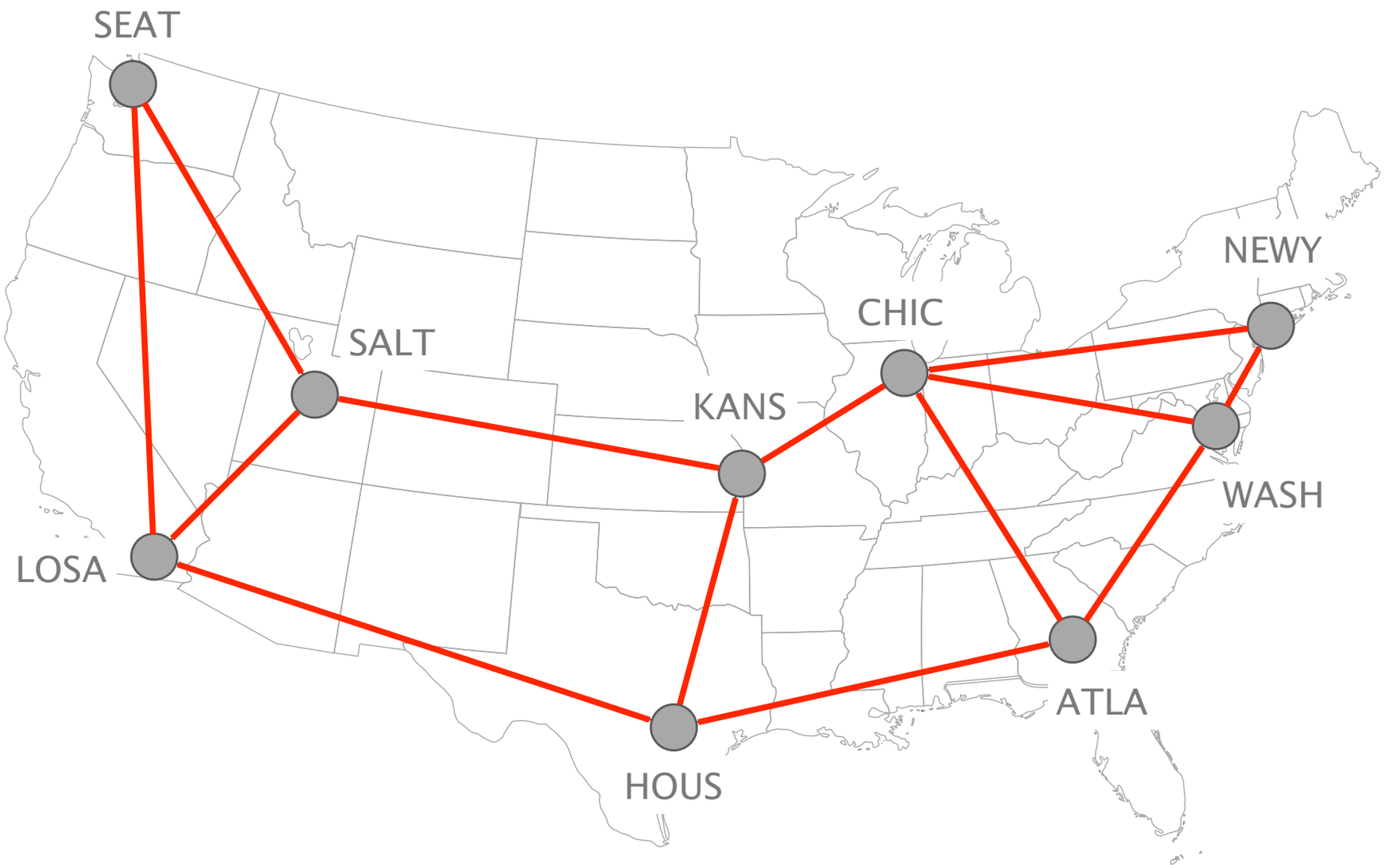
Internet core router

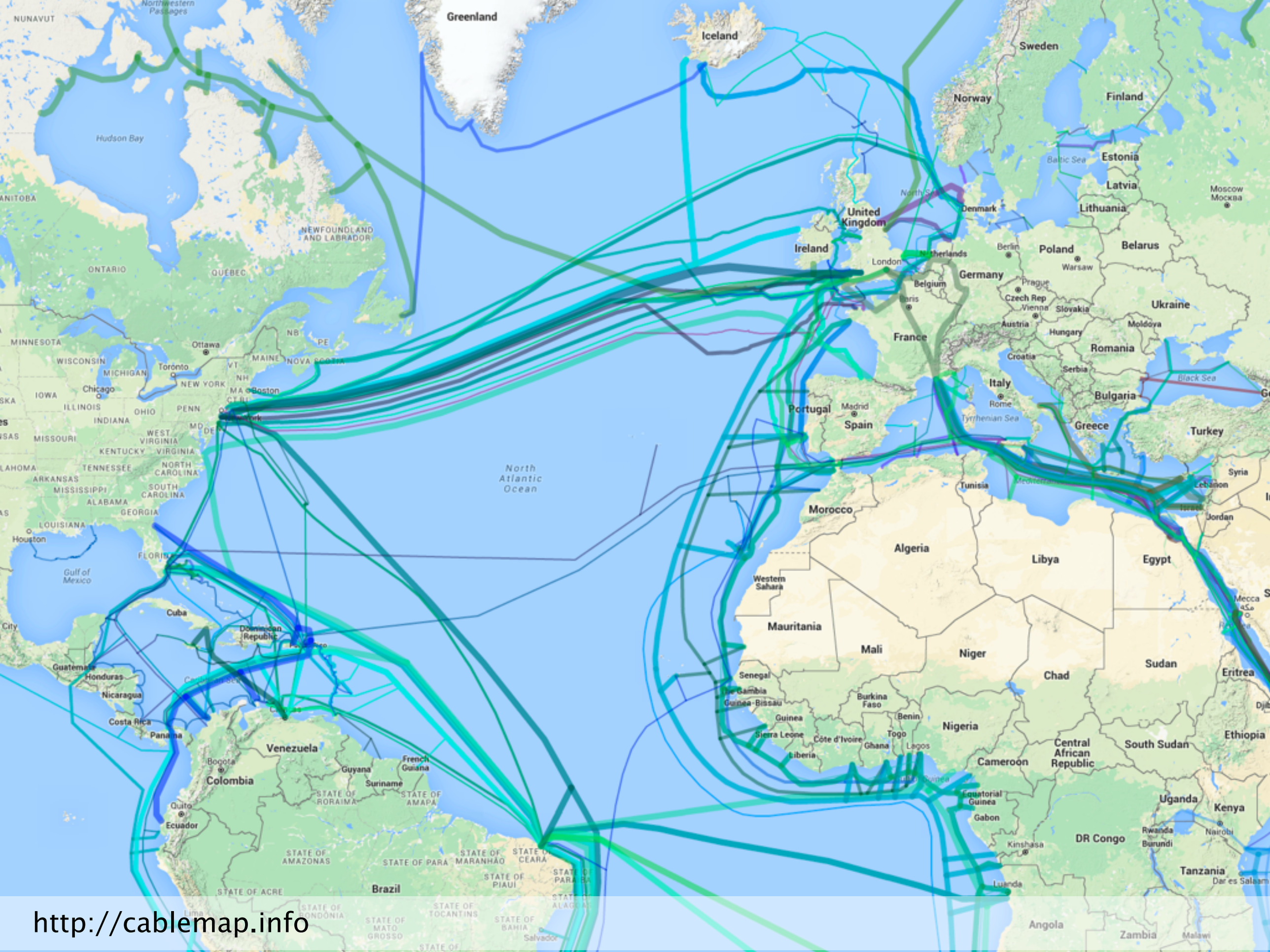
>200cm

700kg

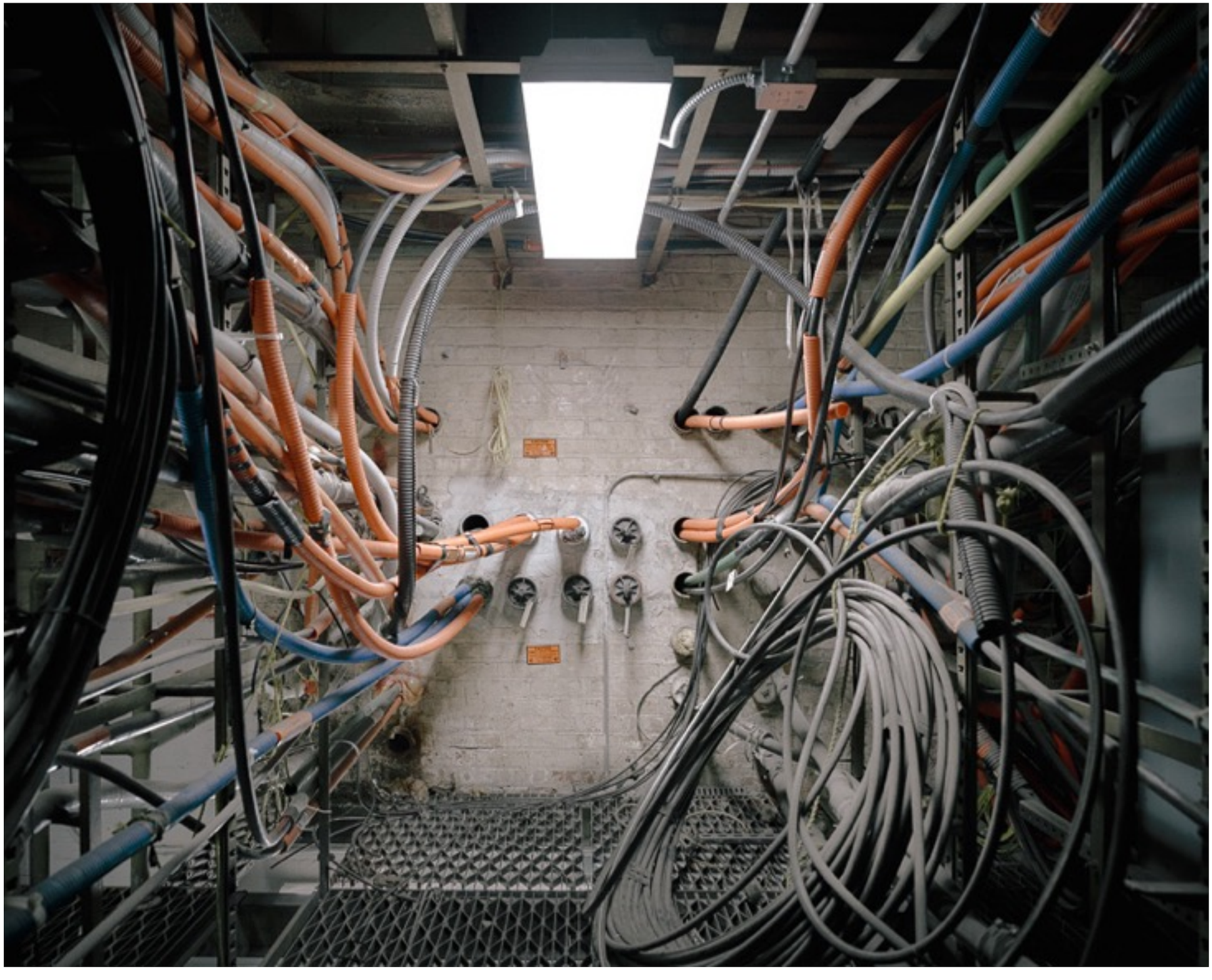
1.2 Tbps

A network is a distributed system composed of routers and links



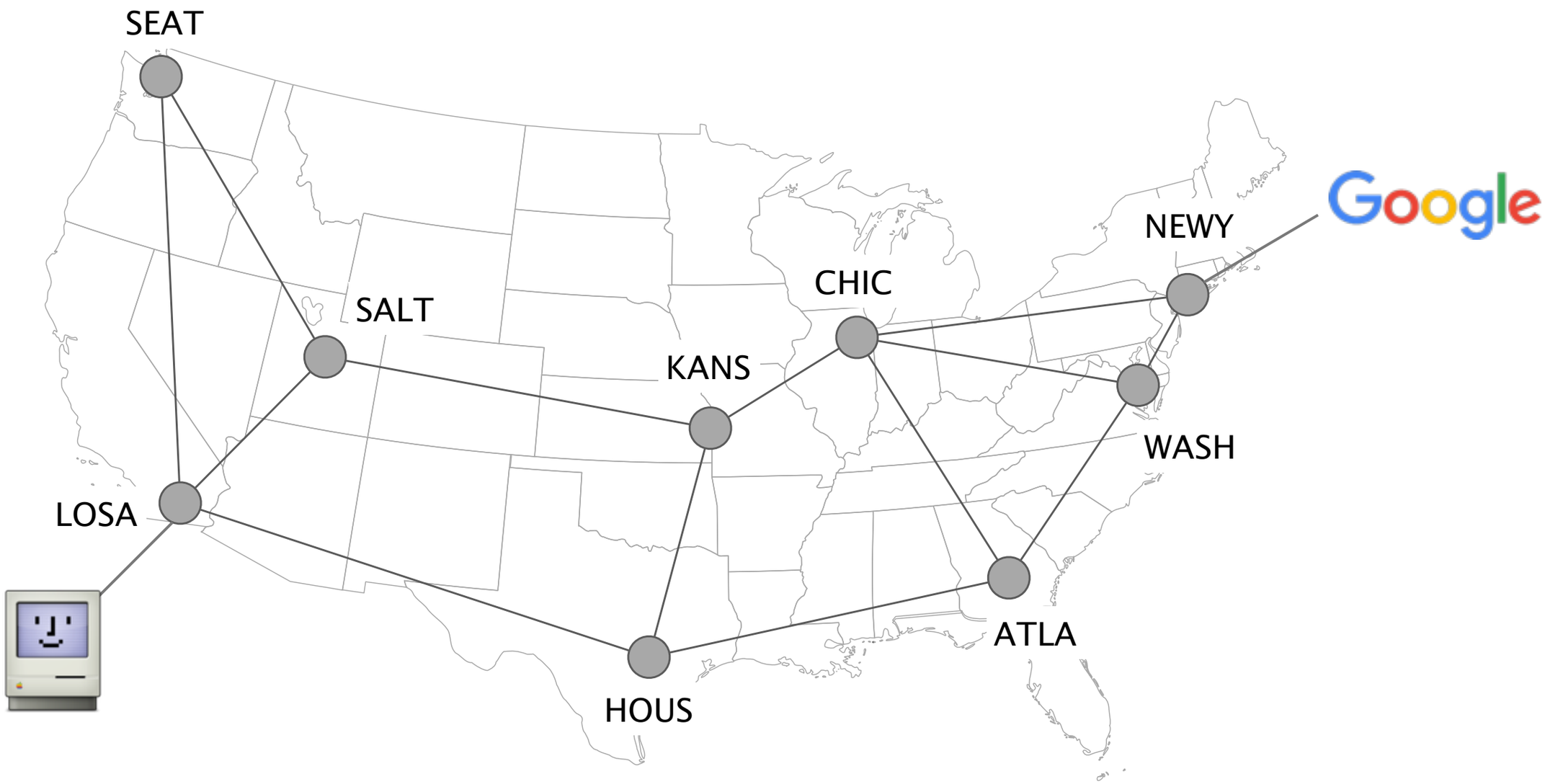


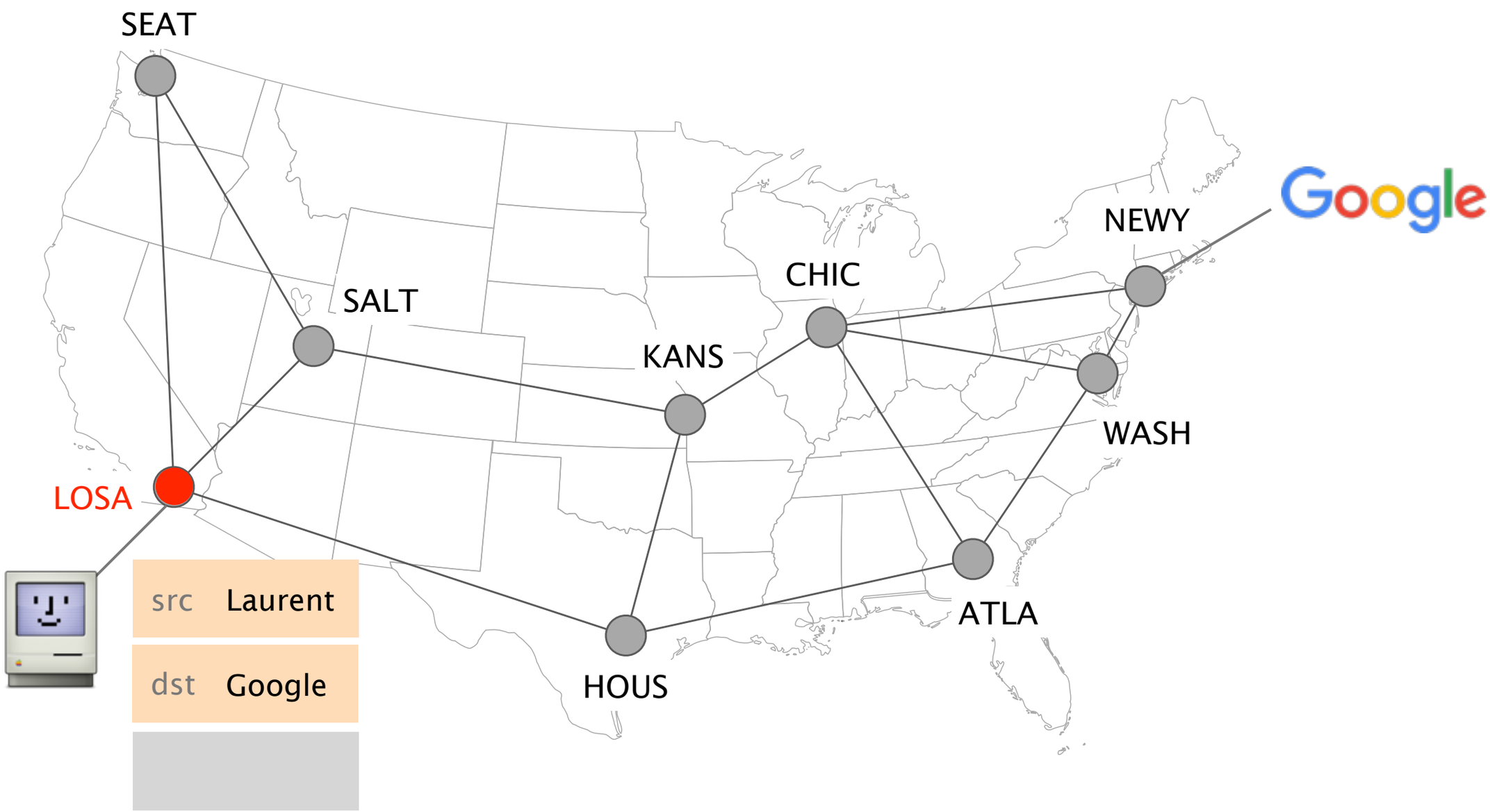
http://cablemap.info

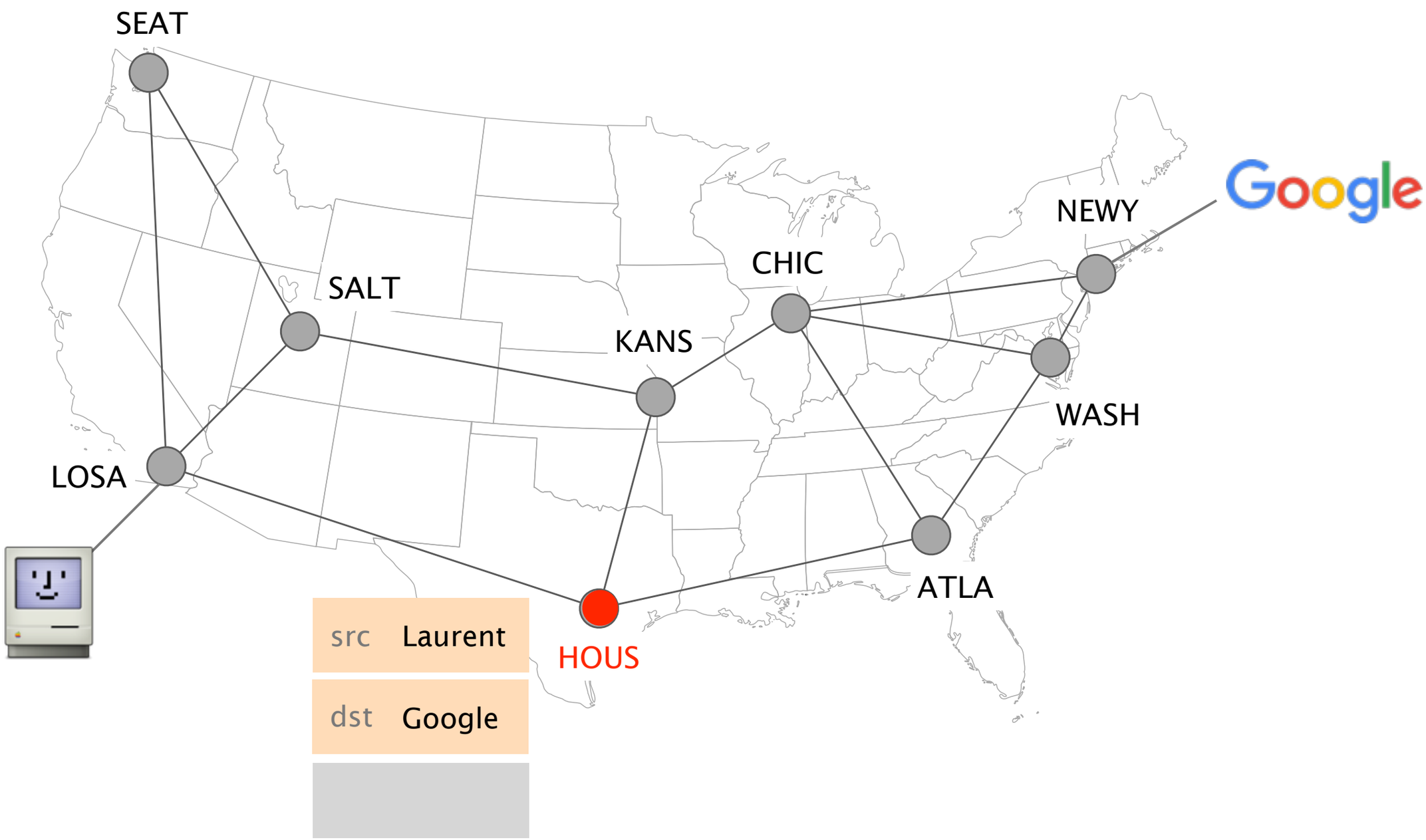


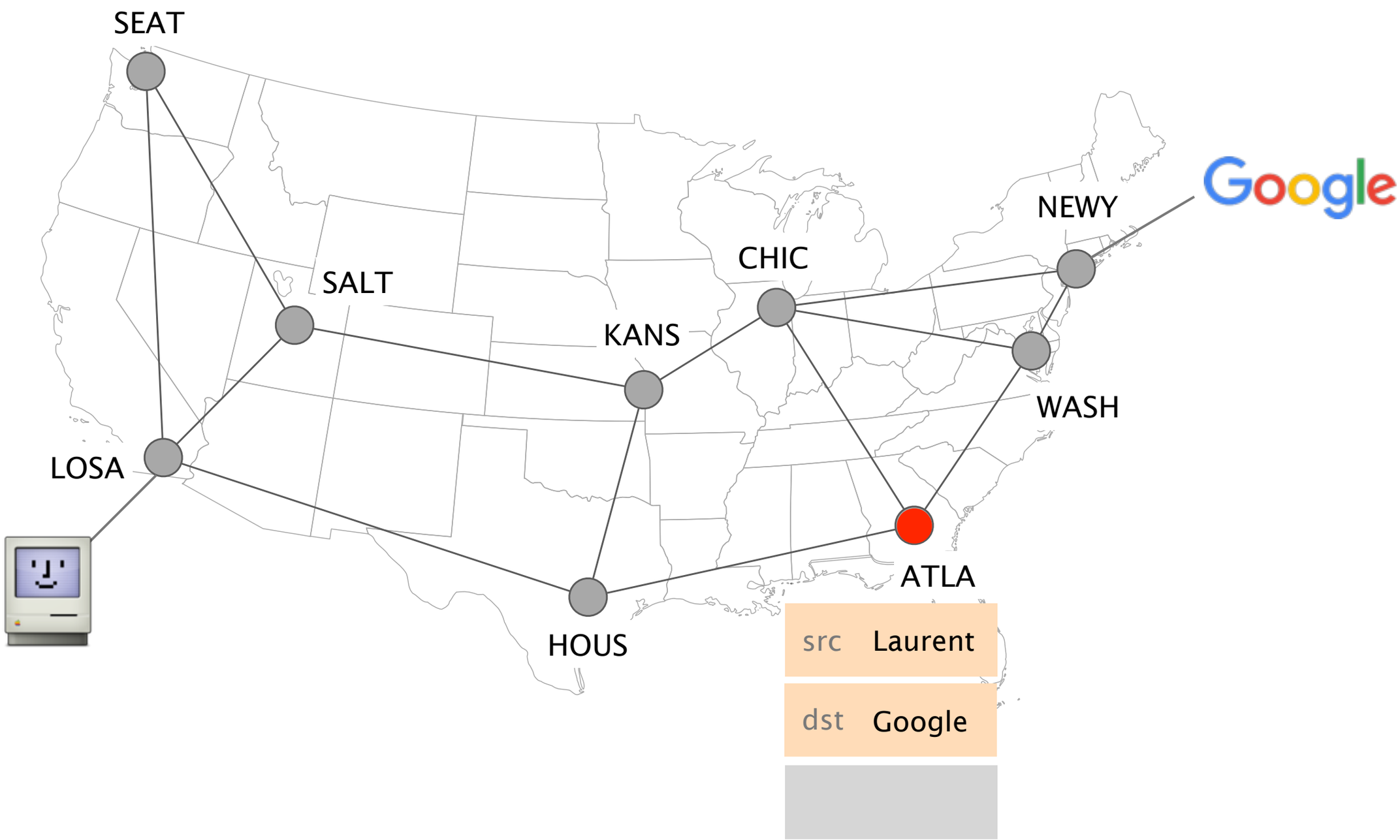
Somewhere in Manhattan… http://www.petergarritano.com/the-internet.html

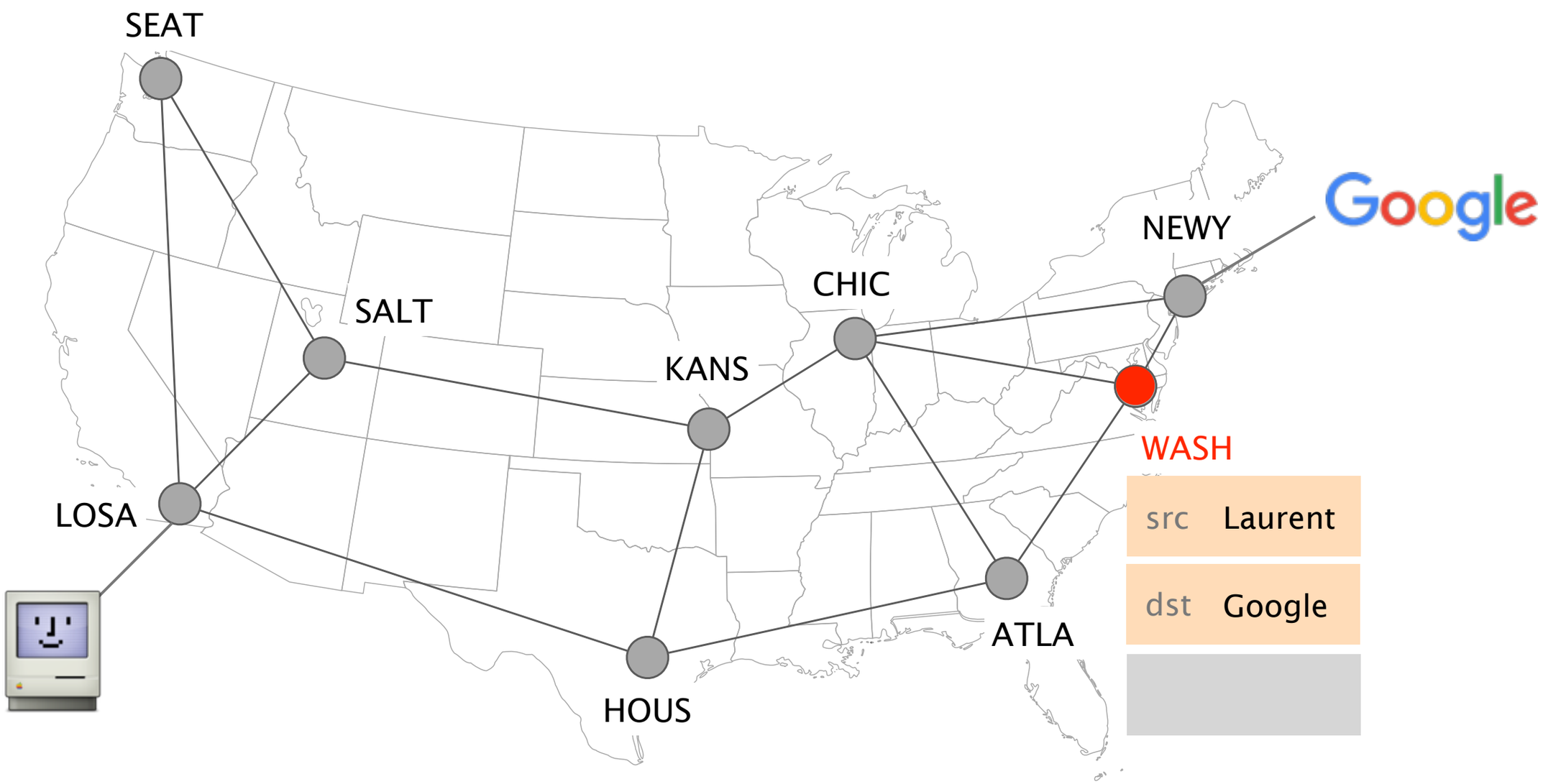
Routers forward IP packets hop-by-hop towards their destination

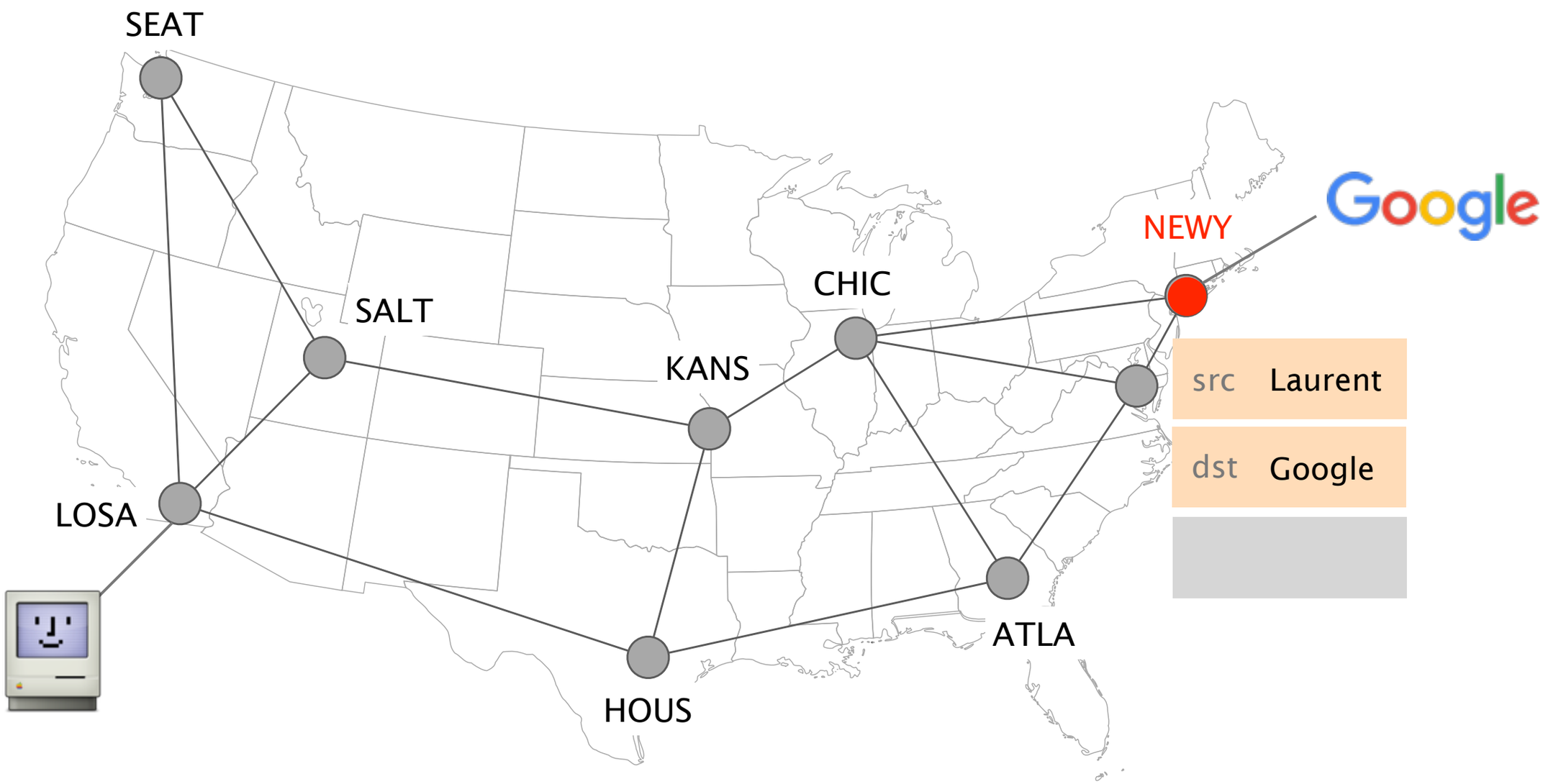


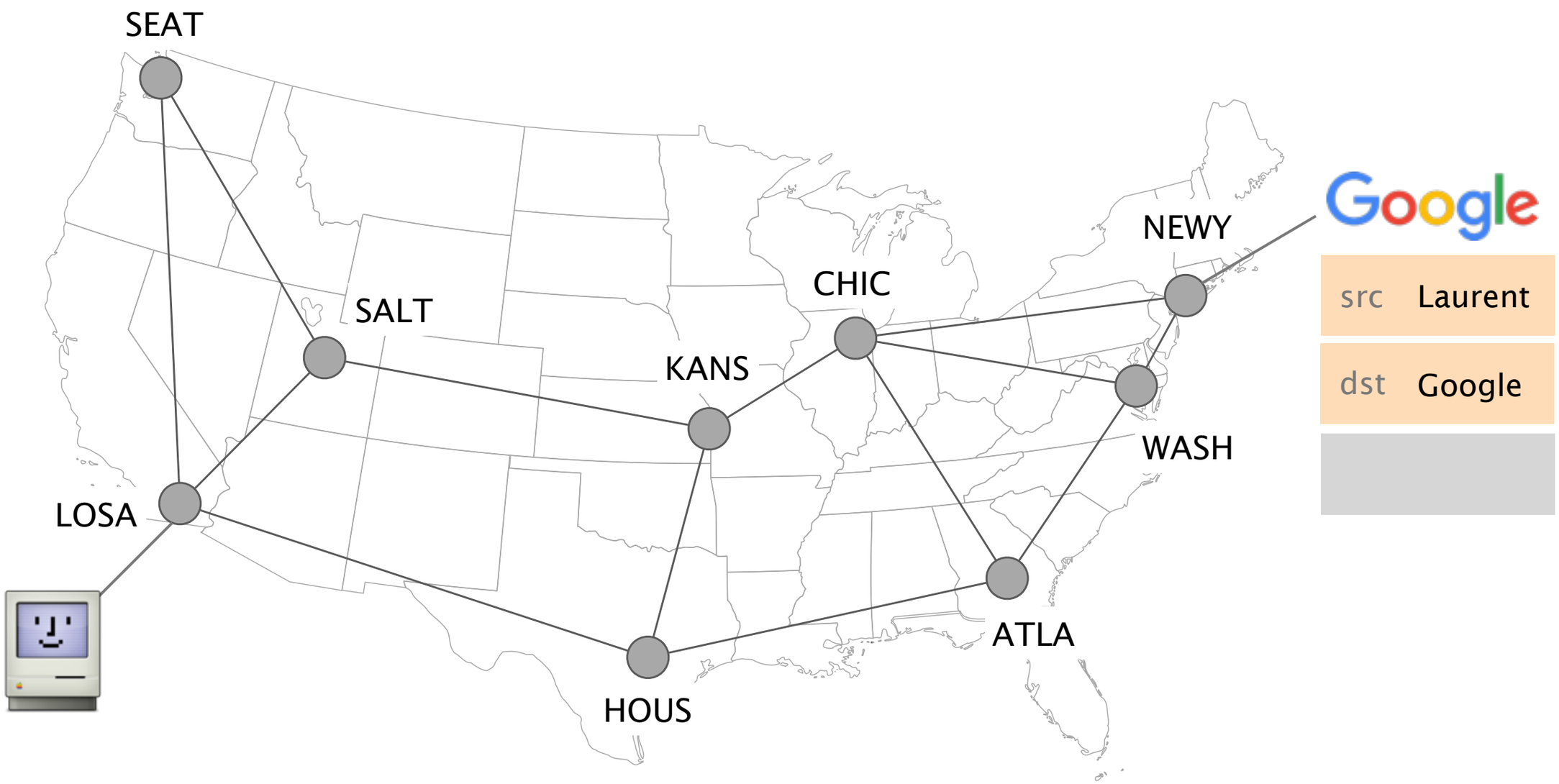

SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
ATLA
HOUS
Google
src Laurent
dst Google


SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
ATLA
HOUS
Google
src Laurent
dst Google

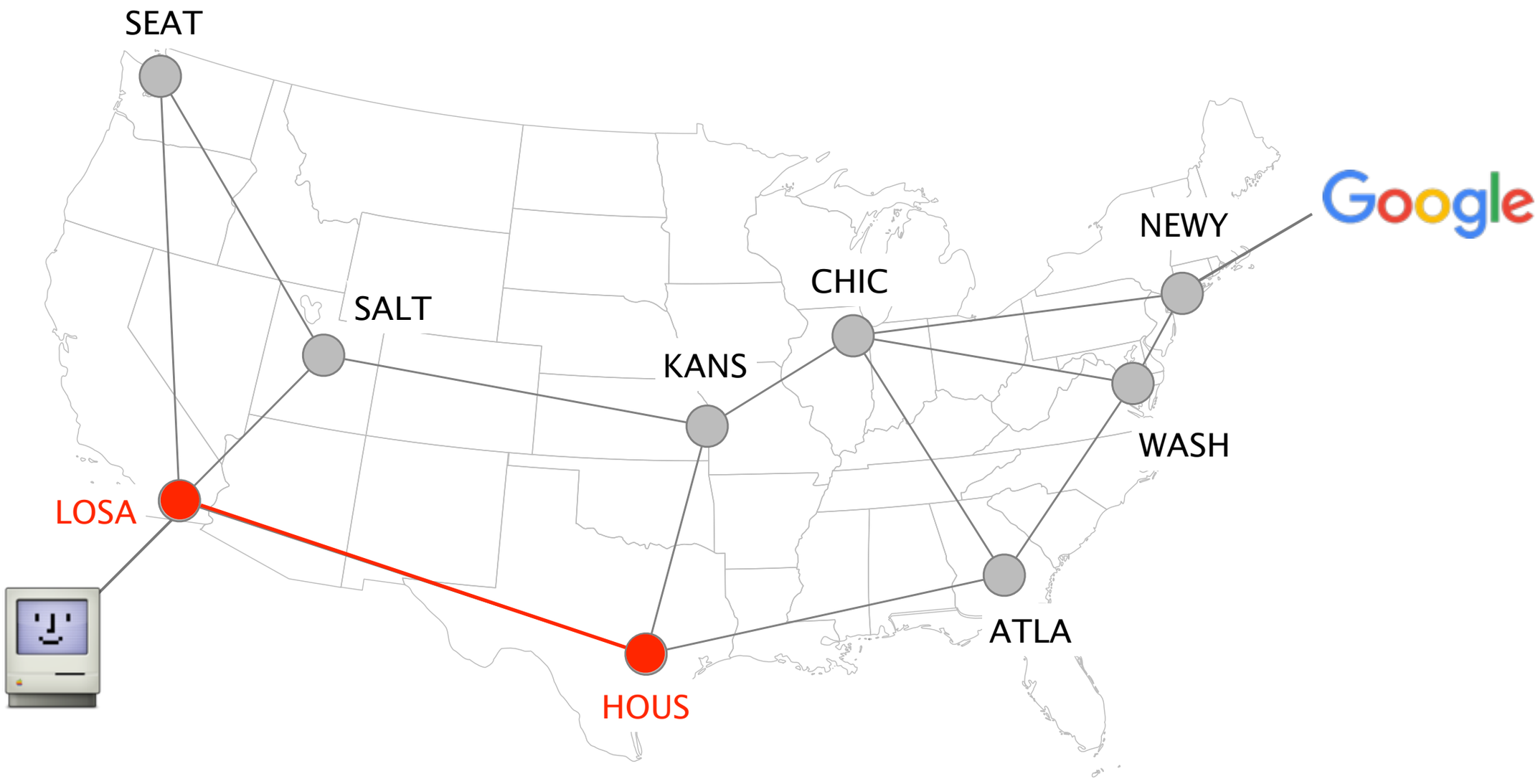

SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
HOUS
ATLA
Google
src Laurent
dst Google


SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
HOUS
ATLA
Google
src Laurent
dst Google


SEAT
SALT
LOSA
KANS
CHIC
NEWY
ATLA
HOUS
Google
src Laurent
dst Google

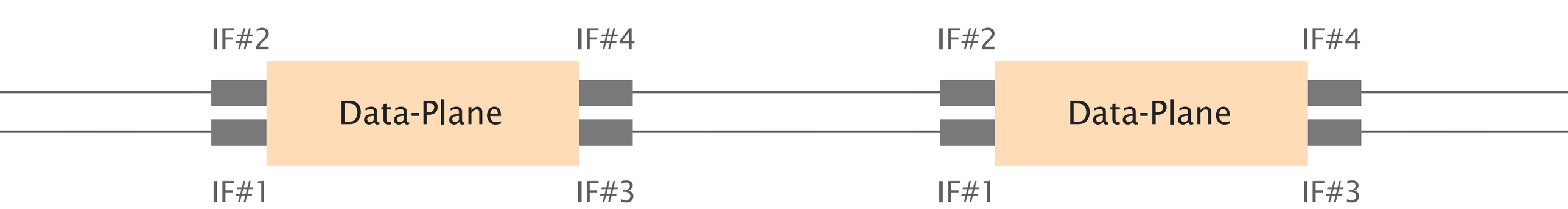

SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
ATLA
HOUS
Google
src Laurent
dst Google

Let’s zoom in on what going on between two routers

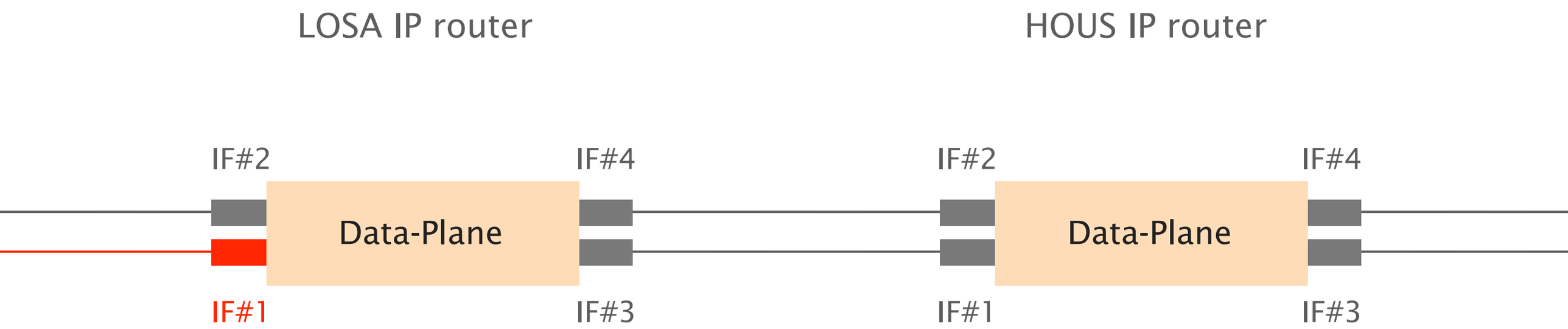


LOSA IP router

HOUS IP router



The forwarding table maps destinations (IP addresses) to output ports



Forwarding table

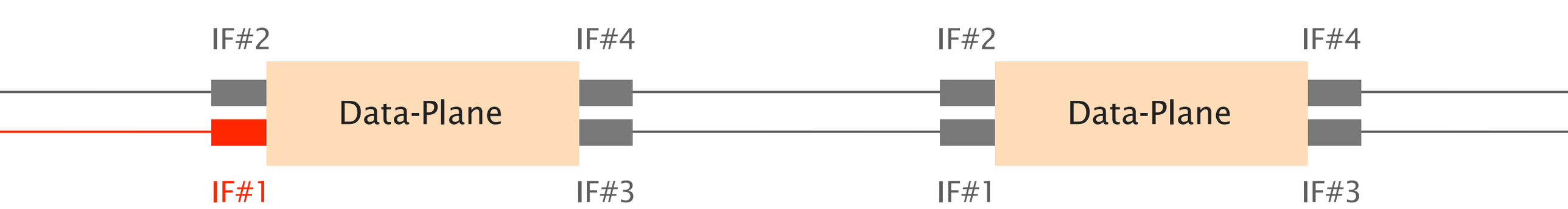
destination	output
Laurent	IF#1
Google	IF#4

src Laurent

dst Google

LOSA IP router

HOUS IP router



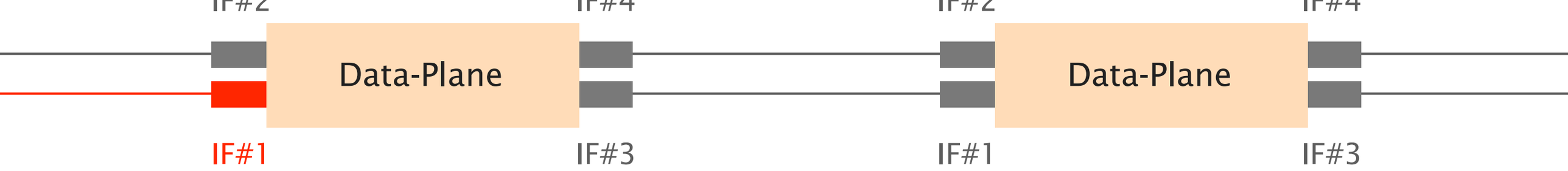
Forwarding table

src	Laurent
dst	Google

destination	output
Laurent	IF#1
Google	IF#4

LOSA IP router

HOUS IP router



Forwarding table

src	Laurent
dst	Google

destination	output
Laurent	IF#1
Google	IF#4

LOSA IP router

HOUS IP router

IF#2

IF#4

IF#2

IF#4

Data-Plane

Data-Plane

IF#1

IF#3

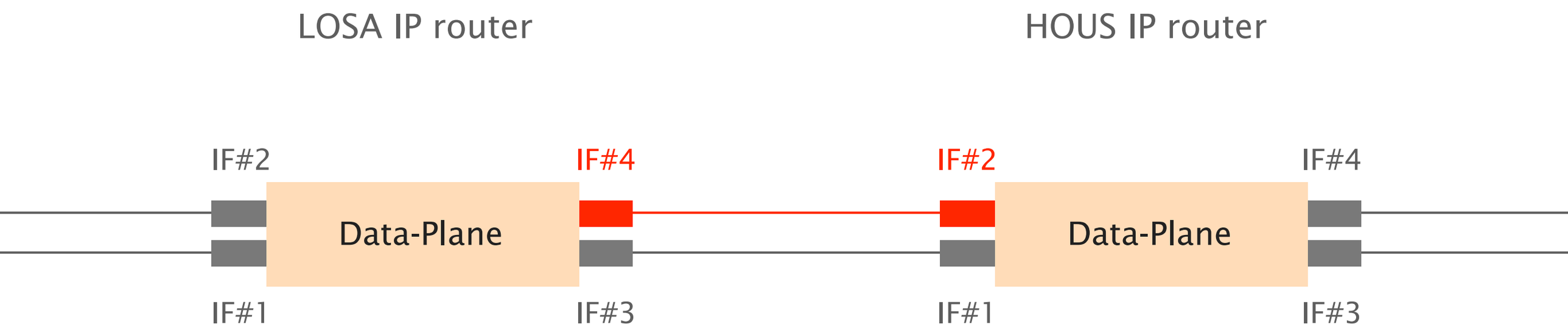
IF#1

IF#3

src Laurent

dst Google

This processes is repeated
at each router, until the destination is reached



src Laurent

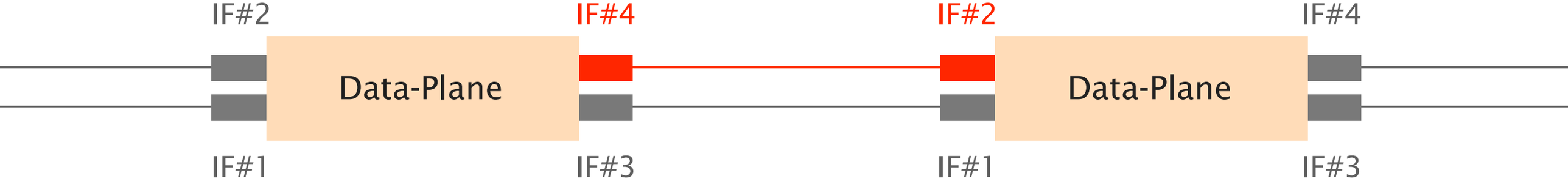
dst Google

Forwarding table

destination	output
Laurent	IF#1
Google	IF#3

LOSA IP router

HOUS IP router



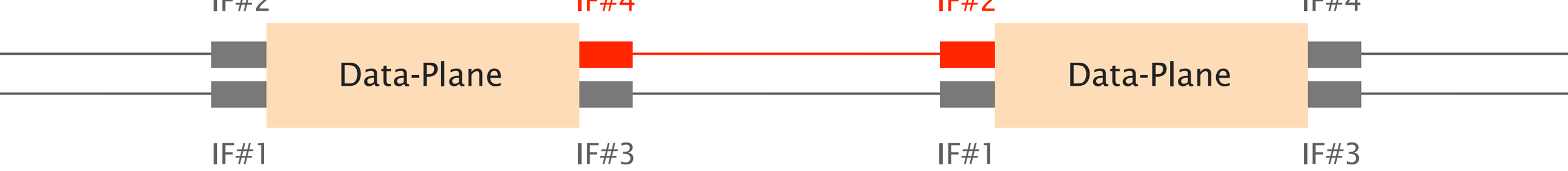
src	Laurent
dst	Google

Forwarding table

destination	output
Laurent	IF#1
Google	IF#3

LOSA IP router

HOUS IP router



src	Laurent
dst	Google

Forwarding table

destination	output
Laurent	IF#1
Google	IF#3

LOSA IP router

HOUS IP router

IF#2

IF#4

IF#2

IF#4

Data-Plane

Data-Plane

IF#1

IF#3

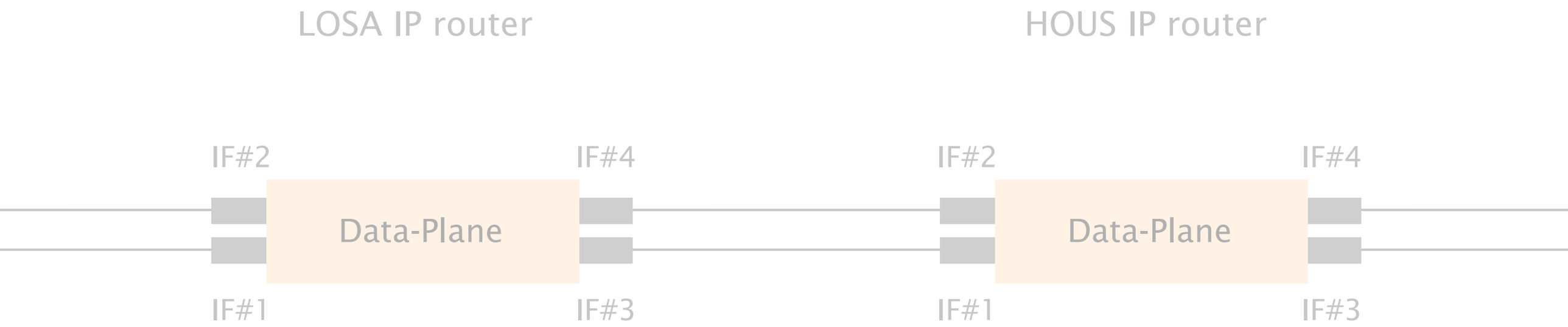
IF#1

IF#3

src Laurent

dst Google

How are these forwarding tables computed and provisioned?



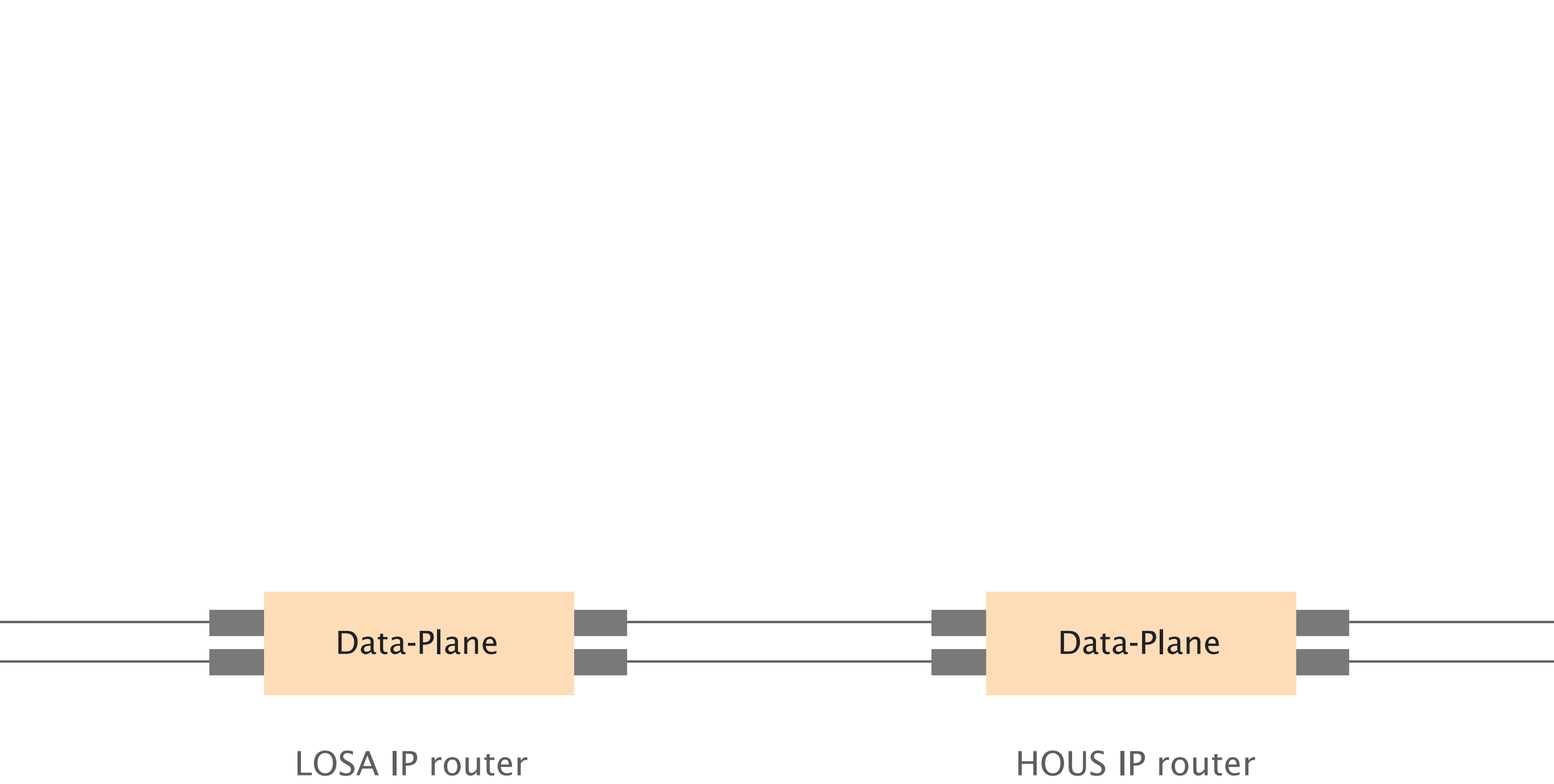
Forwarding table

destination	output
Laurent	IF#1
Google	IF#4

Forwarding table

destination	output
Laurent	IF#1
Google	IF#3

Traditionally, by distributed protocols running on each routers



LOSA IP router

HOUS IP router

Control-Plane

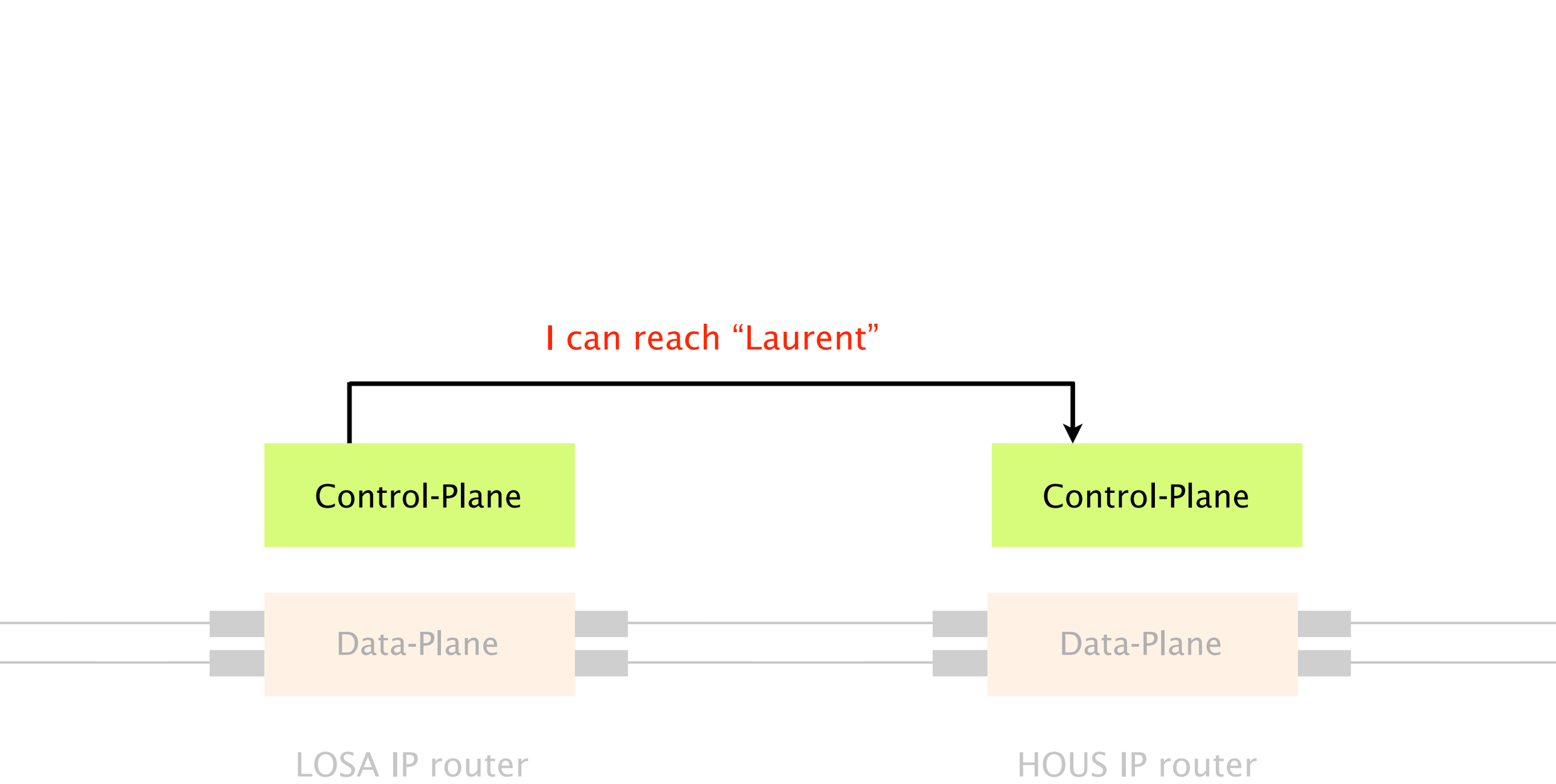
Control-Plane

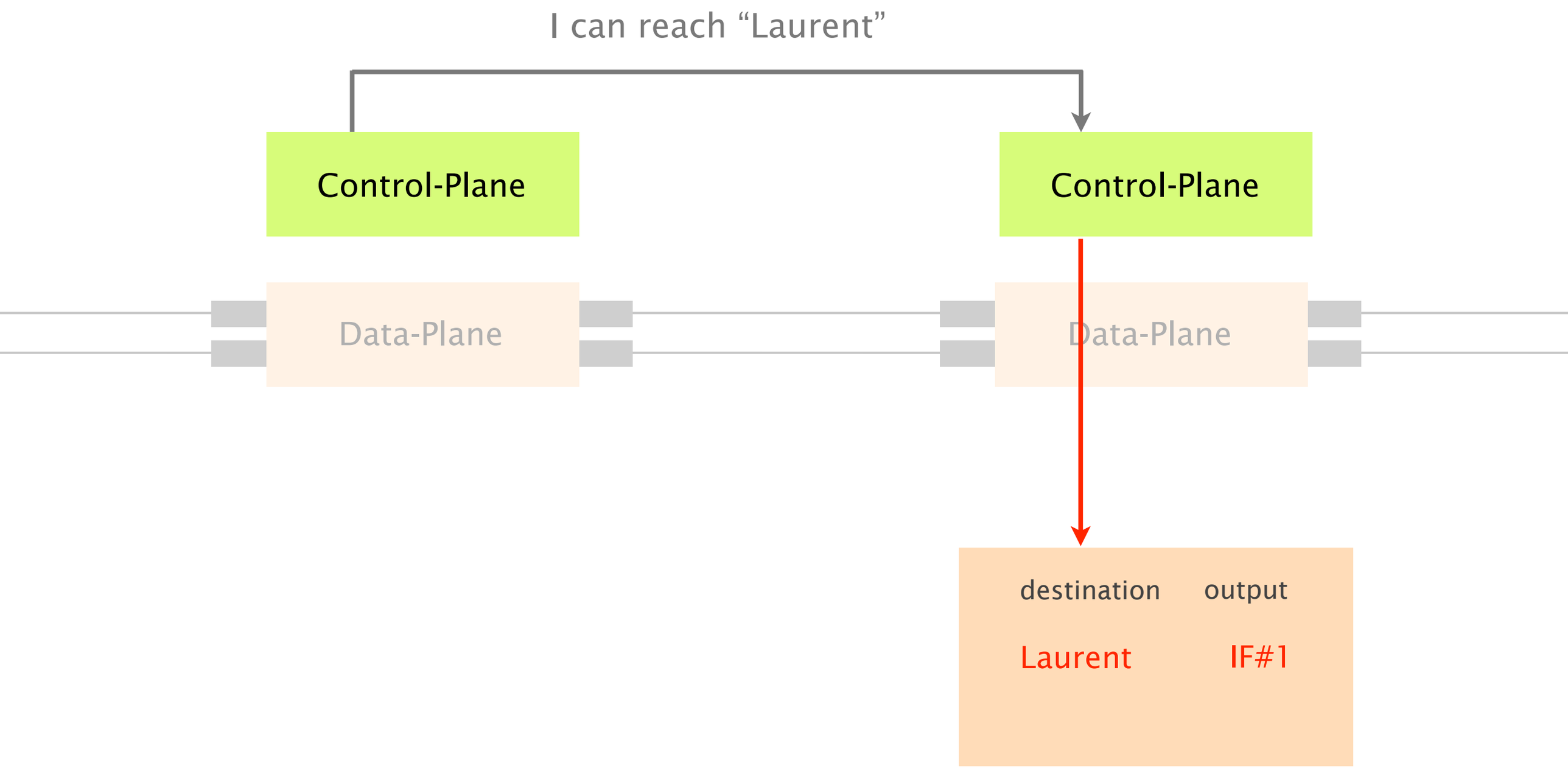
Data-Plane

Data-Plane

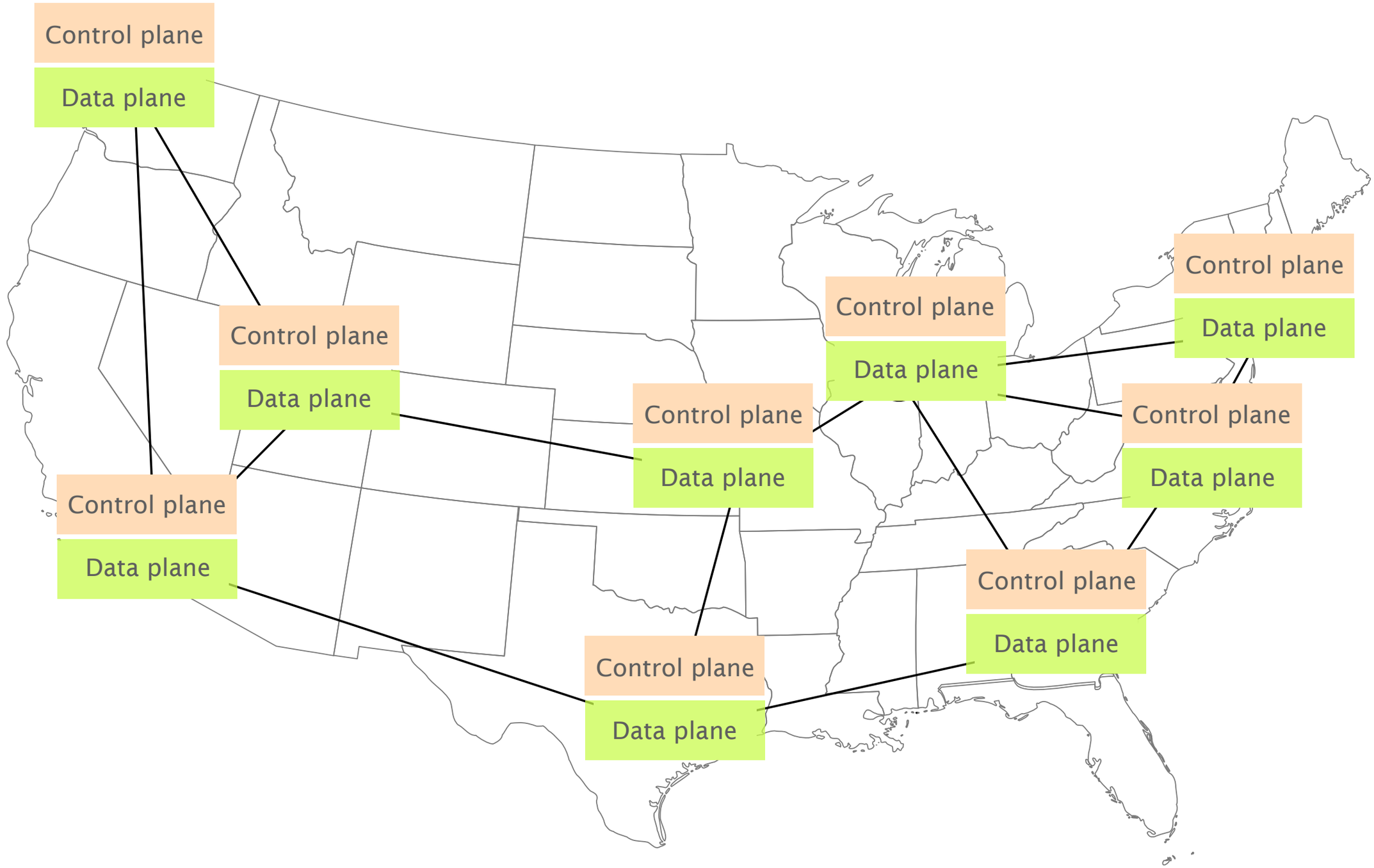
LOSA IP router

HOUS IP router


I can reach “Laurent”
Control-Plane
Control-Plane
Data-Plane
Data-Plane
LOSA IP router
HOUS IP router


I can reach “Laurent”
Control-Plane
Control-Plane
Data-Plane
Data-Plane
destination
output
Laurent
IF#1

Network operators adapt their forwarding table
by configuring the control-plane of each router

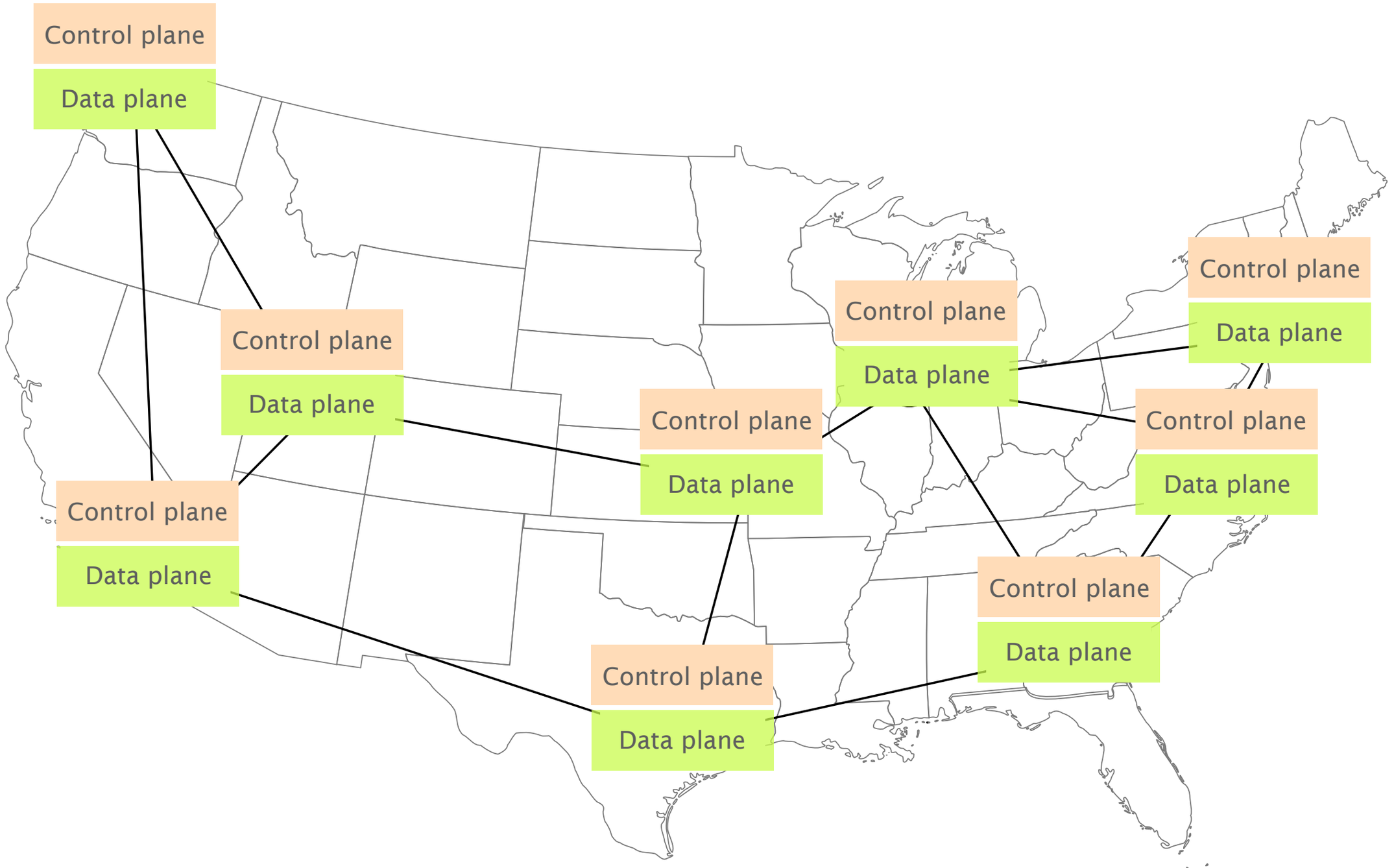


Configuring each router is often done manually,
using arcane low-level, vendor-specific “languages”

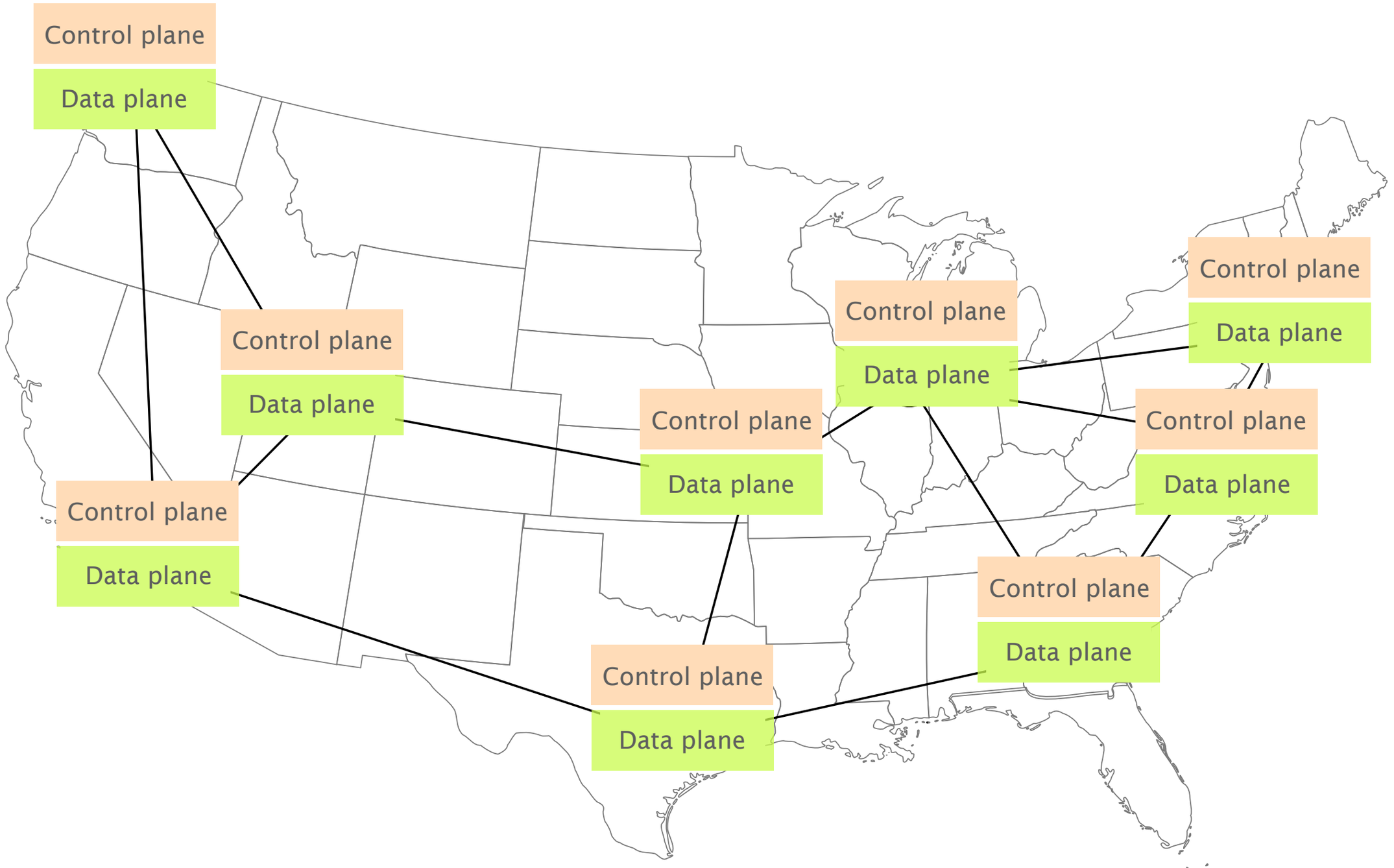
Software-Defined Network

enable network programmability

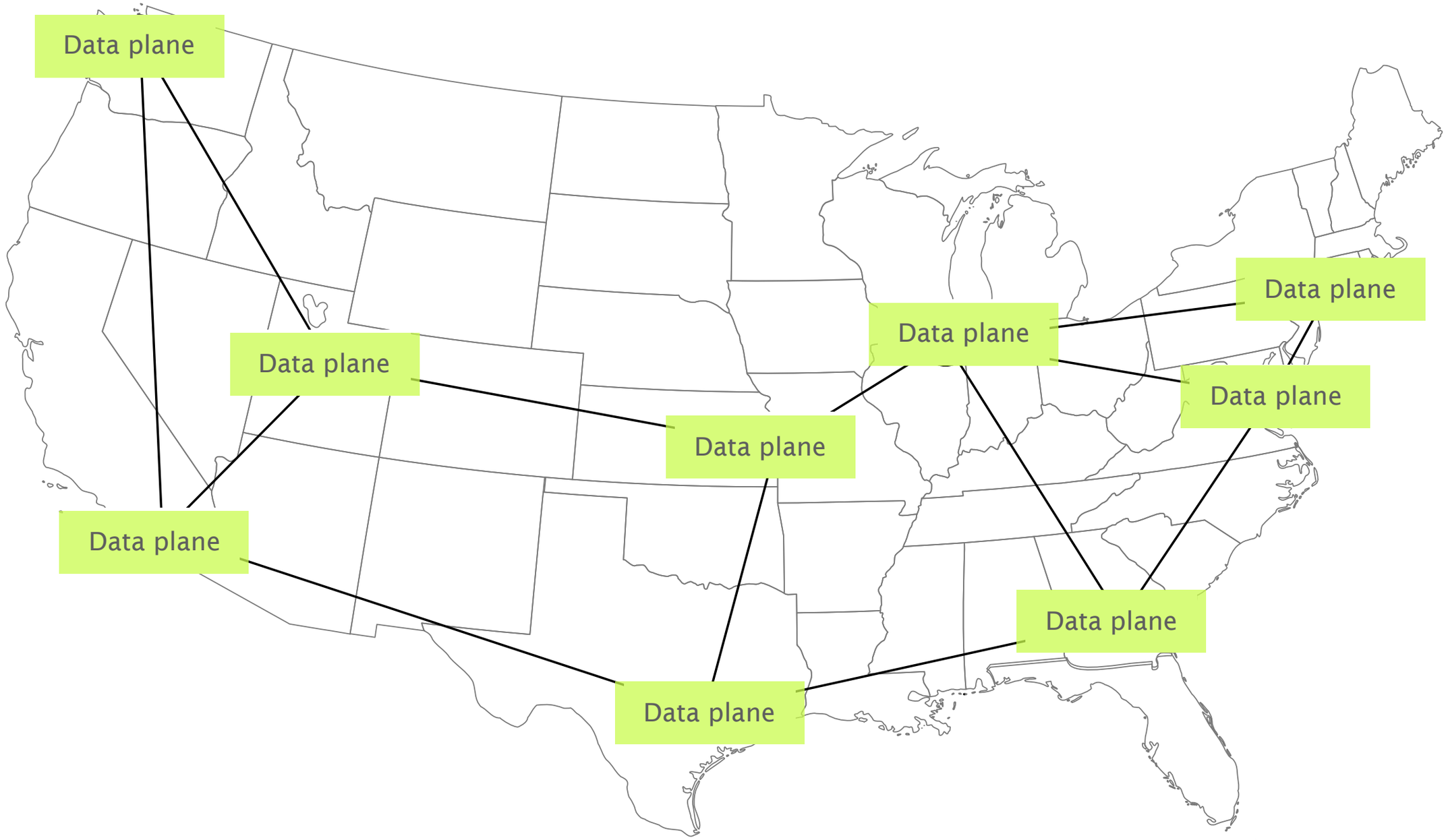
In contrast, SDN simplifies the control of forwarding entries



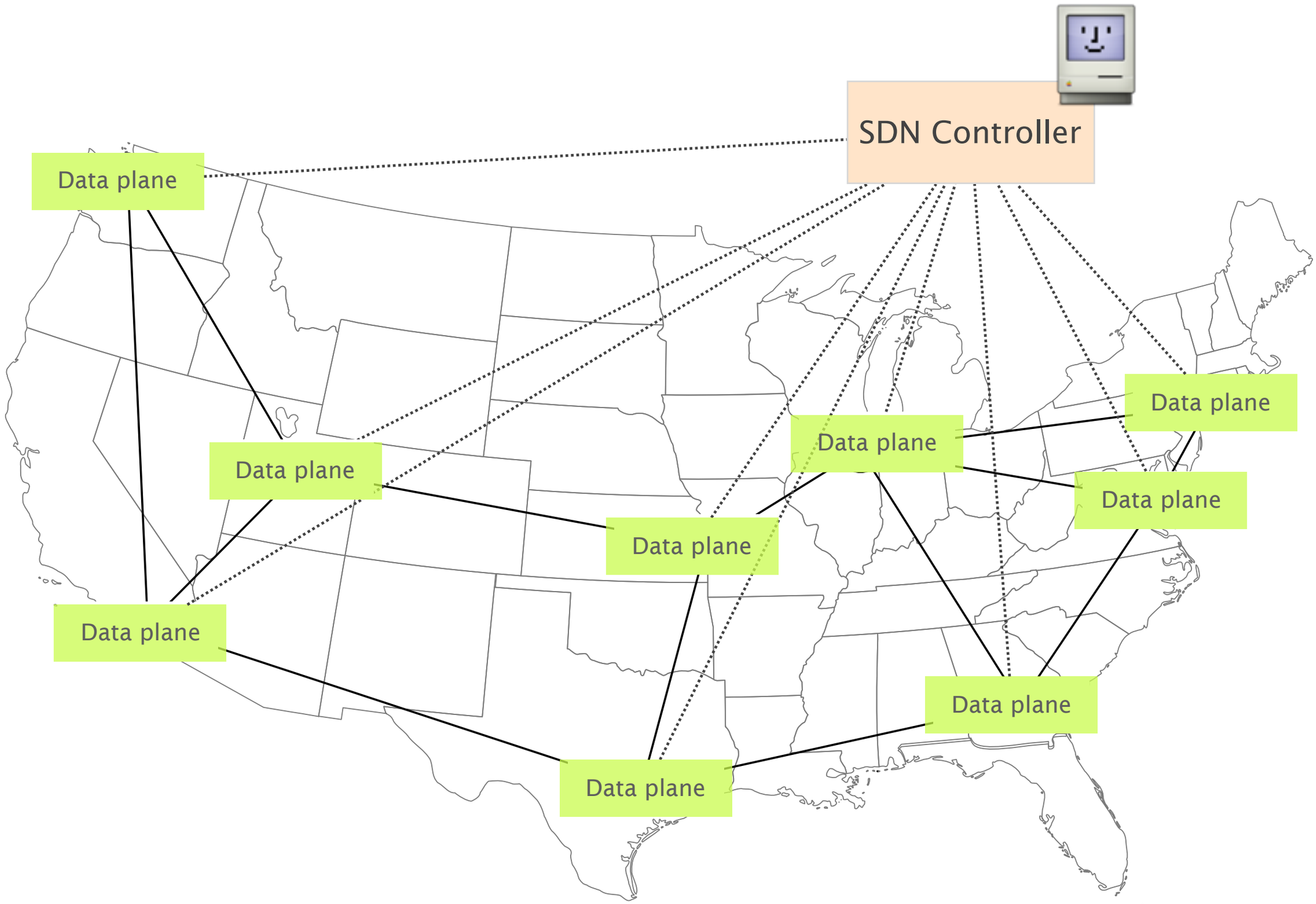
… by removing the intelligence from the equipments



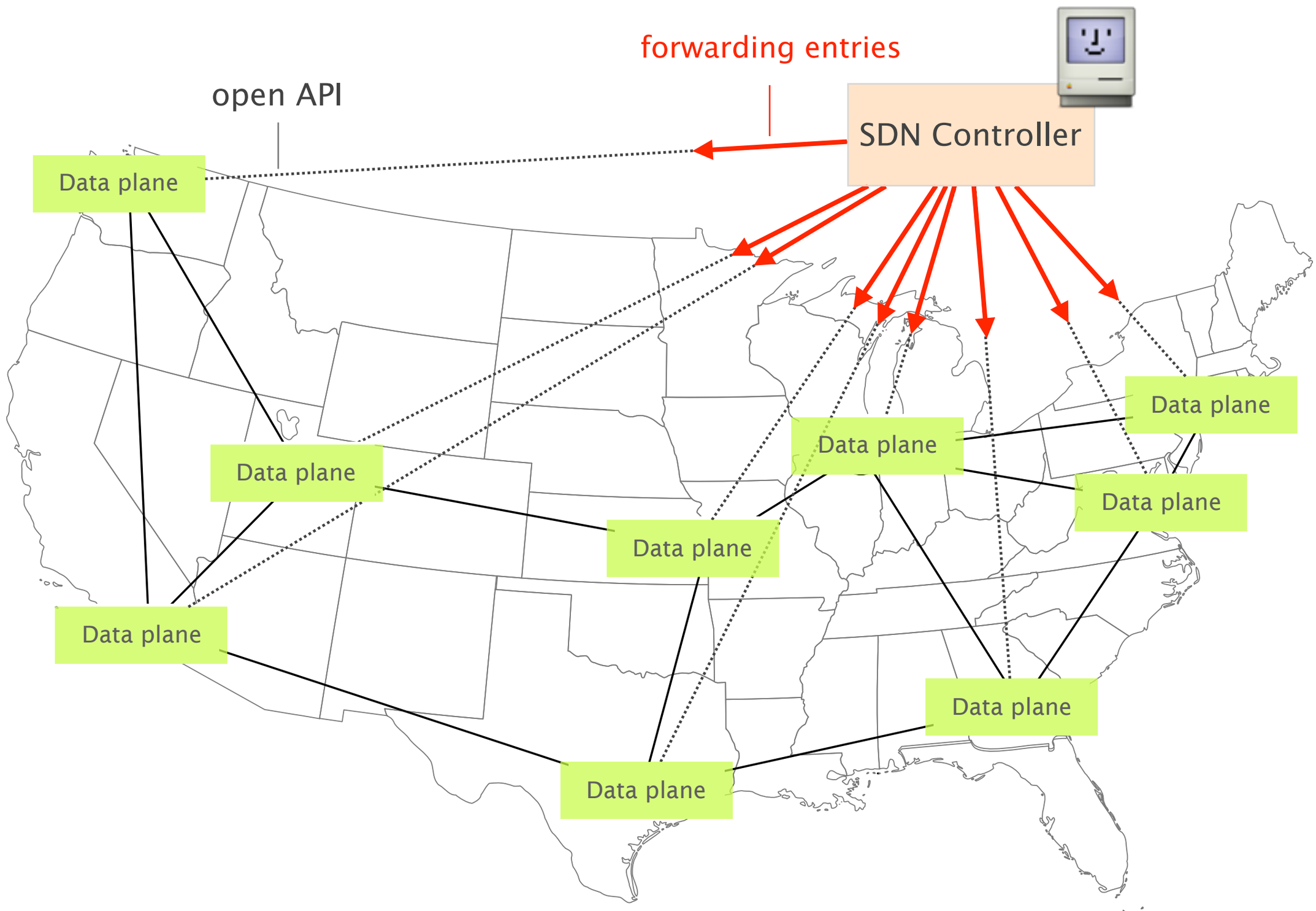
… by removing the intelligence from the equipments



… and centralizing it in a controller
that can run arbitrary programs



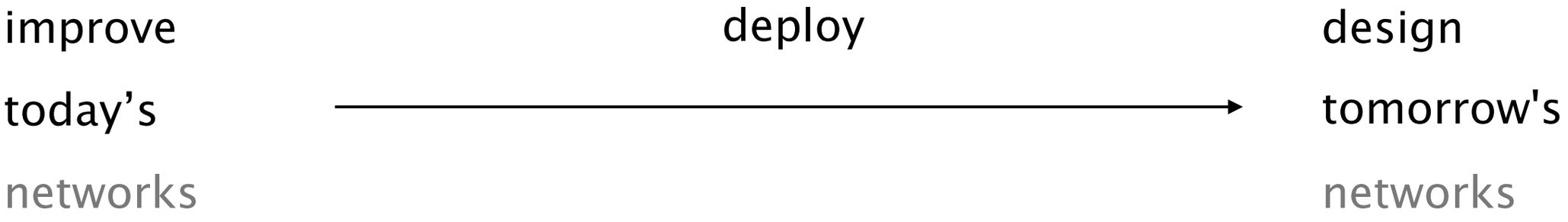
The controller programs entries in the forwarding tables using an open interface



Our focus

these days

Leverage network programability to…



Our focus

these days

Leverage network programability to…

improve

today's

networks

Wouldn’t it be great to program existing network instead of configuring them?

Fibbing

Joint work with: [SIGCOMM'15]

Stefano Vissicchio, Olivier Tilmans and Jennifer Rexford

Fibbing

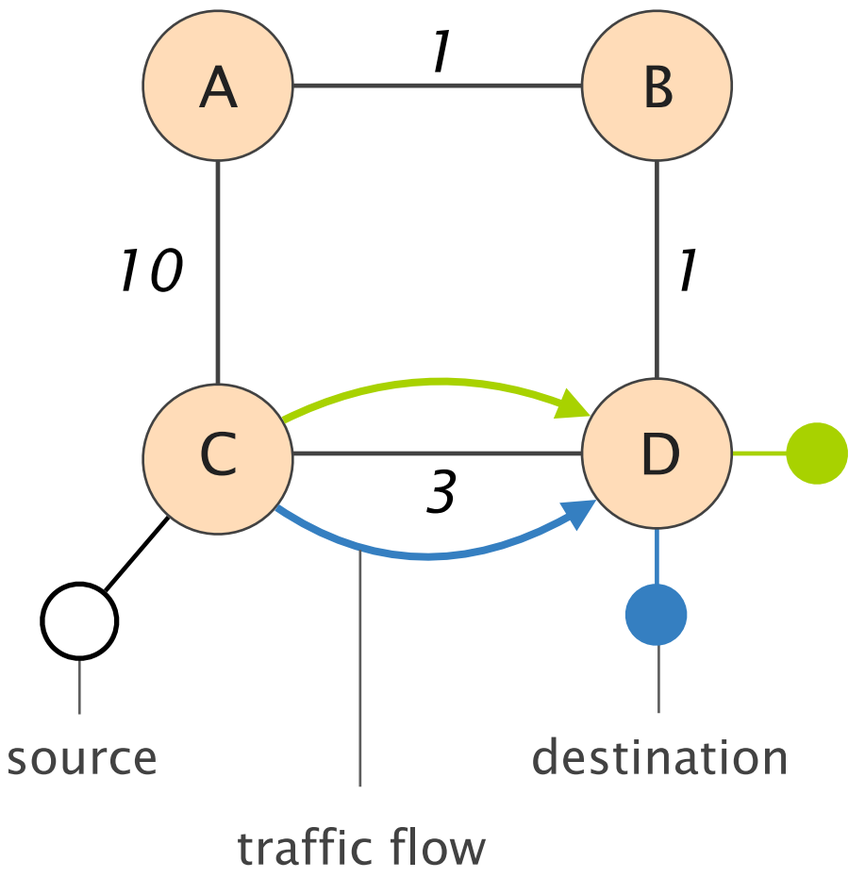
= lying

Fibbing

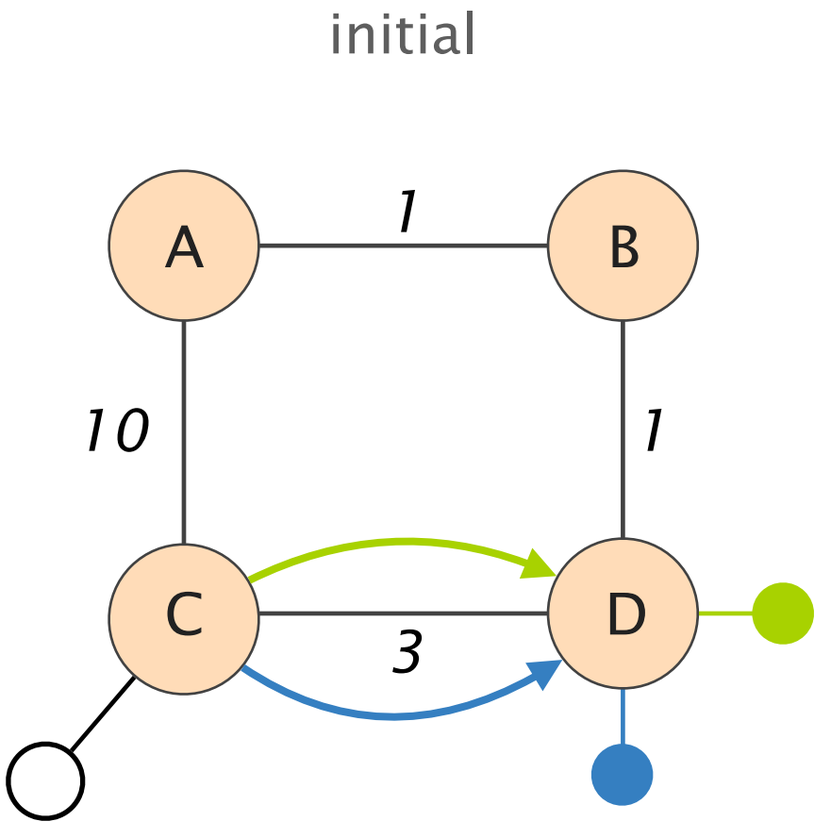
to **control** router's forwarding table

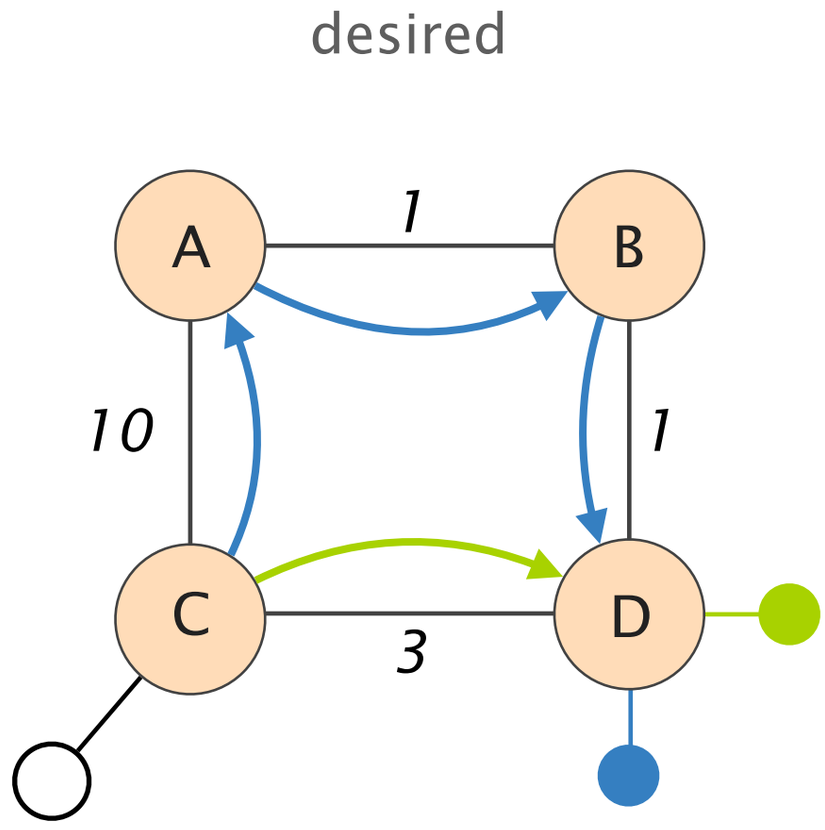
How can lying in a network help?!

Consider this network where a source sends traffic to 2 destinations



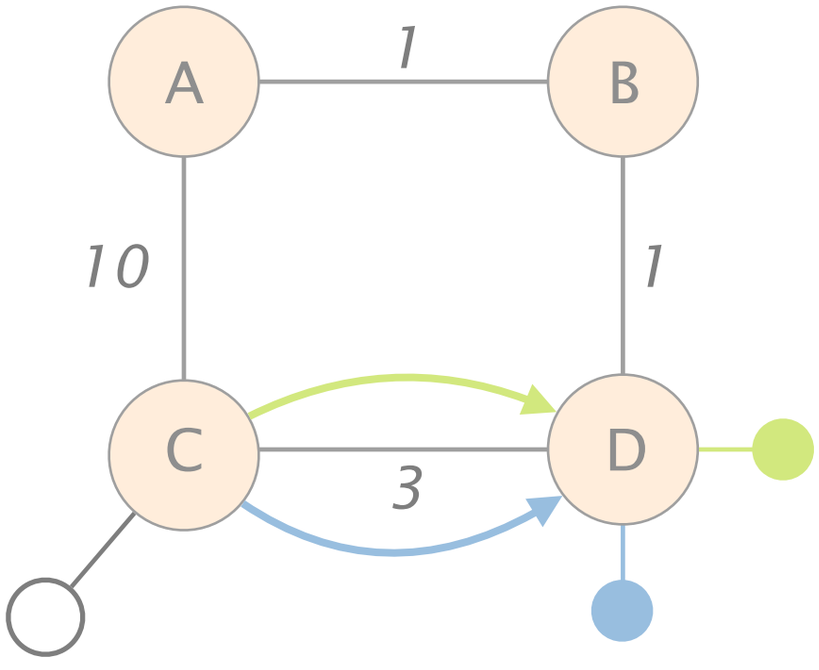
As congestion appears, the operator wants to shift away one flow from (C,D)



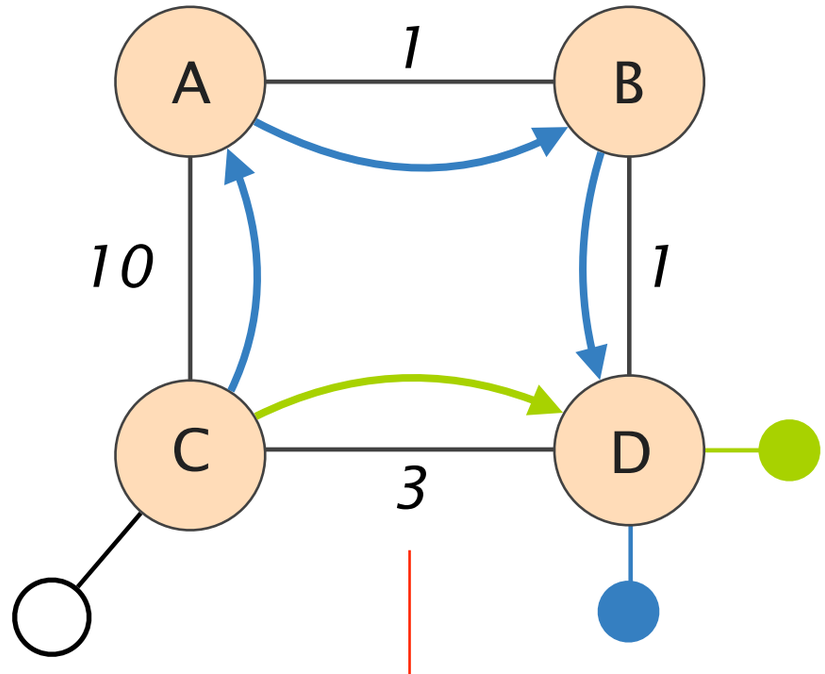


Moving only one flow is impossible though as both destinations are connected to D

initial

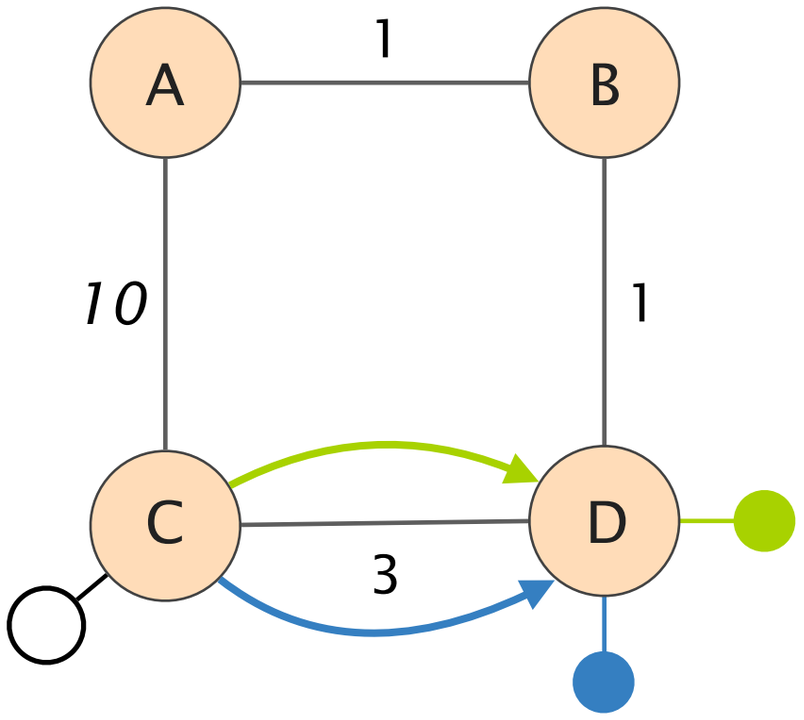


desired

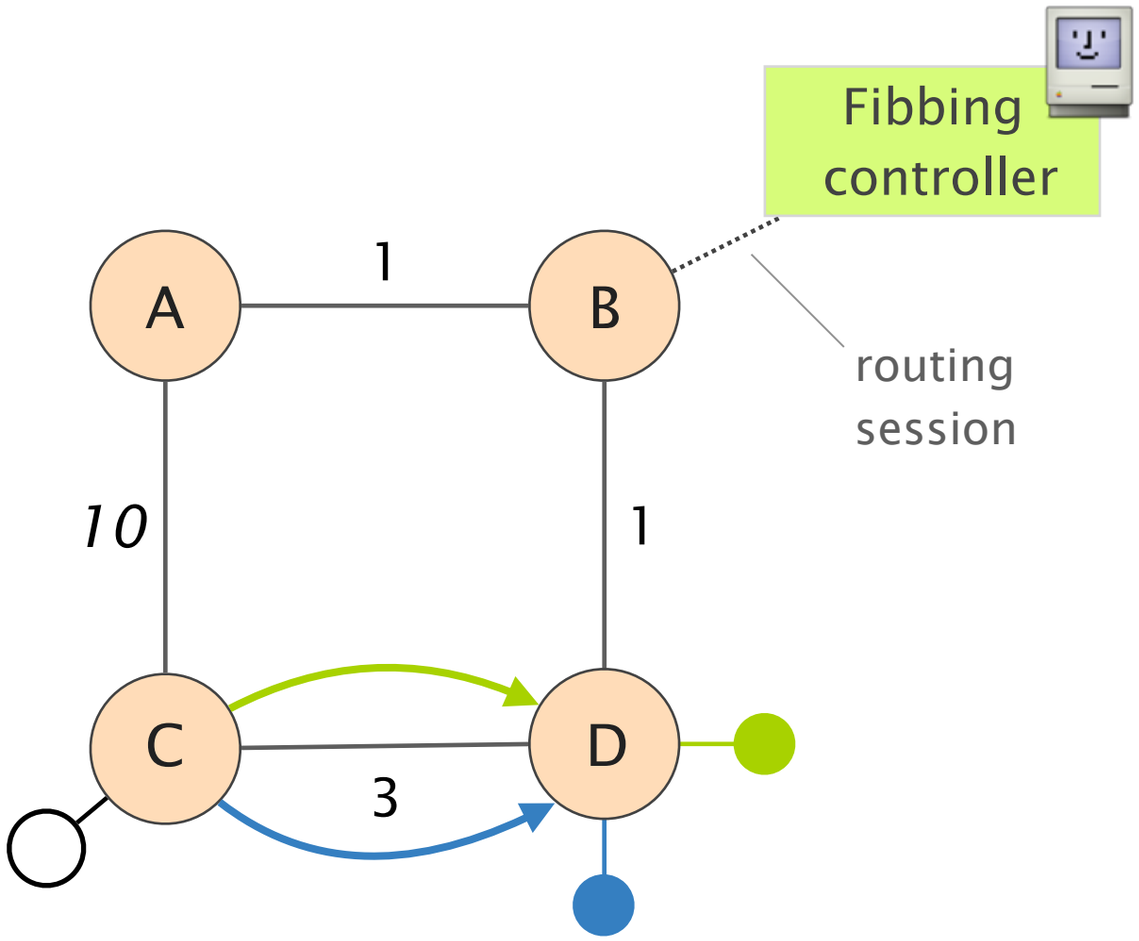


impossible to achieve by reweighing the links

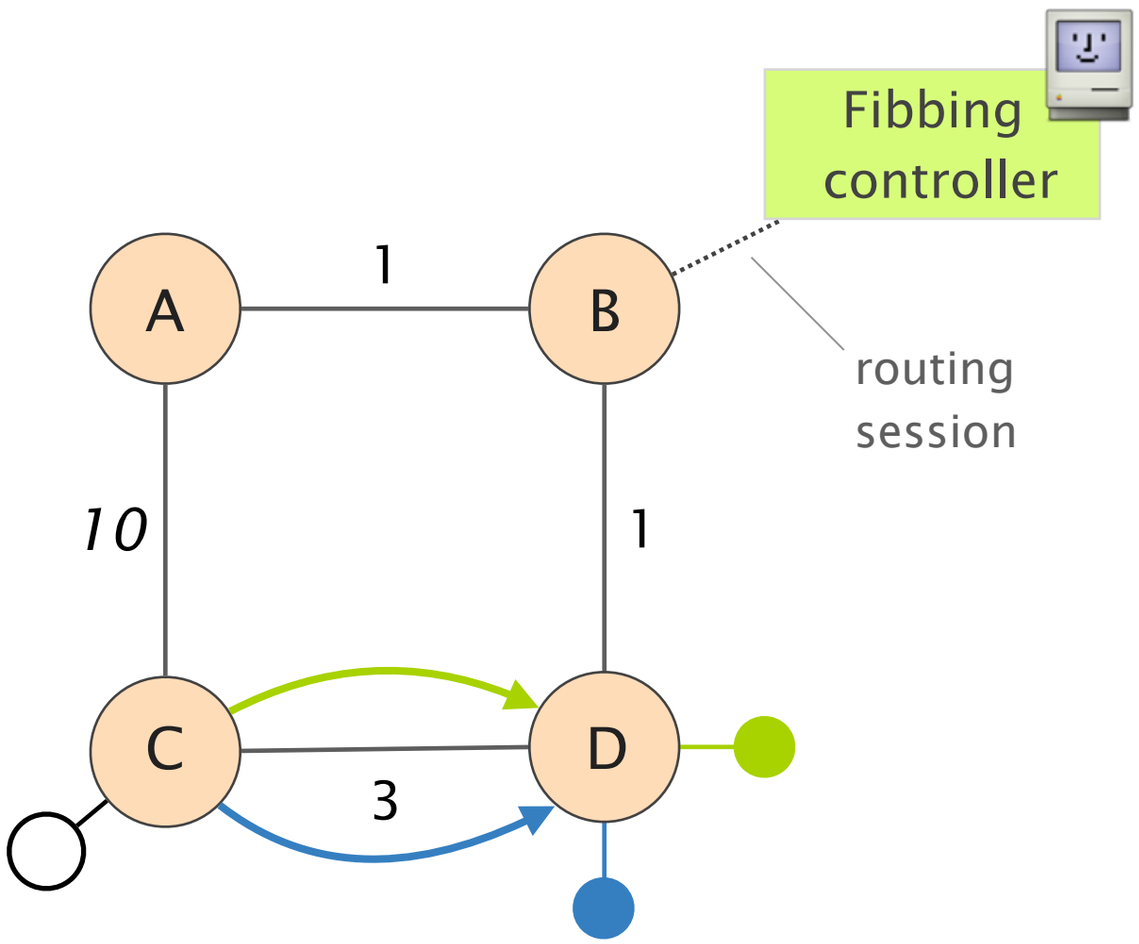
Let’s lie to the router



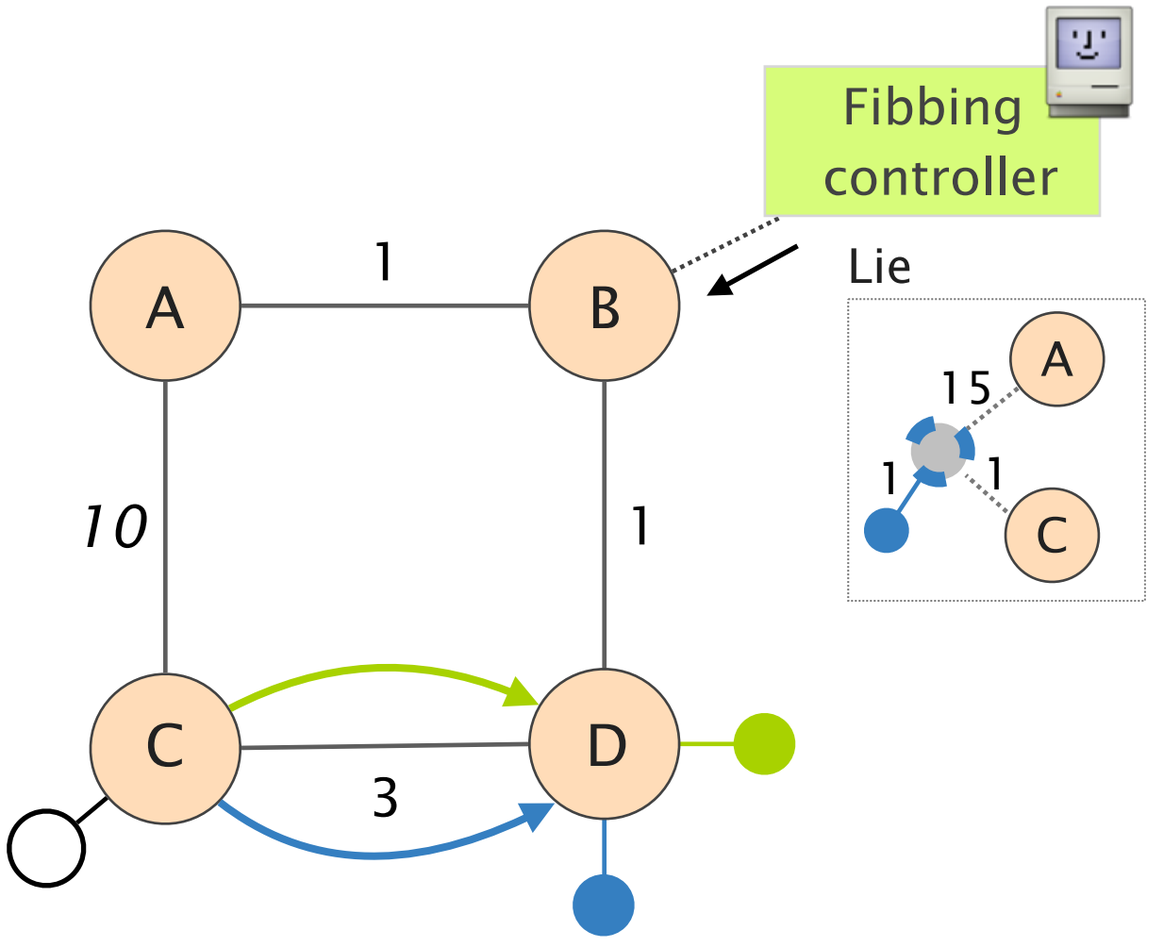
Let's lie to the router



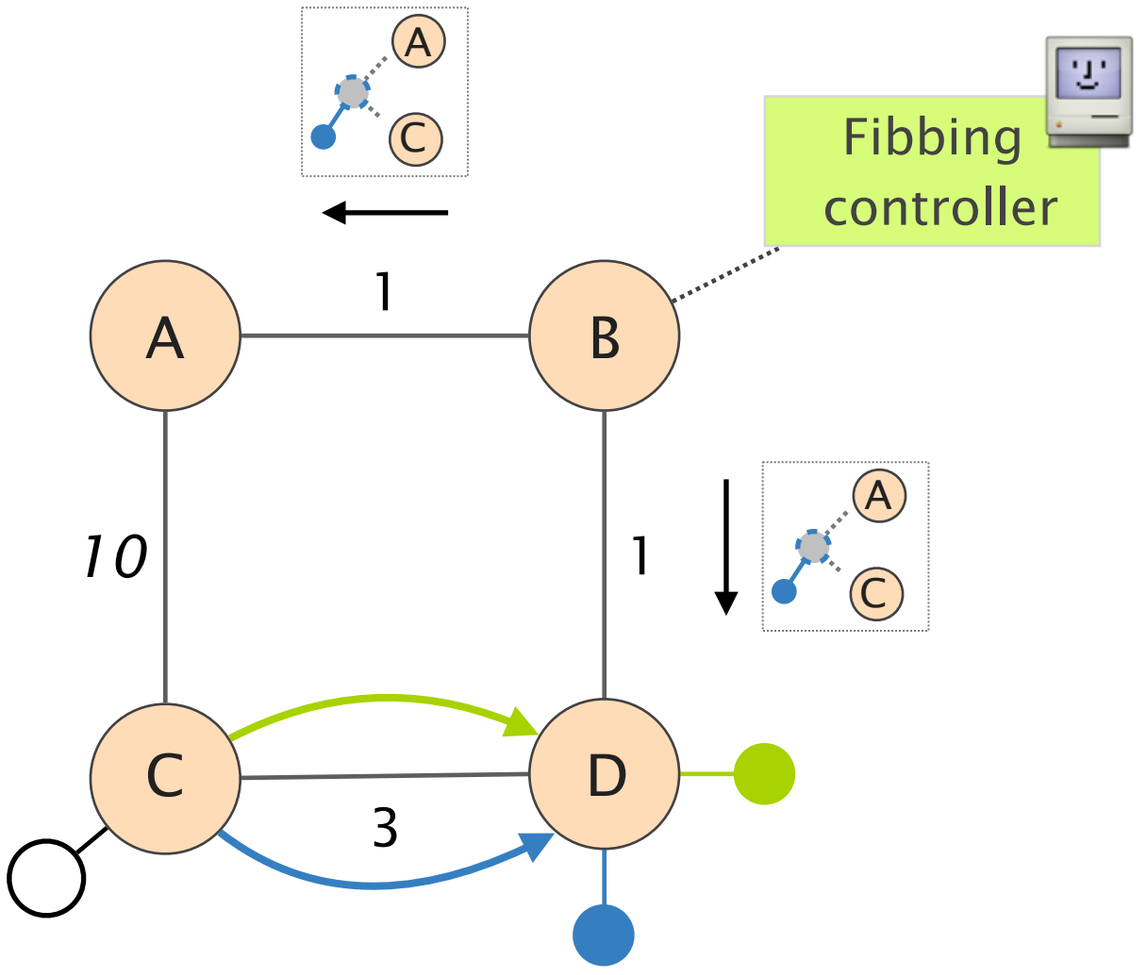
Let's **lie to the router,** by injecting fake nodes, links and destinations



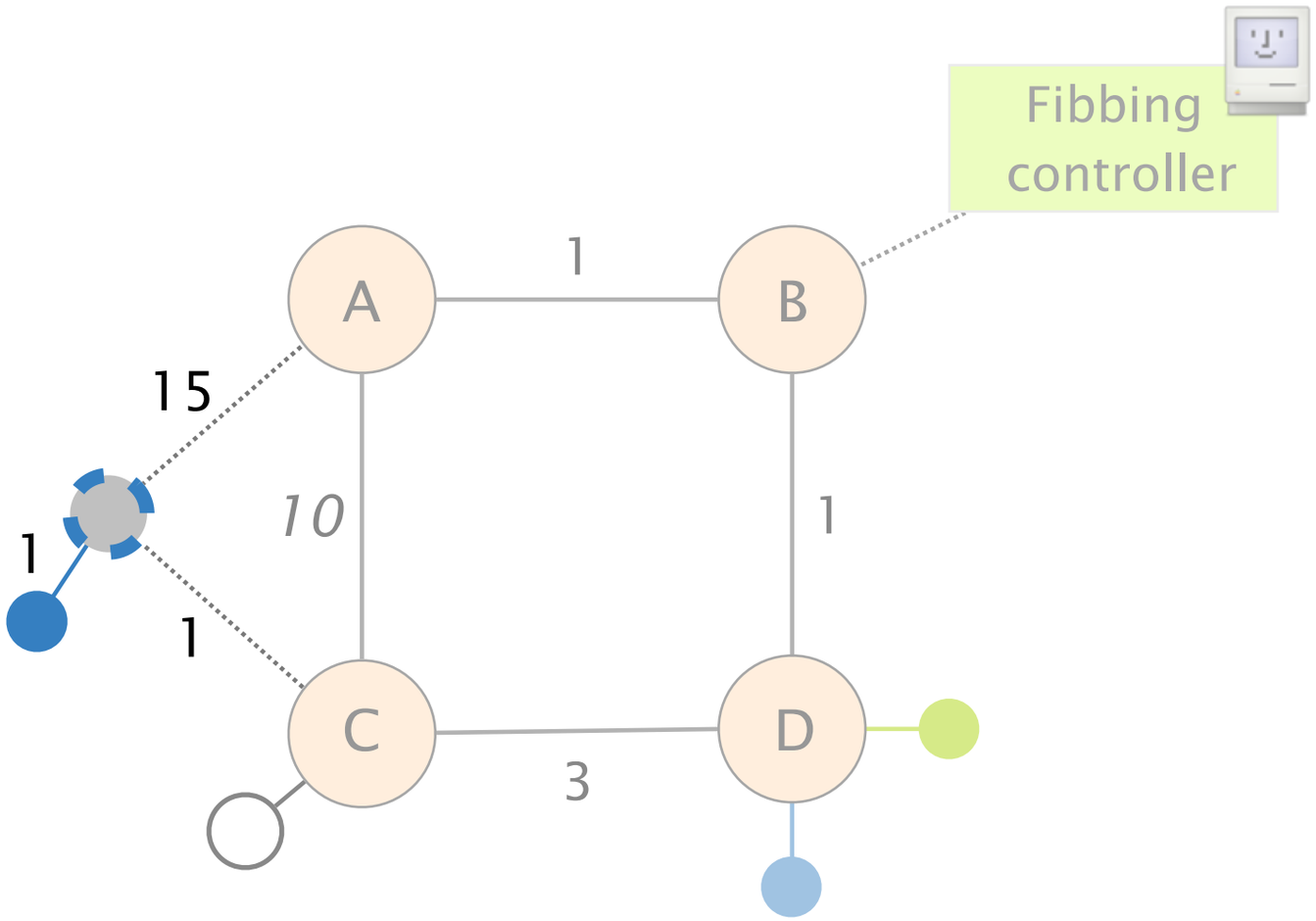
Let's lie to the router, by injecting fake nodes, links and destinations



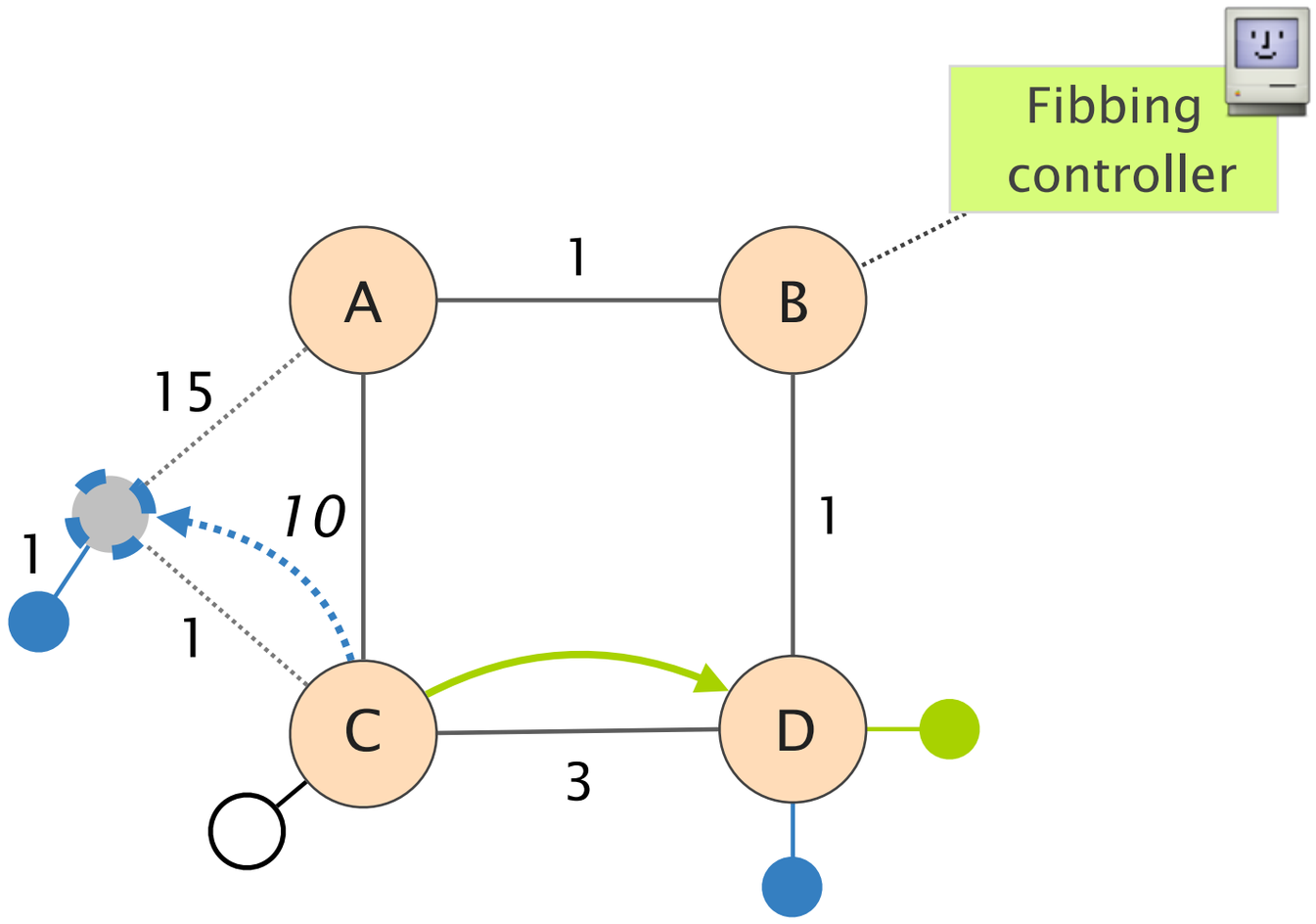
Lies are propagated network-wide by the protocol



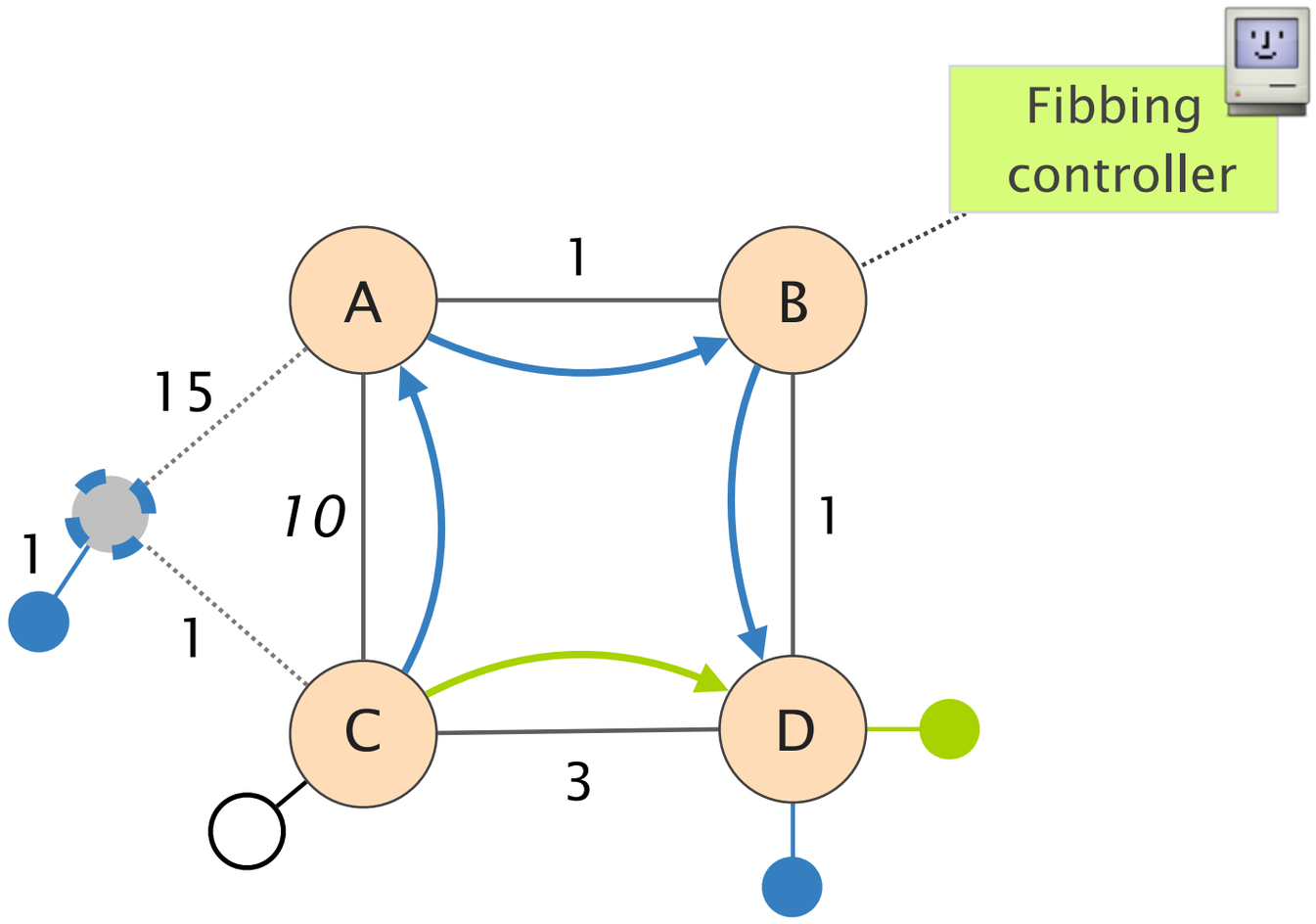
Now all routers see this topology
on which they compute their shortest-paths



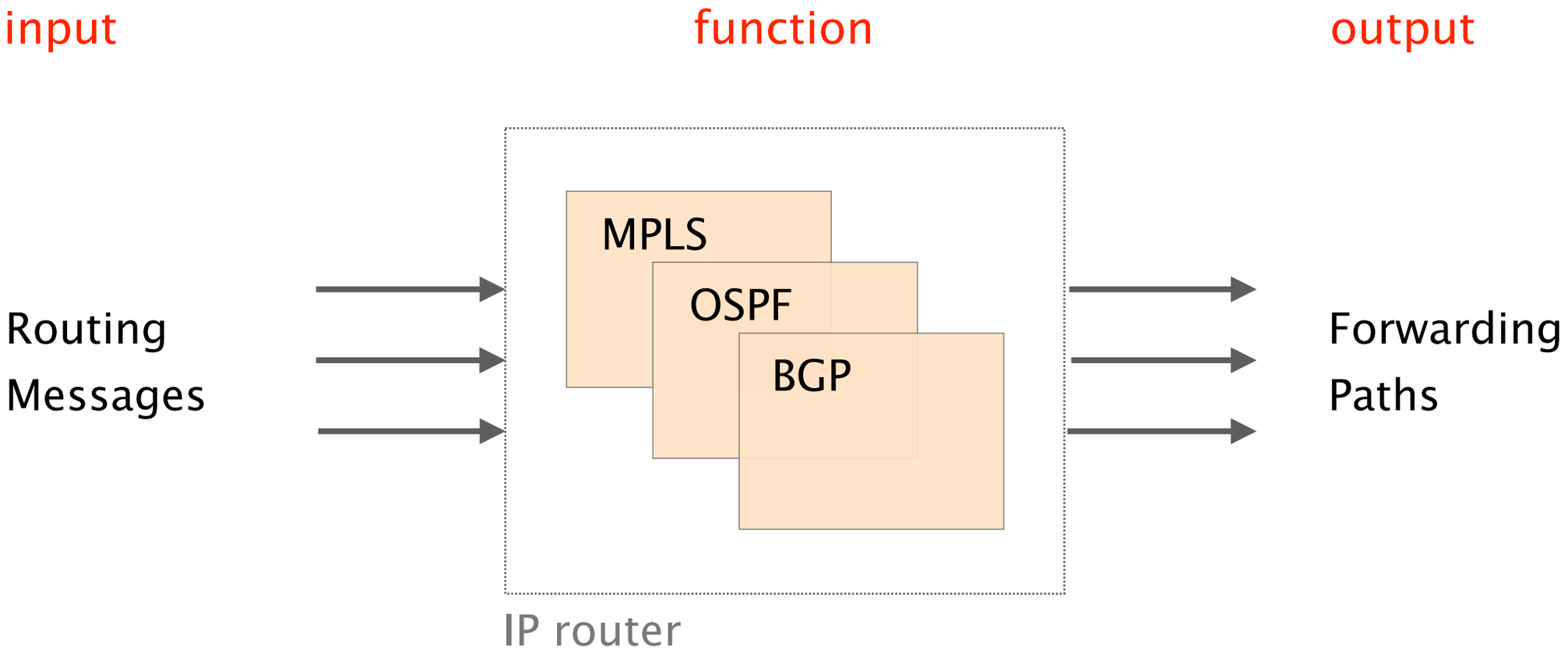
Now, C prefers the virtual node (cost 2)
to reach the blue destination…



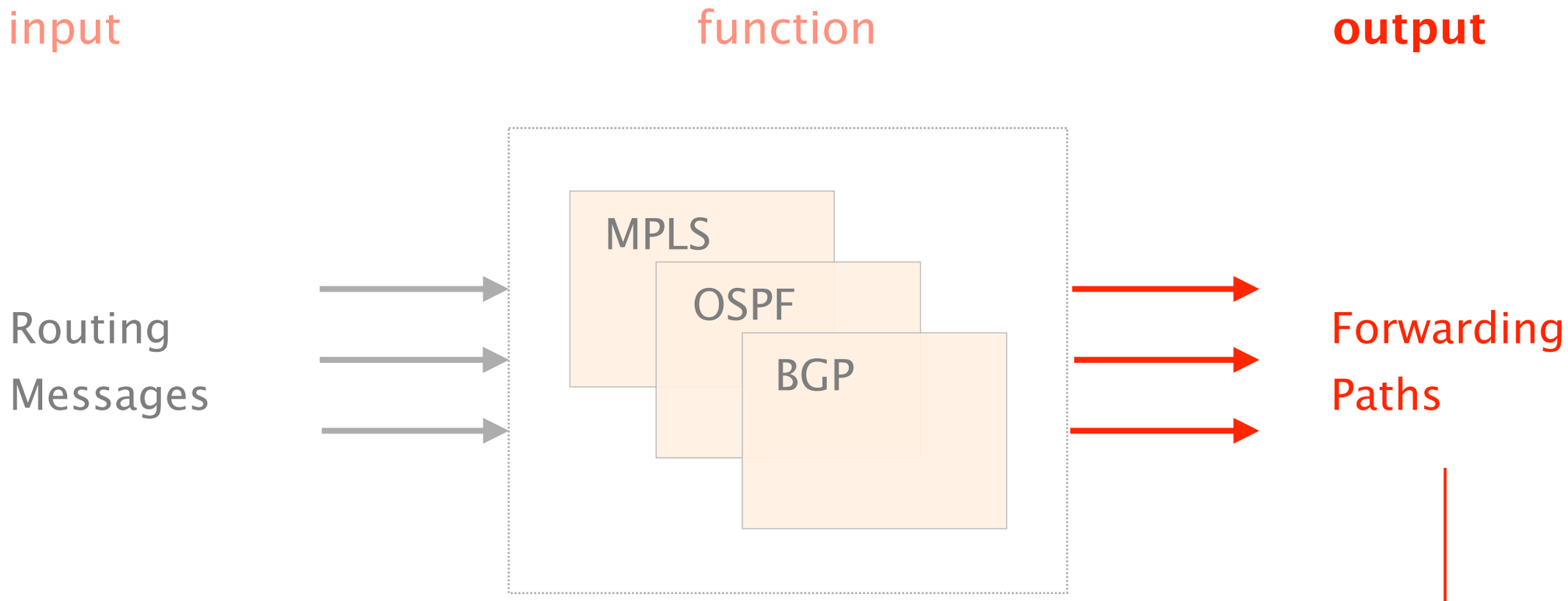
As the virtual node does not really exist,
actual traffic is *physically* sent to A



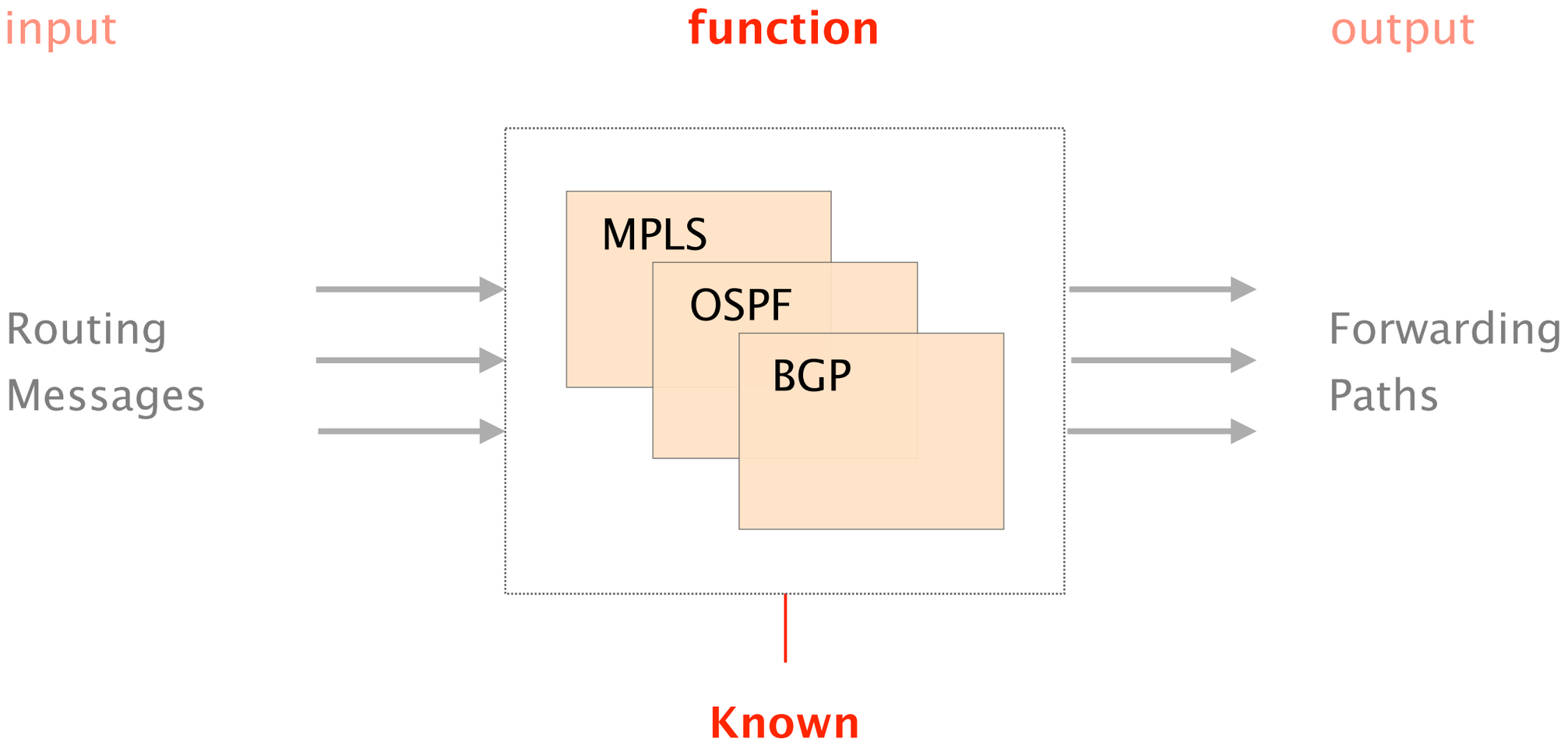
A router control-plane implements a function from routing messages to forwarding paths



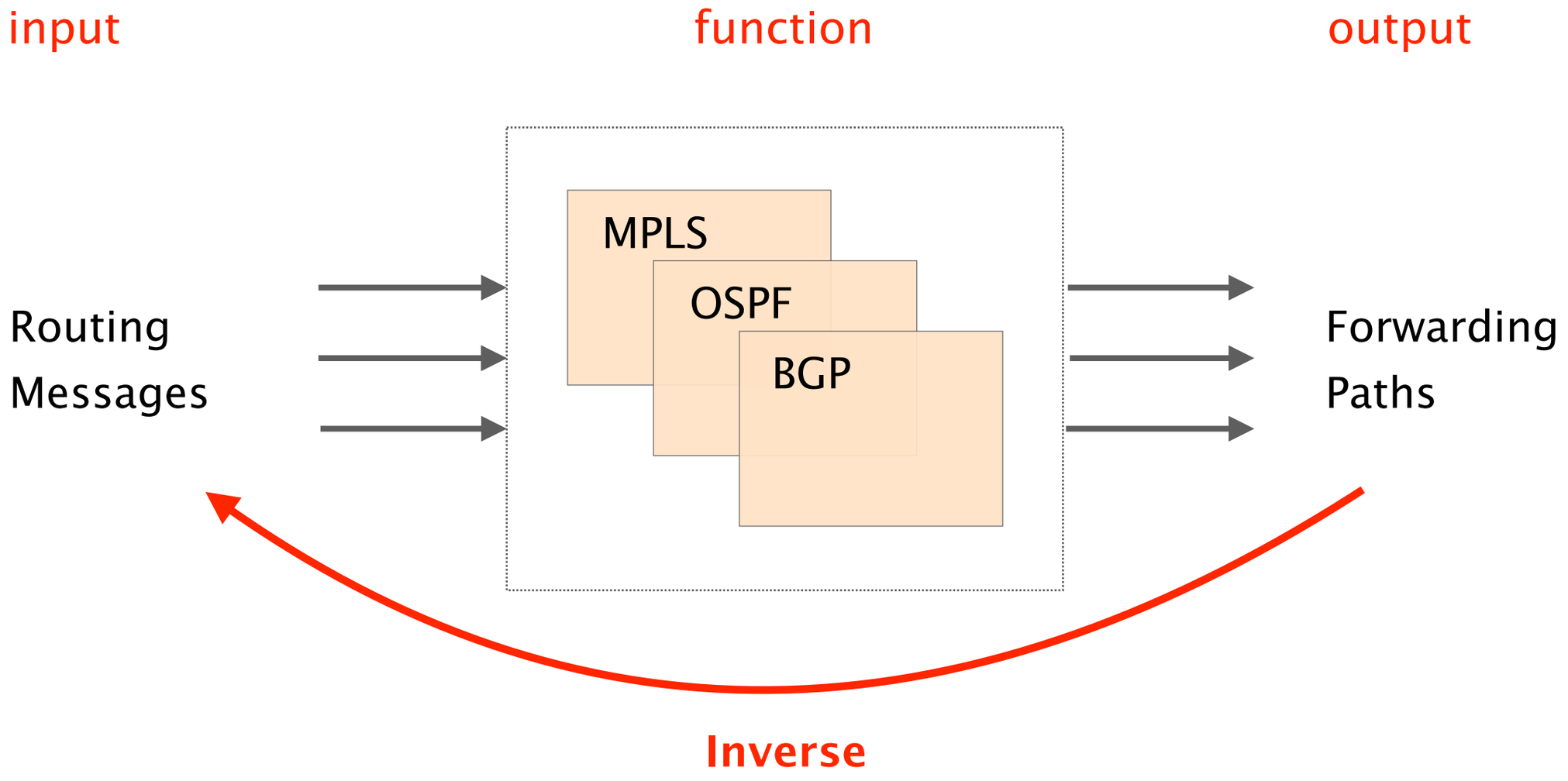
The forwarding paths are known,
provided by the operators or by the controller



The function is known, from the protocols' specification & the configuration



Given a path and a function, Fibbing computes corresponding routing messages by inverting the function



Fibbing is powerful

Fibbing is powerful

Theorem

Fibbing can program

any set of non-contradictory paths

Fibbing is powerful

Theorem

Fibbing can program

any set of non-contradictory paths

Fibbing is powerful

Theorem Fibbing can program any set of non-contradictory paths

- any path is loop-free

  (*e.g.*, [s1, a, b, a, d] is not possible)

- paths are consistent

  (*e.g.* [s1, a, b, d] and [s2, b, a, d] are inconsistent)

Fibbing is fast,
& works in practice

We developed efficient algorithms

polynomial in the # of requirements

Compute and minimize topologies in ms

independently of the size of the network

We tested them against real routers

works on both Cisco and Juniper

Fibbing enables network programmability today, on an existing network

Can be deployed immediately

supported by virtually any network, including ETH's

Simplify controller implementation

most of the heavy work is still done by the routers

Maintain operators' mental model

good old protocols running, easier troubleshooting

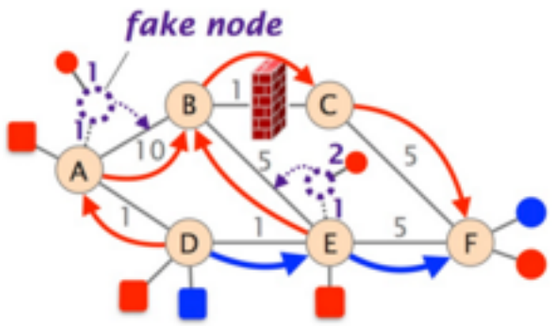
Check out our webpage

http://fibbing.net

Fibbing: Small Lies for Better Networks

Fibbing is an architecture that enables central control over distributed routing. This way, it combines the advantages of SDN (flexibility, expressivity, and manageability) and traditional (robustness, and scalability) approaches.

Fibbing introduces fake nodes and links into an underlying link-state routing protocol, so that routers compute their own forwarding tables based on the augmented topology. Fibbing is expressive, and readily supports flexible load balancing, traffic engineering, and backup routes. Fibbing works with any unmodified routers speaking OSPF.



Fibbing won the Best Paper Award at SIGCOMM 2015!

Read the papers | Look at the presentations | Watch the demo | Get the code

Carlo Daffara @cdaffara · Nov 13

Oh boy. A totally new way of thinking about the basis of SDN. Looks simple and practical. Will test. fibbing.net

2 1

OpenDaylight Project and 2 others follow

Jan-Erik Mångs @jemangs · Nov 17

Fibbing: **Central Control Over Distributed Routing** < Brilliant :-) fibbing.net

AMS-IX and 1 other follow

Michiel Appelman @michielappelman · Aug 21

Interesting concept and cool webpage: fibbing.net – **Central Control Over Distributed Routing**

Olivier Bonaventure Retweeted

ACM SIGCOMM @ACMSIGCOMM · Aug 20

SIGCOMM 2015 best paper award: "**Central Control Over Distributed Routing**" by Vissicchio et. Al., conferences.sigcomm.org/sigcomm/2015/p… #sigcomm2015

Brian Krent @BrianKrent · Nov 13

"Central Control Over Distributed Routing"

fibbing.net/files/sig15.pdf

"Fibbing: Small Lies for Better Networks"

fibbing.net

CSAIL at MIT follows

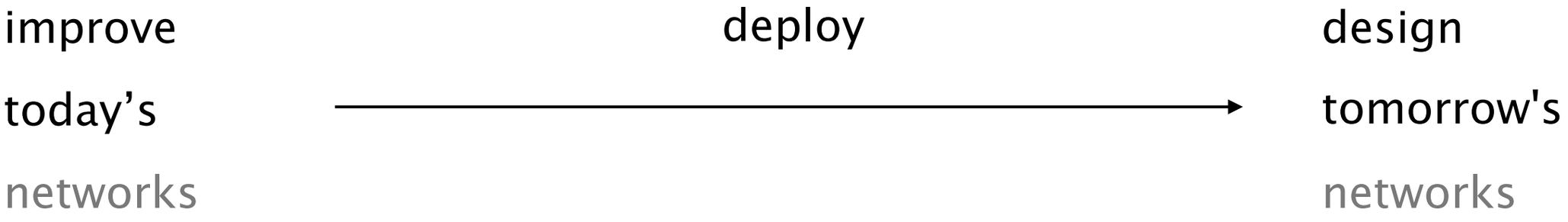
John Evdemon @jevdemon · Nov 13

Fibbing is an architecture that enables **central control over distributed routing**. Interesting idea. fibbing.net

Our focus

these days

Leverage network programability to…

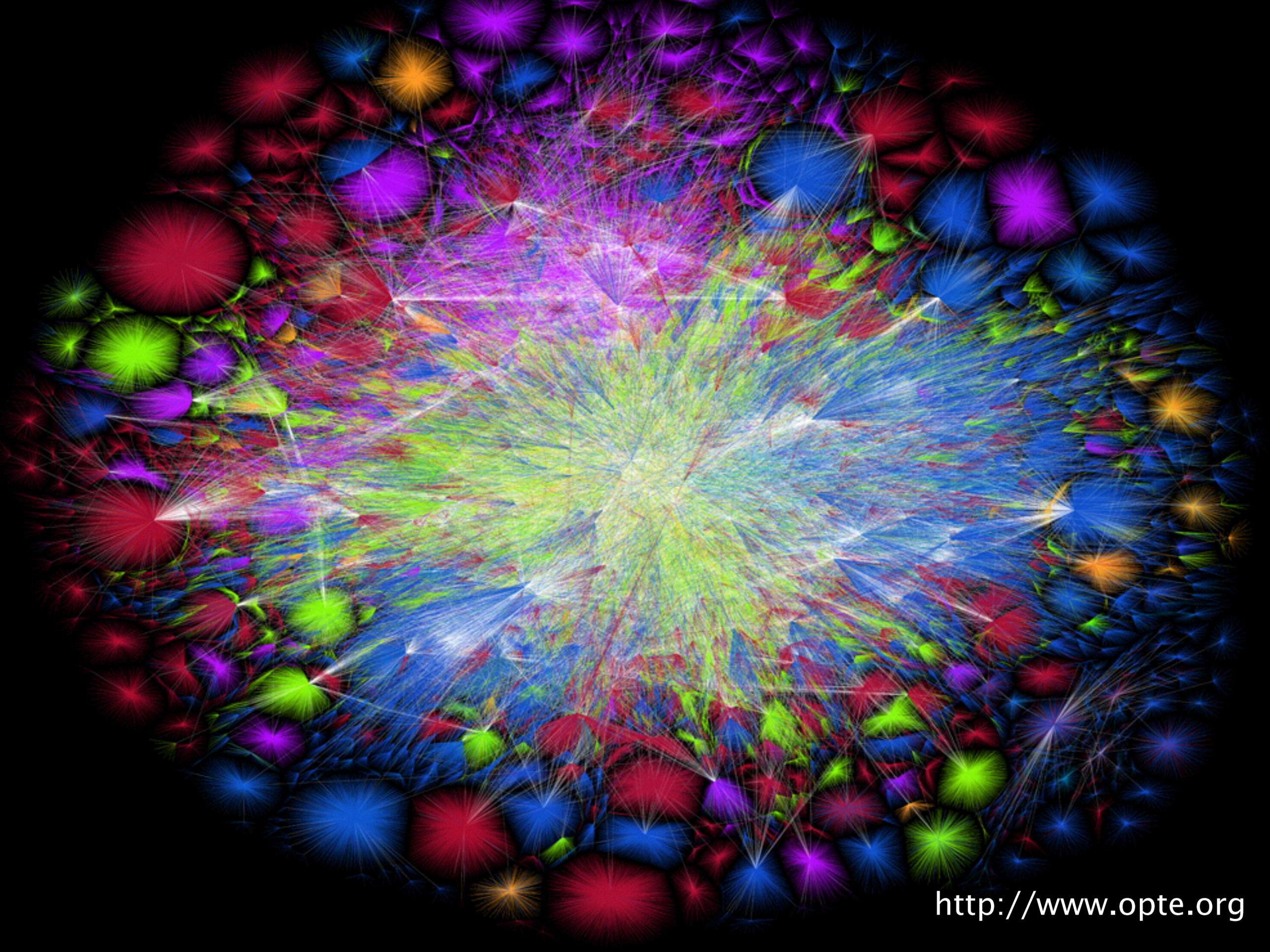


Our focus

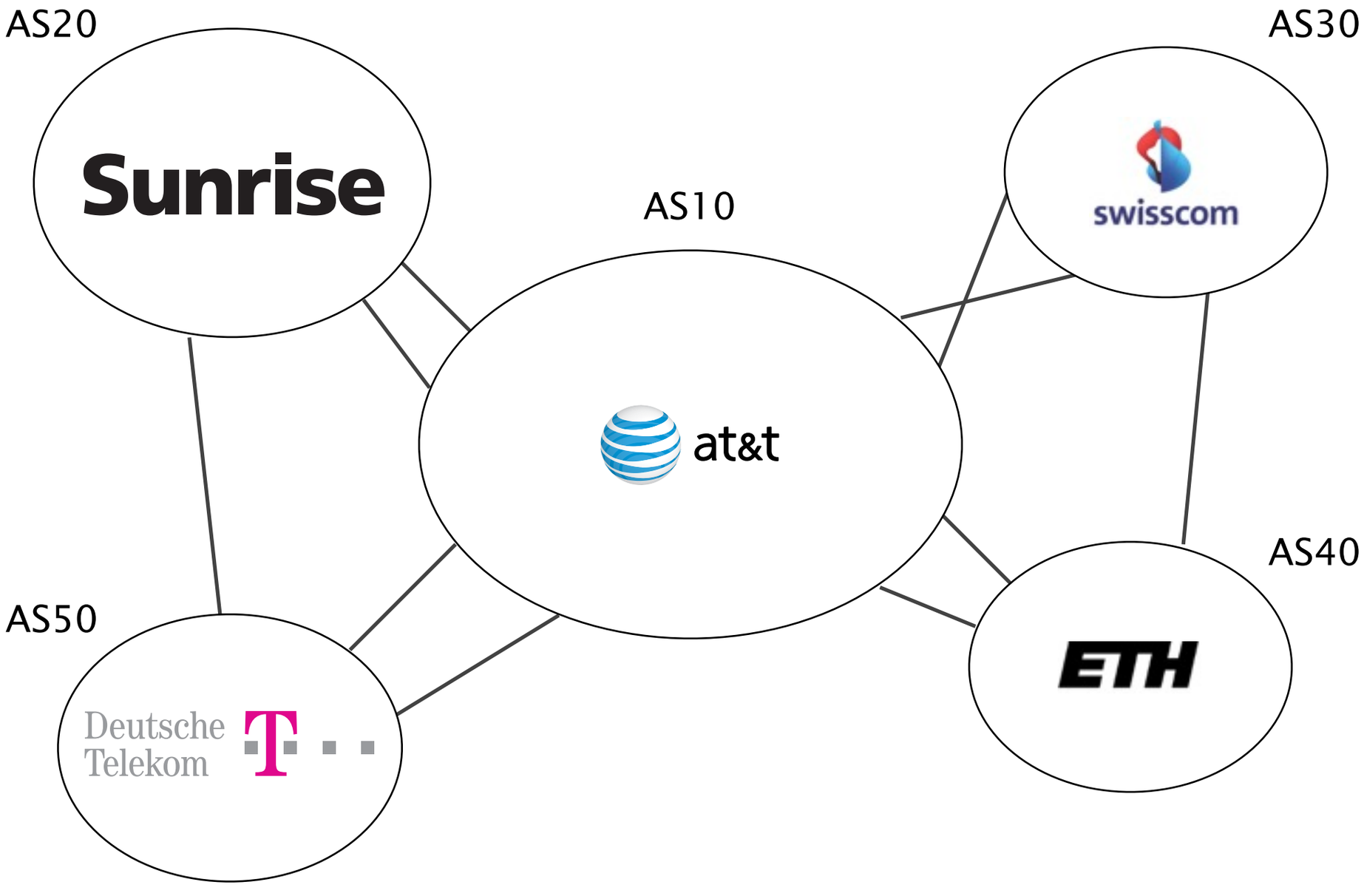
these days

Leverage network programability to…

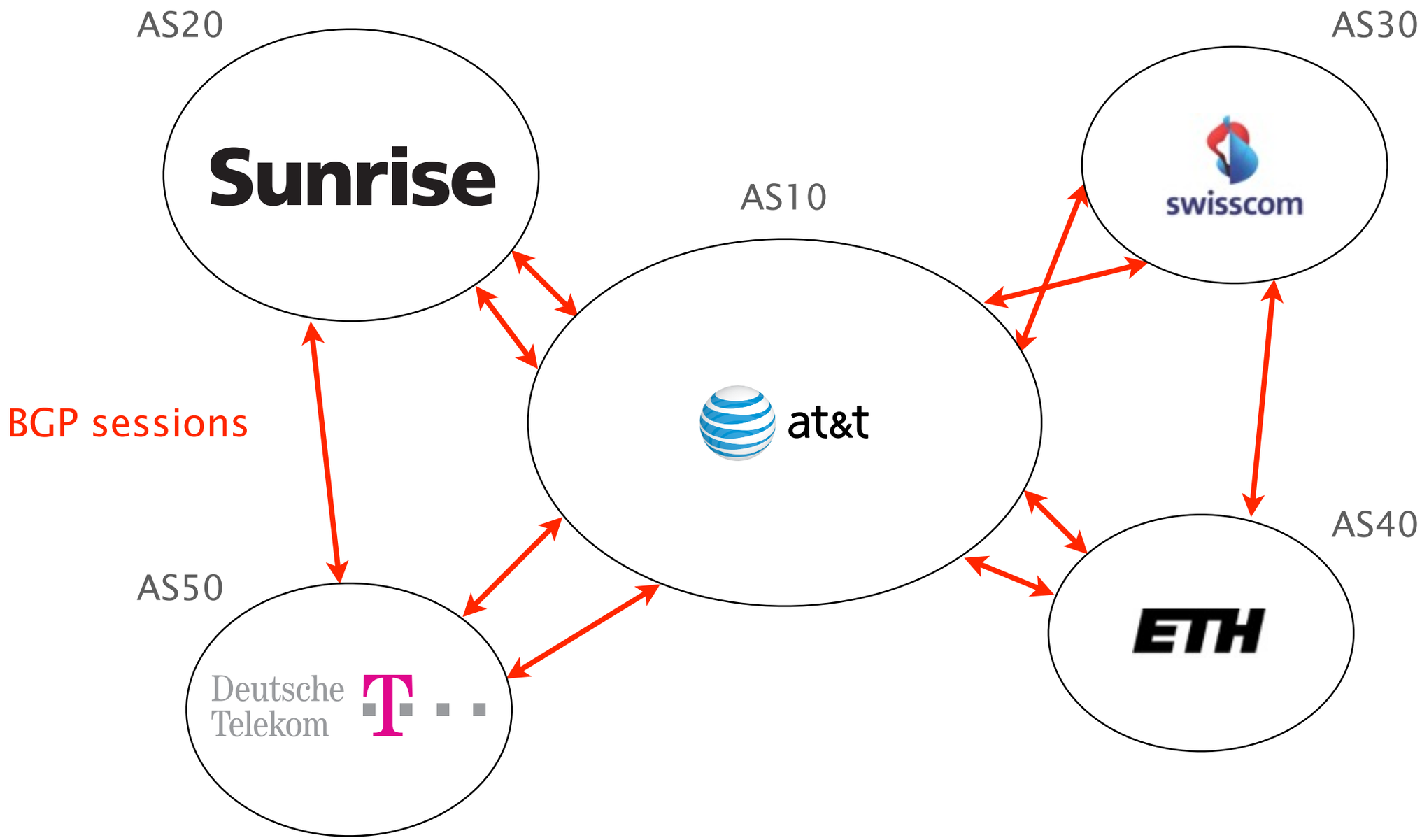
deploy better routing in the Internet


http://www.opte.org

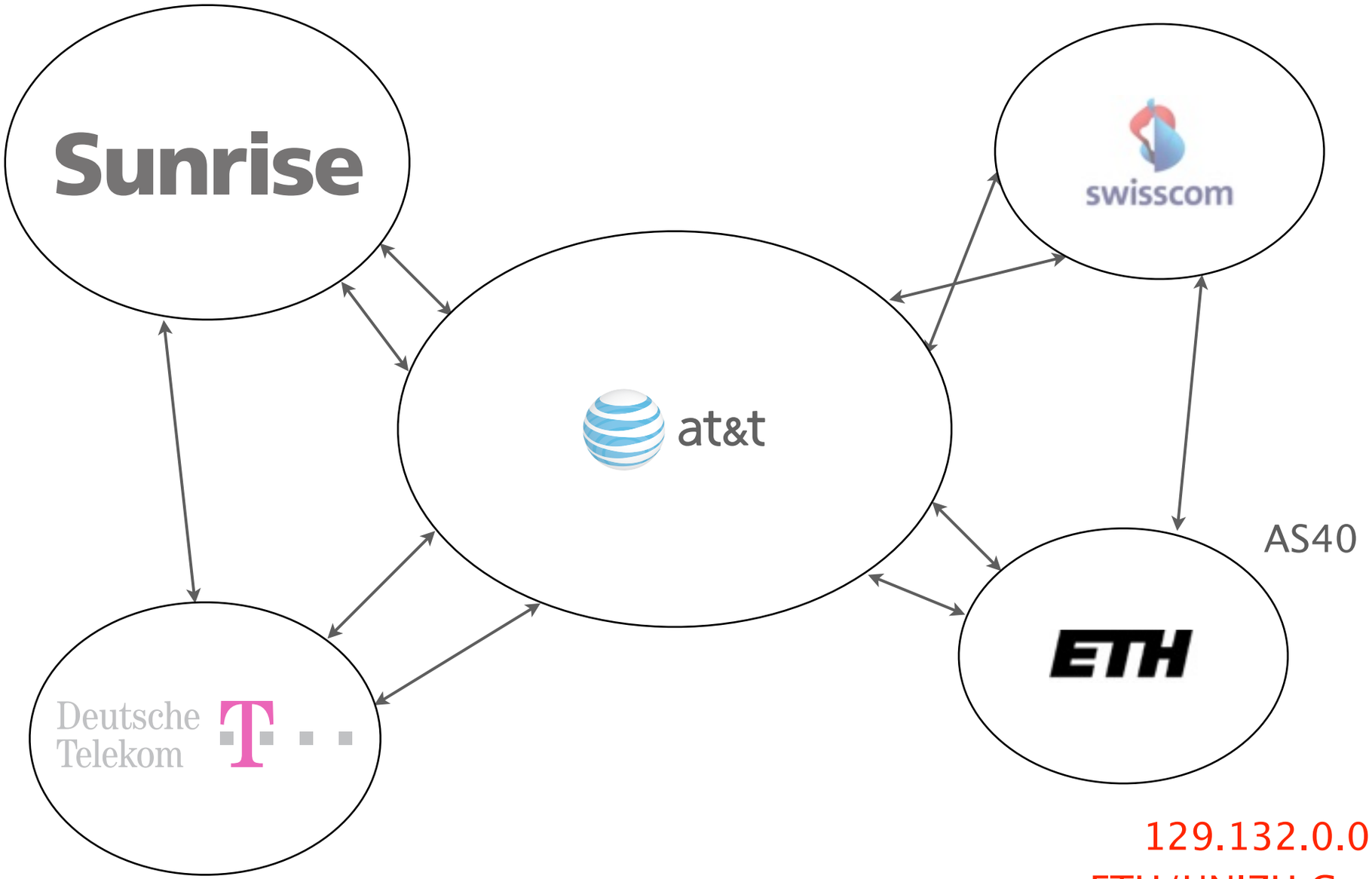
The Internet is a network of >50,000 networks, referred to as Autonomous Systems (AS)



BGP is the routing protocol “glueing” the Internet together

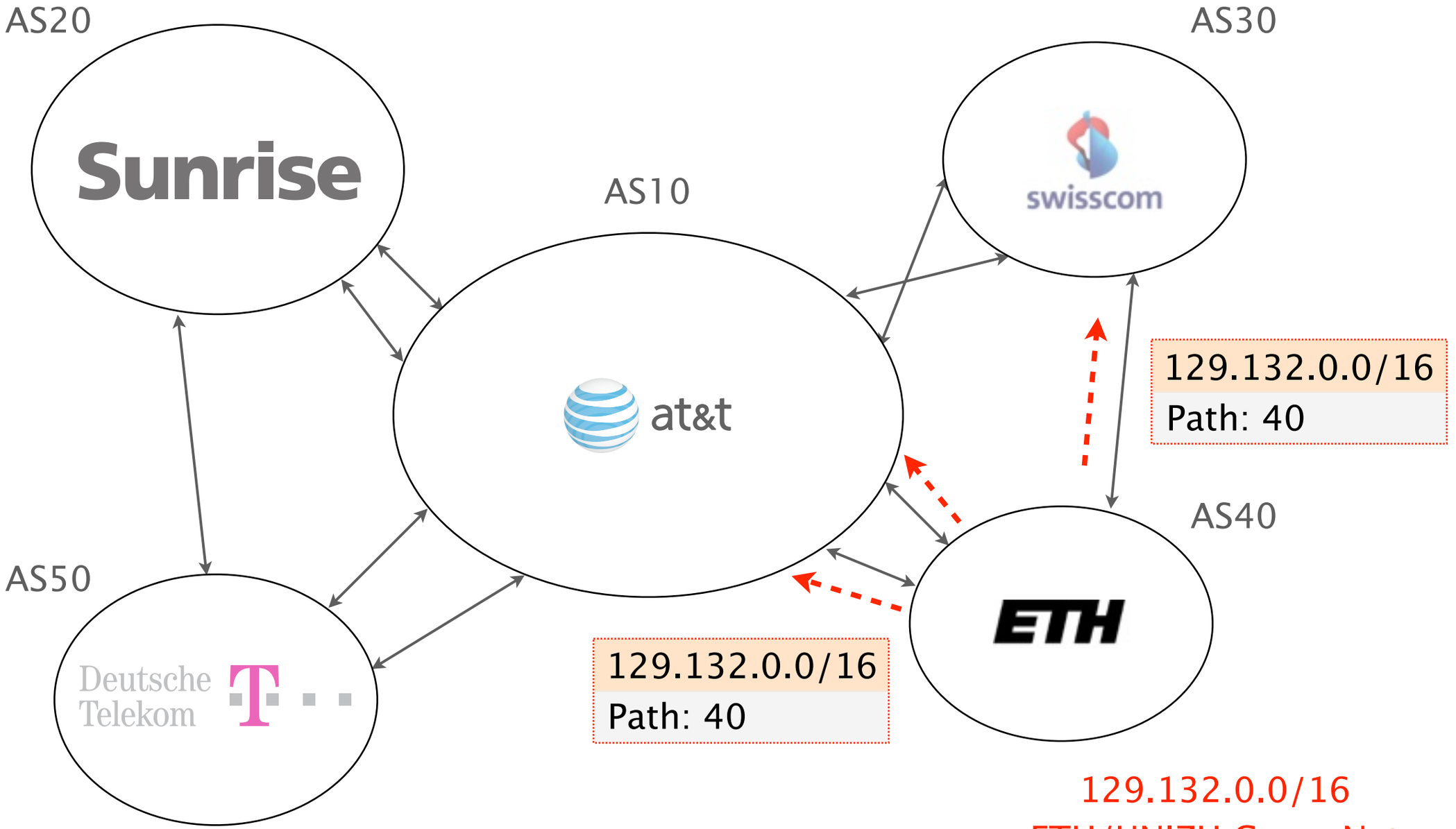


ASes exchange information about
the destinations (IP addresses) they can reach

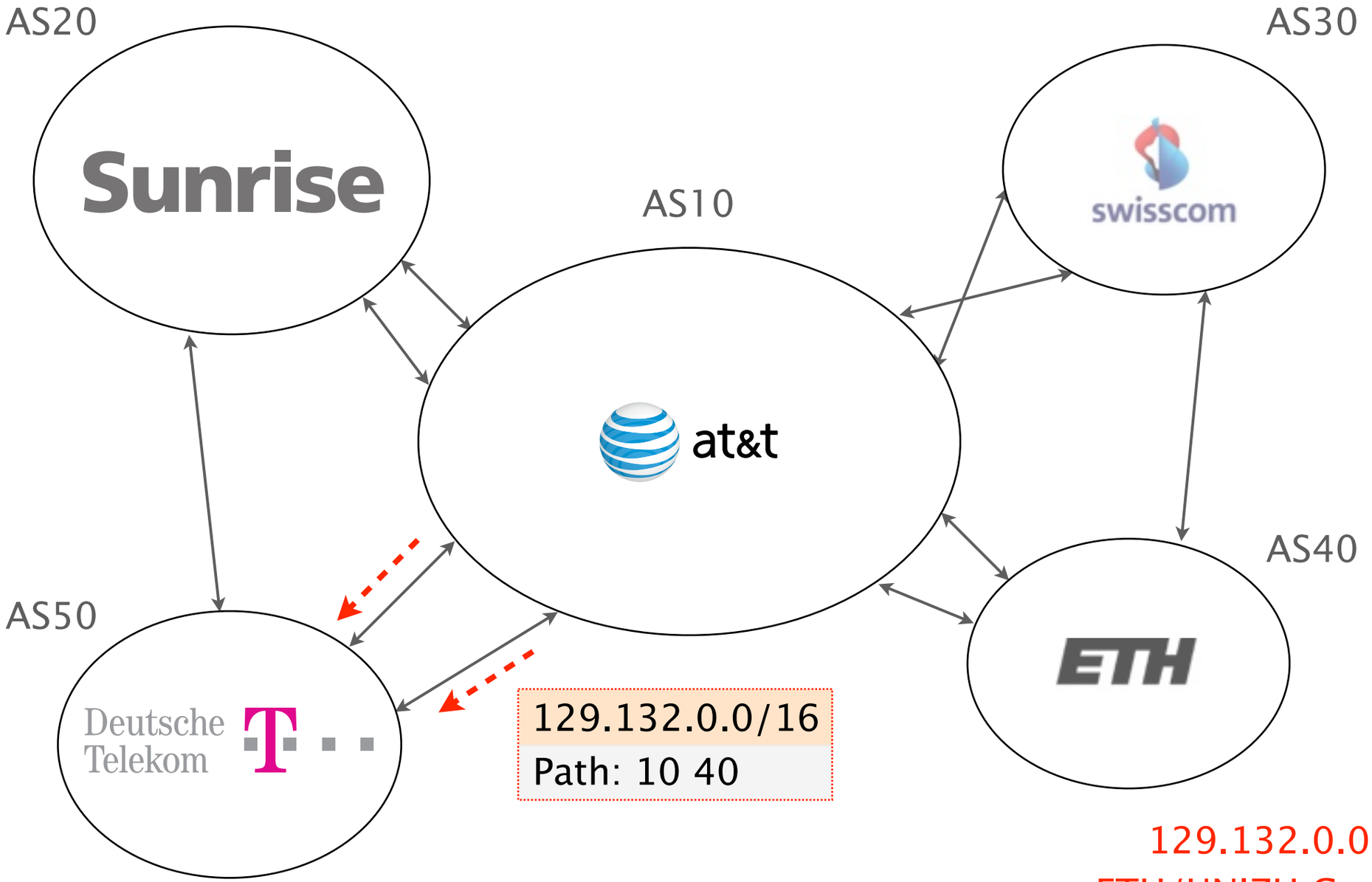


129.132.0.0/16
ETH/UNIZH Camp Net

ASes exchange information about
the destinations (IP addresses) they can reach

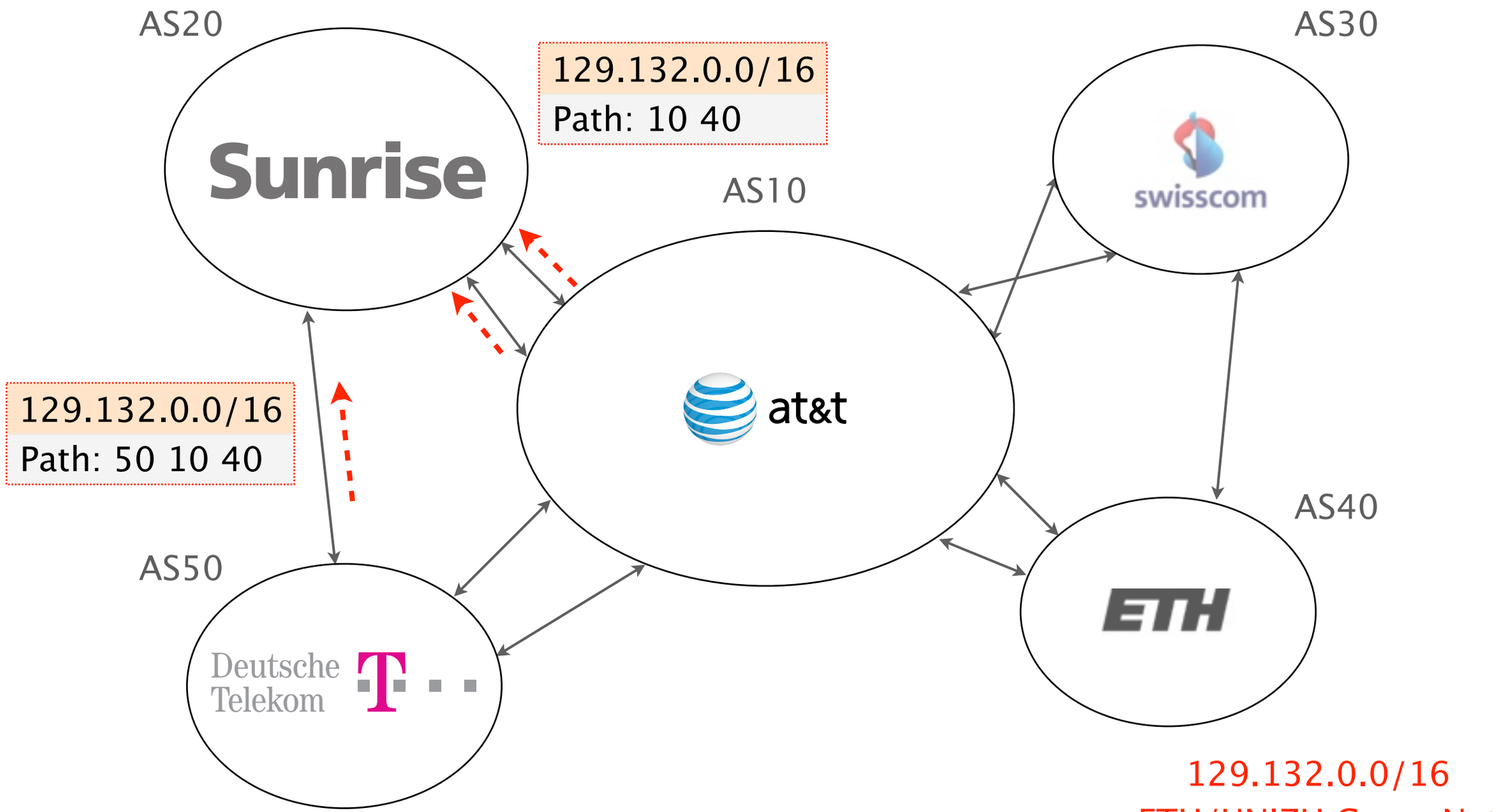


Reachability information is propagated hop-by-hop



129.132.0.0/16
ETH/UNIZH Camp Net

Reachability information is propagated hop-by-hop



Life of a BGP router is made of three consecutive steps

while true:

- receives paths from my neighbors
- select one best path for each destination
- export the best path to my neighbors

BGP is notoriously inflexible
and difficult to manage

BGP is notoriously inflexible
and difficult to manage

Flexibility

Control

Complexity

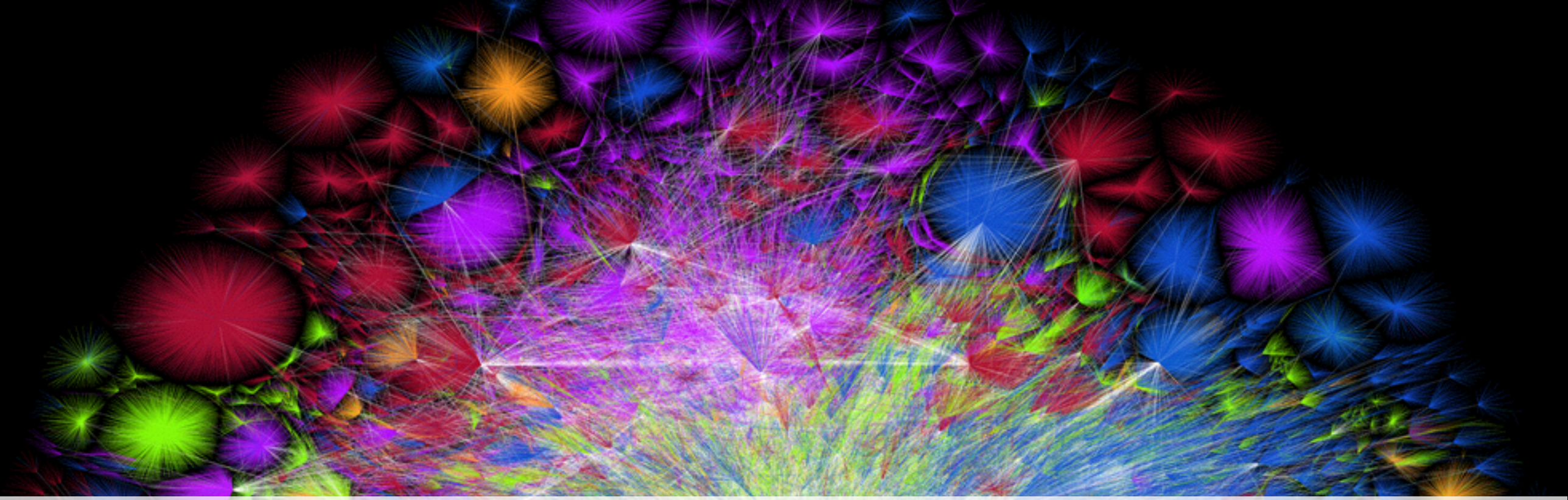
BGP is notoriously inflexible and difficult to manage

	BGP	
Flexibility	limited	destination-based
Control	indirect	configuration-based
	local	
Complexity	high	PSPACE-hard* to know if it converges

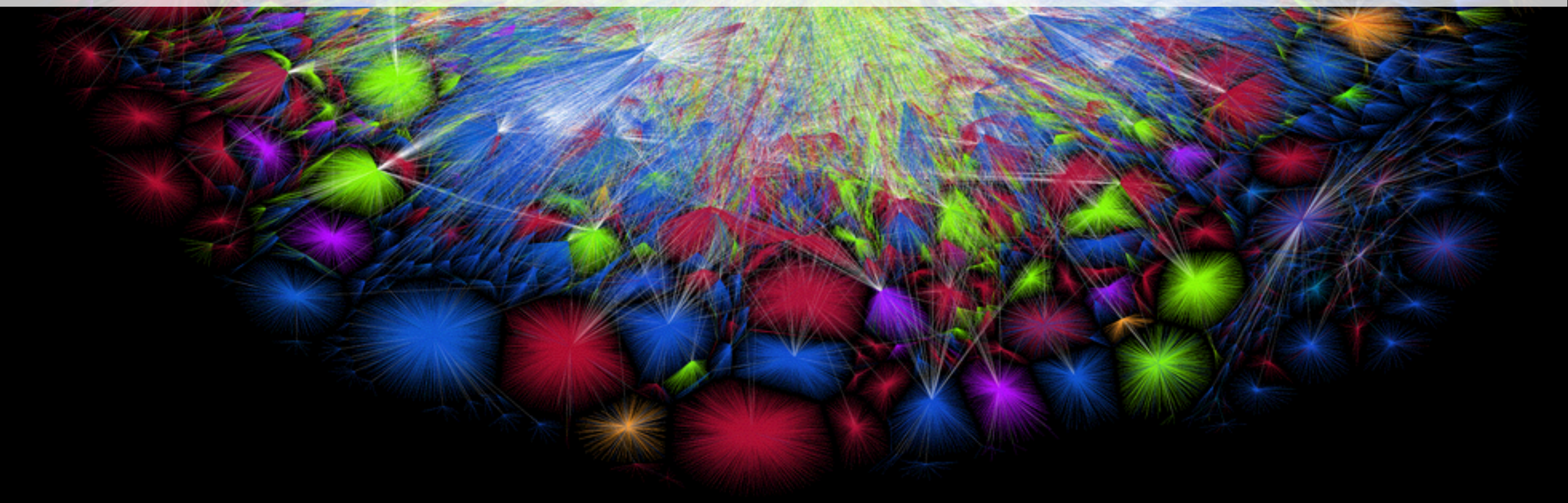
* [CCDVV ICNP'13] Using Routers to Build Logic Circuits: How Powerful is BGP?

SDN can really help!

	BGP	SDN	
Flexibility	limited	high	
Control	indirect	direct	
	local	local & remote	
Complexity	high	simpler	deterministic



How do you deploy SDN in a network composed of 50,000 subnetworks?



How do you deploy SDN in a network composed of 50,000 subnetworks?

you just don't…

Instead, you aim at finding locations where deploying SDN can have the most impact

Instead, you aim at finding locations where deploying SDN can have the most impact

Deploy SDN in locations that

- connect a large number of networks
- carry a large amount of traffic
- are opened to innovation

Internet eXchange Points (IXP) meet all the criteria

Deploy SDN in locations that

- connect a large number of networks
- carry a large amount of traffic
- are opened to innovation

AMS-IX

750 networks

4.3 Tb/s (peak)

BGP Route Server

Mobile peering

Open peering…

https://www.ams-ix.net

A single deployment
can have a large impact

SDX = SDN + IXP

Joint work with: [SIGCOMM'15]

Arpit Gupta, Muhammad Shahbaz, Russ Clark, Ethan Katz-Bassett,
Nick Feamster, Jennifer Rexford and Scott Shenker

SDX = SDN + IXP

Augment the IXP data-plane with SDN capabilities

keeping default forwarding and routing behavior

Enable fine-grained inter domain policies

bringing new features while simplifying operations

SDX = SDN + IXP

Augment the IXP data-plane with SDN capabilities

keeping default forwarding and routing behavior

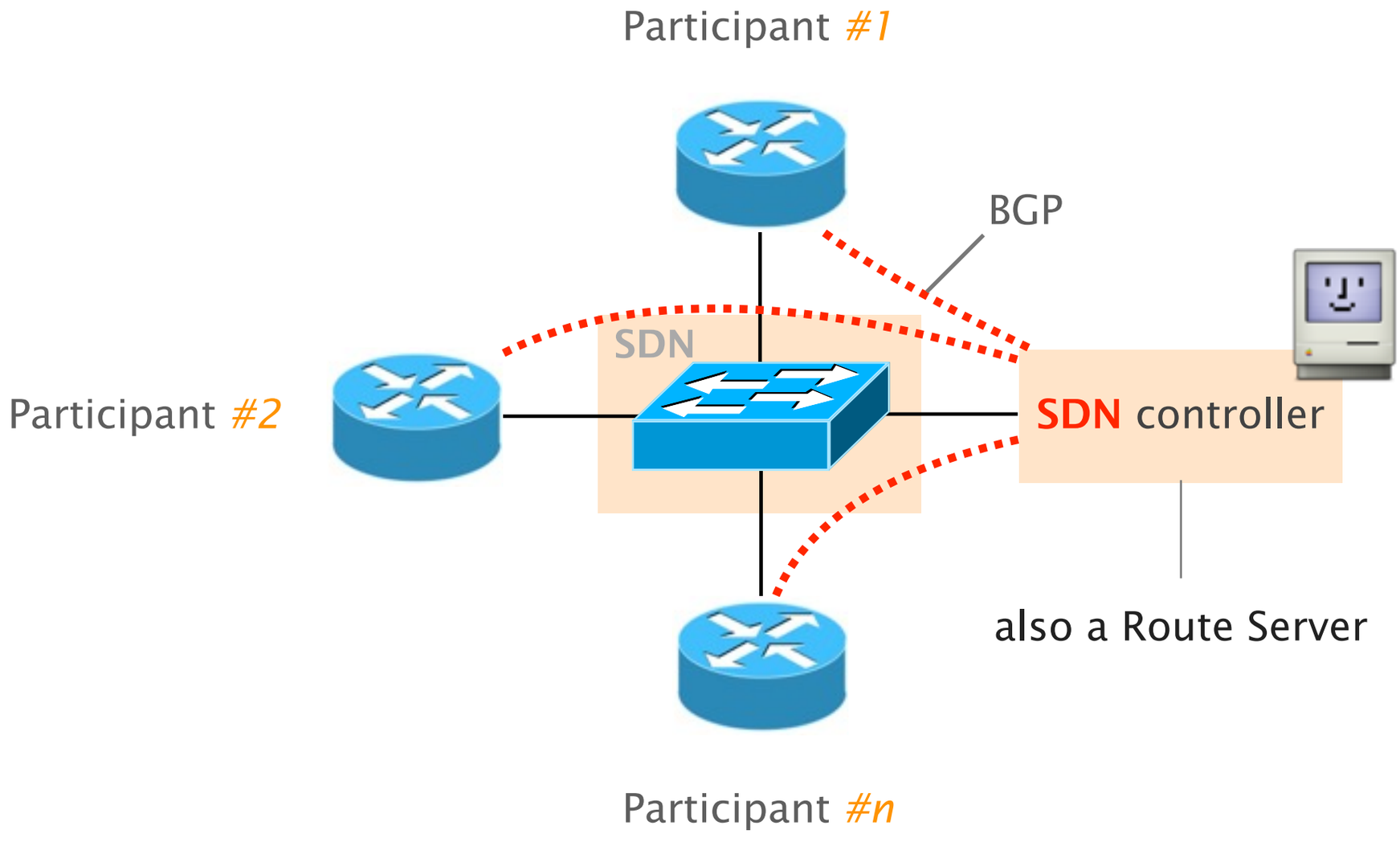
Enable fine-grained inter domain policies

bringing new features while simplifying operations

… with scalability and correctness in mind

supporting the load of a large IXP and resolving conflicts

In a SDX, each participant connects its edge router(s) to a shared SDN-enabled network



Each participant writes policies independently in a high-level language and transmits them to the controller

Participant #1's policy:

```
match(dstip=Google),fwd(1.1)
match(dstip=Yahoo), fwd(1.2)
```

Participant #2's policy:

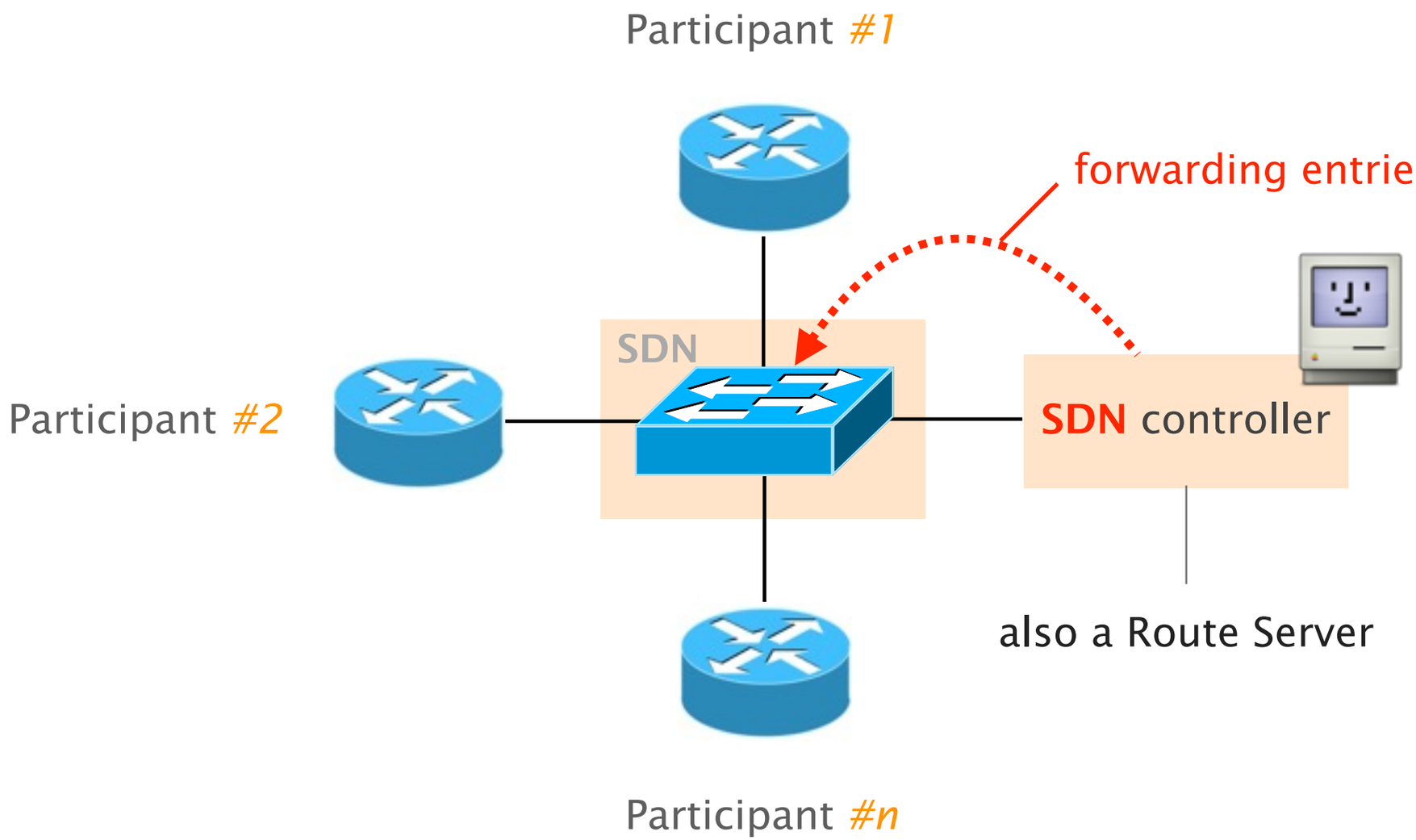
```
match(dstip=ip1), fwd(1)
match(dstip=ip2), fwd(3)
match(dstip=ip3), fwd(5)
```

SDN controller

```
match(dstip=ipX), fwd(n.1)
```

Participant #n's policy

The SDX controller compiles policies to forwarding entries ensuring isolation, scalability and solving conflicts



SDX enables a wide range of novel applications

security	Prevent/block policy violation Prevent participants communication Upstream blocking of attacks
forwarding optimization	Middlebox traffic steering Traffic offloading Inbound Traffic Engineering Fast convergence
peering	Application-specific peering
remote-control	Influence BGP path selection Wide-area load balancing

SDX works today!

We have running code (*)

controller and BGP daemon

We are seeing ongoing deployments

FBI (US), TelX (Atlanta, US)

Many more interested parties

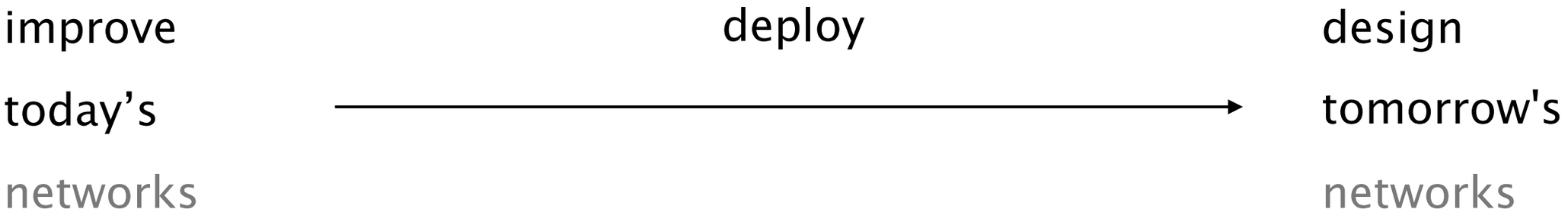
including AT&T, Amazon, Facebook & Google

(*) http://sdx.cs.princeton.edu/

Our focus

these days

Leverage network programability to…



Our focus

these days

Leverage network programability to…

design

tomorrow's

networks

Build controller platforms

Develop new applications

Build controller platforms

Develop new applications

SDN makes network programmability possible,
but not easy…

Challenges

Current SDN interfaces are low-level

remember assembly languages?

Controllers are valuable targets for attackers

take down the brain, not the body

Scalability, reliability & performance are paramount

control loops must complete within a second, all the time

Working on…

management abstractions

Current SDN interfaces are low-level

remember assembly languages?

defense mechanisms

Controllers are valuable targets for attackers

take down the brain, not the body

control optimization

Scalability, reliability & performance are paramount

control loops must complete within a second, all the time

Build controller platforms

Develop new applications

There are many avenues

data-center networks

CREDIT SUISSE

UBS

enterprise networks

cellular (5G) networks

Programmability can radically change the way we do network

"Stars are aligned"

tremendous interests from industry & academia

Tons of interesting research challenges

SDN has only been around for ~6 years

Turn networking into a proper discipline

instead of the current engineering minefield

Improving the Internet

From Fragility to Resilience

Laurent Vanbever

www.vanbever.eu

ETH Zürich

See you at the Apéro!