ECSE-4670: Computer Communication Networks (CCN)

Network Layer

Shivkumar Kalyanaraman: shivkuma@ecse.rpi.edu

Biplab Sikdar: sikdab@rpi.edu

http://www.ecse.rpi.edu/Homepages/shivkuma

Rensselaer

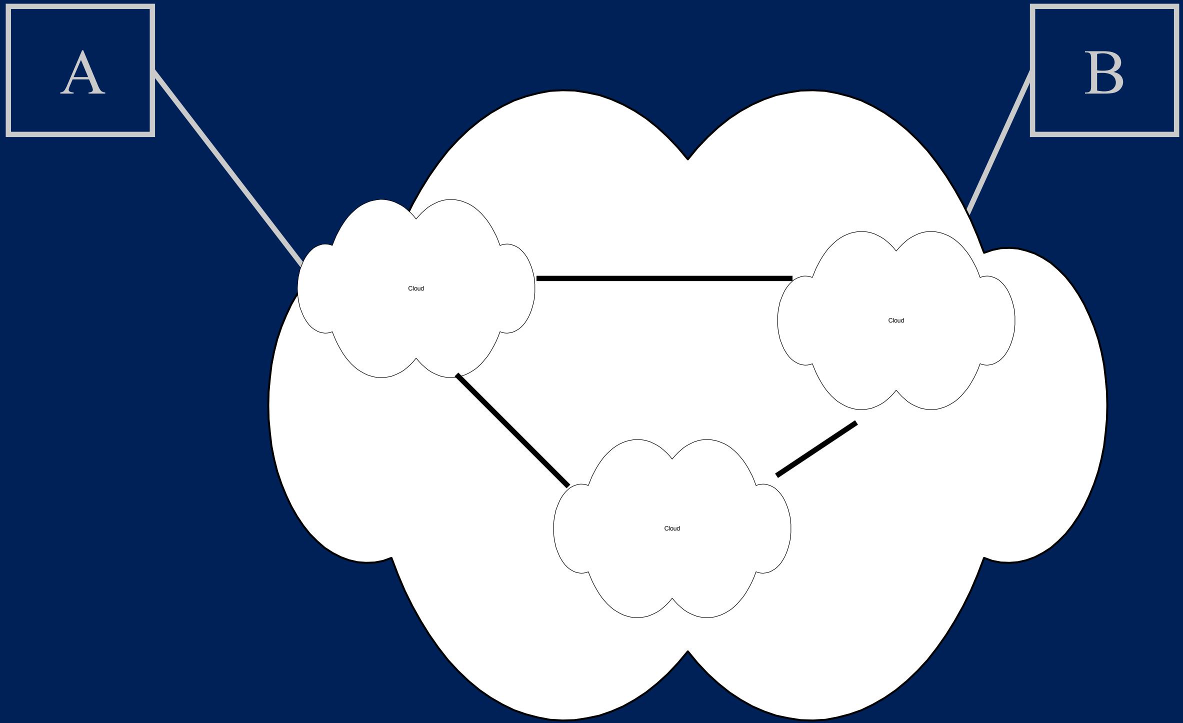
RSVP

- understand principles behind network layer services:
  - Internetworking concepts
  - The network layer
  - IP
  - routing (path selection)
  - how a router works

- Internetworking
- network layer services
- IP
- routing principle: path selection
- hierarchical routing
- Internet routing protocols reliable transfer
  - intra-domain
  - inter-domain
- what's inside a router?

The Internetworking Problem

- **Two nodes communicating across a "*network of networks*"…**
  - **How to <u>transport</u> packets through this heterogeneous mass ?**



The Internetworking Problem

- Problems: ***heterogeneity and scaling***
- ***Heterogeneity***:
  - How to interconnect a large number of disparate *networks* ? (lower layers)
  - How to support a wide variety of *applications* ? (upper layers)
- ***Scaling:***
  - How to support a large number of *end-nodes* and *applications* in this interconnected network ?

Heterogeneity: Solutions

- ***Translation*** (eg: bridges): specify a separate mapping between every pair of protocols
  - (+) No software changes in networks required.
  - (–) Need to specify N mappings when a new lower layer protocol is added to the list
  - (–) When many networks, subset = 0
  - (–) Mapping may be asymmetric

Heterogeneity: Solutions

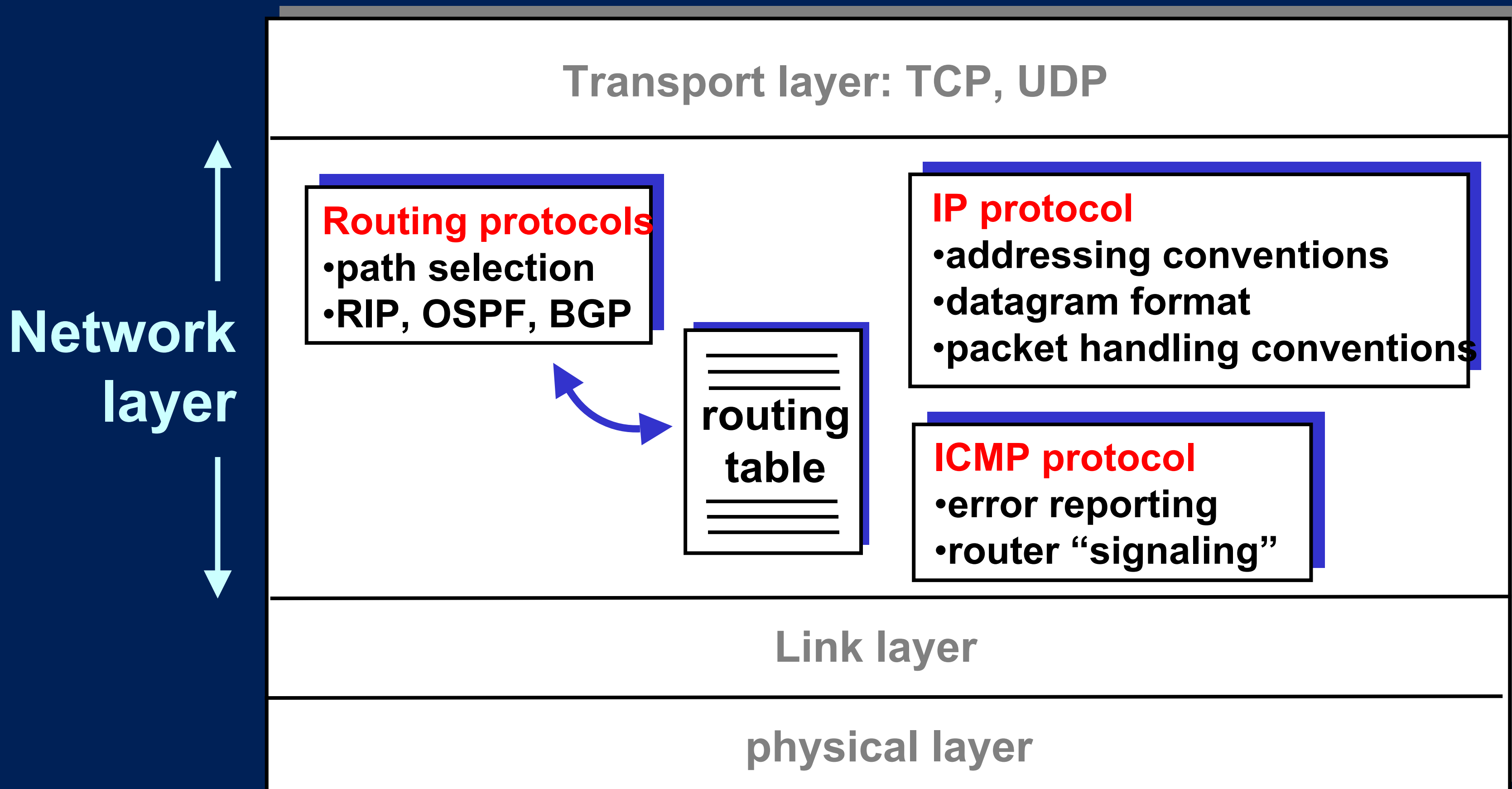
– *Overlay model:* Define a *new* protocol (IP) and map all networks to IP
  - (+) Require *only one mapping* (IP -> new protocol) when a new protocol is added
  - (+) *Global address space* can be created for universal addressibility and scaling
  - (–) Requires *changes* in lower networks (eg: protocol type field for IP)

Heterogeneity: Solutions

- (–) IP has to be ***necessarily simple*** else mapping will be hard.
  - Even in its current form mapping IP to ATM has proven to be really hard.
  - Basis for "best-effort" forwarding
- (–) ***Protocol mapping infrastructure*** needed: address hierarchy, address resolution, fragmentation

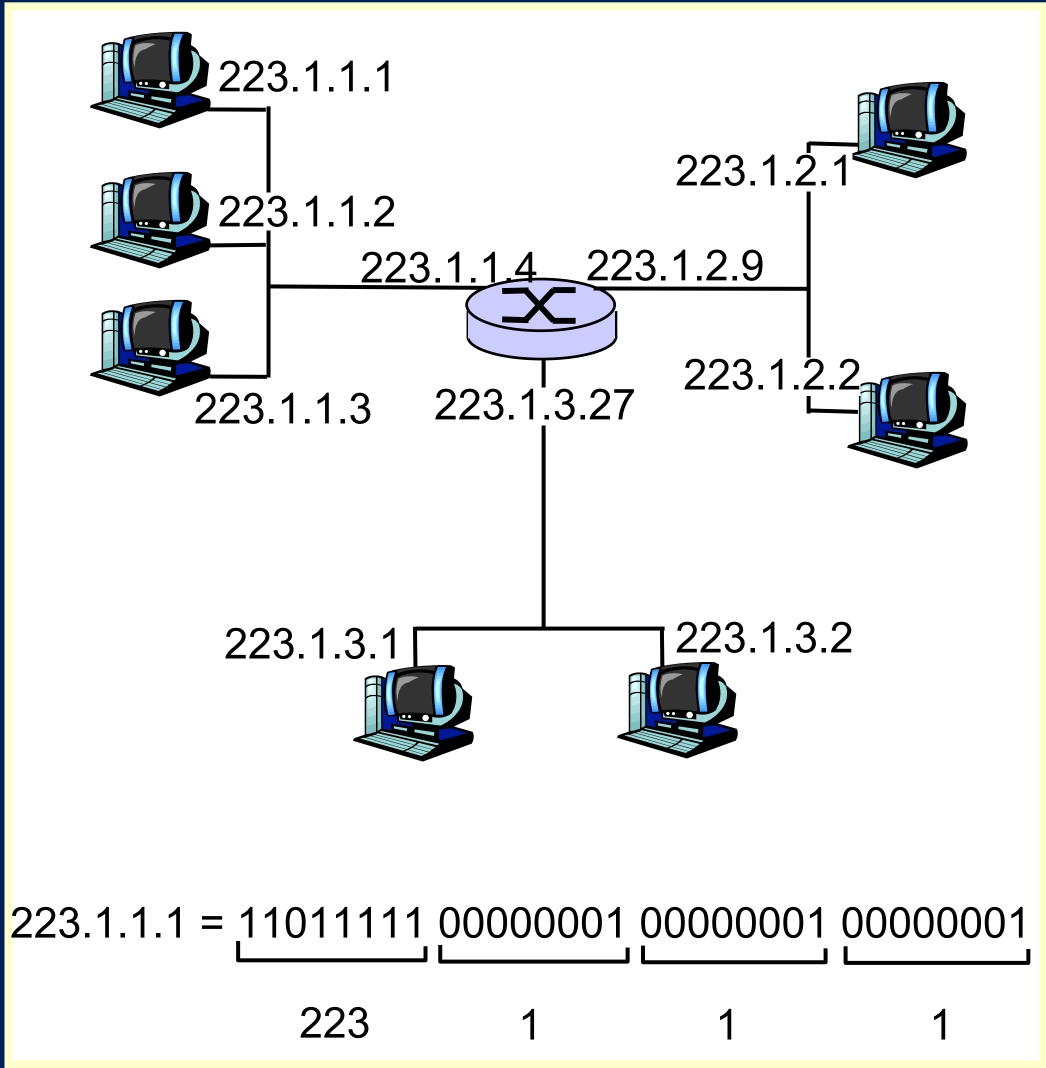
The Internet Network layer

Host, router network layer functions:



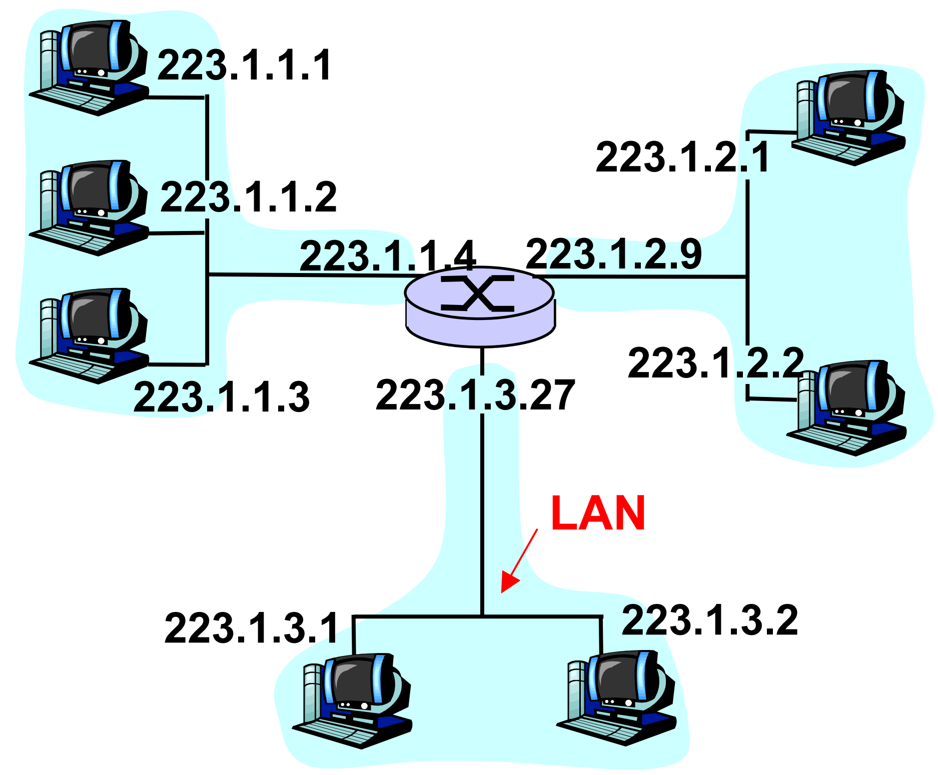
IP Addressing: introduction

- **IP address:** 32-bit identifier for host, router *interface*
- ***interface:*** connection between host, router and physical link
  - router's typically have multiple interfaces
  - host may have multiple interfaces
  - IP addresses associated with interface, not host, router



IP Addressing - 1

- **IP address:**
  - network part (high order bits)
  - host part (low order bits)
- ***What's a network ?*** (from IP address perspective)
  - device interfaces with same network part of IP address
  - can physically reach each other without intervening router

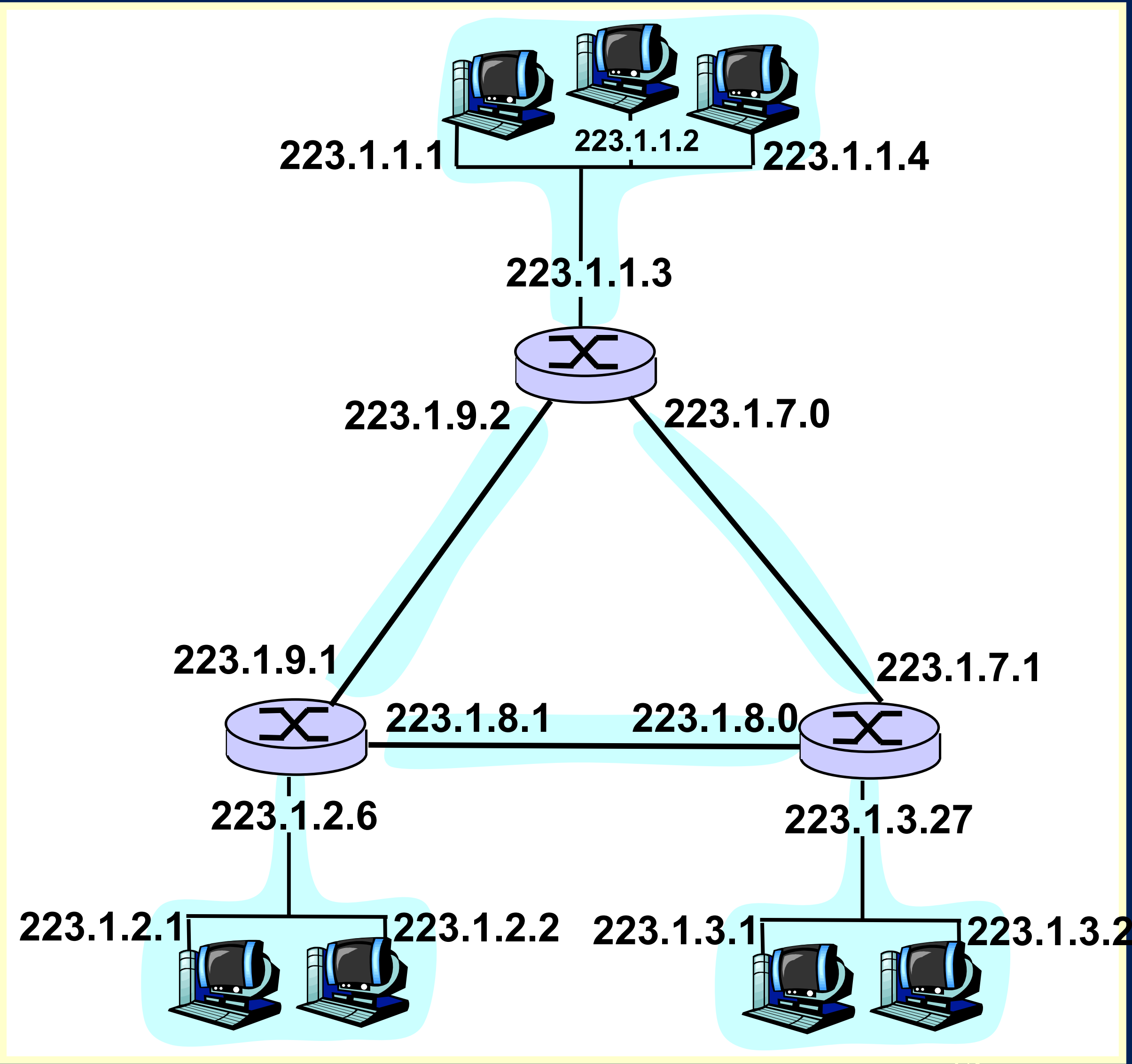


network consisting of 3 IP networks (for IP addresses starting with 223, first 24 bits are network address)

IP Addressing - 2

How to find the networks?

- Detach each interface from router, host
- create "islands of isolated networks

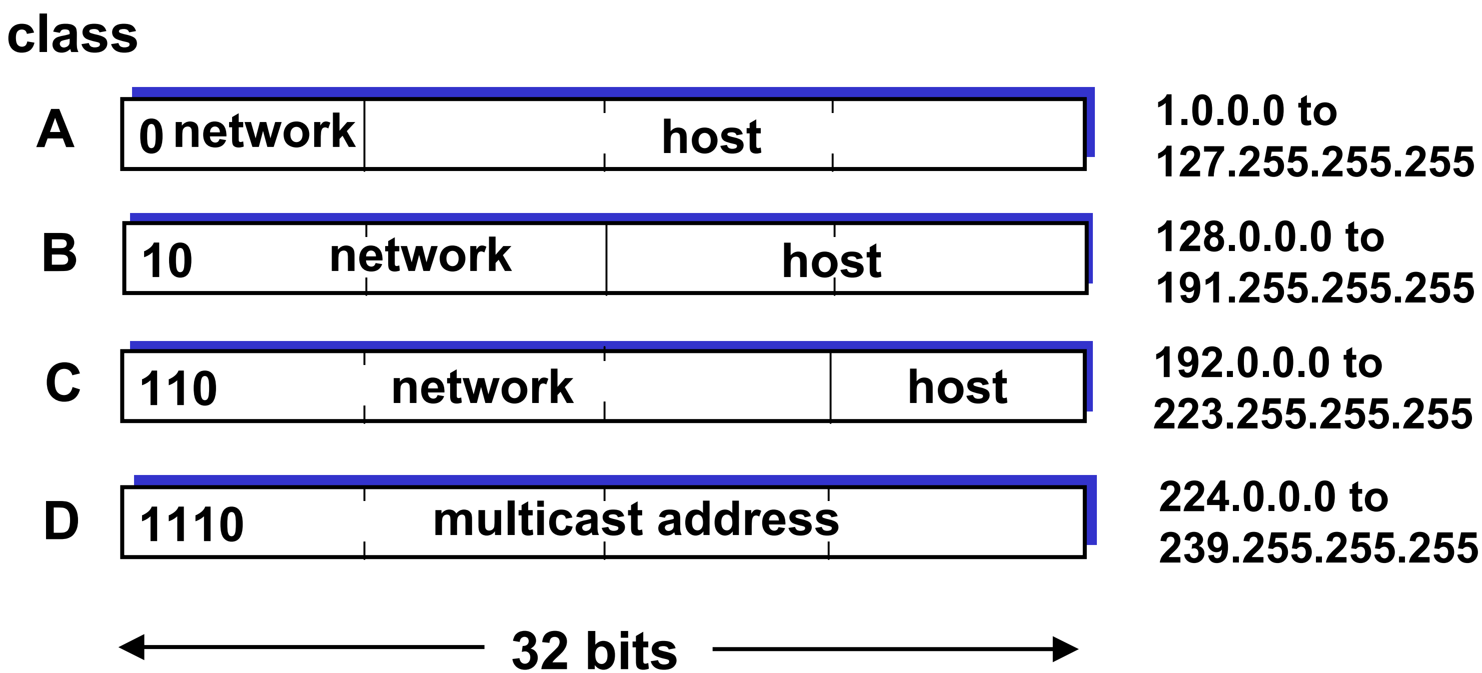


Interconnected system consisting of six networks

IP Addresses

given notion of "network", let's re-examine IP addresses:

"class-full" addressing:



Some Special IP Addresses

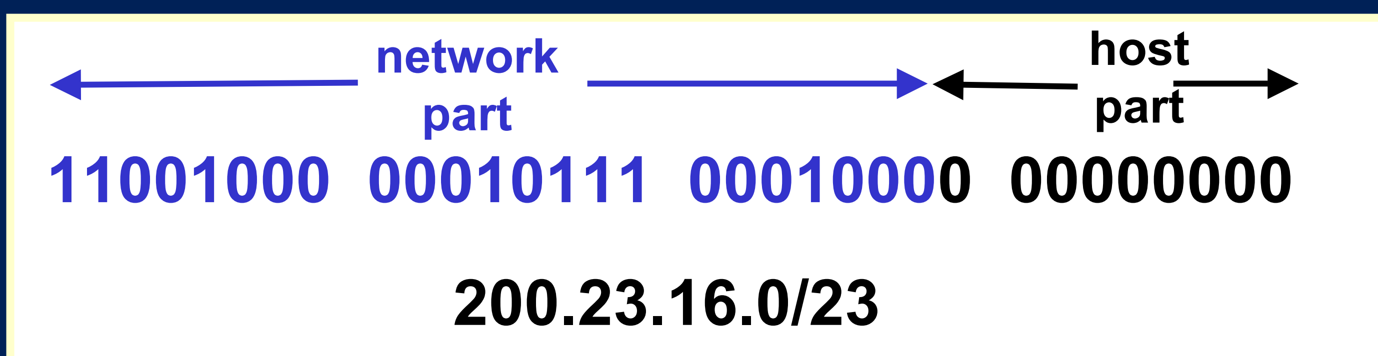
- All-0s $\Rightarrow$ This computer
- All-1s $\Rightarrow$ All hosts on this net (*limited broadcast*: *don't forward out of this net)*
- All-0 *host suffix* $\Rightarrow$ Network Address ('*0'* means '*this'*)
- All-1 *host suffix* $\Rightarrow$ All hosts on the destination net (directed broadcast).
- 127.*.*.* $\Rightarrow$ *Loopback* through IP layer

IP addressing: CIDR - 1

- classful addressing:
  - inefficient use of address space, address space exhaustion
  - e.g., class B net allocated enough addresses for 65K hosts, even if only 2K hosts in that network

IP addressing: CIDR - 2

- CIDR: Classless InterDomain Routing
  - network portion of address of arbitrary length
  - address format: a.b.c.d/x, where x is # bits in network portion of address



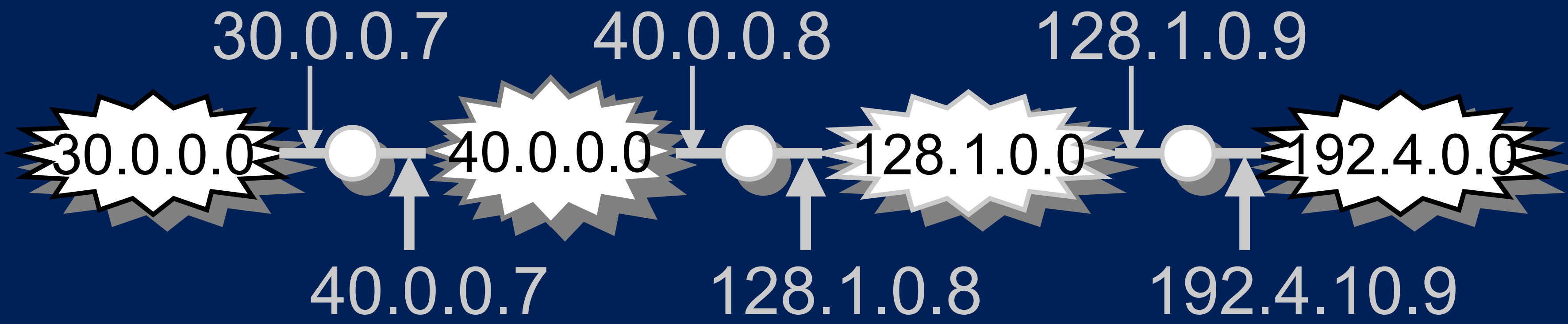
Subnet Addressing

- *External routers* need to store entries only for the "network ID"
- *Internal routers & hosts* use **subnet mask** to identify "subnet ID" and route packets between "subnets" within the "network".
- Eg: Mask: 255.255.255.0 => subnet ID = 8 bits with upto 62 hosts/subnet

Subnet Addressing (Continued)

- Route table lookup:
  - IF ((Mask[i] & Destination Addr) = = Destination[i])
    Forward to NextHop[i]
- Subnet mask can end on any bit.
- Mask must have contiguous 1s followed by contiguous zeros. Routers do not support other types of masks.

Route Table Lookup: Example



Destination	Mask	Next Hop
30.0.0.0	255.0.0.0	40.0.0.7
40.0.0.0	255.0.0.0	Deliver direct
128.1.0.0	255.255.0.0	Deliver direct
192.4.10.0	255.255.255.0	128.1.0.9

IP addresses: how to get one?

Hosts (host portion):

- hard-coded by system admin in a file
- DHCP: Dynamic Host Configuration Protocol: dynamically get address: "plug-and-play"
  - host broadcasts "DHCP discover" msg
  - DHCP server responds with "DHCP offer" msg
  - host requests IP address: "DHCP request" msg
  - DHCP server sends address: "DHCP ack" msg

IP addresses: how to get one?

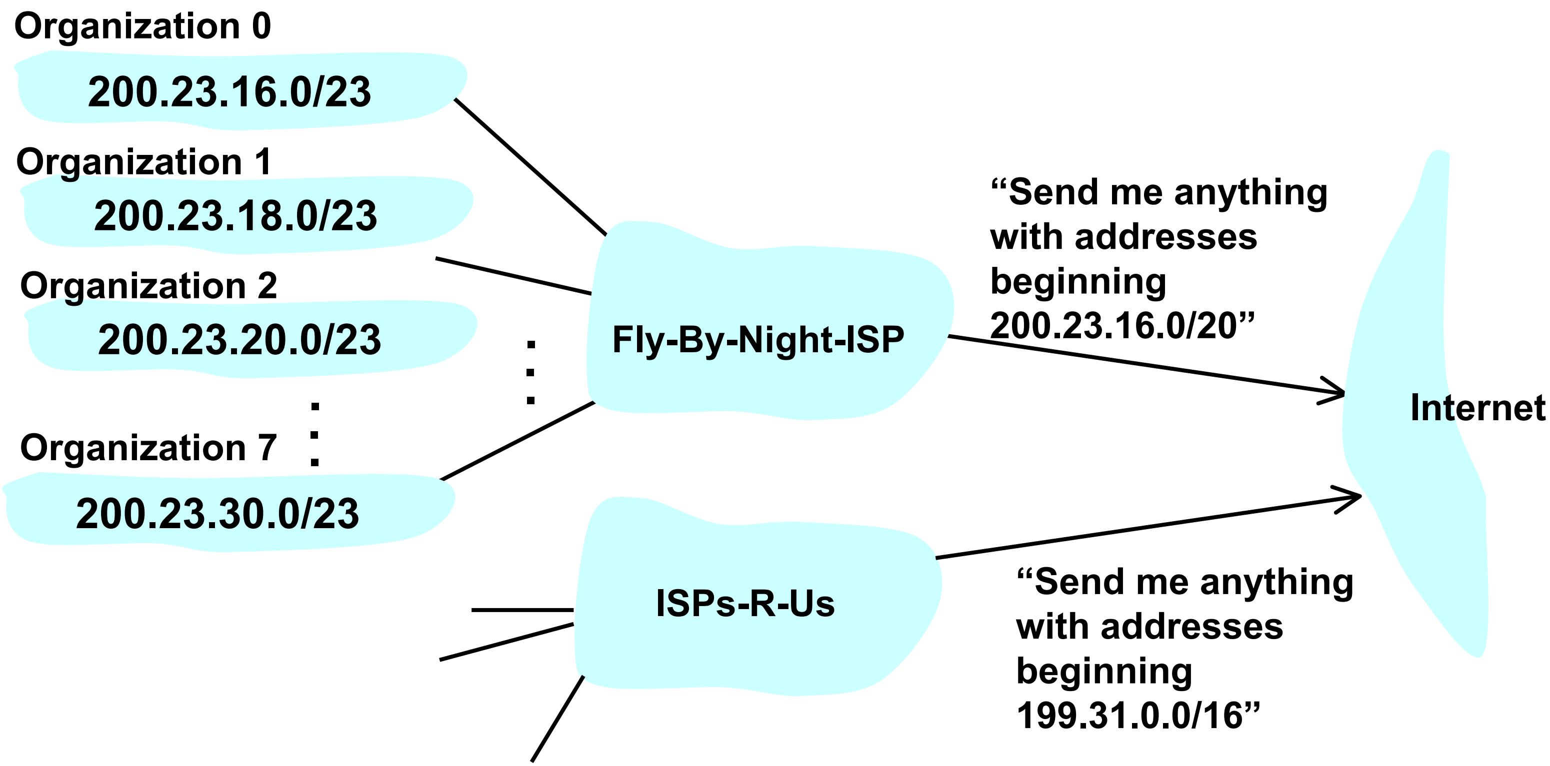
Network (network portion):

- get allocated portion of ISP's address space:

ISP's block	11001000	00010111	00010000	00000000	200.23.16.0/20
Organization 0	11001000	00010111	00010000	00000000	200.23.16.0/23
Organization 1	11001000	00010111	00010010	00000000	200.23.18.0/23
Organization 2	11001000	00010111	00010100	00000000	200.23.20.0/23
...	.....			....	....
Organization 7	11001000	00010111	00011110	00000000	200.23.30.0/23

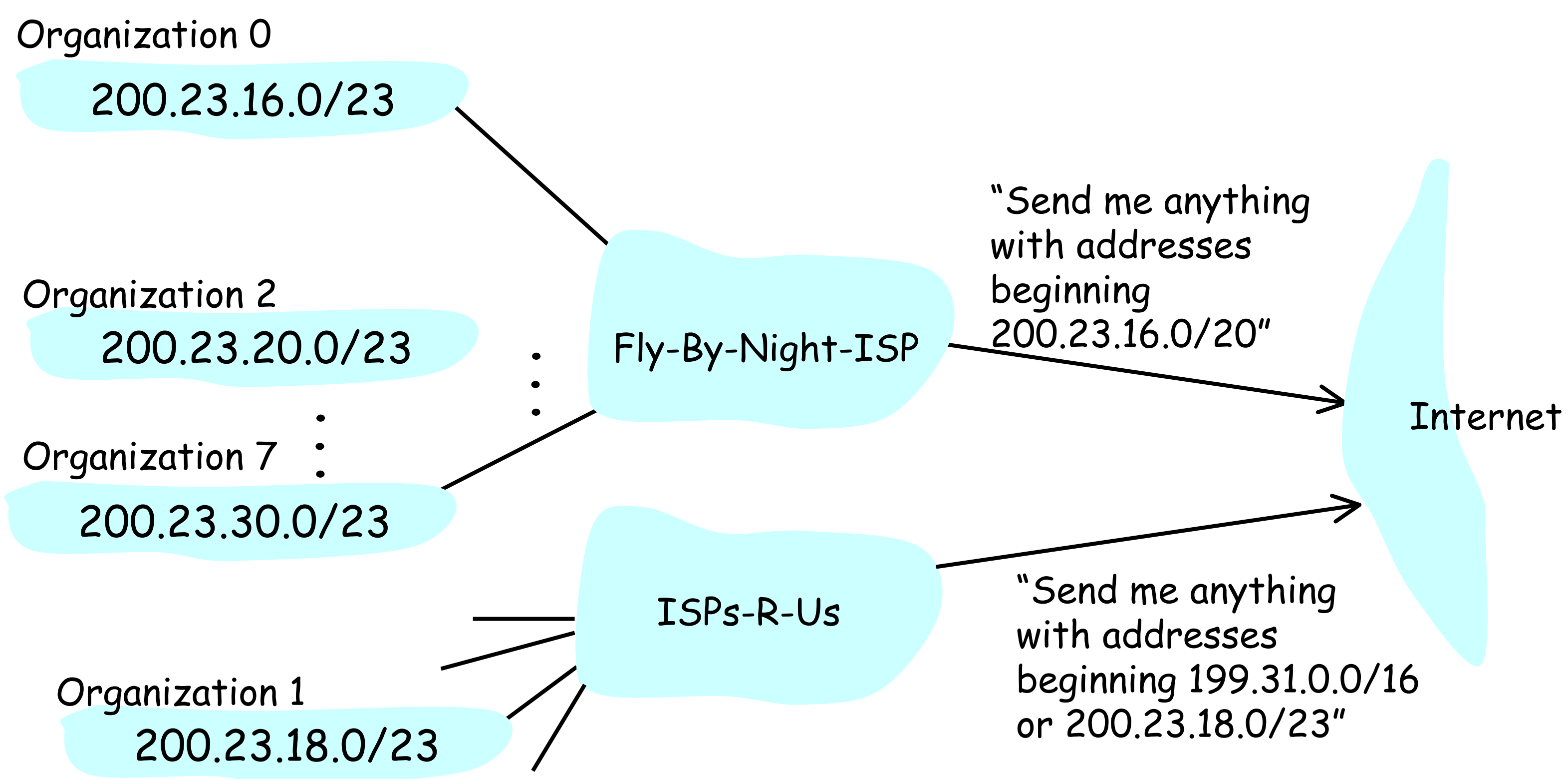
Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical addressing: more specific routes

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



IP addressing: the last word...

Q: How does an ISP get block of addresses?

A: ICANN: Internet Corporation for Assigned Names and Numbers

- allocates addresses
- manages DNS
- assigns domain names, resolves disputes

Getting a datagram from source to dest. - 1

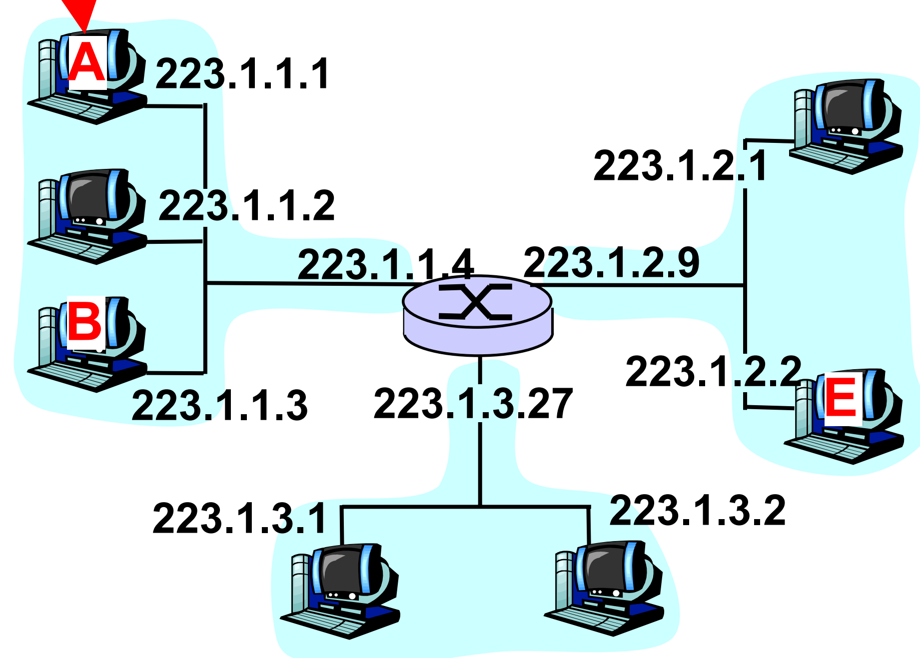
routing table in A

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2

IP datagram:

misc fields	source IP addr	dest IP addr	data

- datagram remains unchanged, as it travels source to destination
- addr fields of interest here



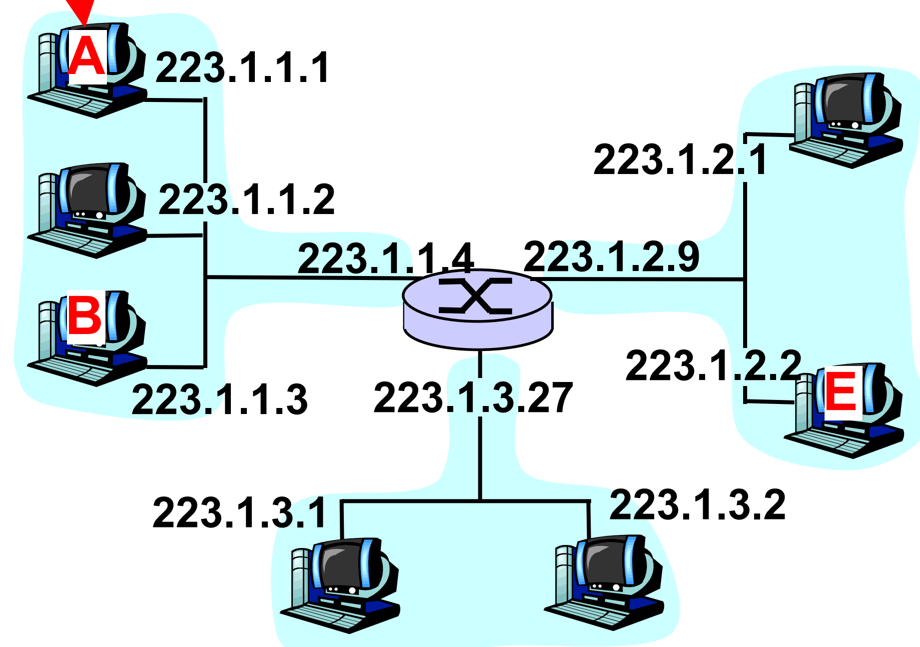
Getting a datagram from source to dest. - 2

misc fields	223.1.1.1	223.1.1.3	data

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2

Starting at A, given IP datagram addressed to B:

- look up net. address of B
- find B is on same net. as A
- link layer will send datagram directly to B inside link-layer frame
  - B and A are directly connected



Getting a datagram from source to dest. - 3

misc fields	223.1.1.1	223.1.2.2	data

Dest. Net.	next router	Nhops
223.1.1		1
223.1.2	223.1.1.4	2
223.1.3	223.1.1.4	2

Starting at A, dest. E:

look up network address of E

E on *different* network

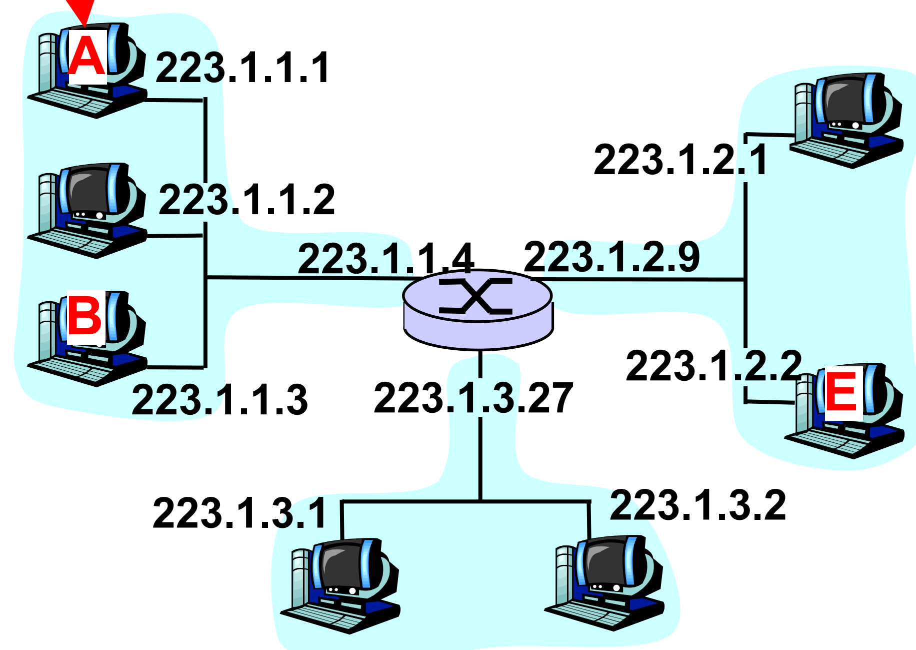
- **A, E not directly attached**

routing table: next hop router to E is 223.1.1.4

link layer sends datagram to router 223.1.1.4 inside link-layer frame

datagram arrives at 223.1.1.4

continued…..



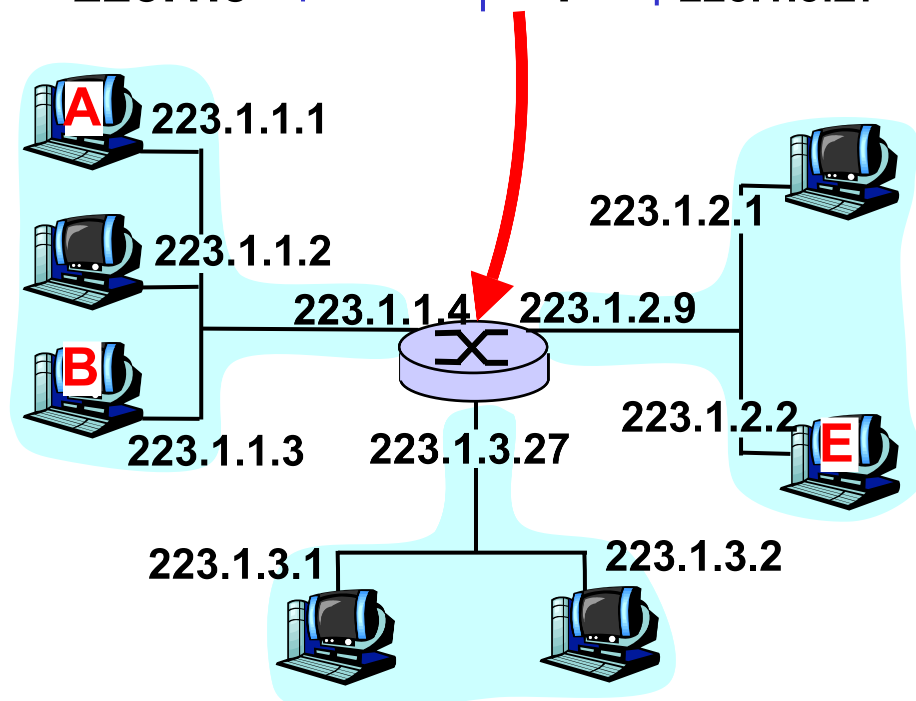
Getting a datagram from source to dest. – 4

misc fields	223.1.1.1	223.1.2.2	data

Dest. network	next router	Nhops	interface
223.1.1	-	1	223.1.1.4
223.1.2	-	1	223.1.2.9
223.1.3	-	1	223.1.3.27

Arriving at 223.1.4, destined for 223.1.2.2

- look up network address of E
- E on *same* network as router's interface 223.1.2.9
  - router, E directly attached
- link layer sends datagram to 223.1.2.2 inside link-layer frame via interface 223.1.2.9
- datagram arrives at 223.1.2.2!!! (hooray!)



IP Features

- Connectionless service
- Addressing
- Data forwarding
- Fragmentation and reassembly
- Supports variable size datagrams
- Best-effort delivery
- Provides only “Send” and “Delivery” services. Error and control messages generated by Internet Control Message Protocol (ICMP)

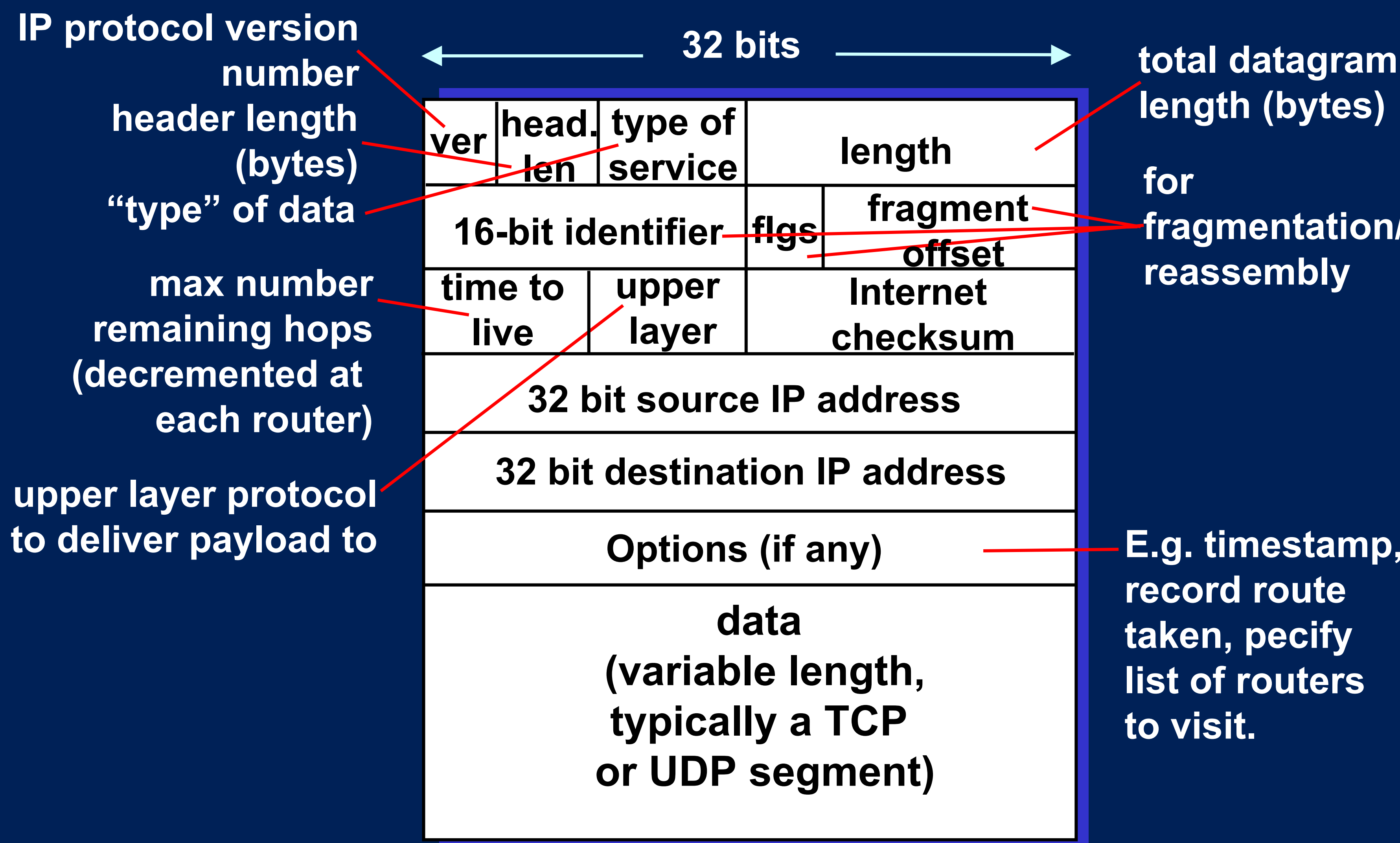
What IP does NOT provide

- End-to-end data reliability & flow control (done by TCP or application layer protocols)
- Sequencing of packets (like TCP)
- Error detection in payload (TCP, UDP or other transport layers)
- Error reporting (ICMP)

What IP does NOT provide (Continued)

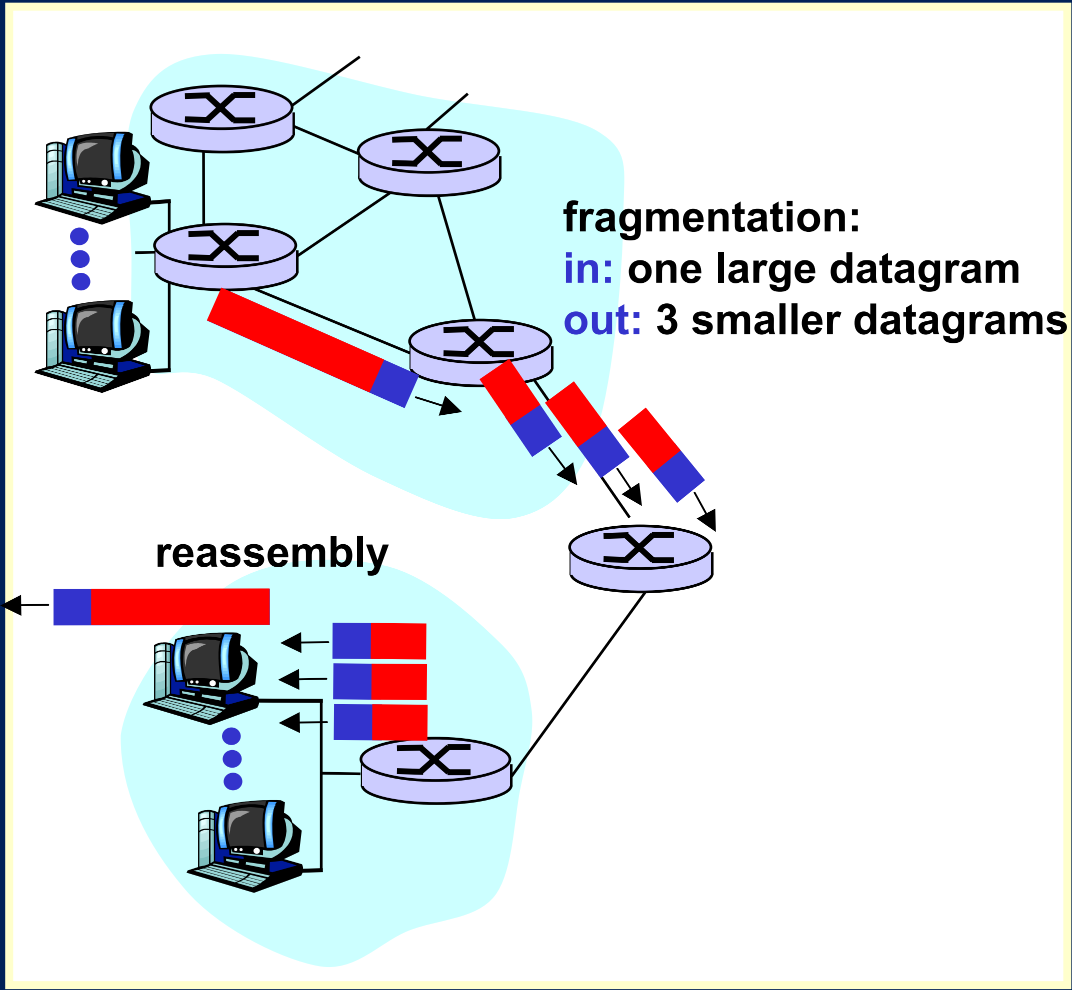
- Setting up route tables (RIP, OSPF, BGP etc)
- Connection setup (it is connectionless)
- Address/Name resolution (ARP, RARP, DNS)
- Configuration (BOOTP, DHCP)
- Multicast (IGMP, MBONE)

IP datagram format

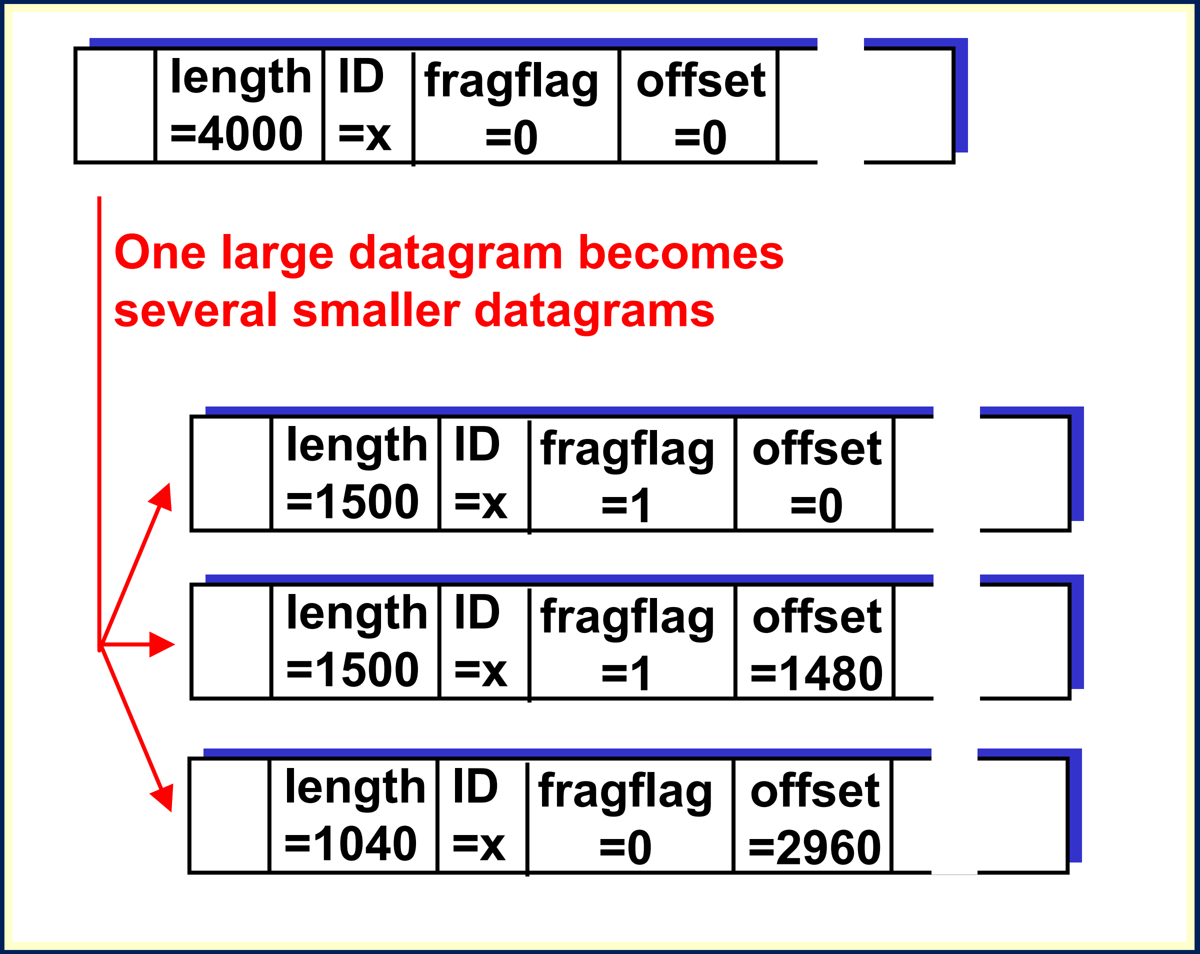


IP Fragmentation & Reassembly - 1

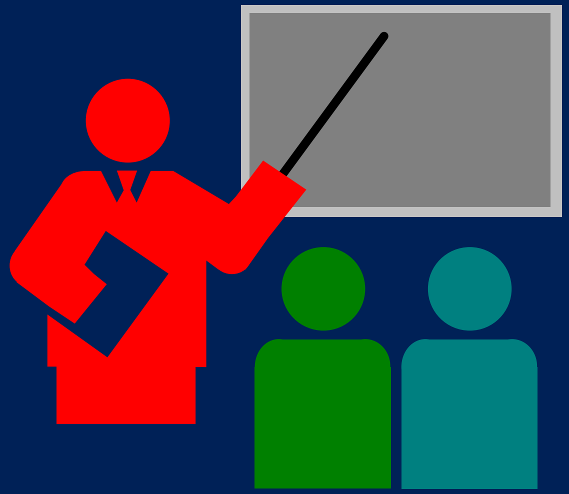
- network links have MTU (max.transfer size) - largest possible link-level frame.
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify,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



IP Fragmentation and Reassembly - 2



Summary



- Internet architectural principles
- IP addressing and header
- Fragmentation/Reassembly, Path MTU discovery