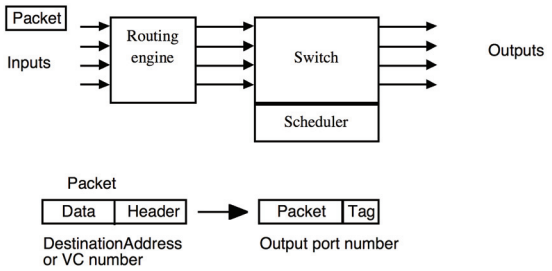
Outline

- Basics of switching
- Blocking
- Interconnection examples
- Complexity
- Recursive constructions

Switching and Routing

- Switching is generally the establishment of connections on a circuit basis
- Routing is generally the forwarding of traffic on a datagram basis
- Routing requires switching but not vice-versa – routing uses connections which are permanently or temporarily set up to in order to forward datagrams (those datagrams may be in circuit form, for instance VPs and VCs)

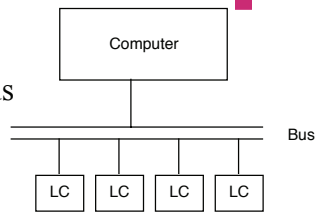
Packet routers



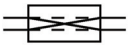
- A packet switch consists of a routing engine (table look-up), a switch scheduler, and a switch fabric.
- The routing engine looks-up the packet address in a routing table and determines which output port to send the packet.
  - Packet is tagged with port number
  - The switch uses the tag to send the packet to the proper output port

MIT

Switch fabrics



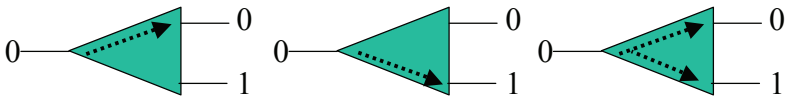
- Simplest switch fabric is simply a shared bus
  - Most of the processing is done in line cards
    - Route table look-up
    - Line cards buffer the packets
    - Line card send packets to proper output
  - Bus bandwidth must be N times LC speed (N ports)
- In general a switch fabric replaces the bus
- Switch fabrics are created from certain building blocks of smaller switches arranged in stages
- Simplest switch is a 2x2 switch, which can be either in the through or crossed position



MIT

Definitions

- A connection state is a mapping from the array of inputs to that of outputs; connections are either point-to-point or multicast
- Basic switch building blocks are:
  - the distributor



  - the concentrator

  - the 2x2 2-state point-to-point switch (switching cell)

MIT

Building up

- Interconnection network: finite collection of nodes together with a set of interconnection lines such that
  - every node is an object with an array of inputs and an array of outputs
  - an interconnection line leads from an output of one node to an input of another node
  - every I/O of a node is incident with at most one interconnection line
  - an I/O is called external if it is not incident with any interconnection line
- A route from an external input to an external output is a chain of distinct $(a_0, b_0, a_1, b_1, \ldots, a_k, b_k)$ where a_0 and b_k are external, b_{j-1} is interconnected to a_j

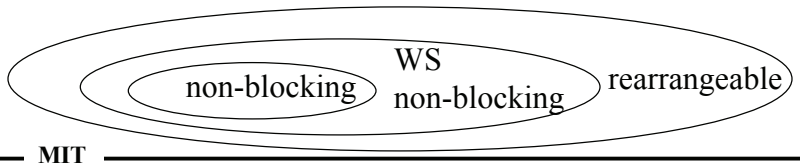
MIT

Building up

- An interconnection network is called a switching network when:
  - every node qualifies to be a switch through proper specification of connection states
  - the network is routable (there exists a route from every external input to every external output)
  - an ordering is specified on external inputs and on external outputs
- Unique routing interconnection networks: all routes from an external input to an external output are parallel, that is (a_0, b_0, a_1, b_1, …, a_k, b_k) and (a_0, b'_0, a'_1, b'_1, …, a'_k, b_k) are such that a_j, a'_j reside on the same nodes and b_j, b'_j reside on the same node
- Otherwise: alternate routing

Blocking

- A mxn unique routing network is called a nonblocking network if for any integer k < min(m,n)+1, any k external inputs, any k external outputs and pairing between these external I/O, there exist k disjoint routes for the matched pairs
- For a routable network, the same property is that ot a rearrangeably nonblocking, or rearrangeable network
- An interconnection network is strictly non-blocking if requests for routes are always granted under the rule of arbitrary route selection, wide-sense non-blocking if there exists an algorithm for route selection that grants all requests



Blocking, Multi-stage networks

- Main connection between rearrangeability and non-blocking property is given by the following theorem:

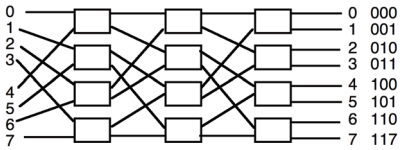
  A switching network composed of non-blocking switches is rearrangeable iff it constructs a non-blocking switch

- A common means of building interconnection networks is to use a multi-stage architecture:
  - every interconnection line is between two stages
  - every external input is on a first-stage node
  - every external output is on a final-stage node
  - nodes within each stage are linearly ordered

MIT

Interconnection networks

- N input, Log(N) stages with N/2 modules per stage
  Example: Omega (shuffle exchange network)

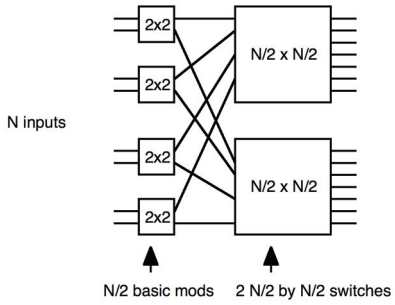


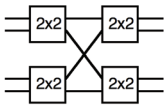
- Notice the order of inputs into a stage is a shuffle of the outputs from the previous stage: (0,4,1,5,2,6,3,7)
- Easily extended to more stages
- Any output can be reached from any input by proper switch settings
  - Not all routes can be done simultaneously
  - Exactly one route between each OD pair

MIT

Interconnection networks

- Another example of a multi-stage interconnection network
- Built using the basic 2x2 switch module
- Recursive construction
  - Construct an N by N switch using two N/2 by N/2 switches and a new stage of N/2 basic (2x2) modules
  - N by N switch has $Log_2(N)$ stages each with N/2 basic (2x2)





Complexity issues

- There are many different parameters that are used to consider the complexity of an interconnection network
- Line complexity: number of interconnection lines
- Node (cell) complexity: number of small nodes (mxn where m < 3 and n < 3)
- Depth: maximum number of nodes on a route (assuming an acyclic interconnection network)
- Entropy of a switch: log of the number of connections states
- What relations exist between complexity and the capabilities of a switch?

Complexity

- The depth of a mxn routable interconnection network is at least max(log(m), log(n)).
- Proof: for a depth d, there are at most 2^d external outputs. Since we have routability, $n < 2^d+1$ and $m < 2^d+1$.
- When a switching network is composed of 2-state switches, the component complexity of the network is at least the entropy of the switch
- Proof: for E the number of switches, there are 2^E ways to form a combination of one connection state in every node. Each combination corresponds to at most one connection state in the node.

Complexity

- When a nxn rearrangeable network is composed of small nodes, its component complexity is at least log(N!)
- Proof: if we take every small node to be replaced by a 2-state point-to-point switch, then we have a non-blocking switch. Thus, there is a different connection state for everyone of the n! one-to-one mapping between the n inputs and the n outputs. We now use the relation for networks composed of 2-state switches.
- Note: using Stirling's formula, we can obtain an approximate simple bound for component complexity

Complexity

- Component complexity:

$$n! \approx \sqrt{2\pi\, n}\left(\frac{n}{e}\right)^n$$

$$\Rightarrow \log(n!) = n\log(n) - 1.44n + \log\left(\frac{n}{2}\right) + \frac{\log(2\pi)}{2}$$

so component complexity is bounded from below by

$$n\log(n) - 1.44n + \Omega(\log(n))$$

- Relation between line and component complexity:

component complexity +mn = line complexity +m + n

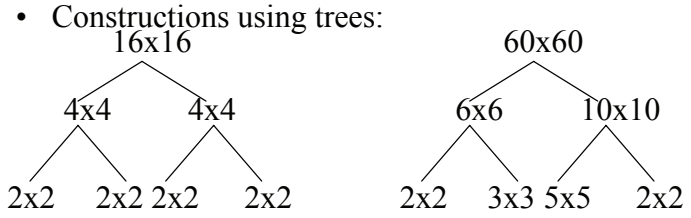
Complexity

- If a mxn nonblocking network is composed of n_{12} 1x2 nodes, n_{21} 2x1 nodes, n_{22} cells, plus possibly crosspoints (edges), then

  $n_{12} + n_{21} + 4\, n_{22} = 2mn - m - n$
- Corollary: a nxn non-blocking network composed of small nodes has component complexity at least $0.5(n^2 - n)$
- Note: directed acyclic graphs can be seen as a special case of a network - a crosspoint network.
- We have basic complexity properties, but how do we build networks?

Recursive 2-stage construction

- 2-stage interconnection with parameters m and n is composed of n mxm input nodes and m nxn output nodes interconnected by a coordinate interchange (static)
- Constructions using trees:



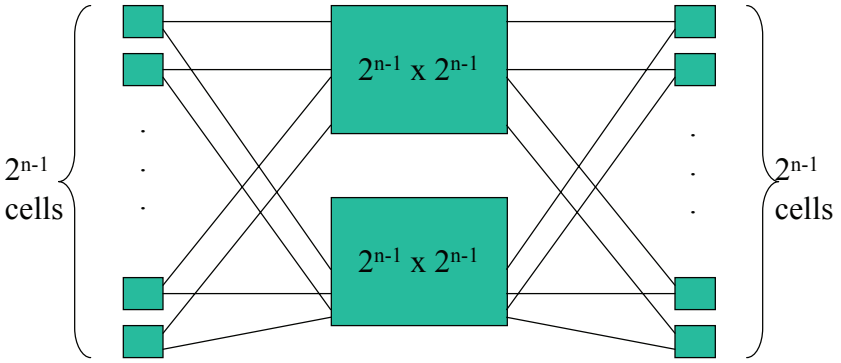
Divide and conquer

- Basic blocks need not be 2x2, trees need not be balanced

MIT

Benes approach

- A three stage approach in which we use as the middle stage two networks of size 2^{n-1} x 2^{n-1} to build a network of size 2^n x 2^n



MIT

Generalized 3-stage approach

- We denote by [nxm, rxp, mxq] the 3-stage network with r nxm input nodes, m rxp middle nodes, p mxq output nodes such that
  - output y of input node x is linked to input x of of middle node y
  - output u of middle node y is linked to input y of output node u
- Rearrangeability theorem: the 3-stage network is rearrangeable iff

$$m \geq \min\{\max\{n, q\}, nr, pq\}$$

- It is strictly non-blocking iff

$$m \geq \min\{n + q - 1, nr, pq\}$$

MIT

Conclusions

- Basic switching theory gives a foundation for the construction of networks to match inputs to outputs
- Complexity results in terms of number of states and hardware complexity effectively characterize the realizable switches under certain parameters
- Recursive constructions give a means of generating networks of increasing size from smaller entities
- To go beyond these approaches, require an algebraic set-up considering permutation matrices

MIT