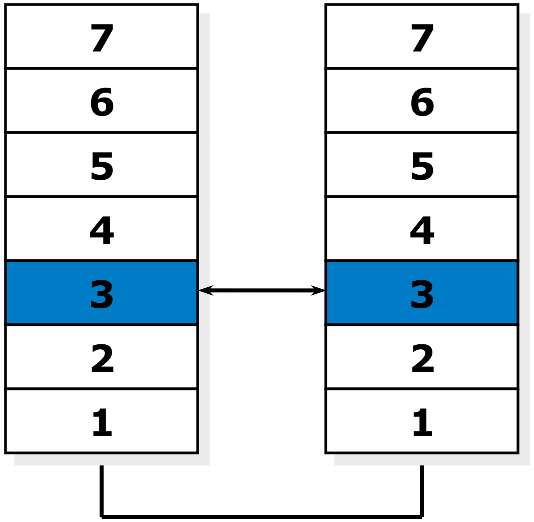
Operating Systems & Computer Networks
Internetworking

- Switches, Routers
- Routing
- Internet Protocol
- Addressing

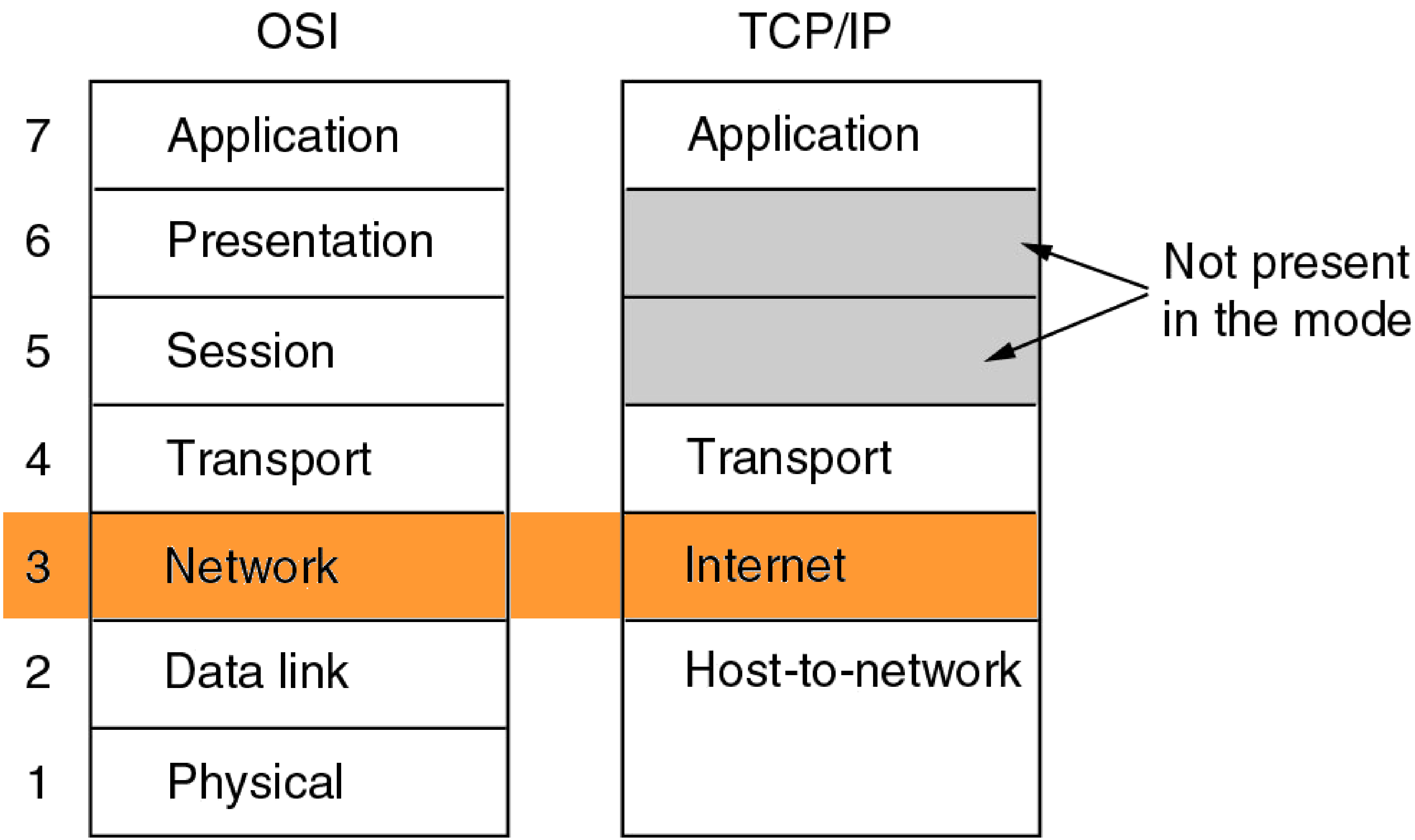


Content (2)

Content (2)

8. Networked Computer & Internet
   - Sockets
   - Internet
   - Layers
   - Protocols
9. Host-to-Network I
   - Physical Layer
   - Media
   - Signals
   - Modems
10. Host-to-Network II
    - Data Link Layer
    - Framing, Flow Control
    - Error Detection / Correction
    - Point-to-Point Protocol
11. Host-to-Network III
    - Topologies
    - Medium Access
    - Local Area Networks
      - Ethernet, WLAN
12. **Internetworking**
    - Switches, Routers
    - Routing
    - Internet Protocol
    - Addressing
13. Transport Layer
    - Protocol Mechanisms
    - TCP, UDP
    - Addressing, Ports

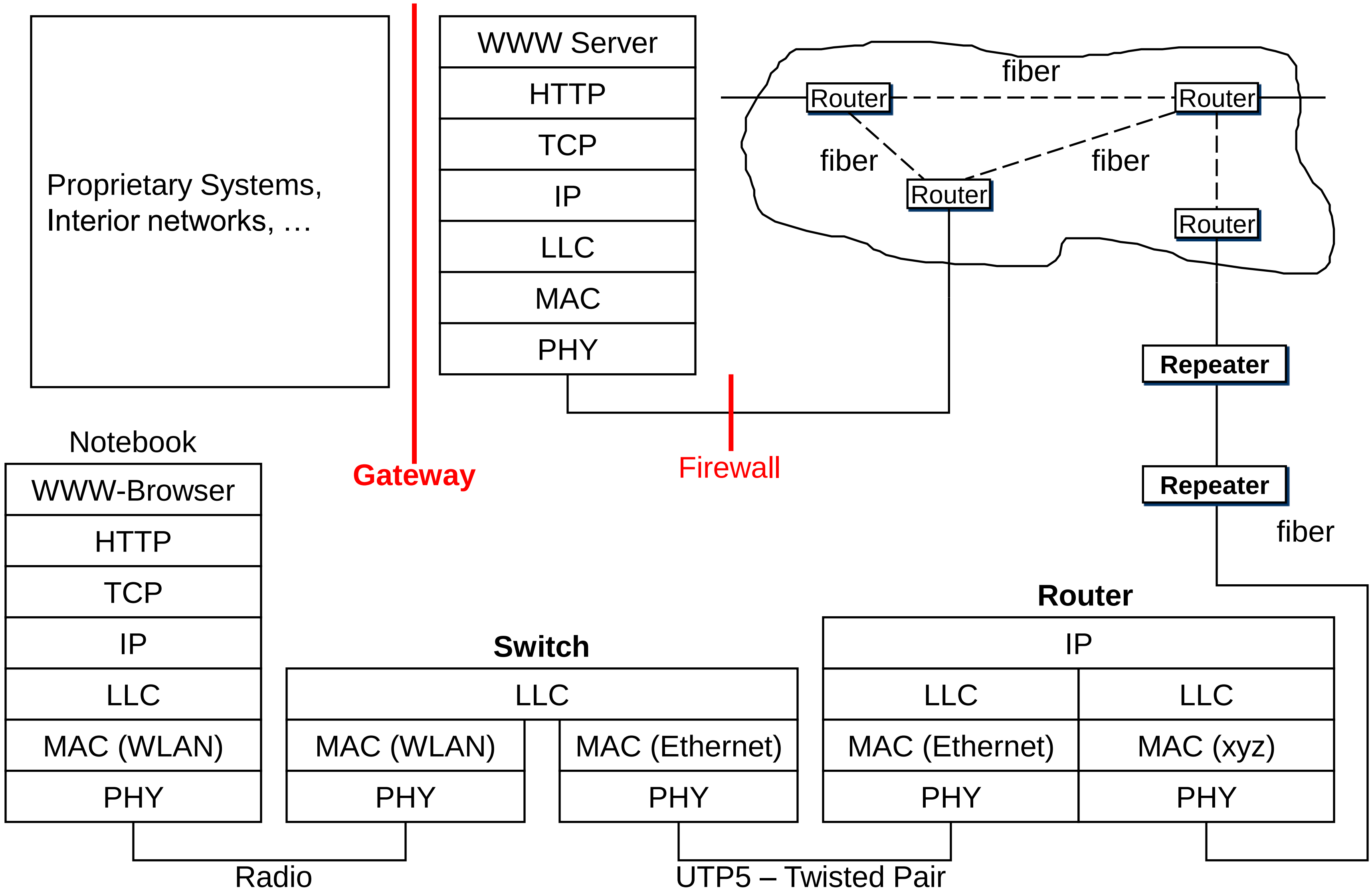
Data Link Layer



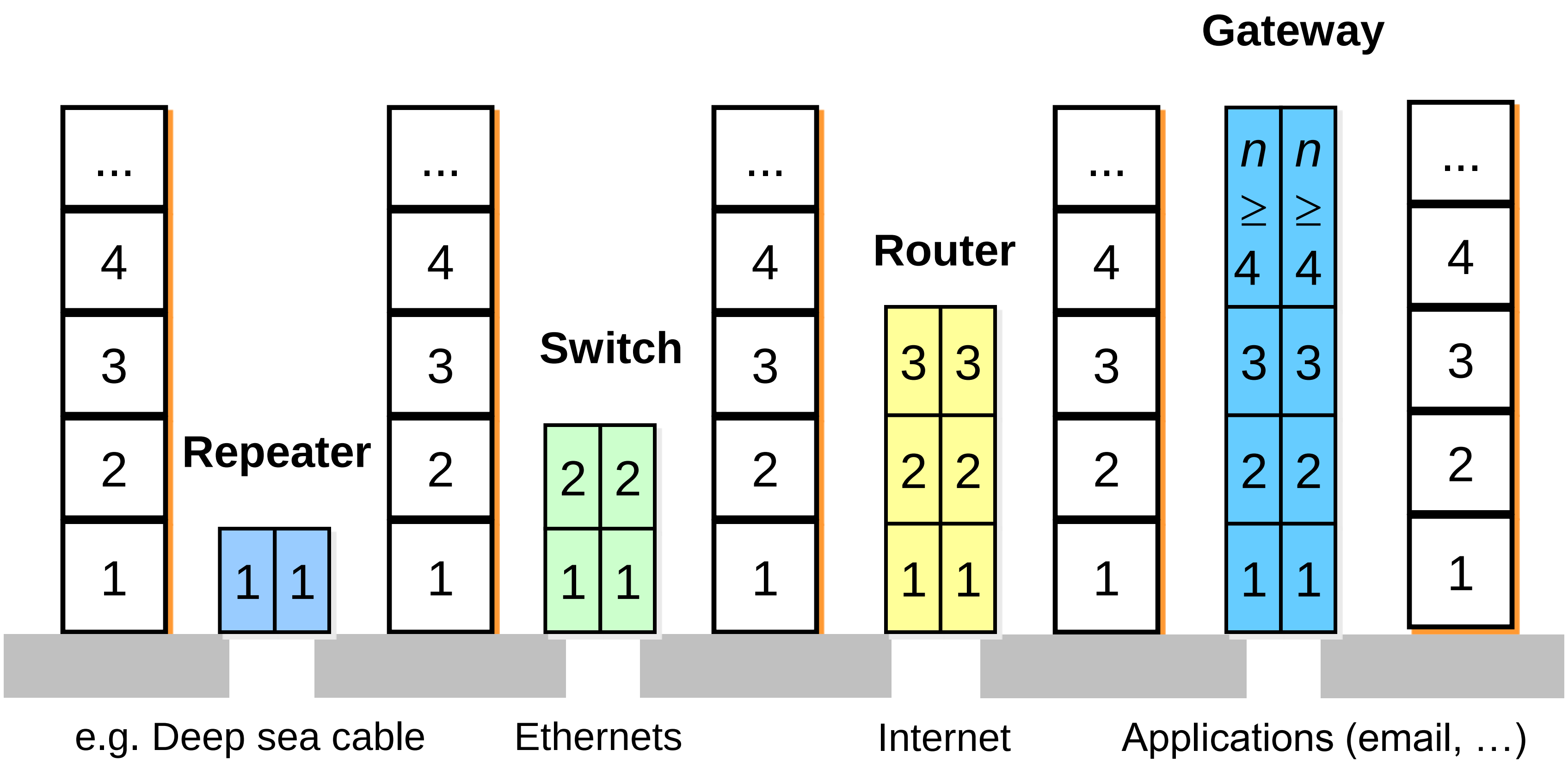
Reasons for Multiple Networks

- Limited number of users/throughput in a single network
- Historical reasons:
    - Different groups started out individually setting up networks
    - Usually heterogeneous
- Geographic distribution of different groups over different buildings, campus, …
    - Impractical/impossible to use a single network because of distance
        - Most MAC protocols set maximum segment length for CSMA/CD
    - Long round-trip delay will negatively influence performance
- Reliability
    - Don’t put all your eggs into one basket
    - “Babbling idiot” problem (isolation of errors)
- Security
    - Contain possible damage caused by promiscuous operation
- Political / business reasons
    - Different authorities, policies, laws, levels of trust, …

Internetworking Units

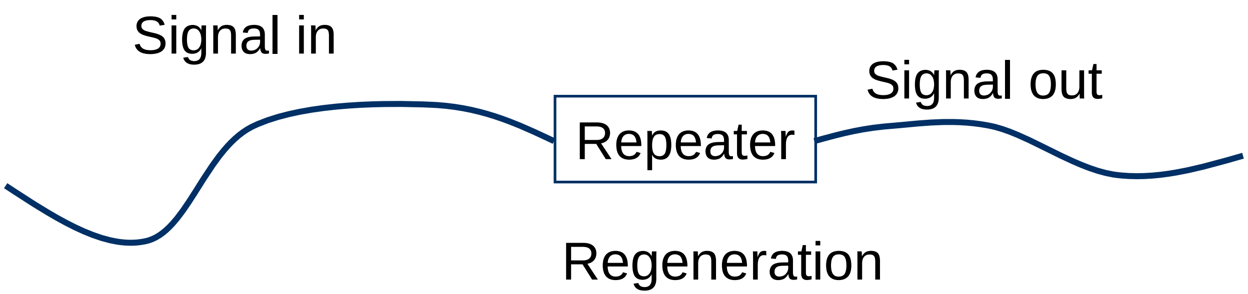


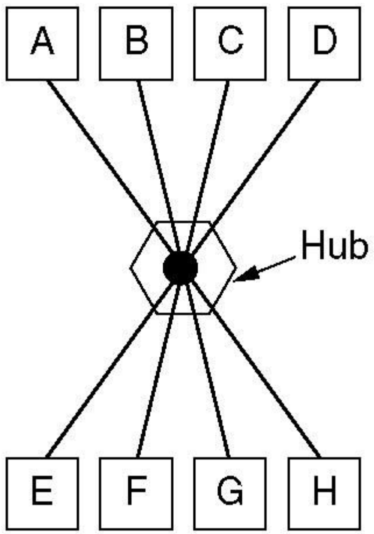
Internetworking Units



Repeater / Hub

- Simplest option: Repeater
  - Physical layer device, connected to two or more cables
  - Amplifies/regenerates arriving signal, puts on other cables
    - Combats attenuation
    - ➢ Signal encodes data (represented by bits)
      - Can be regenerated
      - Opposed to only amplified (which would also amplify noise)
      - ➢ Analog vs. digital transmission
  - Neither understands nor cares about *content (bits)* of packets





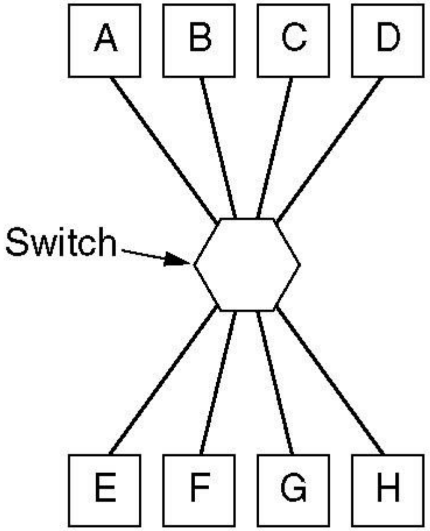
Problems of Physical Layer Solutions

- Physical layer devices, e.g. repeater or hub, do not solve the more interesting problems
  - E.g. no mechanism for handling load, scalability, ...
- Some knowledge of data link layer structure is necessary
  - Ability to understand/inspect content of packets/frames and do something with that knowledge

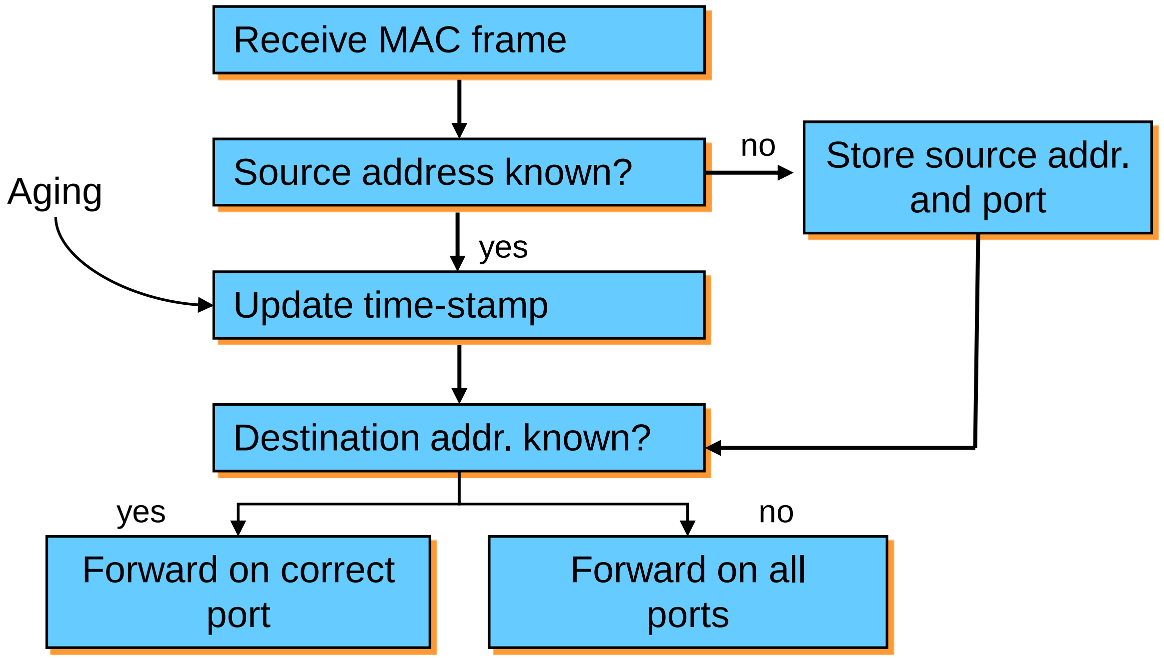
➢ Link-layer devices:
  - Switch: Interconnect several terminals
  - Bridge: Interconnect several networks (of different type)
  - ➢ Nowadays terms sometimes used interchangeably

Switch

- Used to connect several terminals or networks
- Switch inspects arriving packet's MAC addresses and forwards it *only* on correct cable/port
  - Does not bother other terminals
  - Requires data buffer and knowledge *on which* port which terminal is connected
    - Mapping function of MAC address to port
- ➢ How to obtain knowledge about network topology?
  - Observe *from* where packets come to decide how to reach sending terminal
  - ➢ *Backward learning*



Backward Learning – Algorithm



1. Learn address/port mapping from incoming packets
   - Remove expired entries (aging)
2. Forward based on knowledge about destination address
   1. Destination address is known → Forward on correct port
   2. Destination address is unknown → Forward on all ports
      - Only correct receiver will process frame, others will drop it

Routers

- All devices so far either ignored addresses (repeaters, hubs) or worked on MAC-layer addresses (switches, bridges)
- For interconnection outside a single LAN or connection of LANs, these simple addresses are insufficient
  - Unstructured, “flat” addresses do not scale
    - All forwarding devices would need a list of *all* addresses
  - Structured network topologies do not scale
    - World-wide spanning tree is unfeasible

➢ Need more sophisticated addressing structure and devices that operate on it
  - Routers and routing
  - E.g. based on Internet Protocol (IP) addresses

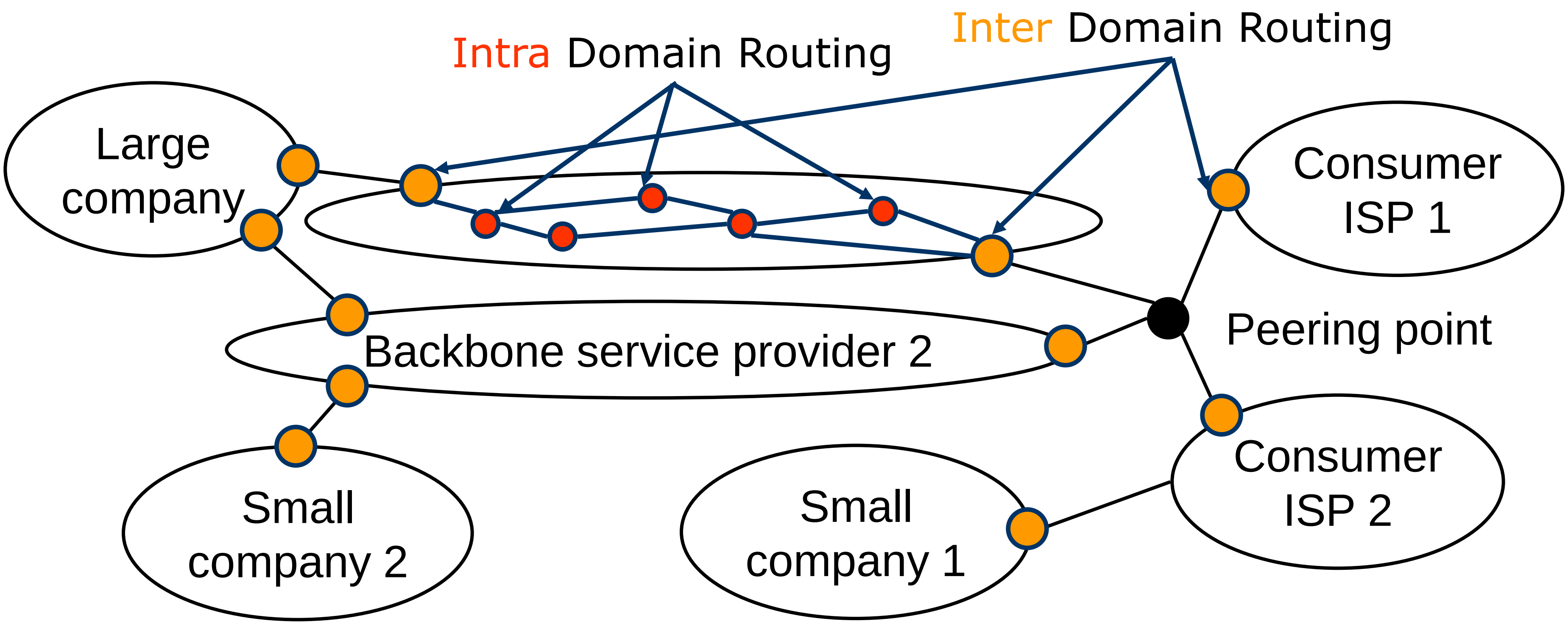
Example: Route to NASA (redone)

```
Z:\>tracert www.nasa.gov

Tracing route to www.nasa.gov.speedera.net [213.61.6.3]
over a maximum of 30 hops:

  1    <1 ms    <1 ms    <1 ms  router-114.inf.fu-berlin.de [160.45.114.1]
  2    <1 ms    <1 ms    <1 ms  zedat.router.fu-berlin.de [160.45.252.181]
  3     1 ms    <1 ms    <1 ms  ice.spine.fu-berlin.de [130.133.98.2]
  4     1 ms    <1 ms    <1 ms  ar-fuberlin1.g-win.dfn.de [188.1.33.33]
  5     1 ms    <1 ms    <1 ms  cr-berlin1-po5-0.g-win.dfn.de [188.1.20.5]
  6     9 ms     9 ms     9 ms  cr-frankfurt1-po9-2.g-win.dfn.de [188.1.18.185]
  7    10 ms     9 ms     9 ms  ir-frankfurt2-po3-0.g-win.dfn.de [188.1.80.38]
  8    10 ms     9 ms     9 ms  DECIX.fe0-0-guy-smiley.FFM.router.COLT.net
                                        [80.81.192.61]
  9    10 ms     9 ms     9 ms  ir1.fra.de.colt.net [213.61.46.70]
 10    11 ms    10 ms     9 ms  ge2-2.ar06.fra.DE.COLT-ISC.NET [213.61.63.8]
 11    11 ms    10 ms    10 ms  213.61.4.141
 12    11 ms    10 ms    10 ms  h-213.61.6.3.host.de.colt.net [213.61.6.3]

Trace complete.
```

- Not all addresses can be resolved to names (see DNS)
- Some requests are redirected to Content Delivery Networks
- Some nodes simply don‘t answer…

The Idea of Internet Routing

- Routing comprises:
  - Updating of routing tables according to routing algorithm
  - Exchange of routing information using routing protocol
  - Forwarding of data based on routing tables and addresses



Autonomous Systems in the IP World

Freie Universität Berlin

- Large organizations can own multiple networks that are under single administrative control
  - ➢ Forming *autonomous system* or *routing domain*
- Autonomous systems form yet another level of aggregating routing information
  - ➢ Give raise to *inter-* and *intra-domain routing*
- Inter-domain routing is hard
  - One organization might not be interested in carrying a competitor's traffic
  - Routing metrics of different domains cannot be compared
    - ➢ Only *reachability* can be expressed
  - Scalability: Currently, inter-domain routers have to know about 150,000-200,000 networks

Intra-domain Routing: OSPF

- The Internet's most prevalent intra-domain (= interior gateway) routing protocol: *Open Shortest Path First* (OSPF)

- Main properties:
  - Open, variety of routing distances, dynamic algorithm
  - Routing based on traffic type (e.g. real-time traffic uses different paths)
  - Load balancing: Also put some packets on the 2nd, 3rd best path
  - Hierarchical routing, some security in place, support tunneled routers in transit networks

- Essential operation: Compute shortest paths on graph abstraction of autonomous system
  - ➢ Link state algorithm

Basic Ideas of Link State Routing

- Distributed, adaptive routing

- Algorithm:
  1. Discovery of new neighbors
     - HELLO packet
  2. Measurement of delay / cost to all neighbors
     - ECHO packet measures round trip time
  3. Creation of link state packets containing all learned data
     - Sender and list of neighbors (including delay, age, ...)
     - Periodic or event triggered update (e.g. upon detecting new neighbors, line failure, ...)
  4. Flooding of packet to all neighbors
     - Flooding, but with enhancements: Duplicate removal, deletion of old packets, ...
  5. Shortest path calculation to all other routers (e.g. Dijkstra)
     - Computing intensive, optimizations exist

Inter-domain Routing: BGPv4

- Routing between domains: *Border Gateway Protocols* (BGP)
- BGP's perspective: Only autonomous systems and their connections
  - Routing complicated by politics, e.g. only route packets for paying customers, …
  - Legal constraints, e.g. traffic originating and ending in Canada must not leave Canada while in transit
- Basic operation: Distance vector protocol
  - Propagate information about reachable networks and distances one hop at a time
    - Each router learns only next step to destination
  - Optimizations in BGP:
    - Not only keep track of cost via a given neighbor, but store entire paths to destination ASs
    - More robust, solves problems like count to infinity, i.e. can handle disconnected networks efficiently

Conclusion: Interconnections

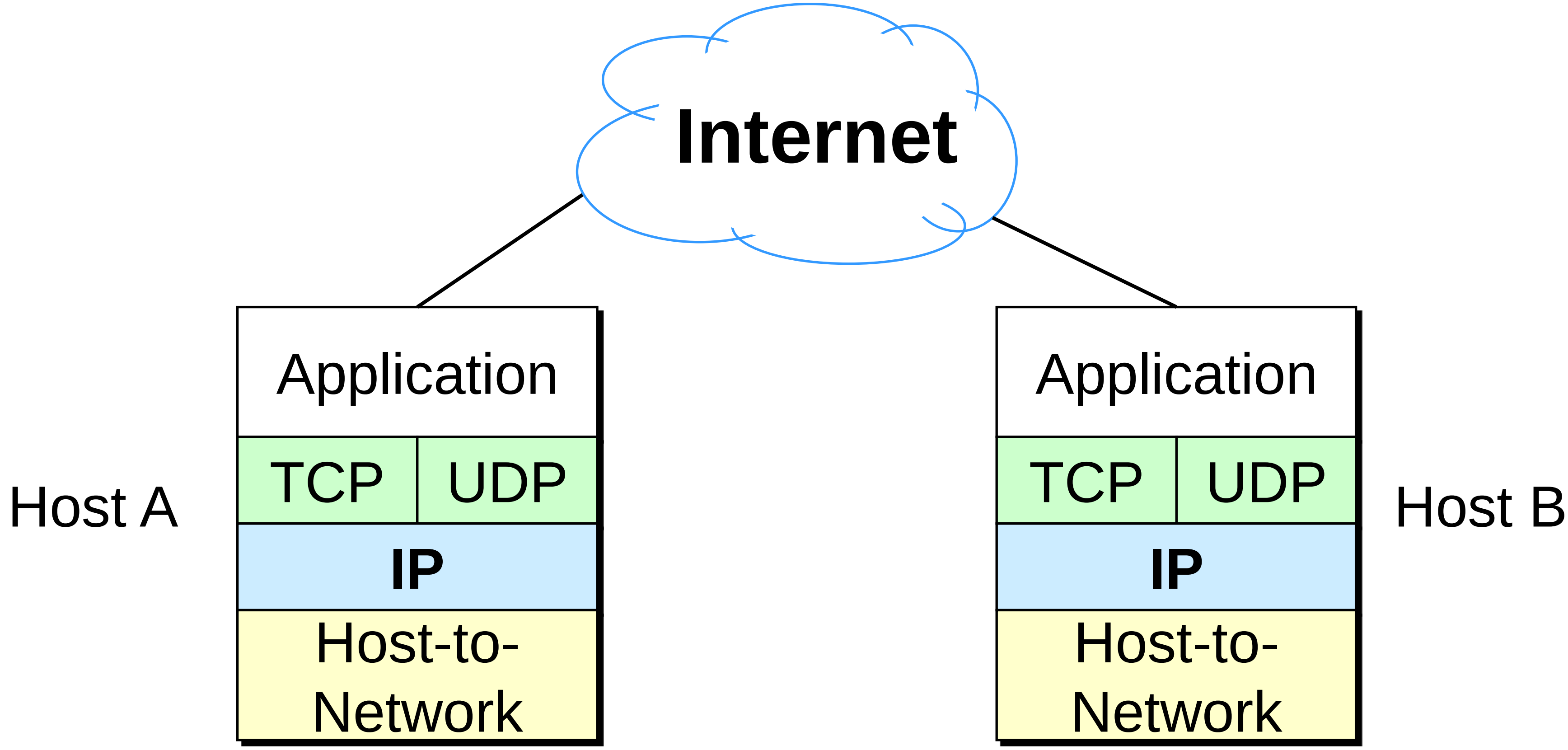
- Single LANs are insufficient to provide communication for all but the simplest installations

- Interconnection of LANs necessary
  - Interconnect on purely physical layer: Repeater, hub
  - Interconnect on data link layer: Bridges, switches
  - Interconnect on network layer: Router
  - Interconnect on higher layer: Gateway

- Problems:
  - Redundant bridges can cause traffic floods; need spanning tree algorithm
  - Simple addresses do not scale; need routers

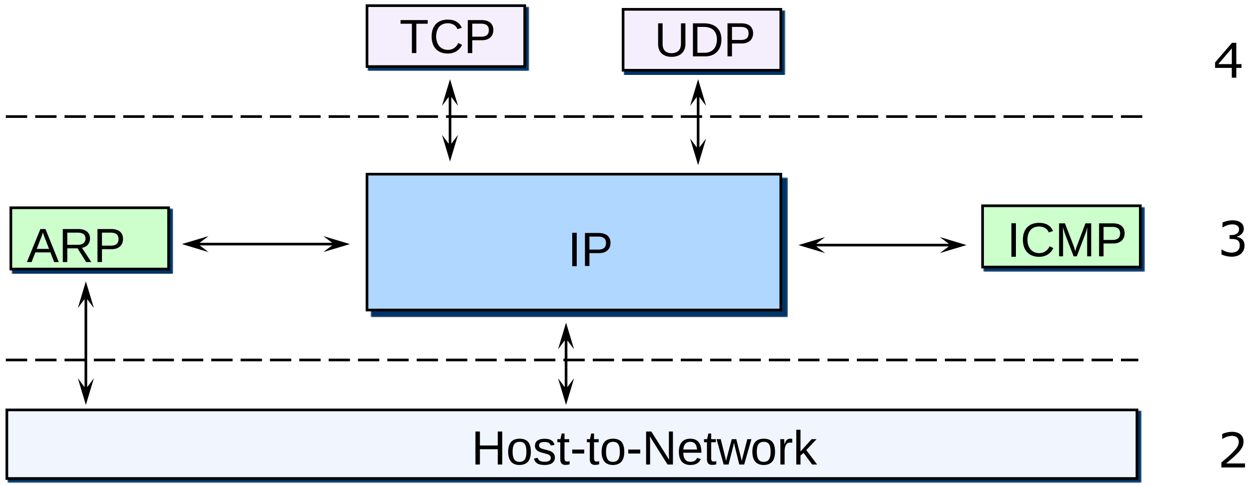
IP

Simplified View of Internet protocols



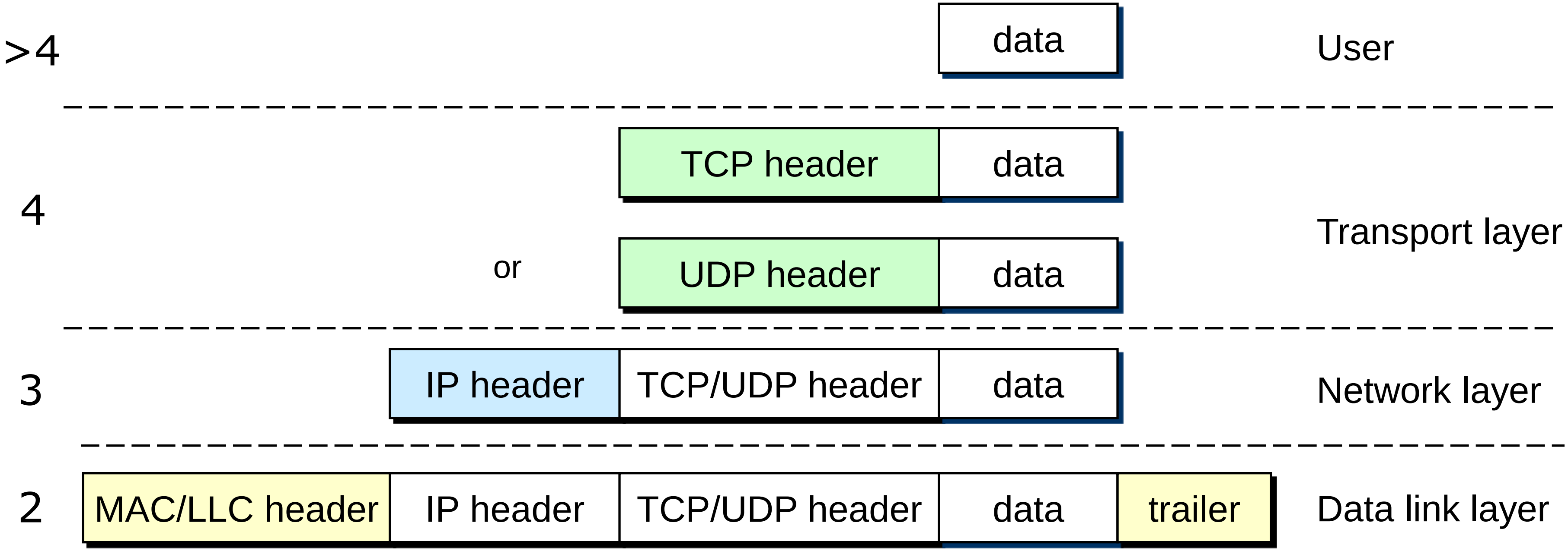
IP and Supporting Protocols

- Transport protocols (Layer 4, TCP or UDP) hand over data together with IP address of receiver to Internet Protocol (IP)
- IP may need to ask Address Resolution Protocol (ARP) for MAC address (Layer 2)
- IP hands over data together with MAC address to Layer 2
- IP forwards data to higher layers (TCP or UDP)
- Internet Control Message Protocol (ICMP) can signal problems during transmission



Data Encapsulation / Decapsulation

- IP forwards data packets through network to receiver
- TCP/UDP add ports (dynamic addresses of processes)
- TCP offers reliable data transmission
- Packets (PDU, protocol data unit) are encapsulated



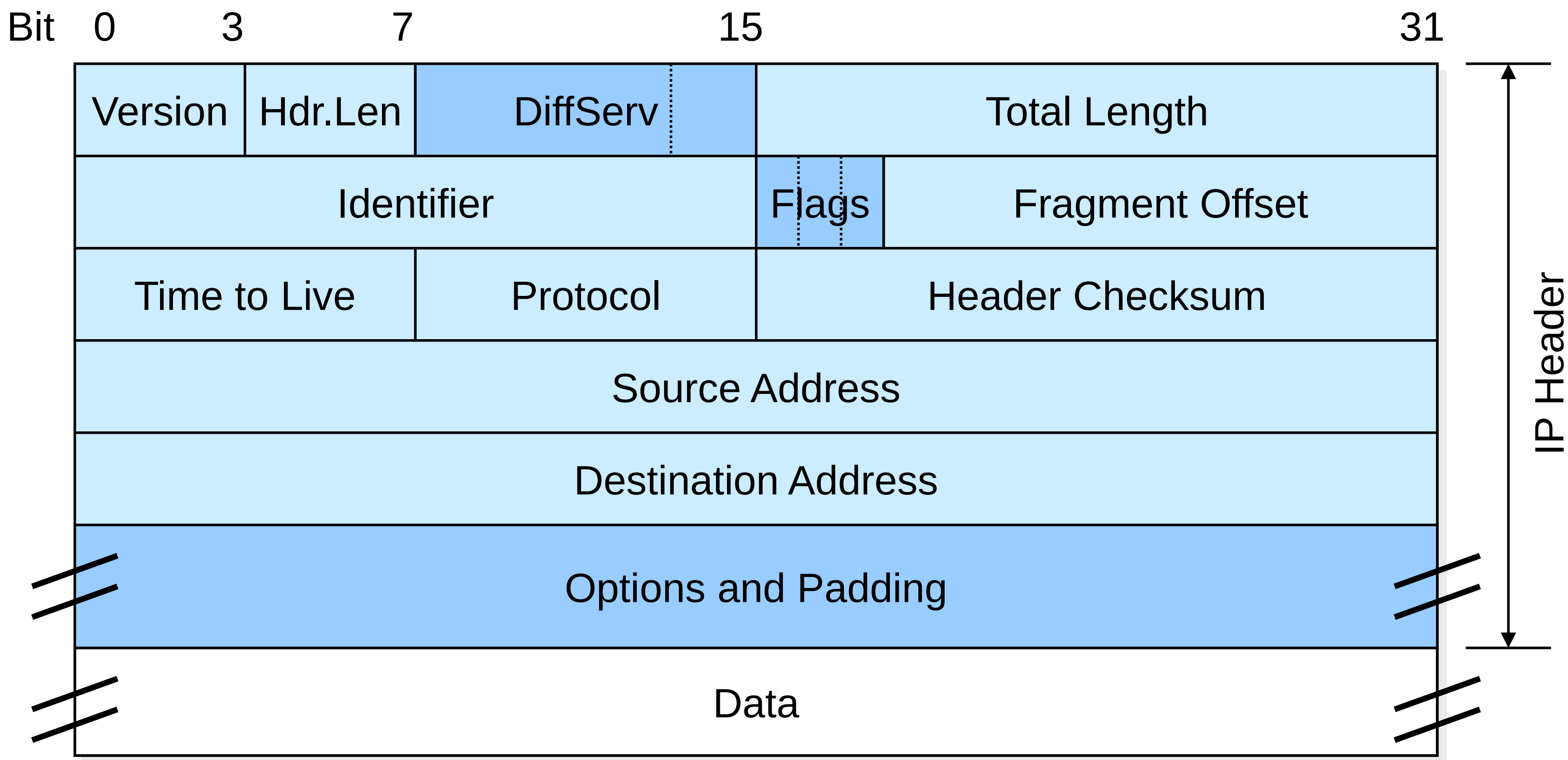
Internet Protocol (IP)

- History:
  - Original development with support of US Department of Defense
  - Already used back in 1969 in APANET
- Tasks
  - Routing support using structured addresses
  - Checking of packet lifetime to avoid routing loops
  - Fragmentation and reassembly
  - Network diagnostics support
- Development
  - Today IP (version 4) is most widely used layer 3 protocol
  - Further development started back in the 80s/90s
    - Project IPng (IP next generation) of the IETF (Internet Engineering Task Force)
  - Result in mid 90s: IPv6, still only rarely used

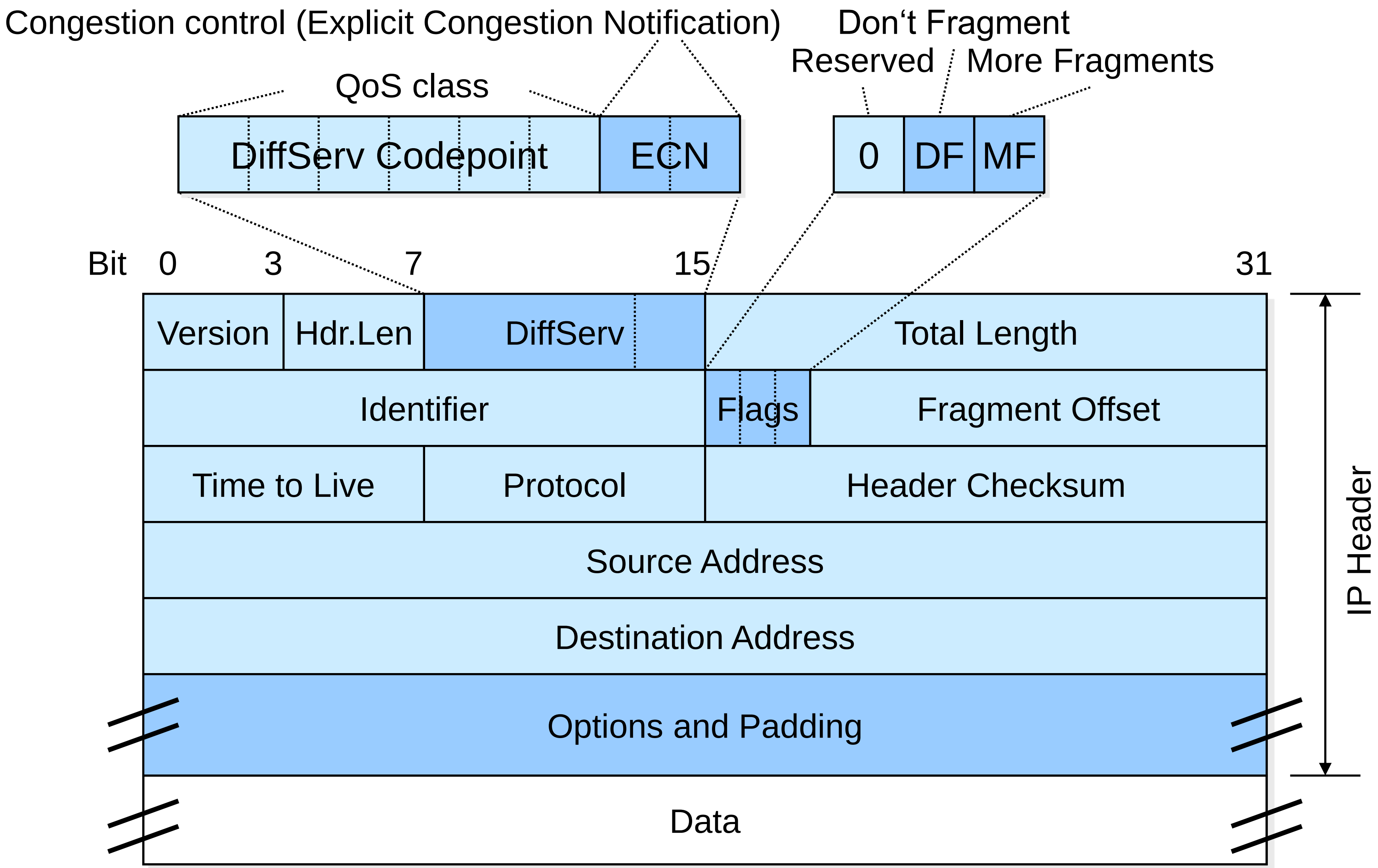
Properties of IP

- Packet oriented
- Connectionless (datagram service)
- Unreliable transmission
  - Datagrams can be lost
  - Datagrams can be duplicated
  - Datagrams can be reordered
  - Datagrams can circle, but solved by Time to Live (TTL) field
  - IP cannot handle Layer 2 errors
  - At least there is ICMP to signal errors
- Routing support via structured addresses
- No flow control (yet, first steps taken)
- Used in private and public networks

IPv4 Datagram



IPv4 Datagram



Structured IP Addresses and Address Classes (Classical View)

0 1 2 4 8 16 24 31

0	Network ID	Host ID

1. Class A: 128 networks, 16M hosts

1.0.0.0 –
127.255.255.255

1	0	Network ID	Host ID

2. Class B: 16k networks, 64k hosts

128.0.0.0 –
191.255.255.255

1	1	0	Network ID	HostID

3. Class C: 2M networks, 256 hosts

192.0.0.0 –
223.255.255.255

1	1	1	0	Multicast address

4. Class D: group communication (Multicast)

224.0.0.0 –
239.255.255.255

1	1	1	1	0	Reserved

5. Class E: reserved for future use

240.0.0.0 –
255.255.255.255

Special IP Addresses

- Some IP addresses are reserved for special uses:

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		This host
0 0 . . . 0 0	Host	A host on this network
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		Broadcast on the local network
Network	1 1 1 1 . . . 1 1 1 1	Broadcast on a distant network
127	(Anything)	Loopback

- Not all of the network/host combinations are available

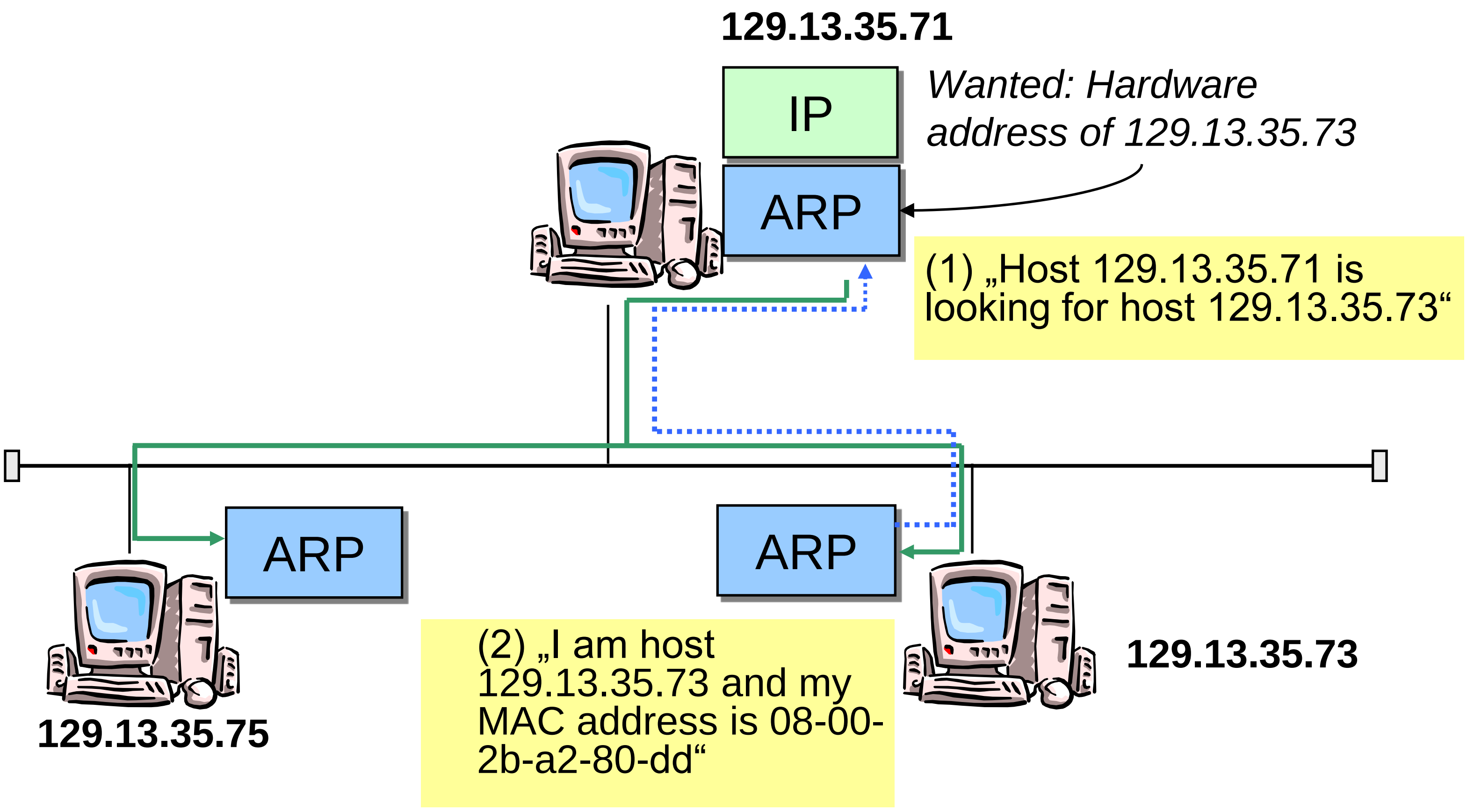
➢ So-called "private" IP addresses
  - Used for internal networks (addresses not routable)
  - Example: 10.0.0.1, 192.168.0.1

Bridging Addressing Gap: ARP

- What happens once a packet arrives at its destination network / LAN?
  - IP address (which is all that is known about destination) needs to be translated into a MAC address that corresponds to the IP address

- Simple solution: Broadcast
  - Broadcast on LAN, asking which node has requested IP address
  - Node answers with its MAC address
  - Router can then forward packet to that MAC address
- *Address Resolution Protocol* (ARP)

Example: ARP


129.13.35.71
IP
ARP
Wanted: Hardware address of 129.13.35.73
(1) „Host 129.13.35.71 is looking for host 129.13.35.73“
ARP
ARP
129.13.35.75
(2) „I am host 129.13.35.73 and my MAC address is 08-00-2b-a2-80-dd“
129.13.35.73

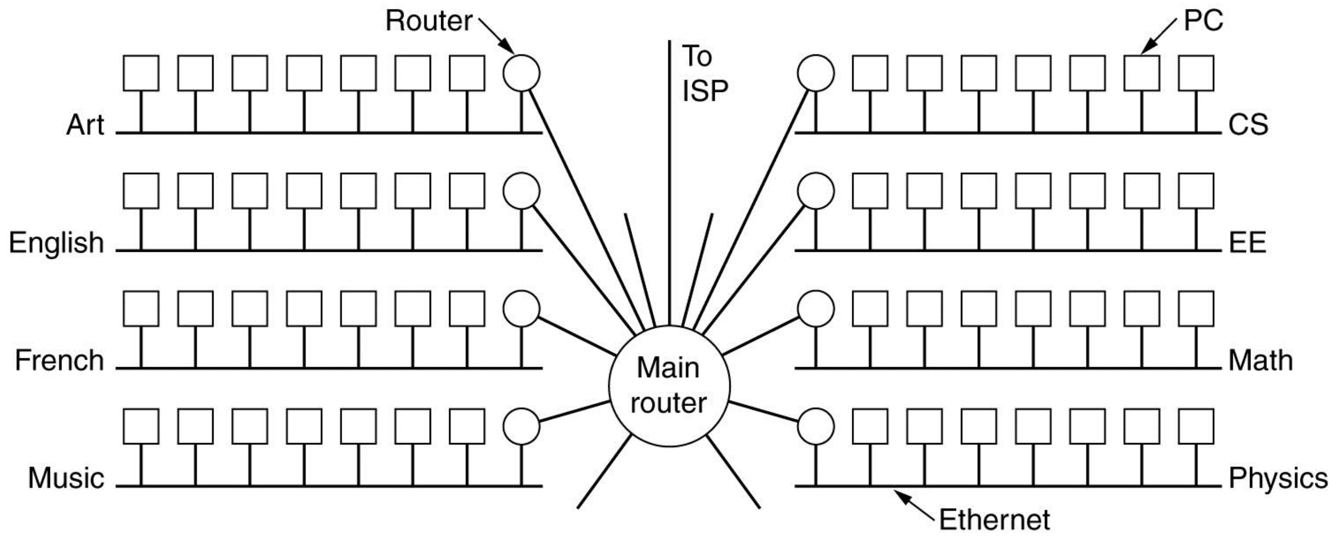
Scalability Problems of IP

- Class A and B networks can contain *many* hosts
  - Too many for a router to easily deal with
  - Additionally, administrative problems in larger networks
  - ➢ Solution: Subnetting, i.e. a network is subdivided into several smaller networks by breaking up the address space

- Network classes waste a lot of addresses
  - Example: Organization with 2000 hosts requires a class B address, wasting 64K-2K ≈ 62.000 host addresses
  - ➢ Solution: Classless addressing ➔ Classless Inter Domain Routing (CIDR)
    - Dynamic boundaries between host/network part of IP address
    - Aggregation on routers to reduce size of global routing table

Subnetting

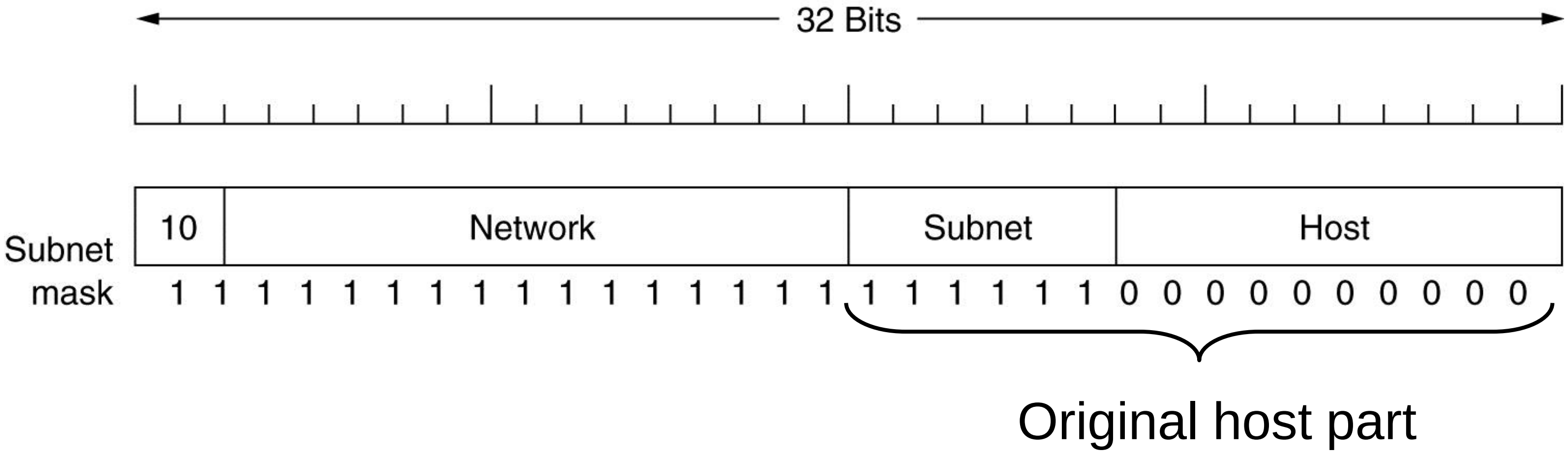
- Suppose an organization has one class B address but is organized into several LANs
  - Example: University with different departments



- Main router should be concerned with whole networks
  - Should not be bothered with all the nodes in each departments
- Obvious case for hierarchical routing and addressing
  - ➢ How to put hierarchies into existing IP addresses?

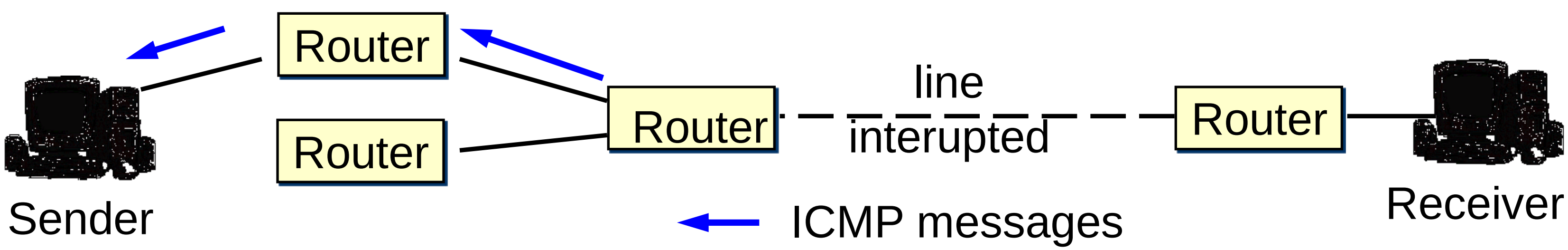
Subnetting – Hierarchies in Addresses

- Manipulating class bits to introduce more hierarchy levels is not practical
- Idea: Have more hierarchy levels implicitly
  - Introduce a *subnet*, represented by “borrowing” bits from host part of IP address
  - Local router has to know where to apply this split
    - Needs a *subnet mask*
  - Represented as x.y.u**/#bits** or as bit pattern needed to mask out the host bits



Controling IP: ICMP

- IP is responsible for (unreliable) data transfer only
- Internet Control Message Protocol (ICMP) is used for error reporting and testing



- Examples:
  - Destination Unreachable
  - Time Exceeded: Time-to-Live field reaches 0
    - Also used when looking up routes using traceroute
  - Echo Request / Reply ("ping")
  - Timestamp Request / Reply

Conclusion: Internet Protocol

- Unreliable datagram transfer

- Needs supporting protocols
  - ARP for mapping IP to MAC address
  - ICMP for error signaling

- Classical addressing wastes addresses
  - Subnetting, subnet masks
  - Classless addressing, CIDR

- Version 4 dominant, version 6 coming (since years…)
  - Much more in Telematics

Content (2)