TDTS04/11: Computer Networks

Instructor: Niklas Carlsson

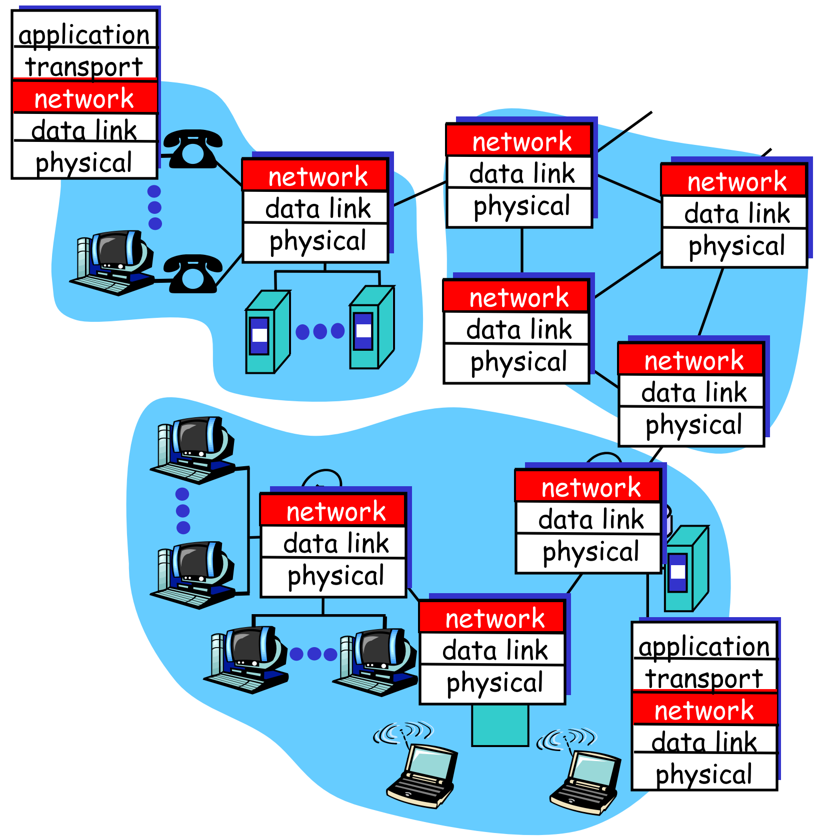
Email: niklas.carlsson@liu.se

Notes derived from "*Computer Networking: A Top Down Approach*", by Jim Kurose and Keith Ross, Addison-Wesley.

The slides are adapted and modified based on slides from the book's companion Web site, as well as modified slides by Anirban Mahanti and Carey Williamson.

Network Layer

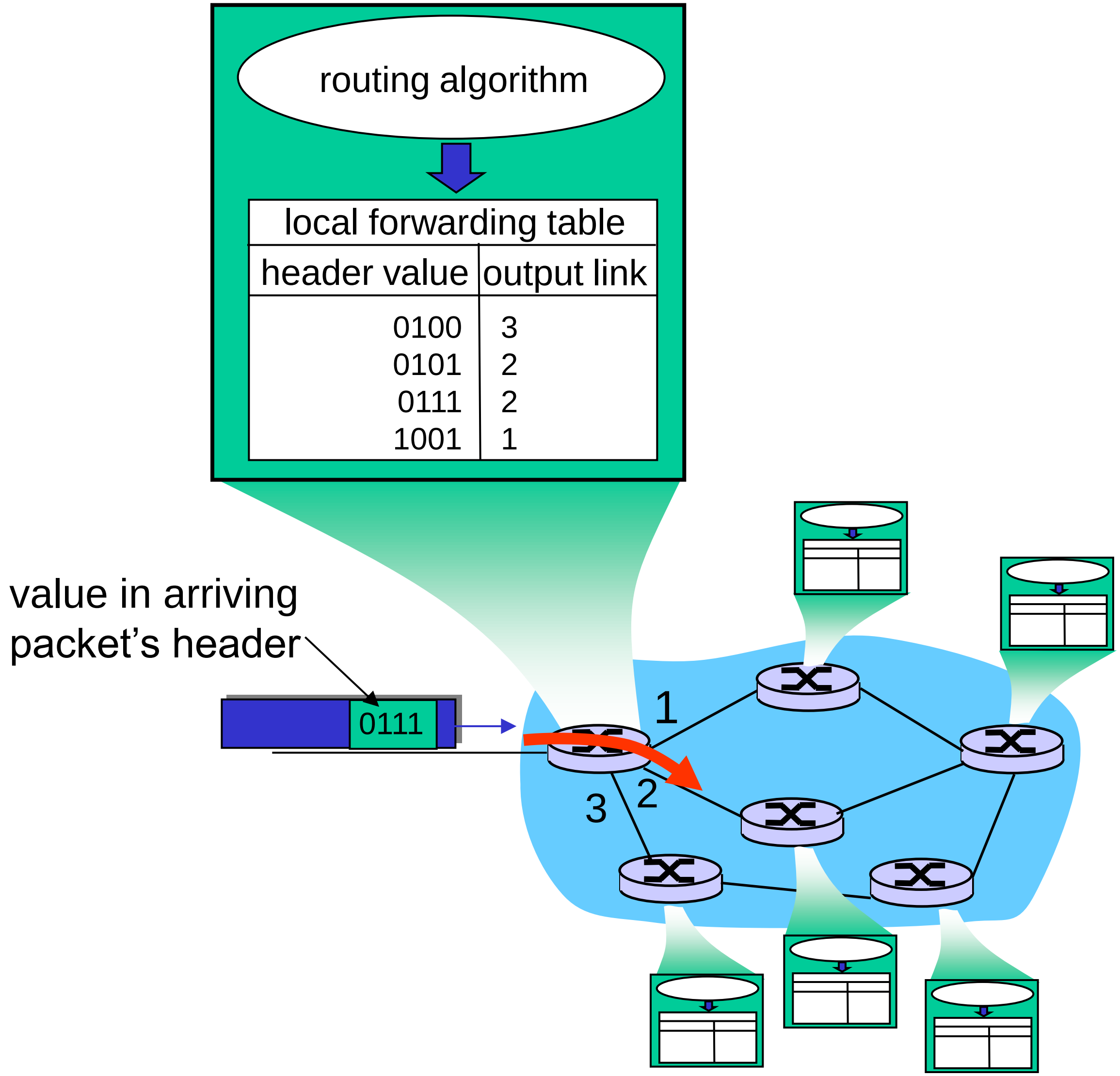
- carries segments from sending to receiving host
- on sending side, encapsulates segments into IP datagrams
- on rcv side, delivers segments to TL
- network layer protocol runs in *every* node (hosts and routers)
- router examines header fields in all IP datagrams passing through it



Key Network-Layer Functions

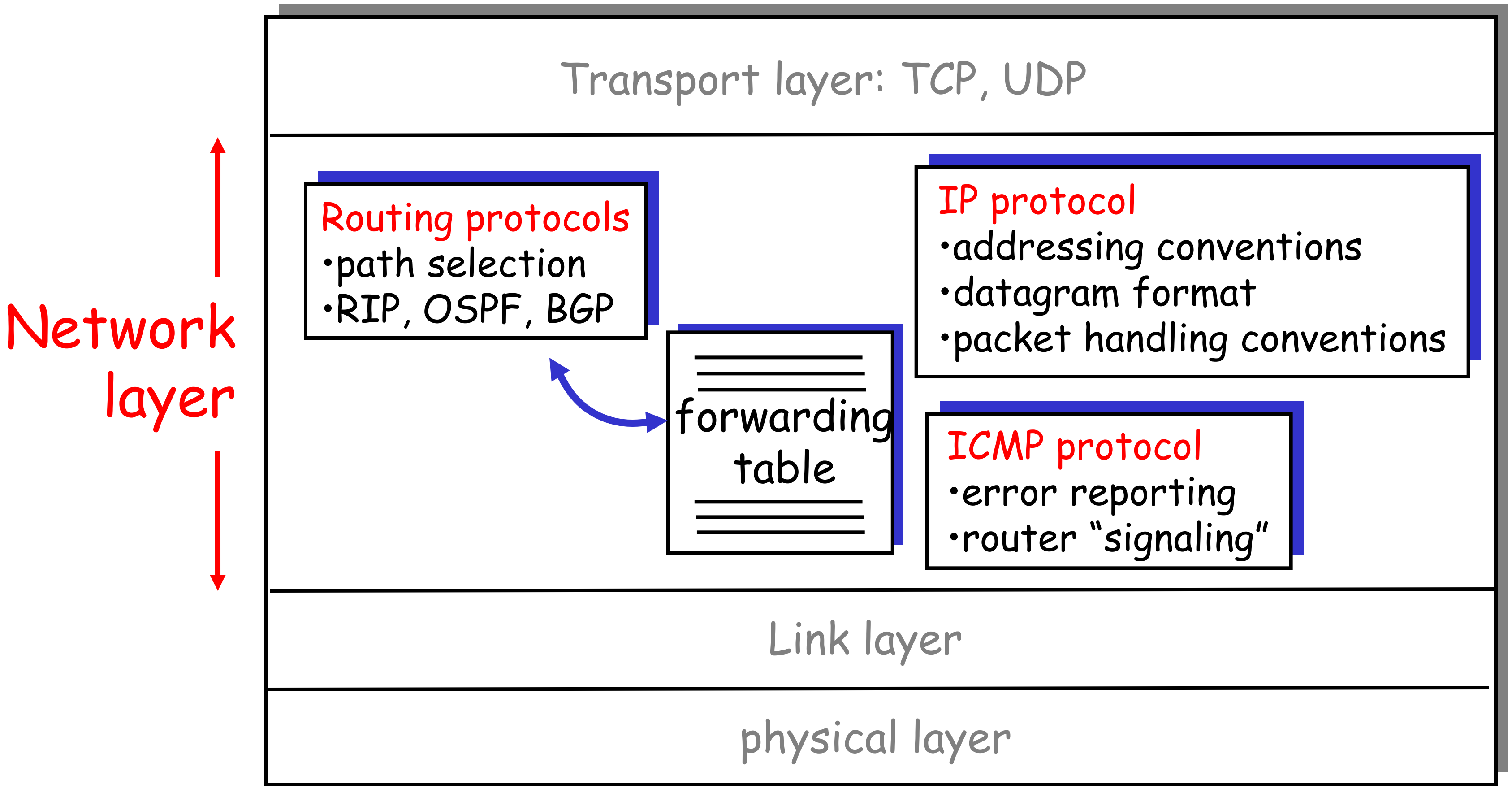
- *forwarding:* move packets from router's input to appropriate router output
- *routing:* determine the path taken by packets as they flow from a sender to a receiver
  - *Routing algorithms* – run at routers to determine "paths";
  - Routers have a forwarding table
    - Destination address-based in Datagram networks
    - Virtual circuit number-based in VC Networks

Interplay between routing and forwarding



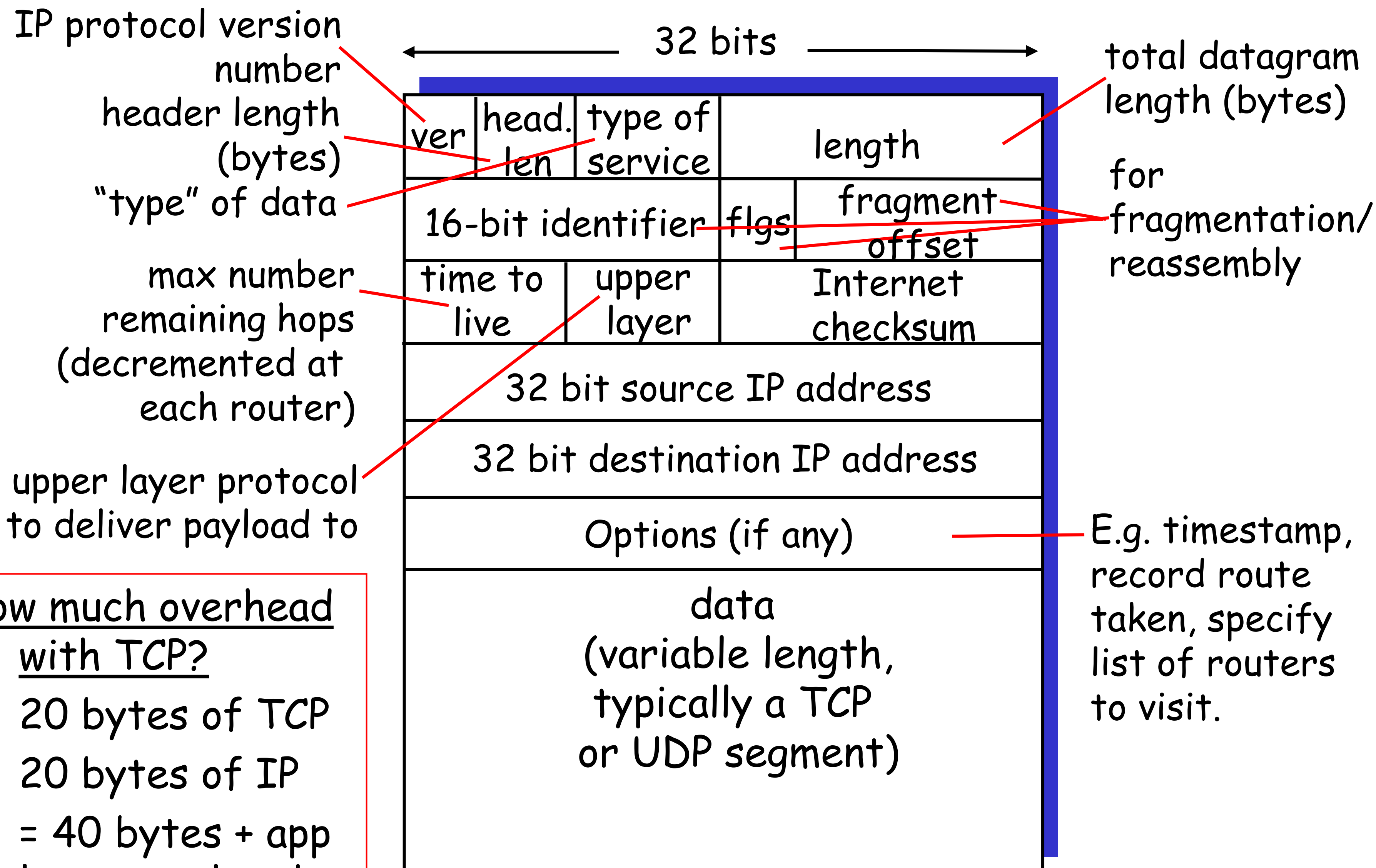
What does the Network layer consist of?

Host, router network layer functions:



The Internet Protocol (IP)

IP datagram format (IPv4)

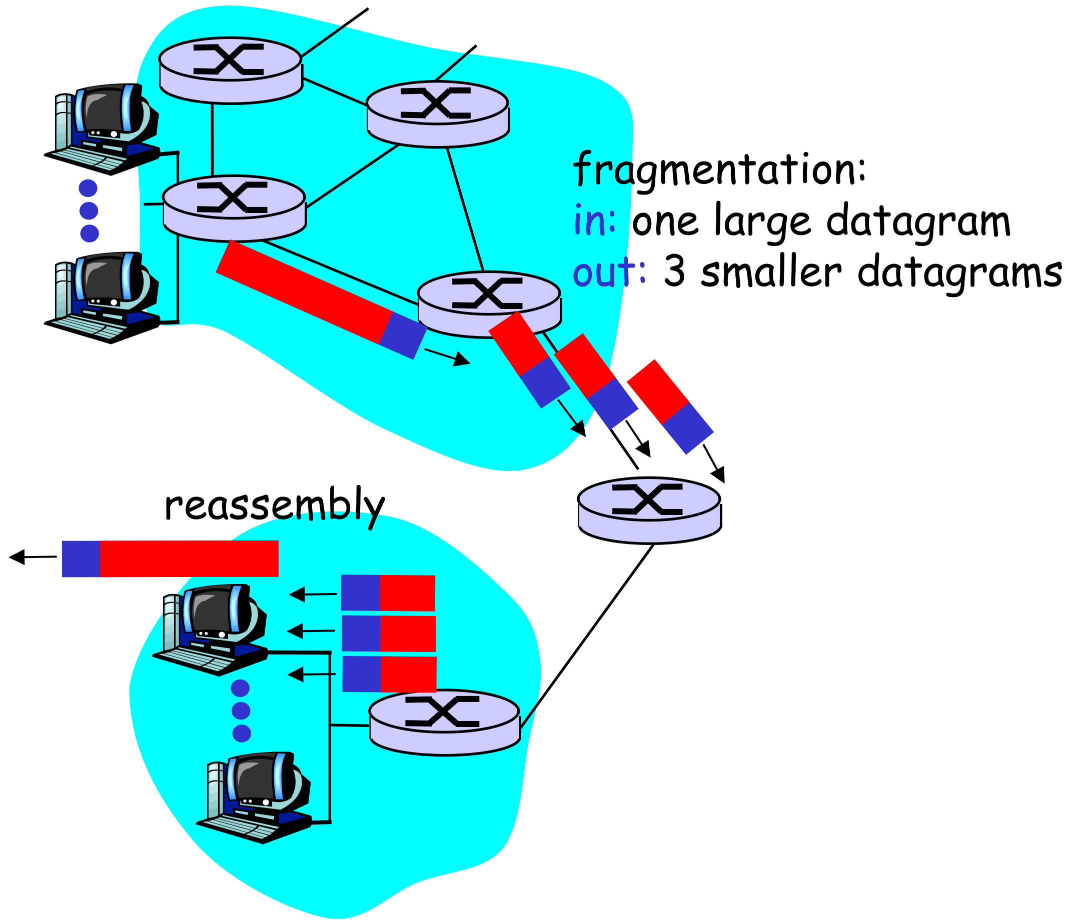


<u>how much overhead with TCP?</u>

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

IP Fragmentation & Reassembly

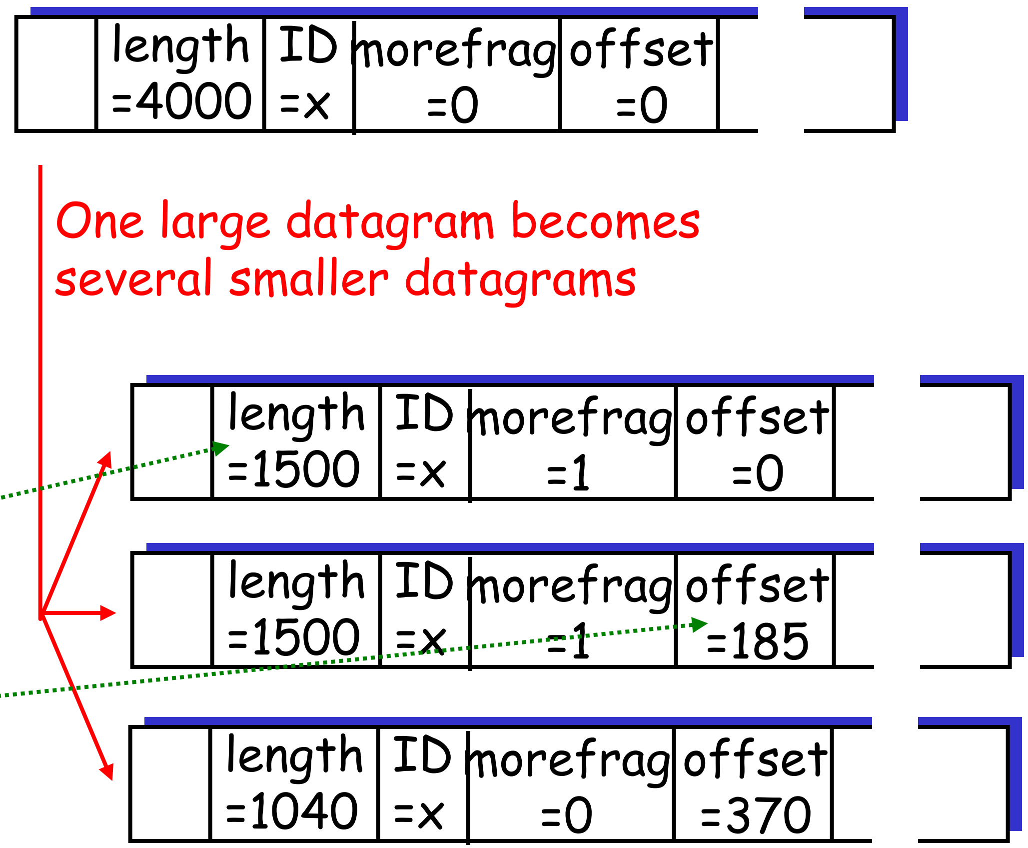
- network links have a limit on the largest possible link-level frame size permitted.
  - MTU: Maximum Transmission Unit
  - different link types, different MTUs
- large IP datagram can be divided ("fragmented") within the network (internetworking)
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



IP Fragmentation and Reassembly

Example

- 4000 byte datagram
- MTU = 1500 bytes



1480 bytes in data field

offset = 1480/8

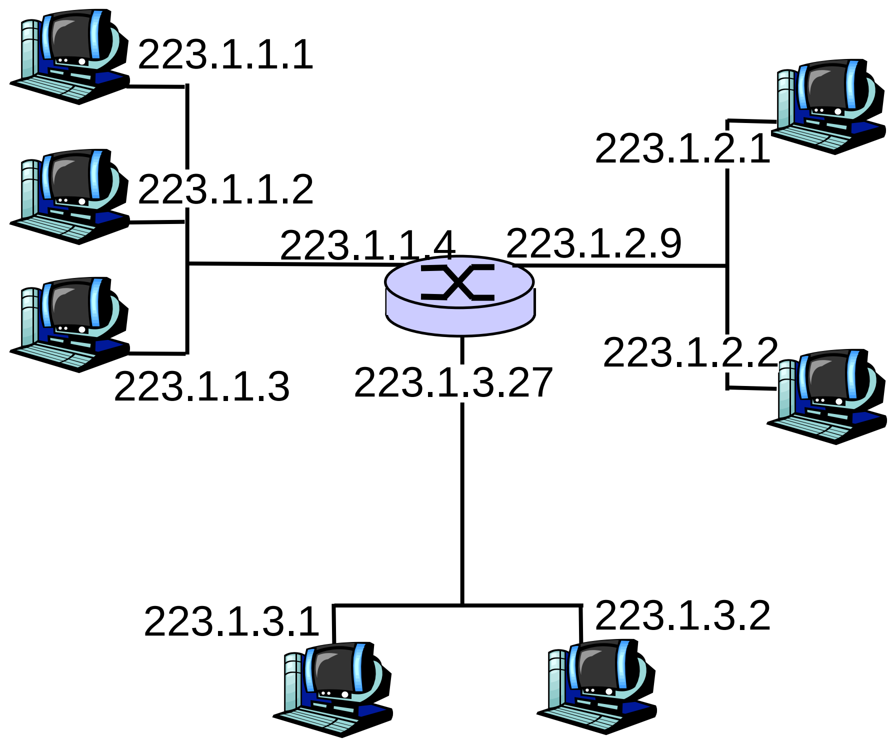
ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
  - error reporting: unreachable host, network, port, protocol
  - echo request/reply (used by ping)
- network-layer "above" IP:
  - ICMP msgs carried in IP datagrams
- ICMP message: type, code plus first 8 bytes of IP datagram causing error

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

IPv4 Addressing

- **IP address:** 32-bit identifier for host, router *interface*
- *interface:* connection between host/router and physical link
  - routers typically have multiple interfaces
  - hosts usually have one, but may have multiple interfaces (multi-homed)
  - IP addresses are associated with each interface

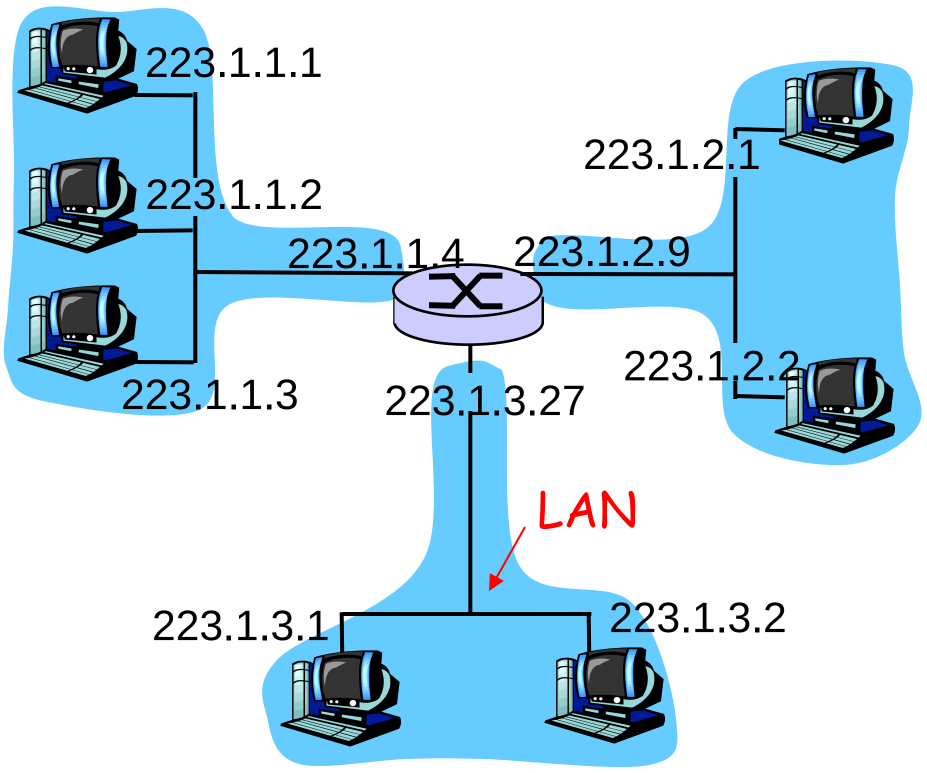


223.1.1.1 = 11011111 00000001 00000001 00000001

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



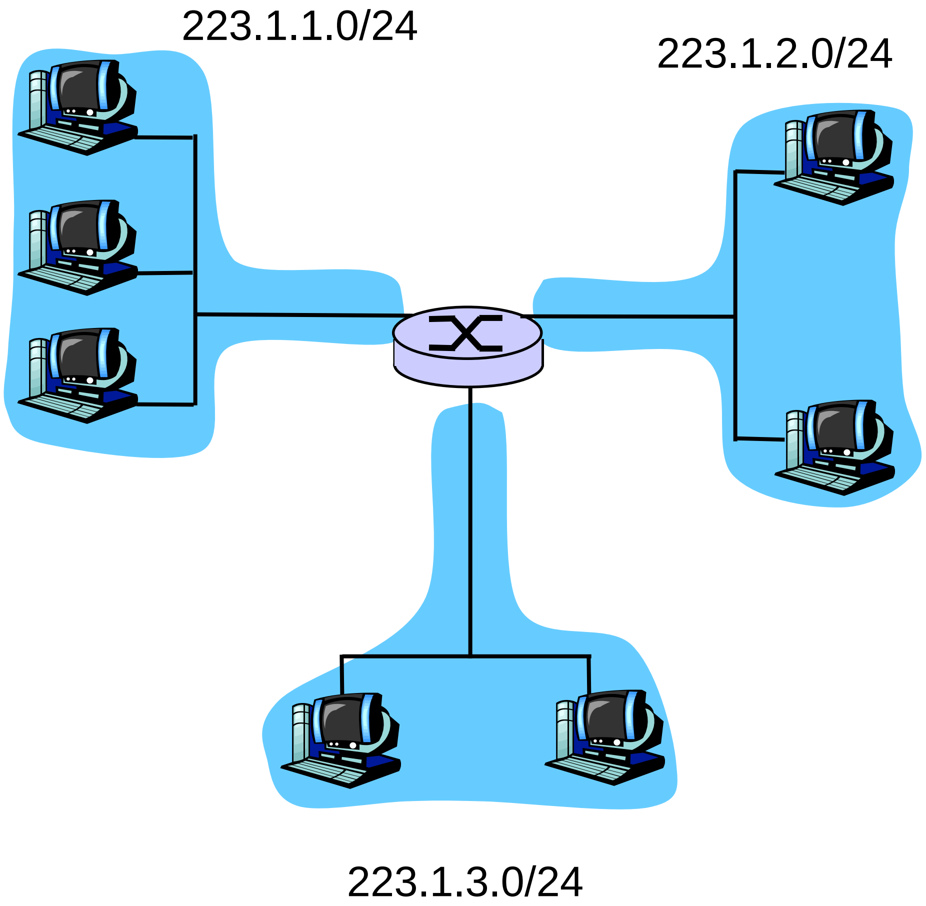
network consisting of 3 subnets

Subnets

Recipe

To determine the subnets:

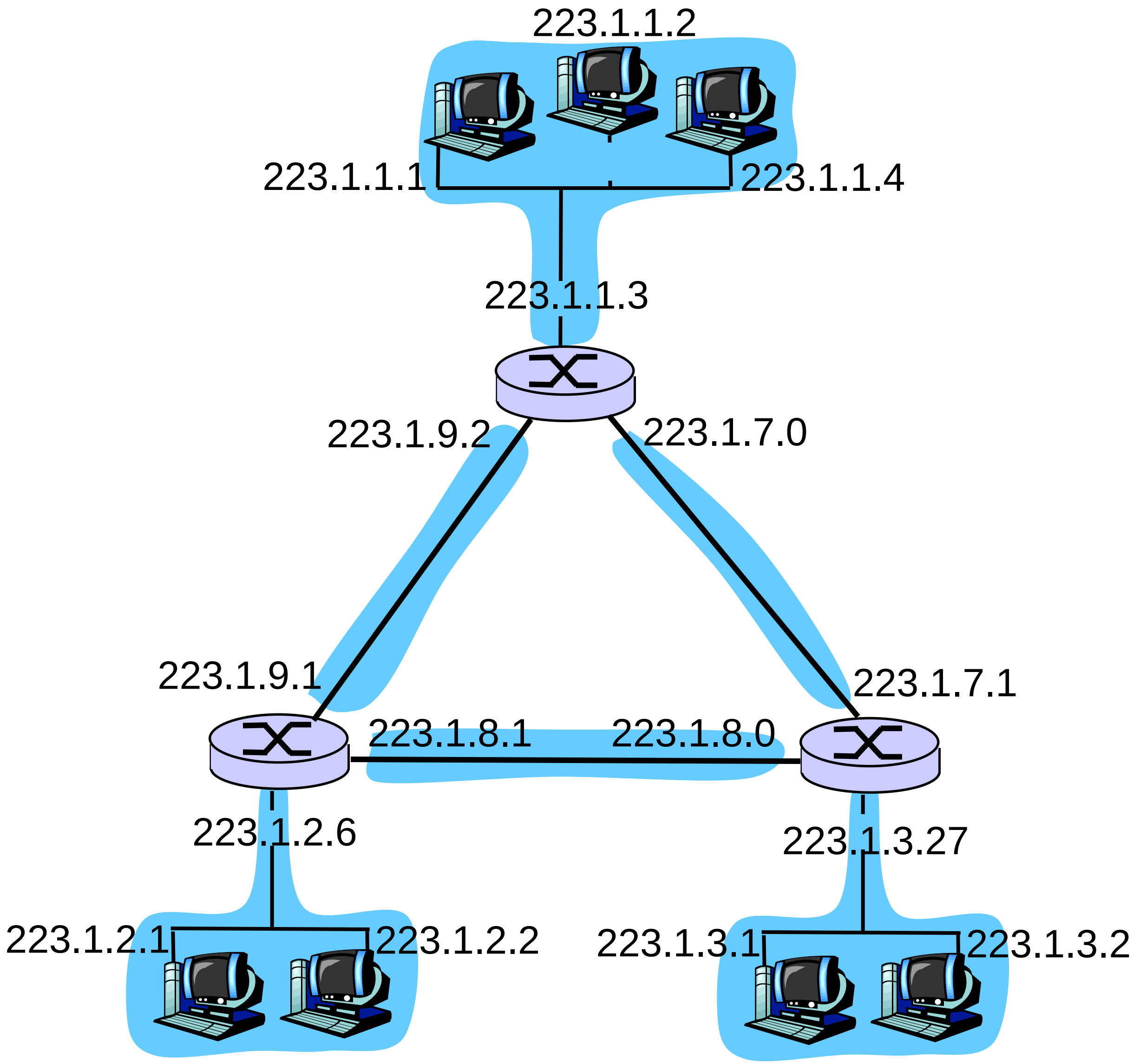
- Detach each interface from its host or router, creating islands of isolated networks.
- Each isolated network is called a subnet.



Subnet mask: /24

Subnets

How many?



Class-based Addressing

- IP addresses consist of:
  - Network part
  - Host part
- IP addresses are divided into five classes: A, B, C, D, and E.
- Problems ??

0	Network (7 bits)	Host (24 bits)

Class A

1	0	Network (14 bits)	Host (16 bits)

Class B

110	Network (21 bits)	Host (8 bits)

Class C

1110	Multicast address

Class D

1111	Future use addresses

Class E

Subnets: Motivation

- The "classful" addressing scheme proposes that the network portion of a IP address uniquely identifies one physical network.
  - Any network with more than 255 hosts needs a class B address. Class B addresses can get exhausted before we have 4 billion hosts!

- Take bits from the host number part to create a "subnet" number ("right sizing").

Addressing in the Internet

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
- Before CIDR, Internet used a class-based addressing scheme where x could be 8, 16, or 24 bits. These corrsp to classes A, B, and C resp.

← subnet part → ← host part →

11001000 00010111 00010000 00000000

200.23.16.0/23

IP addresses: how to get one?

Q: How does *host* get IP address?

IP addresses: how to get one?

Q: How does *host* get IP address?

- hard-coded by system admin in a file
  - Wintel: control-panel->network->configuration->tcp/ip->properties
  - UNIX: /etc/rc.config
- DHCP: Dynamic Host Configuration Protocol: dynamically get address from a server
  - this is becoming very popular

IP addresses: how to get one?

Q: How does *network* get subnet part of IP addr?

A:

IP addresses: how to get one?

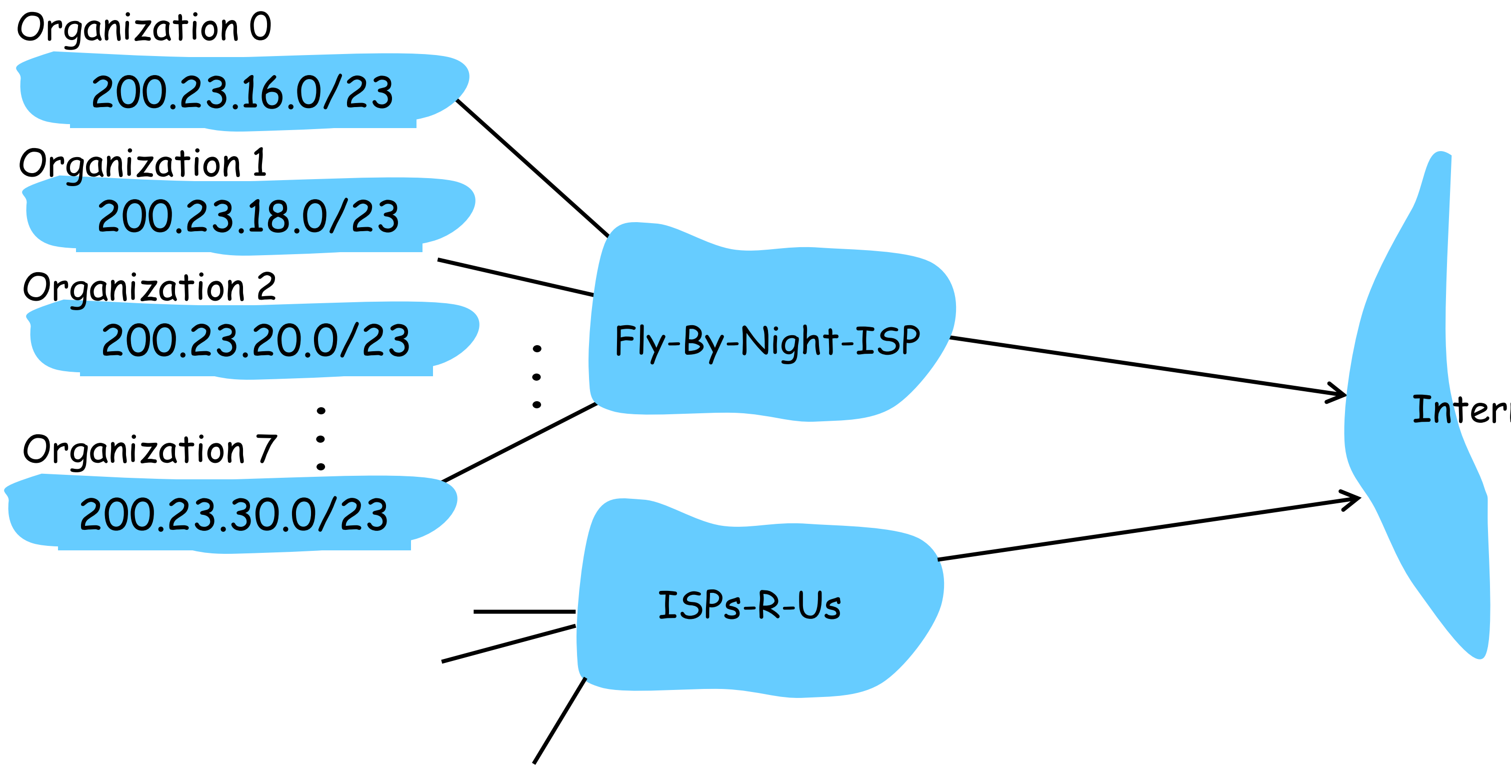
Q: How does *network* get subnet part of IP addr?

A: gets allocated portion of its provider 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

ISP has an address block; it can further divide this block into sub blocks and assign them to subscriber organizations.



IP addressing: the last word...

Q: How does an ISP get block of addresses?

A:

IP addressing: the last word...

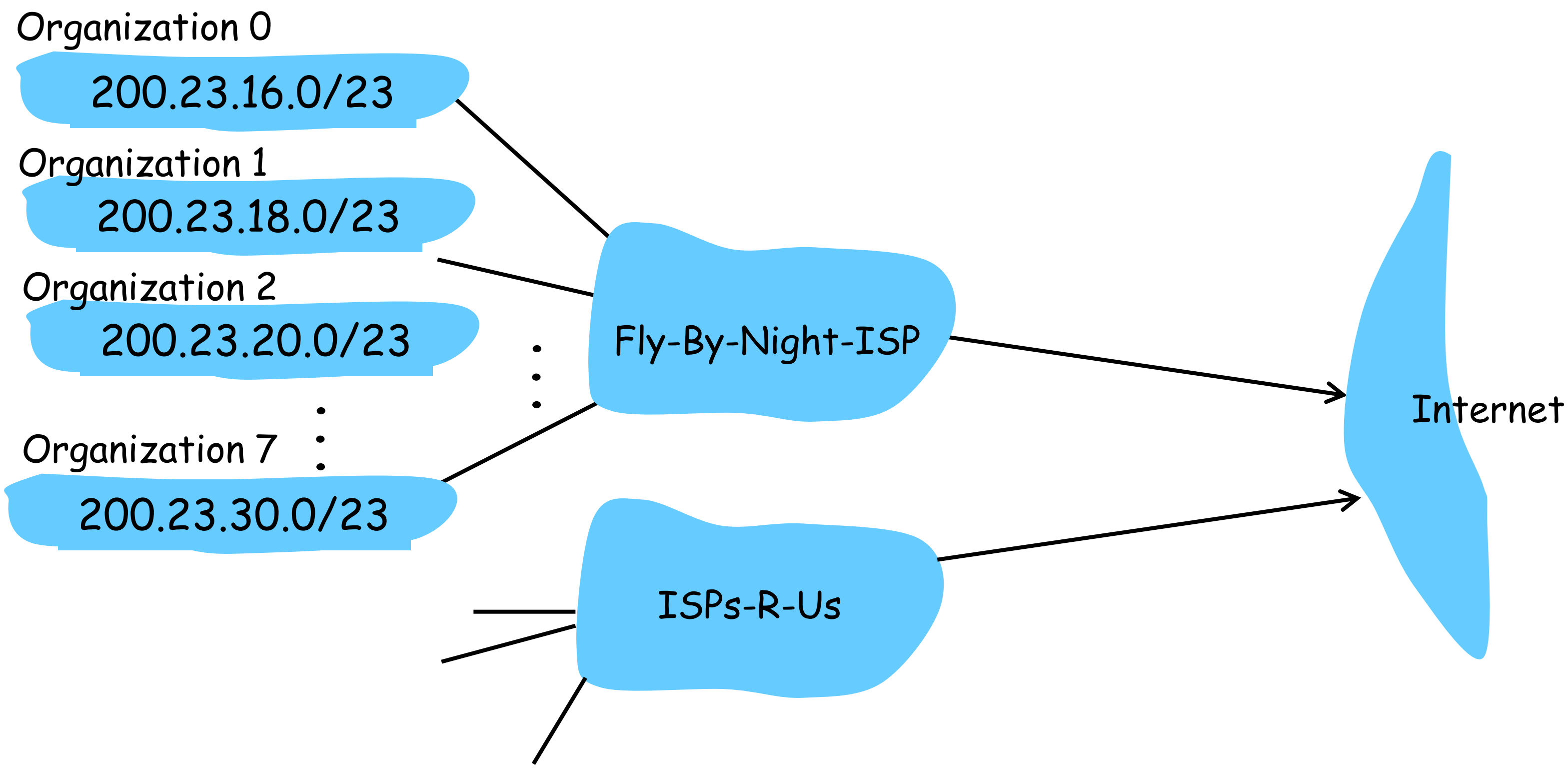
Q: How does an ISP get block of addresses?

A: ICANN: Internet Corporation for Assigned Names and Numbers

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

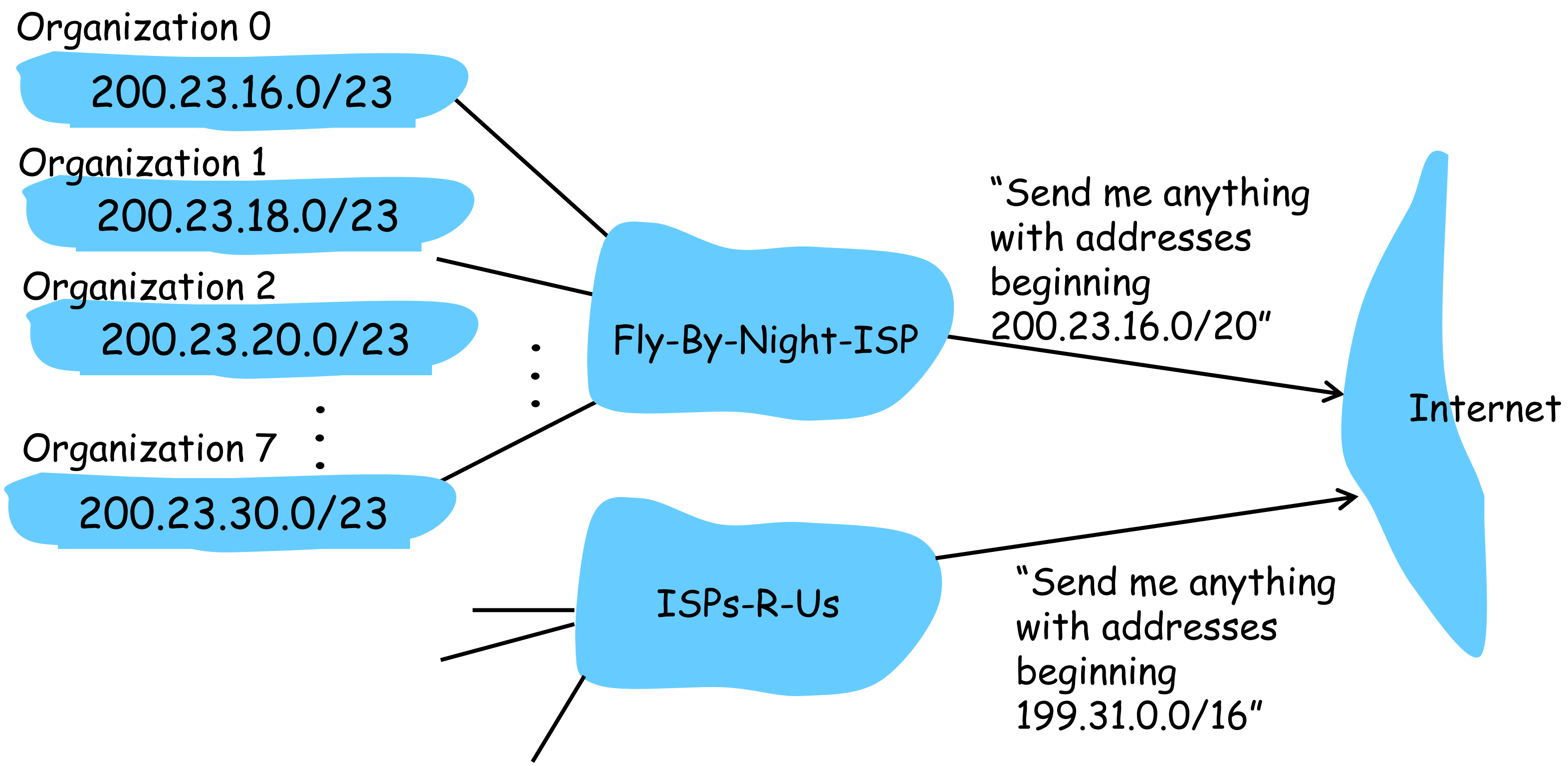
Hierarchical addressing: route aggregation

ISP has an address block; it can further divide this block into sub blocks and assign them to subscriber organizations.



Hierarchical addressing: route aggregation

ISP has an address block; it can further divide this block into sub blocks and assign them to subscriber organizations.



Forwarding: Longest prefix matching

Prefix	Link Interface
200.23.16.0/21	0
200.23.24.0/23	1
200.23.24.0/21	2
otherwise	3

Examples

Dest IP: 200.23.22.161 Which interface?

Dest IP: 200.23.24.172 Which interface?

Forwarding: Longest prefix matching

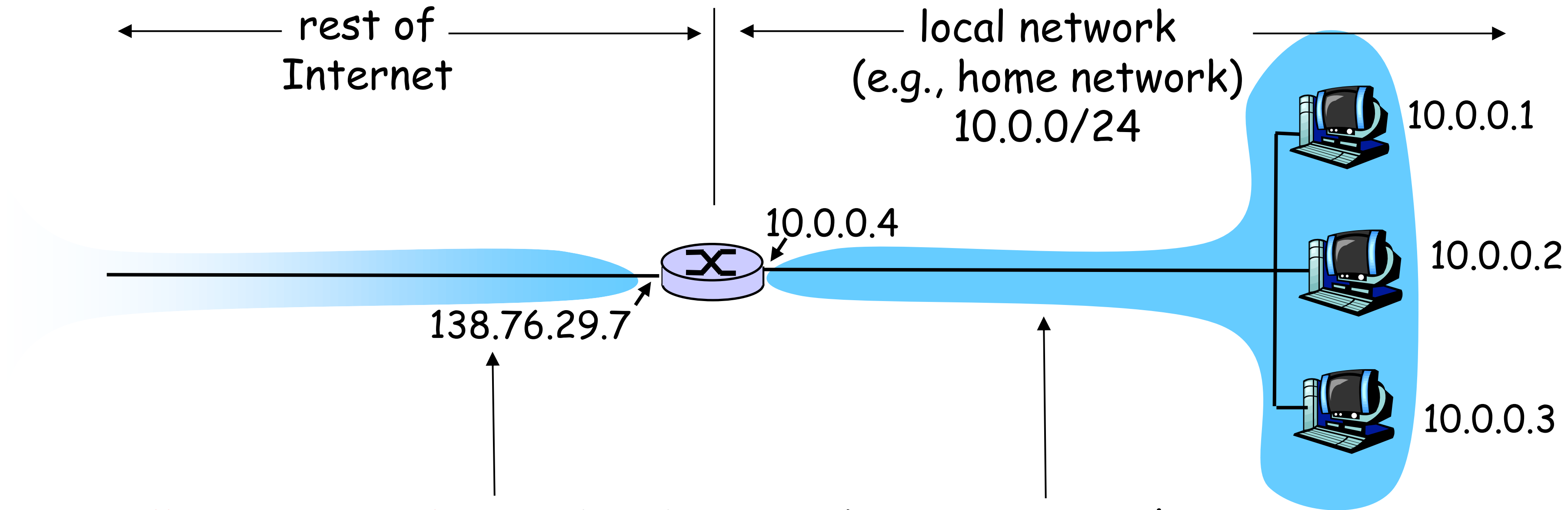
Prefix Match	Link Interface
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

Examples

Dest IP: 11001000 00010111 00010110 10100001 Which interface?

Dest IP: 11001000 00010111 00011000 10101010 Which interface?

NAT: Network Address Translation



All datagrams *leaving* local network have same single source NAT IP address: 138.76.29.7, different source port numbers

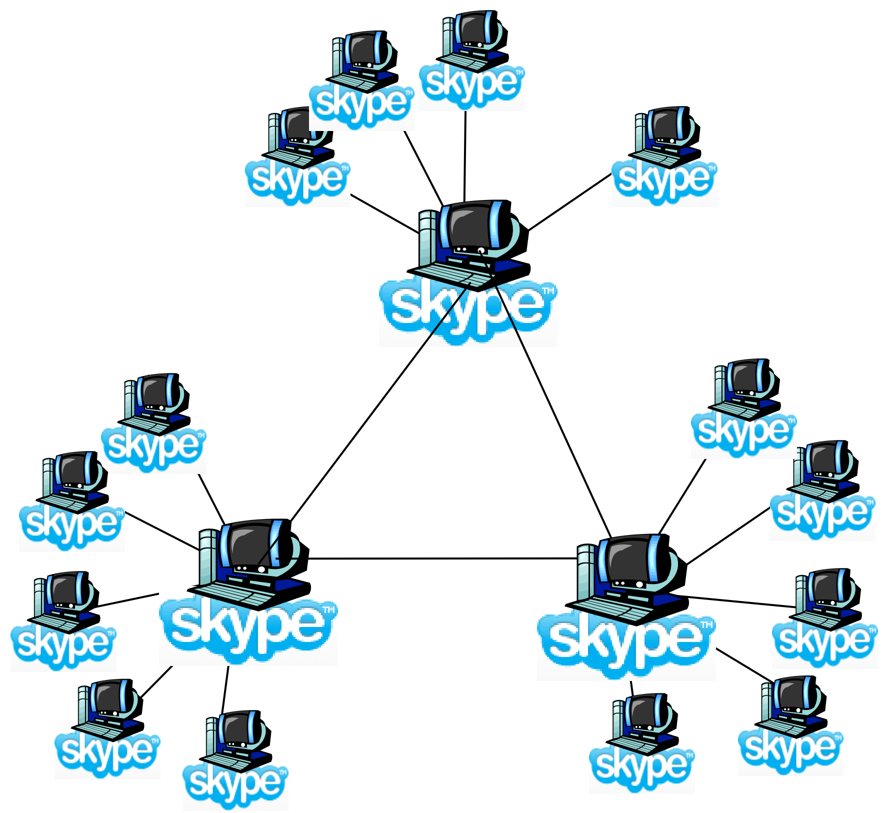
Datagrams with source or destination in this network have 10.0.0/24 address for source, destination (as usual)

NAT: Network Address Translation

- Motivation: local network uses just one IP address as far as outside word is concerned:
  - no need to be allocated range of addresses from ISP: - just one IP address is used for all devices
  - can change addresses of devices in local network without notifying outside world
  - can change ISP without changing addresses of devices in local network
  - devices inside local net not explicitly addressable, visible by outside world (a security plus).

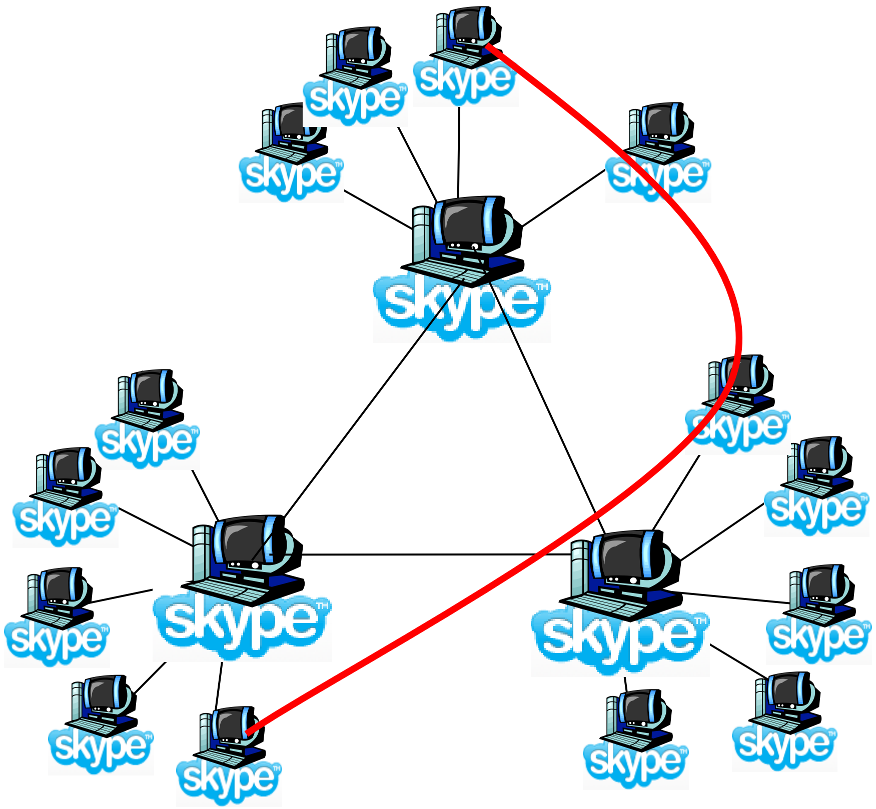
NAT/firewall problems …

- Problem when both Alice and Bob are behind "NATs".
  - NAT prevents an outside peer from initiating a call to insider peer
- Solution:



Peers as relays

- Problem when both Alice and Bob are behind "NATs".
  - NAT prevents an outside peer from initiating a call to insider peer
- Solution:
  - Using Alice's and Bob's SNs, Relay is chosen
  - Each peer initiates session with relay.
  - Peers can now communicate through NATs via relay



NAT: Network Address Translation

- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, eg, P2P applications
  - address shortage should instead be solved by IPv6

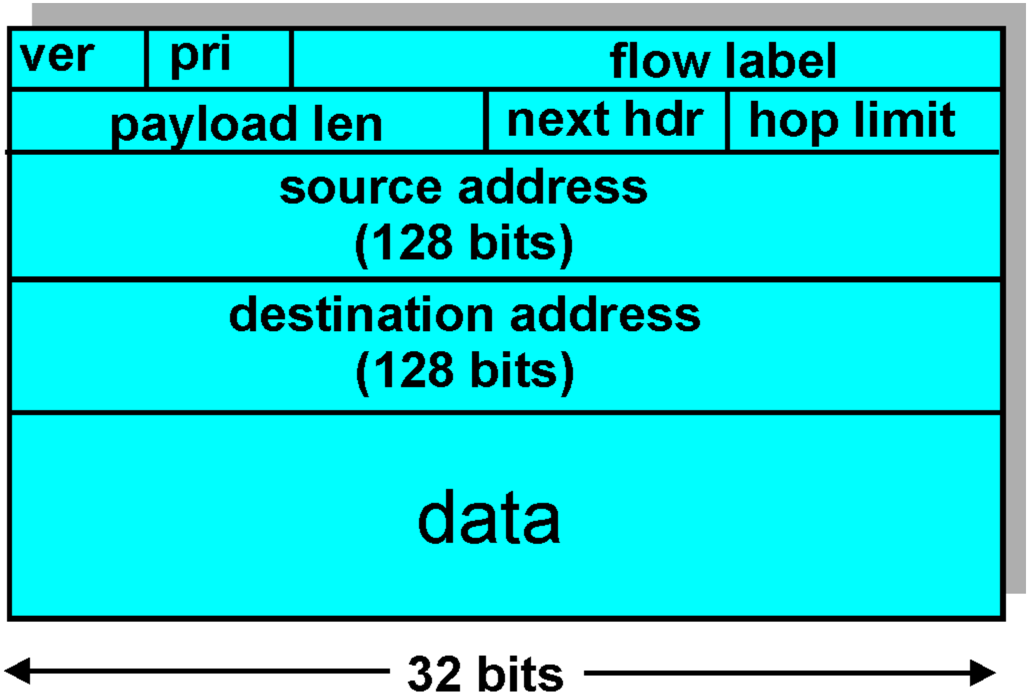
IPv6

- Initial motivation: 32-bit address space soon to be completely allocated.
- Additional motivation:
  - header format helps speed processing/forwarding
  - header changes to facilitate QoS

  IPv6 datagram format:
  - fixed-length 40 byte header
  - no fragmentation allowed

IPv6 Header (Cont)

Priority: identify priority among datagrams in flow
Flow Label: identify datagrams in same "flow."
(concept of "flow" not well defined).
Next header: identify upper layer protocol for data



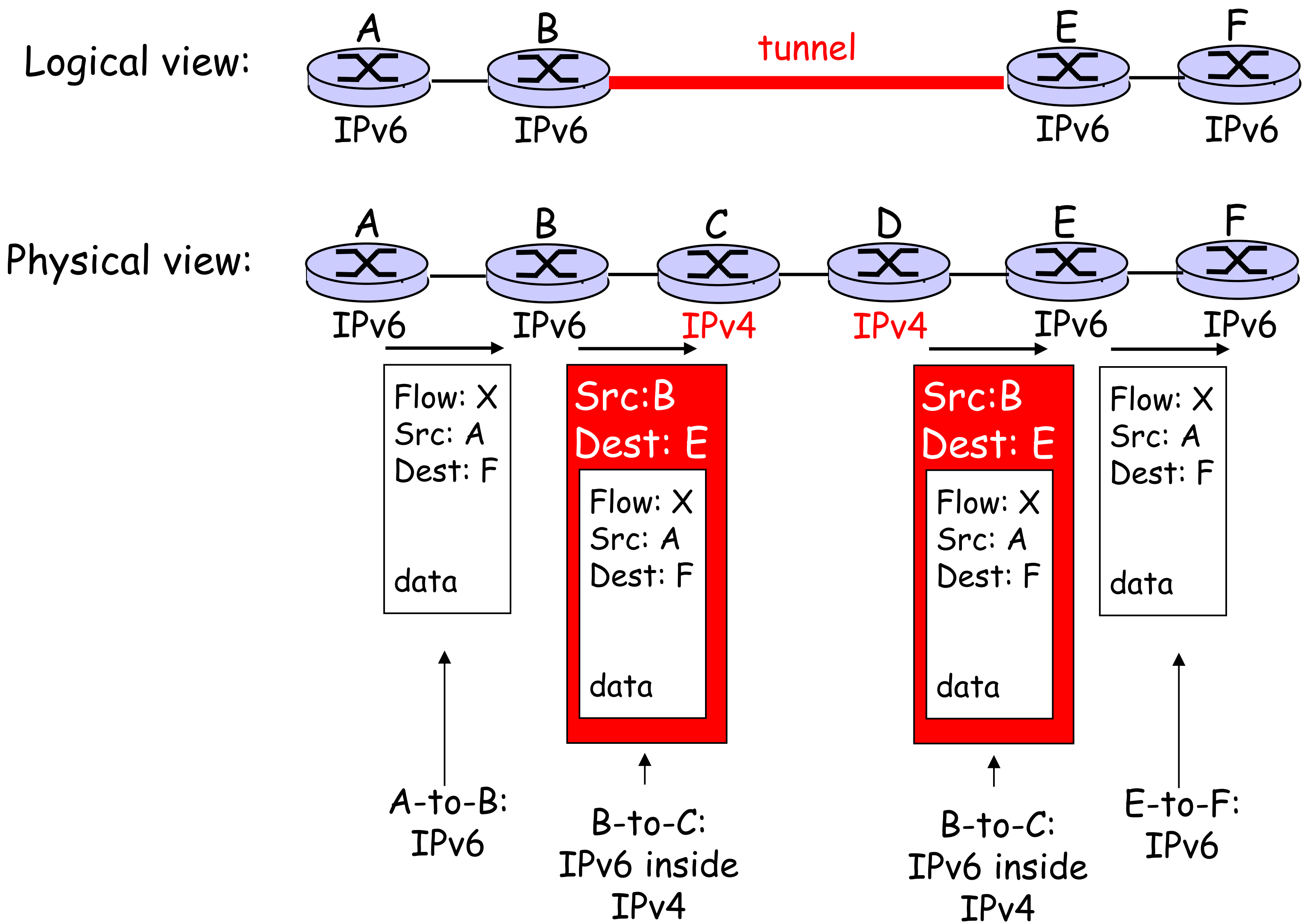
Other Changes from IPv4

- *Checksum*: removed entirely to reduce processing time at each hop
- *Options:* allowed, but outside of header, indicated by “Next Header” field
- *ICMPv6:* new version of ICMP
  - additional message types, e.g. “Packet Too Big”
  - multicast group management functions

Transition From IPv4 To IPv6

- Not all routers can be upgraded simultaneous
  - No "flag days" (when all must switch) reasonable
  - How will the network operate with mixed IPv4 and IPv6 routers?
- *Tunneling:* IPv6 carried as payload in IPv4 datagram among IPv4 routers

Tunneling



Logical view:
A
B
tunnel
E
F
IPv6
IPv6
IPv6
IPv6
Physical view:
A
B
C
D
E
F
IPv6
IPv6
IPv4
IPv4
IPv6
IPv6
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Src:B
Dest: E
Flow: X
Src: A
Dest: F
data
Flow: X
Src: A
Dest: F
data
A-to-B:
IPv6
B-to-C:
IPv6 inside
IPv4
B-to-C:
IPv6 inside
IPv4
E-to-F:
IPv6