COMP 3234B
Computer and Communication Networks

2nd semester 2023-2024

Network Layer (II)

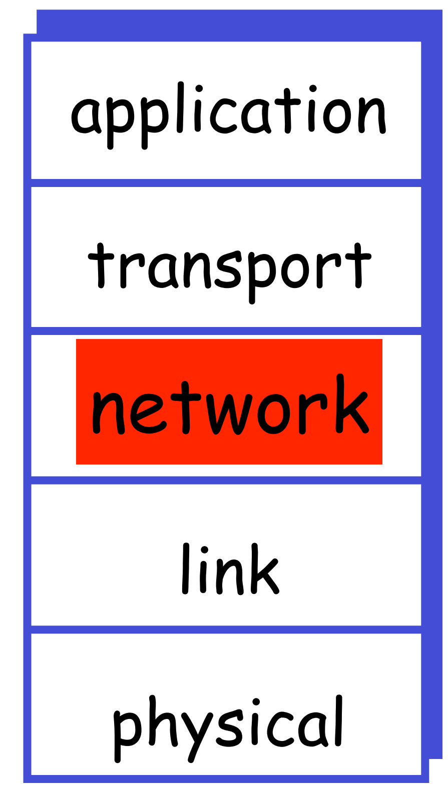
Prof. C Wu

Department of Computer Science
The University of Hong Kong

Roadmap

Network layer

- Principles behind network-layer services (ILO1)
  - forwarding vs. routing
  - network service models
- Router (ILO1)
- IP (ILO2,5)
  - DHCP
  - NAT
- ICMP (ILO2)
- Routing algorithms (ILO3)
- Routing in the Internet (ILO2,3)



IP assignment in the Internet — classful addressing

- In the original approach, IP addresses were divided into pre-defined classes:
  - class A: 8-bit network prefix (16,777,214 hosts)
  - class B: 16-bit network prefix (65,534 hosts)
  - class C: 24-bit network prefix (254 hosts)

Two addresses in each subnet are reserved for special purpose:
— all "0" host bits: used to identify the subnet
— all "1" host bits: used as broadcast address

Rapid depletion of blocks of IP addresses!

IP assignment in the Internet – CIDR

- Classless Interdomain Routing (CIDR): a.b.c.d/x
  - network addresses are allocated in 1-bit increments as opposed to 8-bits in classful network

    E.g., **216.3.128.12/25**

  - x most significant bits: network prefix (subnet portion of IP address)

    an organization is assigned contiguous addresses with a common prefix

    used by routers outside the organization's network

  - other bits: distinguishing interfaces within the organization

    used by routers within the organization's network

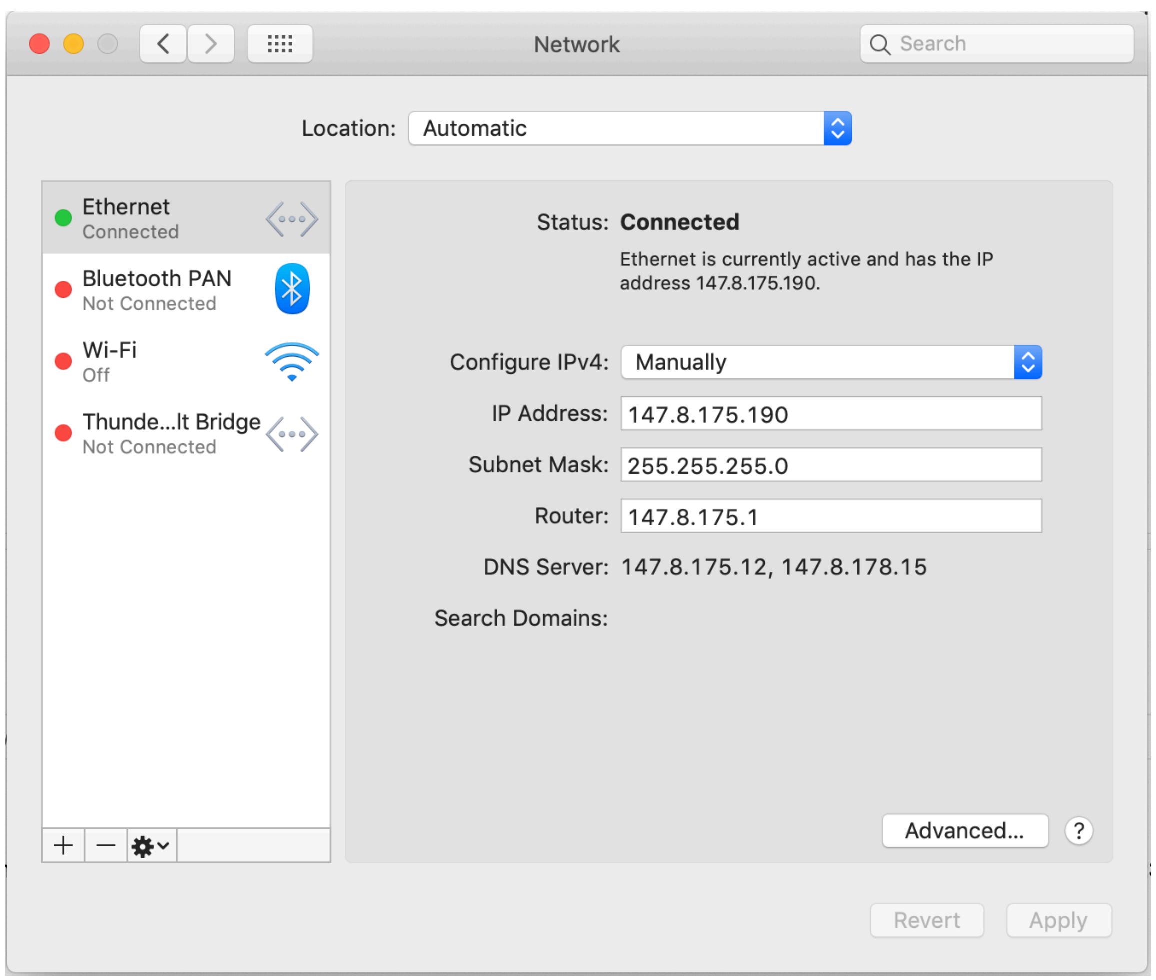
    may have further subnet structure:

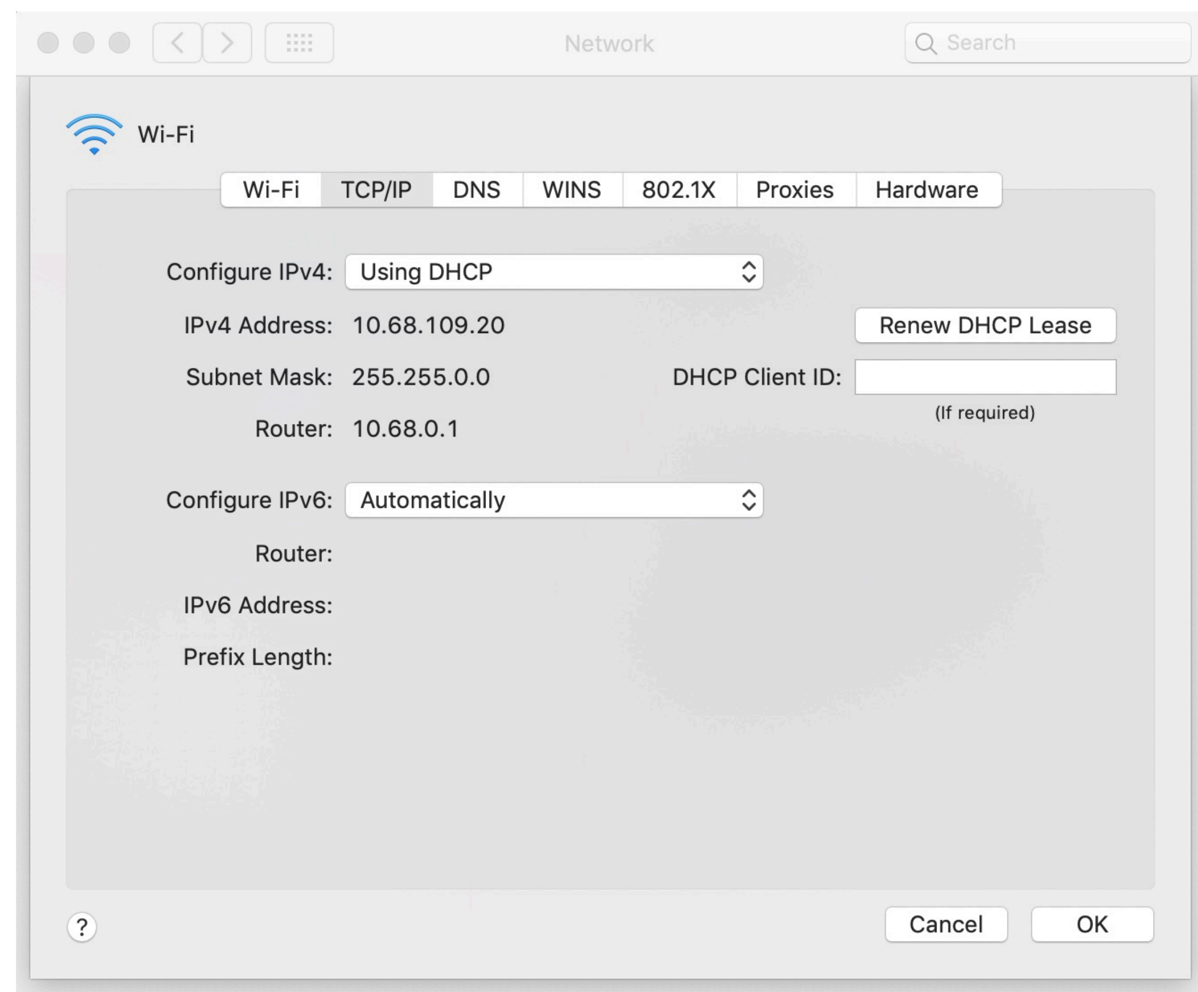
    e. g., a.b.c.d/21: an organization's network;

    a.b.c.d/24: a specific subnet within the organization

How does a host get IP address in its local network?

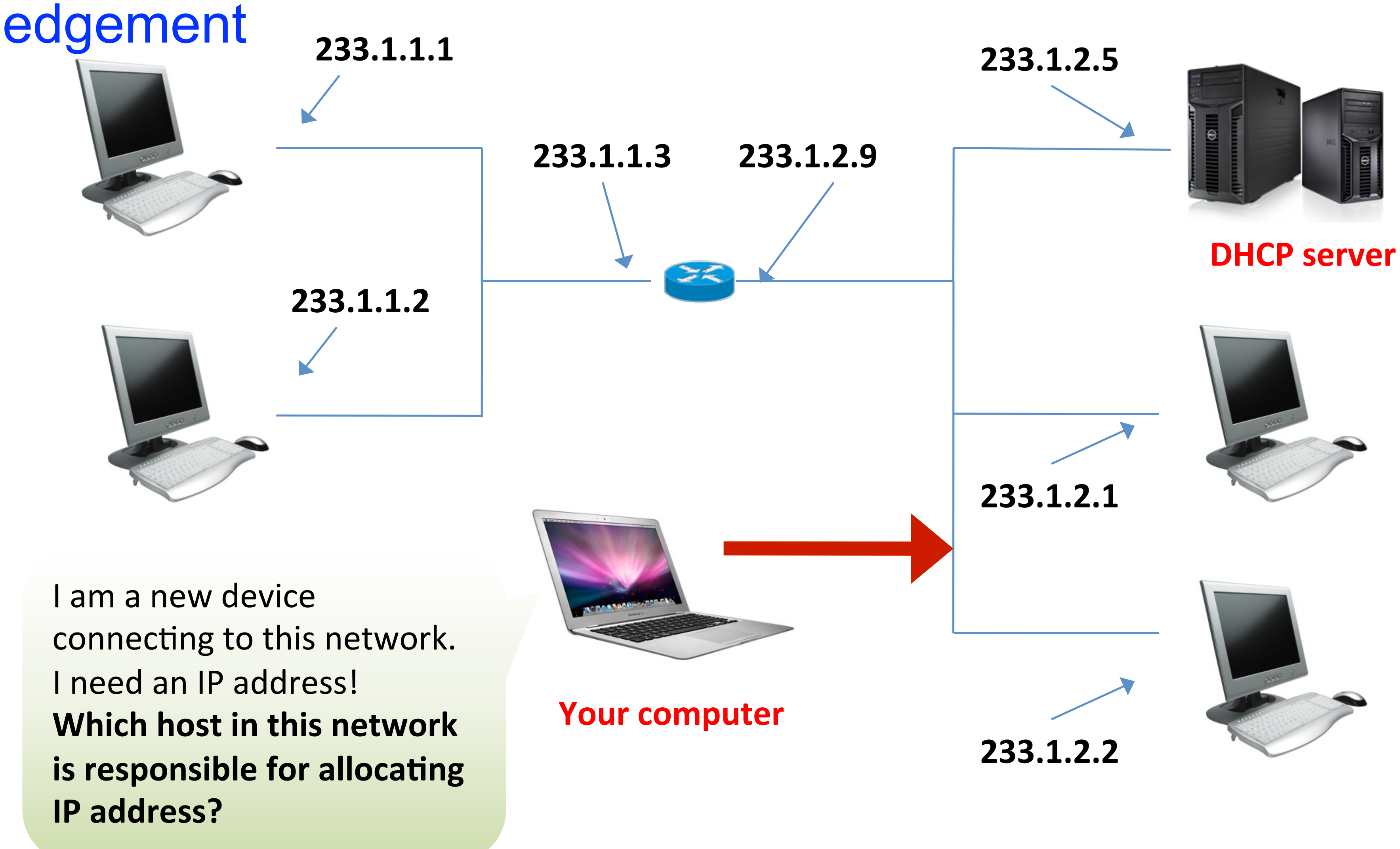
- Hardcode an IP address by system admin (the IP address is allocated from a network admin)
- Dynamically get an IP address from a server (use DHCP)





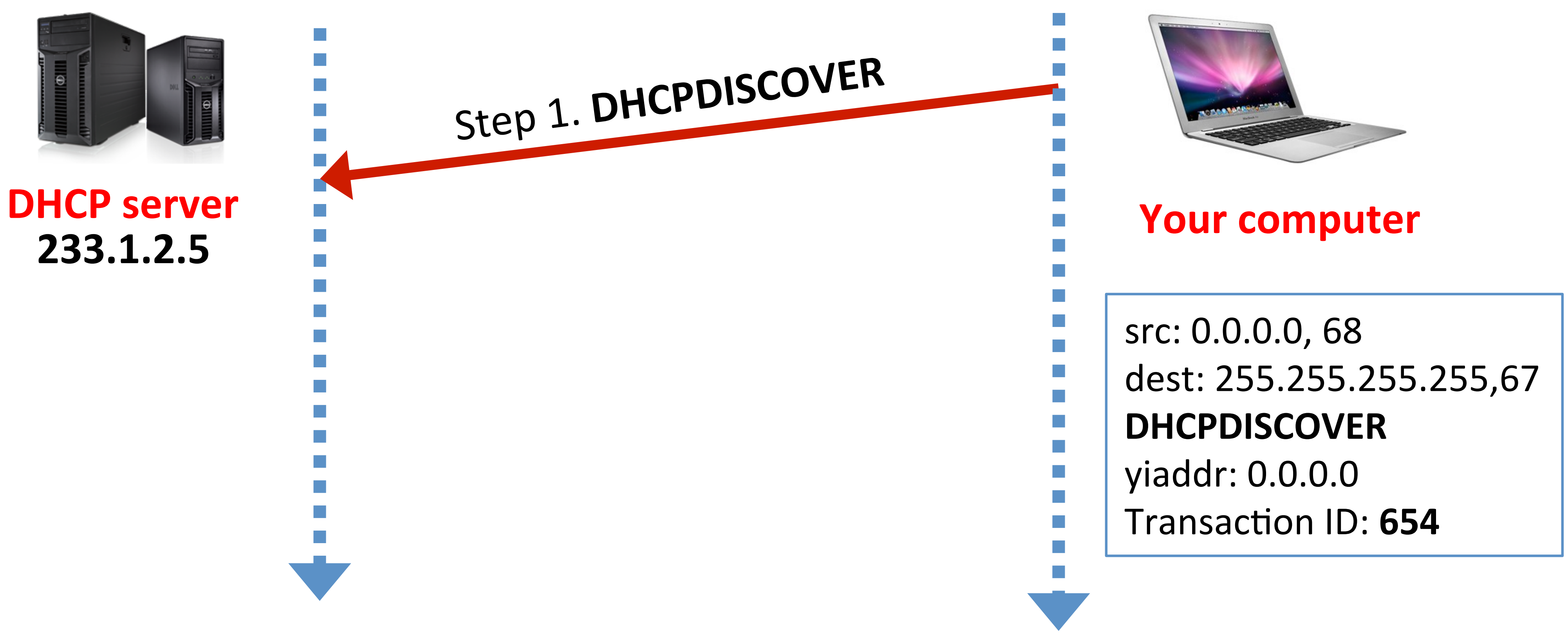
DHCP

- **D**ynamic **H**ost **C**onfiguration **P**rotocol (DHCP) is an application-layer protocol which uses UDP as the transport-layer protocol
- It includes 4 steps
  - Step 1: DHCP server discovery
  - Step 2: DHCP server offer(s)
  - Step 3: DHCP request
  - Step 4: DHCP acknowledgement



Step 1: DHCP server discovery

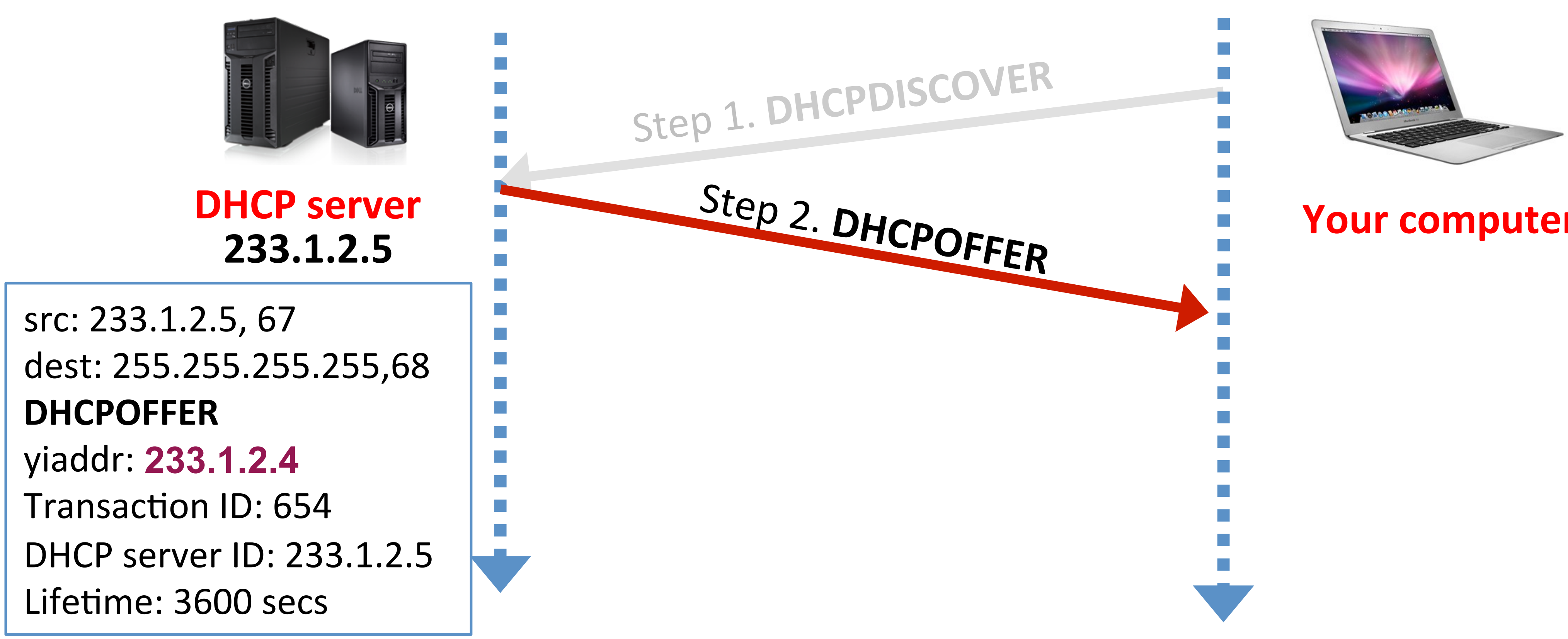
- The DHCP client creates an IP datagram containing its **DHCP discover message** along with the **broadcast** destination IP address of 255.255.255.255 and a source IP address of 0.0.0.0



Step 2: DHCP server offer(s)

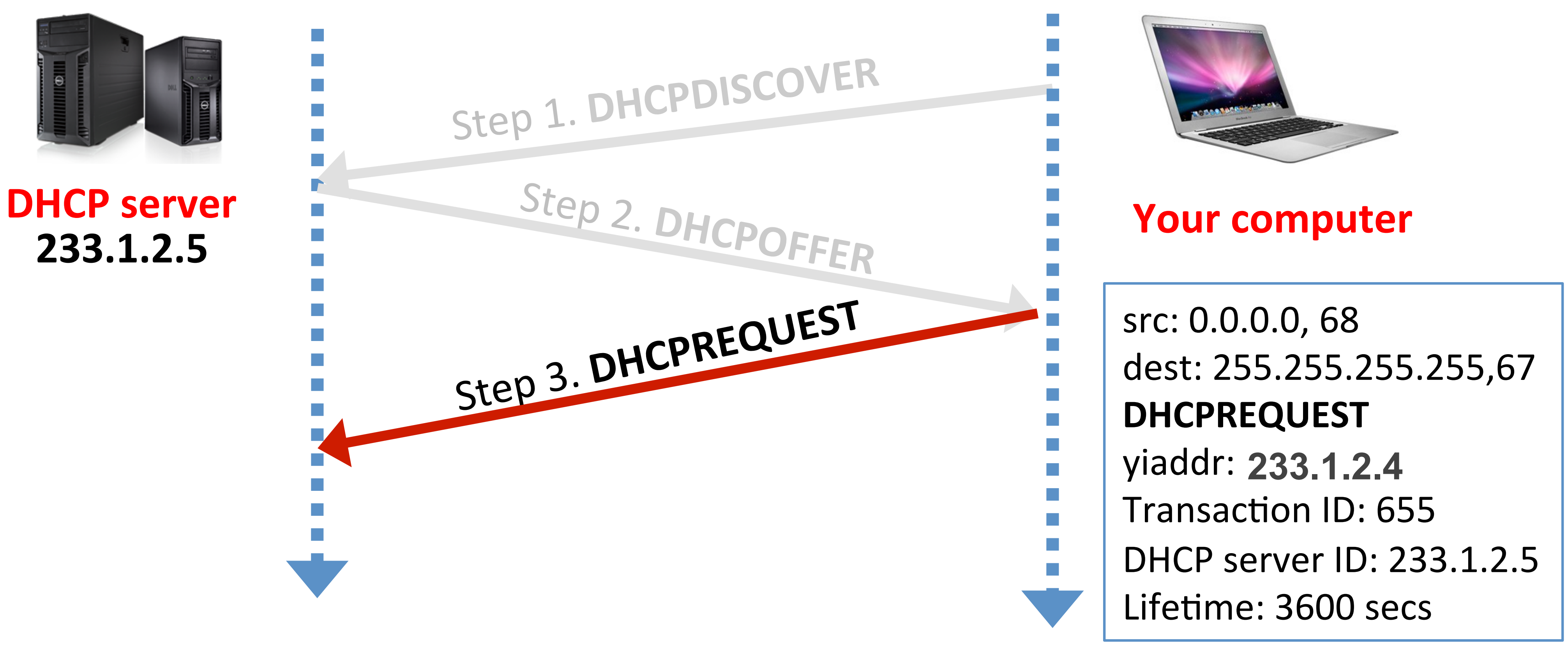
- A DHCP server receiving a DHCP discover message responds to the client with a **DHCP offer message** that is broadcast to all nodes on the subnet, again using the IP broadcast address of 255.255.255.255

- Each server offer message contains the **transaction ID of the received discover message**, the **proposed IP address** for the client, the **subnet mask**, an **IP address lease time,** and possibly other information (address of first-hop router, name and IP address of DNS sever, etc.)



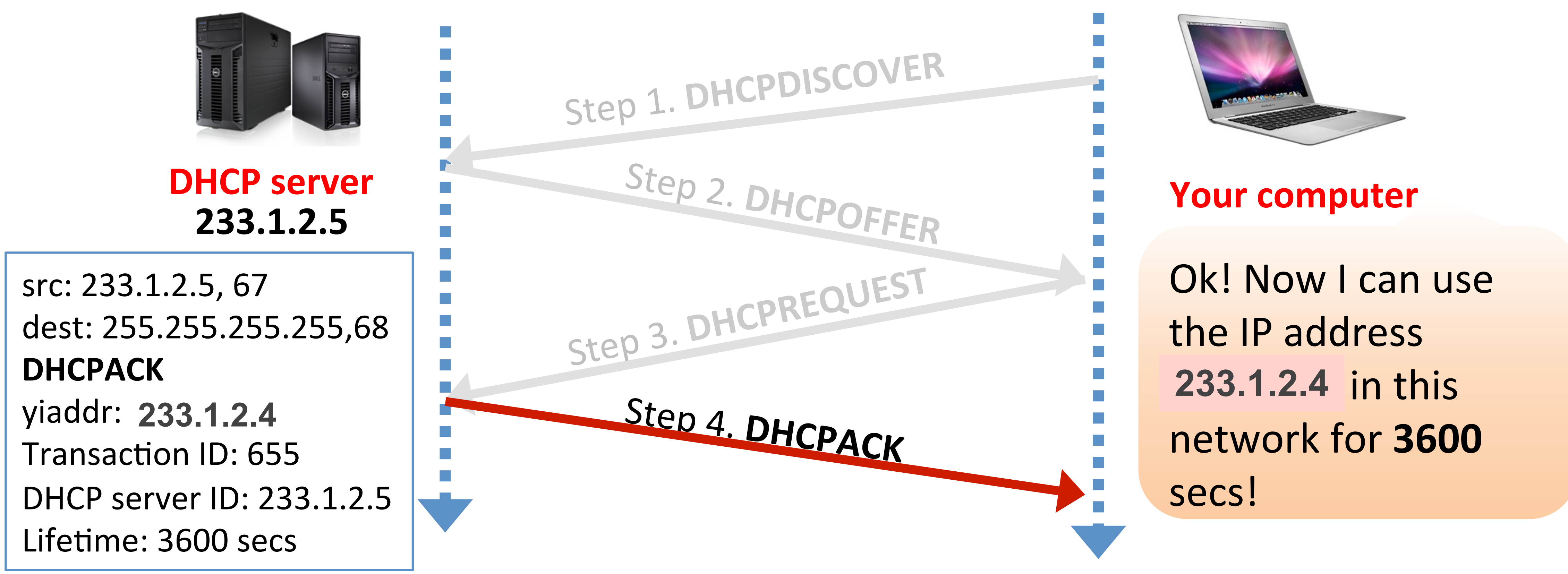
Step 3: DHCP request

- The newly arriving client will choose from among one or more server offers (there can be several DHCP servers on the same subnet) and respond to its selected offer with a **DHCP request message**, echoing back the configuration parameters.



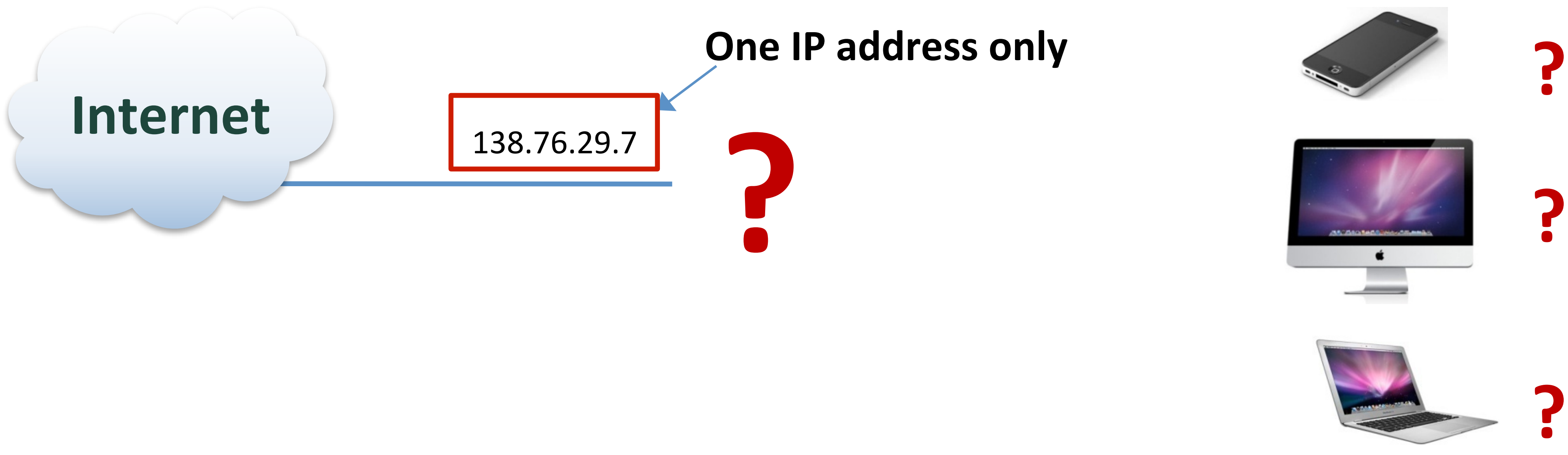
Step 4

- The server responds to the DHCP request message with a **DHCP ACK message**, confirming the requested parameters.



DHCP provides mechanism for client to renew the lease of IP address in use

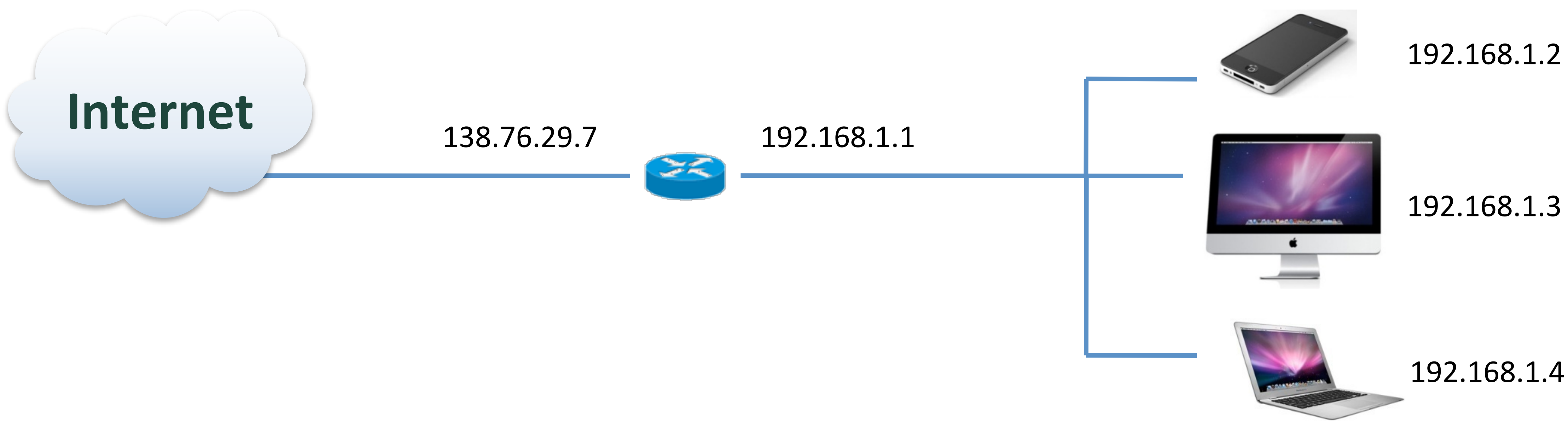
NAT (Network Address Translation)



- Question: if only one IP address is allocated to your home, how can you support multiple devices at home?
  - answer: you need an NAT-enabled router
- The NAT-enabled router has an interface that is part of the home network (LAN on the right) and another interface to the global Internet (WAN on the left)

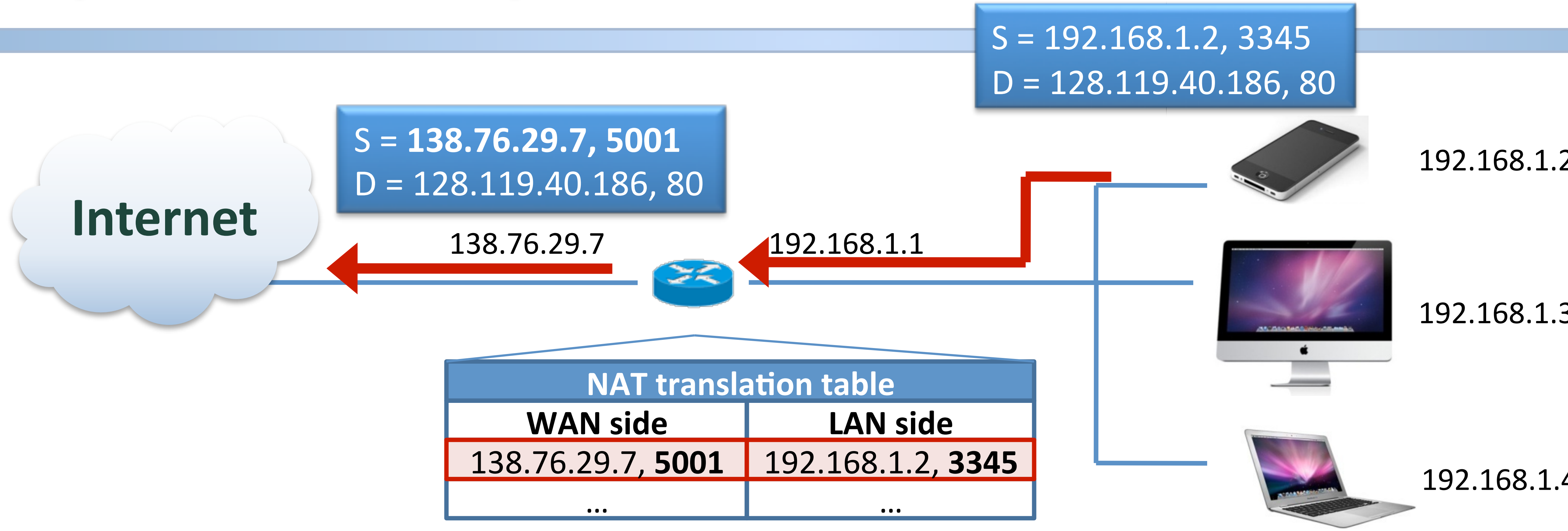
LAN: local area network WAN: wide area network

NAT



- According to RFC1918, the following three blocks of the IP addresses are reserved for **private networks**:
  - 10.0.0.0/8
  - 172.16.0.0/12
  - 192.168.0.0/16
- Each device in the home network is allocated a private IP address

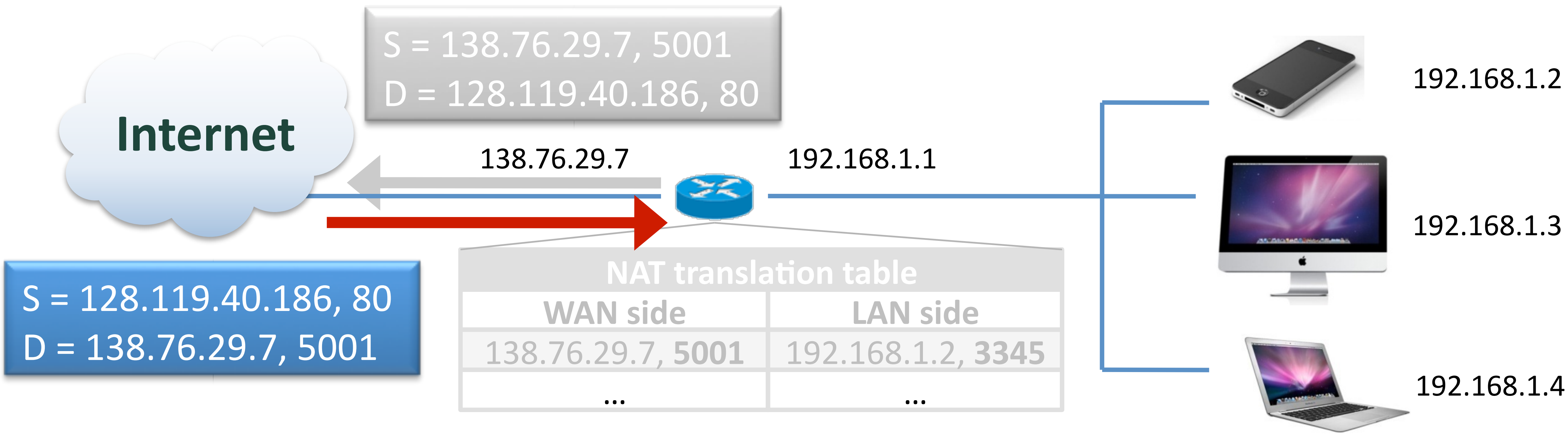
Request from the private network



NAT translation table	
WAN side	**LAN side**
138.76.29.7, **5001**	192.168.1.2, **3345**
…	…

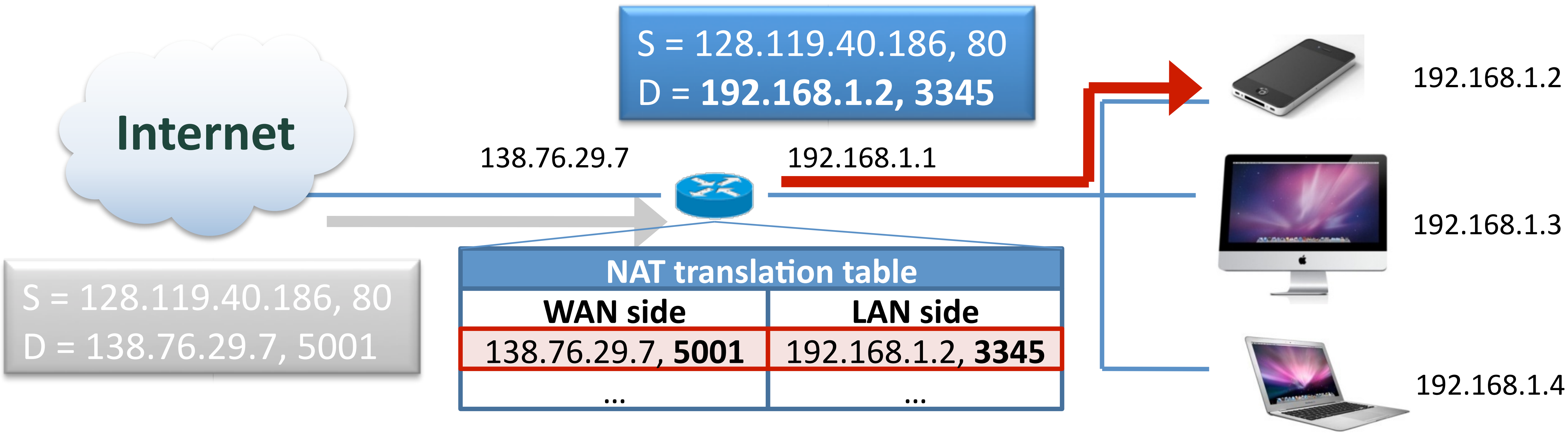
- The NAT-enabled router has an NAT translation table
- When user at host 192.168.1.2 requests a web page on some web server (IP: 128.119.40.186, port: 80), the NAT router:
  - replaces the source IP address with its WAN-side IP address 138.76.29.7
  - replaces the original source port number 3345 with the new source port number 5001
  - adds an entry to the router’s NAT translation table.

Request from the private network



- The web server is unaware that the arriving datagram containing the HTTP request has been manipulated by the NAT router, responds with a datagram whose destination address is 138.76.29.7 and port 5001.

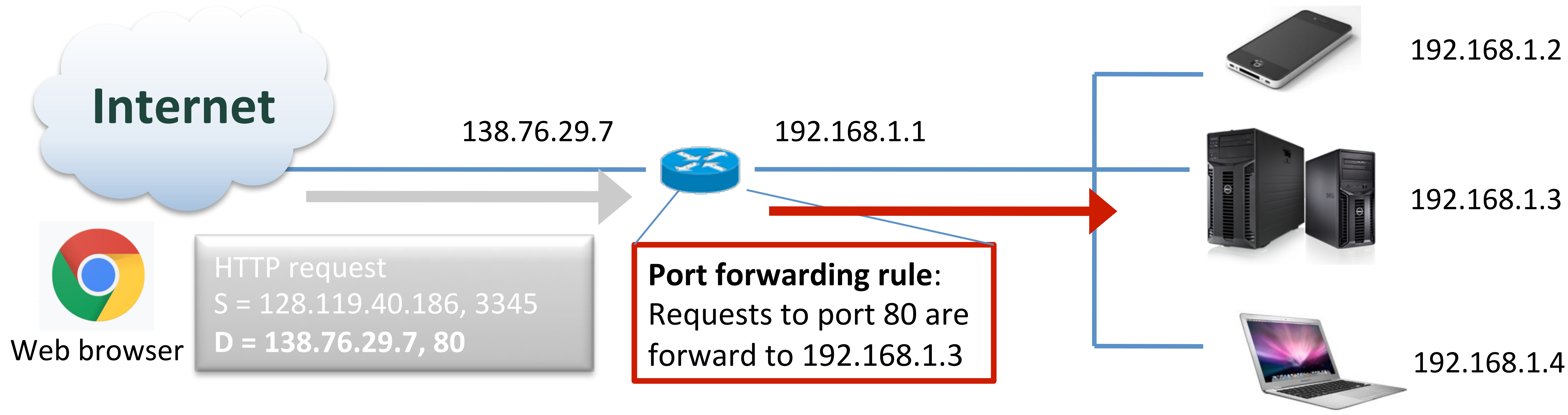
Request from the private network



NAT translation table	
WAN side	**LAN side**
138.76.29.7, **5001**	192.168.1.2, **3345**
…	…

- When the response arrives at the NAT router, the router searches the NAT translation table and **rewrites the datagram's destination address and destination port number**, and forwards the datagram into the home network.

Request from the external network

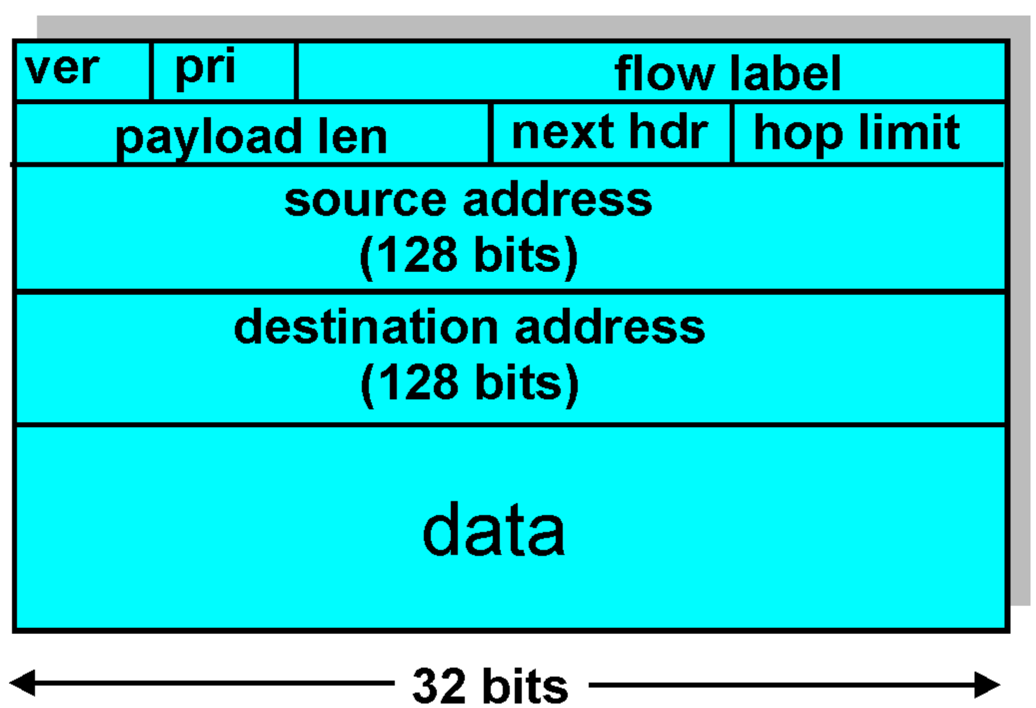


- **Problem:** NAT allows communication where a host on the private network initiates the connection; what if a server is operated in the private network?
- **Port forwarding** (or port mapping) - configuring the NAT router to send all packets received on a particular port to a specific host on the private network
  - e.g., if external hosts need to access a web server (port 80) operating on machine 192.168.1.3, it will be necessary to define a port forwarding rule on the router, redirecting all TCP packets received on port 80 to machine 192.168.1.3

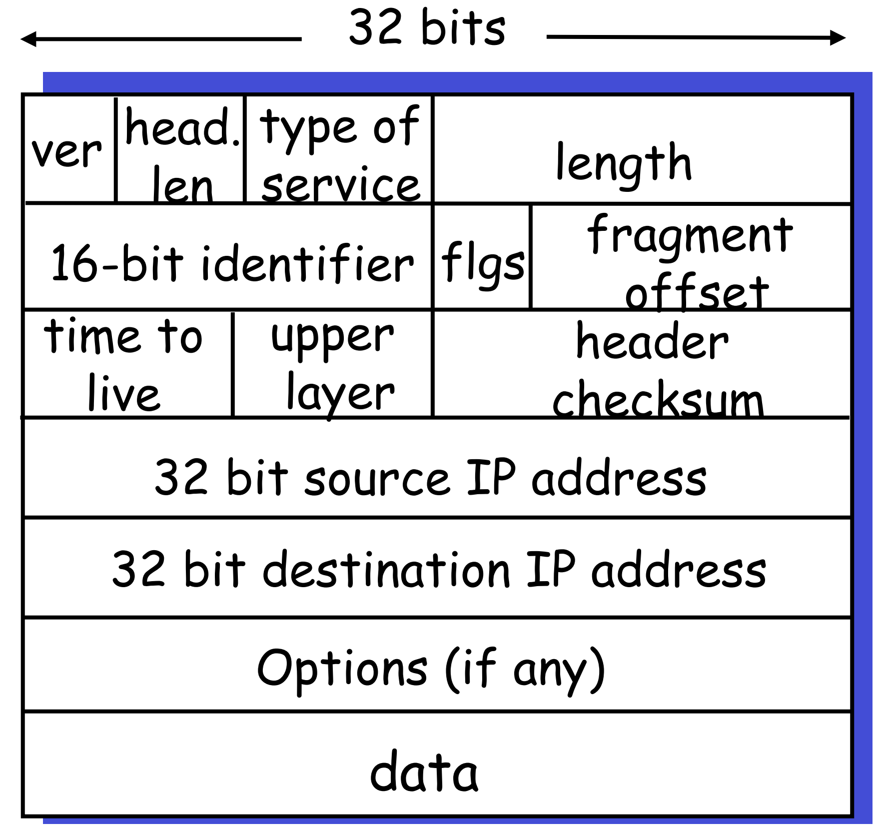
IPv6

- Initial motivation
  - 32-bit address space completely allocated by the Internet Assigned Number Authority (IANA) on Feb. 3, 2011
- Additional motivation:
  - to improve existing IPv4

IPv6 datagram format



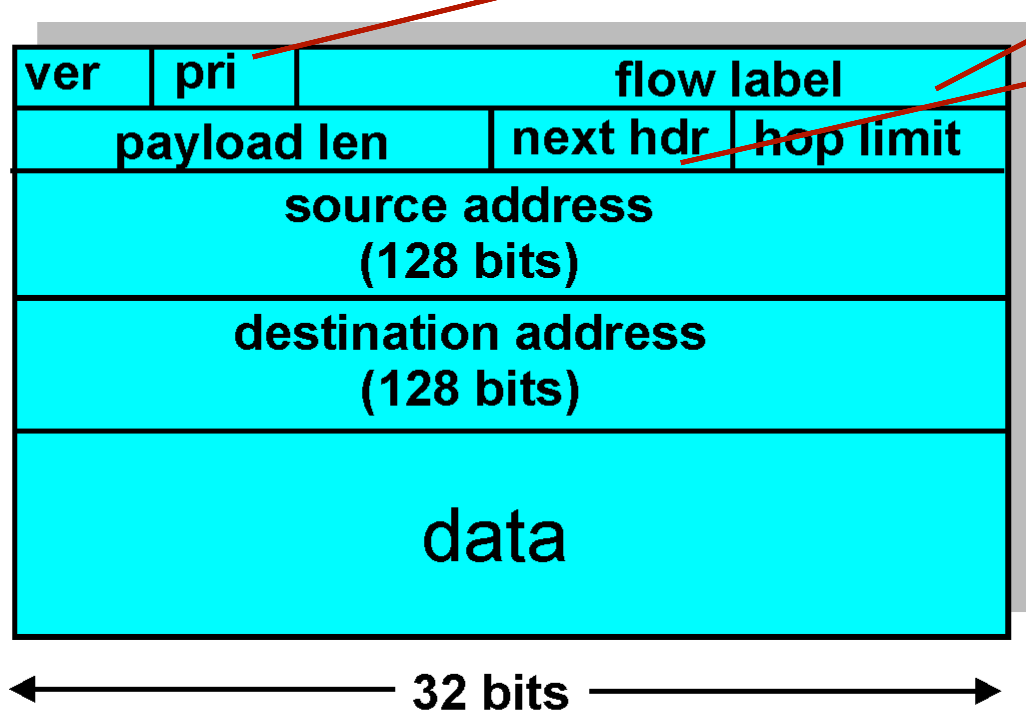
IPv4 datagram format



IPv6 (cont'd)

- Expanded address space: 128 bits 2001:0000:3238:DFE1:0063:0000:0000:FEFB
- Fixed-length 40-byte header
  - allows fast processing of IPV6 datagram
  - no options field (but can be pointed to from "next header" field)
- Flow labeling and priority
  - service differentiation among datagrams belonging to different flows
- no fragmentation/reassembly allowed at the routers
  - can only be done at source/destination hosts
  - to speed up IP forwarding
- no header checksum

IPv6 datagram format



identify priority among datagrams in flow

identify datagrams in same flow

ususally specify upper-layer protocol used by payload; or indicate the type of extension header (if present) immediately following the IPv6 header

checksum: removed entirely to reduce processing time at each hop

Illustrations of Next Header:

```
+---------------+------------------------
|  IPv6 header  | TCP header + data
|               |
| Next Header = |
|      TCP      |
+---------------+------------------------
```

```
+---------------+----------------+------------------------
|  IPv6 header  | Routing header | TCP header + data
|               |                |
| Next Header = |  Next Header = |
|    Routing    |      TCP       |
+---------------+----------------+------------------------
```

```
+---------------+----------------+-----------------+-----------------
|  IPv6 header  | Routing header | Fragment header | fragment of TCP
|               |                |                 |  header + data
| Next Header = |  Next Header = |  Next Header =  |
|    Routing    |    Fragment    |       TCP       |
+---------------+----------------+-----------------+-----------------
```

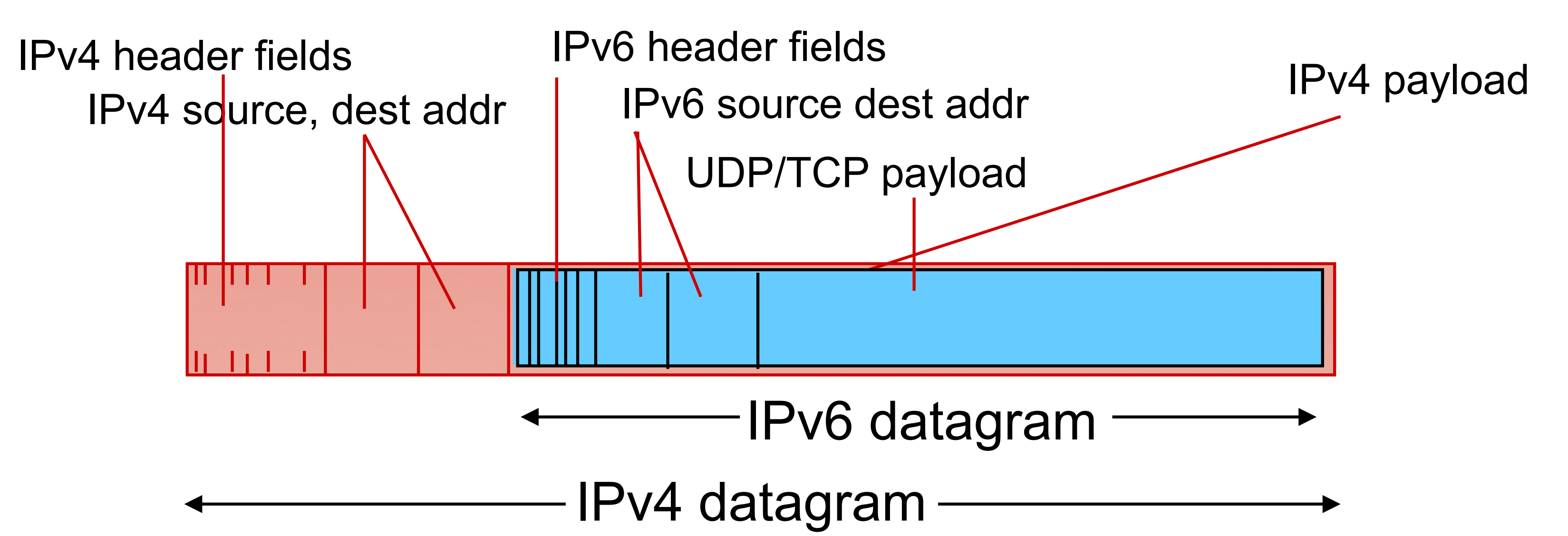
https://datatracker.ietf.org/doc/html/rfc2460

IPv4 —> IPv6

- Not all routers can be upgraded simultaneous
- How will the network operate with mixed IPv4 and IPv6 routers?
  - Dual-stack approach: IPv6 routers also have a complete IPv4 implementation, with both an IPv4 address and an IPv6 address

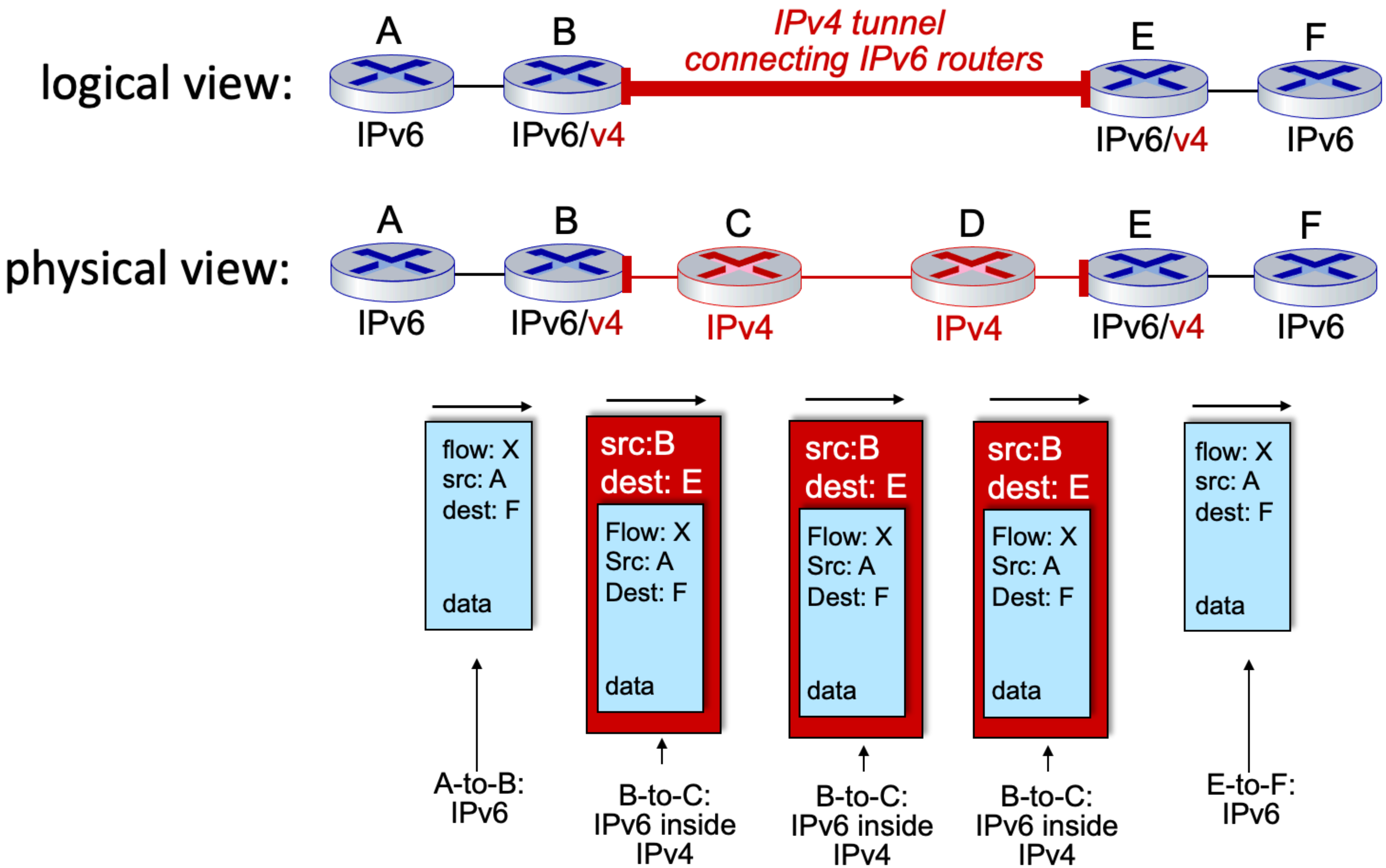
    use IPv4 to communicate with IPv4 routers

    use IPv6 to communicate with IPv6 routers
  - Tunneling: IPv6 carried as payload in IPv4 datagram among IPv4 routers



IPv4 —> IPv6 (cont'd)

Tunneling



IPv6 adoption

- DNS has supported IPv6 since 2008
- All major OSs in use included IPv6 implementation by 2011
- More than 40% of Google users access services via IPv6 (2023)

- Still need a long time for wide deployment, thinking of application-level changes needed (WWW, streaming media, social app, …)

ICMP

- Internet Control Message Protocol
- Used by hosts & routers to communicate network-layer information among each other
  - error reporting

    unreachable host/network/port/protocol
  - echo request, reply

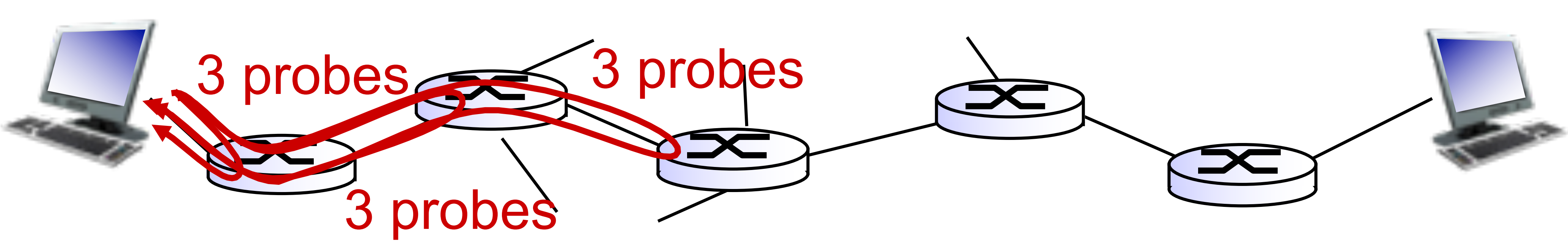
    e.g., used in ping, traceroute
- ICMP message contains
  - type
  - code
  - header and first 8 bytes of IP datagram that caused the ICMP message to be generated
- "Above" IP:
  - ICMP messages carried inside IP datagrams
  - in the IP datagram carrying an ICMP message, the "upper layer protocol" header specifies ICMP

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 - not used)
8	0	echo request (ping)
9	0	router advertisement
10	0	router discovery
11	0	TTL exceeded
12	0	bad IP header

Traceroute using ICMP

- Source host sends series of UDP segments to destination with unlikely UDP port numbers
  - first set has TTL =1
  - second set has TTL=2
  - etc.
- When an IP datagram in nth set arrives at nth router:
  - router discards datagram and sends back to source host an ICMP message (type 11, code 0, including name and IP address of router)
- When ICMP message arrives, source records RTTs and name/IP address of the nth router

stopping:
- UDP segment eventually arrives at destination host;
- destination returns ICMP “port unreachable” message (type 3, code 3);
- source stops after receiving this ICMP message



- Required reading
  - *Computer Networking: A Top-Down Approach* (8th Edition)
    Ch 4.3.2, 4.3.3, 4.3.4, 5.6

- Acknowledgement:
  - Some materials are extracted from the slides created by Prof. Jim F. Kurose and Prof. Keith W. Ross for the textbook.