DATA COMMUNICATOIN NETWORKING

Instructor: Ouldooz Baghban Karimi

Course Book: Computer Networking, A Top-Down Approach
By: Kurose, Ross

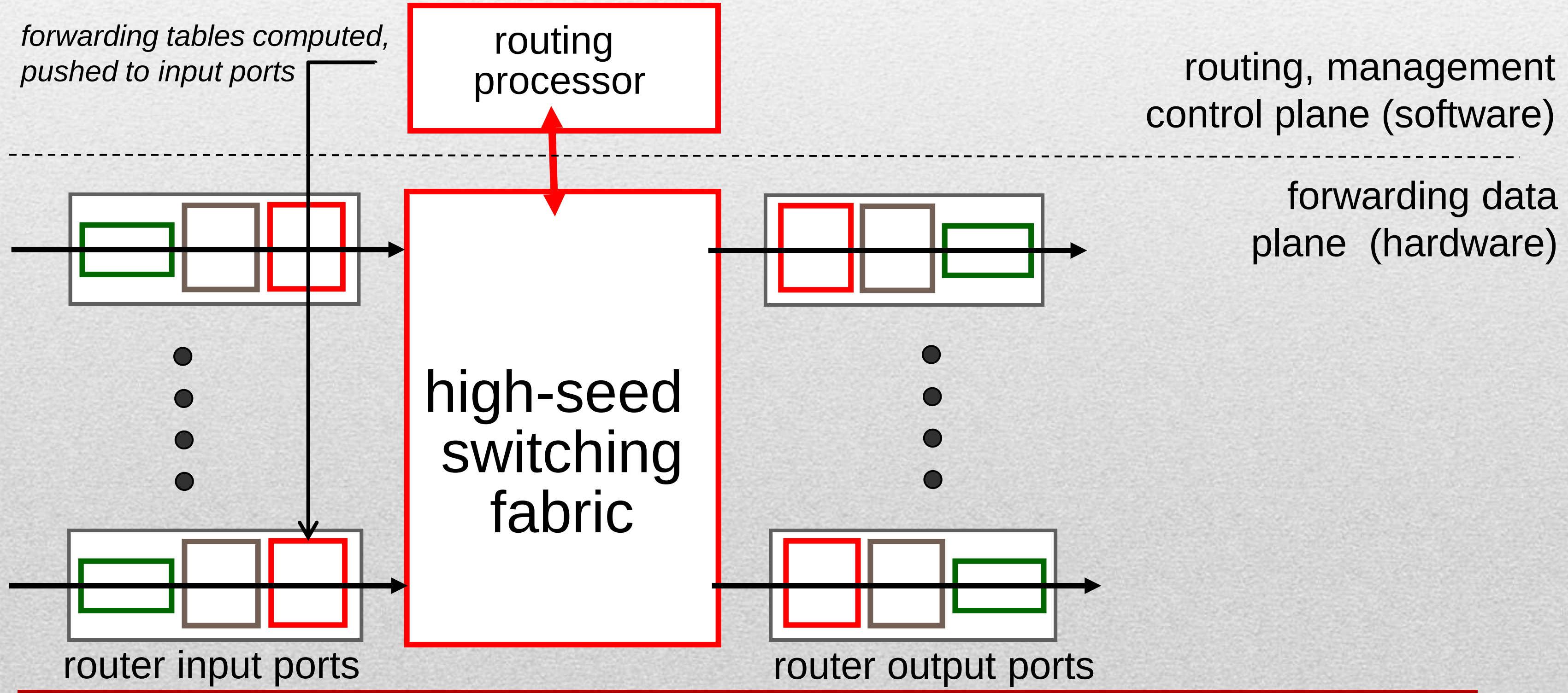
Course Overview

- **Basics of Computer Networks**
  - Internet & Protocol Stack
  - Application Layer
  - Transport Layer
  - Network Layer
  - Data Link Layer
- **Advanced Topics**
  - Case Studies of Computer Networks
  - Internet Applications
  - Network Management
  - Network Security

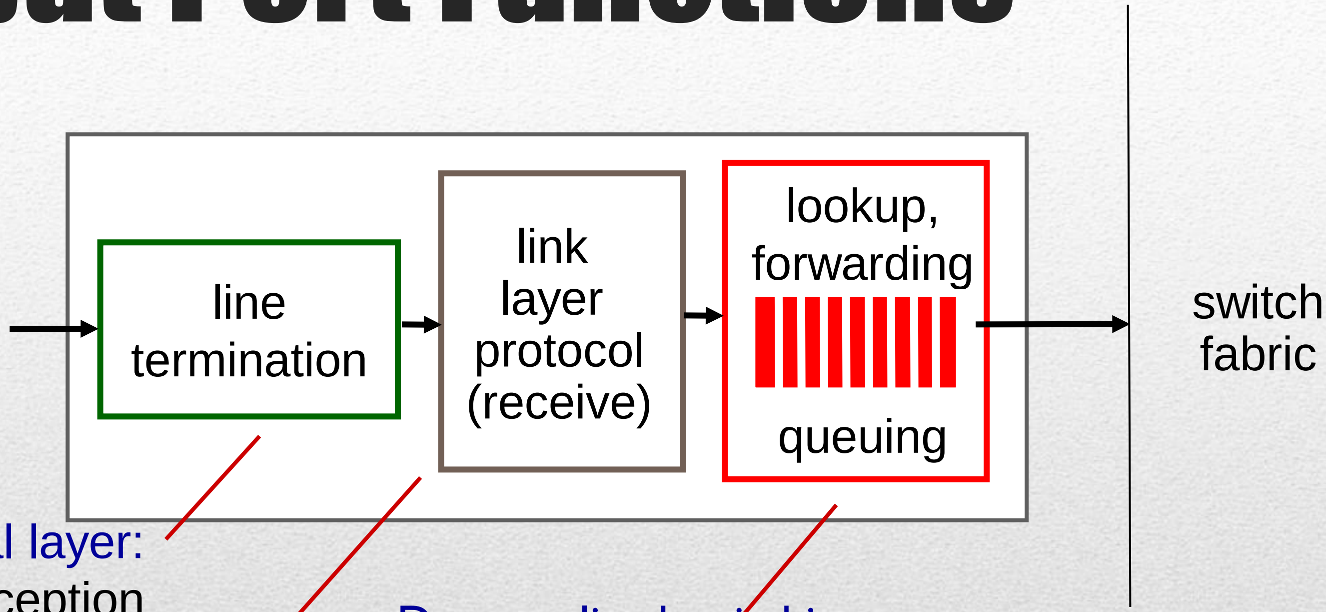
Router Architecture

- Run routing algorithms/protocol
  - RIP, OSPF, BGP

- Forwarding datagrams from incoming to outgoing link



Input Port Functions



Physical layer:
bit-level reception

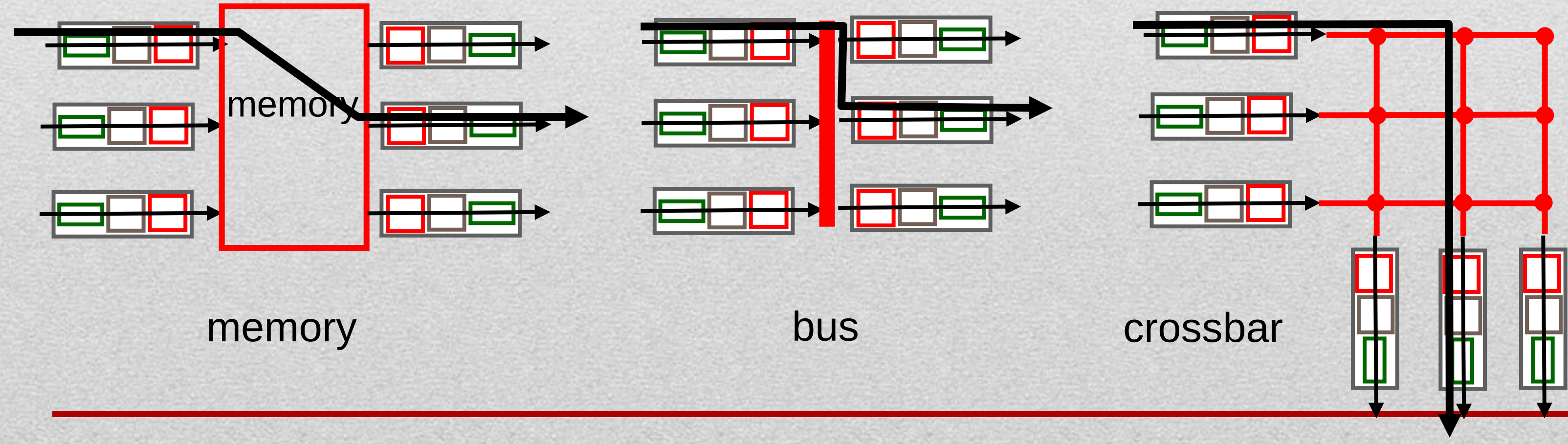
Data link layer:
e.g., Ethernet
see chapter 5

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

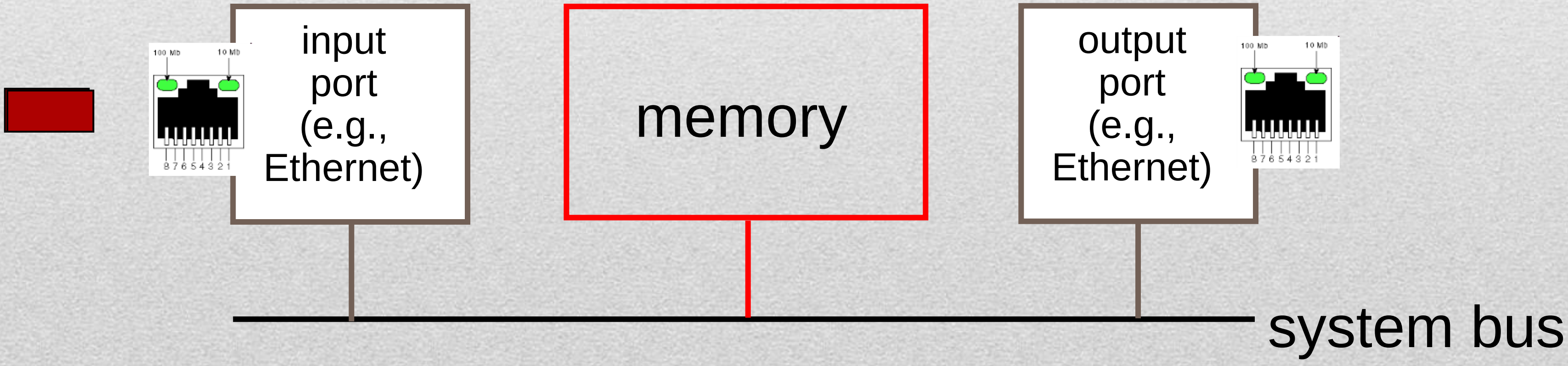
- Transfer packet from input buffer to appropriate output buffer
- Switching rate: rate at which packets can be transfer from inputs to outputs
  - Often measured as multiple of input/output line rate
  - N inputs: switching rate N times line rate desirable
- Three types of switching fabrics



Switching via Memory

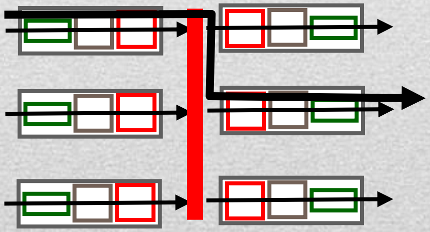
First generation routers:

- Traditional computers with switching under direct control of CPU
- Packet copied to system's memory
- Speed limited by memory bandwidth (2 bus crossings per datagram)



Switching Via A Bus

- Datagram from input port memory to output port memory via a shared bus
- Bus contention
  - Switching speed limited by bus bandwidth
- 32 Gbps bus, Cisco 5600: sufficient speed for access and enterprise routers



bus

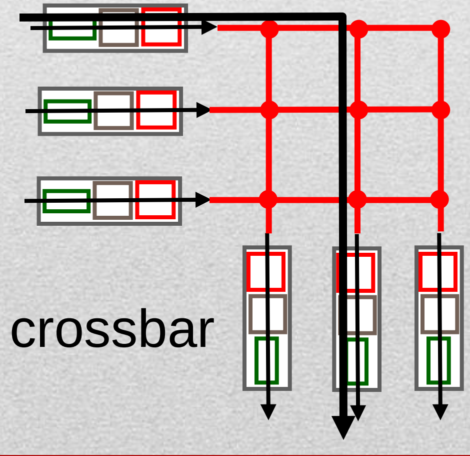
Switching Via Interconnection Network

- Overcome bus bandwidth limitations

- Banyan networks, crossbar, other interconnection nets initially developed to connect processors in multiprocessor

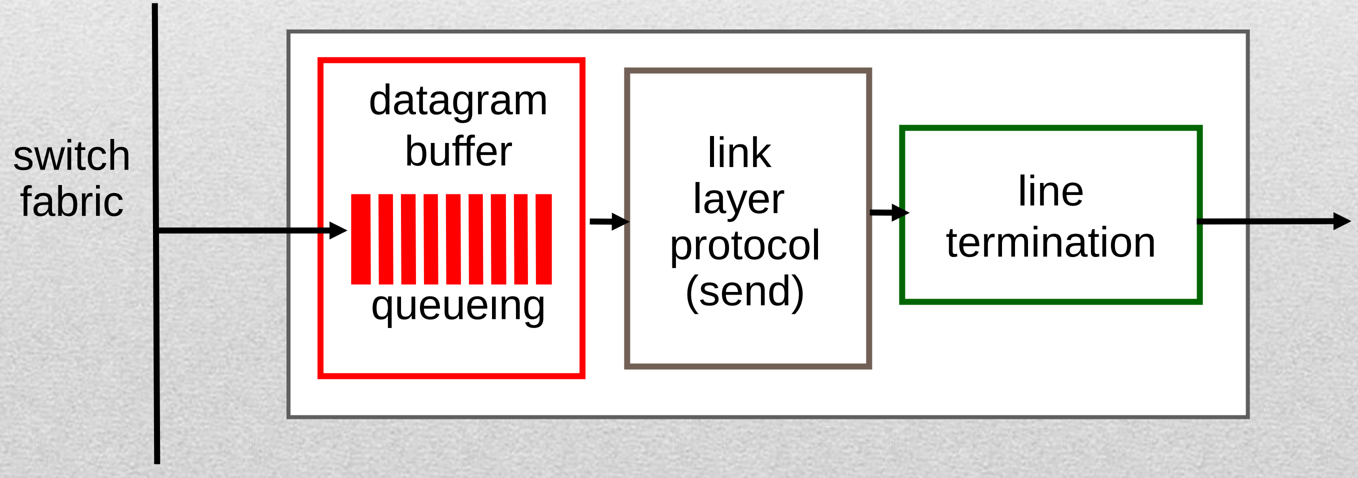
- Advanced design
  - Fragmenting datagram into fixed length cells, switch cells through the fabric.

- Cisco 12000: switches 60 Gbps
  through the interconnection network



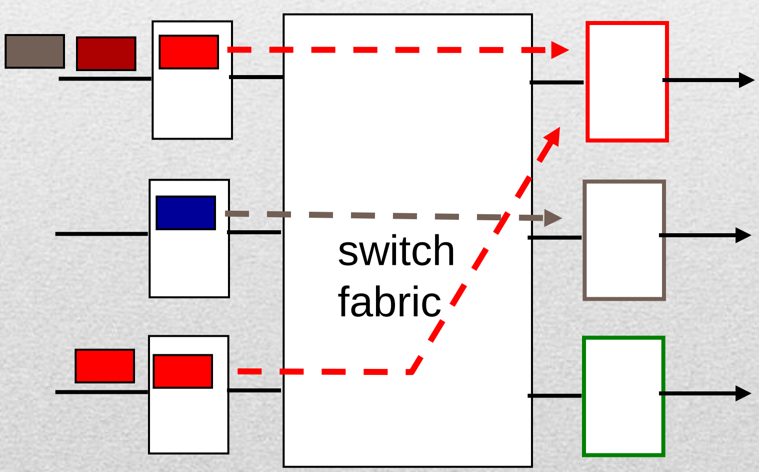
Output Ports

- **Buffering** required when datagrams arrive from fabric faster than the transmission rate
- **Scheduling discipline** chooses among queued datagrams for transmission

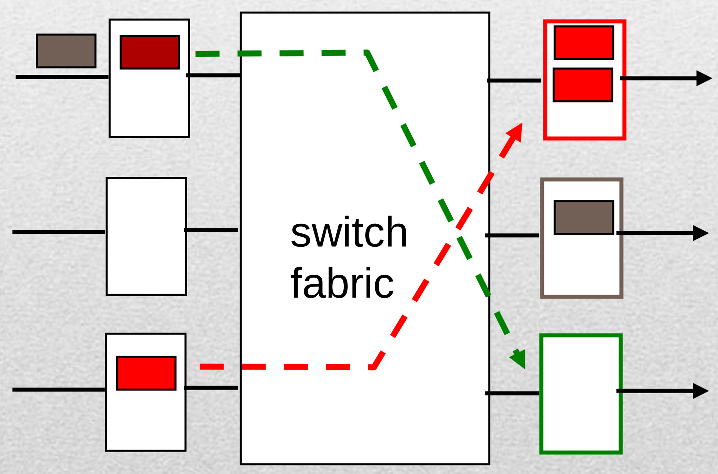


Output Port Queuing

- Buffering when arrival rate via switch exceeds output line speed
- Queuing (delay) and loss due to output port buffer overflow!



at *t*, packets more from input to output



one packet time later

How Much Buffering?

- RFC 3439 rule of thumb: average buffering equal to "typical" RTT (say 250 msec) times link capacity C
  - E.g., C = 10 Gpbs link: 2.5 Gbit buffer

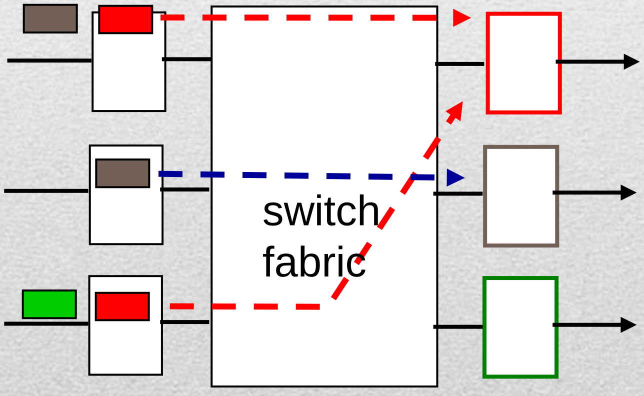
- Recent recommendation: with *N* flows, buffering equal to

$$\frac{RTT \cdot C}{\sqrt{N}}$$

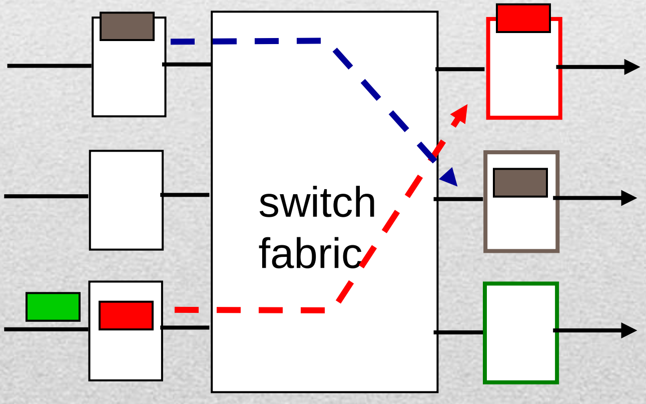
Input Port Queuing

- Fabric slower than input ports combined → queueing may occur at input queues
  - *Queueing delay and loss due to input buffer overflow!*

- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward



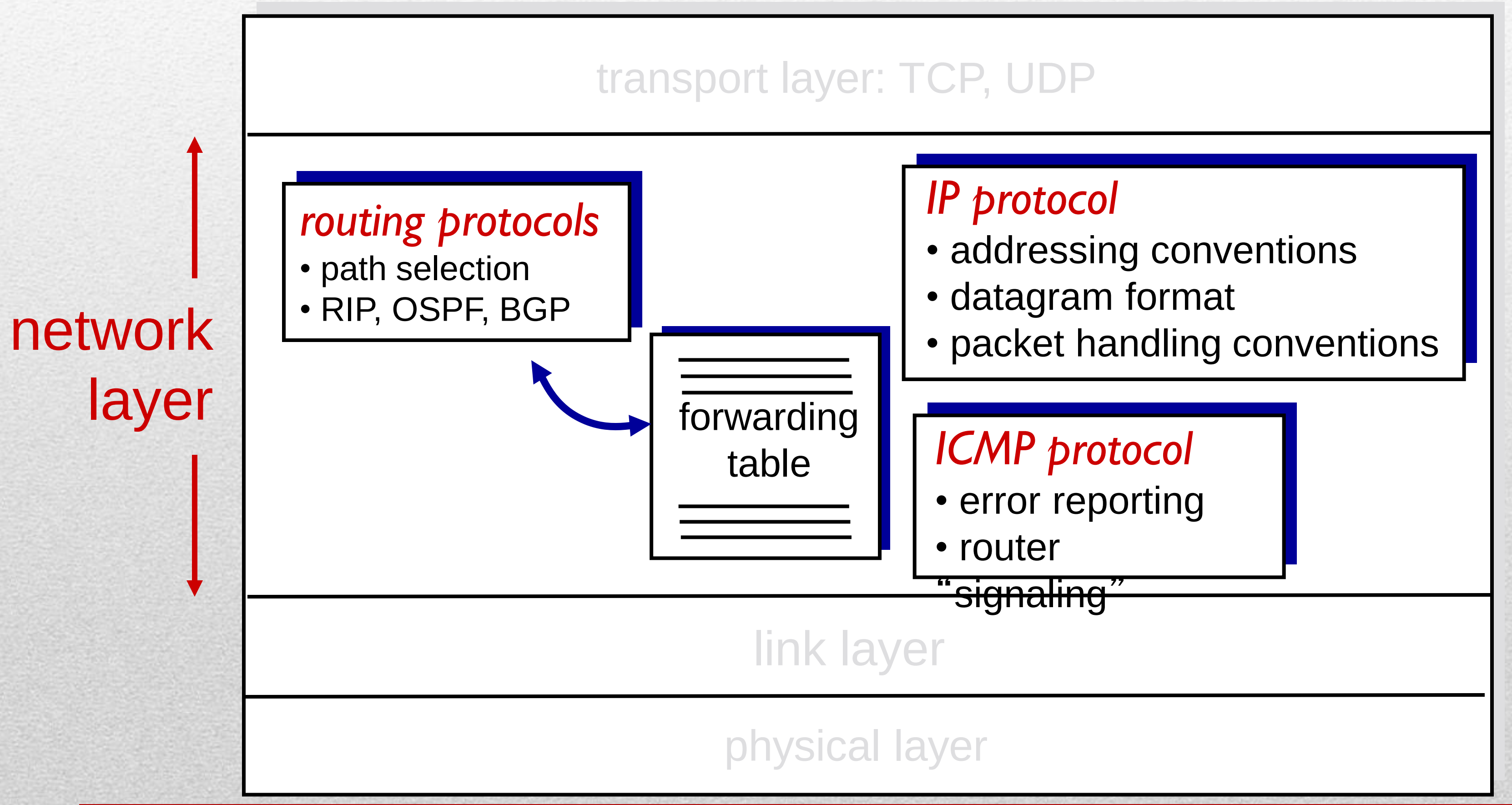
output port contention:
only one red datagram can be transferred.
lower red packet is blocked



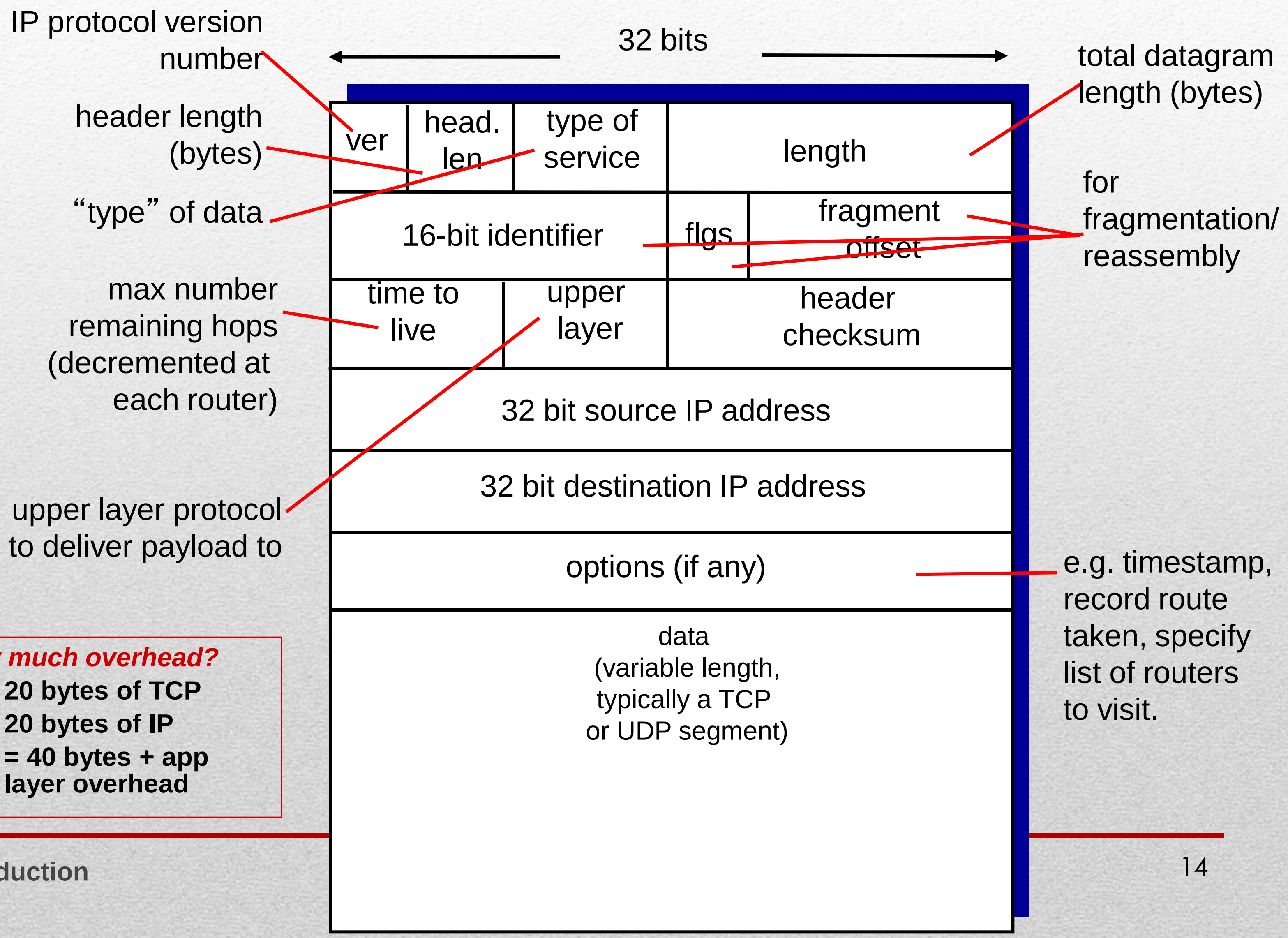
one packet time later: green packet experiences HOL blocking

The Internet Network Layer

Host, router network layer functions:



IP Datagram

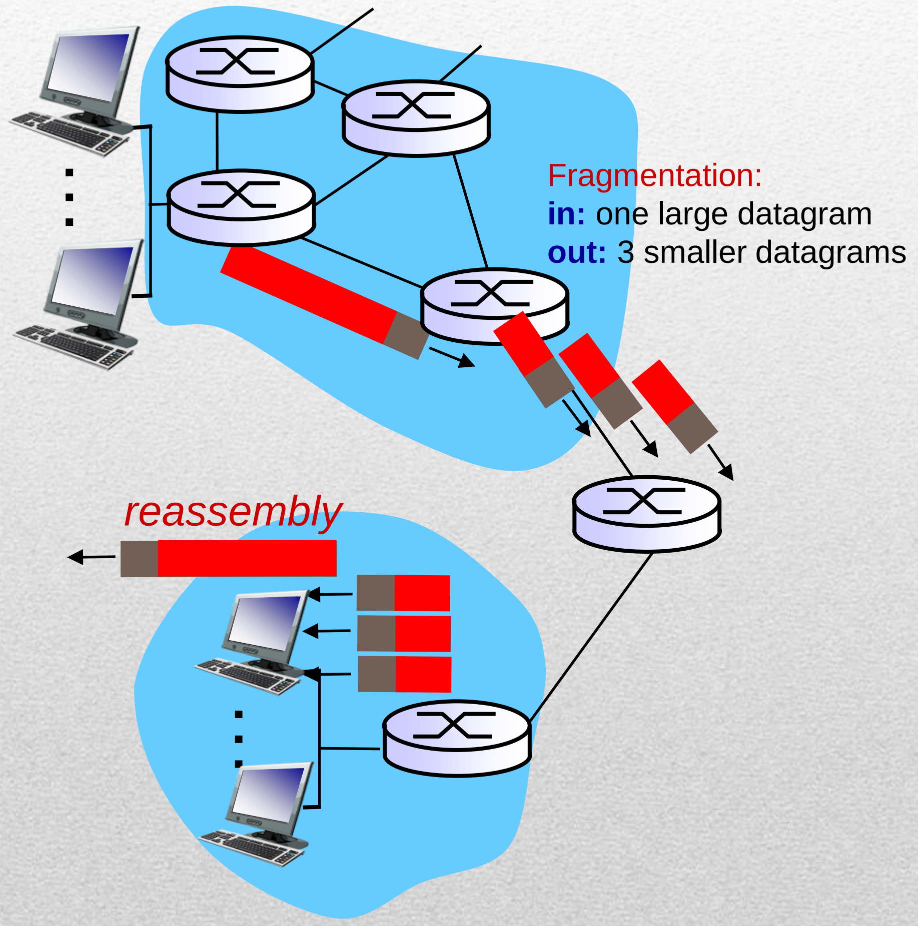


How much overhead?

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

IP Fragmentation & Reassembly

- Network links have MTU (maximum 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