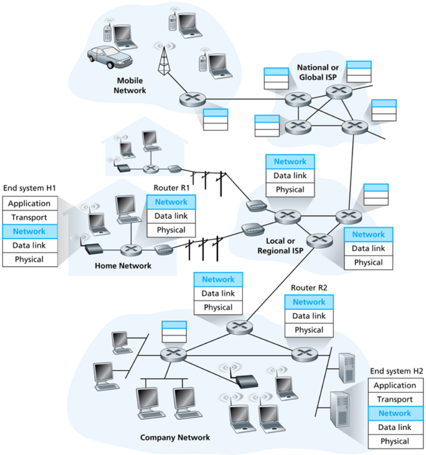
The Network Layer and Routers

Daniel Zappala

CS 460 Computer Networking
Brigham Young University

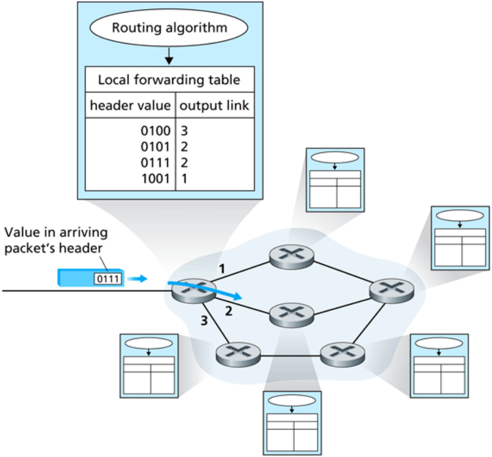
Network Layer



- deliver packets from sending host to receiving host
- must be on every host, router in the Internet – IP defines what it means to be a part of the Internet
- switch: processes only the link layer header
- router: processes the network layer header

Packet Forwarding versus Routing

- forwarding
  - accept packet on incoming interface
  - lookup outgoing interface from forwarding table
  - put packet into queue for outgoing interface
- routing
  - determine the path that packets should take
  - use these paths to create the forwarding table in a given router



Local forwarding table

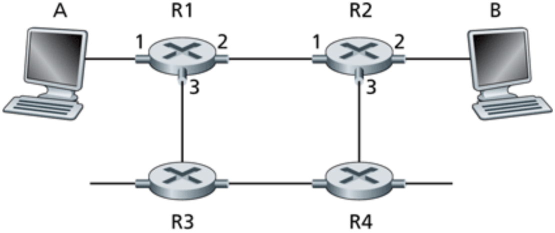
header value	output link
0100	3
0101	2
0111	2
1001	1

How Should the Network Forward Packets?

Virtual Circuit

- setup a fixed path through the network like a telephone circuit
  - assign a flow to a path
  - packets carry virtual circuit identifiers to identify the path, rather than being forwarded based on the destination IP address
- advantages
  - can maintain multiple paths to a destination, each used by different flows
  - can set aside resources for each path, e.g. bandwidth
  - can isolate traffic, e.g. carry voice traffic separate from data traffic

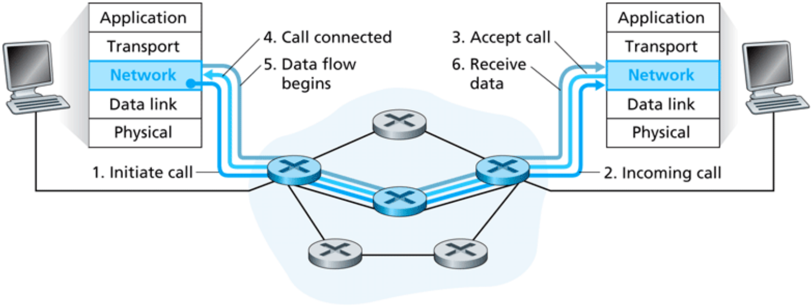
Virtual Circuit Example



- VC table for R1:

Incoming IF	Incoming VC	Outgoing IF	Outgoing VC
1	12	2	22
2	63	1	18
3	7	2	17
1	97	3	87

Virtual Circuit Setup



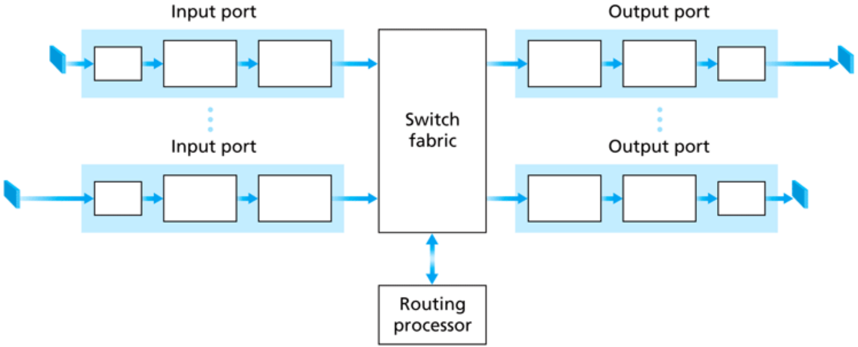
Datagrams

- packets forwarded using only destination address in packet
  - each router sends packet to best next hop toward destination
  - packets sent from same source to same destination may take different paths
- advantages
  - simple
    - no connection setup in network layer
    - no connection state in routers
    - complexity (reliability and transport) at the edges
  - flexible: if a route fails, datagrams can take another path
- disadvantages
  - can only use one route per destination
  - forwarding table can be large
  - no performance guarantees

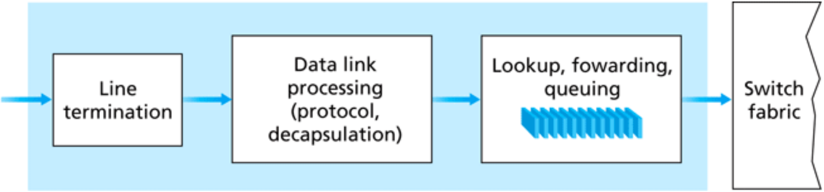
How Do Routers Work?

Routers

- key functions
  - run routing protocol (RIP, OSPF, BGP)
  - forward datagrams from incoming to outgoing link
  - schedule packets in outgoing link queues

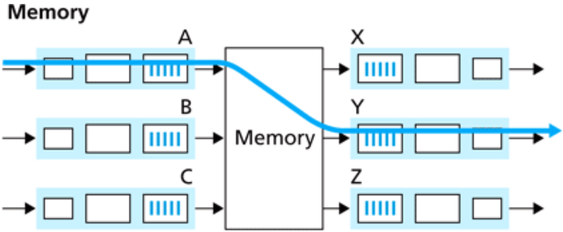


Input Port Processing


Line termination
Data link processing (protocol, decapsulation)
Lookup, fowarding, queuing
Switch fabric

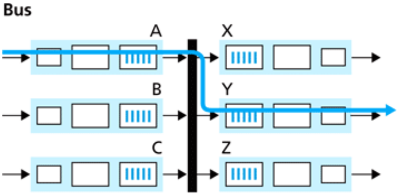
- line termination - physical layer - bits
- data link processing - link layer - frames (e.g. Ethernet)
- lookup, forwarding, queueing - network layer
  - given a datagram destination address, lookup output port using forwarding table stored in input port memory
  - try to complete input port processing at link speed
  - if datagrams arrive faster than processing rate, they may be queued on input port

Switching Fabrics: Memory



- first-generation routers
- switching controlled by a CPU - packet copied to system memory and then out to output port
- memory contention: switching speed limited by memory bandwidth

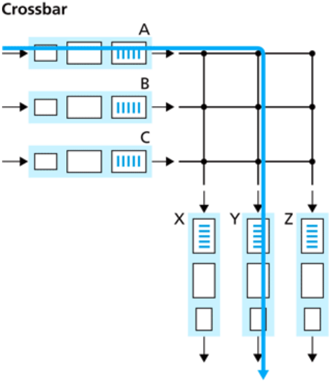
Switching Fabrics: Bus



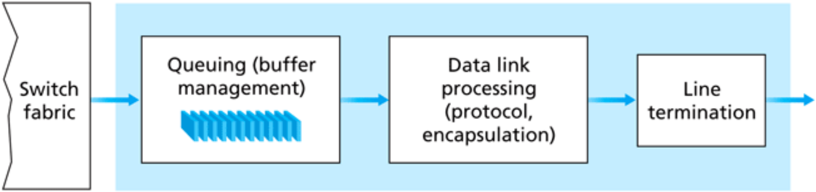
- copy datagram from input port memory to output port memory via shared bus
- bus contention: switching speed limited by bus bandwidth
- Cisco Catalyst 1900: 1 Gbps bus (old product)
  - sufficient speed for access and enterprise routers (not regional or backbone)
  - 3MB memory shared by all ports
  - 14,880 pps to 10-Mbps ports, 148,800 pps to 100-Mbps ports (64-byte packets)

Switching Fabrics: Interconnection Network

- overcome bus bandwidth limitation
- crossbar, Banyan networks, and others
- advanced design: fragment datagram into fixed length cells, switch cells through the fabric (cell switching)
- Cisco 12000 (December 2003)
    - 2.5 - 40 Gbps/slot
    - memory-less crossbar switching matrix
    - distributed processing



Output Port Processing

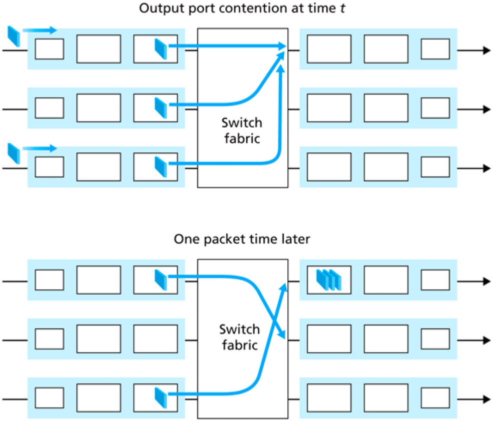

Switch fabric
Queuing (buffer management)
Data link processing (protocol, encapsulation)
Line termination

- buffering required: datagrams may arrive faster than link transmission speed – **delay and loss possible**
- scheduling discipline chooses order in which datagrams are transmitted
  - FIFO: service packets in order they arrive
  - priority queueing: service packets according to priority field in IP header
  - fair queueing: give every TCP flow a fair share of link bandwidth

Active Queue Management

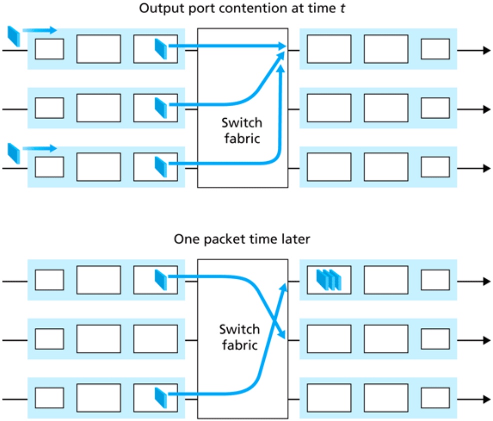
- drop or mark a packet before the queue is full
  - sends a congestion signal to the sending host
  - causes hosts to slow down before buffer becomes full
  - to mark packets, need support from TCP to carry mark back to the source: TCP ECN
- Random Early Detection (RED)
  - use EWMA to track average queue length
  - if queue $length < min_{th}$, let packet in
  - if queue $length > max_{th}$, mark or drop
  - if queue length between min_{th} and max_{th}, mark or drop with probability $= f(length)$
    - this function increases from 0 to 1 as length moves from the minimum to maximum threshold
- AQM keeps the queue size low, which decreases overall queueing delay

Output Port Contention



- in this example, the switch fabric can deliver three packets to an output port in the time it takes for a single packet to be sent at the output port
- results in queueing delay at output port

Head-of-Line Blocking



- in this example, the switch fabric can only deliver one packet a time to an output queue
- even though a second packet in the bottom input queue can be delivered to the middle output queue, it is blocked by a packet ahead of it