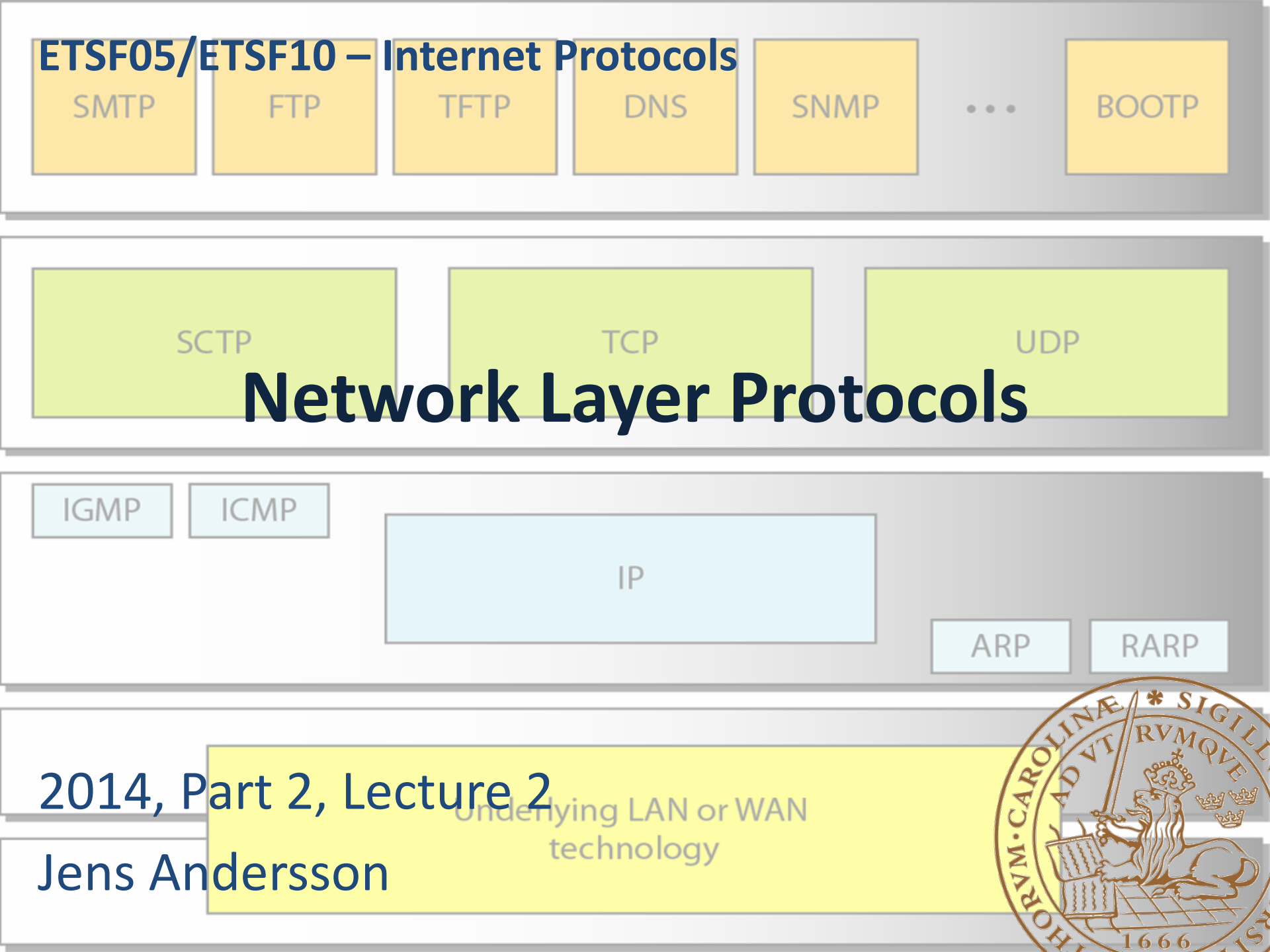
"The requirements for a future all-digital-data distributed network which provides common user service for a wide range of users having different requirements is considered. The use of a standard format message block permits building relatively simple switching mechanisms using an adaptive store-and-forward routing policy to handle all forms of digital data including "real-time" voice. This network rapidly responds to changes in network status."

—On Distributed Communications,
Rand Report RM-3420-PR,
Paul Baran, August 1964

ETSF05/ETSF10 – Internet Protocols



Network Layer Protocols

2014, Part 2, Lecture 2

Jens Andersson

Communication Network
A facility that provides a data transfer service among devices attached to the network.

Internet
A collection of communication networks interconnected by bridges and/or routers.

Intranet
An internet used by a single organization that provides the key Internet applications, especially the World Wide Web. An intranet operates within the organization for internal purposes and can exist as an isolated, self-contained internet, or may have links to the Internet.

Subnetwork
Refers to a constituent network of an internet. This avoids ambiguity because the entire internet, from a user's point of view, is a single network.

End System (ES)
A device attached to one of the networks of an internet that is used to support end-user applications or services.

Intermediate System (IS)
A device used to connect two networks and permit communication between end systems attached to different networks.

Bridge
An IS used to connect two LANs that use similar LAN protocols. The bridge acts as an address filter, picking up packets from one LAN that are intended for a destination on another LAN and passing those packets on. The bridge does not modify the contents of the packets and does not add anything to the packet. The bridge operates at layer 2 of the OSI model.

Router
An IS used to connect two networks that may or may not be similar. The router employs an internet protocol present in each router and each end system of the network. The router operates at layer 3 of the OSI model.

Table 14.1

Internetworking Terms (use as reference)

(Table is on page 453 in the textbook)

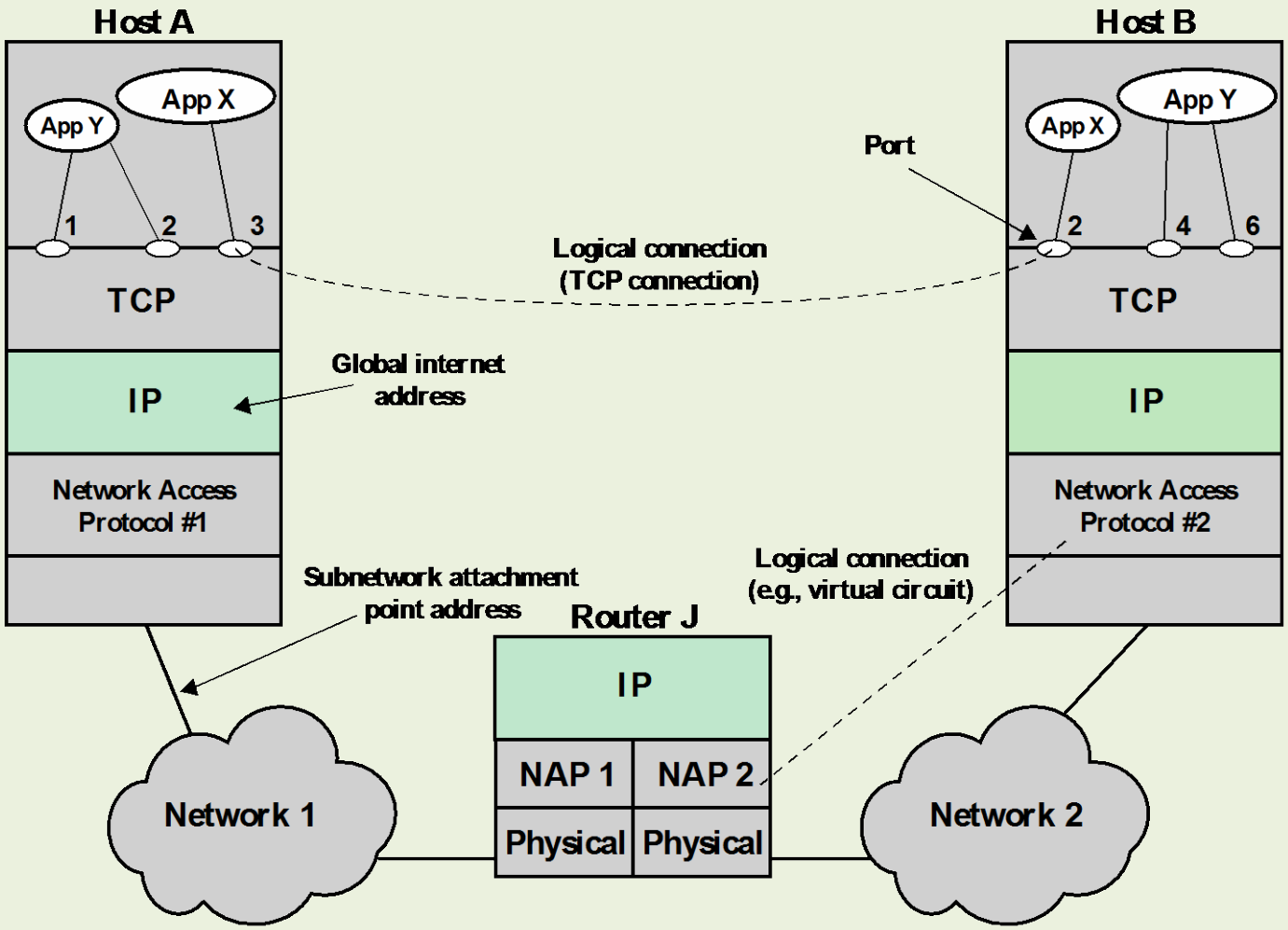


Figure 14.1 TCP/IP Concepts

Connectionless Operation

- Internetworking involves connectionless operation at the level of the Internet Protocol (IP)

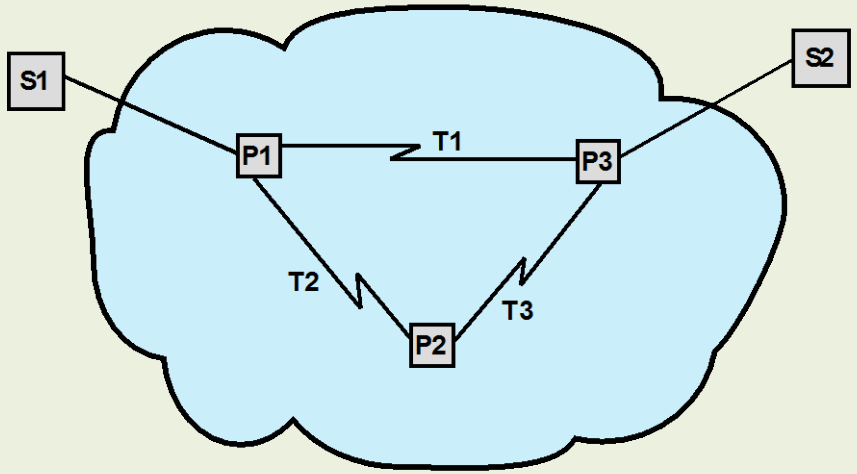
- Initially developed for the DARPA internet project
- Protocol is needed to access a particular network

Connectionless Internetworking

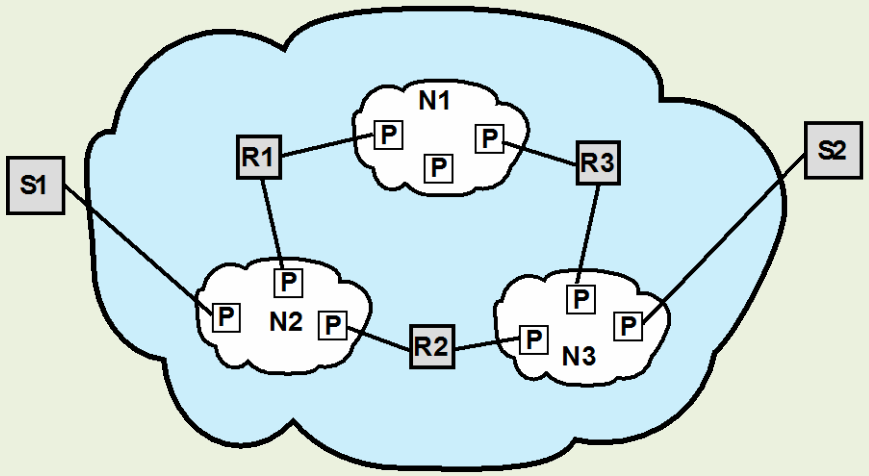
- IP provides a connectionless service between end systems
- Advantages:
  - Is flexible
  - Can be made robust
  - Does not impose unnecessary overhead

IP Design Issues

- Routing
- Datagram lifetime
  - TTL
- **Fragmentation and reassembly**
- Error control
- Flow control



(a) Packet-switching network architecture



(b) Internetwork architecture

Figure 14.3 The Internet as a Network

Internetworking is connecting packet-switching networks!

Fragmentation and Re-assembly

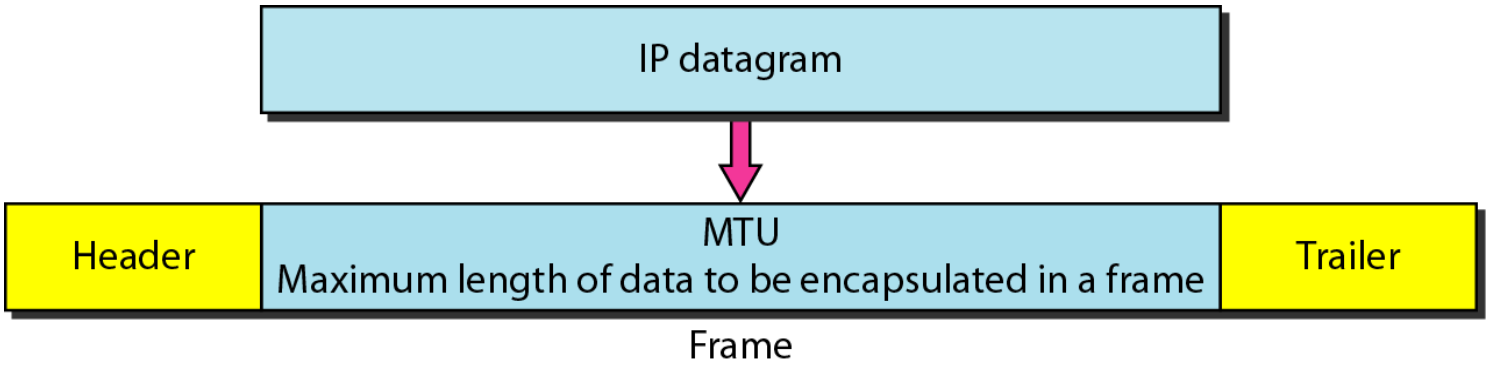
- Protocol exchanges data between two entities
- Lower-level protocols may need to break data up into smaller blocks, called fragmentation
- Reasons for fragmentation:
  - Network only accepts blocks of a certain size
  - Moreefficient error control and smallerretransmission units
  - Fairer access to shared facilities
  - Smaller buffers
- Disadvantages:
  - Smaller buffers
  - More interrupts and processing time

Fragmentation

- Needed when IP datagram size > MTU
- IPv4
  - Performed by the router meeting the problem
- IPv6
  - Performed by the source router only
- Defragmentation by destination host

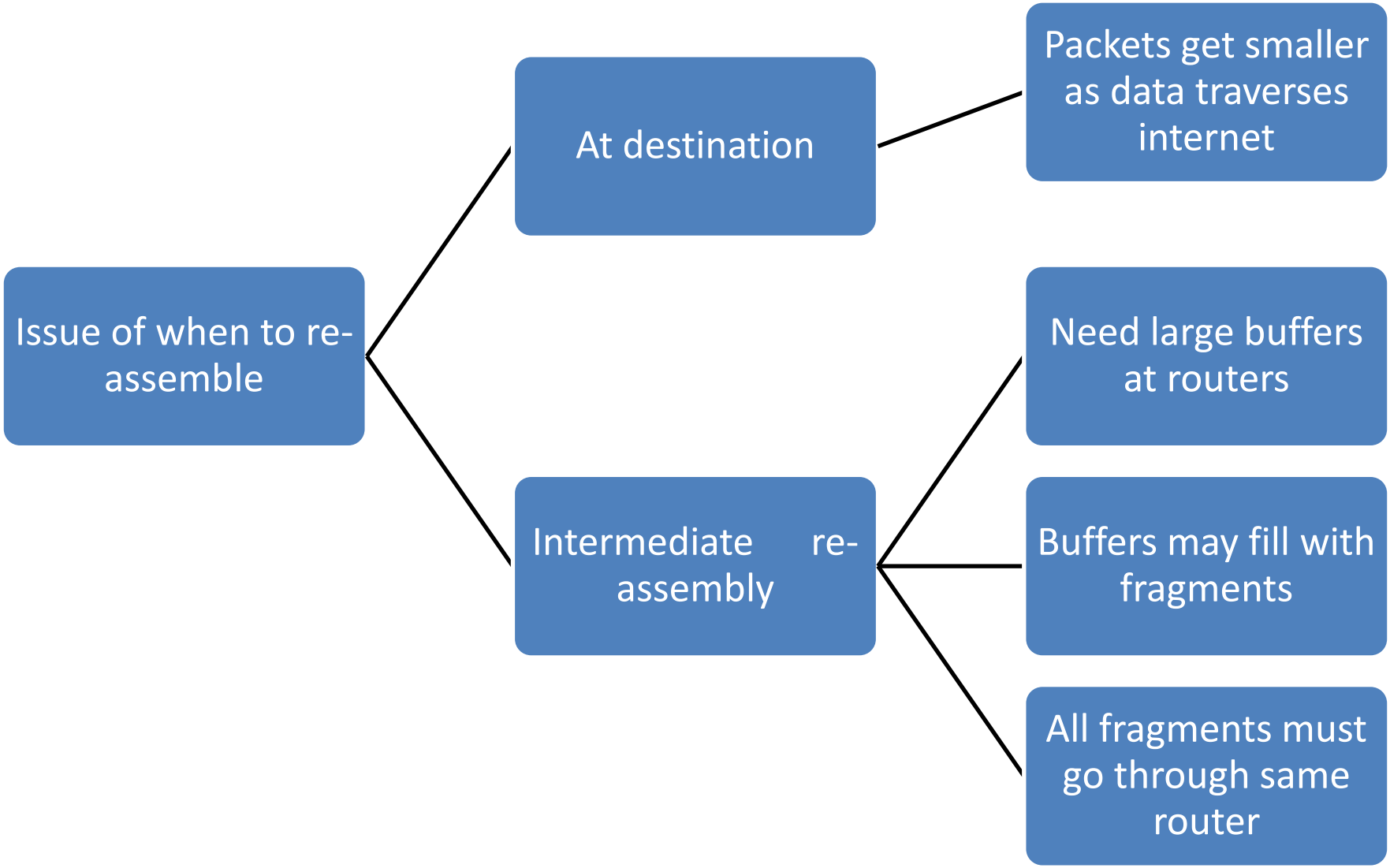
D: Do not fragment
M: More fragments

Maximum datagram size

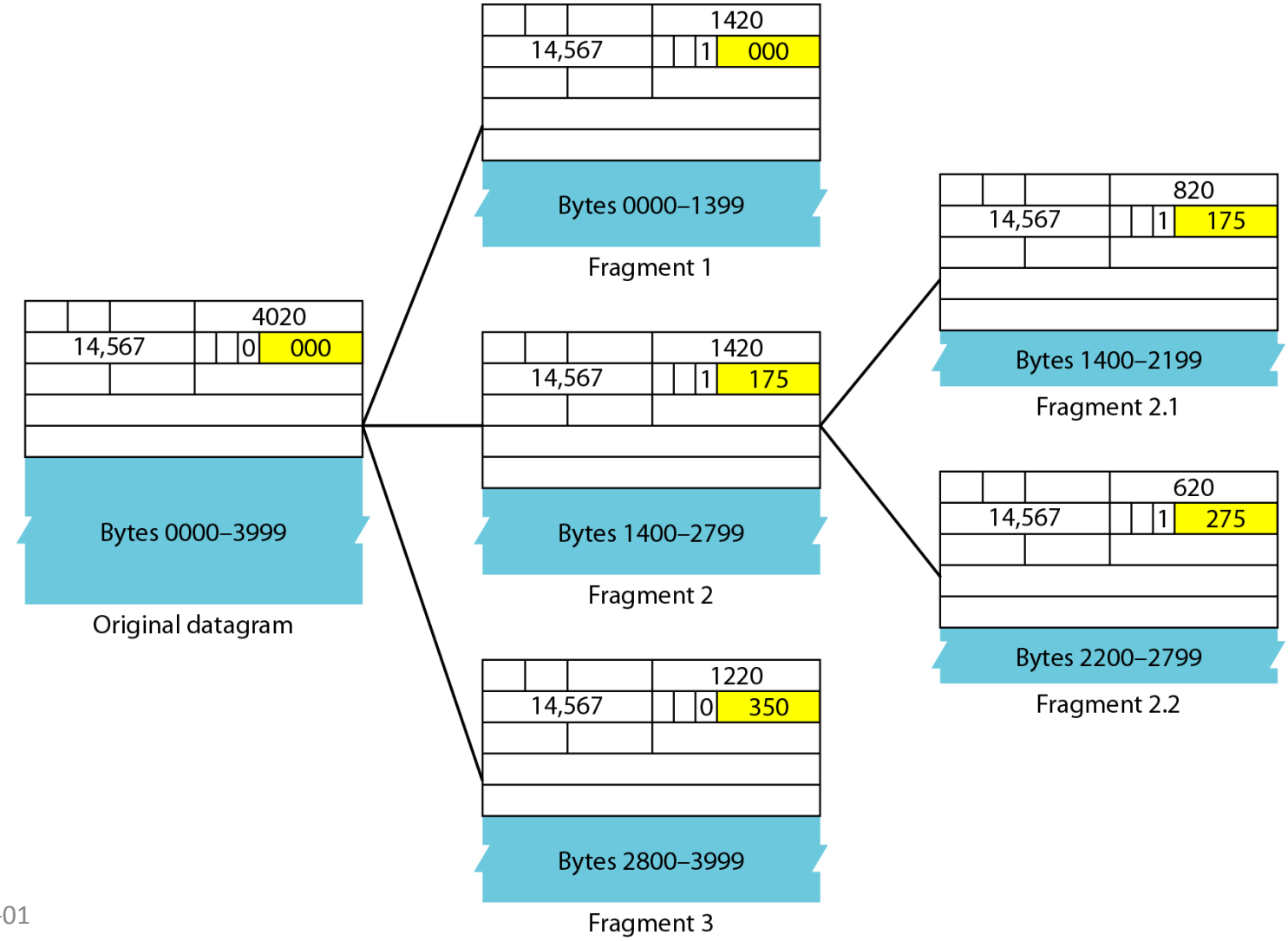


Protocol	*MTU*
Hyperchannel	65,535
Token Ring (16 Mbps)	17,914
Token Ring (4 Mbps)	4,464
FDDI	4,352
Ethernet	1,500
X.25	576
PPP	296

Fragmentation Re-assembly

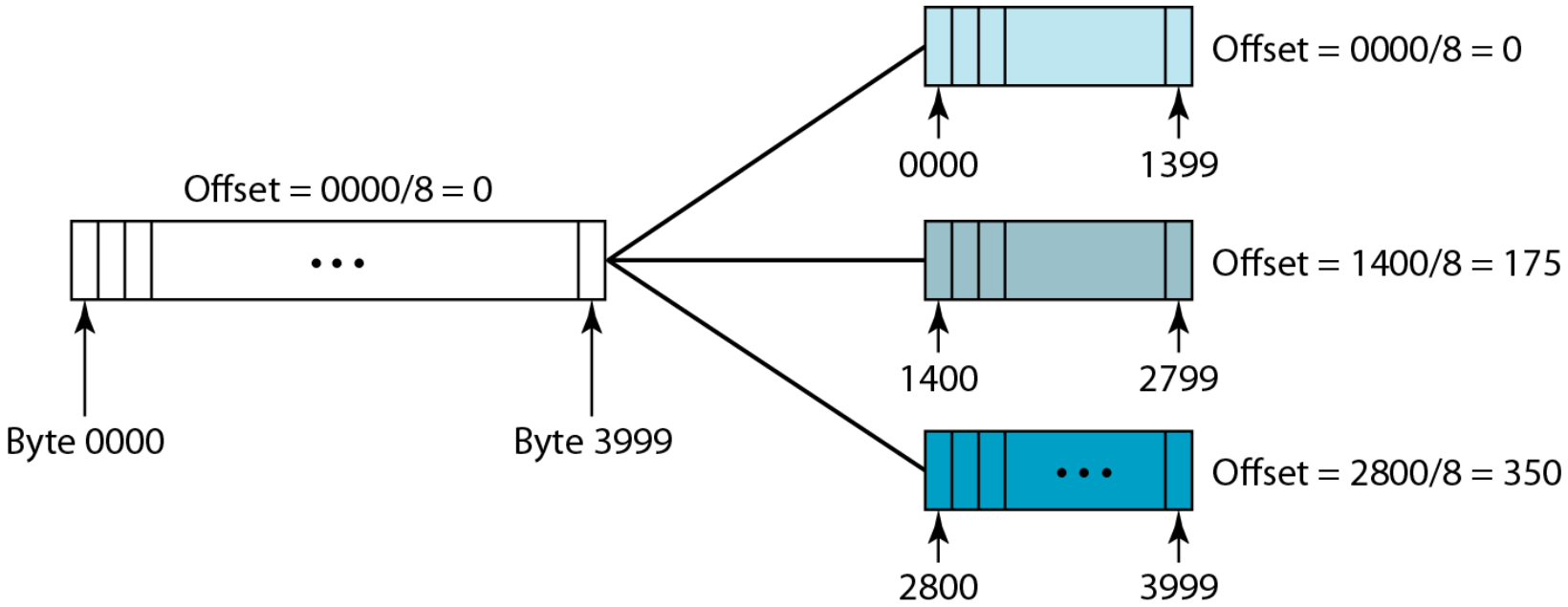


Fragmentation example



Fragmentation offset

- Relative location of fragments
- 13 bits < 16 bits → /8



What with TCP/UDP header?

- Where is a TCP or UDP header in fragments?
- Problem?

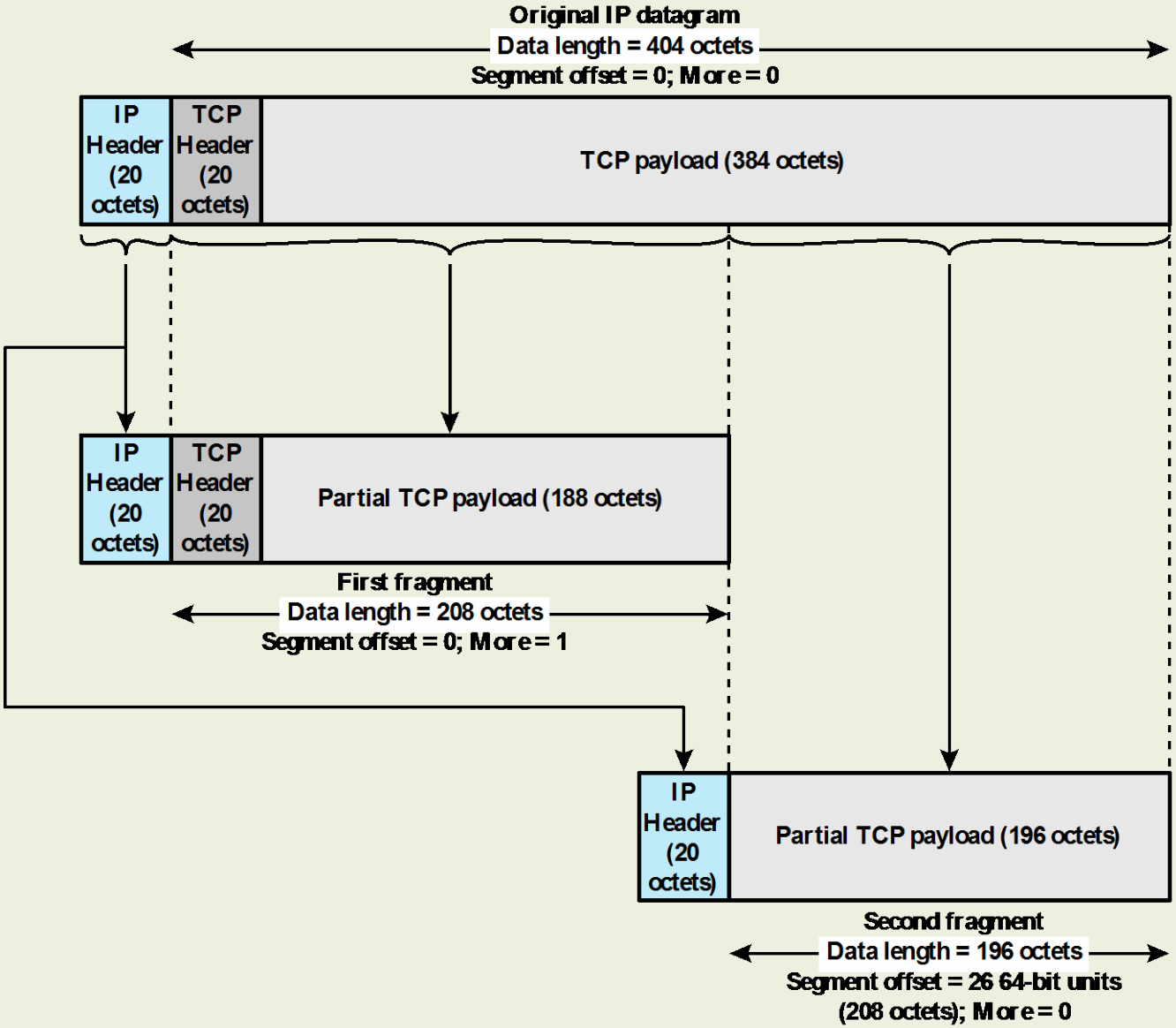


Figure 14.4 Fragmentation Example

Internet Protocol (IP) v4

- Defined in RFC 791
- Part of TCP/IP suite
- Two parts

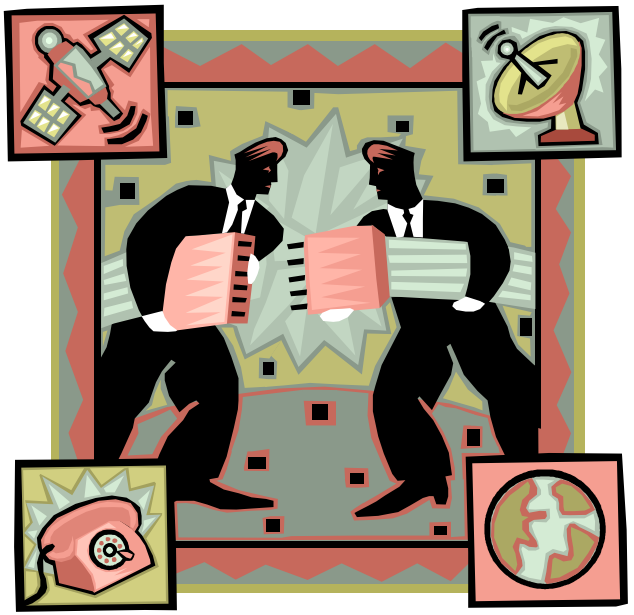
Specification of interface with a higher layer

Specification of actual protocol format and mechanisms

IP Services

- Primitives
  - Specifies functions to be performed
  - Form of primitive implementation dependent
  - **Send-request** transmission of data unit
  - **Deliver-notify** user of arrival of data unit

- Parameters
  - Used to pass data and control information



IP Parameters

- Source and destination addresses
- Protocol
- Type of Service
- Identification
- Don’t fragment indicator
- Time to live
- Data length
- Option data
- User data

IP Options

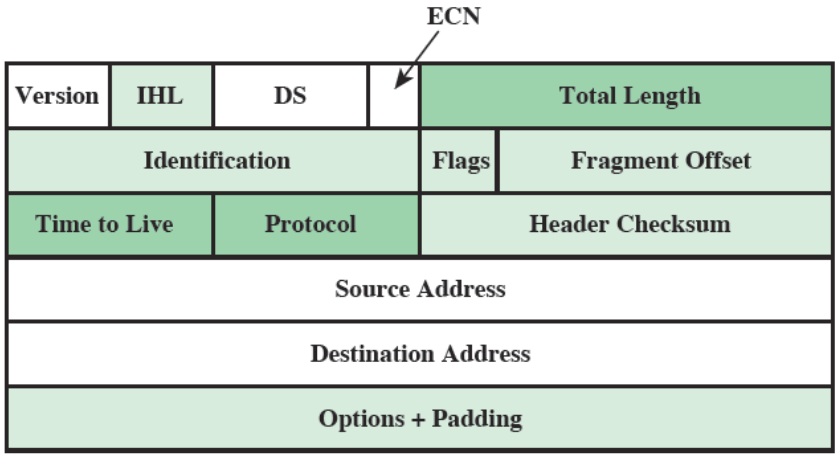
Security

Route recording

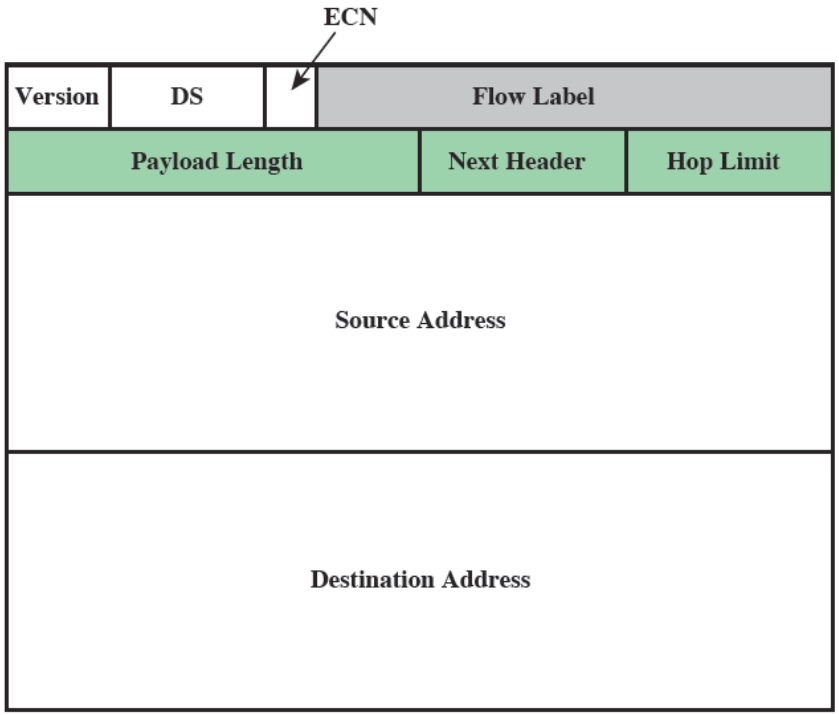
Source routing

Stream identification

Timestamping



(a) IPv4 header



(b) IPv6 header

ECN =Explicit Congestion Notification field

IP Addresses Class A

- Start with binary 0
- Network addresses with a first octet of 0 (binary 0000000) and 127 (binary 01111111) are reserved
- 126 potential Class A network numbers
- Range 1 to 126

IP Addresses Class B

- Start with binary 10
- Range 128 to 191(binary 10000000 to 10111111)
- Second octet also included in network address
- 2^{14} = 16,384 Class B addresses

IP Addresses Class C

- Start with binary 110
- Range 192 to 223
- Second and third octet also part of network address
- 2^{21} = 2,097,152 addresses
- Nearly all allocated
  - See IPv6

Subnets and Subnet Masks

- Allows arbitrary complexity of internetworked LANs within organization
- Insulate overall internet from growth of network numbers and routing complexity
- Site looks to rest of internet like single network
- Each LAN assigned subnet number
- Host portion of address partitioned into subnet number and host number
- Local routers route within subnetted network
- Subnet mask indicates which bits are subnet number and which are host number
- Check Table 14.2

IP Next Generation

Address space exhaustion:

- Two level addressing (network and host) wastes space
- Network addresses used even if not connected
- Growth of networks and the Internet
- Extended use of TCP/IP
- Single address per host

Requirements for new types of service

- Address configuration routing flexibility
- Traffic support

Internet of Things

IPv6 RFCs (use as reference)

- RFC 1752 - Recommendations for the IP Next Generation Protocol
  - Requirements
  - PDU formats
  - Addressing, routing security issues
- RFC 2460 - overall specification
- RFC 4291 - addressing structure

IPv6 Enhancements

- Expanded 128 bit address space
- Improved option mechanism
  - Most not examined by intermediate routes
- Dynamic address assignment
- Increased addressing flexibility
  - Anycast and multicast
- Support for resource allocation
  - Labeled packet flows

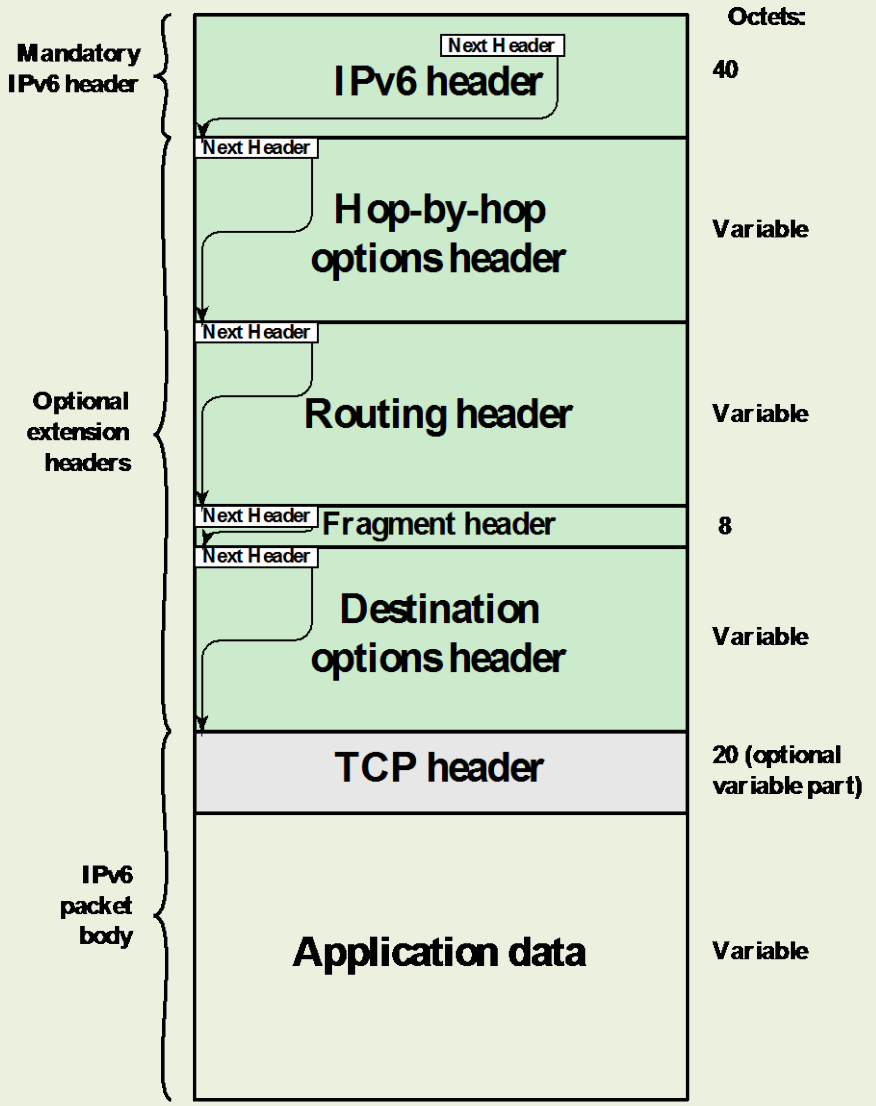
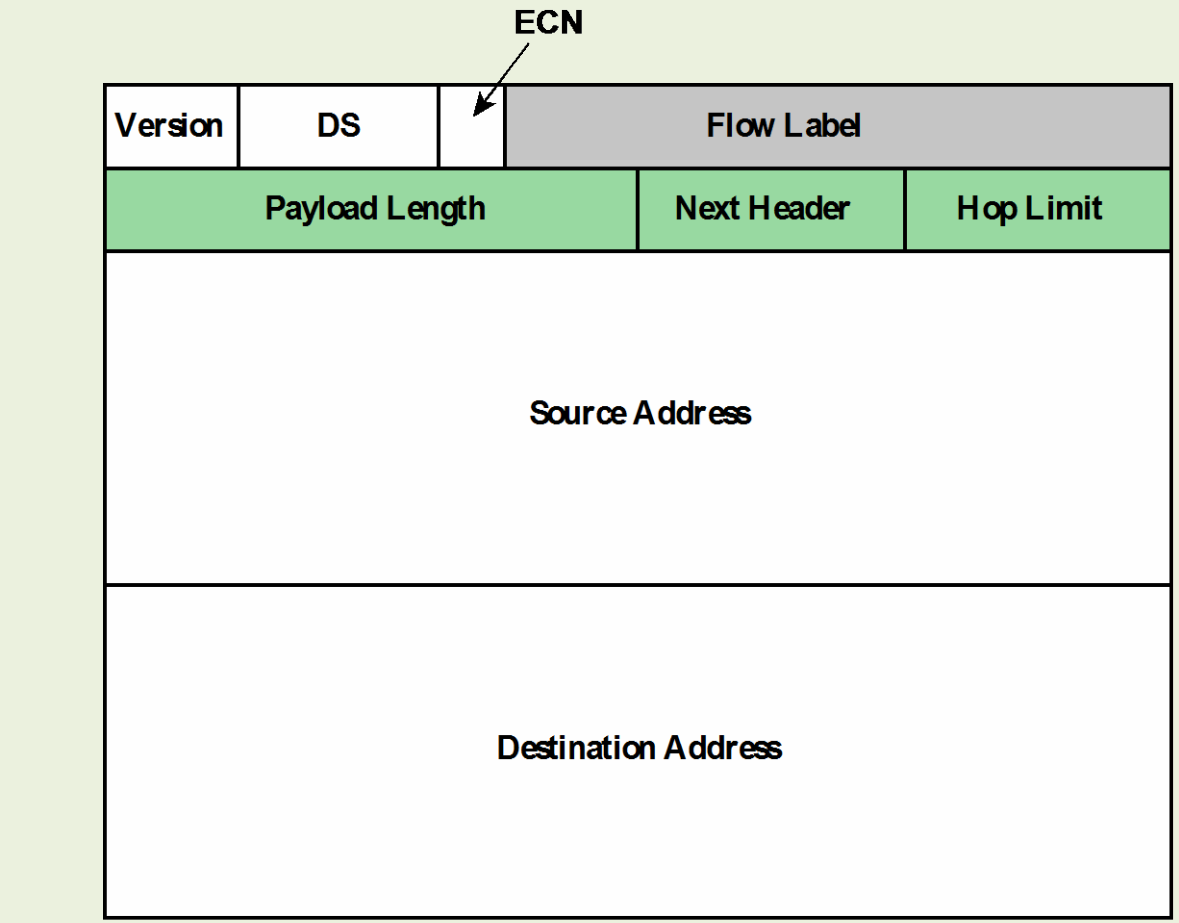


Figure 14.9 IPv6 Packet with Extension Headers (containing a TCP Segment)

IPv4 has option fields as part of single header -> header size varies


ECN
Version
DS
Flow Label
Payload Length
Next Header
Hop Limit
Source Address
Destination Address
(b) IPv6 header
Field name kept from IPv4 to IPv6
Name and position changed in IPv6
Field not kept in IPv6
New field in IPv6

IPv6 Flow Label

- Related sequence of packets
- Special handling
- Identified by source and destination address plus flow label
- Router treats flow as sharing attributes
- May treat flows differently
- Alternative to including all information in every header
- Have requirements on flow label processing

IPv6 Addresses

- 128 bits long
- Assigned to interface
- Single interface may have multiple unicast addresses

Three types of addresses:

- Unicast - single interface address
- Anycast - one of a set of interface addresses
- Multicast - all of a set of interfaces

IPv6 addresses

- 128 bits = 16 bytes
- $2^{128} = 2^{32} \cdot 2^{96} > 3 \cdot 10^{35}$
- Notations

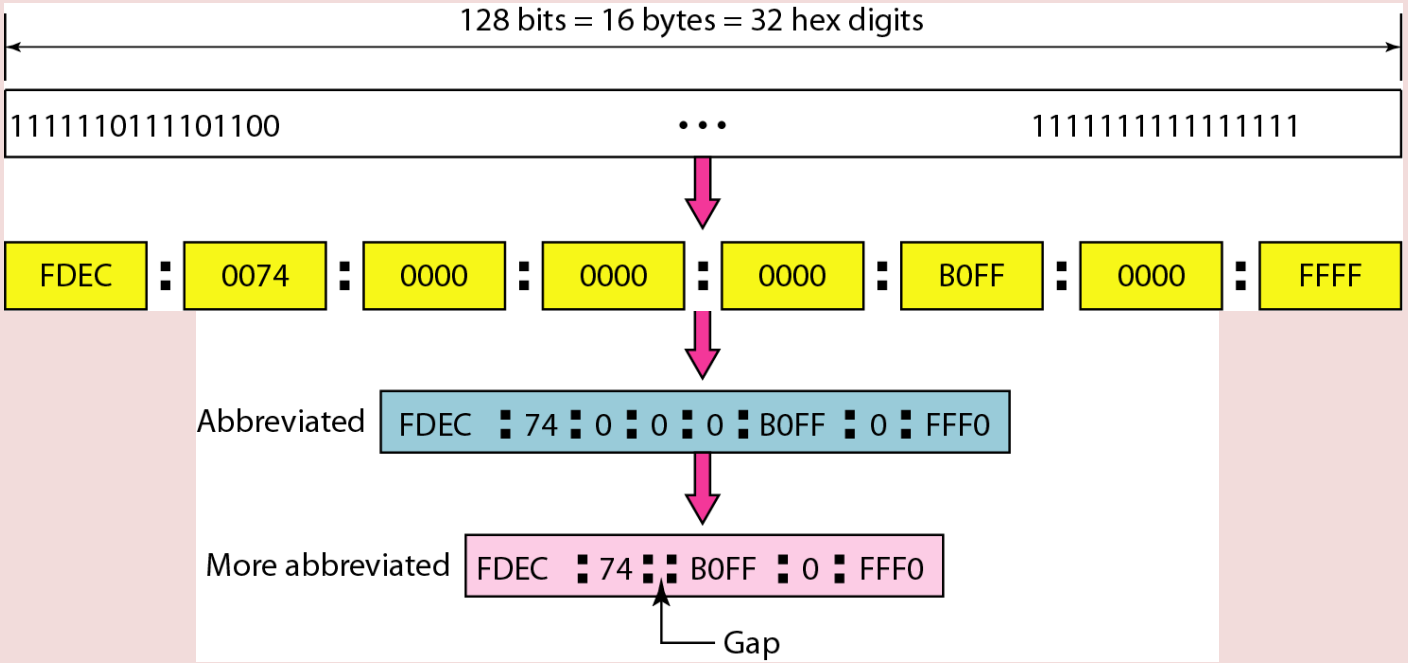
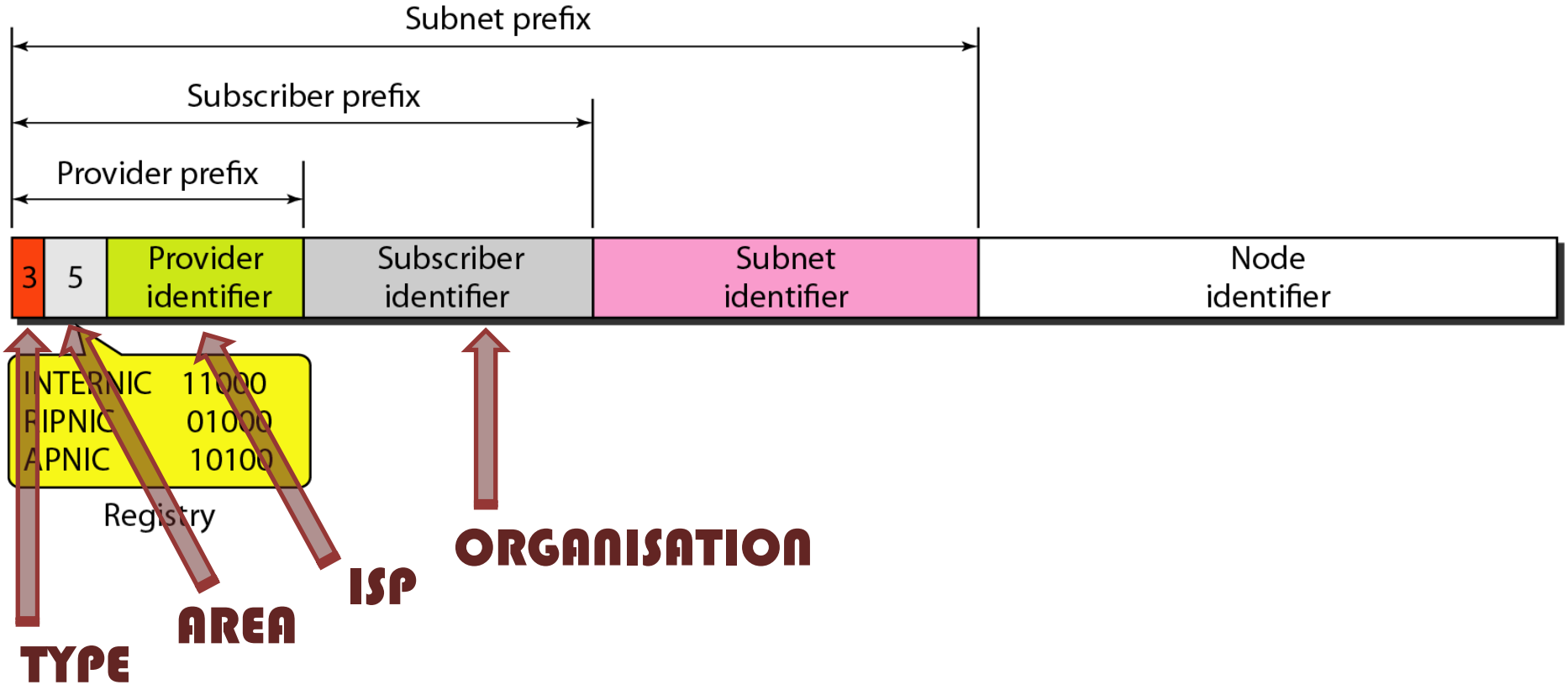


Table 14.3
IPv6 Address Space Usage
(use as reference)

Address Type	Binary Prefix	IPv6 Notation	Fraction of address space
Embedded IPv4 address	00…1111 1111 1111 1111 (96 bits)	::FFFF/96	2^{-96}
Loopback	00…1 (128 bits)	::1/128	2^{-128}
Link-local unicast	1111 1110 10	FE80::/10	1/1024
Multicast	1111 1111	FF00::/8	2/256
Global unicast	Everything else		

Global unicast addresses

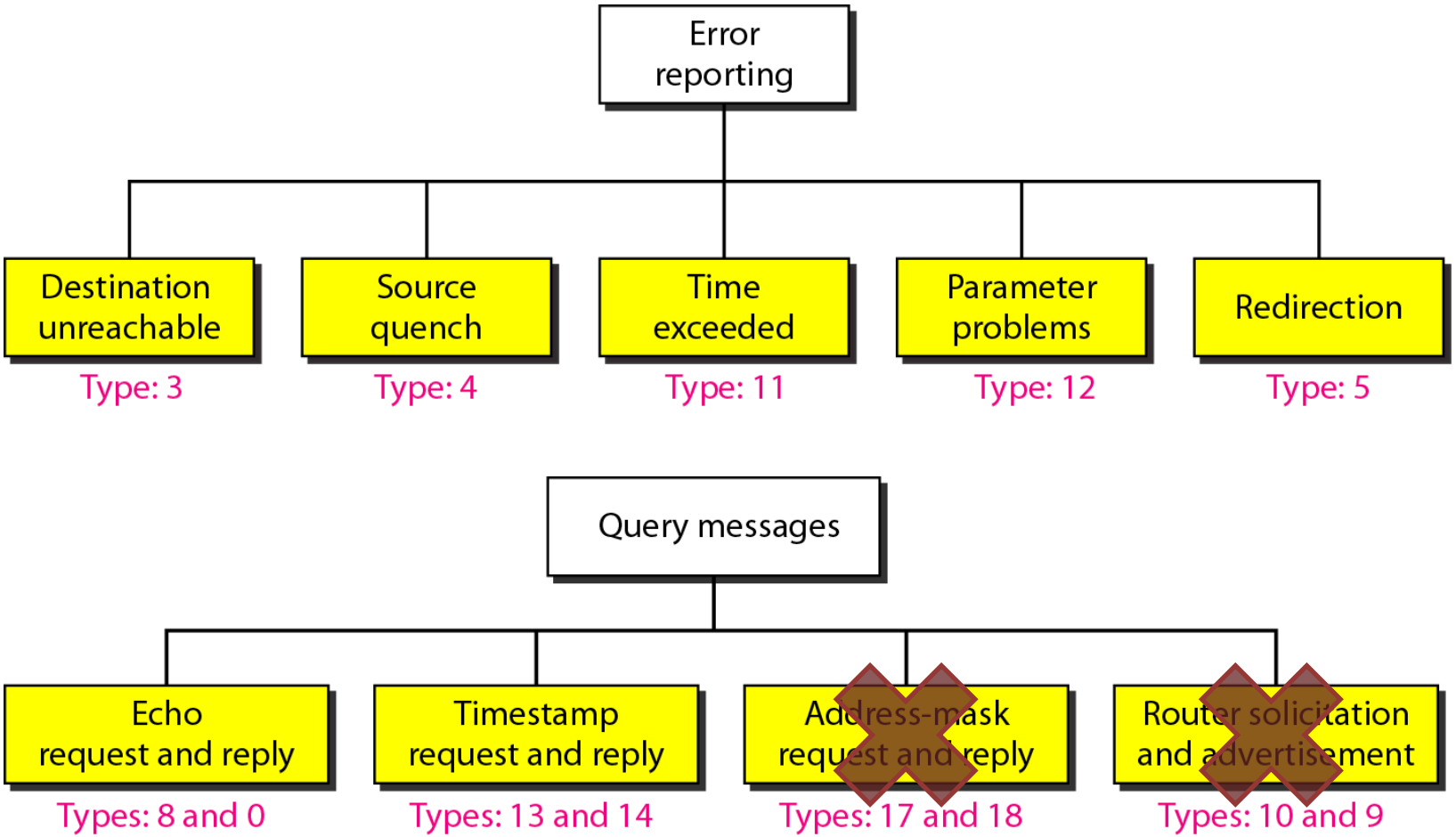
- Identify individual computers



Internet Control Message Protocol (ICMP)

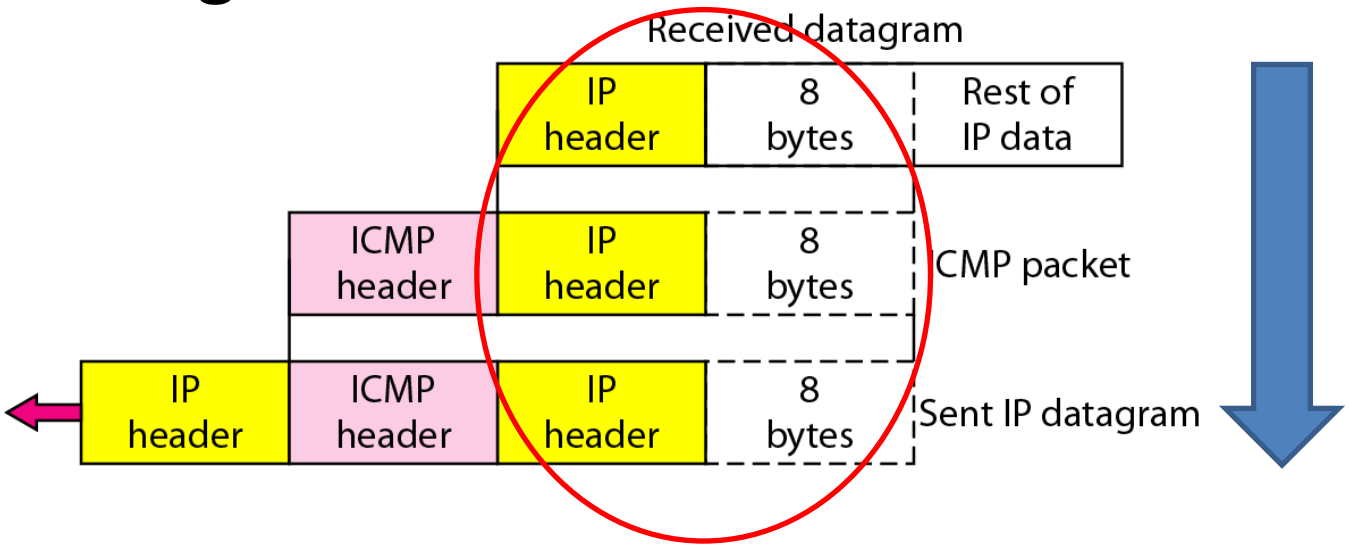
- RFC 792
- Provides a means for transferring messages from routers and other hosts to a host
- Provides feedback about problems
    - Datagram cannot reach its destination
    - Router does not have buffer capacity to forward
    - Router can send traffic on a shorter route
- Encapsulated in IP datagram
  – Hence not reliable

ICMPv4 message types

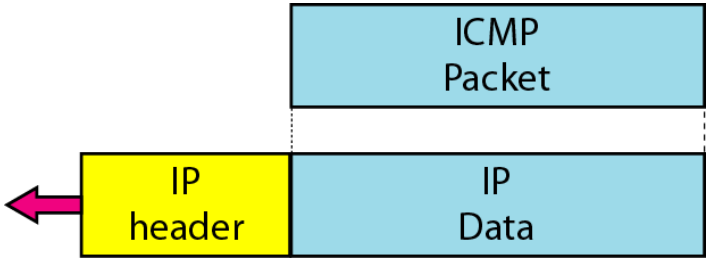


ICMP message formats

- Error reporting



- Query messages



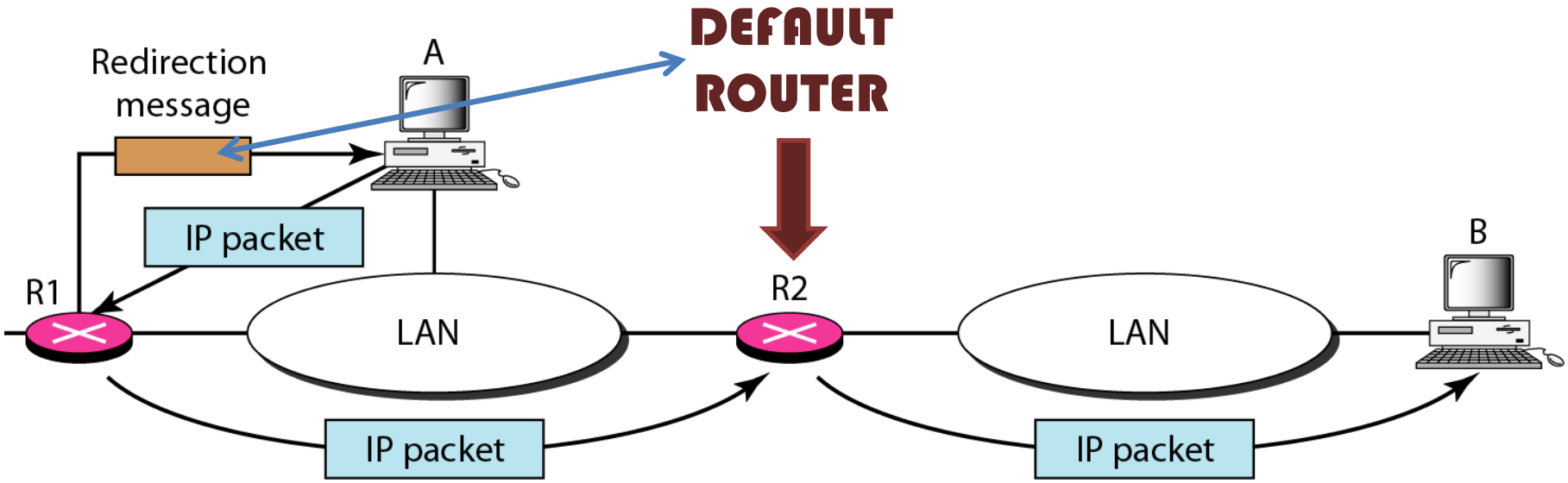
Echo request and reply (query type)

- Is my destination alive?

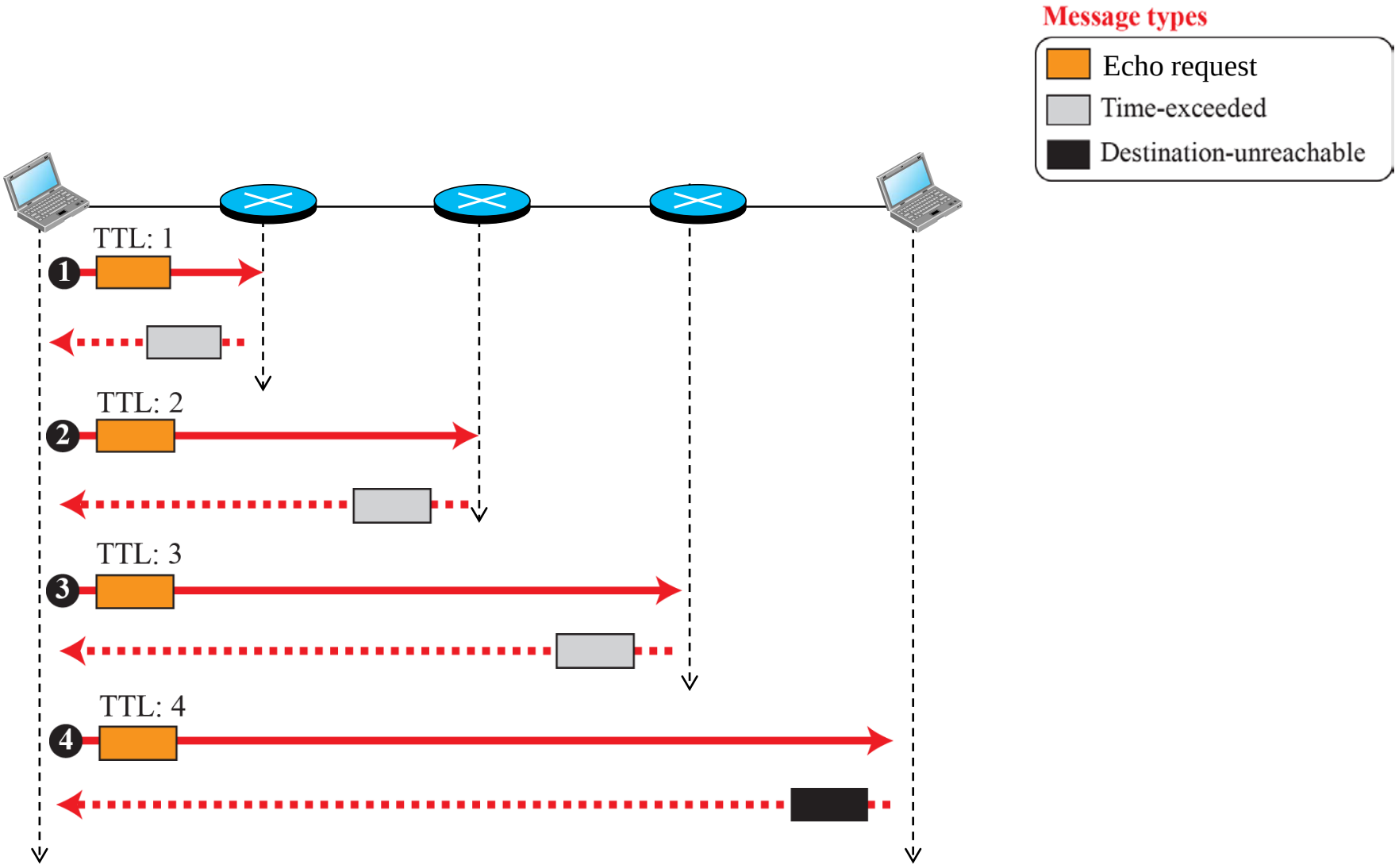
- Network diagnostics
  - IP layer
- Debugging tools
  - Ping
  - Traceroute

Redirection (error reporting type)

- Routing update for hosts
  - Security/reliability?



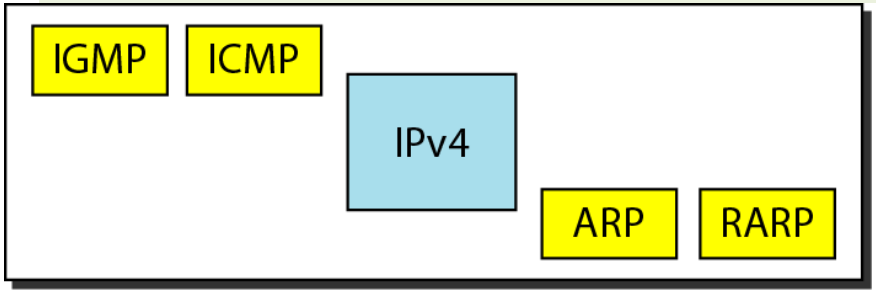
Traceroute


Message types
Echo request
Time-exceeded
Destination-unreachable
TTL: 1
1
TTL: 2
2
TTL: 3
3
TTL: 4
4

Changes to ICMP

ICMPv4

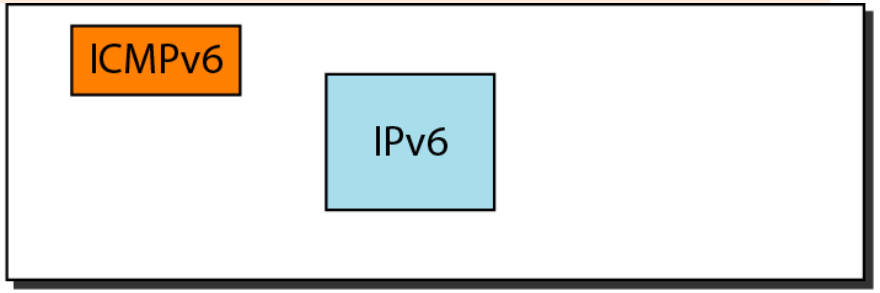
- Some unused functions



Network layer in version 4

ICMPv6

- Same principle
- Some new functions
- Convergence
- Suits IPv6 better



Network layer in version 6

ICMPv6

- Includes ”IPv4 IGMP”
  - Group membership messages
    - Multicast Listener Delivery protocol (MLD)
- Includes ”IPv4 ARP”
  - Part of Neighbor Discovery Protocol (NDP)

Neighbour Discovery Protocol (NDP)

- Router Solicitation/Adverticement
  - Find a router = ”default gateway”
  - Announce router = ”default gateway”
- Neighbour Solicitation/Adverticement
  - Same functionality as IPv4 ARP

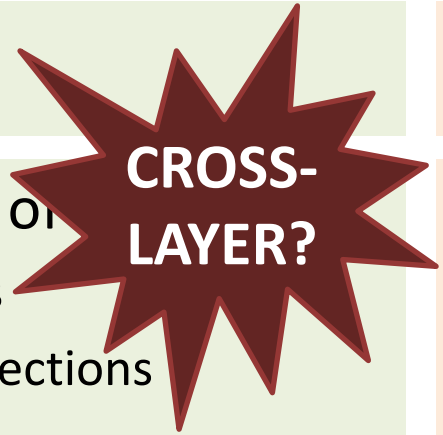
IPv6 Autoconfiguration

- Every NIC has several IPv6 addresses
  - Most have Link Local Address
- Creation of Link Local Address:
  - Use MAC address
  - Prebend with wellknown prefix fe80::/64
  - Check for duplicates

IPv6 and QoS

Flow label

- Identification of
  - TCP sessions
  - Virtual connections
- Processing
  - Flow label table
  - Forwarding table
- Routing
  - Algorithms still necessary
  - But not run for every packet!



Traffic class

- Classification of packets
  - Queueing schemes
  - Relation to delay
- TCP vs. UDP
  - Congestion-controlled
  - Non-congestion-controlled
- Other protocols
  - RTP
  - RSVP

Address Resolution Protocol (ARP)

Need MAC address to send to LAN host

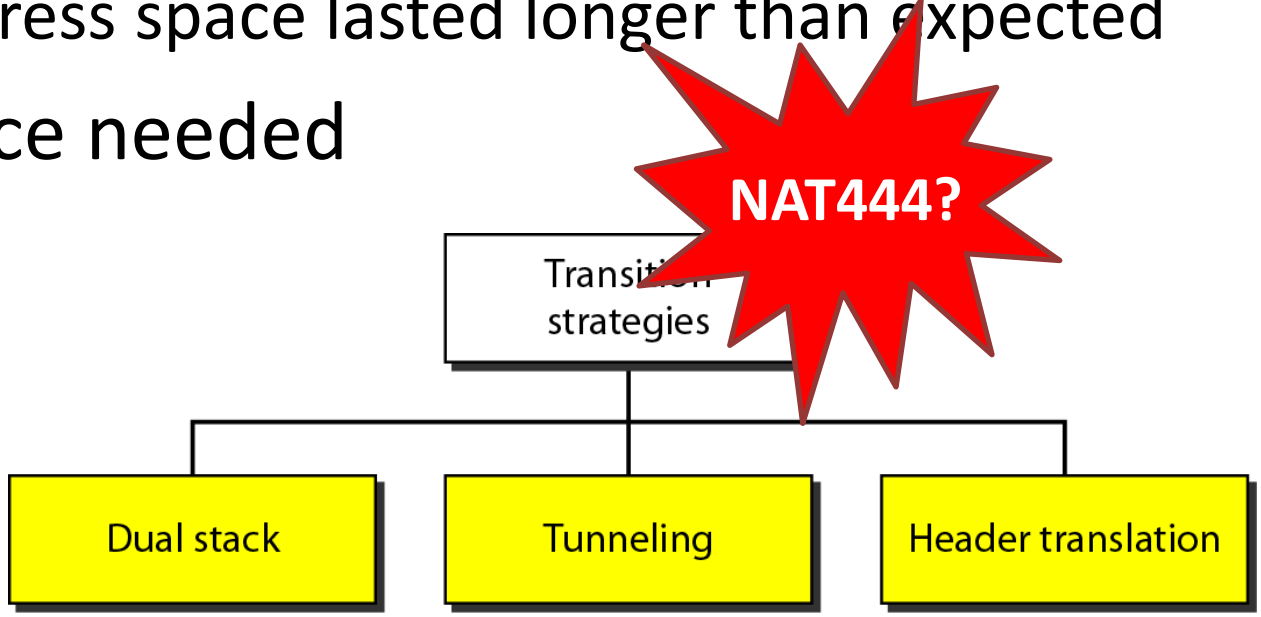
- Manual
- Included in network address
- Use central directory
- Use address resolution protocol

ARP (RFC 826) provides dynamic IP to Ethernet address mapping

- Source broadcasts ARP request
- Destination replies with ARP response

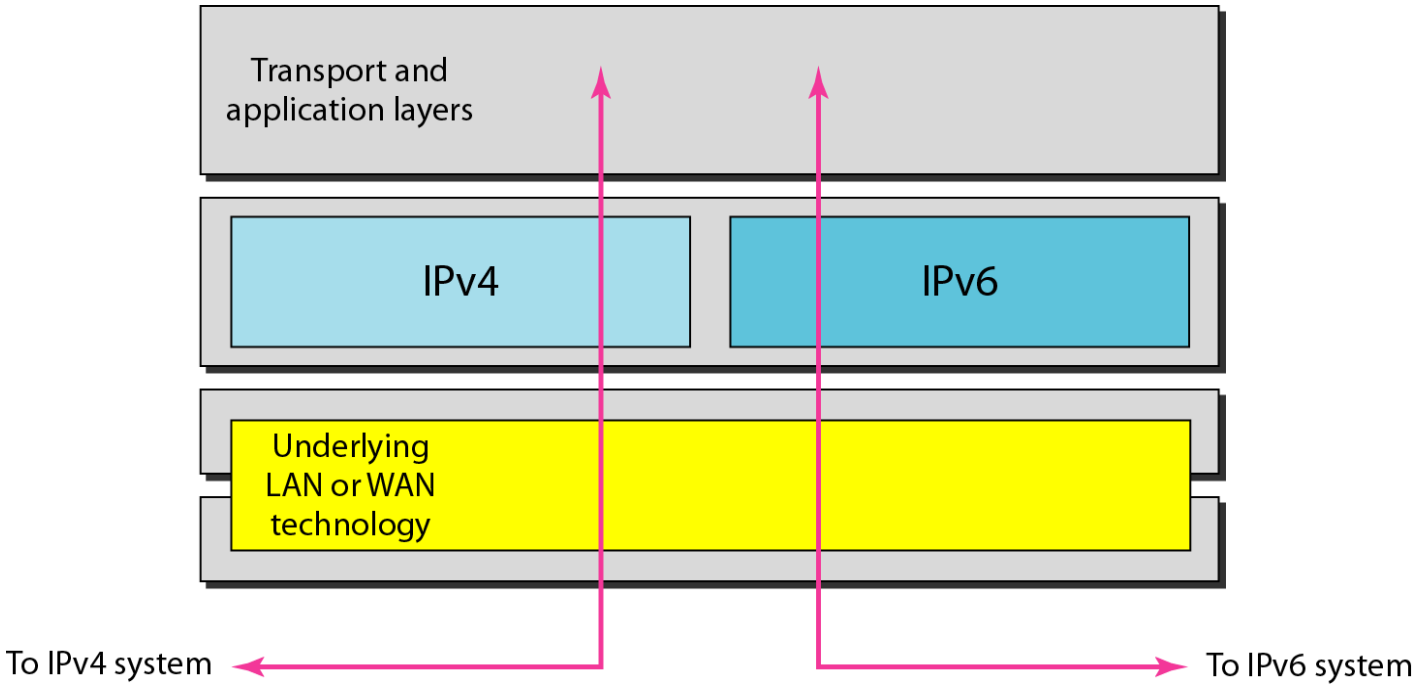
Transition: IPv4 → IPv6

- Cannot happen overnight
  - Too many independent systems
  - Economic cost
  - IPv4 address space lasted longer than expected
- Coexistence needed



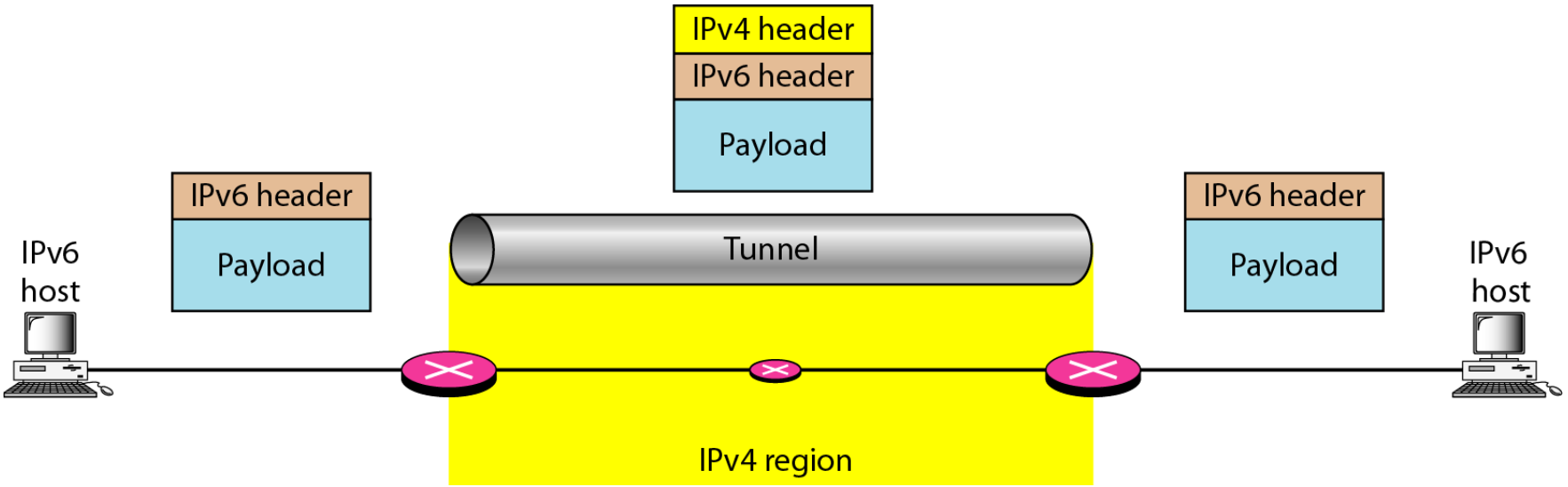
Transition: (1) Dual stack

- Decision based on destination IP



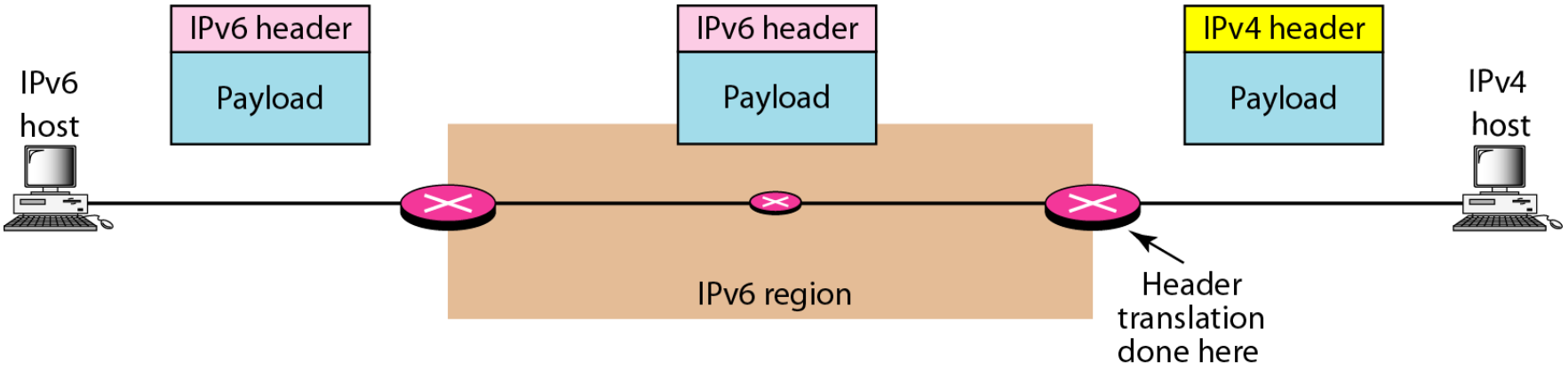
Transition: (2) Tunneling

- A few IPv6 routers



Transition: (3) Header translation

- A few IPv4 routers

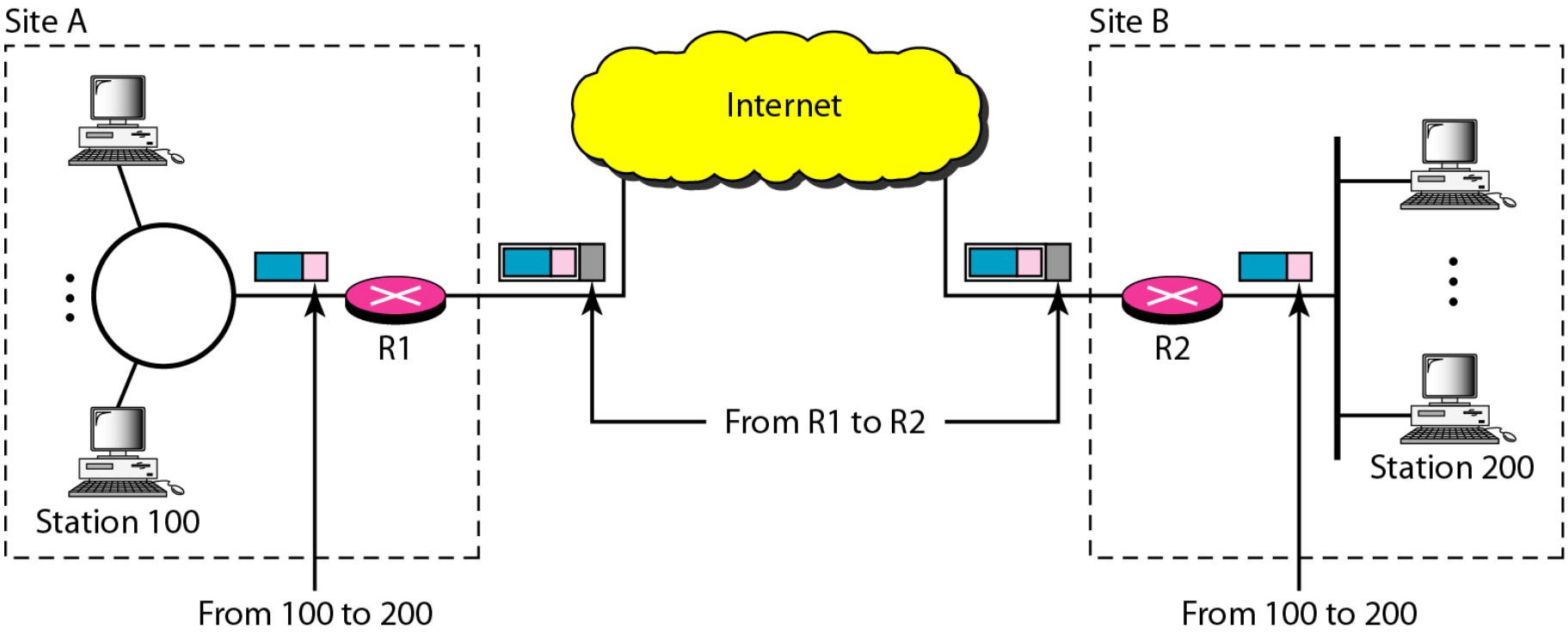


Virtual Private Network (VPN)

- Set of computers interconnected using an unsecure network
    - e.g. linking corporate LANs over Internet
- Using encryption and special protocols to provide security
    - Eavesdropping
    - Entry point for unauthorized users
- Proprietary solutions are problematical
    - Development of IPSec standard

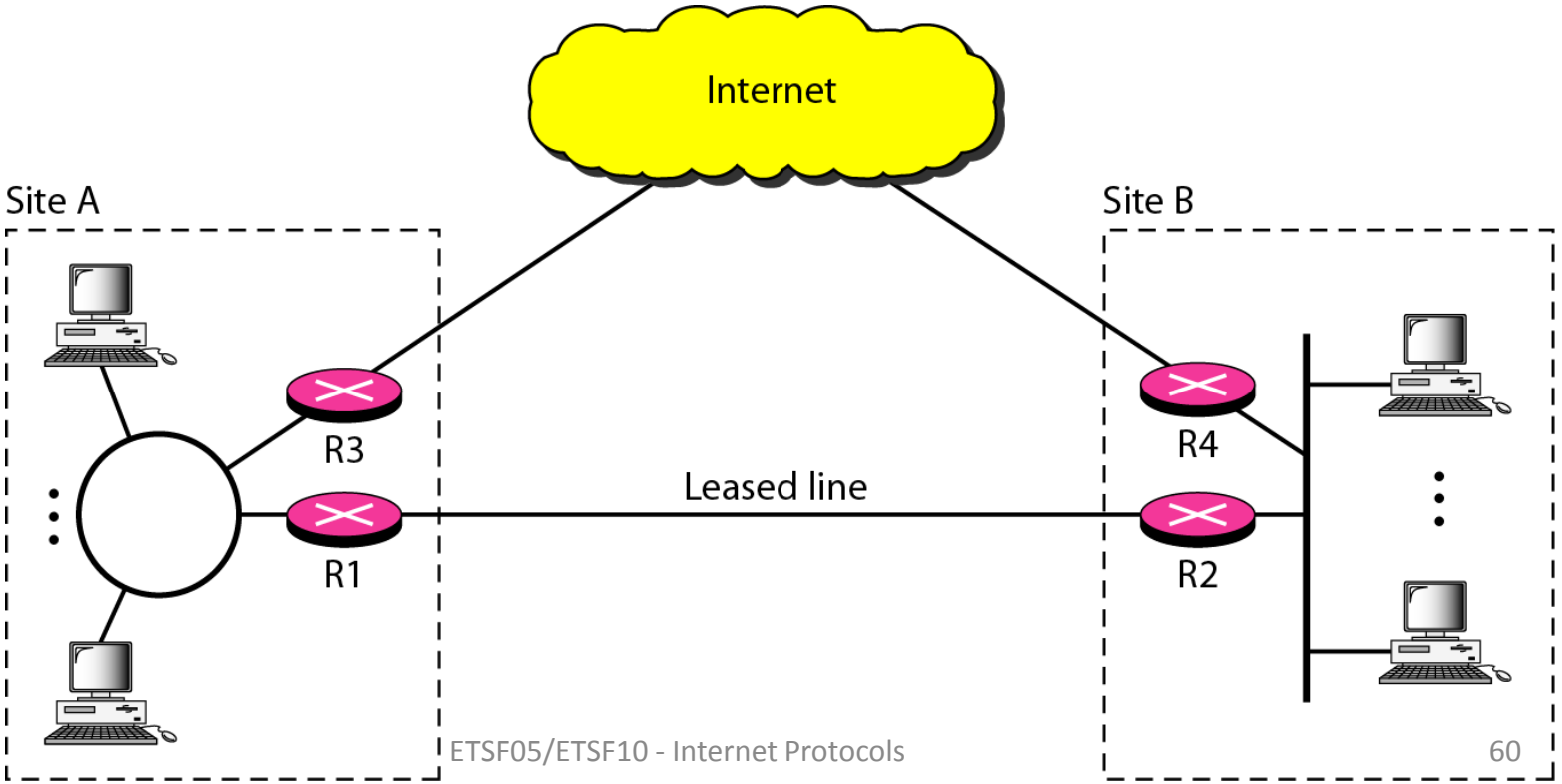
An example VPN

- IPSec between routers



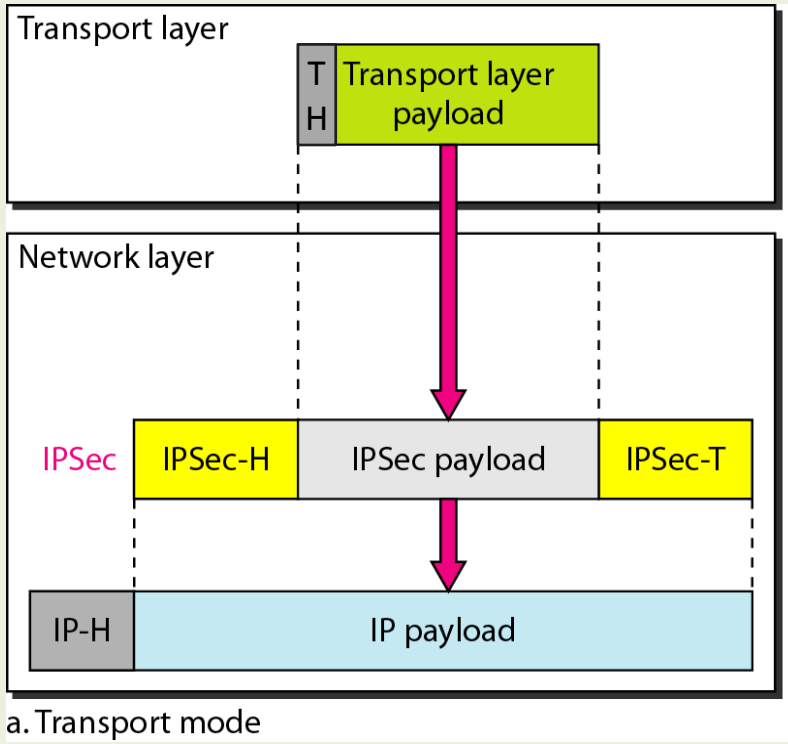
Virtual Private Network (VPN)

- Overlay network
- Alternative to a real private network



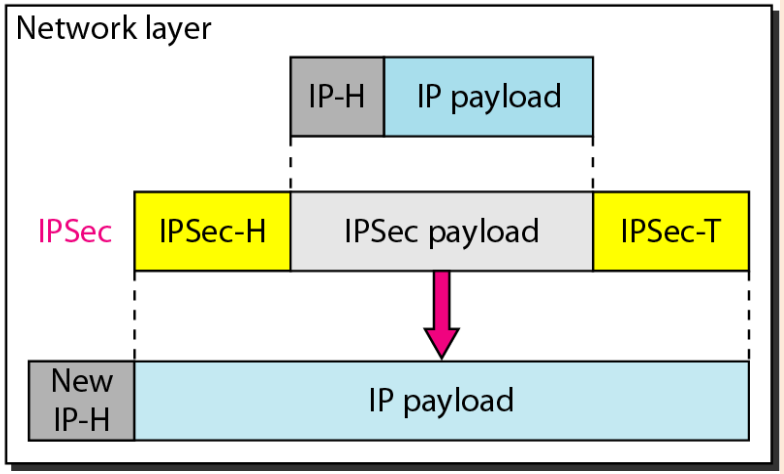
IPSec

Transport mode



a. Transport mode

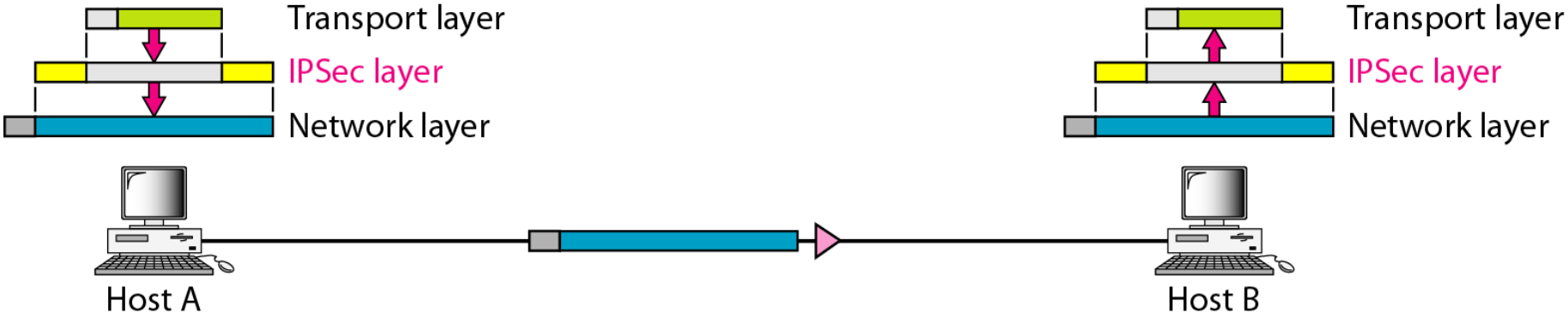
Tunnel mode



b. Tunnel mode

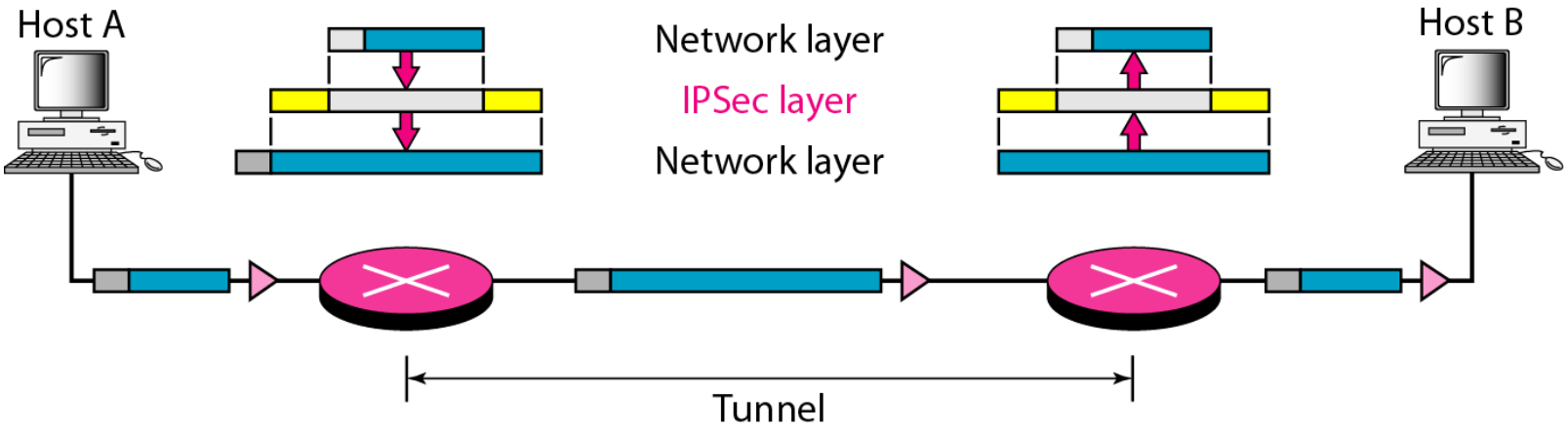
Transport mode in action

- Data protected
- Headers unprotected
  - Addresses fully visible

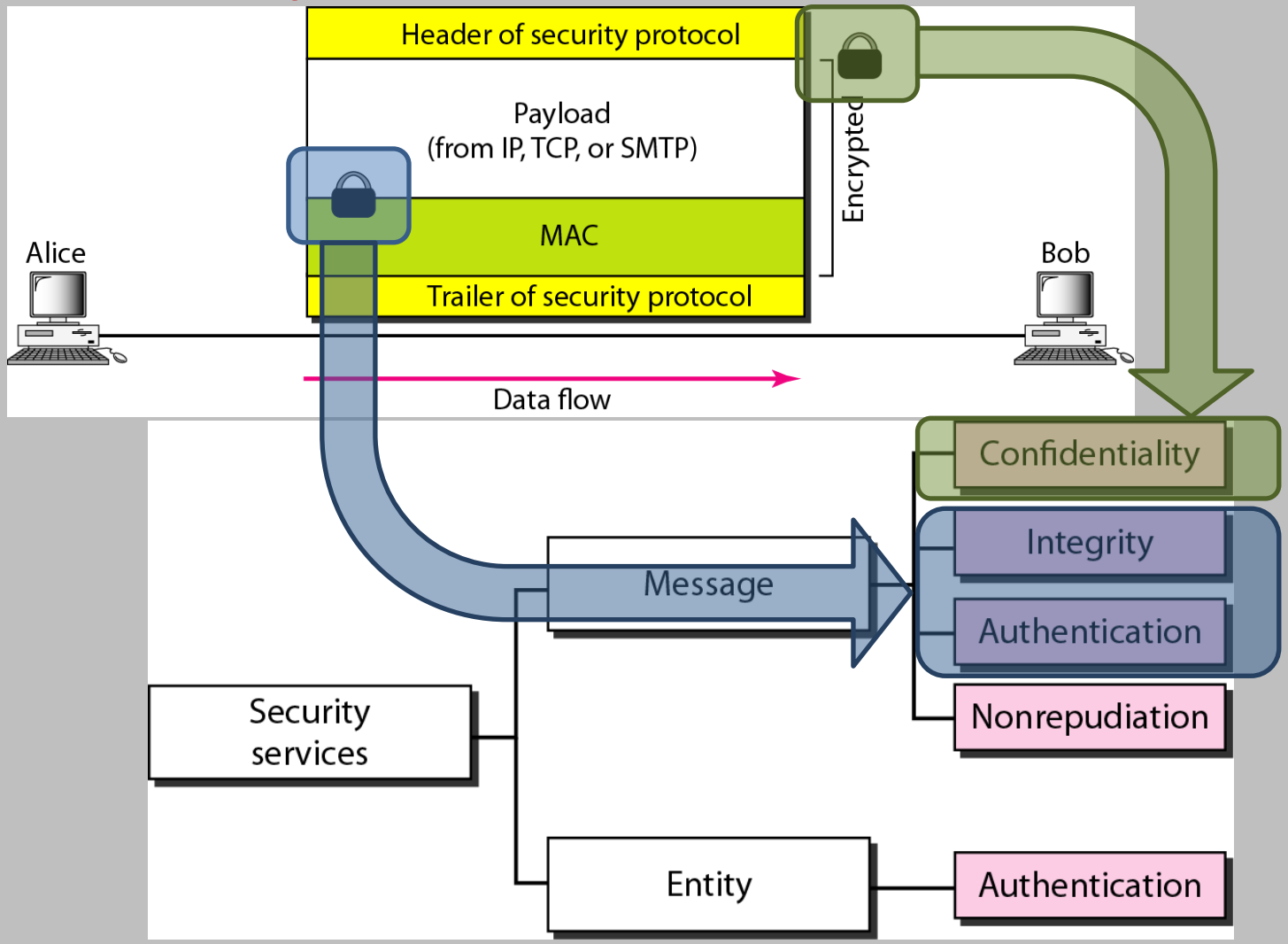


Tunnel mode in action

- Not used between hosts
- Entire packet protected
  - New header inside tunnel

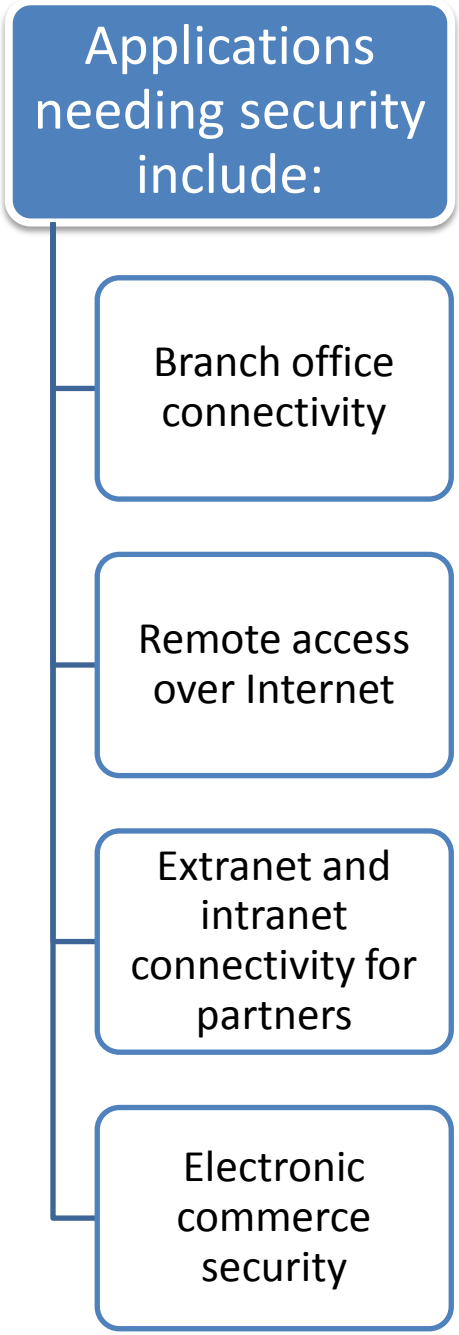


Internet security (discussed in other courses)



IPsec

- RFC 1636 (1994) identified security need
- Encryption and authentication necessary security **features in IPv6**
- Designed **also for use with current IPv4**



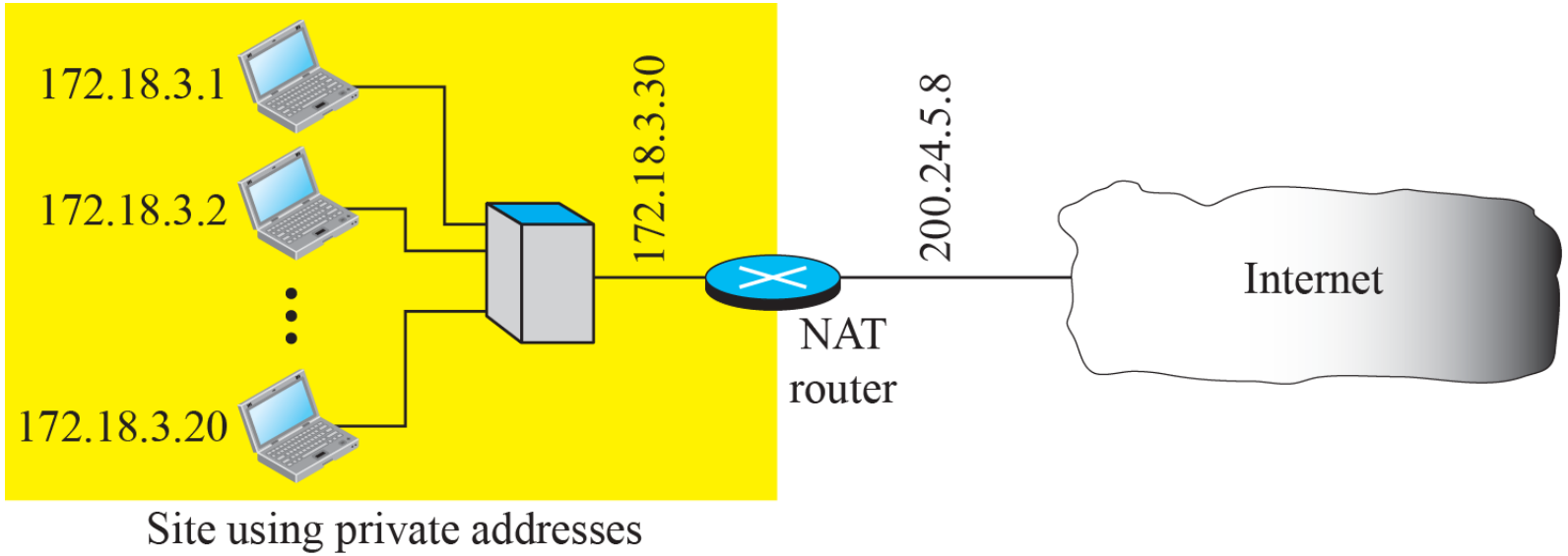
VPN alternatives (bonus material)

- PPTP (Point-to-Point Tunneling Protocol)
- L2TP (Layer 2 Tunneling Protocol)
- SSTP (Secure Socket Tunneling Protocol)
- OpenVPN

- See Wikipedia for information

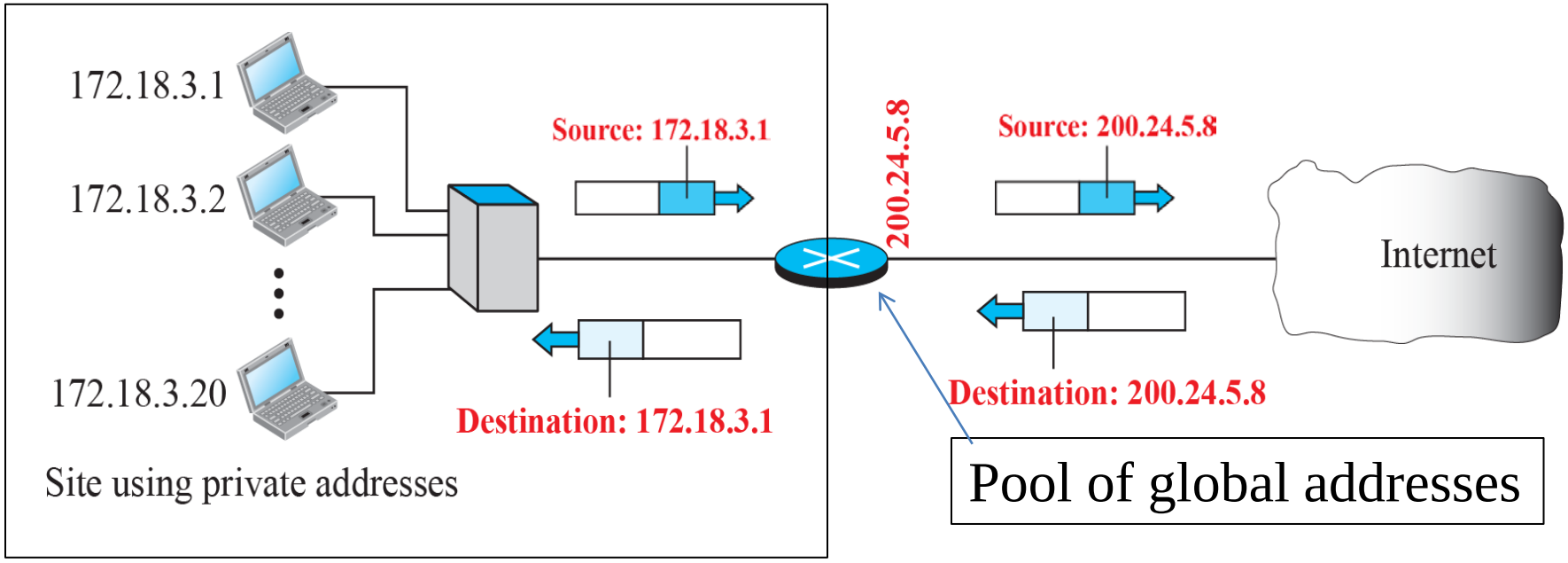
NAT - Network Address Translation

- Sharing of routable addresses (scarse resource)
- Adds some security …



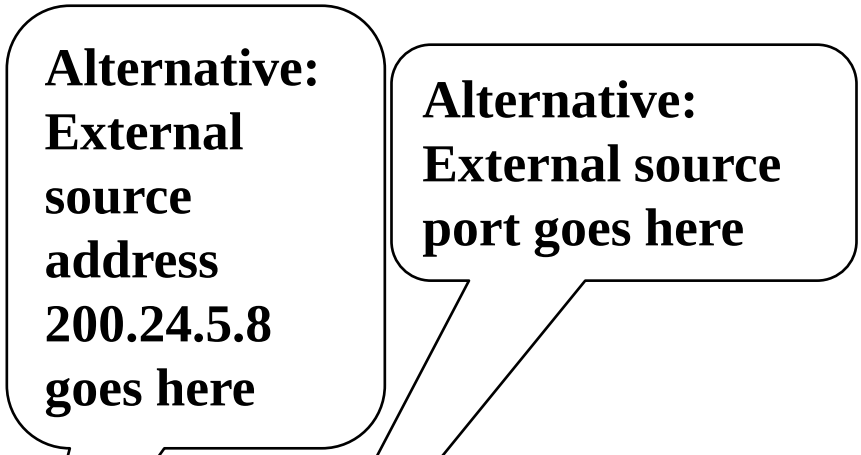
NAT (network address only)

- Change source address on outgoing packets
- Add address pair to active translations table
- Only one internal address per destination



NAT extended

- Add transport layer port



Private address	*Private port*	*External address*	*External port*	*Transport protocol*
172.18.3.1	1400	25.8.3.2	80	TCP
172.18.3.2	1401	25.8.3.2	80	TCP
⋮	⋮	⋮	⋮	⋮

- Normally initiated from inside
- Port forwarding: Setup static entry in table

NAT444, Carrier Grade NAT

- Carrier performs NAT in core
- Benefits?
- Problems?
- Online discussion 2