Networking and Internetworking

Agenda

Types of networks

Network principles

Internet protocols

Case studies

Internet users (june 2020)

WORLD INTERNET USAGE AND POPULATION STATISTICS 2020 Year-Q2 Estimates						
World Regions	Population (2020 Est.)	Population % of World	Internet Users 30 June 2020	Penetration Rate (% Pop.)	Growth 2000-2020	Internet World %
Africa	1,340,598,447	17.2 %	566,138,772	42.2 %	12,441 %	11.7 %
Asia	4,294,516,659	55.1 %	2,525,033,874	58.8 %	2,109 %	52.2 %
Europe	834,995,197	10.7 %	727,848,547	87.2 %	592 %	15.1 %
Latin America / Caribbean	654,287,232	8.4 %	467,817,332	71.5 %	2,489 %	9.7 %
Middle East	260,991,690	3.3 %	184,856,813	70.8 %	5,527 %	3.8 %
North America	368,869,647	4.7 %	332,908,868	90.3 %	208 %	6.9 %
Oceania / Australia	42,690,838	0.5 %	28,917,600	67.7 %	279 %	0.6 %
WORLD TOTAL	7,796,949,710	100.0 %	4,833,521,806	62.0 %	1,239 %	100.0 %

https://www.internetworldstats.com/stats.htm

Principles of networking

- The principles on which computer networks are based include protocol layering, packet switching, routing, and data streaming.
- Internetworking techniques enable heterogeneous networks to be integrated.
- The Internet is the major example; its protocols are almost universally used in distributed systems.
- The addressing and routing schemes used in the Internet have withstood the impact of its enormous growth.
- They have been revisionated to accommodate the growth of users and usages and to meet new application requirements for mobility, security and quality of service

- Principles of networking are counter-intuitive for software developers with limited experience of networks

Distributed systems are based on networks

- Distributed systems are made of software (processes) which use communication infrastructures:
  local area networks, wide area networks, and internetworks.
- The performance, reliability, scalability, mobility, and Quality of Service properties of the underlying networks impact the design and behaviour of distributed systems.
- Technological innovations have resulted in the emergence of wireless networks and high-performance networks with Quality of Service guarantees

Common **false** assumptions on networks

- The network is reliable
- The network is secure
- The network is homogeneous
- The topology does not change
- Latency is zero
- Bandwidth is infinite
- Transport cost is zero
- There is one administrator

When developing non-distributed applications, most of these issues will likely not show up

Network performance

	Example	*Range*	*Bandwidth (Mbps)*	*Latency (ms)*
Wired:				
LAN	Ethernet	1–2 kms	10–10,000	1–10
WAN	IP routing	worldwide	0.010–600	100–500
MAN	ATM	2–50 kms	1–600	10
Internetwork	Internet	worldwide	0.5–600	100–500
Wireless:				
WPAN	Bluetooth (IEEE 802.15.1)	10–30m	0.5–2	5–20
WLAN	WiFi (IEEE 802.11)	0.15–1.5 km	11–108	5–20
WMAN	WiMAX (IEEE 802.16)	5–50 km	1.5–20	5–20
WWAN	3G phone	cell: 1–-5	348–14.4	100–500

km

Network performance main parameters

The primary network performance parameters are the **latency** and the **point-to-point data transfer rate**:

- **Latency** is the delay that occurs after a send operation is executed and before data starts to arrive at the destination computer. It can be measured as the time required to transfer an empty message (we are considering only network latency, which forms a part of the process-to-process latency).
- **Data transfer rate** is the speed at which data can be transferred between two computers in the network once transmission has begun, usually quoted in bits per second.

The time required for a network to transfer a message containing *length* bits between two computers is:

Message transmission time = latency + length / data transfer rate

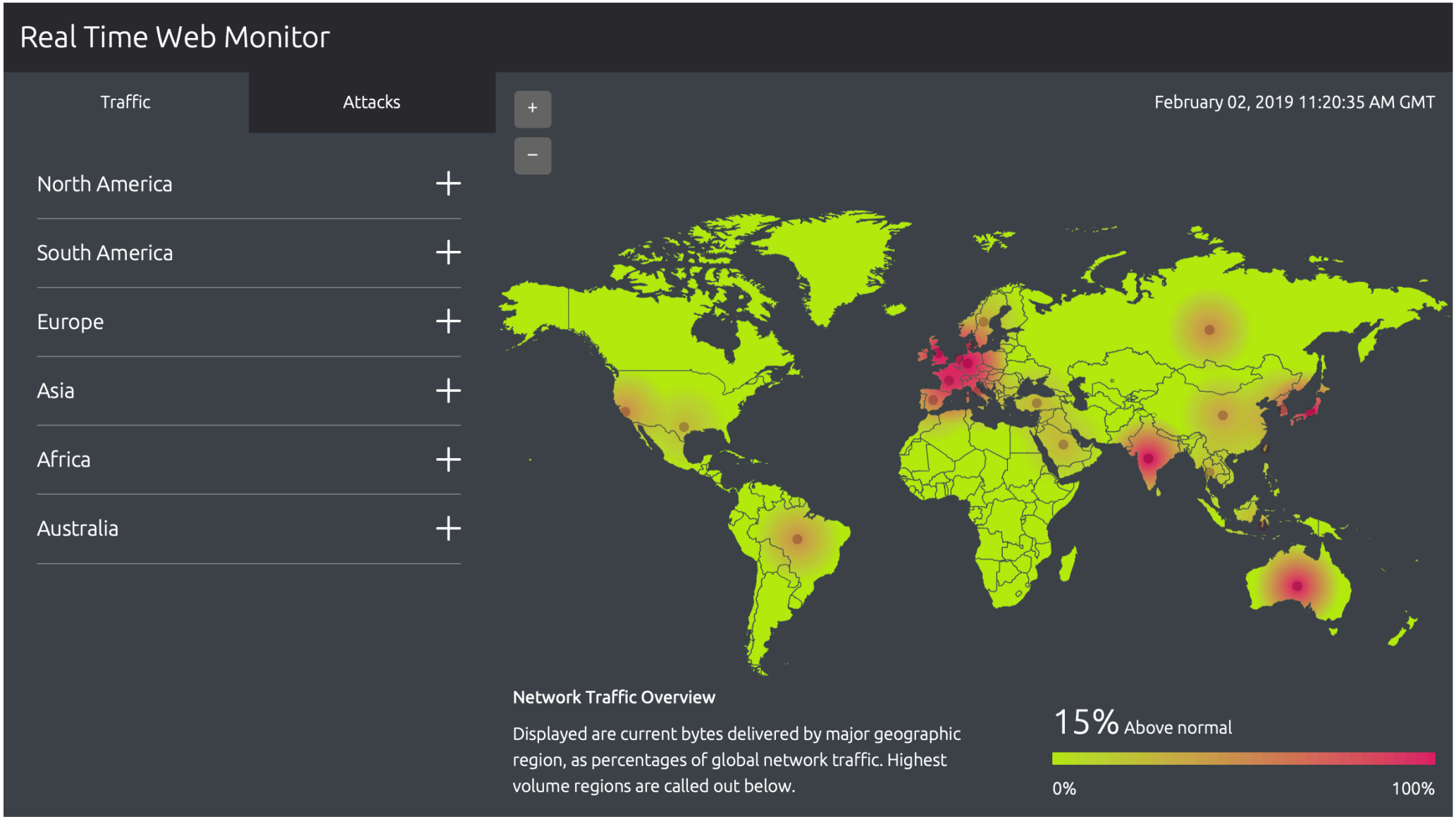
Network parameters in the Internet

On the Internet, round-trip latencies are in the 5–500 ms range, with means of 20–200 ms depending on distance

Requests transmitted across the Internet are 10–100 times slower than those sent on fast local networks.

The bulk of this time difference derives from switching delays at routers and contention for network circuits.

www.akamai.com https://www.akamai.com/us/en/solutions/intelligent-platform/visualizing-akamai/real-time-web-monitor.jsp

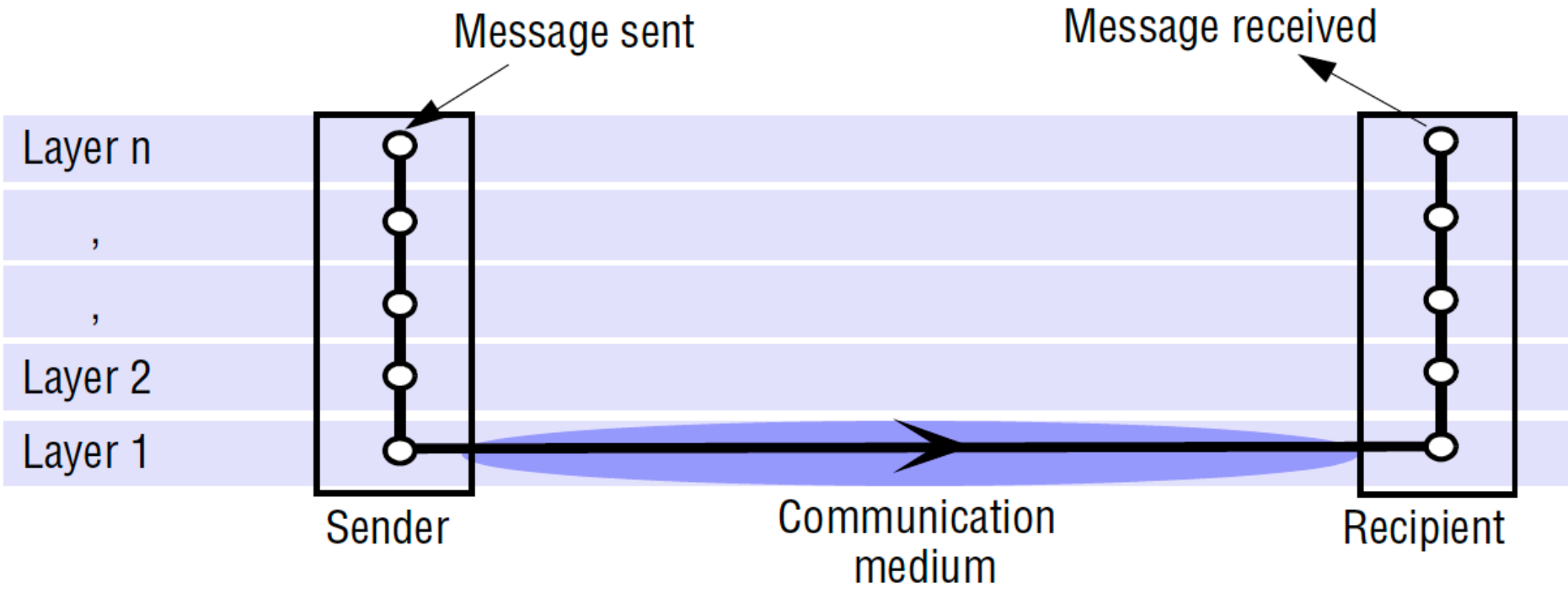


InterProcess Communication (IPC)

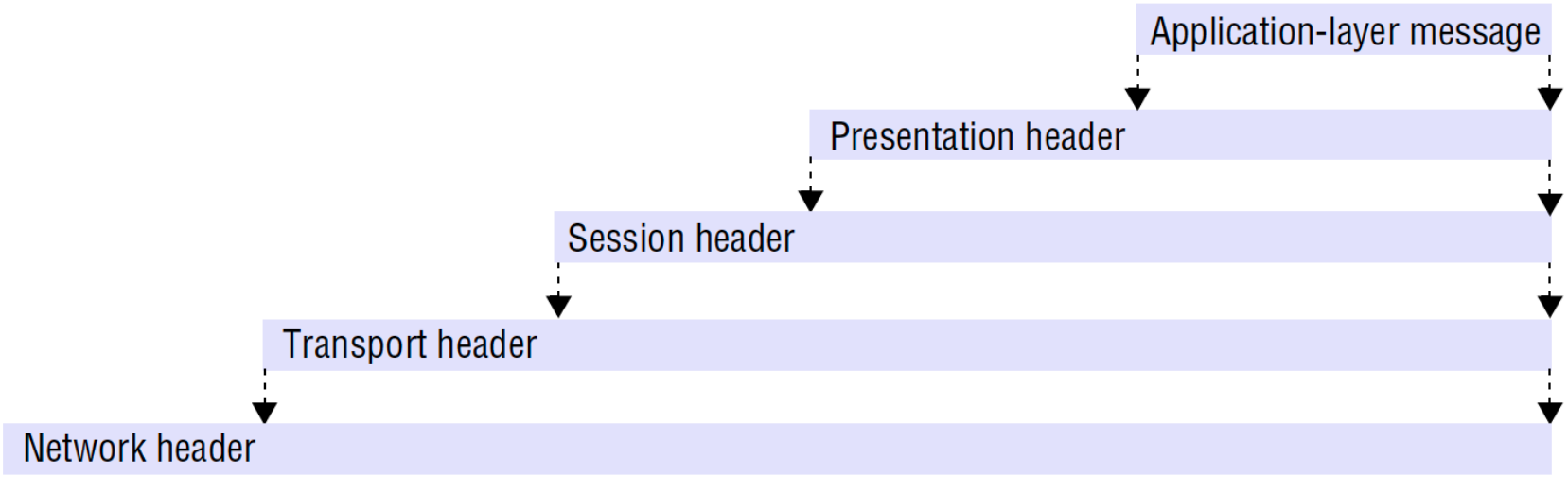
- Question: how do processes on different machines exchange information?
- Answer: with difficulty … ☹
- Established computer network facilities are **too primitive**, resulting in DSs that are too difficult to develop – a new model is required
- IPC is the “heart” of every distributed system.
- Four IPC models are popular:

  RPC; RMI; MOM and Streams

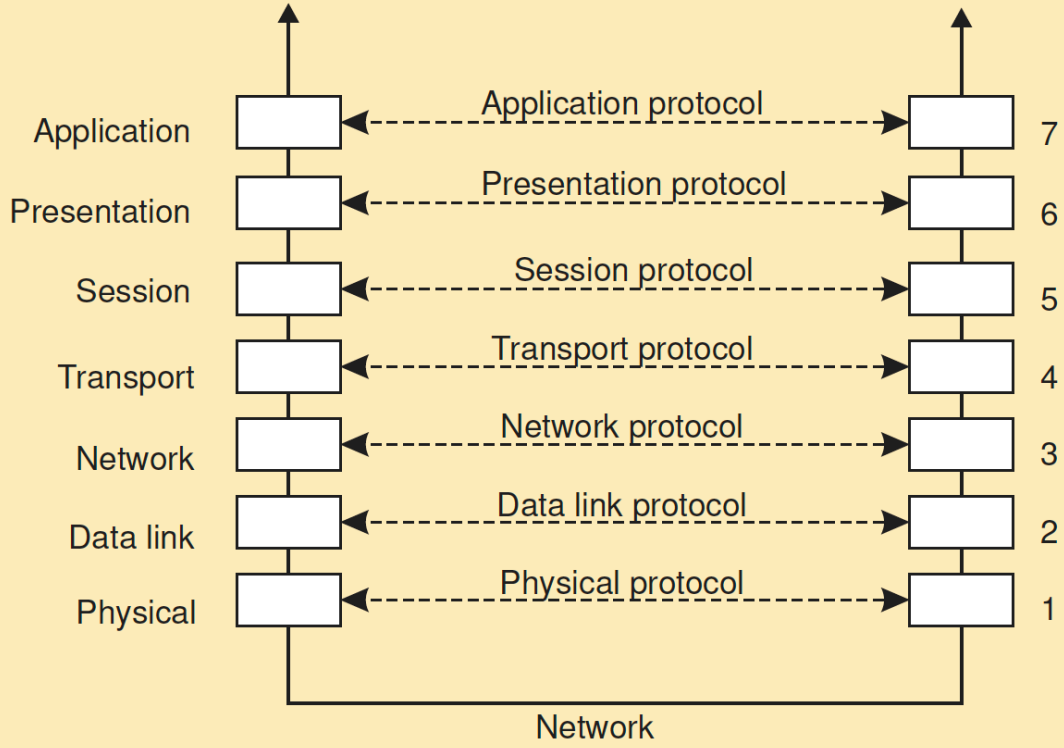
Conceptual layering of protocol software



Encapsulation as it is applied in layered protocols



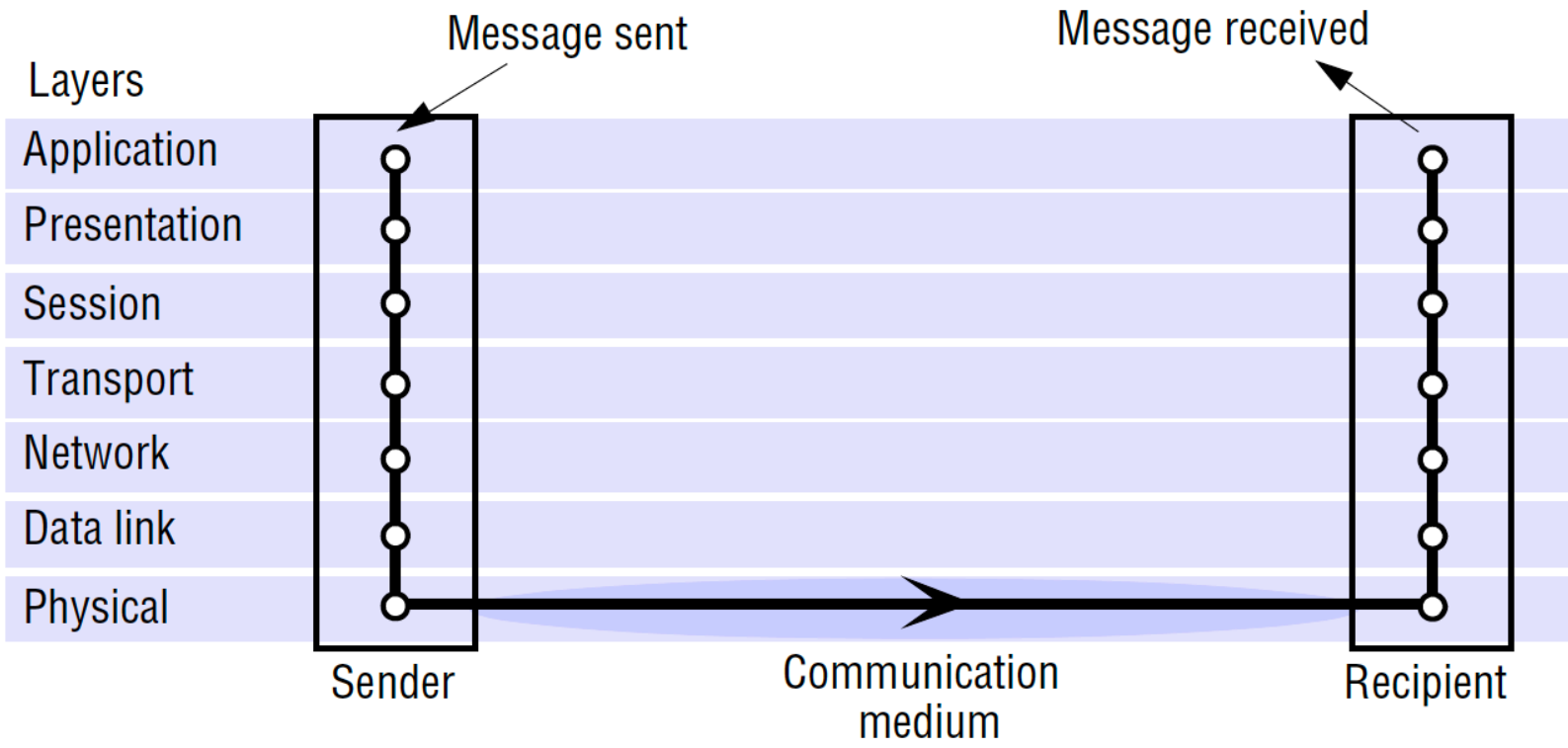
Basic networking model



Drawbacks

- Focus on message-passing only
- Often unneeded or unwanted functionality
- Violates access transparency

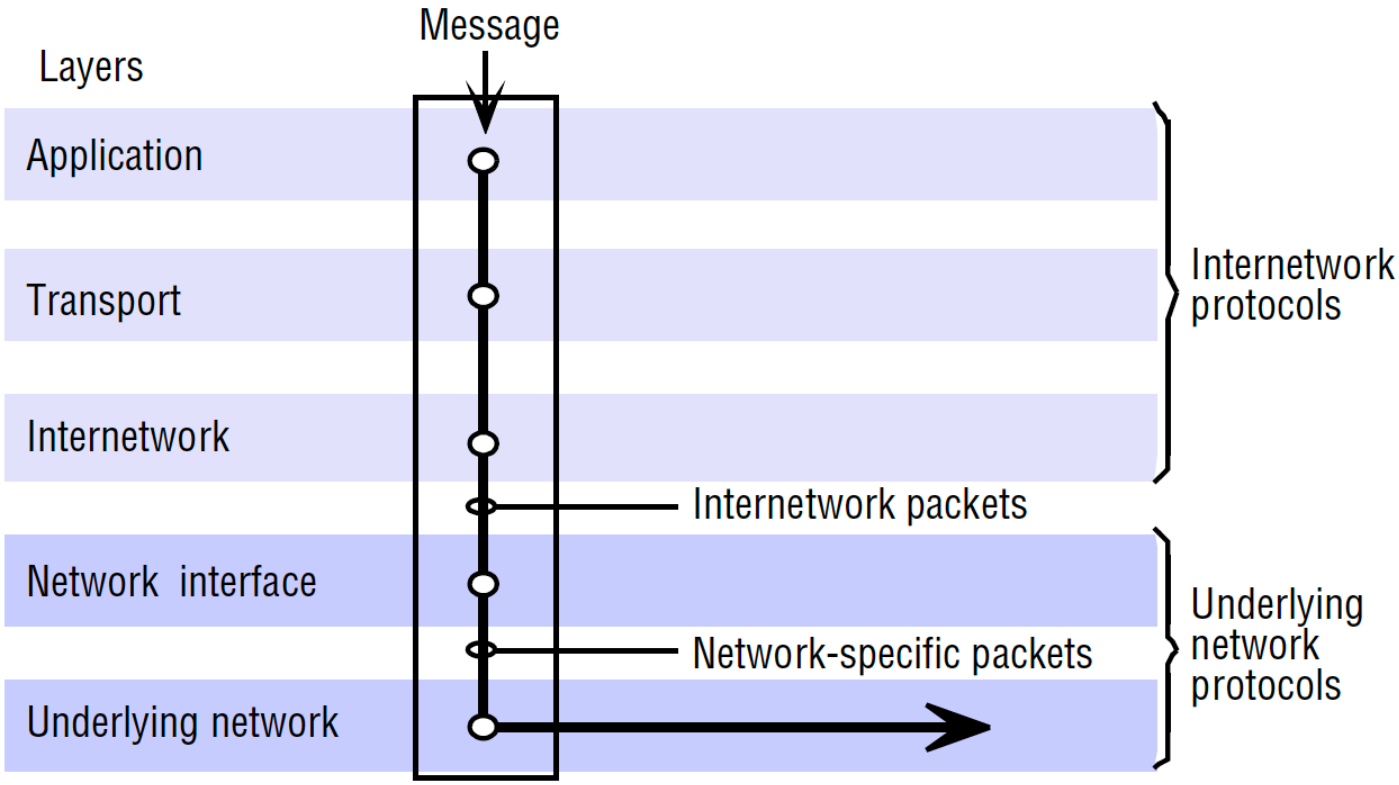
Protocol layers in the ISO Open Systems Interconnection (OSI) model



OSI protocol summary

Layer	*Description*	*Examples*
Application	Protocols that are designed to meet the communication requirements of specific applications, often defining the interface to a service.	HTTP, FTP , SMTP, CORBA IIOP
Presentation	Protocols at this level transmit data in a network representation that is independent of the representations used in individual computers, which may differ. Encryption is also performed in this layer, if required.	Secure Sockets (SSL),CORBA Data Rep.
Session	At this level reliability and adaptation are performed, such as detection of failures and automatic recovery.	
Transport	This is the lowest level at which messages (rather than packets) are handled. Messages are addressed to communication ports attached to processes, Protocols in this layer may be connection-oriented or connectionless.	TCP, UDP
Network	Transfers data packets between computers in a specific network. In a WAN or an internetwork this involves the generation of a route passing through routers. In a single LAN no routing is required.	IP, ATM virtual circuits
Data link	Responsible for transmission of packets between nodes that are directly connected by a physical link. In a WAN transmission is between pairs of routers or between routers and hosts. In a LAN it is between any pair of hosts.	Ethernet MAC, ATM cell transfer, PPP
Physical	The circuits and hardware that drive the network. It transmits sequences of binary data by analogue signalling, using amplitude or frequency modulation of electrical signals (on cable circuits), light signals (on fibre optic circuits) or other electromagnetic signals (on radio and microwave circuits).	Ethernet base- band signalling, ISDN

Internetwork layers



For many distributed systems, the lowest-level interface is that of the network layer

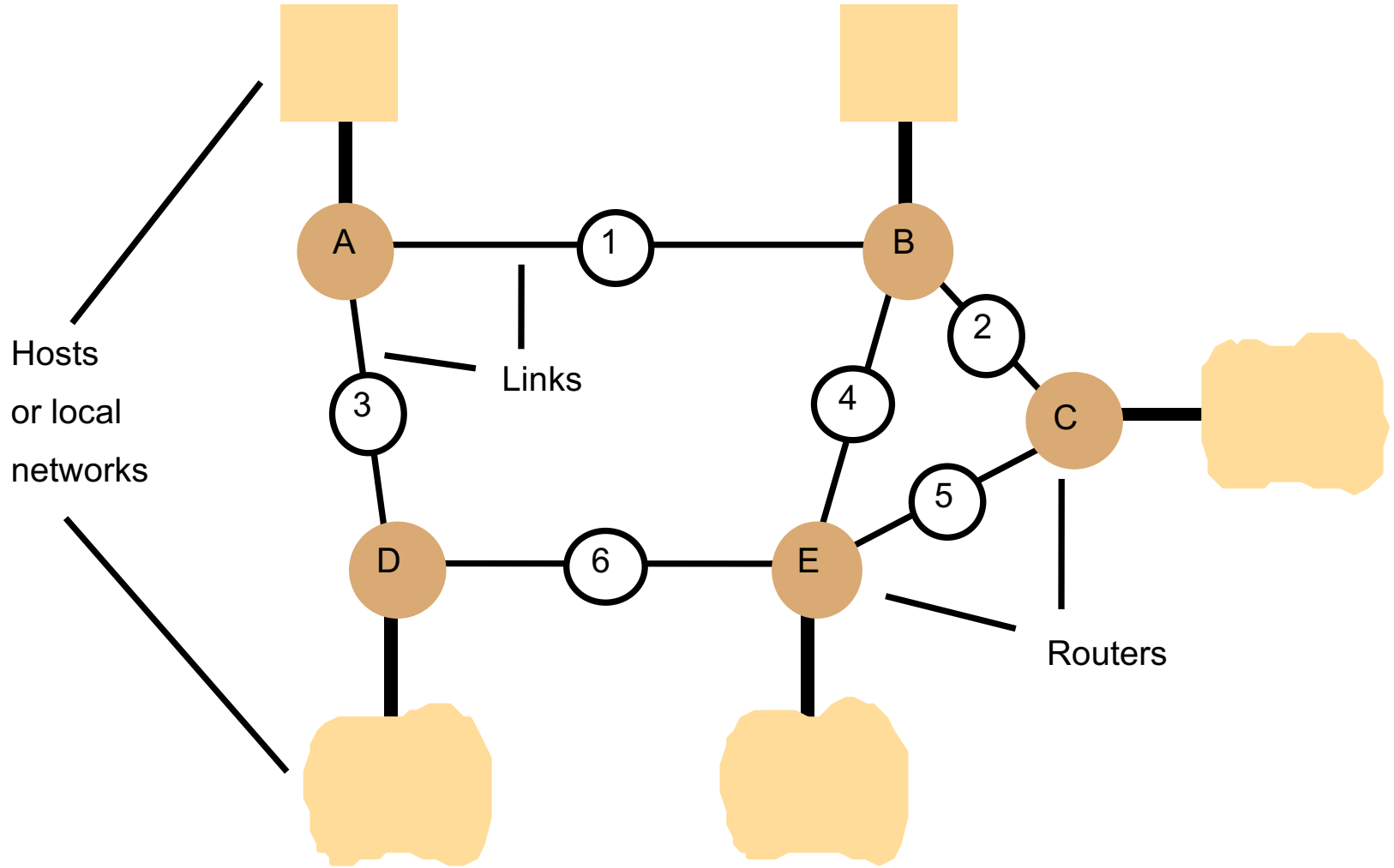
Two approaches to the delivery of packets

Datagram packet delivery: The essential feature of datagram networks is that the delivery of each packet is a 'one-shot' process; no setup is required, and once the packet is delivered the network retains no information about it. In a datagram network a sequence of packets transmitted by a single host to a single destination may follow different routes and when this occurs they may arrive out of sequence.

The Internet's network layer (IP), Ethernet and most wired and wireless local network technologies are based on datagram delivery.

Virtual circuit packet delivery: Some network-level services implement a virtual circuit before packets can pass from a source host A to destination host B. The establishment of a virtual circuit involves the identification of a route from the source to the destination, possibly passing through several intermediate nodes. At each node along the route a table entry is made, indicating which link should be used for the next stage of the route. Once a virtual circuit has been set up, it can be used to transmit any number of packets. The addresses are not needed, because packets are routed at intermediate nodes by reference to the virtual circuit number. When a packet reaches its destination the source can be determined from the virtual circuit number

Example: Routing in a wide area network



Routing algorithms

- The algorithm described in the next slides is a ‘distance vector’ algorithm that is the basis of the link-state algorithm that has been used since 1979 as the main routing algorithm in the Internet.
- Routing in networks is an instance of the problem of path finding in graphs. Bellman’s shortest path algorithm, published before computer networks were developed [Bellman 1957], provides the basis for the distance vector method.
- Bellman’s method was converted into a distributed algorithm suitable for implementation in large networks by Ford and Fulkerson [1962], and protocols based on their work are often referred to as ‘Bellman–Ford’ protocols

The distance vector routing algorithm

- Each row in the routing tables provides the routing information for packets addressed to a given destination.
- The **link** field specifies the outgoing link for packets addressed to the destination.
- The **cost** field is a calculation of the vector distance, or the number of hops to the given destination. For store-and-forward networks with links of similar bandwidth, this gives an estimate of the time for a packet to travel to the destination.
- The cost information stored in the routing tables is not used during packet-routing actions taken by part 1 of the routing algorithm, but it is required for the routing table construction and maintenance actions in part 2.

Routing tables for the network in Figure: Example Routing

Routings from A

To	*Link*	*Cost*
A	local	0
B	1	1
C	1	2
D	3	1
E	1	2

Routings from B

To	*Link*	*Cost*
A	1	1
B	local	0
C	2	1
D	1	2
E	4	1

Routings from C

To	*Link*	*Cost*
A	2	2
B	2	1
C	local	0
D	5	2
E	5	1

Routings from D

To	*Link*	*Cost*
A	3	1
B	3	2
C	6	2
D	local	0
E	6	1

Routings from E

To	*Link*	*Cost*
A	4	2
B	4	1
C	5	1
D	6	1
E	local	0

Pseudo-code for RIP routing algorithm

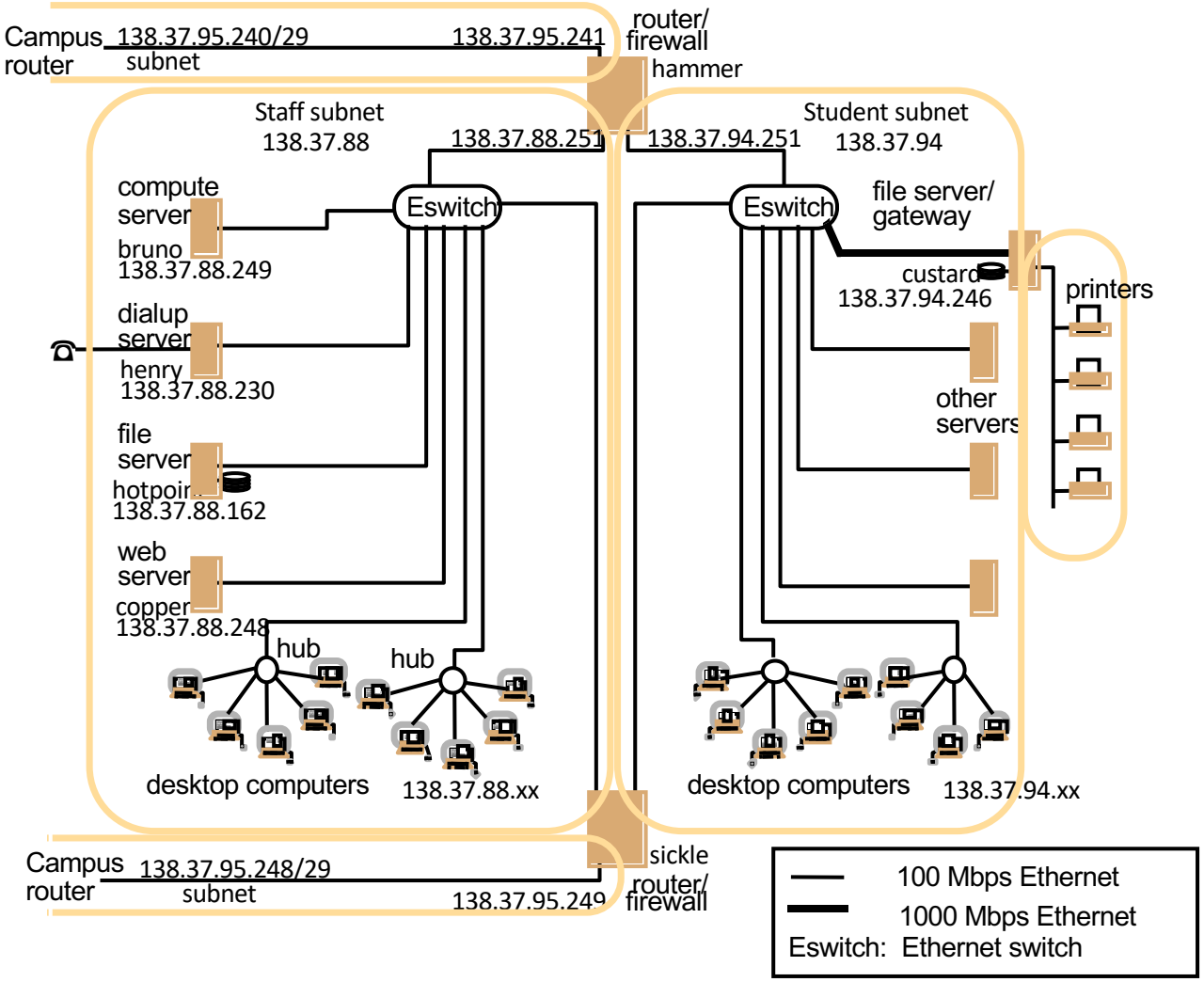
Send: Each *t* seconds or when *Tl* changes, send *Tl* on each non-faulty outgoing link.
Receive: Whenever a routing table *Tr* is received on link *n*:

```
for all rows Rr in Tr {
    if (Rr.link ≠ n) {
        Rr.cost = Rr.cost + 1;
        Rr.link = n;
        if (Rr.destination is not in Tl) add Rr to Tl;    // add new destination to Tl
        else for all rows Rl in Tl {
            if (Rr.destination = Rl.destination and
                    (Rr.cost < Rl.cost or Rl.link = n)) Rl = Rr;
                    // Rr.cost < Rl.cost : remote node has better route
                    // Rl.link = n : remote node is more authoritative
        }
    }
}
```

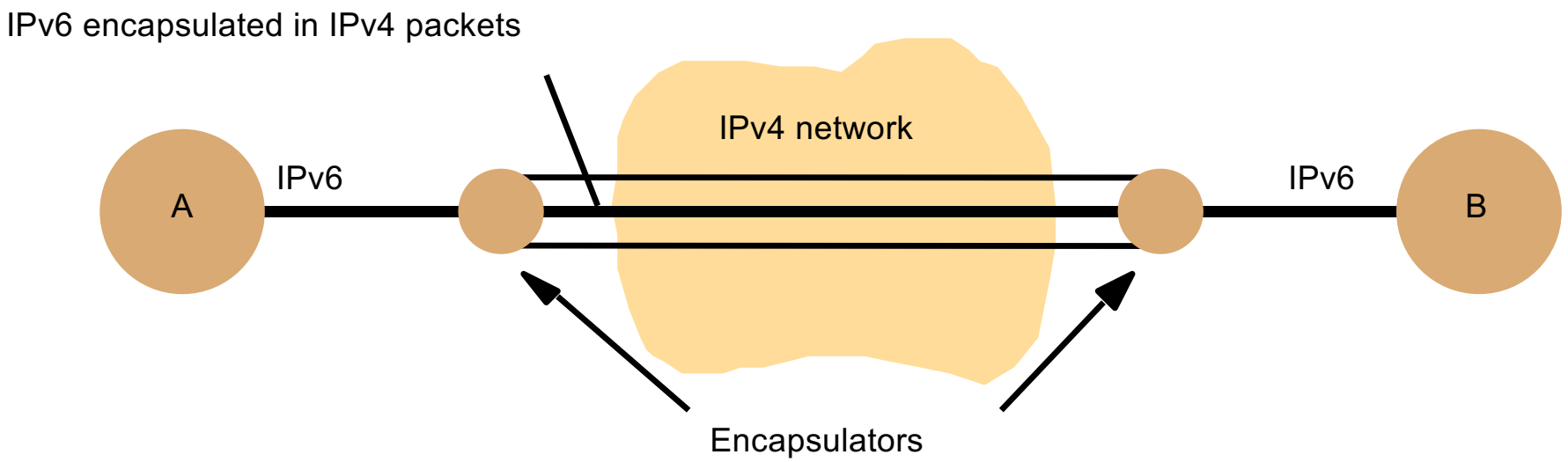
Congestion control

- The capacity of a network is limited by the performance of its communication links and switching nodes.
- When the load at any link or node approaches its capacity, queues will build up at hosts trying to send packets and at intermediate nodes holding packets whose onward transmission is blocked by other traffic.
- If the load continues at the same high level, the queues will continue to grow until they reach the limit of available buffer space. Once this state is reached at a node, the node has no option but to drop further incoming packets.
- the occasional loss of packets at the network level is acceptable and can be remedied by retransmission initiated at higher levels. But if the rate of packet loss and retransmission reaches a substantial level, the effect on the throughput of the network can be devastating

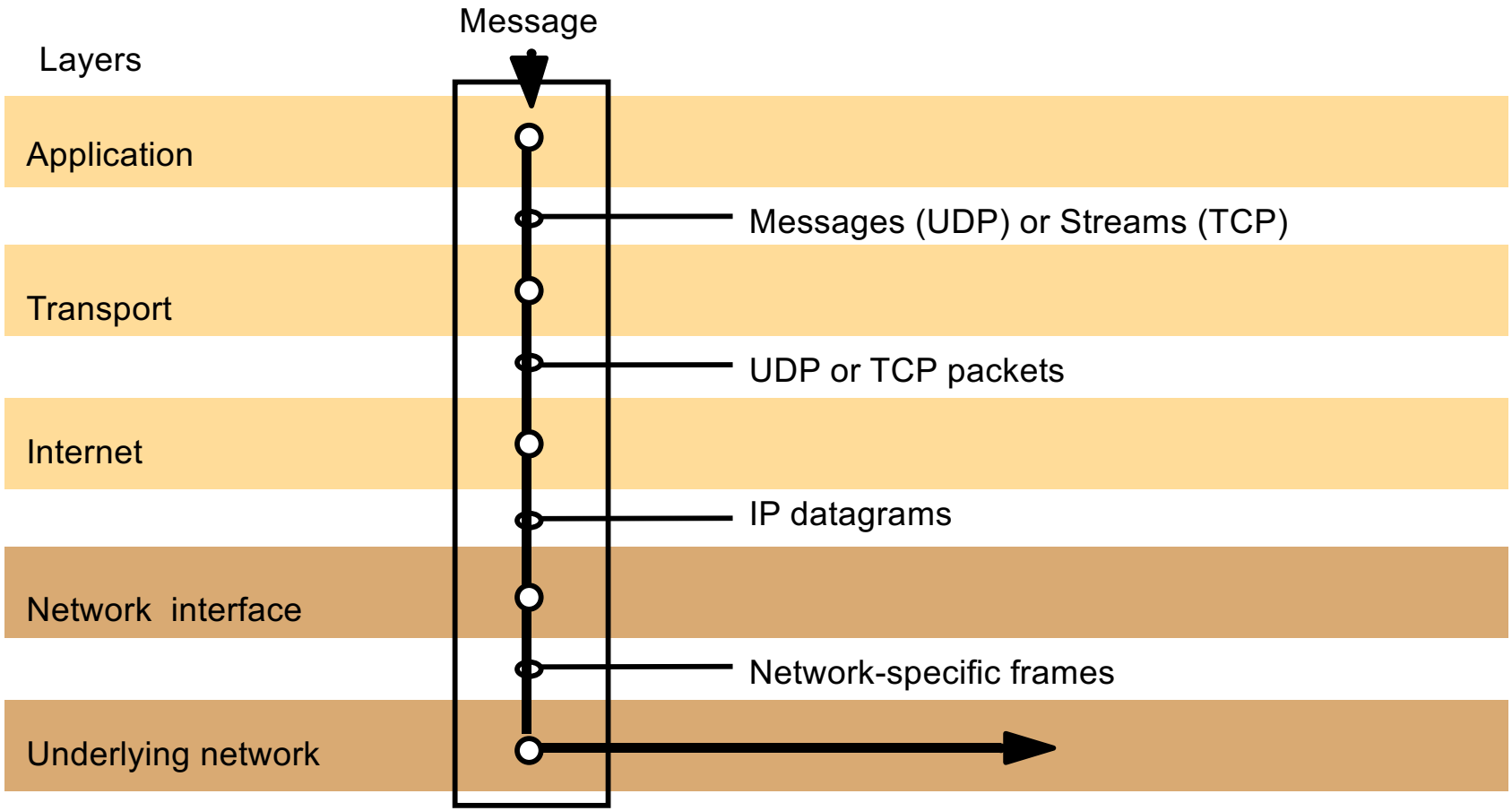
Simplified view of part of a university campus network



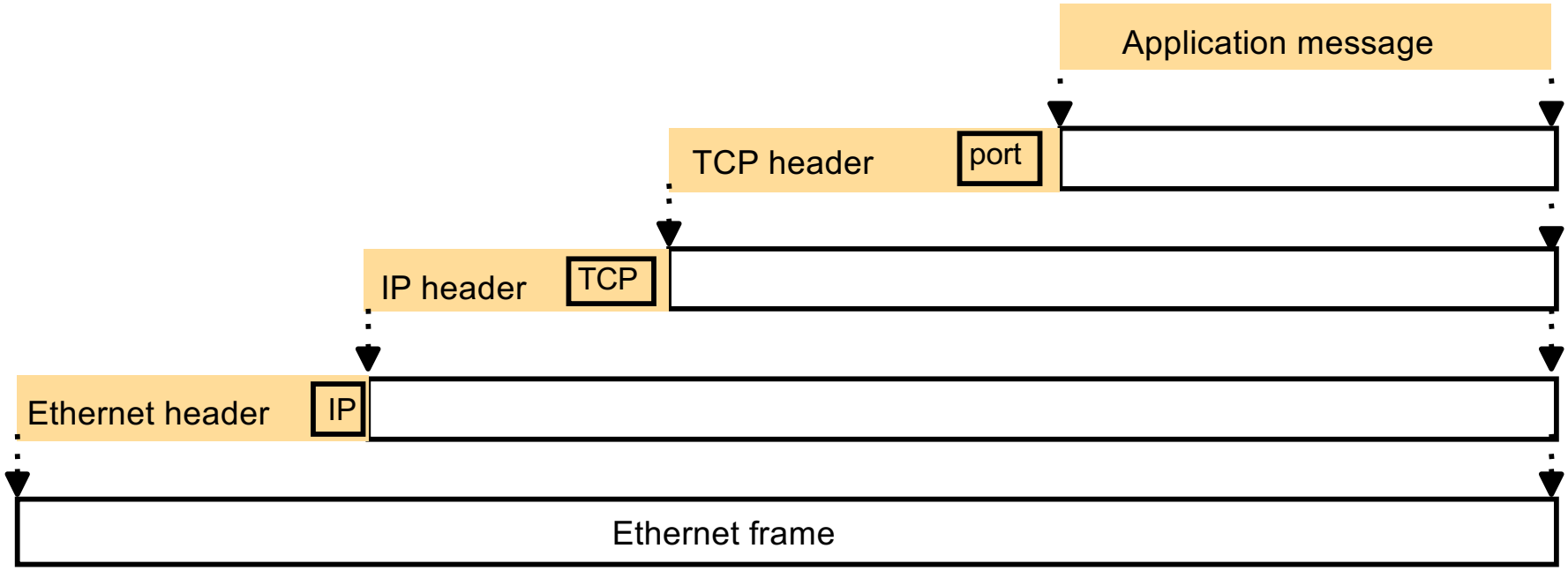
Tunnelling for IPv6 migration



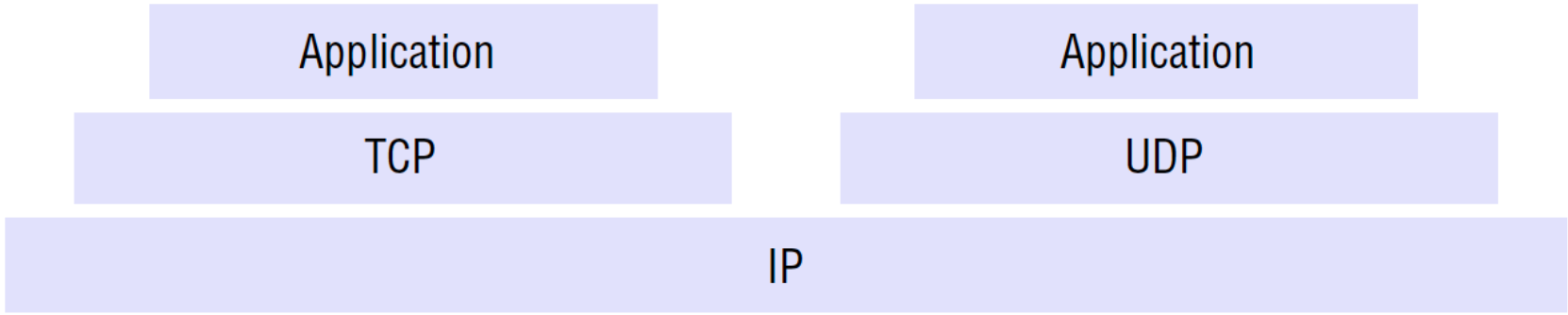
TCP/IP layers



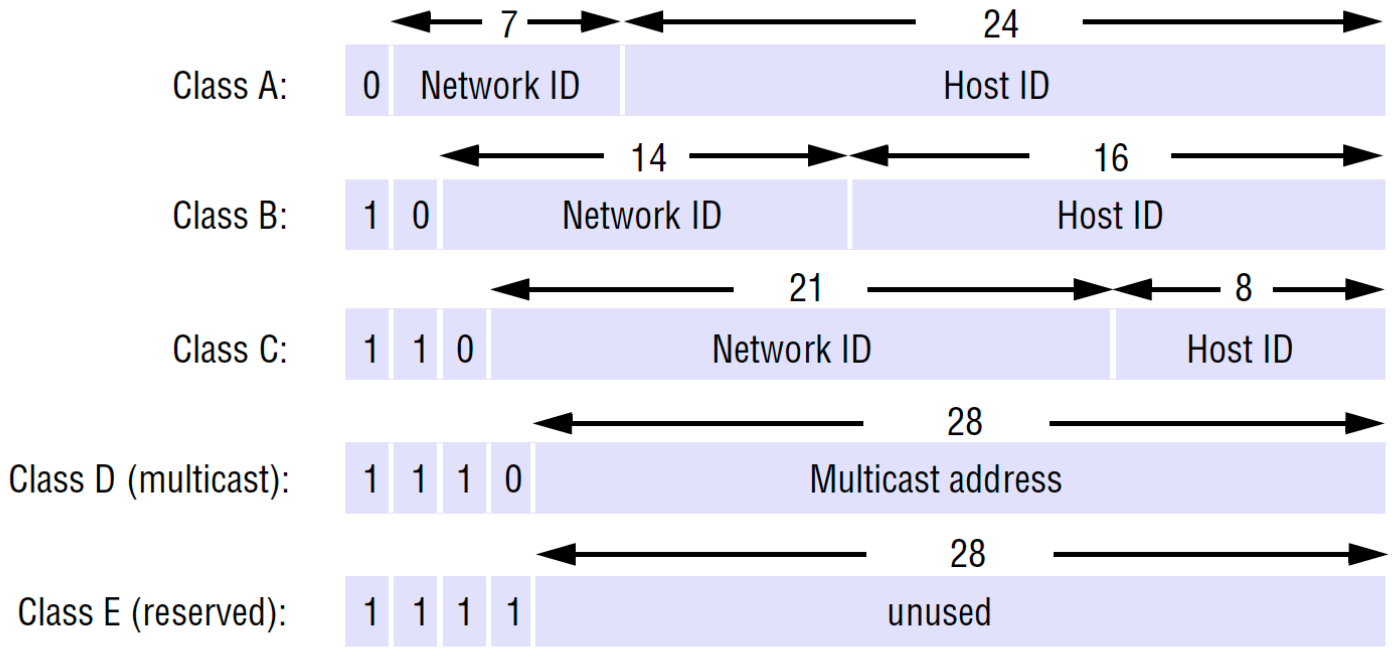
Encapsulation in a message transmitted via TCP over an Ethernet



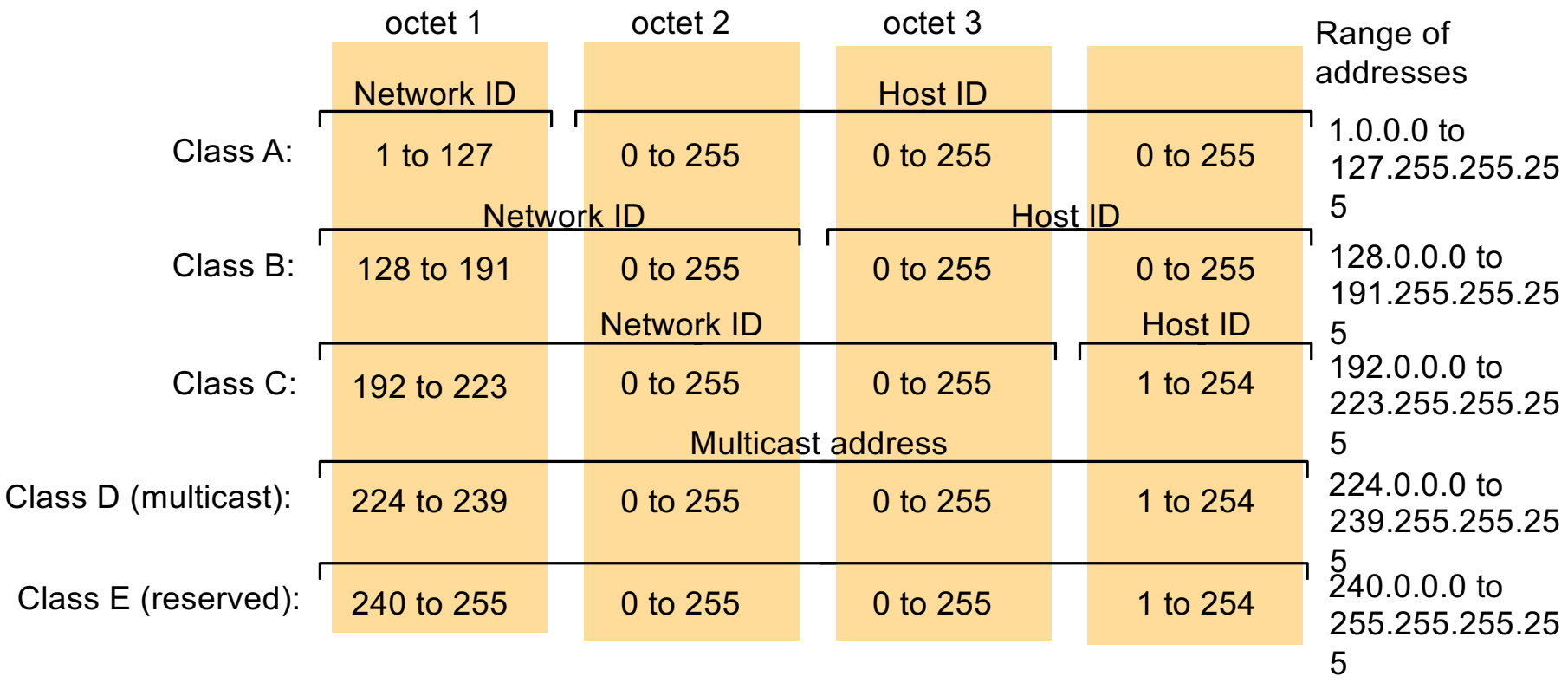
The programmer's conceptual view of a TCP/IP Internet



Internet address structure, showing field sizes in bits



Decimal representation of Internet addresses



Internet addresses

- Internet addresses with host identifiers 0 and all 1s (binary) are used for special purposes. Addresses with the host identifier set to 0 are used to refer to ‘this host’, and a host identifier that is all 1s is used to address a broadcast message to all of the hosts connected to the network specified in the network identifier part of the address.

- Network identifiers are allocated by the Internet Assigned Numbers Authority (IANA) to organizations with networks connected to the Internet. Host identifiers for the computers on each network connected to the Internet are assigned by the managers of the relevant networks.

- Since host addresses include a network identifier, any computer that is connected to more than one network must have separate addresses on each, and whenever a computer is moved to a different network, its Internet address must change. These requirements can lead to substantial administrative overheads, for example in the case of portable computers.

From IP to IPv6

the IP address allocation scheme has not turned out to be very effective.

The main difficulty is that network administrators in user organizations cannot easily predict future growth in their need for host addresses, and they tend to overestimate, requesting Class B addresses when in doubt.

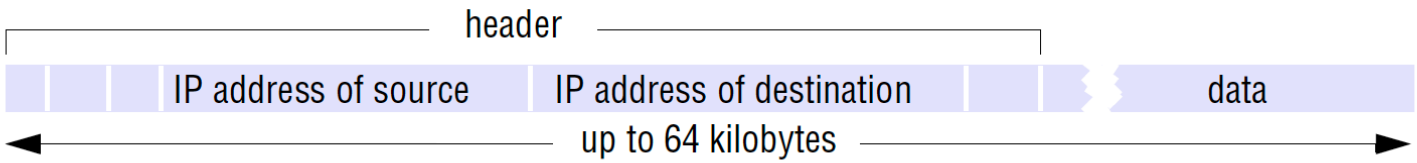
Around 1990 it became evident that based on the rate of allocation at the time, IP addresses were likely to run out around 1996: three steps were taken.

1.The first was to initiate the development of a new IP protocol and addressing scheme, the result of which was the specification of IPv6.

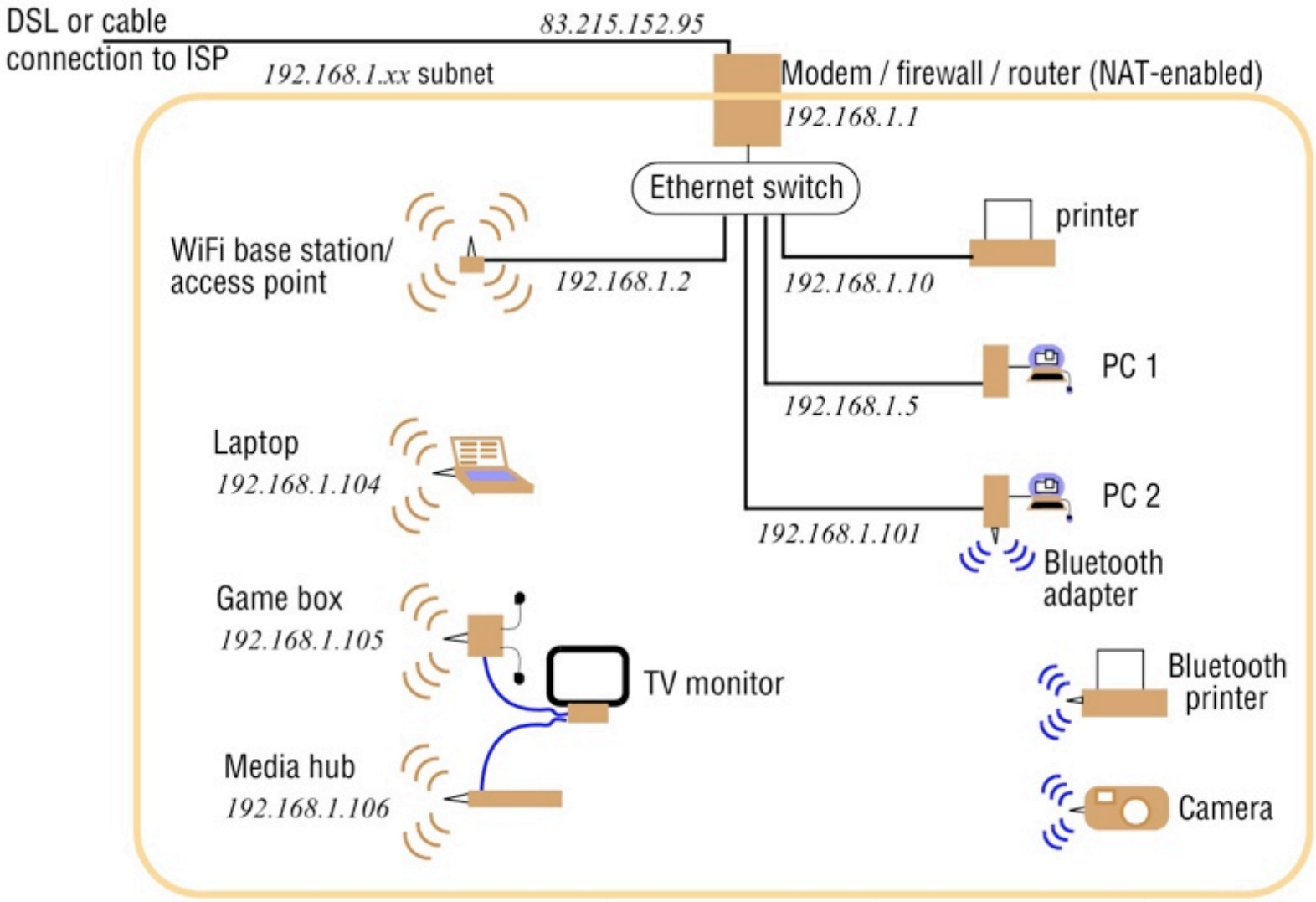
2.The second step was to radically modify the way in which IP addresses were allocated. A new address allocation and routing scheme designed to make more effective use of the IP address space was introduced, called classless interdomain routing (CIDR).

3.The third step was to enable unregistered computers to access the Internet indirectly through routers that implement a Network Address Translation (NAT) scheme.

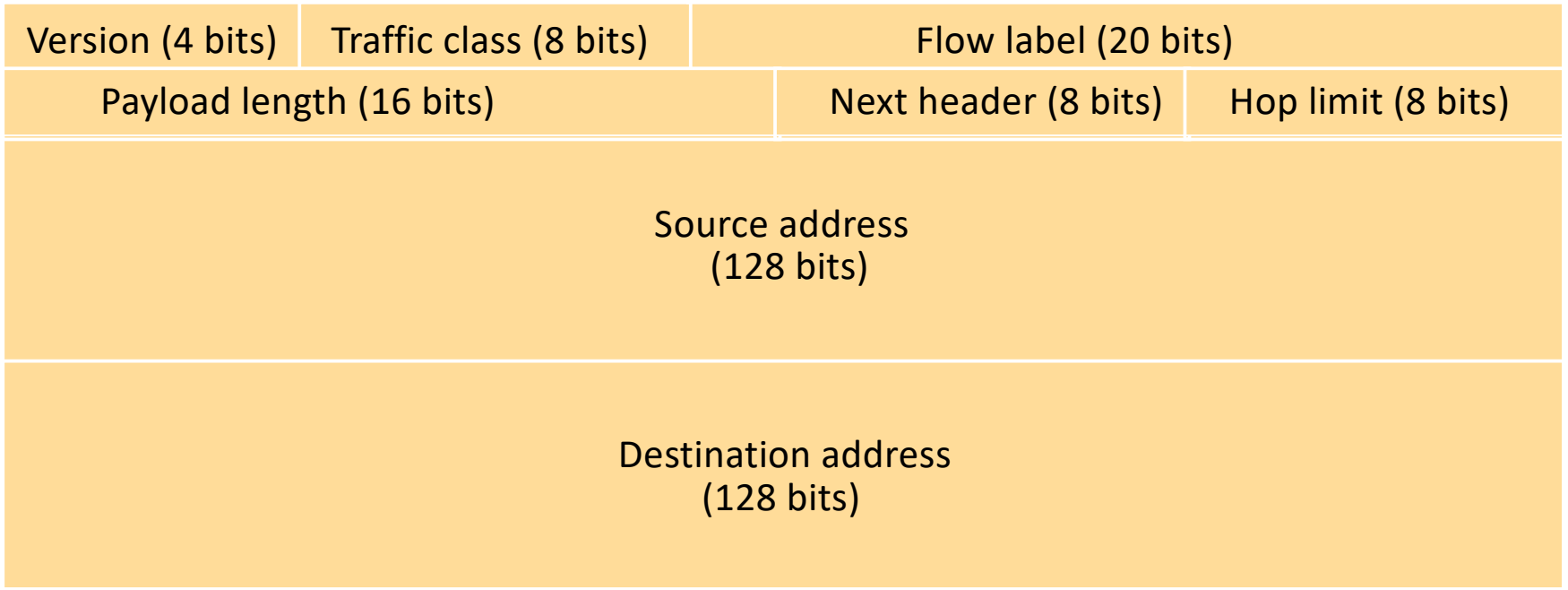
IP packet layout: main components



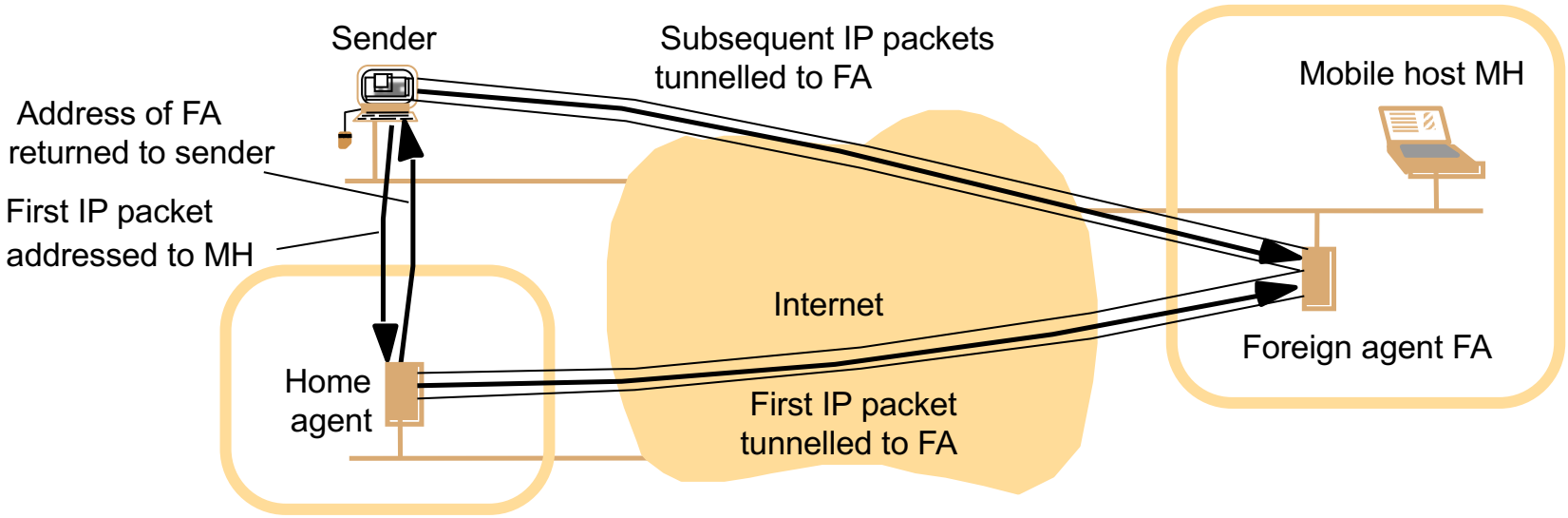
A typical NAT-based home network



IPv6 header layout



The MobileIP routing mechanism

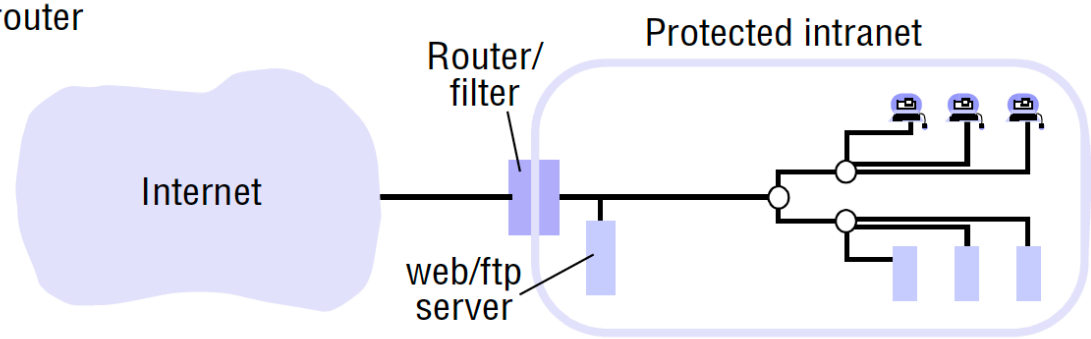


When the mobile host is connected at its home base, packets are routed in the normal way. When it is connected to the Internet elsewhere, two agent processes take responsibility for rerouting: the agents are a home agent (HA) and a foreign agent (FA).

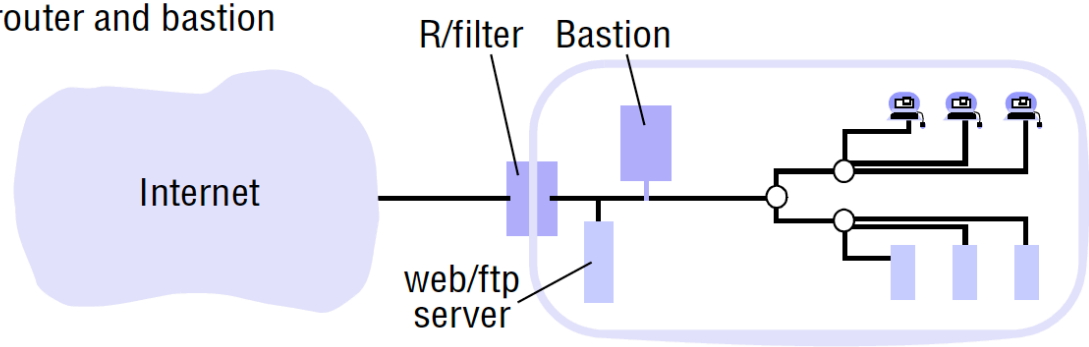
When an IP packet addressed to the mobile host 's home address is received at the home network, it is routed to the HA. The HA then encapsulates the IP packet in a MobileIP packet and sends it to the FA. The FA unpacks the original IP packet and delivers it to the mobile host via the local network to which it is currently attached.

Firewall configurations

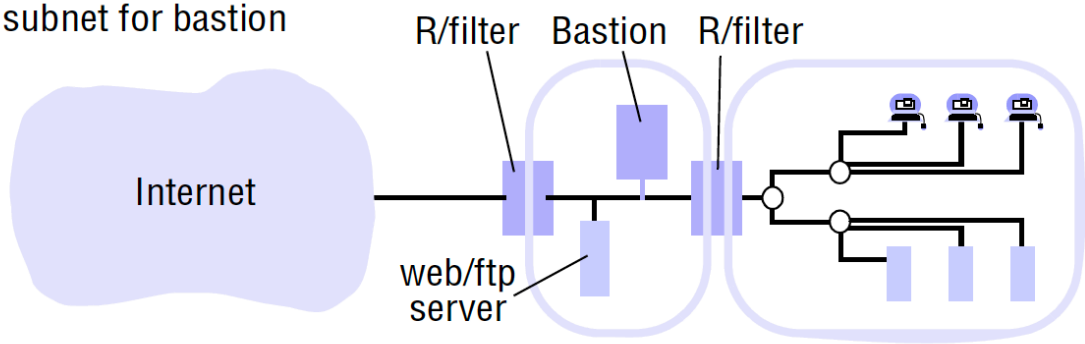
a) Filtering router



b) Filtering router and bastion



c) Screened subnet for bastion



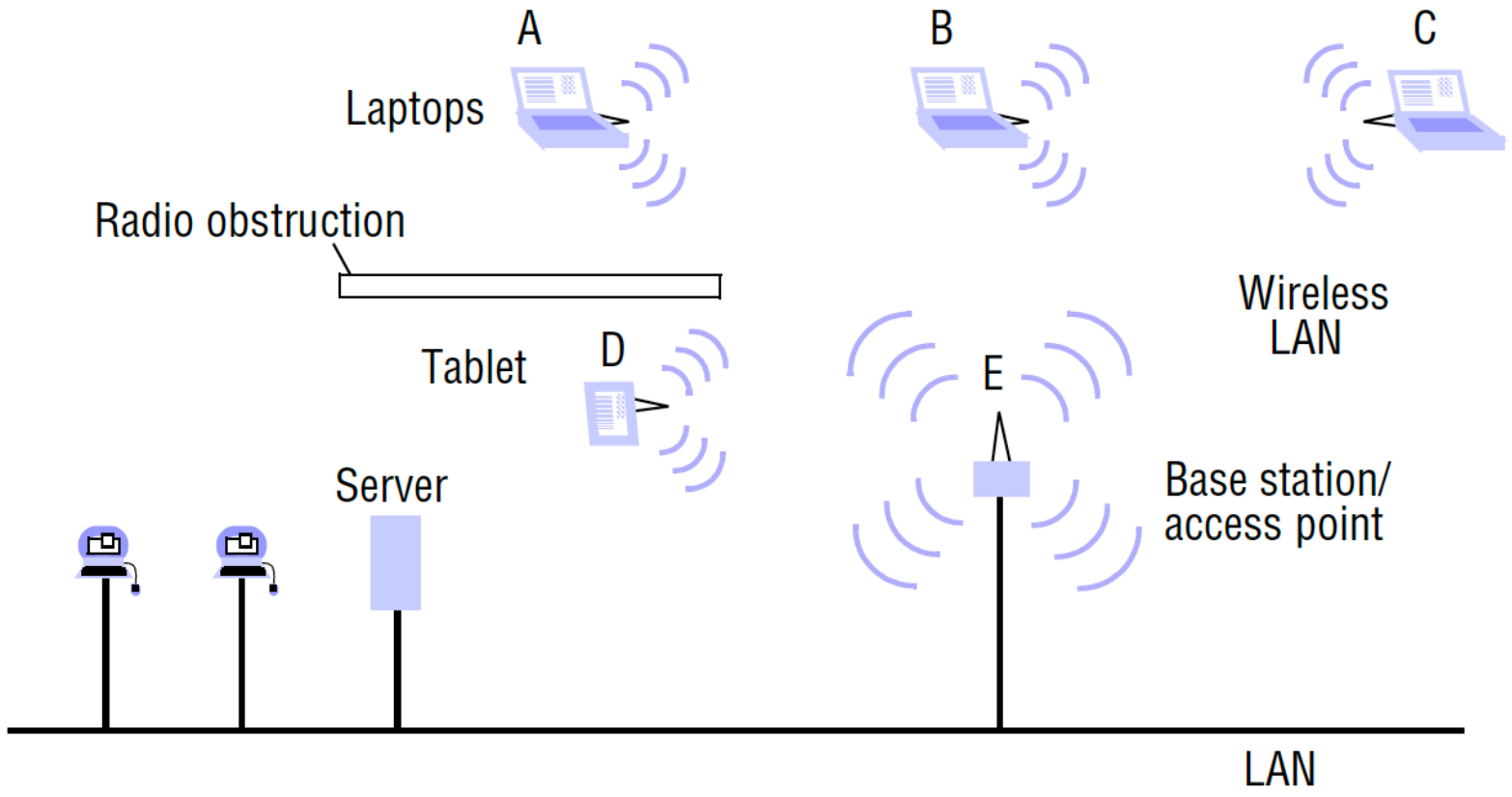
IEEE 802 network standards

IEEE No.	*Name*	*Title*	*Reference*
802.3	Ethernet	CSMA/CD Networks (Ethernet)	[IEEE 1985a]
802.4		Token Bus Networks	[IEEE 1985b]
802.5		Token Ring Networks	[IEEE 1985c]
802.6		Metropolitan Area Networks	[IEEE 1994]
802.11	WiFi	Wireless Local Area Networks	[IEEE 1999]
802.15.1	Bluetooth	Wireless Personal Area Networks	[IEEE 2002]
802.15.4	ZigBee	Wireless Sensor Networks	[IEEE 2003]
802.16	WiMAX	Wireless Metropolitan Area Networks	[IEEE 2004a]

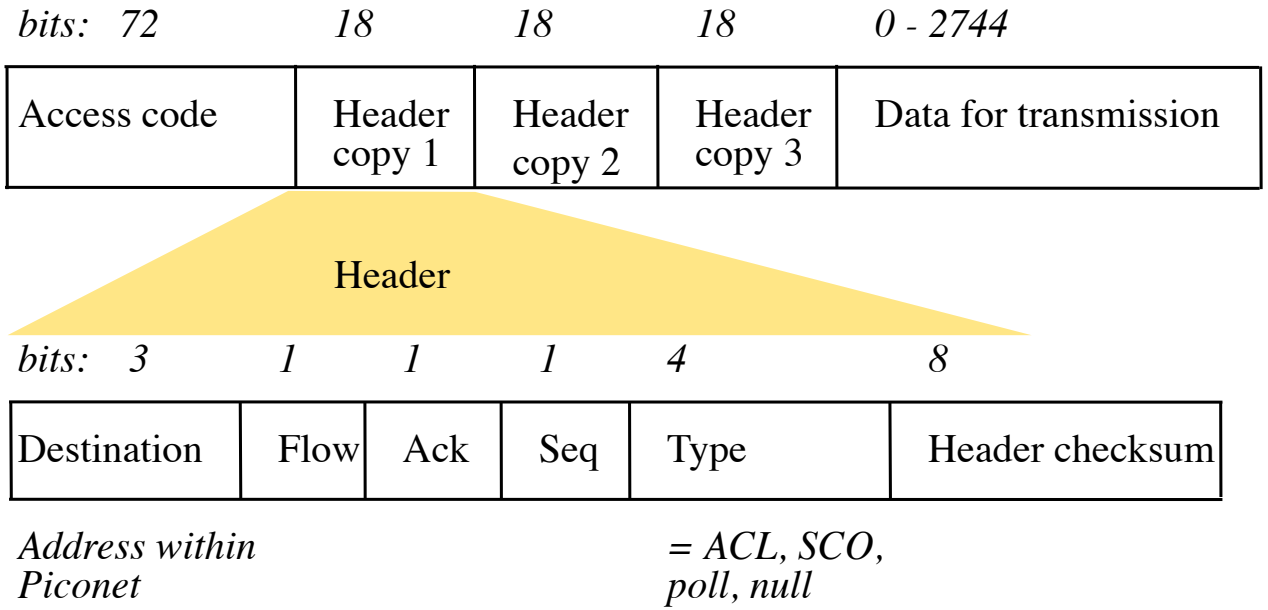
Ethernet ranges and speeds

	10Base5	*10BaseT*	*100BaseT*	*1000BaseT*
Data rate	10 Mbps	10 Mbps	100 Mbps	1000 Mbps
Max. segment lengths:				
Twisted wire (UTP)	100 m	100 m	100 m	25 m
Coaxial cable (STP)	500 m	500 m	500 m	25 m
Multi-mode fibre	2000 m	2000 m	500 m	500 m
Mono-mode fibre	25000 m	25000 m	20000 m	2000 m

Wireless LAN configuration



Bluetooth frame structure



SCO packets (e.g. for voice data) have a 240-bit payload containing 80 bits of data triplicated, filling exactly one timeslot.

Conclusions

- We have summarized the networking principles that are needed as a basis for distributed systems, approaching them from the point of view of a distributed system designer.
- Packet networks and layered protocols provide the basis for communication in distributed systems.
- Local area networks are based on packet broadcasting on a shared medium; Ethernet is the dominant technology
- Wide area networks are based on packet switching to route packets to their destinations through a connected network
- Routing is a key mechanism and a variety of routing algorithms are used, of which the distance-vector method is the most basic but effective.
- Congestion control is needed to prevent overflow of buffers at the receiver and at intermediate nodes.

Exercise

Compare connectionless (UDP) and connection-oriented (TCP) communication for the implementation of each of the following application-level or presentation-level protocols:

i. virtual terminal access (for example, Telnet);

ii. file transfer (for example, FTP);

iii. user location (for example, rwho, finger);

iv. information browsing (for example, HTTP);

v. remote procedure call.

Answer

i. The long duration of sessions, the need for reliability and the unstructured sequences of characters transmitted make connection-oriented communication most suitable for this application. Performance is not critical in this application, so the overheads are of little consequence.

ii. File calls for the transmission of large volumes of data. Connectionless would be ok if error rates are low and the messages can be large, but in the Internet, these requirements aren't met, so TCP is used.

iii. Connectionless is preferable, since messages are short, and a single message is sufficient for each transaction.

iv. Either mode could be used. The volume of data transferred on each transaction can be quite large, so TCP is used in practice.

v. RPC achieves reliability by means of timeouts and re-trys. so connectionless (UDP) communication is often preferred.