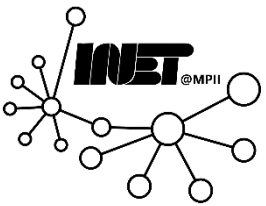
Router Architecture Overview

What's inside a router?

Prof. Anja Feldmann, Ph.D.



What does a Router Look Like?

Ericsson SSR 8020
BNG/BRAS/PGW
- Maximum 16 Tbit/s

Cisco CRS-1
Core Router
- 2.2Tbit/s single chassis
- Up to 322Tbit/s for multichassis

Juniper T4000
Core Router
- 3.8 Tbit/s single chassis

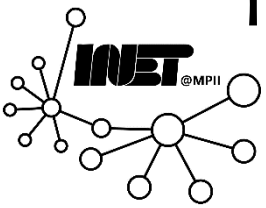
And like this:

Dlink DIR-615 Wireless N 300 Home Router
- LAN: 4x 10/100Mbit/s Ports
- WAN: 1x 10/100Mbit/s Port
- Up to 300 Mbit/s throughput
- WiFi support

Belkin (formerly Cisco, Linksys) N600 DB Wireless Dual-Band N+ Home Router
- LAN: 4x 10/100Mbit/s Ports
- WAN: 1x 10/100Mbit/s Port
- Up to 300 Mbit/s throughput
- WiFi support

Who Makes Core Routers?

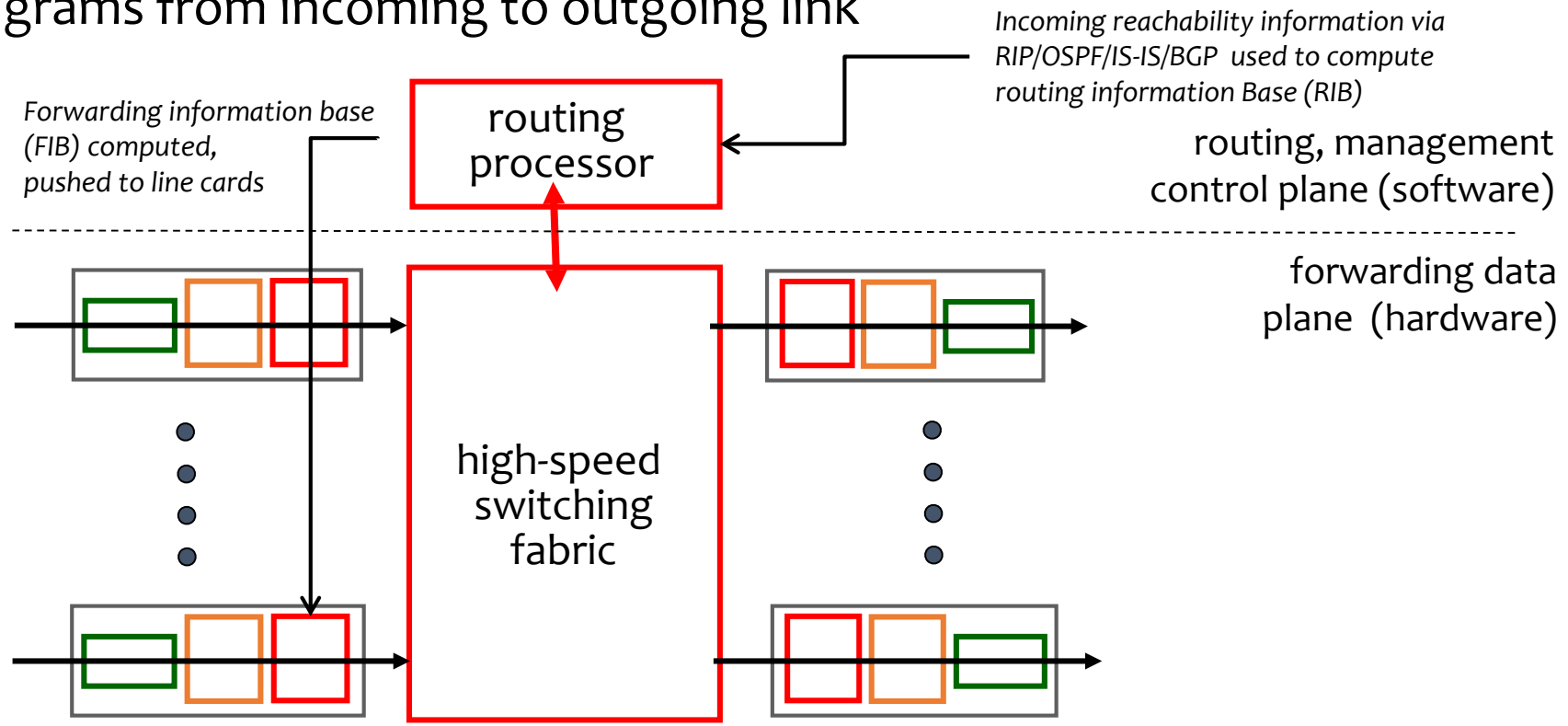
- Cisco
  - CRS (Carrier Router Series)
- Juniper
  - T-series
- Nokia (used to be Alcatel-Lucent)
  - XRS (Extensible Routing System)
- Huawei
  - Netengine
- Others manufacture aggregation/access networking gear, e.g., for edge deployments



Router architecture overview

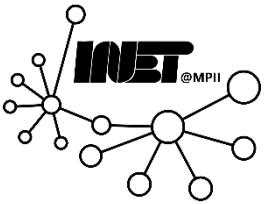
Two key router functions:

- Run routing algorithms/protocol (RIP, OSPF, BGP)
- Forwarding datagrams from incoming to outgoing link



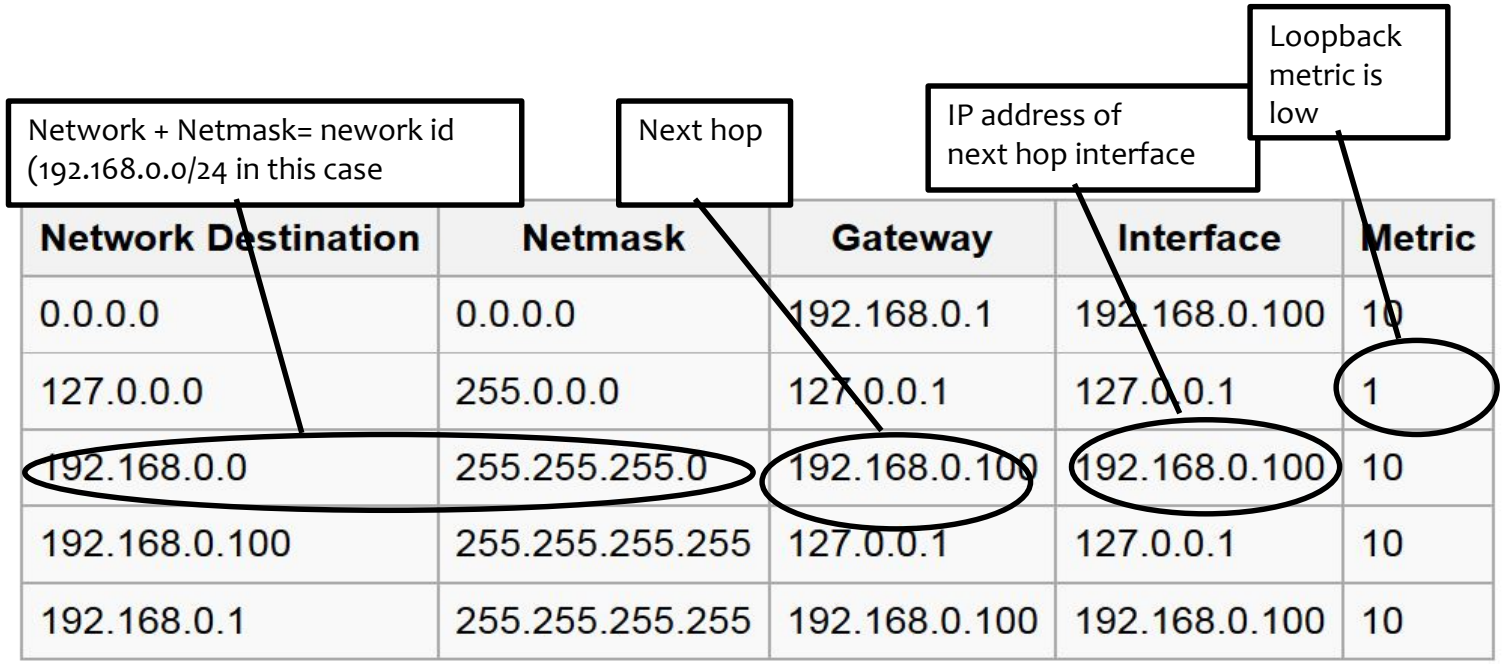
RIB (Routing Information Base)

- Router can contain many different RIBs
  - One for each routing protocol
  - Usually consolidated into one global RIB or into FIB
  - (End system IP addresses (/32s) populated through ARP for default gateway MAC)
- Minimum contents
  - Network id of destination subnet
  - Cost/metric for hop
  - Next hop gateway or end system
- Other information
  - Quality of service, e.g., a U if the link is up
  - Access control lists for security
  - Interface, such as eth0 for first Ethernet line card, etc.



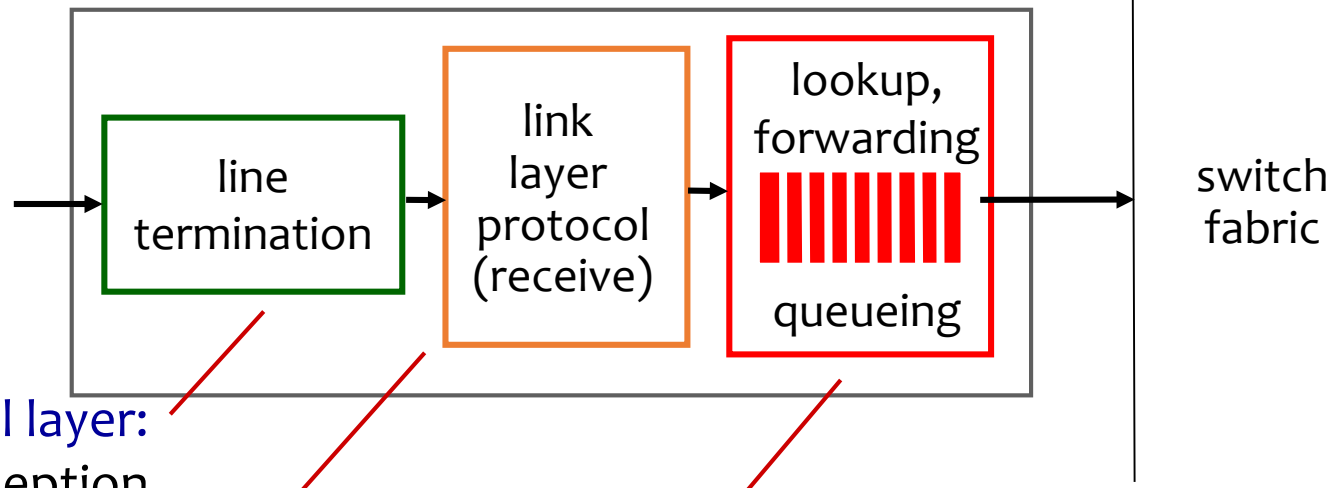
RIB

- Router can contain many different RIBs
- Minimum contents
- Other information



Network Destination	Netmask	Gateway	Interface	Metric
0.0.0.0	0.0.0.0	192.168.0.1	192.168.0.100	10
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
192.168.0.0	255.255.255.0	192.168.0.100	192.168.0.100	10
192.168.0.100	255.255.255.255	127.0.0.1	127.0.0.1	10
192.168.0.1	255.255.255.255	192.168.0.100	192.168.0.100	10

Input port functions



Physical layer:
Bit-level reception

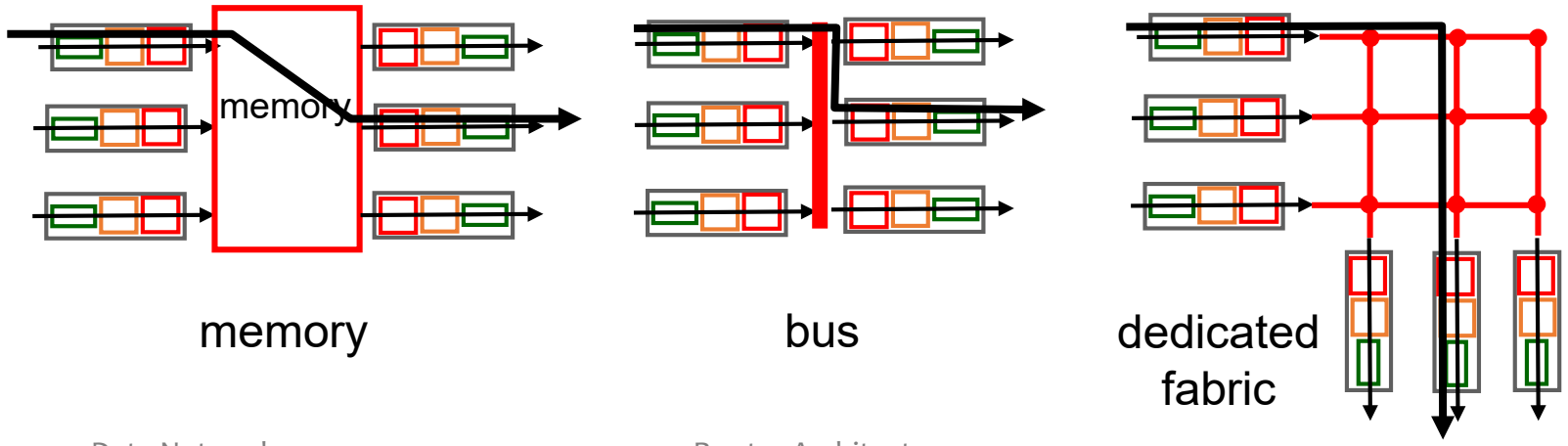
Data link layer:
E.g., Ethernet
see chapter 5

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

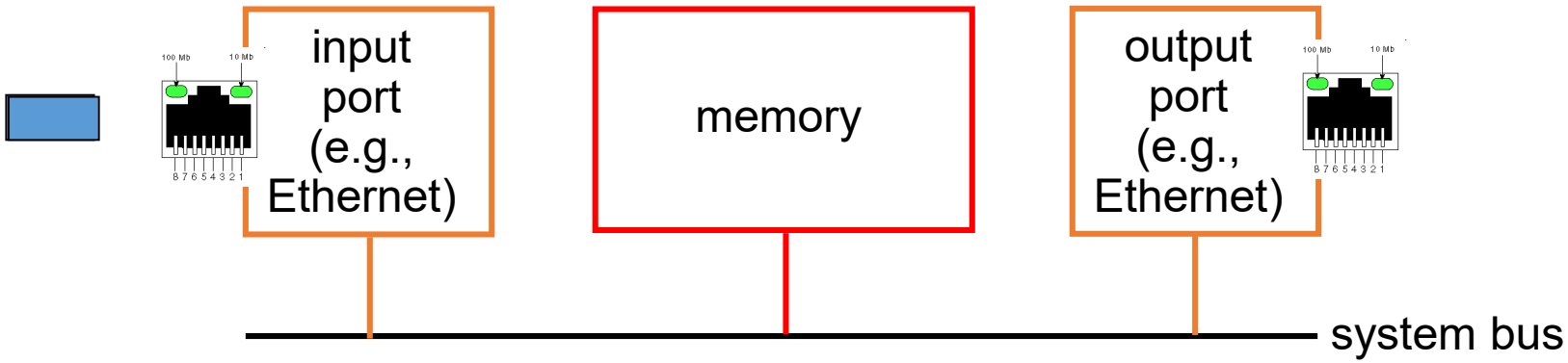
- Transfer packet from input buffer to appropriate output buffer
- Switching rate: Rate at which packets can be transfer from inputs to outputs
  - Often measured as multiple of input/output line rate
  - N inputs: Switching rate N times line rate desirable
- Three types of switching fabrics



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)



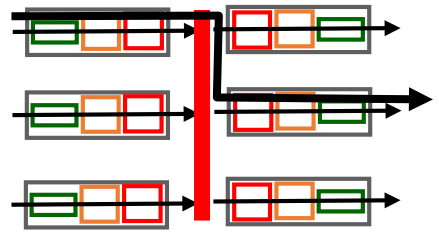
Switching via a bus

Datagram is switched

➞ From input port memory

⬅ To output port memory

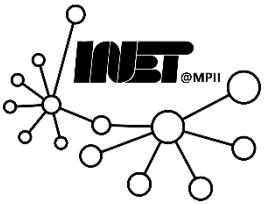
via a shared bus



bus

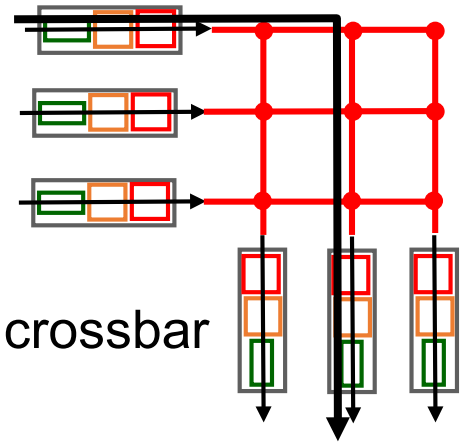
➢*Bus contention:* Switching speed limited by bus bandwidth

E.g. Cisco 5600: 32 Gbps bus
Sufficient speed for access and enterprise routers

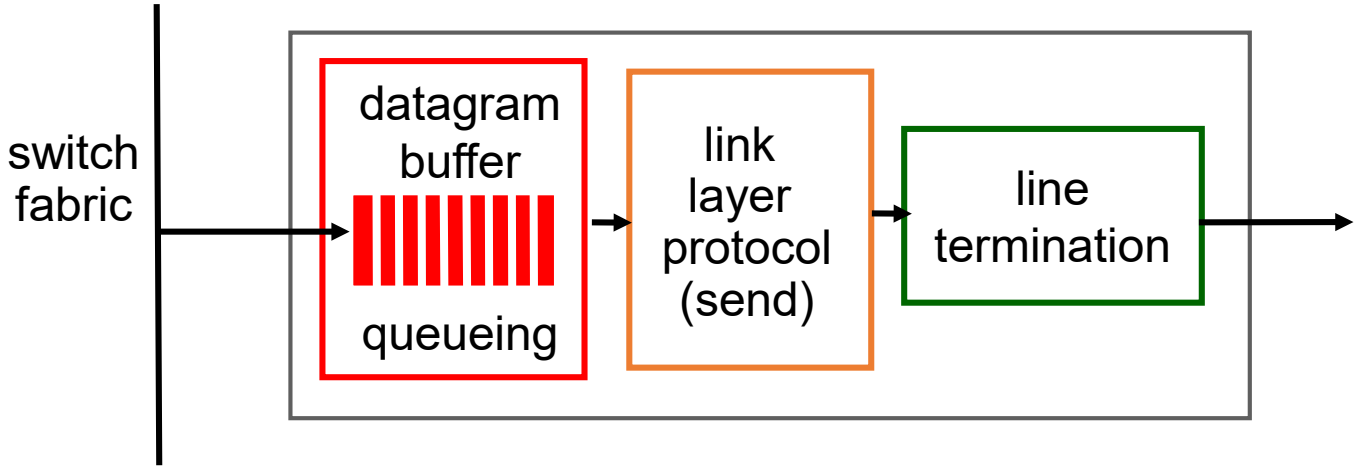


Switching via Dedicated Fabric

- 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
  - Append hardware address of output line card to front of cell
  - Switch cells through the fabric.
- Cisco 12000: Switches 60 Gbps via the interconnection network



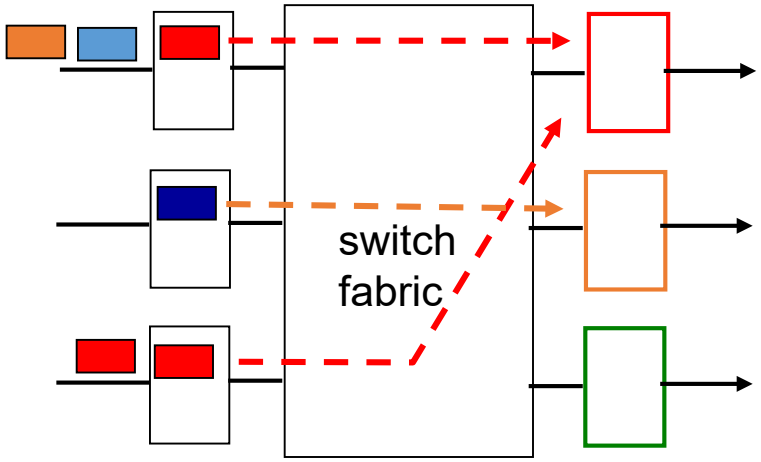
Output ports



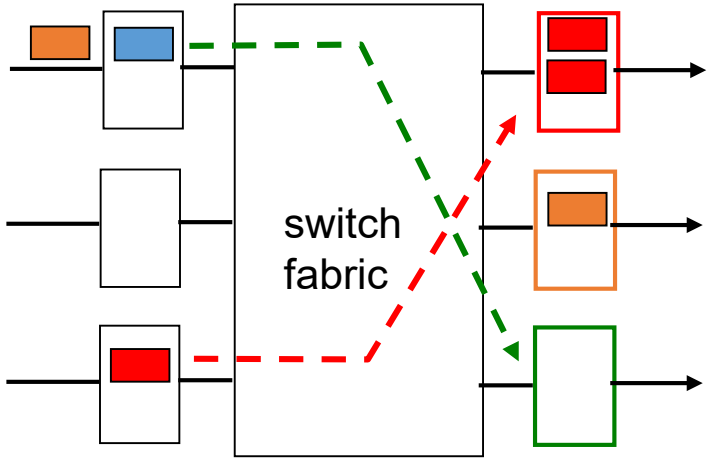
Buffering: Required when datagrams arrive from fabric faster than the transmission rate

Scheduling discipline: Chooses among queued datagrams for transmission

Output port queueing



at *t*, packets move from input to output



one packet per time

Buffering when arrival rate via switch exceeds output line speed

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

How much buffering?

- RFC 3439 rule of thumb: Average buffering:
  "typical" RTT times link capacity C
  - For C = 10 Gpbs link, RTT = 250 msec:

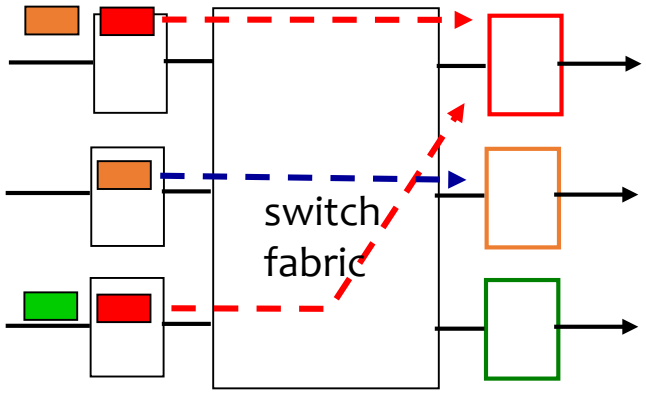
    2.5 Gbit buffer

- Recent recommendation:
  With N flows, buffering equal to:

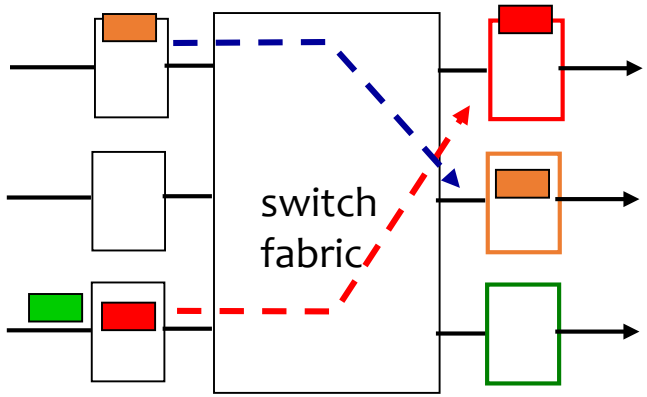
$$\frac{RTT \cdot C}{\sqrt{N}}$$

Input port queuing

- Fabric slower than input ports combined -> queueing 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

Summary

- Routers have evolved from dedicated workstations/PCs to heavy duty industrial equipment costing millions of €s
  - Capable of switching 300+ Tb/s
- Dedicated line cards separate fast path from slow path
  - Slow path through route processor primarily for control plane traffic
- Different hardware approaches to accelerating fast path forwarding
- Queuing and buffering can help or hinder traffic flow