ECE 356/COMPSI 356
Computer Network Architecture

Basic Internetworking (IP)

Monday September 23, 2019

Recap

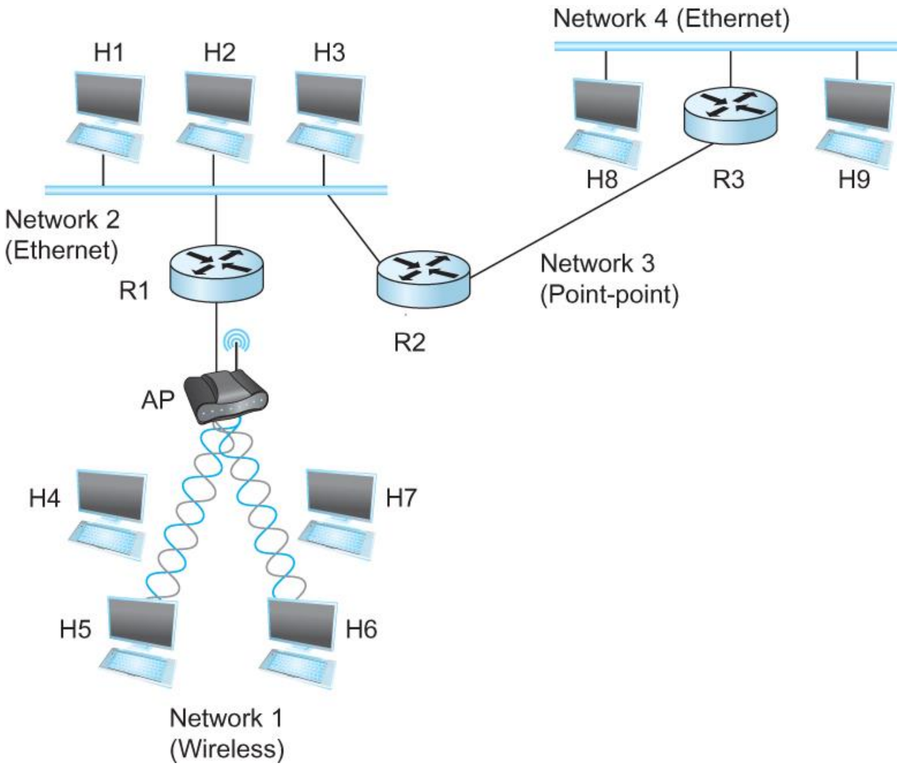
- Last lecture:
  - ➢Approaches to switching
  - ➢Ethernet switches
- Readings for this class: **PD 3.2**

Lecture Outline

- Internet protocol (IP)
- IP header format
- IP addressing
- IP forwarding
  - Forwarding algorithm

Inter-networking

- One level of indirection
  - ➢ Routers interface different networks
- Uniform addressing (IP)
- **Routers** send packets to their destination IP addresses

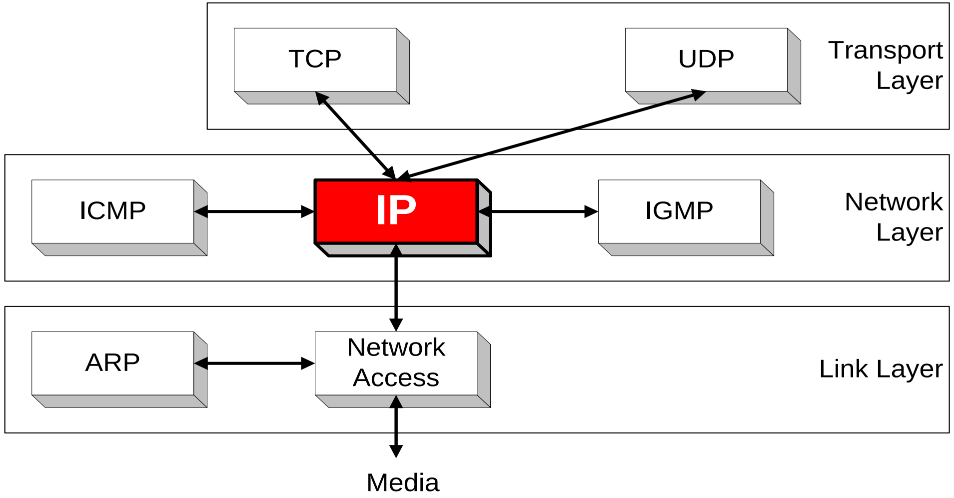


Difference Between a Switch and a Router

- Video posted on Piazza

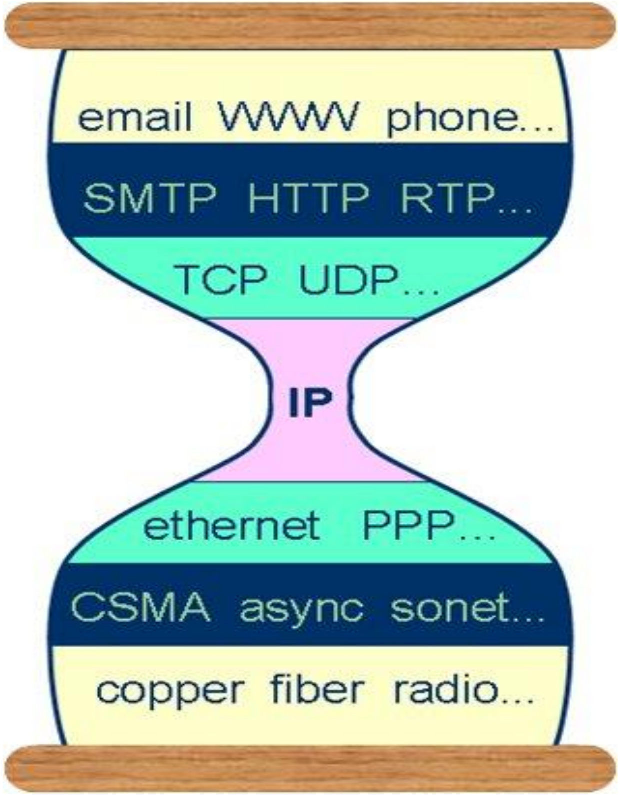
Internet Protocol

- IP (Internet Protocol) is a Network Layer Protocol
- IP's current version is Version 4 (IPv4)
  - RFC 791
- IPv6 is also deployed

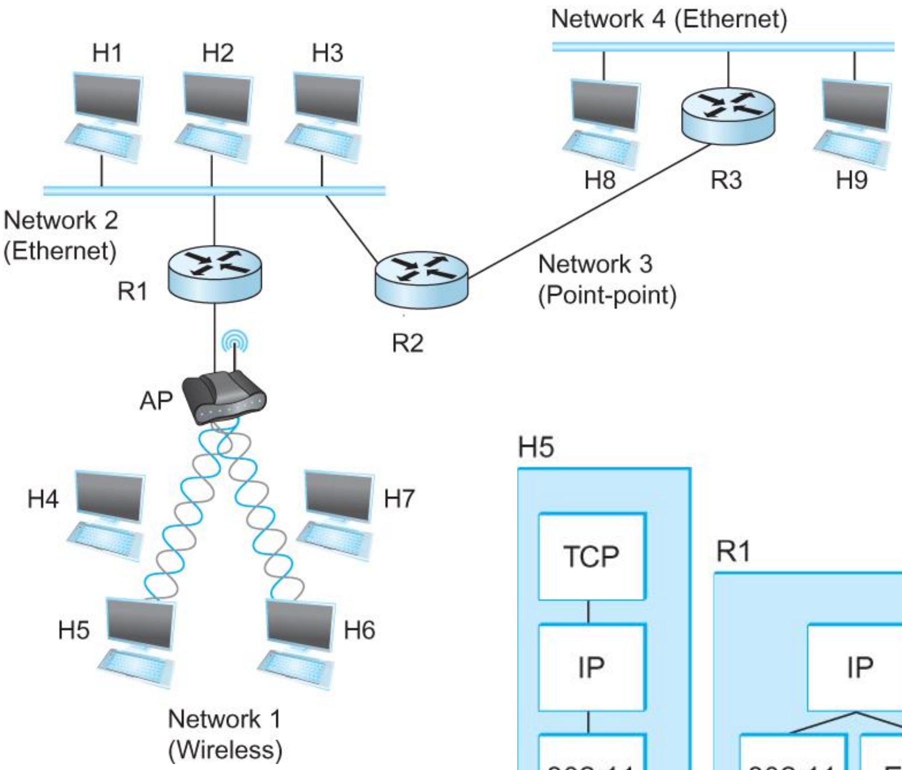


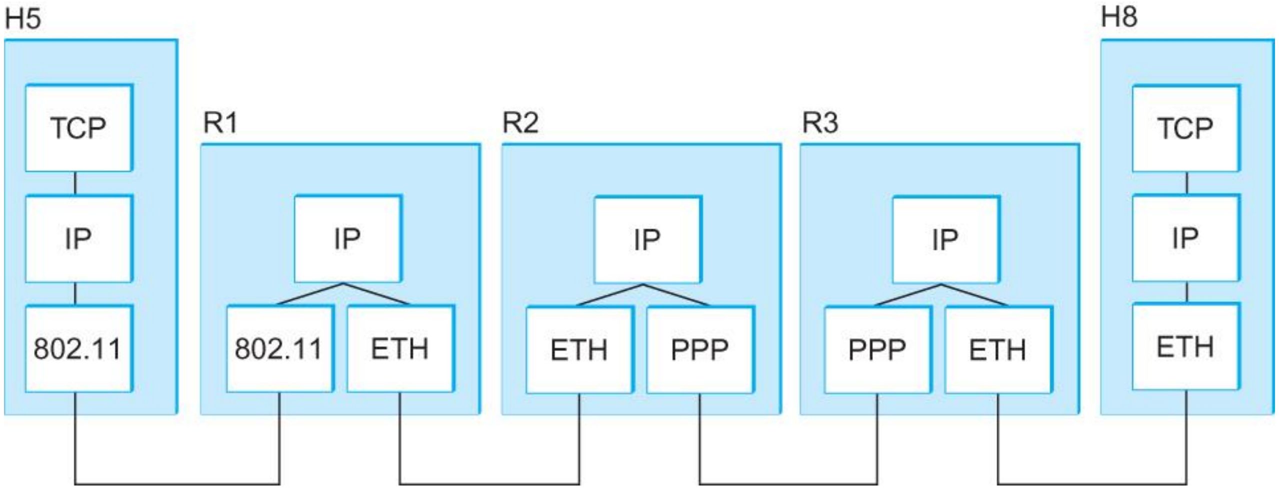
IP: The Thin Waist of the Hourglass

- IP is the *waist of the hourglass* of the Internet protocol architecture
- Multiple higher-layer protocols
- Multiple lower-layer protocols
- Only one protocol at the network layer
- Architecture avoids the $N*M$ problem



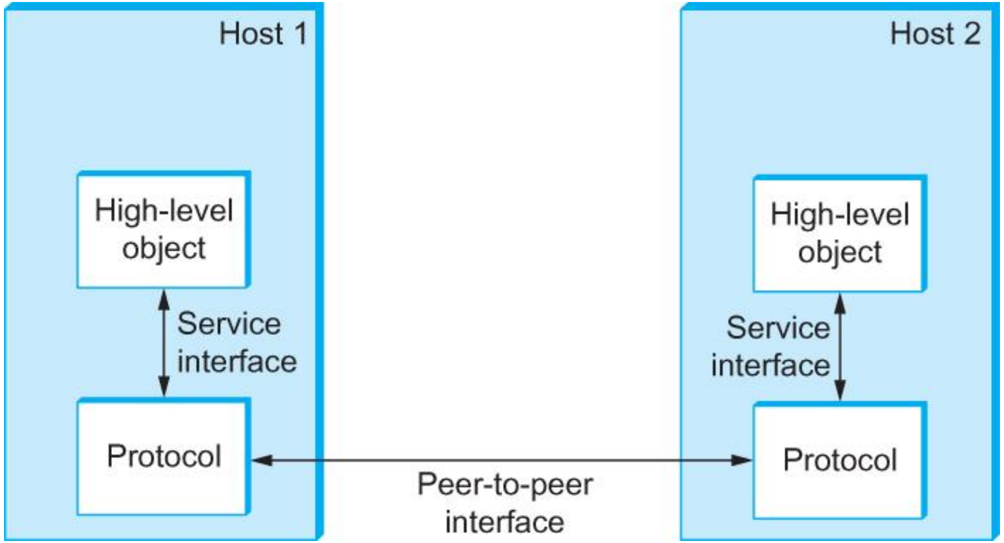
A Simple Internetwork





Recap: Protocol Peer and Service Interfaces

- Each protocol defines two different interfaces
  - Service interface
  - Peer interface



IP Service Model

- Delivery service of IP is minimal
- IP provides an unreliable connectionless best effort service (*datagram service*)
  - Unreliable, connectionless
  - Best effort
- Consequences:
  - Loss, out of order, and duplicates must be handled at the upper layer

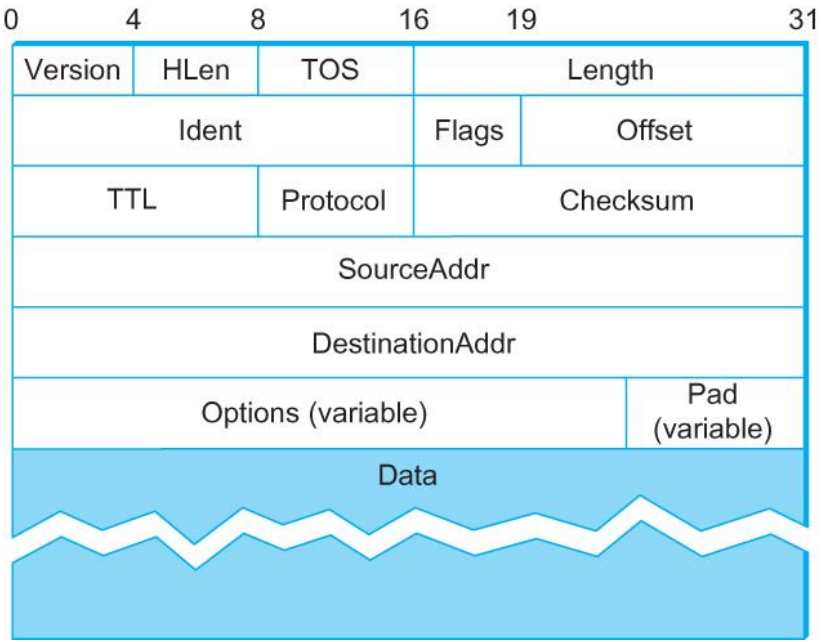
Basic IP Router Functions

- Things you need to understand to do lab2
  - Internet protocol
    - IP header
    - IP addressing
    - IP forwarding
  - Address Resolution Protocol (ARP) – next lecture
  - Error reporting and control – next lecture
    - Internet Control Message Protocol (ICMP)

Lecture Outline

- Internet protocol (IP)
- **IP header format**
- IP addressing
- IP forwarding
  - Forwarding algorithm

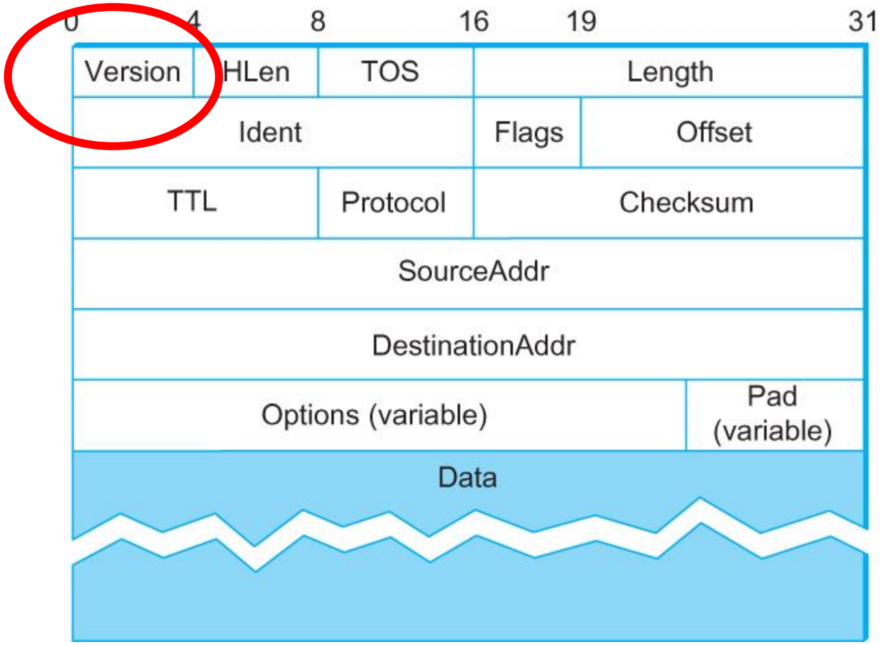
IP Header Format



- 20 bytes fixed length header + variable length options

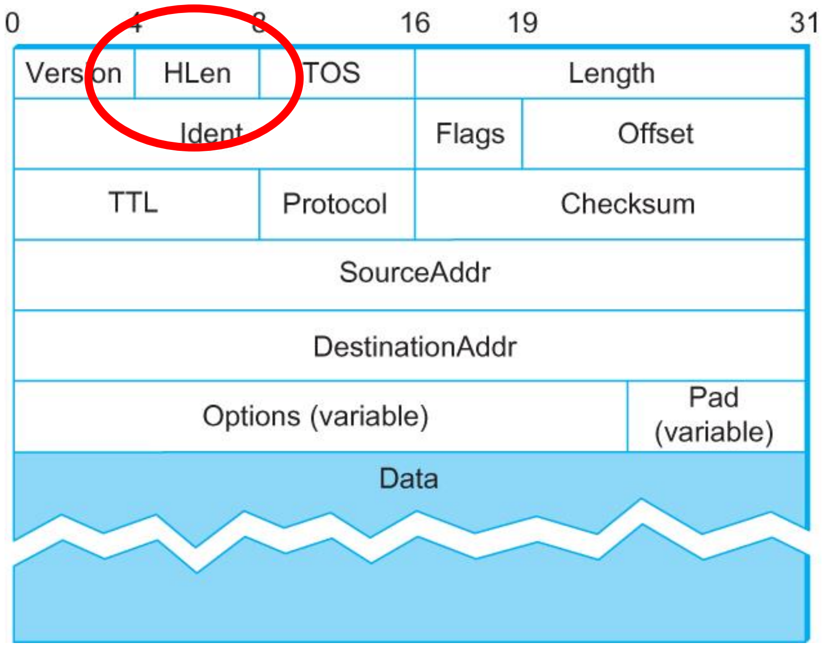
IP Header Format: Version

- Version: v4



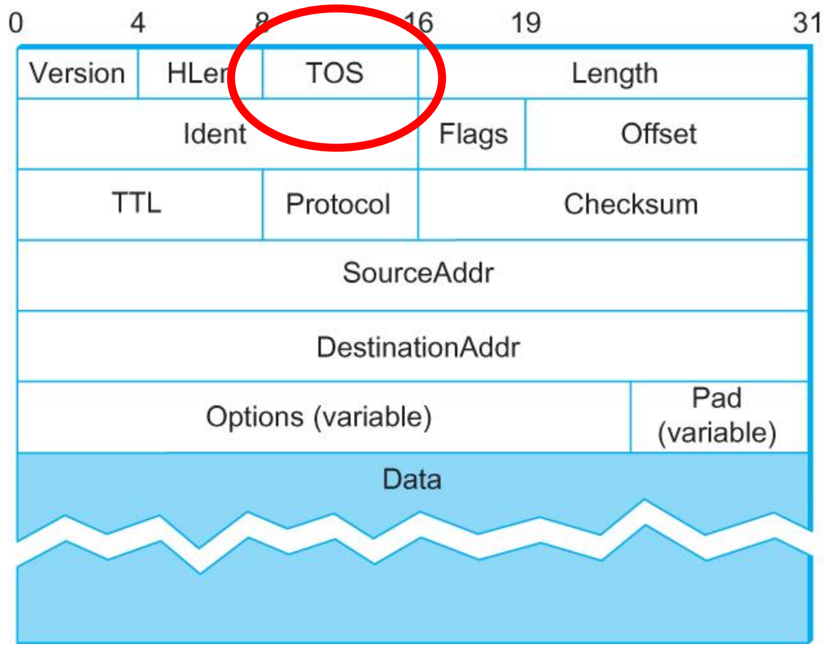
IP Header Format: Header Length

- Internet Header Length (IHL 4 bits): the length of header in 32-bit words
  - ➢ 20 bytes if no options



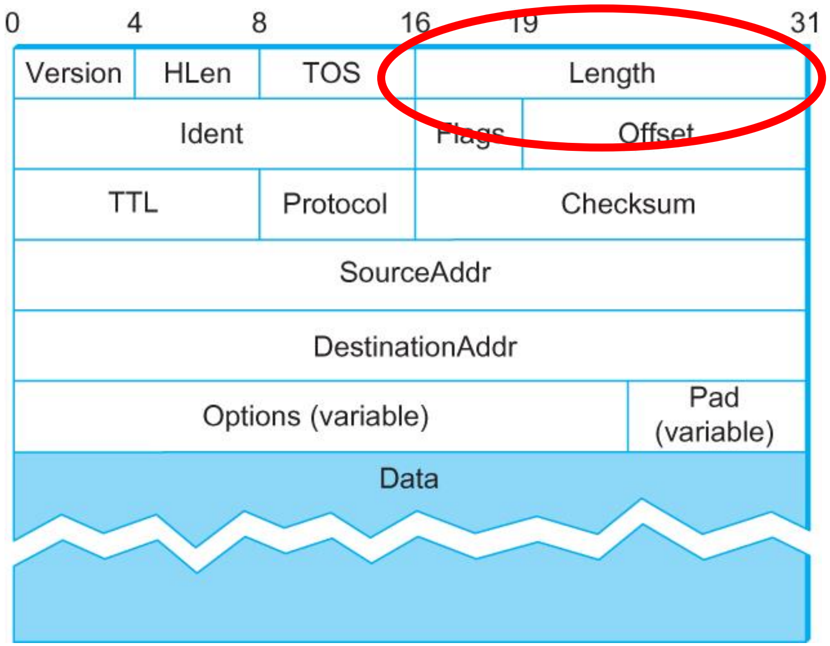
IP Header Format: Type of Service

- Treat packets differently according to application needs
- Real-time, VoIP
- Will discuss later on



IP Header Format: Length

- 16 bits
- Packet length in bytes, including the header
  - Maximum size: 65,535 bytes
  - May not be supported by lower-layer protocols → *fragmentation and reassembly*

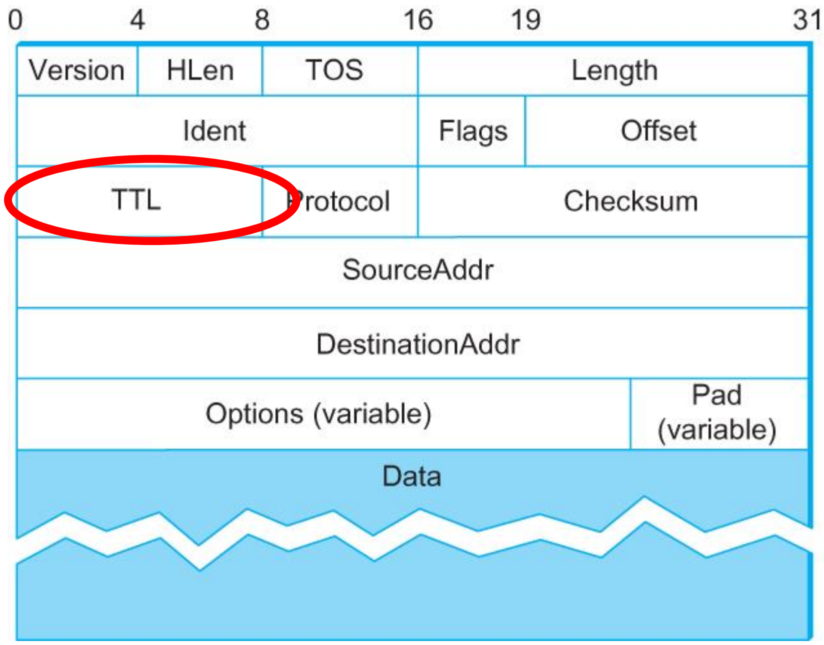


IP Header Format: Fragmentation and Reassembly Fields

- Identification, Flags, Fragment offset
  - Fragmentation and reassembly

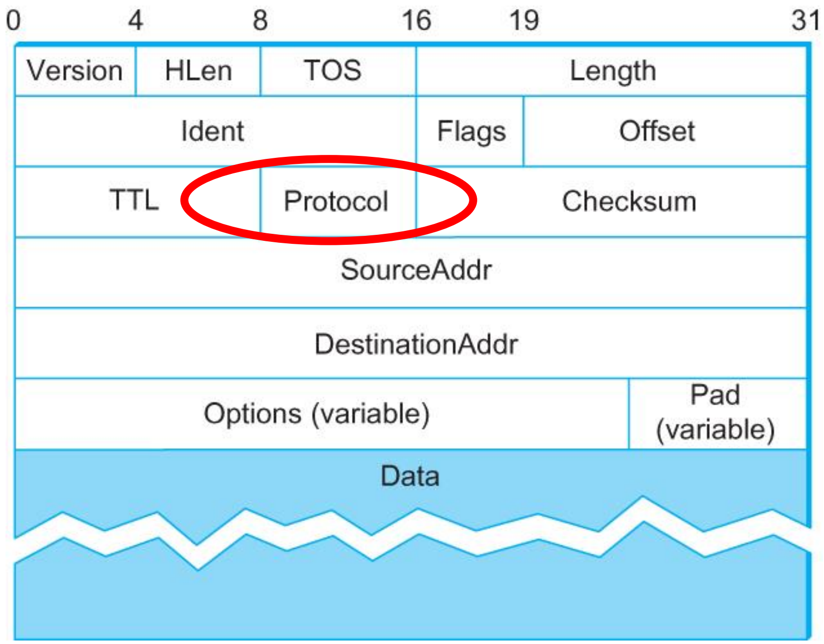
IP Header Format: Time to Live (TTL)

- Specifies the longest path before a datagram is dropped
  - ➢ Ensure that a packet is eventually dropped when a routing loop occurs
- Sender sets the value
  - ➢ Default: 64
- Each router decrements the value by 1
- When the value reaches 0, the datagram is dropped



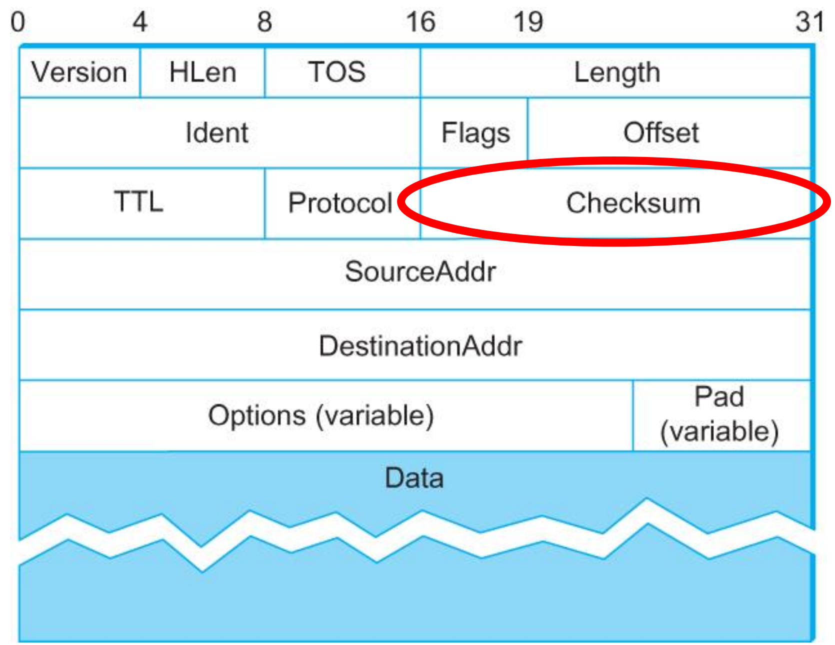
IP Header Format: Protocol

- 1 byte
- Specifies the higher-layer protocol
- De-multiplexing to higher layers
    - ➢ 6: TCP
    - ➢ 17: UDP



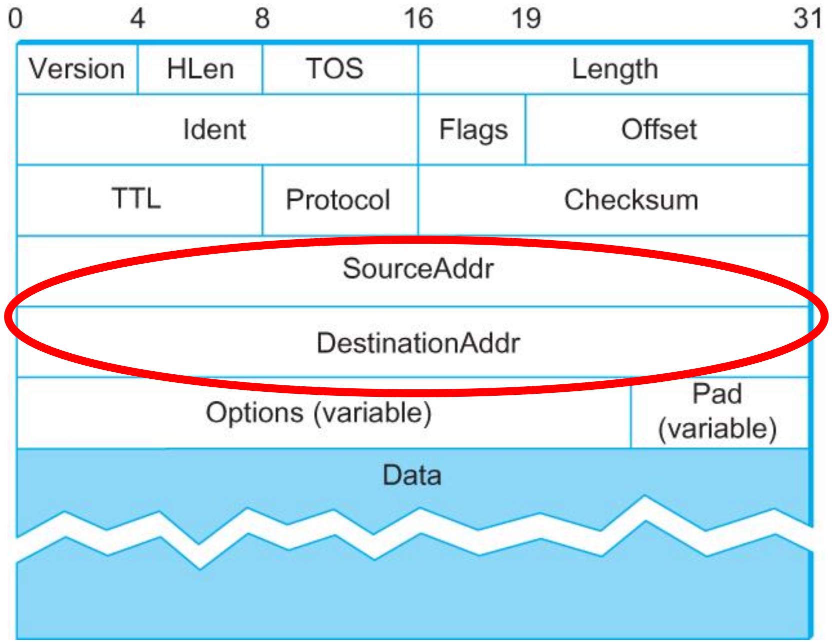
IP Header Format: Header Checksum

- 16 bits
- IP checksum
  - Not as strong as CRC, but easier to calculate in hardware
- Header only
- Must be computed at every hop
  - Why?



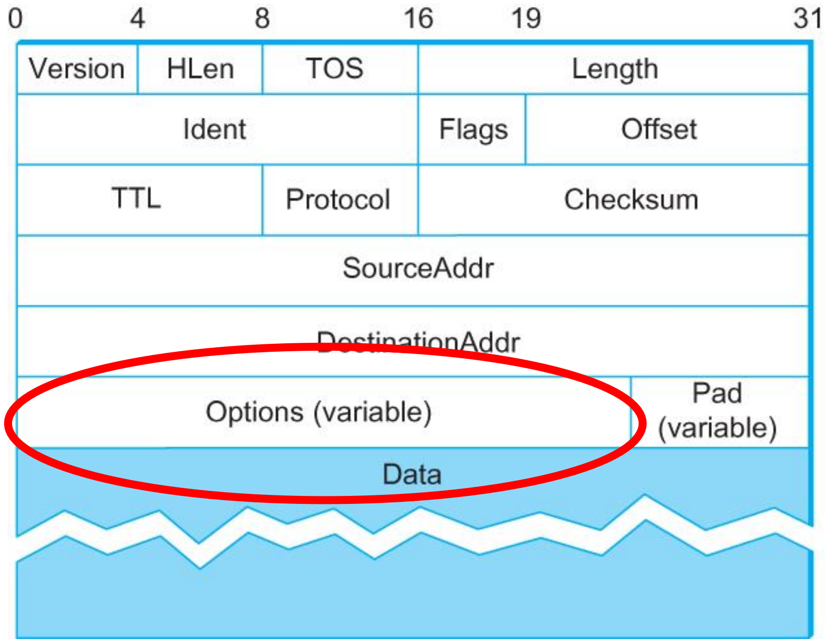
IP Header Format: Addresses

- Source & destination IP addresses
  - ➢ 32 bit address length in IPv4



IP Header Format: Options (1/2)

- Used infrequently
- IP options increase routers processing overhead
- Not included in IPv6

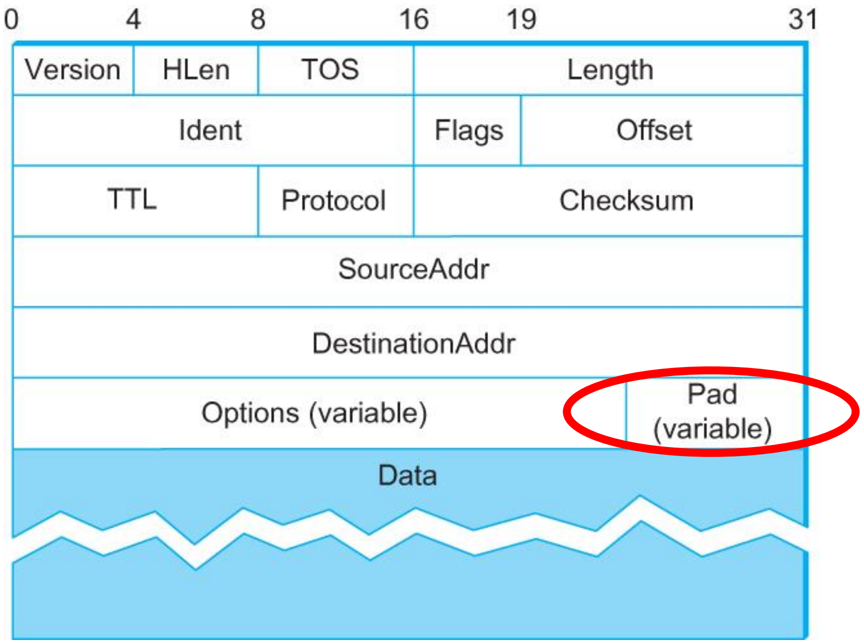


IP Header Format: Options (2/2)

- Record Route: each router that processes the packet adds its IP address to the header
- Timestamp: each router that processes the packet adds its IP address and time to the header
- (loose) Source Routing: specifies a list of routers that must be traversed
- (strict) Source Routing: specifies a list of the only routers that can be traversed

IP Header Format: Padding

- Padding bytes are added to ensure that header ends on a 4-byte boundary



Lecture Outline

- Internet protocol (IP)
- IP header format
- **IP addressing**
- IP forwarding
  - Forwarding algorithm

What is an IP Address?

- An IP address is a *unique global identifier* for a network interface
  - An IP address uniquely identifies a network location
- Routers forwards a packet based on the destination address of the packet
- Uniqueness ensures global reachability

IP Versions

- IPv4 (32-bit)
  - Classful IP addresses (obsolete, but important to understand)
  - Classless inter-domain routing (CIDR) (RFC 854, current standard)
- IP Version 6 addresses (128-bit)

Dotted Decimal Notation

- Each byte is identified by a decimal number in the range [0…255]:

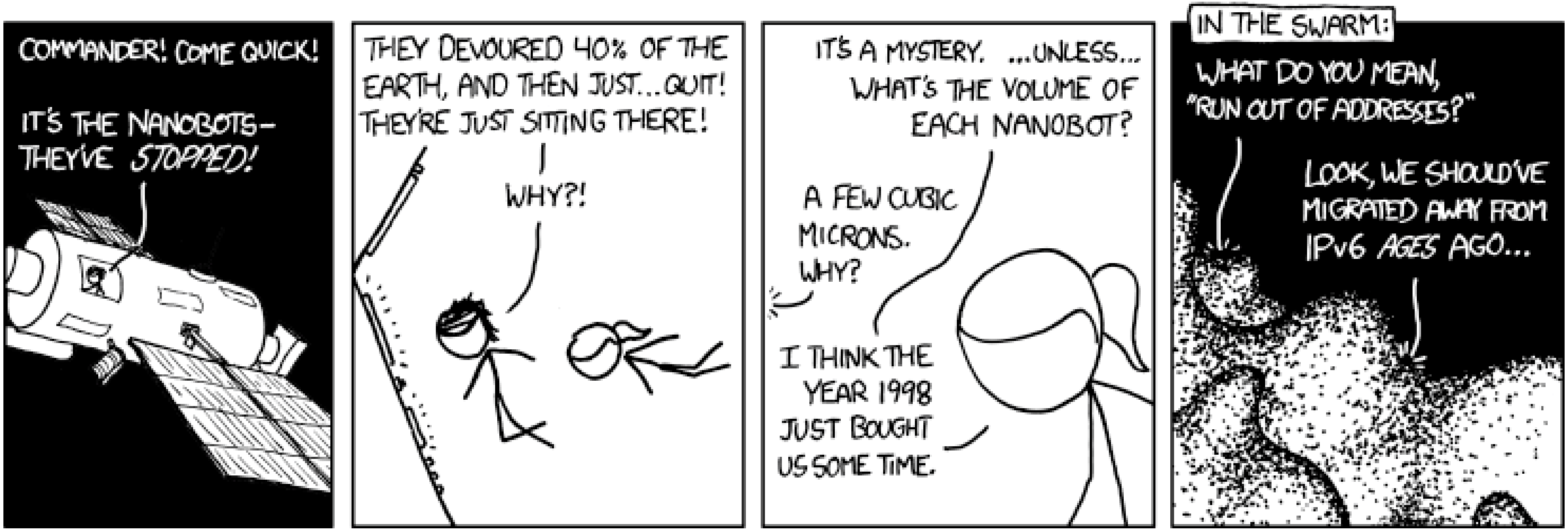
10000000	10001111	10001001	10010000
1st Byte	2nd Byte	3rd Byte	4th Byte
= 128	= 143	= 137	= 144

128.143.137.144

Will We Run Out of IP Addresses?

- Yes!
  - Especially since address space is not well-utilized
- This is the reason for IPv6
  - More on IPv6 in mid-October
- 32-bit address → 128-bit address
  - 4,294,967,296 (~ 4.3 bln) addresses with IPv4
  - 340,282,366,920,938,463,463,374, 607,431,768,211,456 addresses with IPv6. *$5*10^{28}$ addresses per every human*

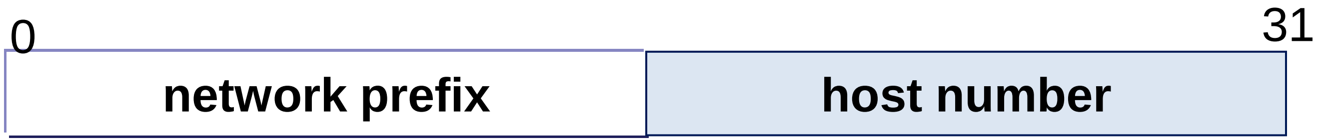
“If the Earth were made entirely out of 1 cubic millimetre grains of sand, then you could give a unique [IPv6] address to each grain in 300 million planets the size of the Earth”



What is the IP Address of google.com ?

- Linux and MAC OS: traceroute google.com
- Windows: tracert google.com

Structure of an IP Address



- An IP address has a structure
  - ➢ Network prefix identifies a network
  - ➢ Host number identifies a specific host interface
- Improves the scalability of routing
  - ➢ Scales better than flat addresses

How Long is a Network Prefix?

- Before 1993: The network prefix is implicitly defined
  - Class-based addressing
- After 1993: The network prefix is indicated by a netmask

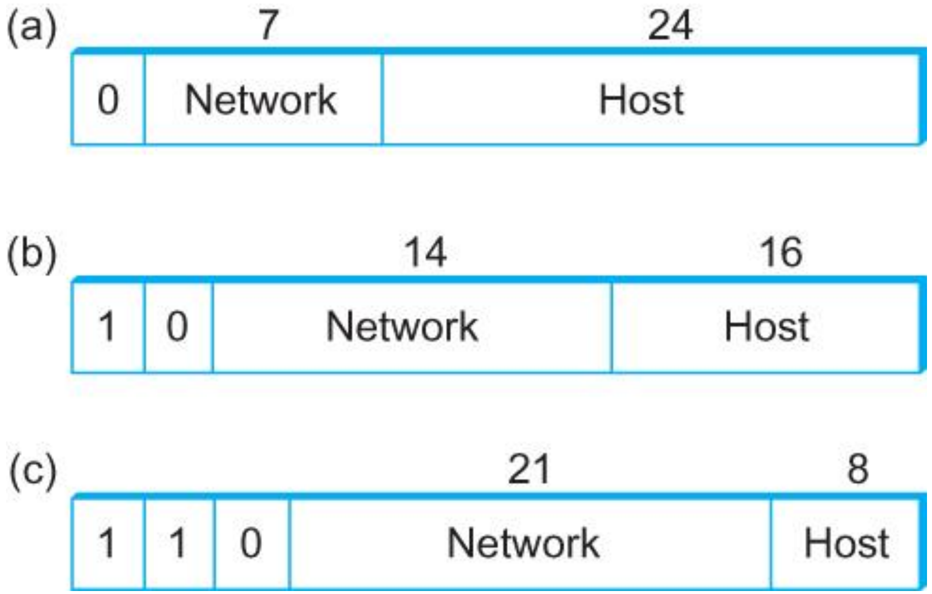
Before 1993: Class-based Addressing

- The Internet address space was divided up into classes:
  - **Class A:** Network prefix is 8 bits long
  - **Class B:** Network prefix is 16 bits long
  - **Class C:** Network prefix is 24 bits long
  - Class D: Multicast address
  - Class E: Reserved

Classful IP Addresses (Until 1993)

- Each IP address contained a key which identifies the class:
  - **Class A:** IP address starts with "0"
  - **Class B:** IP address starts with "10"
  - **Class C:** IP address starts with "110"
  - **Class D:** IP address starts with "1110"
  - **Class E:** IP address starts wit "11110"

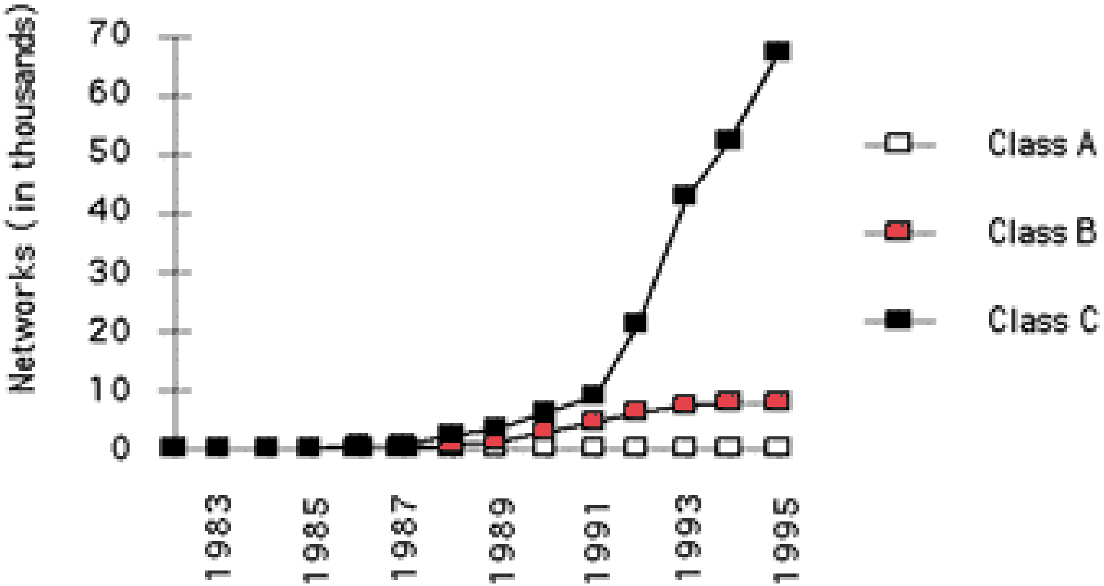
Classful IP Addresses: Class A,B,C Examples



Problems with Classful IP Addresses (1/2)

- Fast growing routing table size
  - Each router must have an entry for every network prefix
  - A,B too large, C too small
    - ~ 2^{21} = 2,097,152 class C networks
  - In 1993, the size of routing tables started to outgrow the capacity of routers
- Local admins must request another network number before installing a new network at their site

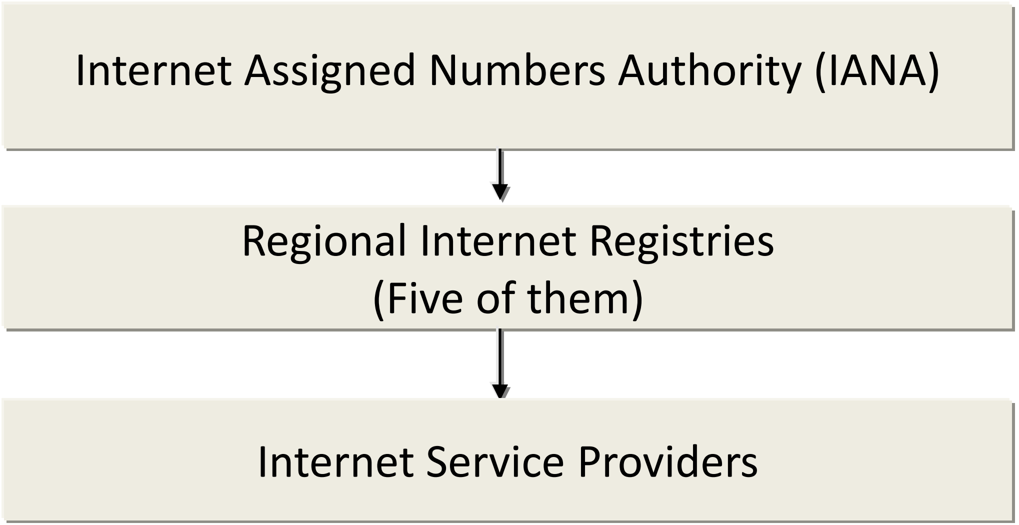
Problems with Classful IP Addresses (2/2)



Solution: Classless Inter-domain Routing (CIDR)

- Network prefix is of *variable length*
  - No rigid class boundary
- Addresses are allocated hierarchically
- Routers can aggregate multiple address prefixes into one routing entry
- Hierarchy is the key

Hierarchical IP Address Allocation



- American Registry for Internet Numbers (ARIN)
- RIPE, APNIC, LACNIC, AfriNIC

CIDR Network Prefix Has Variable Length

	128	143	137	144
Addr	**10000000**	**10001111**	**10001001**	**10010000**

	255	255	255	0
Mask	**11111111**	**11111111**	**1111111**	**00000000**

- A *network mask* specifies the number of bits used to identify a network in an IP address
  - Example above: 24-bit mask
  - Network prefix 128.143.137/24

CIDR Notation

- CIDR notation of an IP address:
  - E.g., 128.143.137.144/24
  - /24 is the prefix length
    - The first 24 bits are the network prefix of the address
    - The remaining 8 bits are available for specific host addresses

CIDR Network Prefix: Another Example

	128	143	31	144
Addr	**10000000**	**10001111**	**00011111**	**10010000**

	255	255	240	0
Mask	**11111111**	**11111111**	**11110000**	**00000000**

- A 20-bit mask example
  - Network prefix 128.143.16/20

CIDR Notation: Blocks of Addresses

- CIDR notation can nicely express blocks of addresses
  - An address block [128.195.0.0, 128.195.255.255] can be represented by an address prefix 128.195.0.0/16
  - How many IP addresses are there in a ***/x*** address block?
    - $2^{(32-x)}$

Example: Address Allocation

- Duke network operators receive a /16 address prefix 152.3.0.0/16 from ARIN
- Allocate address prefixes to three departmental networks
  - ME must have at least 50 hosts
  - ECE and CS must have at least 100 hosts
- Smallest address prefix to each department?

Lecture Outline

- Internet protocol (IP)
- IP header format
- IP addressing
- **IP forwarding**
  - Forwarding algorithm

Forwarding of IP Datagrams

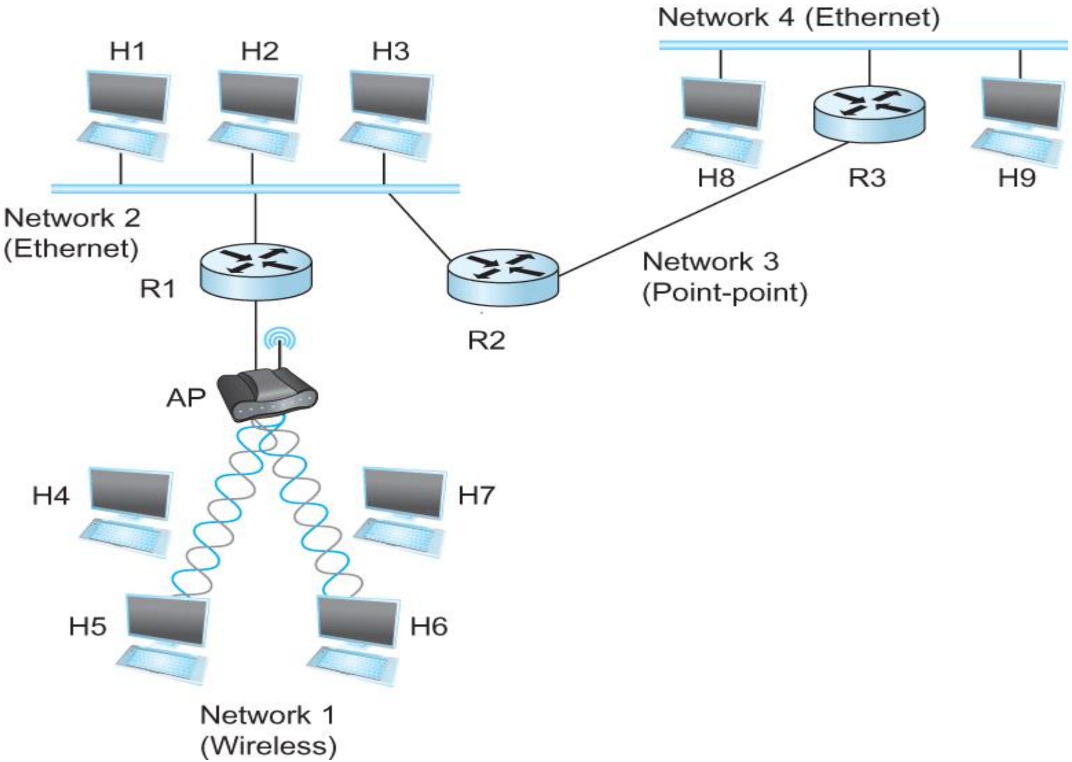
- There are two distinct processes to delivering IP datagrams:
  - *Forwarding* (data plane): How to pass a packet from an input interface to the output interface?
  - *Routing* (control plane): How to find and setup the forwarding tables?

Forwarding: Key Principles

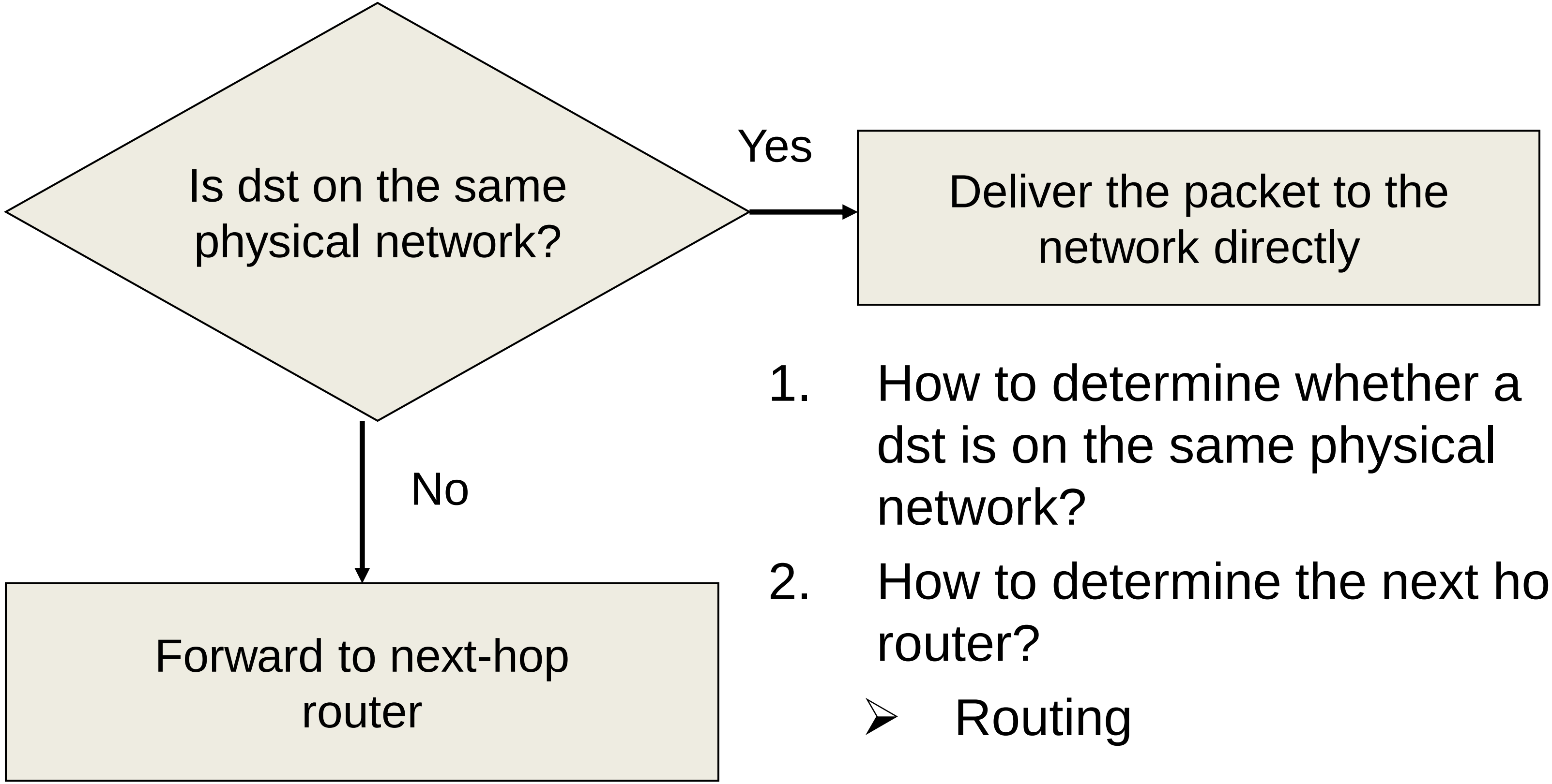
- Each IP datagram contains the IP destination address
- The “network part” of an IP address identifies a single physical network
- All hosts and routers that share the same network part of their address are connected to the same physical network
- Each physical network on the Internet has at least one router that connects this network to other physical networks

Forwarding Basics

- Routers forward according to network prefixes
- All interfaces on the same network have the same network prefixes



Forwarding Algorithm



1. How to determine whether a dst is on the same physical network?
2. How to determine the next hop router?
   - Routing

Detailed Forwarding Algorithm

- If (networkNum == networkNum of one of my interfaces) then
  - ➢ Deliver packet over the interface
- Else
  - ➢ if (NetworkNum is in my forwarding table) then
    - Deliver to the NextHop router
  - ➢ Else
    - Deliver packet to the default router

How Does a Host/Router Determine the Network Number of a Destination Address?

- Destination address & network mask = NetworkNumOfDestination
- If (NetworkNumOfDestination == my NetworkNum) then
  - Send through my direct interfaces

Forwarding Table Lookup

- Forwarding table lookup**:** use the IP destination address as a key to search the routing table
- Result of the lookup is the IP address of a next hop router, and/or the name of a network interface

Destination address	**Next hop/ interface**
network prefix *or* host IP address *or* loopback address *or* default route	IP address of next hop router *or* Name of a network interface

Type of Forwarding Table Entries (1/2)

- **Network route**
  - Destination addresses is a network address (e.g., 10.0.2.0/24)
  - Most entries are network routes
- **Host route**
  - Destination address is an interface address (e.g., 10.0.1.2/32)
  - Used to specify a separate route for certain hosts

Type of Forwarding Table Entries (2/2)

- **Default route**
  - Used when no network or host route matches
- **Loopback address**
  - Routing table for the loopback address (127.0.0.1)
  - The next hop lists the loopback (lo0) interface as outgoing interface

Unified Forwarding Algorithm

- Observation:
  - ➢ A directly connected physical network can be an entry in the forwarding table
  - ➢ A default route can be an entry

1. Look up destination address in the forwarding table using longest prefix match
2. Forward the packet to the next hop indicated by the matched entry

The Longest Prefix Matching Algorithm

1. Search for a match on all 32 bits
2. Search for a match for 31 bits

   …..

32. Search for a match on 0 bits

Host route, loopback entry

→ 32-bit prefix match

Default route is represented as 0.0.0.0/0

→ 0-bit prefix match

Why Longest Prefix Match?

- Longest → smallest network
- Network prefixes may be aggregated

An Example

128.143.71.21

Destination address	Next hop
10.0.0.0/8	eth0
128.143.0.0/16	R2
128.143.64.0/20	R3
128.143.192.0/20	R3
128.143.71.0/24	R4
128.143.71.55/32	R3
0.0.0.0/0 (default)	R5

- The longest prefix match for 128.143.71.21 is for 24 bits with entry 128.143.71.0/24
- Datagram will be sent to R4
- Note how this improves scalability

Lecture Summary

- Internet protocol (IP)
- IP header format
- IP addressing
- IP forwarding
  - Forwarding algorithm

Next Lecture

- IP fragmentation
- Address Resolution Protocol (ARP)
- Internet Control Message Protocol (ICMP)