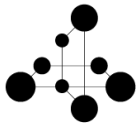


INFORMATION CODING
Linköping University

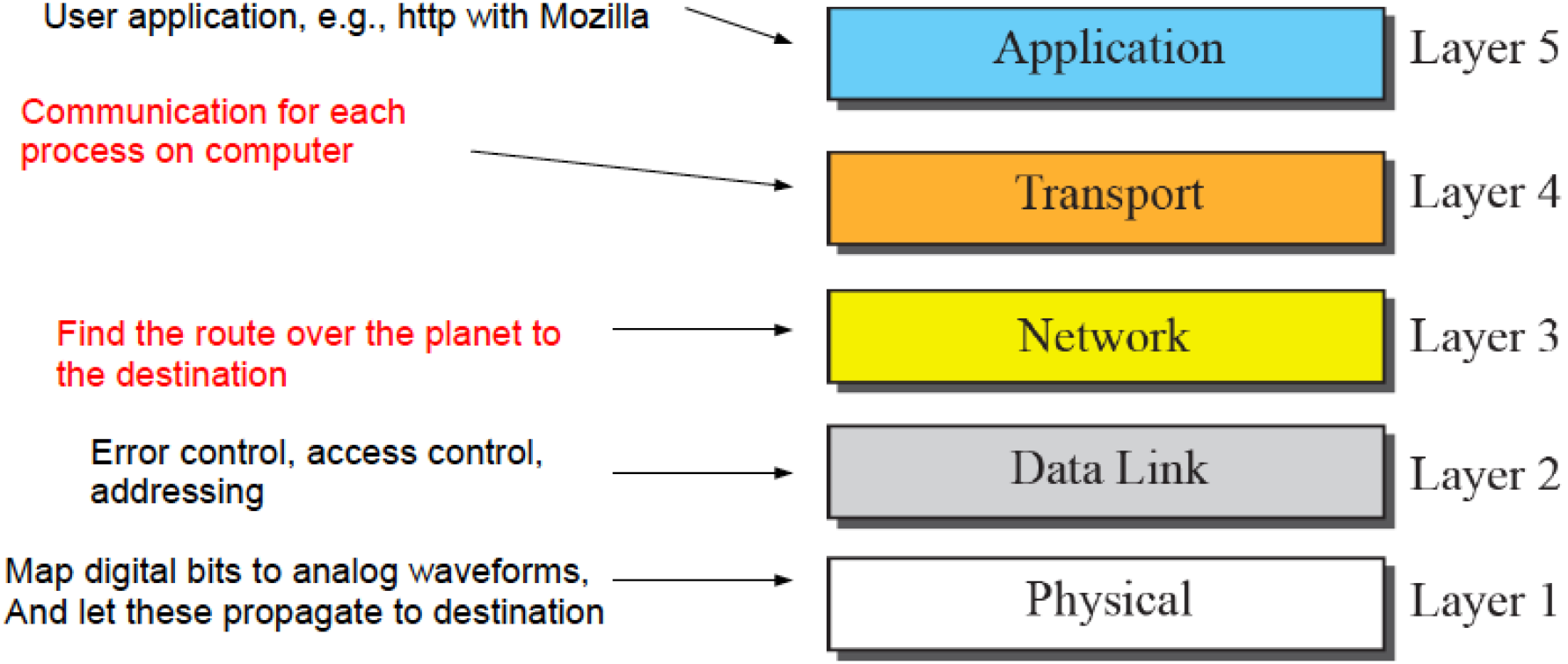
TSIN02 Internetworking

Lecture 3 – Transport and Application layers

Outline

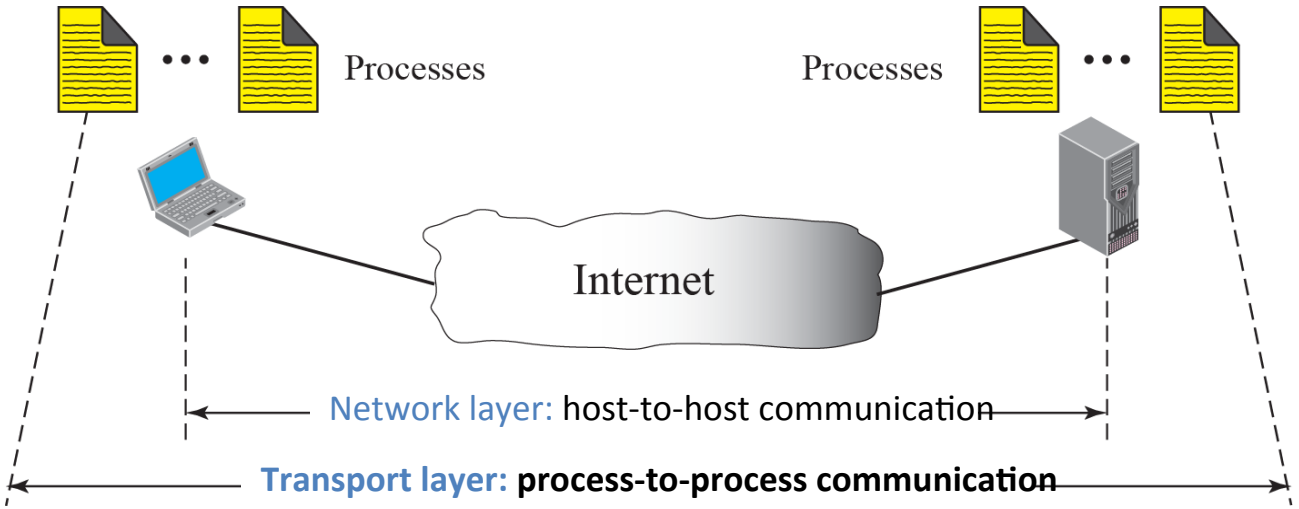
- Layer 4 – Transport layer, UDP and TCP
- Layer 5 – Application layer, DHCP
- Extensions of the basic Internet concept
  - Network Address Translation (NAT)
  - Multiprotocol Label Switching (MPLS)
  - Multicasting
- Overlay networks
  - Peer-to-peer (P2P) vs Client-Server
  - Software-defined networks (SDN)
- Special subnets
  - Mobile IP, LTE
  - Networks for data centers
  - Networks for Internet of things
- IPv6 – some more details

IETF TCP/IP protocol suite



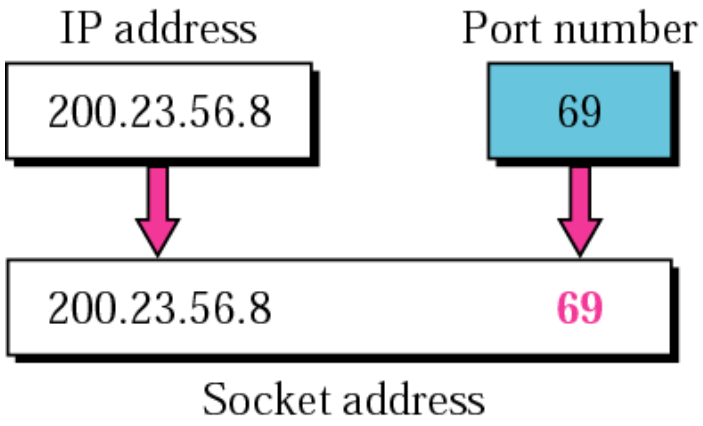
Layer 4: Transport Layer

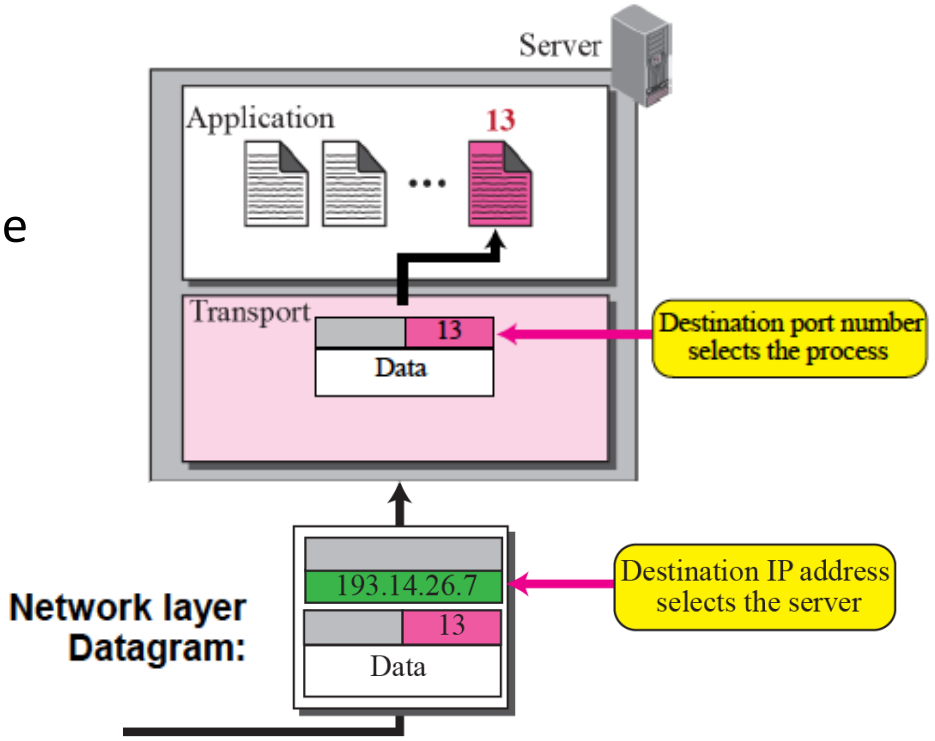
- Layer 4 – No longer part of the network itself, rather "lowest level application" running on host computers.
- Process-to-process communication, i.e. a computer program that is being executed on a computer communicates over the network with a computer program that is being executed on another computer. This is done by adding a *port* number in the IP packet.
- Offers "transport services" to higher level applications
  - UDP: Packet transmission
  - TCP: A connection-oriented service



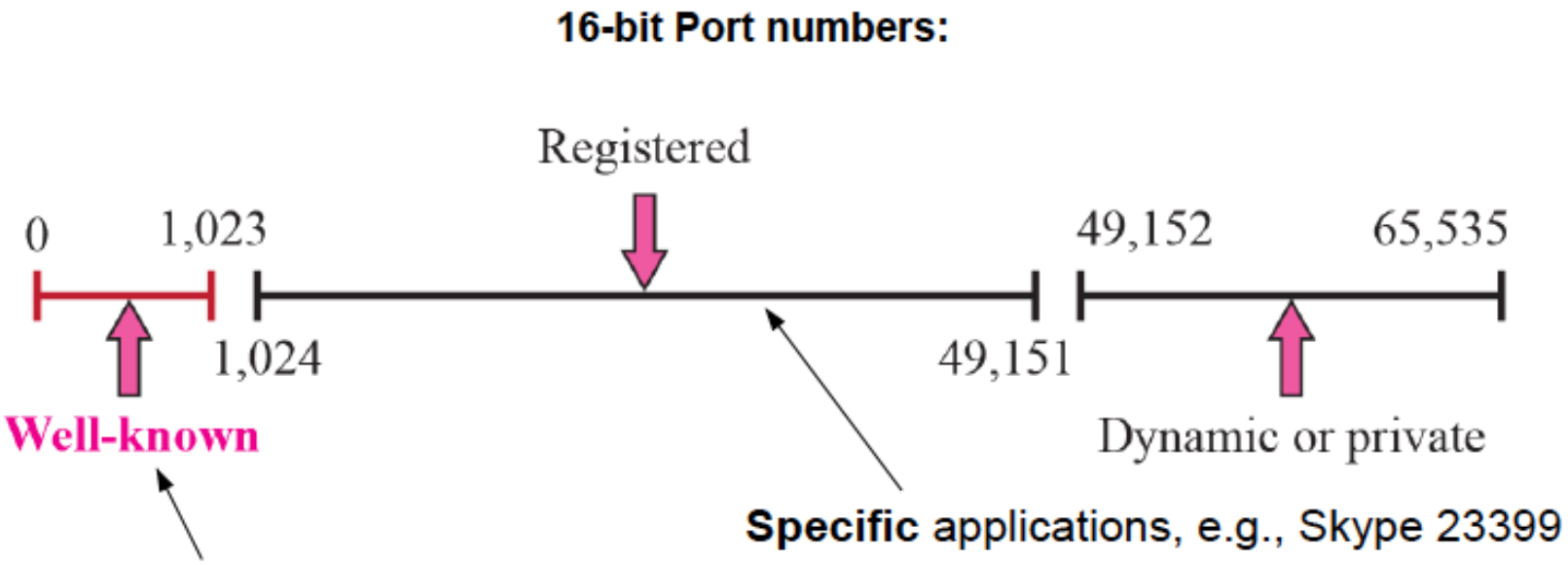
Port addresses

- **Port** number: 16-bit unsigned integer, thus ranging from 0 to 65536
- An IP address, port number pair is called a **socket address**.
- Source and destination processes use a *socket interface* to communicate with each other.





Port Addresses, cont’



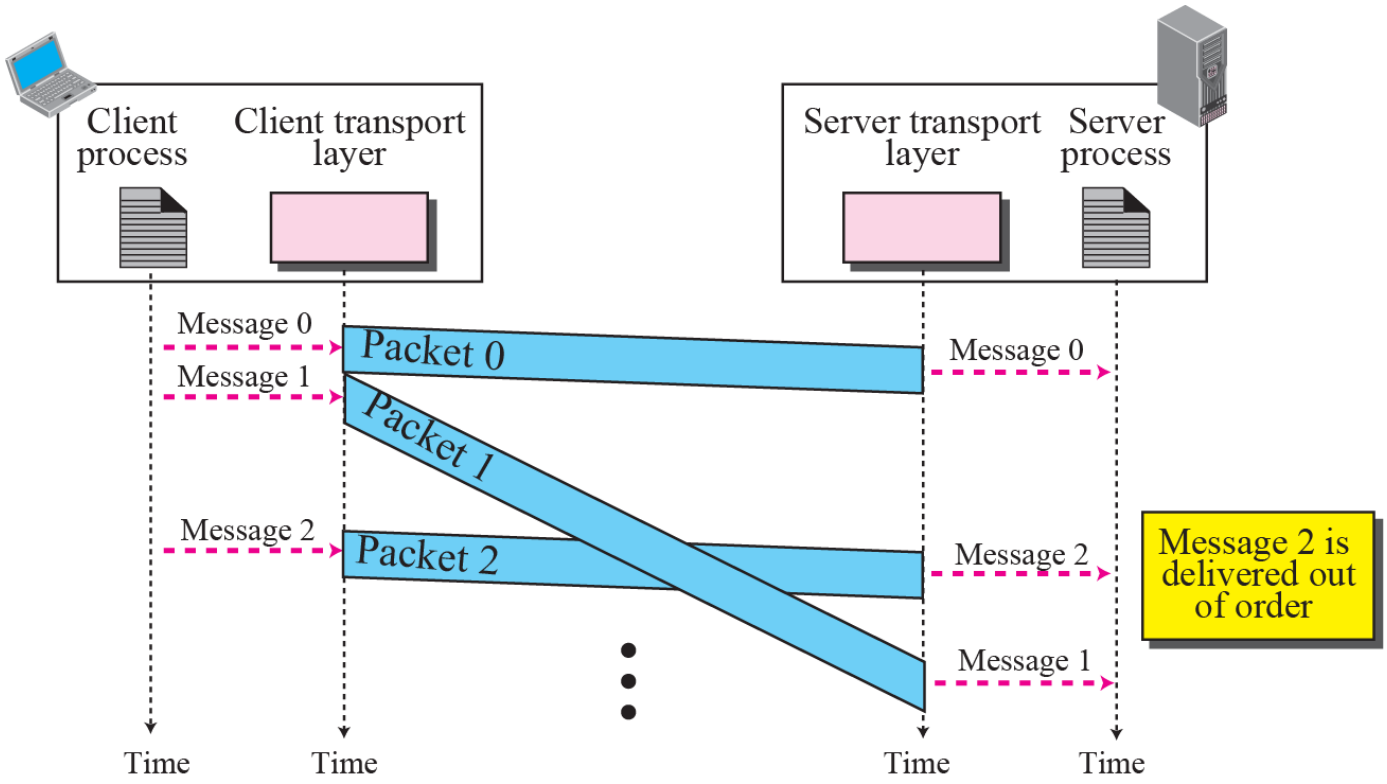
Secure shell (ssh): 22
web (http): 80

User Datagram Protocol (UDP) and Transmission Control Protocol (TCP)

- These are the two main transport protocols
- UDP is connectionless
- TCP is connection-oriented
- TCP is less powerful than a full circuit-switched network layer solution. Delay can not be fully controlled.
- TCP is cheaper/simpler than a full circuit-switched solution in the sense that the routers do not need to keep track of circuits.
- Services not available: delay guarantees, bandwidth guarantees
- TCP transports “segments”, UDP transports “datagrams”

UDP

Connectionless Service:
There is no dependency between transport layer packets at the receiver, e.g.:
- datagrams may disappear
- datagrams may be delayed
- datagrams can arrive out of order



UDP usage

Applications:

- Processes that provide internal flow, congestion, and error control mechanisms =>flexibility.
- Real-time (e.g. video or audio) applications that can not tolerate extended delays caused by ARQ retransmission.
- If only short messages are needed, like with DHCP, we may not need TCP's connection setup.
- TCP longer headers may cause more congestion on the network

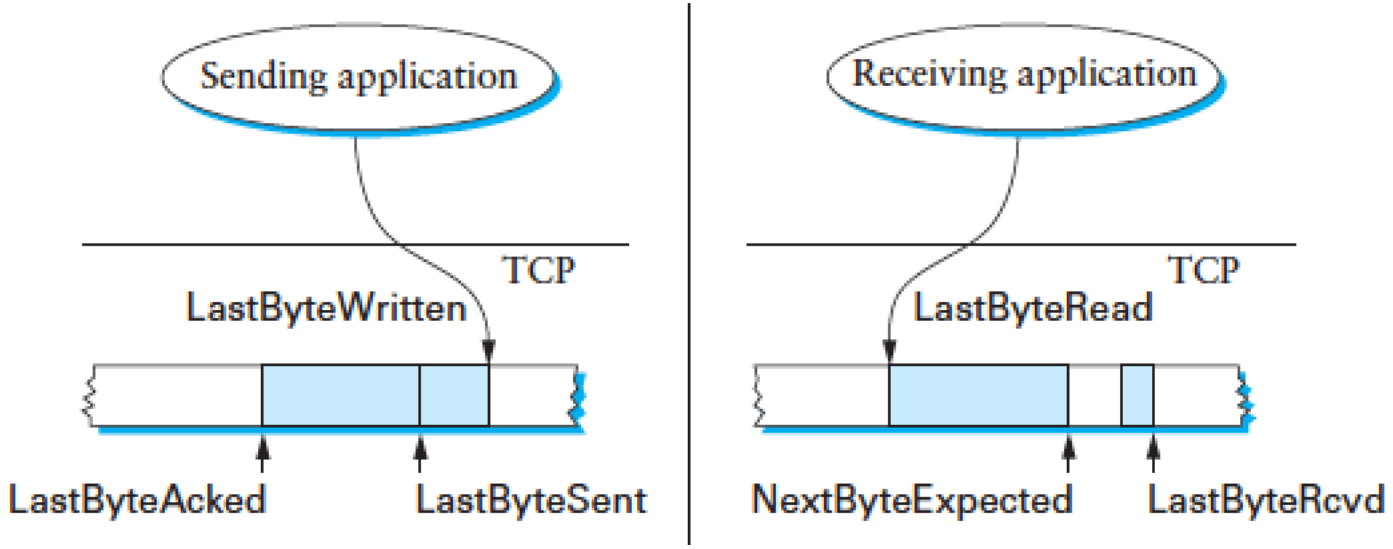
TCP

- TCP is a connection-oriented full duplex, in-order byte stream service.
- TCP is reliable: all TCP segments (packets) are delivered to the receiving application layer, and in order. This is done by means of a byte sequence counter and error control (ARQ).
- TCP has flow control: the sending process does not send more than the receiving process can consume.
- TCP offers congestion control: the network is protected from having higher load than its capacity.

TCP window and ACK mechanism

A buffer ("sliding window") is used both on transmitting side and receiving side to:

- ensure that bytes are delivered to the application in correct order
- ensure that transmitter is not sending more bytes than receiver can handle (flow control)
- allow common ACK for several packets (delayed ACK)



TCP connection-oriented packet delivery

Establish connection (three-way handshake)

- Transmitter: send SYN packet and random sequence number
- Receiver: reply with SYN-ACK packet, incremented sequence number and its own random sequence number
- Transmitter: reply with ACK packet, incremented receiver sequence number

Data transfer (simultaneous send/receive “full duplex”)

- Several transmitted packets can be verified by a single received ACK
- Sequence numbers are updated on both sides for each packet
- A missing ACK leads to retransmission

Terminate connection (three-way handshake)

- Transmitter: send FIN packet
- Receiver: reply with FIN-ACK packet
- Transmitter: reply with ACK packet

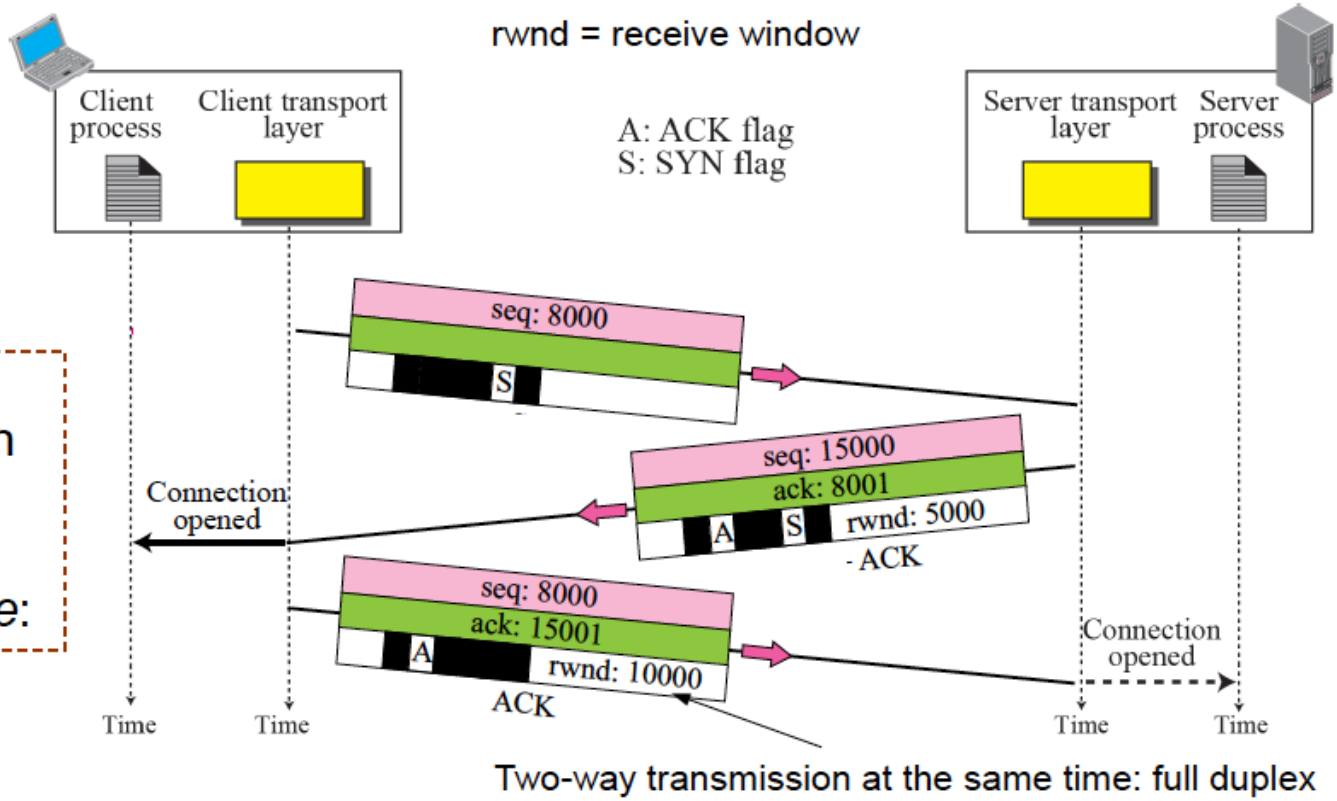
Illustrated graphically in the following 3 slides

TCP connection-oriented packet delivery: connection setup

Three phases:

1. Establish connection **2.** Data transfer **3.** Terminate connection

1. Establish connection using *three-way handshake*:

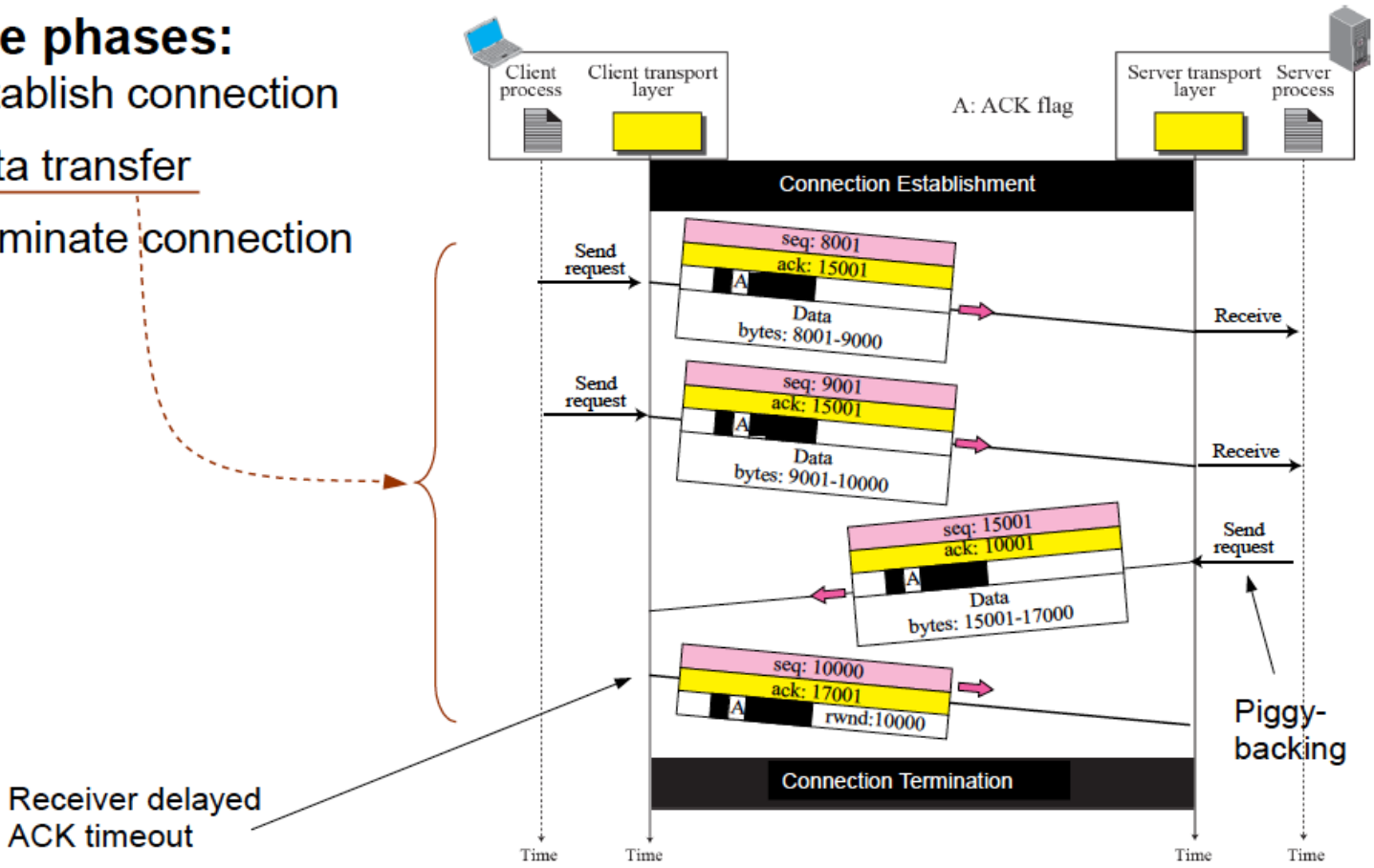


Two-way transmission at the same time: full duplex

TCP connection-oriented packet delivery: data transfer

Three phases:

1. Establish connection
2. Data transfer
3. Terminate connection

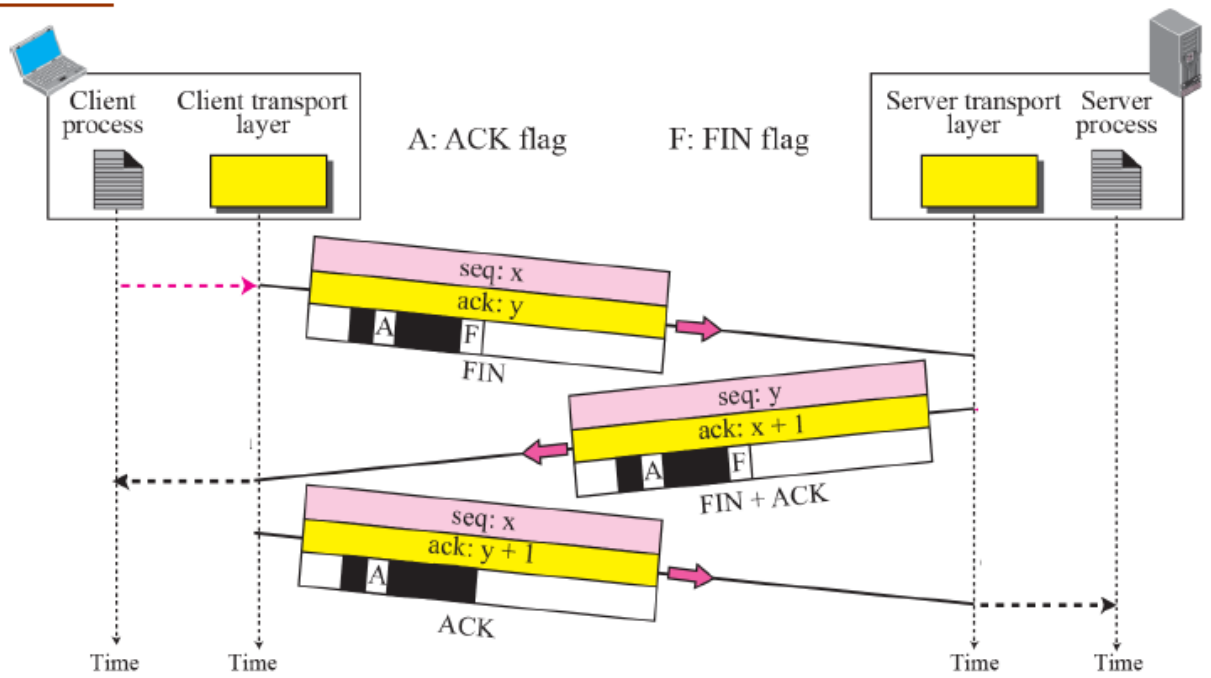


TCP connection-oriented packet delivery: terminate connection

Three phases:

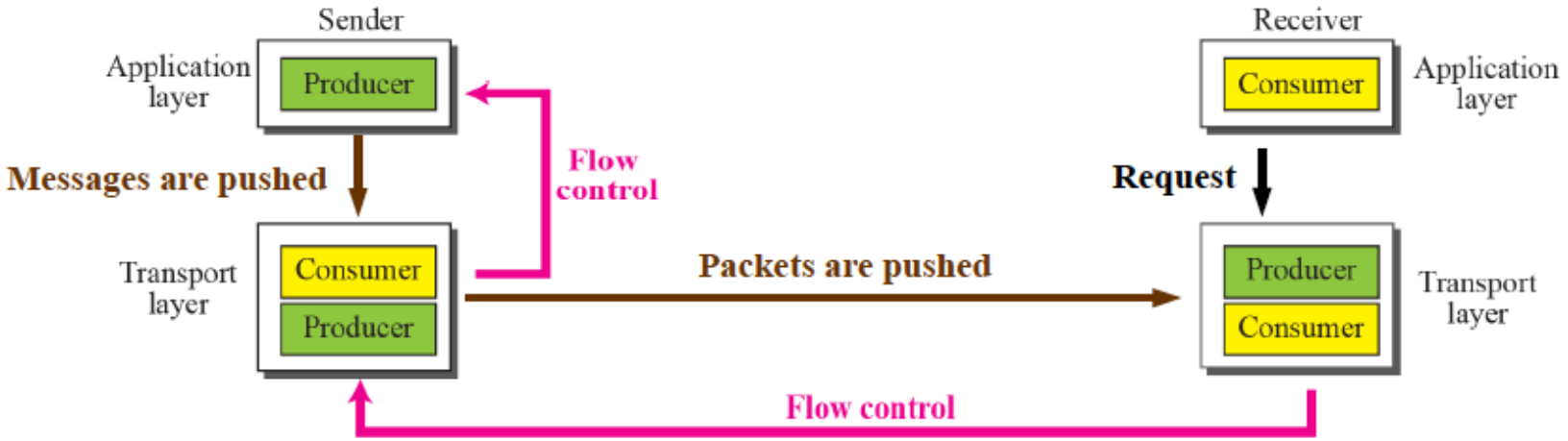
1. Establish connection
2. Data transfer
3. Terminate connection

3. Terminate connection using *three-way handshake*:



TCP flow control

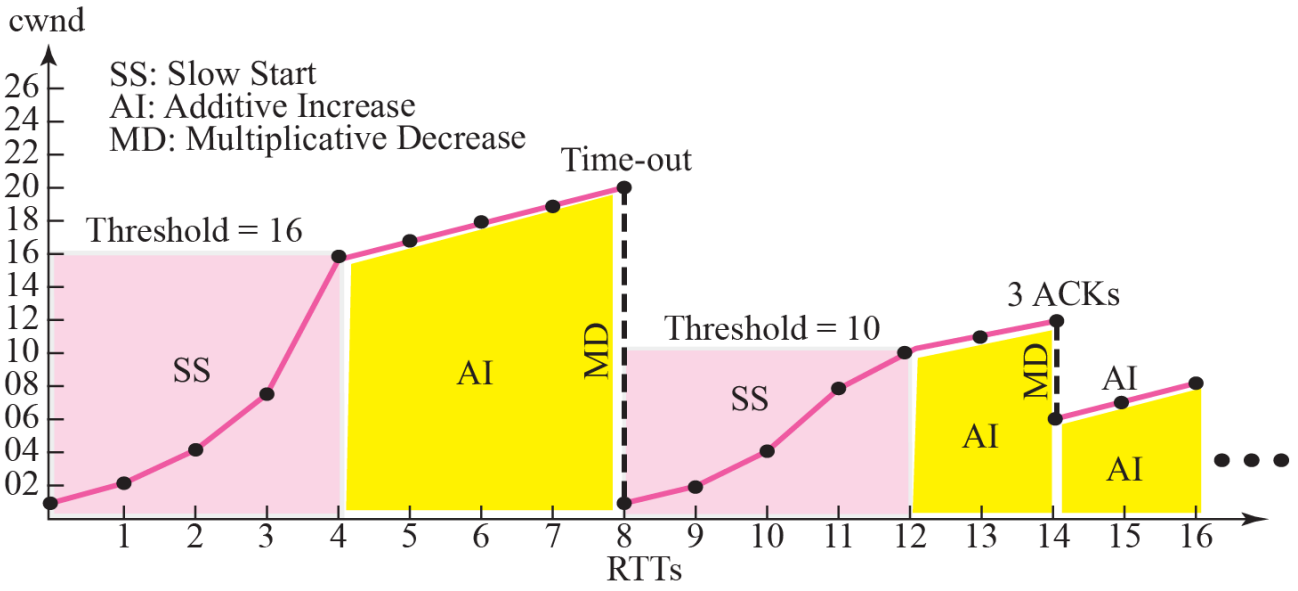
Flow control: messages from receiver to sender telling how fast to transmit



Congestion control

ACKs decide how fast to transmit. More precisely, the network's congestion is measured through the ACKs that arrive/not arrive at sender.

TCP increases its transmit rate (exponentially and linearly) until 3 duplicate ACKs are received. Then it cuts the rate in half. If instead timed-out, the rate is back to start value.



TCP – concluding remark

- TCP is a congestion-aware protocol
- Other protocols which use same, or similar mechanisms to avoid congestion are called *TCP-friendly* protocols.

Other transport protocols

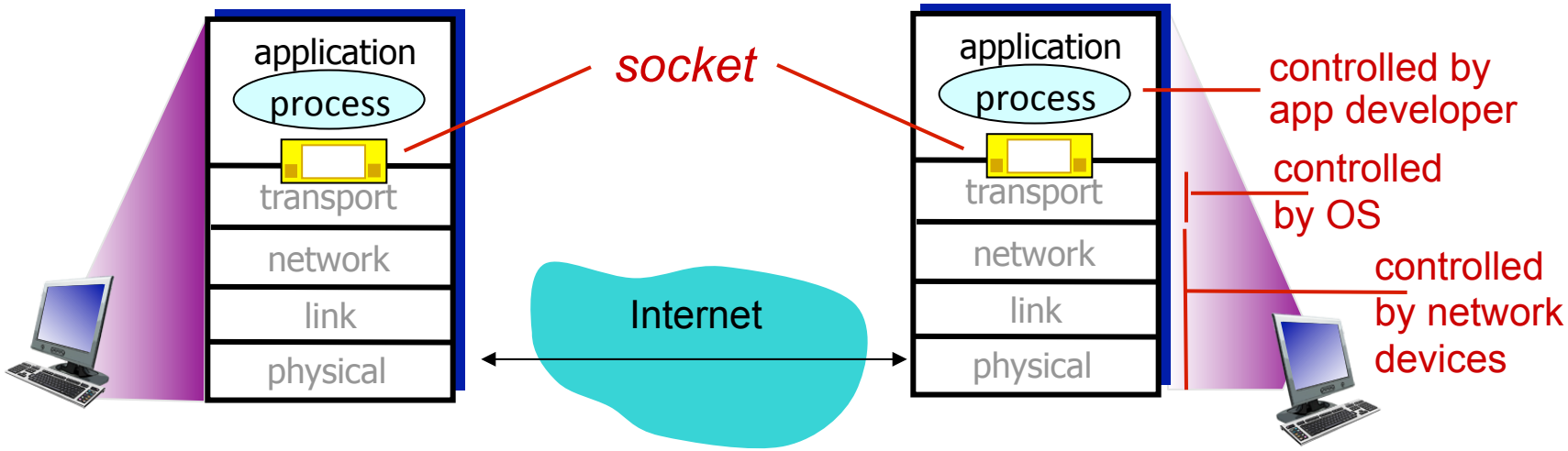
- SCTP - Stream Control Transmission Protocol (TCP-friendly)

- RTP – Realtime Transport Protocol (uses UDP, often viewed as an application layer protocol)

We will return to them when discussing real-time media communication.

Layer 5: Application layer

- Layer 5 protocols are application-specific. Builds on transport layer protocols (mainly UPD, TCP).
- Open protocols: HTTP, FTP, SMTP, DNS, DHCP, SIP, RTP…
- Most of these support the *Client-Server* model of communication
- Proprietary protocol: Skype (uses the *peer-to-peer* model of communication)
- Applications use *sockets calls* to reach the transport layer.



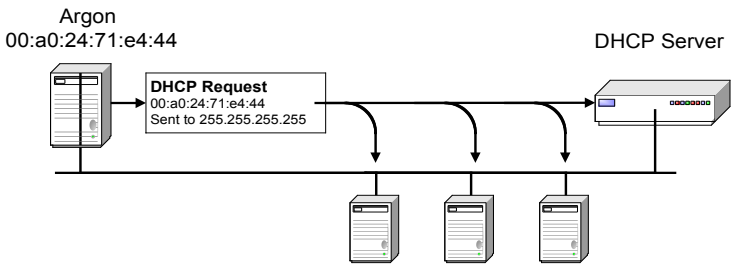
Application protocols that aid network functionality

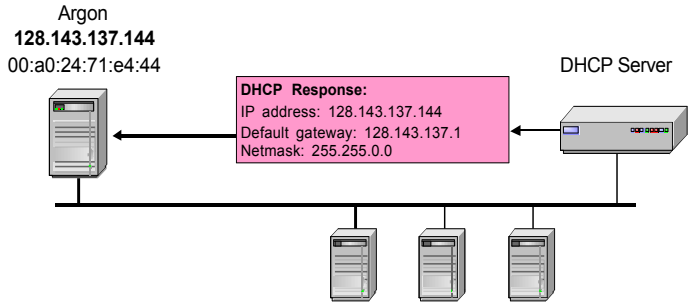
- DHCP – Dynamic Host Configuration Protocol
- DNS – Domain Name Server

DHCP – How to get a temporary IP address

- DHCP assigns an IP address given a physical address, thus performs the reverse of ARP (“RARP”)
- IP addresses are assigned on-demand (saves addresses)
- Avoids manual IP configuration
- Supports mobility of laptops
- Drawback: DNS* cannot be used for DHCP configured hosts

Simplified description:





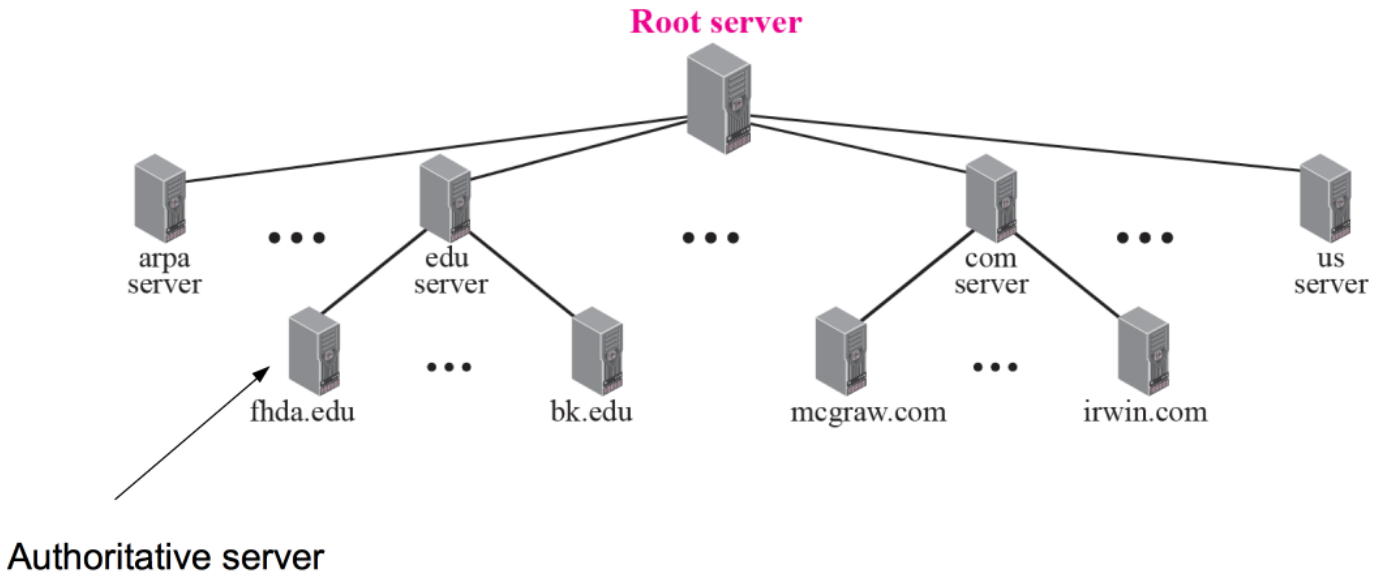
*Soon to be explained

DHCP in more detail

- A temporary host connects to Internet through a local area network. It needs to get an IP address. Using a *DHCP* server it can request for an IP address without manual interaction.
- The host does not know the IP address of the DHCP server that gives out IP addresses. The host broadcasts a **DHCP discover packet** on the local network using link layer addressing. The local network either has a DHCP server or a *DHCP relay agent* that connects to an external DHCP server.
- A DHCP server (or agent) connected to the local network answers with a link layer unicast packet (unicast is possible because the server knows the physical layer address of the host). The answer includes a **DHCP offer packet**, which contains an IP address offer.
- The host sends a DHCP request packet to the DHCP server (unicast is enough this time, since the host now knows IP and MAC addresses of the server). The server acknowledges with a DHCP ACK packet, and the two devices have agreed on the IP address.
- The host also learns the **subnet prefix**, **Domain name system (DNS) server IP, time of lease**, and **default gateway IP**, i.e., first-hop router IP, where to forward packets to Internet. The host now also knows when to use ARP, and default gateway forwarding, when communicating with other computers.

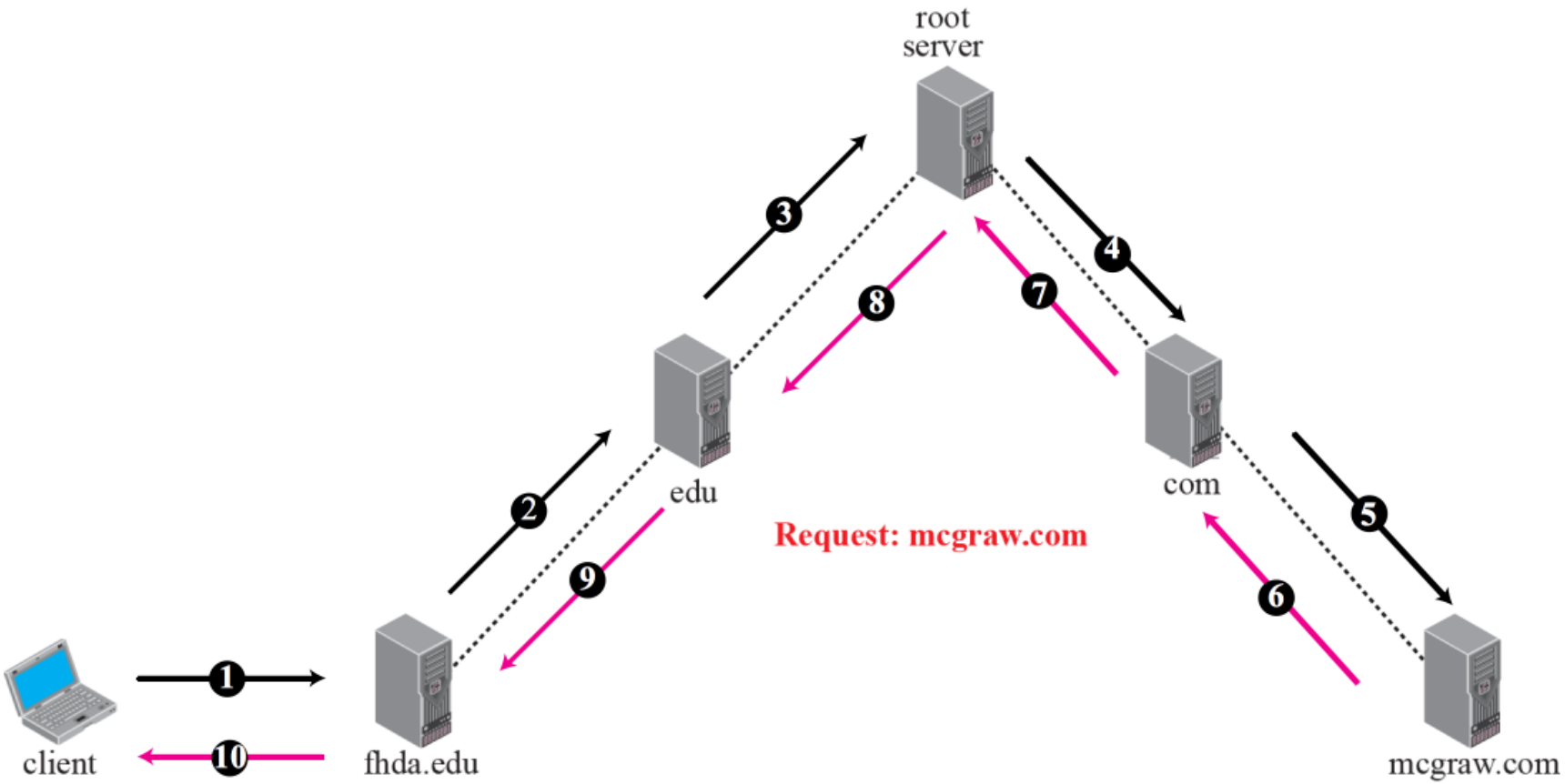
DNS – The Internet's "Yellow Pages"

- Users can't remember IP addresses – there is a need to map symbolic names (www.liu.se) -> IP address
- This is implemented through distributed name servers organized in a tree structure. Top level = redundant set of root servers

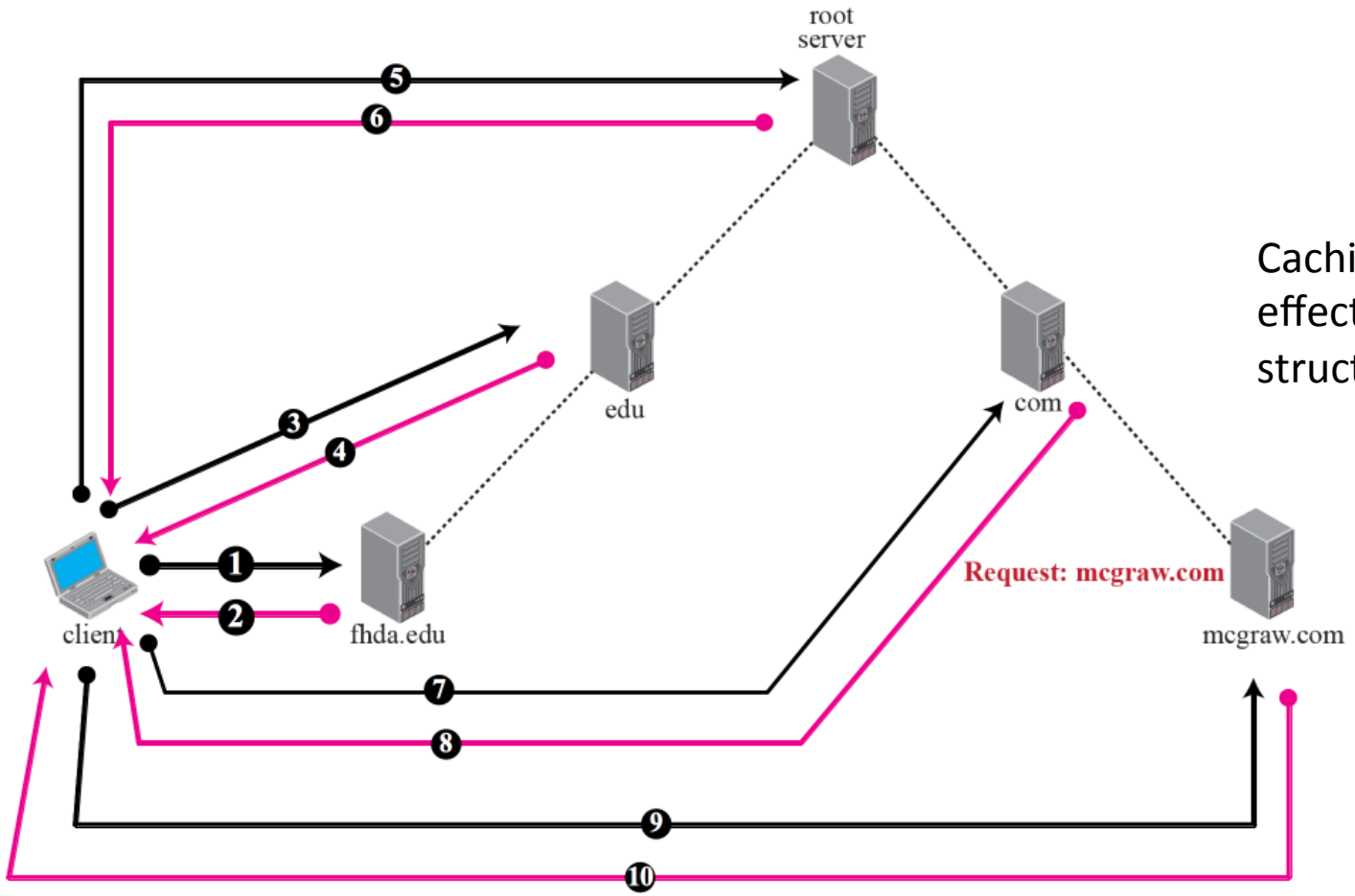


Servers organized as addresses

Recursive resolution



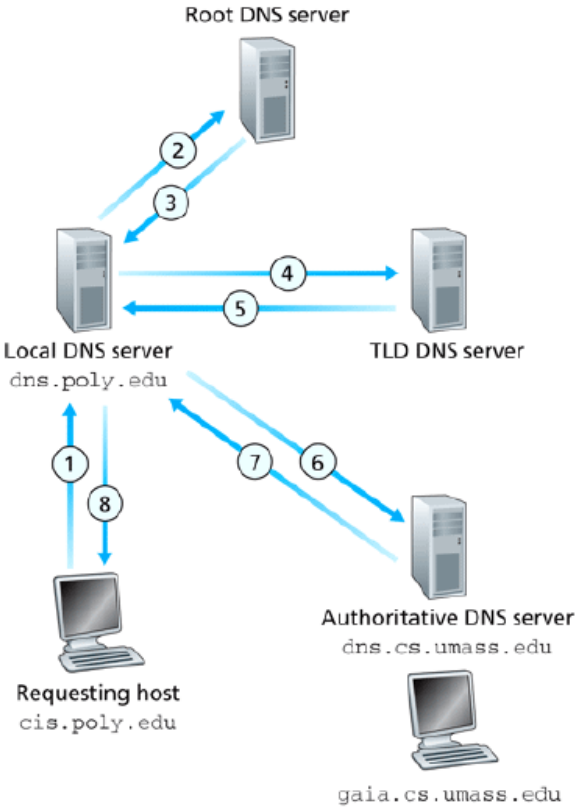
Iterative resolution



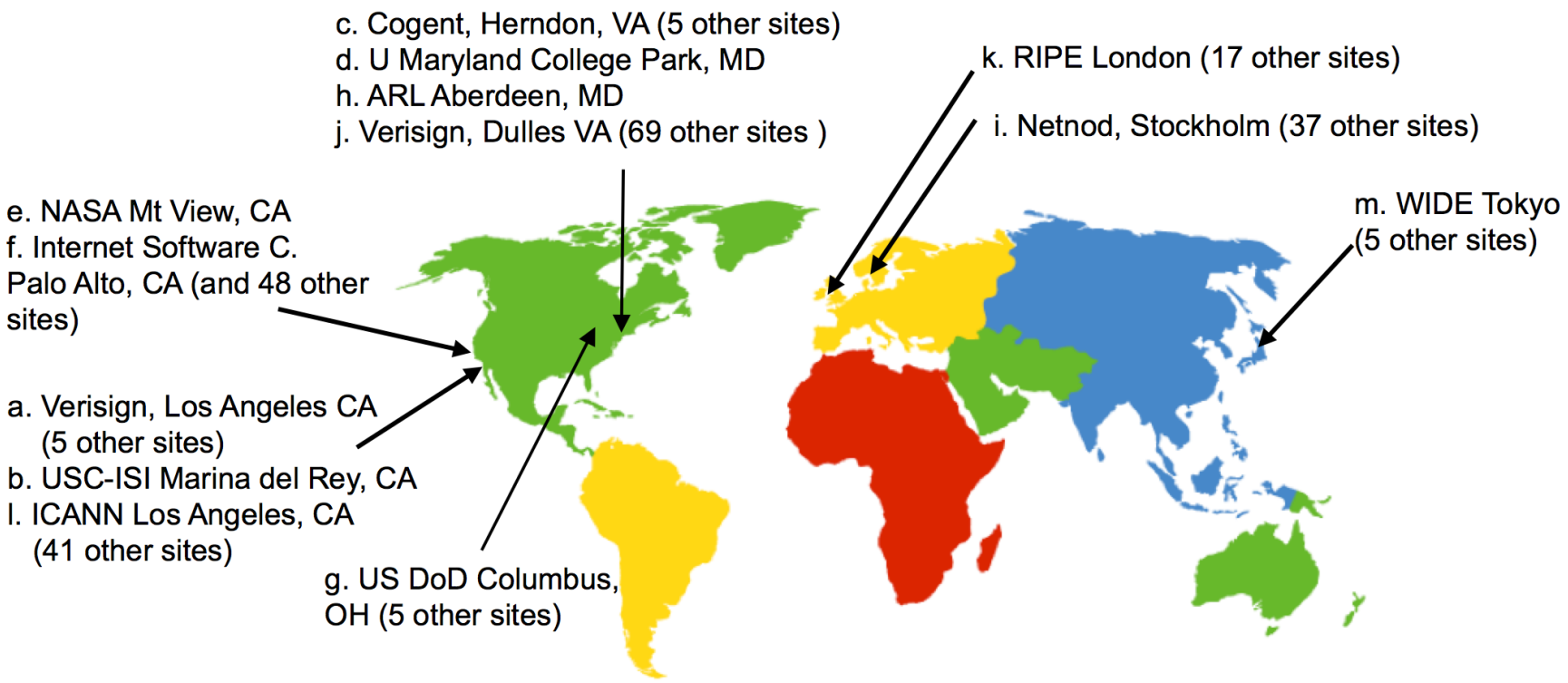
Caching is highly effective in tree structures!

Combined recursive/iterative resolution

- Applications make recursive queries to local DNS server (1)
- Local server queries remote servers non-recursively (2, 4, 6)
  - Aggressively caches result
  - E.g., only contact root on first query ending `.umass.edu`



Root servers



13 root name "servers" worldwide

End of Internet basic introduction!

Extensions of the basic Internet concept

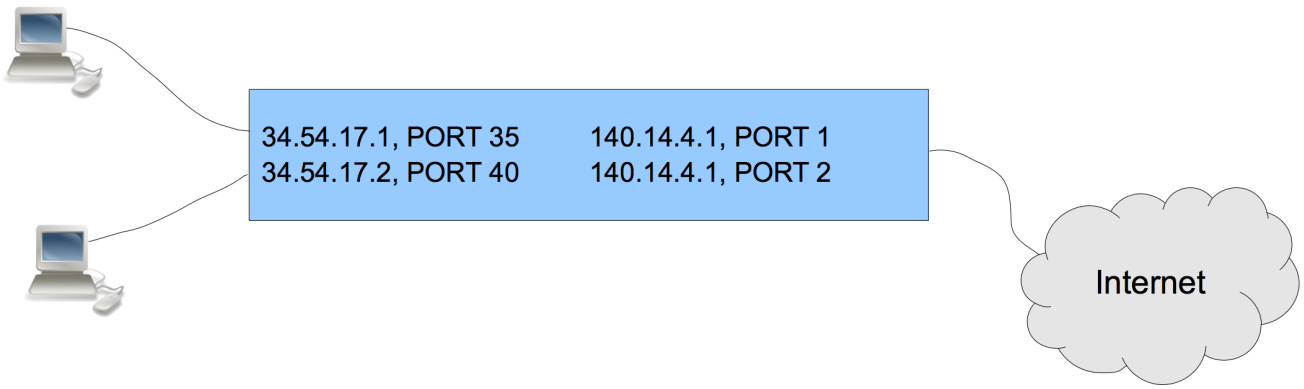
- Network Address Translation (NAT)
- Multiprotocol Label Switching (MPLS)
- Multicast

NAT – Network Address Translation

- There are only 2^{32} IP addresses. We were seeing the end of address space already in mid 1990:s.
- Solution: consider a group of hosts (e.g. on a local network) as processes rather than computers…thus utilize the 16 bit port number as an extension of the IP address!
- The NAT device dynamically assigns mapping between the public IP address and local IP address.
- Additional benefit: improves security as it hides the local computers from the public network (*firewall* functionality).

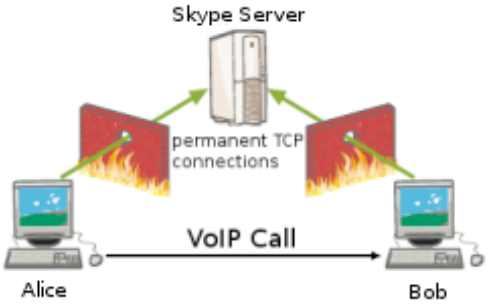
How NAT works

- Maps between the public and local (IP, port) pairs
- Requires knowledge of transport packet format
- NAT device (often part of local router) needs to remember the mapping to allow for incoming response packets
- Breaks end-to-end (node does not know its external IP address

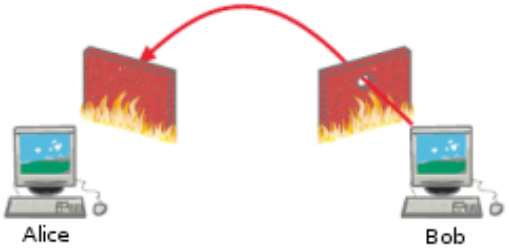


A NAT problem – and the Skype solution

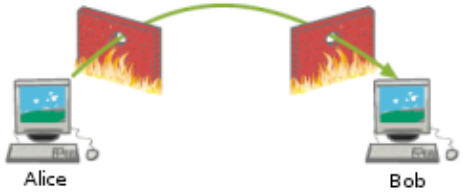
- Outside server cannot initiate communication (such as a TCP request) with an internal host. All traffic has to be initiated from the internal host.
- Skype uses server(s) with public address to “punch holes” in NAT firewalls.
  - All users constantly keep an open connection to a Skype server:



1. Alice tells server that she wants to connect with Bob. Server forwards this request to Bob and includes Alices IP/port address. Alice also gets Bobs IP address.



2. Bob sends a packet to Alice. A hole is punched in Bobs firewall. Packet is disregarded by Alices firewall.



3. Alice can now contact Bob directly using the hole in his firewall. (If Bobs firewall changes port number, a port scan is needed.)

http://churchofbsd.blogspot.se/2011/06/how-skype-gets-round-firewalls.html

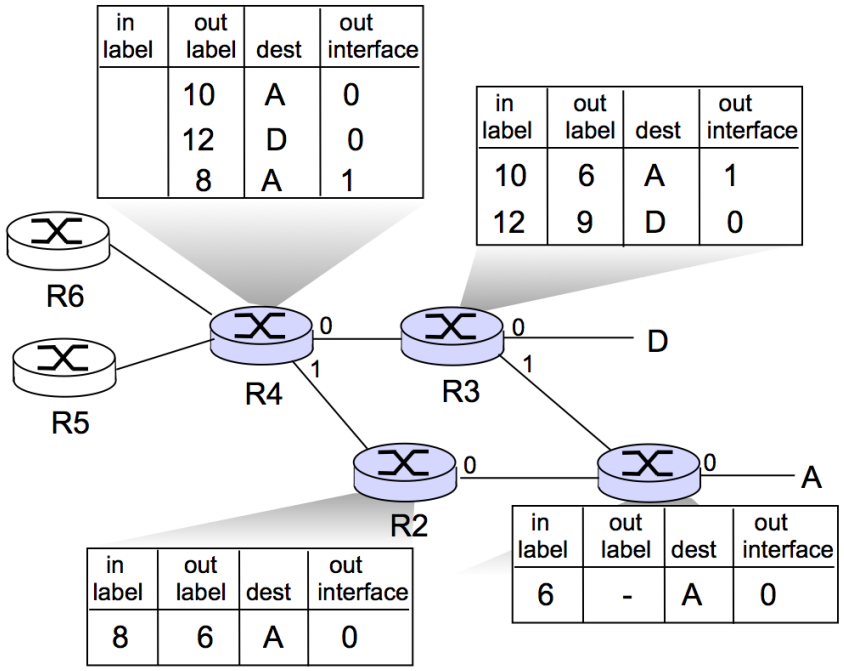
Multi-protocol Label Switching(MPLS)

- MPLS – "Layer 2.5" networking protocol
- Idea is to provide "tunnels" for layer 2 packets through the network by using fix-length address tags (labels)
- Originally intended to reduce IP routing lookups
- Today, the benefits are
  - the ability to control where and how traffic is routed
  - to manage capacity and prevent congestion (improving QoS)
  - to deliver multi-protocol traffic over the same network
  - Improving network resilience through backup route

MPLS – how it works

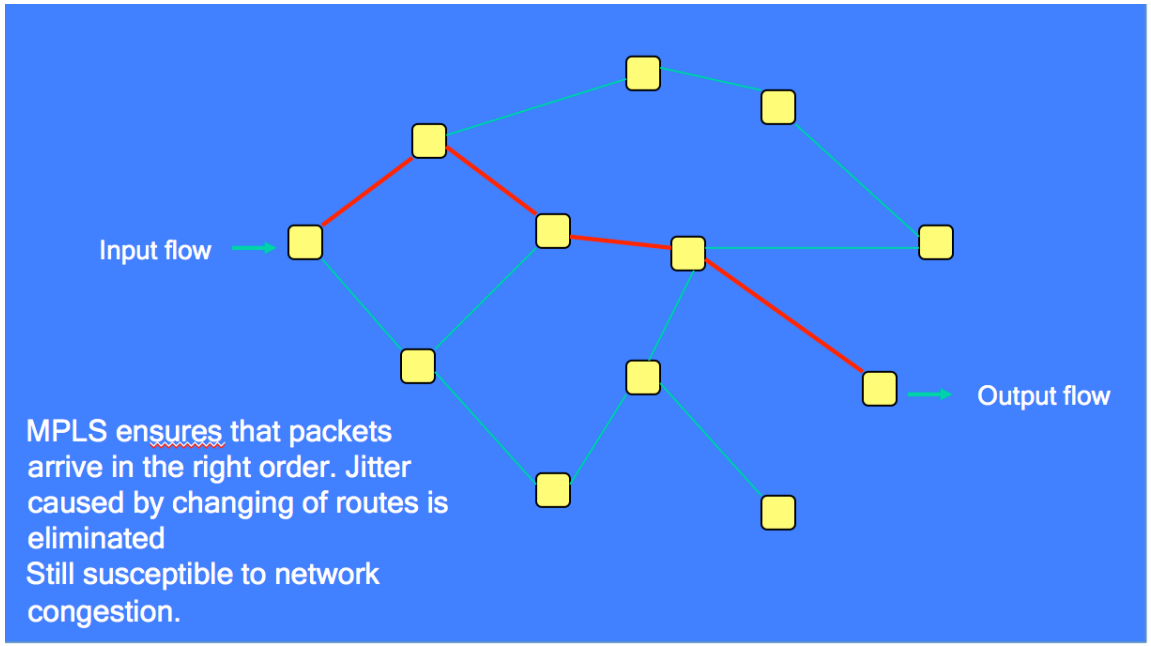
The MPLS packet:

- MPLS-aware routers keep a table that maps the label directly to an output port
- Label is replaced with new label for the next hop



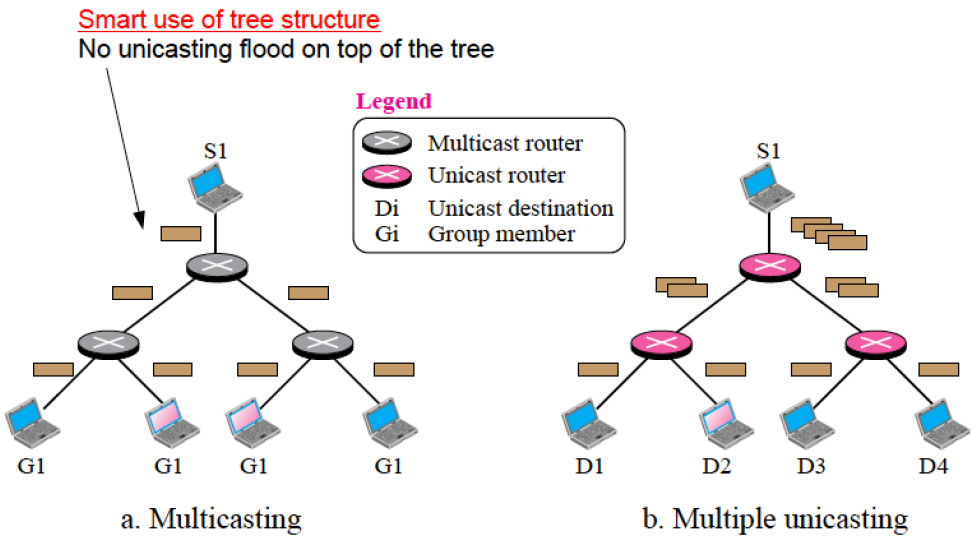
MPLS – how to set the labels

- First router that sees the MPLS packet checks various ways to reach the destination, taking into account not only number of hops but also congestion and available bandwidth on different links.
- All routers along the preferred path are then assigning labels to be used in the transfer of packets belonging to this MPLS session.



Multicast

- IP Multicast allows a stream of (UDP) packets to be sent to many users concurrently.
- Relevant routers copy and transmit the stream in a tree-like structure.
- Users join a “multicast group” by registering with a router that has the stream.
- Source host only needs to transmit one stream instead of one for each user.
- Multicast is commonly used in closed IP networks, however not enabled in the public network due to its complexity and security issues.



a. Multicasting

b. Multiple unicasting

Overlay networks

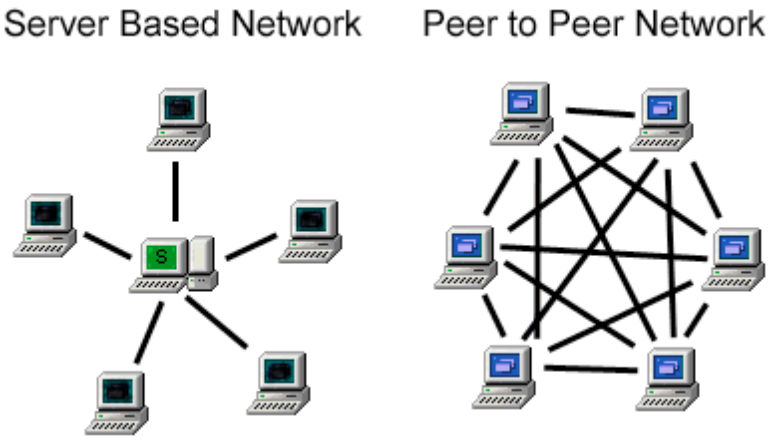
The DNS and multicast systems are example of overlay networks (a network implemented on top of the physical network)

Other overlay networks are:

- Peer-to-peer (P2P)
- Software-defined networks (SDN)

Peer-to-peer (P2P)

- In a peer-to-peer network the hosts act both as clients and as servers. This offloads central servers (e.g. in applications like Skype, Spotify)
- Not only networking resources can be shared but also processing power and storage (e.g. file sharing)
- P2P needs to be implemented at the application layer, no support exists at lower layers.



P2P applications

- File transfer
  - BitTorrent
  - Pando
- Content distribution
  - BBC iPlayer
  - Freecast
  - Tribler
  - blinkx
- Privacy protection
  - Tor
- Real time communication
  - Skype
  - Spotify
- Web search engines
  - YaCy
  - FAROO

Software-defined networks (SDN)

- SDN is a concept in which the network elements are programmable on a fairly low level. Thus, routing and other network functions can be altered as part of the users' need for new functionality.
- Typically, the control plane and the data plane are separated to avoid unnecessary complications.
- The first network designed according to this principle was SOFTNET , a fully user programmable network designed and operated by Swedish radio amateurs in the early 1980:s*.

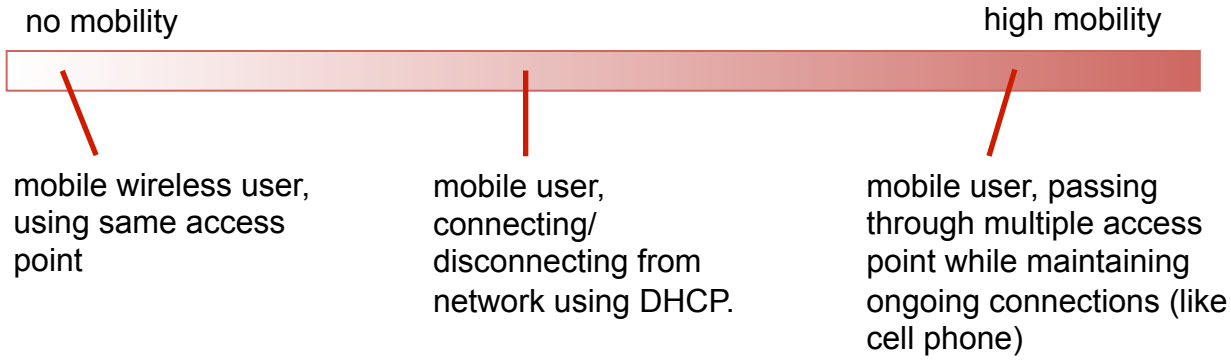
*J.Zander, R. Forchheimer, The SOFTNET project: A retrospect EUROCON 88, Stockholm 1988, 10.1109/EURCON.1988.11172

Special subnets

- Mobile IP, LTE
- Networks for *Internet-of-things*
- Networks for data centers (will be discussed in lecture 7)

Mobile IP, LTE

- We have learned that finding a host on Internet is relatively easy since the subnets are hierarchically ordered. Routing and forwarding group hosts into subnets, which save calculations and storage.
- However, when hosts are mobile, keeping track of them gets difficult if they are to keep their IP addresses (routing and forwarding => exceptions). This does not scale! (If one host/IP moves from a subnet to another, should all routers in the Internet have to keep track of this?)

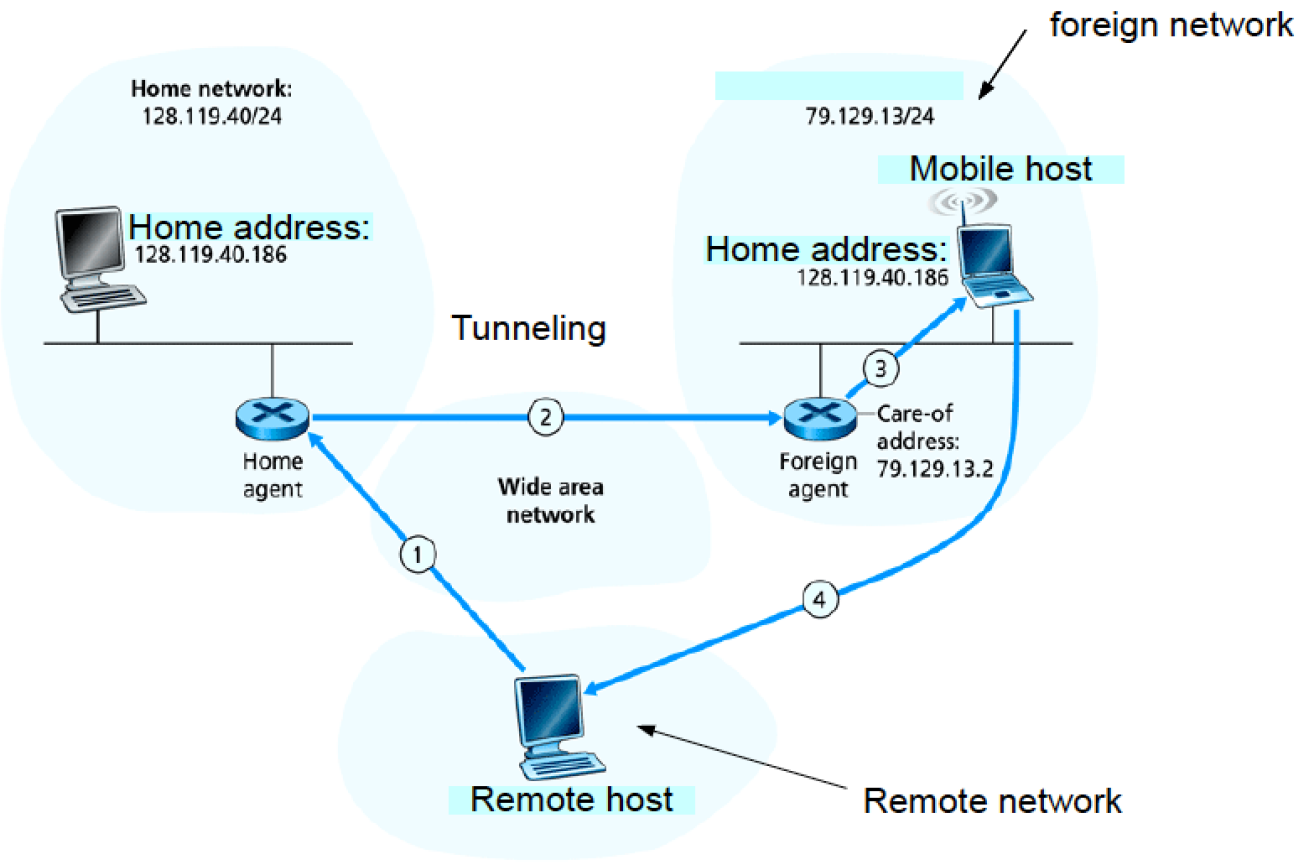


Mobil IP – two solutions

- *indirect routing:* communication from a remote host to mobile host goes through *home agent*, then forwarded to mobile host

- *direct routing:* remote host gets foreign address of mobile host, then sends directly to mobile host

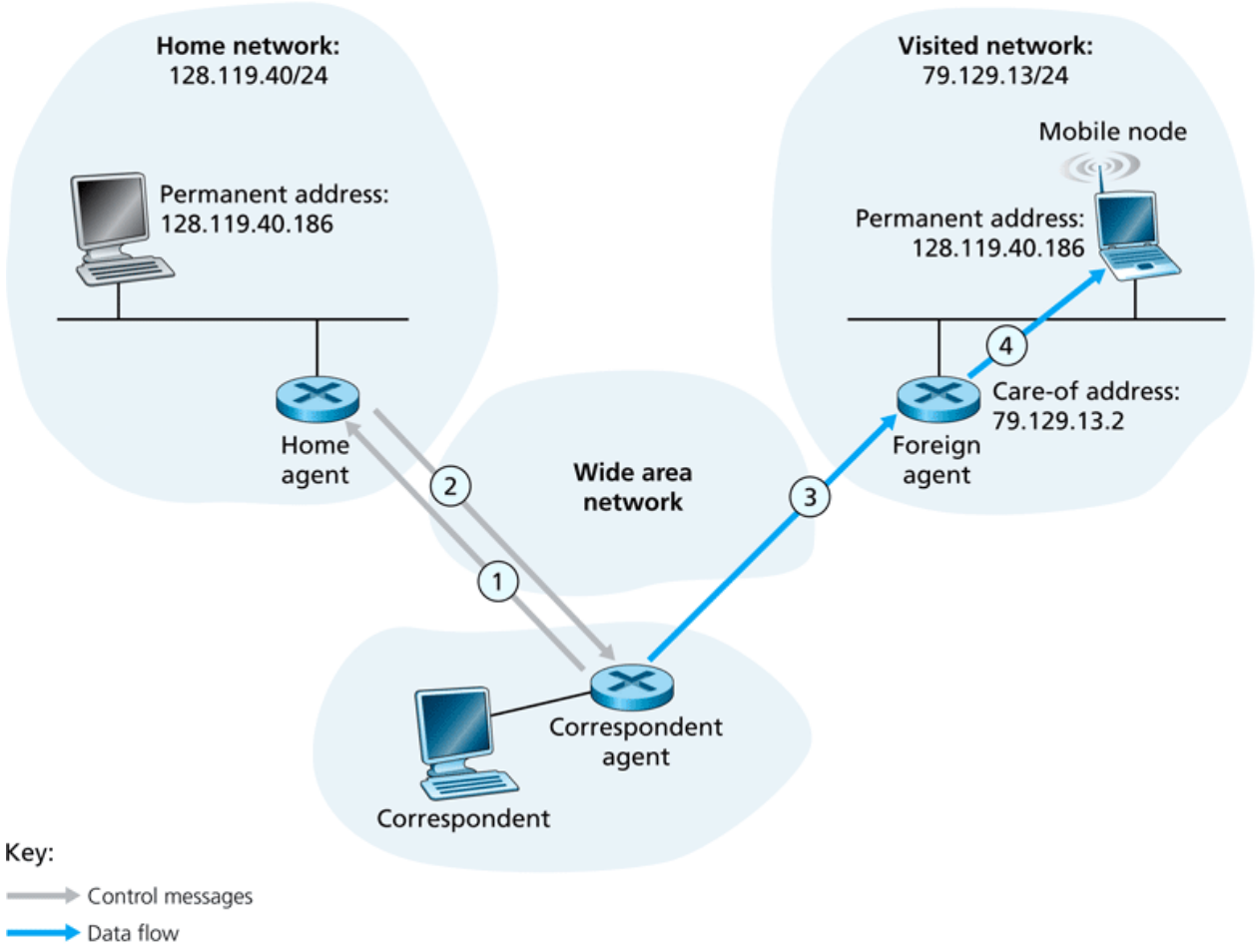
Indirect routing ("triangle" routing)



Indirect routing - comments

- Mobile host uses two addresses:
  - permanent address: used by remote hosts
  - care-of-address: used by home agent for forwarding of packets
- If mobile host moves to another network:
  - registers with new network
  - informs home agent
  - packets forwarded to new “care-of- address”
  - transparent to the remote host!
- Drawback: all packets via home agent
- Inefficient when both remote host and mobile host are in the same network (*double-crossing*)

Direct routing

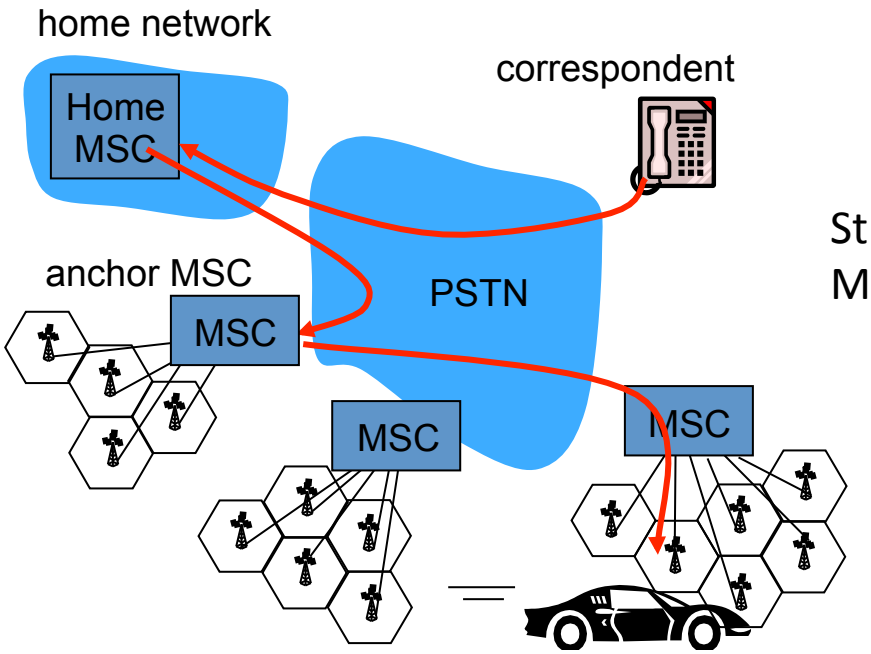


Direct routing - comments

- Direct routing overcomes the triangle routing problem.
- However, direct routing is non-transparent to the remote user, must get care-of-address from home agent.
- What happens if mobile host changes visited network?
  - Proposal: triangle routing through foreign agent in first visited network.

LTE

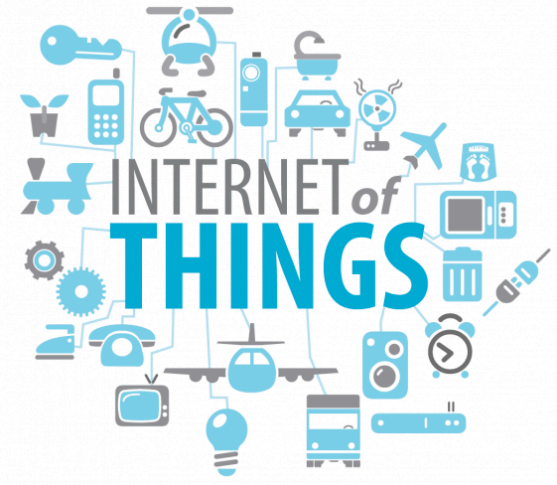
- LTE ("Long Term Evolution", 4G) uses indirect routing over IPv4 networks and a combination of indirect and direct routing over IPv6 networks.



Structure of cellular network.
MSC: Mobile Switching Center

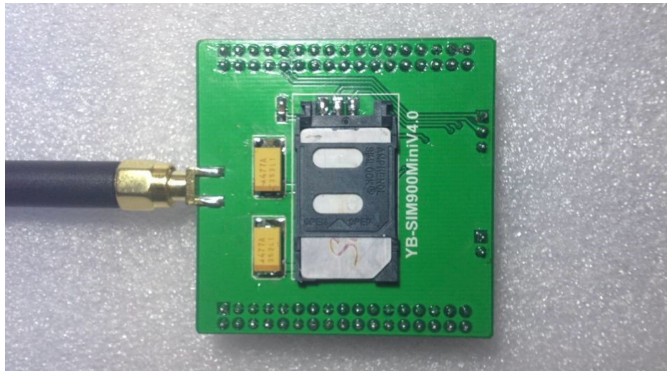
Internet-of-things (IoT)

- It is envisaged that by the year 2020 there will be more than 50 billion devices connected to Internet.
- This will require a new addressing scheme and dedicated access networks.
- Network types:
  - Geographically distributed devices (M2M communication)
  - PAN (Personal Area network): Communication between laptop, pad, mobile phone
  - BAN (Body Area Network) : Communication between devices close to, or inside the body
- Examples of access technologies:
  - Cellular (for geographically distributed devices)
  - Bluetooth (1-24 Mb/s, 10-100m)
  - Wireless USB (110-480 Mb/s, 3-10m)
  - ZigBee (250 kb/s, 10-100m, very low power)
  - Z-Wave (10 – 100 kb/s, 30 m, low power)



Cellular modems – M2M

- Dedicated phone modems that use GSM (GPRS), 3G or LTE to connect to a cellular network
- In the telecom world, these applications are termed “M2M” (machine-to-machine).

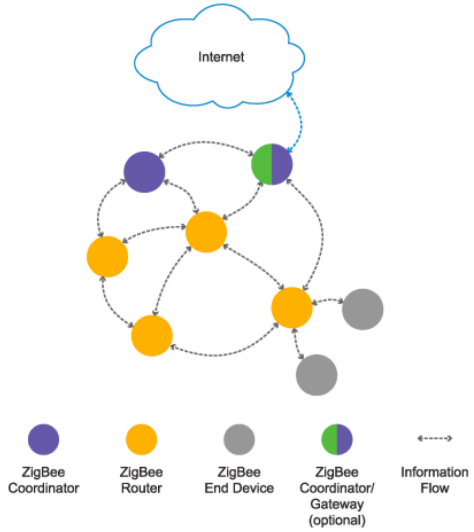


Cellular modem

Example: Electricity monitor (Eseye Ltd)

ZigBee

- ZigBee is an open industry standard to connect sensors and other small devices into a local wireless network.
- The standard requires that devices can run at least 2 years on their battery.
- Uses MAC layer addressing
- Theoretically up to 64000 devices on the same network (in practice around 500)



Example of a small ZigBee network.

- "Coordinators" control the integrity and security of the network,
- "Routers" extend the network,
- End devices are e.g. light switches, sensors.

Network picture : *ZigBee Alliance*

IPv6 : supports IoT

- 128-bit addresses. $2^{128} \approx 3.4*10^{38}$
- New notation (hexadecimal colon notation)
  e.g. FDEC:BA33:0000:0000:FFCD:03F1:0000:0001
  With compression of 0:s:FDEC:BA33::FFCD:3F1:0:1
- IPv4 compatible, example: ::130.24.24.18
- CIDR notation is supported: FDEC::14AB:0:AC6C/60
  (i.e.: 60 bit network prefix, 68 bit host suffix)

IPv6 -> IPv4 transition

Three ways:

- Dual stack: All hosts implement both IPv4 and IPv6
- Tunneling: Two computers want to communicate using IPv6 over a region that uses IPv4. The IPv6 packet is encapsulated inside an IPv4 packet
- Header translation: Translate the IPv6 header to an IPv4 header before delivery on an IPv4 network.

IPv6 – towards virtual circuit

IPv6 defines a label field for identifying virtual connections (faster processing than source and destination IP addresses).

- Coming lectures:
  - The core network: Optical communication

End of general introduction to Internet

  - Some network economics
  - Clouds and Data centers
  - Real-time and media applications
  - Security