CS 43: Computer Networks

18: Network Layer, IP

November 12, 2018

SWARTHMORE COLLEGE

Reading Quiz

The Network Layer!

Application: the application (e.g., the Web, Email)

Transport: end-to-end connections, reliability

Network: routing

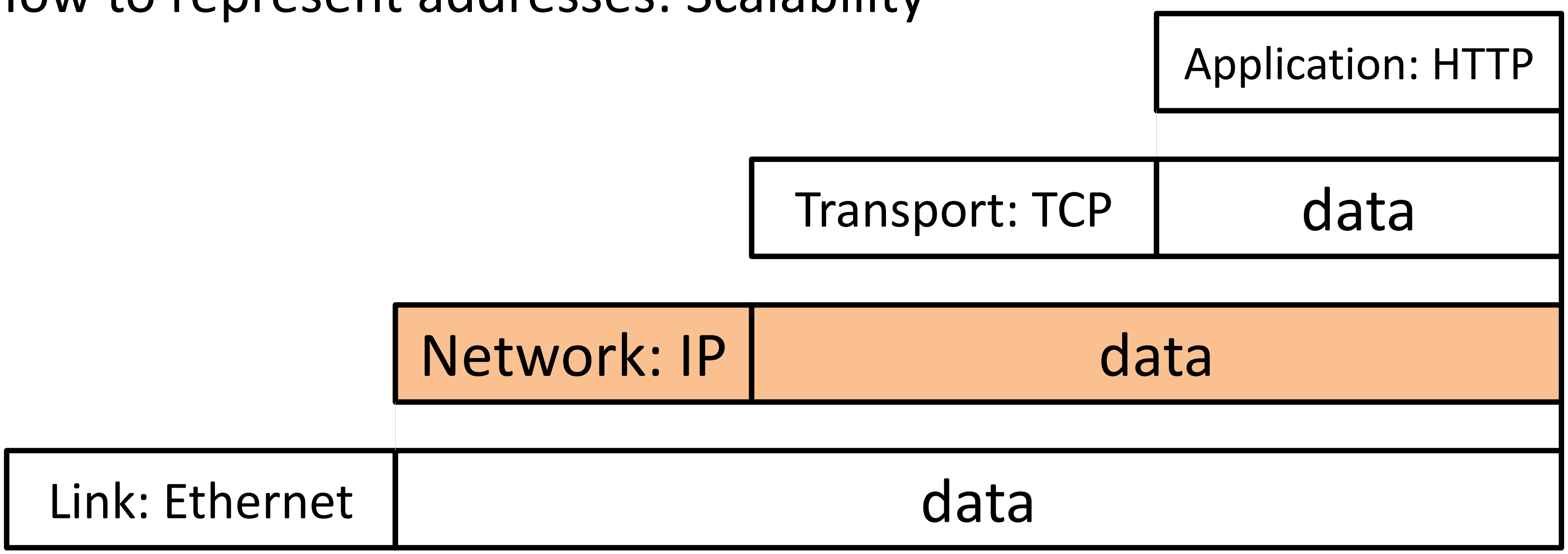
Link (data-link): framing, error detection

Physical: 1’s and 0’s/bits across a medium (copper, the air, fiber)

Network Layer

- Function: Route packets end-to-end on a network, through multiple hops

- Key challenge
  - How to route packets: Convergence
  - How to represent addresses: Scalability



Network Layer Functions

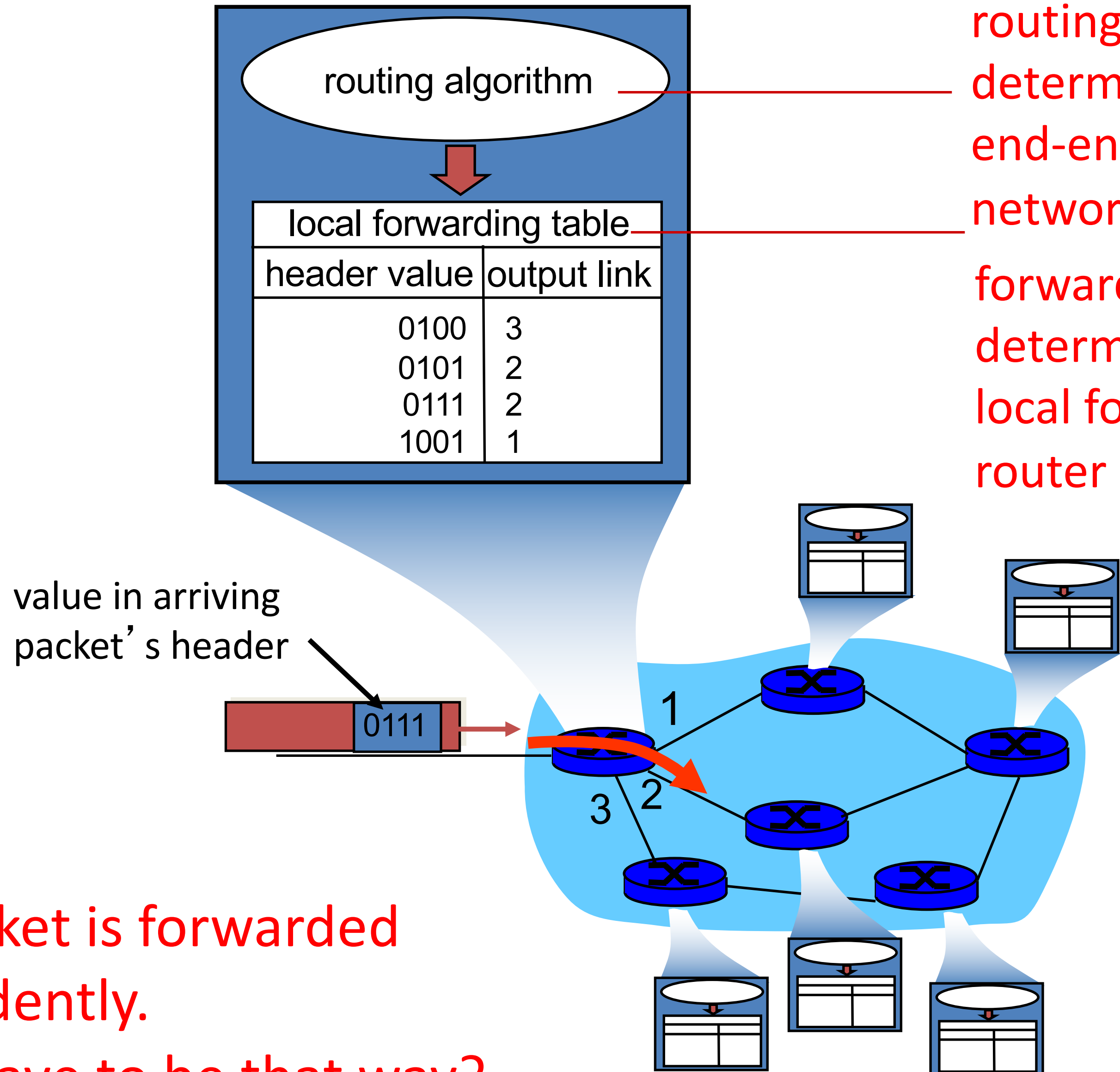
- Forwarding: move packets from router's input to appropriate router output
  - Look up in a table
- Routing: determine route taken by packets from source to destination.
  - Populating the table

Datagram vs. "Virtual Circuit"

- Datagram network provides network-layer connectionless service (packet switching)

- Virtual-circuit network provides network-layer connection service (like circuit switching)

Interplay between routing and forwarding

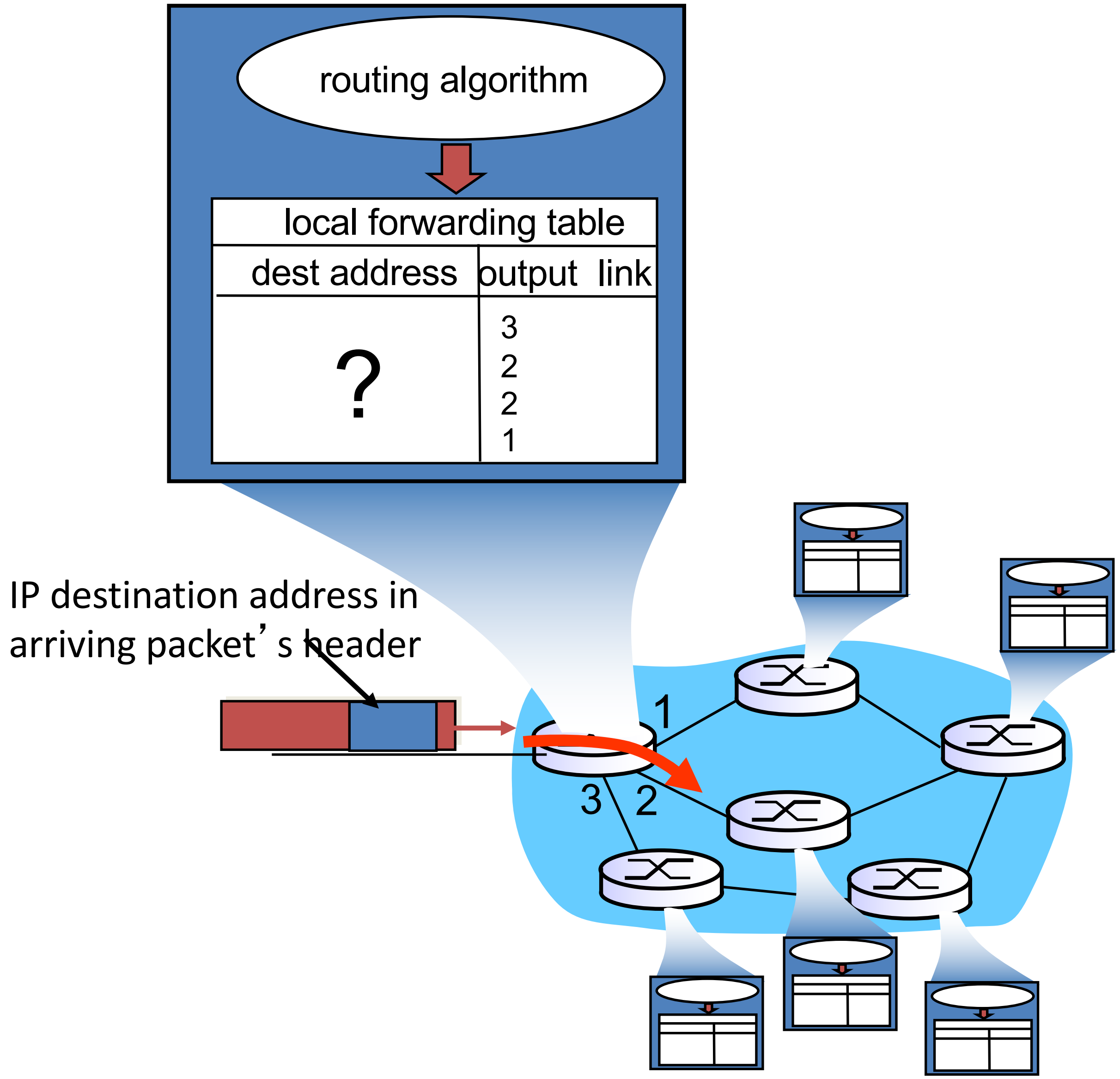


routing algorithm determines end-end-path through network

forwarding table determines local forwarding at this router

Each packet is forwarded independently.
Does it have to be that way?

Datagram forwarding table



How should we populate a router's forwarding table?

A. A person should add entries to the table.

B. A program external to the router should add entries to the table.

C. Routers should communicate with each other to add entries to their tables.

D. Some other mechanism.

Routers exchange state (we'll save the what and when for later). They decide, for each destination, how to get there, and build a lookup structure for their forwarding table. What should they build?

A. A list – scan for the destination.

B. A hash table – look up the destination.

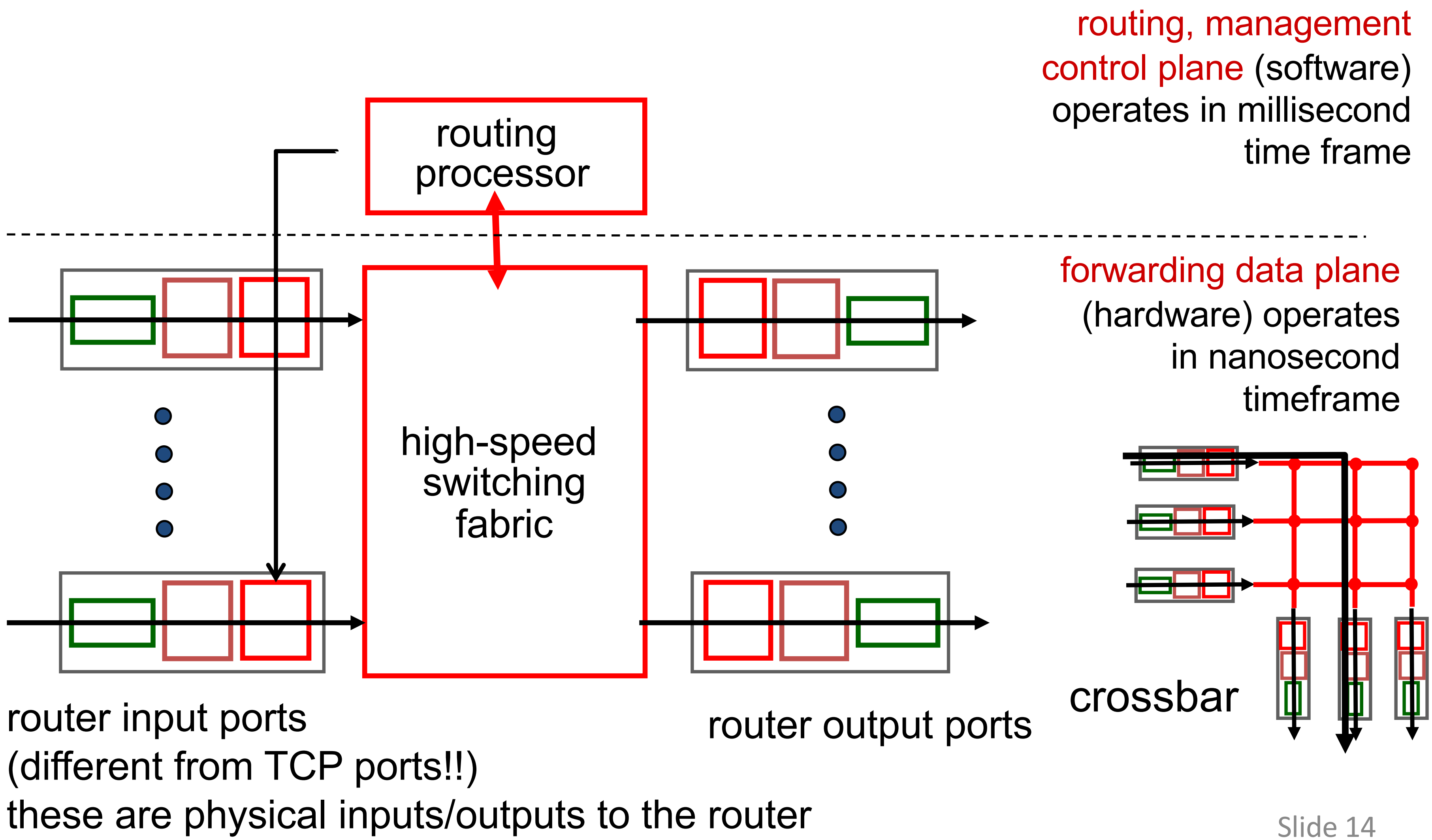
C. A tree – Follow branches that lead to the destination.

D. Some other software structure.

E. We can't do this in software, we need special hardware.

Aside: router architecture overview

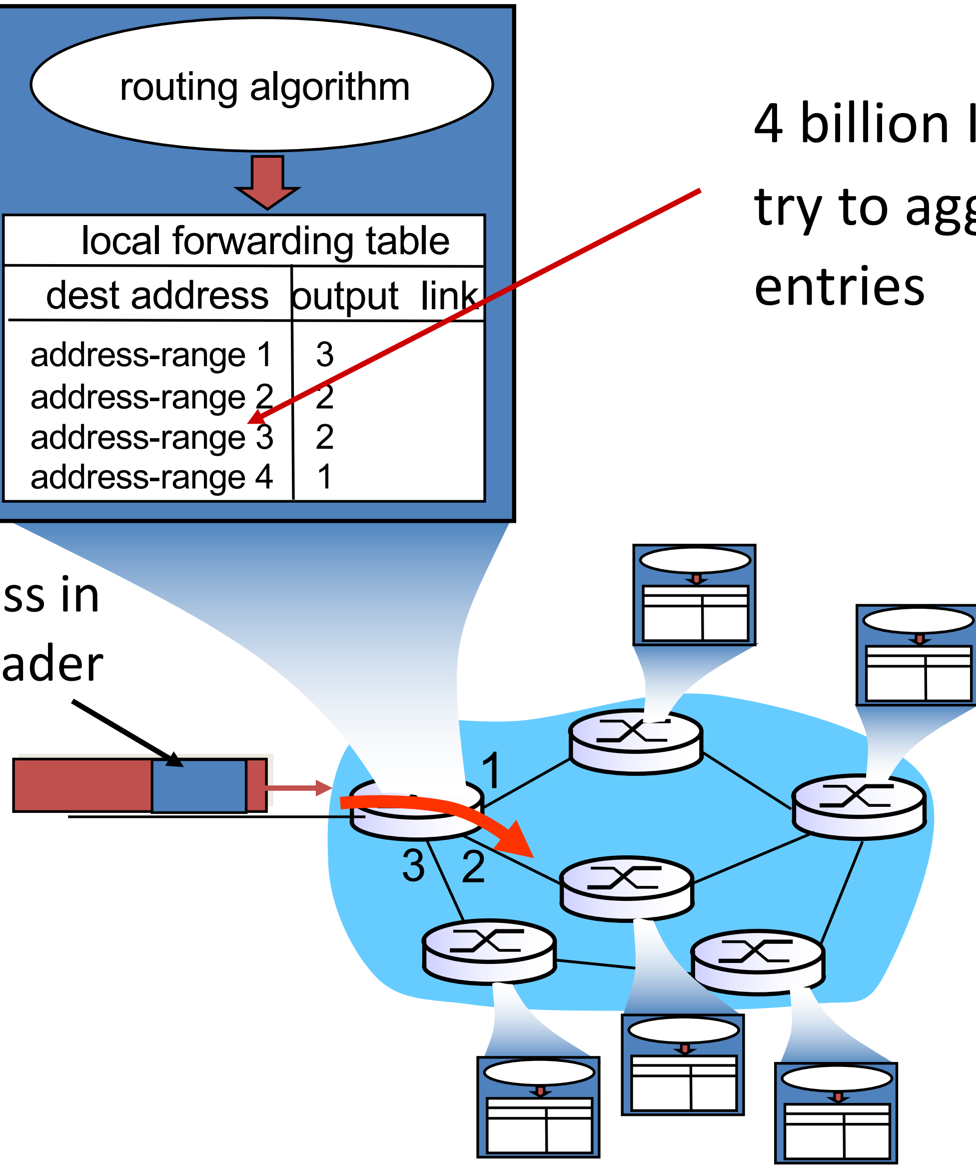
- high-level view of generic router architecture:



Datagram forwarding table

4 billion IP addresses, try to aggregate table entries

IP destination address in arriving packet's header



IP Addressing

- IP: 32-bit addresses
  - Usually written in dotted notation, e.g. 192.168.21.76
  - Each number is a byte
  - Stored in Big Endian order (network byte order)

	0	8	16	24	31
Decimal	192	168	21	76	
Hex	C0	A8	15	4C	
Binary	11000000	10101000	00010101	01001100	

Datagram forwarding table

Destination Address Range	Link Interface
200.23.16.* through 200.23.23.*	0
200.23.24.0 through 200.23.24.255	1
200.23.25.* through 200.23.31.*	2
Otherwise (default gateway)	3

Q: but what happens if ranges don't divide up so nicely?

Datagram forwarding table

Destination Address Range	Link Interface
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
Otherwise (default gateway)	3

Q: but what happens if ranges don't divide up so nicely?

Longest prefix matching

In a forwarding table entry, use the longest address prefix that matches destination address.

Destination IP Address Range	Link interface
<upper 16 bit> 000100*** *********	0
<upper 16 bit> 00011000 *********	1
<upper 16 bit> 00011***** *********	2
Otherwise (default gateway)	3

DA: **<upper 16 bits> 00011000** 10101010

DA: **<upper 16 bits> 00010**110 10100001

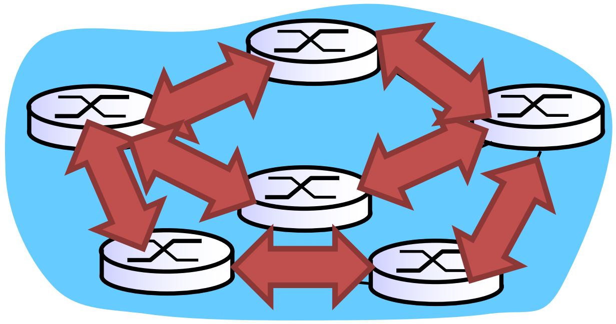
which interface?

Routing

Traditional

- Routers run a routing protocol to exchange state.

- Use state to build up the forwarding table.



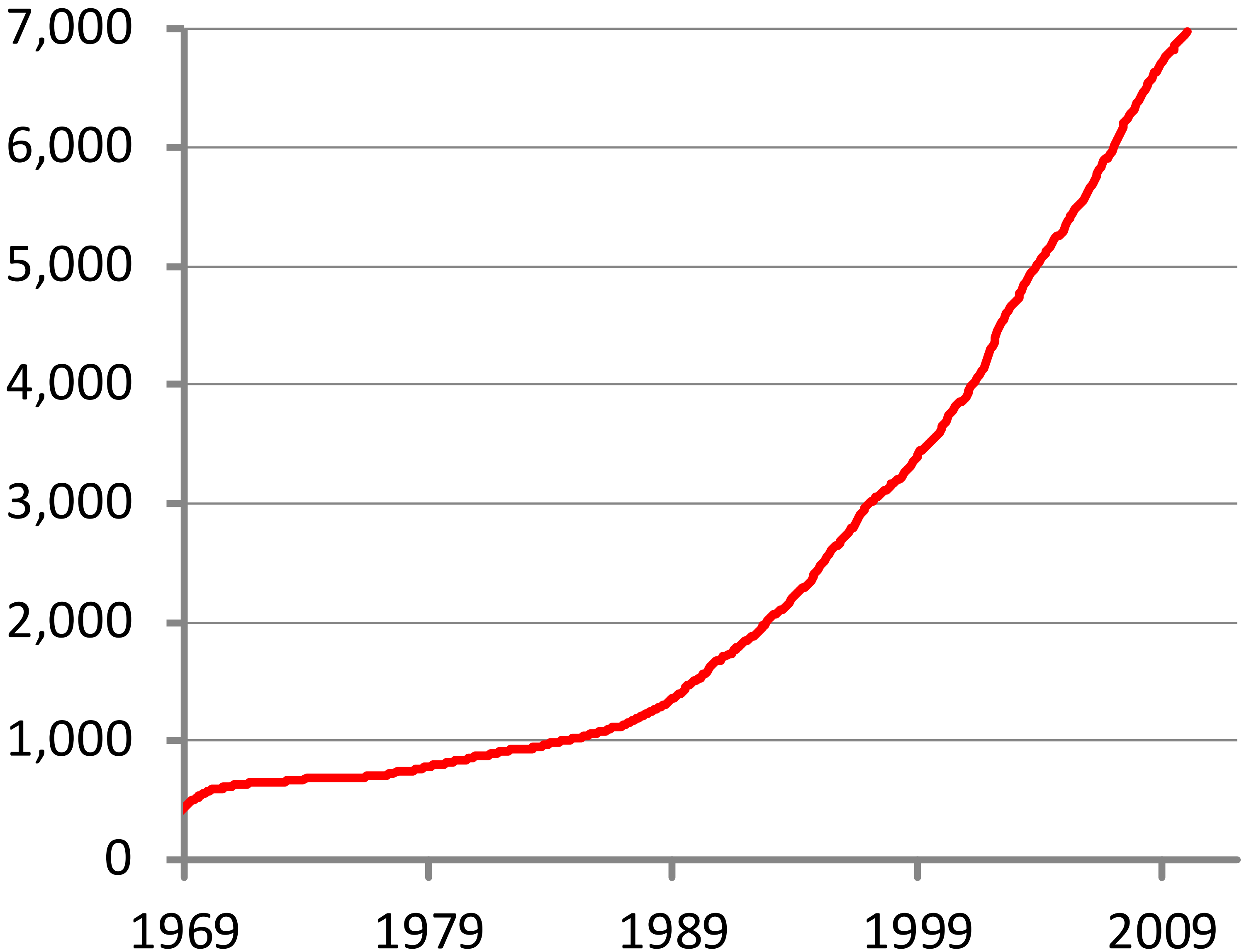
What services would we like a router to implement?

A. Basic connectivity: route packets to destination
B. Find policy-compliant paths (keep ISPs happy)
C. Traffic engineering
D. Impose limits on what can be accessed on the Internet vs. local ISP
E. All of the above

Nice things to have..

- Traffic engineering:
  - Want to avoid persistent overloads on links
  - Choose routes to spread traffic load across links
- Access Control:
  - Limit access to backend database machines.
  - Firewalls
- Network measurement

Number of published Internet Standards

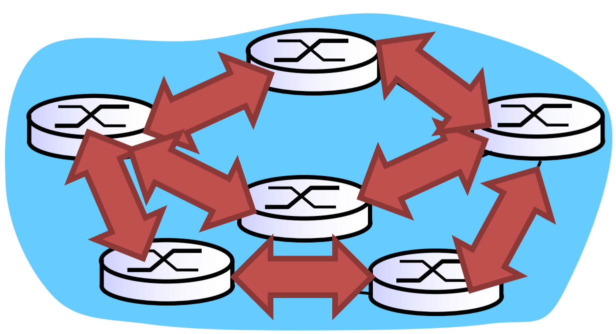


Graph from Nick McKeown

Routing

<u>Traditional</u>

- Routers run a routing protocol to exchange state.
- Use state to build up the forwarding table.

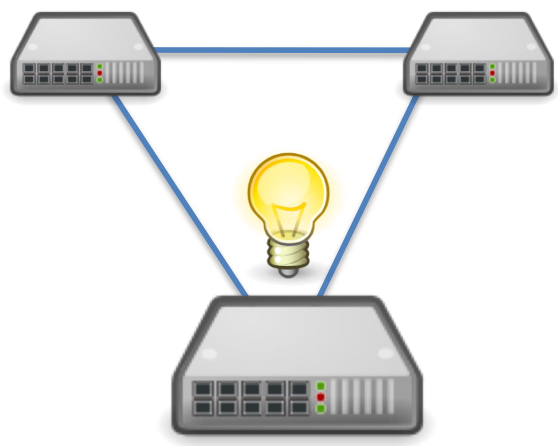


<u>Software-Defined</u>

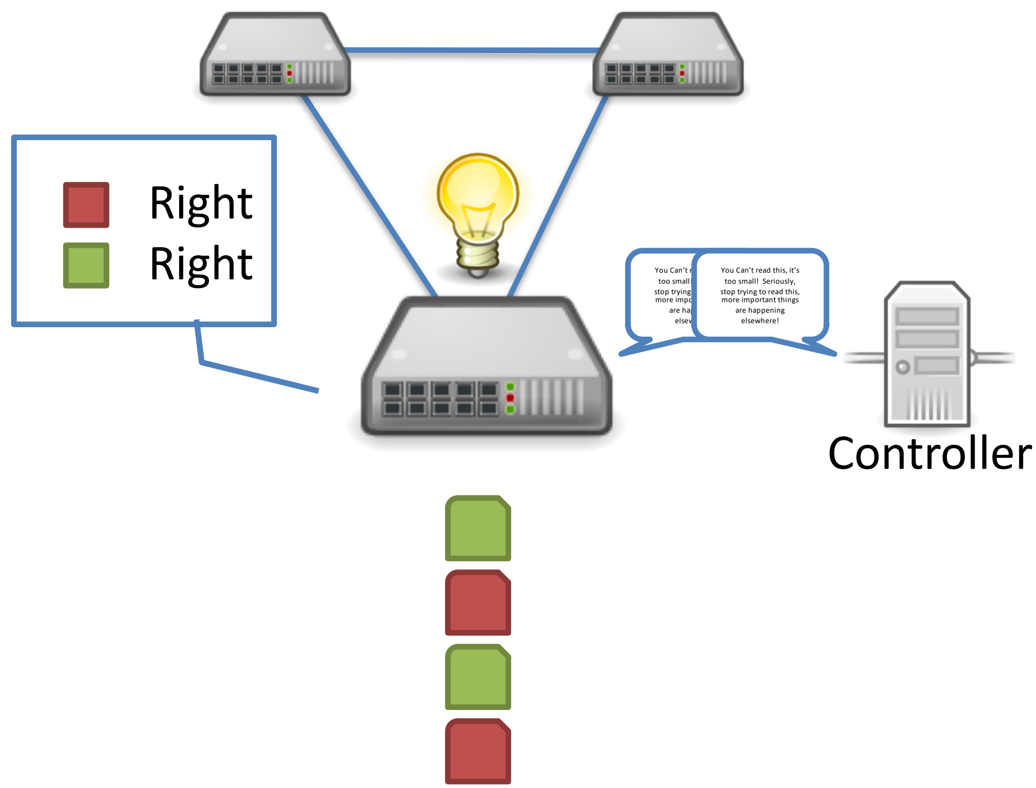
- Routers are dumb, just do what they're told.
- Controller service explicitly tells each router what to do.
- Rare on the Internet, hot topic in data centers.

Software-Defined Networking (SDN)

Traditional Hardware



SDN Hardware



Summary

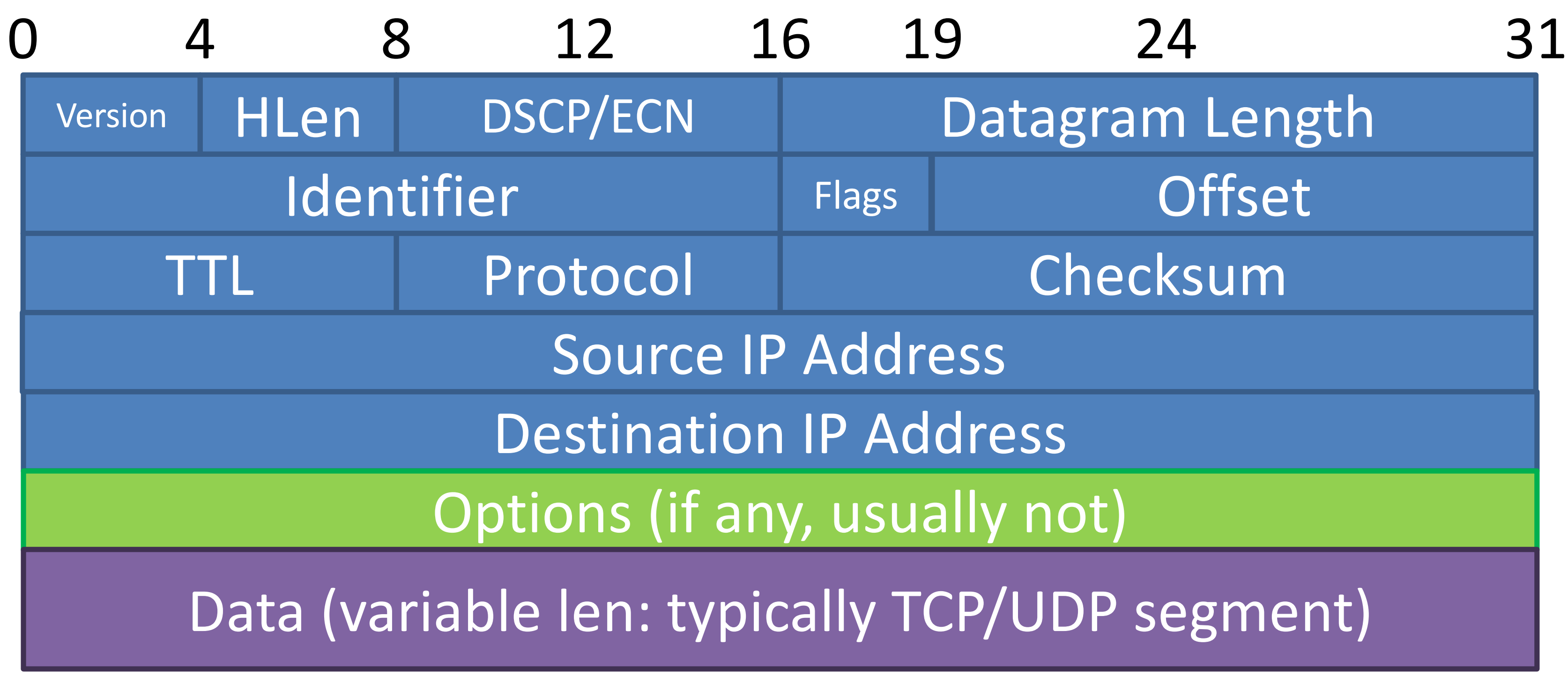
- On the Internet, best-effort packet switching is the norm
- Forwarding: move packets from router's input to appropriate router output: Look up in a table
- Routing: determine route taken by packets from source to destination: Populating the table
- Hardware helps with quick forwarding using longest prefix matching.

Today

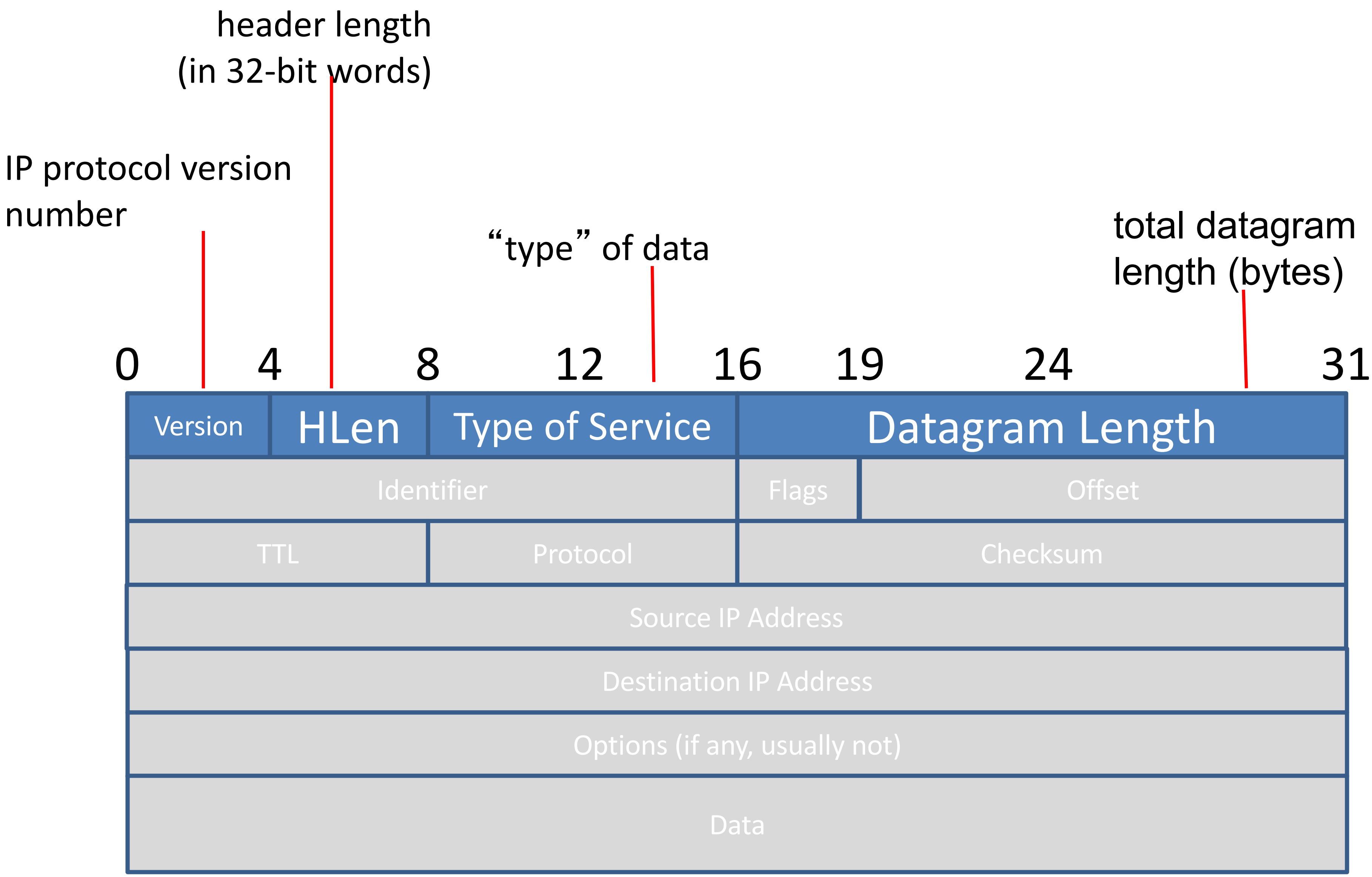
- IP header format
- Subnets and IP addressing
  - CIDR
  - Route aggregation
- DHCP: Assigning an IP address to an interface
- Fragmentation

IP Datagrams

- IP Datagrams are like a letter
  - Totally self-contained
  - Include all necessary addressing information
  - No advanced setup of connections or circuits

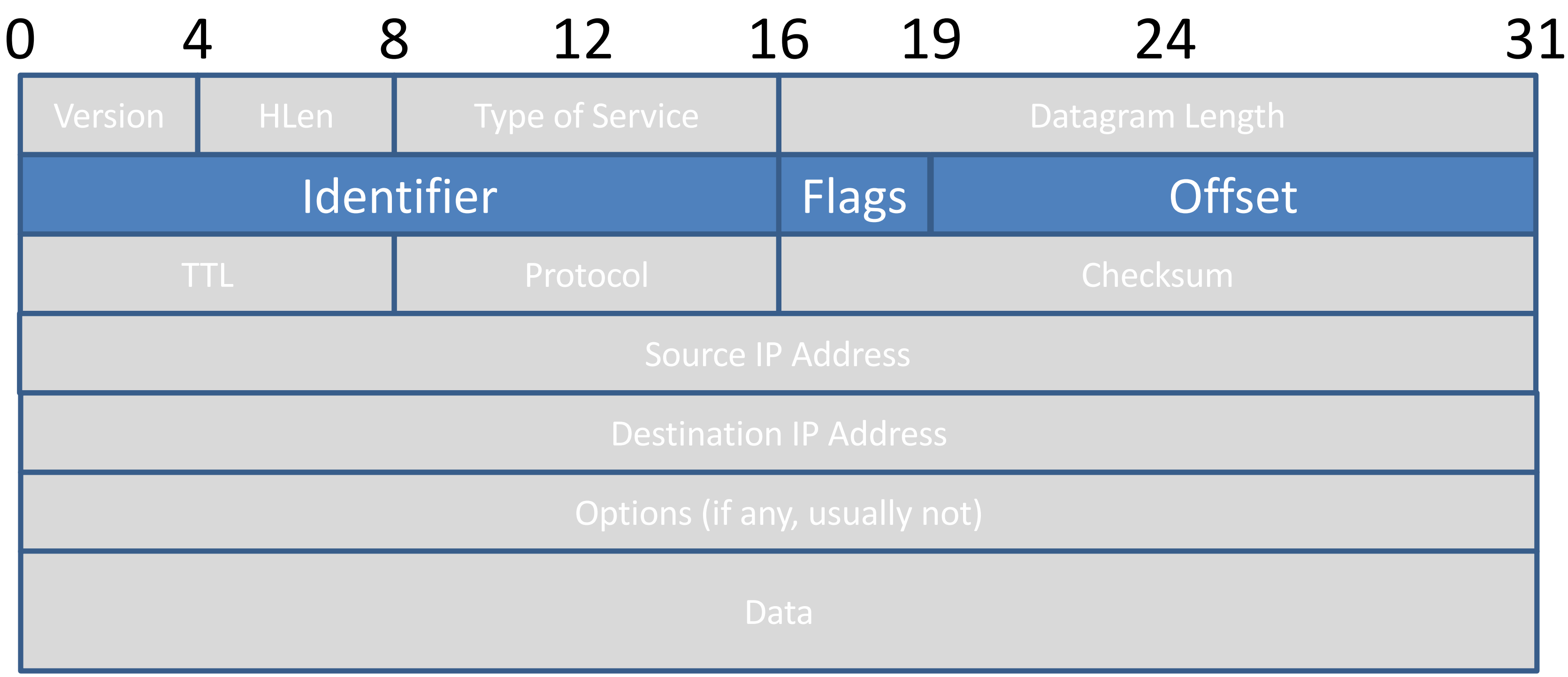


IP Datagram Format

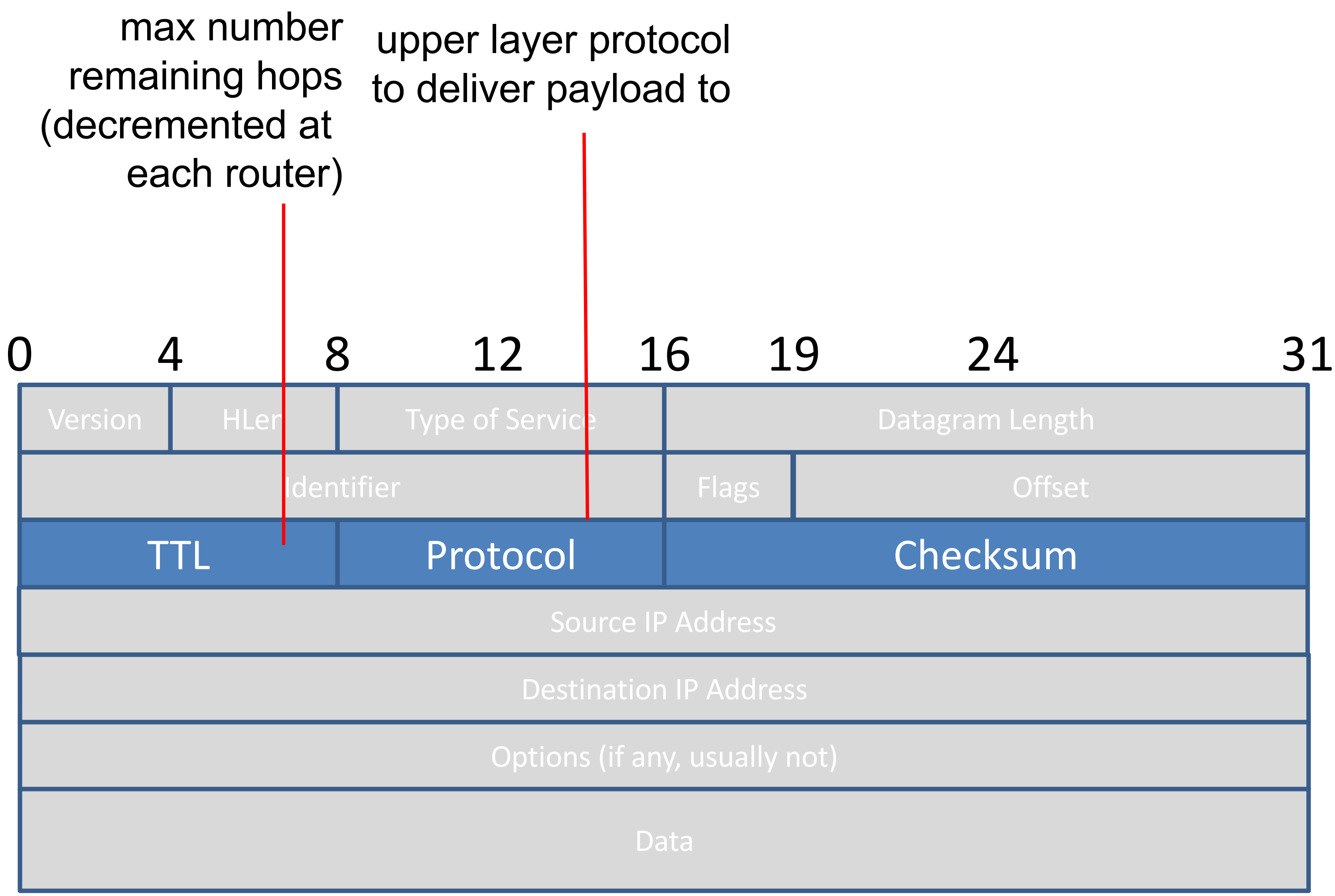


IP Datagram Format

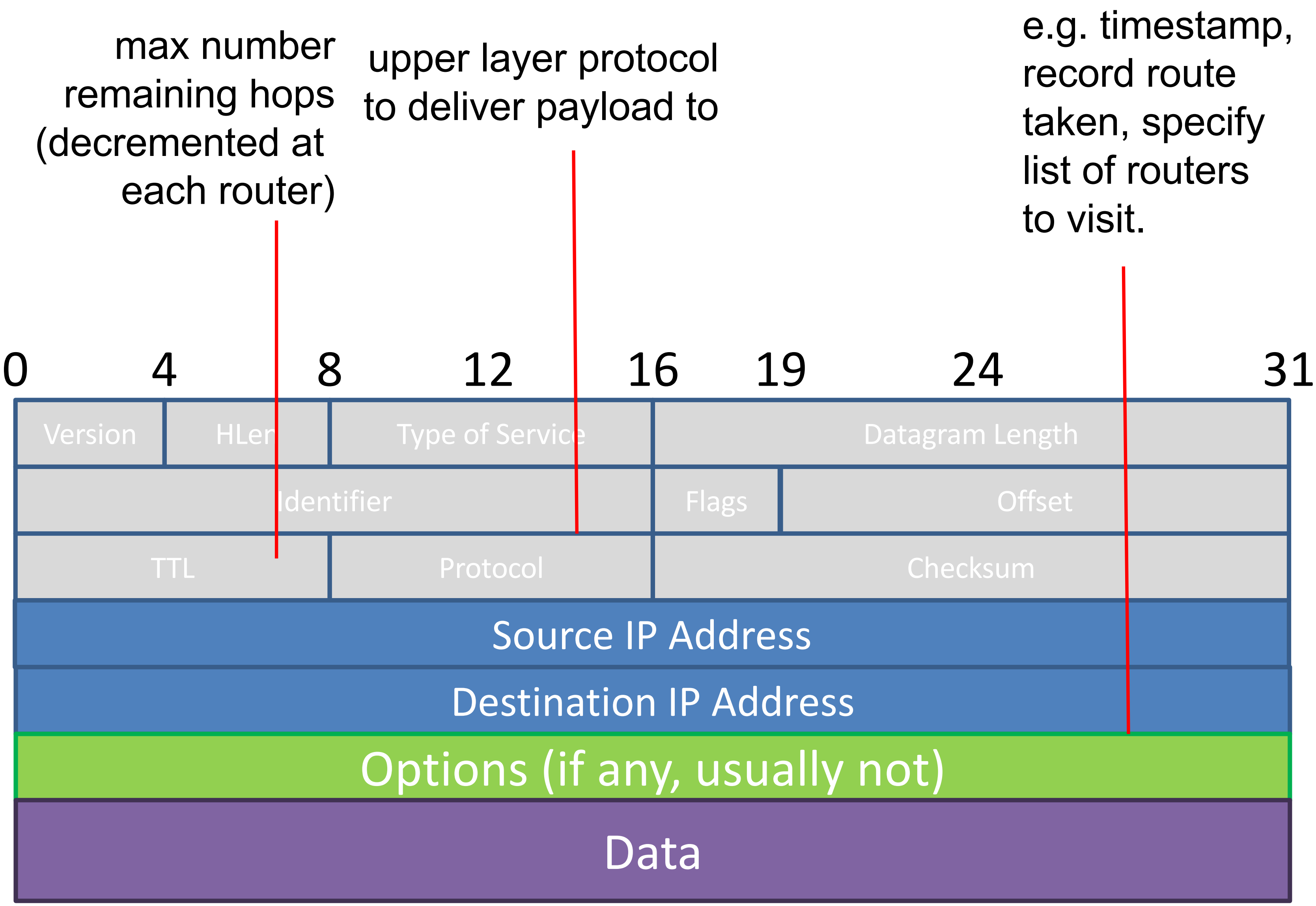
- fragmentation/ reassembly: Identifier, Flags, Offset



IP Datagram Format



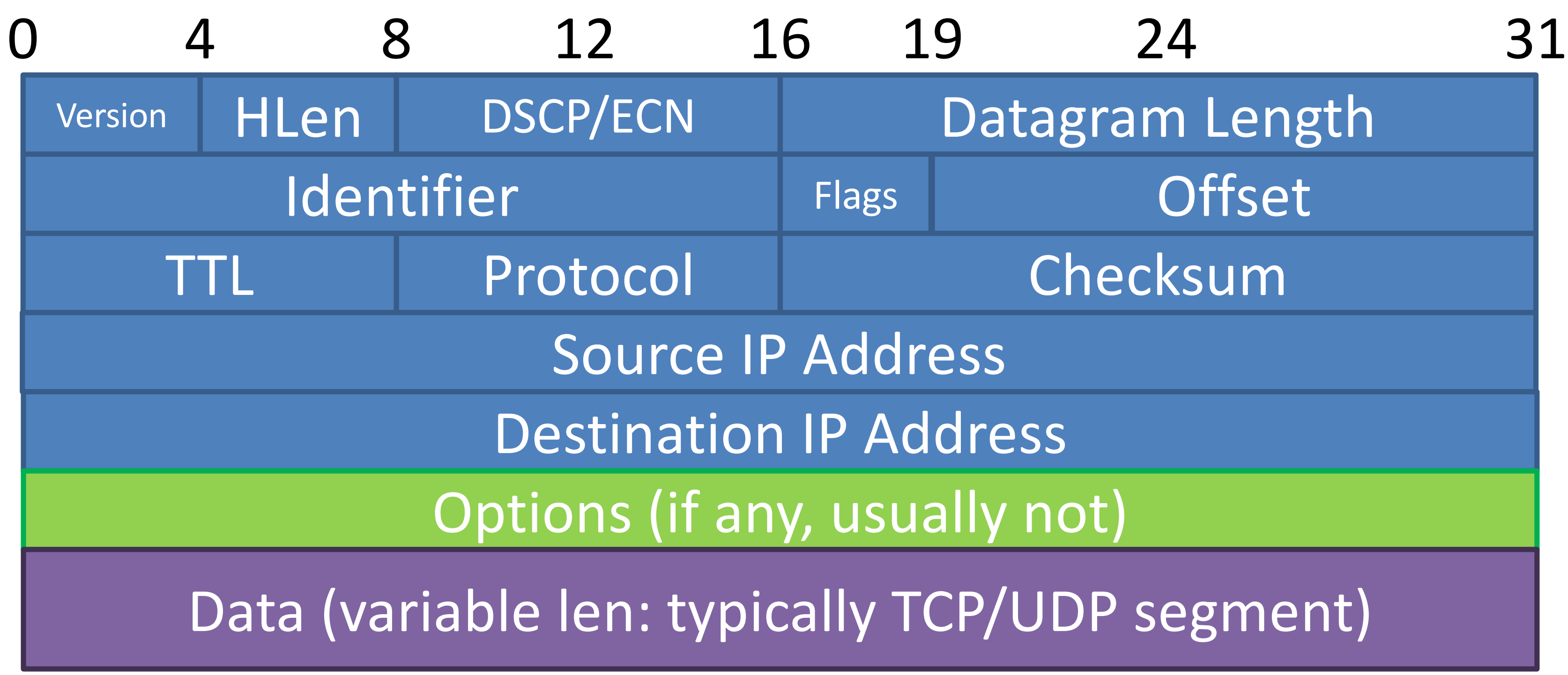
IP Datagram Format



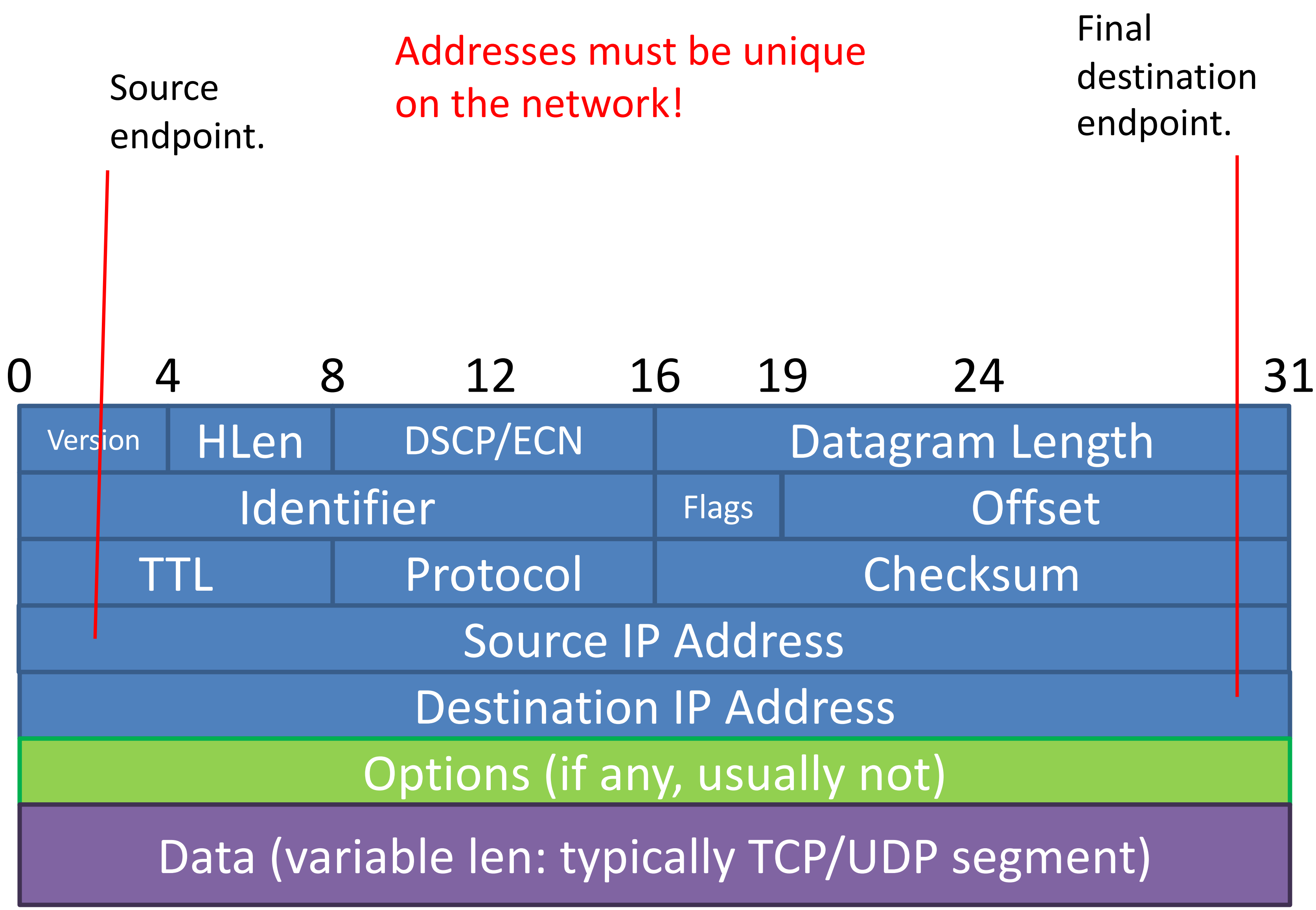
IP Datagrams

how much overhead?

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



IP Datagrams

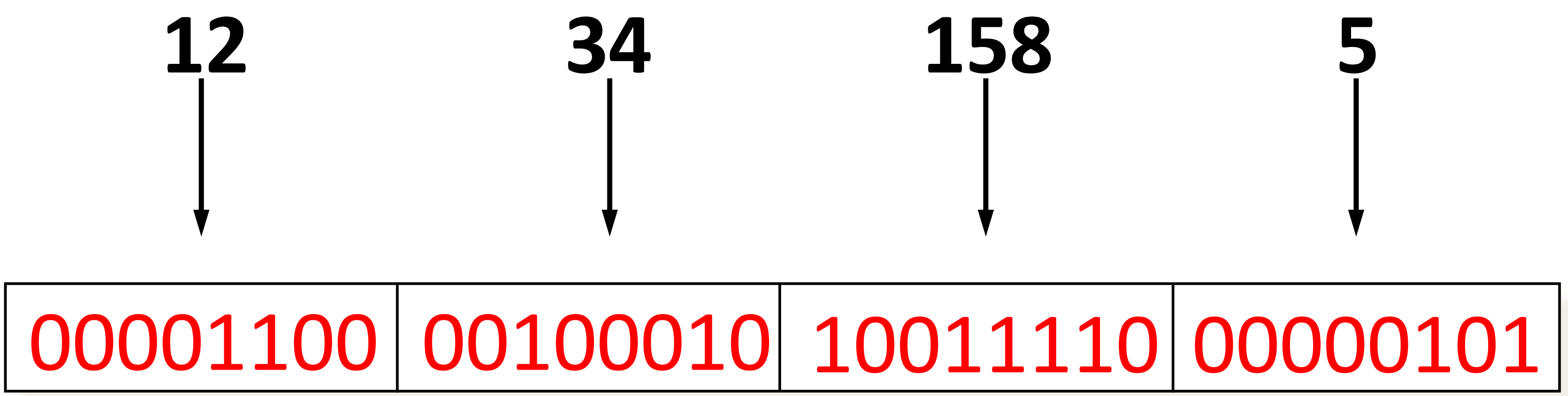


What's in a name?

- Host name: web.cs.swarthmore.edu
  - Domain: registrar for each top-level domain (e.g., .edu)
  - Host name: local administrator assigns to each host
- IP addresses: 130.58.68.164
  - Prefixes: ICANN, regional Internet registries, and ISPs
  - Hosts: static configuration, or dynamic using DHCP
- MAC addresses: D8:D3:85:94:5F:1E
  - OIDs: assigned to vendors by the IEEE
  - Adapters: assigned by the vendor from its block

IP Address (IPv4)

- A unique 32-bit unsigned integer value
- Identifies an interface (on a host, on a router, …)
- Represented in dotted-quad/octet notation



IP Addresses

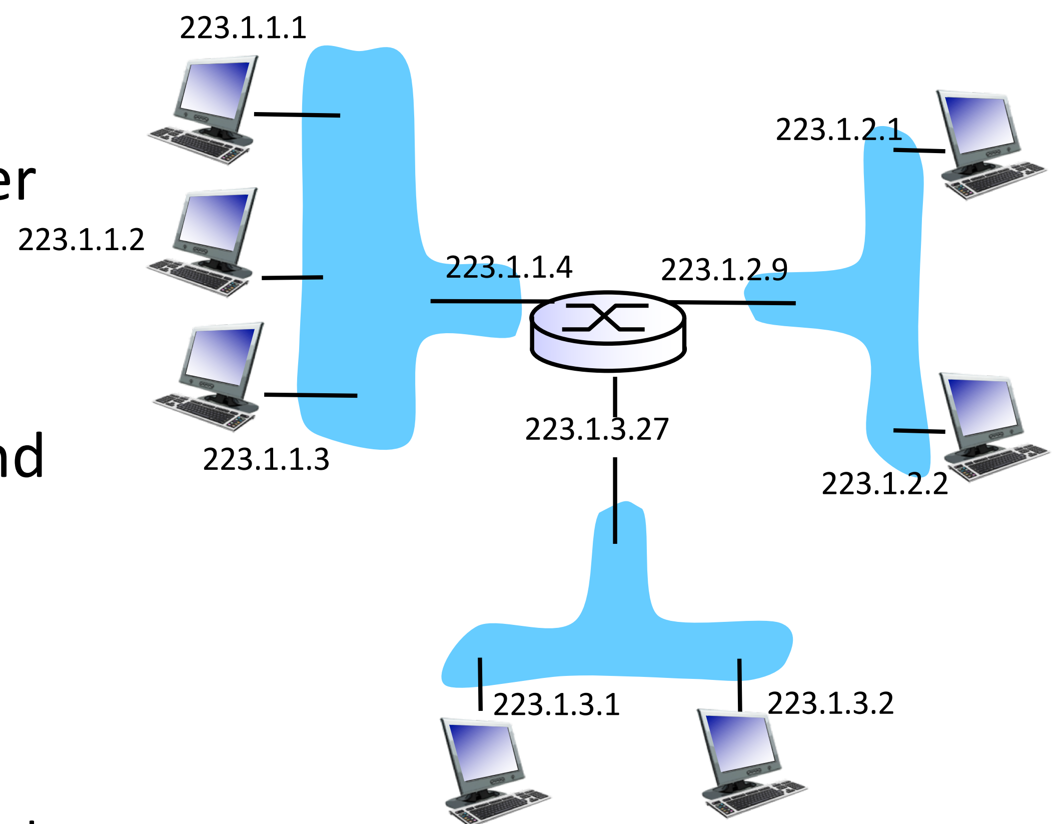
- 2^{32} => 4,294,967,296 possible addresses.

- In the early 80's, that's a lot!
  - Population was ~4.5 billion.

- Now…not so much.
  - Population > 7 billion.

Network Interfaces

- IP address: 32-bit identifier for host, router *interface*
- interface: connection between host/router and physical link
  - router's typically have multiple interfaces
  - host typically has one or two interfaces (e.g., wired Ethernet, wireless 802.11)
- IP addresses associated with each interface



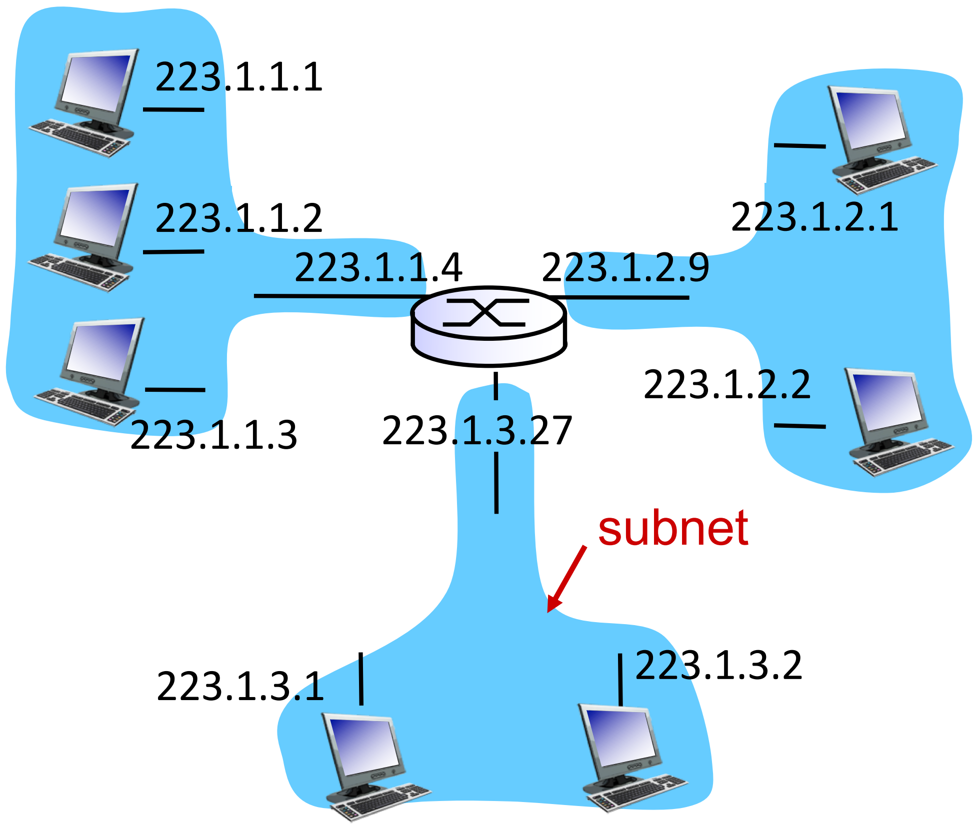
223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

Subnets

- IP address:
  - subnet part - high order bits
  - host part - low order bits

- what's a subnet?
  - device interfaces with same subnet part of IP address
  - can physically reach each other without intervening router
  - On the same link layer



network consisting of 3 subnets

Assigning Addresses

- IANA – Internet Assigned Numbers Authority
  - (Run by Jon Postel until 1988)
  - Now a part of ICANN
  - ARIN: North America
  - RIPE: Europe

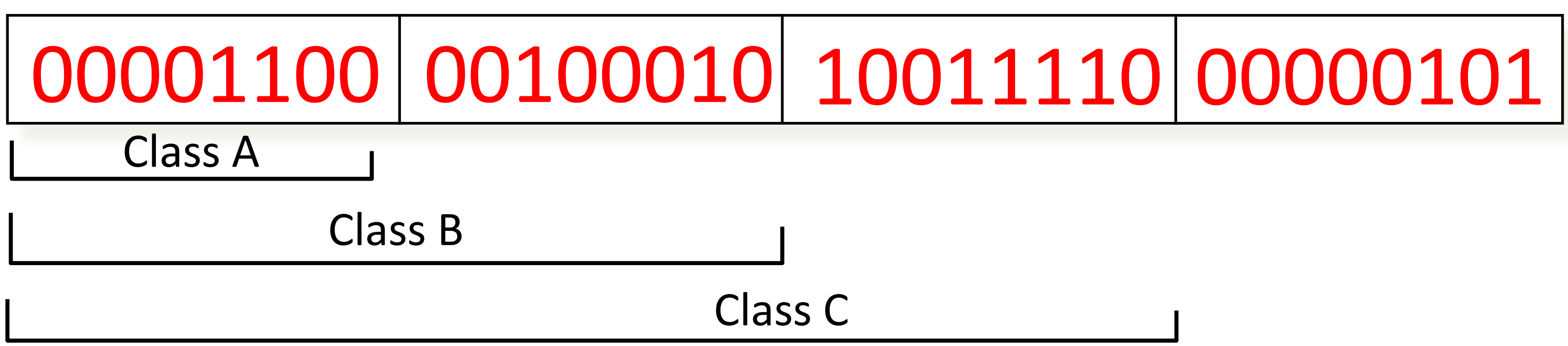
- ICANN: Internet Corporation for Assigned Names and Numbers
  - Manages IP addresses, DNS, resolves disputes

Who gets an address? How many?

- Back in the old days, you called up Jon Postel
  - "How many addresses do you need?"
  - "Here you go! I may have rounded a bit."

Who gets an address? How many?

- Classful Addressing
  - Class A: 8-bit prefix, 24 bits for hosts (16,777,216)
  - Class B: 16-bit prefix, 16 bits for hosts (65,536)
  - Class C: 24-bit prefix, 8 bits for hosts (256)



CIDR

- Classless Interdomain Routing
  - Prefix (subnet) length is no longer fixed
  - (Can be division of bits rather than just 8/24, 16/16, and 24/8)

Why do we give out addresses in CIDR blocks? How many of these statements are true? (Which ones?)

- It requires fewer resources at routers.
- It requires fewer resources at end hosts.
- It reduces the number of block allocations that need to be managed.
- It better utilizes the IP address space.

A – 0, B – 1, C – 2, D – 3, E – 4

CIDR

- Classless Interdomain Routing
  - Prefix (subnet) length is no longer fixed
  - Address blocks come with a **subnet mask**

- Subnet mask written in two ways:
  - Dotted decimal: 255.255.240.0
  - /20
  - Both mean:

    11111111 11111111 11110000 00000000

/20

CIDR

- Addresses divided into two pieces:
  - Prefix portion (network address)
  - Host portion

- Given an IP address and mask,
  we can determine:
  - The prefix (network address) by ANDing
  - The broadcast address by ORing inverted mask

Network Address (Subnet Address)

- E.g., 230.8.1.3/18   /18 => mask is 255.255.192.0

11100110	**00001000**	**00000001**	**00000011**	IP address
11111111	**11111111**	**11000000**	**00000000**	/18 Subnet mask

Network Address (Subnet Address)

- E.g., 230.8.1.3/18 /18 => mask is 255.255.192.0

11100110	**00001000**	**00000001**	**00000011**	IP address
11111111	**11111111**	**11000000**	**00000000**	/18 Subnet mask
11100110	**00001000**	**00000000**	**00000000**	and the two

Network address advertised by router: 230.8.0.0

Why might a device care about its "Network Address"?

- Answers the question: is the destination on the same subnet as me?
- Address + subnet mask -> Network address
- If destination is on same network:
  - Send directly to them
- Else:
  - Send to gateway router

Broadcast Address

- E.g., 230.8.1.3/18

11100110	**00001000**	**00000001**	**00000011**	IP address
~~11111111~~	**~~11111111~~**	**~~11000000~~**	**~~00000000~~**	/18 Subnet mask
00000000	**00000000**	**00111111**	**11111111**	complement of the subnet mask

Broadcast Address

- E.g., 230.8.1.3/18

11100110	**00001000**	**00000001**	**00000011**	IP address
00000000	**00000000**	**00111111**	**11111111**	complement of the subnet mask

Broadcast Address

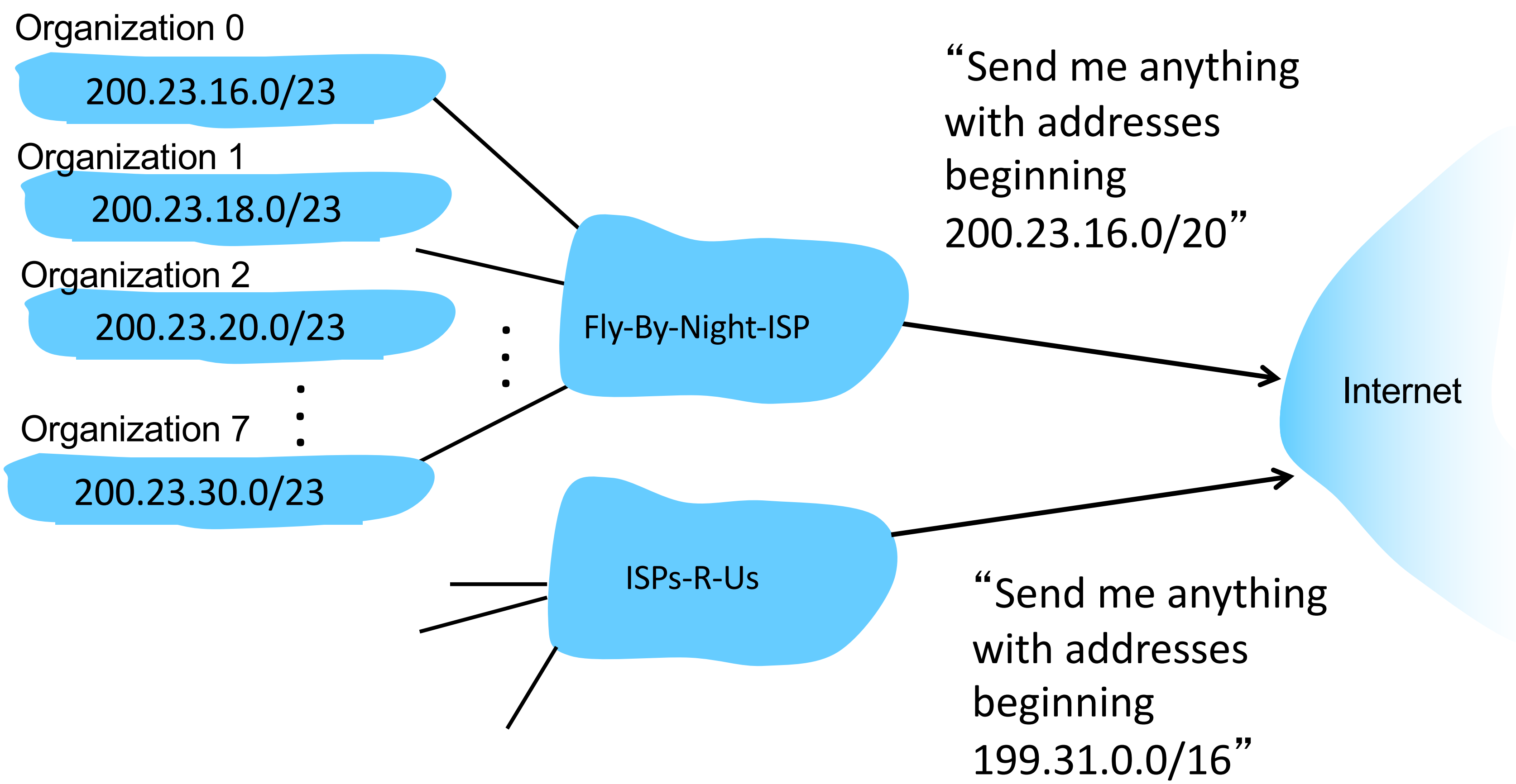
- E.g., 230.8.1.3/18

11100110	**00001000**	**00000001**	**00000011**	IP address
00000000	**00000000**	**00111111**	**11111111**	complement of the subnet mask
11100110	**00001000**	**00111111**	**11111111**	OR of the IP address and the complement of the subnet mask

Broadcast address: 230.8.63.255

Hierarchical Addressing: Route Aggregation

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical Addressing: Route Aggregation

"Send me anything with addresses beginning 200.23.16.0/20" translates to the following:

	200	23	16	0	
200.23.16.0/20 =	11001000	00010111	00010000	00000000	/20 Prefix

/20 prefix

200.23.16.0 = 11001000 00010111 00010000 00000000

|

200.23.31.255 = 11001000 00010111 00011111 11111111

Range represented by the /20 prefix

/20 prefix contains the range of IP addresses that match the the first 20 bits, and can have any value for the remaining 12 bits in the range of :

[first 20 bits] 0000 00000000

[first 20 bits] 1111 11111111

A total of 2^12 = 4,096 IP addresses

Route aggregation in Fly-By-Night ISP

Fly-By-Night-ISP

200.23.16.0/20 = 11001000 00010111 00010000 00000000

Individual Organizations: All of these organizations IP addresses lie withinFly-by-Night's /20 prefix (first 20 bits are the same)

- they more specifically match on the three more bits to form a /23 prefix (first 23 bits of all IP addresses within their organization are the same).
- The last 9 (32-23) bits provide 2^9 = 512 unique IP addresses within each organization.

/23 prefixes

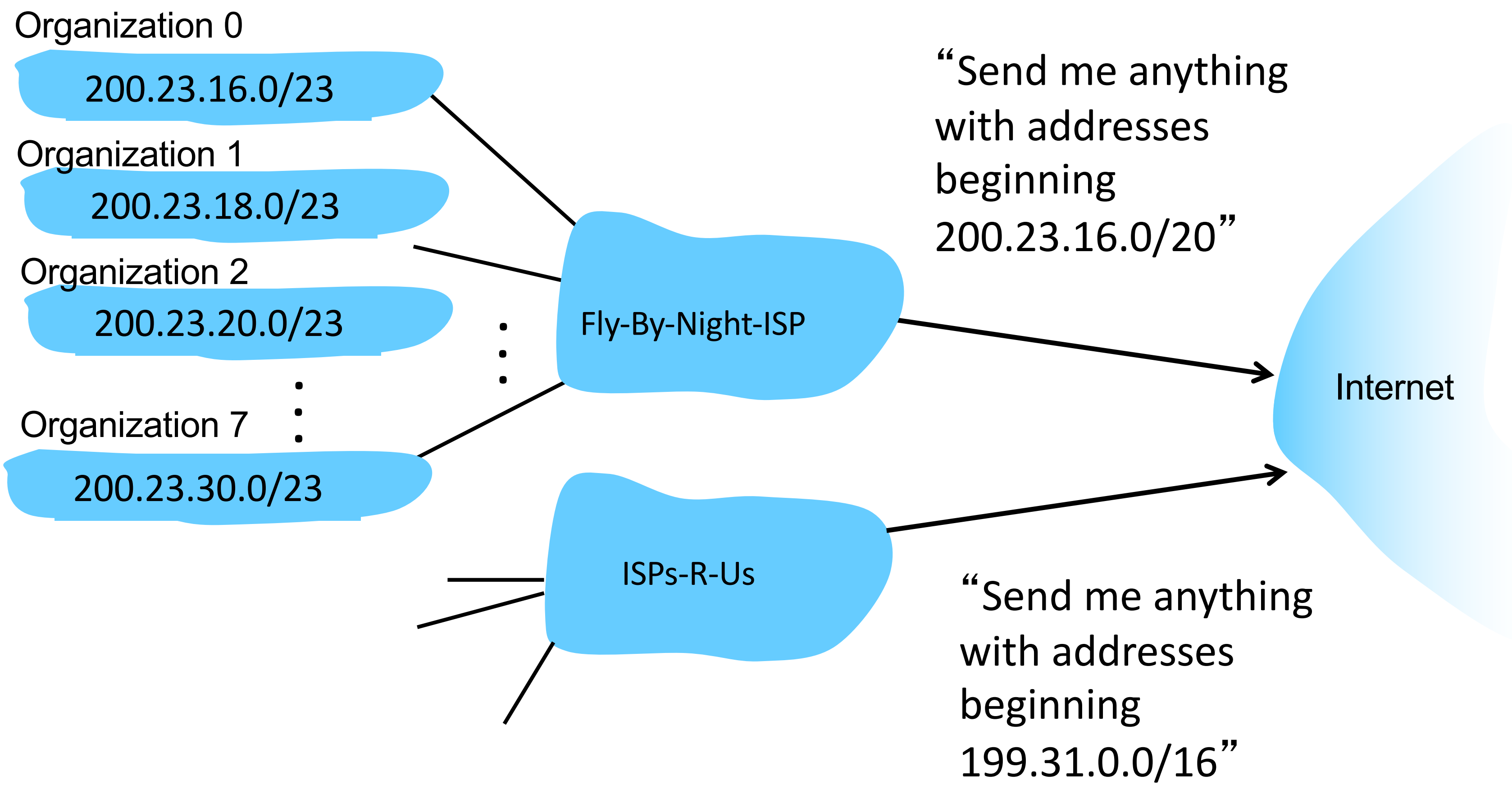
200.23.16.0/23 = 11001000 00010111 00010000 00000000

200.23.18.0/23 = 11001000 00010111 00010010 00000000

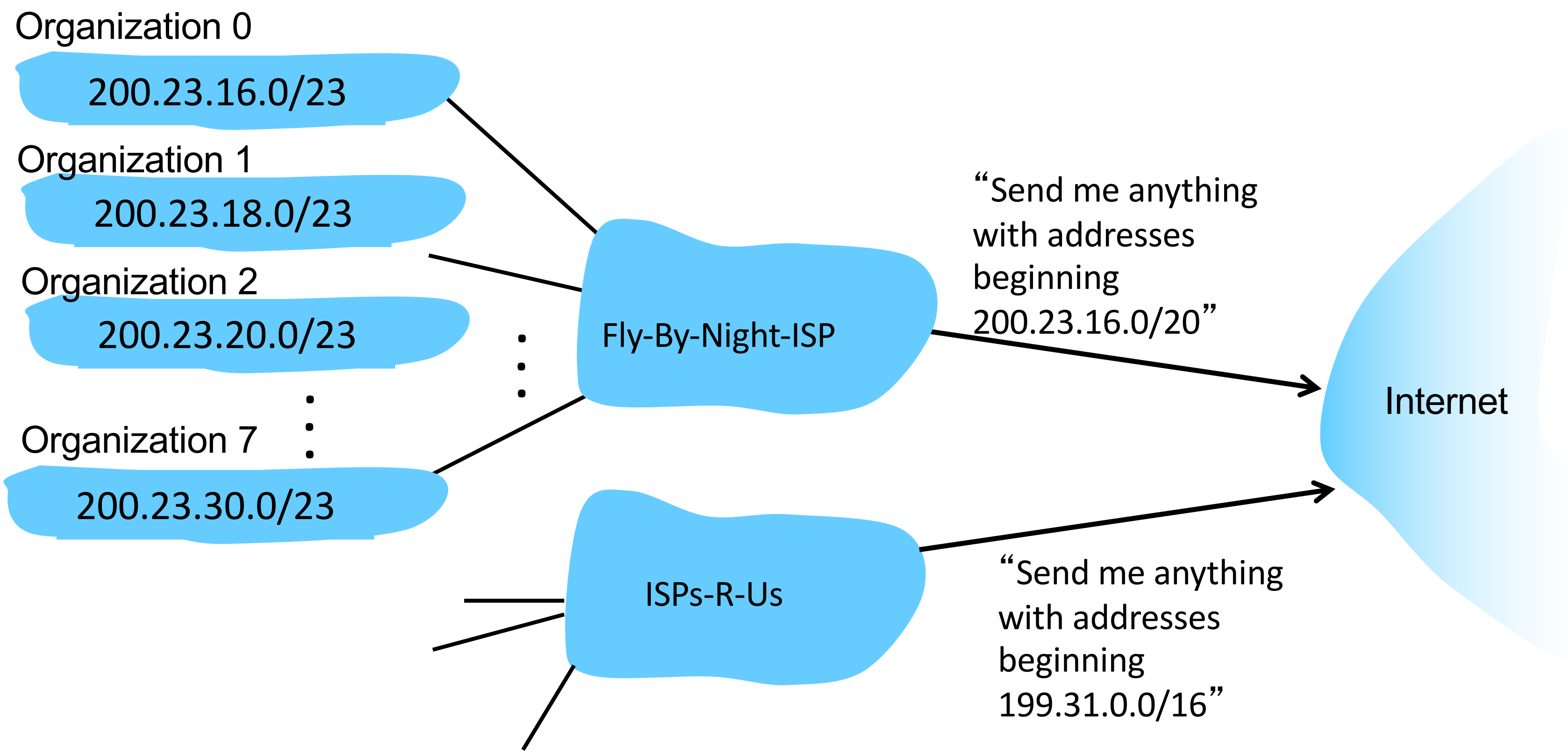
200.23.20.0/23 = 11001000 00010111 00010100 00000000

200.23.30.0/23 = 11001000 00010111 00011110 00000000

What should we do if organization 1 decides to switch to ISPs-R-Us?



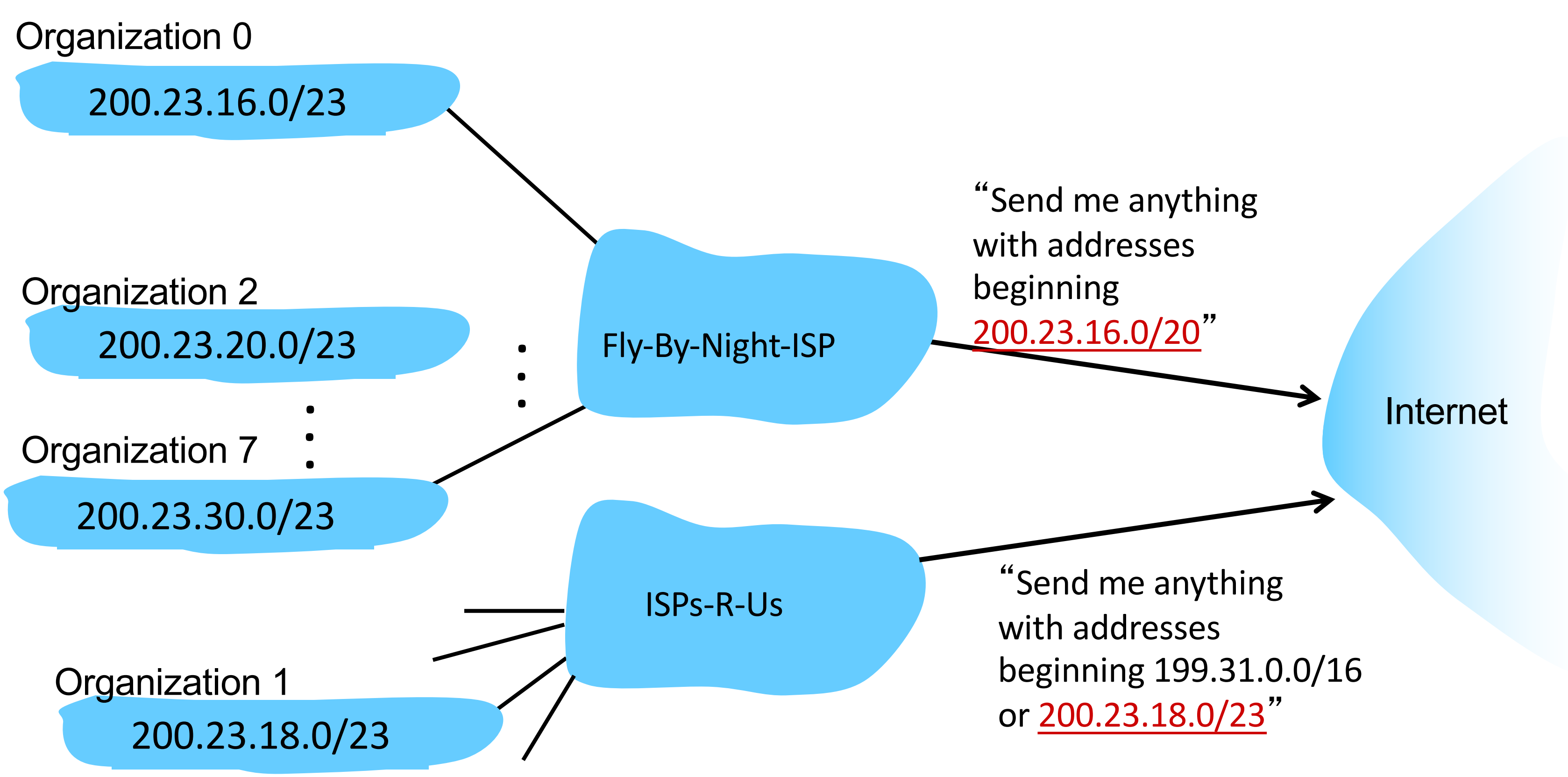
What should we do if organization 1 decides to switch to ISPs-R-Us?



A. Move 200.23.18.0/23 to ISPs-R-Us (and break up Fly-By-Night's /20 block).
B. Give new addresses to Organization 1 (and force them to change all their addresses).
C. Some other solution.

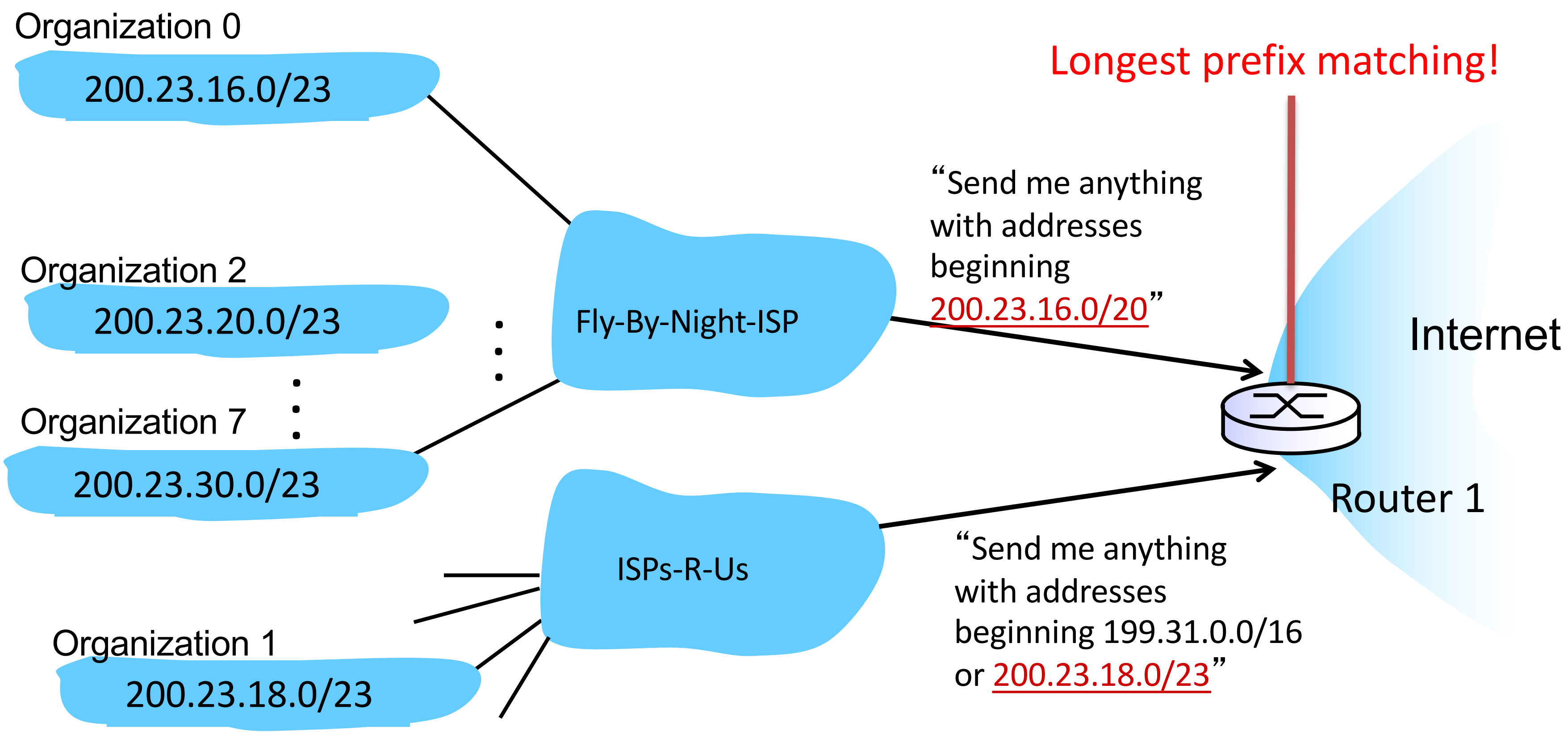
Hierarchical addressing: More Specific Routes

ISPs-R-Us has a more specific route to Organization 1



Hierarchical addressing: More Specific Routes

ISPs-R-Us has a more specific route to Organization 1

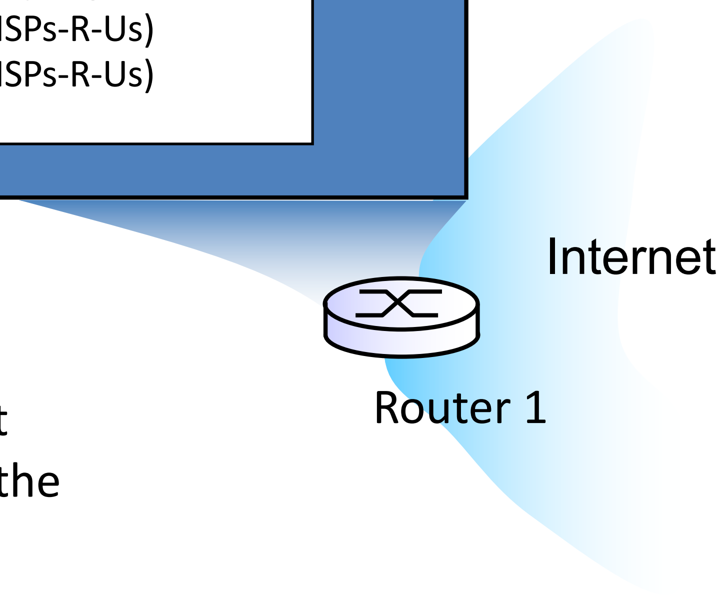


Longest Prefix Matching at Router 1

routing algorithm

local forwarding table

dest address	output link
200.23.16.0/20 = 11001000 00010111 00010000 00000000	(to Fly-by-Night ISP)
199.31.0.0/16 = 11000111 00011111 00000000 00000000	(to ISPs-R-Us)
200.23.18.0/23 = 11001000 00010111 00010010 00000000	(to ISPs-R-Us)



Now, when an incoming packet addressed with destination address 200.23.18.5 arrives – this address belongs to Organization 1 and the packet will be matched using longest prefix matching and will be routed to ISPs-R-Us rather than the Fly-by-Night ISP.

How does an end host get an IP address?

- Static IP: hard-coded
  - Windows: control-panel->network->configuration->tcp/ip->properties
  - UNIX: /etc/rc.config

- DHCP: Dynamic Host Configuration Protocol: dynamically get address from as server
  - "plug-and-play"