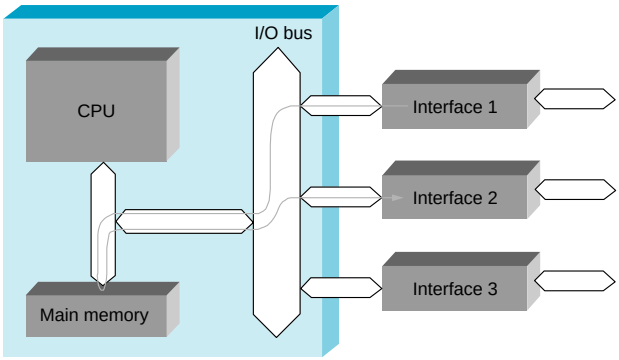
Switch Construction

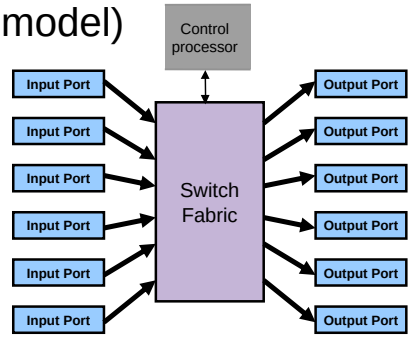
Workstation-Based

- Aggregate bandwidth
  - 1/2 of the I/O bus bandwidth
  - capacity shared among all hosts connected to switch
  - example: 1Gbps bus can support 5 x 100Mbps ports (in theory)
- Packets-per-second
  - must be able to switch small packets
  - 300,000 packets-per-second is achievable



Switching Hardware

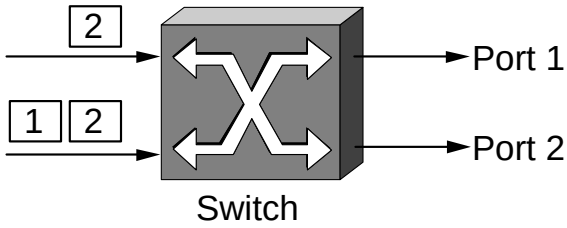
- Design Goals
  - Throughput (depends on traffic model)
  - Scalability (a function of *n*)



- Ports
  - Circuit management (e.g., map VCIs, switch datagrams)
  - buffering (input and/or output)
- Fabric
  - As simple as possible
  - Sometimes do buffering (internal)

Buffering

- Wherever contention is possible
  - Input port (contend for fabric)
  - Internal (contend for output port)
  - Output port (contend for link)
- Head-of-Line Blocking
  - Input buffering



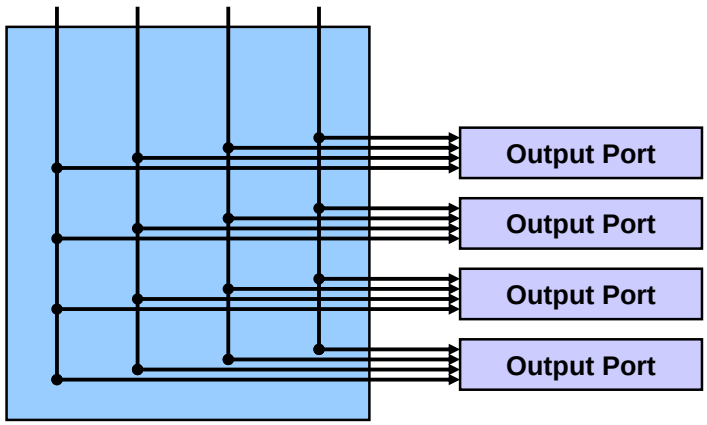
Switch Design

- Crossbar switches
- Banyan Networks
- Batcher Networks
- Sunshine Switch

Crossbar Switch

- Every input port is connected to every output port
  - NxN
- Output ports
  - Complexity scales as $O(N^2)$

Crossbar Switch



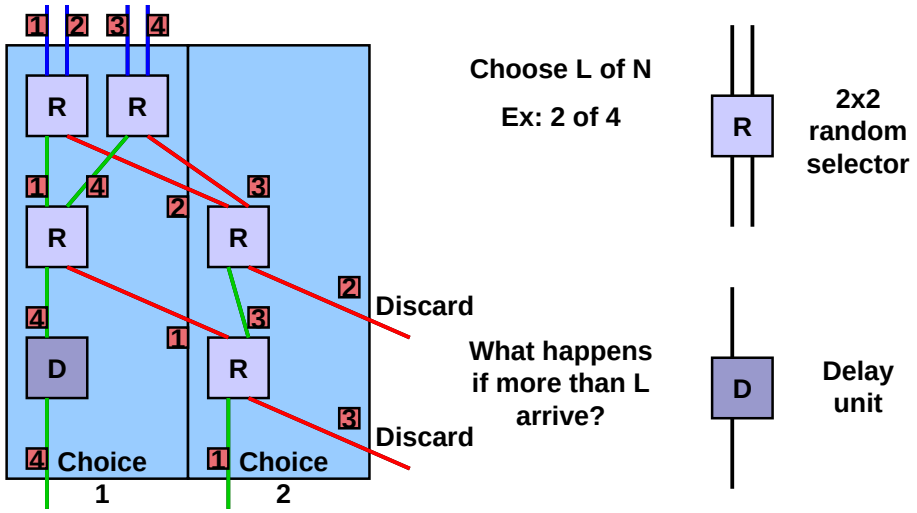
Knockout Switch

- Assumption:
  - It is unlikely that N inputs will have packets destined for the same output port
- Pick L from N packets at a port
  - Output port maintains L cyclic buffers
  - Shifter places up to L packets in one cycle
  - Each buffer gets only one packet
  - Output port uses round-robin between buffers
  - Arrival order is maintained
- Problem
  - Hot spots
- Output ports scale as O(N)

Knockout Switch

- Output port design
  - Packet filters
    - Recognize packets destined for a specific port
  - Concentrator
    - Selects up to L packets from those destined for this port
    - Discards excess packets
  - Queue
    - Length L

Knockout Switch

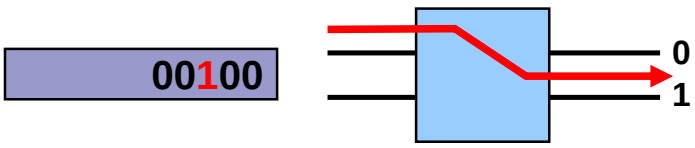


Self-Routing Fabrics

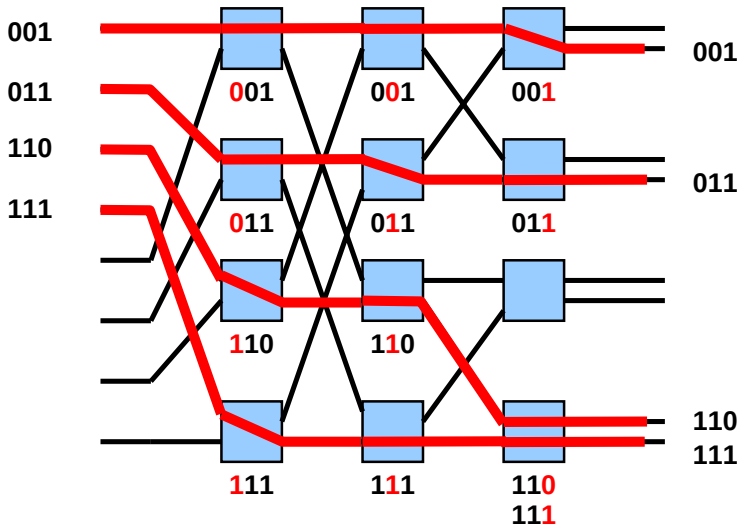
- Idea
  - Use source routing on “network” in switch
  - Input port attaches output port number as header
  - Fabric routes packet based on output port
- Types
  - Banyan Network
  - Batcher-Banyan Network
  - Sunshine Switch

Banyan Network

- A network of 2x2 switches
  - Each element routes to output 0 or 1 based on packet header
  - A switch at stage i looks at bit i in the header



Banyan Network



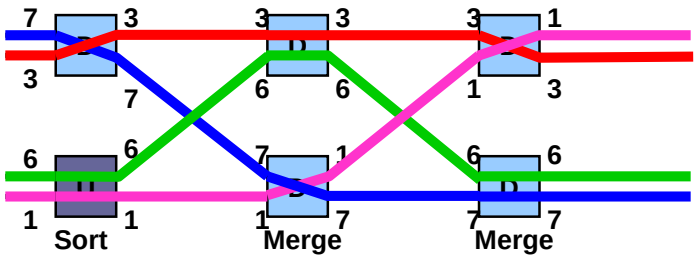
Banyan Network

- Perfect Shuffle
  - N inputs requires $\log_2 N$ stages of N/2 switching elements
  - Complexity on order of $N \log_2 N$
- Collisions
  - If two packets arrive at the same switch destined for the same output port, a collision will occur
  - If all packets are sorted in ascending order upon arrival to a banyan network, no collisions will occur!

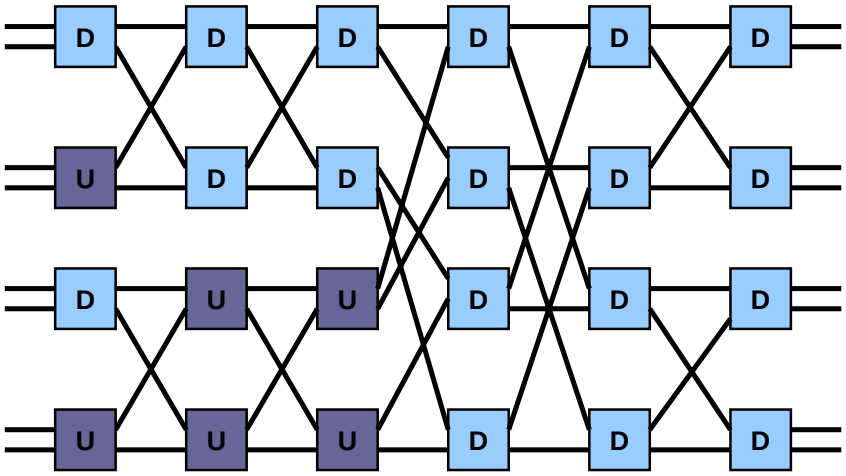
Batcher Network

- Performs merge sort
- A network of 2x2 switches
  - Each element routes to output 0 or 1 based on packet header
  - A switch at stage i looks at the whole header
  - Two types of switches
    - Up switch
      - Sends higher number to top output (0)
    - Down switch
      - Sends higher number to bottom output (1)

Batcher Network



Batcher Network



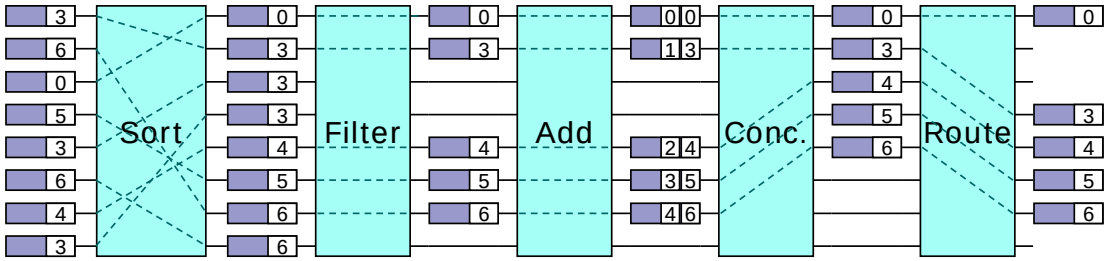
Batcher Network

- How it really works
  - Merger is presented with a pair of sorted lists, one in ascending order, one in descending order
  - First stage of merger sends packets to the correct half of the network
  - Second stage sends them to the correct quarter
- Size
  - N/2 switches per stage
  - $\log_2 N \times (1 + \log_2 N)/2$ stages
  - Complexity = $N \log_2^2 N$

Batcher-Banyan Network

- Idea
  - Attach a batcher network back-to-back with a banyan network
  - Arbitrary unique permutations can be routed without contention
  - Two packets destined for same output port still collide!

Batcher-Banyan Switch Architecture

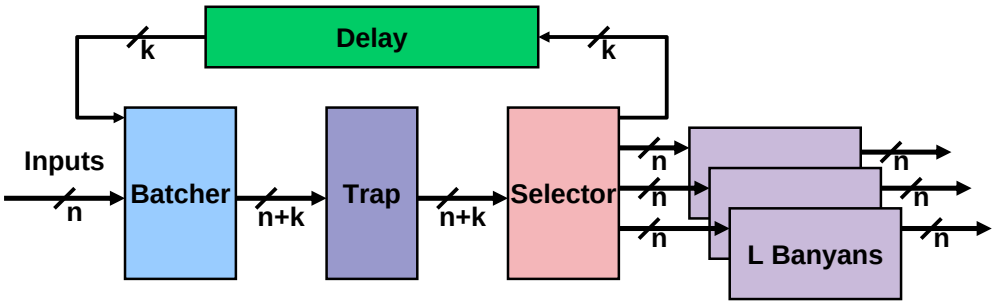


- Simple components with no buffering.
  - filter eliminates duplicates by comparing consecutive addresses and returns ack to inputs
  - adder computes and inserts “rank” of cells
  - concentrator uses rank as output address
  - routing network delivers to output
- Adder, concentrator and routing network all have $\log_2 n$ stages (conc. is reverse banyan, routing net. is banyan)

Sunshine Switch

- Sunshine Switch
  - Like a knockout switch
    - Can handle up to L packets per output port
  - Recirculates overflow packets
    - If more than L packets arrive for any output port in one cycle
- Elements
  - Multiple Banyan networks
    - Enables multiple packets per output port
  - Delay Box
    - Excess (K) packets are recirculated and resubmitted to the switch
  - Batcher network
    - N new packets
    - K delayed packets
  - Trap
    - Identifies packets destined for banyan
    - Identifies excess packets
  - Selector
    - Routes multiple packets for same output on separate banyans

Sunshine Switch



References

- Robin Kravets, “Switching Hardware”, UIUC, CS/ECE-438.