Chapter 4: outline

4.1 introduction

4.2 virtual circuit and datagram networks

4.3 what's inside a router

4.4 IP: Internet Protocol

- datagram format
- IPv4 addressing
- ICMP
- IPv6

4.5 routing algorithms

- link state
- distance vector
- hierarchical routing

4.6 routing in the Internet

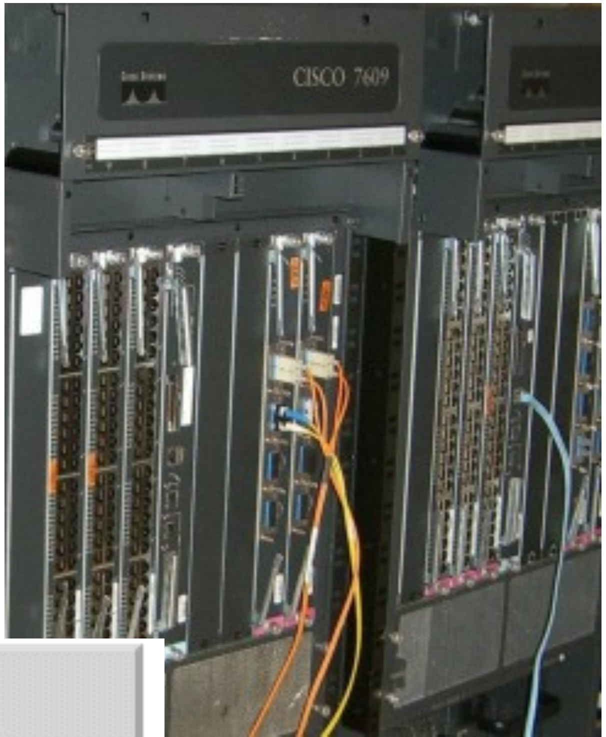
- RIP
- OSPF
- BGP

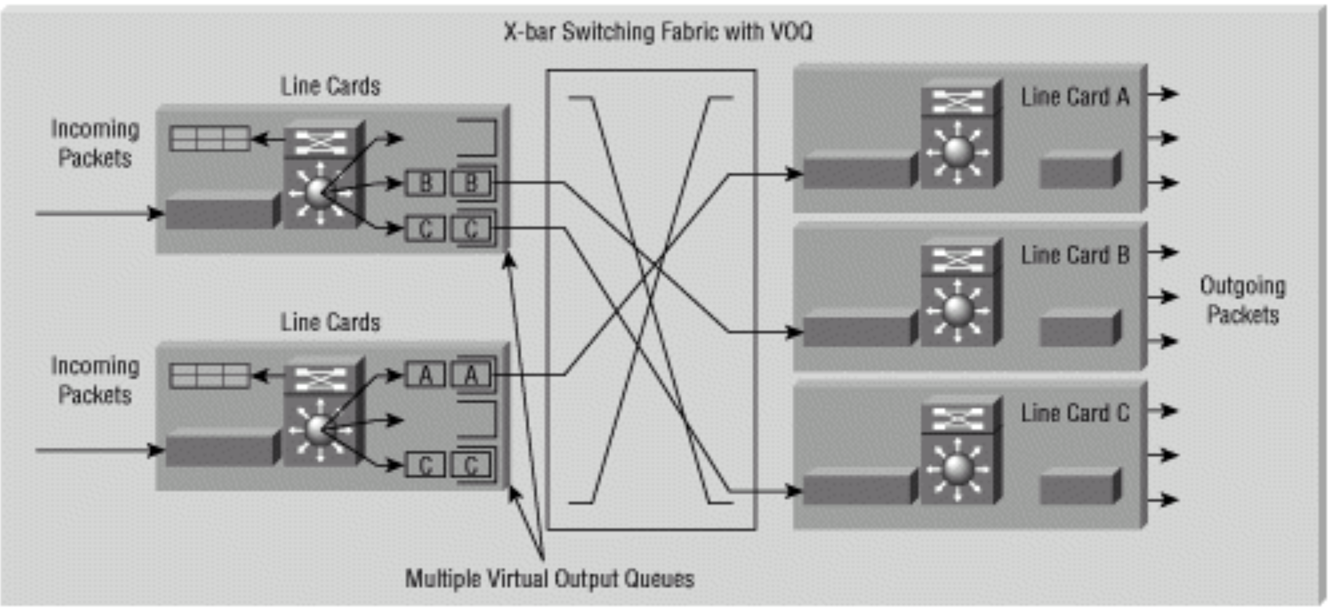
4.7 broadcast and multicast routing

Router

two key router functions:

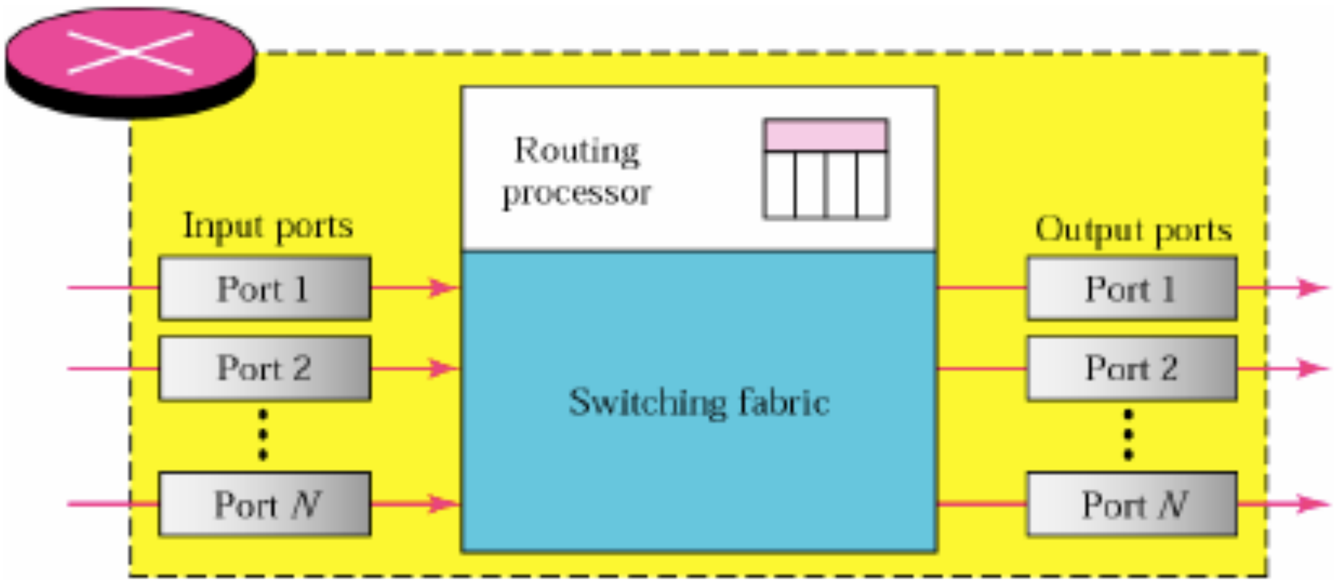
- run routing algorithms/protocol (RIP, OSPF, BGP)
- *forwarding* datagrams from incoming to outgoing link





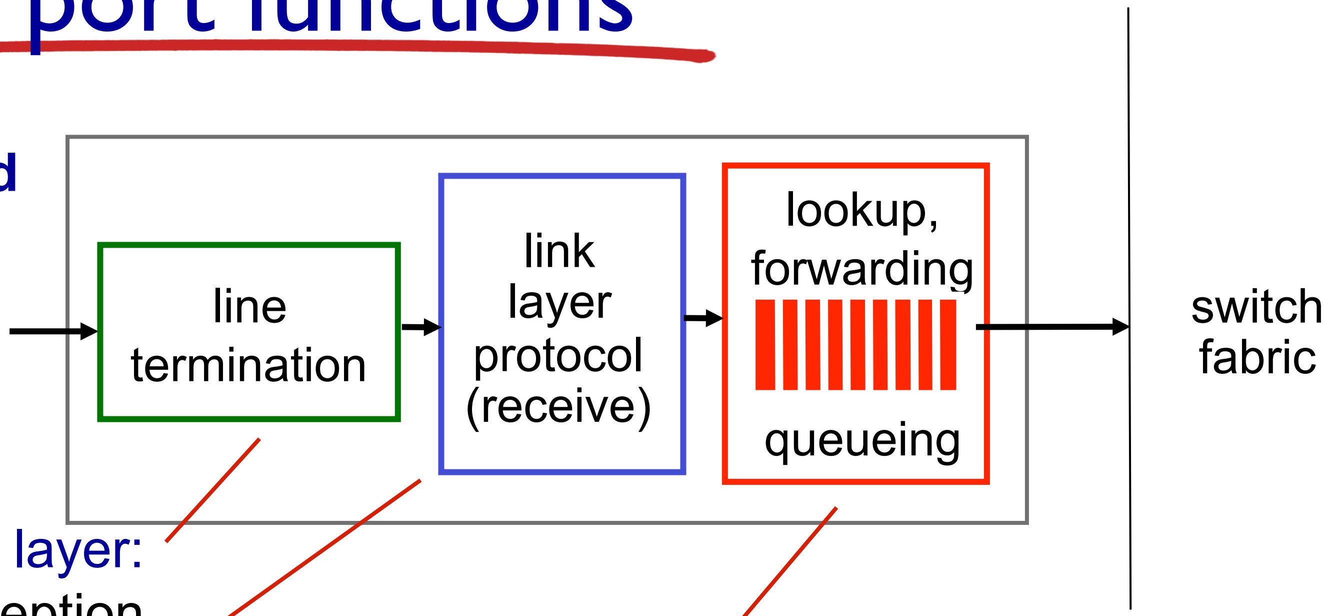
Router architecture overview

- Main components:
  - Input ports/Interfaces
  - Switching fabric
  - Output ports/Interfaces
  - Routing processor: (1)executing routing protocol, (2)maintaining routing information, forwarding tables, etc.



Input port functions

Line card



physical layer:
bit-level reception

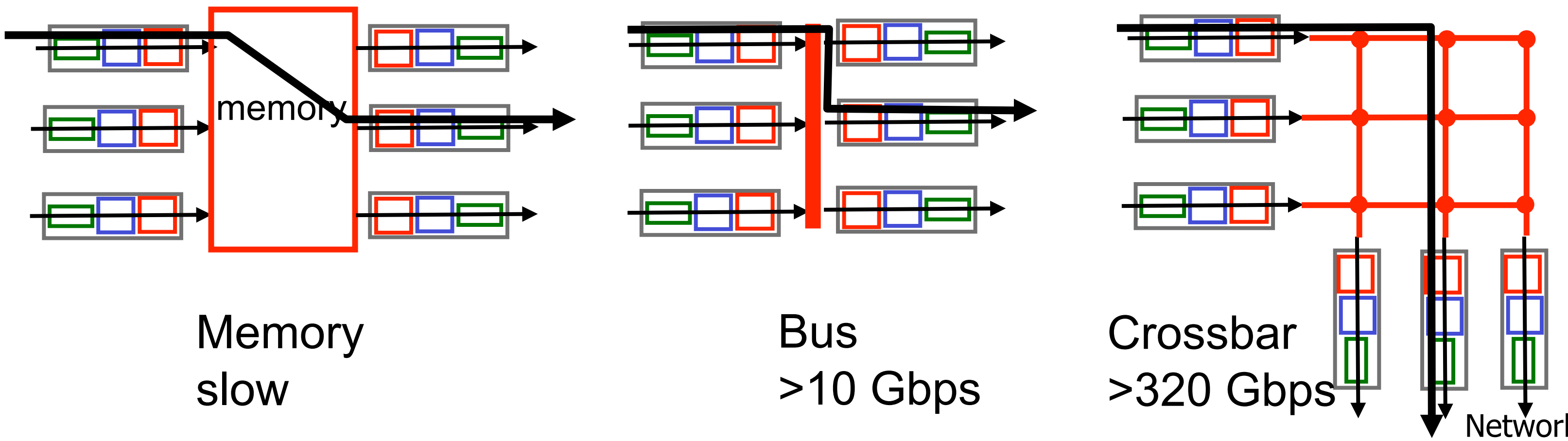
data link layer:
decapsulation, error checking, etc

Network layer – decentralized switching:

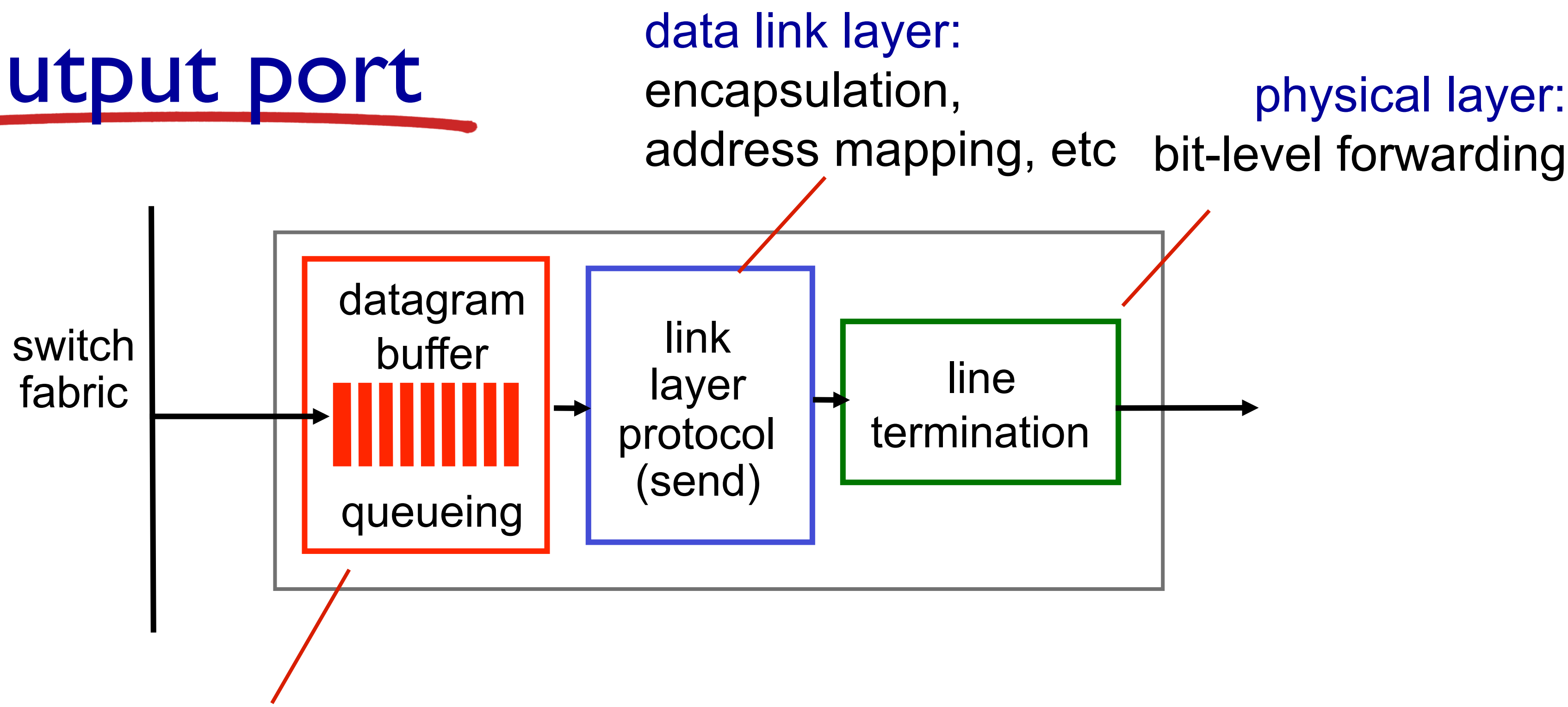
- Packet forwarding = decide which output line to forward each packet based on packet header.
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

Switching fabrics

- Switching fabric function – transfer packets between input and output line cards
- Types of switching fabric
  - Via memory: datagram is received through input port, stored in memory, then send to output port – slow.
  - Via a bus: datagram is sent directly from input to output via a shared bus – does not scale well
  - Via a crossbar: interconnection network consisting of 2N busses that interconnect N input and N output



Output port



- *buffering* required when datagrams arrive from fabric faster than the transmission rate
- *Buffer management* decide when and which packets to drop if there is not enough memory to store all income packets
- *scheduling discipline* decide which packet, of those queued to send out next

Chapter 4: outline

Internet Protocol (IP)

- Host-to-host network-layer delivery protocol for the Internet with following properties
  - Connectionless service – each packet is handled independently
  - Best-effort delivery service
    1. Does its best to deliver packet to its destination, but with no guarantees
    2. Limited error control - only error detection, corrupted packets are discarded
    3. No flow control
  - Must be paired with a reliable transport – (TCP) and/or application-layer protocol to ensure reliability

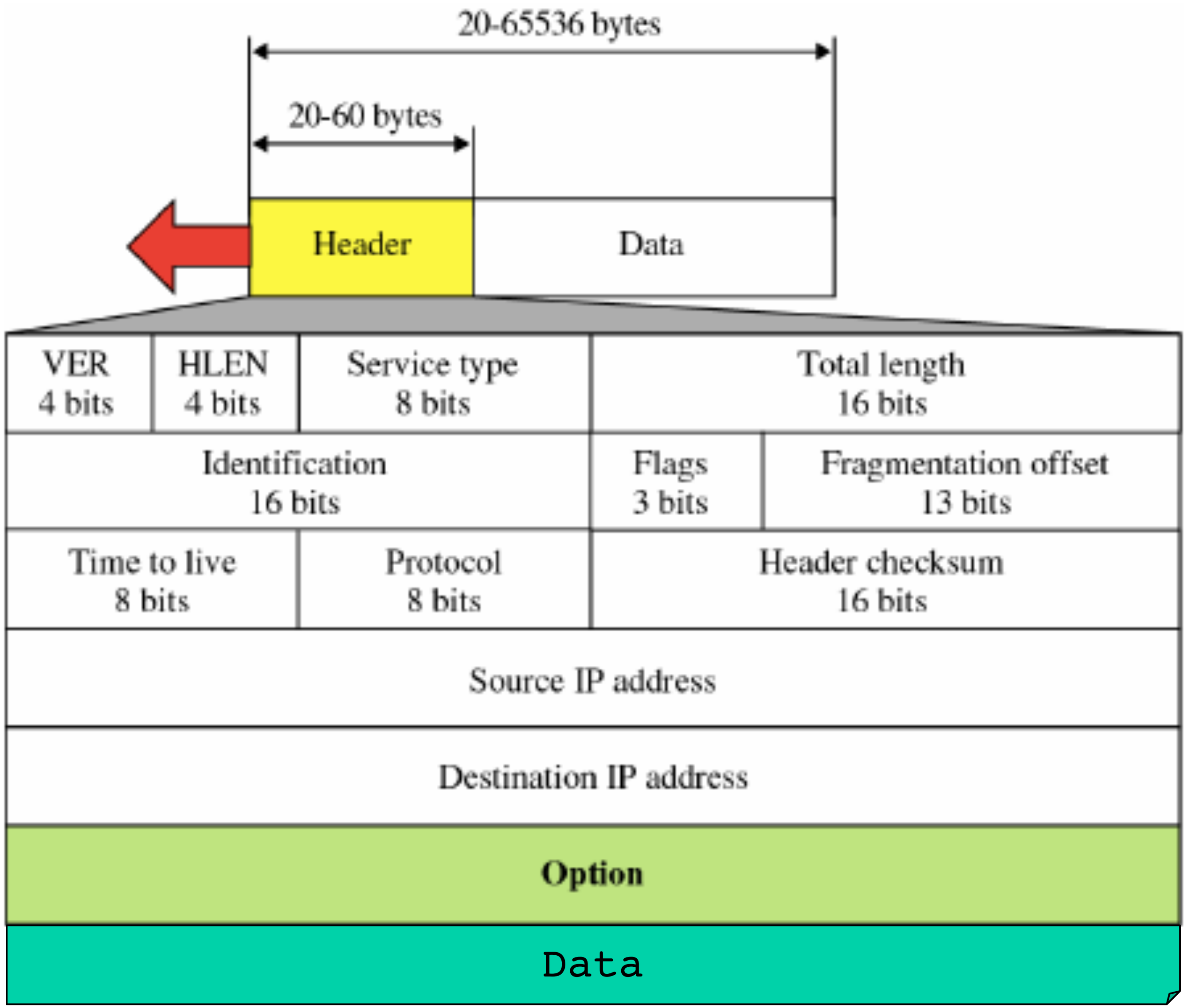
IP Versions

- IPv4, IPv6, Mobile IP
  - IPv4 – version currently in wide use (formalized in 1981)
  - IPv6 – new version created to correct some of significant problems of IPv4 such as exhaustion of address space (formalized in 1996)
  - Mobile IP – enhanced version of IPv4 which supports IP in mobile environments (formalized in 1996)

IP datagram format

- Datagram – IP packet = variable length packet consisting of *header* and *data*
  - Header – 20 to 60 bytes in length, contains information essential to routing and delivery
  - Data – length determined by Maximum Transmission Unit (MTU) of link layer protocol (theoretically between 20 to 65536 bytes)

IP datagram format



IP Datagram Fields

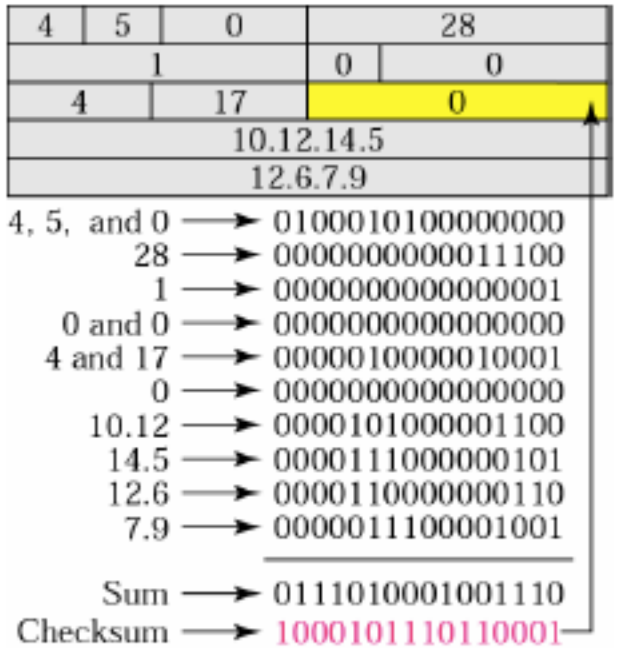
- **Version number** – 4-bit field, specifies IP protocol version of the datagram (IPv4 or IPv6)
  - Different versions of IP use different datagram formats
  - By looking at version number router can determine how to interpret remainder of datagram
- **Header length** – 4-bit field, defines total length of datagram header in 4-byte words
  - When there are no options header length is 20 → HLEN = 5
- **Service type** – 8-bit field, allows different types of datagram to be distinguished from each other based on their associated/requested QoS.

IP Datagram Fields (cont.)

- Time-To-Live (TTL) – 8-bit field, controls maximum number of hops visited by datagram and/or time spend in the network
  - Field is decremented by one each time datagram is processed by a router – *when TTL reaches 0, datagram must be dropped*.
  - Ensures that (1) datagram does not circulate/loop forever, or (2) to limit its journey, e.g. LAN only: TTL=1.
- Protocol – 8-bit field, indicates specific higher-level protocol that uses the services of IP layer (IP datagram can encapsulate data from a number of higher-layer protocols
  - Used only at final destination to facilitate demultiplexing
  - Protocol number is glue that binds network and transport layer (similar to port number that binds transport and appl. layers)
  - Values: 1 – ICMP, 2 – IGMP, 6 – TCP, 17 – UDP, 89 - OSPF

IP Datagram Fields (cont.)

- Header checksum – 16-bit field, aids in detecting errors in header only!
  - Checksum must be recomputed and stored again at each router as TTL and some options fields may change.
  - Router discard datagrams for which an error is detected.
  - Checksum calculation:
    1) Divide header into 16-bit sections – checksum field itself is set to 0
    2) Sum all sections using 1s complement arithmetic

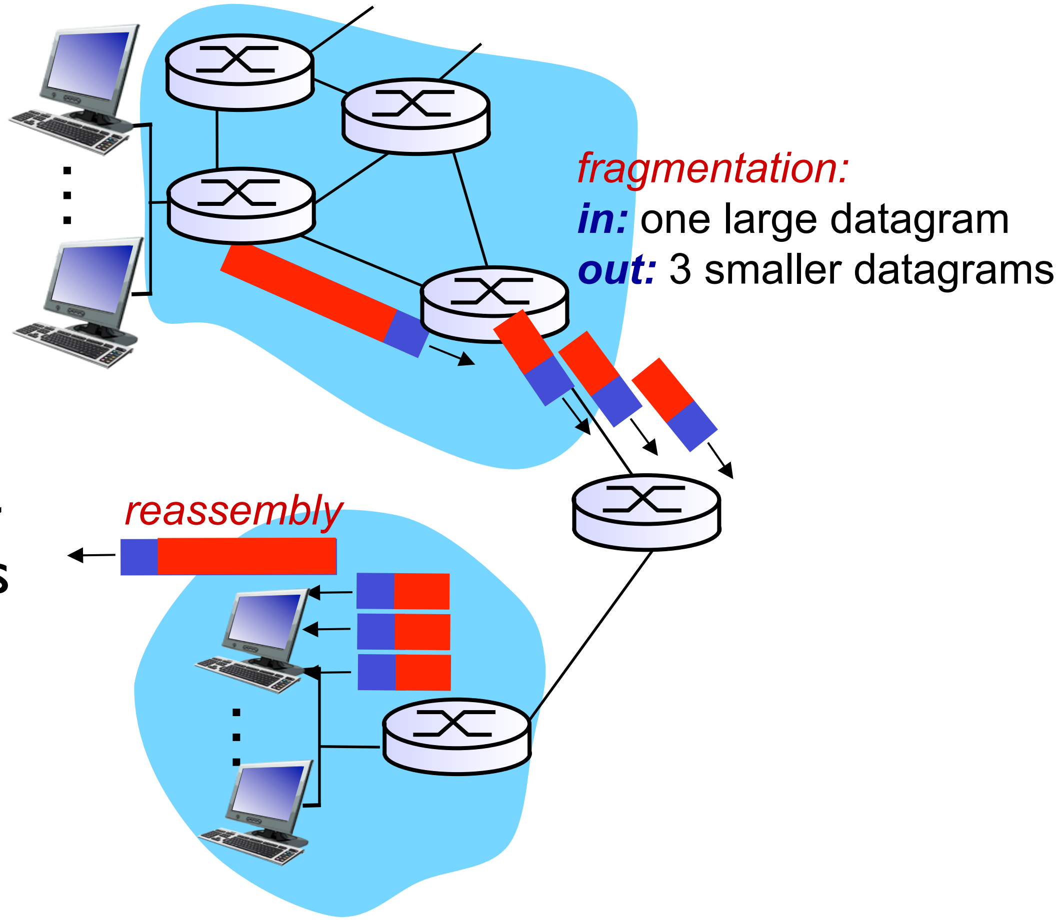


IP Datagram Fields (cont.)

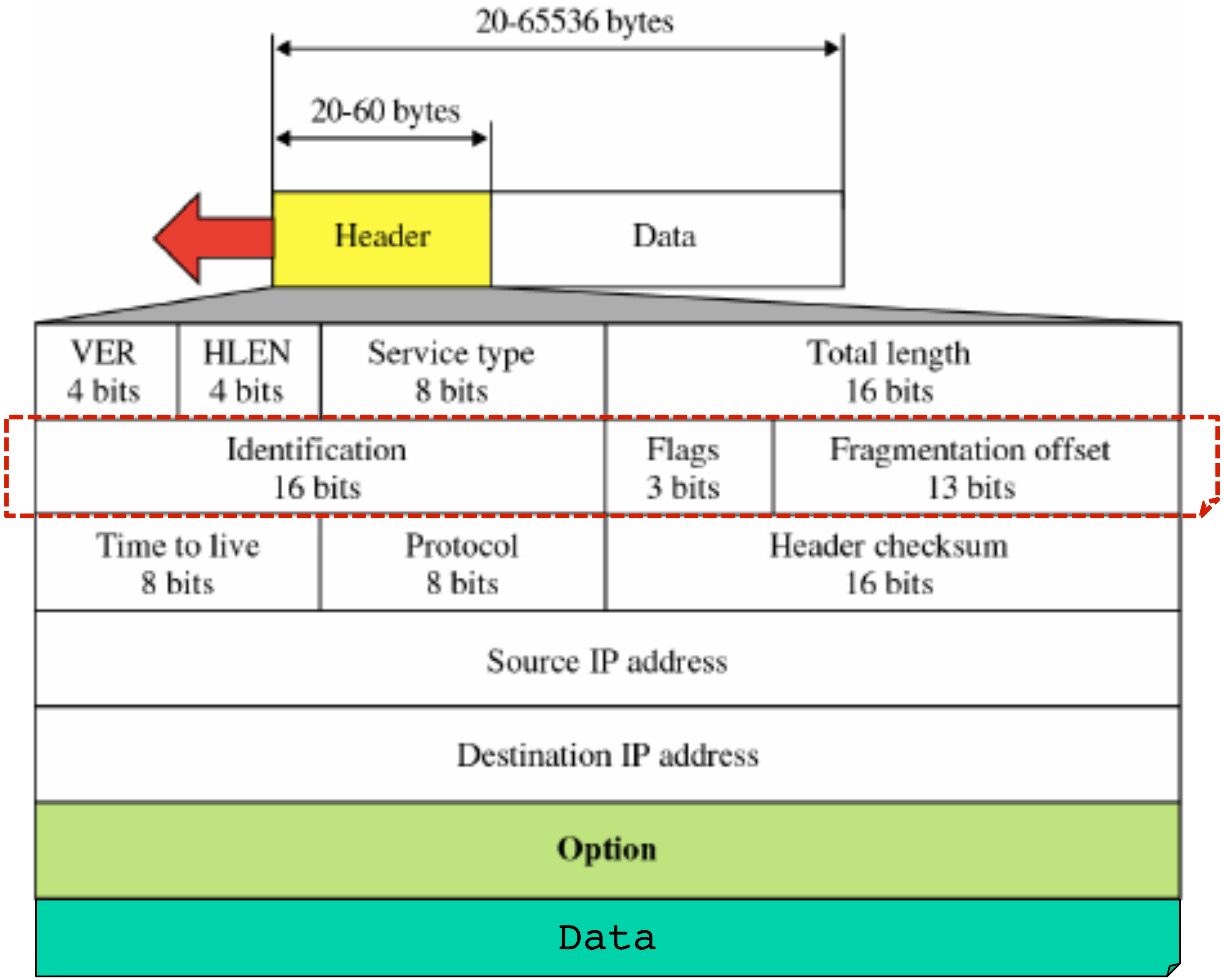
- Source and destination IP address – 32-bit field, must remain unchanged until IP datagram reaches its final destination
- Options – 32-bit fields, not required for every datagram, allows expansion of IP header for special purposes
  - Seldom used
  - Options were dropped in IPv6 header
- Data (payload) – it usually contains the transport layer segment (TCP or UDP) to be delivered to the destination. It can carry other types of data, such as ICMP (Internet Control Message Protocol) messages.

IP fragmentation, reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



IP fragmentation, reassembly (cont.)



IP fragmentation, reassembly (cont.)

- Identification – 16-bit field, uniquely identifies datagram originating from source host
  - To guarantee uniqueness, IP uses counter to label each datagram
  - When IP sends a datagram, it copies current counter value to identification field, and increases counter by one
  - When datagram is fragmented, identification field is copied into all fragments
  - Identification number helps destination in reassembling datagram
- Flags – 3-bit field
  - 1^{st} bit is reserved
  - 2^{nd} bit: “do not fragment” bit, 1= no fragment
  - 3^{rd} bit: “more fragment” bit, 1=not last fragment, 0=last one

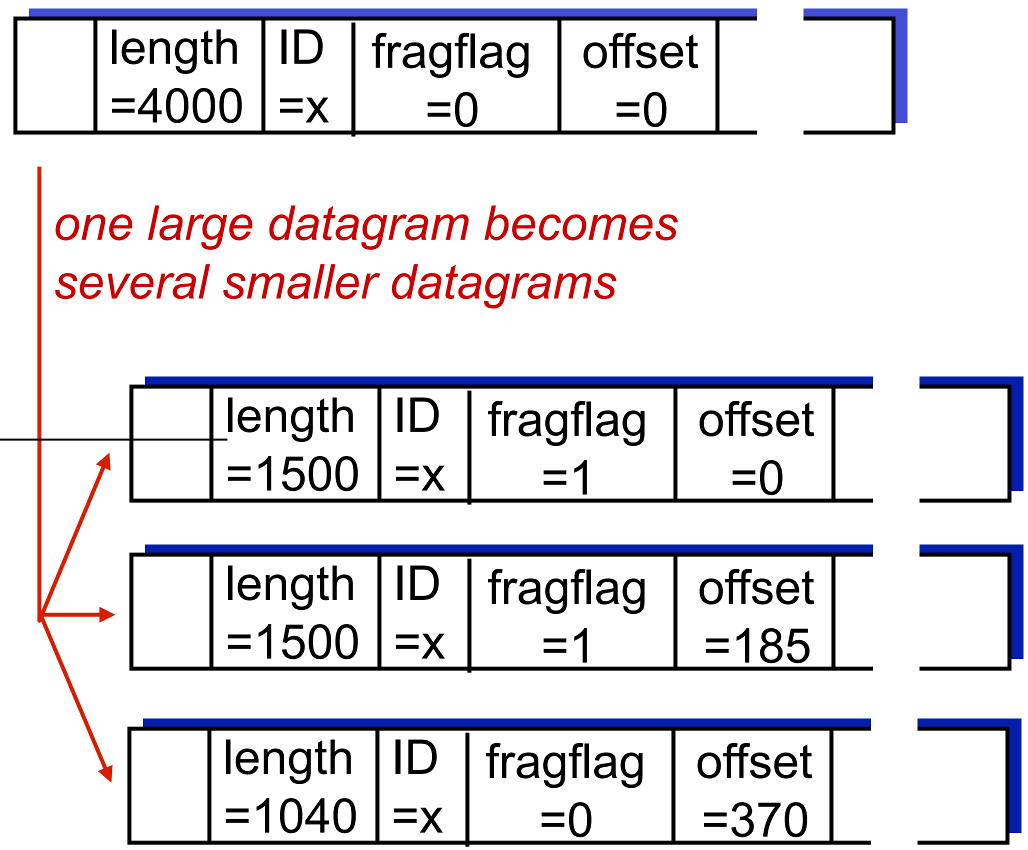
D: Do not fragment
M: More fragments

IP fragmentation, reassembly (cont.)

- Fragmentation offset – 13-bit field, shows relative position of fragment data with respect to whole datagram
  - The offset is measured in units of 8 bytes

example:

- 4000 byte datagram
- MTU = 1500 bytes



1480 bytes in data field

offset = 1480/8

offset = 2960/8

Chapter 4: outline

4.1 introduction

4.2 virtual circuit and datagram networks

4.3 what's inside a router

4.4 IP: Internet Protocol

- datagram format
- IPv4 addressing
- ICMP
- IPv6

4.5 routing algorithms

- link state
- distance vector
- hierarchical routing

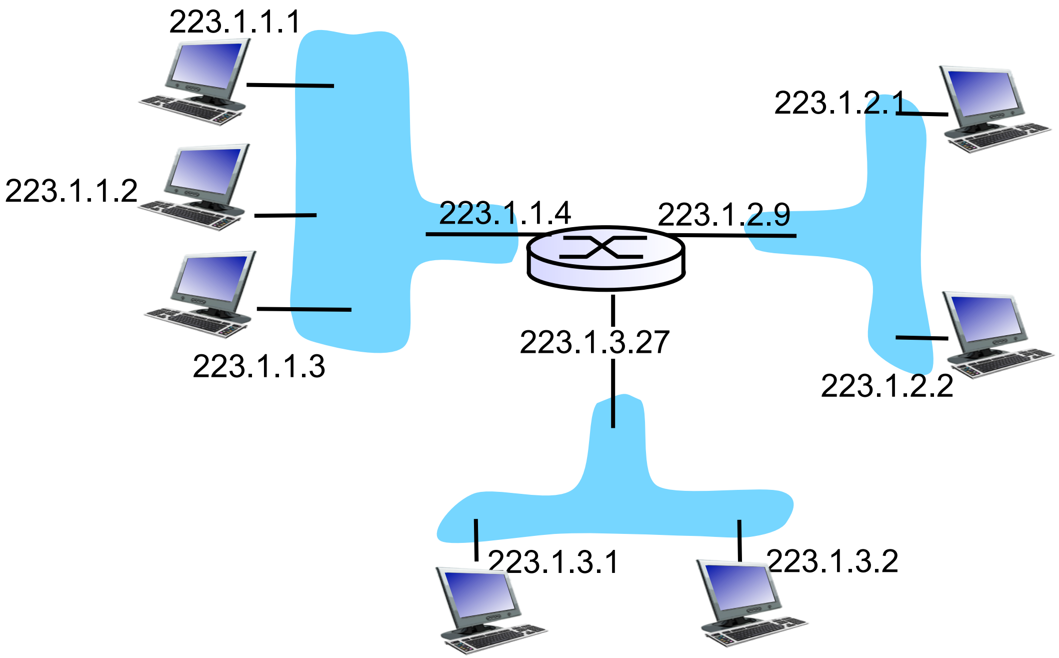
4.6 routing in the Internet

- RIP
- OSPF
- BGP

4.7 broadcast and multicast routing

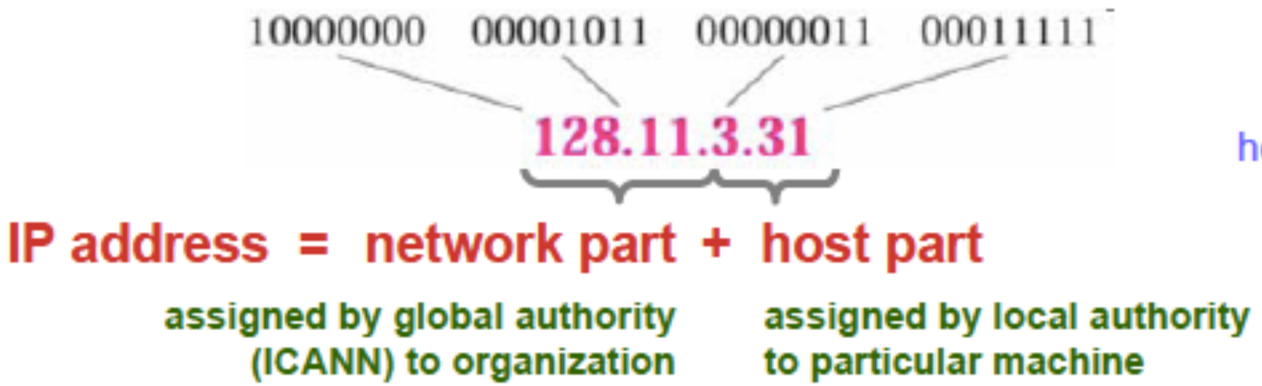
IP addressing

- *IP address:* uniquely and universally identifies each device connect to the network
  - IP Address: 3-bit (4-byte) binary address that identifies a host/router interface to the Internet
  - Two devices on the Internet can never have the same address at the same time; But, a single device can have two IP addresses if it is connected to the Internet via two networks
  - Routers typically have multiple interfaces, e.g. multiple IP addresses



IP addressing (cont.)

- *IP address: Binaay Notation* 32-bit/4-byte representation with a space inserted between each octet (byte). There are about 4 billions possible IP addresses.
- *IP address: Decimal Notation:* 4-number decimal representation with a decimal dot separating the numbers
  - Each decimal number, [0,255], corresponding to a byte



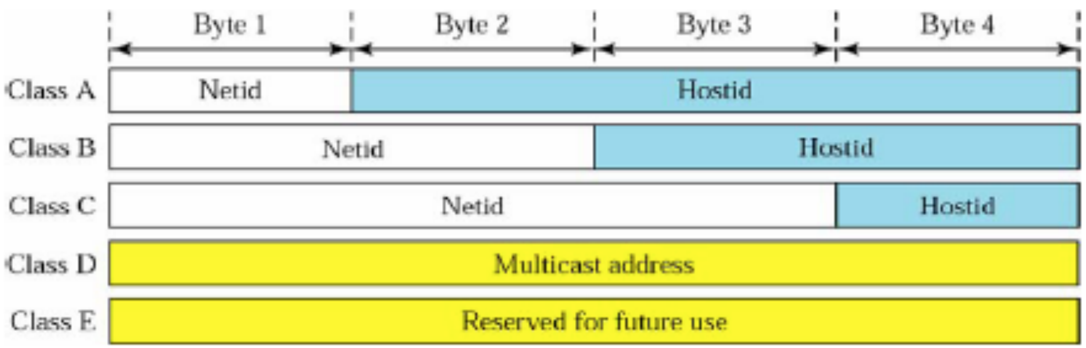
How many bits go to network and how many to host part!?

Classful and Classless IP addressing

- Originally, IP addressing used the concept of classes. This architecture is called classful addressing.
- In the mid 1990s, a new architecture – classless addressing, was introduced.
- Classless Addressing known as CIDR "Classless InterDomain Routing" addressing – removes class privileges to compensate for address depletion
- CIDR is used for Internet address assignment

Classful IP addressing

- Supports addressing of different size networks by dividing address space into 5 classes: A, B, C, D, E
  - *An IP address in classes A, B, and C is divided into Netid and Hostid*



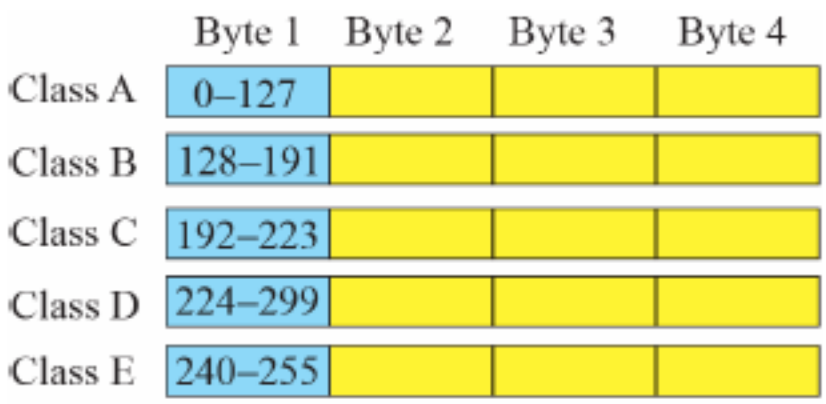
While many class A and B addresses are wasted, the number of addresses in class C is smaller than the needs of most organizations.

How do we know if an IP address is a class-A / B or C!?

Classful IP addressing (cont.)

- Recognizing classes
  - Binary Notation – first few bits of an IP address in binary notation immediately identify the class of the given address
  - Decimal Notation – each class has a specific range of numbers in decimal notation – it is enough to look at the first number to determine the class

Binary notation



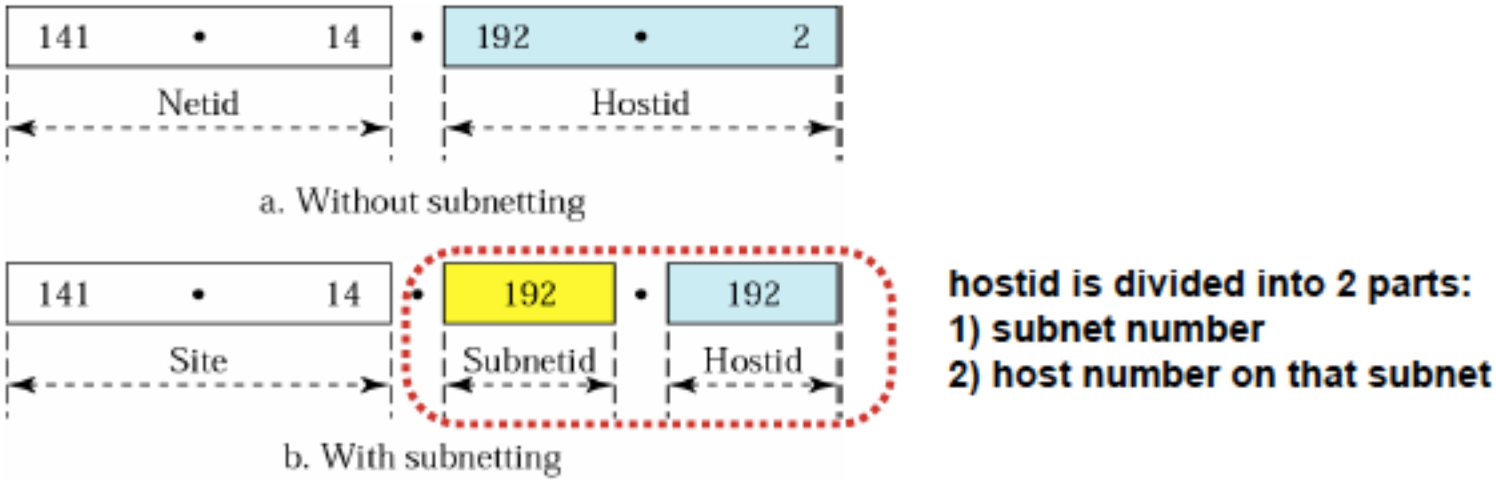
Dotted-decimal notation

Classful IP addressing (cont.)

- Disadvantages of classful network addressing
  - Lack of a class to support medium-sized organizations
    - Class C which supports 254 hosts – too small
    - Class B which supports 65534 hosts – too large
  - A premature depletion of class B addresses has already occurred
    - In the early days of the Internet, addresses were freely assigned to those who asked for them without concerns about the eventual depletion of the IP address space
- Two existing mechanisms for overcoming the limitations of classful addressing:
  - Subnetting - if an organization gets assigned a "big" block of IP addresses how to distribute them among multiple LAN
  - Supernetting – how an organization can combine several class C blocks to create a larger range of address

Subnets

- Network divided into several smaller subnetworks each having its own subnetwork address
  - Internally, each subnetwork is recognized by its subnetwork address; to the rest of the Internet all subnetoworks still appear as a single network
- Organization of address space in a subnetted network
  - A number of HostID bits are borrowed for subnet identification
  - With *m* borrowed bits, 2^m subnets can be created
  - Number of hosts in each subnet: $2^{\text{Hostid-}m}$



Classless addressing: CIDR

CIDR: Classless InterDomain Routing

- subnet portion of address of arbitrary length
- address format: a.b.c.d/x, where x is # bits in subnet portion of address

subnet part ←→ host part

11001000 00010111 00010000 00000000

200.23.16.0/23