TSN: Lecture 5
Packet Switches

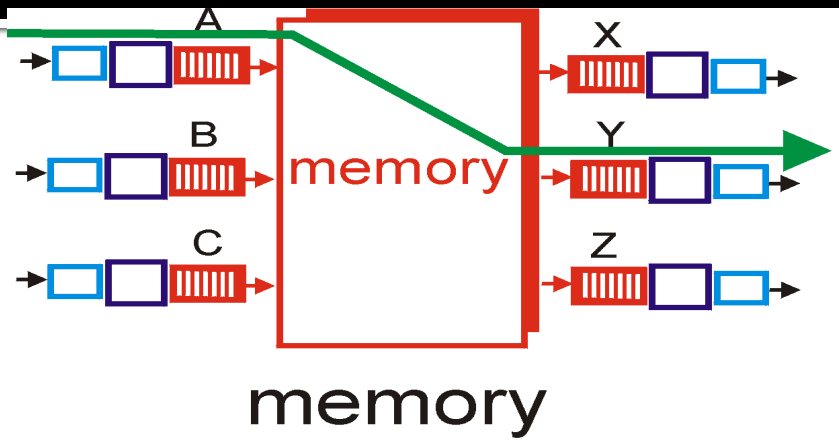
Topics Covered

- Packet Switching
- First Generation Routers
- Second Generation Routers
- Third Generation Routers

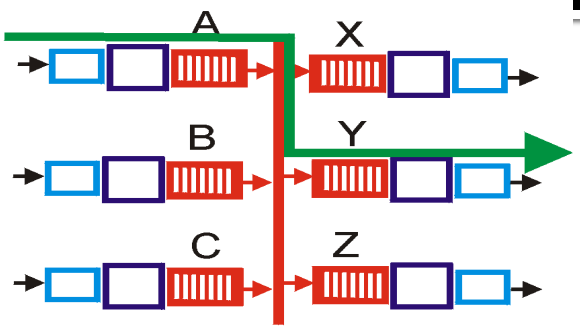
Three generations of packet switches

- Different trade-offs between cost and performance
- Represent evolution in switching capacity
- All three generations are represented in current products
  - Cisco, Juniper, Nortel, Alcatel and many others

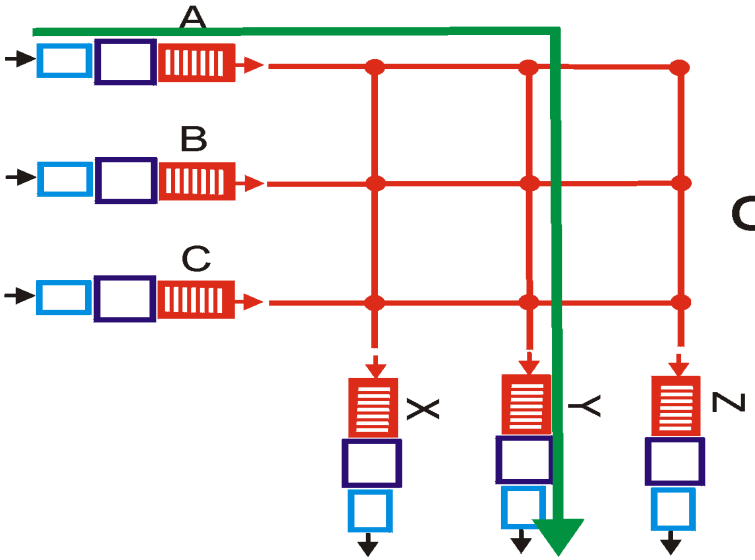
Three types of switching fabrics



memory



bus

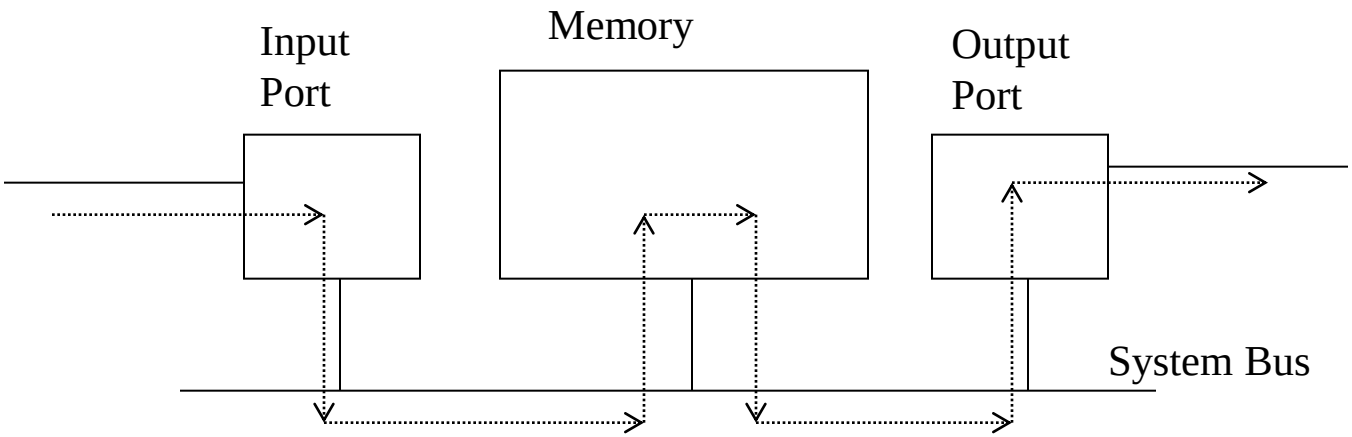


crossbar

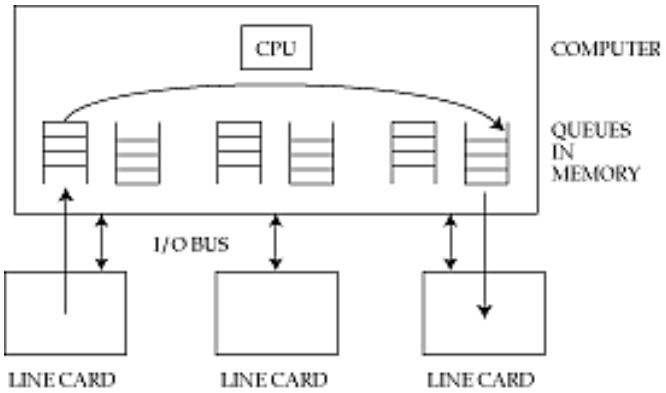
Switching Via Memory

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)

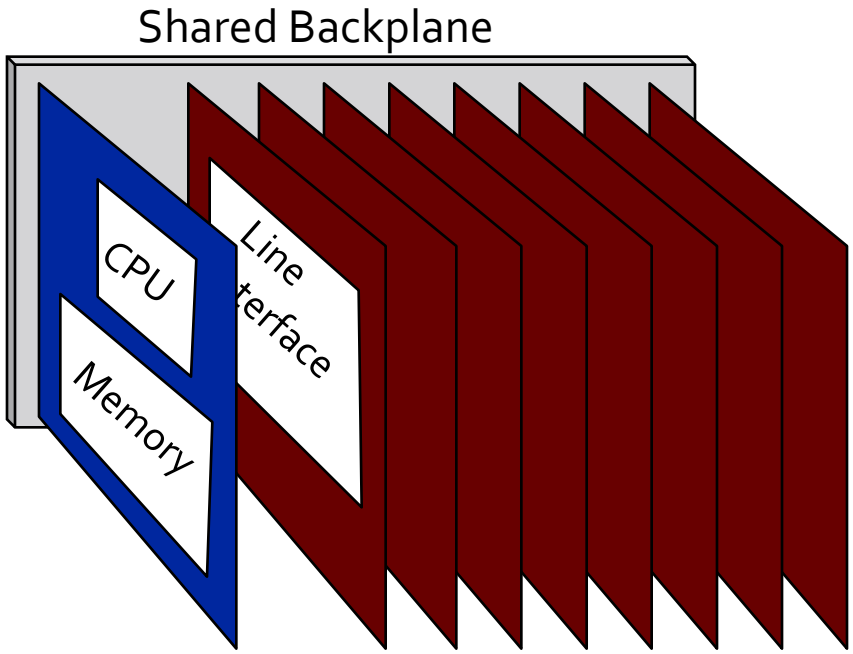


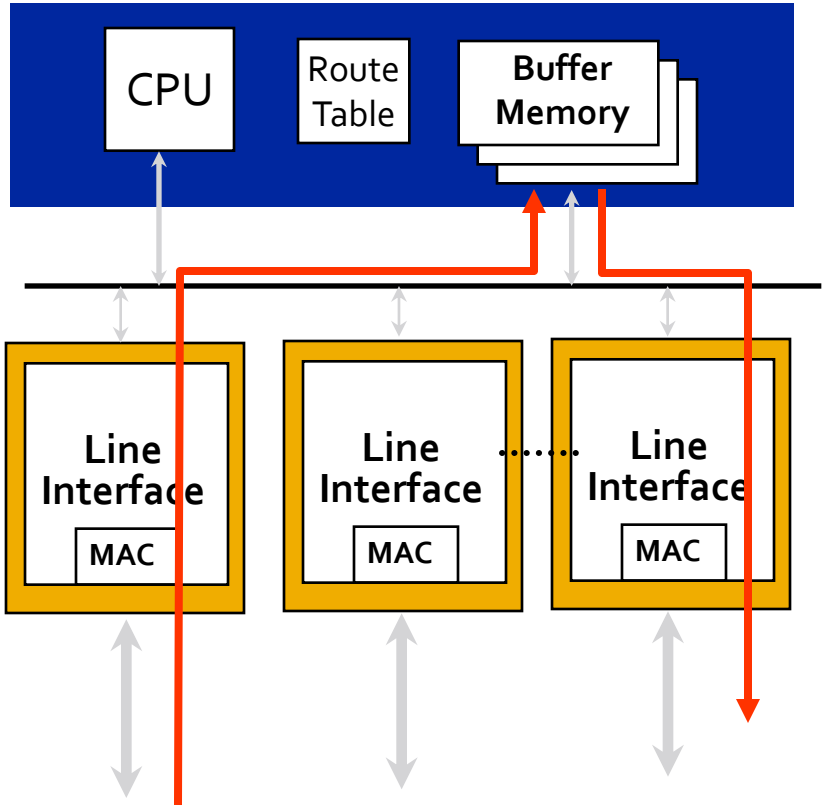
First generation switch



- A computer with multiple line cards
- Processor periodically polls inputs (or is interrupted)
- Most Ethernet switches and cheap packet routers
- Bottleneck can be CPU, host-adaptor or I/O bus, depending on the traffic scenario
- Line card can be cheap (no CPUs on line cards!!)

First Generation Routers





Typically <0.5Gb/s aggregate capacity

Example

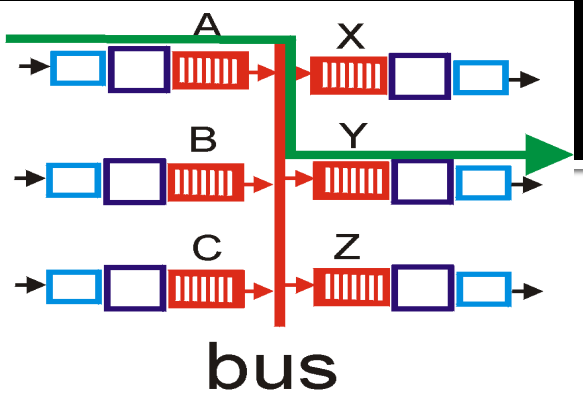
- Assume instruction takes one clock cycle (=7.52 ns)
- Mean packet size 500 bytes
- Interrupt takes 10 microseconds, word (4 bytes) access takes 50 ns
- Per-packet processing time (routing table lookup and others) takes 200 instructions = 1.504 µs

- Copy loop (one word copy)

```
register <- memory[read_ptr]
memory [write_ptr] <- register
read_ptr <- read_ptr + 4
write_ptr <- write_ptr + 4
counter <- counter -1
if (counter not 0) branch to top of loop
```

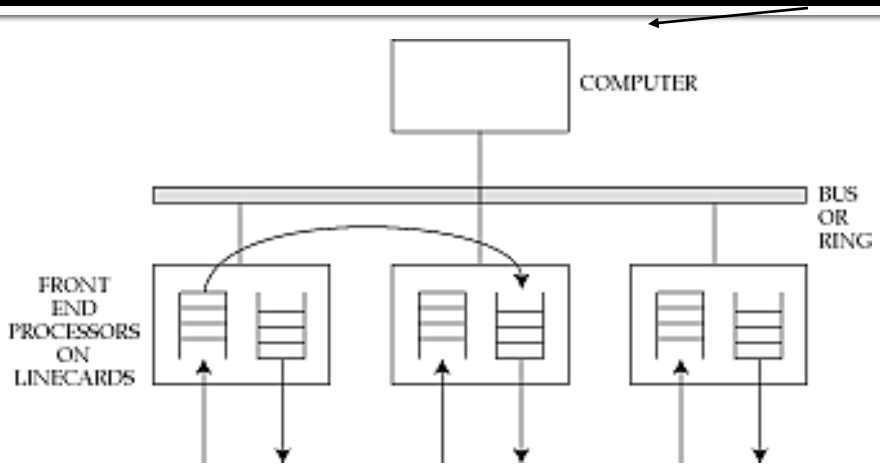
- 4 instructions + 2 memory accesses = 130.08 ns
- Copying packet takes 500/4 *130.08 = 16.26 µs; interrupt 10 µs
- Total time = 16.26+10+1.504=27.764 µs => speed is 144.1 Mbps
- How many Ethernet ports (10Mbps, 100Mbps) can be support??
  - Linux, Windows can do this now!!

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
- 1 Gbps bus, Cisco 1900: sufficient speed for access and enterprise routers (not regional or backbone)

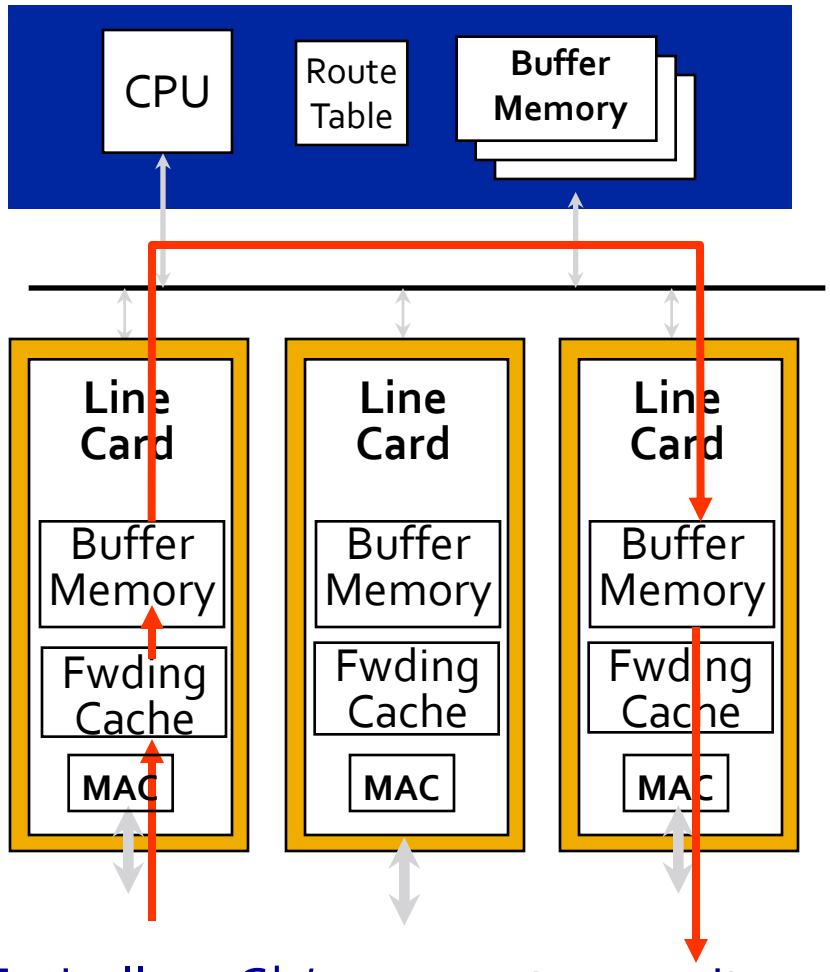
Second generation switch



Routing Protocols
Routing Tables

- Port mapping intelligence in line cards (processor based)
- Ring or Bus based backplane to connect line cards
  - Bottleneck -> performance impact (discuss bus and ring)
- Lookup cache on line cards (for better performance)
- For switch, cross connect table (port mapping entries) only changed when calls are setup / torn down
- For datagram router, ask the control processor if the entry is not found in local forwarding table or automatically updated.

Second Generation Routers



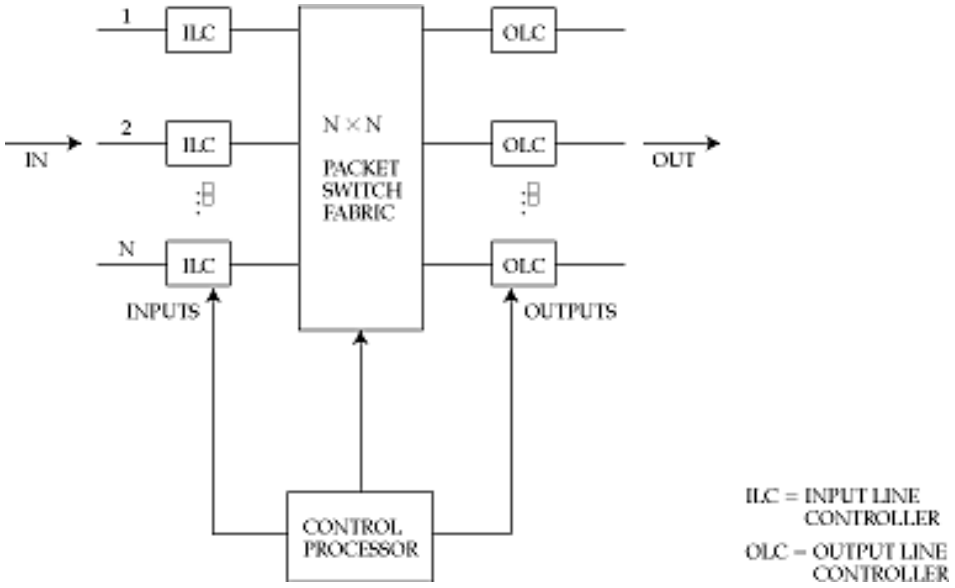
Typically <5Gb/s aggregate capacity

Switching Via An Interconnection Network

- overcome bus bandwidth limitations
- Banyan networks, other interconnection nets initially developed to connect processors in multiprocessor
- Advanced design: fragmenting datagram into fixed length cells, switch cells through the fabric.
  - Segmentation and Reassembly (SAR)
    - Discuss overheads
- Cisco 12000: switches Gbps through the interconnection network

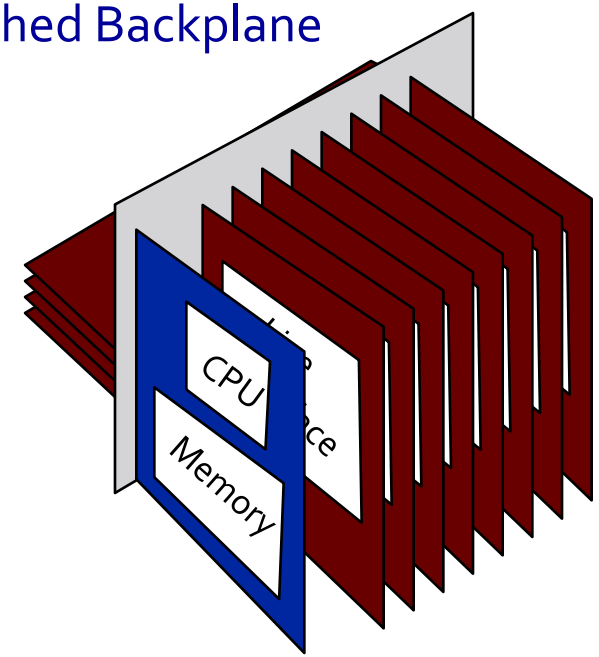
Third generation switches

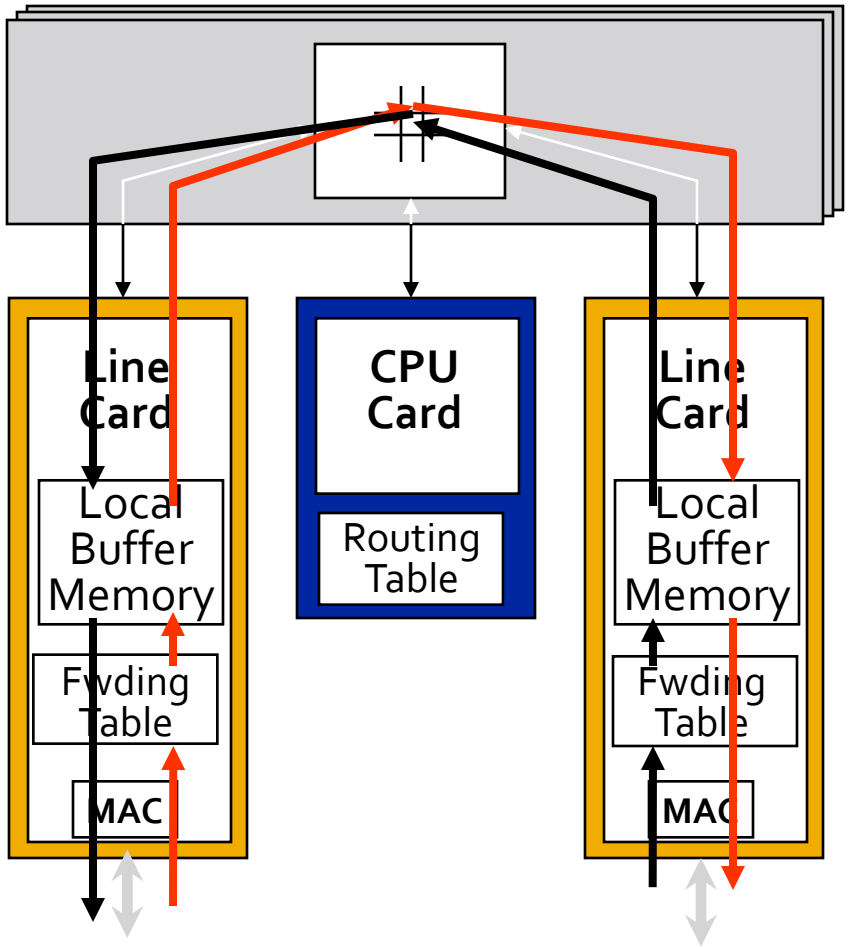
- Bottleneck in second generation switch is the bus (or ring)
- Third generation switch provides parallel paths using a switch fabric



Third Generation Routers

Switched Backplane



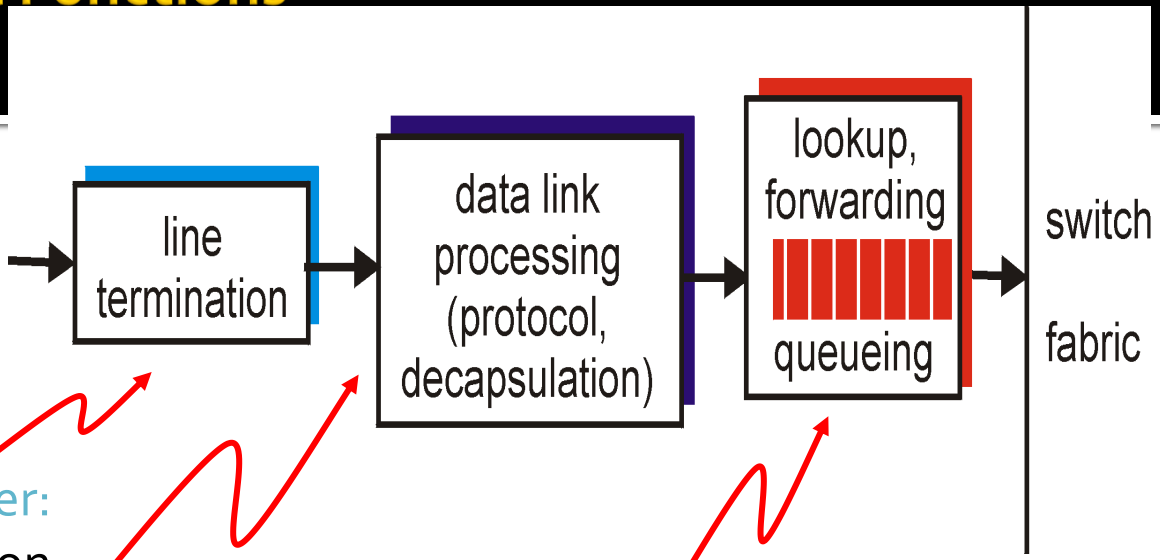


Typically <50Gb/s aggregate capacity

Third generation (contd.)

- Features
  - self-routing fabric
  - output buffer is a point of contention
    - unless we *arbitrate* access to fabric
  - potential for unlimited scaling, as long as we can resolve contention for output buffer

Input Port Functions



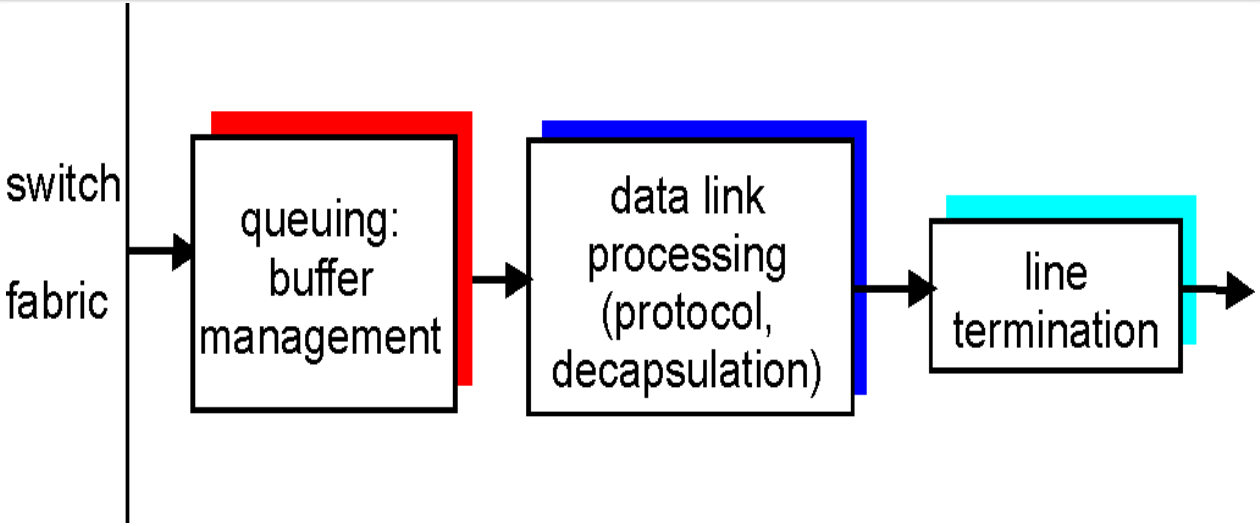
Physical layer:
bit-level reception

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

Decentralized switching:

- given datagram dest., lookup output port using forwarding table in input port memory
- goal: complete input port processing at 'line speed'
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

Output Ports



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