Høyskolen Kristiania

Digital Technology

TK1104

Guest lecturer: Peyman Teymoori

peymant@ifi.uio.no

Lecturer: Toktam Ramezanifarkhani

Toktam.Ramezanifarkhani@kristiania.no

Toktamr@ifi.uio.no

Computer Networks – Network Layer

Learning objectives

Høyskolen Kristiania

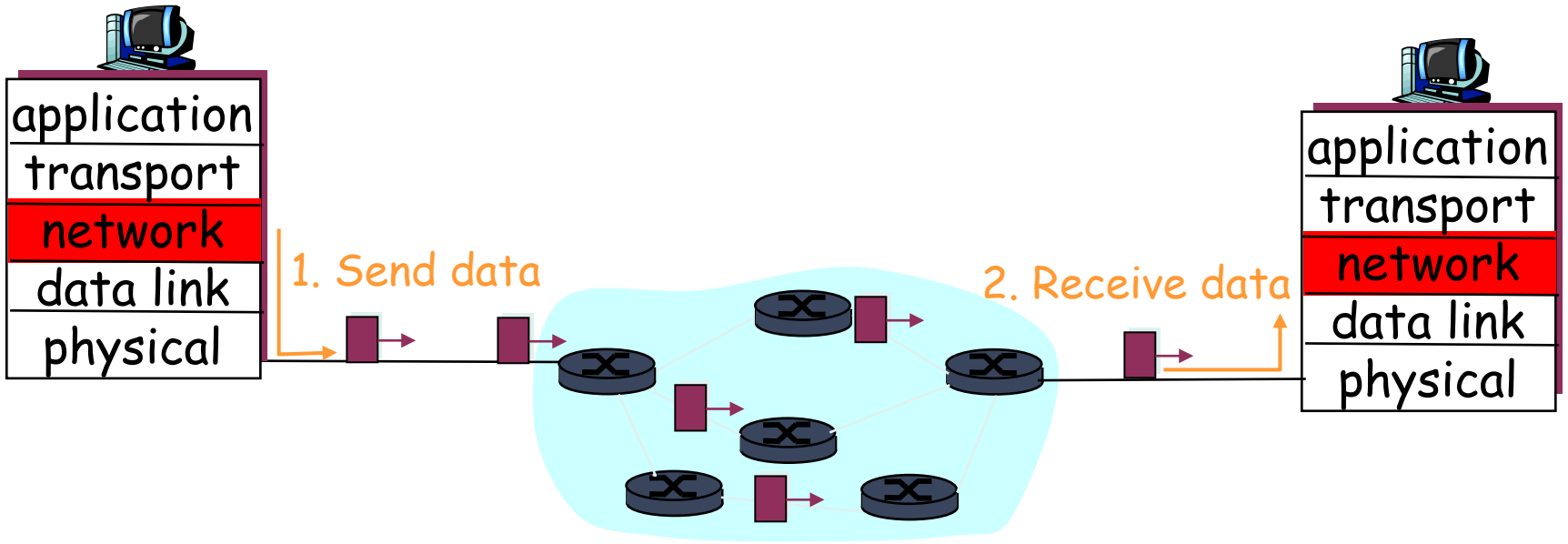
- To learn about addresses on the Internet
  - IP, subnets
- To learn debug the network using ICMP
  - ping, trace route
- To understand how NAT (Network Address Translation) works
- To get to use some basic networking commands in the terminal / prompt

Layer´s name	designation of transmission unit	Most important tasks / functions Example of protocols / standards
Applikasjonslaget	Melding (Message)	Støtte nettverksapplikasjoner Ex: HTTP, DNS, FTP, SMTP, POP3….
Transport layer	Segment	transport of application layer messages between client and server pages of an application: including mux / demux, different levels of reliability and more .. Ex: TCP, UDP ...
The network layer	Datagram	routing of datagram from / to host through the network core Ex: IP (v4 and v6) ICMP, RIP, OSPF, BGP
The data line layer	Frame	(Reliable) delivery of frame from neighbor node to neighbor node. Ex: Ethernet II, FDDI, IEEE 802.11
Physical	Bit	(Code and) Move single bit between communication partners. Ex: 10BaseT,

The network layer

Datagram network (Internet)

- No pre-setup at the network layer level
- The routers do not care about the condition of the route
  - Stateless routers
- The packets are routed based on the recipient ID
  - the packets between the same sender and receiver can follow different routes
- Data can be lost in a router queue if the buffer is not large enough

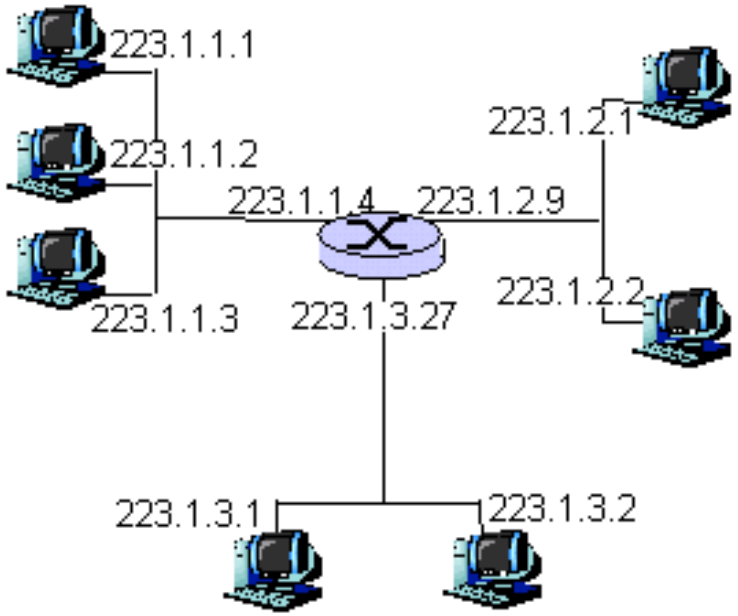


Høyskolen Kristiania

Internet Protocol v4

IPv4 addressing

- IPv4 address: 32-bit «id» for each host and router interface (adapter)
- A host can have multiple interfaces
- A router usually has several connections, each with its own interface
- IP address belongs to each interface



223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Exercise: what is my (local) IP address?

Høyskolen Kristiania

- on Windows, go to command prompt and enter
  - ipconfig

- on mac OS, go to terminal and enter
  - ipconfig getifaddr en0
  - ipconfig getoption en0 subnet_mask

- on Android/iPhone
  - Click on the WiFi icon and then, click on the connected one!

Write it down somewhere!

IPv4 addresses: allocation

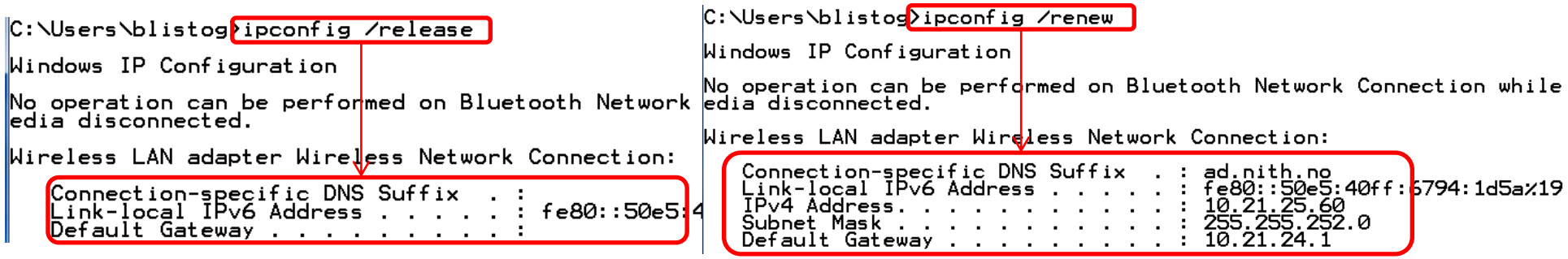
Høyskolen Kristiania

- For hosts in LAN
  - Can be set manually / statically
  - Dynamic Host Configuration Protocol(DHCP)
- For networks
  - Is assigned its share of the ISP's assigned address space
- For Internet Service Provider (ISP)
  - International Agency (ICANN) assigns addresses, manages DNS, assigns domain names and resolves disputes.
  - “Continent registrar”: RIPE distributes IP addresses and AS numbers to Europe and others.

Dynamic Host Configuration Protocol

Høyskolen Kristiania

- Each DHCP server has a set of possible addresses (pool)
- Sets the address dynamically with "plug-and-play"
- Host transmits : `DHCP discover`
- DHCP server responds: `DHCP offer`
- Host transmits : `DHCP request`
- DHCP server transmits: IP address and other network parameters (eg DNS server) + `DHCP ack`
- Hosting is set up with these values

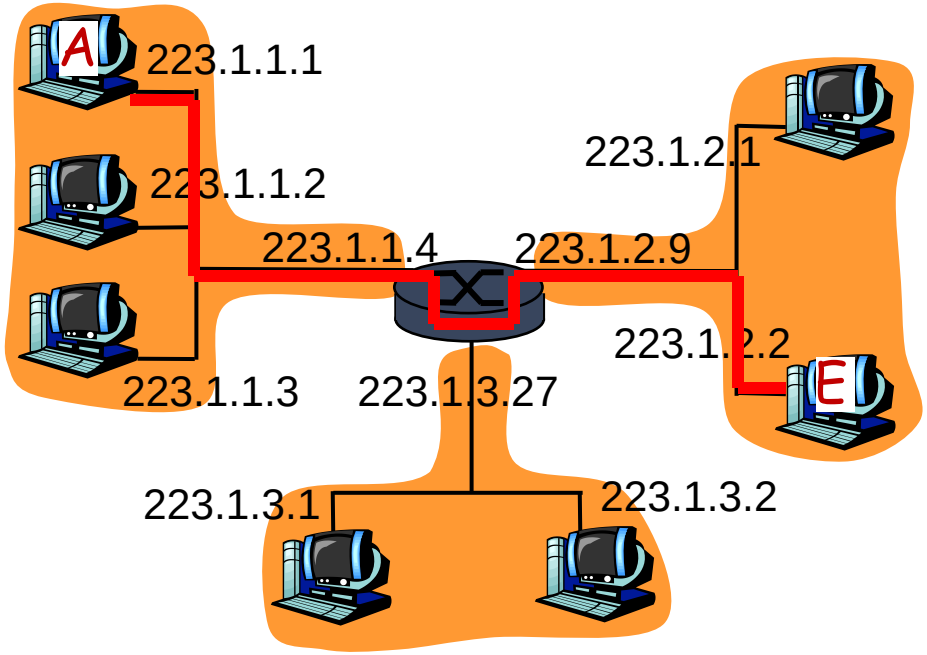


Datagram from sender to recipient

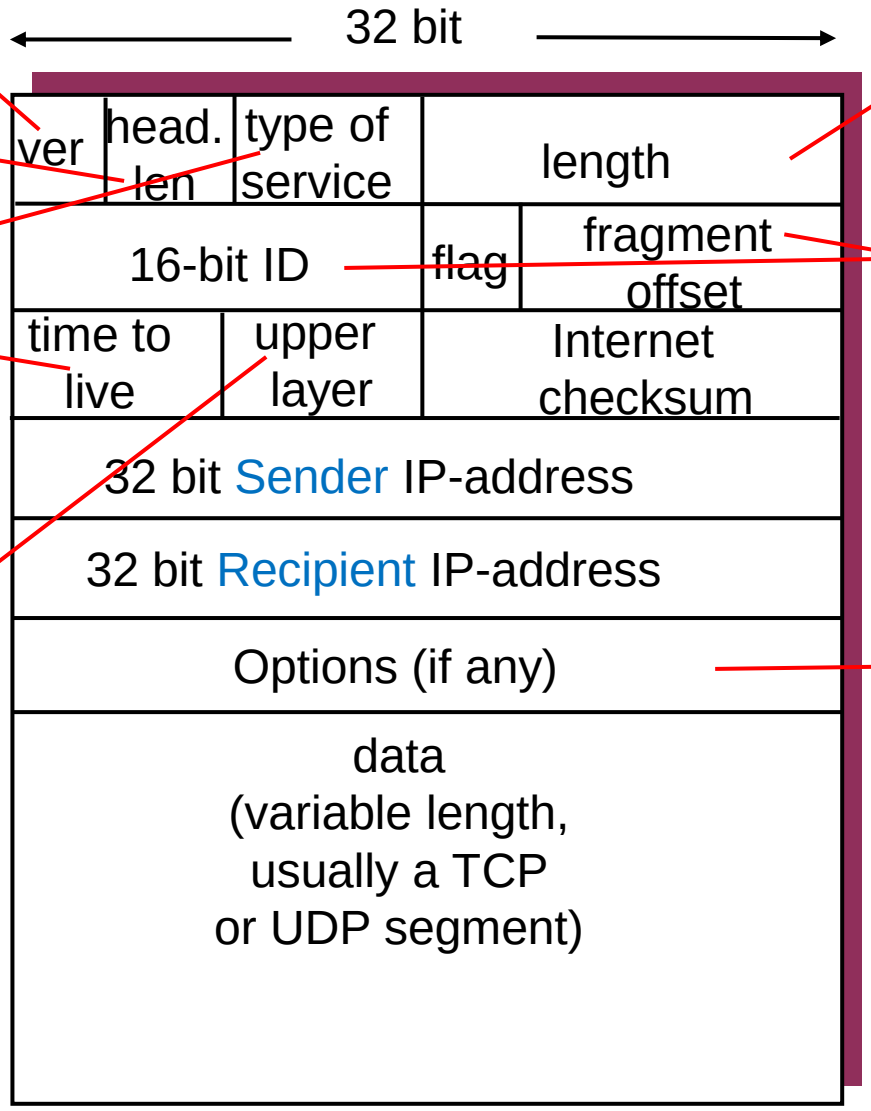
Høyskolen Kristiania

Sender A, recipient E

- Find the network address of E
- E on another network
- A, E not directly connected
- Routing table: the next jump router to E is 223.1.1.4
- The link layer sends the datagram to router 223.1.1.4 in the link layer frame
- Datagram arrives 223.1.1.4
- E on the same network as 223.1.2.9
- Datagram is sent to 223.1.2.2



IPv4 datagram-format



IP protocol version

header length (byte)

"type" data

max number remaining jump (decremented in each router)

the protocol to be had payload (TCP, UDP)

total datagram-length (bytes)

for fragmentation / composition

For example. timestamp, record route, specify list of routers you have to stop by

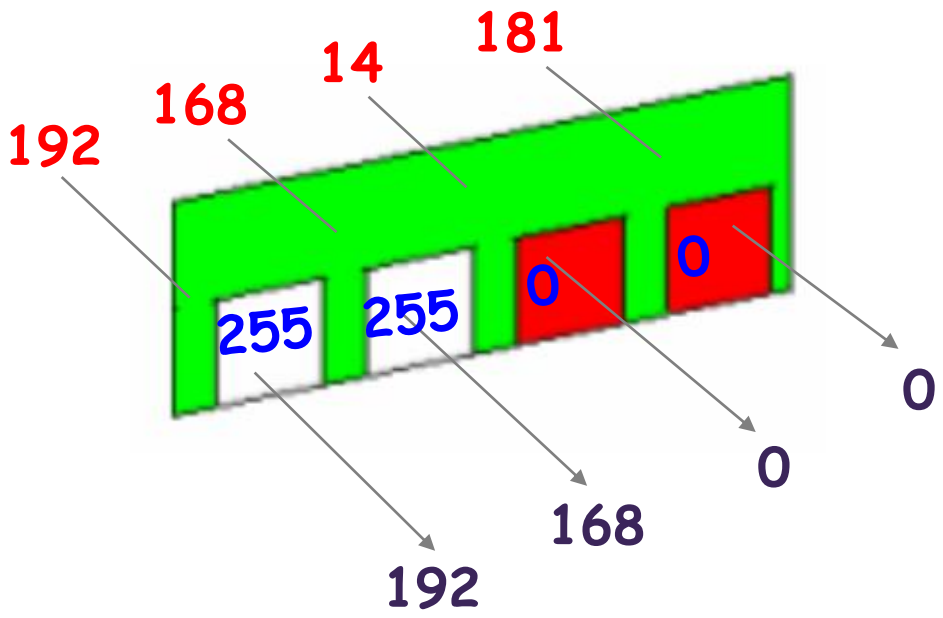
<u>how much overhead with TCP?</u>

- Min 20 byte for TCP
- Min 20 byte for IP
- = 40 byte + app. lags overhead

The subnet mask

Høyskolen Kristiania

- The subnet mask indicates which bits are PREFIX and which are HOST
- A subnet mask is a bitmask applied to an IP address.
  - Address 192.168.14.181, mask 255.255.0.0



	192.168.14.181
AND	255.255.0.0
=	192.168.0.0

Ex: Which network?

Assume 10.21.26.184 with subnet mask 255.255.252.0

To which network does it belong?

```
  10 . 21.0001 10 10.1011 1000
&
  255.255.1111 11 00.0000 0000
-------------------------------
  10 . 21.0001 10 00.0000 0000
```

22 bits for **prefix**, **10** bits for **host**

The network is **10.21.24.0/22**

Lowest address is **10.21.24.1**

Broadcast is **10.21.27.255**

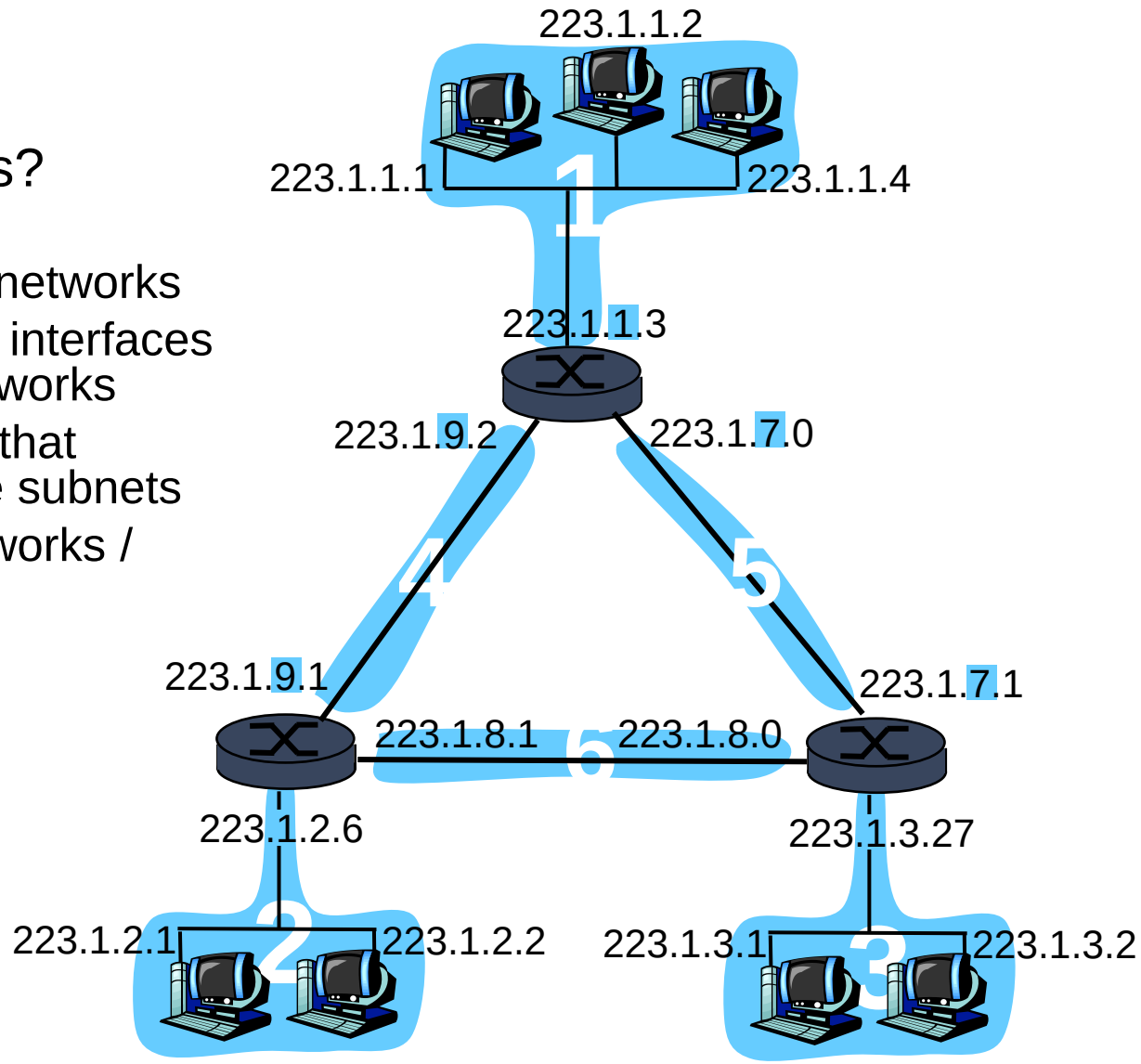
all host bits set to 1 !!!

Special IP addresses

- When booting, a machine can identify itself with a temporary address
  - 0.0.0.0 (default route)
- *Loopback means addressing oneself*
  - 127.0.0.1
  - Many systems use 127.0.0.0/8
- The lowest (network) and highest (broadcast) address are thus not used in ordinary addressing of hosts or routers

Subnet

- How many subnets?
  - 3 subnets / local networks
  - The routers have interfaces in different IP networks
  - 3 "link networks" that connect the three subnets
  - 6 different IP networks / subnets



Høyskolen Kristiania

Internet Control Message Protocol (ICMP)

ICMP - Internet Control Message Protocol

- Used by host, router and gateway
  - Error reporting
  - Echo request / response (`ping`)
- ICMP-message
  - Type, code and first 8 bytes in the datagram with the error
- `ping` and `tracert` often utilizes ICMP

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ping

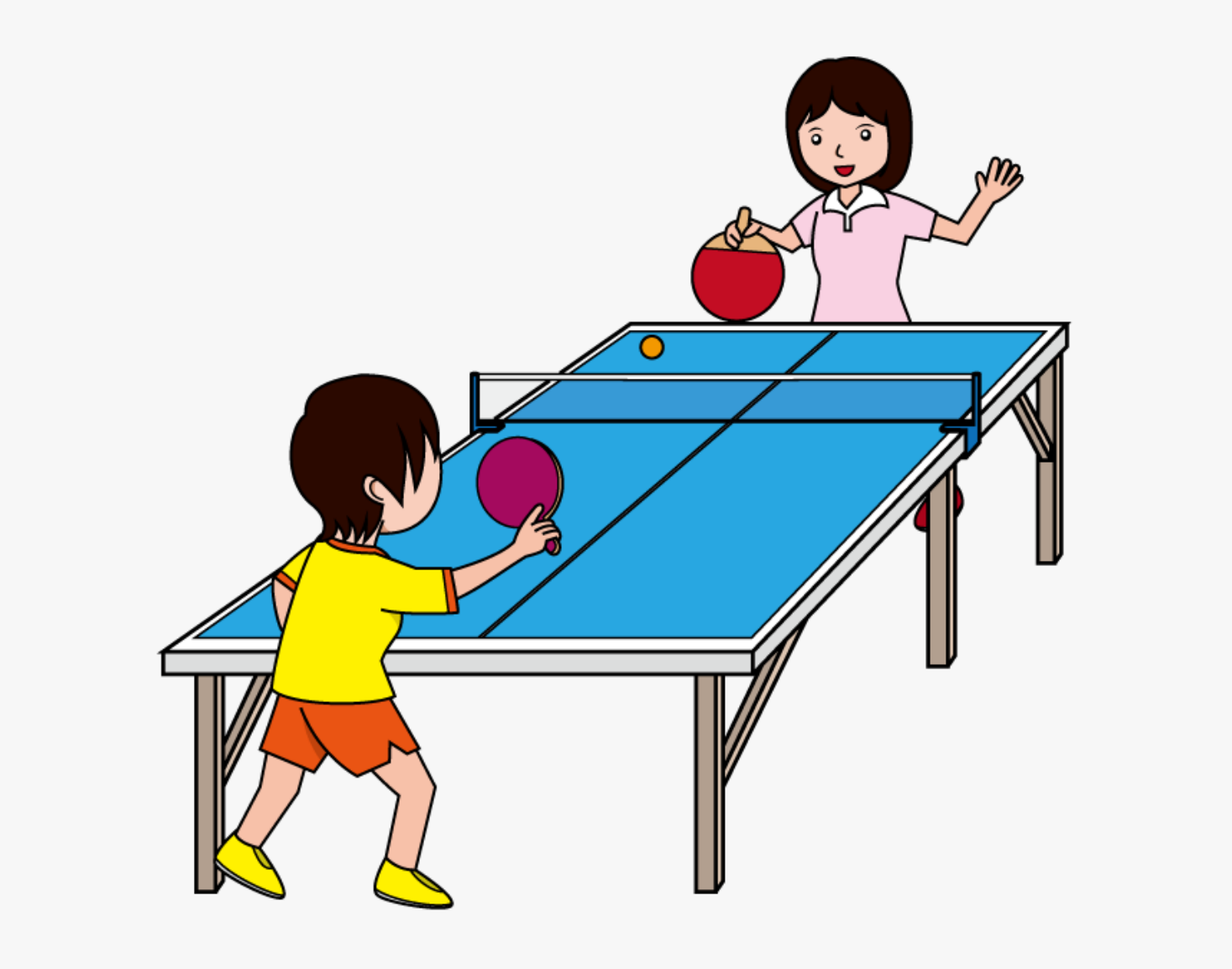
Høyskolen Kristiania

- Sends an ICMP echo package to the specified address
- Useful for checking if the IP address exists and can be reached.

```
~->ping vg.no
PING vg.no (195.88.55.16) 56(84) bytes of data.
64 bytes from www.vg.no (195.88.55.16): icmp_req=1 ttl=251 time=0.801 ms
64 bytes from www.vg.no (195.88.55.16): icmp_req=2 ttl=251 time=0.817 ms
64 bytes from www.vg.no (195.88.55.16): icmp_req=3 ttl=251 time=0.824 ms
^C
--- vg.no ping statistics ---
3 packets transmitted, 3 received, 0% packet loss, time 2003ms
rtt min/avg/max/mdev = 0.801/0.814/0.824/0.009 ms
```

```
C:\>ping vg.no

Pinging vg.no [2001:67c:21e0::16] with 32 bytes of data:
Reply from 2001:67c:21e0::16: time<1ms
Reply from 2001:67c:21e0::16: time<1ms
Reply from 2001:67c:21e0::16: time<1ms
Reply from 2001:67c:21e0::16: time<1ms

Ping statistics for 2001:67c:21e0::16:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 0ms, Maximum = 0ms, Average = 0ms
```

Exercise: ping your classmate



tracert (Windows) / traceroute (Linux/mac)

Høyskolen Kristiania

- Application that sets the TTL field in the IP header first to 1, then 2, then 3 etc.
- Then triggers ICMP type 11 feedback from each router along the way
- Useful for checking where on the route delays / problems may have occurred

```
~->traceroute google.com
traceroute to google.com (173.194.32.51), 30 hops max, 60 byte packets
 1  stolav-gw4.uninett.no (158.36.84.169)  1.990 ms  1.947 ms  2.133 ms
 2  stolav-gw2.uninett.no (128.39.230.137)  2.118 ms  2.102 ms  2.187 ms
 3  dk-uni.nordu.net (109.105.102.25)  10.464 ms  10.453 ms  10.442 ms
 4  se-tug.nordu.net (109.105.97.9)  18.141 ms  18.129 ms  18.114 ms
 5  se-tug2.nordu.net (109.105.97.18)  18.100 ms  18.089 ms  18.074 ms
 6  google-gw.nordu.net (109.105.98.6)  18.059 ms  17.214 ms  17.207 ms
 7  216.239.43.122 (216.239.43.122)  17.770 ms  17.751 ms  17.747 ms
 8  216.239.43.255 (216.239.43.255)  18.858 ms  18.854 ms  18.850 ms
 9  arn06s02-in-f19.1e100.net (173.194.32.51)  17.992 ms  18.352 ms  18.335 ms
```

```
C:\>tracert google.com

Tracing route to google.com [173.194.32.51]
over a maximum of 30 hops:

  1    <1 ms    <1 ms    <1 ms  stolav-gw4.uninett.no [158.36.84.169]
  2    <1 ms     1 ms    <1 ms  stolav-gw2.uninett.no [128.39.230.137]
  3     9 ms     9 ms     9 ms  dk-uni.nordu.net [109.105.102.25]
  4    16 ms    16 ms    16 ms  se-tug.nordu.net [109.105.97.9]
  5    16 ms    16 ms    17 ms  se-tug2.nordu.net [109.105.97.18]
  6    17 ms    16 ms    17 ms  google-gw.nordu.net [109.105.98.6]
  7    17 ms    17 ms    17 ms  216.239.43.122
  8    18 ms    17 ms    17 ms  216.239.43.255
  9    17 ms    17 ms    17 ms  arn06s02-in-f19.1e100.net [173.194.32.51]

Trace complete.
```

Høyskolen Kristiania

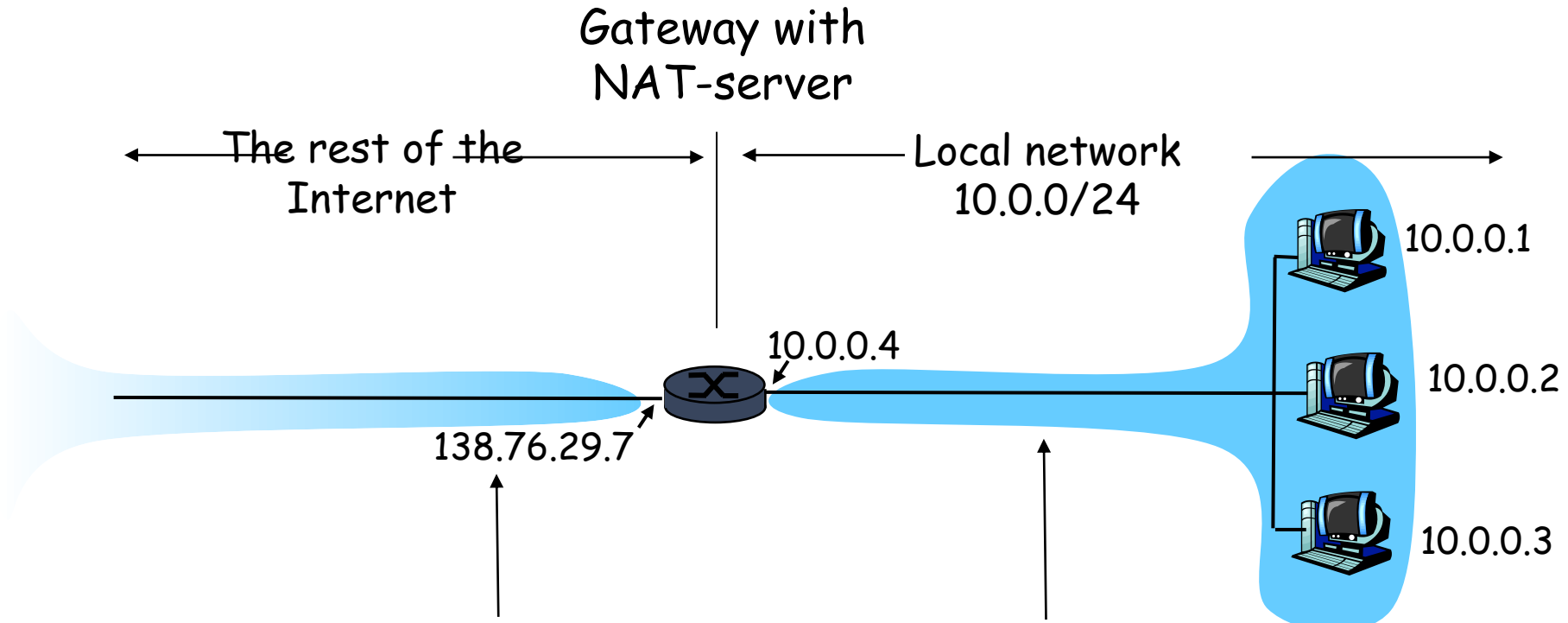
Network Address Translation

NAT: Network Address Translation

Høyskolen Kristiania

- Why?: The LAN has only one / a few IP addresses from the Internet perspective:
- ISP does not have to assign an address range:
  - only one / a few IP address (es) for an entire organization's network
- Can change addresses within LAN without having to inform the outside world about it
- Can change ISP without having to change addresses in the LAN
- Equipment in the LAN is not directly addressable or visible to outsiders (better security)

NAT: Network Address Translation

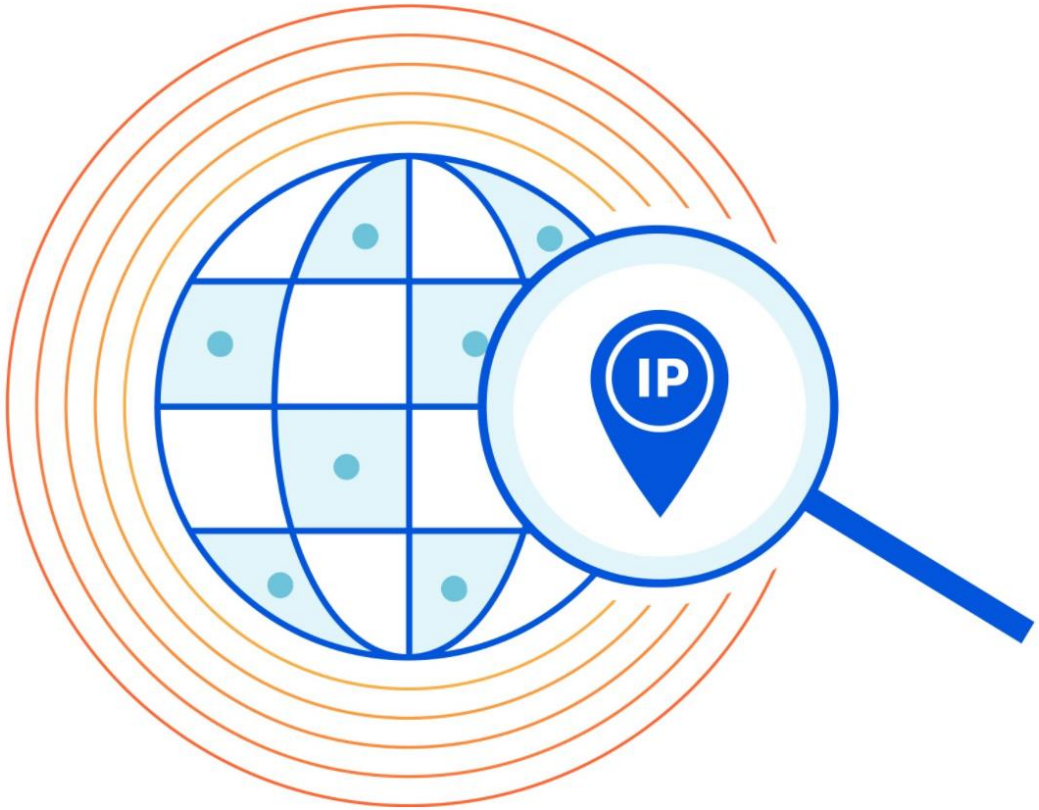


All datagrams *leaving* the LAN has the *same* sender IP address: e.g. 138.76.29.7, Different sender port numbers

Datagram sender or receiver within this network has 10.0.0 / 24 address for source, target (as usual) User (typically) PRIVATE ADDRESSES (10.x.x.x, 192.168.x.x, ..)

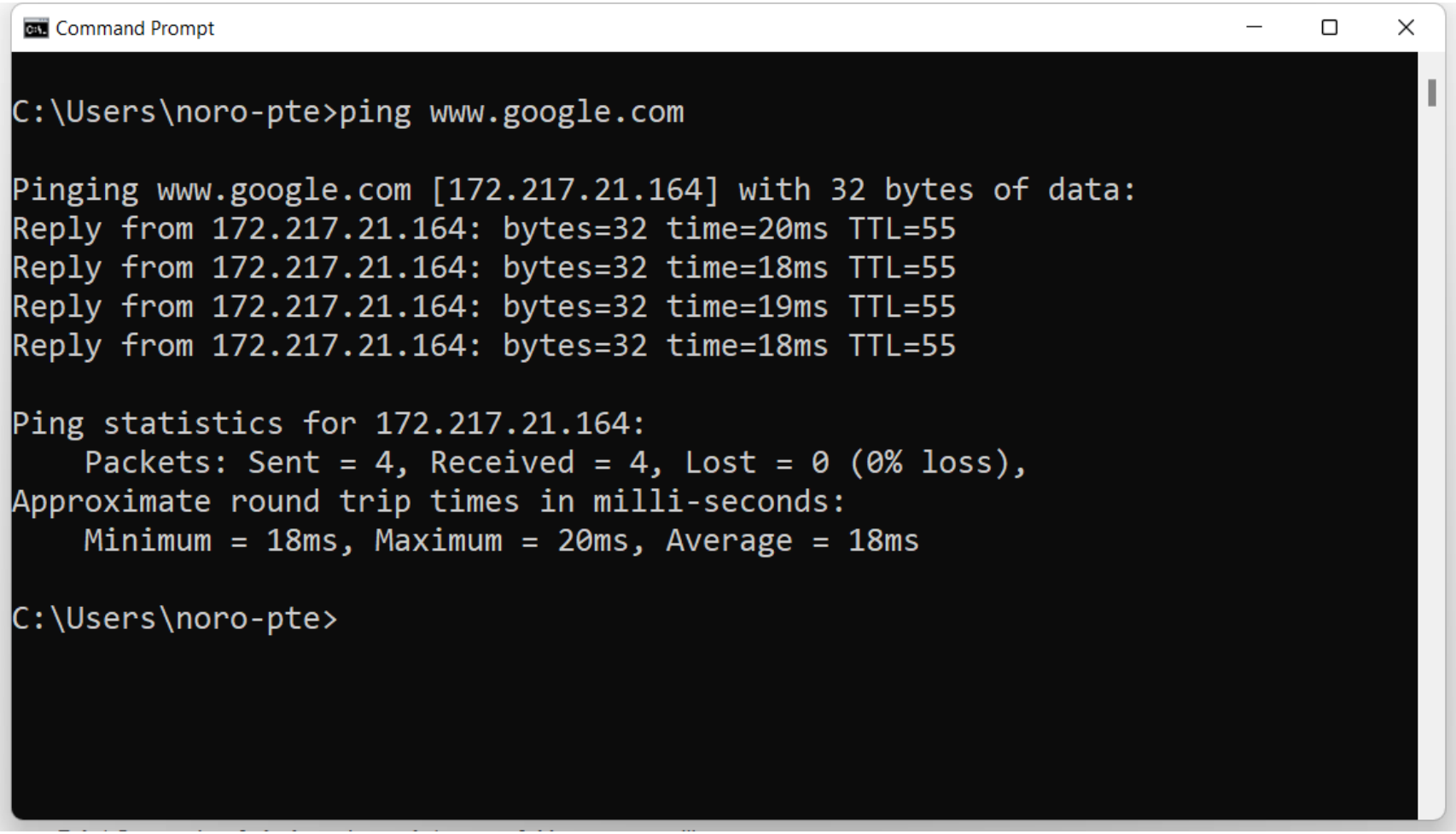
What is my public IP address?

- Go to whatismyipaddress.com



Ping (Windows/Linux)

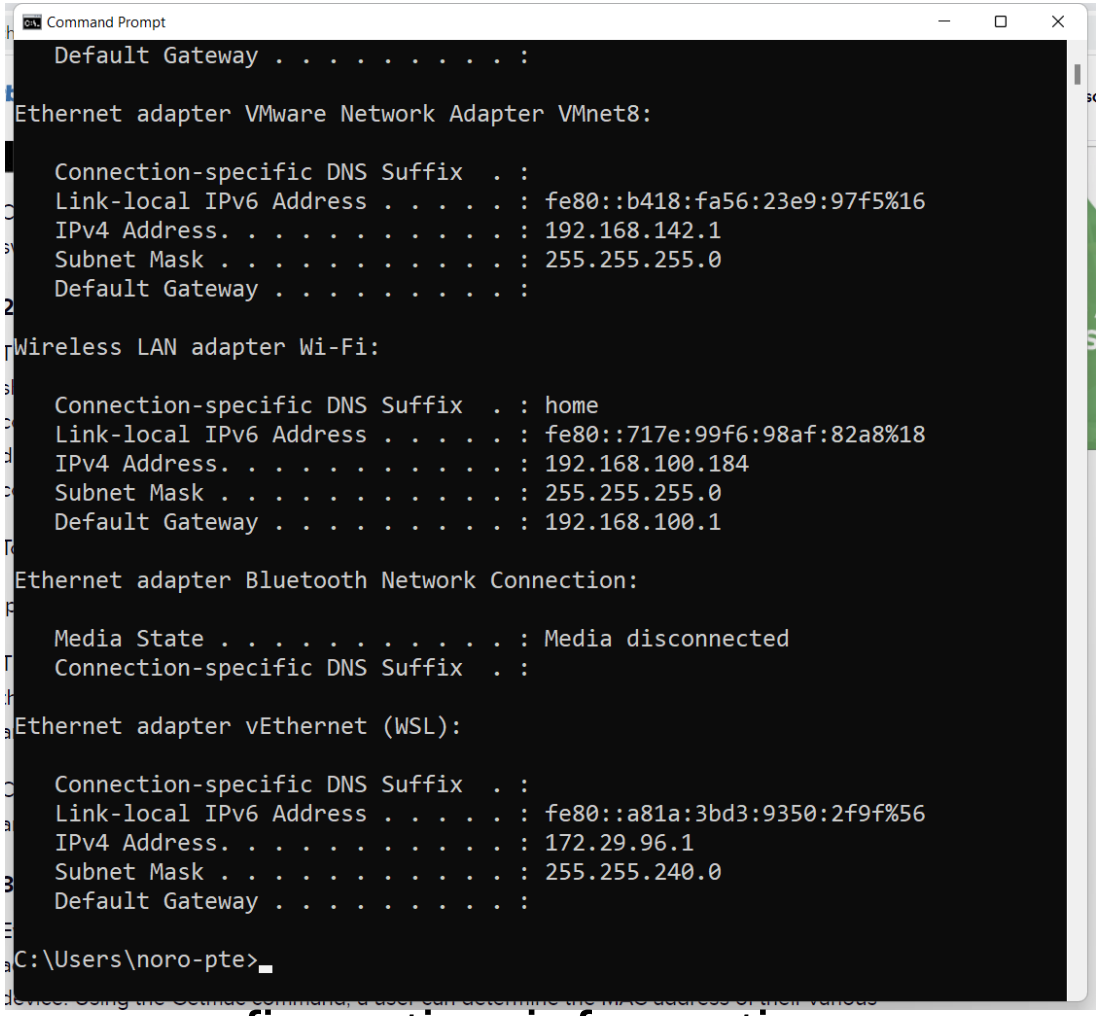
Høyskolen Kristiania



- Test the reachability of a device on a network

ipconfig (Windows)

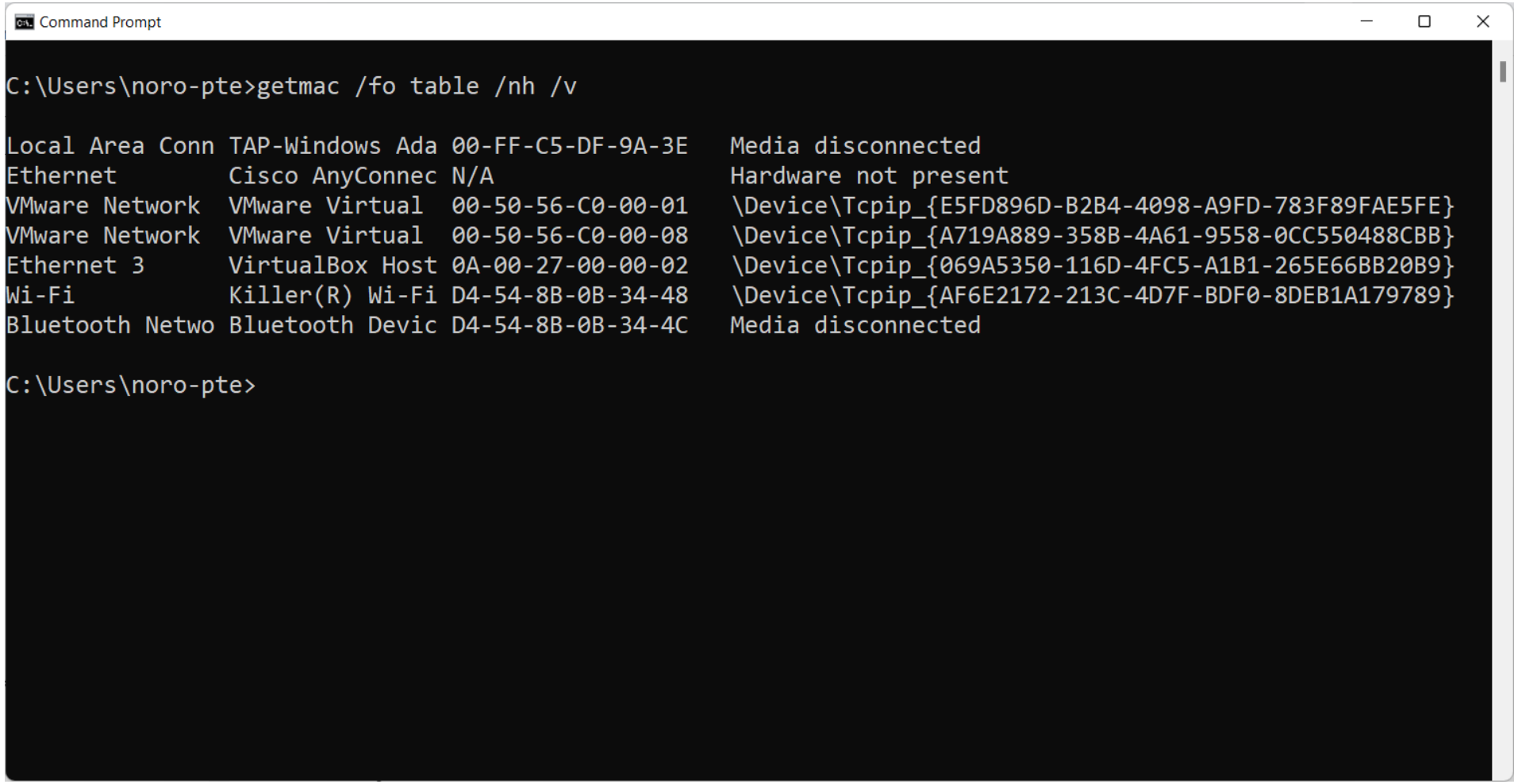
Høyskolen Kristiania



- Displays basic IP address configuration information

getmac (Windows), ip a (Linux)

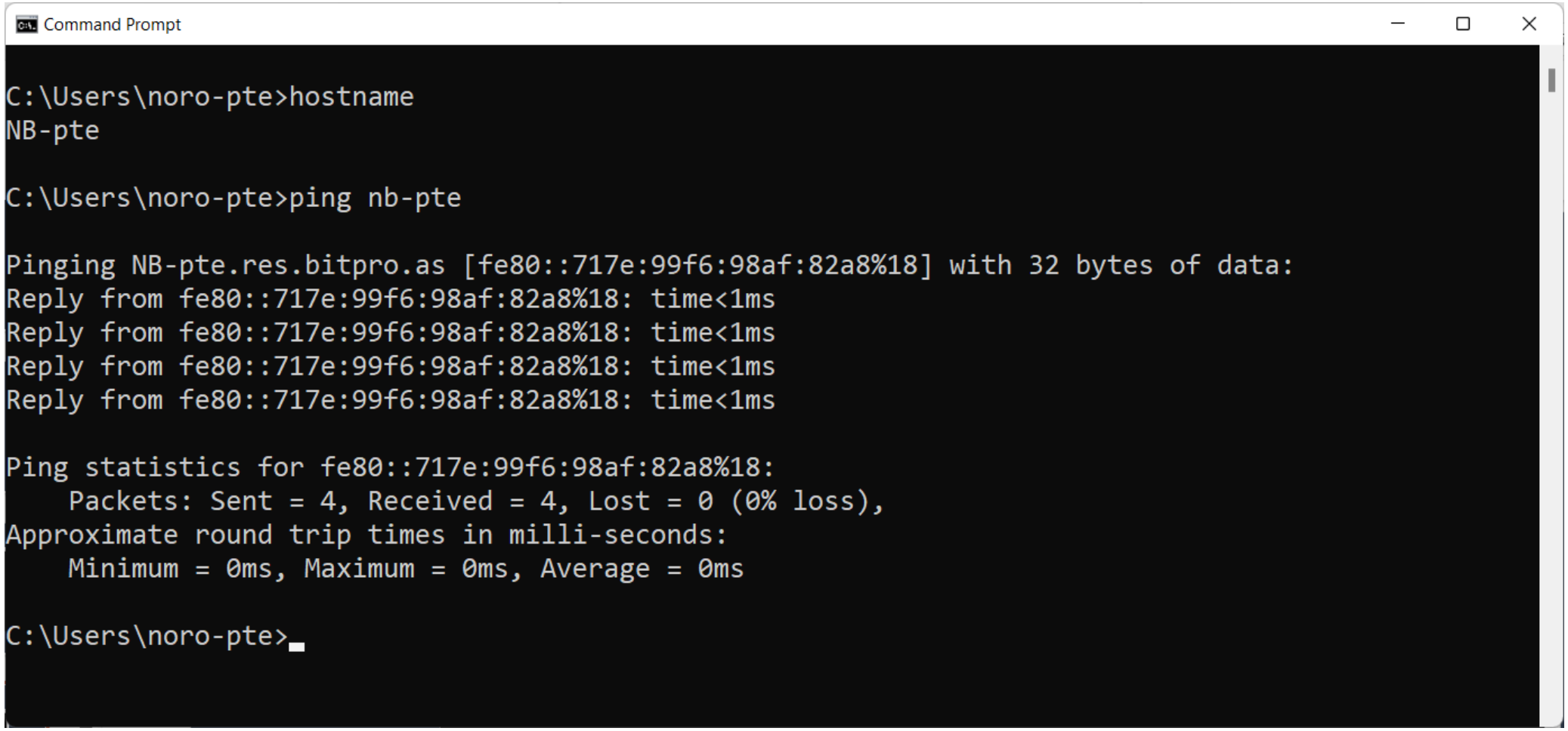
Høyskolen Kristiania



- Displays the MAC address of various network devices

hostname (Windows/Linux)

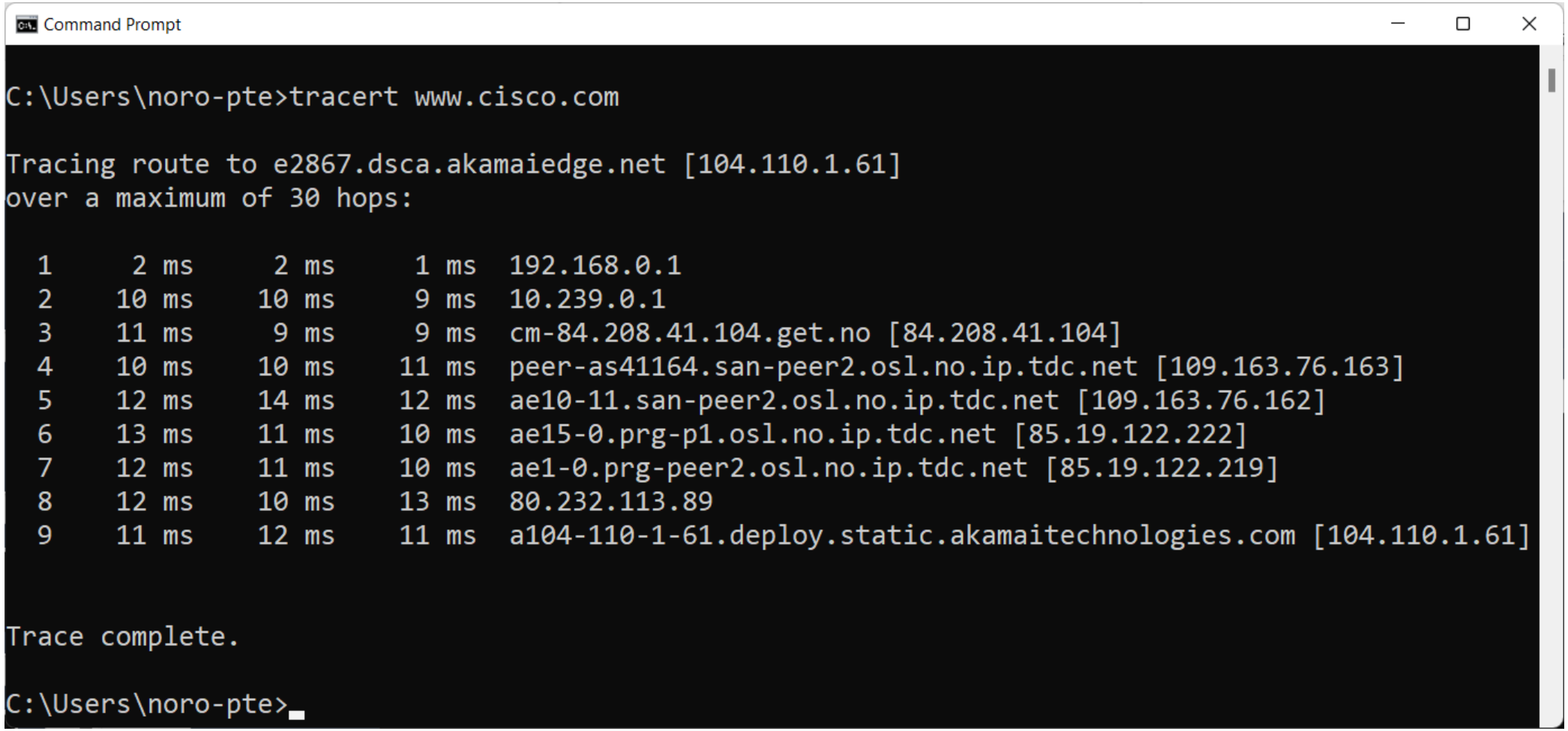
Høyskolen Kristiania



- Displays the current name of your computer

tracert (Windows), traceroot (Linux)

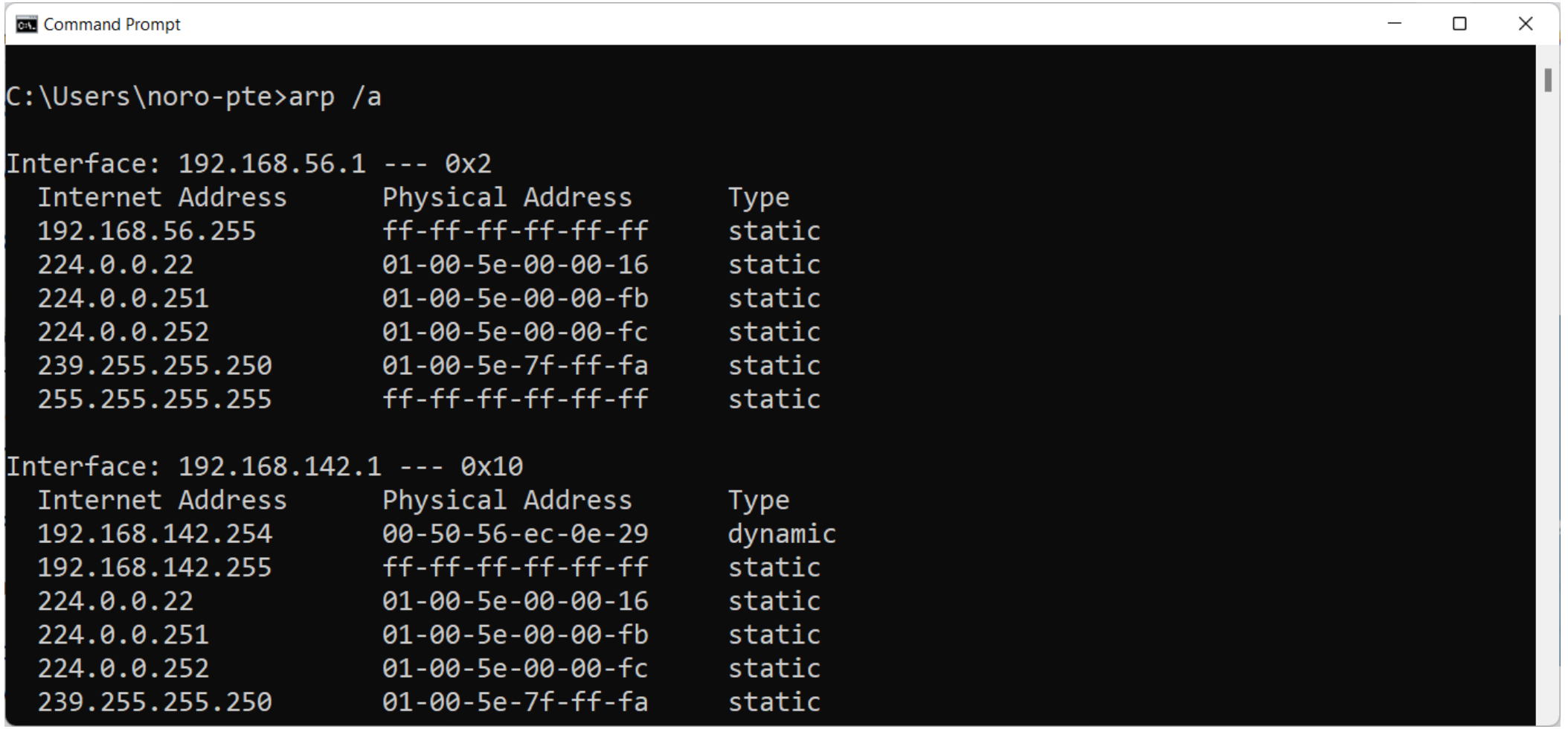
Høyskolen Kristiania



- Traces the route a data packet takes before reaching its destination

arp (Windows/Linux)

Høyskolen Kristiania



- Displays entries in the Address Resolution Protocol (ARP) cache

Høyskolen Kristiania

End

Summary

What should you know?

Høyskolen Kristiania

- Explain what an IPv4 address and subnet mask are
- Find out which IP network a given IP address belongs to, and calculate the broadcast address
- Recognize localhost, private and self-configured addresses (IPv4).
- Explain the role of the Standard Gateway

What should you know?

Høyskolen Kristiania

- Describe how DHCP works
- Describe how NAT (NAT/PAT) works
- Describe the role to routers and routing-tabeles
- use `ipconfig` (`ifconfig`), `ping`, `tracert` (`traceroute`) and `netstat -r` (`route print`)

Exercises

Høyskolen Kristiania

- Text assignment set

- Practical exercises
  - ipconfig /release, ipconfig /renew (Linux/Mac: ifconfig)
  - ping
  - tracert (Linux/Mac: traceroute)
  - netstat –r
  - Wireshark – now look at the network layer
  - Wireshark - go through the latest tasks with HTTP, FTP etc; but look at the network layer

Høyskolen Kristiania

For optional self-study

For those who want to learn some topics more in depth to understand it better, here are some extra topics related to today's teaching, it must be expected some personal work to understand these topics.

There will be no questions on the exam from these, and this is therefore not considered to be part of the syllabus.

Virtual connection(VC)

Høyskolen Kristiania

Common on the network layer is the switched connection, but it is possible to set up a fixed connection:

- Much like a telephone connection
- Must establish a connection (route) first
- Each packet contains the connection's route ID
- Each router (stopover) maintains the state of each route
  - The transport layer only sees each end of a route
- Data is transferred when the route is established
- The route is "torn down" when the transfer is complete

Forwardingtable: example

Høyskolen Kristiania

Address area for recipient		Link interface
11001000 00010111 00010000 00000000 **11001000 00010111 00010**111 11111111	to	0
11001000 00010111 00011000 00000000 **11001000 00010111 00011**000 11111111	to	1
11001000 00010111 00011001 00000000 **11001000 00010111 00011**111 11111111	to	2
	or	3

Is it something wrong here?

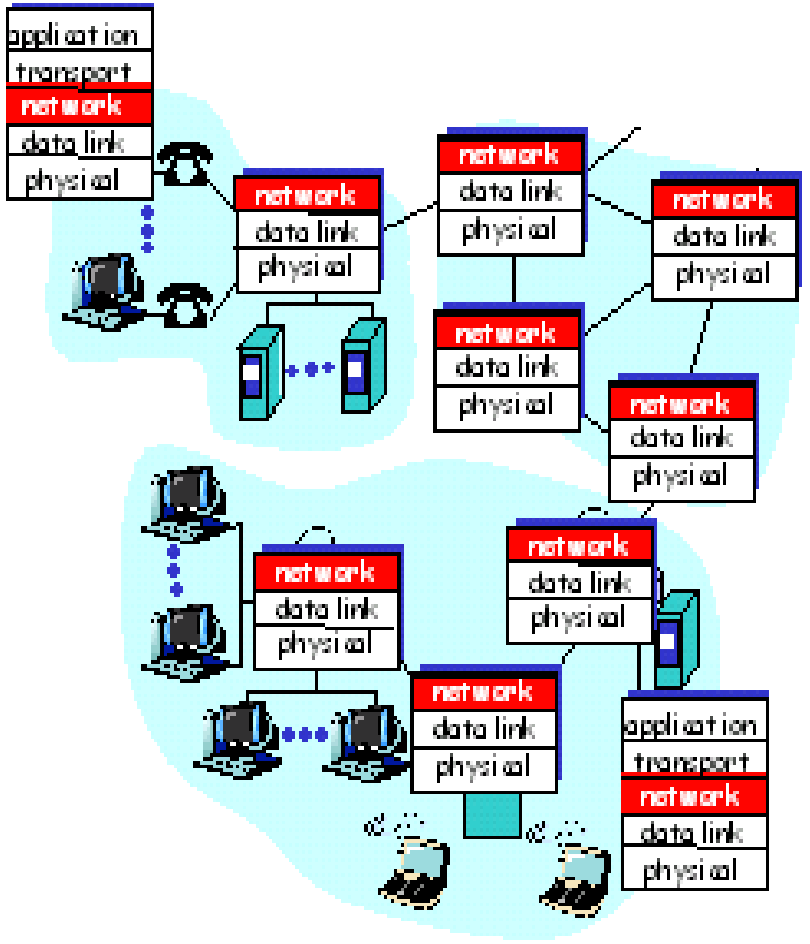
Høyskolen Kristiania

Wireless LAN adapter WLANUSB:

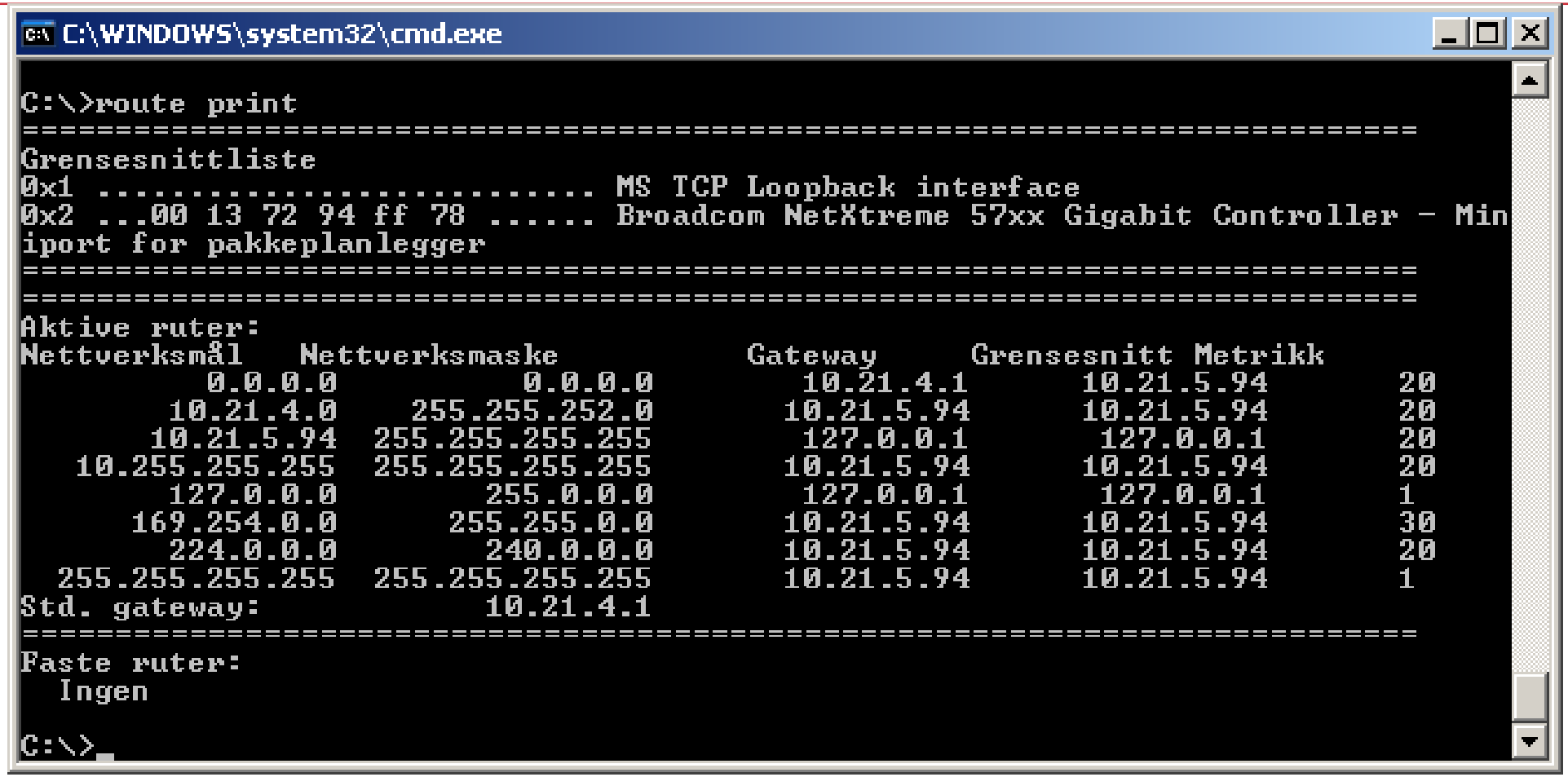
Connection-specific DNS Suffix . : ad.nith.no
Description . . . . . : D-Link DWA-140 Wireless N USB Adapter(rev.B3)
Physical Address. . . . . : B8-A3-86-90-50-E8
DHCP Enabled. . . . . : Yes
Autoconfiguration Enabled . . . . : Yes
Link-local IPv6 Address . . . . . : fe80::50e5:40ff:6794:1d5a%16 (Preferred)
IPv4 Address. . . . . : 10.21.**30**.228 (Preferred)
Subnet Mask . . . . . : 255.255.252.0
Lease Obtained. . . . . : 6. november 2013 15:25:30
Lease Expires . . . . . : 6. november 2013 23:25:30
Default Gateway . . . . . : 10.21.28.1
DHCP Server . . . . . : 1.1.1.1
DHCPv6 IAID . . . . . : 548971398
DHCPv6 Client DUID. . . . . : 00-01-00-01-14-6A-F2-0B-D8-D3-85-77-A0-3F
DNS Servers . . . . . : 158.36.131.10
Primary WINS Server . . . . . : 158.36.131.10
NetBIOS over Tcpip. . . . . : Enabled

The network layer

- Moves packets from sender to recipient
- Network protocol also on each stopover
- Routing from sender to receiver
- Switching packets from router input side to router output side
- If necessary, the router call setup is defined for the entire route before the packet is sent



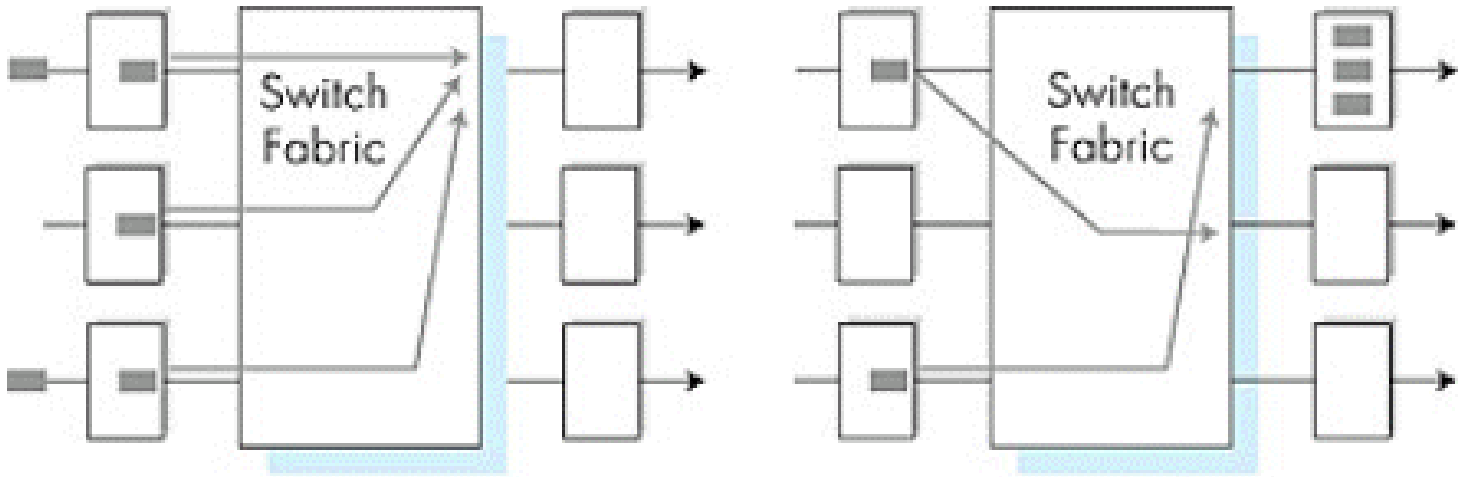
route print



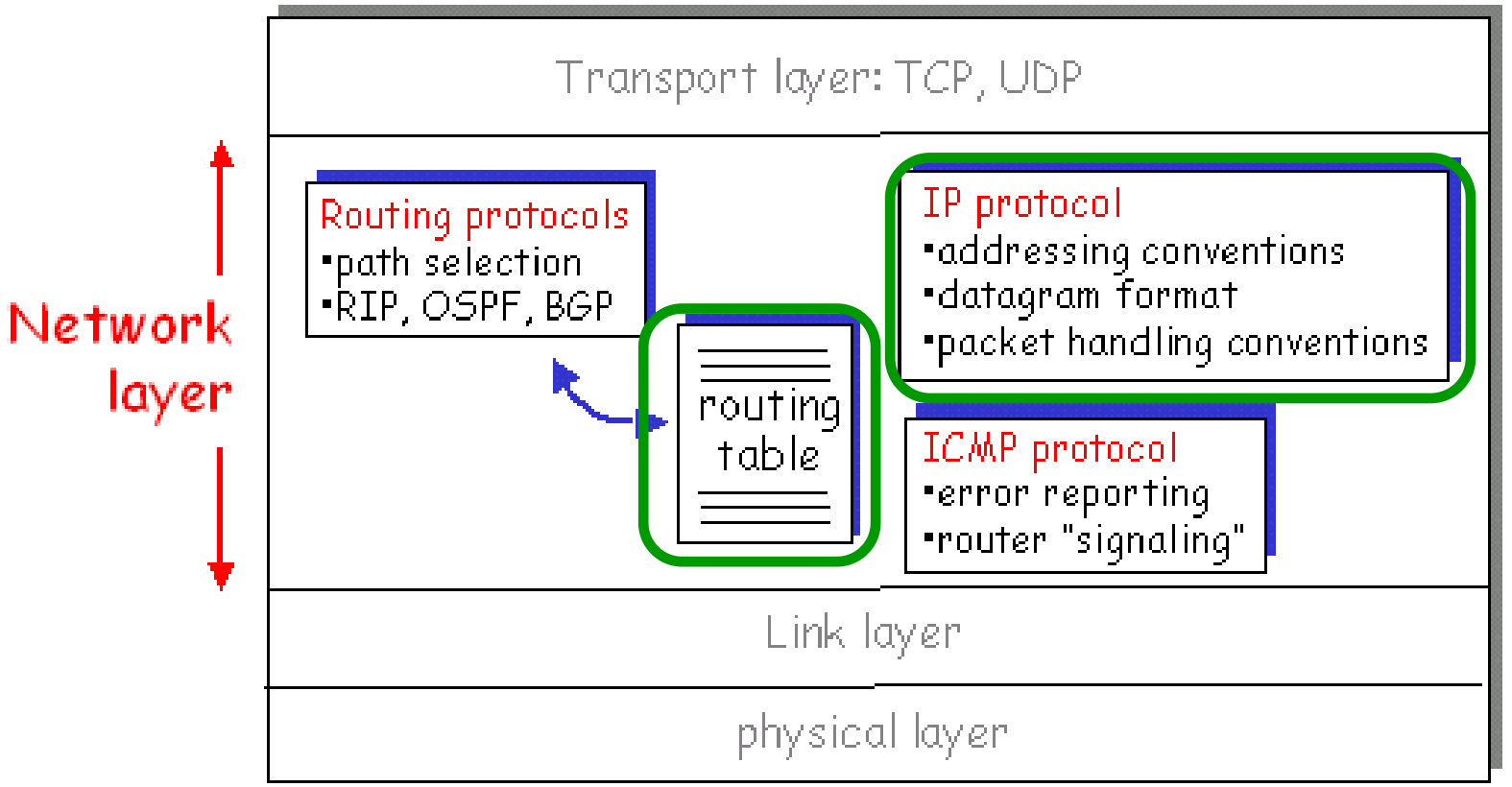
Can also use `netstat -r`

Queuing at the outgoing gate

- Data can be lost in the queue if the buffer is not large enough



Internet network layer



ipconfig (ifconfig)

Høyskolen Kristiania

- `ipconfig` shows the network parameters of the interfaces / adapters

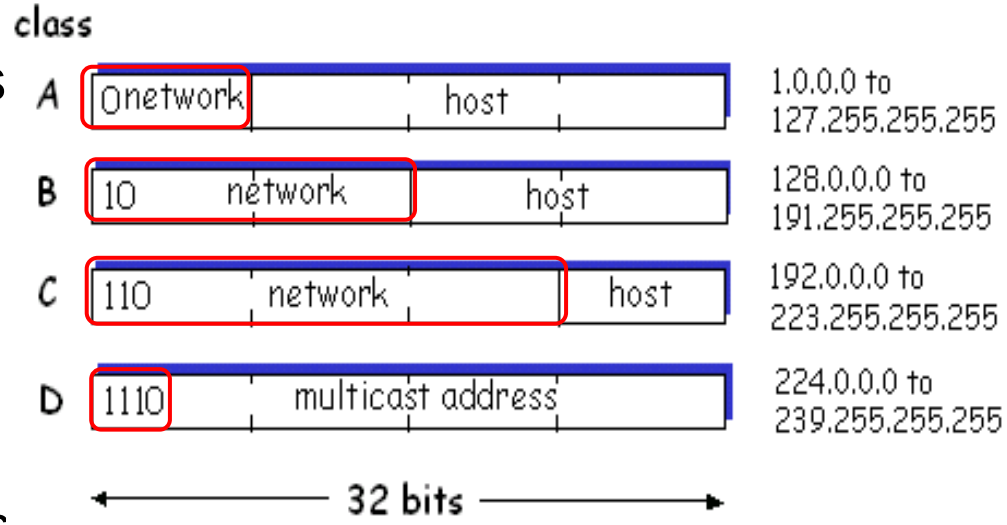
```
C:\Users\blistog>ipconfig

Windows IP Configuration

Ethernet adapter e0:

   Connection-specific DNS Suffix  . :
   IPv6 Address. . . . . . . . . . . : 2001:700:2e00::51
   Link-local IPv6 Address . . . . . : fe80::b46c:b98f:85ec:dba0%12
   IPv4 Address. . . . . . . . . . . : 158.36.131.51
   Subnet Mask . . . . . . . . . . . : 255.255.255.128
   Default Gateway . . . . . . . . . : 2001:700:2e00::1
                                       158.36.131.1

Ethernet adapter Bluetooth Network Connection:

   Media State . . . . . . . . . . . : Media disconnected
   Connection-specific DNS Suffix  . :

Tunnel adapter isatap.{84470FB0-16A7-4A31-B6B9-8AECF834115C}:

   Media State . . . . . . . . . . . : Media disconnected
   Connection-specific DNS Suffix  . :

Tunnel adapter Teredo Tunneling Pseudo-Interface:

   Connection-specific DNS Suffix  . :
   IPv6 Address. . . . . . . . . . . : 2001:0:5ef5:73b8:3826:3138:61db:7ccc
   Link-local IPv6 Address . . . . . : fe80::3826:3138:61db:7ccc%15
   Default Gateway . . . . . . . . . :

Tunnel adapter isatap.{C682A4AE-3FEC-4053-8456-5A377FD0FF55}:

   Media State . . . . . . . . . . . : Media disconnected
   Connection-specific DNS Suffix  . :
```

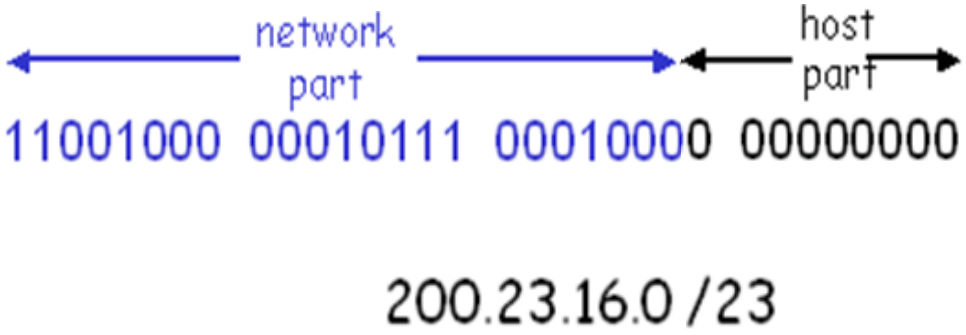
IPv4 Address & Subnet Mask = Network Prefix from which it is routed; Std Gateway = the way out into the Internet

Techniques for sending IPv6 through IPv4 networks

IP addresses: classes and CIDR

- Originally divided into 6 different classes, each with its own predefined prefix - length
- Classification of addresses became too "rigid"
  - A class may contain (many) unused addresses
- Class A, B, C are common addresses, D is multicast, E is reserved for research, and 127. * is a reserved "class" for loopback
- Classless Inter-Domain Routing (CIDR)
  - The network part has any length, x
  - Format a.b.c.d / x





DHCP

- How does the DHCP server know where to send your network parameters (IP, subnet mask, std gw, DNS, etc.)?
  - Your machine broadcasts (MAC address: FF-FF-FF-FF-FF-FF-FF) the first request in the LAN
  - If there is a DHCP server there, then it responds with an offer of IP m.m.
  - The rest can then take place at Nettverkslaget
  - Sets a period for which you "lease" the parameters
  - Must be renewed when the lease expires.

IP Address & Subnet Mask = IP Network

Høyskolen Kristiania

- Machines / adapters must belong to the same IP network in order to send directly to each other
  - 10.21.3.5 / 255.255.254.0 can send directly to 10.21.2.255 / 255.255.254.0
  - 10.21.3.5 / 255.255.255.0 must send via gateway (router) for å nå 10.21.2.255 / 255.255.255.0
- **The prefix** is determined by the IP address and the subnet mask, and it is this that determines whether you belong to the same IP network or not.

IP addressing: CIDR

- “Classfull” addressing (A, B, C, D, ..):
  - inefficient use of address spaces, quickly runs out of available addresses
  - eg: a class B network has enough addresses for 65,000 machines, even if it is only e.g. 2000 machines in the network
- CIDR: Classless InterDomain Routing
  - Network part (prefix) of the address is of arbitrary length
  - address format: a.b.c.d/x, where x is the number of bits in the network part of the addres.

200.23.16.0/23

Special IP addresses

Høyskolen Kristiania

- Some IP addresses are reserved for special use
  - Private addresses
  - Documentation
  - Self-configuration
  - Broadcast
  - Multicast
  - Network address (entire local IP network)
  - Temporary addressing
  - Loopback (myself)
- See RFC 1166

Special IP addresses (2)

- *Private addresses are used only within a WAN*
  - cannot be routed outside the LAN / WAN
  - dropped automatically by Internet routers
- Provides flexibility for organizations internally
- The same address can also have external IP (NAT)

IPv4 addresses	Network
10.0.0.0 – 10.255.255.255	1 class A network
172.16.0.0 – 172.31.255.255	16 class B network
192.168.0.0 – 192.168.255.255	65536 class C network

RFC 1918

Special IP-addresses (3)

- In documentation, use addresses that are not used anywhere else
  - 192.0.**2**.0/24
  - 198.51.100.0/24
  - 203.0.113.0/24
- When self-configuring the IP address, the DHCP server may be unavailable
  - then uses an "automatic", special address:
    - **169.254.1.0 – 169.254.254.254** (/16)
    - These are also not route-only
    - Most often these can be interpreted as that there are problems with contacting DHCP server, or that you do not have access to LAN

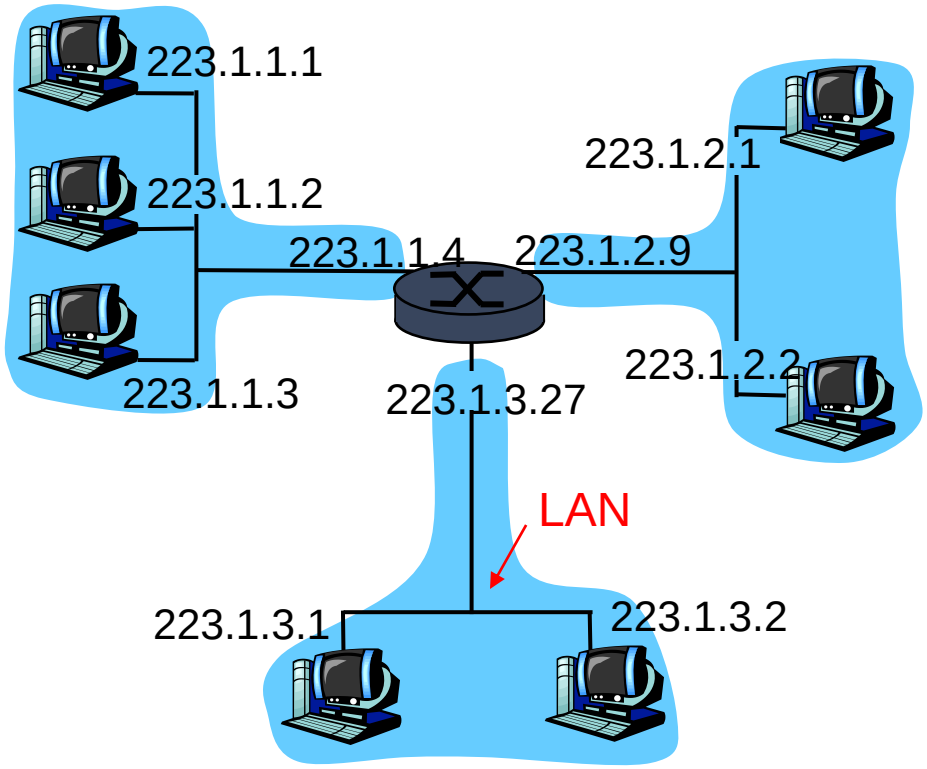
Special IP addresses (5)

Høyskolen Kristiania

- *Multicast is the same as broadcasting limited to a group of nodes in a list (located on the router)*
- 224.0.0.0 – 239.255.255.255

Subnet

- IP addresses - two parts:
  - subnet part (most significant bits, left end bits)
  - machine part (least significant bits, bits at the right end)
- *What is a subnet?*
  - interface with equal subnet part of the IP address
  - can reach each other physically without going through routes

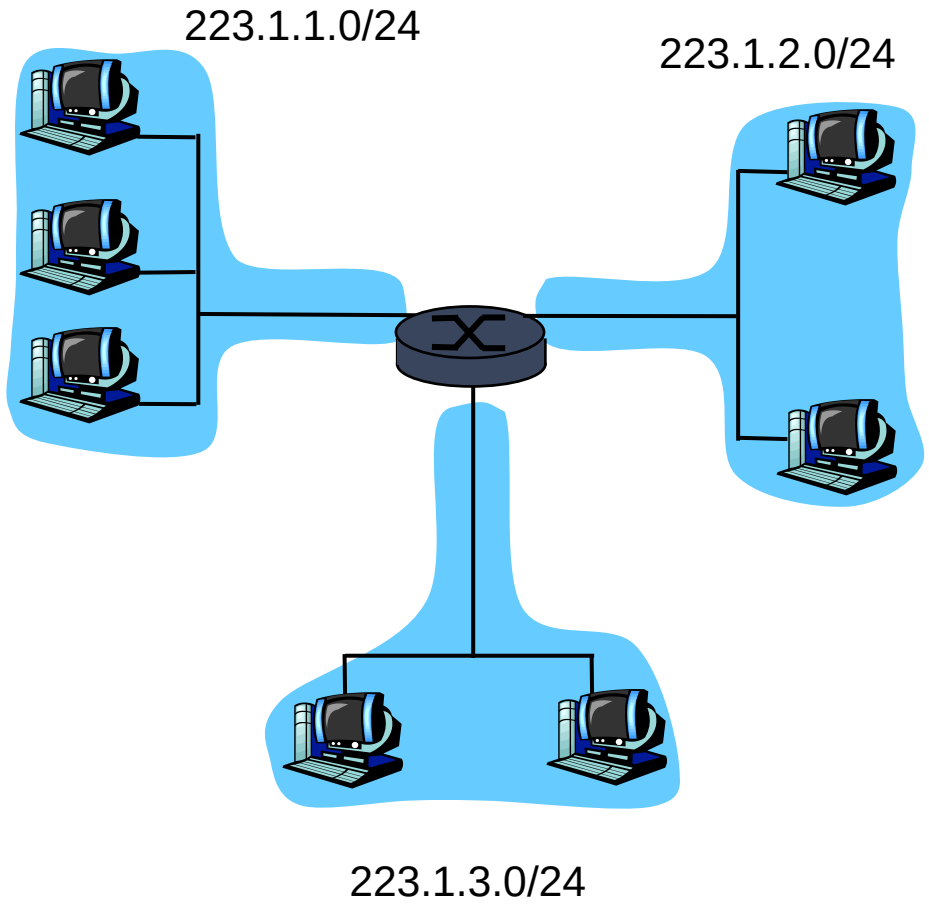


network consisting of 3 IP networks

Subnet

Recipe

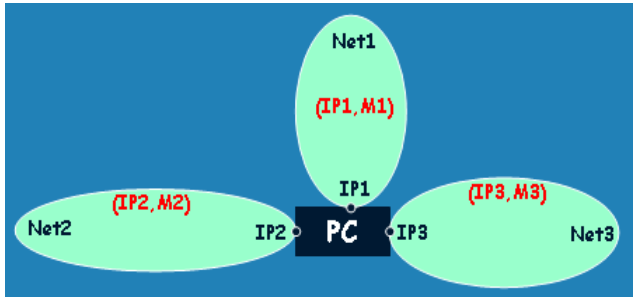
- To find the subnets, disconnect each interface from its machine or router so that we get islands of isolated networks. Each isolated network is then called a subnet.
- Machines on different subnets must then have a router in between to get in touch with each other



Subnet mask: / 24

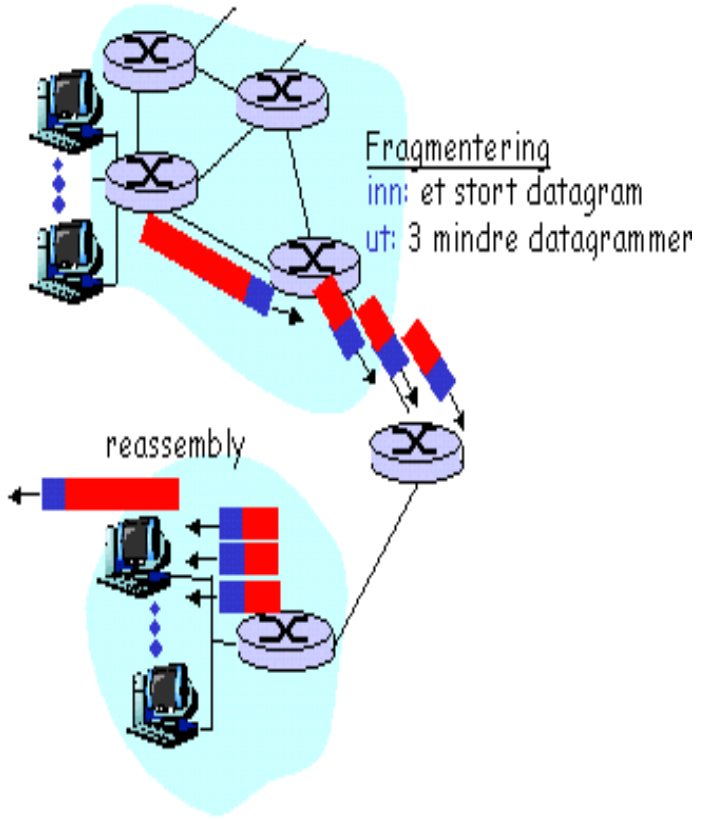
Broadcast

- When broadcasting (general query for a service) you can address either the local address space (subnet) or the entire IP network
- Limited broadcasting
  - 255.255.255.255
- Network broadcasting
  - Uses the web part of the address
  - 192.0.2.235/24 will use 192.0.2.255
  - 192.0.2.5/27 will use 192.0.2.31
  - 11000000000000000000000010000111111

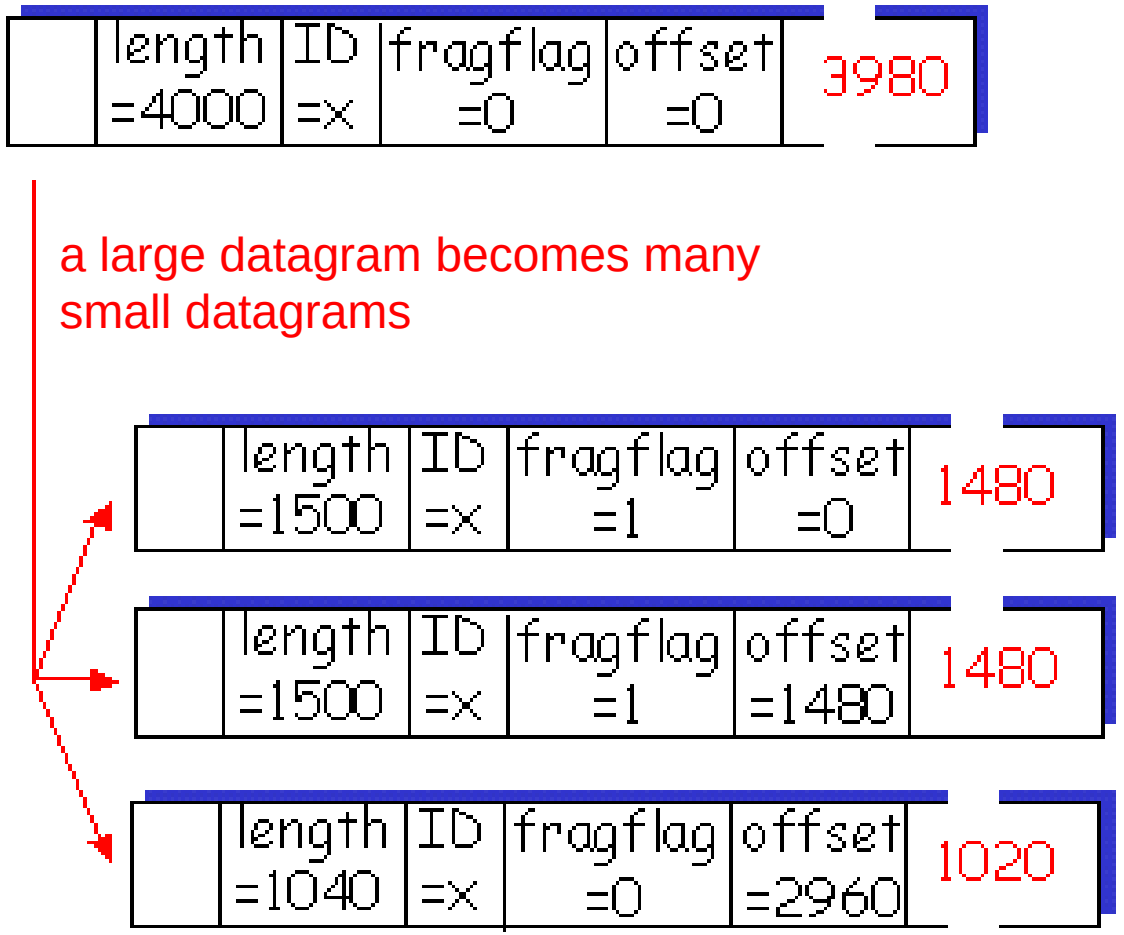


IP fragmentation

- Networks have package size restrictions, Maximum Transfer Unit (MTU)
  - All routers must handle a minimum of 576 bytes MTU, so it has become defacto packet size
  - Most routers have MTU 1500 bytes (incl ethernet header, IP header, TCP header, etc)
- Large datagrams are divided into smaller, independent but coherent datagrams
  - Fragmentation flag
  - Offset
- Assembled at the recipient
  - One fails, all fails
  - TCP resends the entire datagram



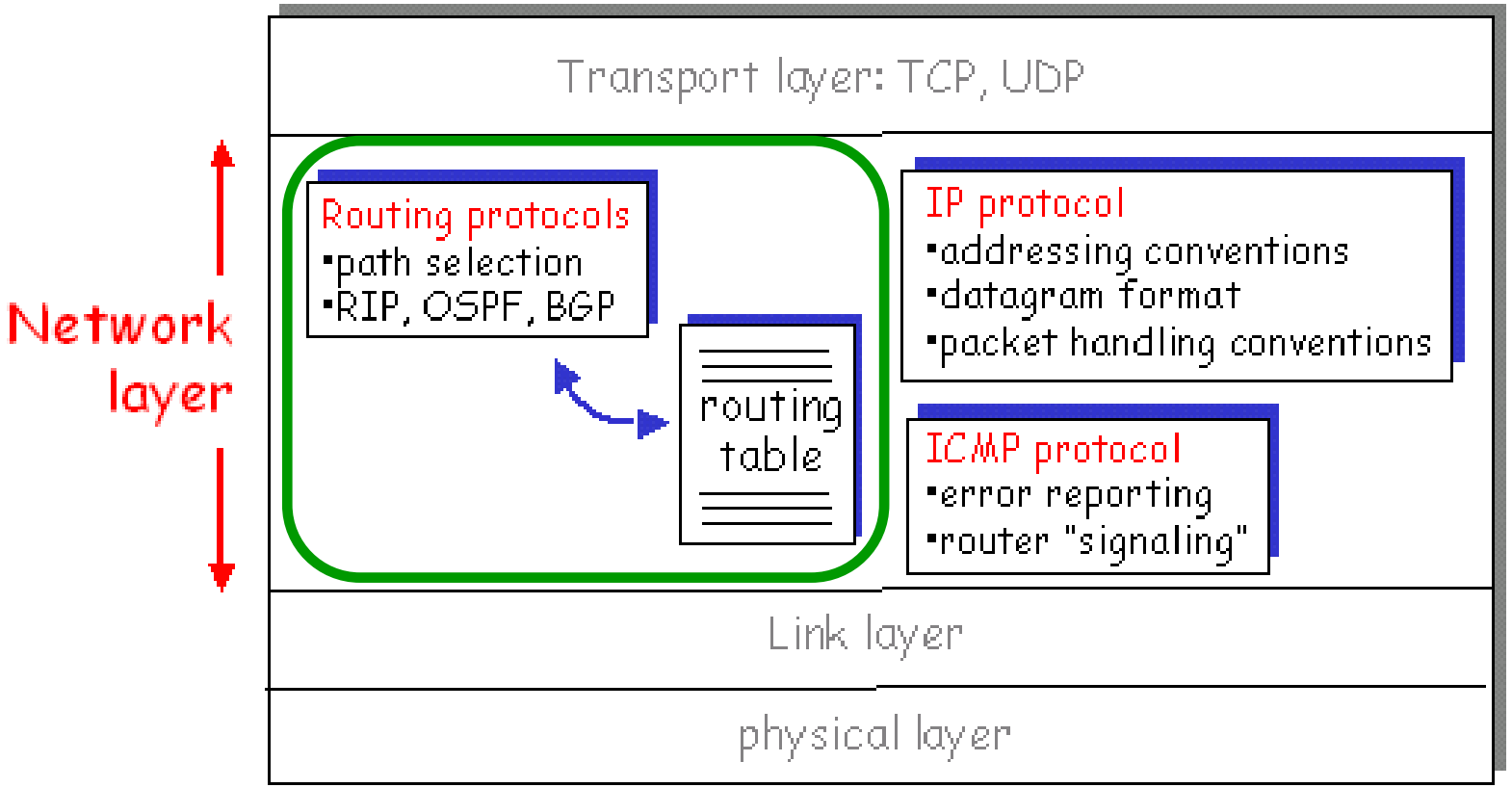
IP fragmentering



Høyskolen Kristiania

Routing in the trunk network

Internet network team



NAT: Network Address Translation

Høyskolen Kristiania

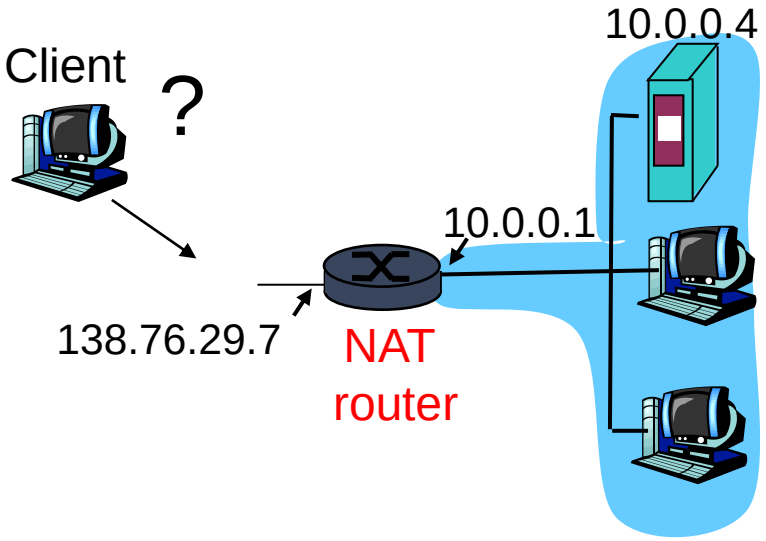
Implementation:

NAT-router:

- *outgoing datagrams: replace* senders IP address and port number with NAT IP address and new port number
  - . . . responding machines will then use the NAT IP address and the new port number as the recipient address.
- *remember (in NAT translation table)* each (sender´s IP address, port number) to (NAT IP address, new port number) translation pair
- *incoming datagrams: replace* NAT IP address and the new port number in recipient fields with the corresponding sender IP address and port number stored in NAT table

NAT traversal problem

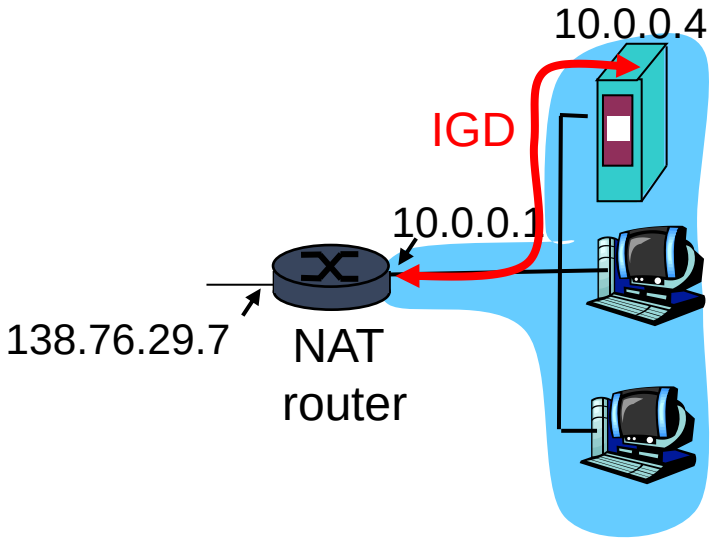
- External client wants to server with address 10.0.0.4
  - the server address 10.0.0.4 is local on LANdet (the client can not use it as the recipient address)
  - Only an externally visible NAT address: 138.76.29.7
- Solution 1: Static configure NAT to forward incoming connection requests to a specific port on the server
  - For example ,, (123.76.29.7, port 2500) always to 10.0.0.4 port 25000



NAT traversal problem

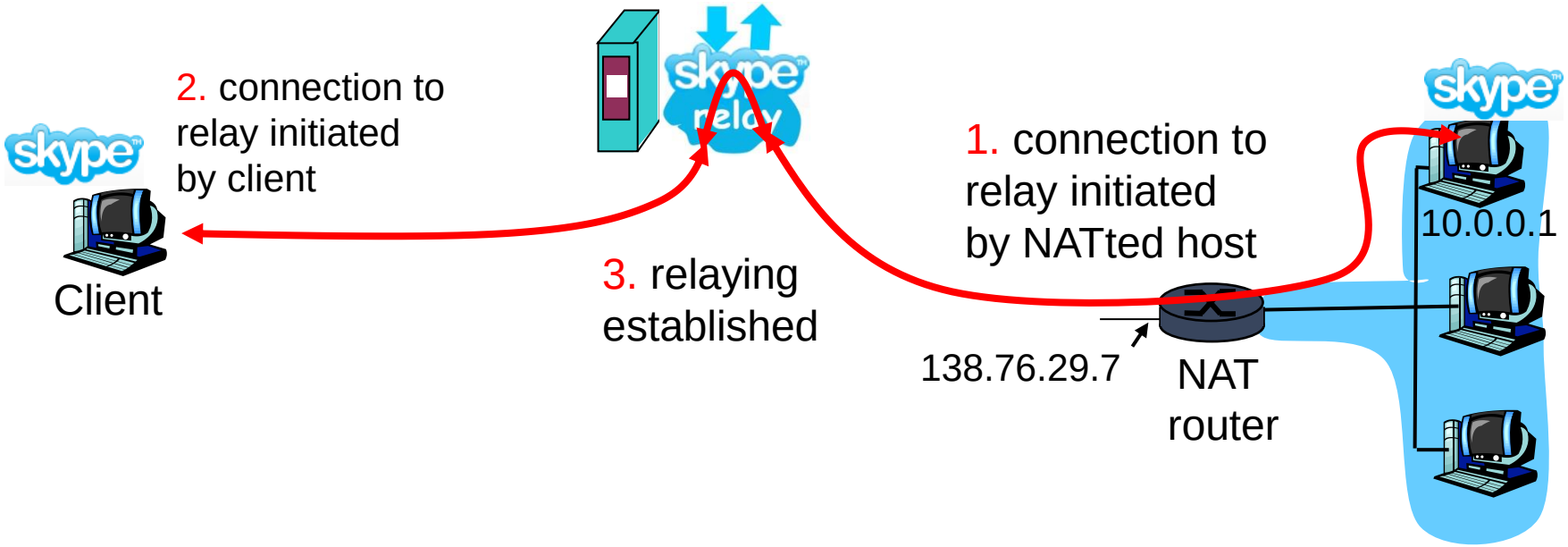
- Solution 2: Universal Plug and Play (UPnP) Internet Gateway Device (IGD) protocol. Allows the NAT machine to:
  - learn public IP address (138.76.29.7)
  - Add / remove port mappings (with lease times) on router

  ie, automate static NAT “port map configuration”

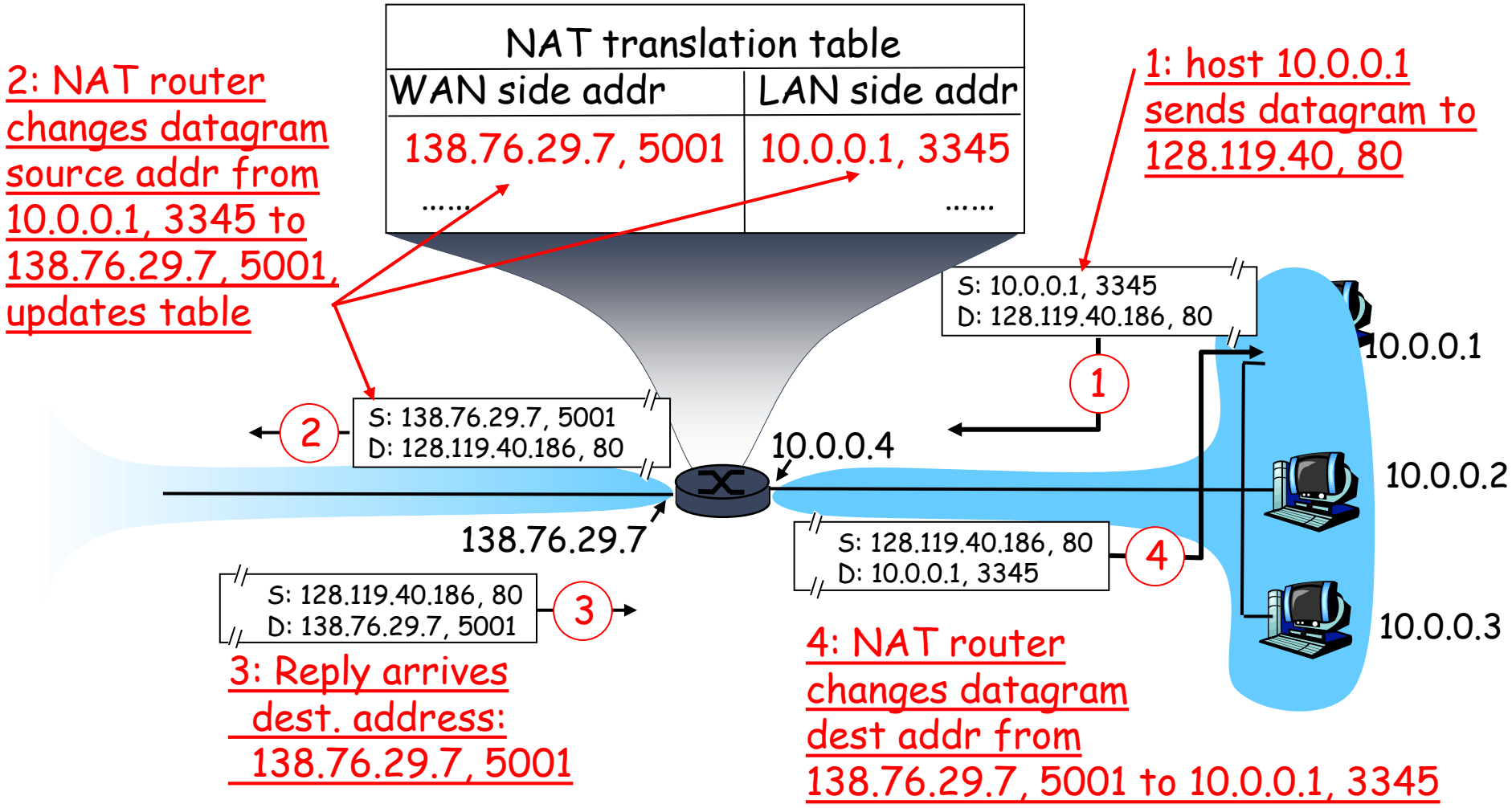


NAT traversal problem

- Solution 3: relaying (e.g Skype)
  - The NAT client established connection to relay
  - External client connects to relay
  - relay forwards packets between two connections

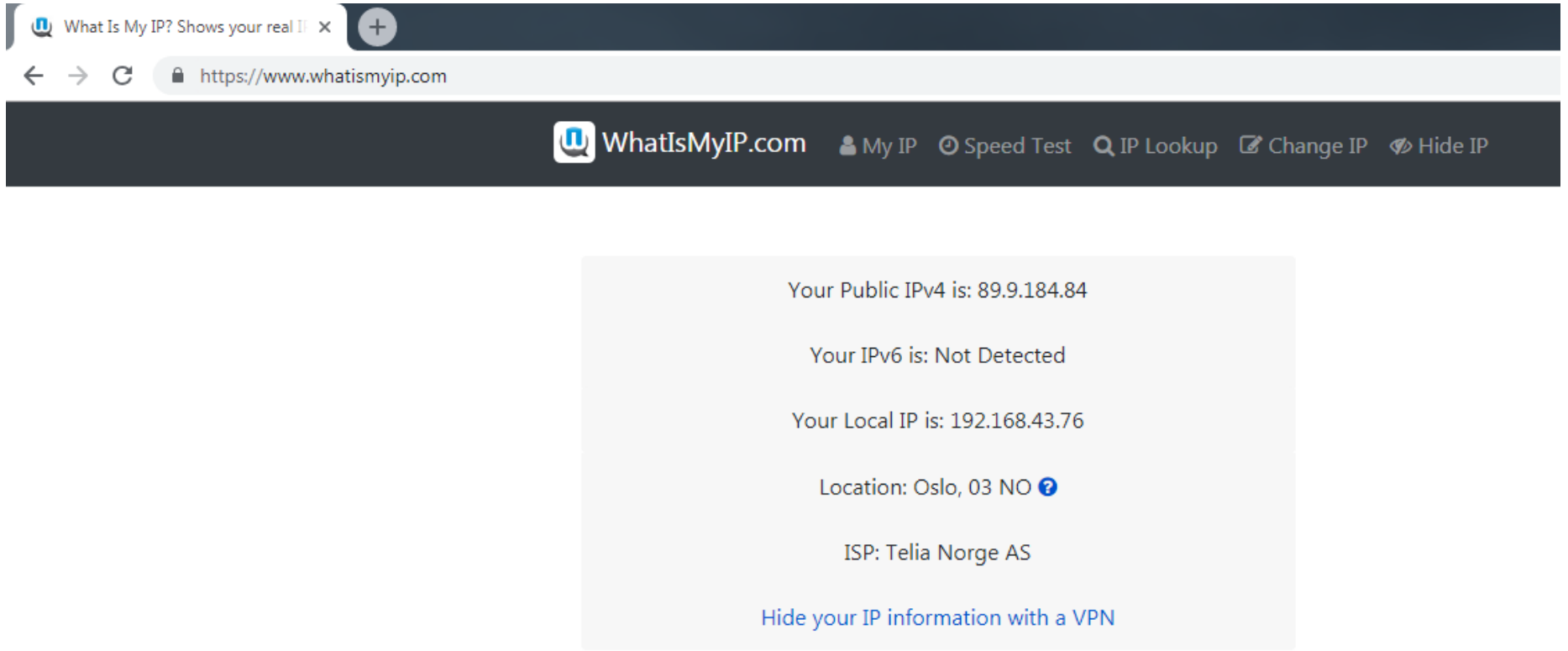


NAT: Network Address Translation



Find your remote IP

Høyskolen Kristiania



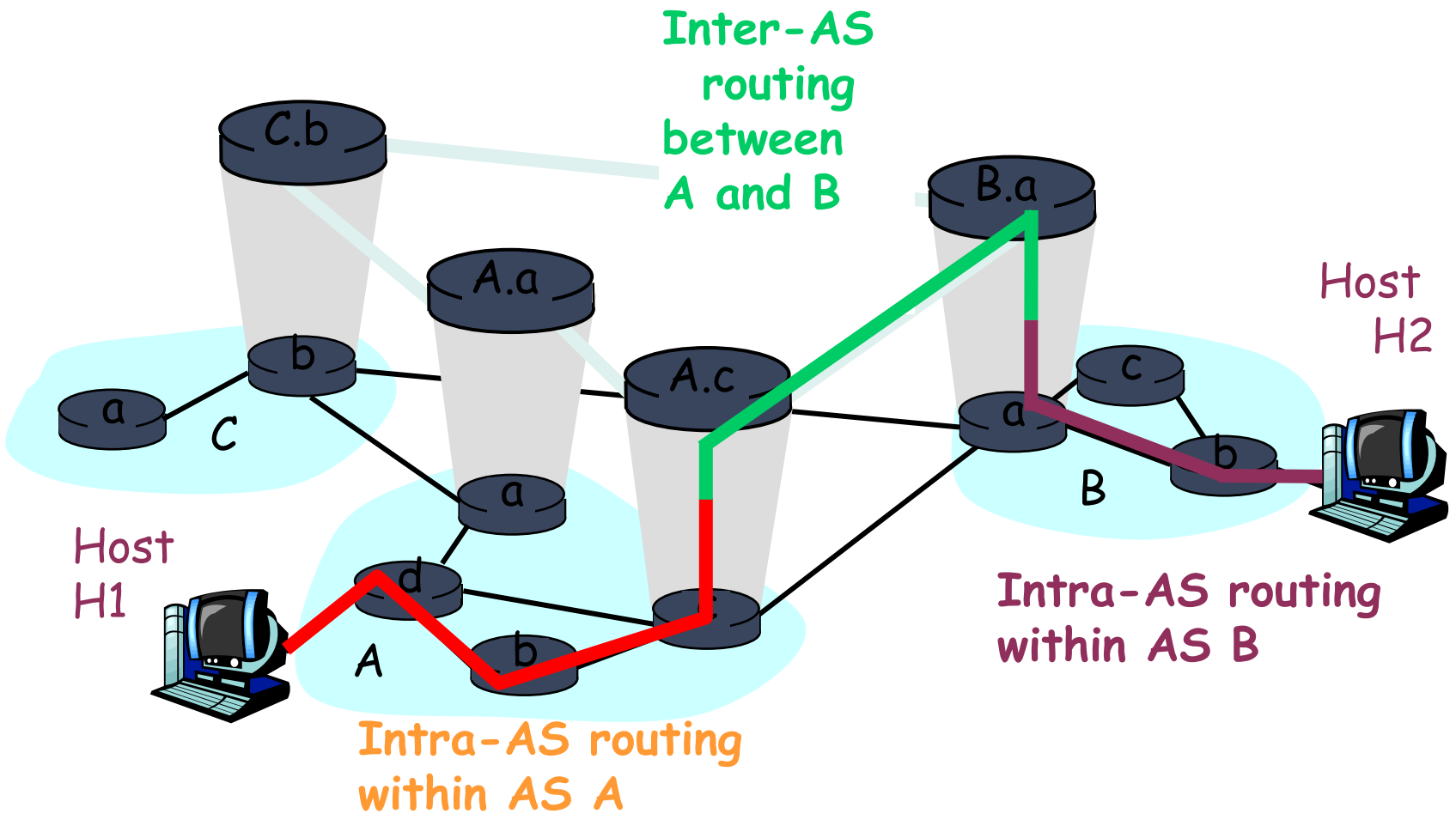
Routing

- Routing is about setting up a routing table for each individual router
- The table indicates which interface a datagram should be forwarded from
- All host machines have routing tables, but they are small
- Routers in the backbone network have tables with ~ 200000 lines!
  - Each line is an IP network that the router has found one **cheapest** way to

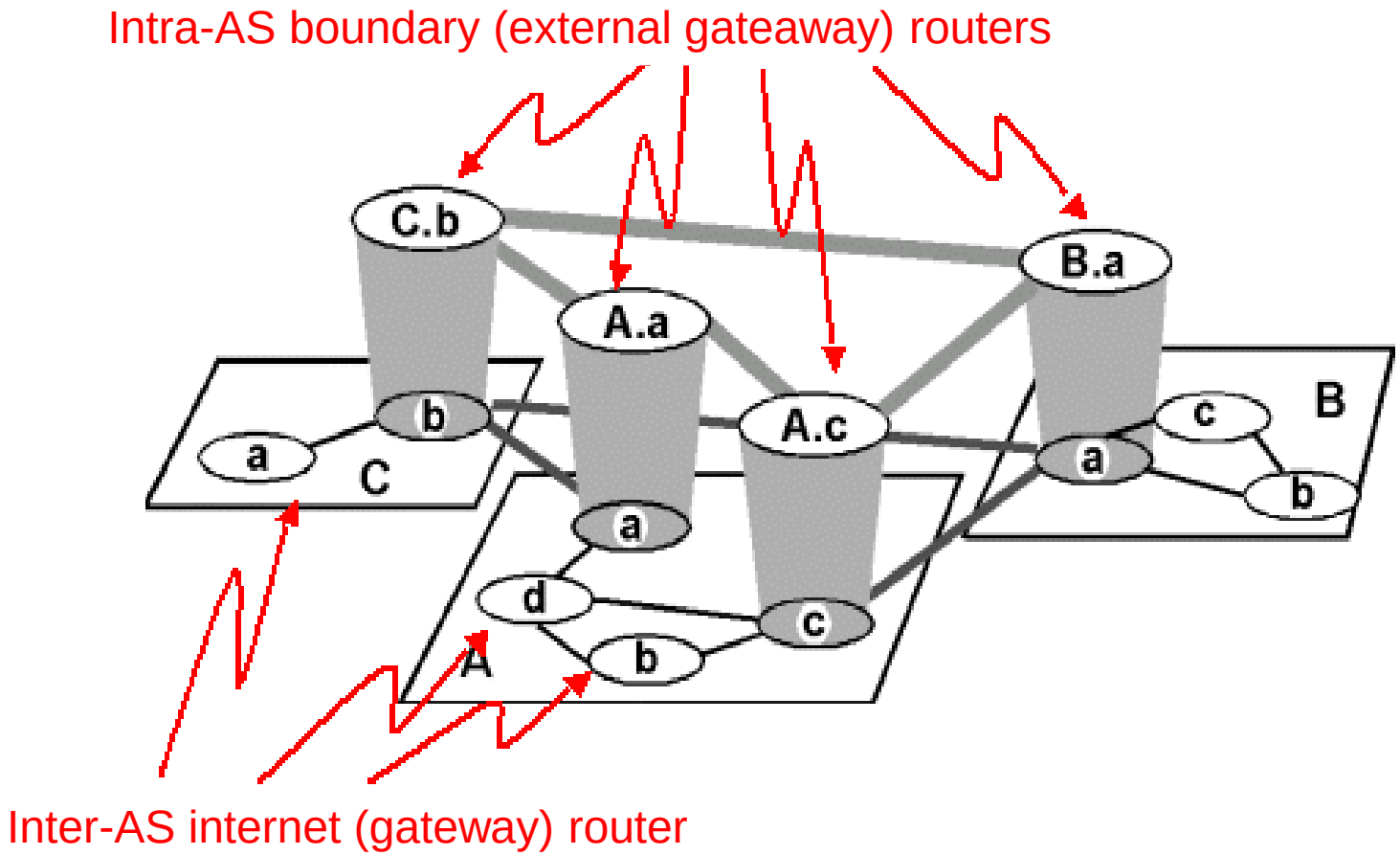
Hierarchical routing

- The internet is large and complex
- Internet = network of networks
- Divides the routers into regions
  - **Autonomous systems** (AS)
- Routers in the same AS run the same routing protocol (intra-AS protocol)
- Gateway performs inter-AS routing between different ASs
  - Uses BGP as a routing protocol

Intra-AS and Inter-AS routing



Internet AS hierarchy

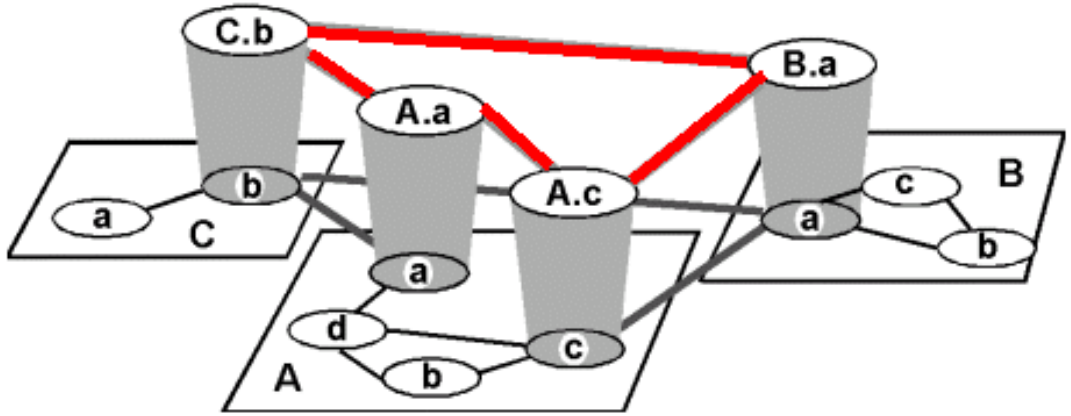


Intra-AS routing

- Also called Interior Gateway Protocols (IGP)
- The most common are
  - Routing Information Protocol (RIP)
  - Open Shortest Path First (OSPF)
  - Interior Gateway Routing Protocol (IGRP)
    - Cisco proprietary
    - EIGRP further development of IGRP
- Implemented in programs on the routers
  - Exchanges routing information with other routers within AS / WAN

Internet **inter**-AS routing

- Border Gateway Protocol (**BGPv4**) is the standard of the Internet
- Also routes from AS-Nummer
- Uses Path Vector protocol
  - Finds the cheapest way out from "neighbor gossip"
  - Saves the "AS route" (path) to the recipient



BGP - Border Gateway Protocol

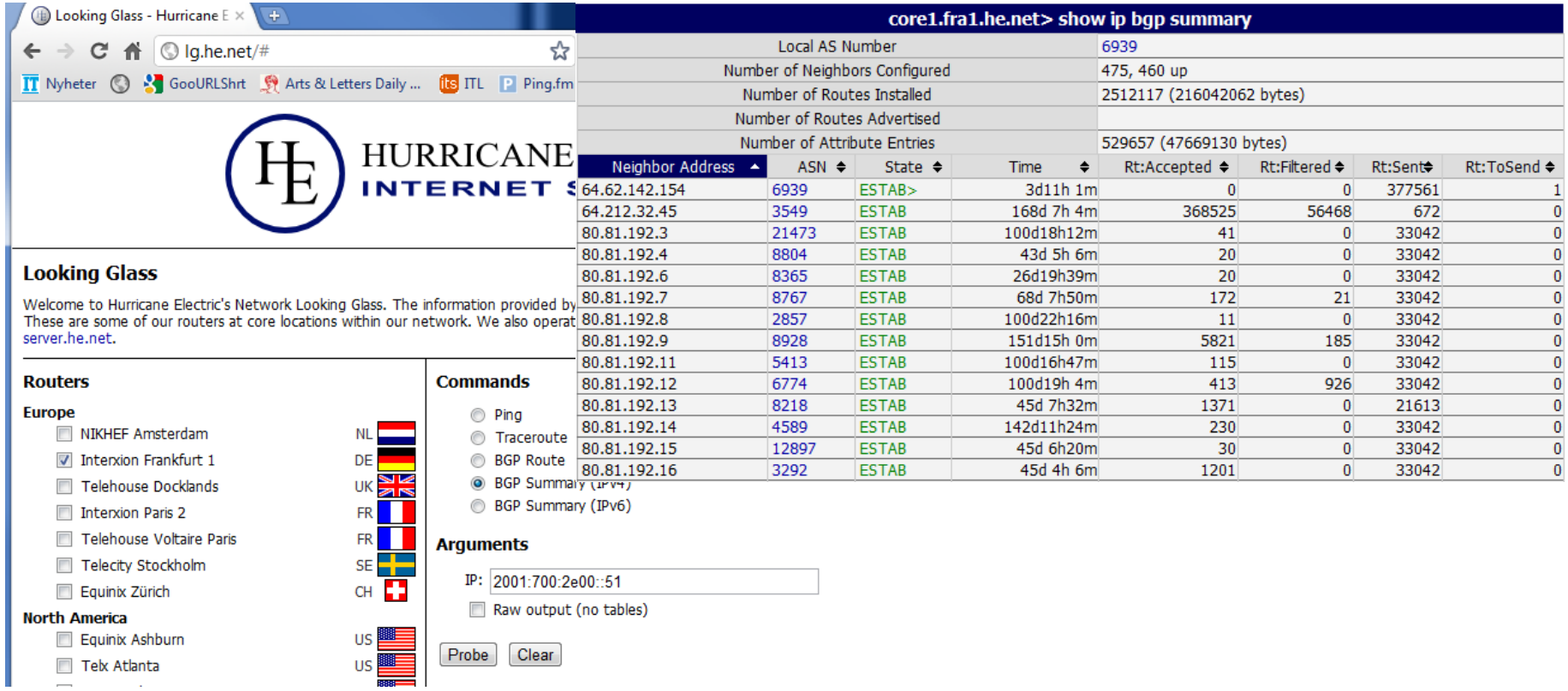
- When gateway X sends its route proposal to node Z over to gateway Y, the following can happen:
  - The route is not accepted (cost, policy, reliability ...)
  - The route is accepted and then used from Y to Z
  - Y sends its updated route cost to its «other neighbors»
- Uses TCP between routers.
- 4 types of messages
  - OPEN: Opens TCP connection and confirms sender
  - UPDATE: Opens a new route or closes an old one
  - KEEPALIVE: Keeps the connection alive without UPDATE
  - NOTIFICATION: Sends error message. Closes connection

Why are different Intra- and Inter-AS routing used?

- Policy
  - Inter-AS: Wants control over the routing between the areas; must take into account peering agreements and prices
  - Intra-AS: Uniform control of routing within the area; primarily looking for efficiency and load balancing
- Scale
  - Hierarchical routing saves table space and reduces the amount of updates
- Performance
  - Inter-AS: Policy can be more important than performance
  - Intra-AS: Focuses (most often) on performance

Looking glass routere

- Some "kind" companies let you see trunk network routers from the "inside" e.g. http://lg.he.net



Høyskolen Kristiania

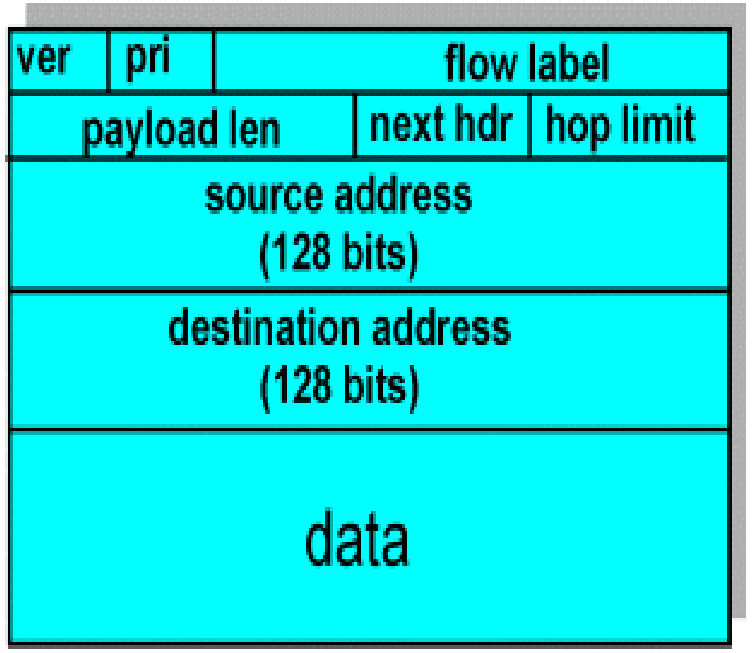
Internet Protocol v6

IPv4 → IPv6

Høyskolen Kristiania

- The colossal growth of the Internet requires more and more addresses
- Feb 3 In 2011, the latest IPv4 address blocks were assigned to the Continental Registers!
- IPv6 increases the address field from 32 to 128 bits
- Simplifies header format
  - Removes the possibility of fragmentation
  - Removes the checksum.
- Simplifies addressing and data flow
- Fixed header length = 40 bytes

IPv6 header

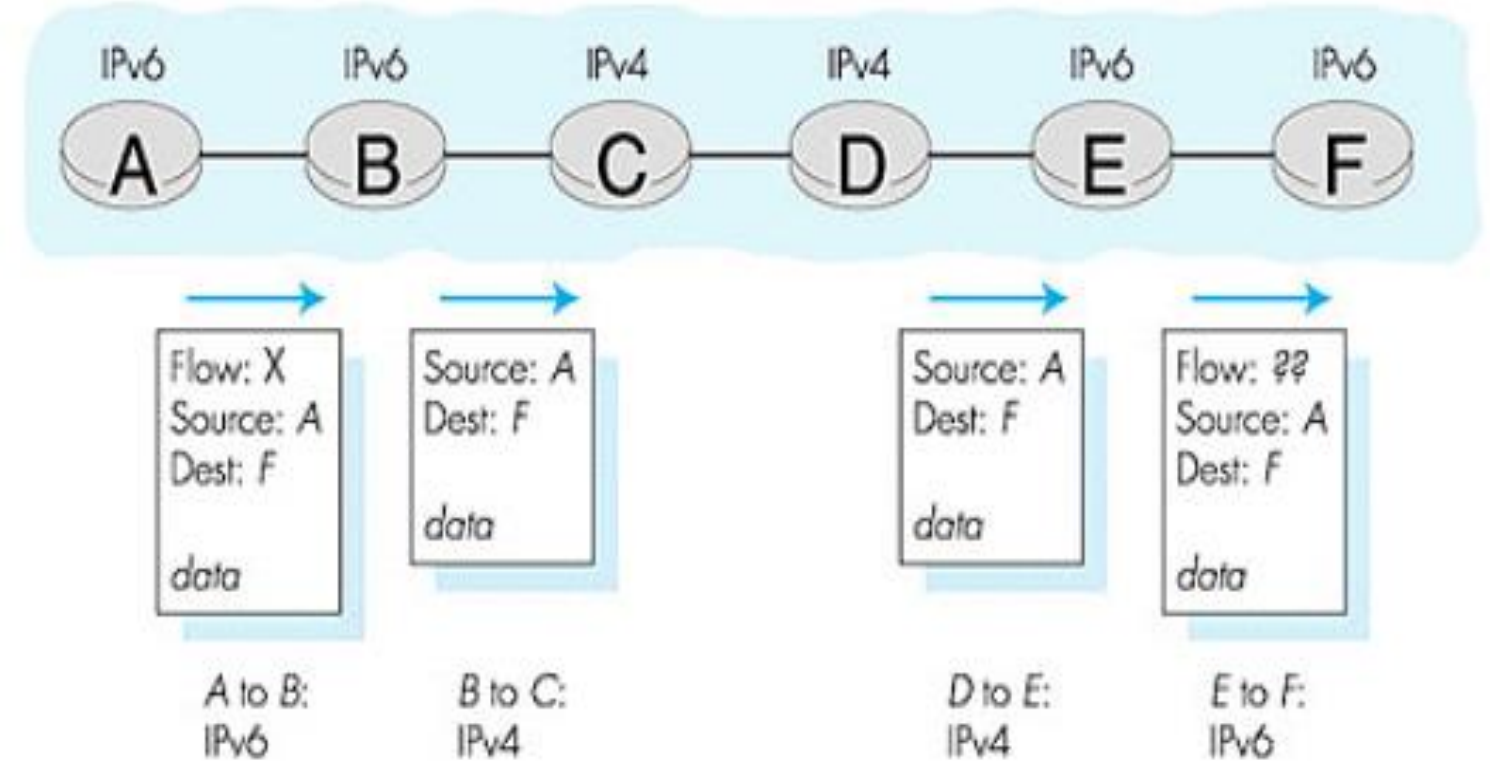


- Version: 0110
- Traffic Class
  - Prioritering innad i en datastrøm
- Flow Label
  - QoS
  - Ensure differentiation in service quality
  - Something vaguely defined
    - variable router support
- Payload length
  - Number of bytes with payload
- Next Header
  - Protocol at the level above the stack (UDP, TCP,…?)
  - Can / will also be Header extensions such as **IPSec**
- Hop Limit
  - Corresponds to TTL as practiced IPv4
- DATA

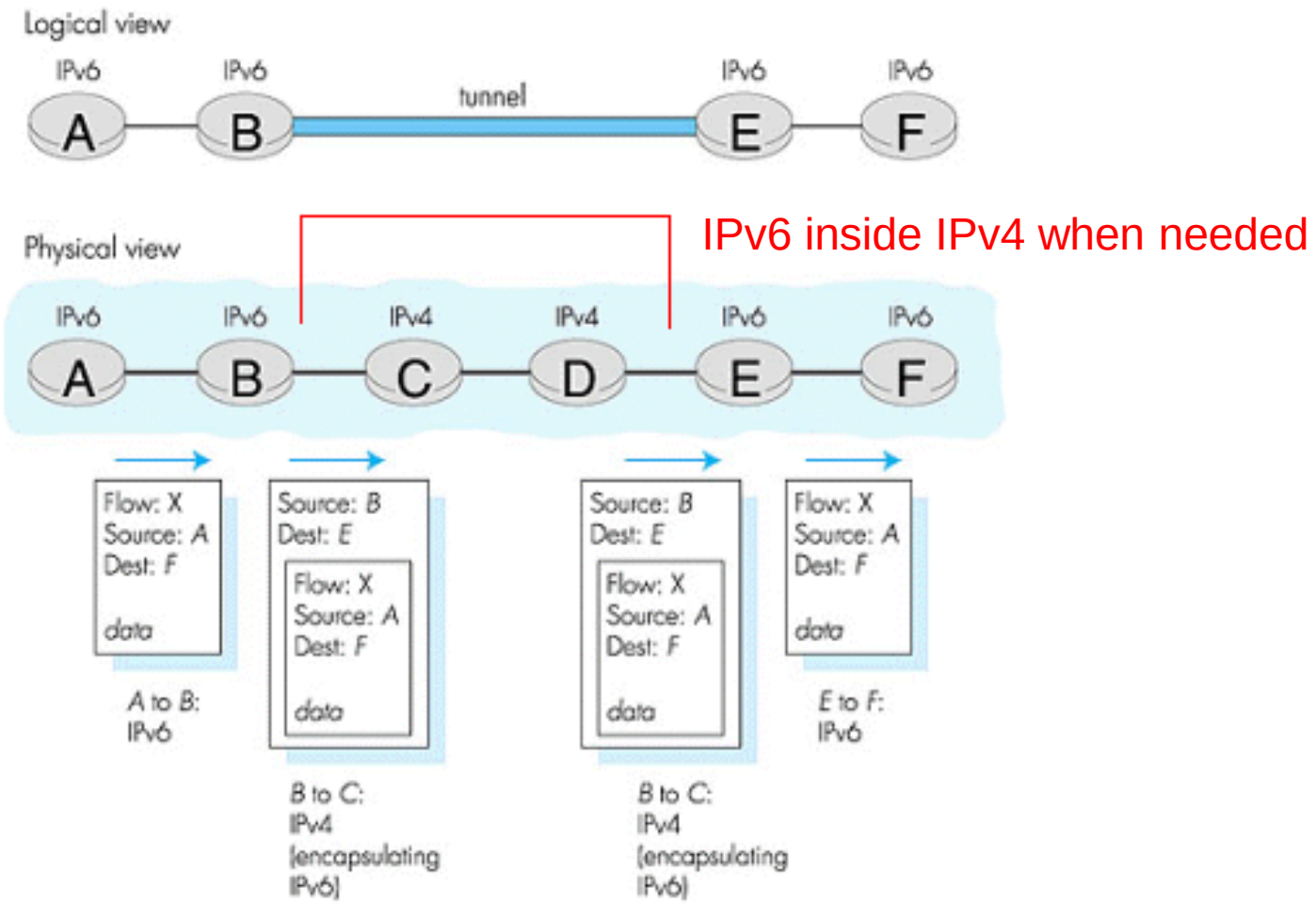
Transition from IPv4 to IPv6

- Impossible to upgrade all routers at once
- Transition period where both types of routers must be able to operate together
- Two (three) solutions are in use:
  - Dual stack
    - The router understands both protocols and translates between them
    - NAT64 / DNS64, (SLAAC)
  - Tunneling
    - An IPv4 router treats IPv6 datagram as data and leaves the datagram untouched (wrapped in IPv4)
    - 6in4, 6rd, Teredo, ISATAP

Dual stack


IPv6
IPv6
IPv4
IPv4
IPv6
IPv6
A
B
C
D
E
F
Flow: X
Source: A
Dest: F
data
Source: A
Dest: F
data
Source: A
Dest: F
data
Flow: ??
Source: A
Dest: F
data
A to B:
IPv6
B to C:
IPv4
D to E:
IPv4
E to F:
IPv6

Tunneling


Logical view
IPv6
IPv6
tunnel
IPv6
IPv6
A
B
E
F
Physical view
IPv6 inside IPv4 when needed
IPv6
IPv6
IPv4
IPv4
IPv6
IPv6
A
B
C
D
E
F
Flow: X
Source: A
Dest: F
data
A to B:
IPv6
Source: B
Dest: E
Flow: X
Source: A
Dest: F
data
B to C:
IPv4
(encapsulating
IPv6)
Source: B
Dest: E
Flow: X
Source: A
Dest: F
data
B to C:
IPv4
(encapsulating
IPv6)
Flow: X
Source: A
Dest: F
data
E to F:
IPv6

IPv6 notation

Høyskolen Kristiania

- 128 bits (16 bytes) are written hexadecimal in 8 groups of 2 bytes
- `IPCONFIG /all` gives for example:

```
Ethernet-kort eth0:

        Tilkoblingsspesifikt DNS-suffiks : oslo.nith.no
        Beskrivelse  . . . . . . . . . . : Intel(R) PRO/1000 PL Network Connecti
on
        Fysisk adresse . . . . . . . . . : 00-0E-7B-98-F8-A1
        DHCP aktivert. . . . . . . . . . : Ja
        Automatisk konfigurasjon aktivert: Ja
        IP-adresse . . . . . . . . . . . : 10.21.11.173
        Nettverksmaske . . . . . . . . . : 255.255.255.0
        IP-adresse . . . . . . . . . . . : fe80::20e:7bff:fe98:f8a1%5
        Standard gateway . . . . . . . . : 10.21.11.1
        DHCP-server. . . . . . . . . . . : 10.21.11.20
        DNS-servere. . . . . . . . . . . : 10.21.4.131
                                           10.21.21.101
```

- % 5 is Win adapter no. (not really part of the standard)
- fe80 :: = fe80: 0000: 0000: 0000 = web prefix
- :: is a minimum of four zeros, here 12 from the rest of the address
- 020e: 7bff: fe98: f8a1 is based on the MAC address

IPv6 Notation

- Uses CIDR to set the network prefix
- Some special addresses
  - :: /128 corresponds to 0.0.0.0 and is not used other than internally on the node
  - ::1 corresponds to 127.0.0.1 (localhost)
- 2000::/3 is Global Unicast
    - 2001 :: / 32 is awarded to ISPs that typically distribute / 48 and / 64 networks to customers
    - 2002 :: / 16 will be "old internet" (IANA)
      - 6to4 - Allocated to RIPE o.l.
  - 2001:**db8**::/**32** used in documentation
    - Note: initial zero is not written
  - **fe80**: /64 is a local link address. Corresponds in many ways to 169.254.X.X addresses (taken if IPv6-capable router is not available)
  - ::ffff:/96 used on IPv4 transition addresses, which get the format: ::ffff:192.0.2.114

TYPES OF ADDRESSES

Høyskolen Kristiania

- Broadcast will not continue from v4
  - Uses multicast instead.
- Three types of addresses:
  - Unicast
    - Single address
  - Anycast
    - First to take courage
    - Determined by the router
  - Multicast
    - To a predefined group (site)
    - Format: FF00 :: / 8 (Eg FF02 :: 1 all nodes on the same link)
    - Uses MLD (Multicast Listener Discovery) and ND (Neighbor Discovery) protocols.

IPv6

- ARP is going away
- DHCP is no longer needed
  - but DHCPv6 can be used to share DNS server and in managed networks (LAN / WAN)
  - Replaced by IPv6 message exchange to find gateway that assigns scope: global address based on MAC address
- New ICMPv6
  - Can "replace" DHCP setup of routers and the like.
  - Many new types of messages
- New versions of RIP, OSPF and other routing protocols.

SLAAC

- StateLess Address AutoConfiguration
  - Will automatically set up the IPv6 network for you
- Typically for use in local and home networks with IPv6 capable router and ISP offering IPv6
- Uses ICMPv6 to find router
  - is assigned an IPv6 address and other parameters by the router

Ipconfig / all (ifconfig -l) «tells»

Høyskolen Kristiania

```
Wireless LAN adapter WLANUSB:

   Connection-specific DNS Suffix  . : ad.nith.no
   Description . . . . . . . . . . . : D-Link DWA-140 Wireless N USB Adapter(rev.B3)
   Physical Address. . . . . . . . . : B8-A3-86-90-50-E8
   DHCP Enabled. . . . . . . . . . . : Yes
   Autoconfiguration Enabled . . . . : Yes
   Link-local IPv6 Address . . . . . : fe80::50e5:40ff:6794:1d5a%16(Preferred)
   IPv4 Address. . . . . . . . . . . : 10.21.26.56(Preferred)
   Subnet Mask . . . . . . . . . . . : 255.255.252.0
   Lease Obtained. . . . . . . . . . : 19. november 2013 14:29:14
   Lease Expires . . . . . . . . . . : 19. november 2013 22:29:14
   Default Gateway . . . . . . . . . : 10.21.24.1
   DHCP Server . . . . . . . . . . . : 1.1.1.1
   DHCPv6 IAID . . . . . . . . . . . : 548971398
   DHCPv6 Client DUID. . . . . . . . : 00-01-00-01-14-6A-F2-0B-D8-D3-85-77-A0-3F
   DNS Servers . . . . . . . . . . . : 158.36.131.10
   Primary WINS Server . . . . . . . : 158.36.131.10
   NetBIOS over Tcpip. . . . . . . . : Enabled

                           Tunnel adapter Teredo Tunneling Pseudo-Interface:

                              Media State . . . . . . . . . . . : Media disconnected
                              Connection-specific DNS Suffix  . :
                              Description . . . . . . . . . . . : Teredo Tunneling Pseudo-Interface
                              Physical Address. . . . . . . . . : 00-00-00-00-00-00-00-E0
                              DHCP Enabled. . . . . . . . . . . : No
                              Autoconfiguration Enabled . . . . : Yes
```

IPv6 in reality

- In Windows, OSX and Linux (kernel 2.6.x) IPv6 is part of the "standard package"
- In Vista, Win7 & 8 IPv6 is enabled by default, OSX (Free BSD) is also enabled by default since 10.4 (?)
- In reality, you have to rely on tunneling through IPv4 and a good deal of manual work
  - It will (maybe) not last (very ..) long…
  - uPnP is geared towards IPv6 e.g.

IPv6 problems per nowadays

Høyskolen Kristiania

- Tunneling by 6in4 turns out to be stopped by many firewalls because it enters a new version code (41) in the IPv4 header
- Only a few network administrators know much about IPv6..
- Many ISPs have not made any preparations at all…
- Only a few years ago DNS AAAA records became available in root servers.
- Many applications are unable to handle anything other than IPv4 sockets.
- +++