Communication Networks

Spring 2019

Laurent Vanbever

nsg.ee.ethz.ch

ETH Zürich (D-ITET)

May 20 2019

Last Monday on

Communication Networks

IPv6

next generation of
Internet addressing

programmable
networks

next generation of
network devices

IPv6

programmable networks

next generation of Internet addressing

Network Address Translation (NAT)

Sharing a single (public) address between hosts

Port numbers (transport layer) are used to distinguish

One of the main reasons why we can still use IPv4

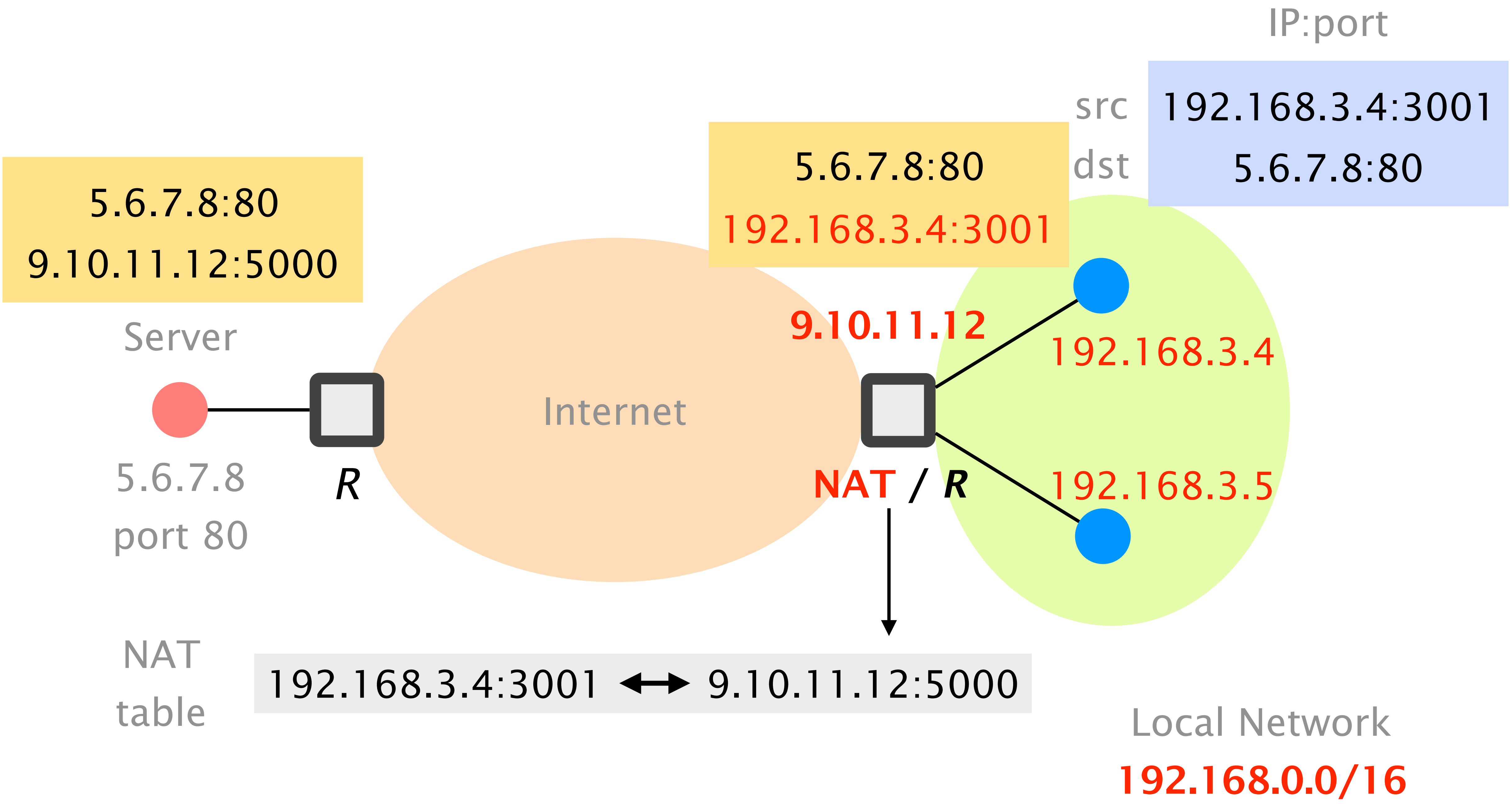
Saved us from address depletion

Violates the general end-to-end principle of the Internet

A NAT box adds a layer of indirection

The Internet with NAT

Hosts behind NAT get a private address



IP:port

src 192.168.3.4:3001

dst 5.6.7.8:80

5.6.7.8:80

192.168.3.4:3001

5.6.7.8:80

9.10.11.12:5000

Server

5.6.7.8

port 80

R

Internet

9.10.11.12

NAT / R

192.168.3.4

192.168.3.5

NAT table

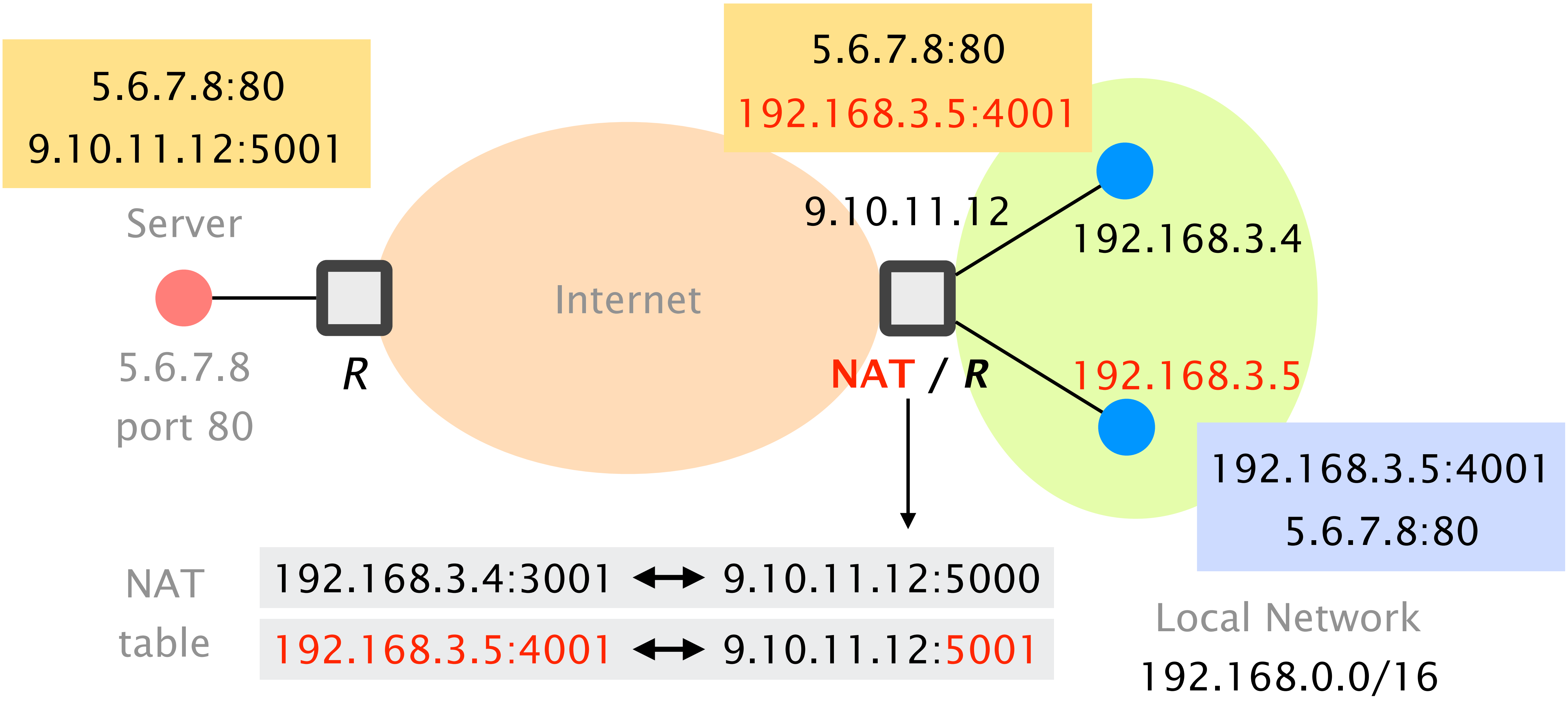
192.168.3.4:3001 ⟷ 9.10.11.12:5000

Local Network

192.168.0.0/16

The Internet with NAT

The port numbers are used to multiplex single addresses



NAT table

192.168.3.4:3001	↔	9.10.11.12:5000
192.168.3.5:4001	↔	9.10.11.12:5001

Let's talk about **IPv6**

IPv6 addresses are encoded in 128 bits

Notation

8 groups of 16 bits each separated by colons (:)

Each group is written as four hexadecimal digits

Simplification

Leading zeros in any group are removed

One section of zeros is replaced by a double colon (::)

Normally the longest section

Examples

1080:0:0:0:8:800:200C:417A → 1080::8:800:200C:417A

FF01:0:0:0:0:0:0:0101 → FF01::101

0:0:0:0:0:0:0:1 → ::1

There are three types of IPv6 addresses: unicast, anycast, and multicast

Unicast

Identifies a single interface

Packets are delivered to this specific interface

Anycast

Identifies a set of interfaces

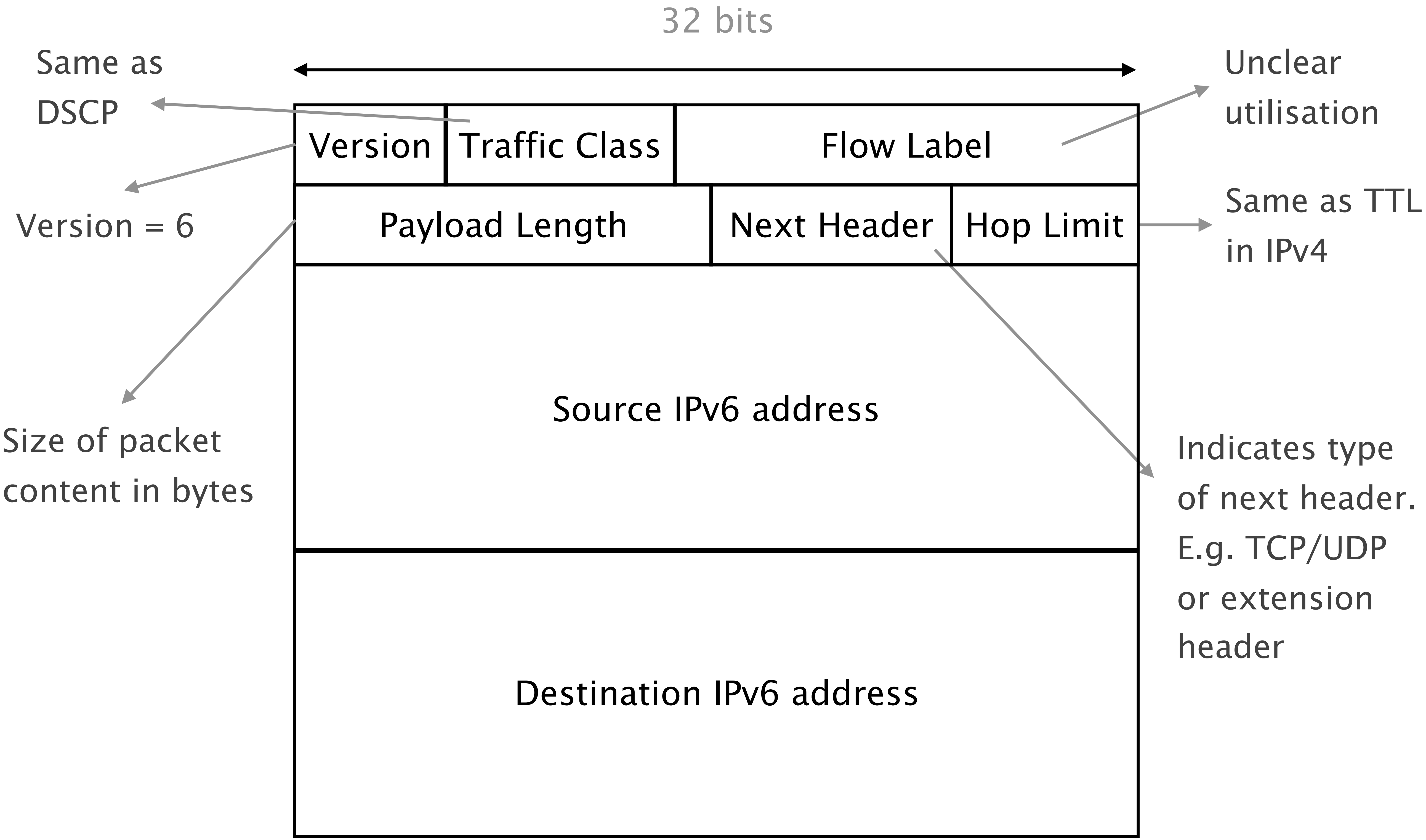
Packets are delivered to the "nearest" interface

Multicast

Identifies a set of interfaces

Packets are delivered to **all** interfaces

The IPv6 packet header format



32 bits

Version	Traffic Class	Flow Label	
Payload Length		Next Header	Hop Limit
Source IPv6 address			
Destination IPv6 address			

Same as DSCP

Version = 6

Size of packet content in bytes

Unclear utilisation

Same as TTL in IPv4

Indicates type of next header. E.g. TCP/UDP or extension header

How can a node obtain its IPv6 address(es)?

Manual configuration

As in the project, e.g. with ifconfig

From a server by using DHCPv6

Similar to the IPv4 version

Automatically

Using its link-local address and neighbor discovery

IPv6 autoconfiguration to find link-local address

Consider an end-system which has just started, it needs an IPv6 address to send ICMPv6 messages

Ethernet (MAC): 0800:200C:417A

Link-local: FE80::M_{64}(800:200C:417A)

M_{64}: 64-bit representation of the MAC address

Neighbor solicitation for FE80::M_{64}(800:200C:417A)

If **no** answer, the created link-local address is valid

IPv6 autoconfiguration
to obtain the IPv6 prefix of subnet

Routers periodically advertise the prefix

Sent to all end-systems: FF02::1

The advertisements can contain:

IPv6 prefix and length

Network MTU to use

Maximum hop limit to use

Lifetime of the default router

How long generated addresses are preferred

IPv6 autoconfiguration
to build global unicast address

Ethernet (MAC): 0800:200C:417A

Prefix: 2001:6a8:3080:1::/64

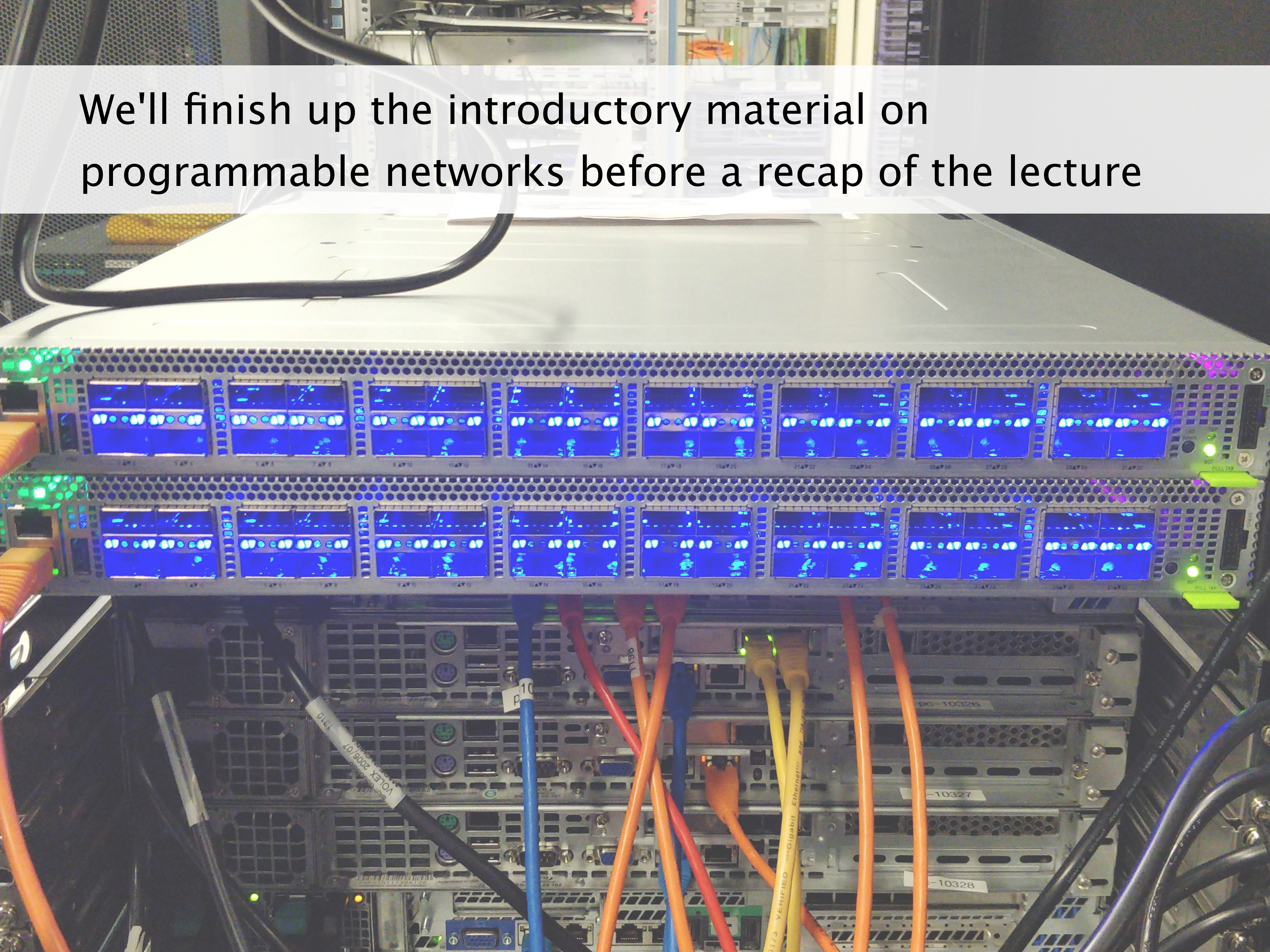
Global unicast:

2001:6a8:3080:1:M_{64}(800:200C:417A)

contains MAC address of host

Today on

Communication Networks



We'll finish up the introductory material on programmable networks before a recap of the lecture

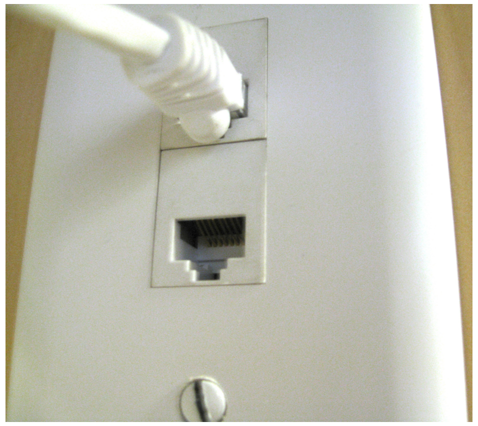
Communication Networks

So what?!

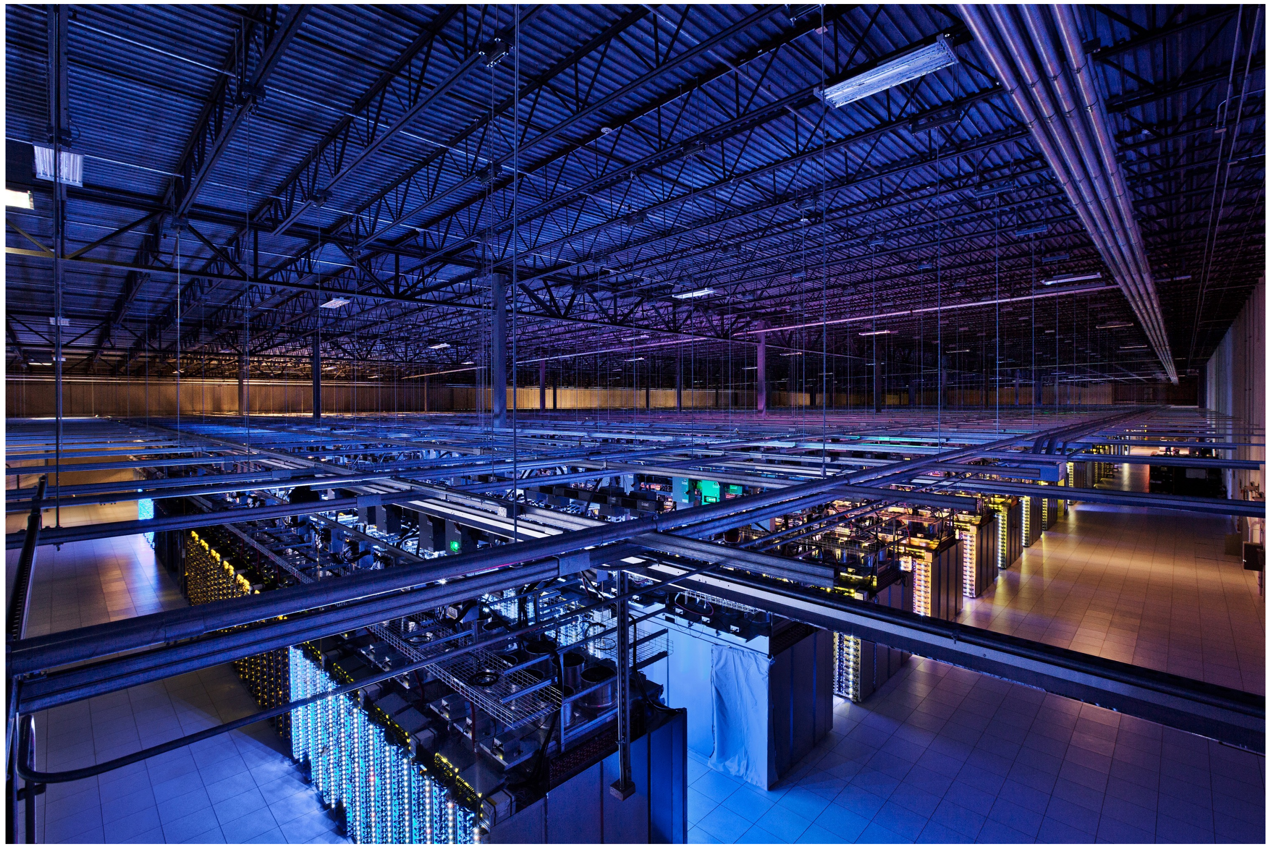
Google | google.com/datacenters

Knowledge

Understand how the Internet works and why



from your network plug…



…to Google's data-center

List any

technologies, principles, applications…

used after typing in:

> www.google.ch

and pressing enter in your browser

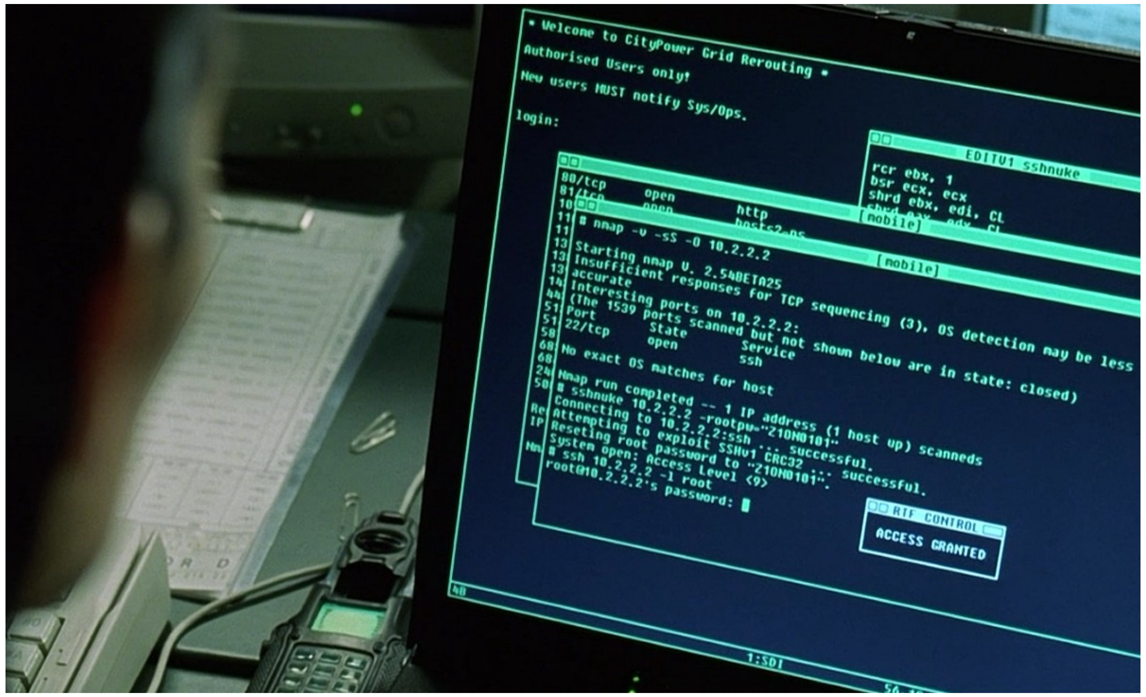
Insight

Key concepts and problems in Networking

Naming Layering Routing Reliability Sharing

Skill

Build, operate and configure networks



```
* Welcome to CityPower Grid Rerouting *
Authorised Users only!
New users MUST notify Sys/Ops.

login:
```

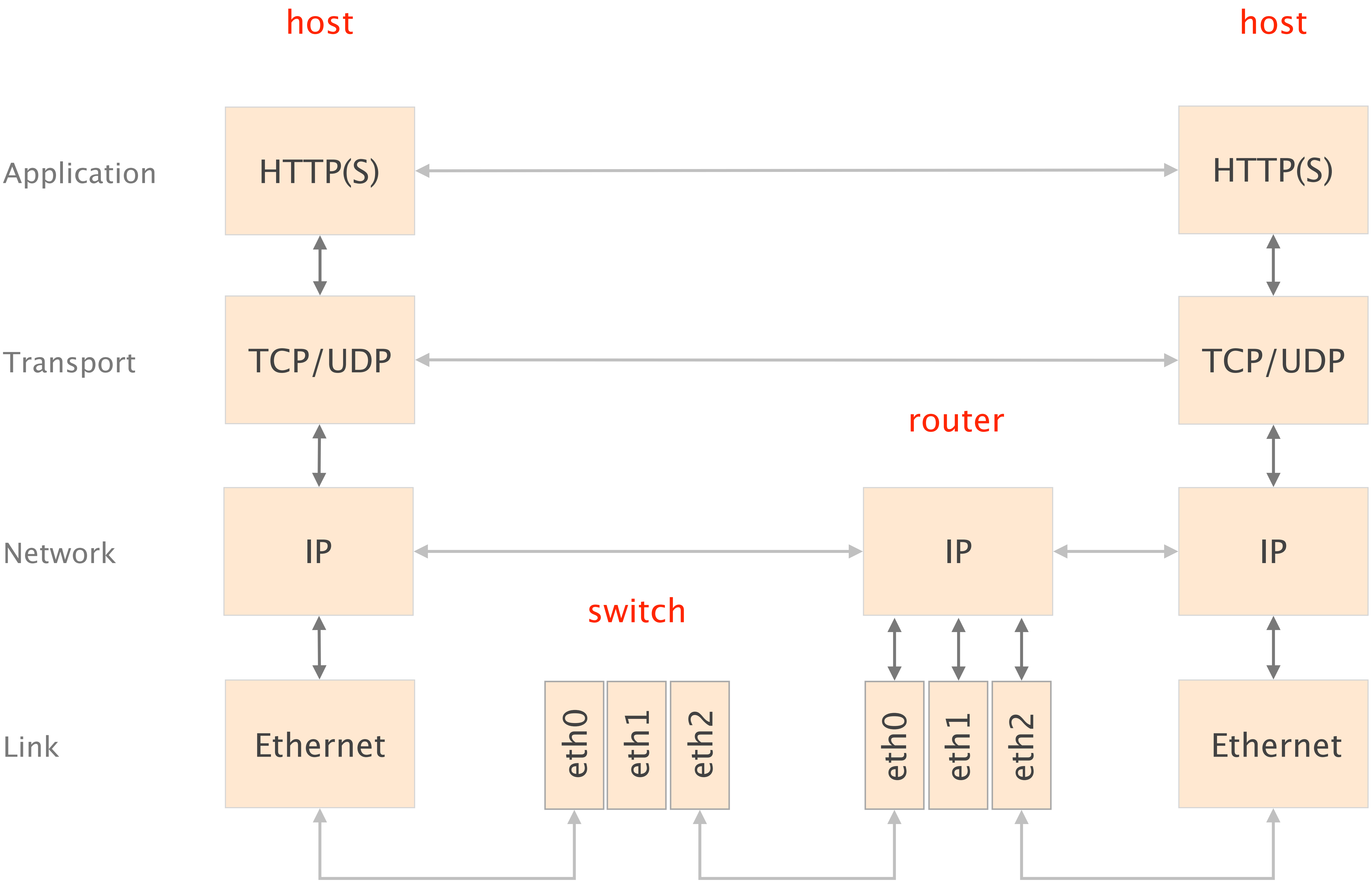
```
EDITV1 sshnuke
rcr ebx, 1
bsr ecx, ecx
shrd ebx, edi, CL
```

```
# nmap -v -sS -O 10.2.2.2
Starting nmap V. 2.54BETA25
Insufficient responses for TCP sequencing (3), OS detection may be less
accurate
Interesting ports on 10.2.2.2:
(The 1539 ports scanned but not shown below are in state: closed)
Port       State       Service
22/tcp     open        ssh

No exact OS matches for host

Nmap run completed -- 1 IP address (1 host up) scanneds
# sshnuke 10.2.2.2 -rootpw="Z10N0101"
Connecting to 10.2.2.2:ssh ... successful.
Attempting to exploit SSHv1 CRC32 ... successful.
Reseting root password to "Z10N0101".
System open: Access Level <9>
# ssh 10.2.2.2 -l root
root@10.2.2.2's password: █
```

Trinity using a port scanner (nmap) in Matrix Reloaded™

The Internet is organized as layers, providing a set of services

	layer	service provided
L5	Application	network access
L4	Transport	end-to-end delivery (reliable or not)
L3	Network	global best-effort delivery
L2	Link	local best-effort delivery
L1	Physical	physical transfer of bits


host
host
Application
HTTP(S)
HTTP(S)
Transport
TCP/UDP
TCP/UDP
router
Network
IP
IP
IP
switch
Link
Ethernet
eth0
eth1
eth2
eth0
eth1
eth2
Ethernet

We started with the fundamentals of routing and reliable transport

	Application	network access
L4	Transport	end-to-end delivery (reliable or not)
L3	Network	global best-effort delivery
	Link	local best-effort delivery
	Physical	physical transfer of bits

We saw three ways to compute valid routing state

	Intuition	Example
#1	Use tree-like topologies	Spanning-tree
#2	Rely on a global network view	Link-State SDN
#3	Rely on distributed computation	Distance-Vector BGP

We saw how to design a reliable transport protocol

plus, you're implementing one right now

goals

correctness	ensure data is delivered, in order, and untouched
timeliness	minimize time until data is transferred
efficiency	optimal use of bandwidth
fairness	play well with other concurrent communications

In each case, we explored the rationale behind each protocol and why they came to be

Why did the protocols end up looking like this?

minimum set of features required

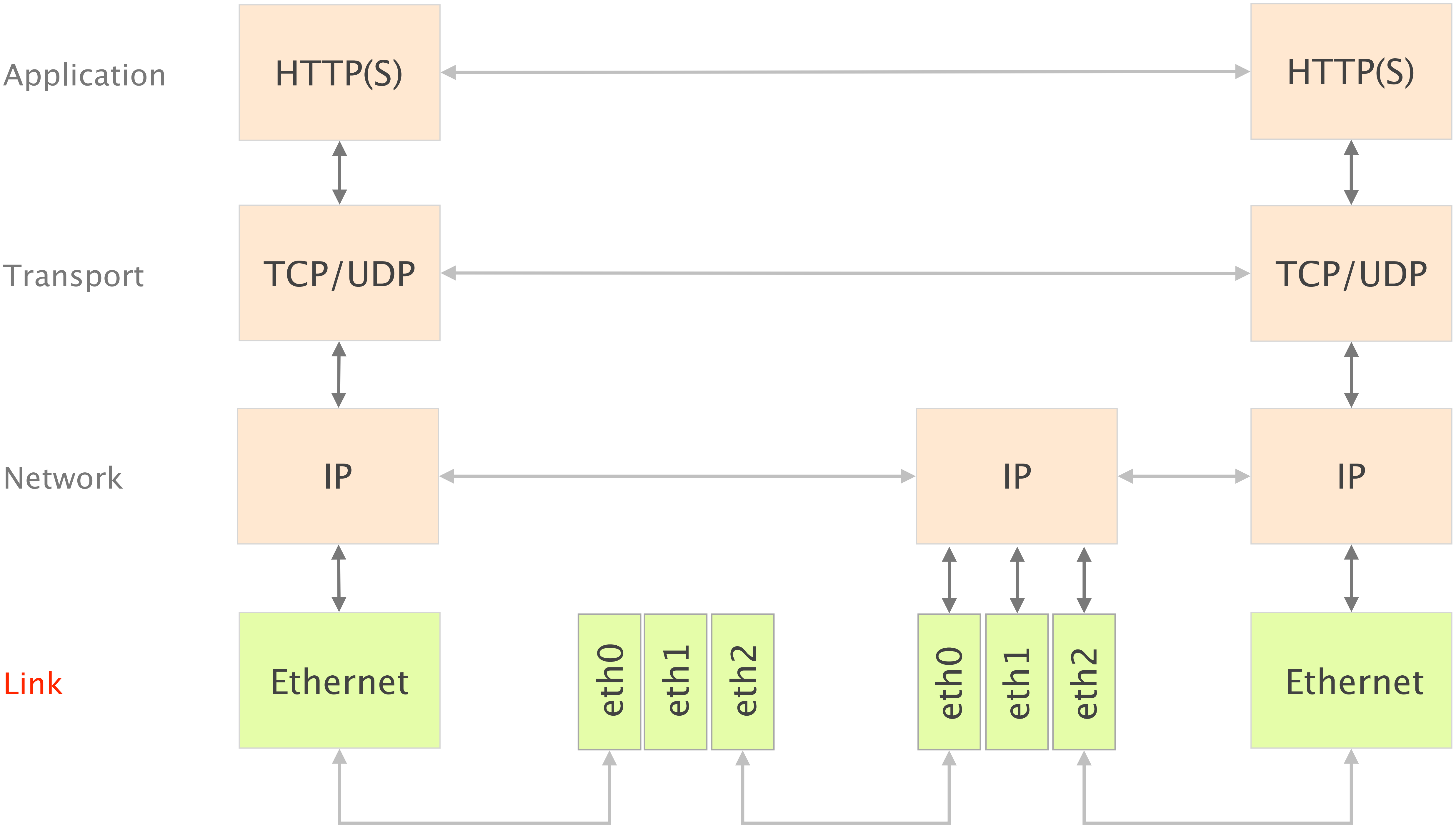
What tradeoffs do they achieve?

efficiency, cost,…

When is one design more adapted than another?

packet switching *vs* circuit switching, DV *vs* LS,…

We then climbed up the layers, starting from layer 2



Communication Networks

Part 2: The Link Layer

#1 What is a link?

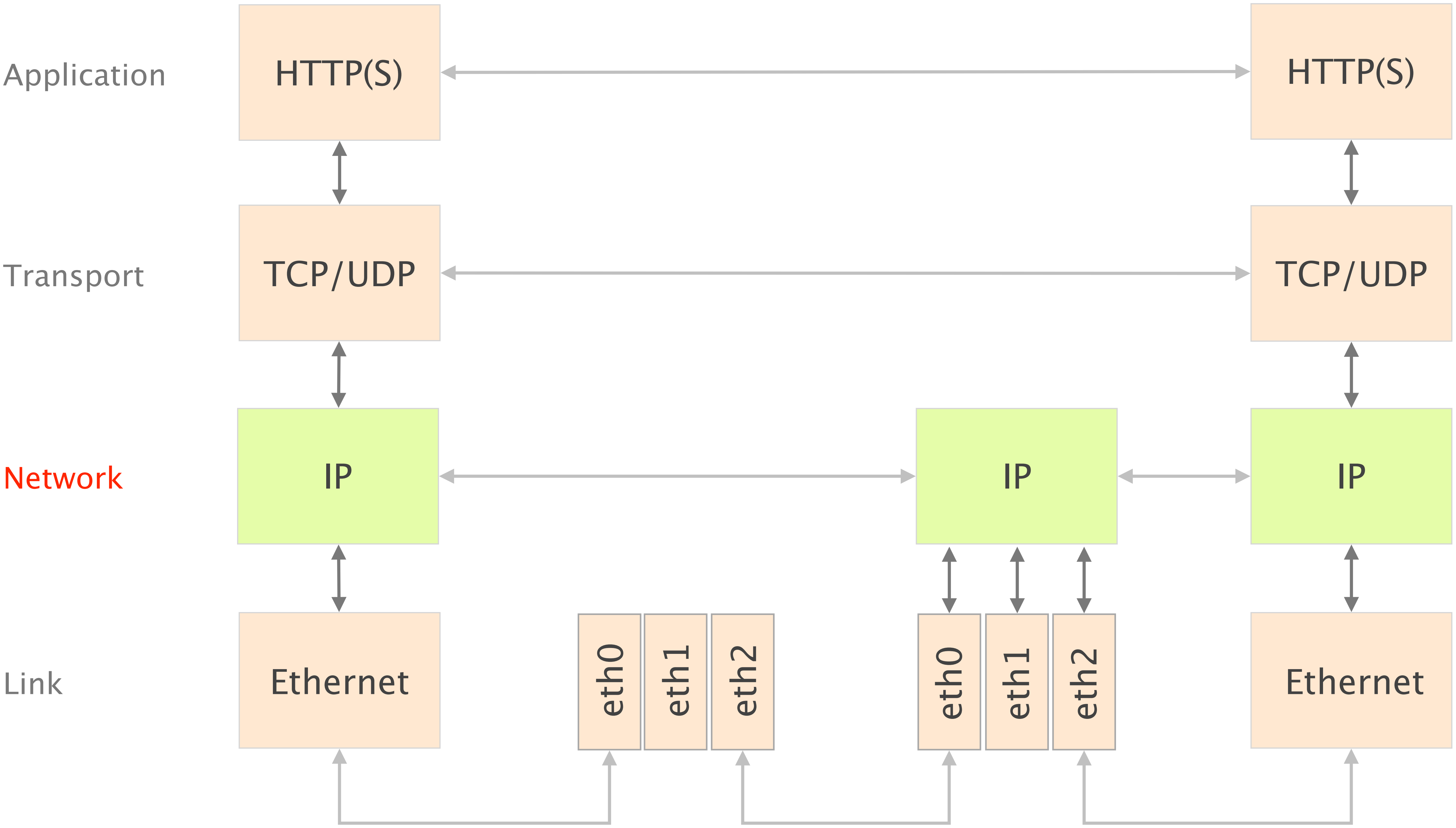
#2 How do we identify link adapters?

#3 How do we share a network medium?

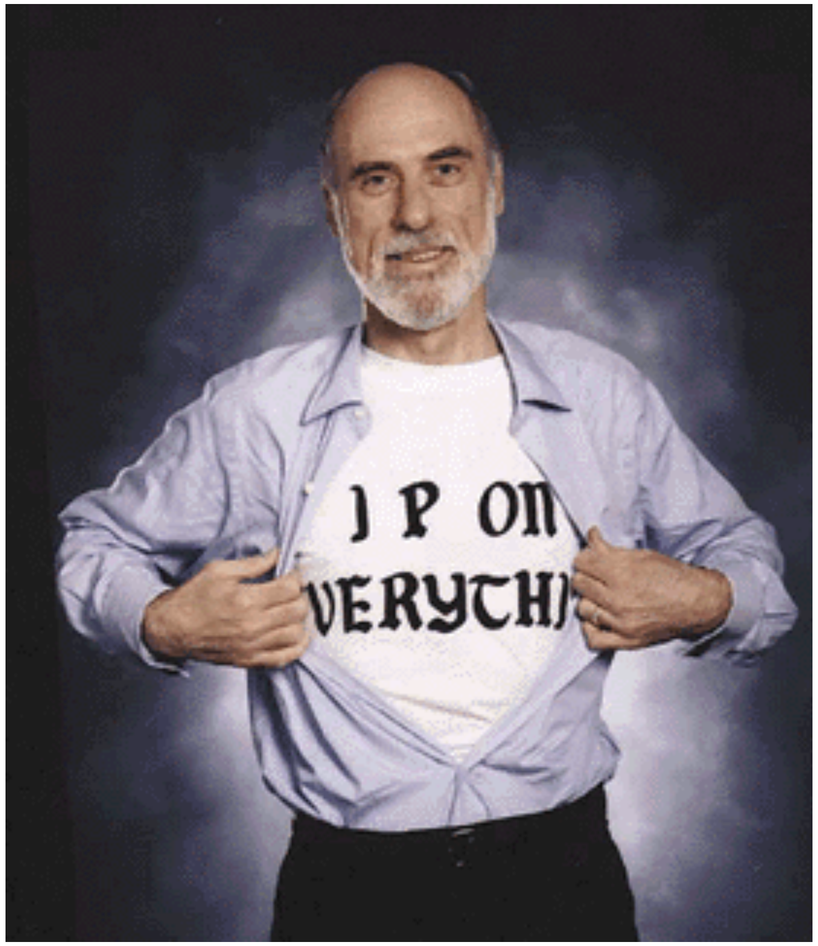
#4 What is Ethernet?

#5 How do we interconnect segments at the link layer?

We then spent multiple weeks on layer 3



Internet Protocol and Forwarding



source: Boardwatch Magazine

1 **IP addresses**

use, structure, allocation

2 **IP forwarding**

longest prefix match rule

3 **IP header**

IPv4 and IPv6, wire format

We also talked about **IPv6**

Internet routing

from here to there, and back

1 **Intra-domain routing**

Link-state protocols

Distance-vector protocols

2 **Inter-domain routing**

Path-vector protocols

Internet routing comes into two flavors: *intra-* and *inter-domain* routing

inter-domain
routing

Find paths between networks

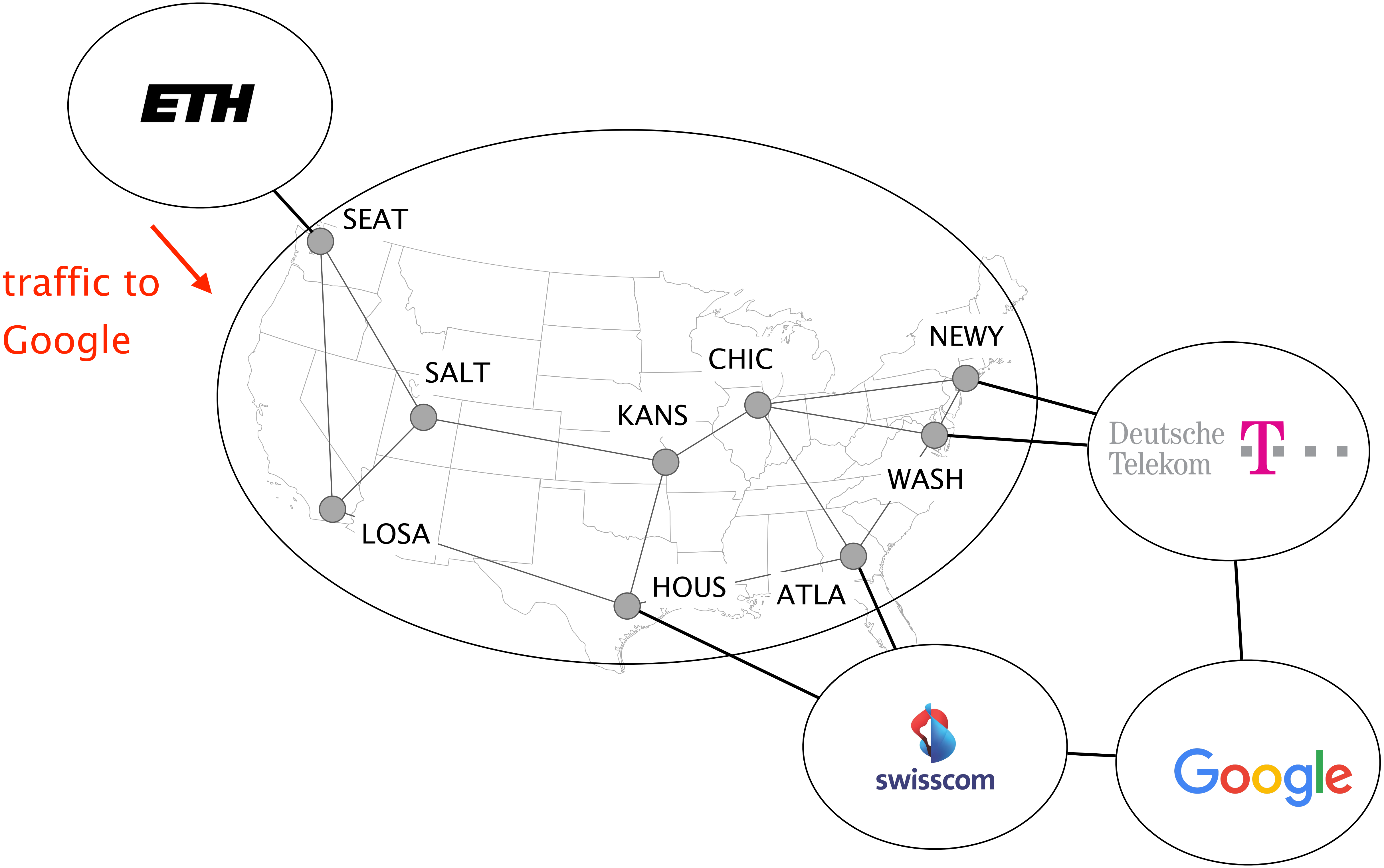
intra-domain
routing

Find paths within a network

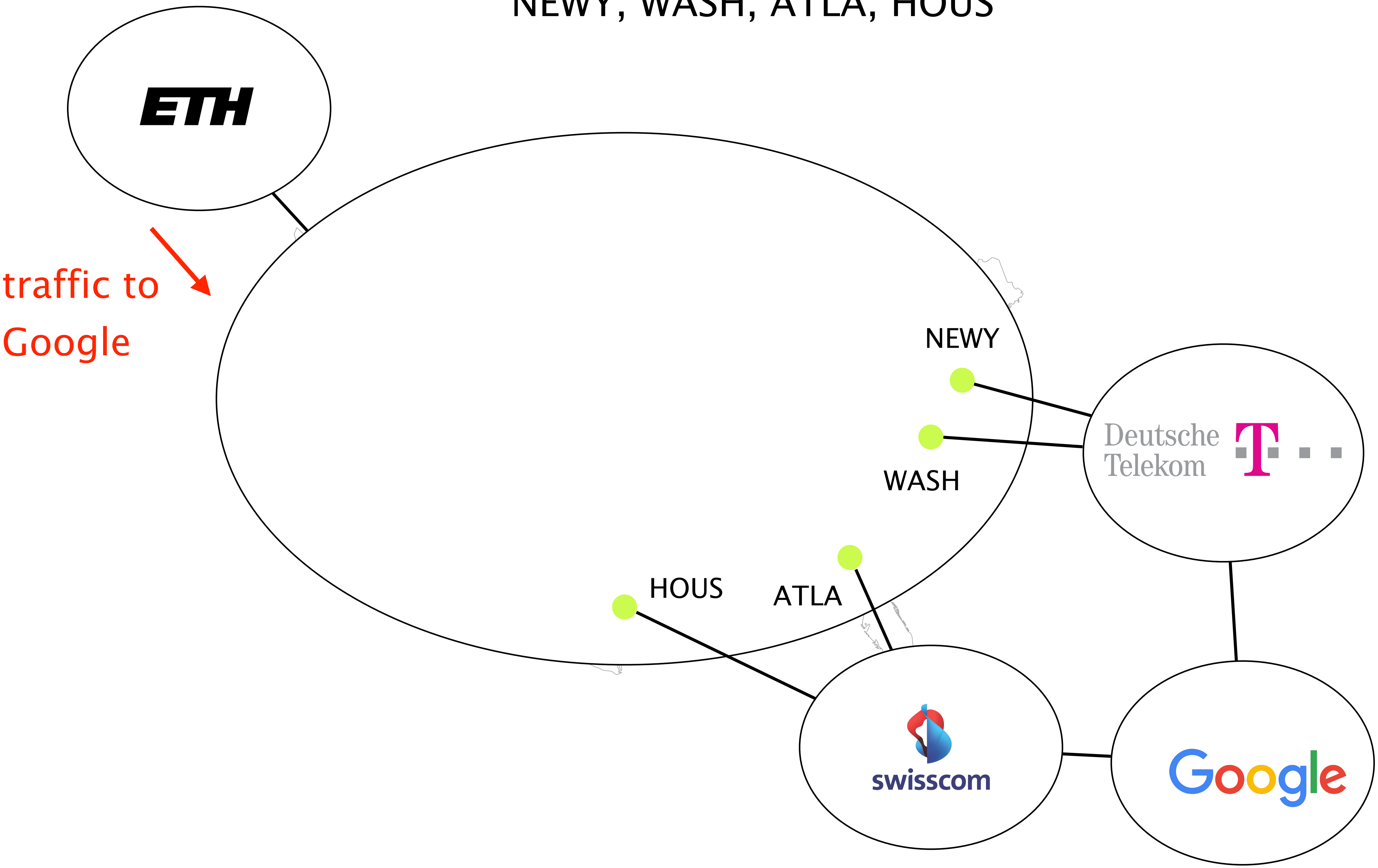
inter-domain
routing

intra-domain
routing

Find paths between networks


ETH
traffic to Google
SEAT
SALT
KANS
CHIC
NEWY
WASH
LOSA
HOUS
ATLA
Deutsche Telekom
T
swisscom
Google

Google can be reached via NEWY, WASH, ATLA, HOUS

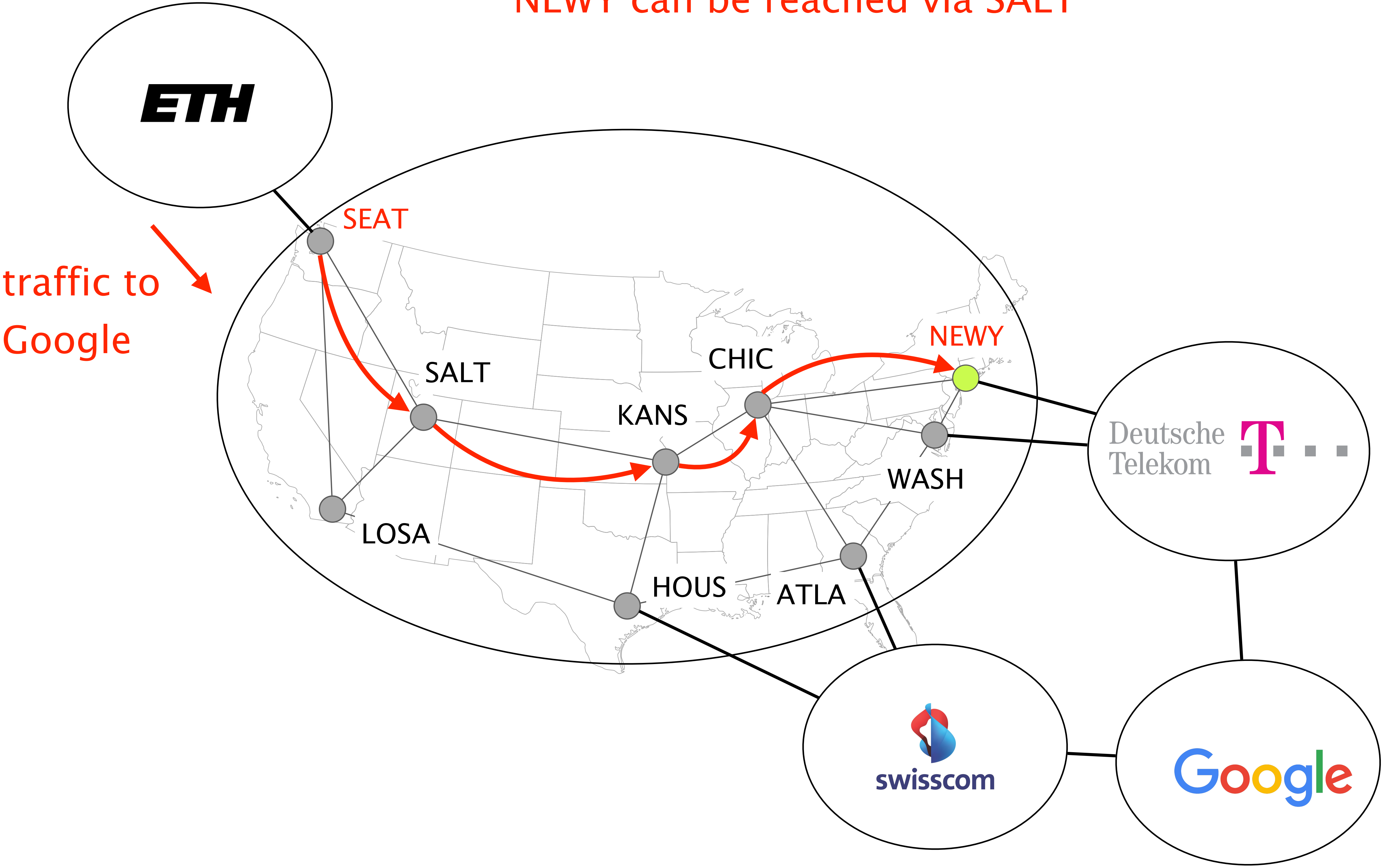


inter-domain
routing

intra-domain
routing

Find paths within a network

NEWY can be reached via SALT



Border Gateway Protocol

policies and more

1 **BGP Policies**

Follow the money

2 **Protocol**

How does it work?

3 **Problems**

security, performance, …

Business relationships conditions
route selection

For a destination *p*, prefer routes coming from

- customers over
- peers over
- providers

route type

Business relationships conditions

route exportation

		send to		
		customer	peer	provider
from	customer			
	peer			
	provider			

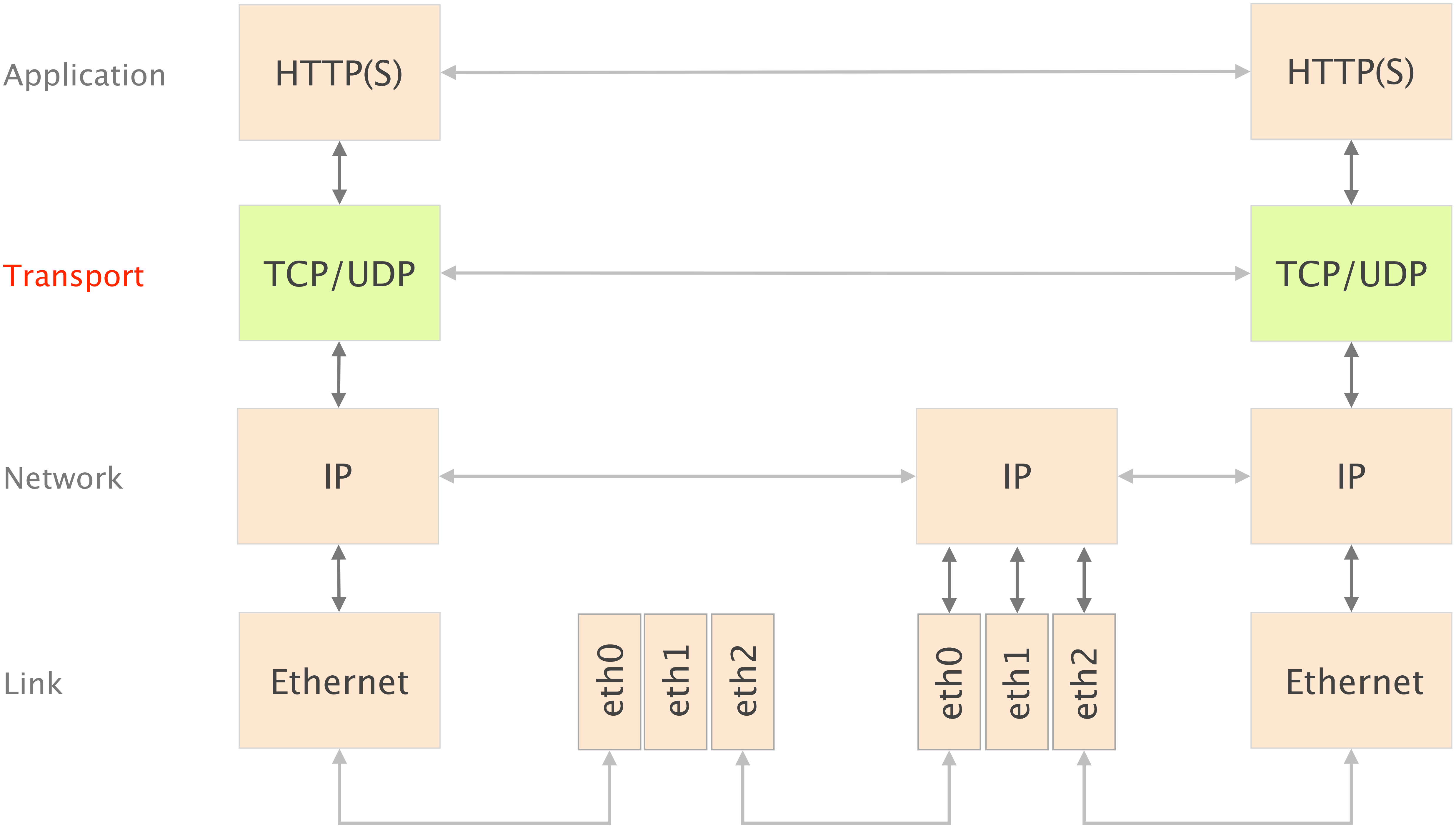
Routes coming from customers are propagated to everyone else

		send to		
		customer	peer	provider
	customer	✓	✓	✓
from	peer			
	provider			

Routes coming from peers and providers are only propagated to customers

		send to		
		customer	peer	provider
from	customer	✓	✓	✓
	peer	✓	-	-
	provider	✓	-	-

4 = 3+1


Application
HTTP(S)
HTTP(S)
Transport
TCP/UDP
TCP/UDP
Network
IP
IP
IP
Link
Ethernet
eth0
eth1
eth2
eth0
eth1
eth2
Ethernet

We looked at the requirements and implementation of transport protocols (UDP/TCP)

Data delivering, to the *correct* application

- IP just points towards next protocol
- *Transport needs to demultiplex incoming data (ports)*

Files or bytestreams abstractions for the applications

- Network deals with packets
- *Transport layer needs to translate between them*

Reliable transfer (if needed)

Not overloading the receiver

Not overloading the network

We then looked at Congestion Control and how it solves three fundamental problems

#1 bandwidth estimation

How to adjust the bandwidth of a single flow to the bottleneck bandwidth?

could be 1 Mbps or 1 Gbps…

#2 bandwidth adaptation

How to adjust the bandwidth of a single flow to variation of the bottleneck bandwidth?

#3 fairness

How to share bandwidth “fairly" among flows, without overloading the network

… by combining two key mechanisms

detecting
congestion

reacting to
congestion

We finally looked at
what’s running on top of all this …

DNS

Web

google.ch ⟷ 172.217.16.131

http://www.google.ch

We finally looked at what's running on top of all this …

Video Streaming

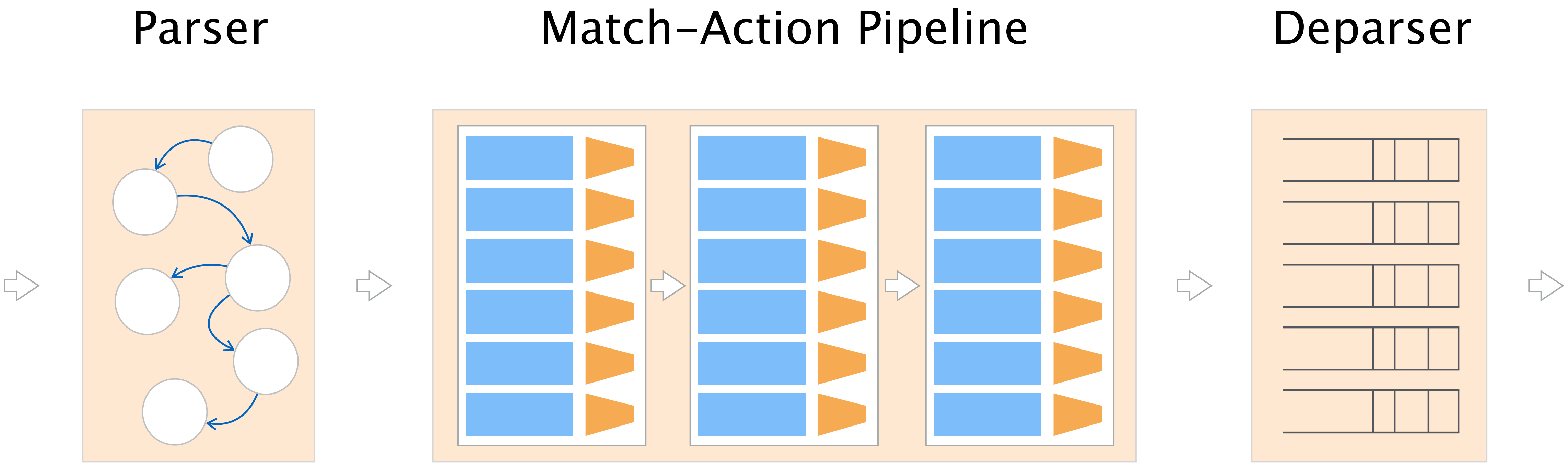
HTTP-based

E-mail

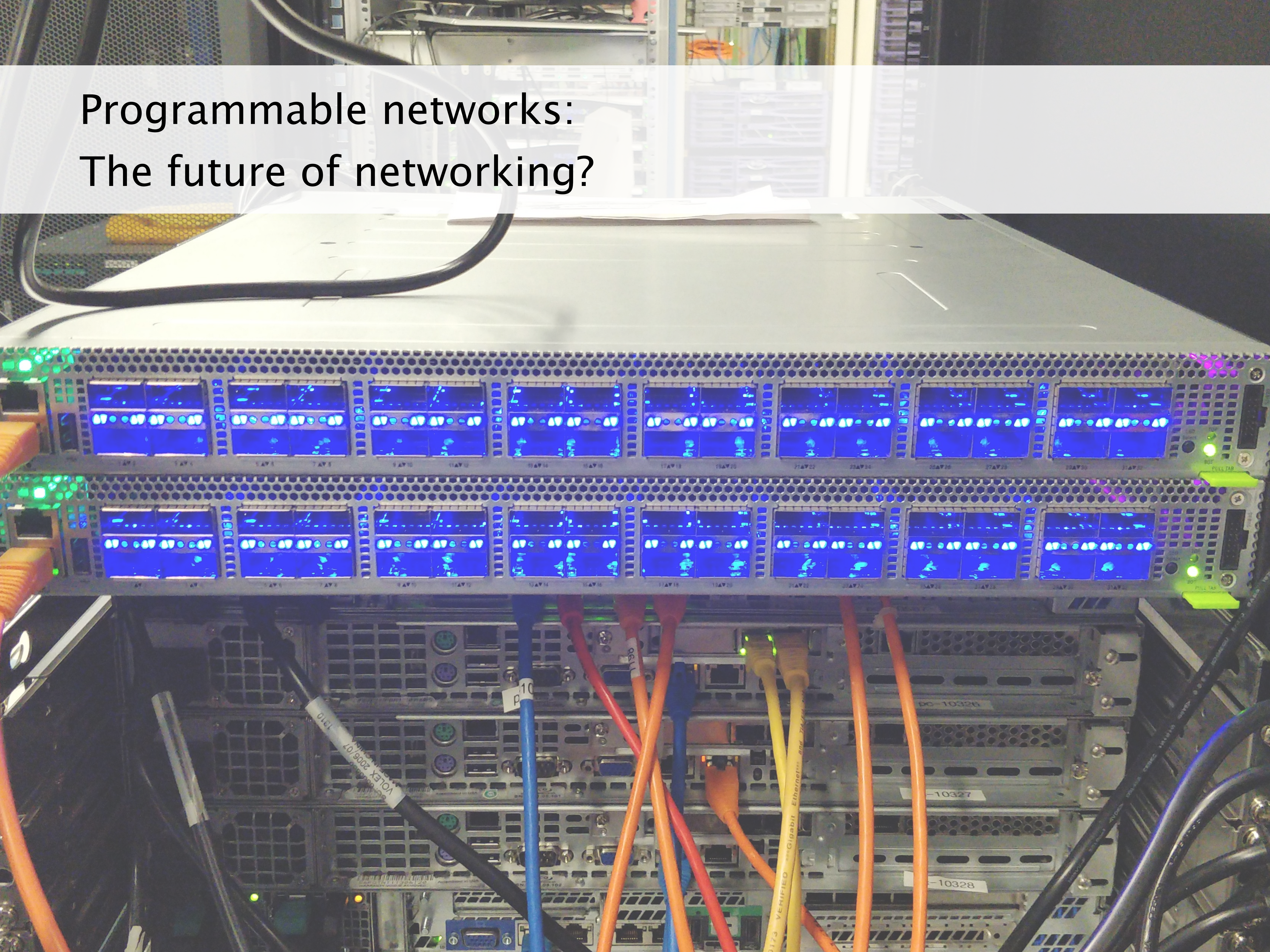
MX, SMTP, POP, IMAP

We finally spoke about network programmability

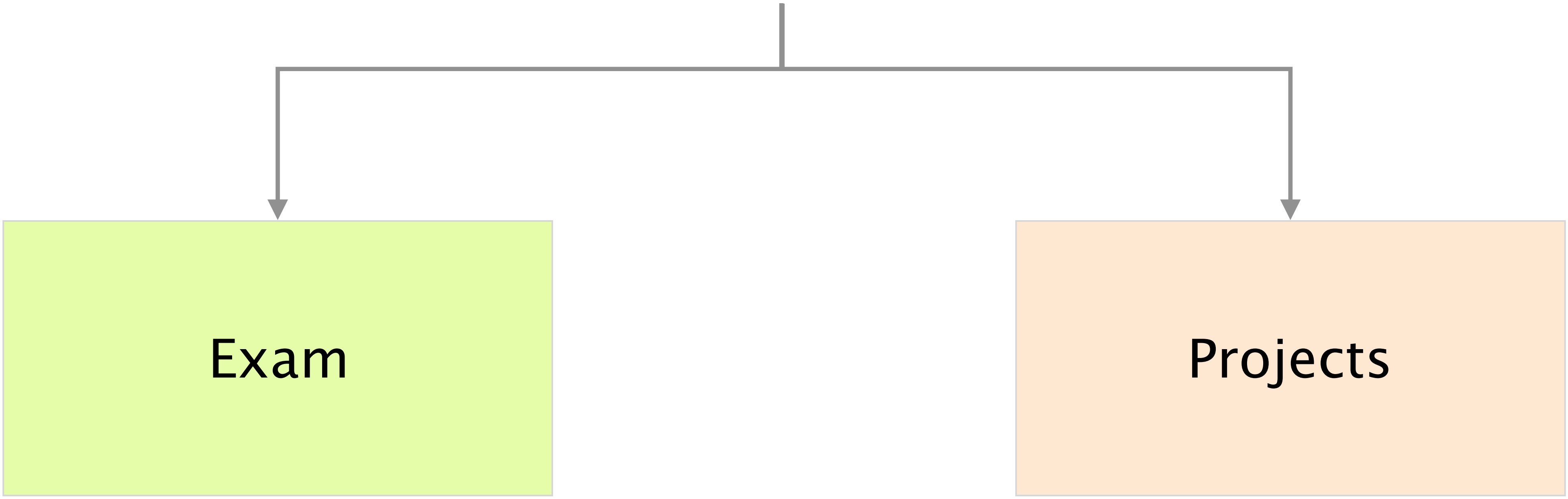
Protocol Independent Switch Architecture (PISA) for high-speed programmable packet forwarding



Programmable networks:
The future of networking?



Your final grade



Exam

Projects

80%

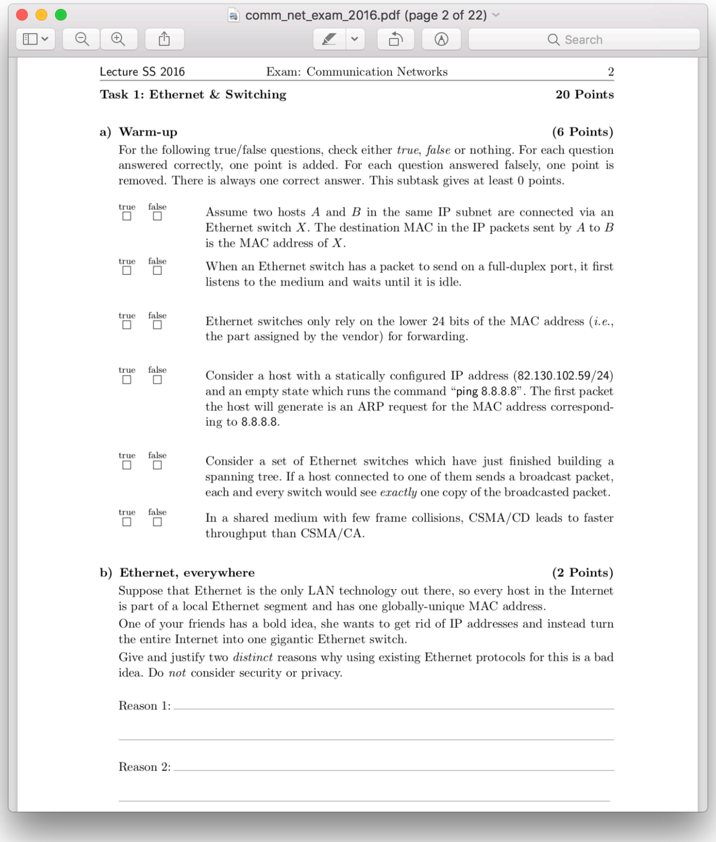
written, open book

20%

The exam will be open book, most of the questions will be open-ended, with some multiple choices

verify your understanding of the material

Make sure you can do *all* the exercises, including the ones in previous exams



Lecture SS 2016 Exam: Communication Networks 2

Task 1: Ethernet & Switching **20 Points**

a) Warm-up **(6 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered falsely, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true □ false □ Assume two hosts A and B in the same IP subnet are connected via an Ethernet switch X. The destination MAC in the IP packets sent by A to B is the MAC address of X.

true □ false □ When an Ethernet switch has a packet to send on a full-duplex port, it first listens to the medium and waits until it is idle.

true □ false □ Ethernet switches only rely on the lower 24 bits of the MAC address (*i.e.*, the part assigned by the vendor) for forwarding.

true □ false □ Consider a host with a statically configured IP address (82.130.102.59/24) and an empty state which runs the command "ping 8.8.8.8". The first packet the host will generate is an ARP request for the MAC address corresponding to 8.8.8.8.

true □ false □ Consider a set of Ethernet switches which have just finished building a spanning tree. If a host connected to one of them sends a broadcast packet, each and every switch would see *exactly* one copy of the broadcasted packet.

true □ false □ In a shared medium with few frame collisions, CSMA/CD leads to faster throughput than CSMA/CA.

b) Ethernet, everywhere **(2 Points)**

Suppose that Ethernet is the only LAN technology out there, so every host in the Internet is part of a local Ethernet segment and has one globally-unique MAC address.

One of your friends has a bold idea, she wants to get rid of IP addresses and instead turn the entire Internet into one gigantic Ethernet switch.

Give and justify two *distinct* reasons why using existing Ethernet protocols for this is a bad idea. Do *not* consider security or privacy.

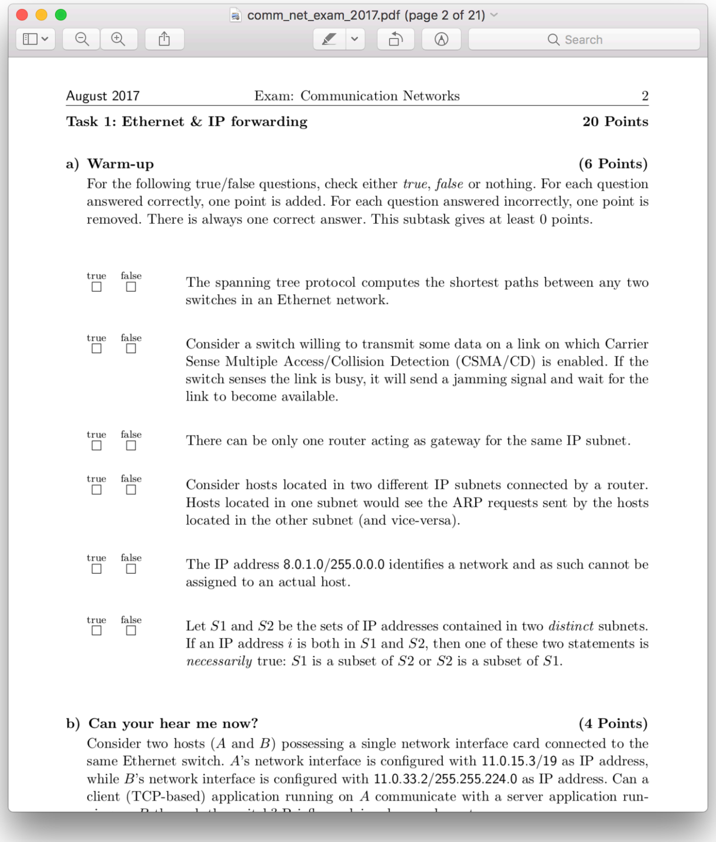
Reason 1: ______________________________

____

Reason 2: ______________________________

____

Millesime **2016**



August 2017 Exam: Communication Networks 2

Task 1: Ethernet & IP forwarding **20 Points**

a) Warm-up **(6 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true □ false □ The spanning tree protocol computes the shortest paths between any two switches in an Ethernet network.

true □ false □ Consider a switch willing to transmit some data on a link on which Carrier Sense Multiple Access/Collision Detection (CSMA/CD) is enabled. If the switch senses the link is busy, it will send a jamming signal and wait for the link to become available.

true □ false □ There can be only one router acting as gateway for the same IP subnet.

true □ false □ Consider hosts located in two different IP subnets connected by a router. Hosts located in one subnet would see the ARP requests sent by the hosts located in the other subnet (and vice-versa).

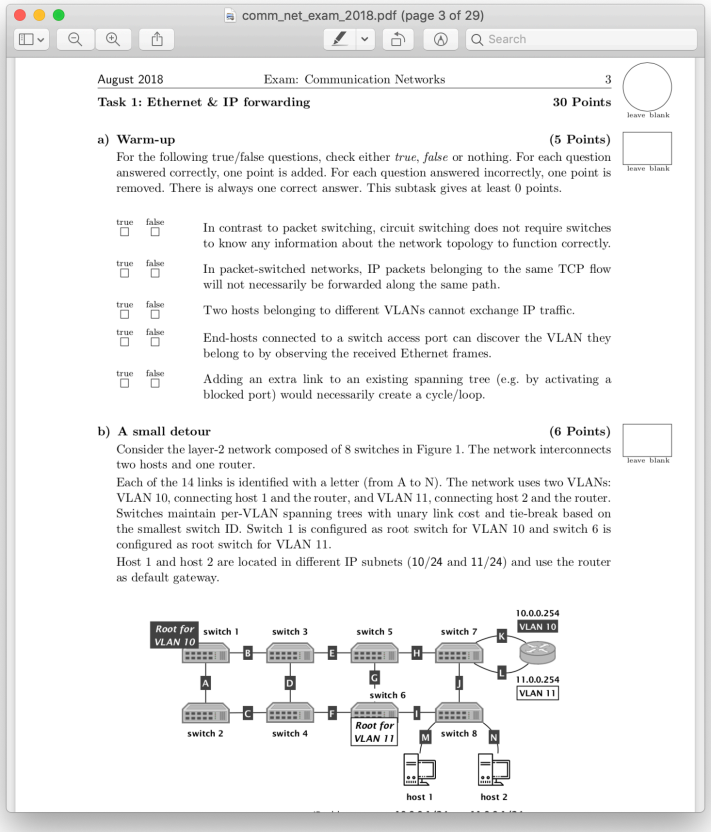
true □ false □ The IP address 8.0.1.0/255.0.0.0 identifies a network and as such cannot be assigned to an actual host.

true □ false □ Let $S1$ and $S2$ be the sets of IP addresses contained in two *distinct* subnets. If an IP address i is both in $S1$ and $S2$, then one of these two statements is *necessarily* true: $S1$ is a subset of $S2$ or $S2$ is a subset of $S1$.

b) Can your hear me now? **(4 Points)**

Consider two hosts (A and B) possessing a single network interface card connected to the same Ethernet switch. A's network interface is configured with 11.0.15.3/19 as IP address, while B's network interface is configured with 11.0.33.2/255.255.224.0 as IP address. Can a client (TCP-based) application running on A communicate with a server application run-

Millesime **2017**



August 2018 Exam: Communication Networks 3

Task 1: Ethernet & IP forwarding **30 Points**

a) Warm-up **(5 Points)**

For the following true/false questions, check either *true*, *false* or nothing. For each question answered correctly, one point is added. For each question answered incorrectly, one point is removed. There is always one correct answer. This subtask gives at least 0 points.

true □ false □ In contrast to packet switching, circuit switching does not require switches to know any information about the network topology to function correctly.

true □ false □ In packet-switched networks, IP packets belonging to the same TCP flow will not necessarily be forwarded along the same path.

true □ false □ Two hosts belonging to different VLANs cannot exchange IP traffic.

true □ false □ End-hosts connected to a switch access port can discover the VLAN they belong to by observing the received Ethernet frames.

true □ false □ Adding an extra link to an existing spanning tree (e.g. by activating a blocked port) would necessarily create a cycle/loop.

b) A small detour **(6 Points)**

Consider the layer-2 network composed of 8 switches in Figure 1. The network interconnects two hosts and one router.

Each of the 14 links is identified with a letter (from A to N). The network uses two VLANs: VLAN 10, connecting host 1 and the router, and VLAN 11, connecting host 2 and the router. Switches maintain per-VLAN spanning trees with unary link cost and tie-break based on the smallest switch ID. Switch 1 is configured as root switch for VLAN 10 and switch 6 is configured as root switch for VLAN 11.

Host 1 and host 2 are located in different IP subnets (10/24 and 11/24) and use the router as default gateway.

Millesime **2018**

https://comm-net.ethz.ch/#tab-exam

Don't forget the assignments, *they matter*

No programming question no Python at the exam

but we could ask you to describe a procedure in English

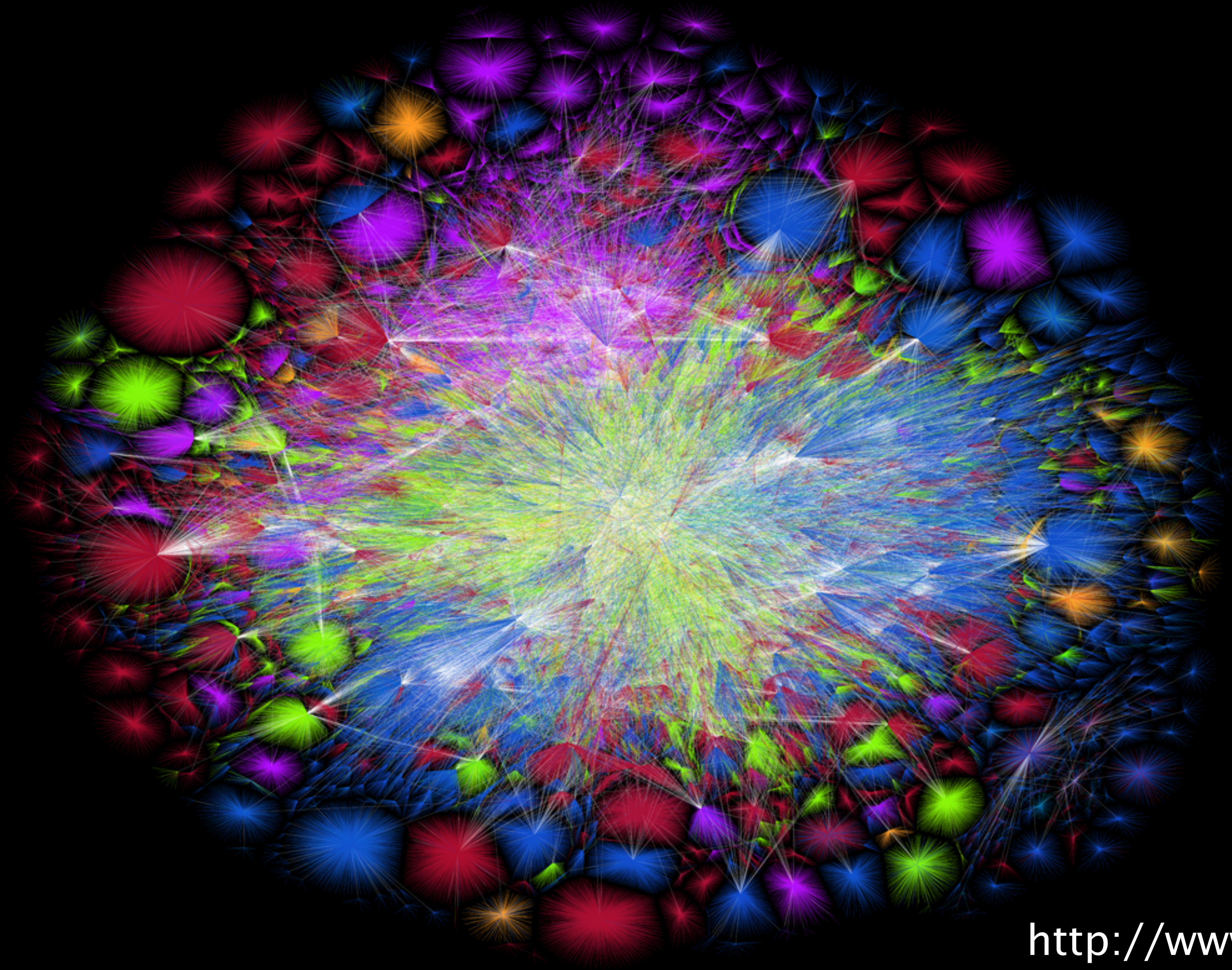
What would you change in your solution to achieve *X*?

No configuration question no Quagga at the exam

but we could ask you to describe a configuration in English

How would you realize policy *X*?

Now you (better) understand this!



http://www.opte.org

Communication Networks

Spring 2019

Laurent Vanbever

nsg.ee.ethz.ch

ETH Zürich (D-ITET)

May 20 2019