High performance parallel systems

Lecture 7 – Computer Networks

Kenneth Skovhede, NBI, 2020-12-08

UNIVERSITY OF COPENHAGEN

A layered approach

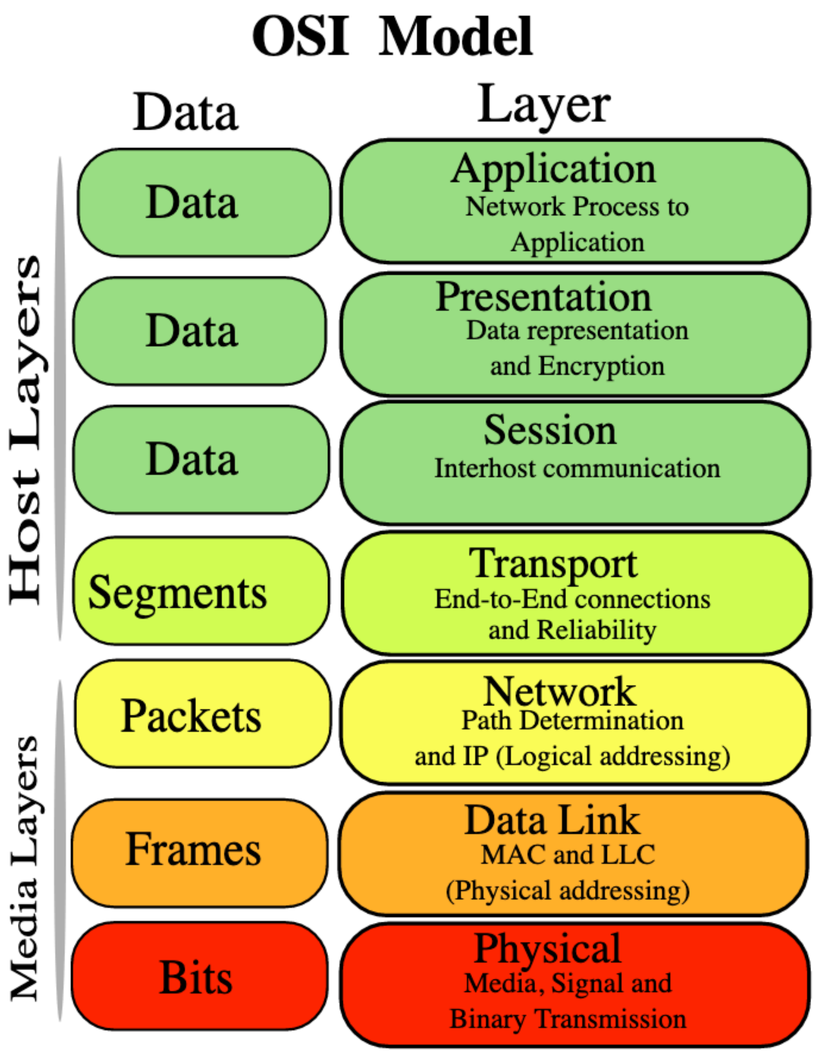


Image from: https://commons.wikimedia.org/wiki/File:Osi-model-jb.svg

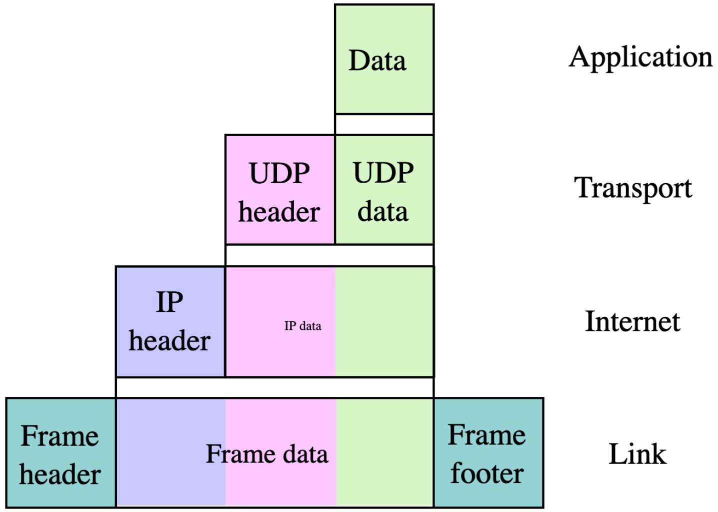
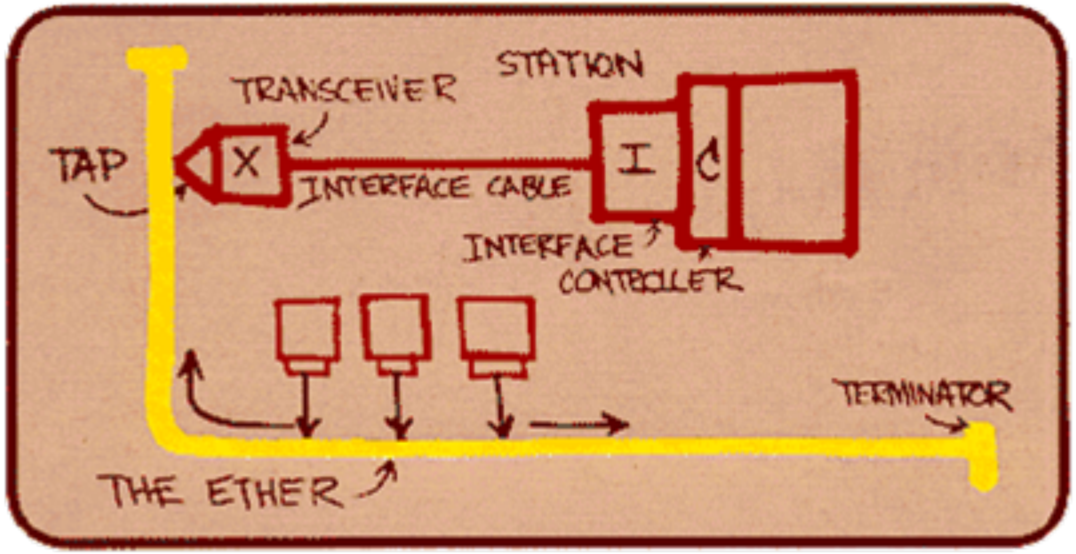


Image from: https://commons.wikimedia.org/wiki/File:UDP_encapsulation.svg

Ethernet – From Coax to present day



Drawing by Robert M. Metcalfe in 1976

Sharing a cable

https://www.youtube.com/watch?v=ajh1eZUVuCk

Carrier-sense multiple access with collision avoidance

CSMA/CA

Using exponential backoff

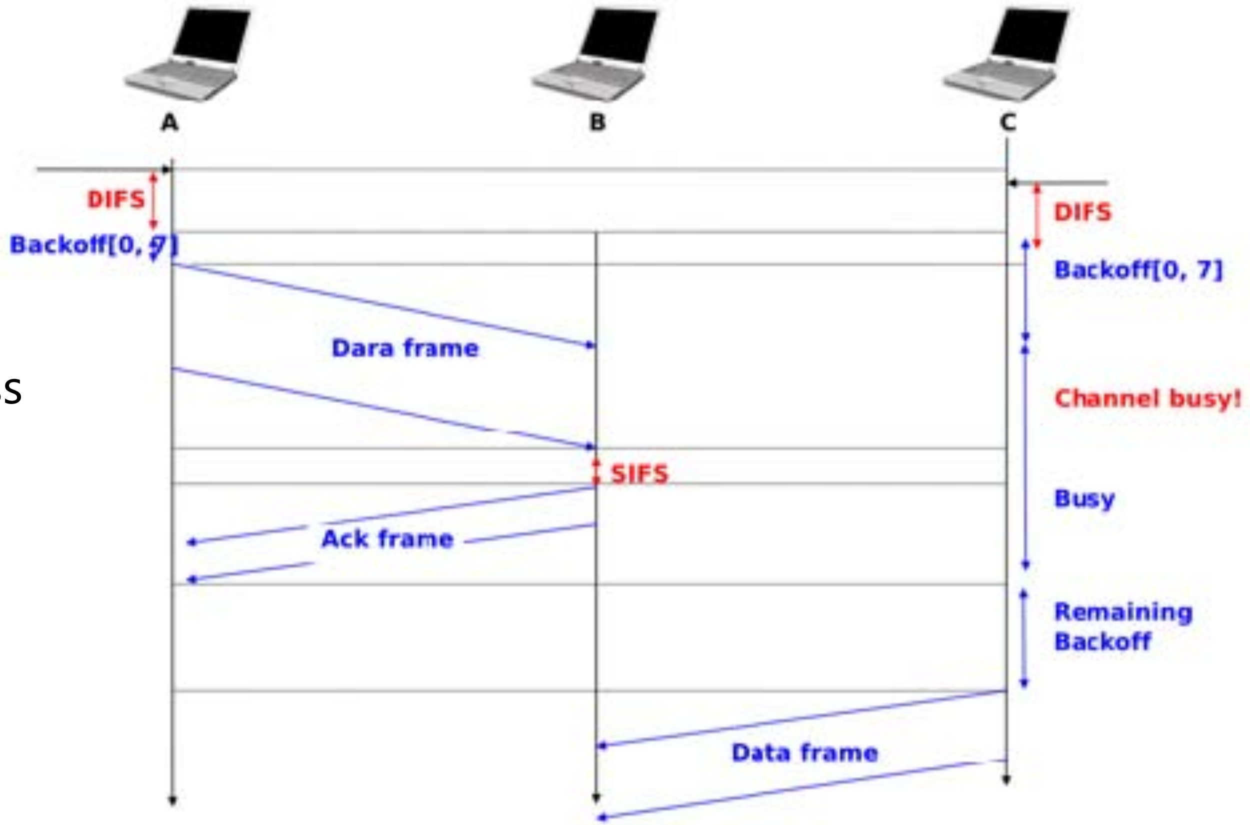
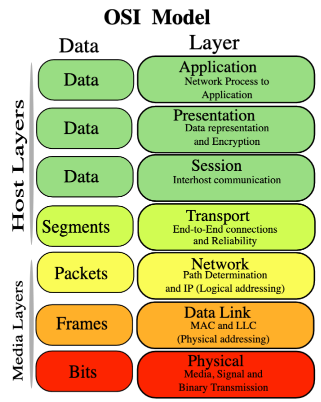
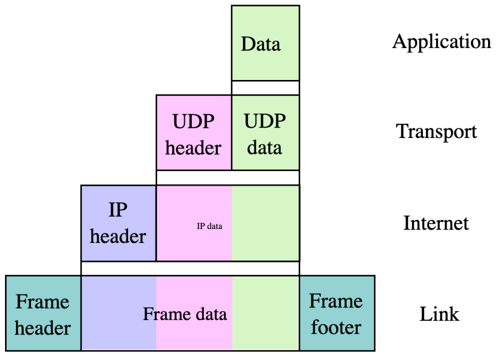


Image from: http://www.opentextbooks.org.hk/ditatopic/3611

Sending a packet over ethernet cable





Preamble
AA AA AA AA AA AA AB

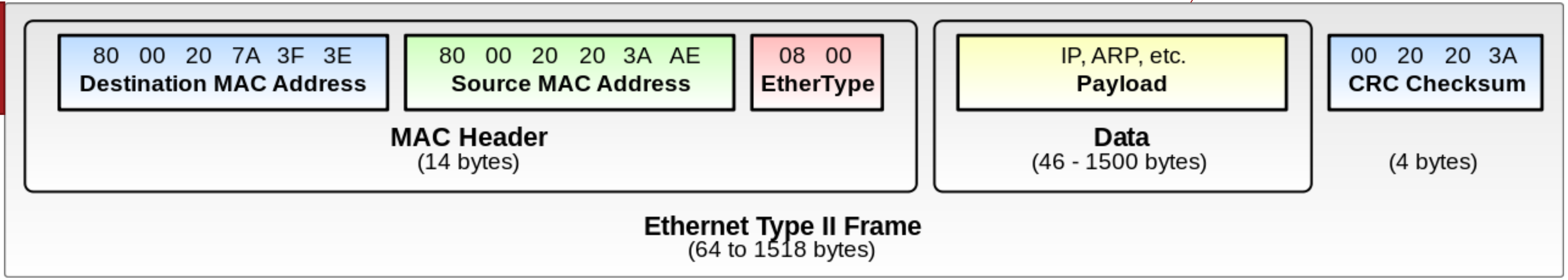
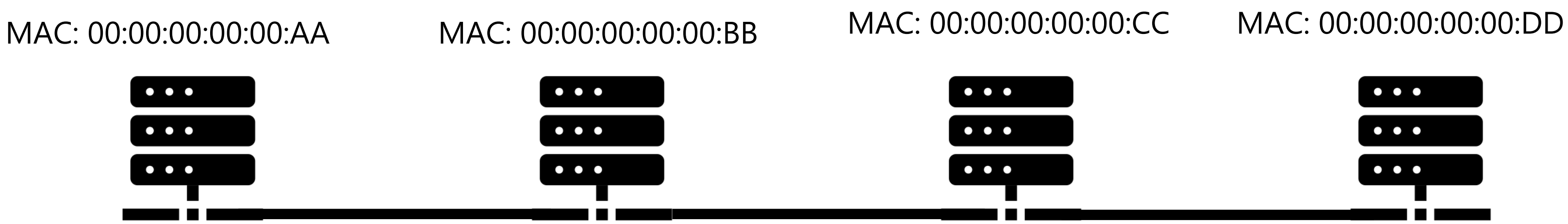


Image from: https://commons.wikimedia.org/wiki/File:Ethernet_Type_II_Frame_format.svg

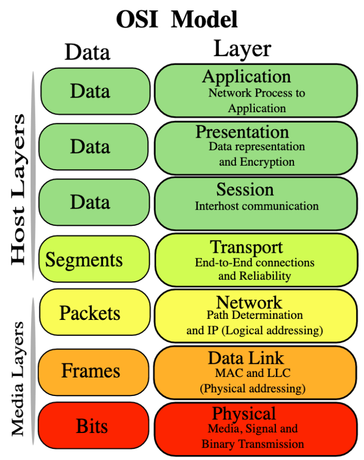
10101010 10101010 10101010 10101010 10101010 10101010 10101010 10101011

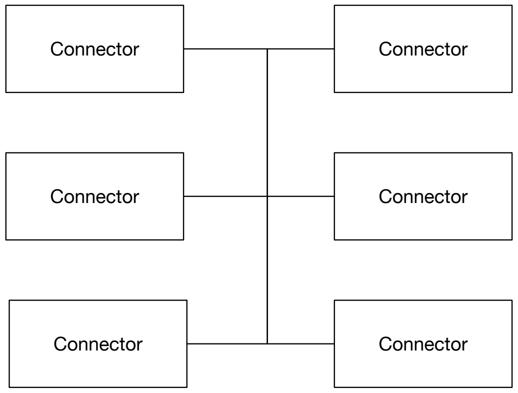
Sending messages on a local network

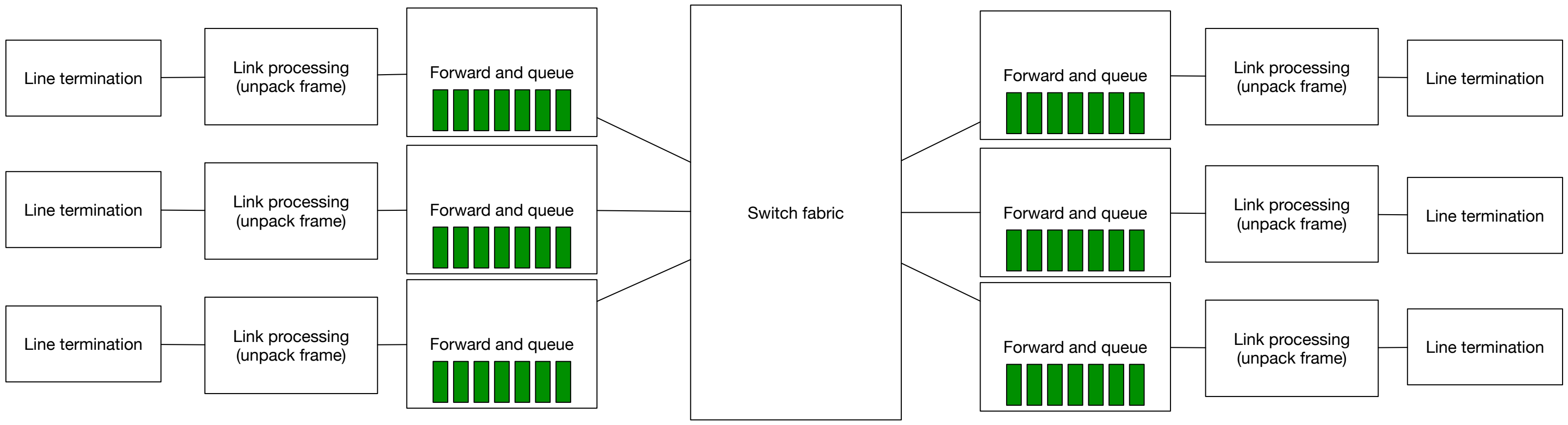



```
FROM:     ...:AA
TO:       ...:CC
Payload:  ......
```

Connecting ethernet – Switches and Hubs







Ethernet – One protocol fits all (speeds)

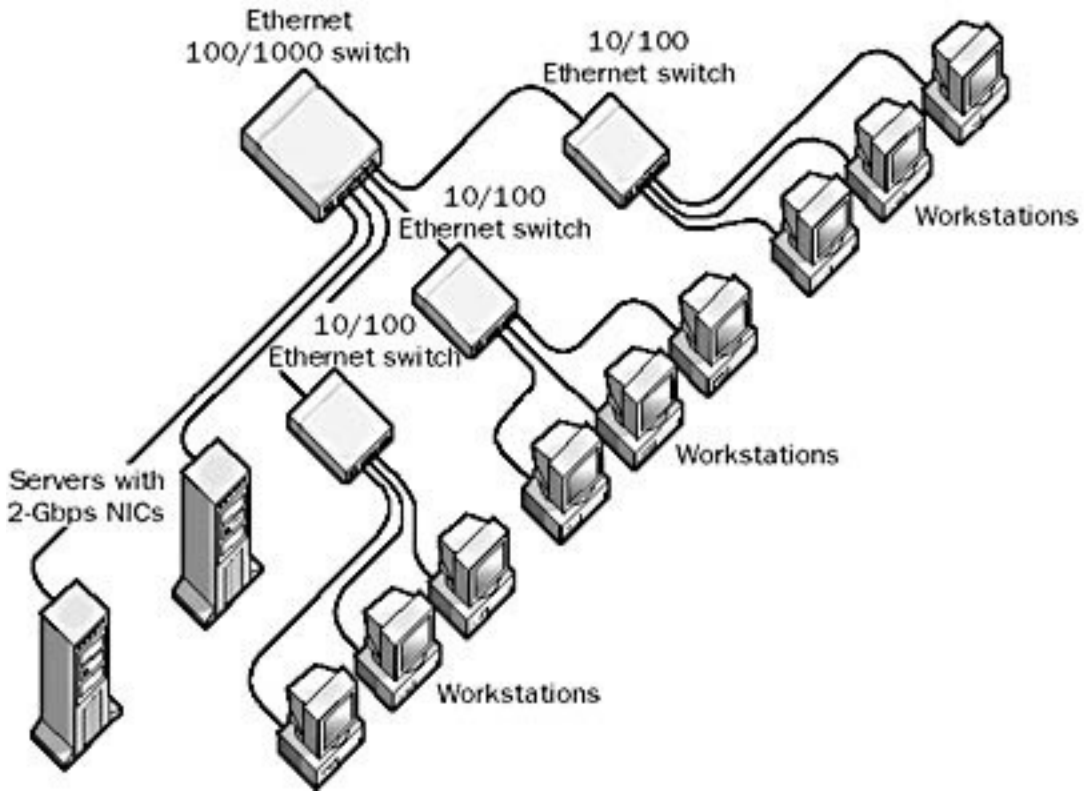
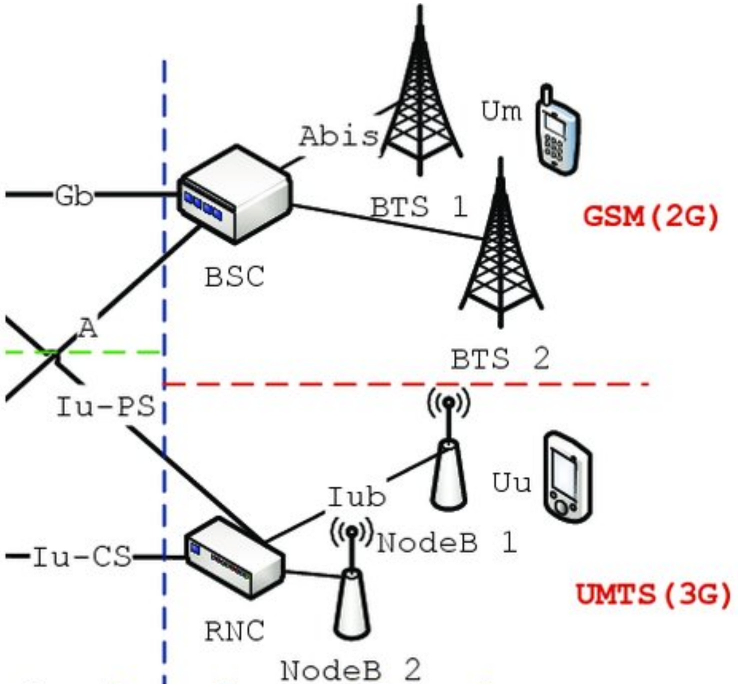
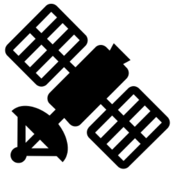
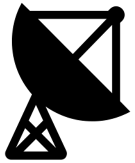


Image from: https://networkencyclopedia.com/gigabit-ethernet/

Alternatives to Ethernet: WiFi, 3/4/5G, etc







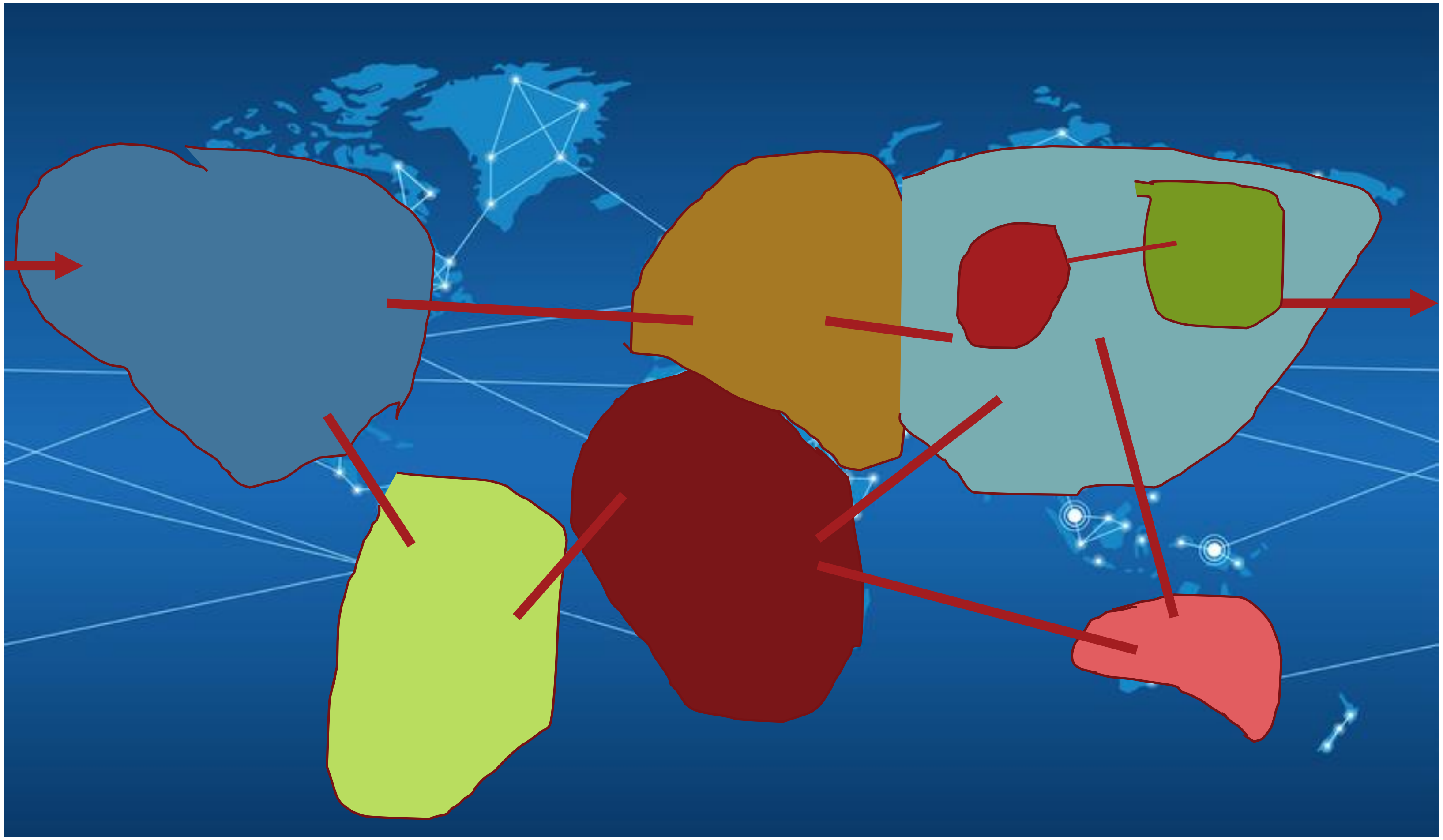
CSMA/CA

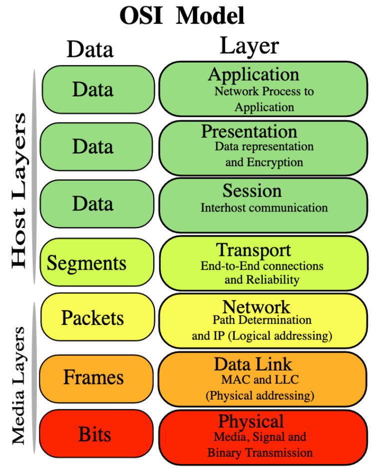
From: DOI: 10.1109/ISCIT.2012.6381046

Internet Protocol

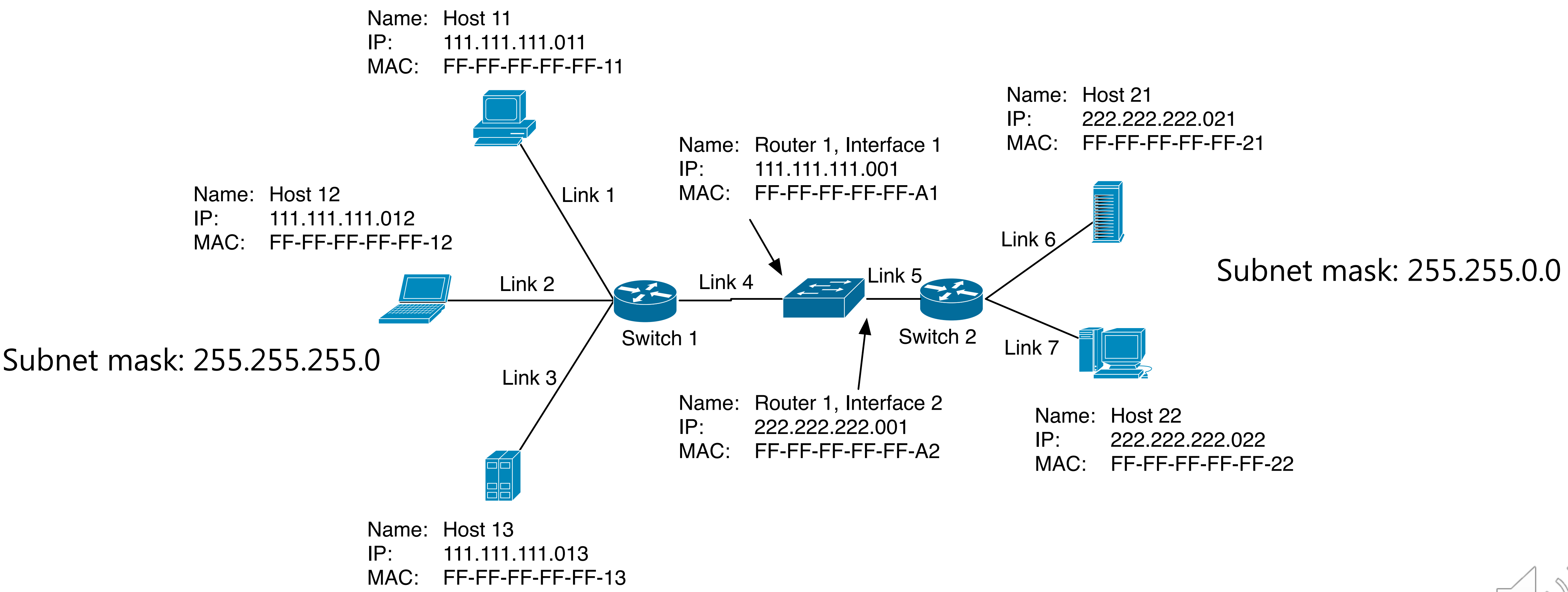
Image from: https://www.vecteezy.com/vector-art/376244-seamless-map-of-the-global-network-system

Internet Protocol

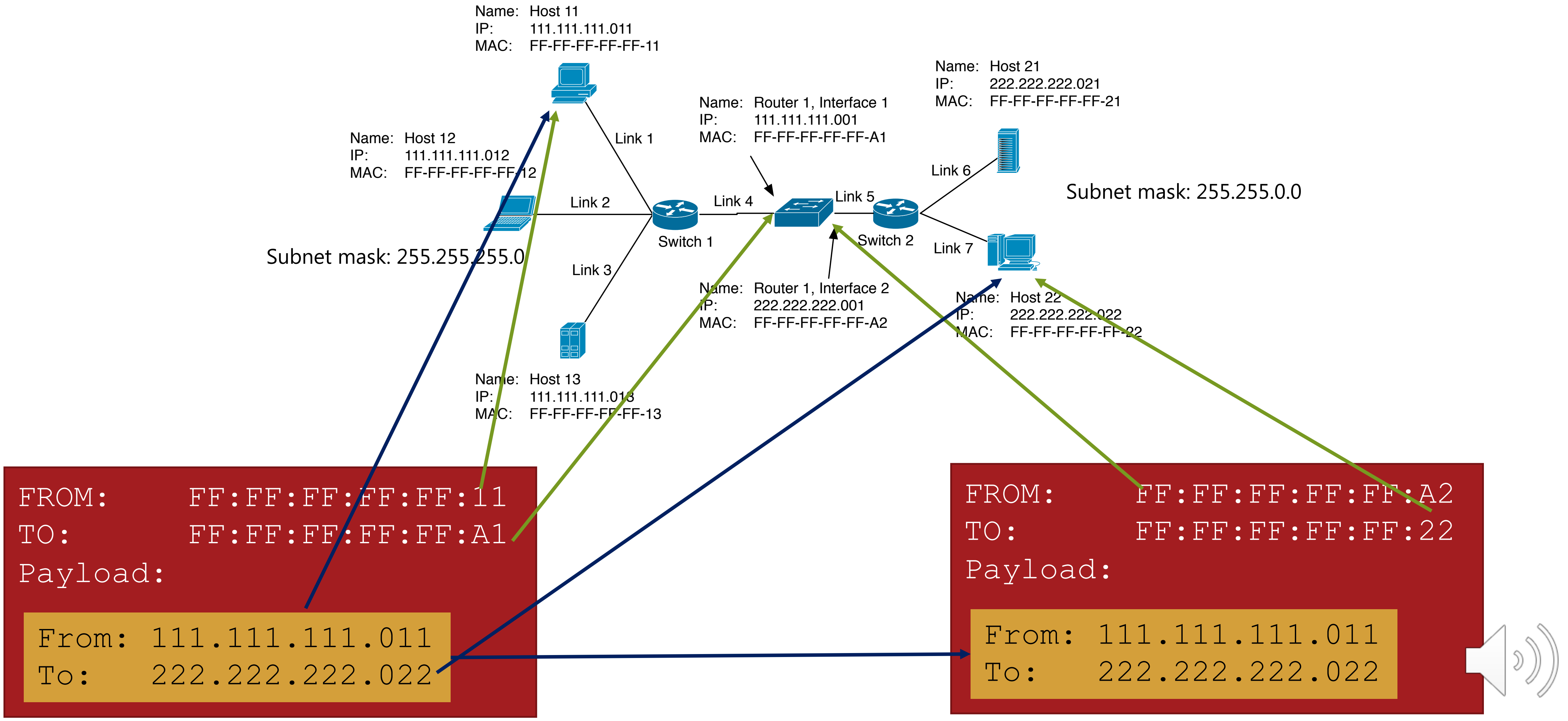




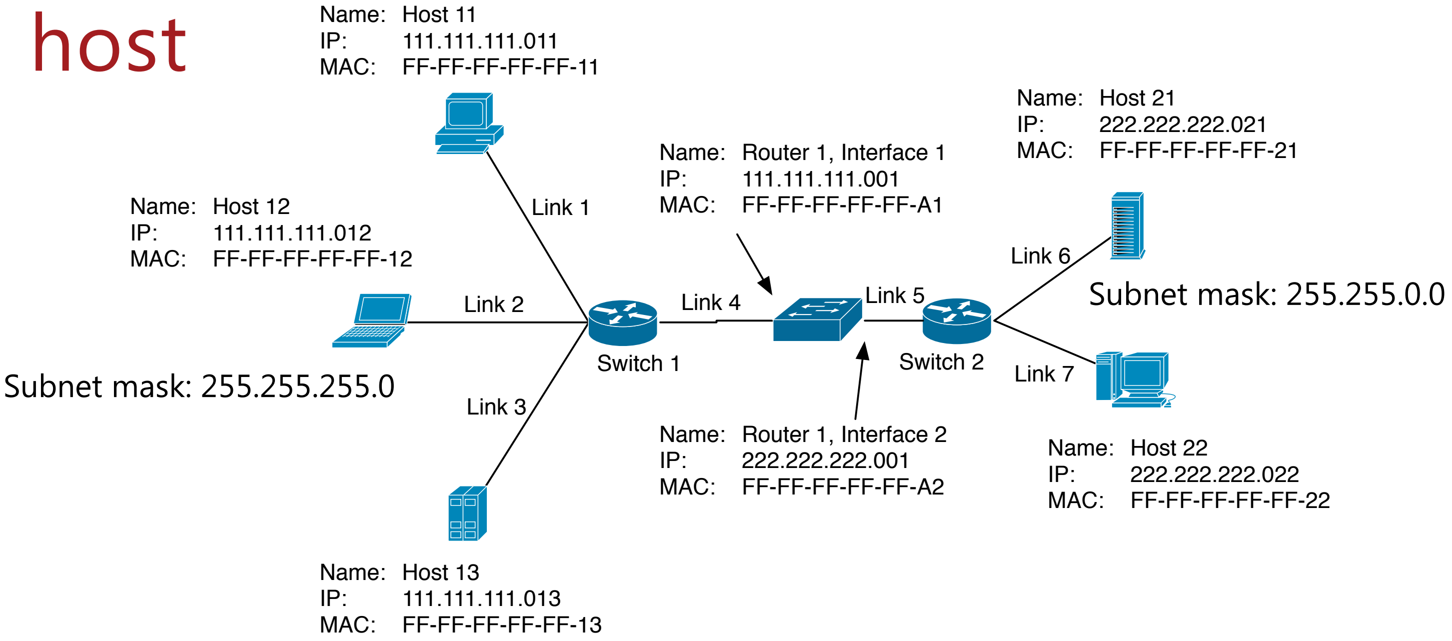
Introducing routers and IP addresses



Routing with subnets



Routing as seen from a host



Own IP:	111.111.111.011	111.111.111.011
Target IP:	111.111.111.012	222.222.222.022
Subnet mask:	255.255.255.0	255.255.255.0
Own masked:	111.111.111.0	111.111.111.0
Target masked:	111.111.111.0	222.222.222.0
	Own subnet	*Other subnet*

Special IP addresses

CIDR block	Subnet mask	Comment
10.0.0.0/8	255.0.0.0	Private network
127.0.0.0/8	255.0.0.0	Local host - 127.0.0.1
172.16.0.0/12	255.240.0.0	Private network
192.168.0.0/16	255.255.0.0	Private network

Try it yourself – IP addresses

A host has IP 192.168.0.2 and subnet mask 255.255.255.240.
For each of these target IPs figure out if it sends to the router:

- 192.168.0.12
- 192.168.0.63
- 192.168.1.03

How many hosts can be in a subnet with these ranges:

- 10.0.0.0/8
- 172.16.0.0/12
- 192.168.0.0/16
- 255.0.0.0

What is the smallest IP and largest IP in these ranges:

- 10.0.0.0/8
- 172.16.0.0/12
- 192.168.0.0/16
- 88.5.6.7 subnet mask 255.0.0.0

Try it yourself – IP addresses

A host has IP 192.168.0.2 and subnet mask 255.255.255.240.
For each of these target IPs figure out if it sends to the router:

- 192.168.0.12 Local
- 192.168.0.63 Router
- 192.168.1.03 Router

240	=	0b1111 0000
2	=	0b0000 ~~0010~~
12	=	0b0000 1100
63	=	0b0011 1111
3	=	0b0000 0011

How many hosts can be in a subnet with these ranges:

- 10.0.0.0/8 $2^{(32-8)}$ = 16.777.216 (- 2)
- 172.16.0.0/12 $2^{(32-12)}$ = 1.048.576 (- 2)
- 192.168.0.0/16 $2^{(32-16)}$ = 65.536 (- 2)
- 255.0.0.0 2^{24} = 16.777.216 (- 2)

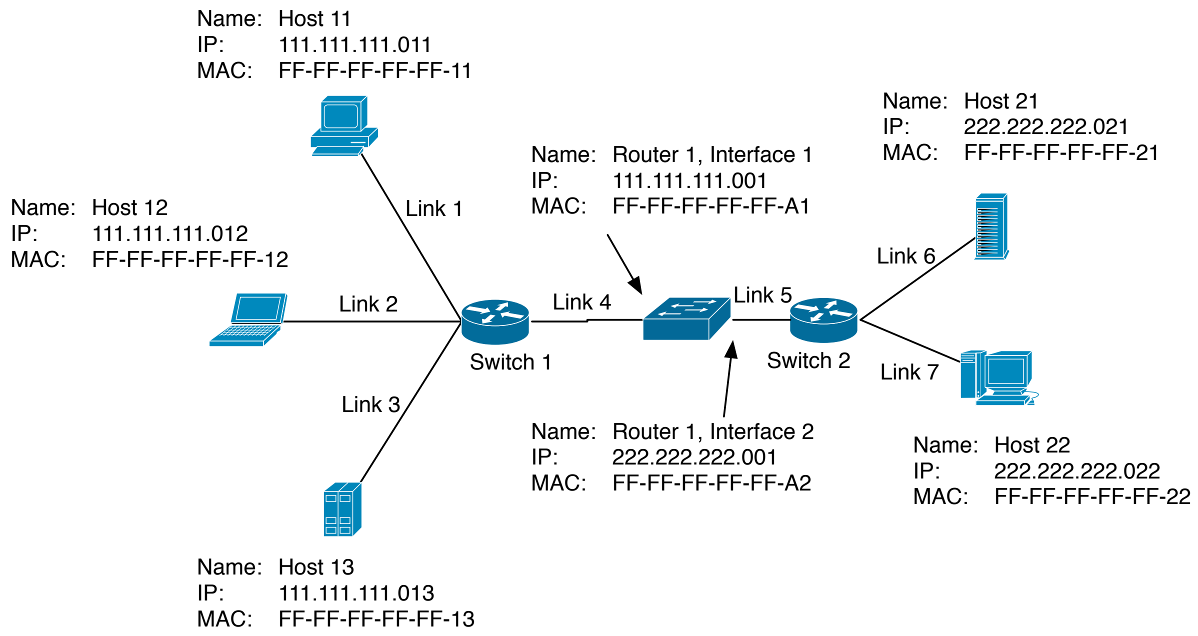
What is the smallest IP and largest IP in these ranges:

- 10.0.0.0/8
- 172.16.0.0/12
- 192.168.0.0/16
- 88.5.6.7 subnet mask 255.0.0.0

10.0.0.0 – 10.255.255.255
172.16.0.0 – 172.31.255.255
192.168.0.0 - 192.168.255.255
88.0.0.0 – 88.255.255.255

DHCP – Dynamic setup

Response contains:
- IP Address
- Subnet mask
- Gateway IP
- DNS hosts
- Network parameters



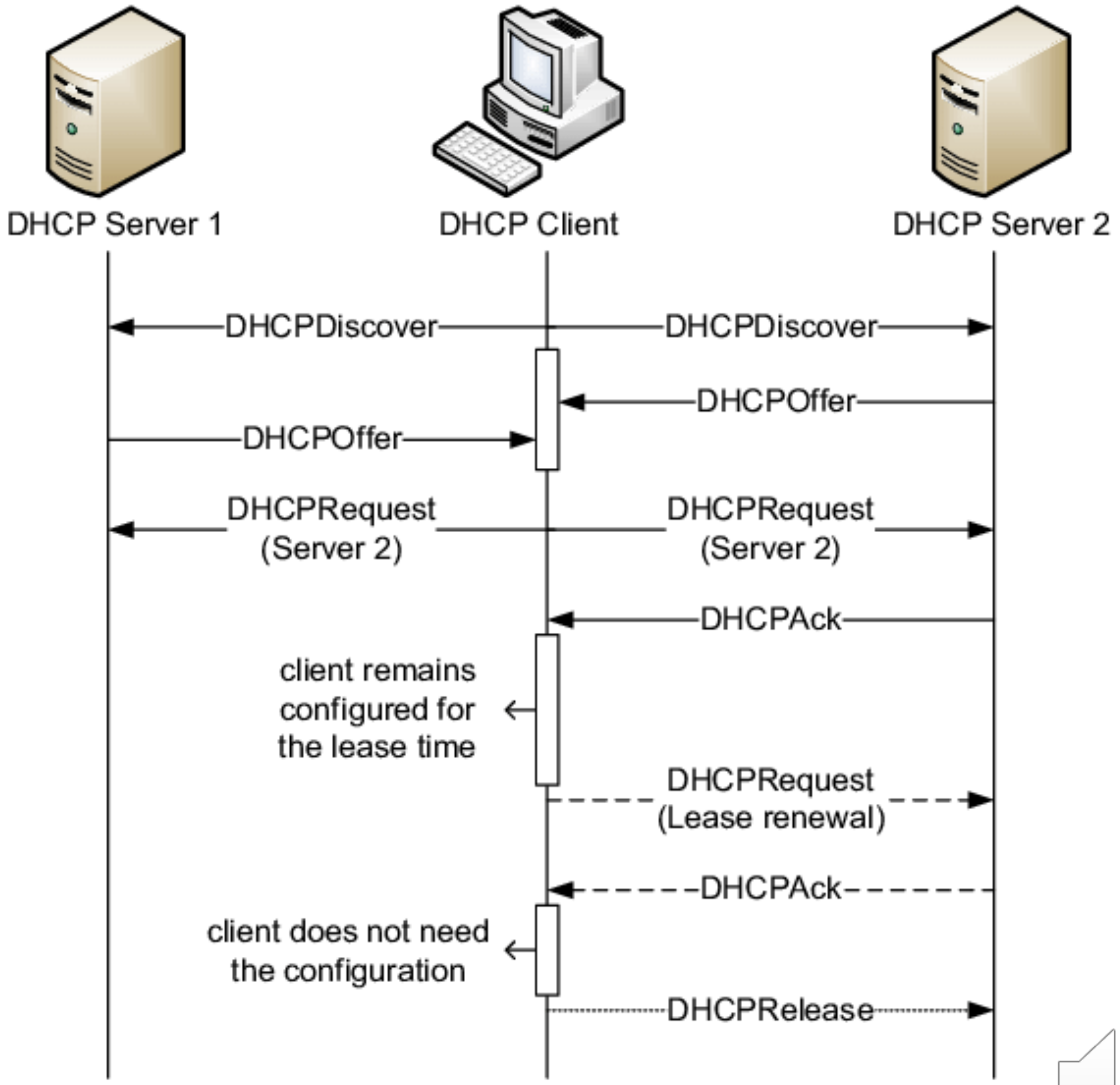
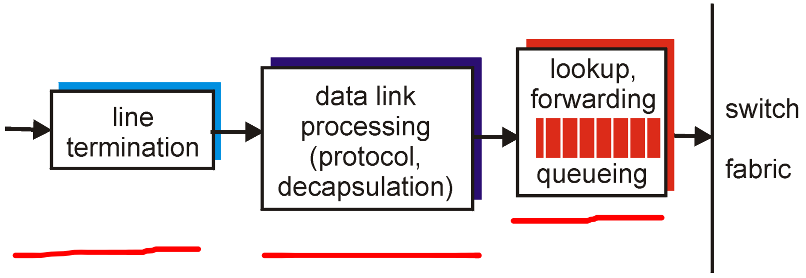
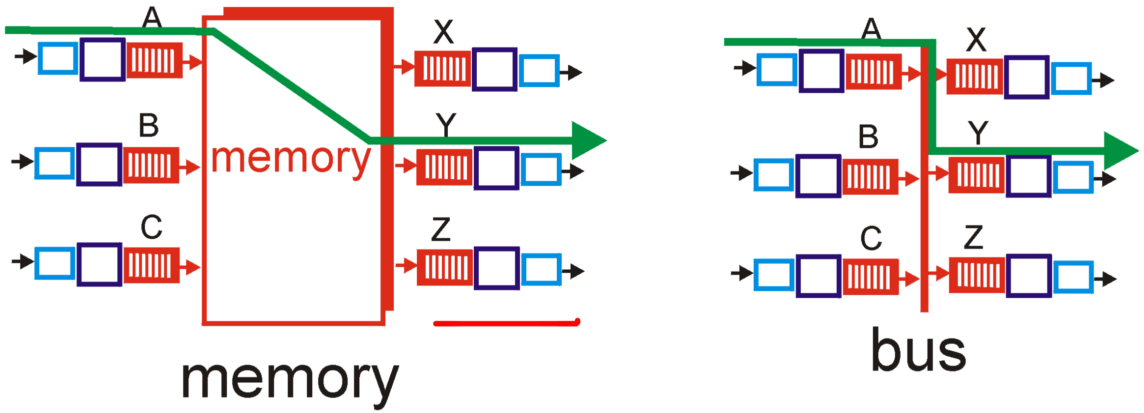
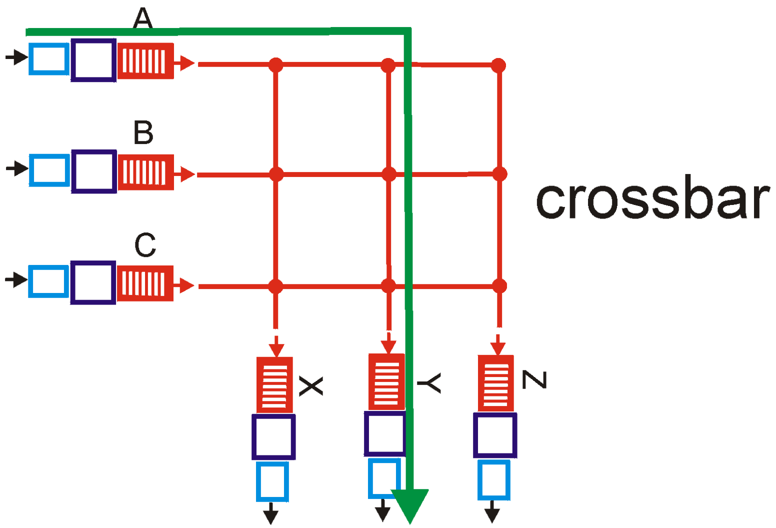


Image from: DOI: [10.1016/j.cose.2013.03.004]

Routing an IP package





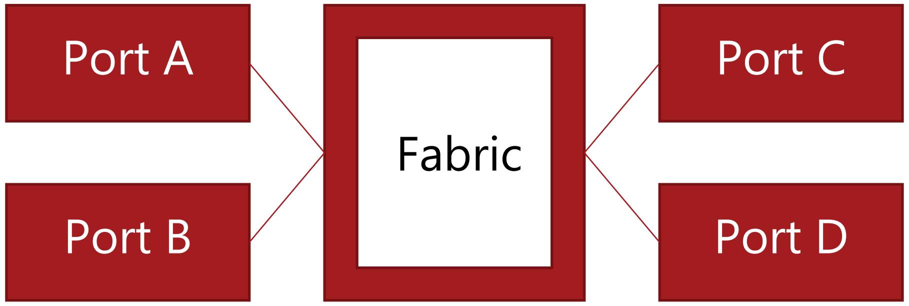


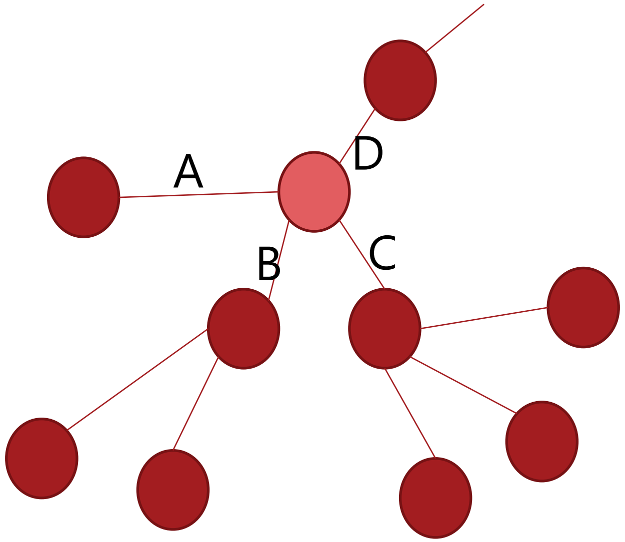
Drops packages on overload
May split (fragment) packages

Images from: http://www2.ic.uff.br/~michael/kr1999/4-network/4_06-inside.htm

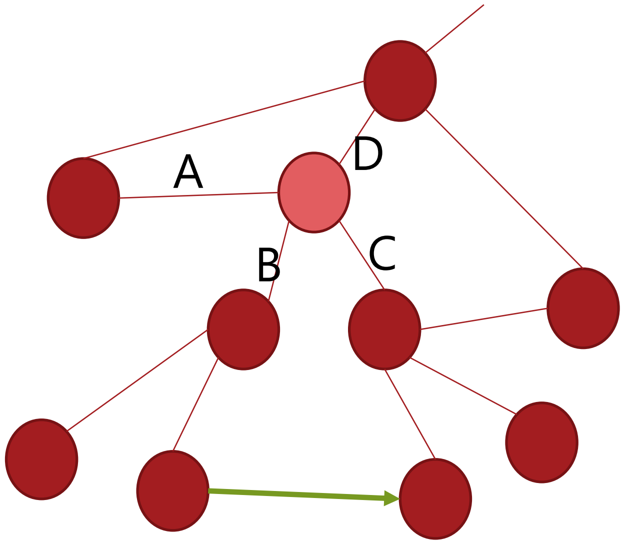
Routing from a router perspective

A = 1.0.0.0/24
B = 2.0.0.0/24
C = 3.0.0.0/24
D = others

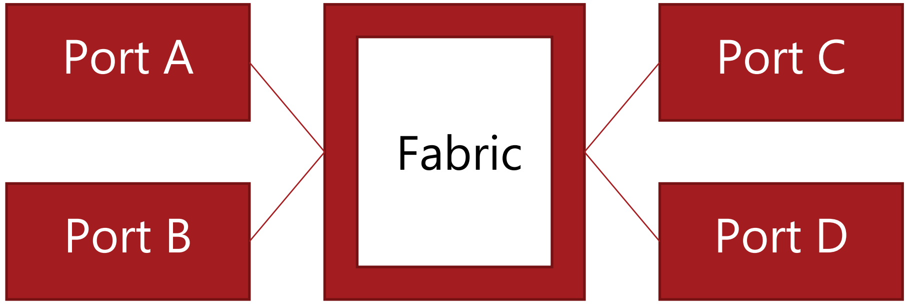




Routing from a router perspective



1.0.0.0/24 = A	0000 0001 0000 0000 0000 0000 0000 0000
2.0.0.0/24 = B	0000 0010 0000 0000 0000 0000 0000 0000
3.0.0.0/24 = C	0000 0011 0000 0000 0000 0000 0000 0000
2.2.3.0/8 = C	0000 0010 0000 0010 0000 0011 0000 0000
D = others	



Longest matching prefix rule

Allows dynamic updating of the routes

Packets in a sequence may use different routes

IPv6 arriving soon?

IPv4	IPv6
32 bits address, 2^{32} ~= $4*10^6$	128 bits address, 2^{128} ~= $3.4*10^{32}$
Dotted decimal: **123.456.789.012**	Hex: **0123:4567:89ab:cdef:0123:4567:89ab:cdef**
Allows fragmentation	Does not allow fragmentation
Allows QoS	Allows QoS

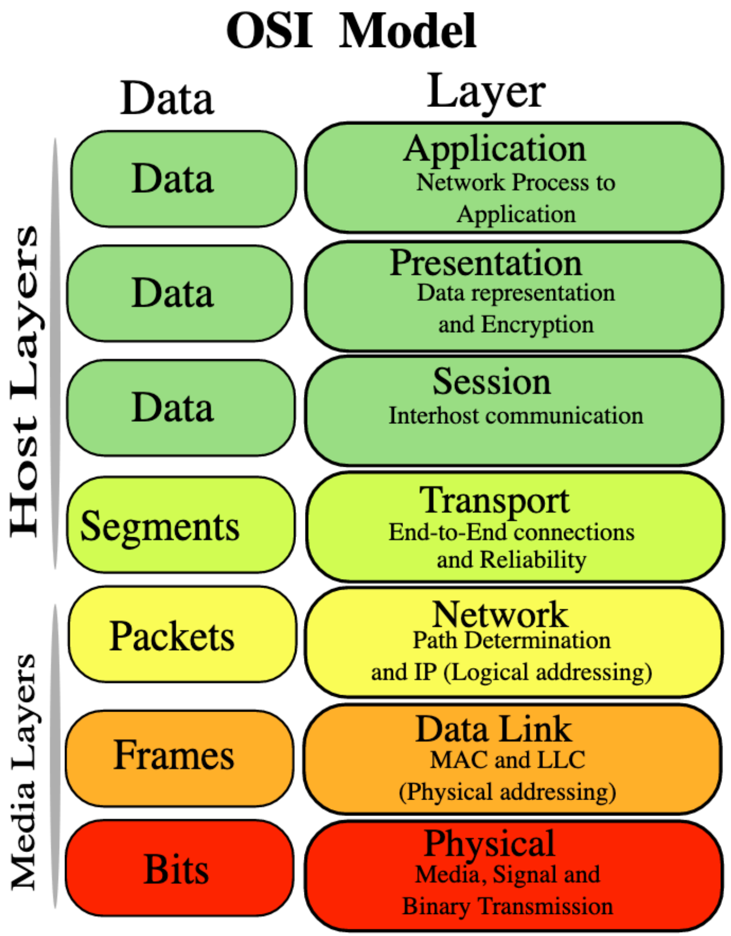
IPv4 cannot address IPv6 host

- Proposed fixed with NAT-like router translating, e.g. 1.1.1.1 => 1:1:1:1:0:0:0:0

IPv4 can be tunneled inside IPv6

Transport layer – TCP and UDP

Mostly using TCP, sometimes UDP and rarely others

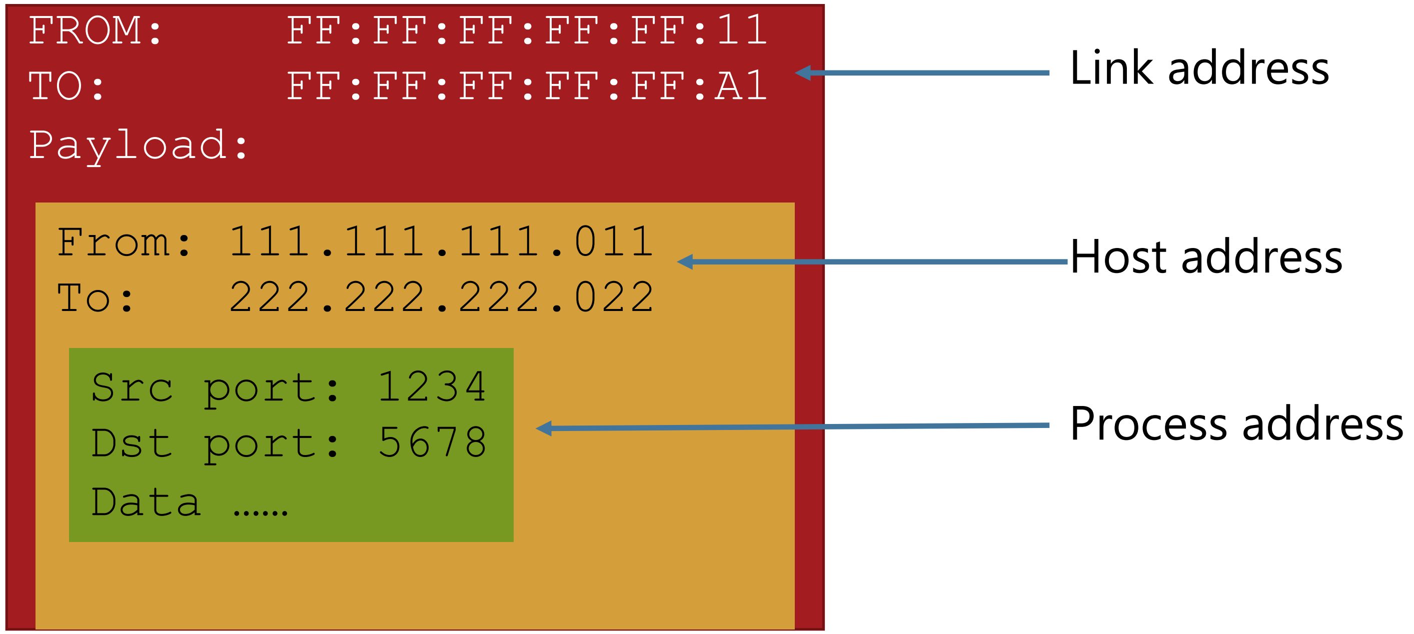


UDP segments

UDP segment header consist of 4 fields each 2 bytes (16 bits)

- Source port
- Destination port
- Length in bytes (excluding header)
- Checksum

Data bytes follow



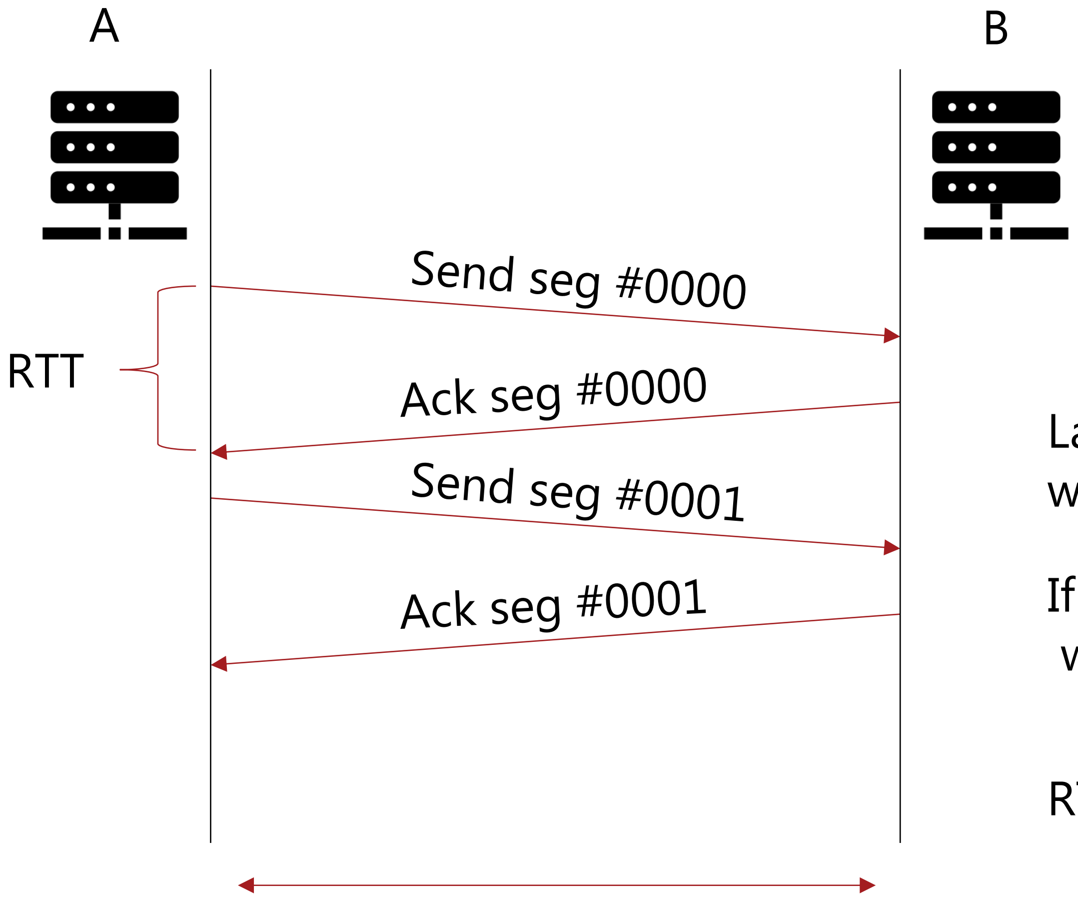
TCP – Internet workhorse

Described in RFC 675 in 1974 by Vinton Cerf & others

- Uses ports like in UDP
- Has guarantees for delivery
- Supports out-of-order delivery

Intermingled with IP specification to form TCP/IP

TCP – Starting out

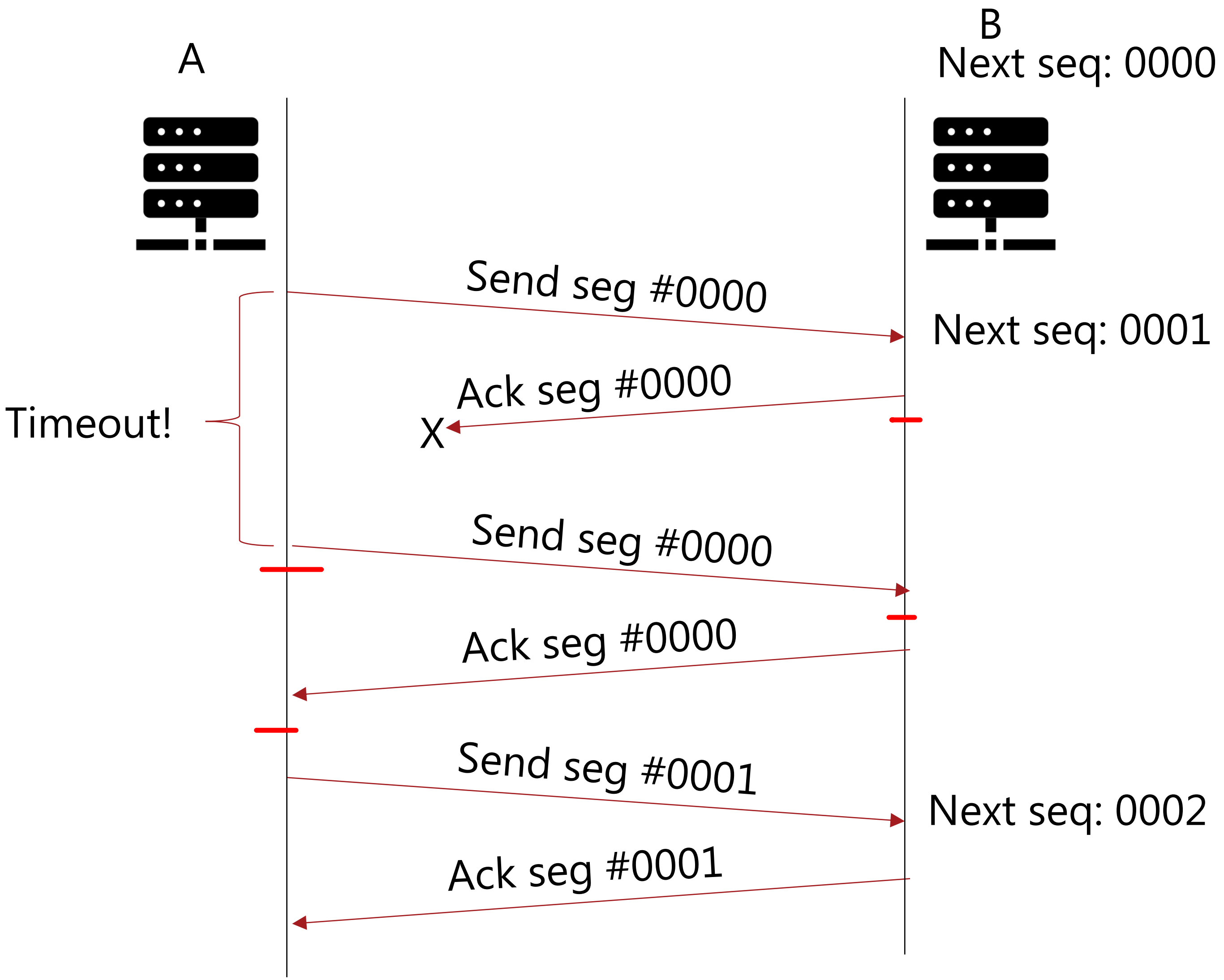


Latency, usually given as round-trip time, RTT
where latency = RTT/2

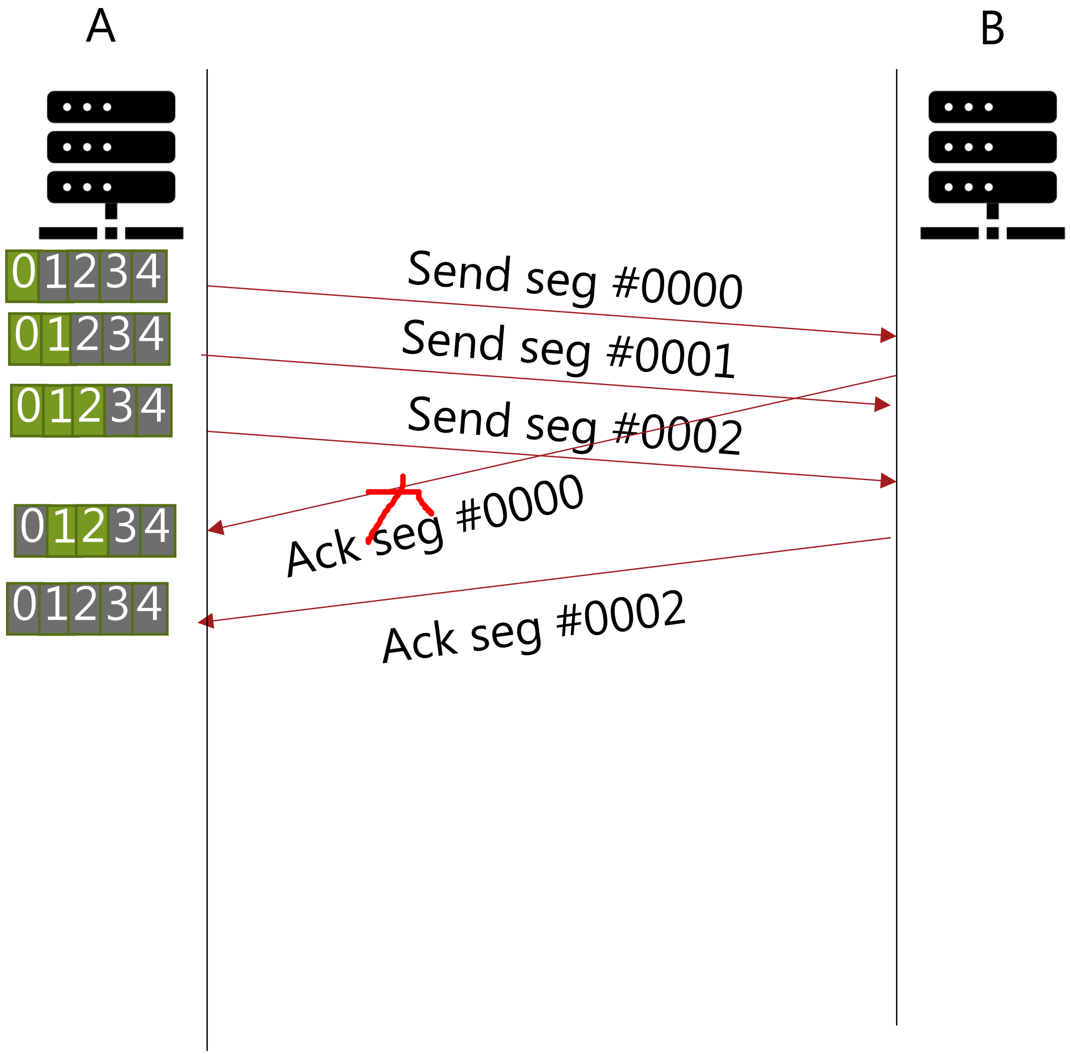
If RTT = 1 sec and package size is 10KiB,
we get at most 10KiB/s

RTT=10ms, MTU=1500 => ~146 KiB/s

TCP – Handling package loss

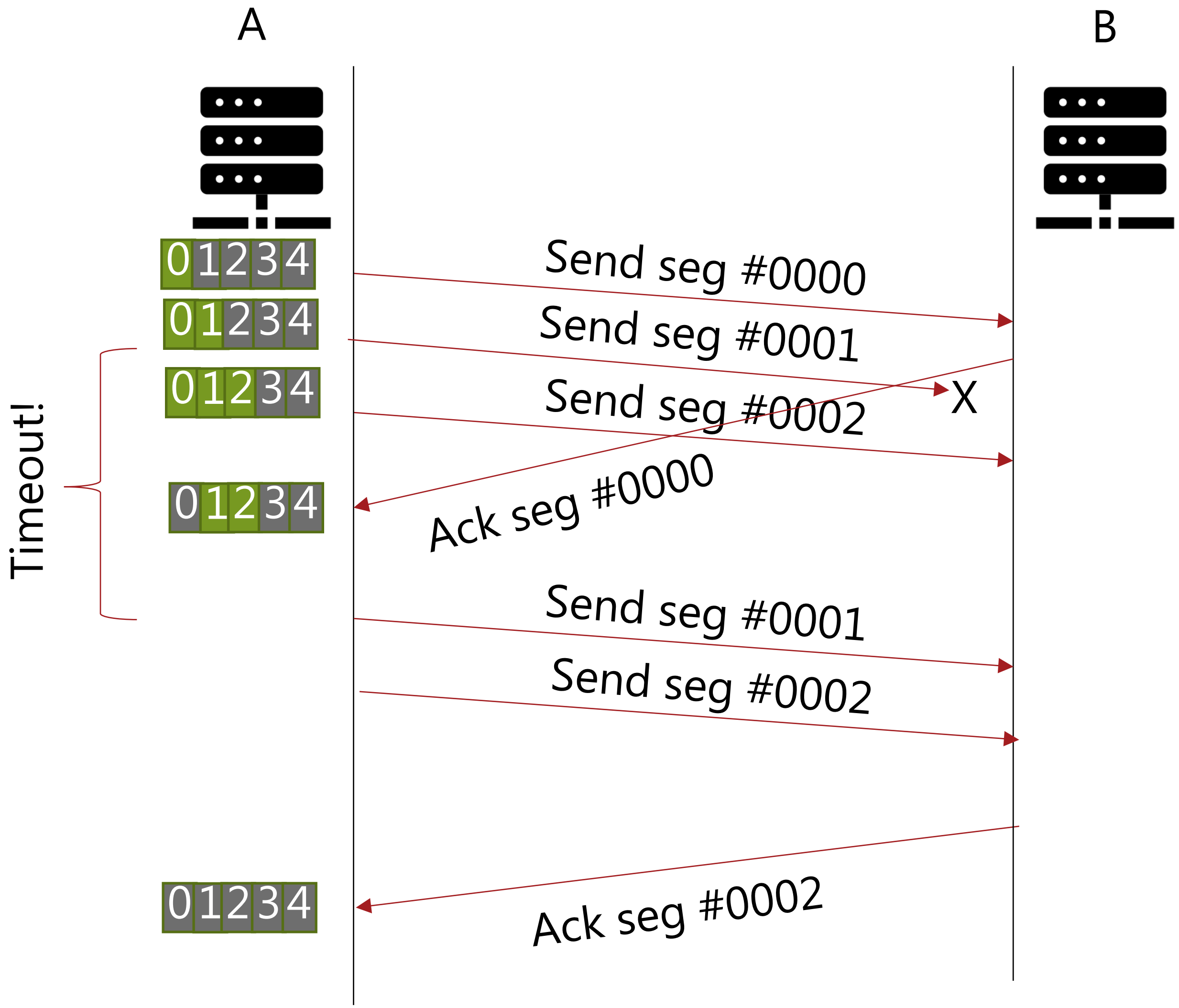


TCP – Send & Receive Windows

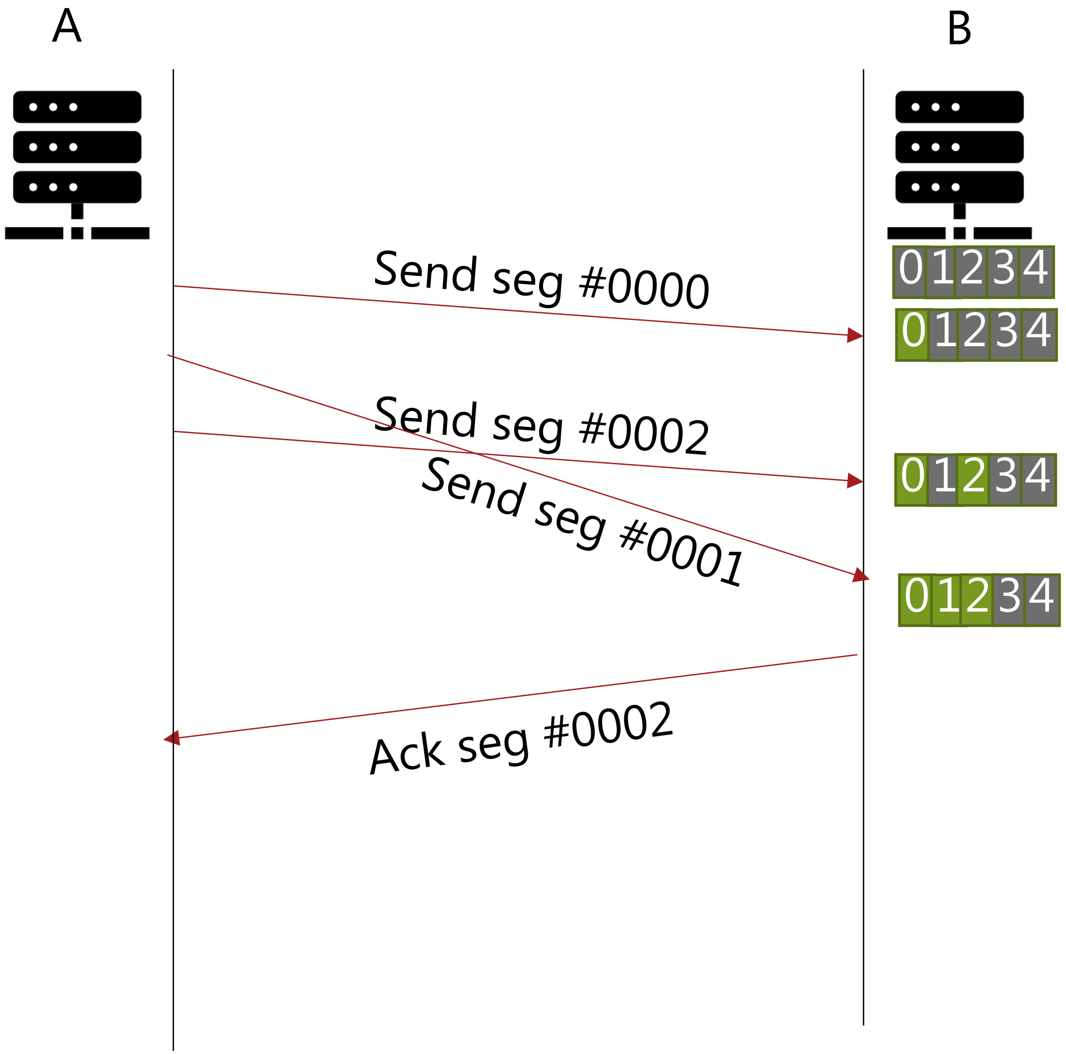


Cumulative ACK

TCP – Send Window



TCP – Receive Window



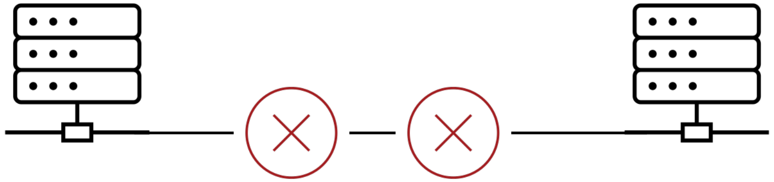
Triple ACK for same SEQ would indicate single missing package

TCP – Flow & Congestion Control

Related concepts, but different goals and implementations

Flow control
Avoid sending data that receiver cannot store

- Receiver sends buffer space indication with ACK
- Sender throttles based on projected buffer space
- When buffer space is 0, sender waits for space
    - After a timeout it sends one segment to test



Congestion control
Avoid clogging the network

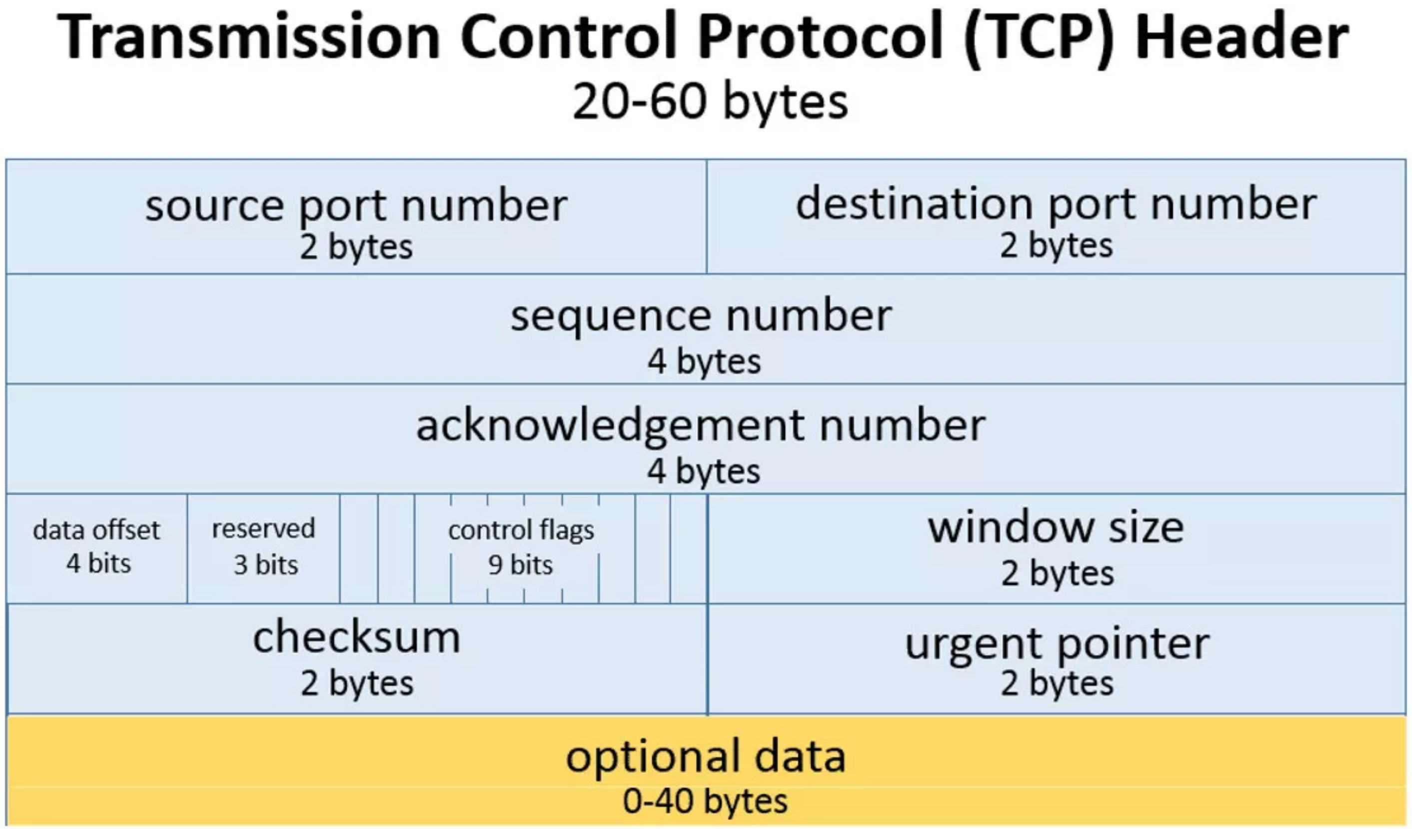
- Receiver is not actively involved
- Sender measures network events
    - can use package loss (Tahoe)
    - or use ACK delay (Vegas)
- Sender throttles based on estimated network congestion

TCP header

FROM: FF:FF:FF:FF:FF:11
TO: FF:FF:FF:FF:FF:A1
Payload:

From: 111.111.111.011
To: 222.222.222.022

Src port: 1234
Dst port: 5678
Data ……



TCP/IP and NAT

Internal host	External port
10.0.0.1:6789	4455

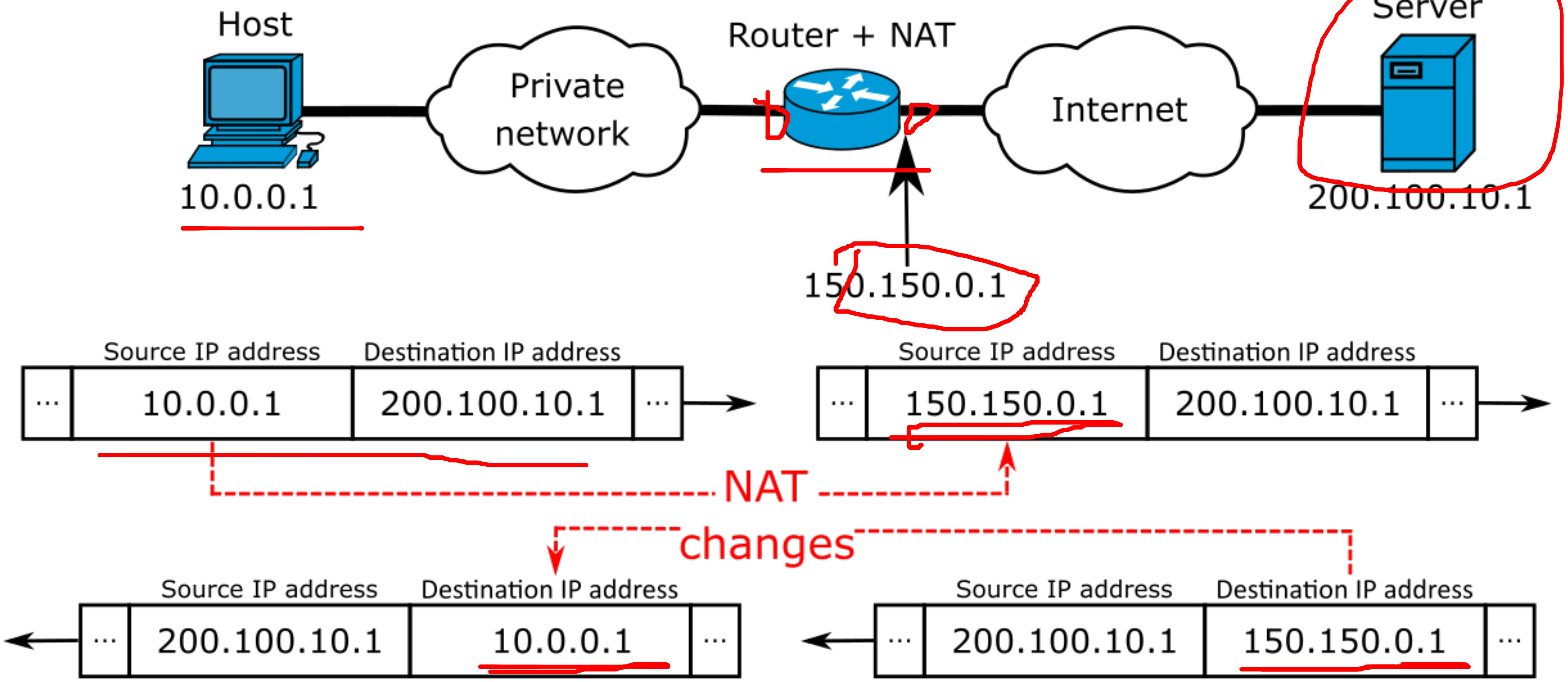
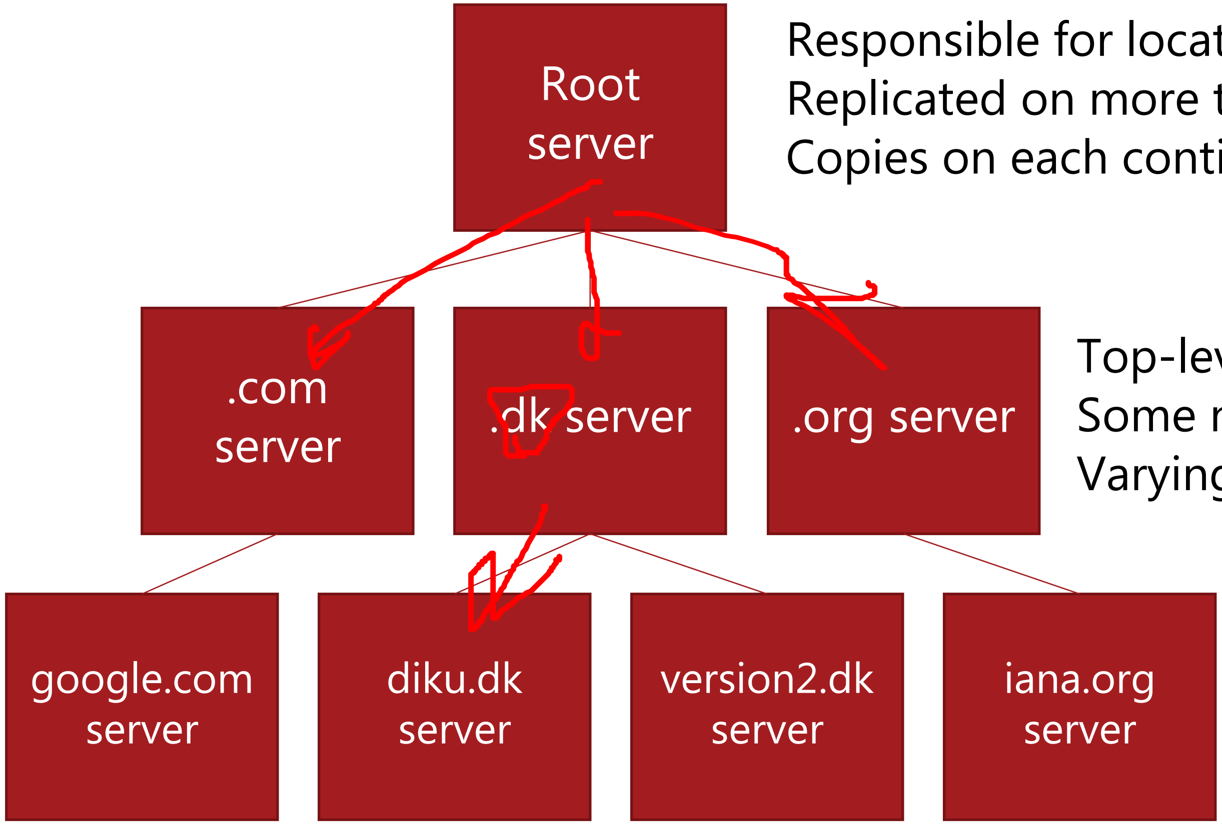


Image from: https://commons.wikimedia.org/wiki/File:NAT_Concept-en.svg

DNS – domain names

You can consider DNS a phone book for machines



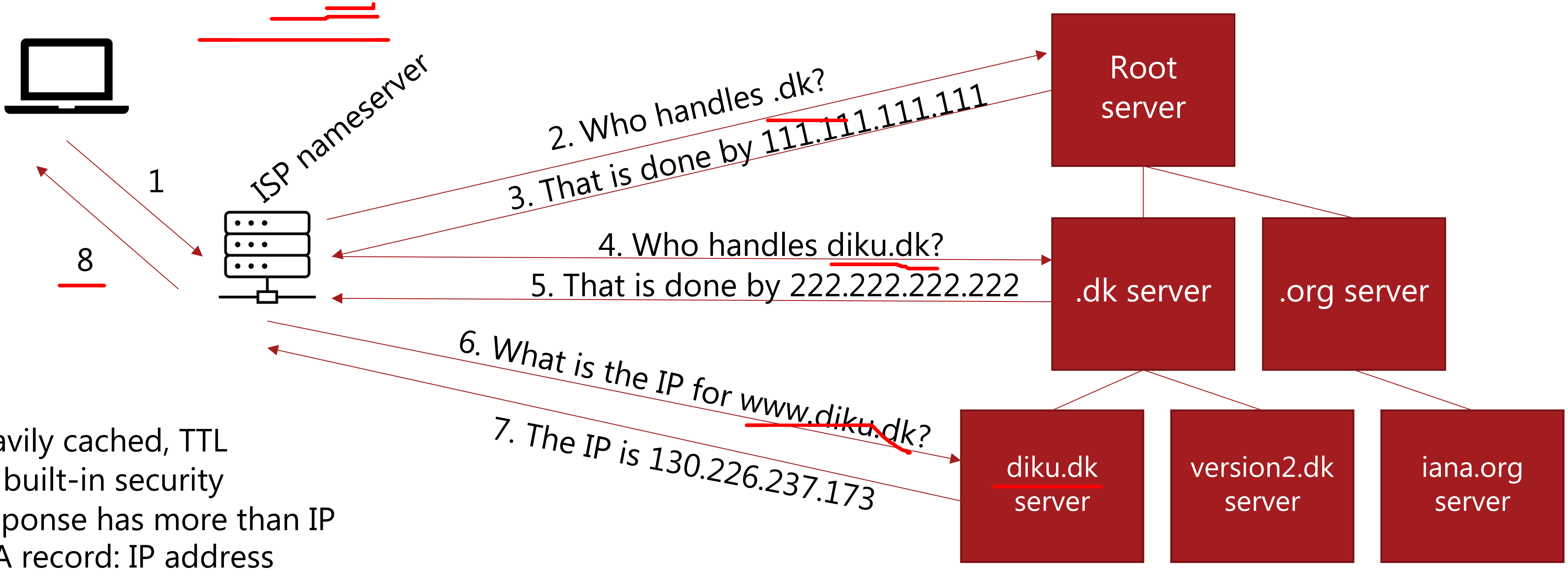
Responsible for locating top-level servers
Replicated on more that 700 nodes
Copies on each continent

Top-level servers
Some national, some private
Varying redundancy levels

Domain servers return host information, e.g. www.diku.dk = 130.226.237.1

DNS – a lookup process

I need the IP for www.diku.dk



Heavily cached, TTL
No built-in security
Response has more than IP

- A record: IP address
- AAA record: IPv6 address
- MX record: email server
- …

Applications on the network

HTTP / Web based

Custom protocol