Queen Mary
University of London

ECS524
Network layer

Prof. Steve Uhlig
steve.uhlig@qmul.ac.uk
Office: Eng202

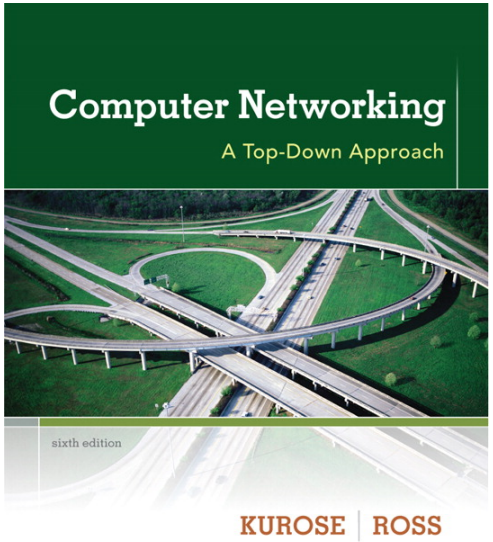
Dr. Felix Cuadrado
felix.cuadrado@qmul.ac.uk
Office: E205

Slides

Queen Mary University of London

Disclaimer:
Some of the slides' content is borrowed directly from those provided by the authors of the textbook. They are available from

http://www-net.cs.umass.edu/kurose-ross-ppt-6e



Computer Networking: A Top Down Approach
6th edition
Jim Kurose, Keith Ross
Addison-Wesley
March 2012

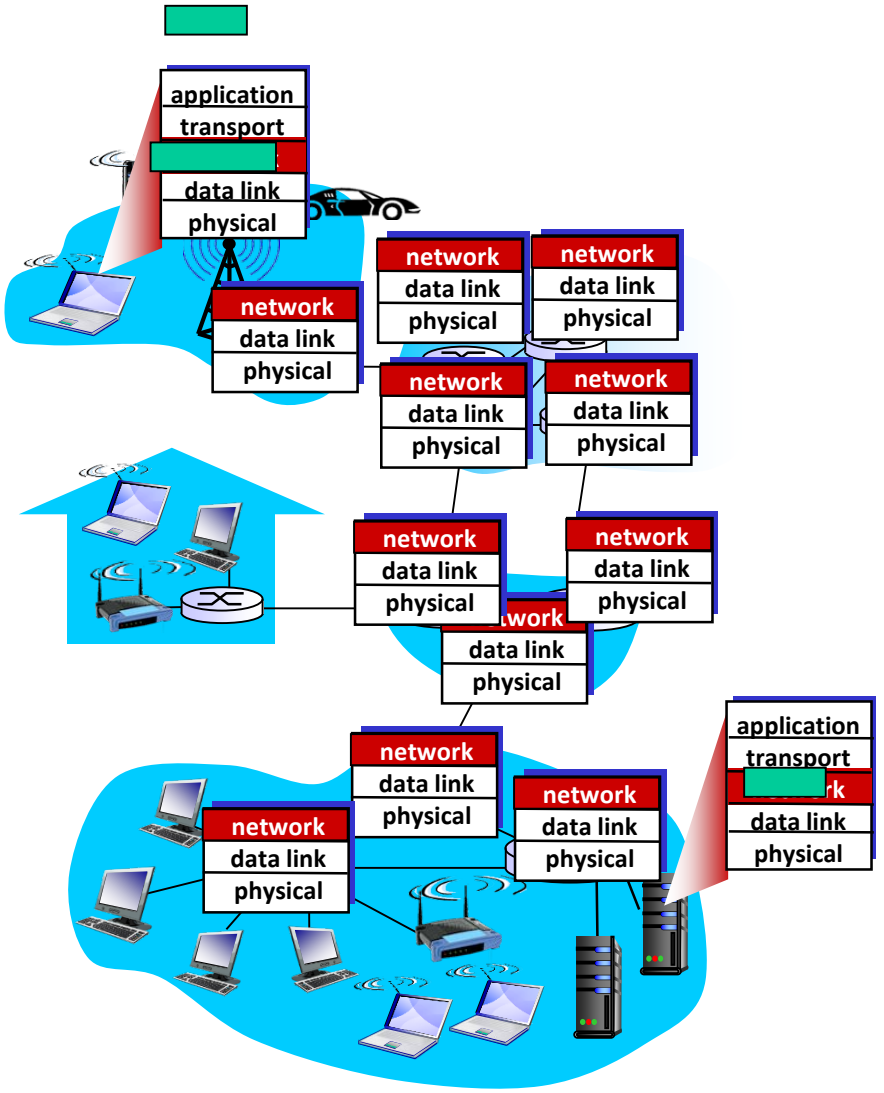
The Network layer

Queen Mary University of London

- **Introduction**
- IP Addressing
- IP routers
- IP
- Routing: concepts
- Routing: practice

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



Datagram networks

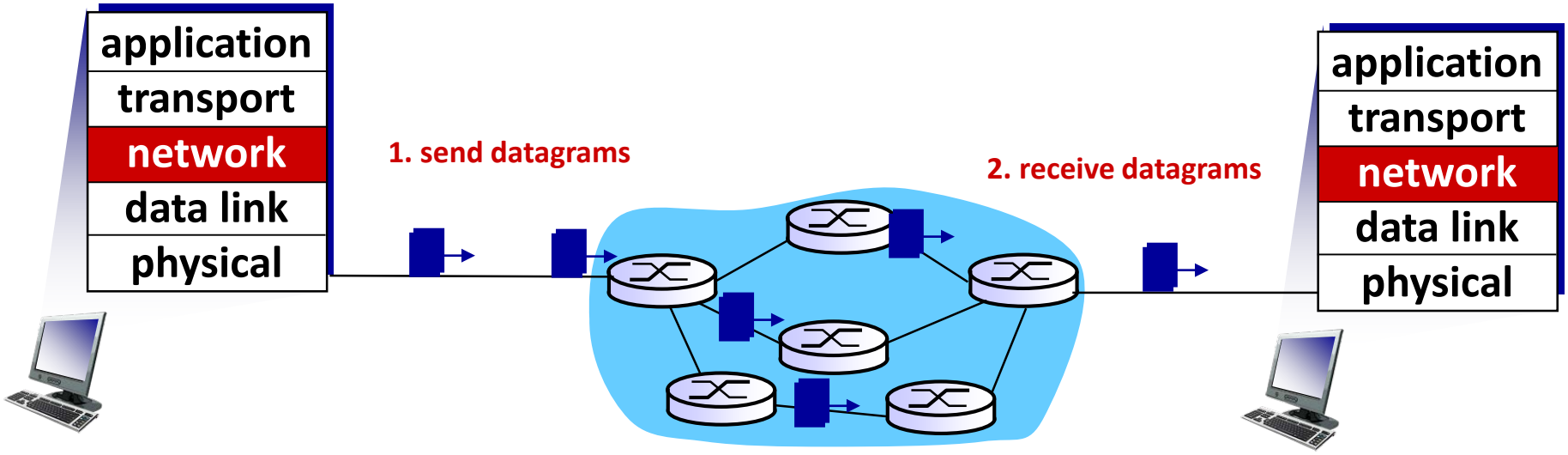
Queen Mary University of London

no call setup at network layer

routers: no state about end-to-end connections

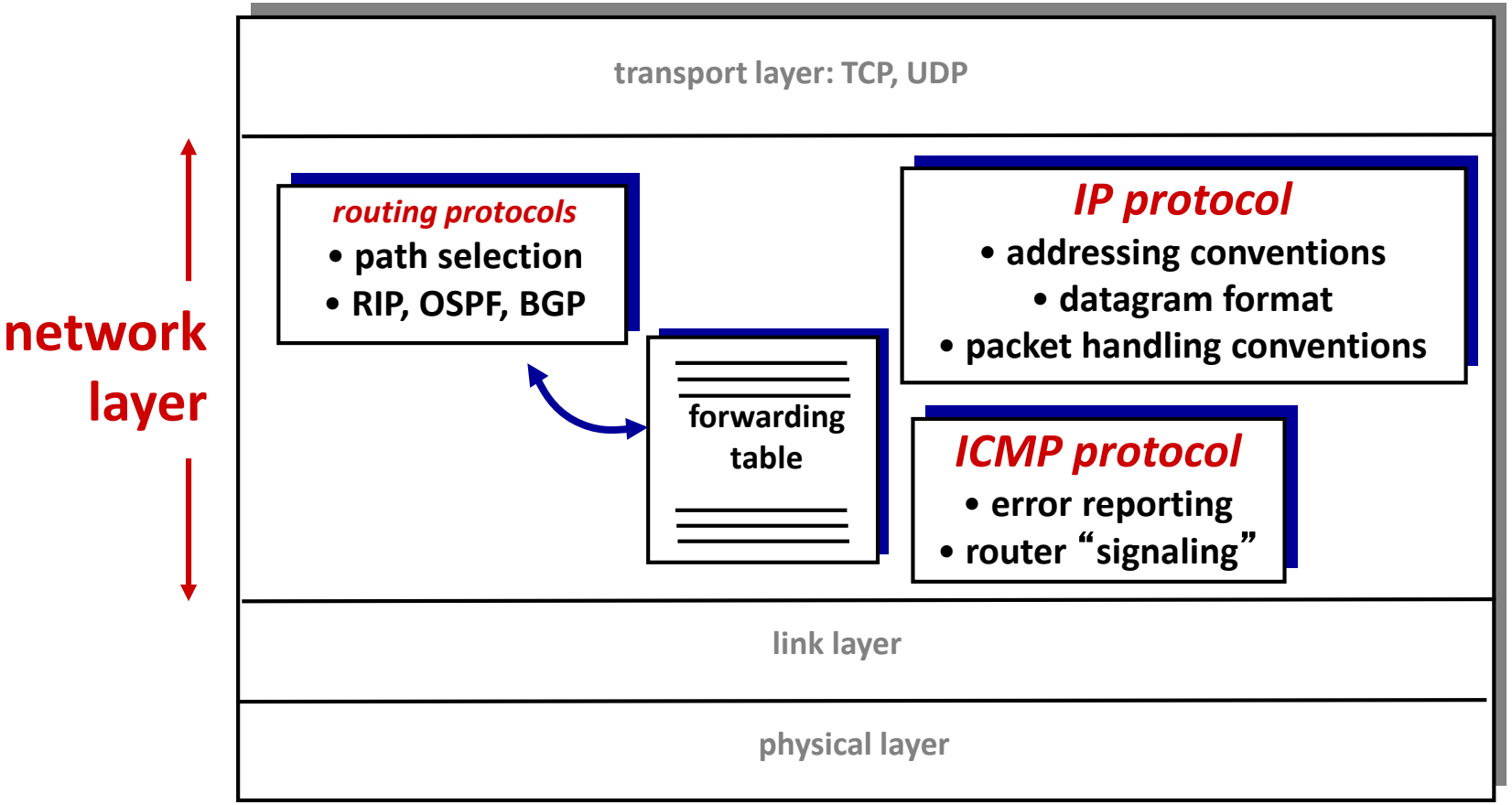
no network-level concept of “connection”

packets forwarded using destination host address



The Internet network layer

- Host, router network layer functions:



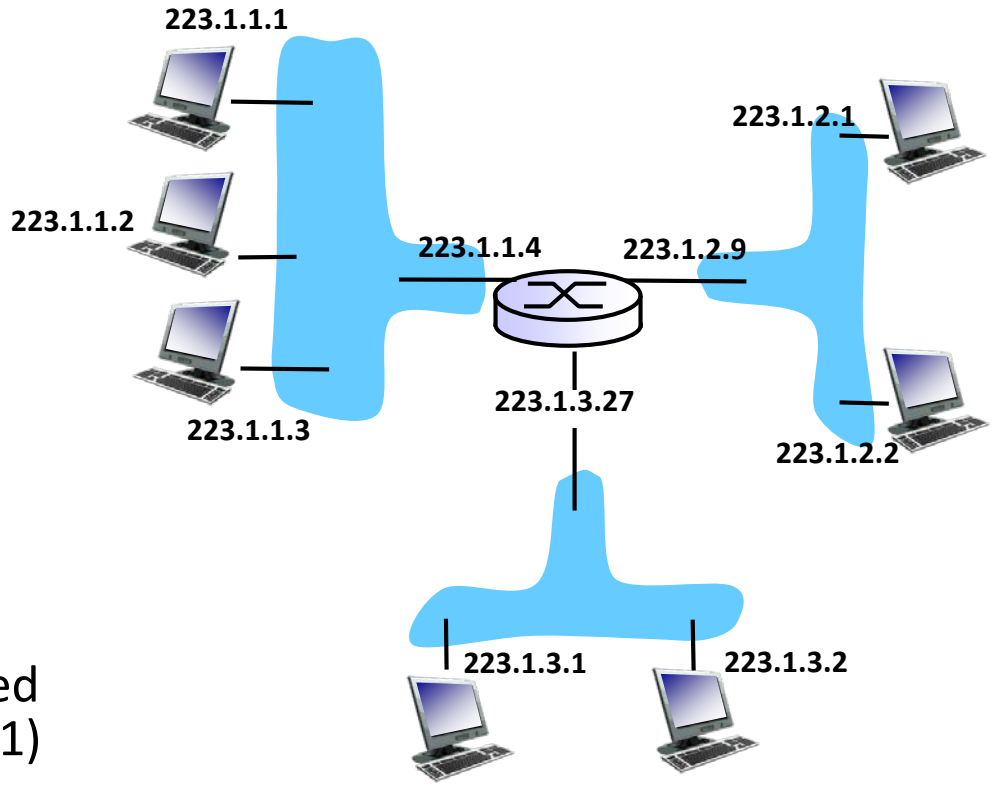
The Network layer

Queen Mary University of London

- Introduction
- **IP Addressing**
- IP routers
- IP
- Routing: concepts
- Routing: practice

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 typically has one or two interfaces (e.g., wired Ethernet, wireless 802.11)
- *IP addresses associated with each interface*

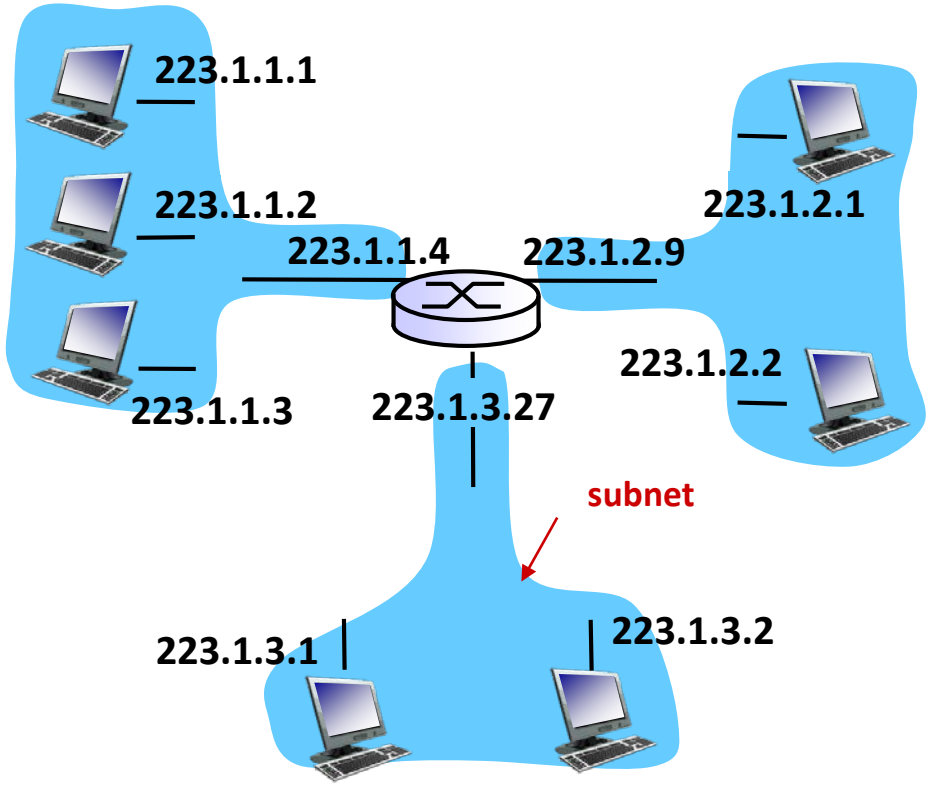


223.1.1.1 = 11011111 00000001 00000001 00000001

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*

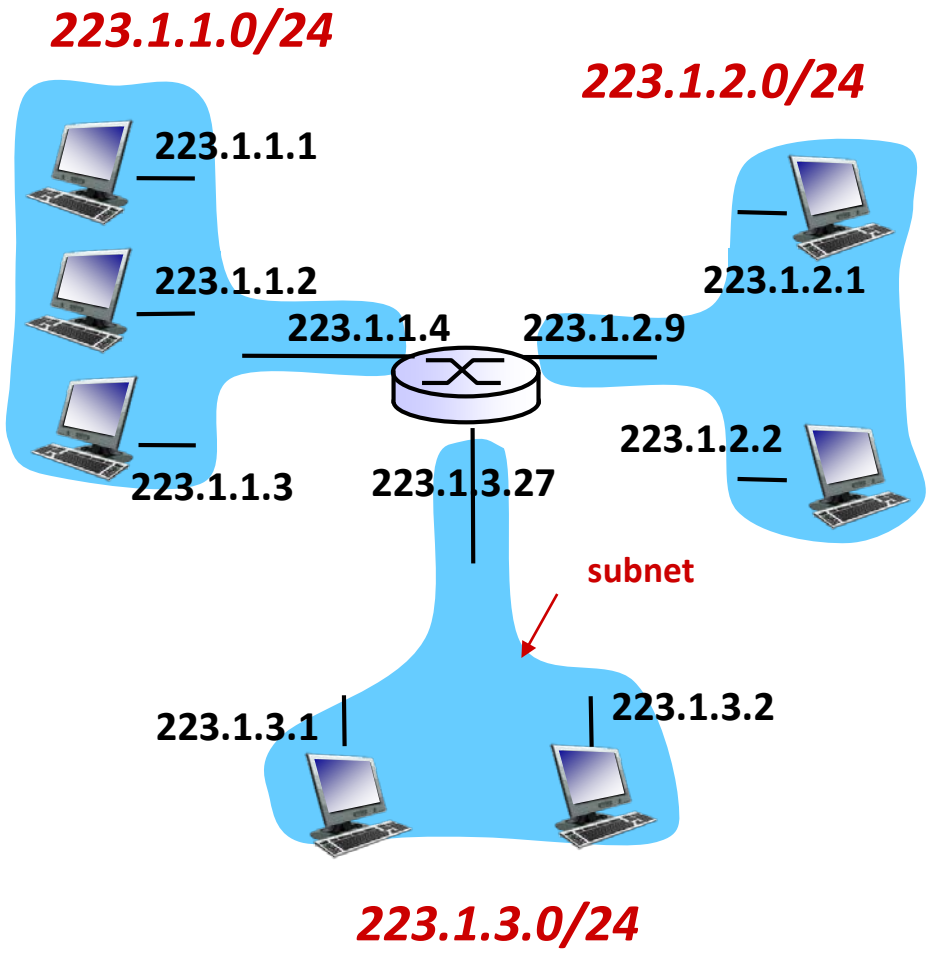


network consisting of 3 subnets

Subnets

Recipe

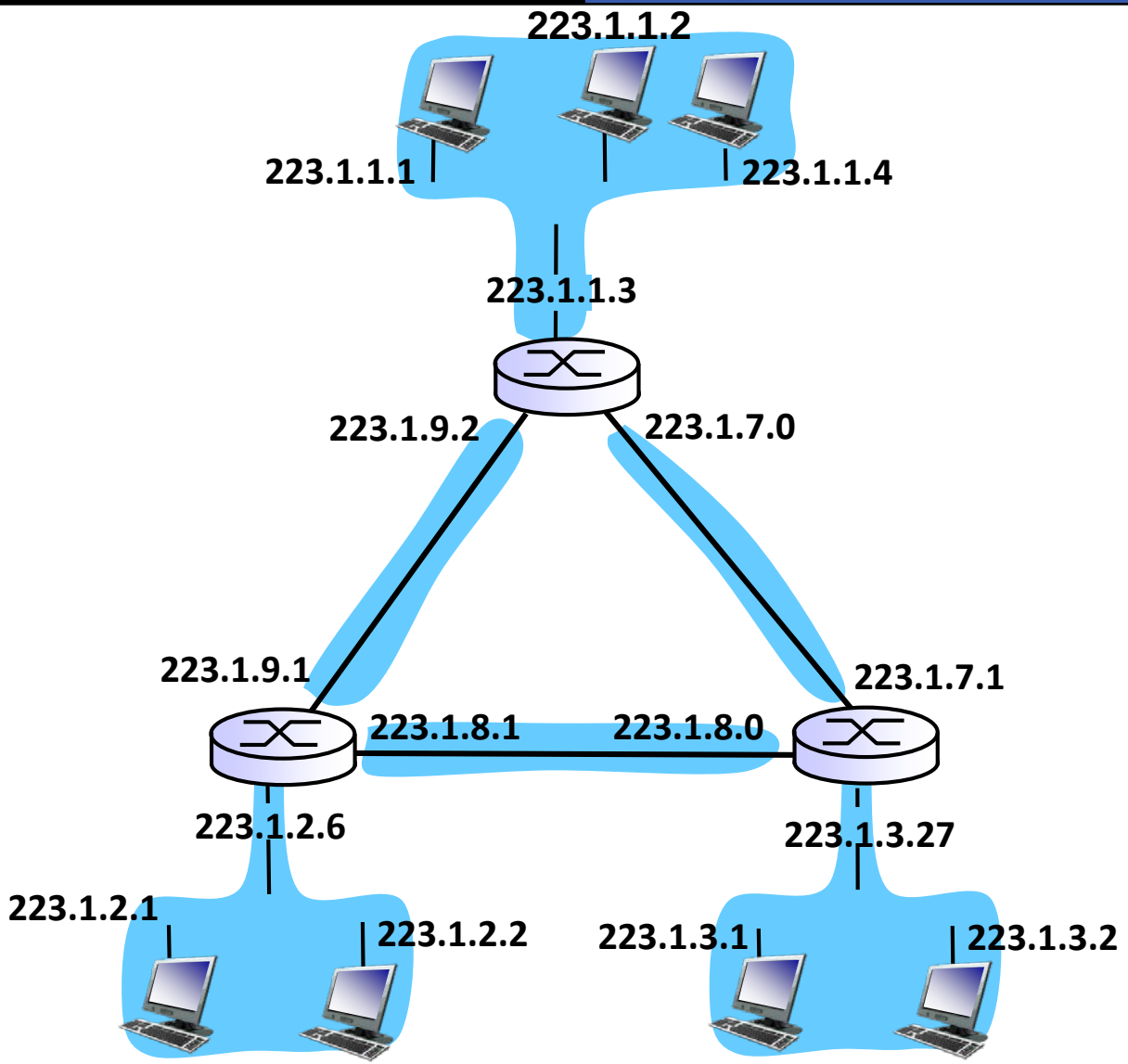
- to determine the subnets, detach each interface from its host or router, creating islands of isolated networks
- each isolated network is called a *subnet*



subnet mask: /24

Subnets

How many?



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

11001000 00010111 00010000 00000000

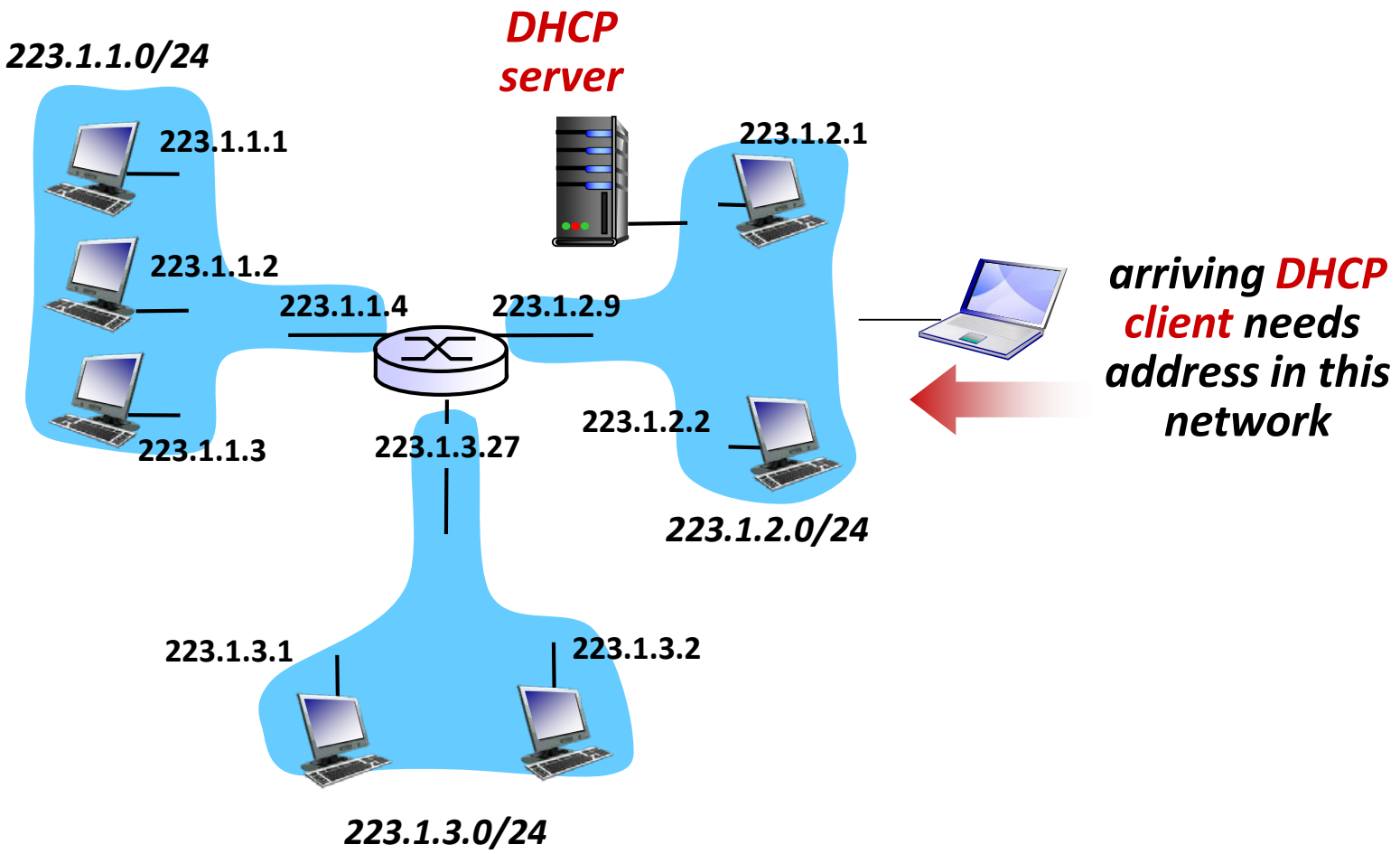
200.23.16.0/23

IP address: how to get one?

Q: How does a *host* get an IP address?

- Hard-coded by system admin in a file
  - 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"

DHCP client-server scenario

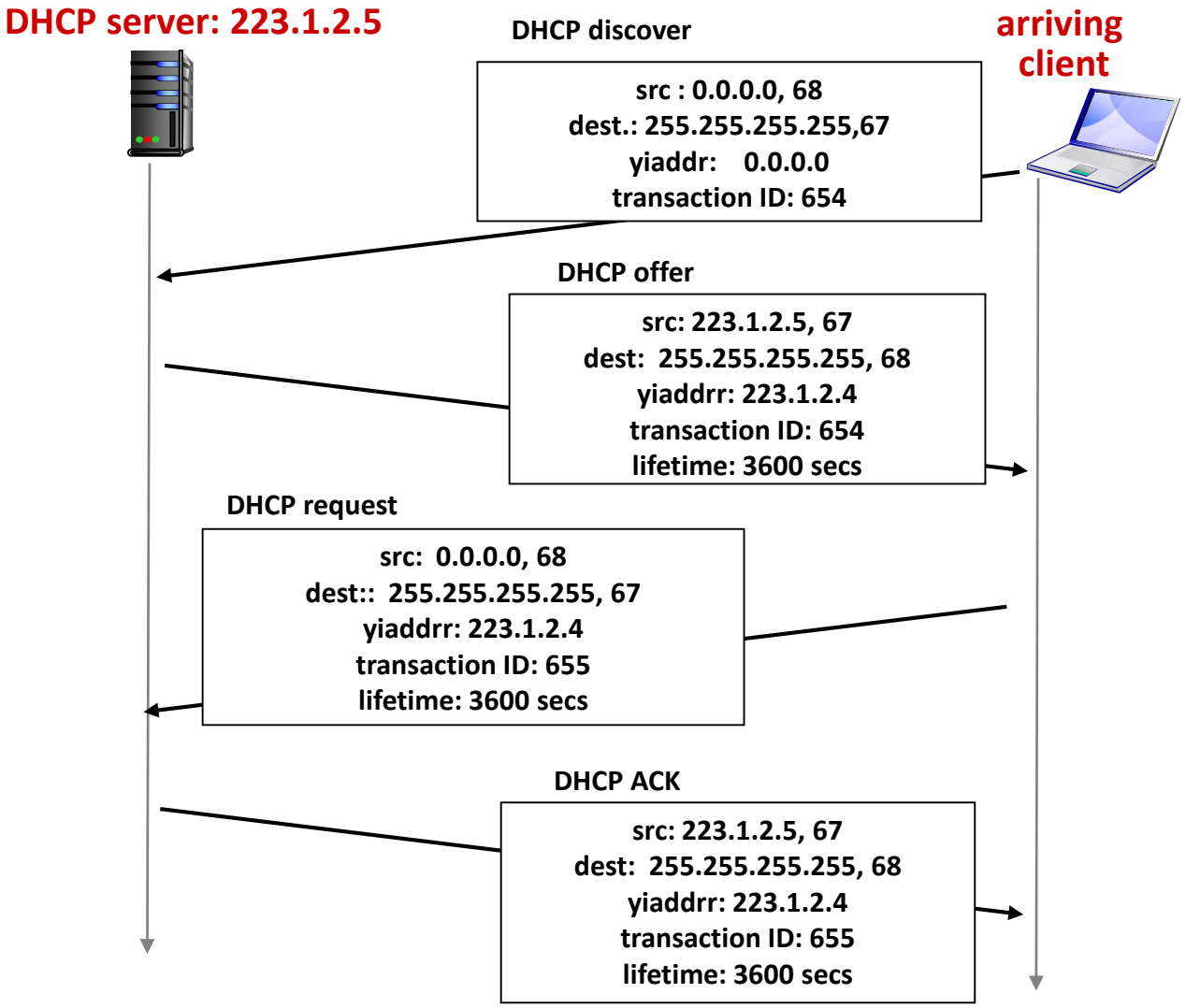


DHCP: Dynamic Host Configuration Protocol

Queen Mary University of London

- *Goal:* allow host to *dynamically* obtain its IP address from network server when it joins network
  - can renew its lease on address in use
  - allows reuse of addresses (only hold address while connected/"on")
  - support for mobile users who want to join network (more shortly)
- *DHCP overview:*
  - host broadcasts "DHCP discover" msg [optional]
  - DHCP server responds with "DHCP offer" msg [optional]
  - host requests IP address: "DHCP request" msg
  - DHCP server sends address: "DHCP ack" msg

DHCP client-server scenario



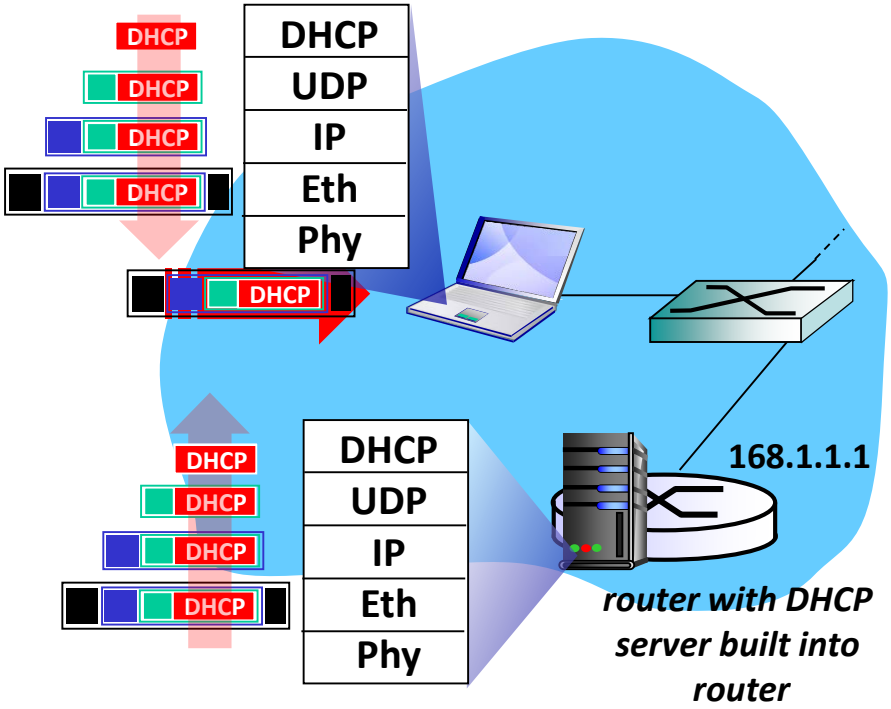
DHCP: more than IP addresses

Queen Mary University of London

DHCP can return more than just allocated IP address on subnet:

- address of first-hop router for client
- name and IP address of DNS server
- network mask (indicating network versus host portion of address)

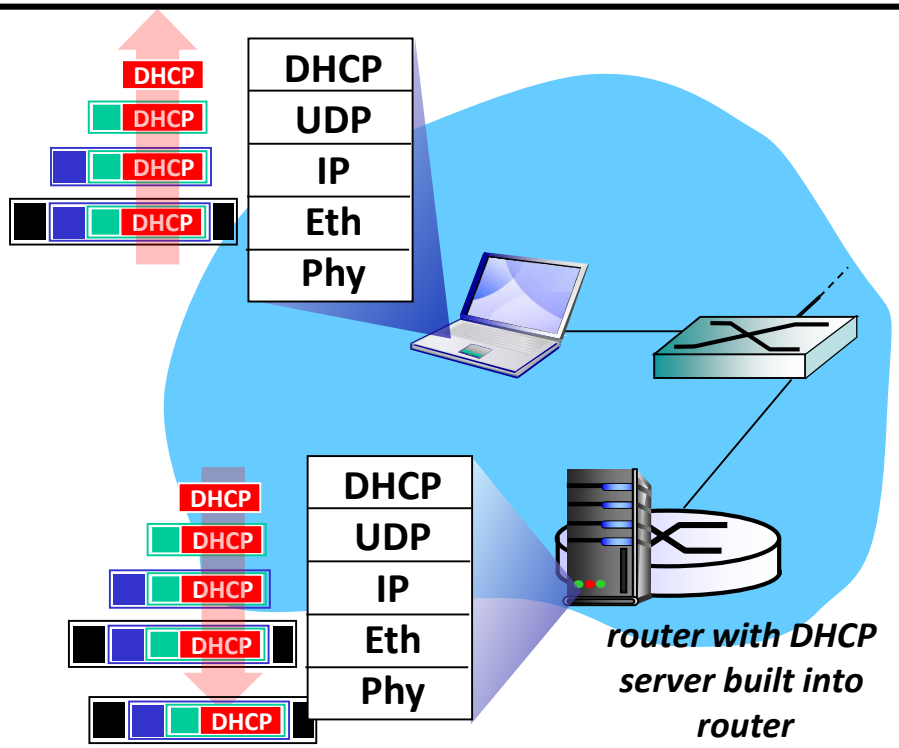
DHCP: example



Connecting laptop needs its IP address, addr of first-hop router, addr of DNS server: use DHCP

- **DHCP request encapsulated in UDP, encapsulated in IP, encapsulated in 802.1 Ethernet**
- **Ethernet frame broadcast (dest: FFFFFFFFFFFF) on LAN, received at router running DHCP server**
- **Ethernet demuxed to IP demuxed, UDP demuxed to DHCP**

DHCP: example



DHCP server formulates DHCP ACK containing client's IP address, IP address of first-hop router for client, name & IP address of DNS server

- **Encapsulation of DHCP server, frame forwarded to client, demuxing up to DHCP at client**
- **Client now knows its IP address, name and IP address of DNS server, IP address of its first-hop router**

IP addresses: how to get one?

Queen Mary University of London

Q: how does *network* get subnet part of IP address?

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**

IP addressing: the last word...

Q: how does an ISP get block of addresses?

A: ICANN: Internet Corporation for Assigned Names and Numbers http://www.icann.org/

allocates addresses

manages DNS

assigns domain names, resolves disputes

The Network layer

Queen Mary University of London

- Introduction
- Addressing
- **IP routers**
- IP
- Routing: concepts
- Routing: practice

Two key network-layer functions

forwarding: move packets from router's input to appropriate router output

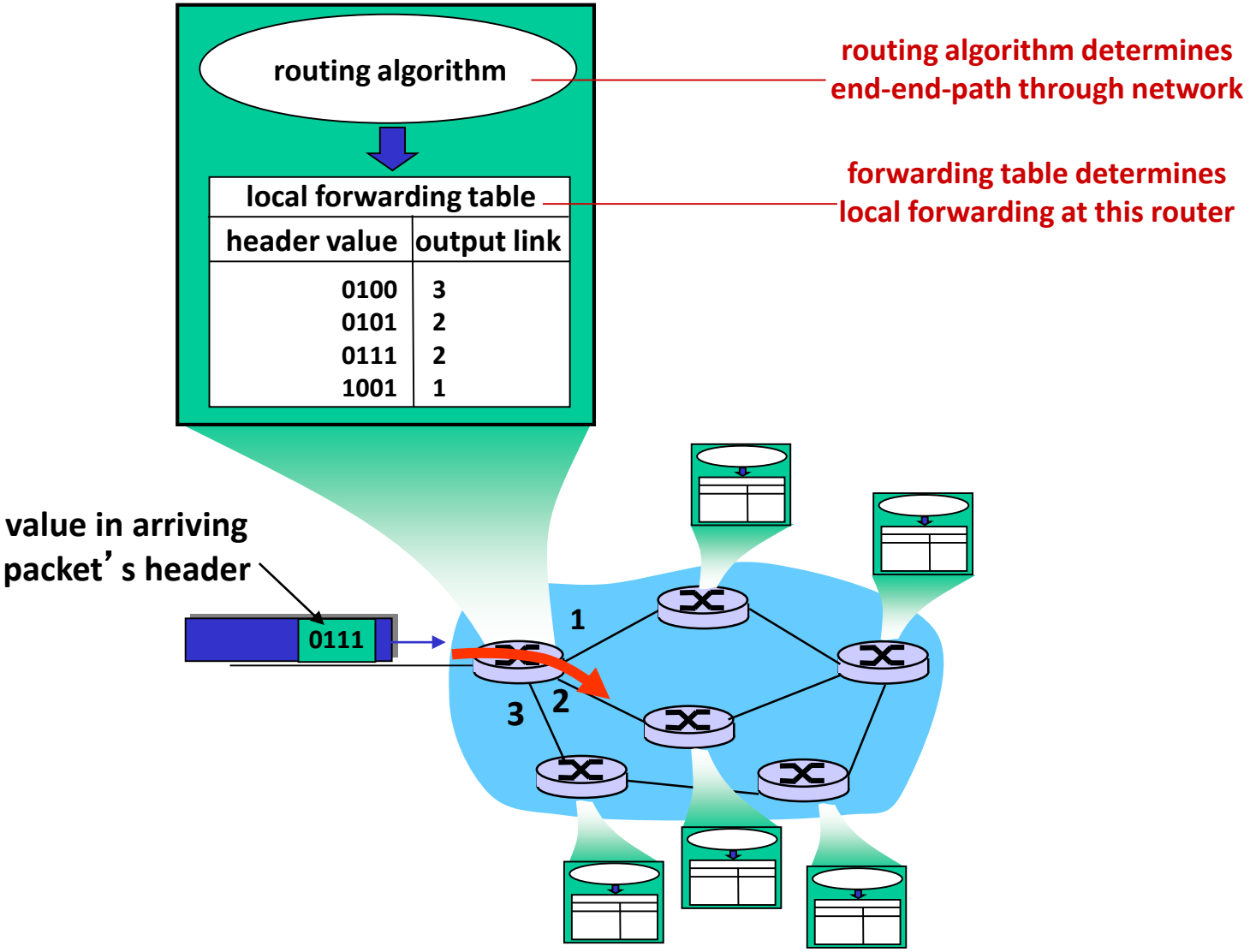
routing: determine route taken by packets from source to dest.

routing algorithms

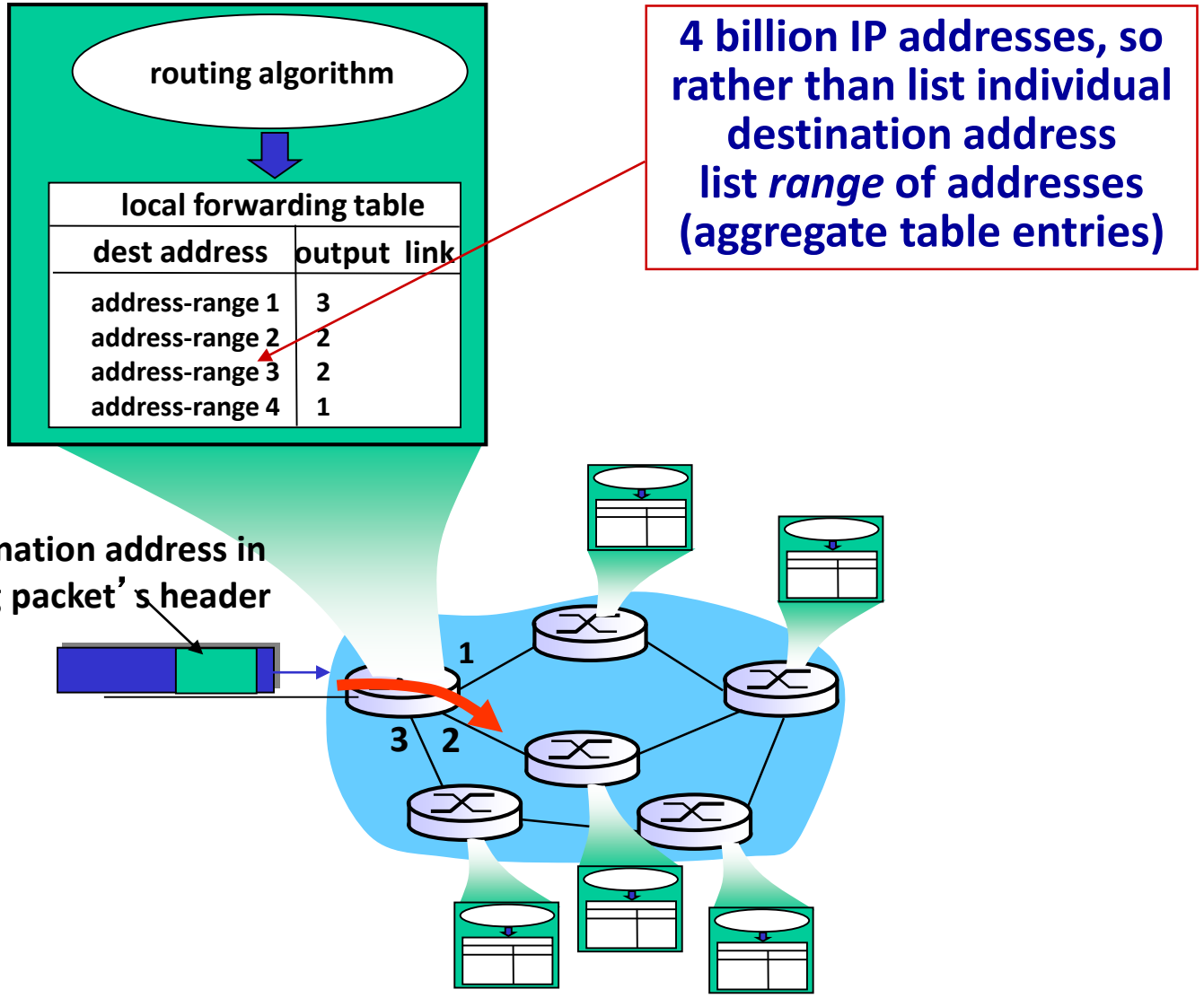
analogy:

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

Interplay between routing and forwarding



Datagram forwarding table



Datagram forwarding table

Queen Mary University of London

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	3

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

Longest prefix matching

Queen Mary University of London

longest prefix matching

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** *********	0
11001000 00010111 00011000 *********	1
11001000 00010111 00011*** *********	2
otherwise	3

examples:

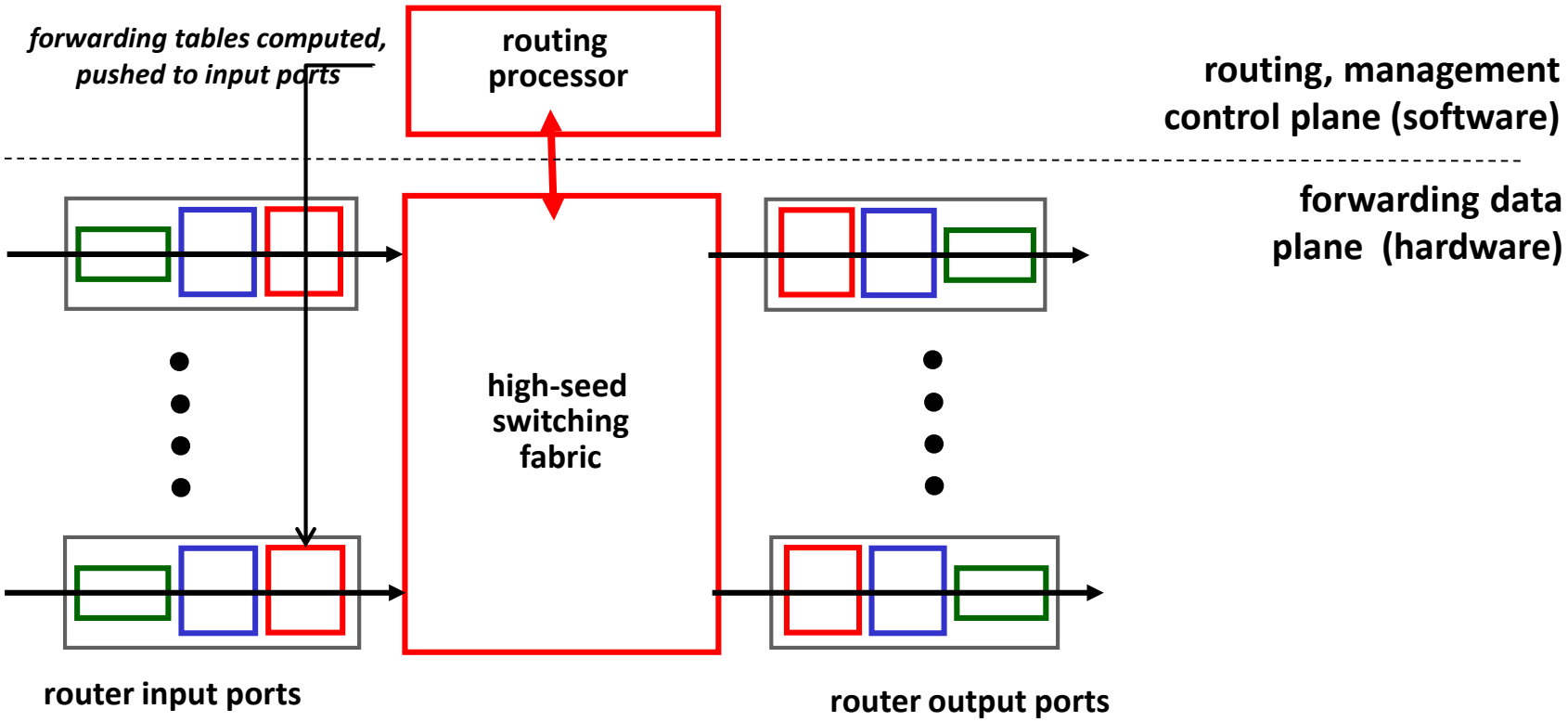
DA: 11001000 00010111 00010110 10100001 — **which interface?**

DA: 11001000 00010111 00011000 10101010 — **which interface?**

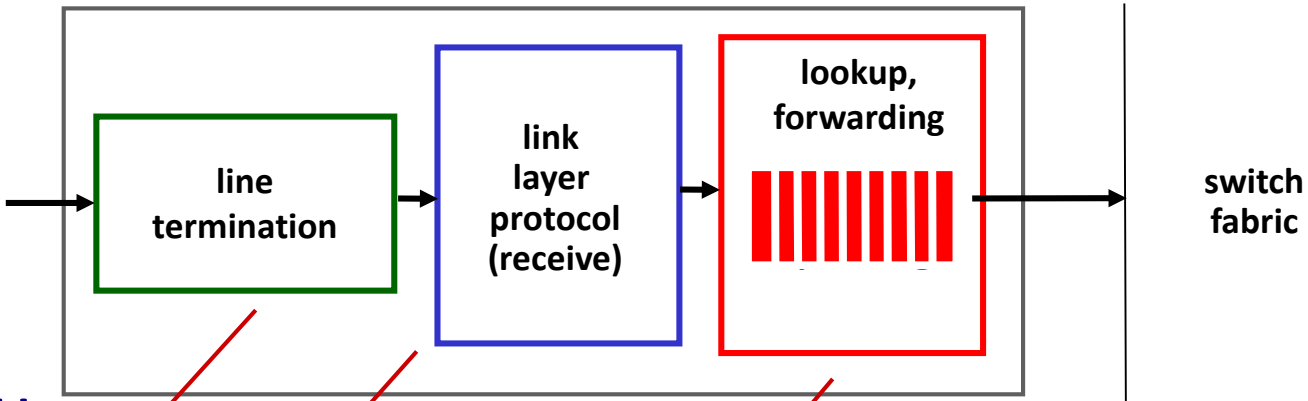
Router architecture overview

Two key router functions:

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



Input port functions



physical layer:
bit-level reception

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

decentralized switching:

given datagram dest., lookup output port using forwarding table in input port memory *("match plus action")*

goal: complete input port processing at 'line speed'

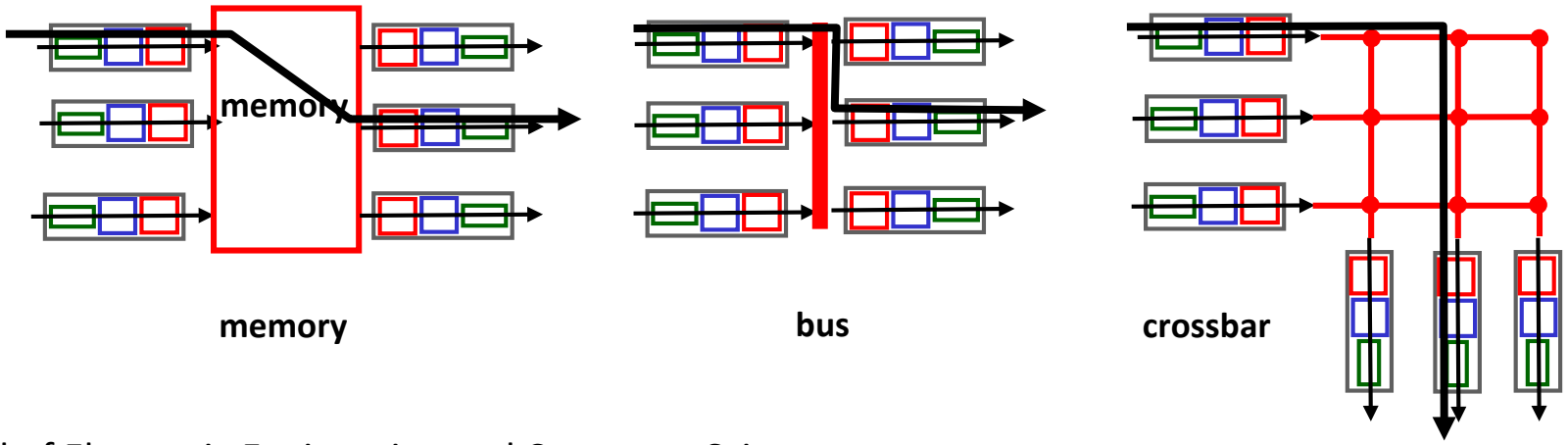
queuing: if datagrams arrive faster than forwarding rate into switch fabric

Switching fabrics

- Transfer packet from input buffer to appropriate output buffer
- Switching rate: rate at which packets can be transfer from inputs to outputs

  often measured as multiple of input/output line rate

  N inputs: switching rate N times line rate desirable
- Three types of switching fabrics

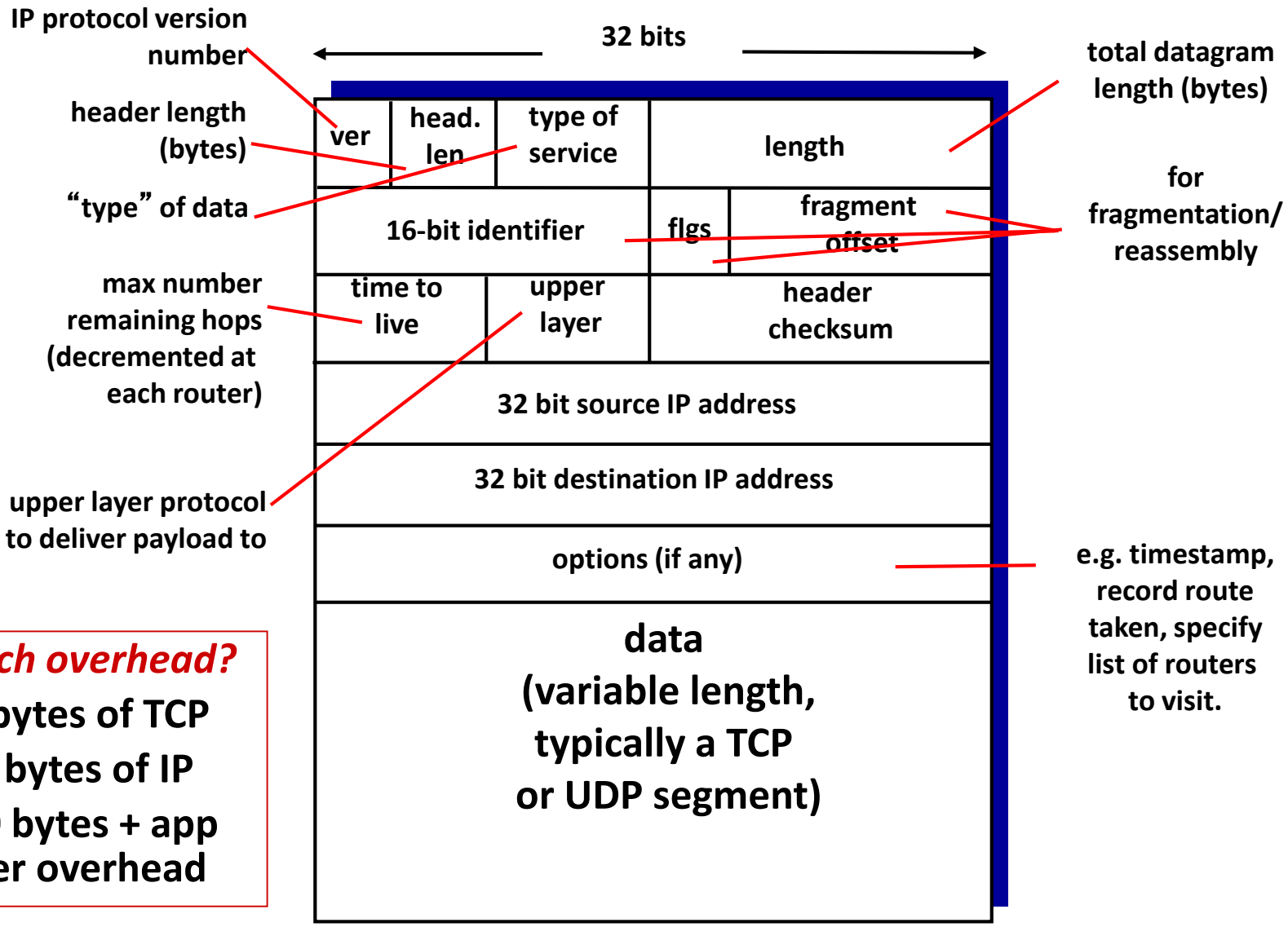


The Network layer

Queen Mary University of London

- Introduction
- IP Addressing
- IP router
- **IP**
- Routing: concepts
- Routing: practice

IP datagram format



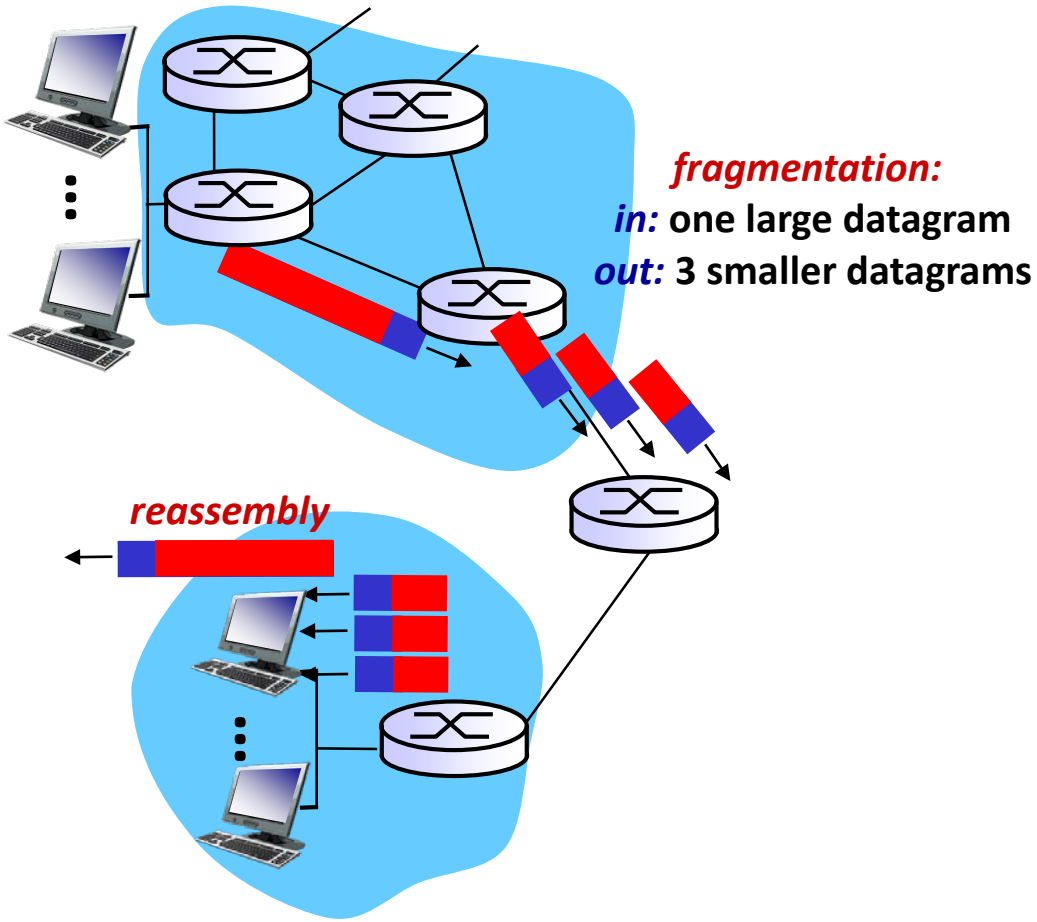
how much overhead?

- 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

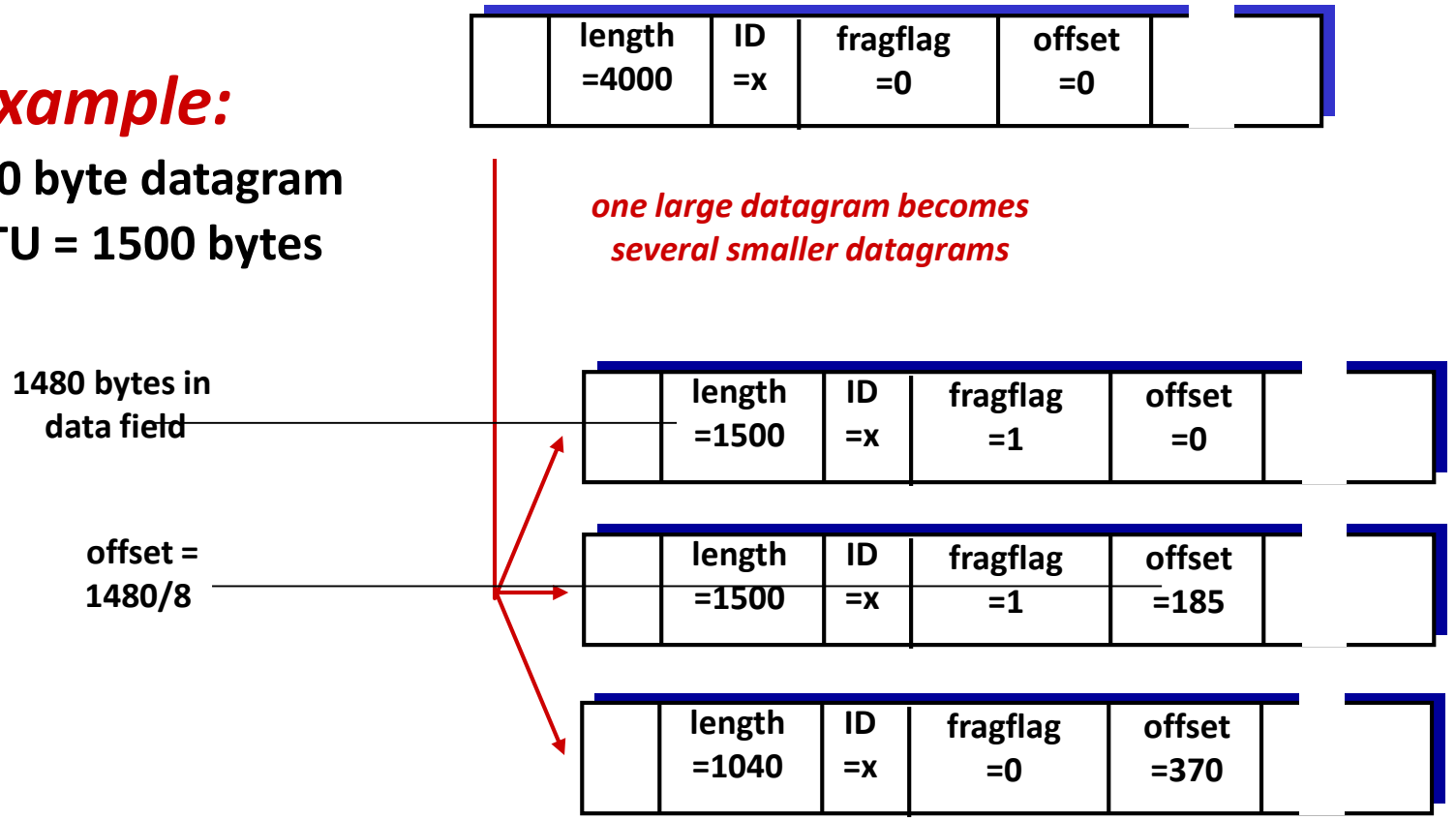
  different link types, different MTUs
- Large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - reassembled" only at final destination
  - IP header bits used to identify, order related fragments



IP fragmentation, reassembly

example:

- 4000 byte datagram
- MTU = 1500 bytes



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

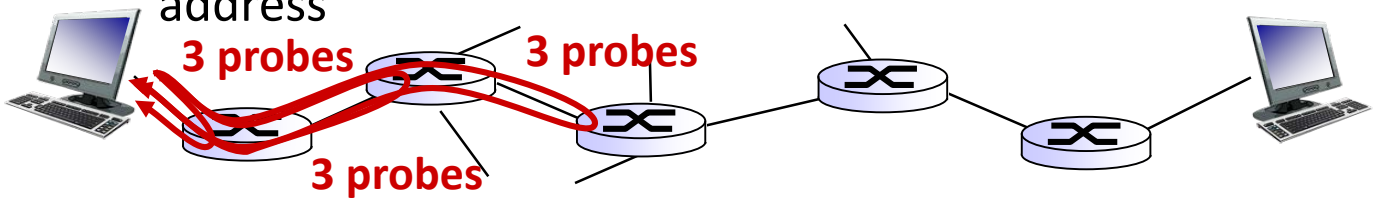
Traceroute and ICMP

- Source sends series of UDP segments to dest
  - first set has TTL =1
  - second set has TTL=2, etc.
  - unlikely port number
- When *n*th set of datagrams arrives to nth router:
  - router discards datagrams
  - and sends source ICMP messages (type 11, code 0)
  - ICMP messages includes name of router & IP address

- When ICMP messages arrives, source records RTTs

stopping criteria:

- **UDP segment eventually arrives at destination host**
- **Destination returns ICMP "port unreachable" message (type 3, code 3)**
  - **source stops**



The Network layer

Queen Mary University of London

- Introduction
- IP router
- IP
- **Routing: concepts**
- Routing: practice