Network layer – data plane

CE 352, Computer Networks

Salem Al-Agtash

Lecture 12

Slides are adapted from Computer Networking: A Top Down Approach, 7th Edition © J.F Kurose and K.W. Ross

Recall (Course objectives)

Develop the ability to:

- Describe the layered network architecture, the services provided by each layer, and the respective protocols (Application,..) that deliver these services, including wireless protocols.
- Compare and contrast mechanisms used to improve network performance.
- Describe security threats and mechanisms and protocols used to protect network communications.
- Explain how network mechanisms and protocols (Application,..) have evolved to accommodate the growth of the internet and the use of new applications and technology.
- Write network software using sockets.

Recall (Layers)

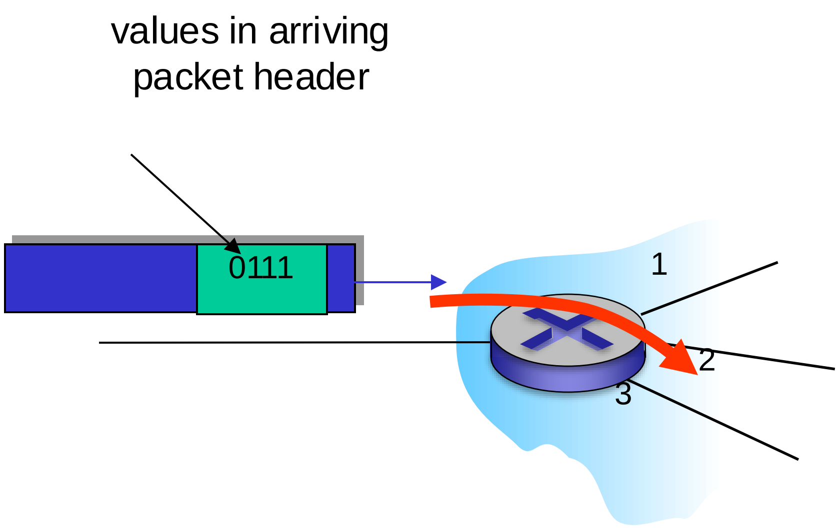
- Protocols, reference model, layers
- Internet protocol stack
- Application layer, Transport layer

Today

- Network Layer
  - Internals
  - Forwarding
  - Addressing
  - IP protocol details
  - Routing
  - Much more …

Application	Web Browser (HTTP) Email (SMTP, IMAP, POP3) Remote Login (Telnet) Name Resolution (DNS) File Transfer (FTP, TFTP)	Bits
Transport	Transmission Control Protocol (TCP) User Datagram Protocol (UDP)	Port: 16
Network	Internet Protocol IP, SNMP, ICMP, DHCP OSPF, BGP	IPv4: 32 OR IPv6: 128
Data Link	Ethernet Token Ring FDDI WAN Protocols	MAC : 48
Physical	Copper Twisted Pair Fiber Optic Radio	

Network layer: data plane, control plane



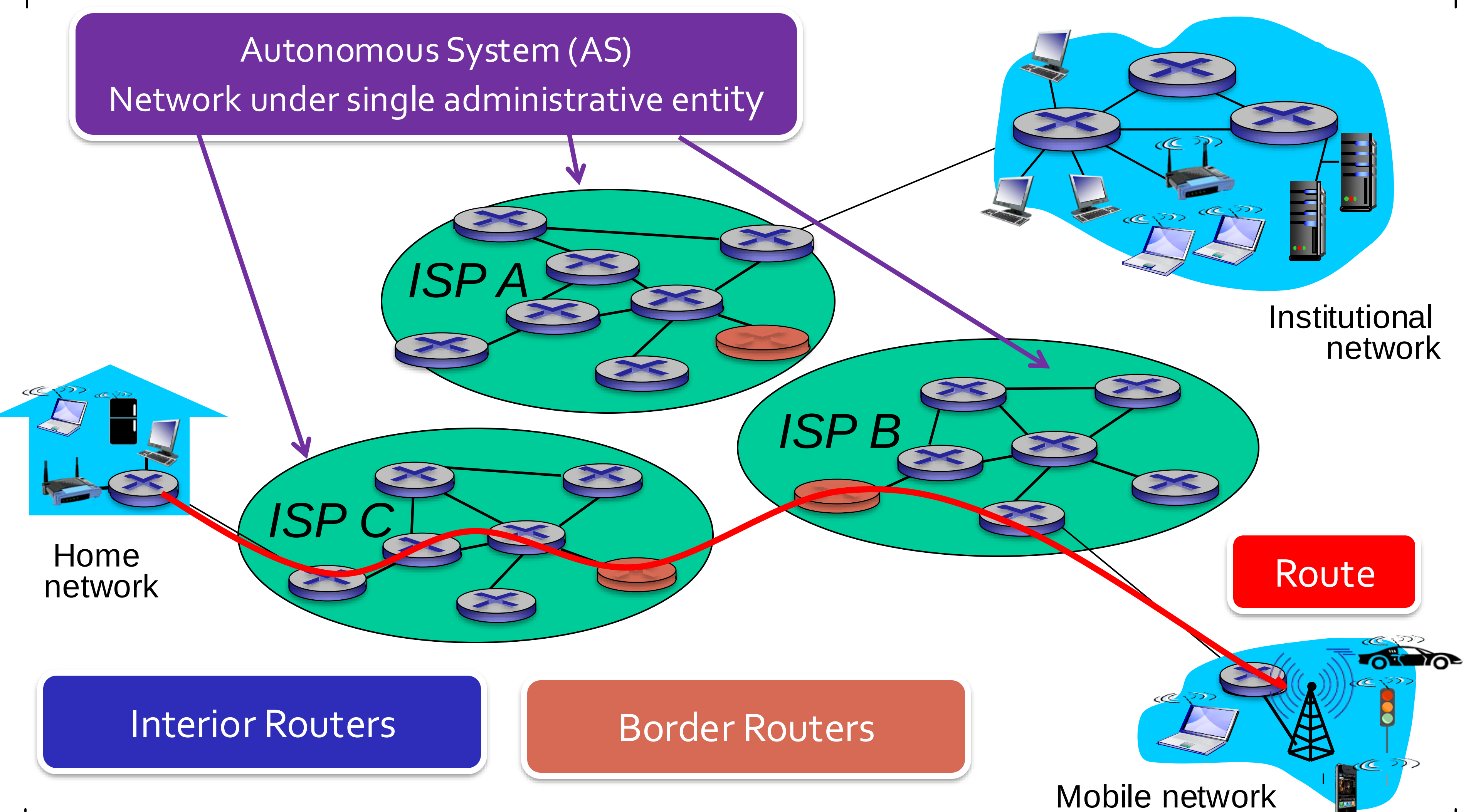
Data plane (chapter 4) : [Forwarding]

- local, per-router function
- determines how datagram arriving on router input port is forwarded to router output port
- forwarding function

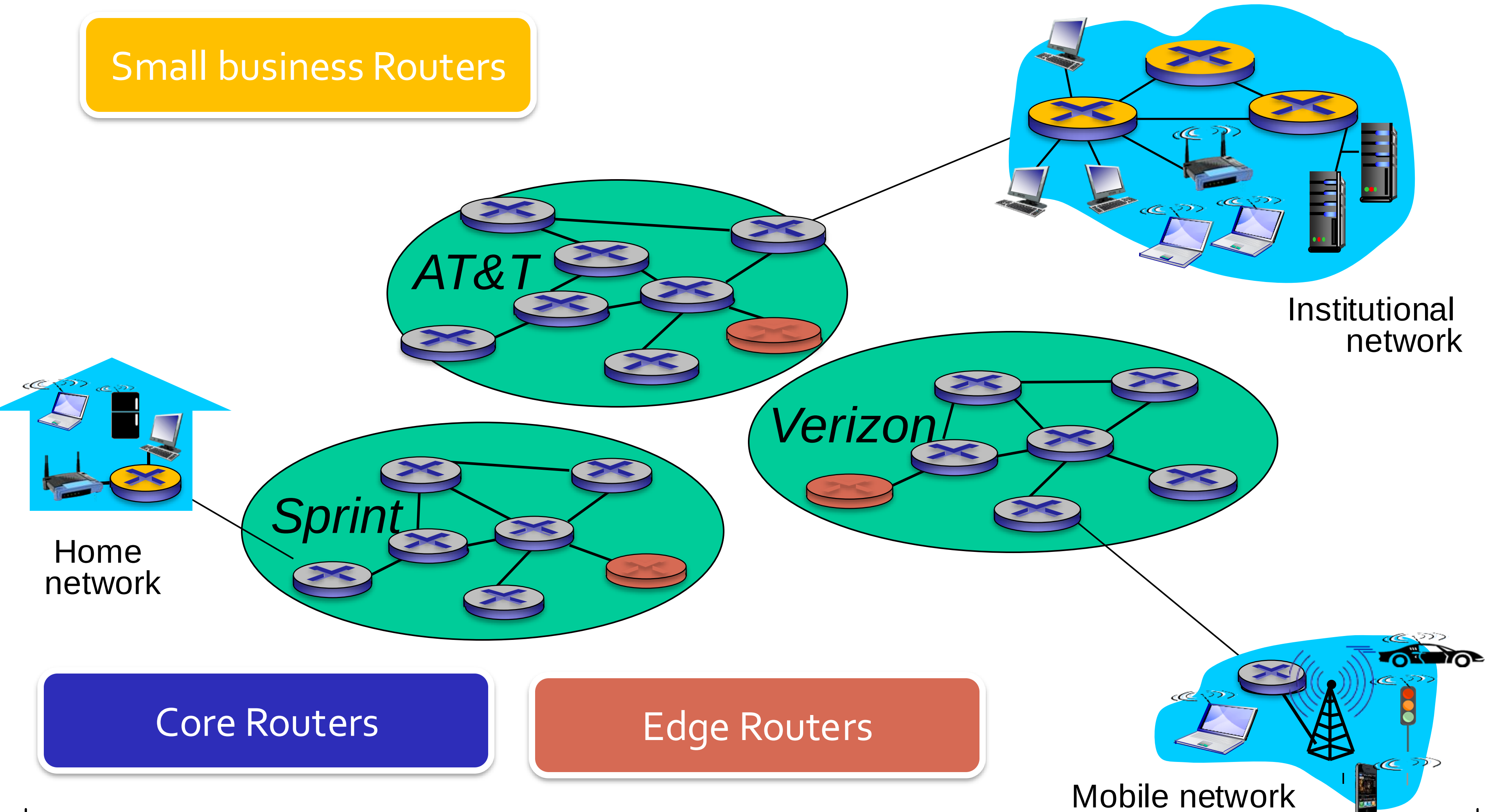
Control plane (chapter 5): [Routing]

- network-wide logic
- determines how datagram is routed among routers along end-end path from source host to destination host
- two control-plane approaches:
  - *traditional routing algorithms:* implemented in routers
  - *software-defined networking (SDN)*: implemented in (remote) servers

Important context

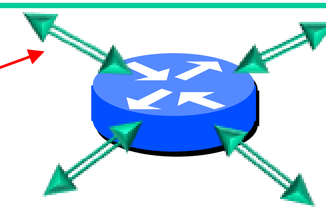


Technologies



Routers (network/ layer 3 device)

- Building blocks of the Internet infrastructure
- Vendors: Cisco, Juniper, Alcatel-Lucent, Huawei
- Number of ports (interfaces) and and line rate (bps)



- Core routers

Cisco CRS - 8

10/40/100 Gbps)

Juniper T4000

10/40 Gbps)

- Edge routers

Cisco ASR 1006

1/10 Gbps)

Juniper M120

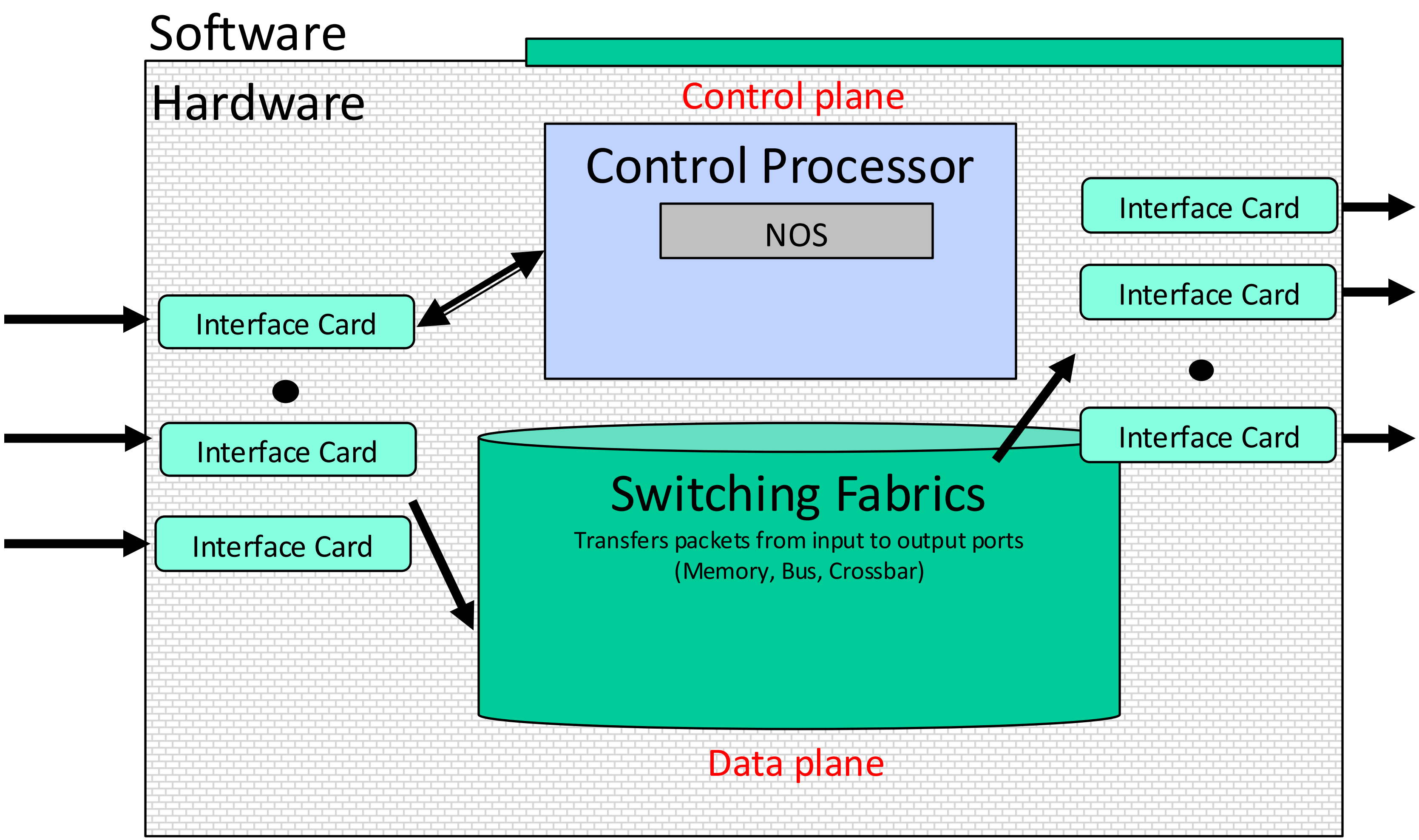
2.5/10 Gbps)

- Small business routers

Cisco 3945E

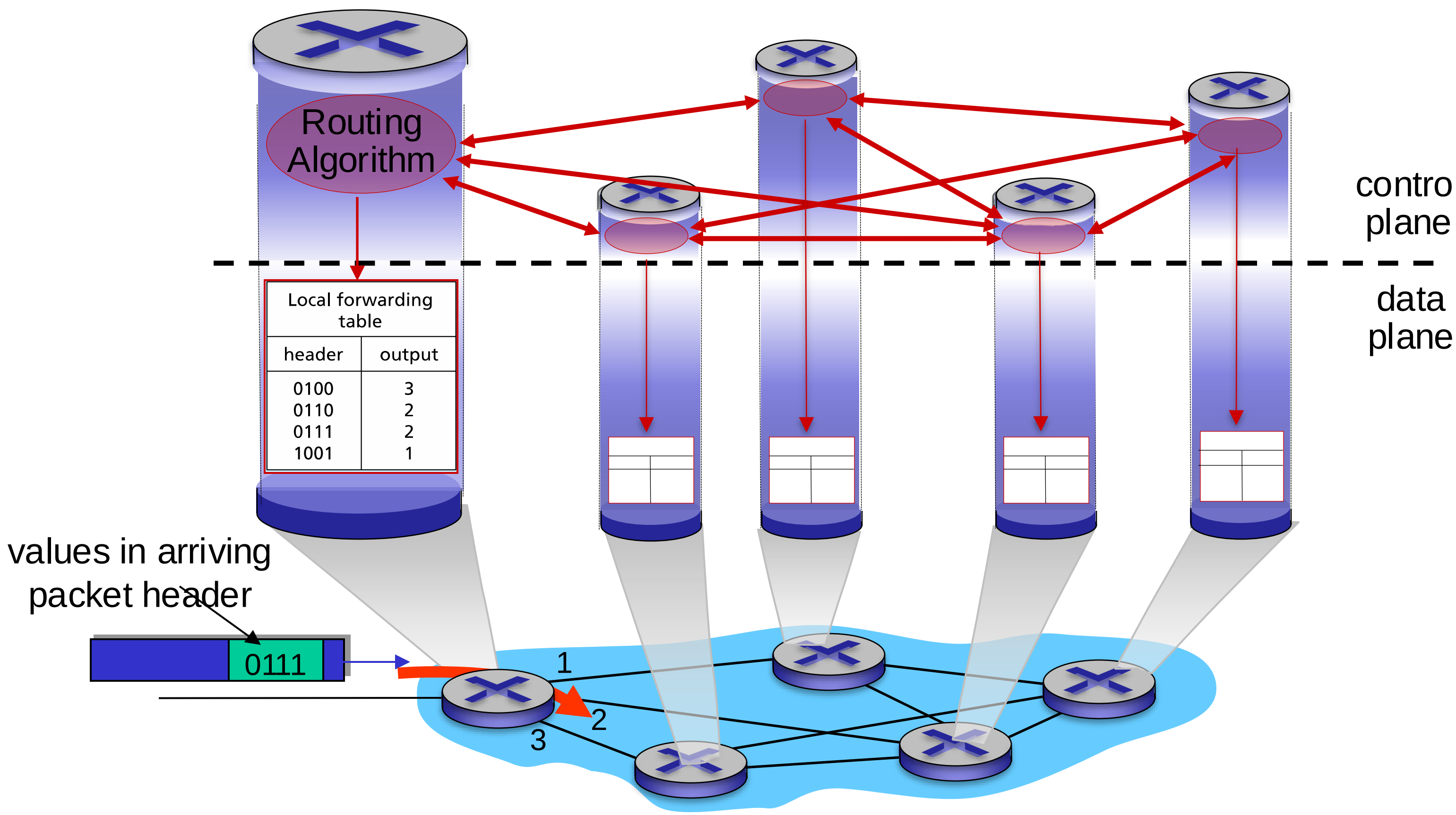
10/100/1000 Mbps)

Inside routers



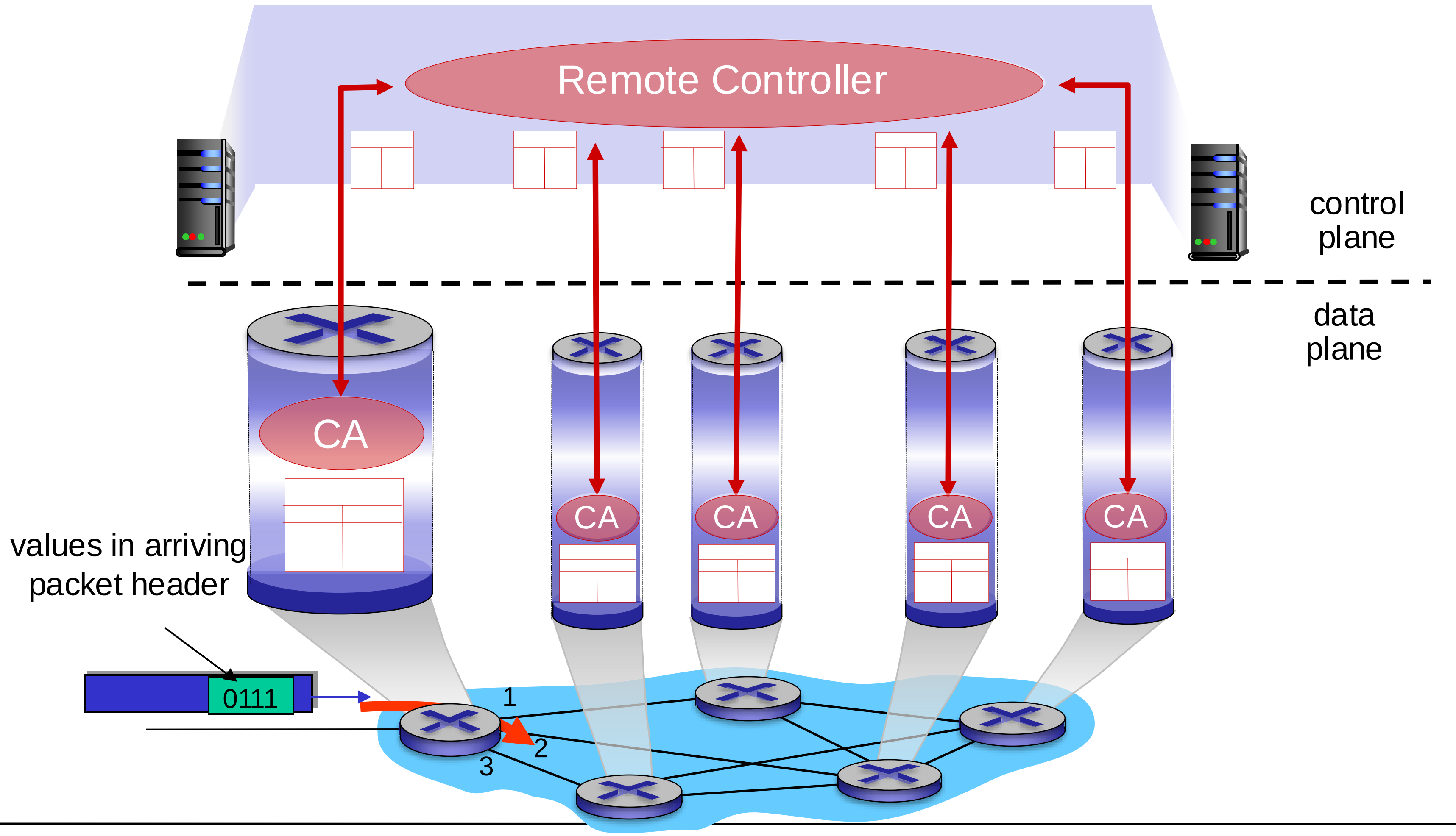
Per-router control plane

Individual routing algorithm components *in each and every router* interact in the control plane

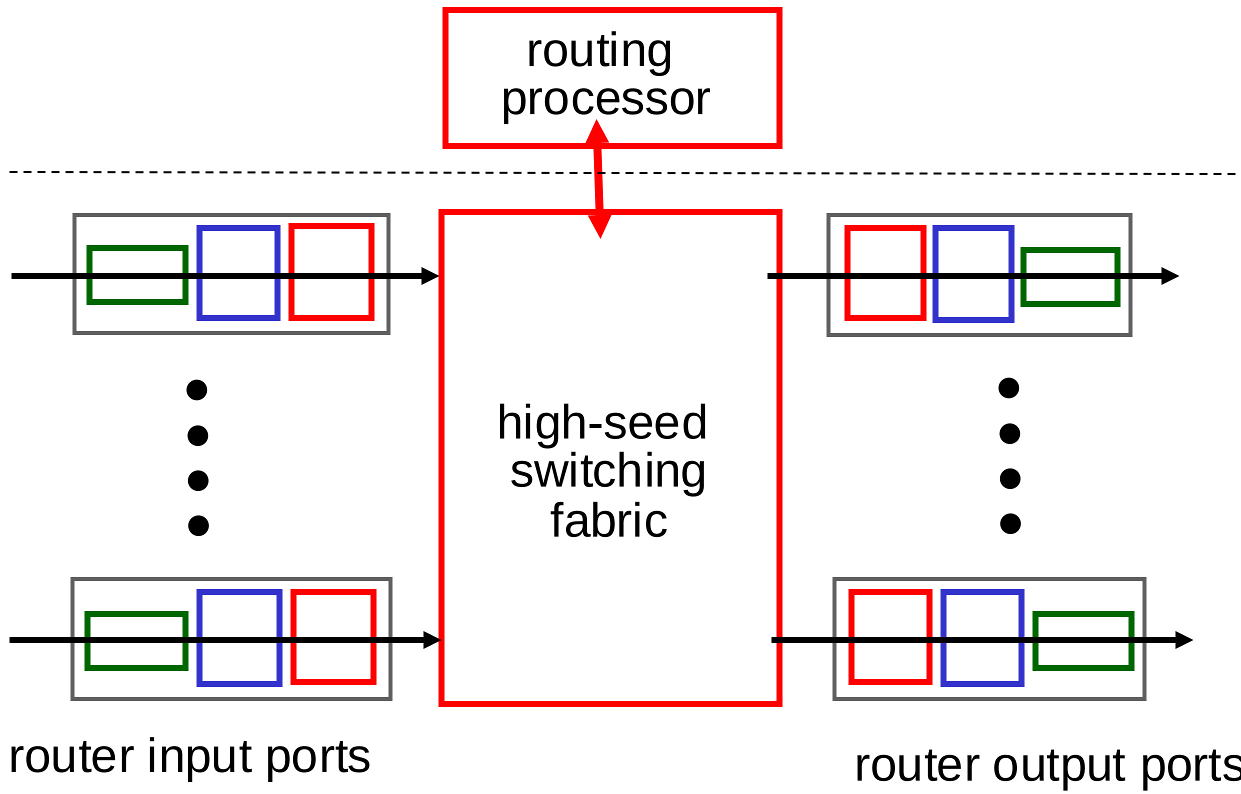


Logically centralized control plane

A distinct (typically remote) controller interacts with local control agents (CAs)



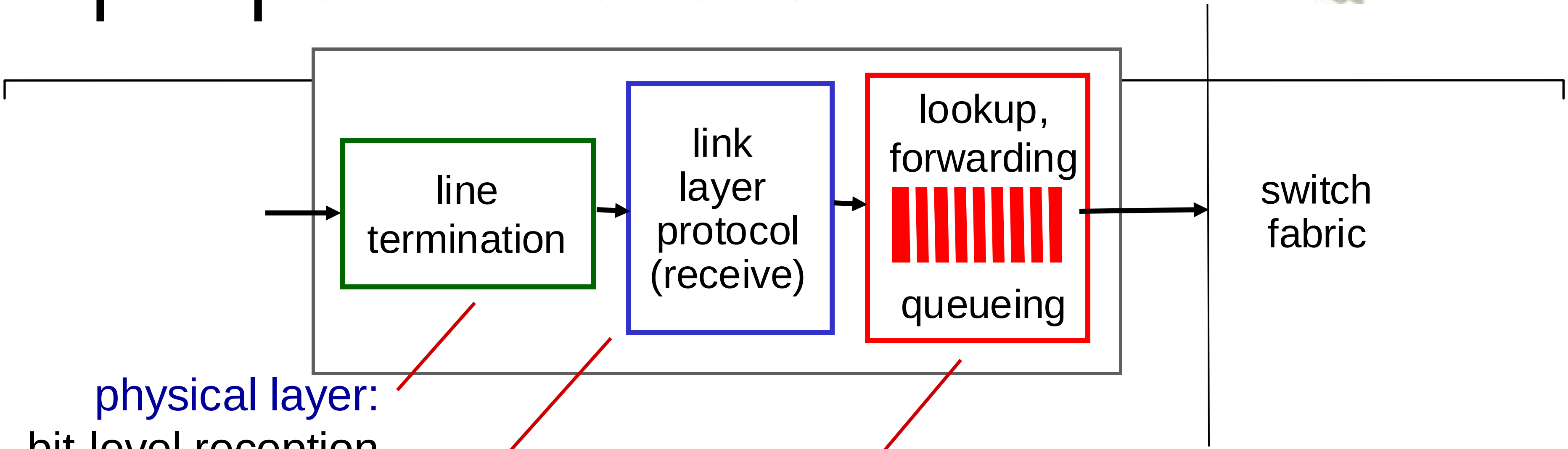
Router high-level details



routing, management
control plane (software)
operates in millisecond
time frame

forwarding data plane
(hardware) operates
in nanosecond
timeframe

Input port functions



physical layer:
bit-level reception

data link layer:
e.g., Ethernet
(Chapter 6)

decentralized switching:

- Using header field values, lookup output port using forwarding table in input port memory *("match plus action")*
- Goal: complete input port processing at 'line speed' -- *destination-based or generalized forwarding*
- Queuing: if datagrams arrive faster than forwarding rate into switch fabric

Destination-based forwarding

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	3

Prefix matching

longest prefix matching

when looking for forwarding table entry for given destination address, use *longest* address prefix (network ID) 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?

Prefix matching

longest prefix matching

when looking for forwarding table entry for given destination address, use *longest* address prefix (network ID) 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? 0

DA: 11001000 00010111 00011000 10101010 which interface? 1

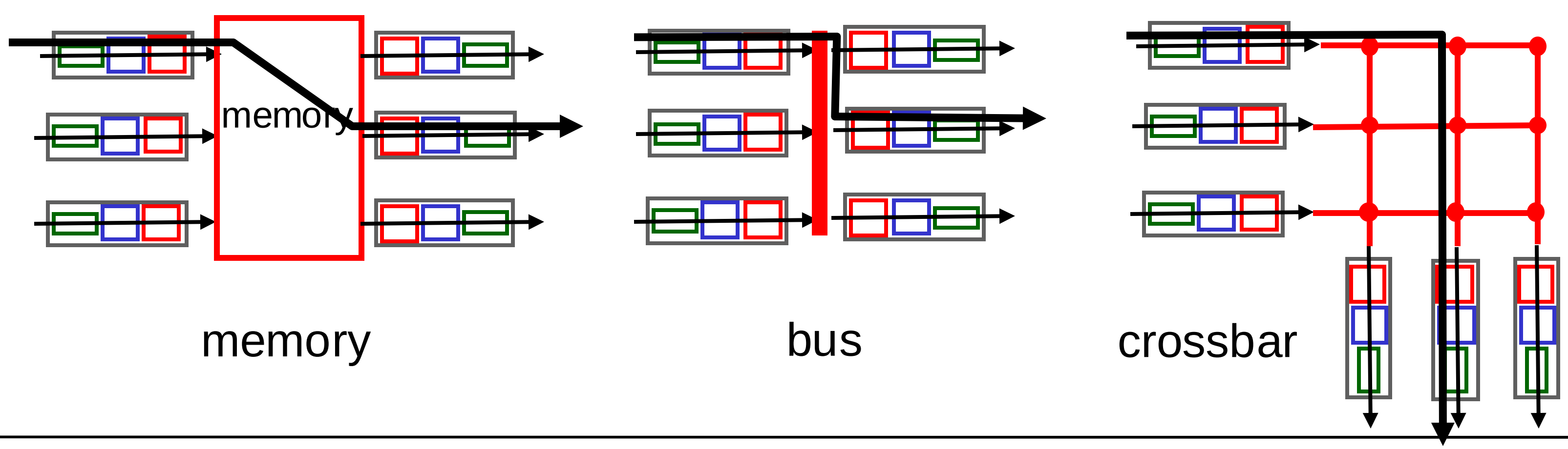
Switching fabrics

Transferring packet from input buffer to appropriate output buffer

Switching rate: rate at which packets can be transferred from inputs to outputs

- often measured as multiple of input/output line rate
- N inputs: switching rate is then N times line rate

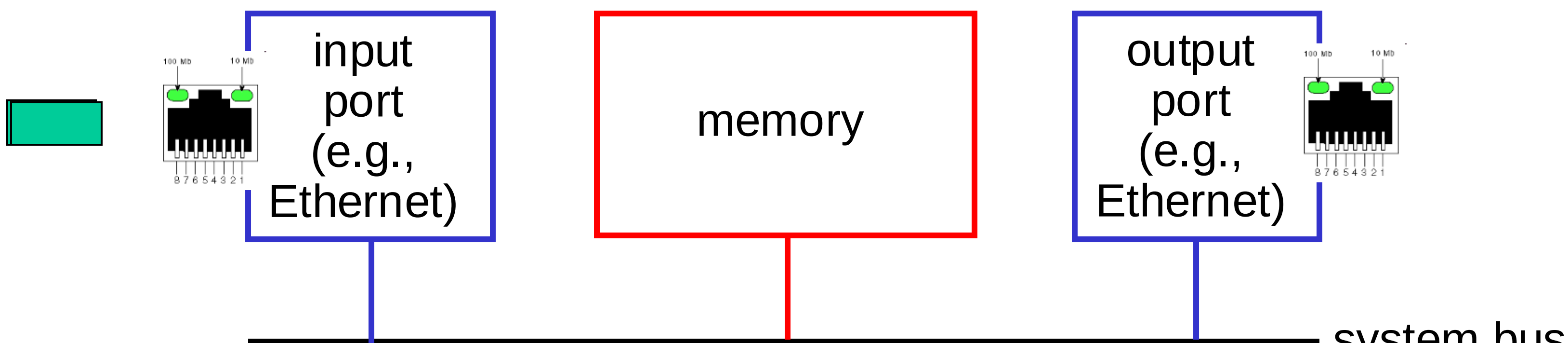
Three types of switching fabrics



Switching via memory

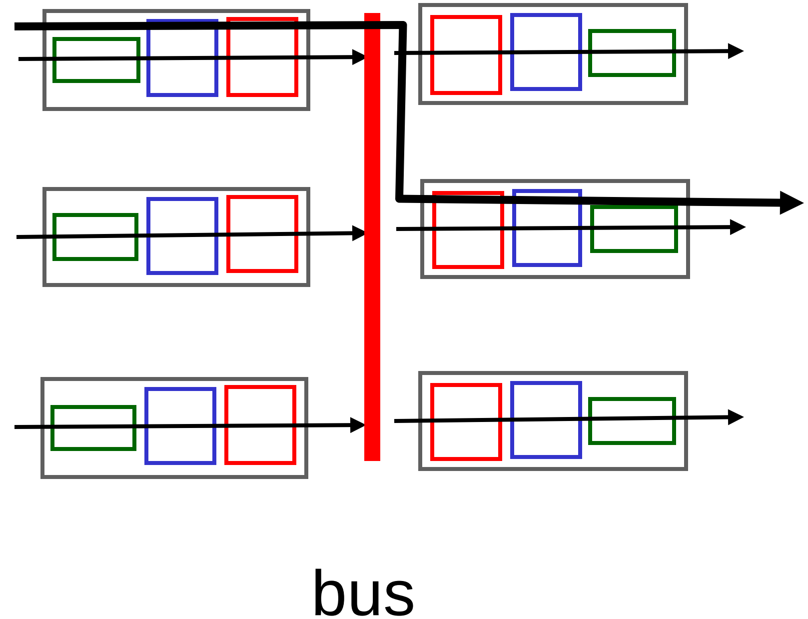
First generation routers:

- Traditional computers with switching under direct control of CPU
- Packet copied to system's memory
- Speed limited by memory bandwidth



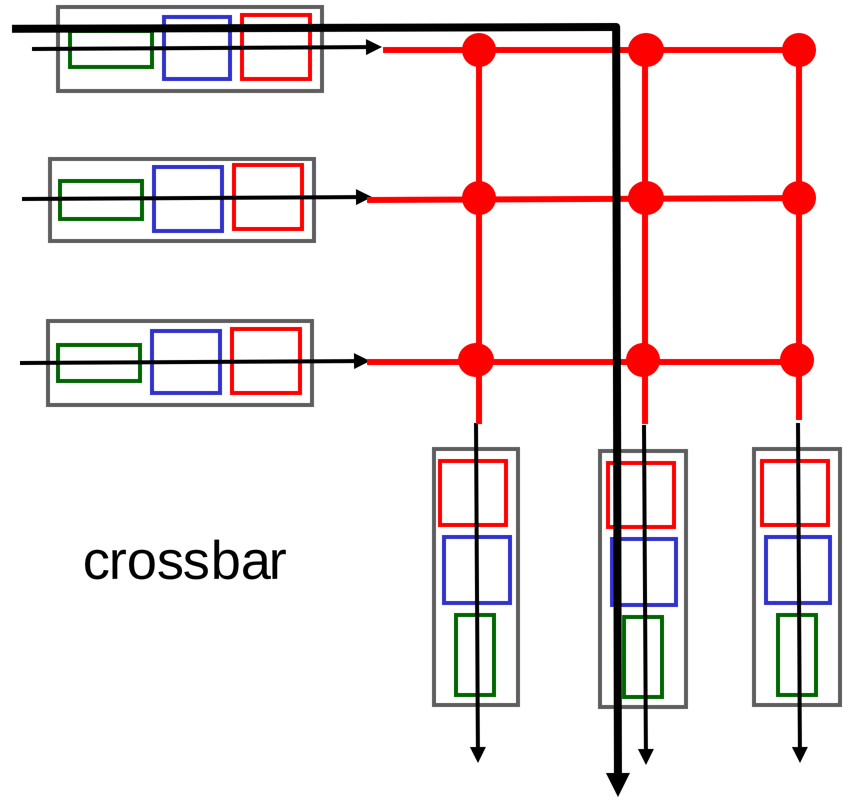
Switching via a bus

- 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



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

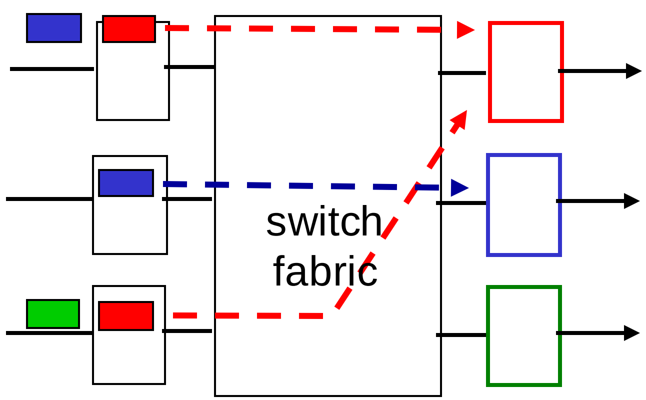


Input port queuing

fabric slower than input ports combined -> queueing may occur at input queues

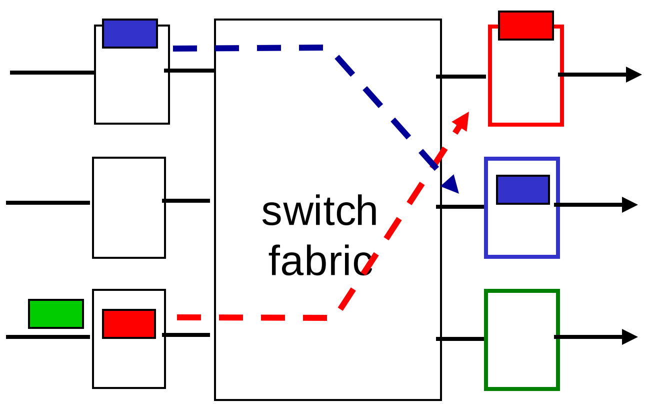
- *queueing delay and loss due to input buffer overflow!*

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



output port contention:

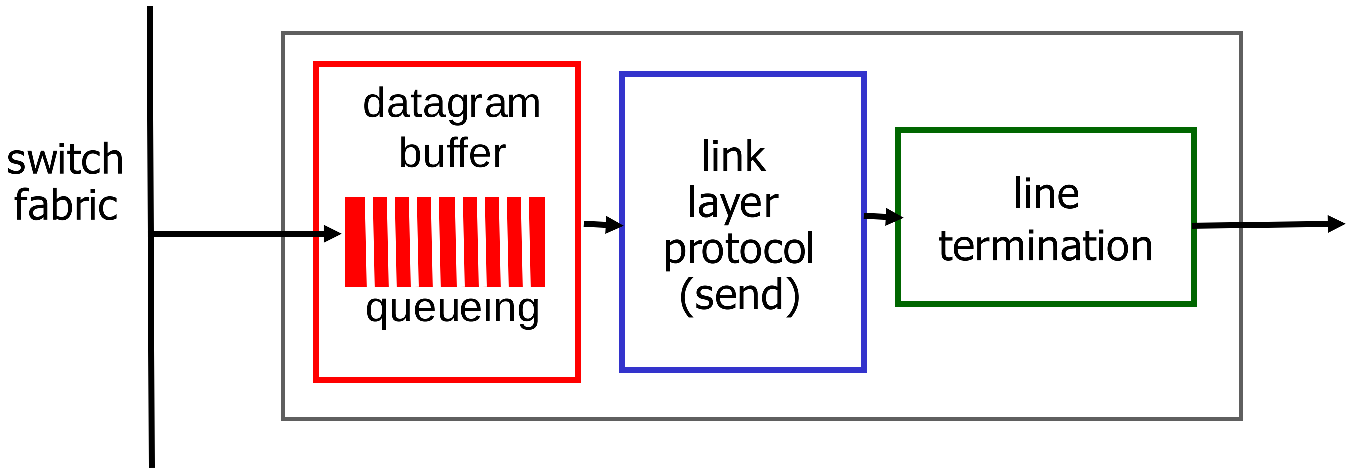
- only one red datagram can be transferred.
- *lower red packet is blocked*



one packet time later:

- green packet experiences HOL blocking

Output ports



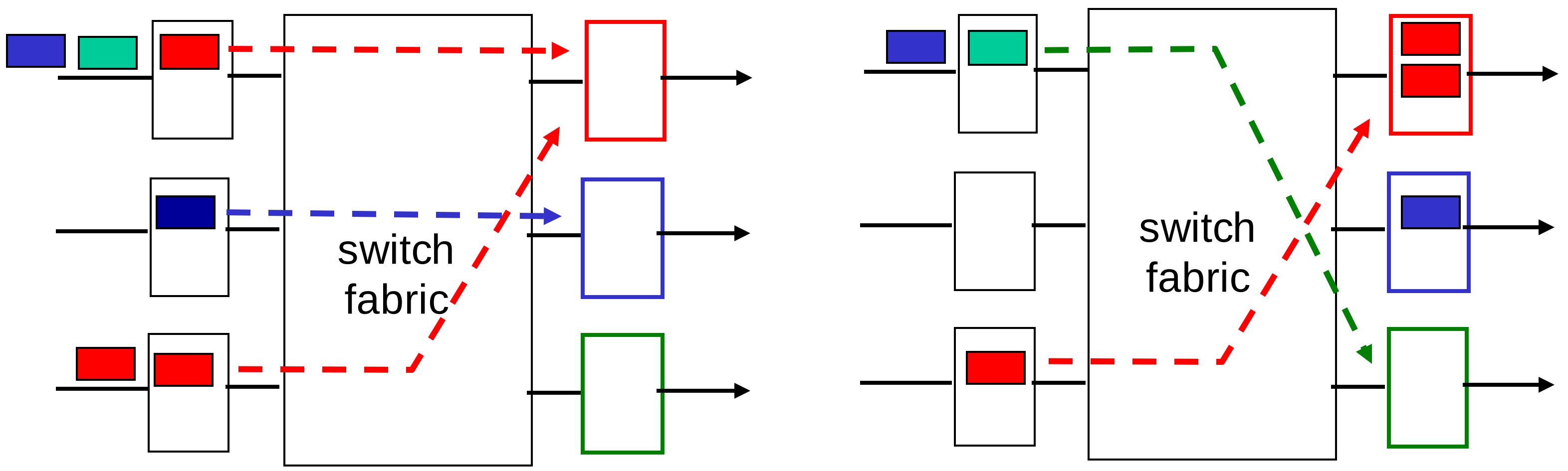
- *Buffering* required when datagrams arrive from fabric faster than the transmission rate

Datagram (packets) can be lost due to congestion, lack of buffers

- *Scheduling discipline* chooses among queued datagrams for transmission

Priority scheduling – who gets best performance, network neutrality

Output port queueing



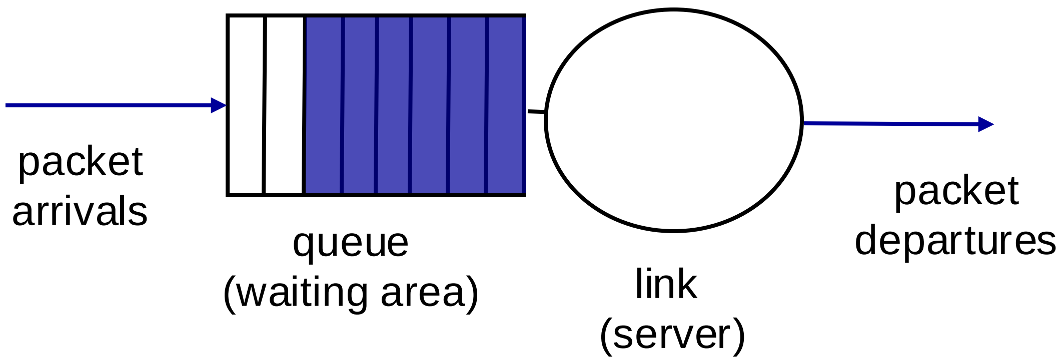
at *t*, packets more
from input to output

one packet time later

buffering when arrival rate via switch exceeds output line speed

queueing (delay) and loss due to output port buffer overflow!

Scheduling mechanisms: FIFO

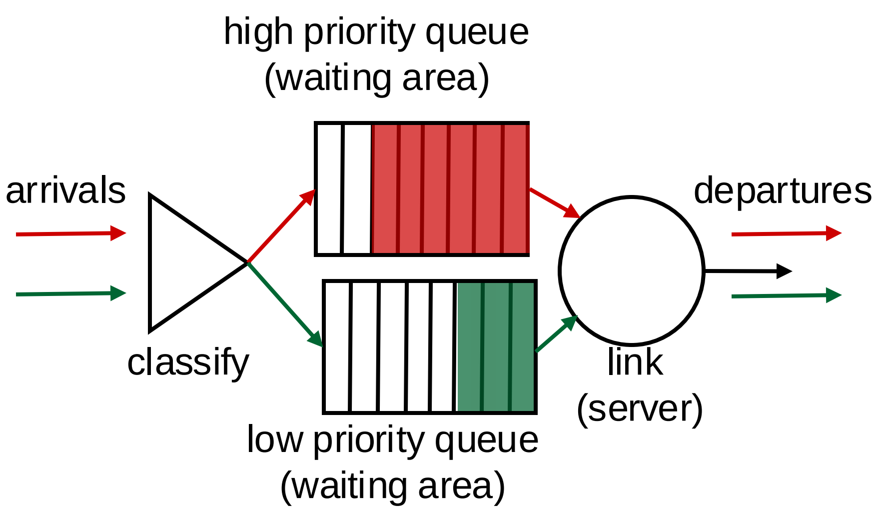


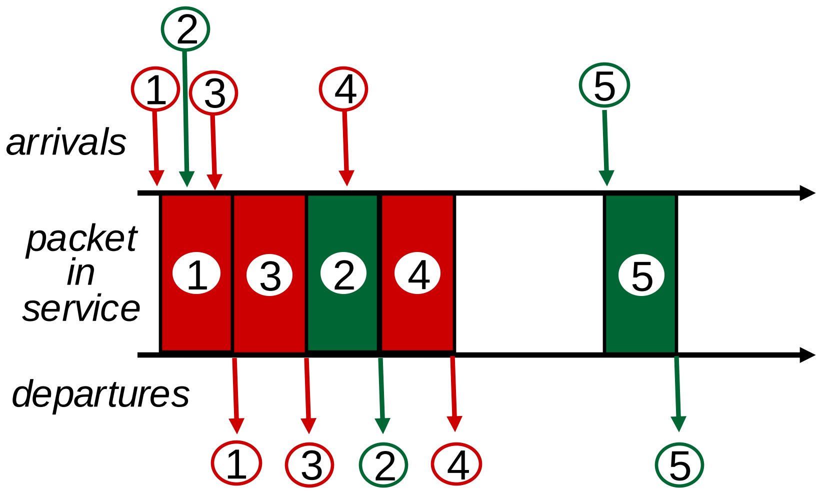
scheduling: choose next packet to send on link

FIFO (first in first out) scheduling: send in order of arrival to queue

- *discard policy:* if packet arrives to full queue: who to discard?
  - *tail drop:* drop arriving packet
  - *priority:* drop/remove on priority basis
  - *random:* drop/remove randomly

Scheduling policies: priority



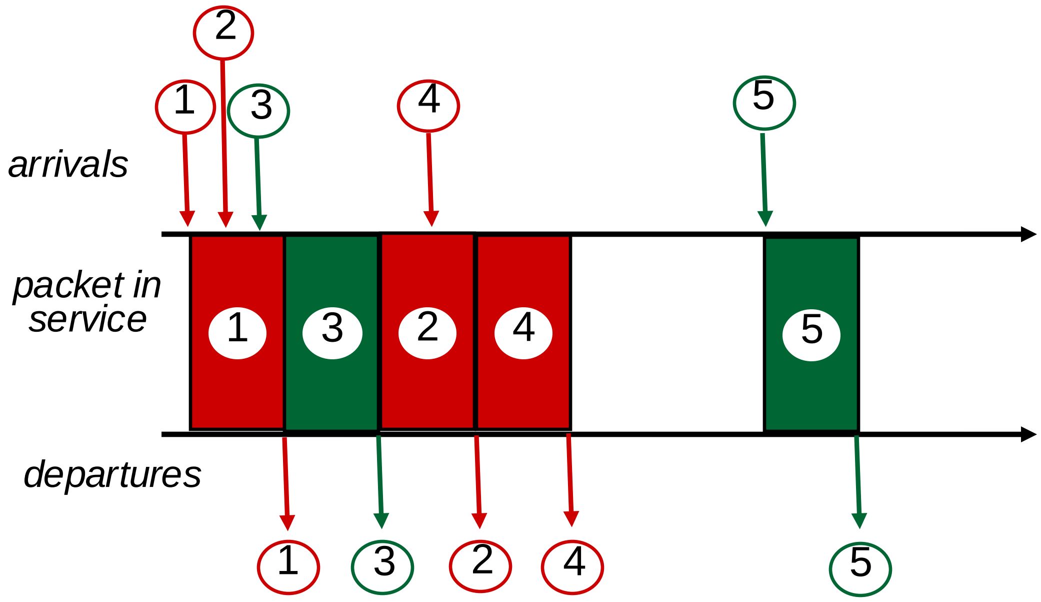


priority scheduling: send highest priority queued packet

multiple *classes*, with different priorities

- class may depend on header info, e.g. IP source/dest, port numbers, etc.

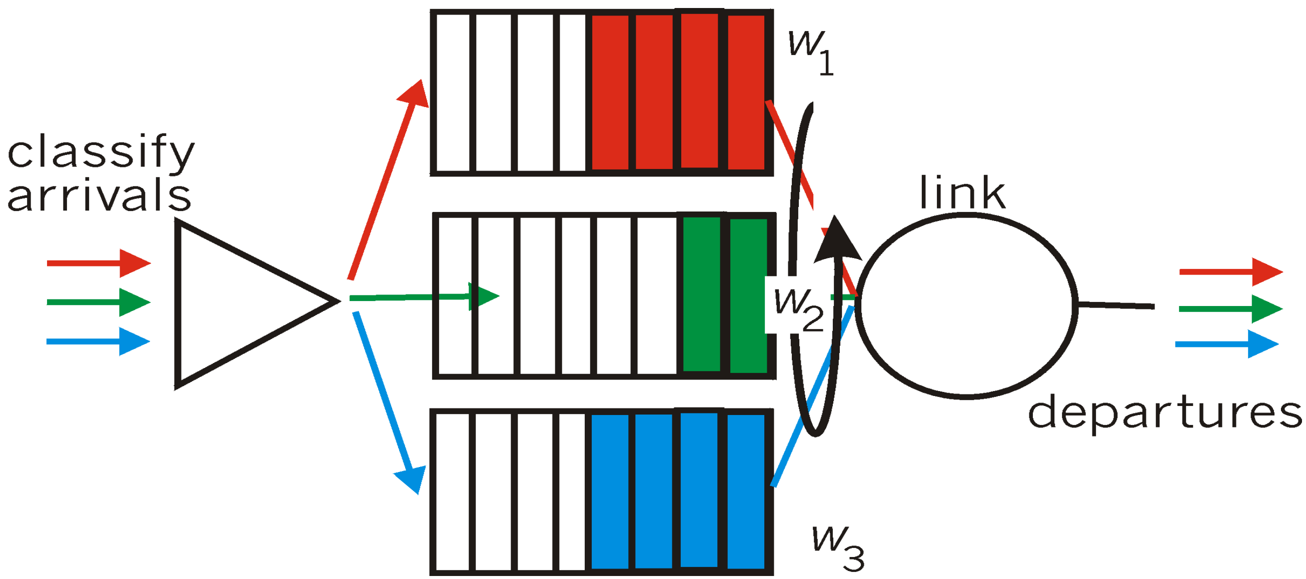
Scheduling policies: RR



Round Robin (RR) scheduling:

- multiple classes
- cyclically scan class queues, sending one complete packet from each class (if available)

Scheduling policies: WFQ



Weighted Fair Queuing (WFQ):

generalized Round Robin

each class gets weighted amount of service in each cycle

Summary

Today:

- Network layer
  - context
  - technologies
  - control plane
  - data plane
- Input/output port functions
- Switching fabrics
- Queueing and scheduling

Canvas discussion:

- Reflection
- Exit ticket

Next time:

- read 4.3 of K&R (IP addressing)
- follow on Canvas! material and announcements

Any questions?