Computer Networks

Wenzhong Li, Chen Tian

Nanjing University

Material with thanks to James F. Kurose, Mosharaf Chowdhury, and other colleagues.

Chapter 3. Network Layer

- Network Layer Functions
- IP Protocol Basic
- IP Protocol Suit
- Routing Fundamentals
- Internet Routing Protocols
- IP Multicasting

- NAT
- ARP
- DHCP
- ICMP
- Mobile IP
- IPv6

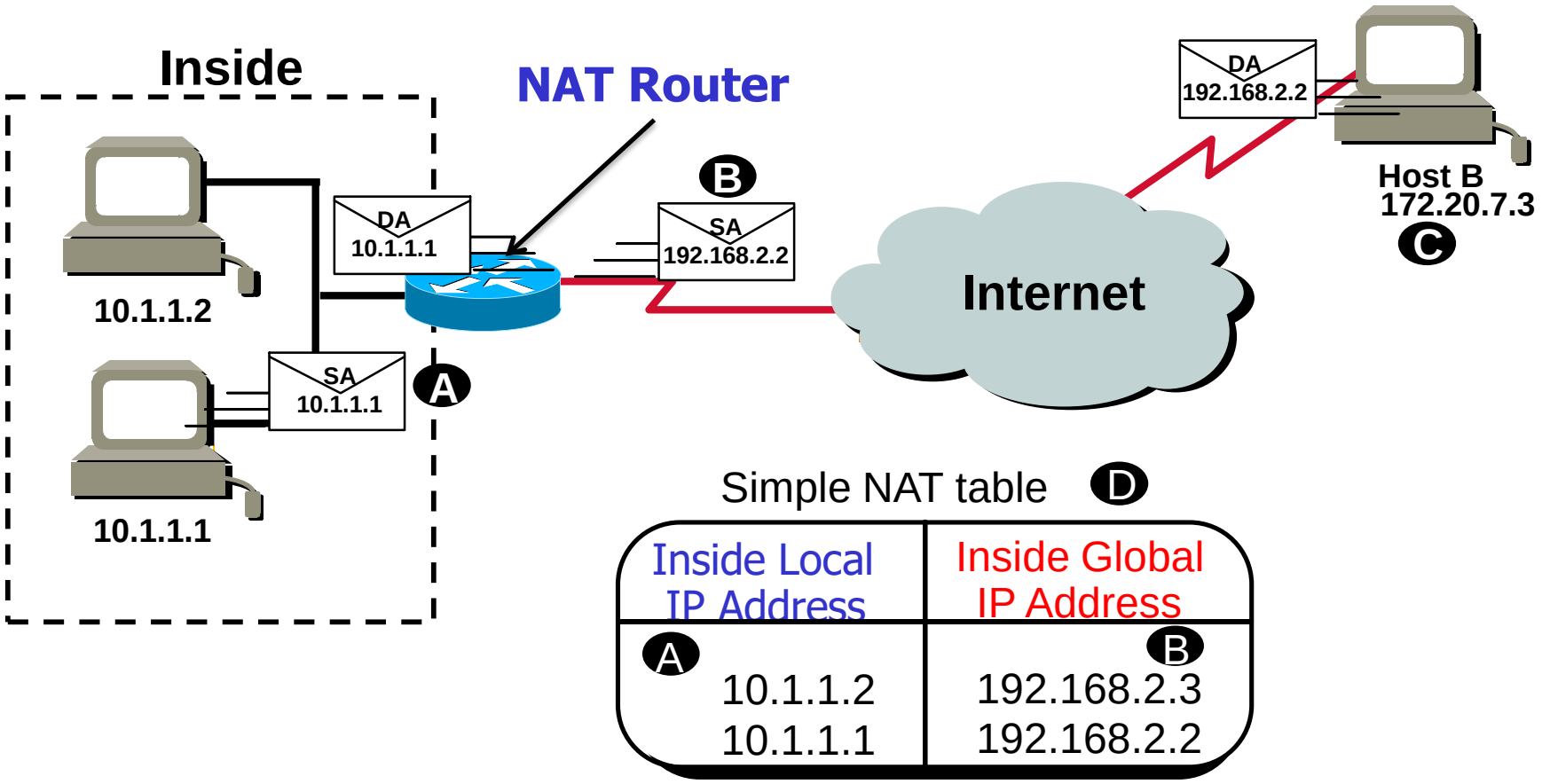
Network Address Translation

- NAT
  - Enables different sets of IP addresses for internal and external traffic
  - The IP address translations occur where the Intranet interfaces with the broader Internet
- Purposes
  - Acts as a firewall by hiding internal IP addresses
  - Enables an enterprise (organization) to use more internal IP addresses
  - Isolate the (organization / ISP) changes

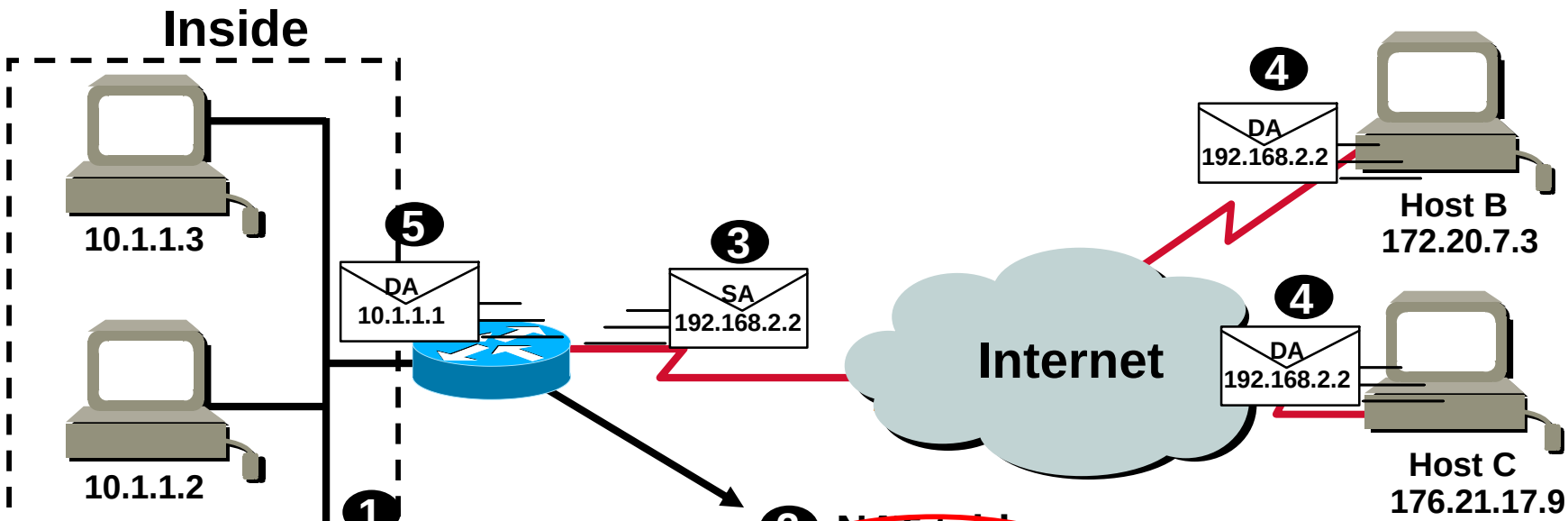
3 Types of NAT

- Static NAT
  - A private IP address is mapped to one reserved public IP address
  - Usually for server hosts in Intranet
- Dynamic NAT
  - The NAT router keeps a pool of registered IP addresses, and assign to private IP addresses on demand
  - Usually for client PCs in Intranet
- Single-Address NAT/Overloading/Masquerading/Network Address Port Translation (NAPT)

Illustration of NAT



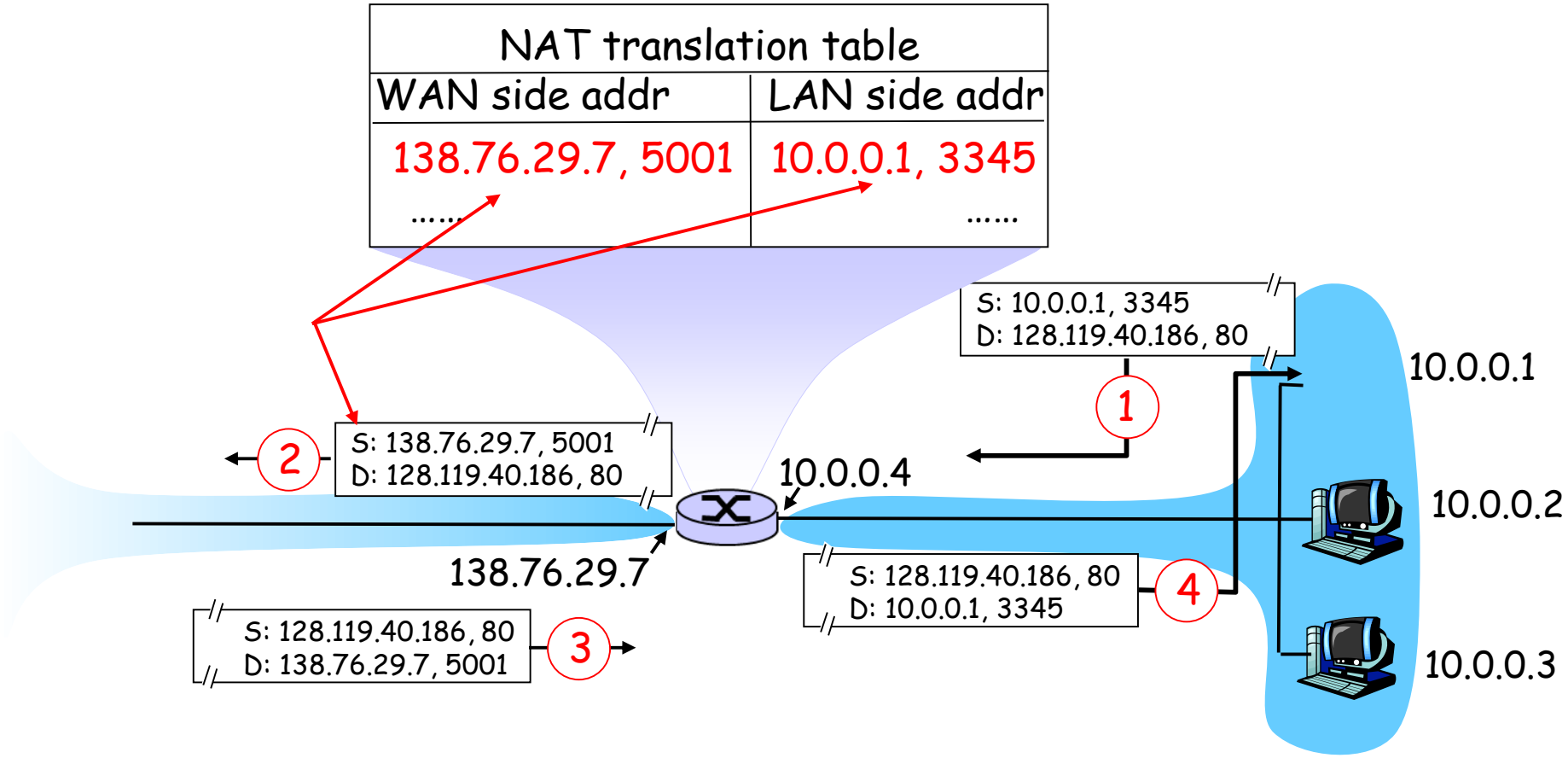
Overloading Global Address



2 NAT table

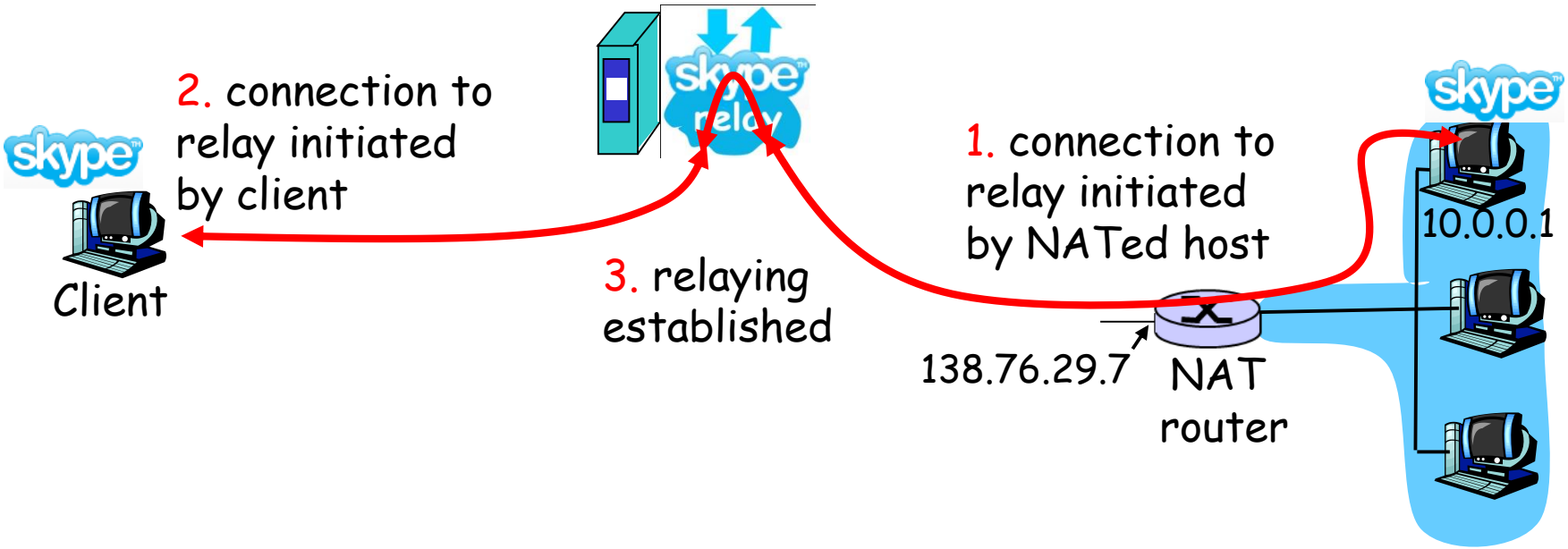
Protocol 10.1.1.1	Inside Local IP Address: Port	Inside Global IP Address: Port	Outside Global IP Address: Port
TCP	10.1.1.3:1723	192.168.2.2:1492	176.21.17.9:23
TCP	10.1.1.2:1723	192.168.2.2:1723	176.21.17.9:23
TCP	10.1.1.1:1024	192.168.2.2:1024	172.20.7.3:23

Network Address Translation



NAT is Controversial

- Addresses changes from time to time
  - E.g. must be taken into account by P2P applications
- Relaying in Skype
  - NATed supernodes establishes connection to relay
  - External client connects to relay
  - Relay bridges packets between 2 connections



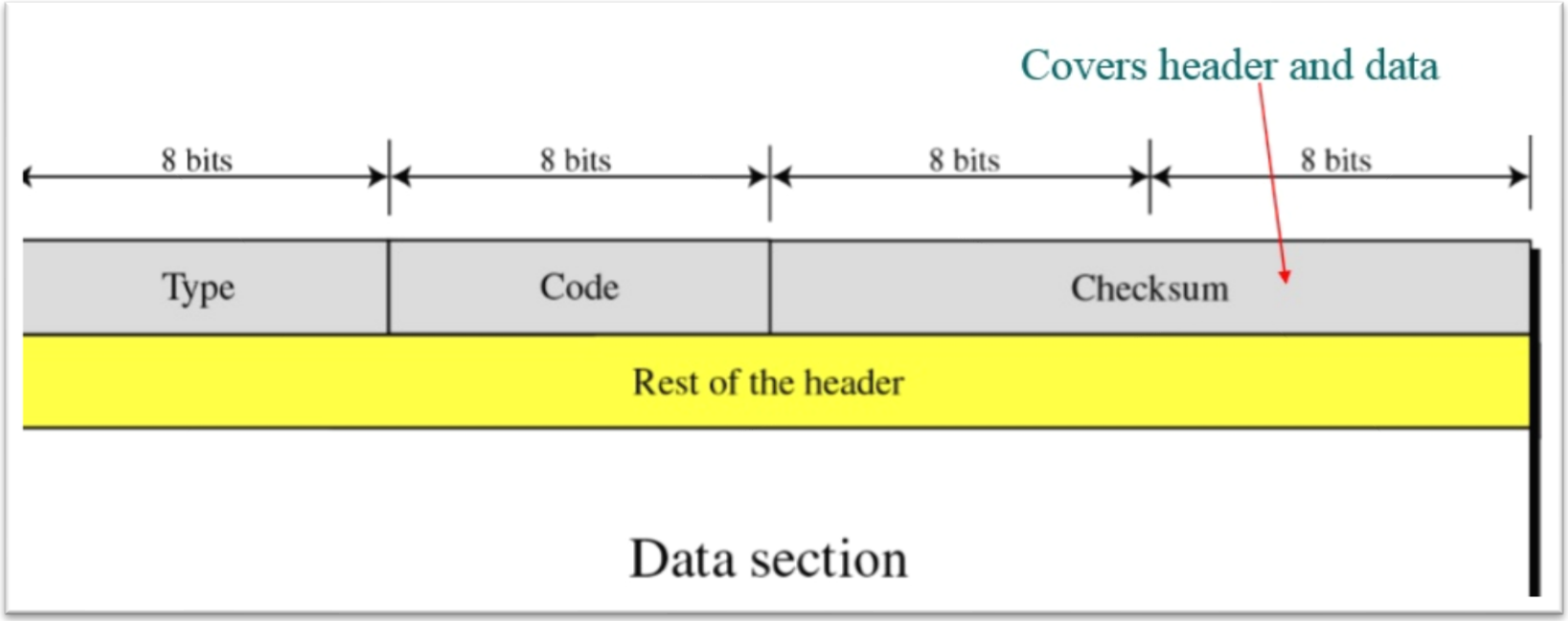
IP protocol suits

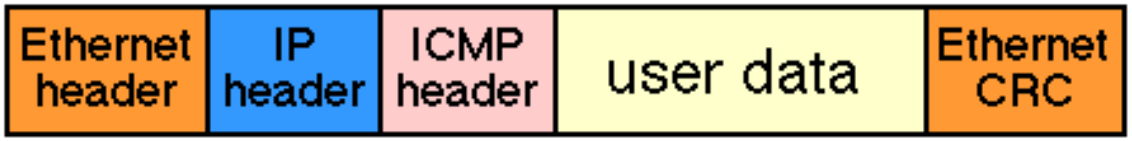
- ARP (Address Resolution Protocol)
- DHCP (Dynamic Host Configuration Protocol)
- ICMP (Internet Control Message Protocol)
- Mobile IP (IP Mobility Support Protocol)
- IPv6 (Internet Protocol Version 6)

ICMP

- Internet Control Message Protocol (RFC 792)
- Transfer of error and control msgs among routers and hosts
  - Echo request and reply to facilitate diagnostic
  - Feedback about problems, e.g. time to live expired, unreachable host
- Encapsulated in IP datagram
  - Protocol type = 1
  - Not reliable

ICMP Message Format



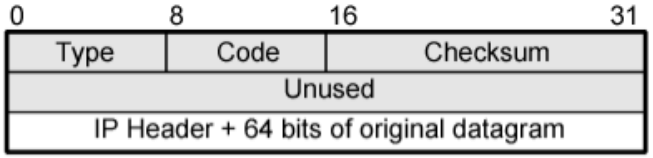


ICMP Message Types

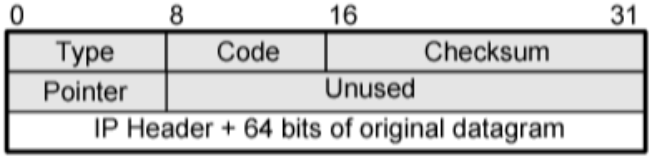
- ICMP Messages
  - Error Reports
    - Destination unreachable
    - Source quench (congestion control)
    - Parameters problem
    - Redirection
  - Request/Reply
    - Echo request/reply
    - Timestamp request/reply
    - Address mask request/reply
    - Router discovery/advertisement

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	Parameter unintelligible
13	0	timestamp
14	0	timestamp reply
15	0	address mask request
16	0	address mask reply

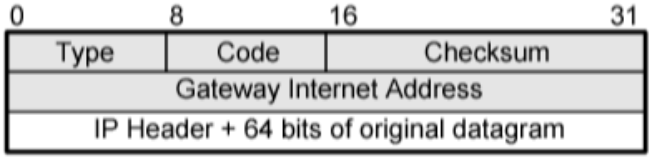
Some ICMP Message Formats



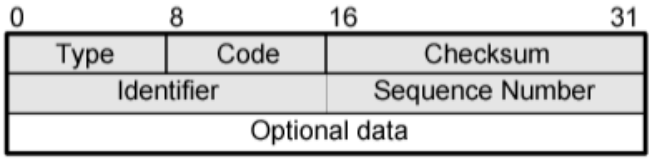
(a) Destination Unreachable; Time Exceeded; Source Quench



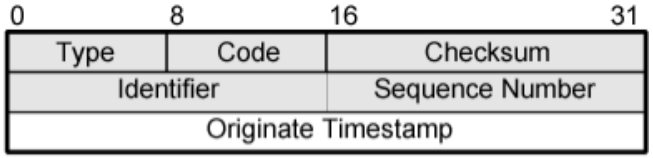
(b) Parameter Problem



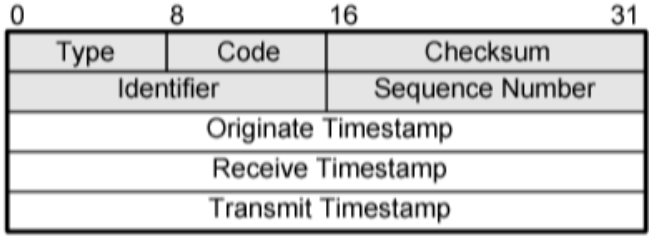
(c) Redirect



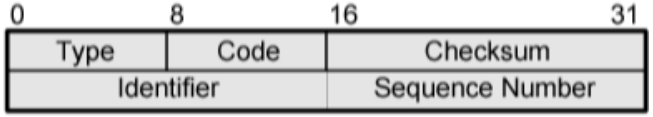
(d) Echo, Echo Reply



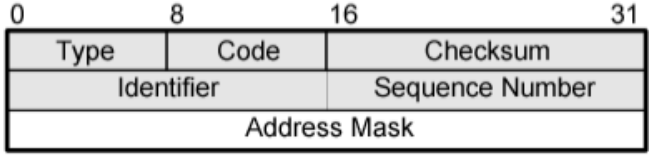
(e) Timestamp



(f) Timestamp Reply



(g) Address Mask Request



(h) Address Mask Reply

Using ICMP – Ping

- Test destination reachability
- Source sends echo request to a remote host or router
- If remote system receives the ICMP packet, it sends back an echo reply to source
- The ping utility may further do
  - Calculate round-trip time
  - Count the number of hops to destination (use TTL)

Traceroute

www.traceroute.org

traceroute: gaia.cs.umass.edu to www.eurecom.fr

Three delay measurements from gaia.cs.umass.edu to cs-gw.cs.umass.edu

```
1  cs-gw (128.119.240.254)  1 ms  1 ms  2 ms
2  border1-rt-fa5-1-0.gw.umass.edu (128.119.3.145)  1 ms  1 ms  2 ms
3  cht-vbns.gw.umass.edu (128.119.3.130)  6 ms 5 ms 5 ms
4  jn1-at1-0-0-19.wor.vbns.net (204.147.132.129)  16 ms 11 ms 13 ms
5  jn1-so7-0-0-0.wae.vbns.net (204.147.136.136)  21 ms 18 ms 18 ms
6  abilene-vbns.abilene.ucaid.edu (198.32.11.9)  22 ms  18 ms  22 ms
7  nycm-wash.abilene.ucaid.edu (198.32.8.46)  22 ms  22 ms  22 ms
8  62.40.103.253 (62.40.103.253)  104 ms 109 ms 106 ms
9  de2-1.de1.de.geant.net (62.40.96.129)  109 ms 102 ms 104 ms
10  de.fr1.fr.geant.net (62.40.96.50)  113 ms 121 ms 114 ms
11  renater-gw.fr1.fr.geant.net (62.40.103.54)  112 ms  114 ms  112 ms
12  nio-n2.cssi.renater.fr (193.51.206.13)  111 ms  114 ms  116 ms
13  nice.cssi.renater.fr (195.220.98.102)  123 ms  125 ms  124 ms
14  r3t2-nice.cssi.renater.fr (195.220.98.110)  126 ms  126 ms  124 ms
15  eurecom-valbonne.r3t2.ft.net (193.48.50.54)  135 ms  128 ms  133 ms
16  194.214.211.25 (194.214.211.25)  126 ms  128 ms  126 ms
17  * * *
18  * * *
19  fantasia.eurecom.fr (193.55.113.142)  132 ms  128 ms  136 ms
```

trans-oceanic link

* means no response (probe lost, router not replying)

Using ICMP – Traceroute

- Measures the number of hops required to reach a destination
- Source sends 1st IP (UDP) packet with the TTL value set to 1
- The first router decrements the TTL to 0, discards the packet, sends a TTL expired back
- Source calculates RTT, and repeat 3 times
- Source sends 2nd IP packet with the TTL set to 2
- The second router will send back a TTL expired
- Source calculates RTT, and repeat 3 times
- Source repeats this with increasing TTL until destination is reached (or host unreachable)
- May suffer from dynamic routing (how?)

Using ICMP – Path MTU

- Determines the minimum MTU along the path to destination
- Source sends a large IP packet with don't fragment bit set
- If packet too large, relevant router will send back a parameter unintelligible
- Source decrements the packet length accordingly and tries again
- Until the packet reaches destination without ICMP error message
- Also suffer from dynamic routing

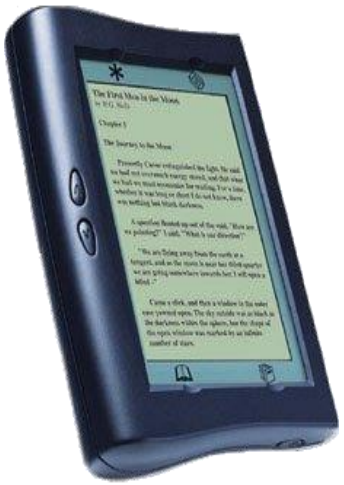
Mobile IP

Mobile IP

- Mobile IP standard
  - Approved by the Internet Engineering Steering Group (IESG) in June 1996
  - Published as a proposed standard by the Internet Engineering Tasks force (IETF) in November 1996

- Developed in order to cope with the increasing popularity of PDA's and Laptop's

Mobile Devices



Need for Mobile IP

- Datagram moved from one network to the other by routers, which use destination's IP addresses
- IP address is divided into two parts: <netID, hostID>

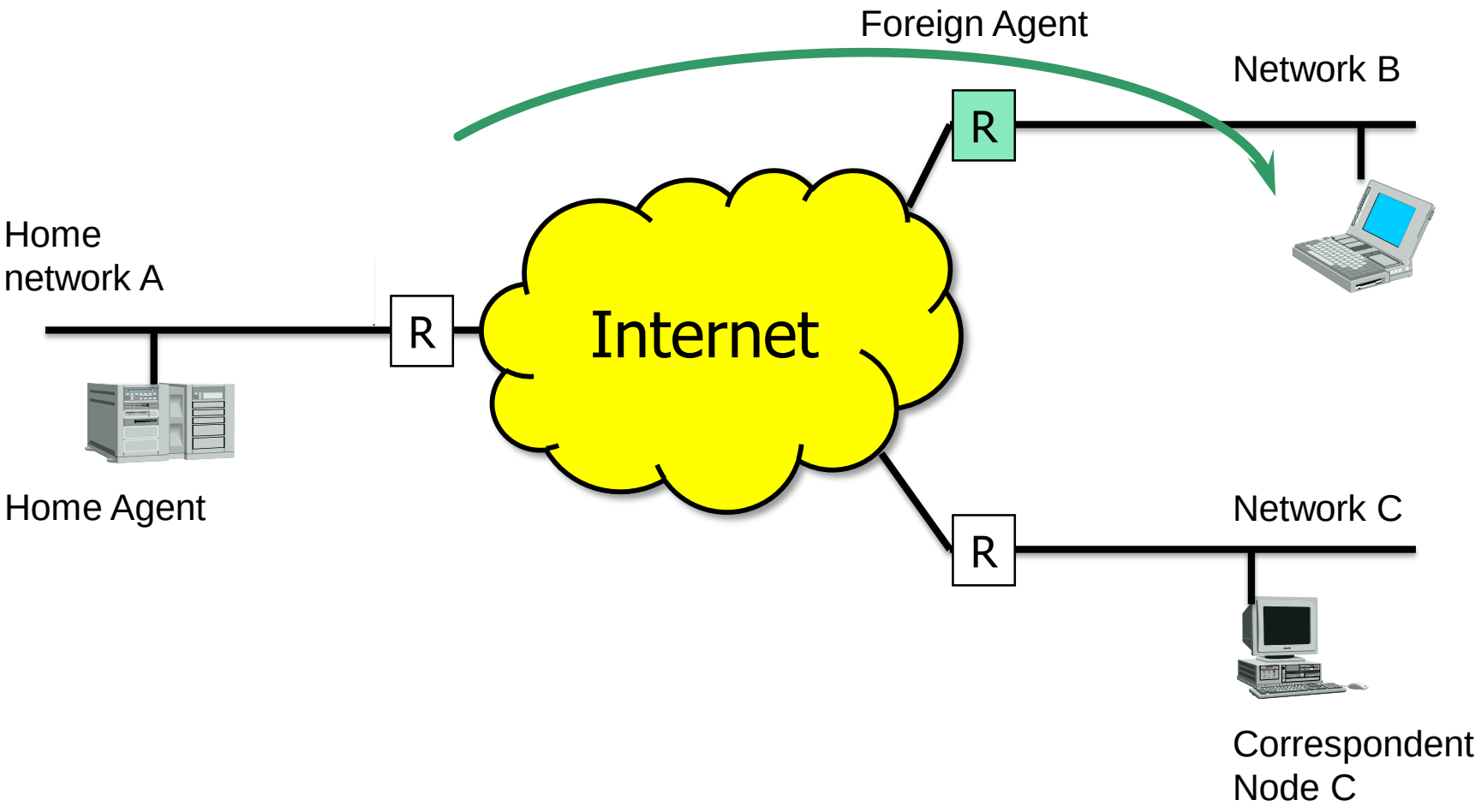
- Most applications over the Internet are supported by TCP connections
- TCP uses IP address and port number for routing and delivery

Need for Mobile IP

- As a mobile device moves from one network to the other, its IP address changes dynamically
- Thus the TCP connection needs to restart any ongoing communications each time it moves

- Mobile IP is to deal with the problem of dynamically varying IP addresses
- No need to change the TCP, i.e. IP address of the mobile device is pretend to be unchanged

An Illustration

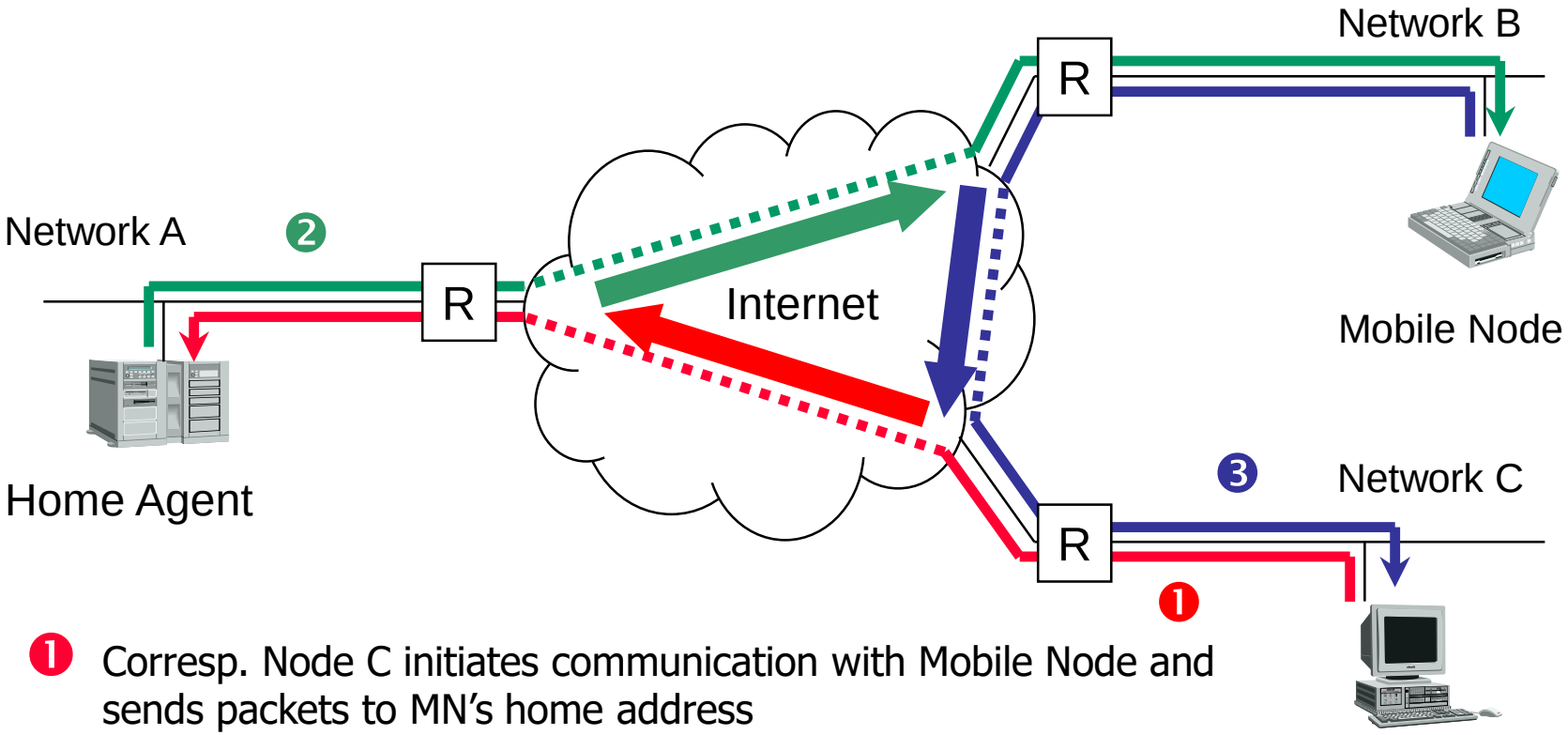


Different Entities

移动节点
通信节点
归属代理
外部代理

- Mobile Node
  - A host that may change its point of attachment from one network to the other
- Correspondent Node
  - A host that sends a packet addressed to a mobile node
- Home Agent
  - A node on the home network that maintains a list of registered mobile nodes
- Foreign Agent
  - A router on a foreign network that assists a mobile node in delivering datagram

Triangle Routing



1. Corresp. Node C initiates communication with Mobile Node and sends packets to MN's home address
2. Home Agent intercepts packets and forward them to the Mobile Node (proxy functionality)
3. Mobile Node replies directly to Corresp. Node C

The Protocol

- Mobile IP includes 3 capabilities
  - Discovery
  - Registration
  - Tunneling

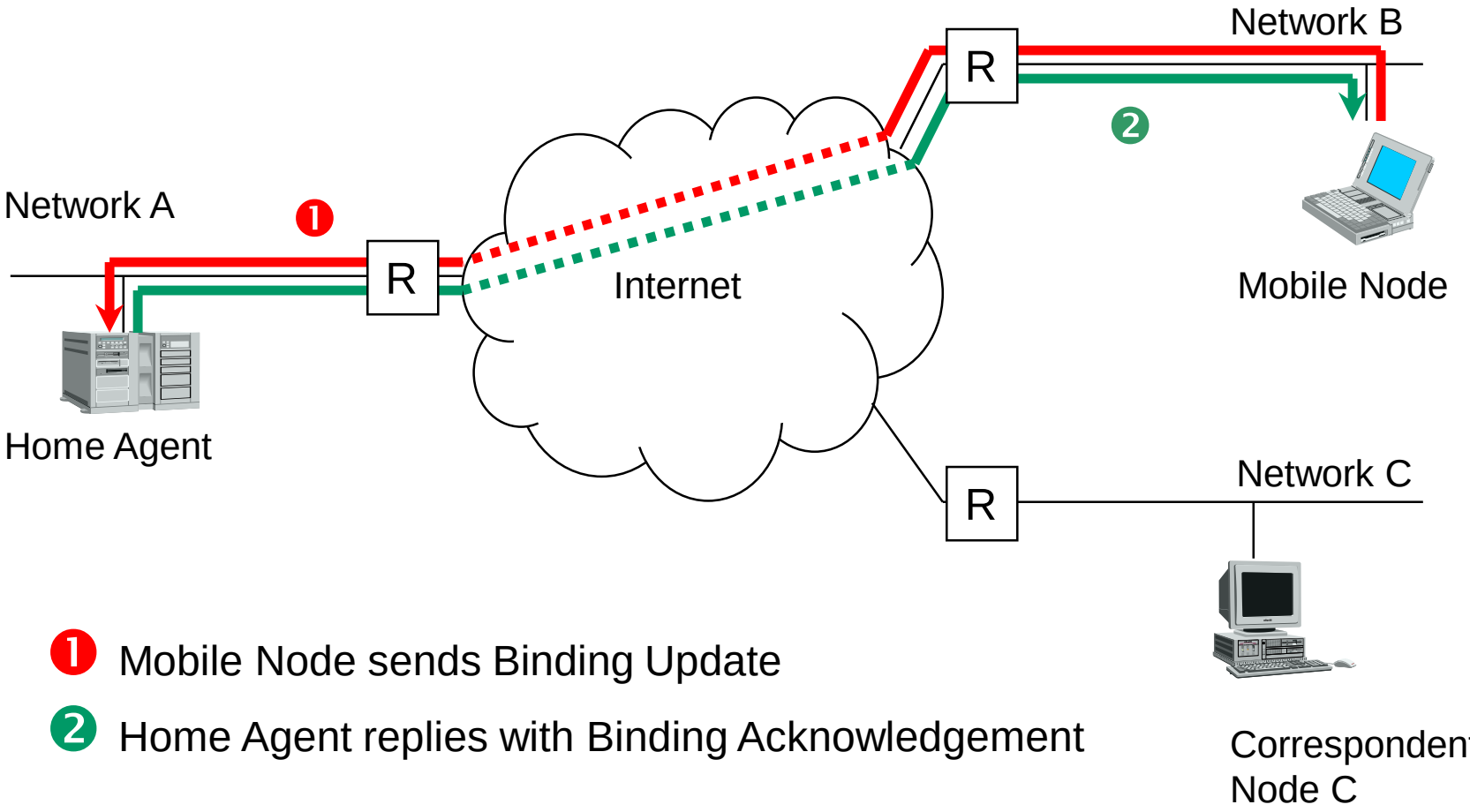
Discovery

- Mobile (Foreign) Agents
  - Send ICMP router advertisements with mobility agent advertisement extension periodically informing its presence
- Mobile node
  - Optionally request an advertisement from an agent
  - Or simply wait for the next advertisement

Registration

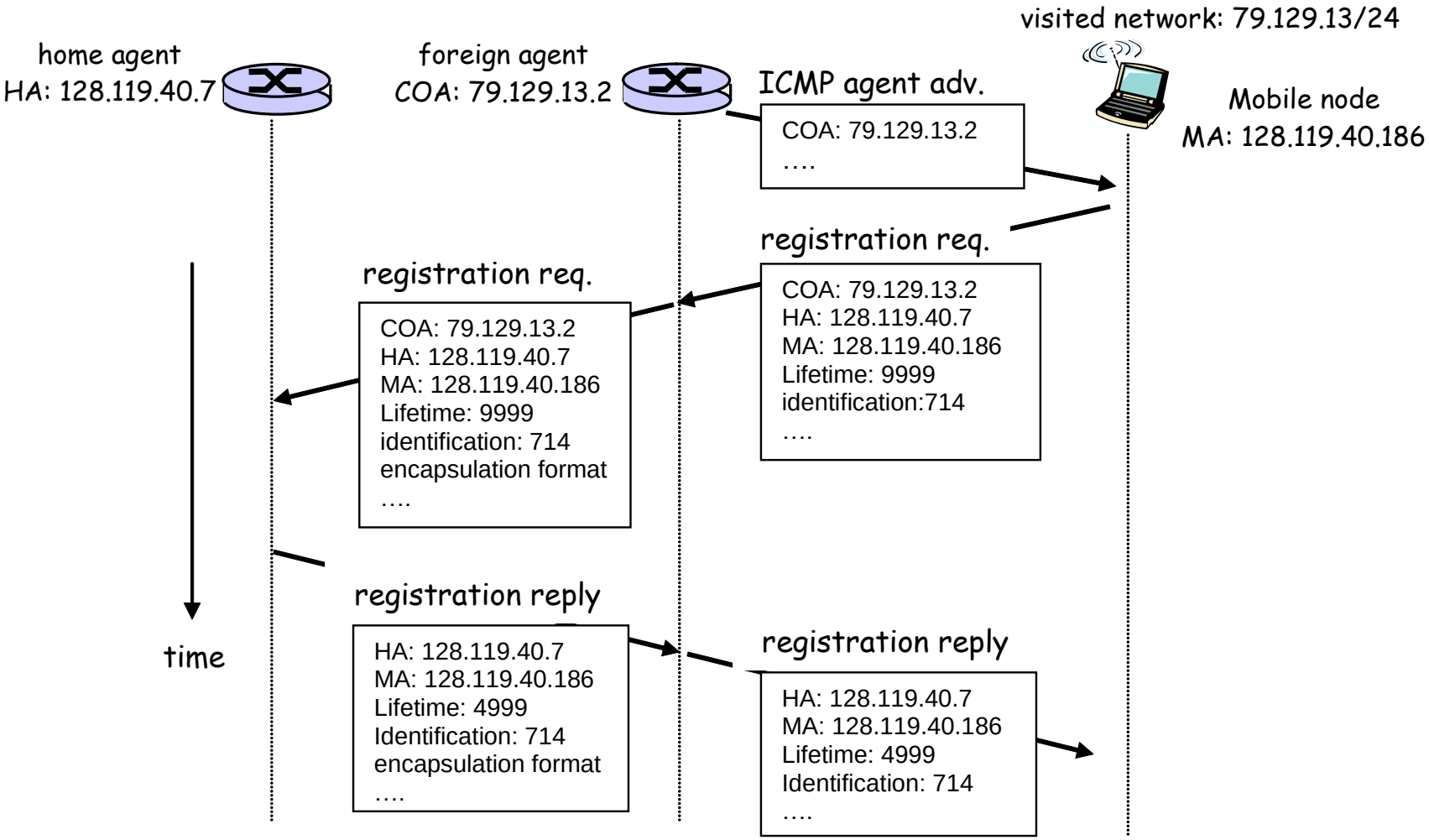
- Mobile node
  - Acquires a Care-of-Address from the foreign agent
  - Requests its home agent to forward its data packets to the foreign agent
- 4 steps
  - Mobile node sends registration request to the foreign agent
  - Foreign agent relays this request to the home agent
  - Home agent sends registration reply to the foreign agent
  - Foreign agent relays this reply to the mobile node

Registration of Mobile Node



1. Mobile Node sends Binding Update
2. Home Agent replies with Binding Acknowledgement

A Registration Example



Tunneling

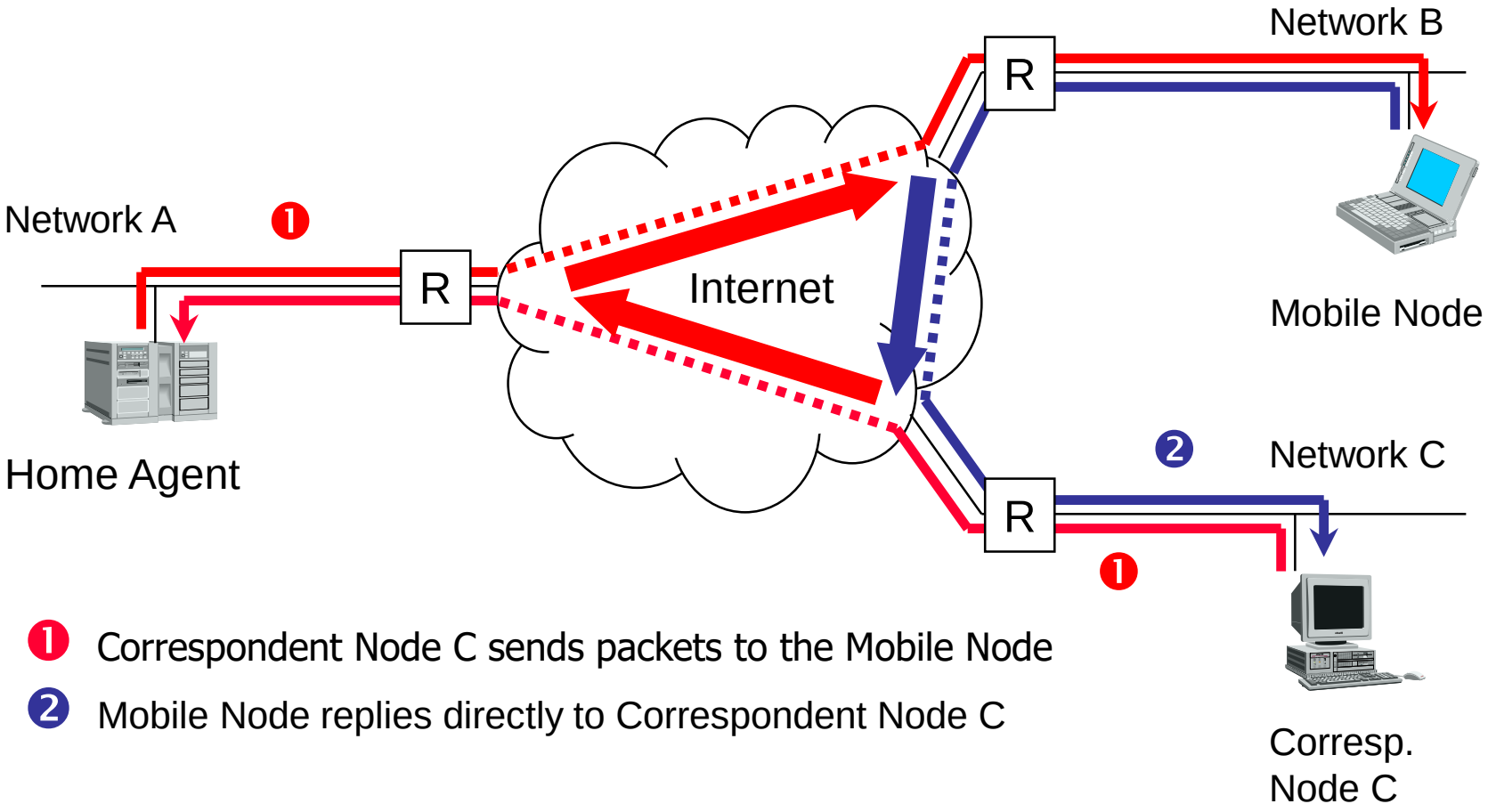
- After registration, an IP tunnel is set up
  - Between the home agent and care-of-address of the mobile node
  - Home agent broadcasts gratuitous ARP request which binds the mobile nodes IP address to the home agents MAC address
  - Thus home agent receives packets destined to the mobile node, and forwards the packets to the foreign agent through the IP tunnel

Tunneling

- For a correspondent node
  - Assumes the reply from the mobile node is coming from its home network
  - Continues to send the packet to the home agent

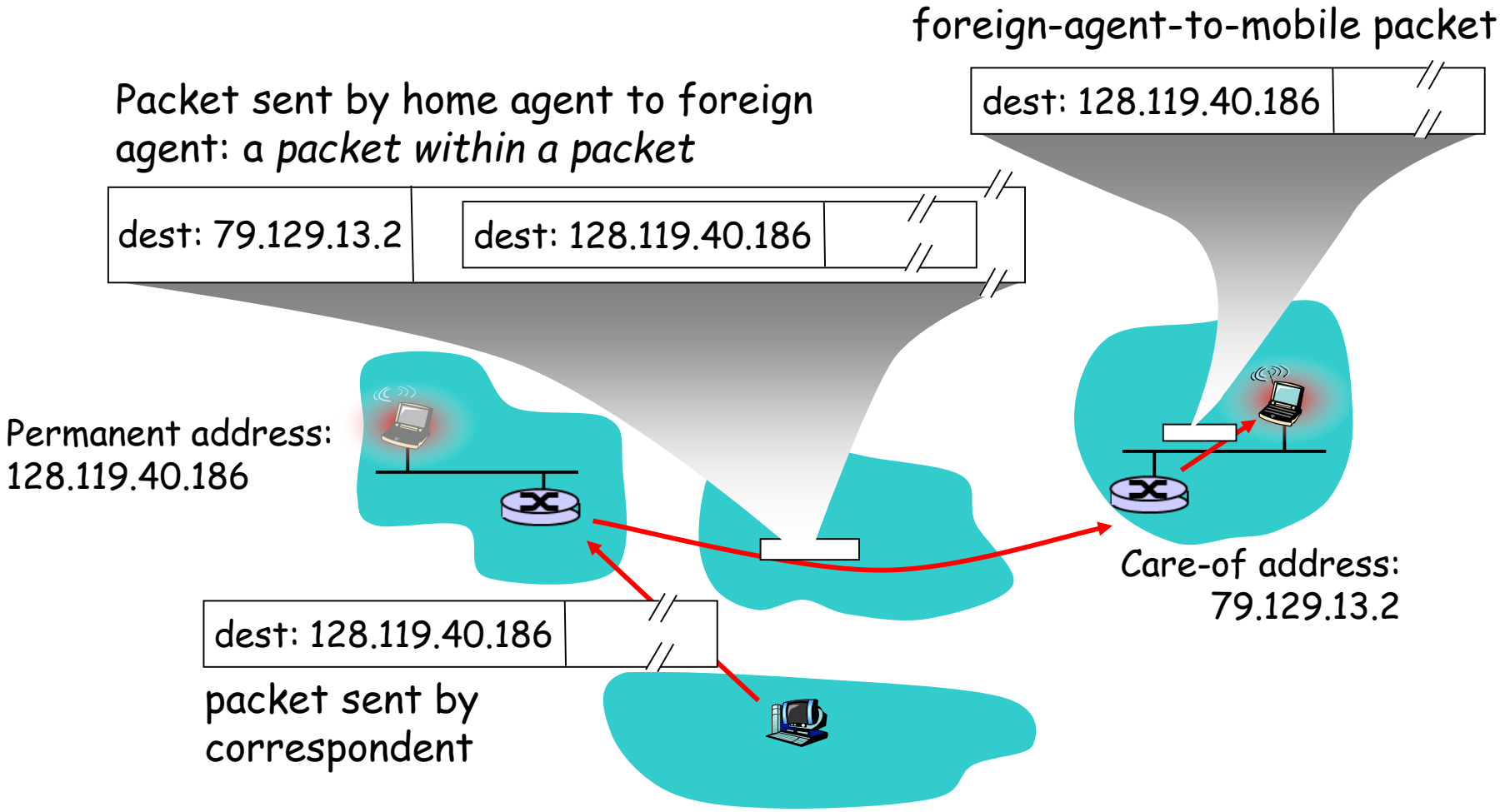
- Thus the TCP connection is maintained without changing the MN's IP address

IP Tunneling



1. Correspondent Node C sends packets to the Mobile Node
2. Mobile Node replies directly to Correspondent Node C

Indirect Routing



IPv6

IPv6

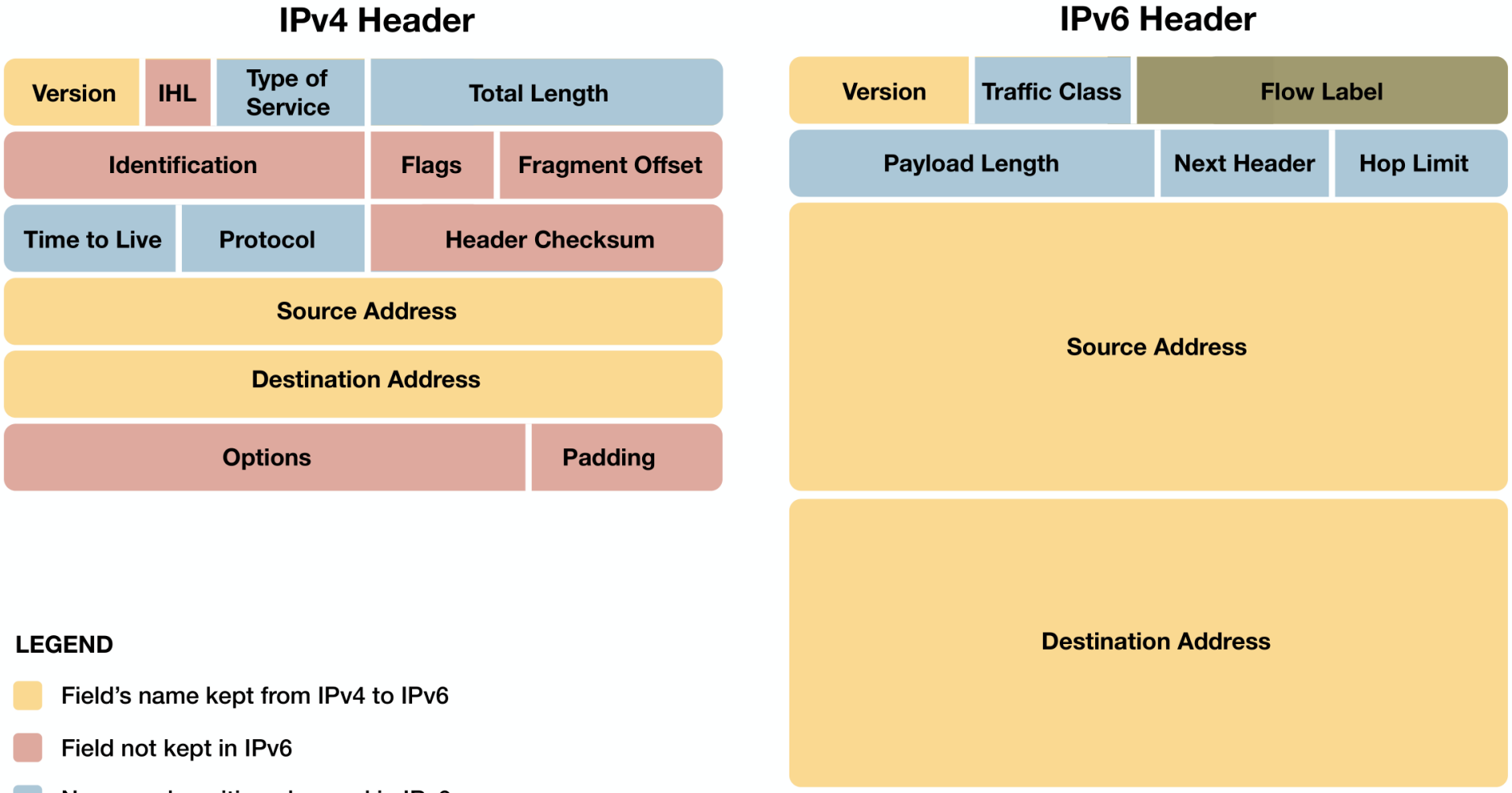
- Initial motivation: address space exhaustion
  - Rapid growth of networks and the Internet
  - 32-bit address space (esp. net address) soon to be completely allocated
- Additional motivation
  - New header format helps speed processing and forwarding
  - Header changes to facilitate QOS
  - No fragmentation at router
  - New address mode: route to "best" of several replicated servers

IPv6 RFCs

- 1752 – Recommendations for the IP Next Generation Protocol
- 2460 – Overall specification
- 2373 – addressing structure

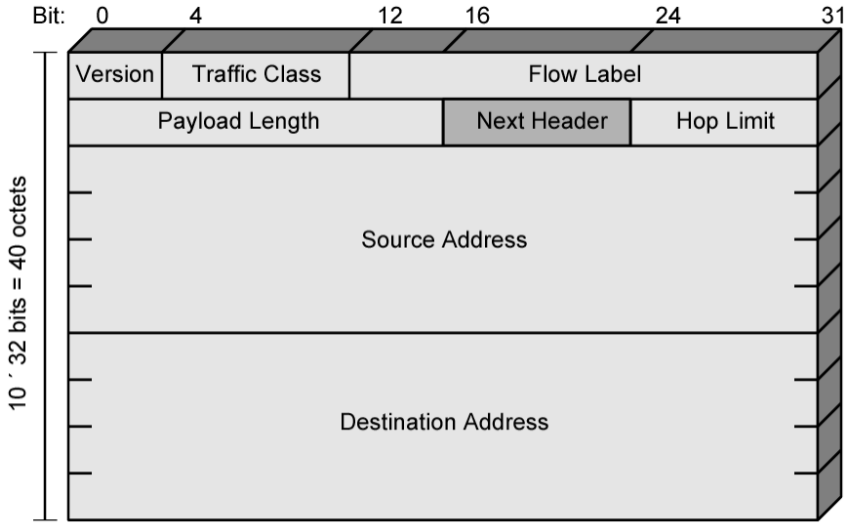
- Others (www.rfc-editor.org)
  - 1981 – Path MTU Discovery for IPv6
  - 2401 – Security Architecture for the Internet Protocol
  - 2402 – IP Authentication Header
  - 2406 – IP Encapsulating Security Protocol (ESP)
  - 2463 – ICMP for IPv6
  - …

IPv6 Header VS IPv4 Header

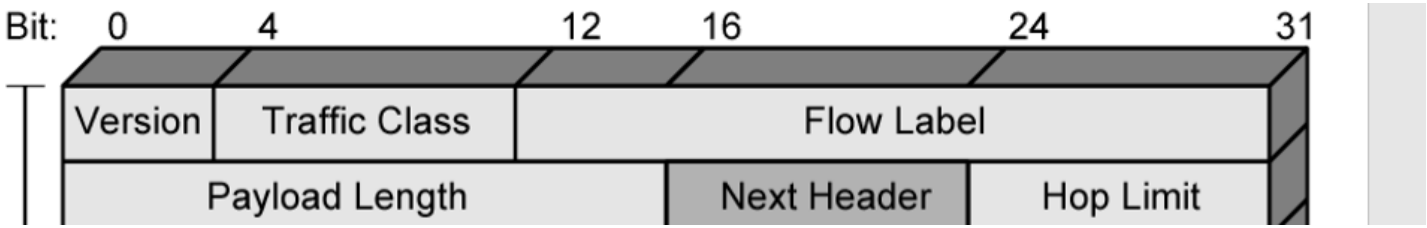


IPv6 Header Fields

- Version (4 bits): 6
- Traffic Class (8 bits)
  - Classes or priorities of packet, identify QoS
- Flow Label (20 bits)
  - Identify datagrams in the same "flow"
- Payload length (16 bits)
  - Includes all extension headers plus user data
- Next Header (8 bits)
  - Identifies type of the next header
  - Extension or next layer up
- Source / Destination Address (128 bits)

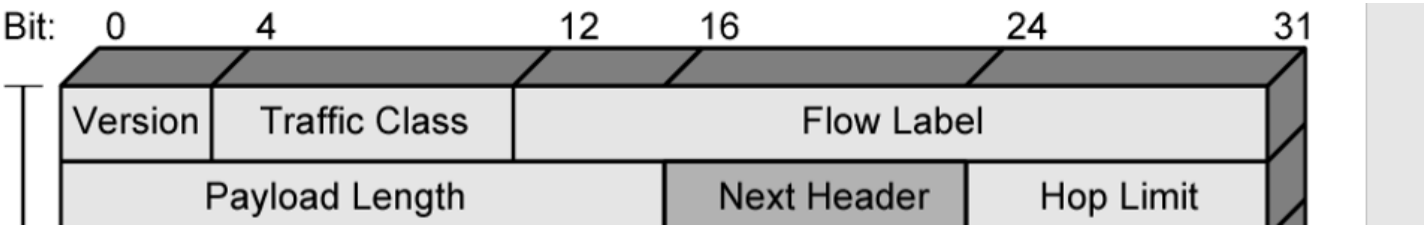


Traffic Class



- The 8-bit field in the IPv6 header is available for use by originating nodes and/or forwarding routers to identify and distinguish between different classes or priorities of IPv6 packets.
  - E.g., used as the codepoint in DiffServ
- General requirements
  - Service interface must provide means for upper-layer protocol to supply the value of traffic class
  - Value of traffic class can be changed by source, forwarder, receiver
  - An upper-layer protocol should not assume the value of traffic class in a packet has not been changed.

IPv6 Flow

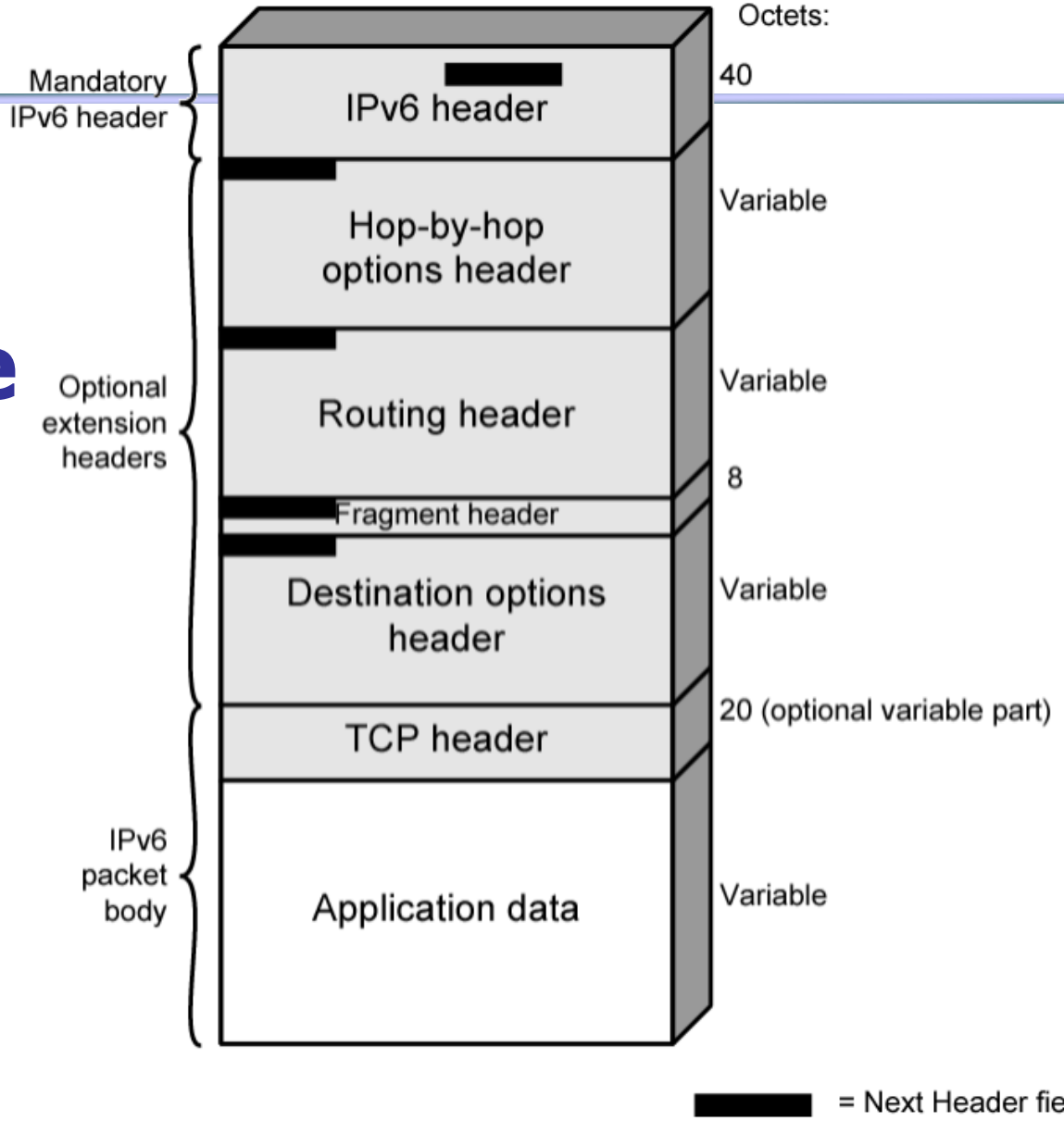


- A sequence of packets sent from a particular source to a particular destination
- From hosts point of view
  - Generated from one application and have the same transfer service requirements
  - May comprise a single or multiple TCP connections
  - One application may generate a single flow or multiple flows
- From routers point of view
  - Share attributes that affect how these packets are handled by the router
  - e.g. routing, resource allocation, discard requirements, accounting, and security

Flow Label

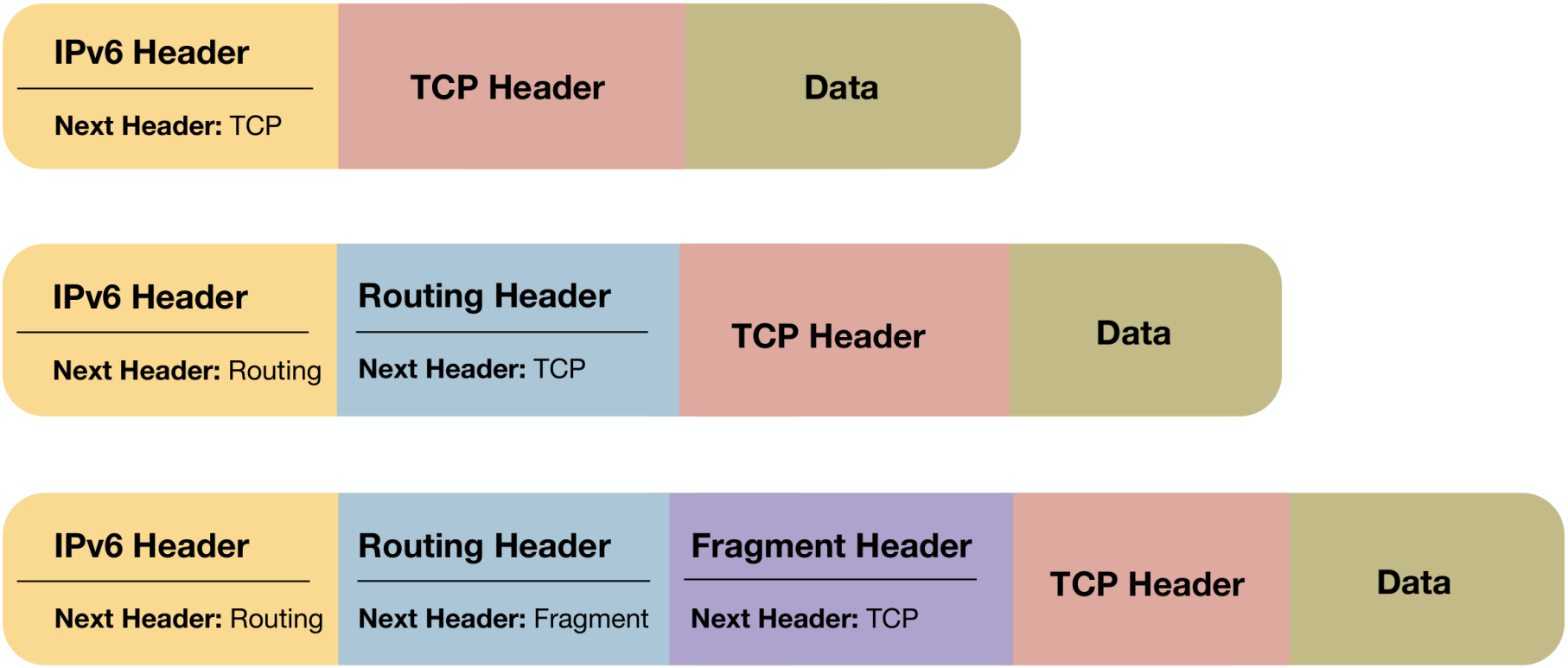
- A flow is uniquely identified by the combination of
  - Source and destination address
  - A non-zero 20-bit Flow Label
- Flow requirements are defined prior to flow commencement
  - Then a unique Flow Label is assigned to the flow
- Router decide how to route and process the packet by
  - Simply looking up the Flow Label in a table and without examining the rest of the header

IPv6 Header Structure



Extension Headers

- Optional fields go into extension headers
- Daisy-chained after the main header



Extension Headers

Appeared in order

- Hop-by-Hop Options: Require processing at each router
- Routing: Source routing
- Fragment: source fragmentation

- Authentication
- Encapsulating security payload

- Destination options: handle at destination

Fragmentation

- Routers don't fragment packets with IPv6
  - More efficient handling of packets in the core
  - Fragmentation is being done by host

- If a packet is too big for next hop:
  - "Packet too big" error message
  - This is an ICMPv6 message
  - Filtering ICMPv6 causes problems

Broadcast

- IPv6 has no broadcast
- There is an “all nodes” multicast group
  - ff02::1

- Disadvantages of broadcast:
  - It wakes up all nodes
  - Only a few devices are involved
  - Can create broadcast storms

Neighbor Discovery

- IPv6 has no ARP
- Replacement is called Neighbor Discovery
  - Uses ICMPv6
  - Uses Multicast

- Neighbor Discovery is used by nodes:
  - For address resolution
  - To find neighboring routers
  - To track address changes
  - To check neighbor reachability
  - To do Duplicate Address Detection

IPv6 Addresses

- 128 bits long, assigned to interface

FEDC : BA98 : 7654 : 3210 : FEDC : BA98 : 7654 : 3210
1080 : 0 : 0 : 0 : 8 : 800 : 200C : 417A

- Single interface may have multiple unicast addresses

- 3 types of address defined
  - Unicast, Multicast, Anycast

Example IPv6 Addresses

- Different IPv6 addresses
  - A unicast address
    - 1080:0:0:0:8:800:200C:417A, simplified as 1080::8:800:200C:417A
  - A multicast address
    - FF01:0:0:0:0:0:0:101, simplified as FF01::101
  - The loopback address
    - 0:0:0:0:0:0:0:1, simplified as ::1
  - Unspecified addresses
    - 0:0:0:0:0:0:0:0, simplified as ::
- IPv4 address → IPv6 address
  - x:x:x:x:x:x:d.d.d.d, 2 possible ways
  - 0:0:0:0:0:0:13.1.68.3, simplified as ::13.1.68.3
  - 0:0:0:0:0:FFFF:129.144.52.38, simplified as ::FFFF:129.144.52.38

Summary of Header Changes

- 40 bytes
- Address increased from 32 to 128 bits
- Fragmentation and options fields removed from base header
- Header checksum removed
- Header length is only payload (because fixed length header
- New flow label field
- TOS→> Traffic Class
- Protocol -> Next Header(extension headers)
- Time To Live-> Hop Limit
- Alignment changed to 64 bits

Advantages of IPv6 over IPv4

- Expanded addressing capabilities
  - 128 bit
  - Scalability of multicast addresses
  - Anycast – delivered to one of a set of nodes
  - Address auto-configuration
- Improved option mechanism
  - Separate optional headers between IPv6 header and transport layer header
  - Most are not examined by intermediate routers
  - Easier to extend options
  - Checksum removed to further reduce processing time at each router

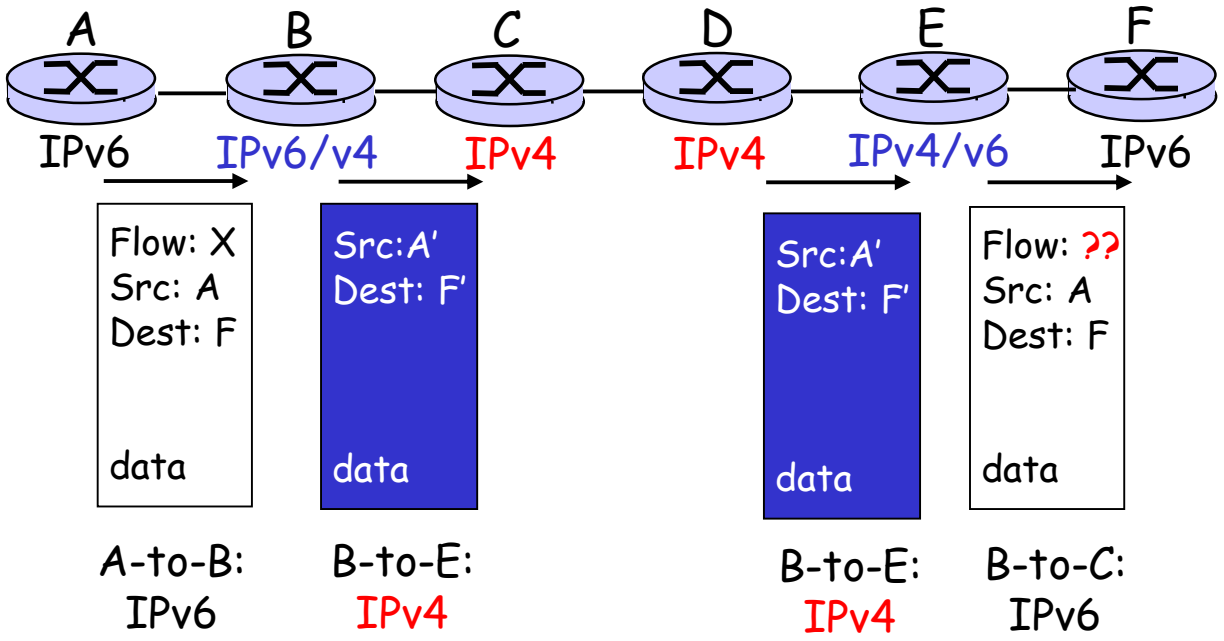
Advantages of IPv6 over IPv4

- Support for resource allocation
  - Uses traffic class
  - Grouping packets to particular traffic flow
  - Allows QoS handling other than best-effort, e.g. real-time video
- More efficient and robust mobility mechanism
- More security: Built-in, strong IP-layer encryption and authentication

Transition From IPv4 To IPv6

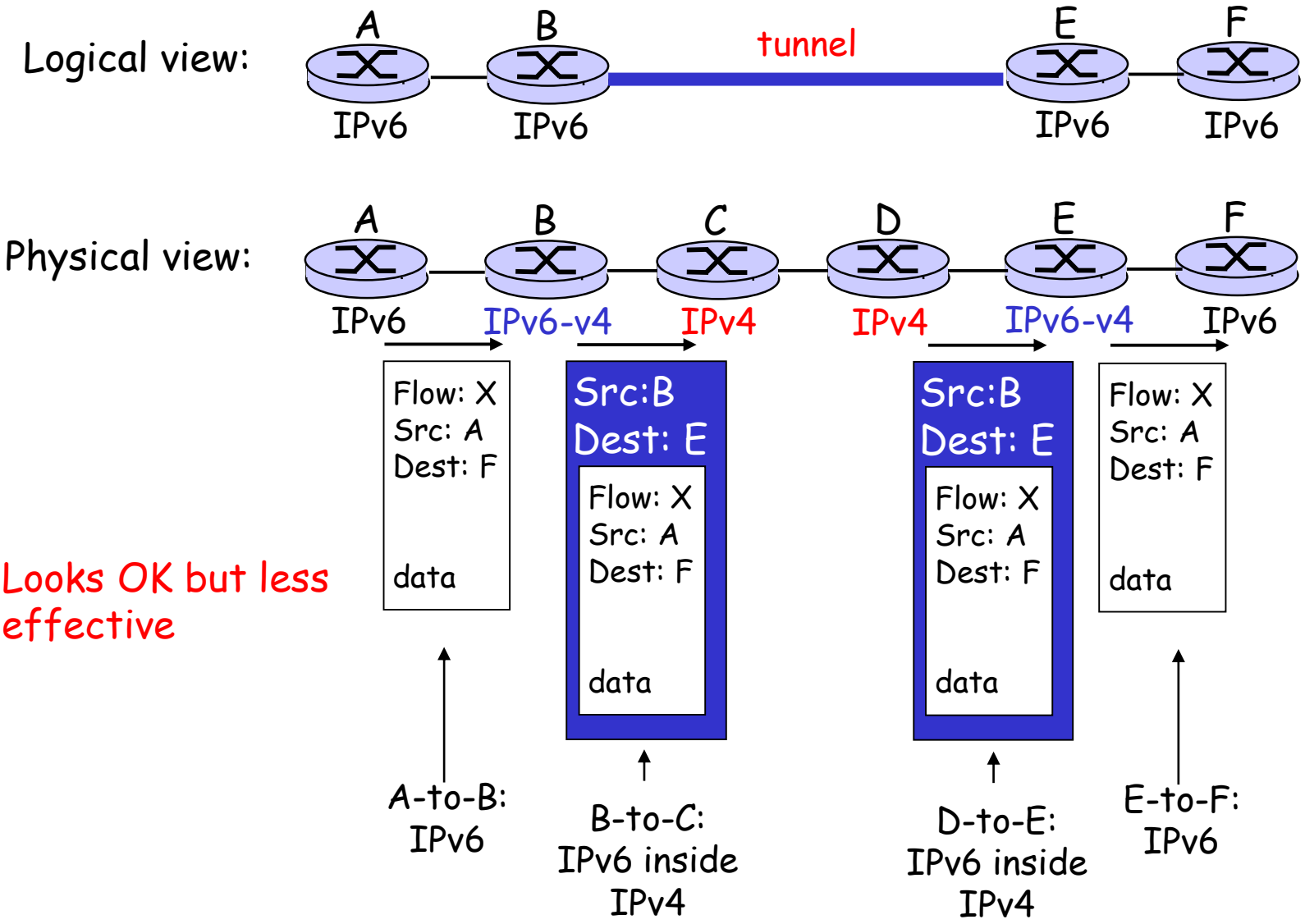
- Not all routers can be upgraded simultaneously
  - How will the network operate with mixed IPv4 and IPv6 routers
- Two proposed approaches
  - Dual Stack – some routers with dual stack (IPv6, IPv4) can translate between formats
  - Tunneling – IPv6 carried as payload in IPv4 datagram among IPv4 routers

Dual Stack Approach



- Address translation between IPv4 and IPv6 is needed
- Some IPv6 features is lost

Tunneling


Logical view:
A
B
tunnel
E
F
IPv6
IPv6
IPv6
IPv6
Physical view:
A
B
C
D
E
F
IPv6
IPv6-v4
IPv4
IPv4
IPv6-v4
IPv6
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Flow: X
Src: A
Dest: F
data
Looks OK but less effective
A-to-B:
IPv6
B-to-C:
IPv6 inside
IPv4
D-to-E:
IPv6 inside
IPv4
E-to-F:
IPv6

Summary

- NAT原理及优缺点
- ARP地址解析原理和流程
- DHCP动态地址获取的过程
- ICMP
  - 用于发送出错信息
  - Ping和traceroute的实现原理
- Mobile IP
  - 移动终端，归属代理，外部代理，隧道
  - 三角路由
- IPv6
  - 地址格式
  - 和IPv4的异同，优缺点？
  - V4和V6的融合