Network Layer: Control/data plane, addressing, routers

CS 352, Lecture 10

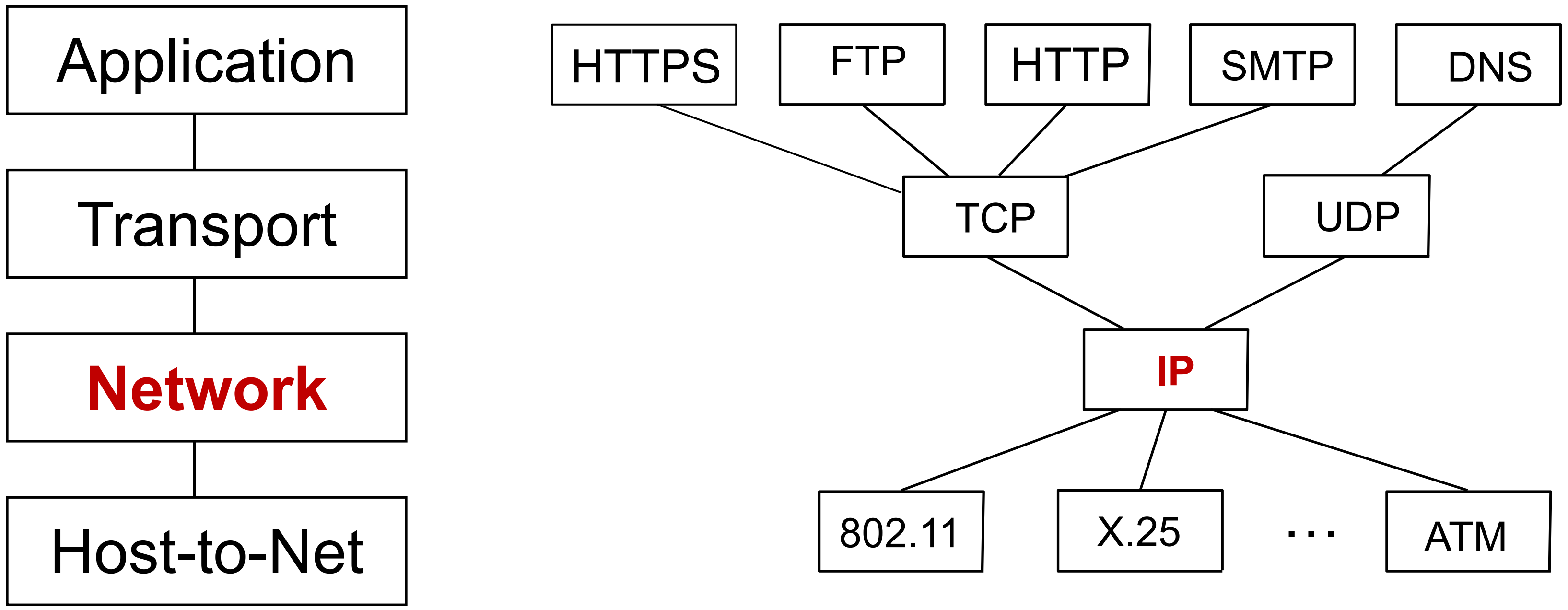
http://www.cs.rutgers.edu/~sn624/352-S19

Srinivas Narayana

(heavily adapted from slides by Prof. Badri Nath and the textbook authors)

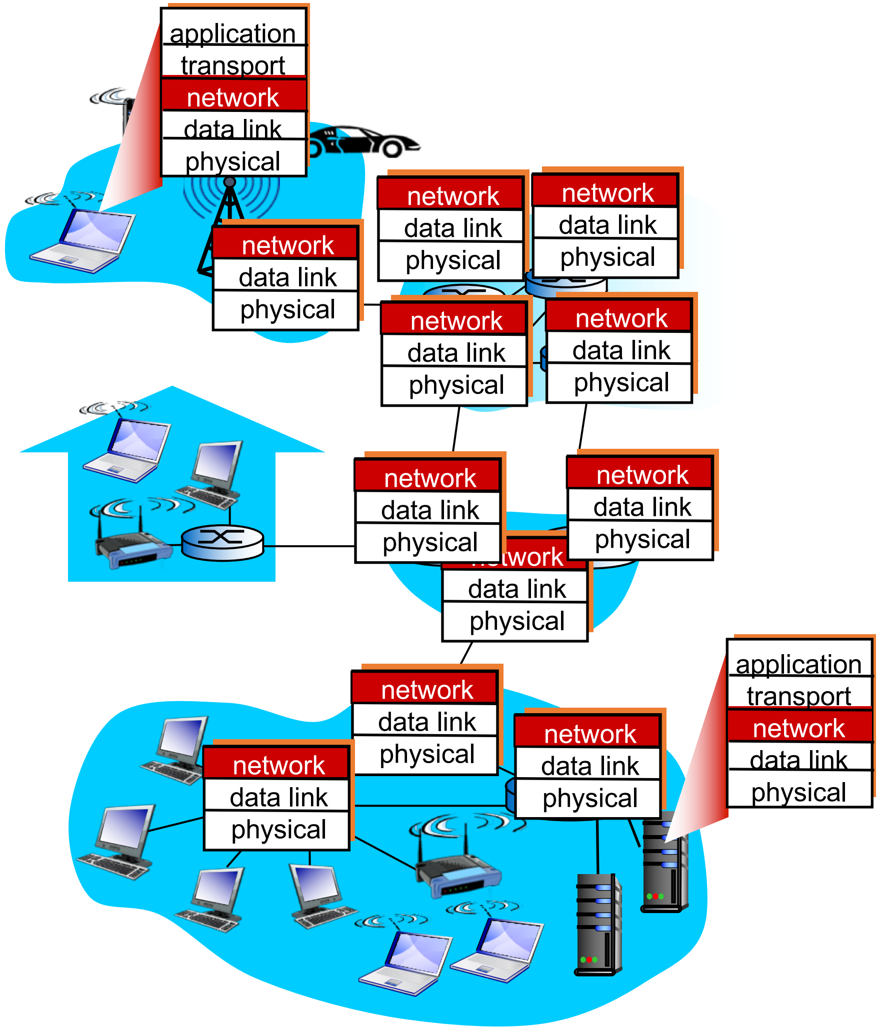
RUTGERS
UNIVERSITY | NEW BRUNSWICK

Next: Network layer



Network layer

- transport segment from sending to receiving host
- on sending side encapsulates segments into datagrams
- on receiving side, delivers segments to transport layer
- network layer protocols in *every* host, router
- router examines header fields in all IP datagrams passing through it



Two key network-layer functions

network-layer functions:

- *forwarding:* move packets from router's input to appropriate router output
- *routing:* determine route taken by packets from source to destination
  - *routing algorithms*

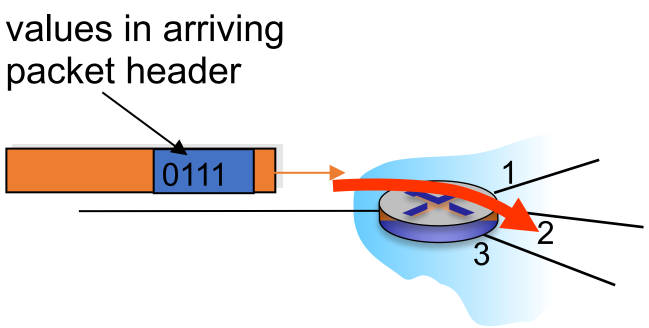
analogy: taking a trip

- *forwarding:* process of getting through single interchange
- *routing:* process of planning trip from source to destination

Network layer: data plane, control plane

Data plane

- local, per-router function
- determines how datagram arriving on router input port is forwarded to router output port
- forwarding function



Control plane

- network-wide logic
- determines how datagram is routed among routers along end-end path from source host to destination host
- two control-plane approaches:
  - *traditional routing algorithms:* implemented in routers
  - *software-defined networking (SDN)*: implemented in (remote) servers

IP Addressing

IPv4 Addresses

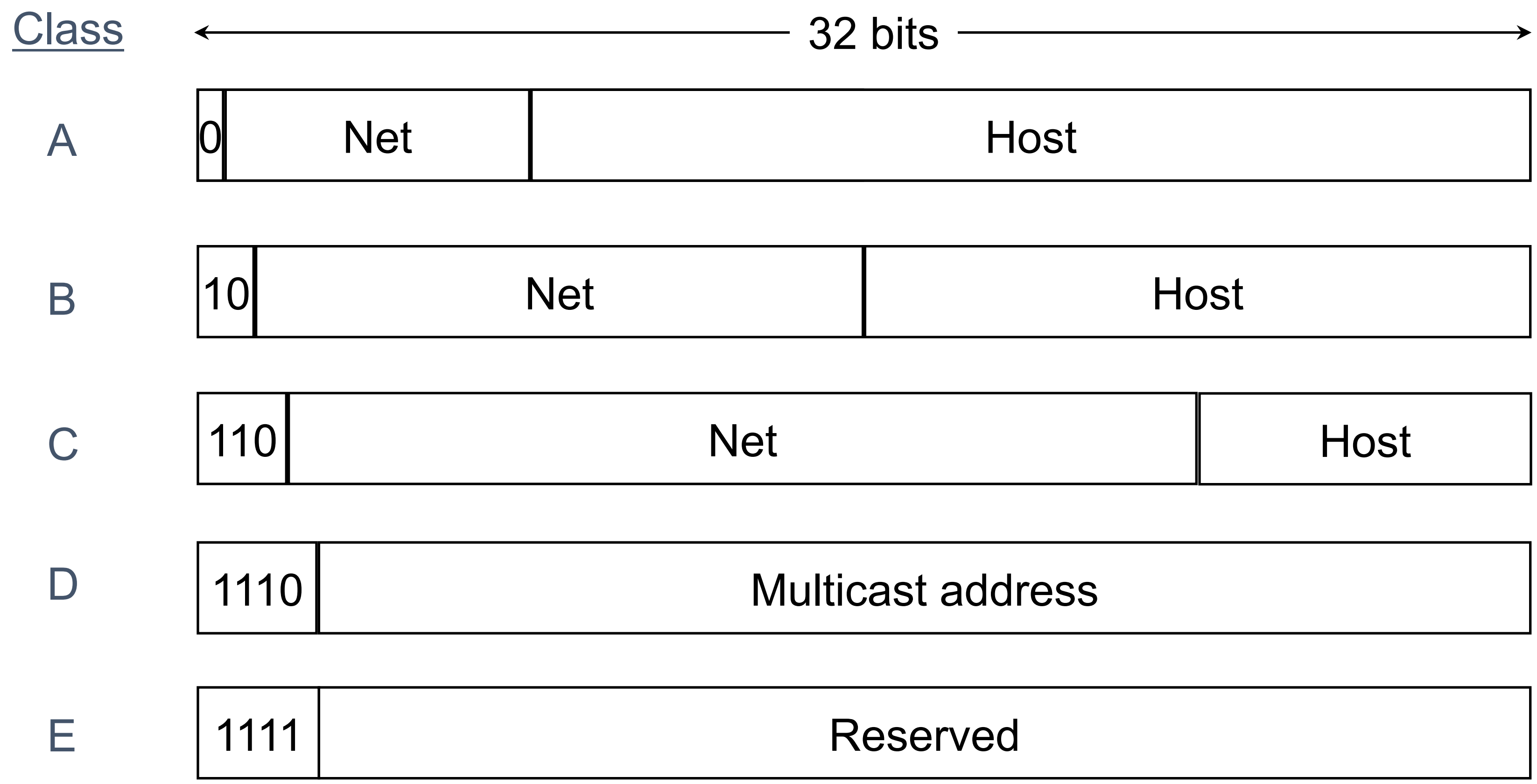
- 32 bits long
- Identifier for host, router *interface*
- Notation:
  - Each byte is written in decimal in MSB order, separated by dots
  - Example: 128.195.1.80 stands for the 32-bit IP address

    10000000 11000011 00000001 01010000

Types of IPv4 Addresses

- Unicast Address
  - Destination is a single host
- Multicast address
  - Destination is a group of hosts
- Broadcast address
  - 255.255.255.255
  - Destination is all hosts

IPv4 Address Classes (old)



IP Address Classes

- Class A:
  - For very large organizations
  - 16 million hosts allowed
- Class B:
  - For large organizations
  - 65 thousand hosts allowed
- Class C
  - For small organizations
  - 255 hosts allowed
- Class D
  - Multicast addresses
  - No network/host hierarchy

Problems with Class-based Routing

- Too many small networks requiring multiple class C addresses
- Running out of class B addresses, not enough nets in class A
- Addressing strategy must allow for greater diversity of network sizes

IP 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

CIDR

- An ISP can obtain a block of addresses and partition this further to its customers
  - Say an ISP has 200.8.0.0/16 address (65K addresses).
  - He has another customer who needs only 64 addresses starting from 200.8.4.128
  - Then that block can be specified as 200.8.4.128/26

Subnetting

Example: Class B address with 8-bit subnetting

	16 bits	8 bits	8 bits
	Network id	Subnet id	Host id
Example Address:	165.230	.24	.8

Subnet Masks

Subnet masks allow hosts to determine if another IP address is on the same subnet or the same network

	16 bits	8 bits	8 bits
	Network id	Subnet id	Host id
Mask:	1111111111111111	11111111	00000000
	255.255	.255	.0

Subnet Masks *(cont'd)*

Assume IP addresses A and B share subnet mask M.

Are IP addresses A and B on the same subnet?

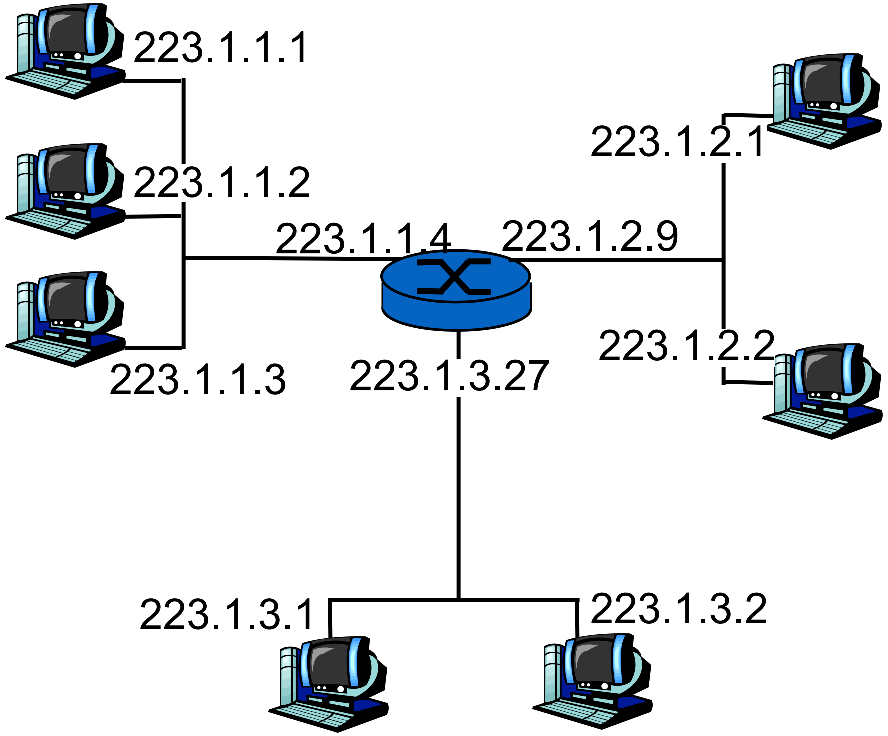
1. Compute logical AND (A & M).
2. Compute logical AND (B & M).
3. If (A & M) == (B & M) then A and B are on the same subnet.

Example: A and B are class B addresses

A = 165.230.82.52
B = 165.230.24.93
M = 255.255.255.0

Same (classful) network?
Same subnet?

Example of IP Addressing in a network

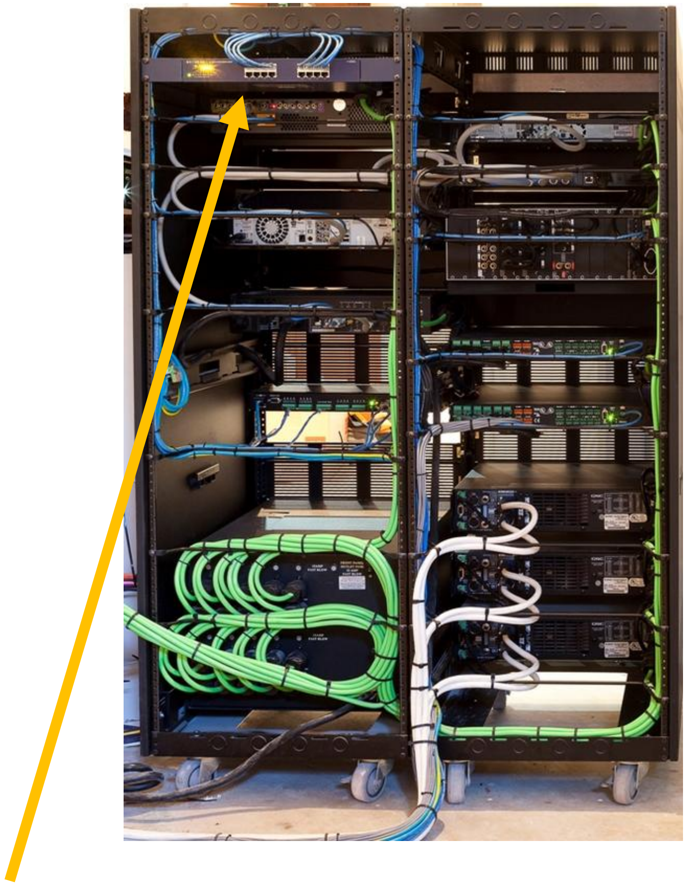


What’s inside a router?

What do routers look like?

Access routers

Core router

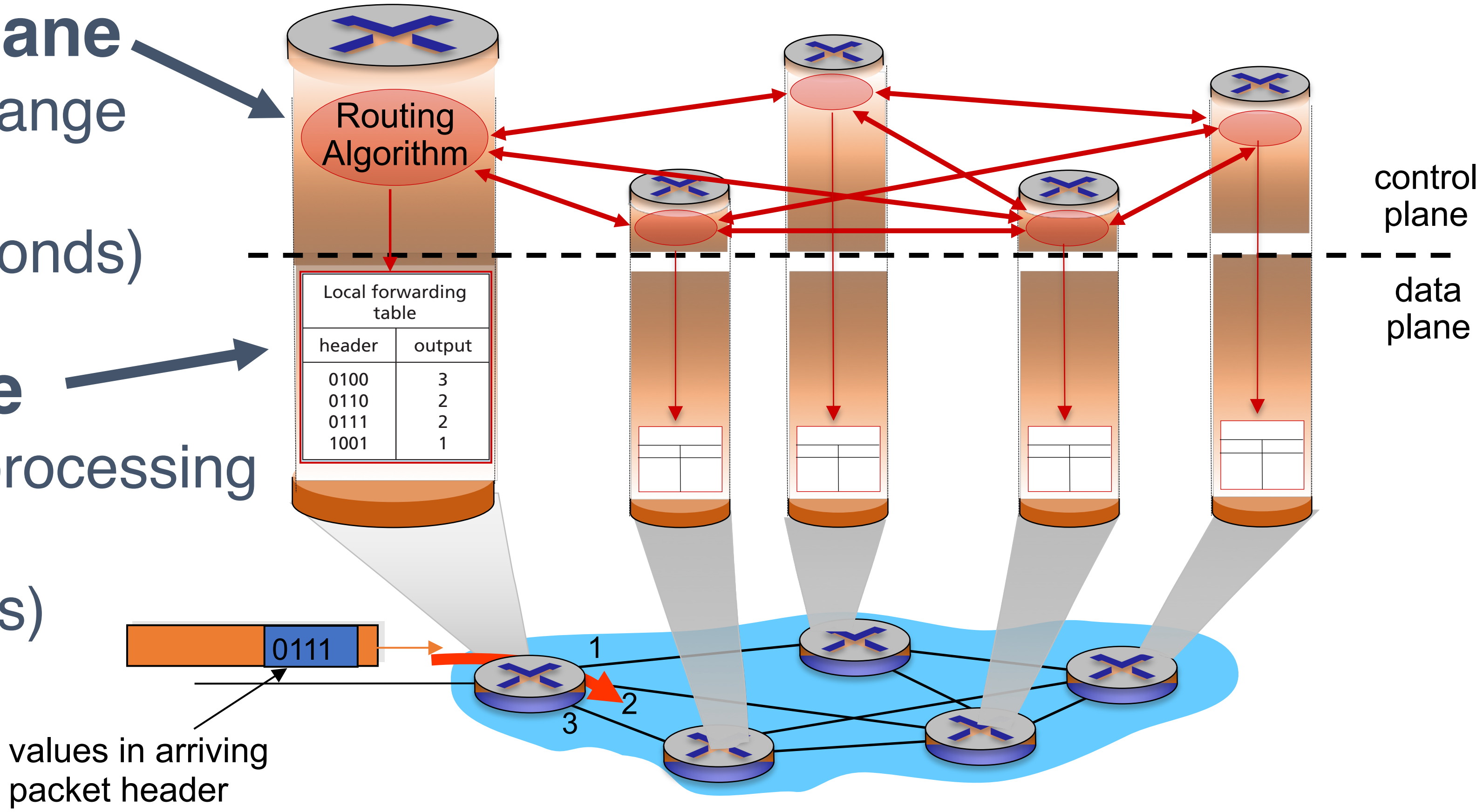


Data center top-of-rack switch

Basic components: Control & Data Planes

Control plane
per route-change processing
(~ a few seconds)

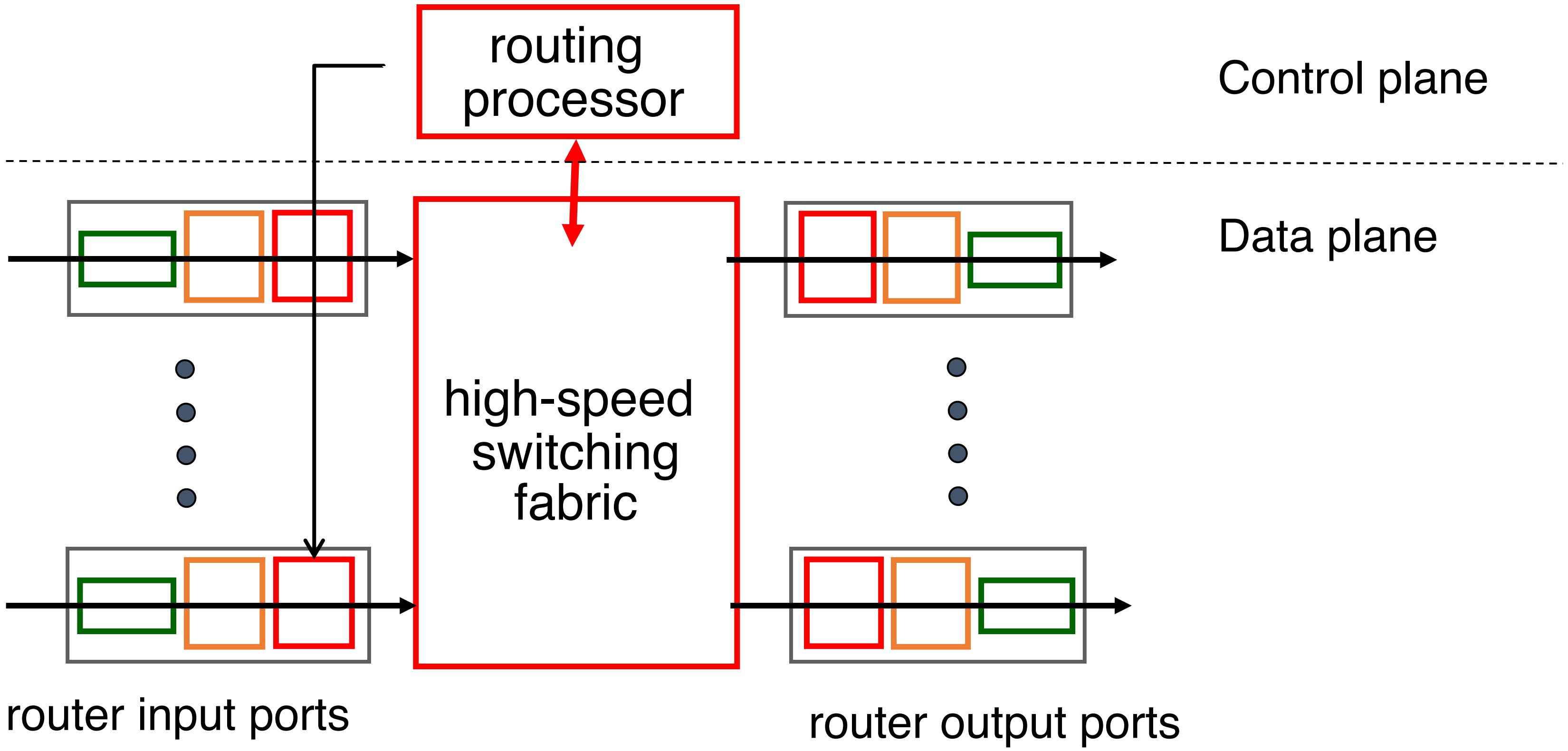
Data plane
per-packet processing
(~ tens of nanoseconds)



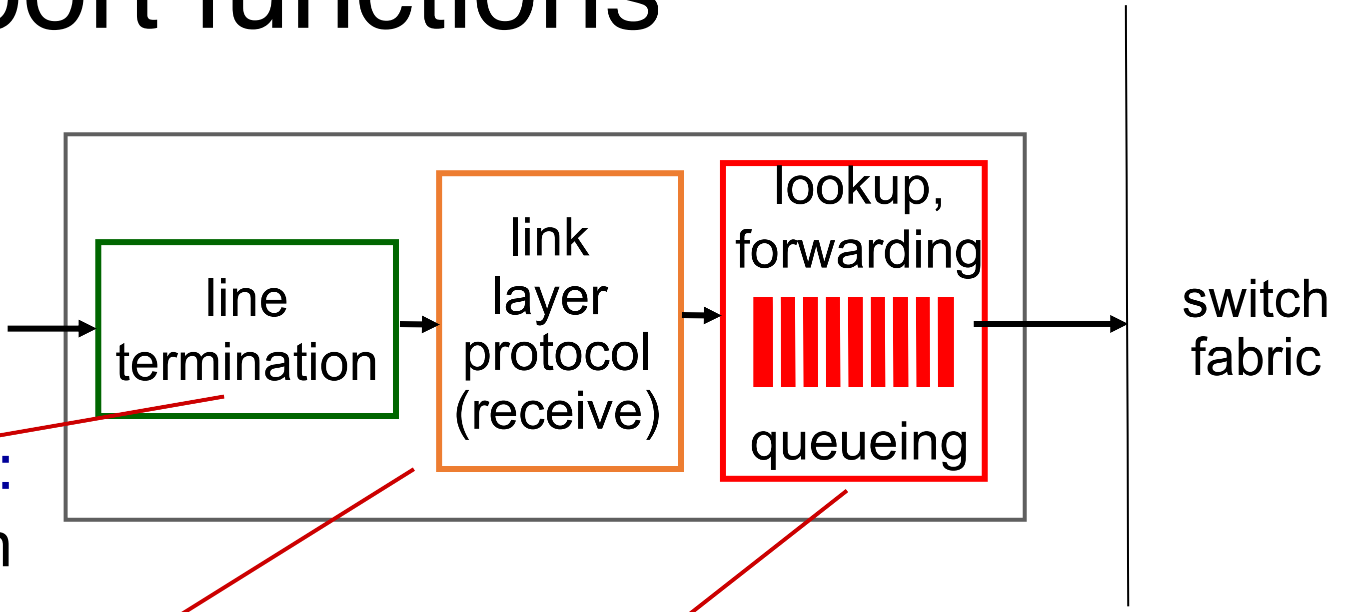
Traditionally:

Individual routing algorithm components *in each and every router* interact in the control plane

Router architecture overview



Input port functions



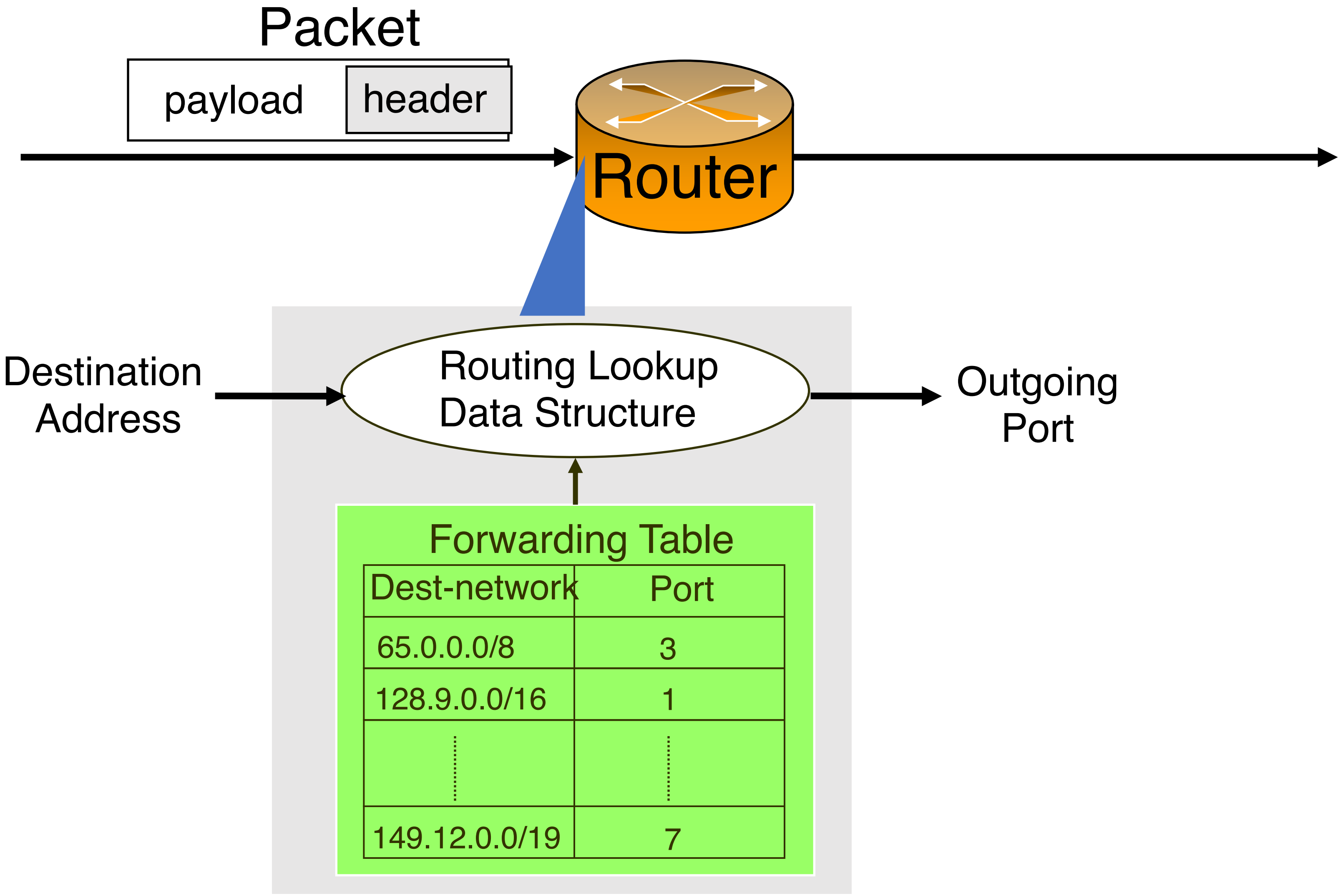
physical layer:
bit-level reception

data link layer:
e.g., Ethernet
(We'll see this later in the course)

Switching:

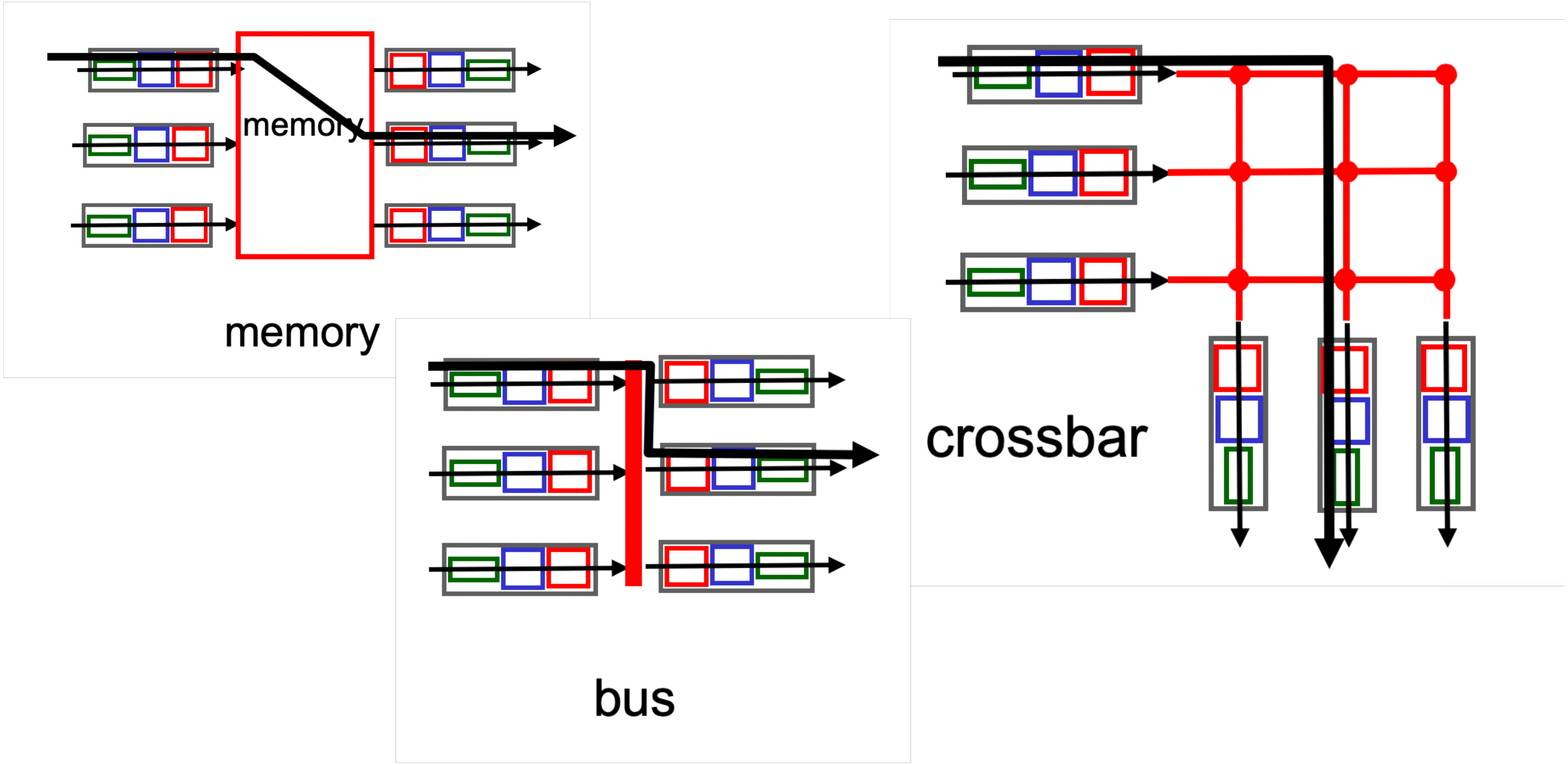
- using header field values, lookup output port using forwarding table in input port memory *("match plus action")*
- goal: complete input port processing at 'line speed'
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

Destination-based Forwarding in the Internet

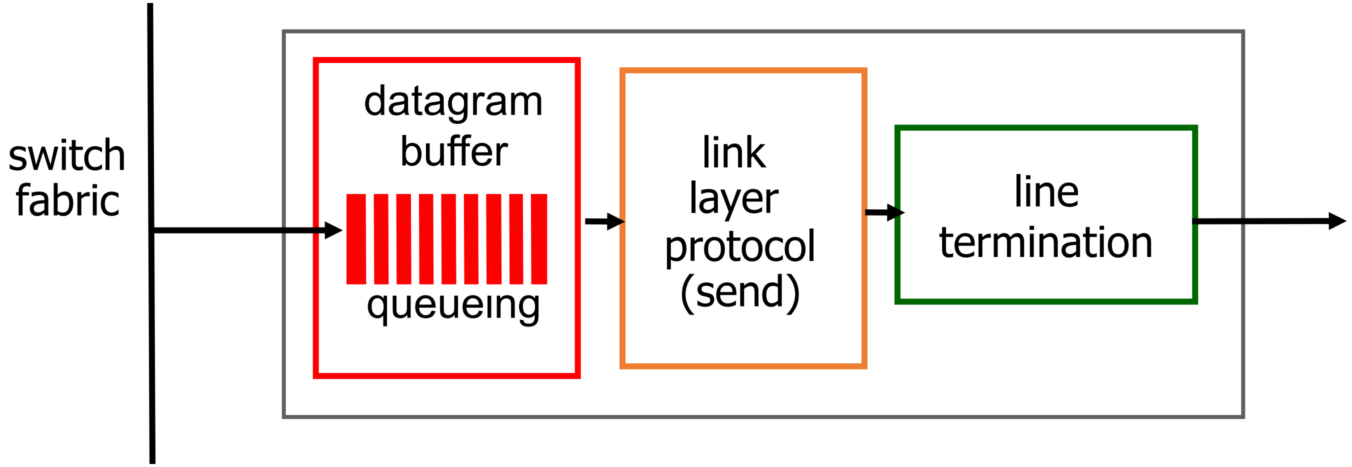


Dest-network	Port
65.0.0.0/8	3
128.9.0.0/16	1
⋮	⋮
149.12.0.0/19	7

Three types of switching fabrics



Output Ports



- *Buffering* required when datagrams arrive from fabric faster than the transmission rate
  - Important implication: if buffers filled up, packets are dropped!

- *Scheduling discipline* chooses among queued datagrams for transmission
  - Important implication: Who gets priority is chosen by the scheduler

Prefixes and IP lookup

Example Forwarding Table

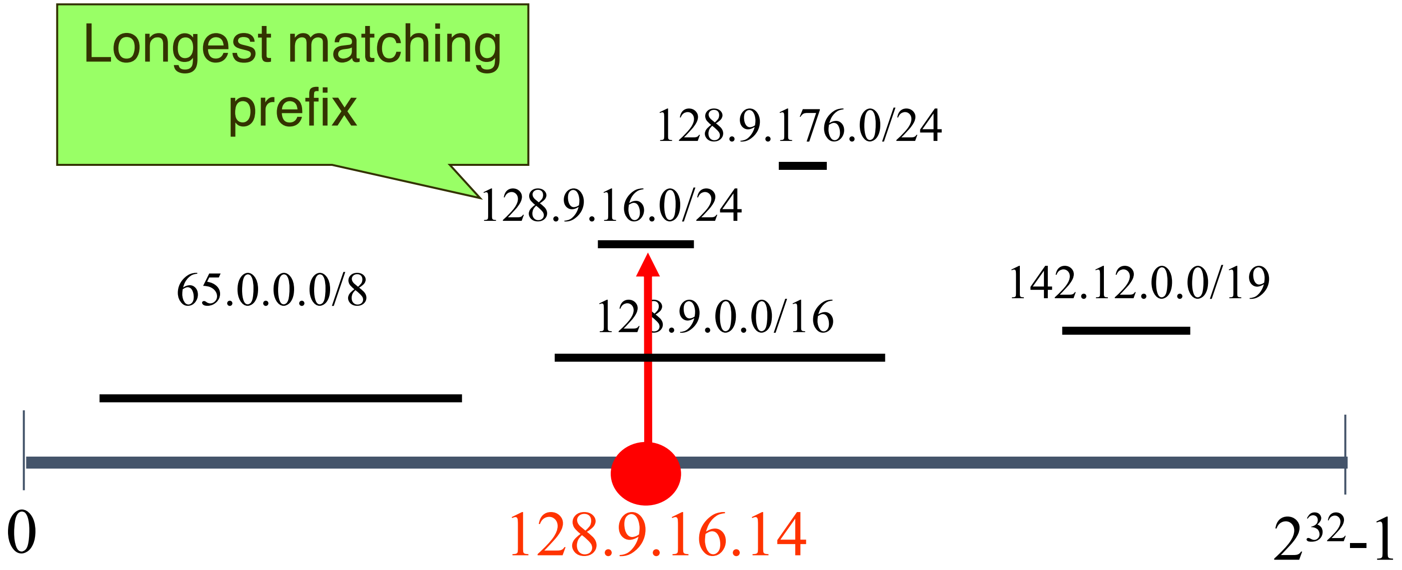
Destination IP Prefix	Outgoing Port
65.0.0.0/8	3
128.9.0.0/16	1
65.0.0.128/25	4
142.12.0.0/19	7

Prefix length

IP prefix: 0-32 bits

Longest prefix match

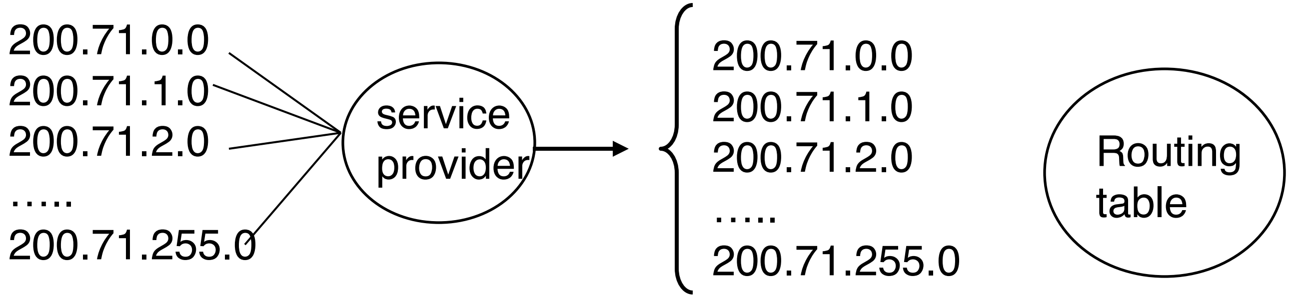
Prefixes can Overlap



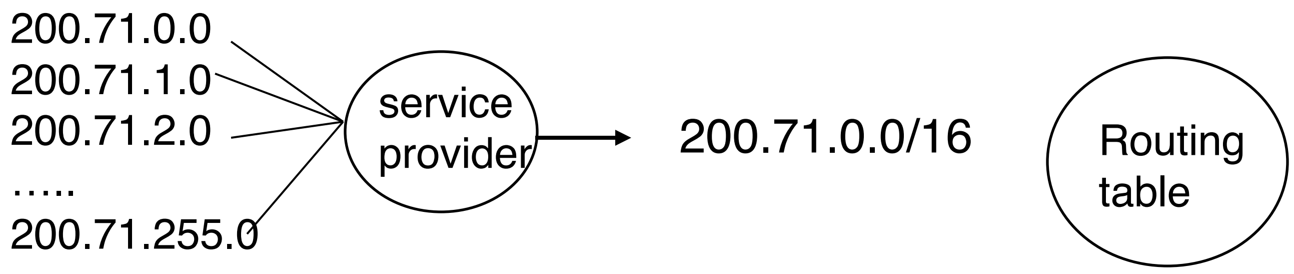
Routing lookup: Find the longest matching prefix (the most specific route) among all prefixes that match the destination address.

Reducing Routing Table Size

Without CIDR:

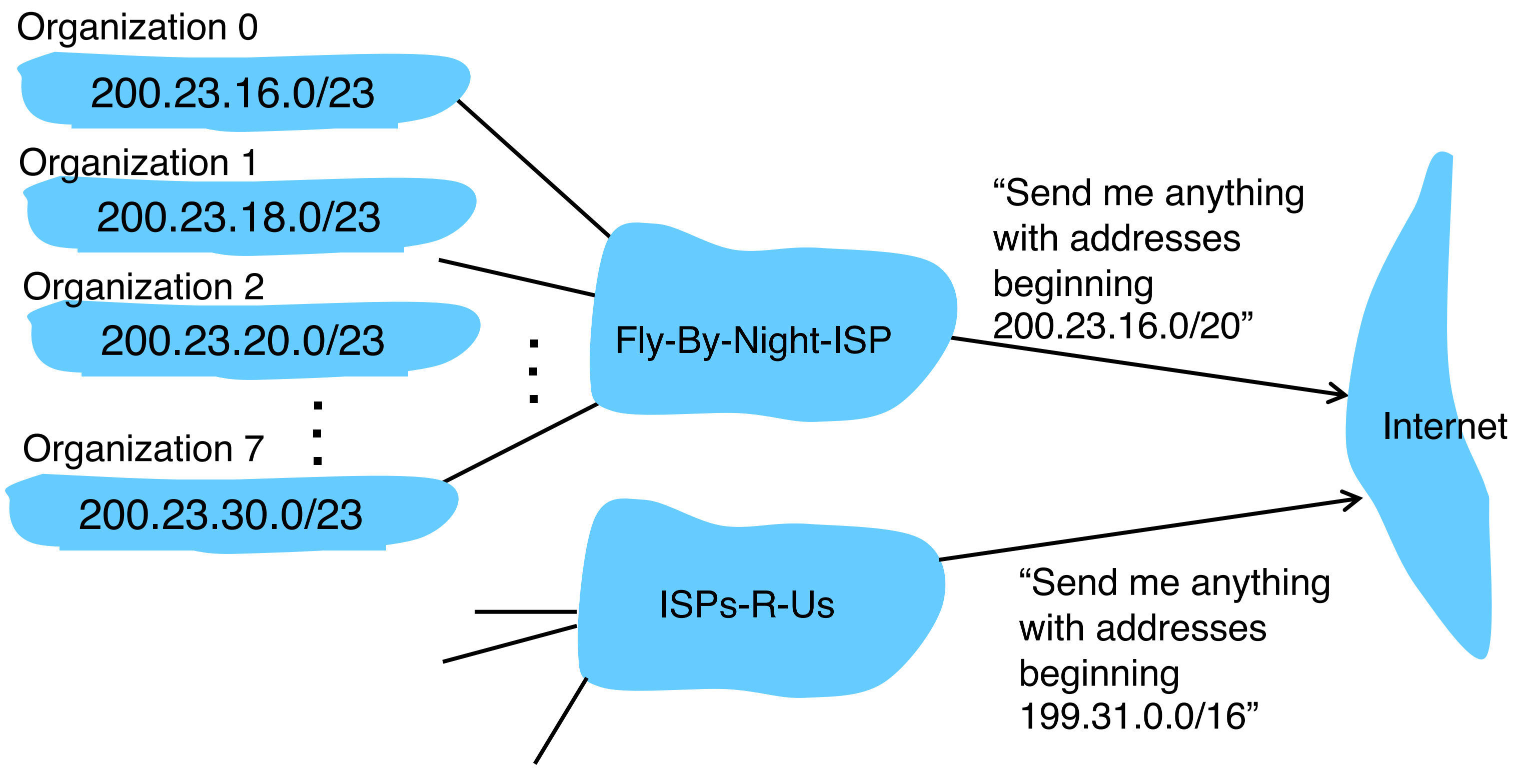


With CIDR:



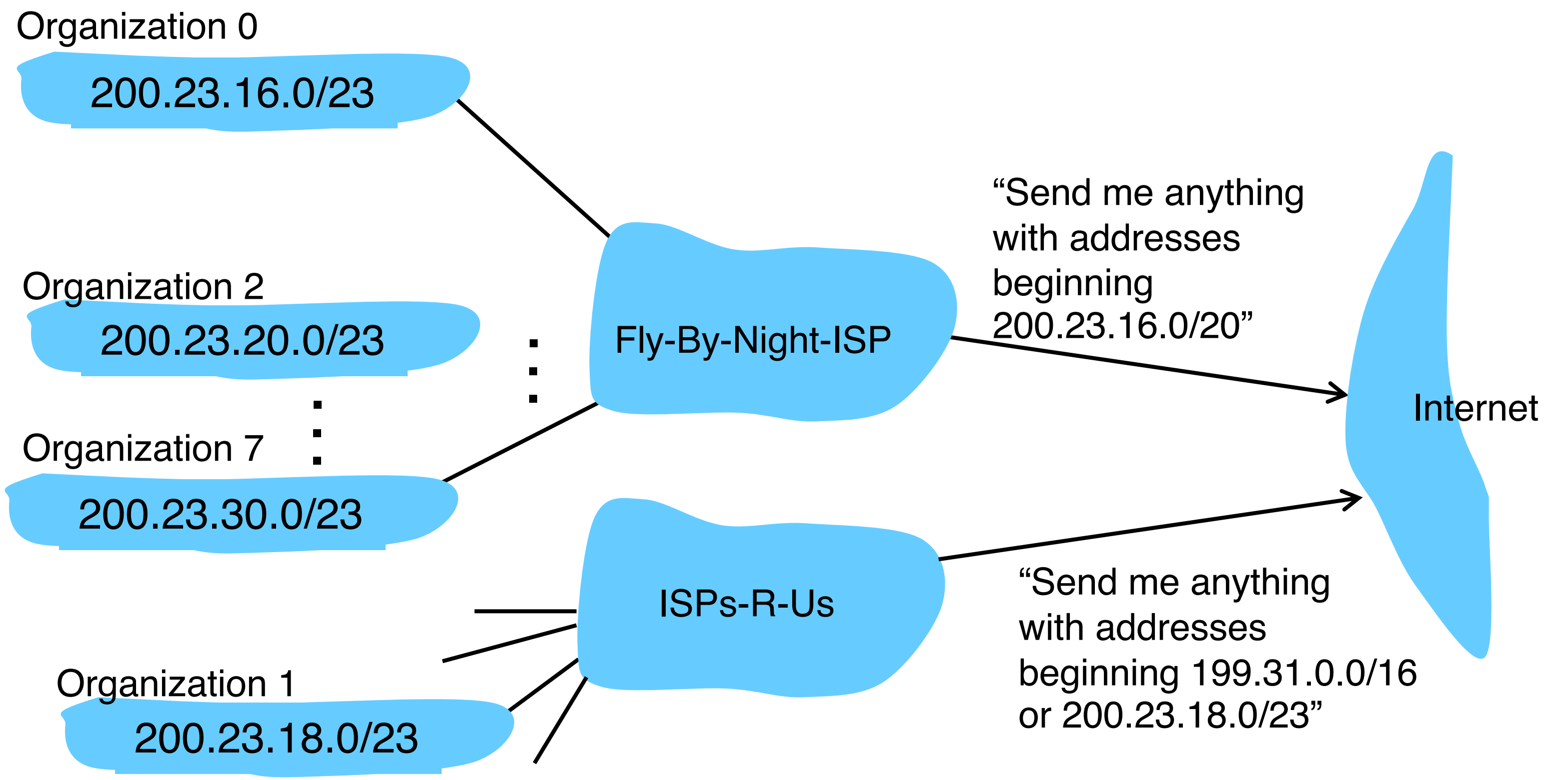
Hierarchical addressing: Route aggregation

Efficient advertisement of routing information!



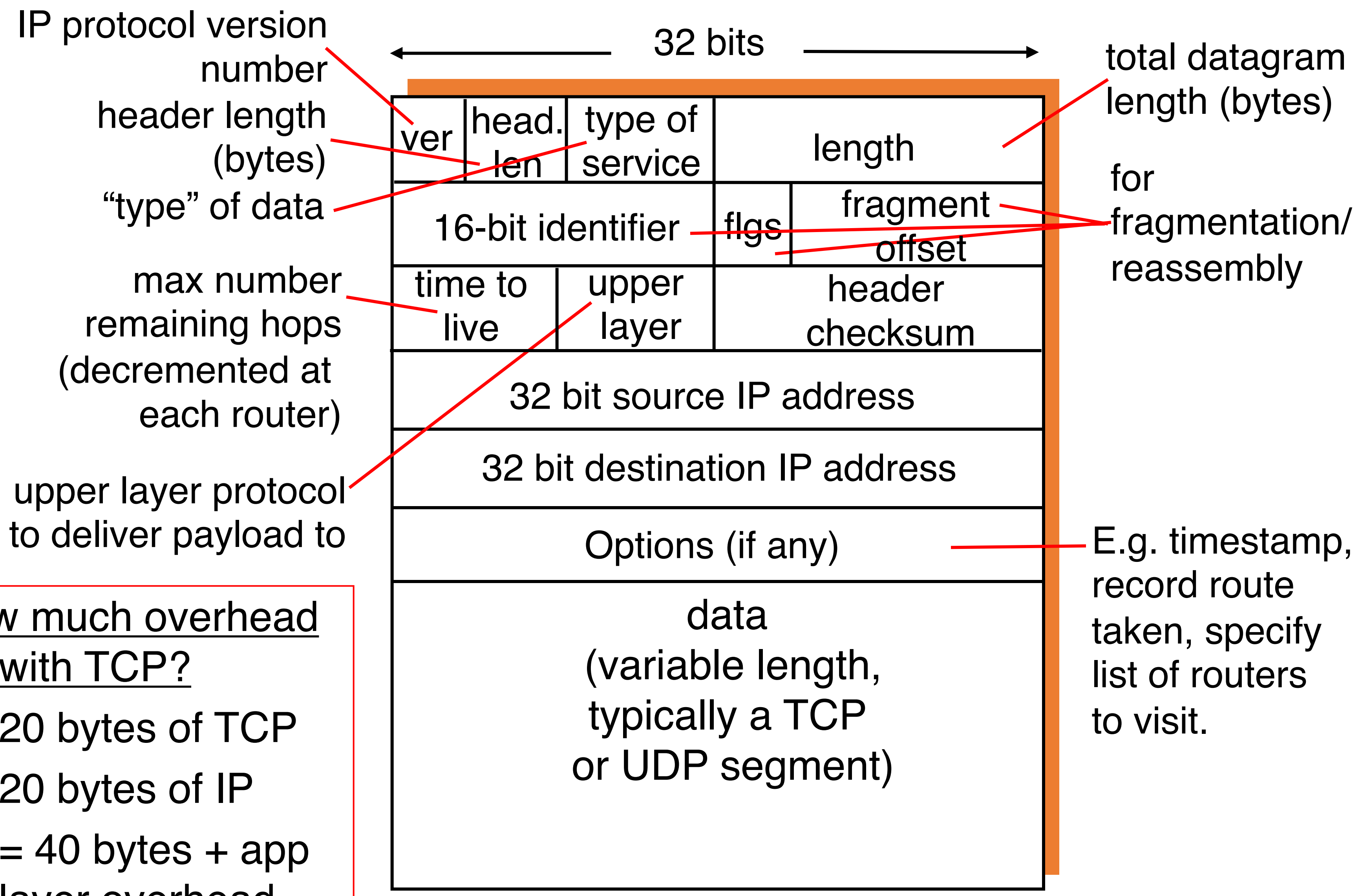
LPM: Announcing more specific routes

ISPs-R-Us has a more specific route to Organization 1
Longest prefix match will be used to route IP packets



The Internet Protocol (IP)

IP datagram format

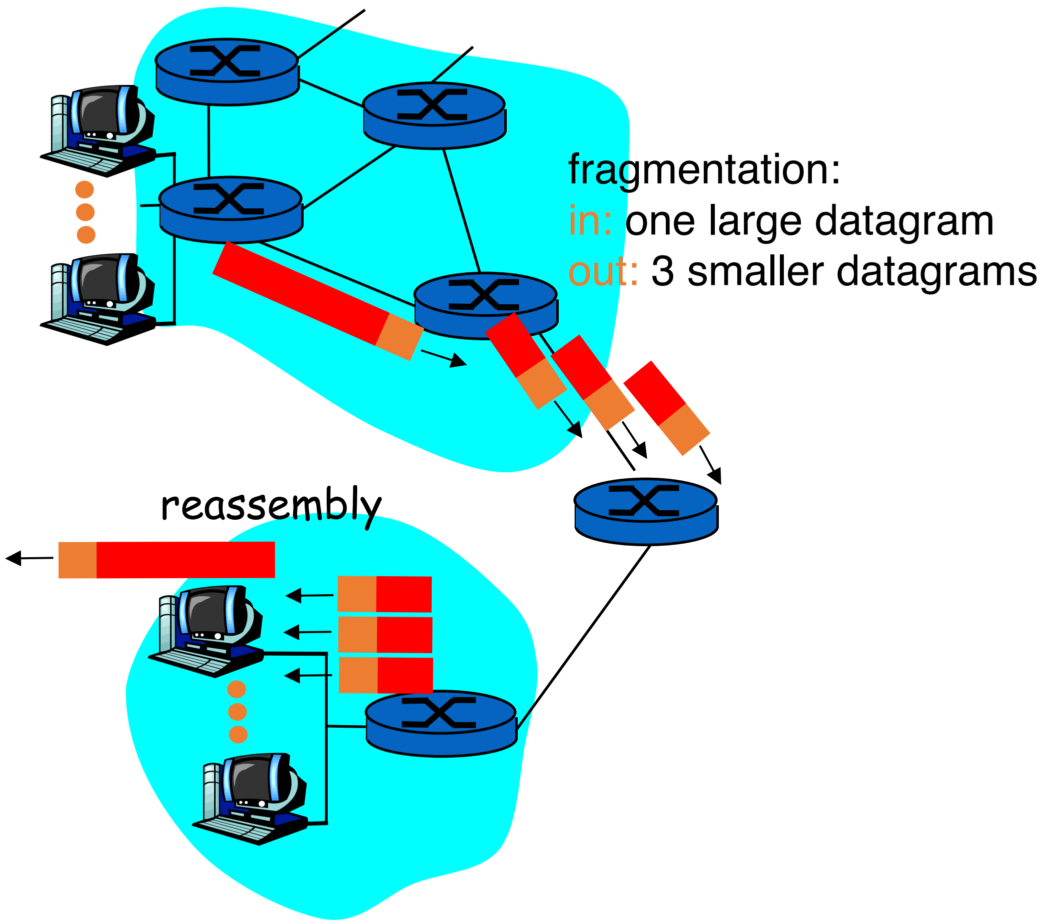


how much overhead with TCP?

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

IP Fragmentation & Reassembly

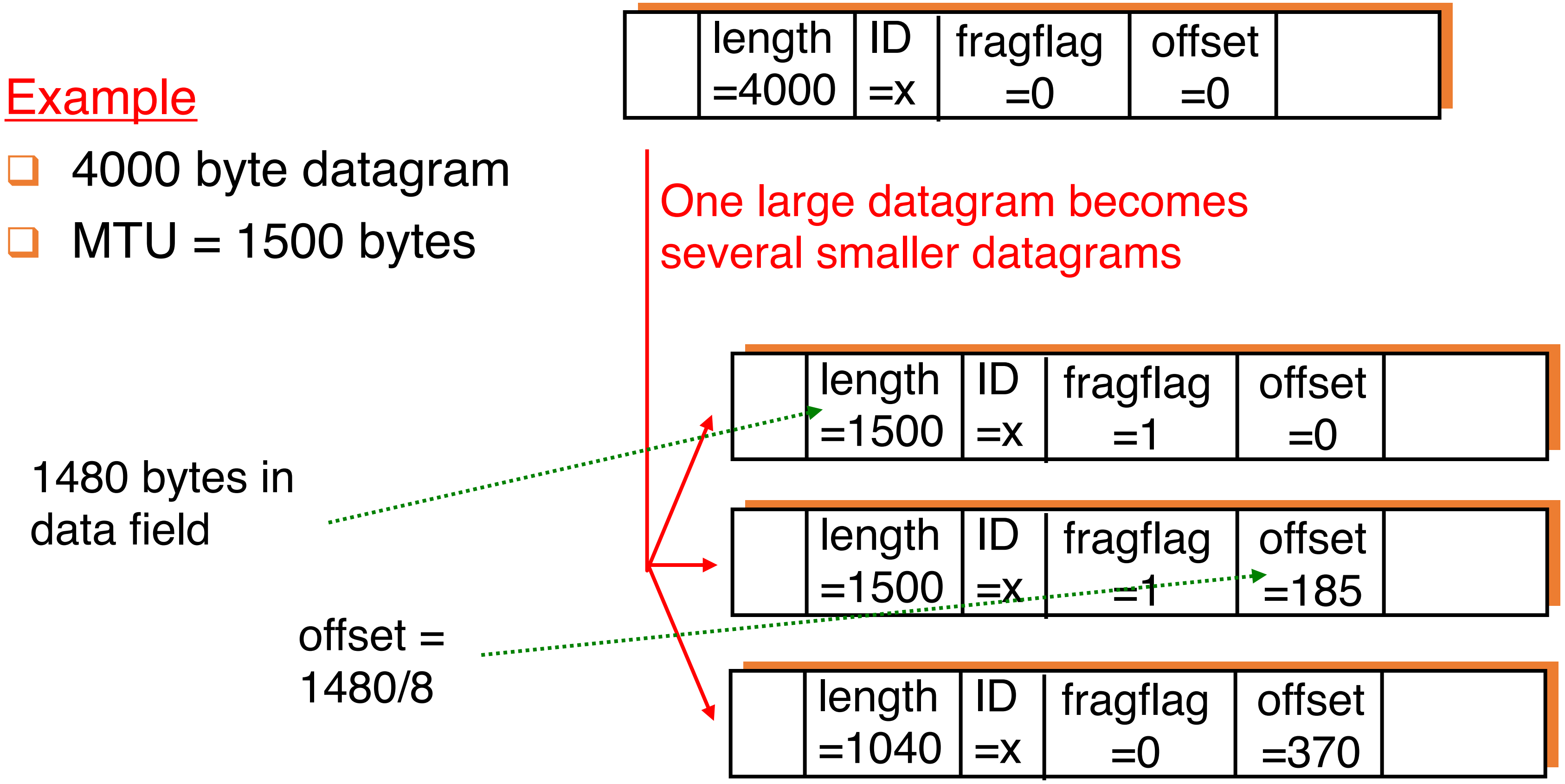
- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify,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



IP Fragmentation and Reassembly

Example

- 4000 byte datagram
- MTU = 1500 bytes

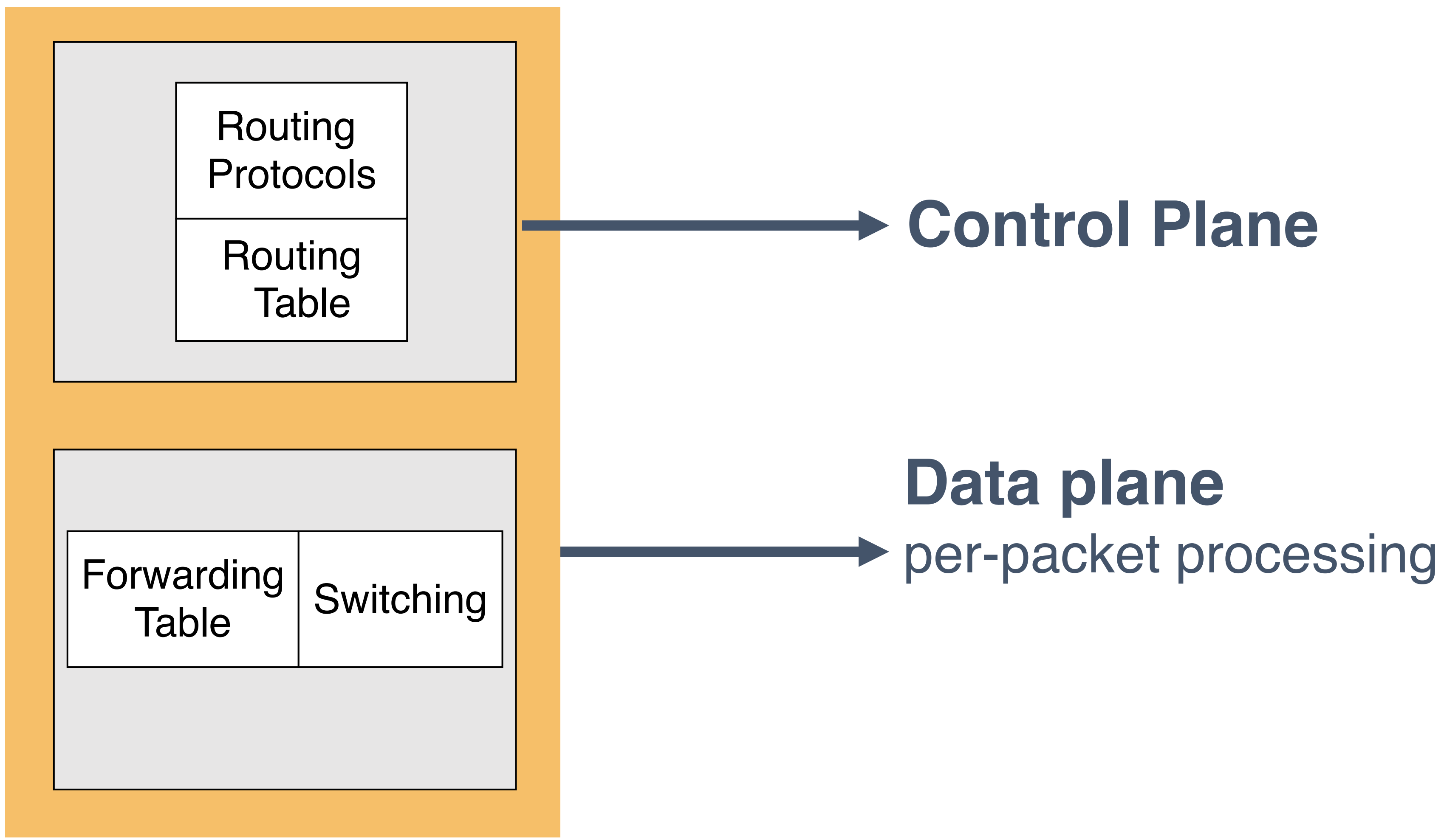


IP Address Hierarchy

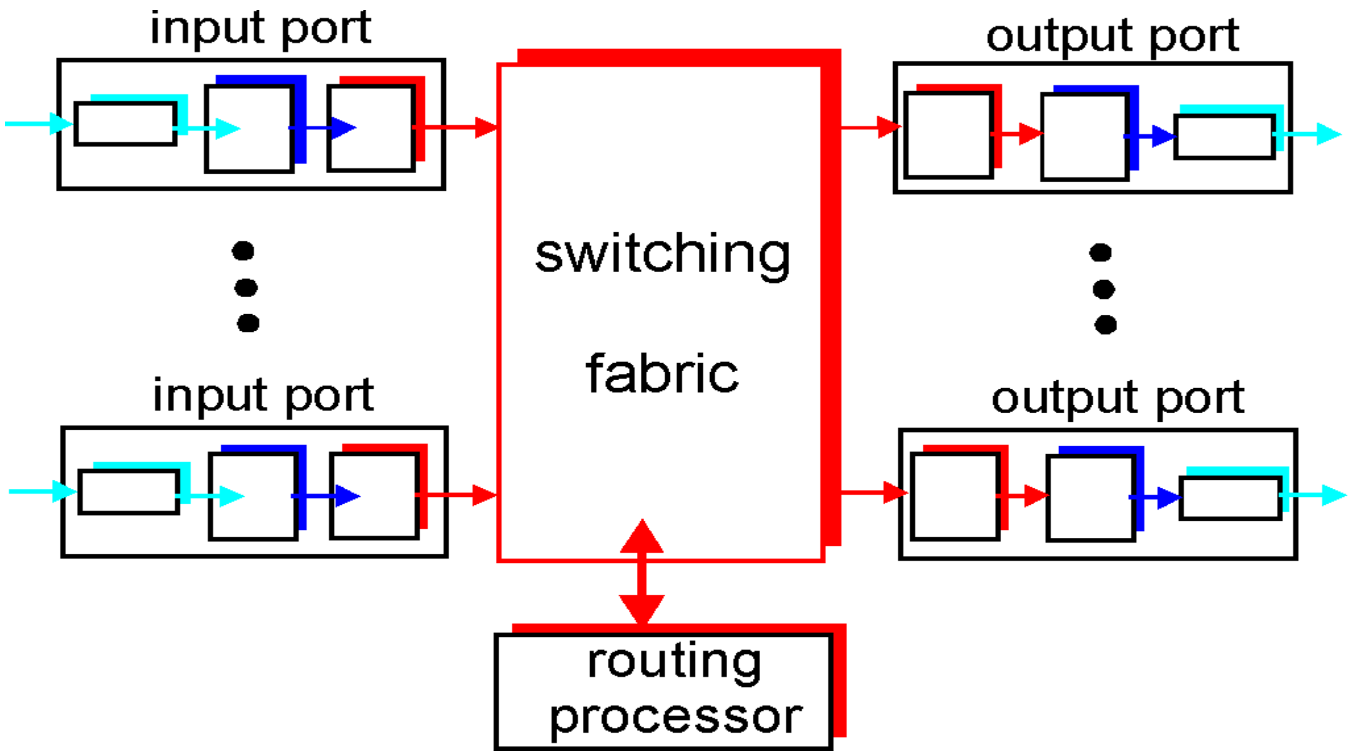
- Class A, B, C addresses support two levels of hierarchy

- However, the host portion can be further split into “subnets” by the address class owner
  - more than 2 levels of hierarchy

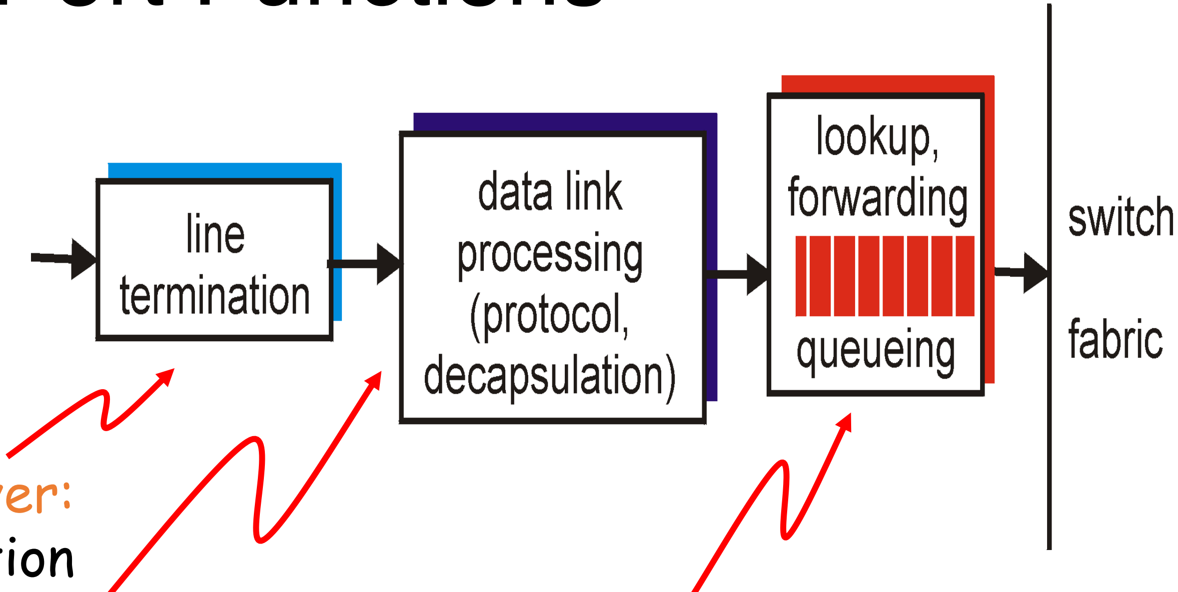
Basic Components



Router Architecture Overview



Input Port Functions



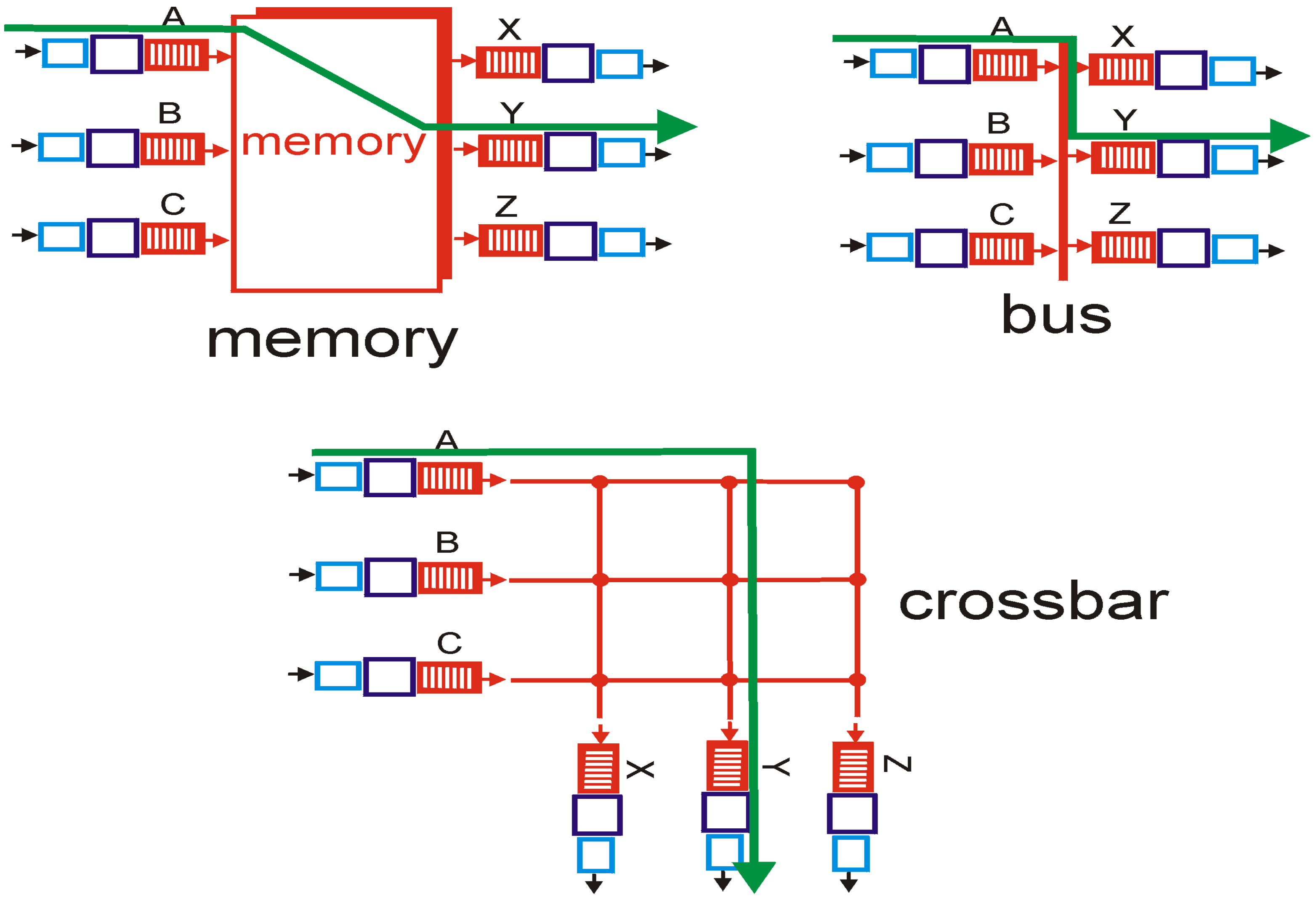
Physical layer:
bit-level reception

Data link layer:
e.g., Ethernet
see chapter 5

Decentralized switching:

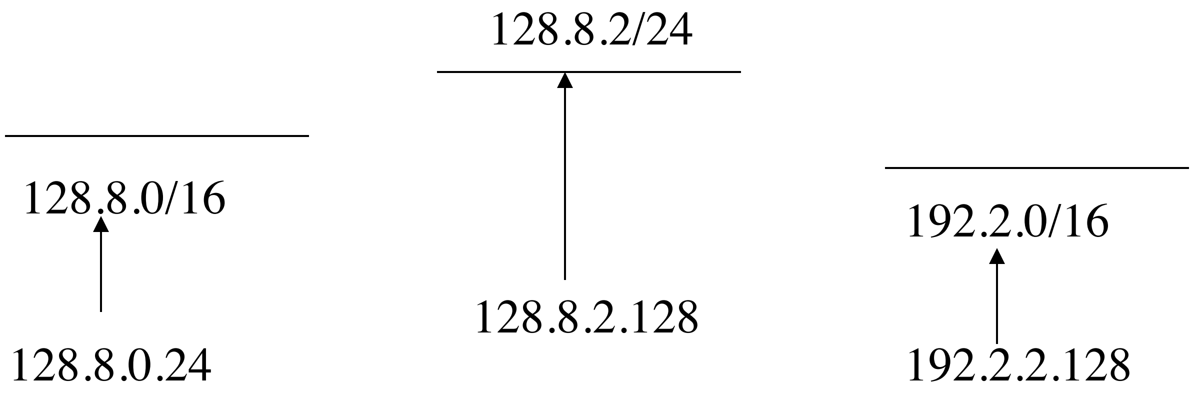
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

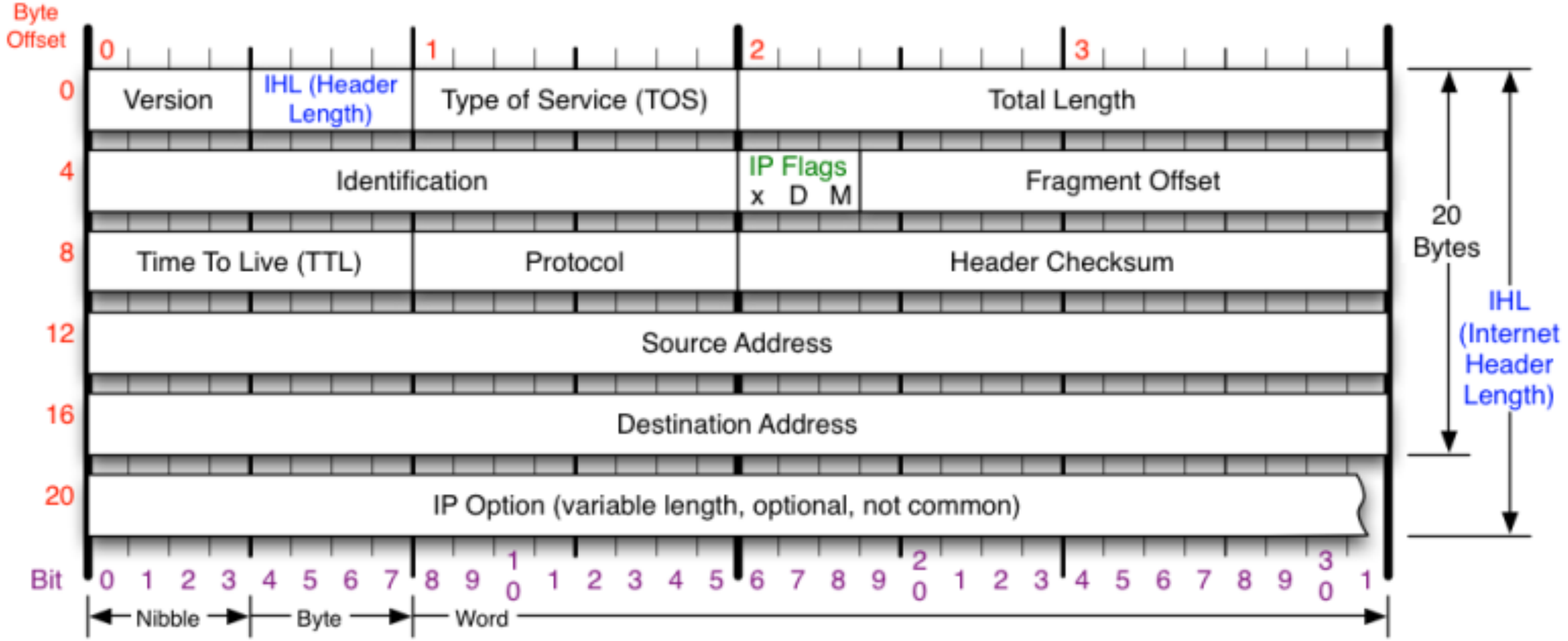
Three types of switching fabrics



Longest prefix match

- With CIDR, route entries are prefixes <prefix, CIDR mask>
- Can be aggregated
- We need to find the longest matching prefix that matches the destination address
- Need to search all prefixes of all length (in order) and among prefixes of the same length





Version

Version of IP Protocol. 4 and 6 are valid. This diagram represents version 4 structure only.

Header Length

Number of 32-bit words in TCP header, minimum value of 5. Multiply by 4 to get byte count.

Protocol

IP Protocol ID. Including (but not limited to):

1 ICMP	17 UDP	57 SKIP
2 IGMP	47 GRE	88 EIGRP
6 TCP	50 ESP	89 OSPF
9 IGRP	51 AH	115 L2TP

Total Length

Total length of IP datagram, or IP fragment if fragmented. Measured in Bytes.

Fragment Offset

Fragment offset from start of IP datagram. Measured in 8 byte (2 words, 64 bits) increments. If IP datagram is fragmented, fragment size (Total Length) must be a multiple of 8 bytes.

Header Checksum

Checksum of entire IP header

IP Flags

x D M

x 0x80 reserved (evil bit)
D 0x40 Do Not Fragment
M 0x20 More Fragments follow

RFC 791

Please refer to RFC 791 for the complete Internet Protocol (IP) Specification.

The Internet Protocol (IP)

- Provides delivery of packets from one host to any other host in the Internet
- Internet packets are called “datagrams” and may be up to 64 kilobytes in length
  - although they are typically much smaller