Network Layer: Protocols (Part 2)

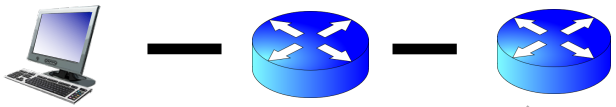
Lecture 21

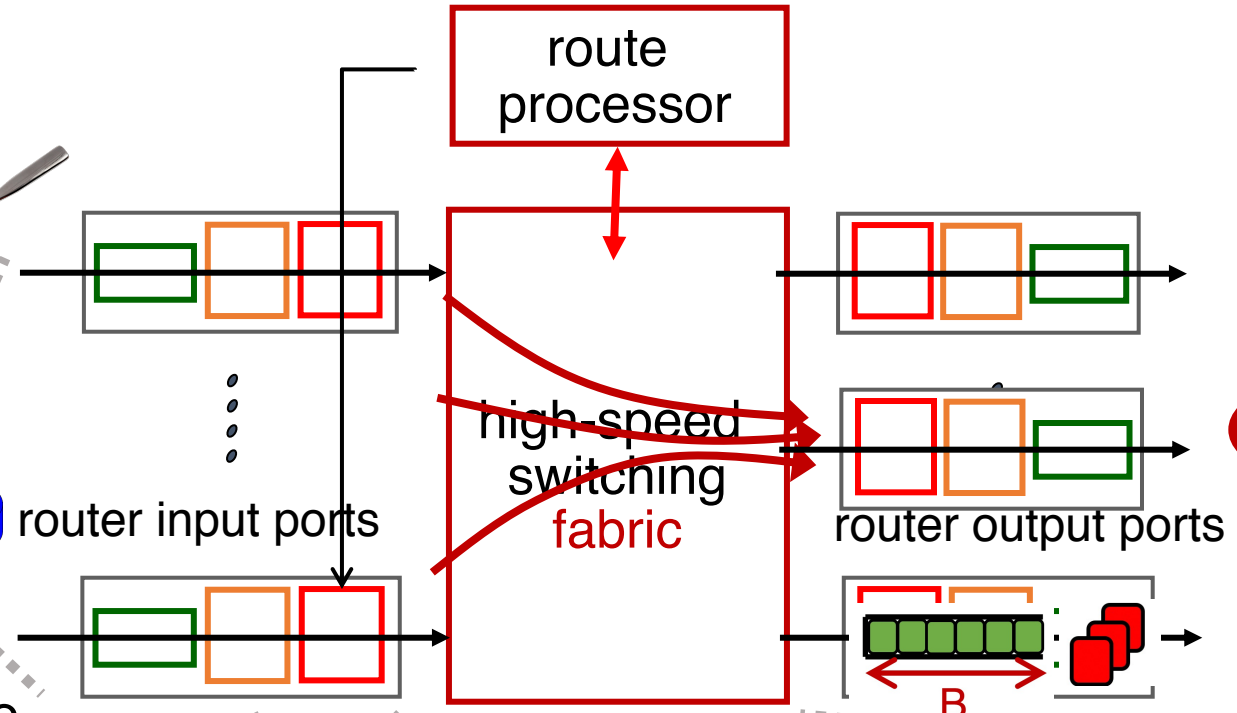
http://www.cs.rutgers.edu/~sn624/352-S22

Srinivas Narayana

RUTGERS
UNIVERSITY | NEW BRUNSWICK

Quick recap of concepts





In nonblocking fabrics, queues form only due to contention for output port.

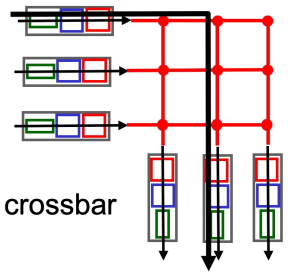
Output port contention is fundamental; unavoidable.

Scheduling and buffer management crucial.

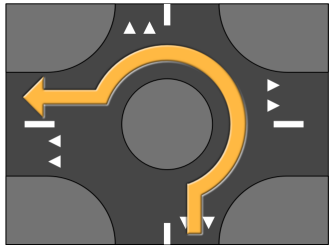
Forwarding Table

Longest prefix matching on dst IP address

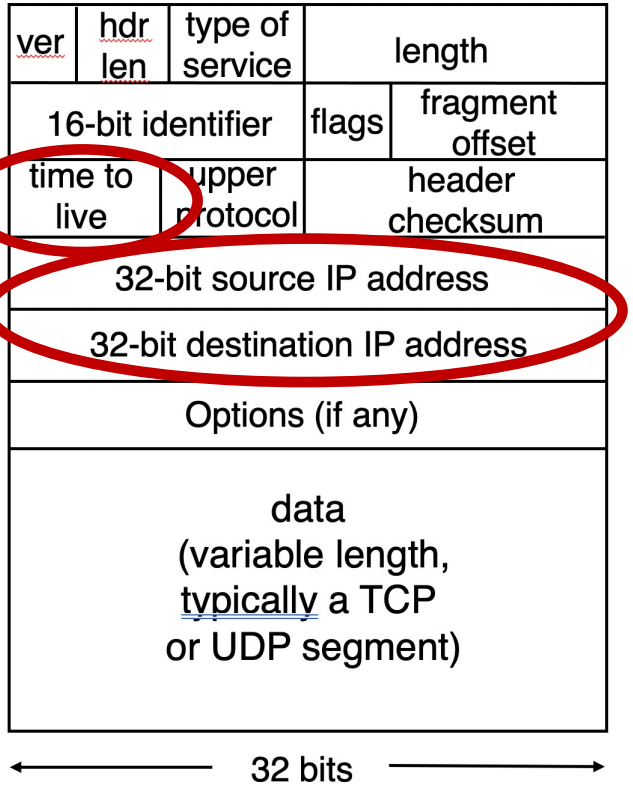
Nonblocking



Blocking



IPv4 datagram



Mapping routes

- `ip route` (Linux)
  - Try `ip route get ____`
- `route` (Mac OS X)

The network layer is all about reachability. Every protocol we'll see solves a sub-problem.

How does an endpoint get an address?

DHCP

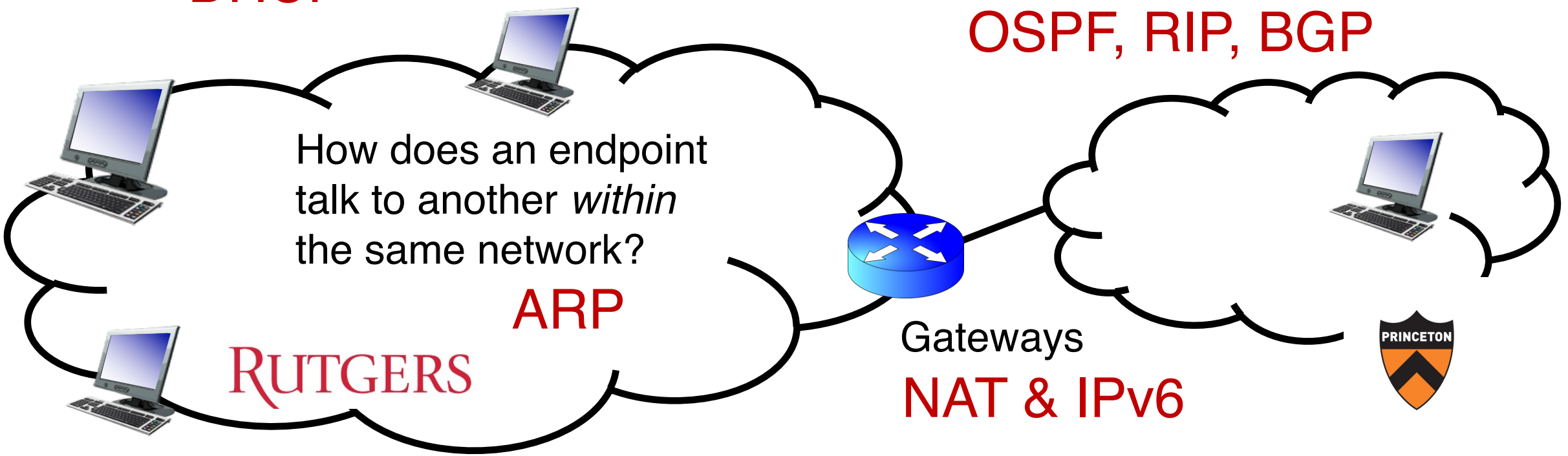
Debugging?

ICMP

How does an endpoint talk to another *outside* its network?

Routing protocols

OSPF, RIP, BGP



Dynamic Host Configuration Protocol (DHCP)

How does an endpoint get its IP addr?

- One possibility: hard-code the IP address on the endpoint
  - e.g., a system admin writing addresses in a file
  - Linux: /etc/network/interfaces
  - Mac OS X (10.14.6): system preferences > Network > name of interface > advanced > TCP/IP > “Manually”

- Another possibility: dynamically receive an address “from the network”
  - DHCP: Dynamic Host Configuration Protocol
  - Provide plug-and-play functionality for endpoints (e.g., phones, laptops)

Many similar bootstrapping problems

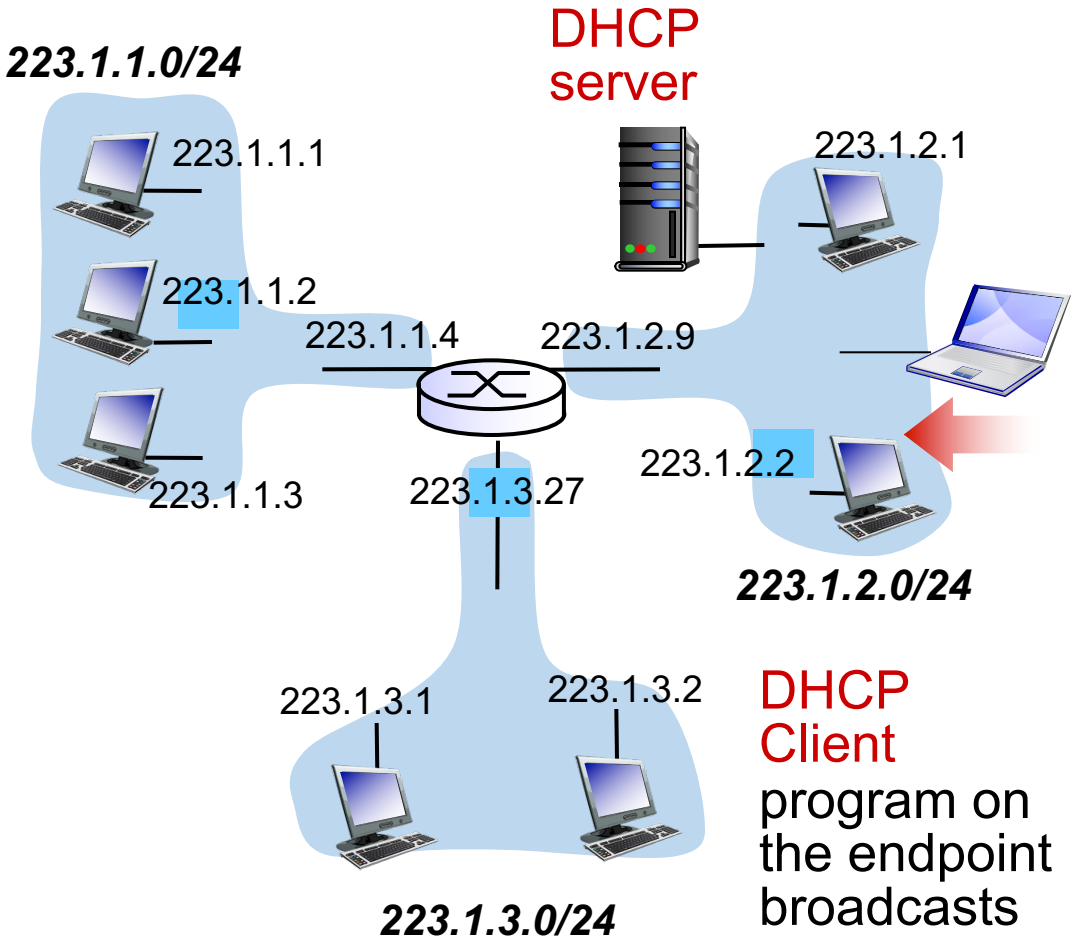
- How does a host get its IP address?
- How does a host know its local DNS server?
- How does a host know its netmask?
  - i.e., so that it can know which other hosts are in the same network
  - Note: the details how A and B talk to each other changes significantly when A and B are in the same network vs. different network
- How does a host know how to reach other networks?
  - i.e., which router is at the “border” of the current network?
  - This router is also called the gateway router: crucial for an endpoint to communicate with another endpoint external to the network

How DHCP works

- An endpoint that just joined a network knows nothing about it
  - Endpoint doesn't even have an IP address for its point of attachment
- We solved a similar bootstrapping problem before:
  - Domain Name Service (DNS) to retrieve addresses
- Often, it makes little sense to have the endpoint contact a "known" server to receive an IP address
  - E.g., connecting to a brand-new network you've never been in
- The only idea that really works is to ask everyone
  - Broadcast a "query"

How DHCP works

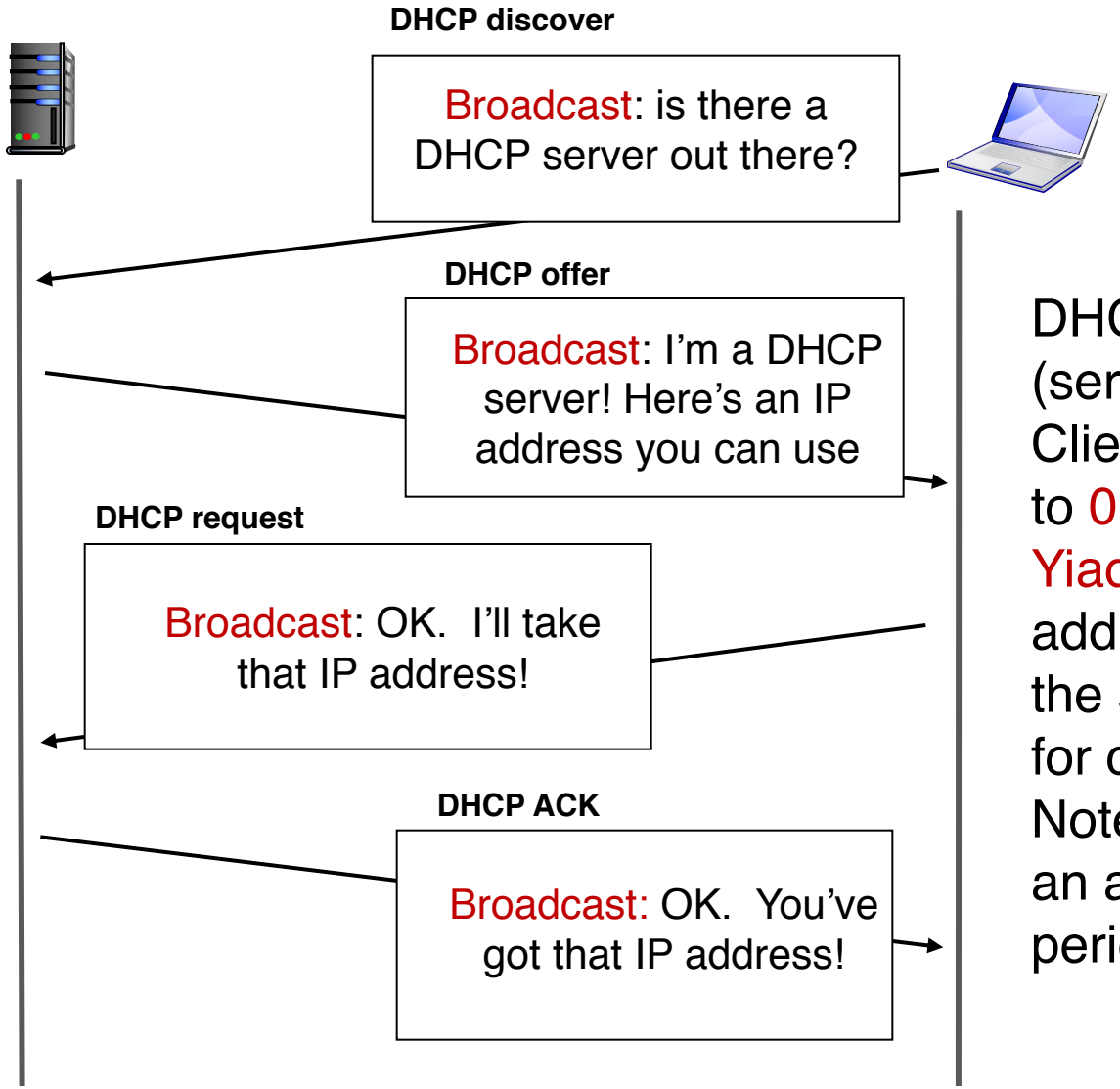
- DHCP allows a host to dynamically obtain its IP address from a server on a network when it joins the network
- DHCP can allow a host to be mobile across different networks, obtaining IP addresses as needed
- DHCP uses leases on addresses
  - Host must renew lease periodically
  - Allows network to reuse an IP with an expired lease, reclaiming addresses from inactive hosts



DHCP Client program on the endpoint broadcasts message requesting an IP address.

DHCP client-server scenario

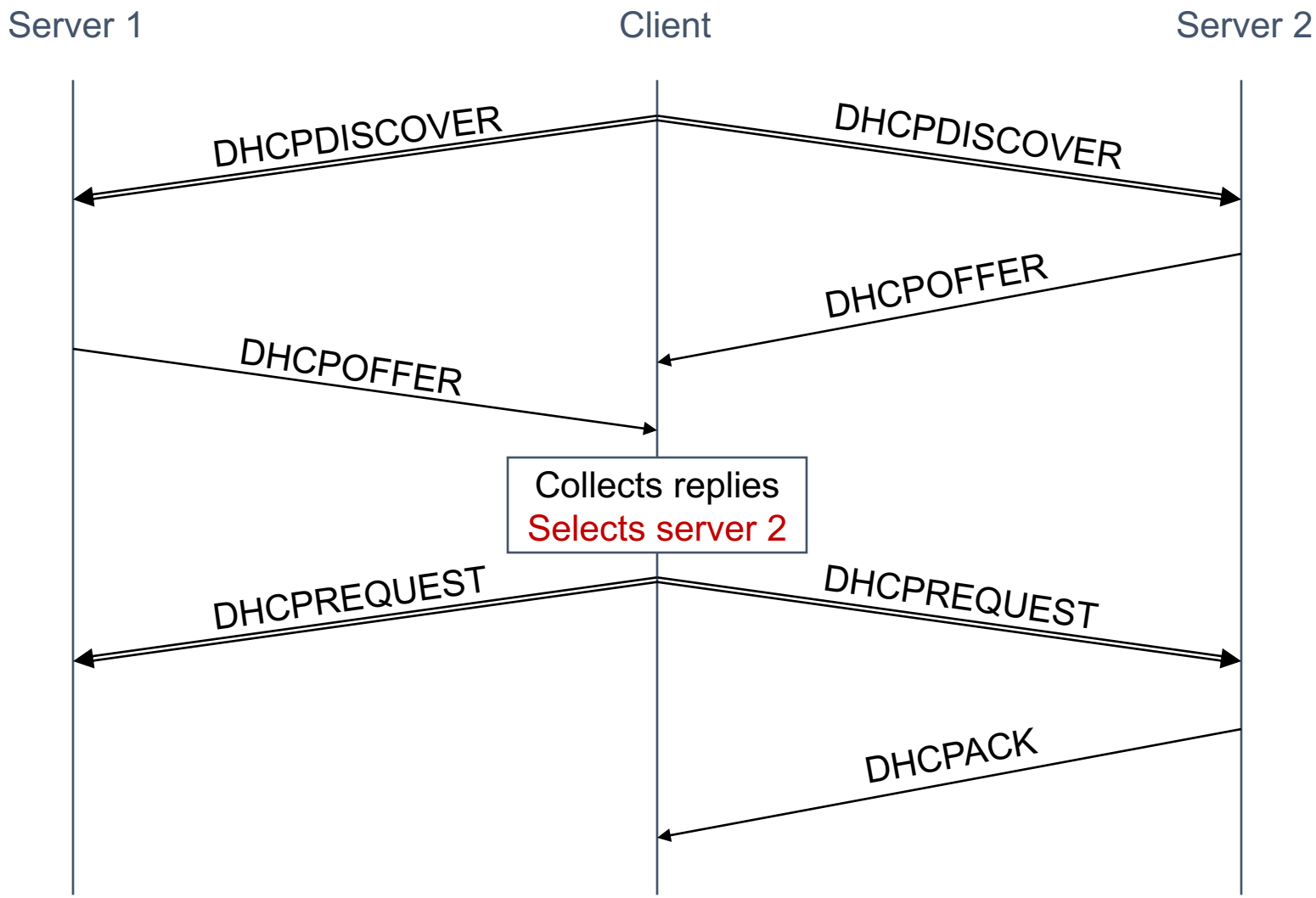
DHCP server:
223.1.2.5



223.1.2.4
Arriving client

DHCP runs on UDP ports 67 (server) and 68 (client)
Client's initial IP address is set to 0.0.0.0
Yiaddr stands for "your IP address" – an address value the server sends to the client for consideration
Note that the IP allocation has an associated lifetime (lease period)

Multiple DHCP servers can coexist



DHCP returns more than an IP address

- Name and IP address of the local DNS server
- Netmask of the IP network the host is on
  - Useful to know whether another endpoint is inside or outside the current IP network
- Address of the gateway router to enable the endpoint to reach other IP networks

Your home router runs DHCP

- Likely, your home devices (laptops, tablets, phones) are all using DHCP-assigned IP addresses
- The DHCP server is running on the control processor of your home's access router (e.g., WiFi router)
- You can access the DHCP client program on Linux using the command `dhclient`

Summary of DHCP

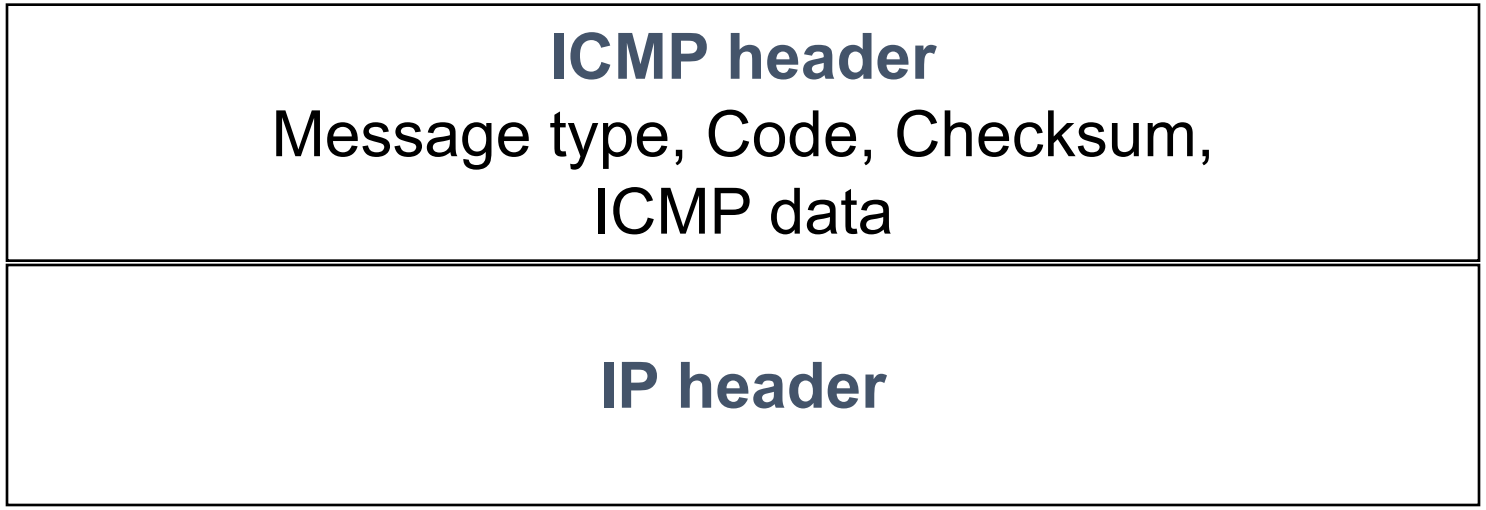
- Want endpoints to have plug and play functionality
  - Avoid tedious manual configuration of IP addresses and other information
- DHCP: a general bootstrapping mechanism for critical information required for network layer functionality
- Hosts can be simple: receive information from DHCP servers by broadcasting over the network

Internet Control Message Protocol (ICMP)

Internet Control Message Protocol

- A protocol for troubleshooting and diagnostics
- Works over IP: unreliable delivery of packets
- Some functions of ICMP:
  - Determine reachability and network errors
  - Specify that packets have been in the network for too long

ICMP message



https://en.wikipedia.org/wiki/Internet_Control_Message_Protocol#Control_messages

Specific uses of ICMP

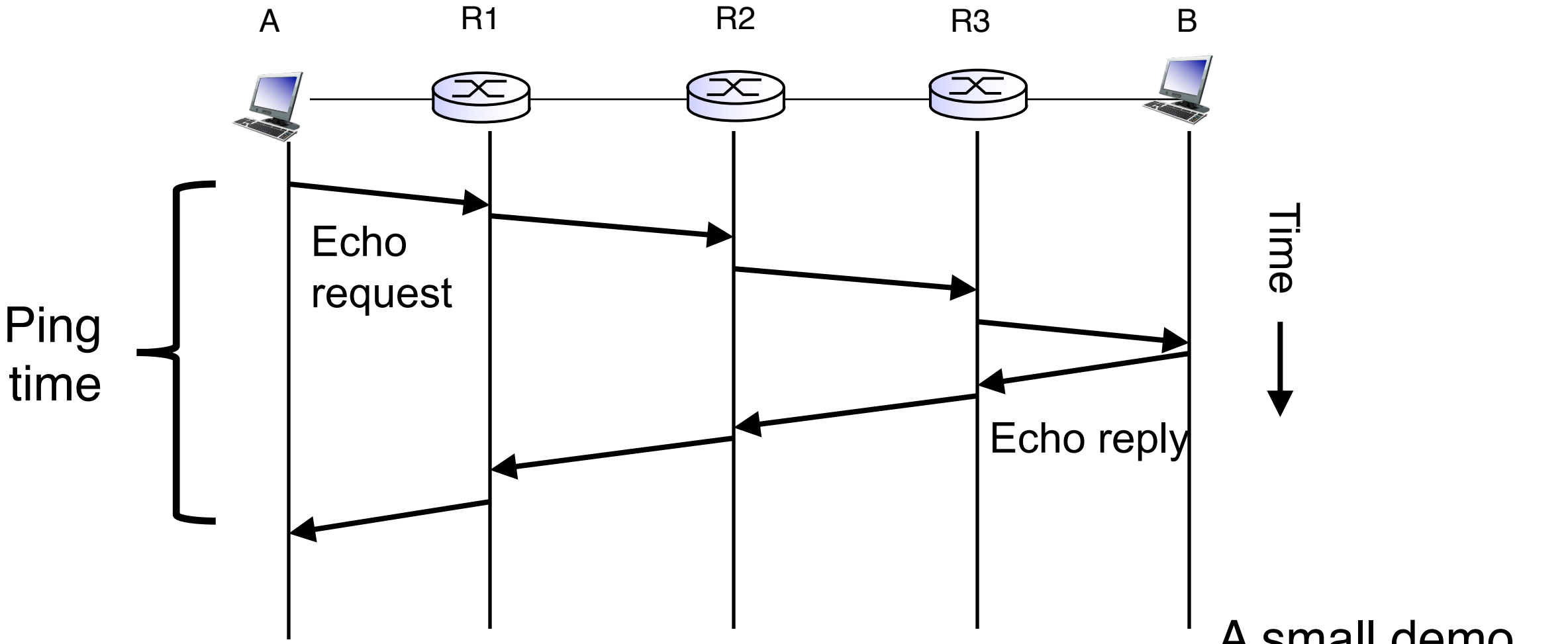
- Echo request reply
  - Check remotely if an endpoint is alive and connected
  - *Without* running an app remotely or controlling that endpoint
- An unreachable destination
  - Invalid address and/or port
- Knowing if packet's IP time-to-live expired
  - Example, due to routing loops

- Look at two tools built using ICMP: ping and traceroute

Ping

- Uses ICMP echo request (type=8, code=0) and reply (type=0, code=0)
- Source sends ICMP echo request message to dst address
- Destination network stack replies with an ICMP echo reply message
- Source can calculate round trip time (RTT) of packets
- If no echo reply comes back, then the destination is unreachable
- Don’t need to have a server program running on the other side
  - In general, the remote endpoint can be completely outside your control

Ping



A small demo…

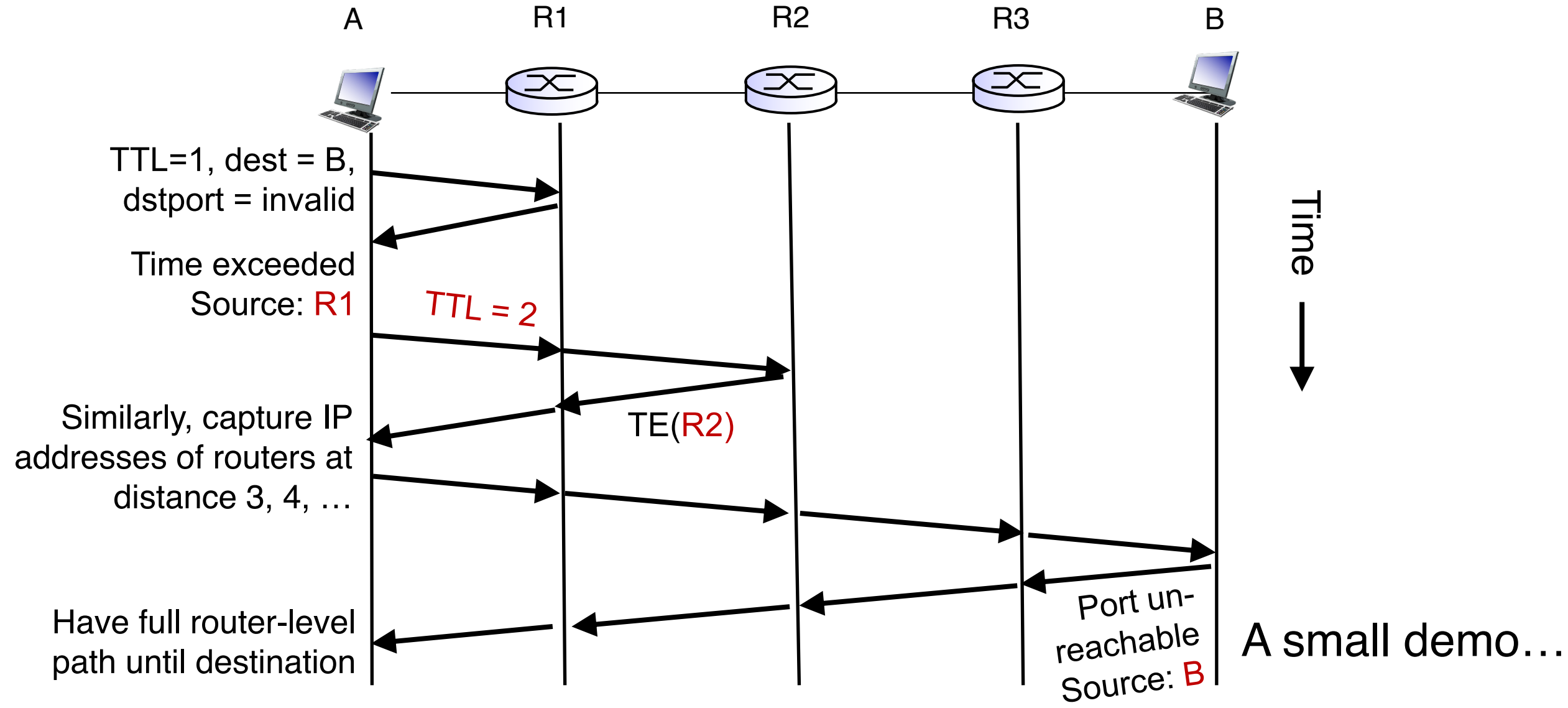
Traceroute

- A tool that can record the router-level path taken by packets
- A clever use of the IP time-to-live (TTL) field
- In general, when a router receives an IP packet, it decrements the TTL field on the packet
  - A failsafe mechanism to ensure packets don't keep taking up network resources for too long
- If a router receives a packet with TTL=0, it sends an ICMP time exceeded message (type=11, code=0) to the source endpoint

Traceroute

- Traceroute sends multiple packets to a destination endpoint
- But it progressively increases the TTL on those packets: 1, 2, ...
- Every time a time exceeded message is received, record the router's IP address
- Process repeated until the destination endpoint is reached
- If the packet reaches the destination endpoint (i.e.: TTL is high enough), then the endpoint sends a port unreachable message

Traceroute



A small demo…

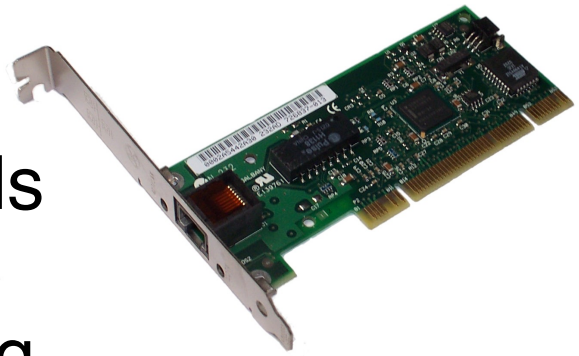
Summary of ICMP

- A protocol for network diagnostics and troubleshooting
- Two useful tools: ping and traceroute
- Ping: test connectivity to a machine totally outside your control
  - Use ICMP echo request and reply
- Traceroute: determine router-level path to a remote endpoint
  - A smart use of the TTL field in the IP header

Address Resolution Protocol

Background: Let's peek into the link layer

- Each network adapter has a hardware address or a MAC address
  - E.g., the Wi-Fi adapter on your laptop has one
- Assigned by the manufacturer, not expected to vary over time
  - Think about it as an identifier for the device
- To communicate over a single link, a sender needs the destination hardware address
- Directory mechanisms like DNS and bootstrapping mechanisms like DHCP provide IP addresses
- Given an IP address, how does an endpoint find the hardware address?



Address Resolution Protocol (ARP)

- ARP solves the following problem. Given an IP, find the machine's hardware address
  - IP → MAC resolution
- All endpoints that are looked up are expected to be within the same network
- Hence, address resolution can use broadcast:
  - We don't need to develop directory mechanisms like DNS
  - Send (ARP) queries to everyone, asking for a MAC given an IP

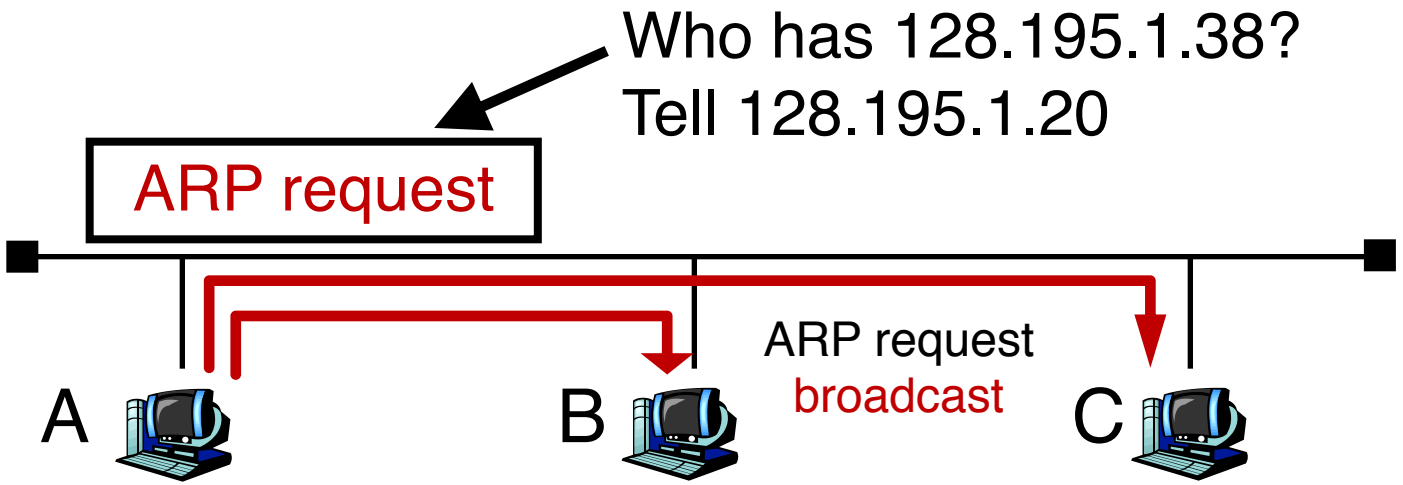
ARP packet format

- Hardware type: link-layer protocol
  - Example: Ethernet (1)
- Hardware address length:
  - Example: Ethernet = 6 bytes
- Protocol Type: network-layer protocol
  - Example: IPv4 (0x0800)
- Protocol address length
  - Example: IPv4 = 4 bytes
- Operation:
  - ARP request: 1, reply: 2
- Sender’s addresses
- Address to be resolved (or response)

Internet Protocol (IPv4) over Ethernet ARP packet

Octet offset	0	1
0	Hardware type (HTYPE)	
2	Protocol type (PTYPE)	
4	Hardware address length (HLEN)	Protocol address length (PLEN)
6	Operation (OPER)	
8	Sender hardware address (SHA) (first 2 bytes)	
10	(next 2 bytes)	
12	(last 2 bytes)	
14	Sender protocol address (SPA) (first 2 bytes)	
16	(last 2 bytes)	
18	Target hardware address (THA) (first 2 bytes)	
20	(next 2 bytes)	
22	(last 2 bytes)	
24	Target protocol address (TPA) (first 2 bytes)	
26	(last 2 bytes)	

ARP operation



Ethernet Address:
05:23:f4:3d:e1:04
IP Address:
128.195.1.20

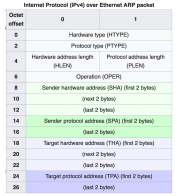
Wants to transmit
to 128.195.1.38

Ethernet Address:
12:04:2c:6e:11:9c
IP Address:
128.195.1.122

Different target
IP address:
ignore ARP

Ethernet Address:
98:22:ee:f1:90:1a
IP Address:
128.195.1.38

Matching target
IP: send reply



Hardware type: Ethernet
Protocol type: IPv4
Hardware addr length: 6
Protocol addr length: 4
Operation: 2 (reply)
Sender hardware addr:
05:23:f4:3d:e1:04
Sender protocol addr:
128.195.1.20
Target HW addr:
98:22:ee:f1:90:1a
Target protocol addr:
128.195.1.38

Communicating outside the local net?

- Suppose endpoint A wants to communicate with endpoint B that is in a different network
- ARP broadcast outside the local network is too expensive
  - How does one limit the scope of the broadcast? Internet-wide?
- Besides, the hardware address format used by B's network might be different from that of A's network!
- ARPs are not meaningful across network boundaries
- Communicating to a network-external endpoint just means sending the packet to the gateway router
  - Host can know that a destination is external using IP addr and netmask
  - Host can talk to the gateway using DHCP (to get IP) and ARP (to get MAC)

Summary of ARP

- A useful mechanism to allow hosts inside a network to communicate:
- ARP protocol helps resolve IP addresses into MAC addresses using a broadcast mechanism
- Communication outside the local network requires ARP-ing for and sending packets to the gateway