ECE/COMPSCI 356 Computer Network Architecture

Lecture 10: The Internet Protocol (IP)

Neil Gong

neil.gong@duke.edu

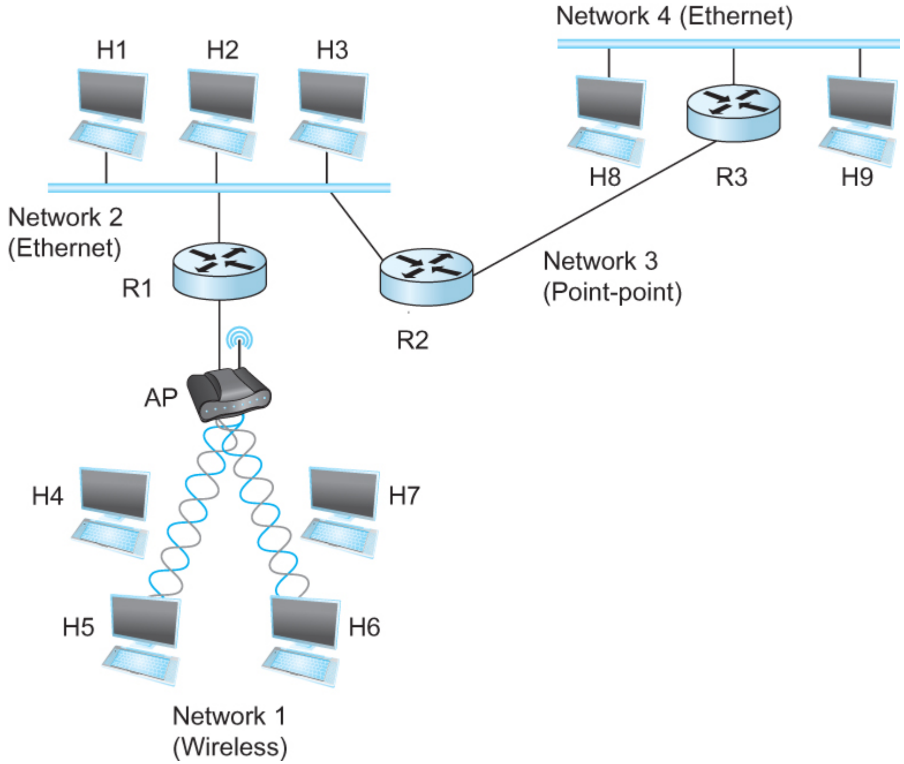
Slides credit: Xiaowei Yang, PD

Roadmap

- IP header format
- IP addressing
- IP forwarding

Internetworking

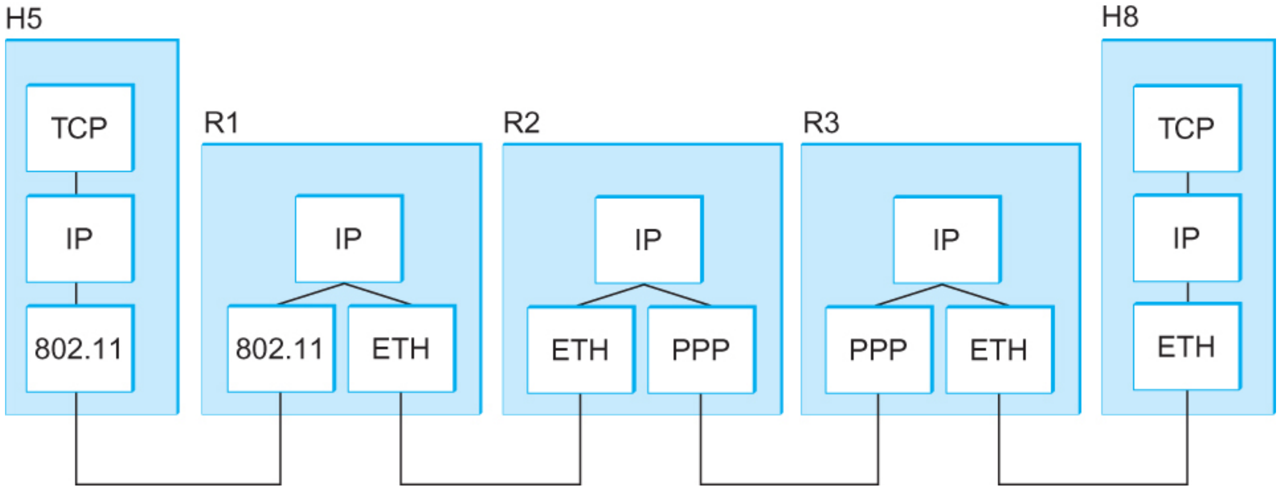
- What is internetwork
  - An arbitrary collection of networks interconnected to provide some sort of host-host to packet delivery service



A simple internetwork where H represents hosts and R represents routers

Internetworking

- What is IP
  - IP stands for Internet Protocol
  - Key tool used today to build scalable, heterogeneous internetworks
  - It runs on all the nodes in a collection of networks and defines the infrastructure that allows these nodes and networks to function as a single logical internetwork



A simple internetwork showing the protocol layers

IP Service Model

- Packet Delivery Model
  - Connectionless model for data delivery
  - Best-effort delivery (unreliable service)
    - packets are lost
    - packets are delivered out of order
    - duplicate copies of a packet are delivered
    - packets can be delayed for a long time
- Global Addressing Scheme
  - Provides a way to identify all hosts in the network

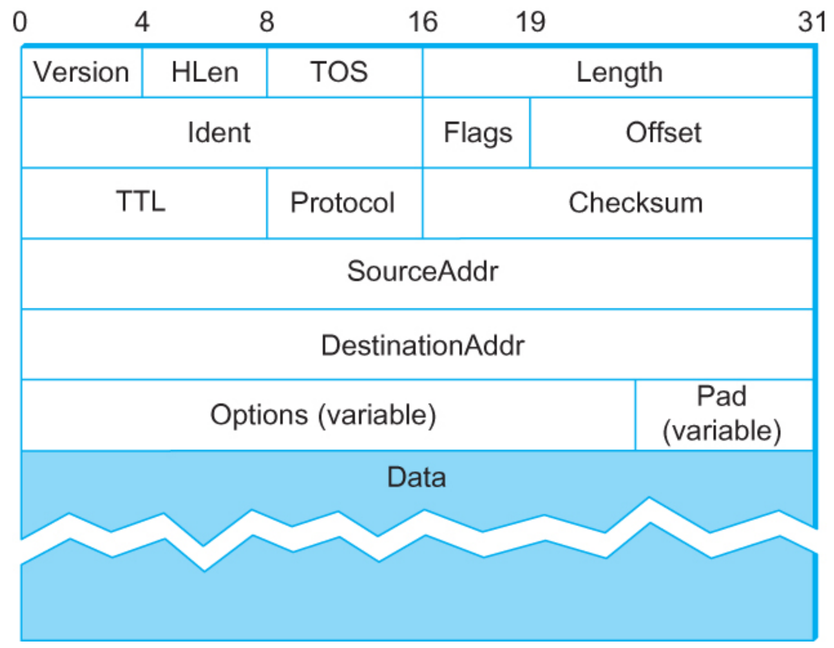
Basic IP functions

- Things you need to understand to do lab2
  - Internet protocol
    - IP header
    - IP addressing
    - IP forwarding
  - Address resolution protocol
  - Error reporting and control
    - Internet Control Message Protocol

Basic IP functions

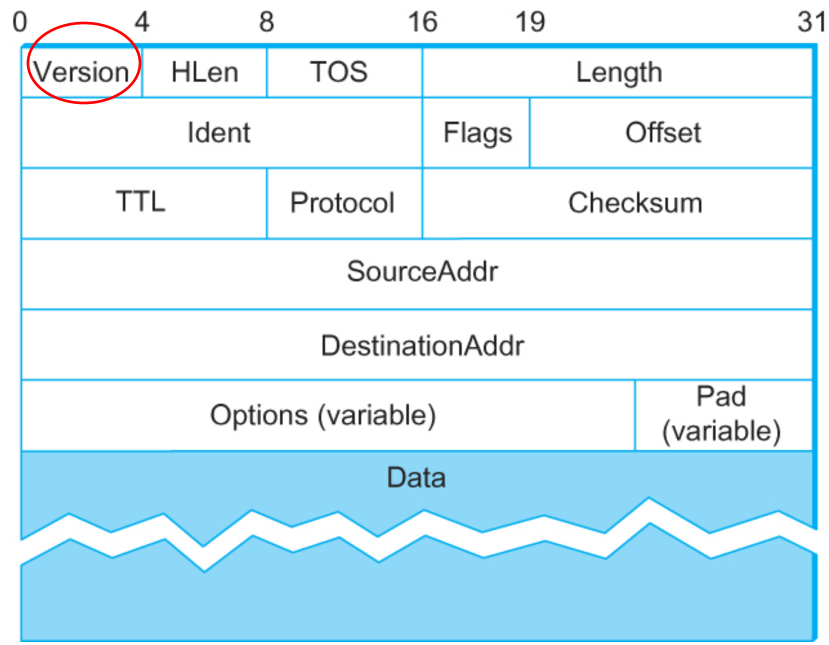
- Things you need to understand to do lab2
  - Internet protocol
    - **IP header**
    - IP addressing
    - IP forwarding
  - Address resolution protocol
  - Error reporting and control
    - Internet Control Message Protocol

IP header format



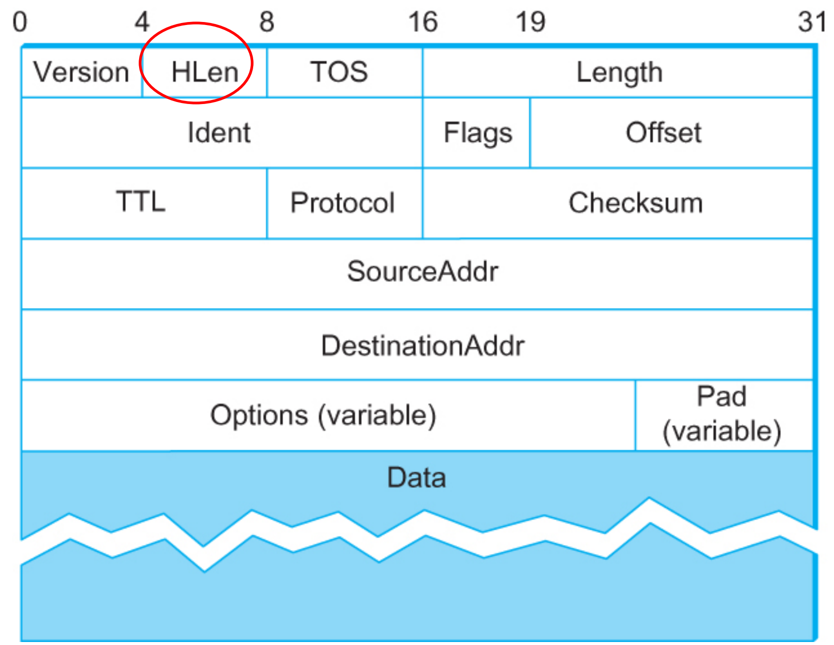
- 20 bytes fixed length header + variable-length Options

IP header format



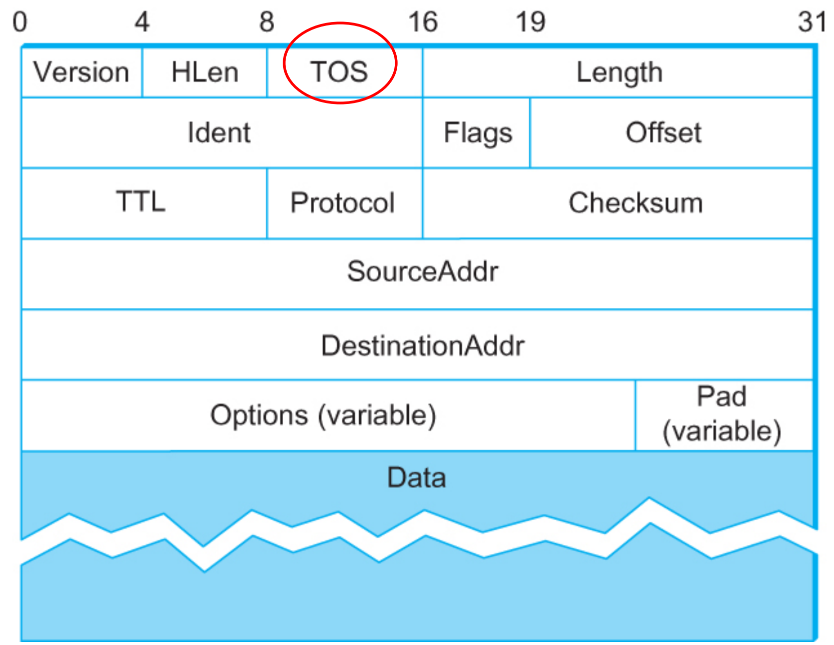
- **Version (4 bits):** 4

IP header format



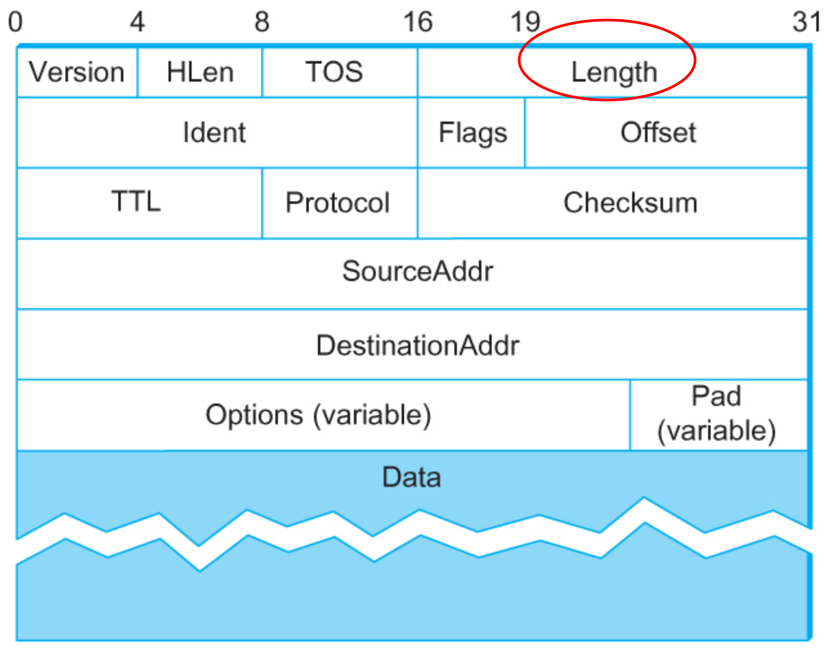
- **HLen (4bits):** the length of header in 32-bit words
  - Maximum header length?

IP header format



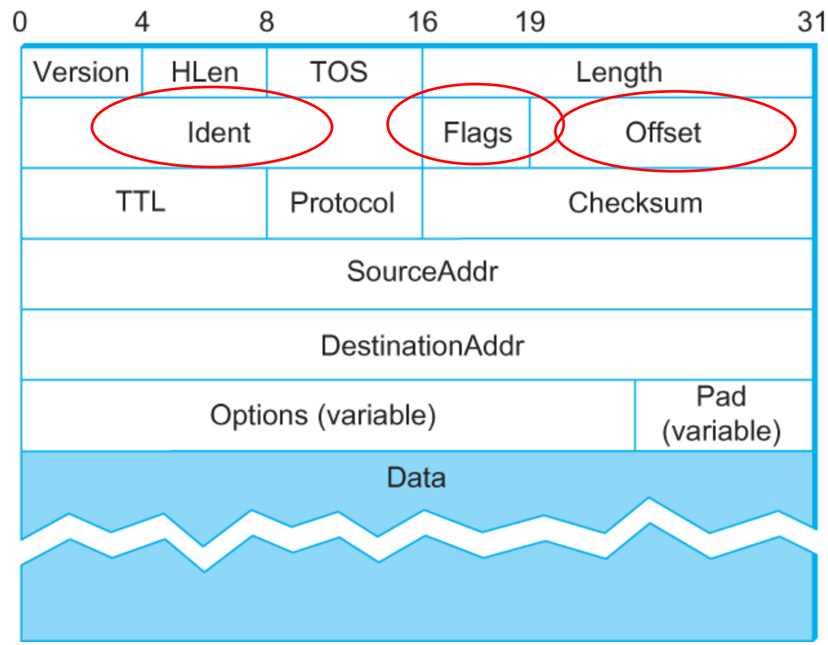
- **TOS (8):** type of service (not widely used)
  - How to treat a packet, e.g., low delay

IP header format



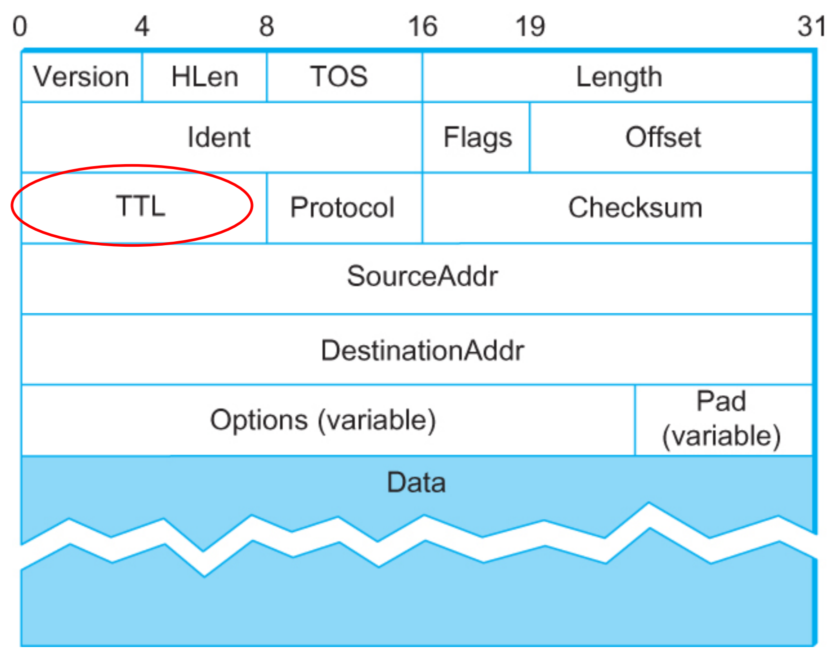
- **Length (16 bits):** packet length in bytes, including the header
  - 65,535 bytes
  - Fragmentation and reassembly

IP header format



- **Ident, Flags, Offset**
  - Used for fragmentation and assembly

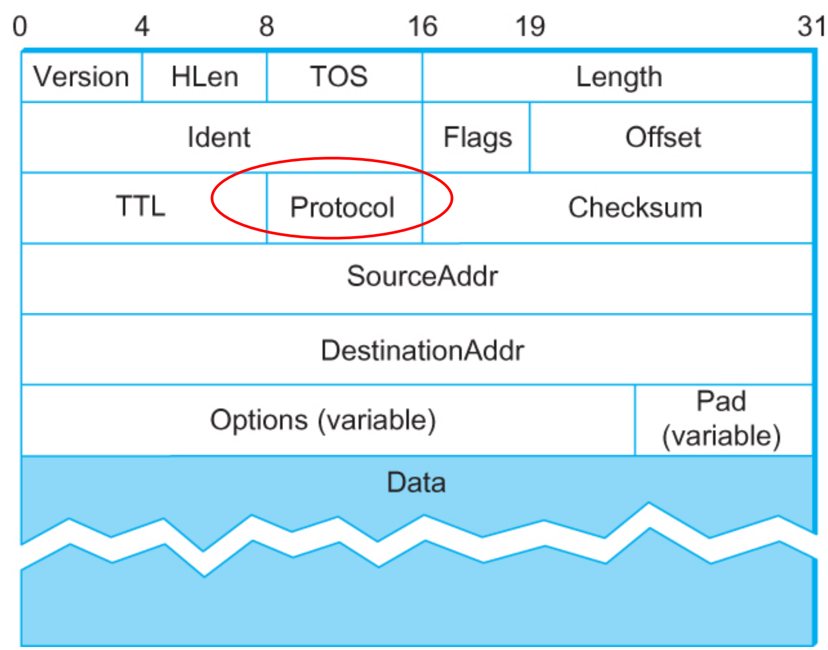
IP header format



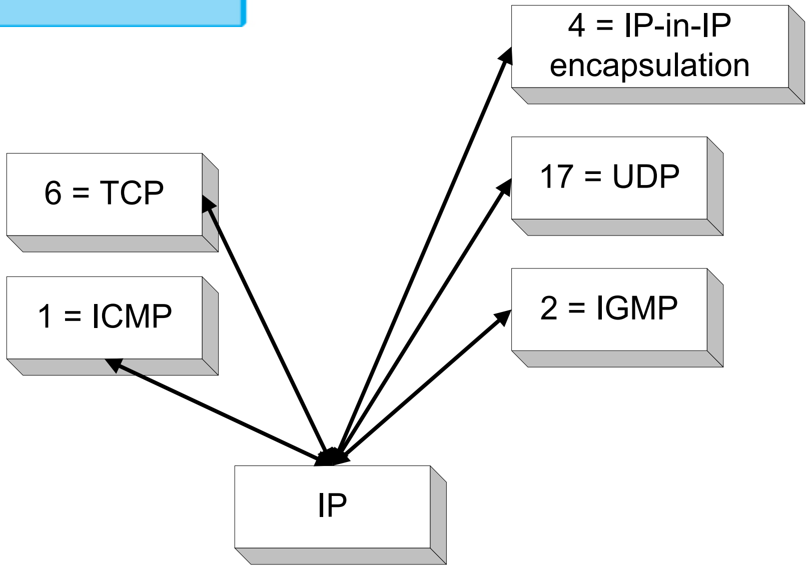
- **Time To Live (TTL) (1byte):**
  - Specifies the longest path before a datagram is dropped
  - Role of TTL field: Ensure that a packet is eventually dropped when a routing loop occurs

  Used as follows:
  - Sender sets the value (e.g., 64)
  - Each router decrements the value by 1
  - When the value reaches 0, the datagram is dropped

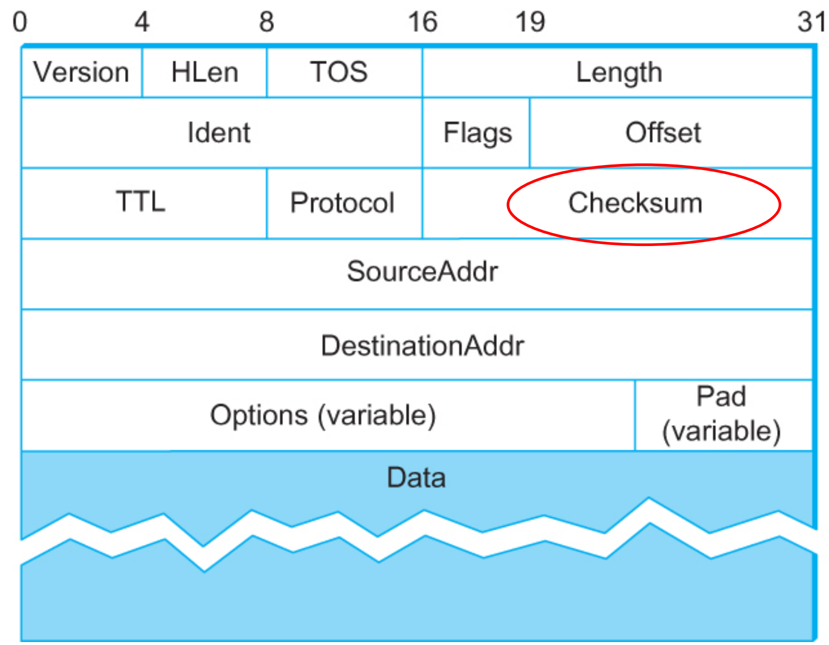
IP header format



- **Protocol (1 byte):**
  - Specifies the higher-layer protocol.
  - De-multiplexing to higher layers.



IP header format



- **Header checksum (16 bits):** header checksum
  - Header only
  - Computed at every hop

Internet Checksum Algorithm

- Sender
  - Consider the header as a sequence of 16-bit integers.
  - Add them together using 16-bit ones complement arithmetic (explained next slide) and then take the ones complement of the result.
  - That 16-bit number is the checksum
- Receiver
  - Perform the same calculation on the received header and compare the result with the received checksum
  - If any transmitted header, including the checksum itself, is corrupted, then the results will not match, so the receiver knows that an error occurred

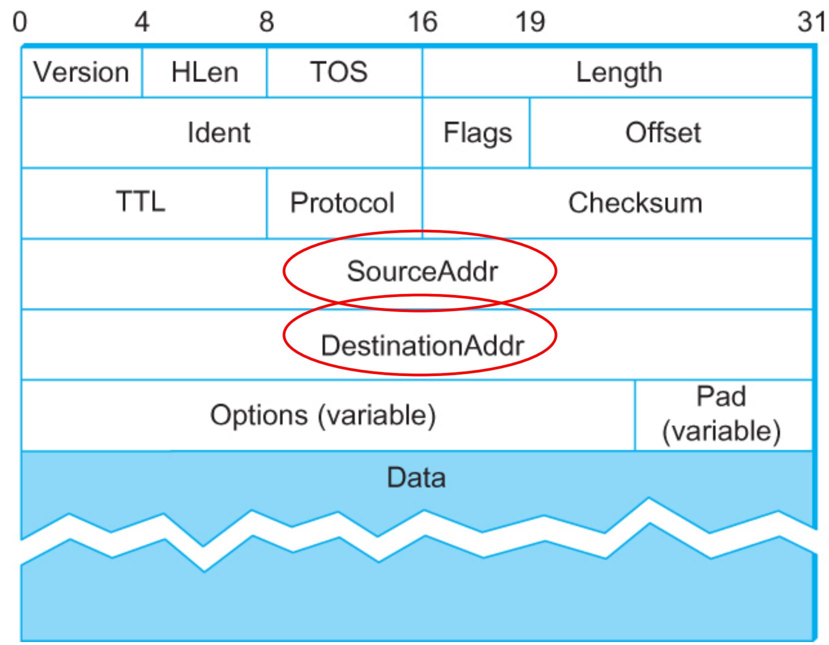
Internet Checksum Algorithm

- When adding numbers in ones complement arithmetic, a carryout from the most significant bit needs to be added to the result.

Internet Checksum Algorithm

- If we add 1010 and 1100 ignoring the carry, we get 0110
- In ones complement arithmetic, the fact that this operation caused a carry from the most significant bit causes us to increment the result, giving 0111

IP header format



- **SourceAddr, DestinationAddr**
  - Source & destination IP addresses

IP header format

- **Options:**
  - 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 options increase routers processing overhead

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

Overview

- IP header format
- **IP addressing**
- IP forwarding

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 forward a packet based on the destination address of the packet

- Uniqueness ensures global reachability

IP versions

- IPv4 (32-bit)
  - Classful IP addresses (obsolete)
  - 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

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**

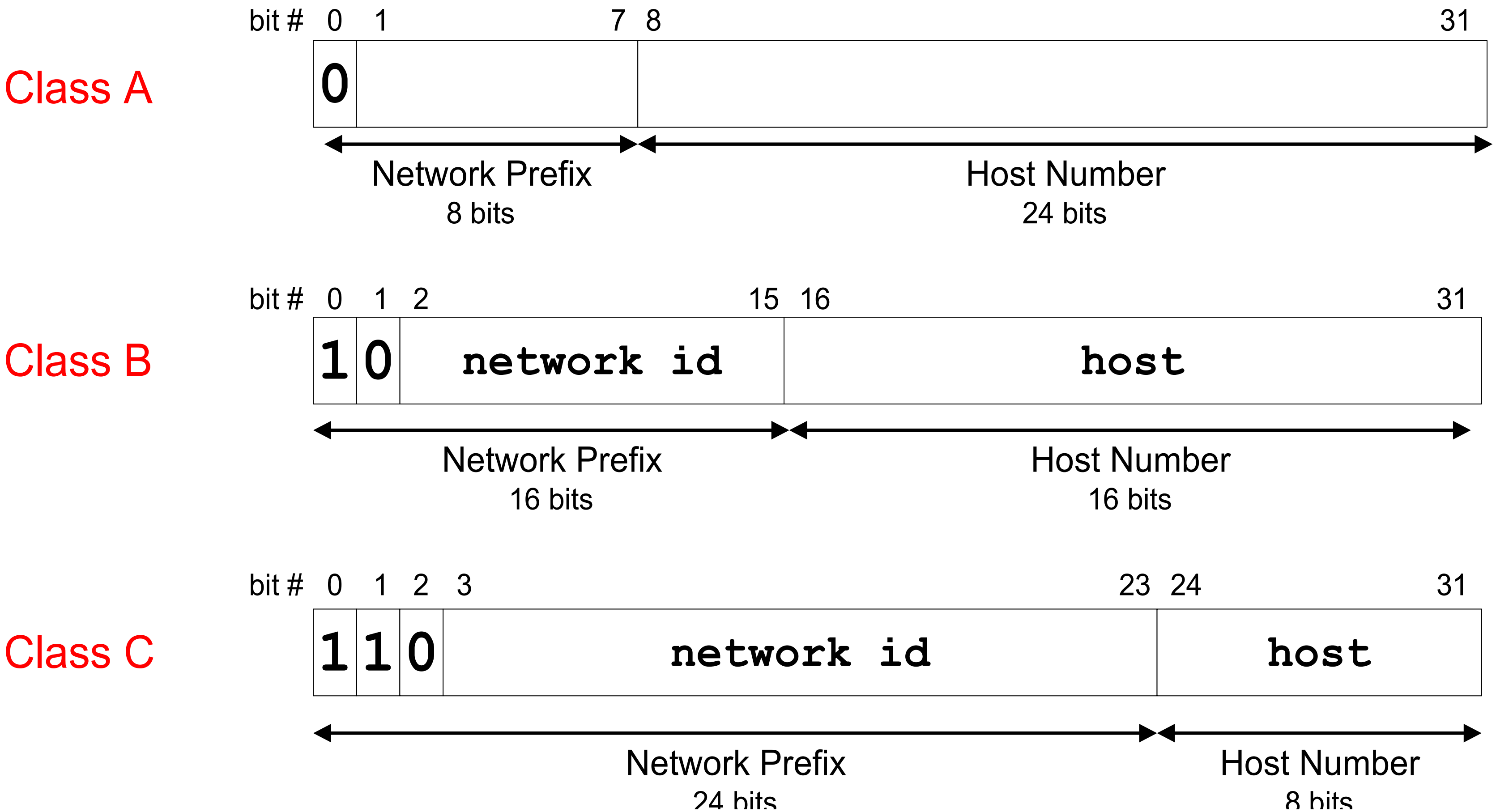
Before 1993: Class-based addressing

- 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 is multicast address
  - Class E is 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 (before 1993)



Classful IP Addresses (before 1993)

Class D

bit #	0	1	2	3	4 … 31
	1	1	1	0	multicast group id

Class E

bit #	0	1	2	3	4	5 … 31
	1	1	1	1	0	(reserved for future use)

Problems with Classful IP Addresses

- A,B too large, C too small

- Local admins must request another network number before installing a new network at their site

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

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 prefix bits used to identify a network in an IP address.

CIDR notation

- CIDR notation of an IP address:
  - 128.143.137.144/24
  - /24 is the prefix length. It states that the first 24 bits are the network prefix of the address (and the remaining 8 bits are available for specific host 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?
    - $2^{(32-x)}$

Overview

- IP header format
- IP addressing
- **IP forwarding**

Forwarding of IP datagrams

- There are two distinct processes for delivering IP datagrams:

  1. **Forwarding (data plane):** How to pass a packet from an input interface to the output interface?
  2. **Routing (control plane):** How to find and setup the forwarding tables?

Key points

- 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

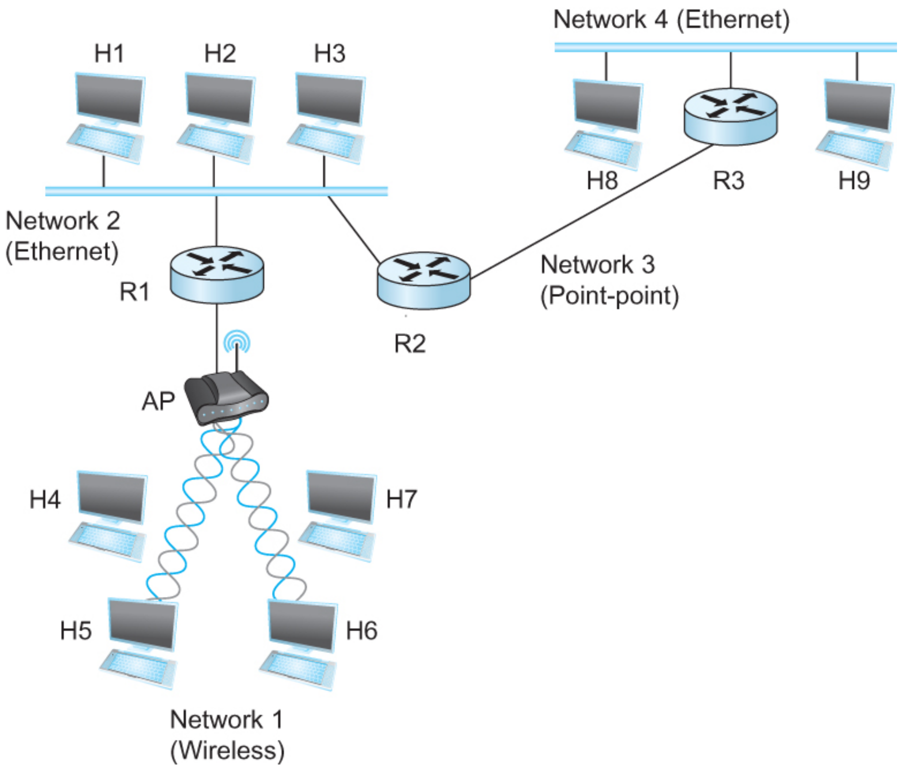
Forwarding basics

- Routers forward packets according to network prefixes
  - network prefixes are called NetworkNum
  - Each router has a forwarding table
    - Entry: (NetworkNum, NextHop)

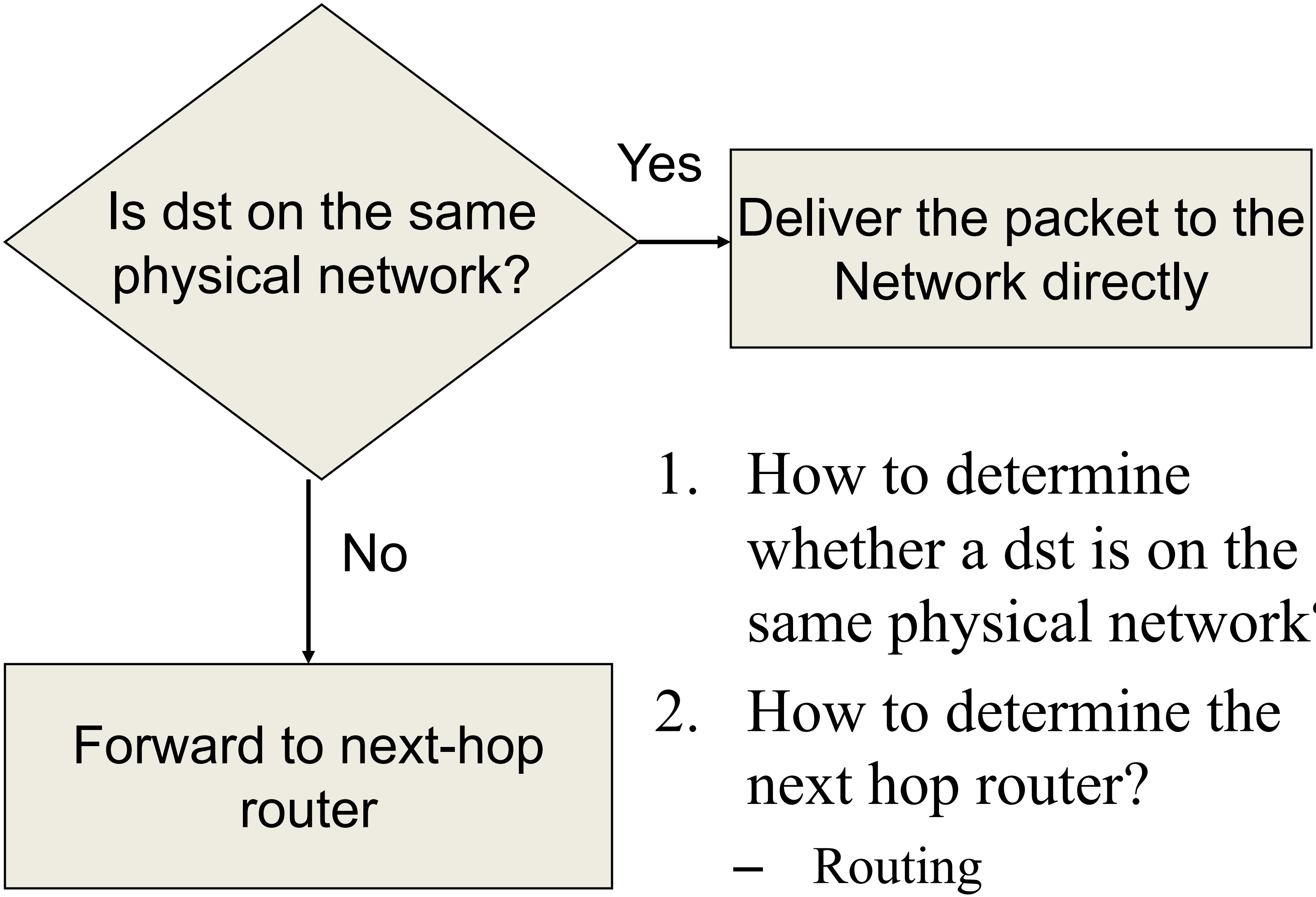
Forwarding table – an example

Forwarding table for R2

NetworkNum	NextHop
1	R1
2	Interface 1
3	Interface 0
4	R3



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 to determine the NetworkNum of a destination address?

- Destination address & network mask = NetworkNum

	128	143	137	144
Addr	10000000	10001111	10001001	10010000
	255	255	255	0
Mask	11111111	11111111	1111111	00000000

- Equivalent: NetworkNum=network prefix bits

Type of forwarding table entries

- **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
- **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 interface as outgoing interface

Unified forwarding algorithm

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

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

Summary

- IP header format
- IP addressing
- IP forwarding