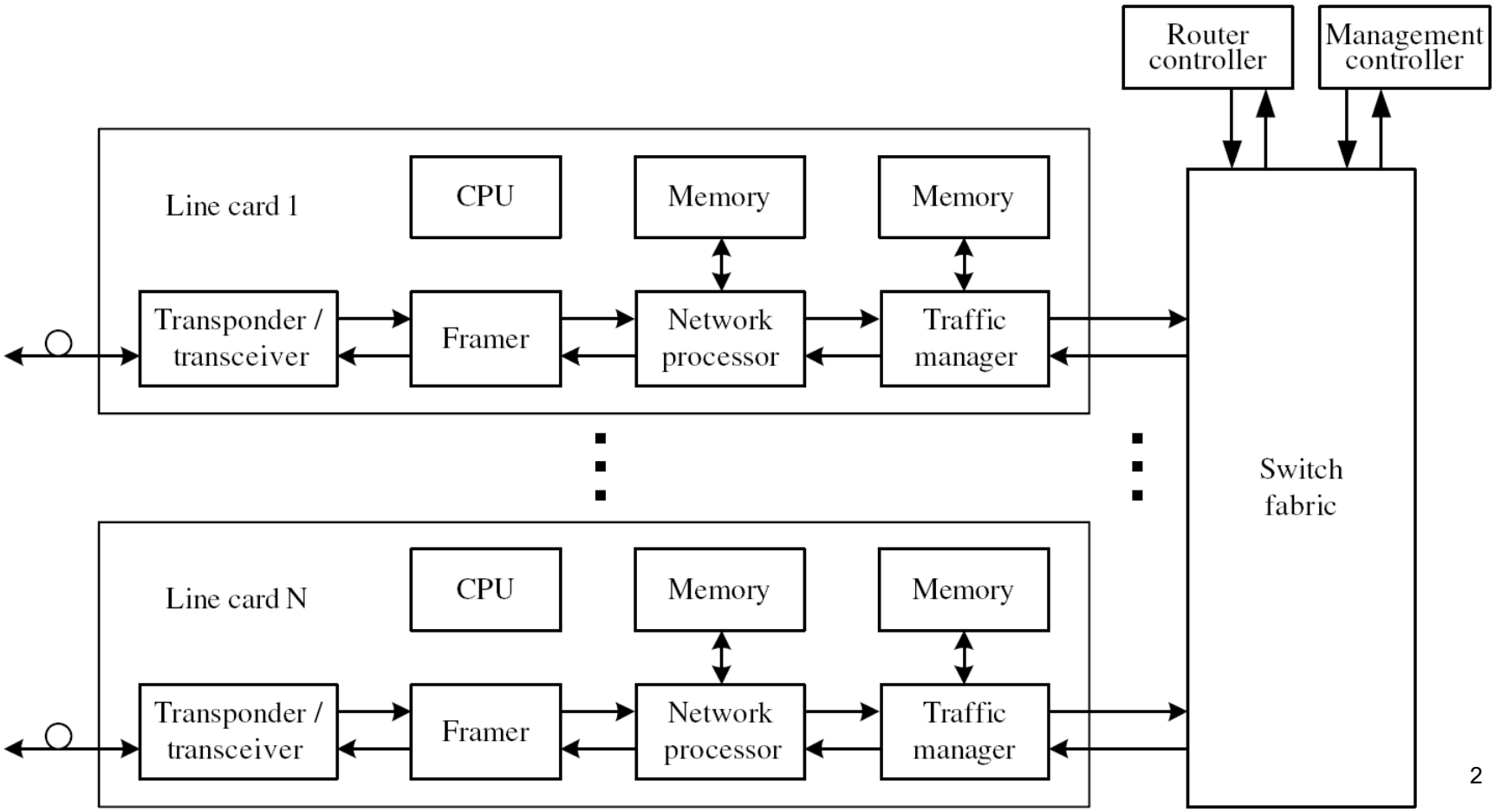
Basics of Packet Switching

Masoud Sabaei

Associate professor

Department of Computer Engineering,
Amirkabir University of Technology

Typical Router Architecture

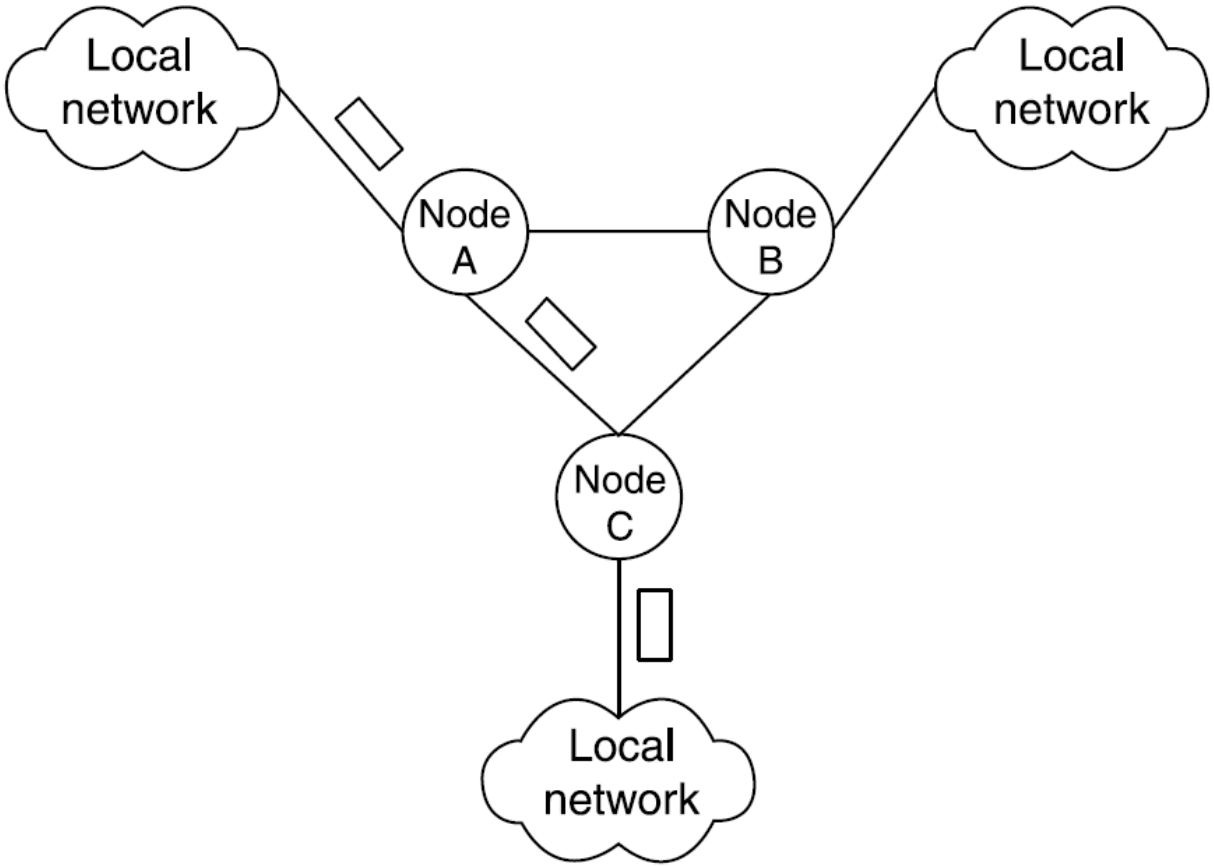


Fundamental Switching Concept

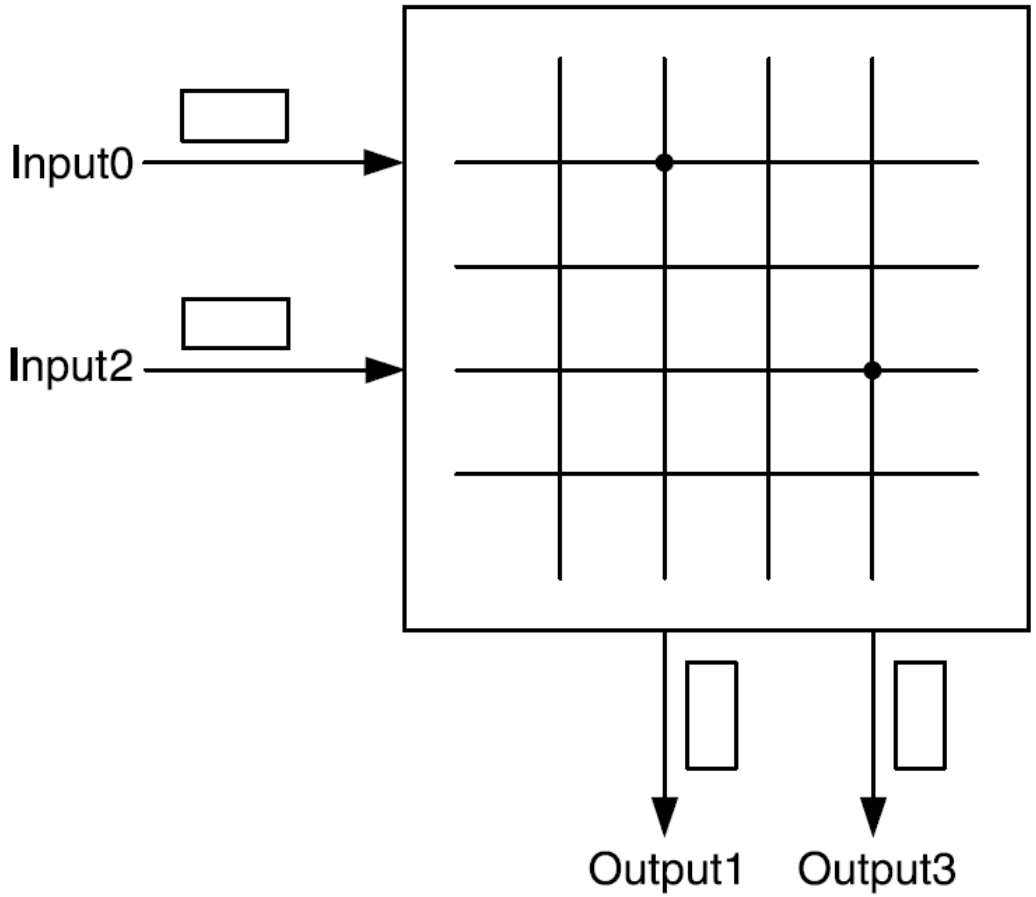
- Switching and Routing,
- Unicast and Multicast,
- Throughput and Speedup,
- Blocking and Output Contention,
- Cell-mode switching and Packet-mode switching.

Basics of Packet Switching

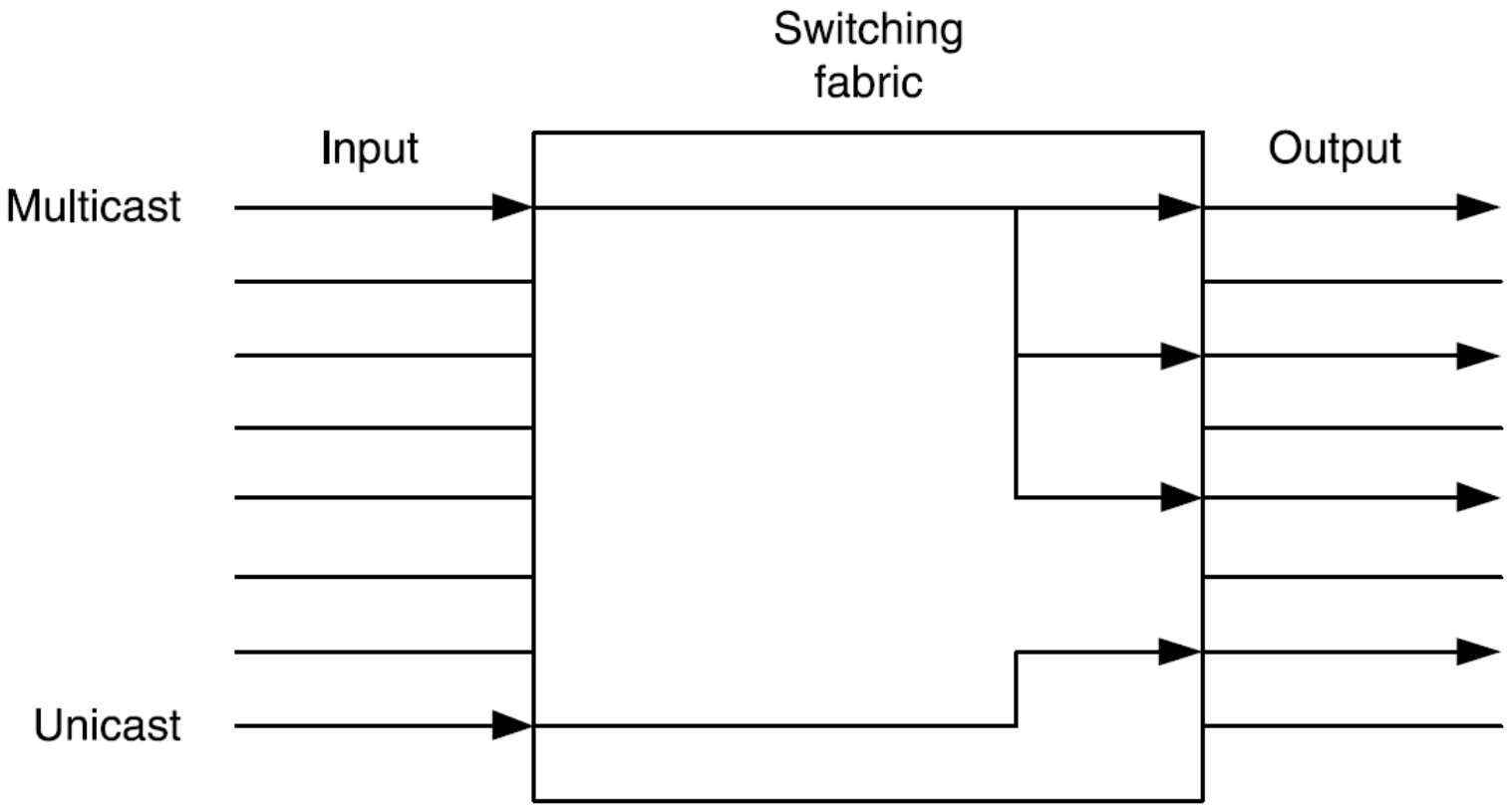
Routing



Switching



Unicast and Multicast



Throughput and Speedup

- Throughput

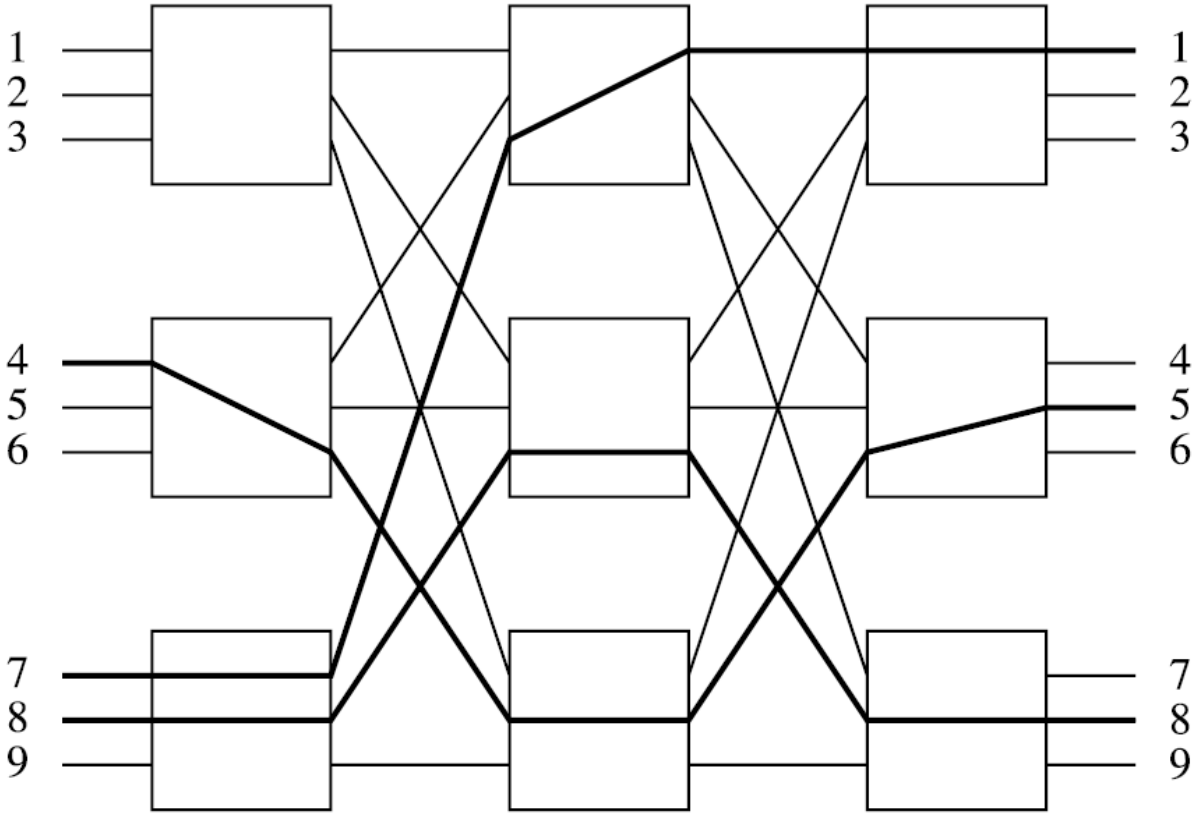
  The throughput of a switch fabric is defined as the ratio of the average aggregated output rate to the average aggregated input rate when all the input ports carry 100% traffic at line rate. It is a positive value no more than one.

- Speedup

  A speedup of *k* means that the internal forwarding rate of the switch fabric is *k* times the input line rate. So when a speedup exceeds one, buffers must be used at output ports.

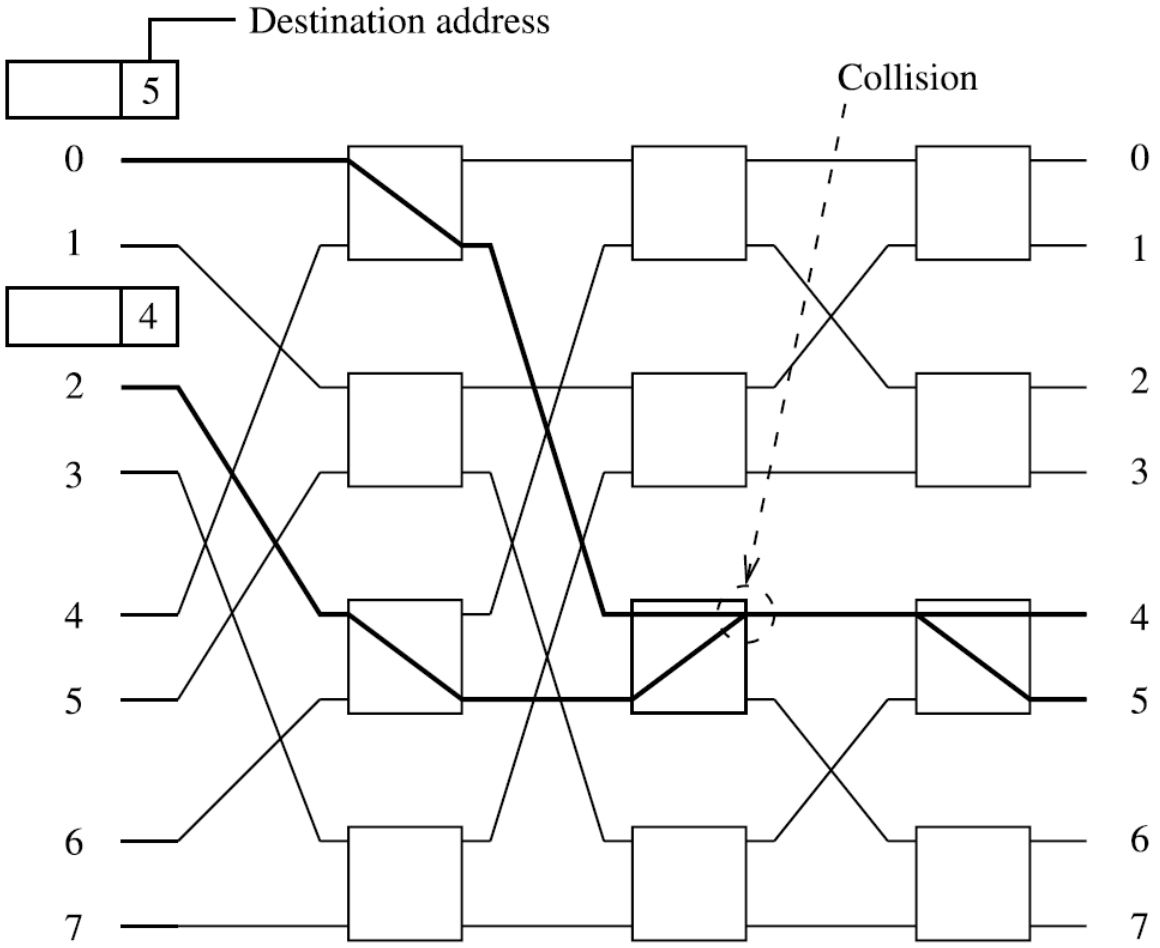
Basics of Packet Switching

(Internal) Blocking

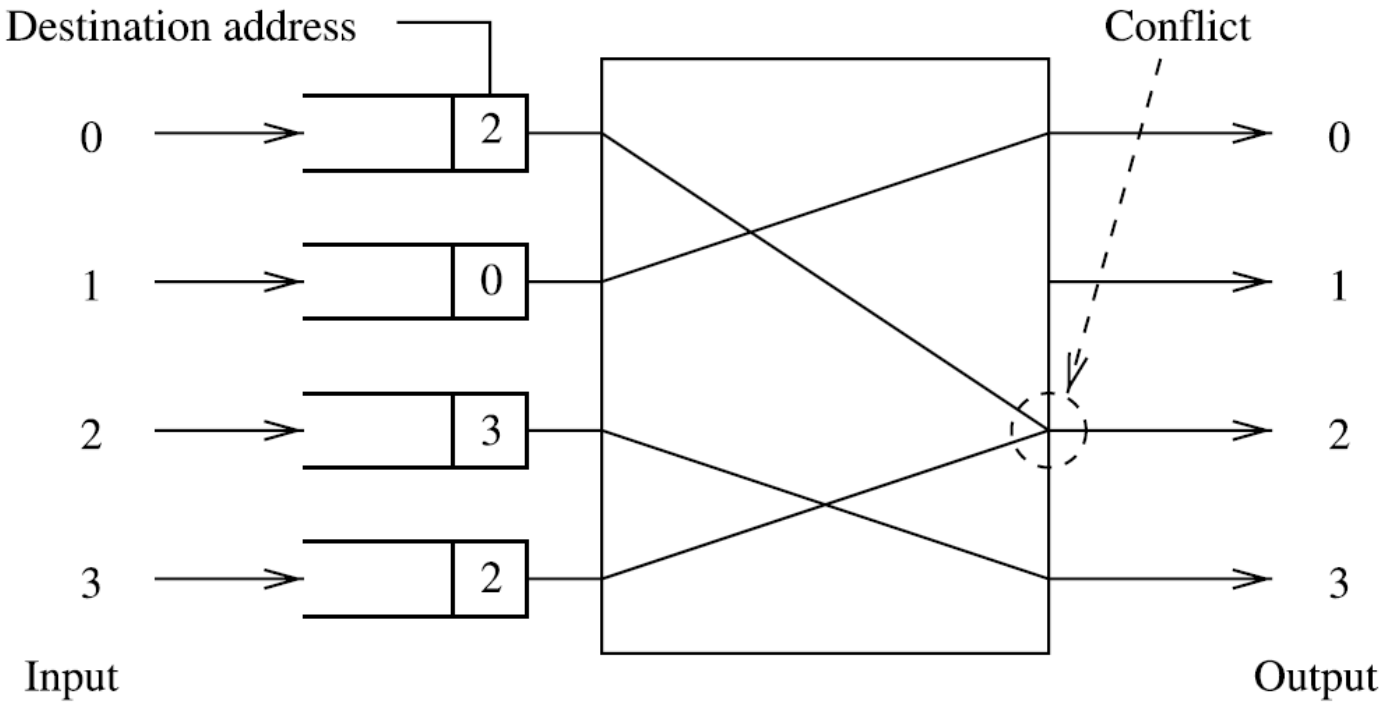


Basics of Packet Switching

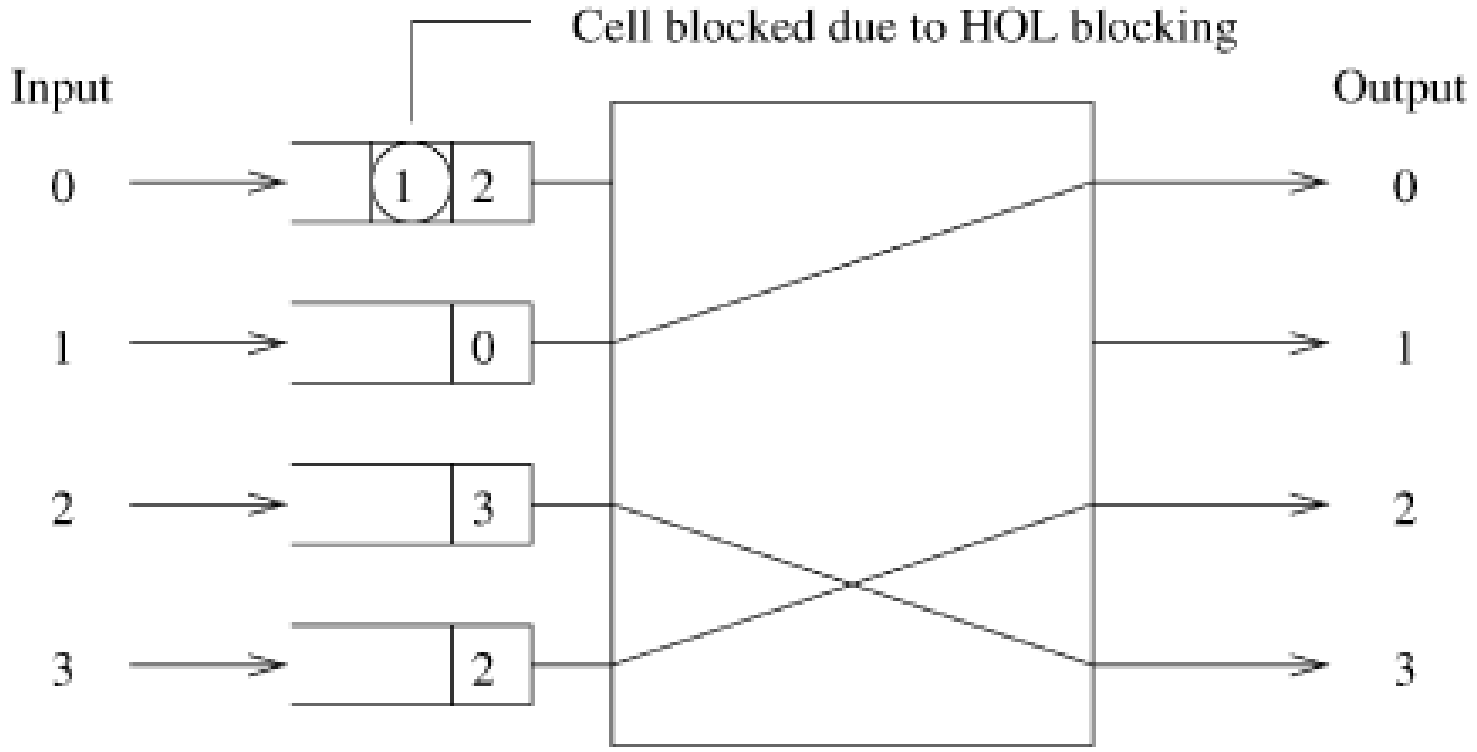
(Internal) Blocking



Output Port Contention

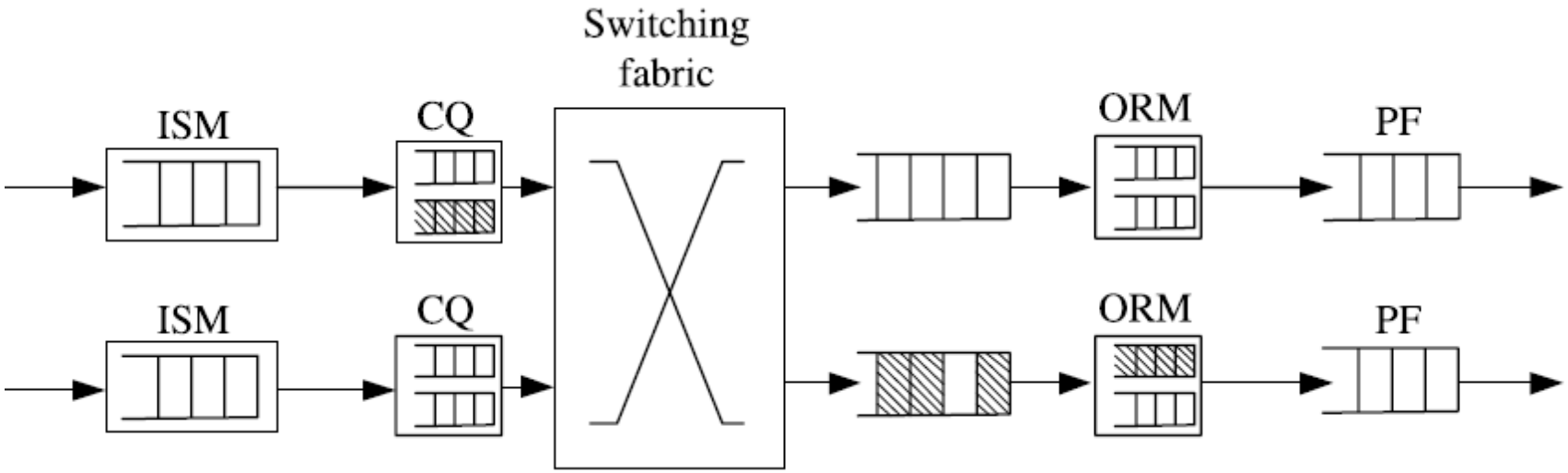


Head of Line (HOL) Blocking



Basics of Packet Switching

Cell-mode switching and Packet-mode switching



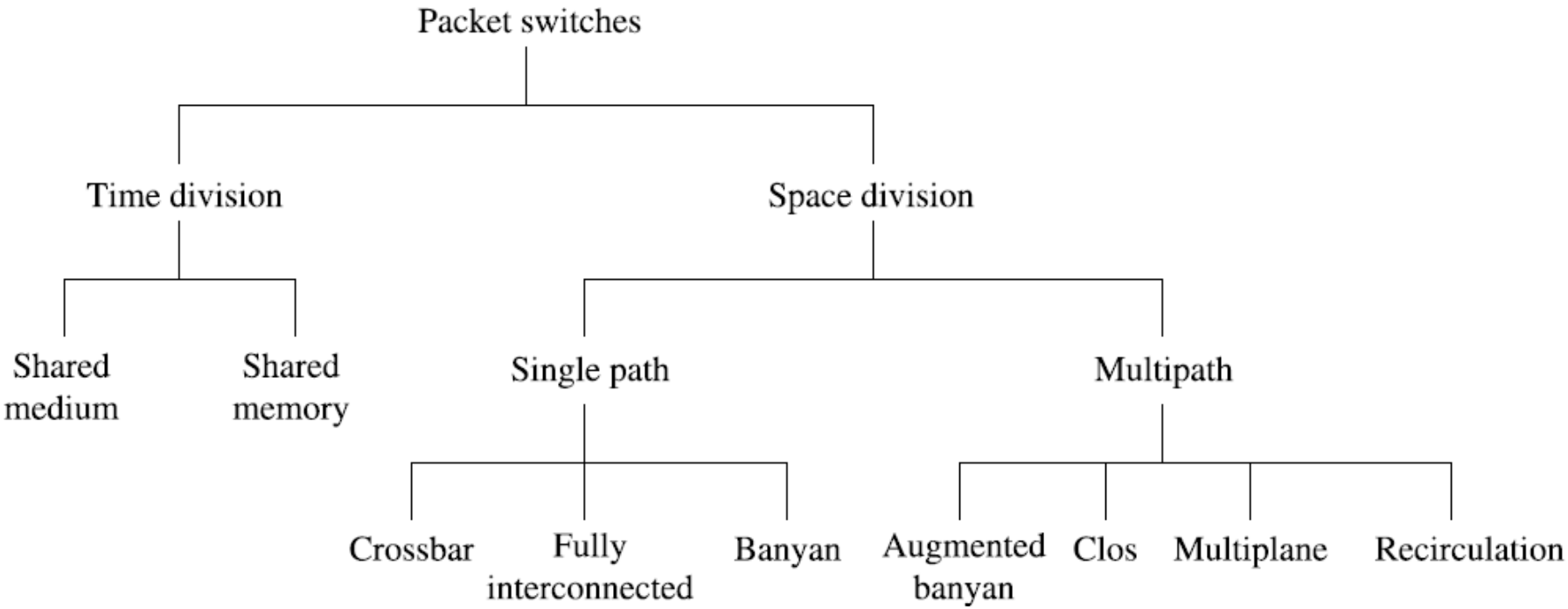
ISM: Input segmentation module

ORM: Output reassembly module

CQ: Virtual output cell queue

PF: Packet FIFO queue

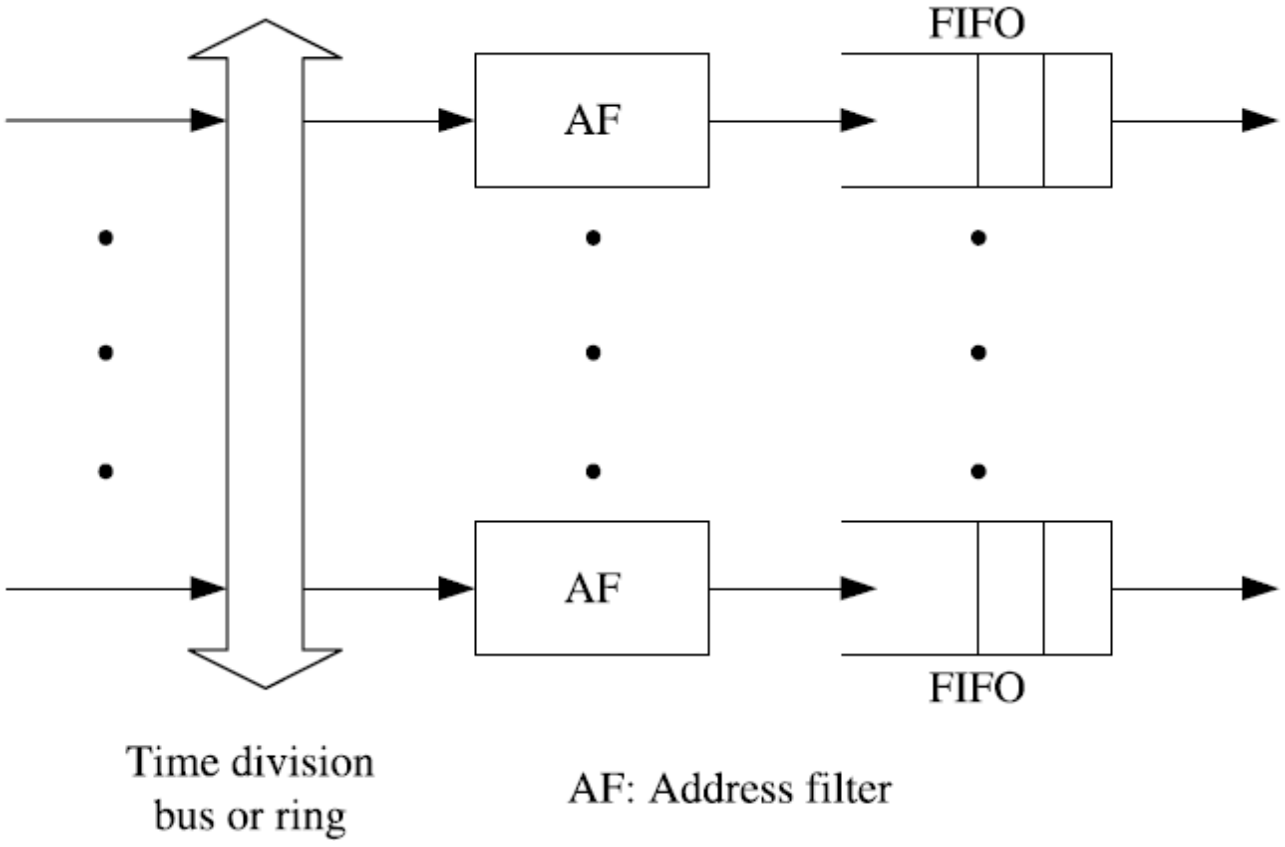
Switch Fabric Classification



Time Division Switching

- Shared-Medium Switch
- Shared-Memory Switch

Shared-Medium Switch



Shared-Medium Switch

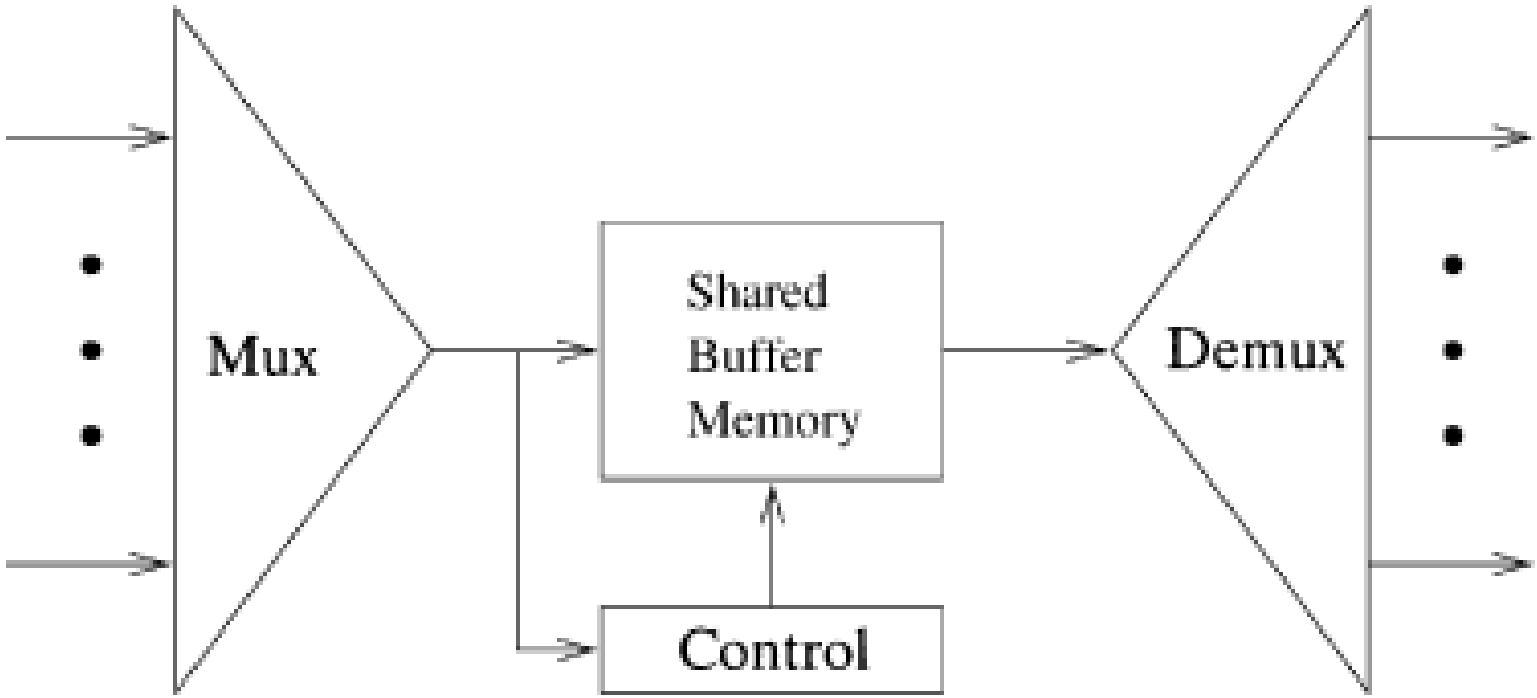
$$(N+1) \leq \frac{T_{\text{cell}}}{T_{\text{mem}}}$$

For example, the cell slot time for a 64-byte packet at a 10G interface is

$$T_{\text{cell}} = \frac{\text{Packetsize}}{\text{throughput}} = (64 \times 8)/(10 \times 10^9) = 51.2\,\text{ns}.$$

If T_{mem} = 4 ns, a total of $(51.2/4) - 1 = 11$ ports can be supported.

Shared-Memory Switch

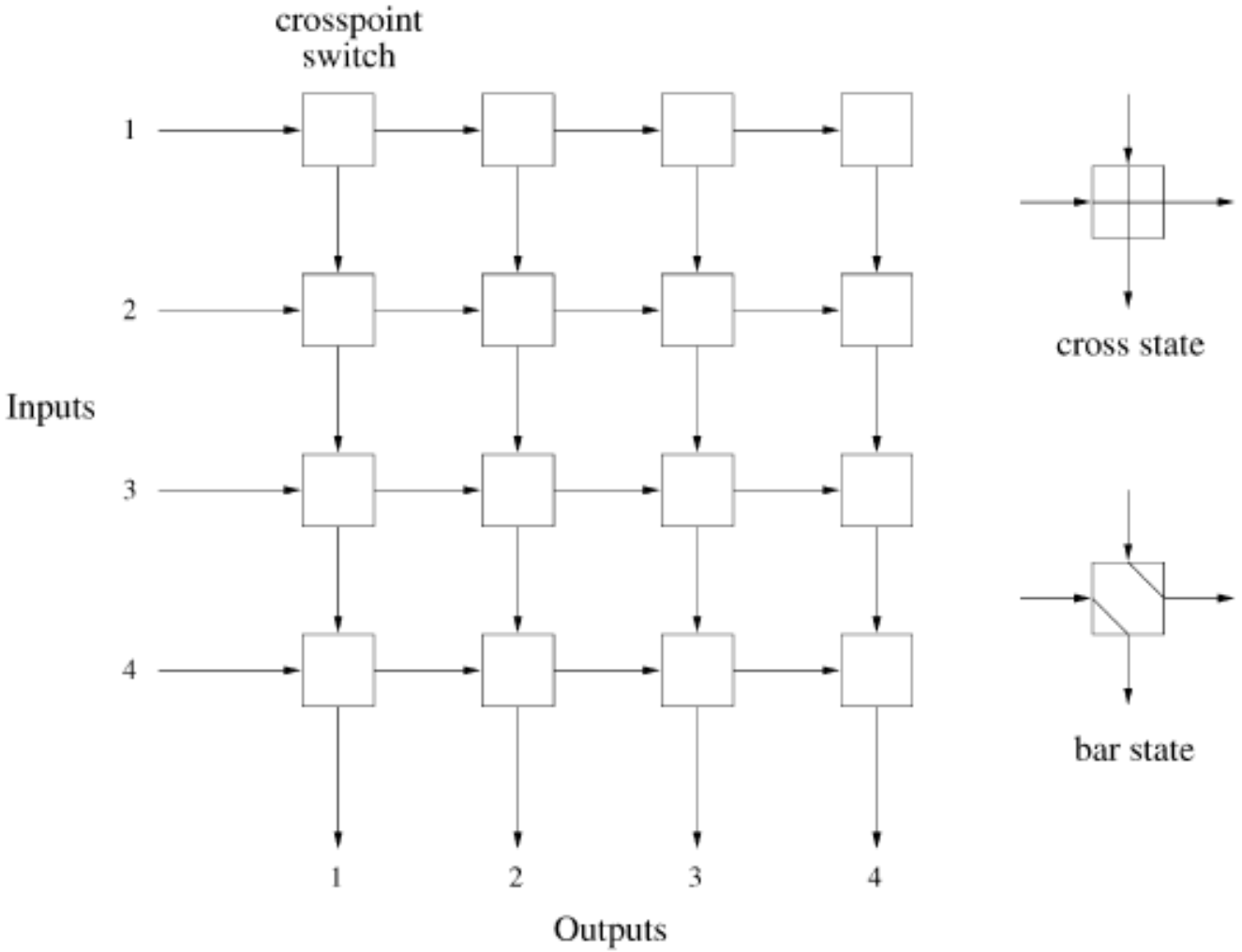


Mux : Multiplexer Demux : Demultiplier

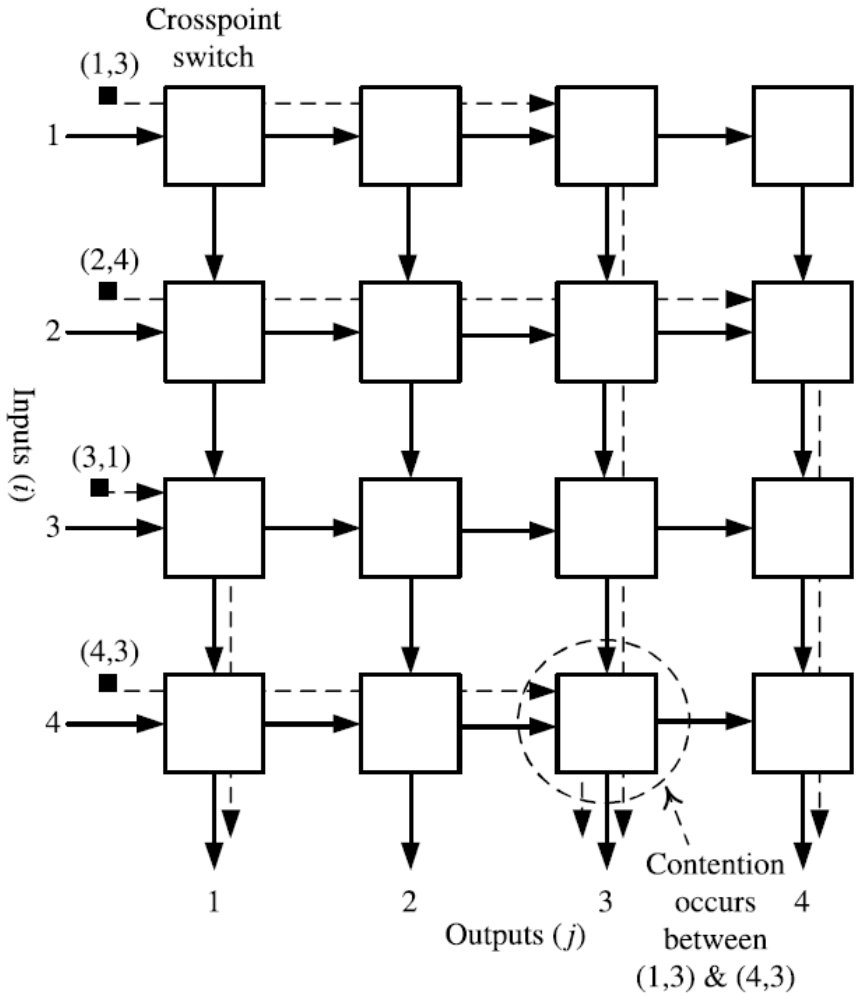
Space division Switching

- **Single-Path Switches**
  - Crossbar Switches
  - Fully Interconnected Switches
  - Banyan-Based Switches
- **Multiple-Path Switches**
  - Augmented Banyan Switches
  - Clos Switches
  - Multiplane Switches
  - Recirculation Switches

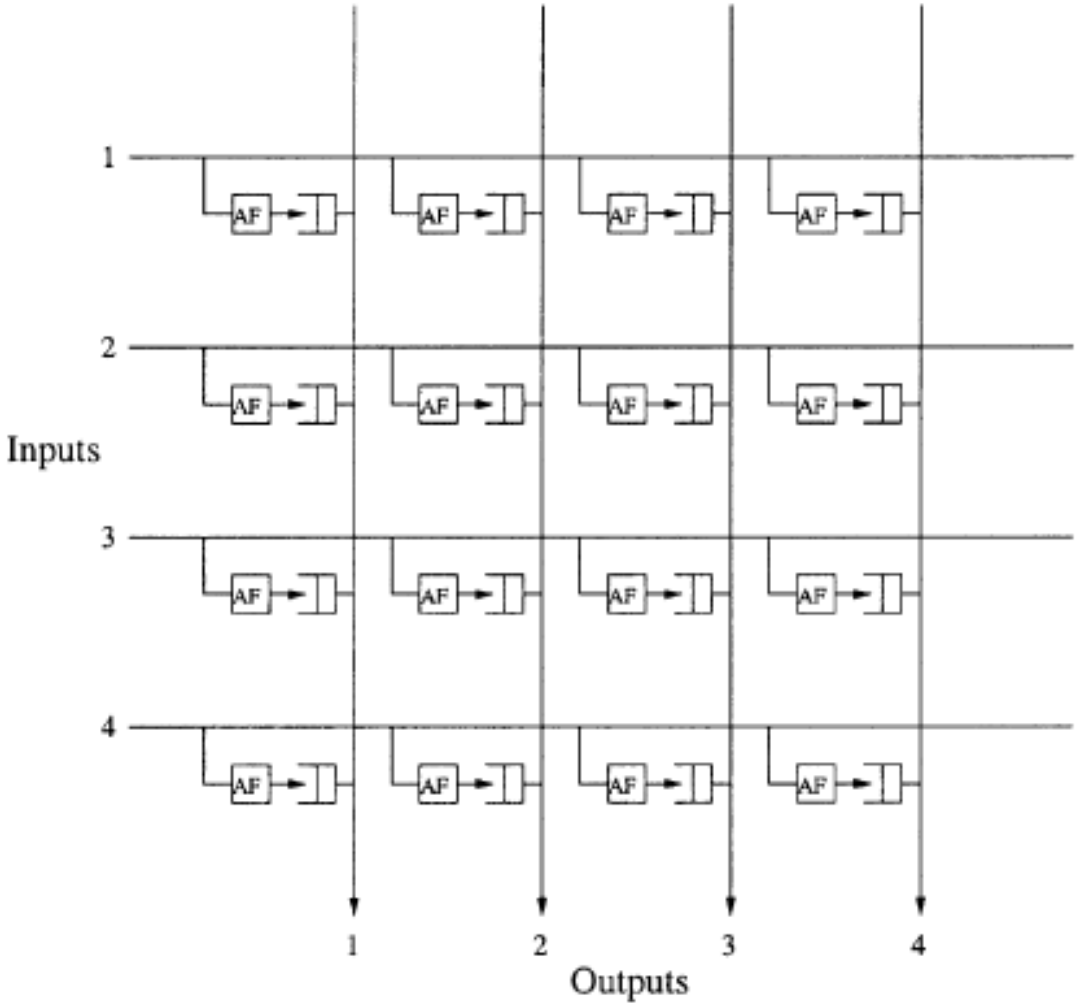
Space Division - Crossbar Switches (1)



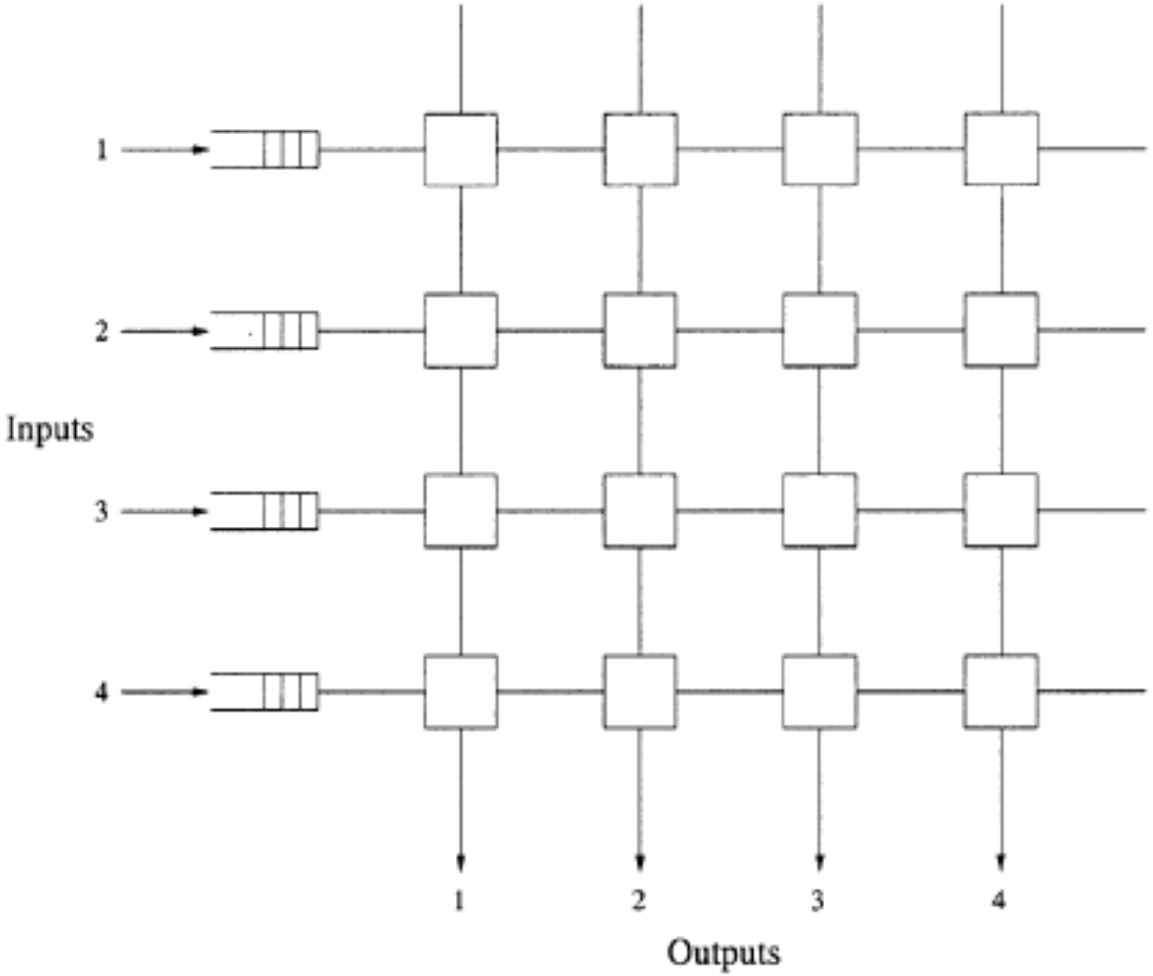
Space Division - Crossbar Switches (2)



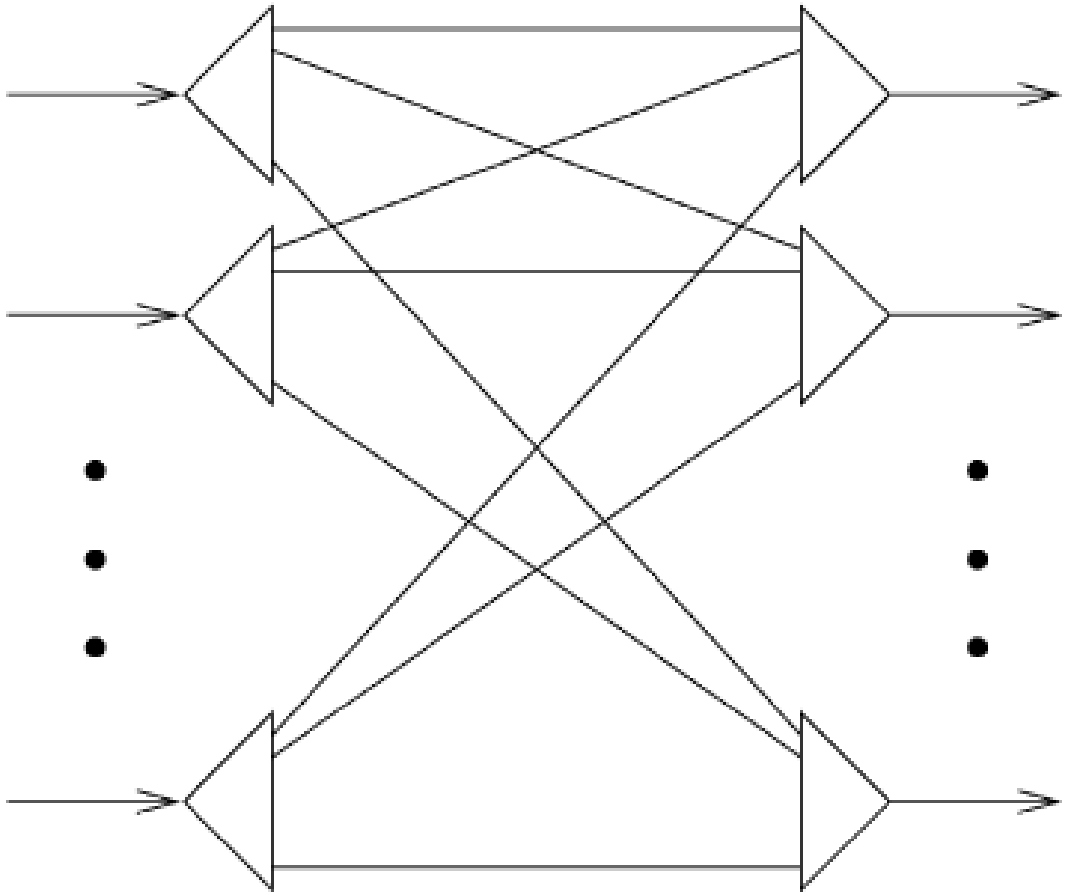
Space Division - Crossbar Switches (3)



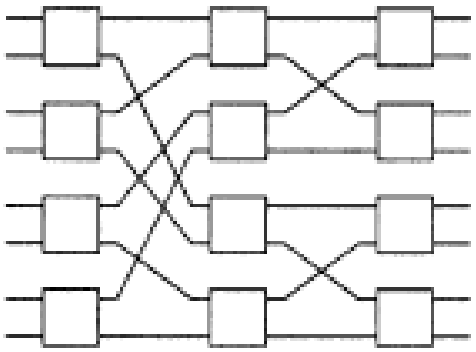
Space Division - Crossbar Switches (4)



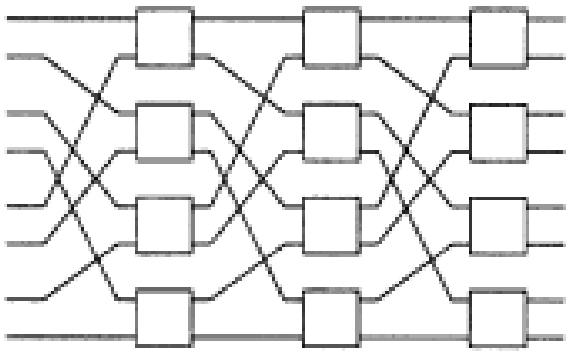
Space Division - Fully Interconnected Switches



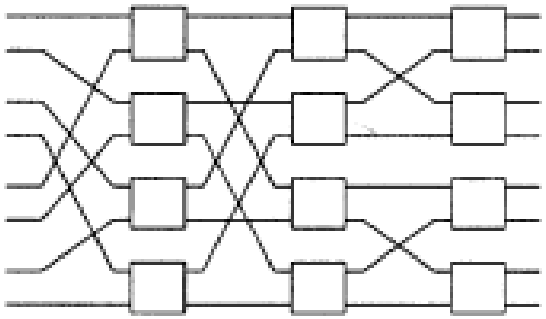
Space Division - Banyan-Based Switches



(a) Delta network



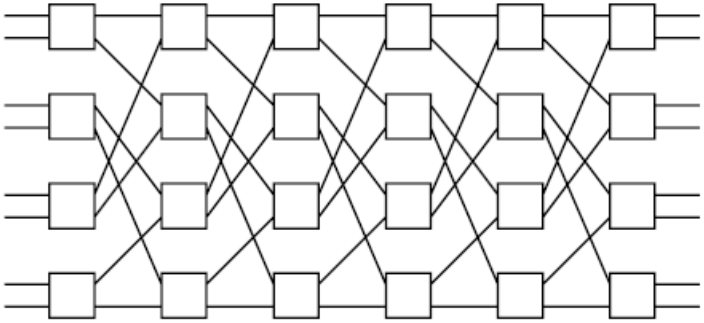
(b) Omega network



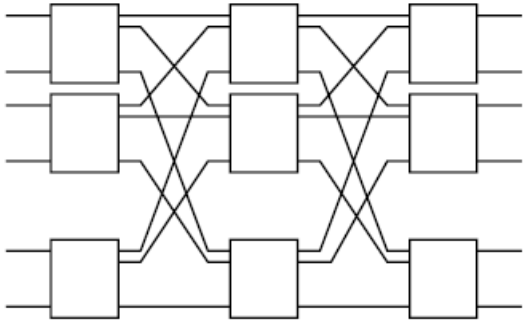
(c) Banyan network

: 2 × 2 Switch Element (SE)

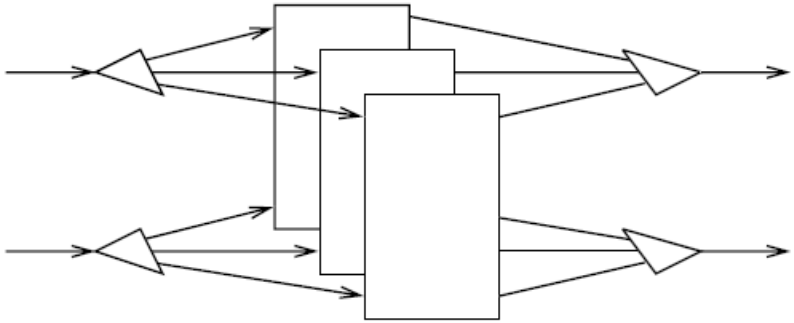
Space Division - Multiple-Path Switches



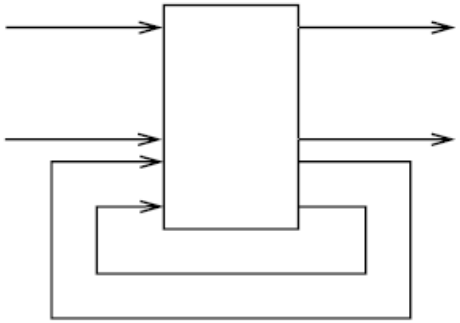
(a)



(b)

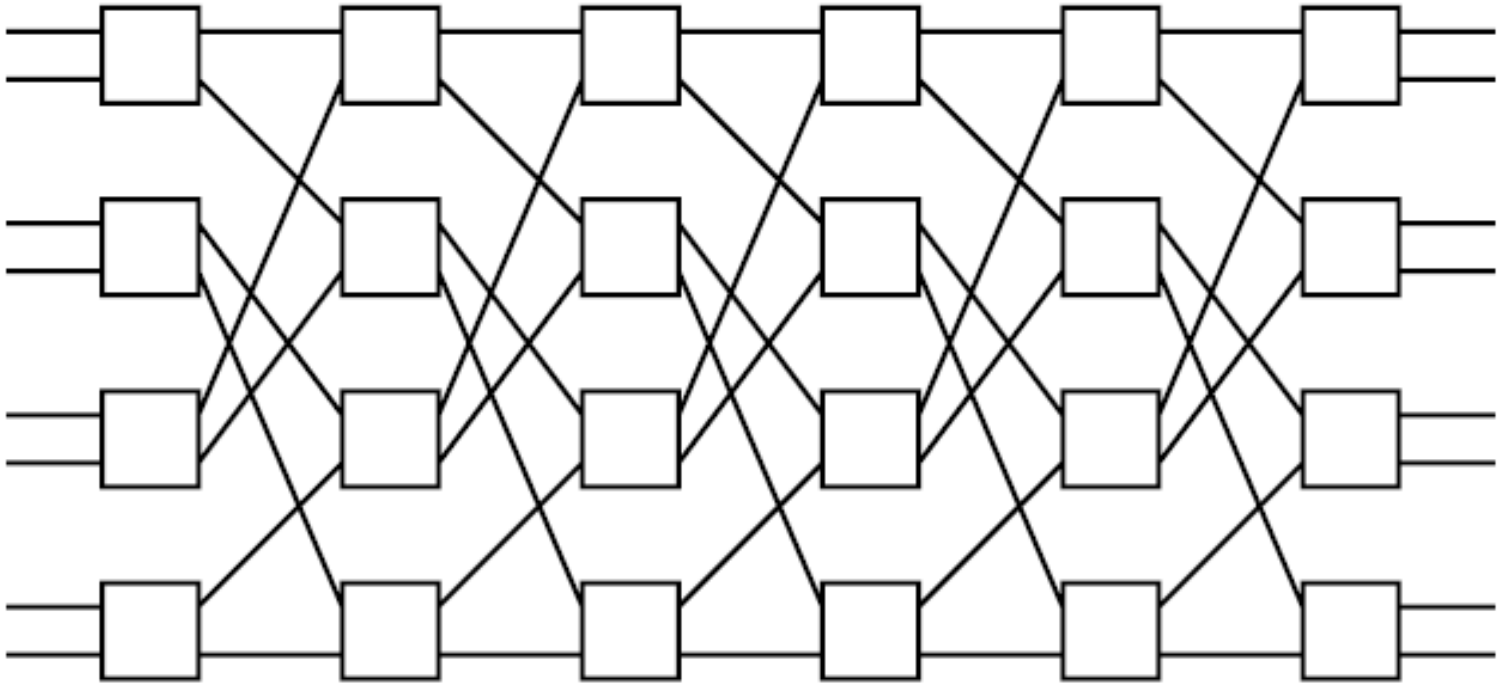


(c)

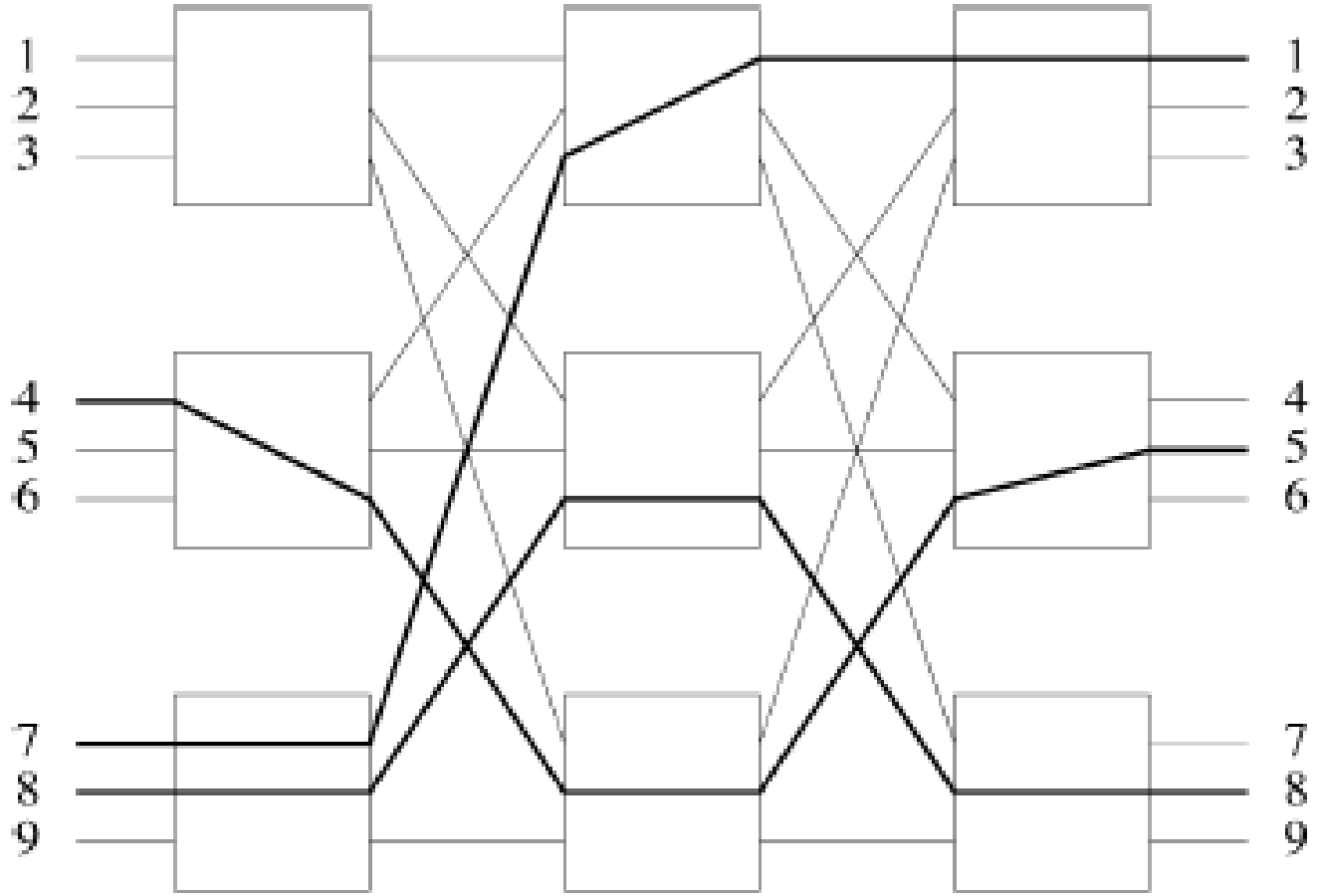


(d)

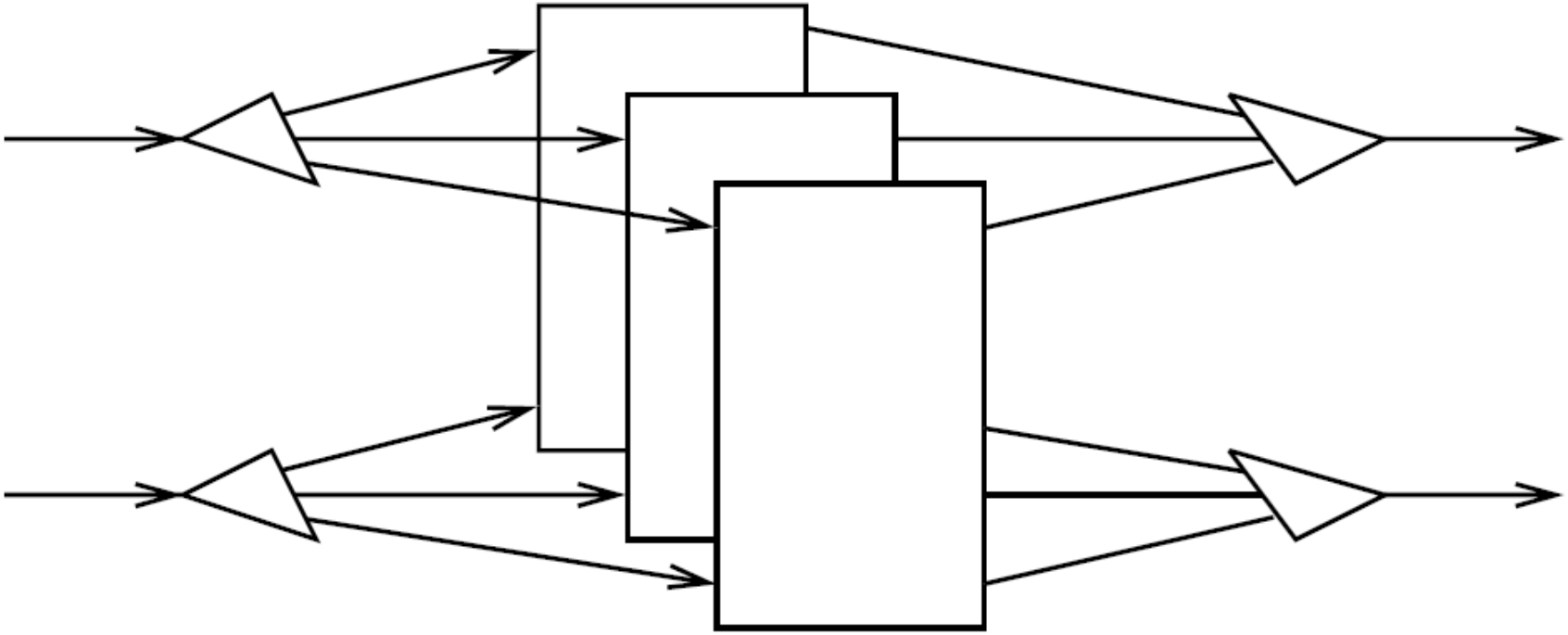
Space Division - Augmented Banyan Switches



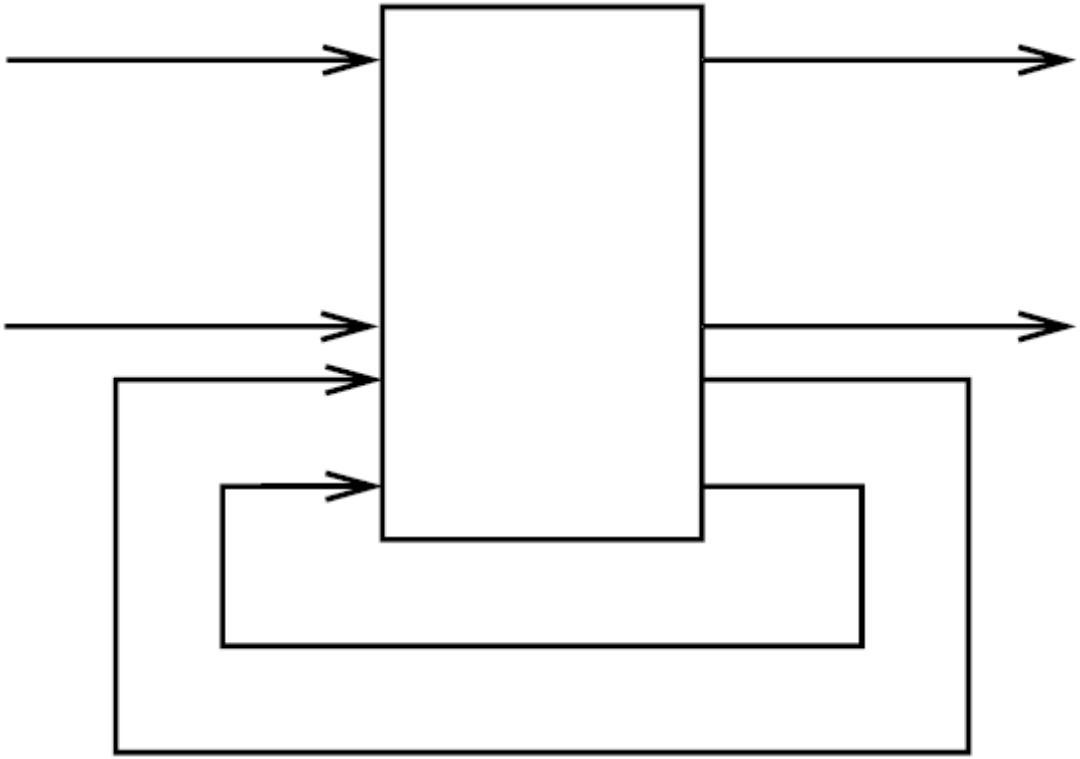
Space Division - Three-Stage Clos Switches



Space Division - Multiplane Switches



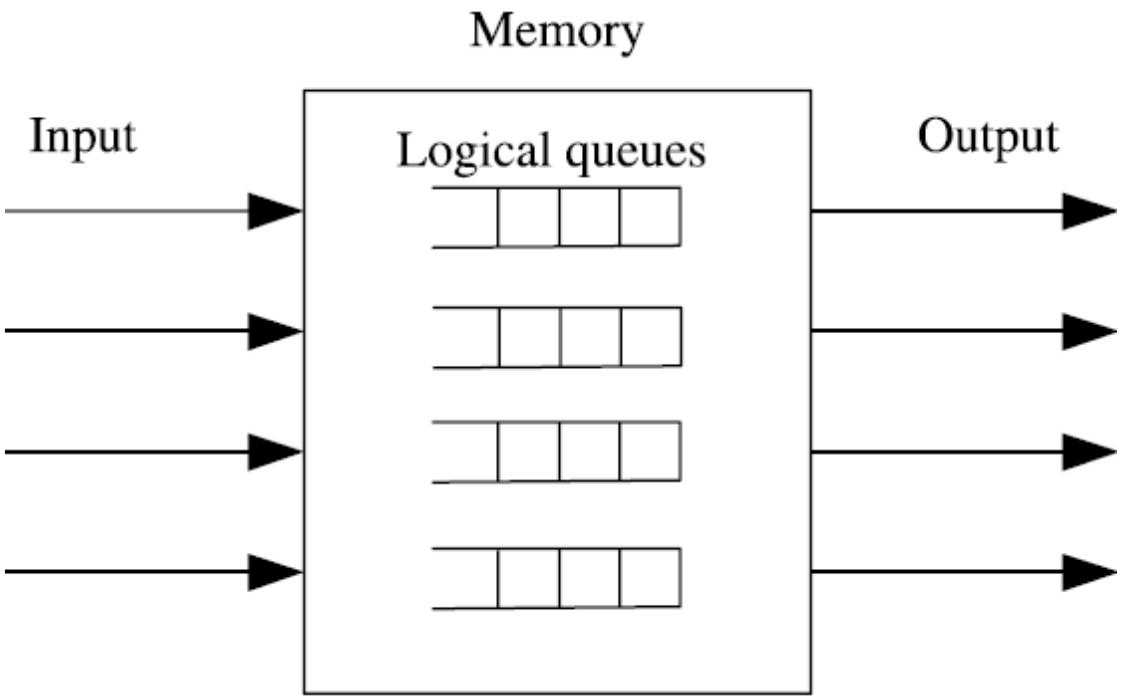
Space Division -Recirculation Switches



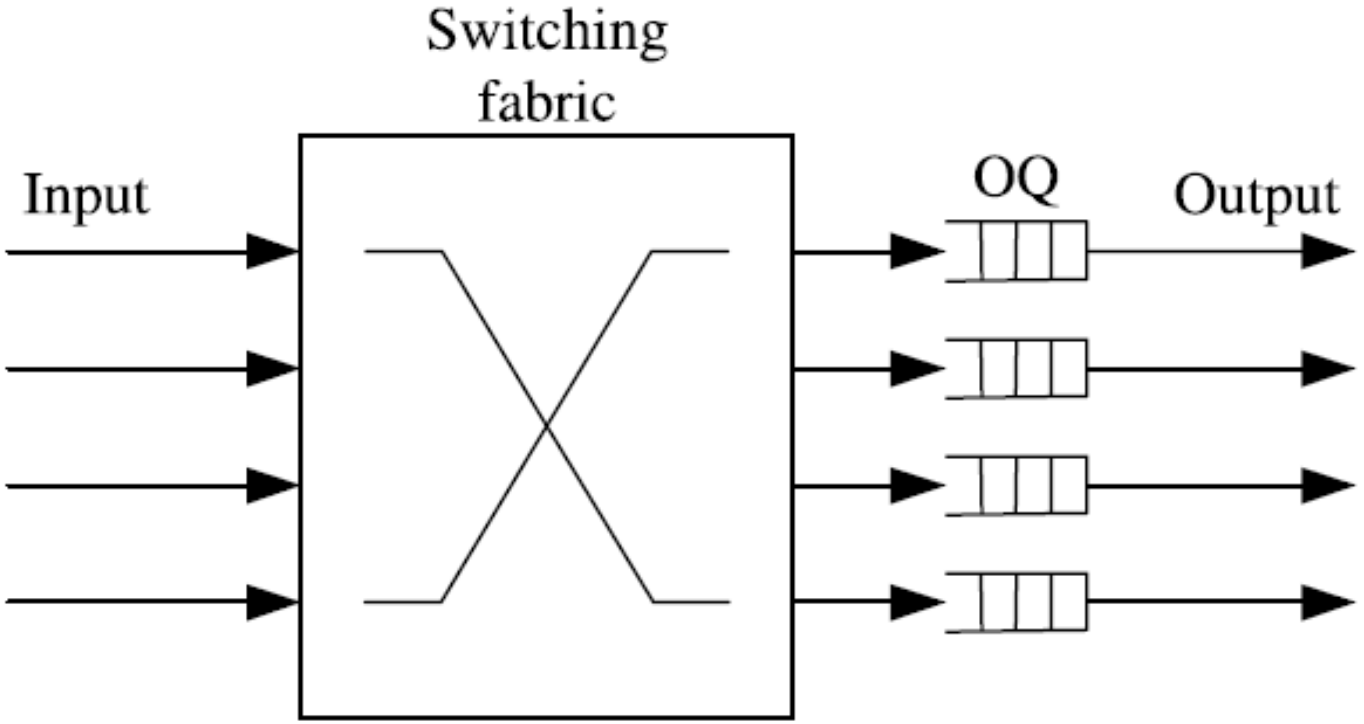
Buffering Strategies

- Shared-Buffer Queueing,
- Output Queueing,
- Input Queueing,
- Virtual-Output-Queueing (VOQ),
- Combined Input and Output Queueing,
- Crosspoint Queueing.

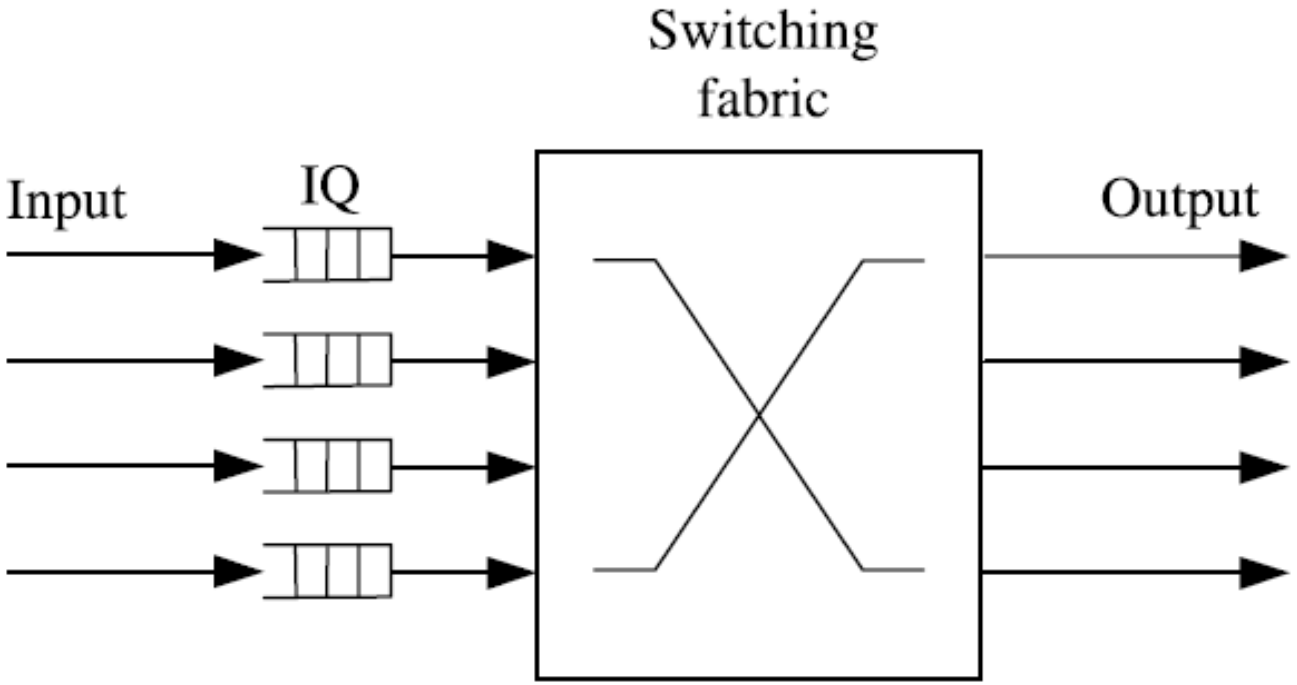
Shared-Memory Queueing



Output Queueing (OQ)



Input Queueing

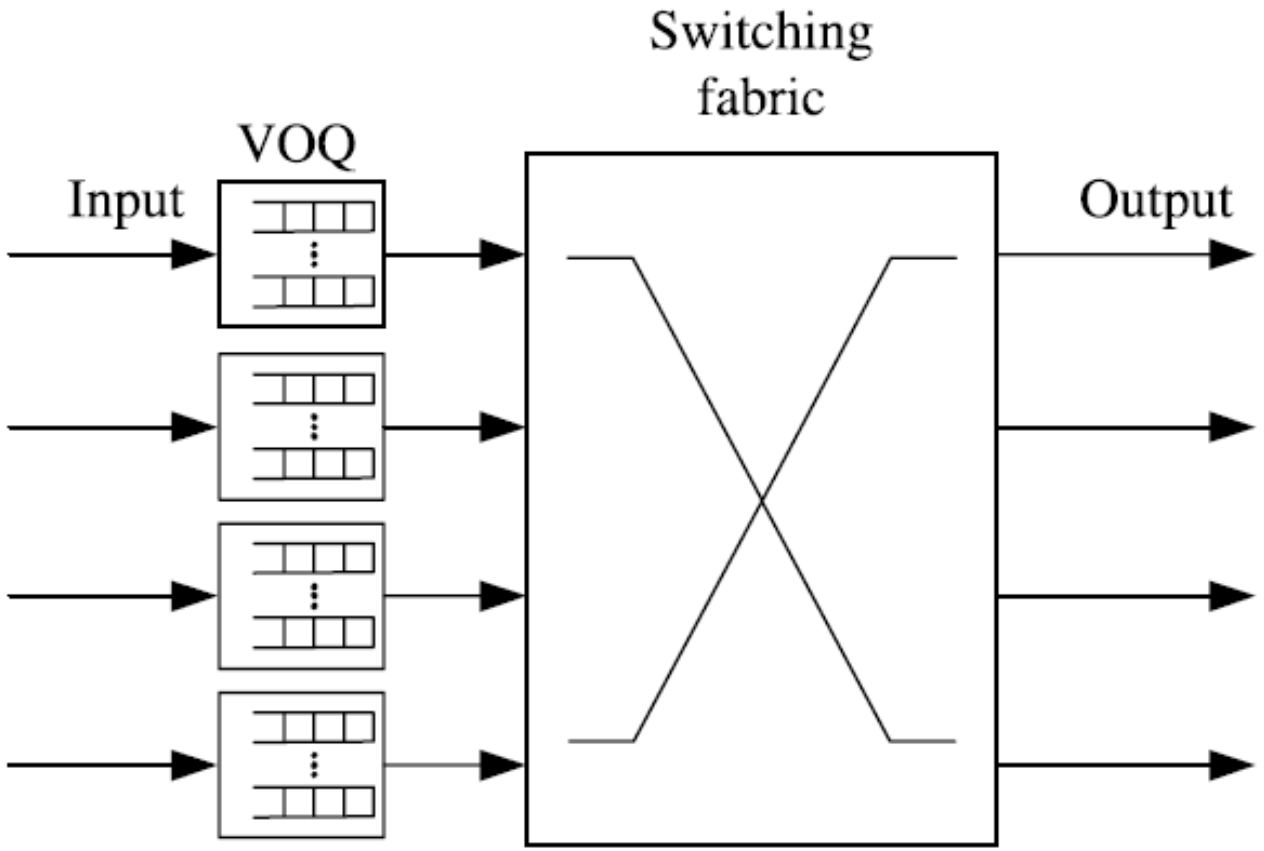


Input-Buffered Switches

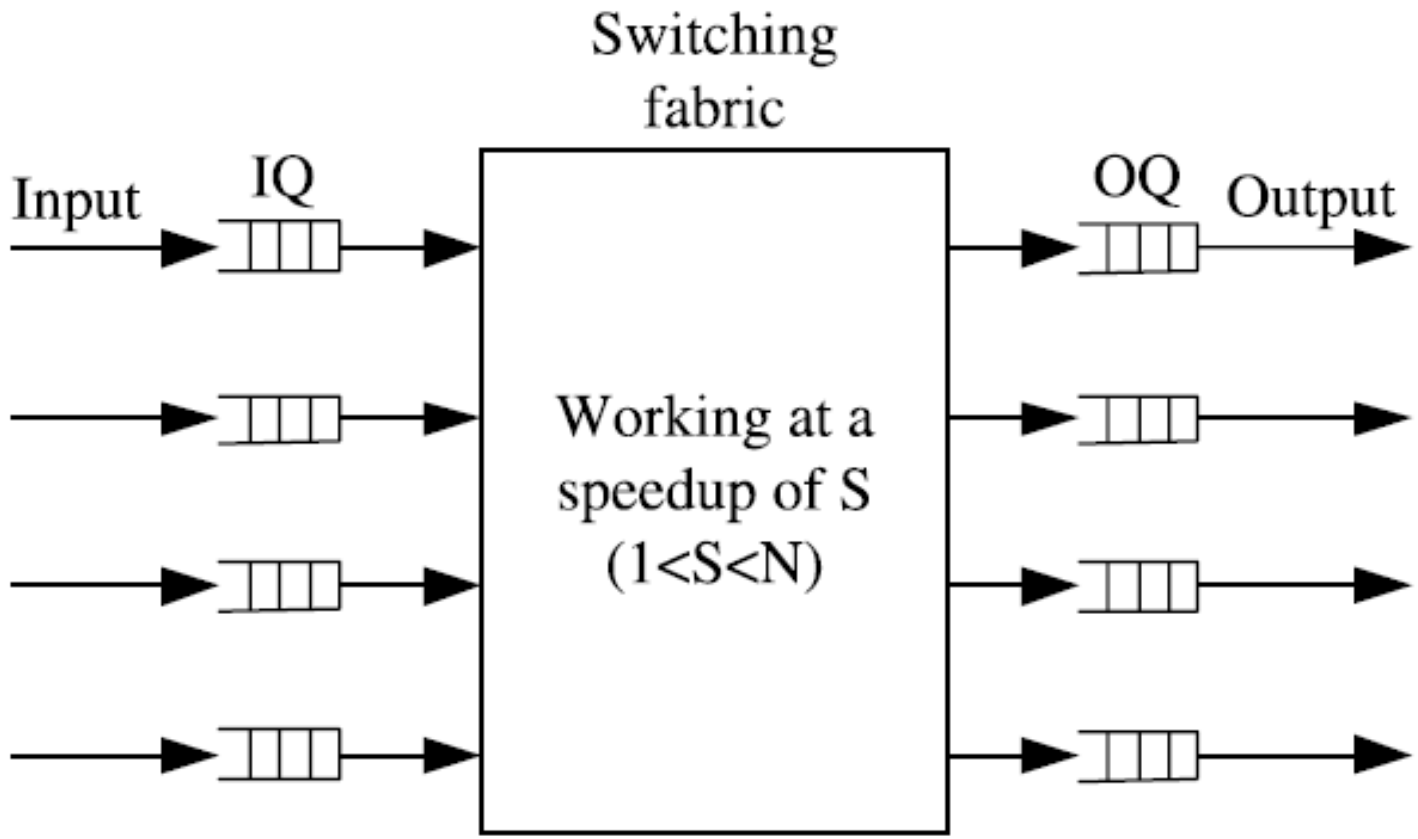
- **HOL blocking problem:**
- The throughput limitation is 58.6% for uniform traffic.

- By using *Windowing* technique, the throughput will be increased. For instance, by increasing the window size to two, the maximum throughput is increased to 70%.

Virtual Output Queuing (VOQ)

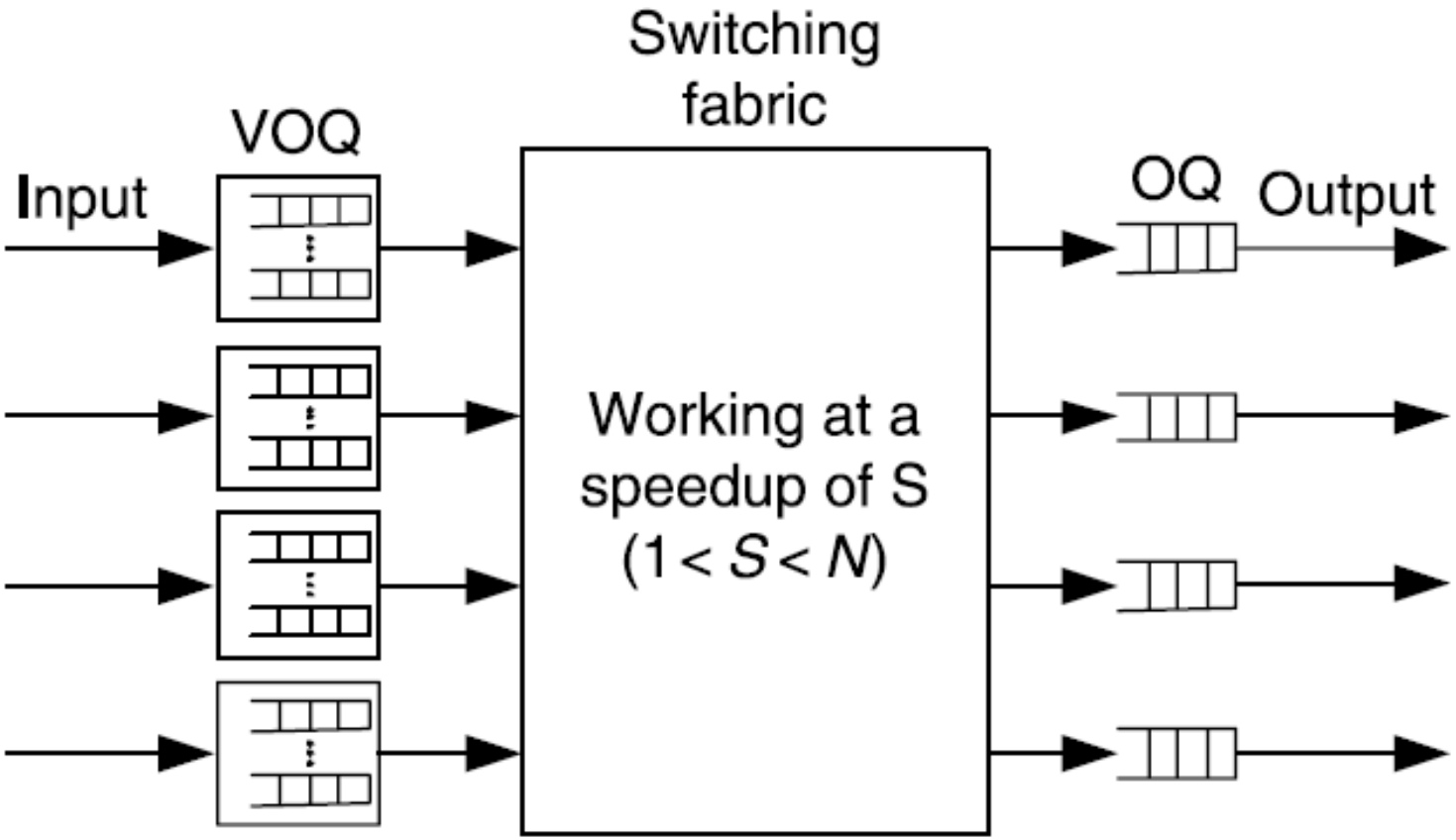


Combined Input and Output Queueing

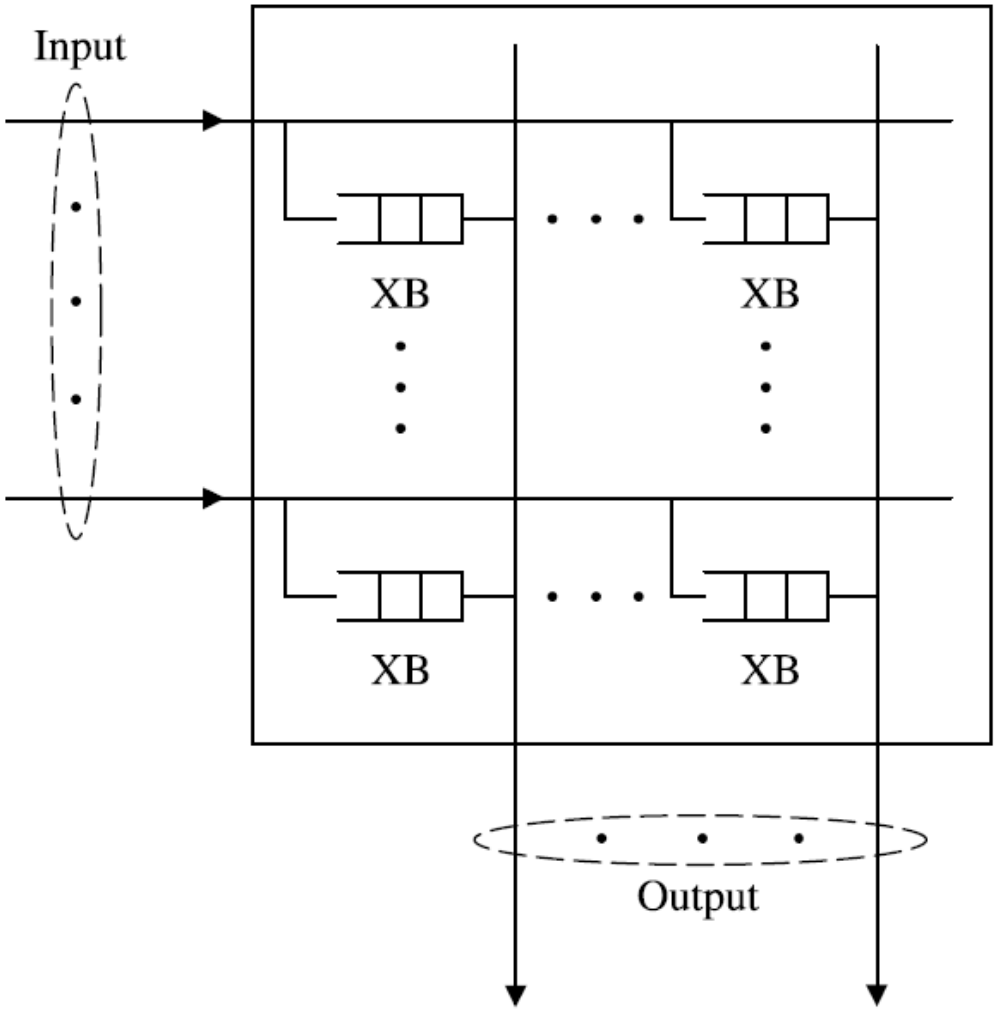


Basics of Packet Switching

Combined Input and Output Queueing with VOQ



Crosspoint Queueing

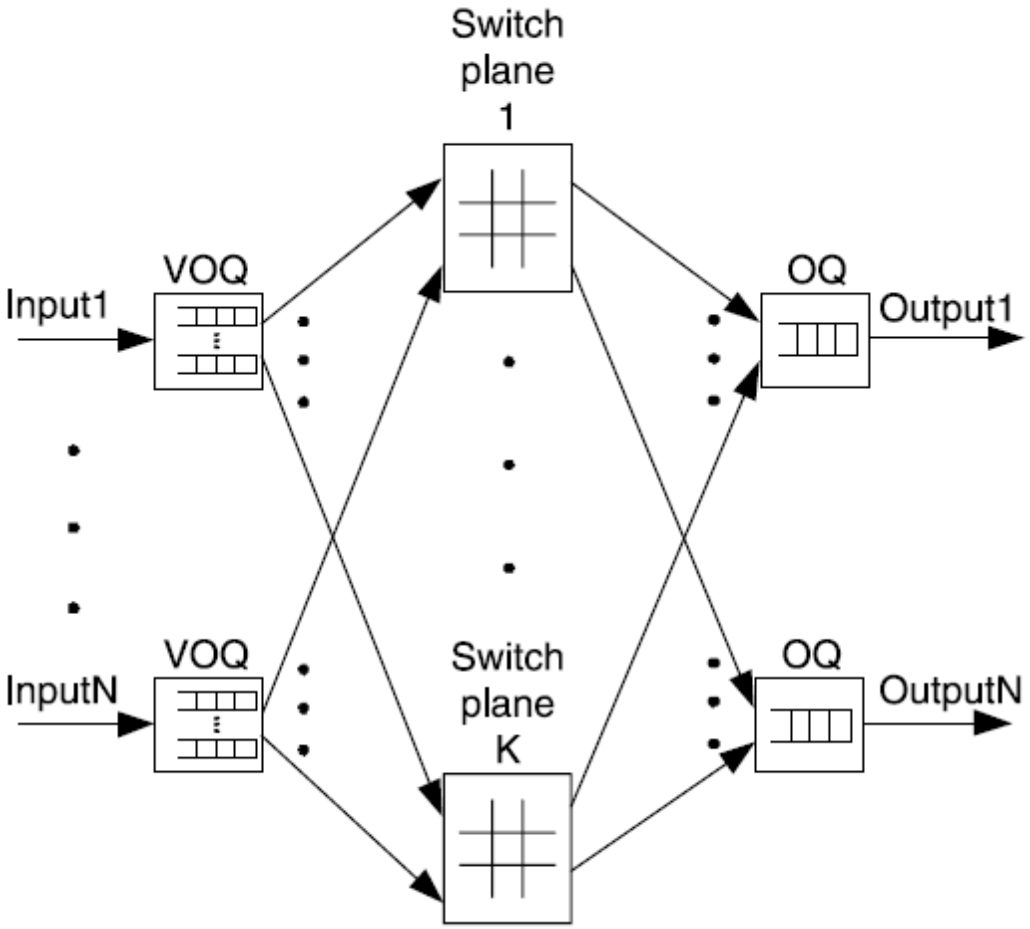


Basics of Packet Switching

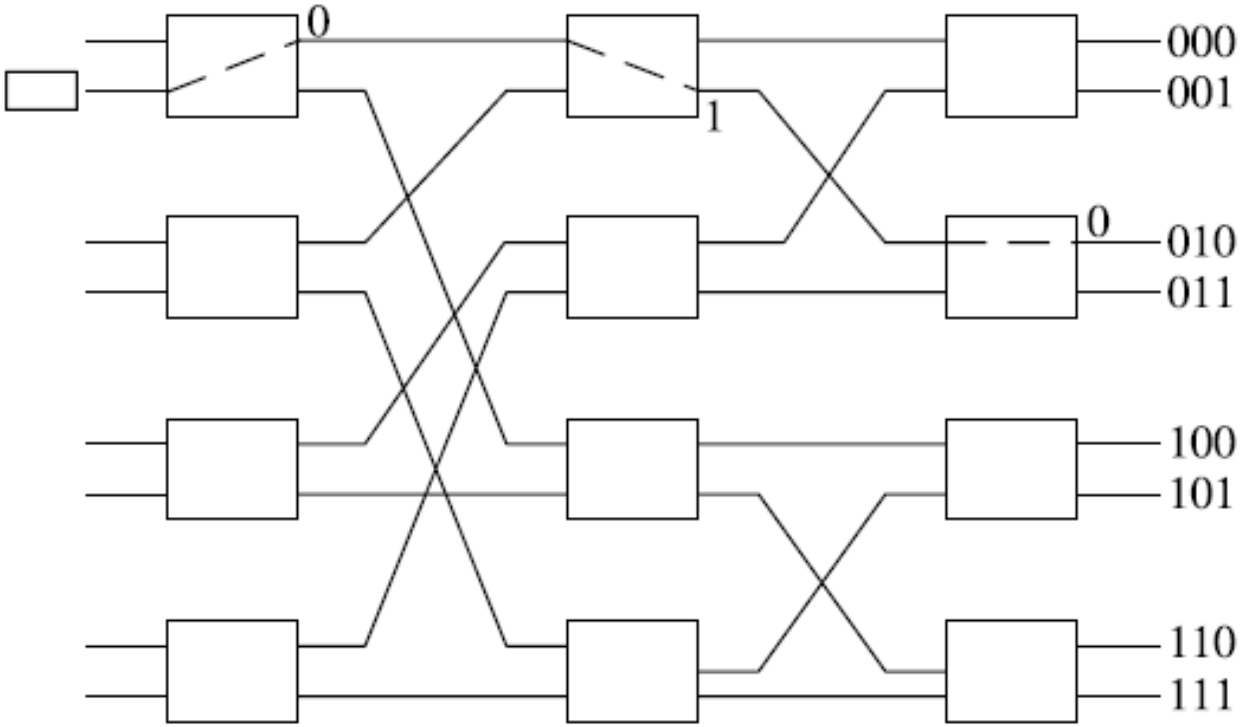
Multiplane Switching and Multistage Switching

- **Parallel Packet Switch,**
- **Banyan-Based Switch,**
- **CLOS Switches.**

Parallel Packet Switch

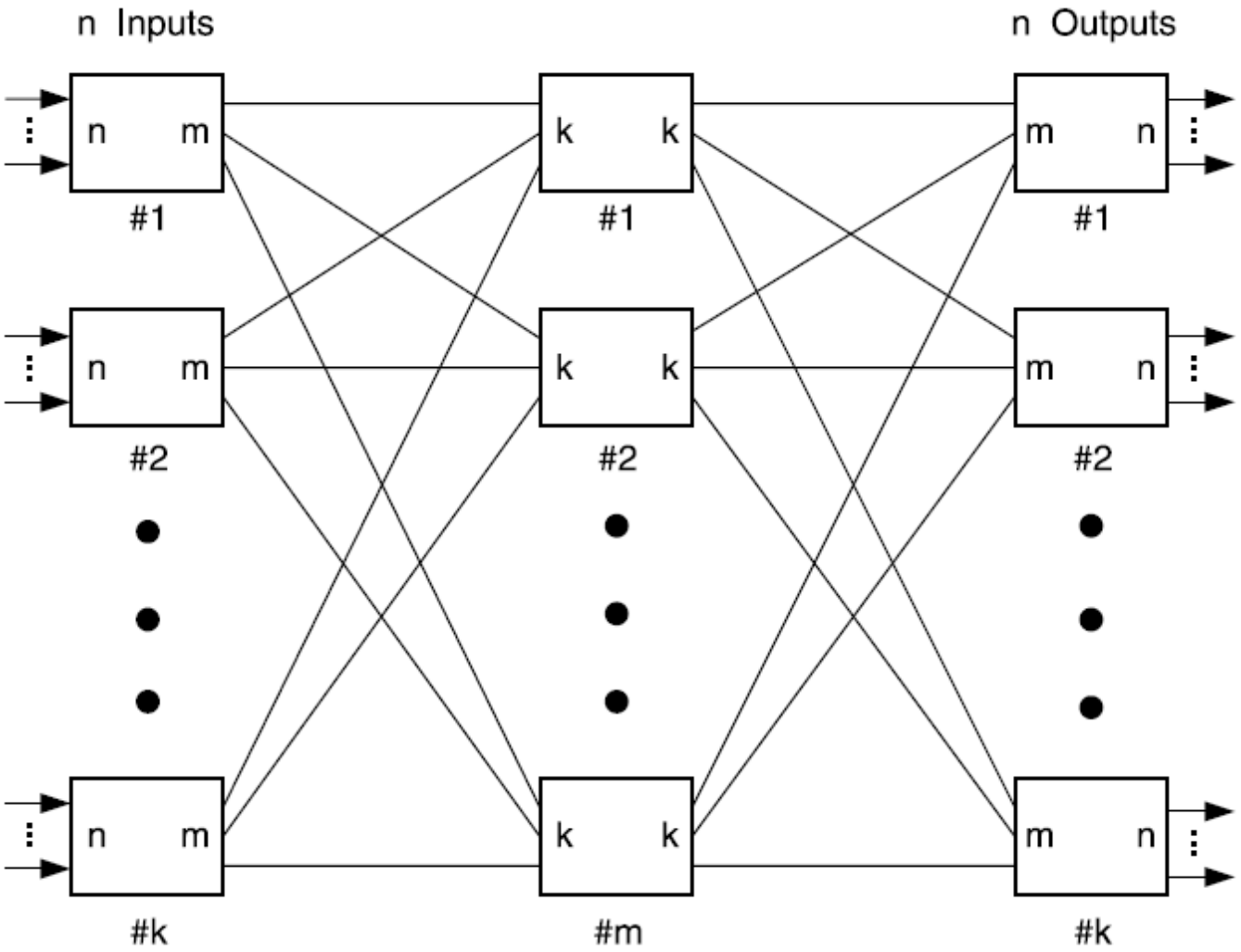


Banyan-Based Switch



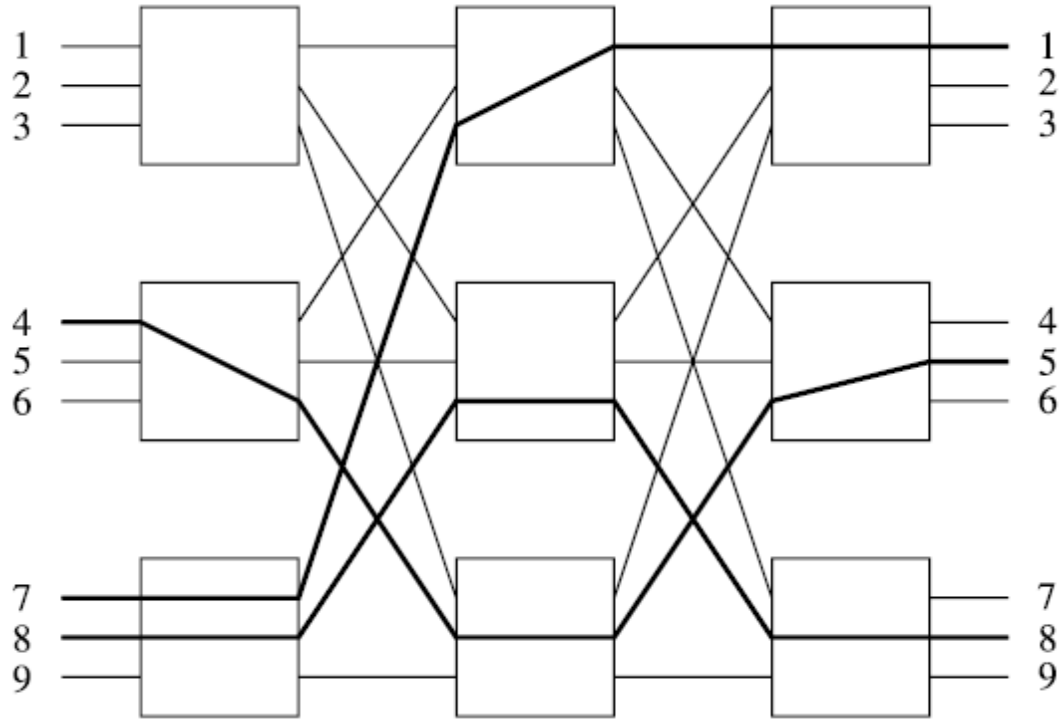
Basics of Packet Switching

CLOS Switching



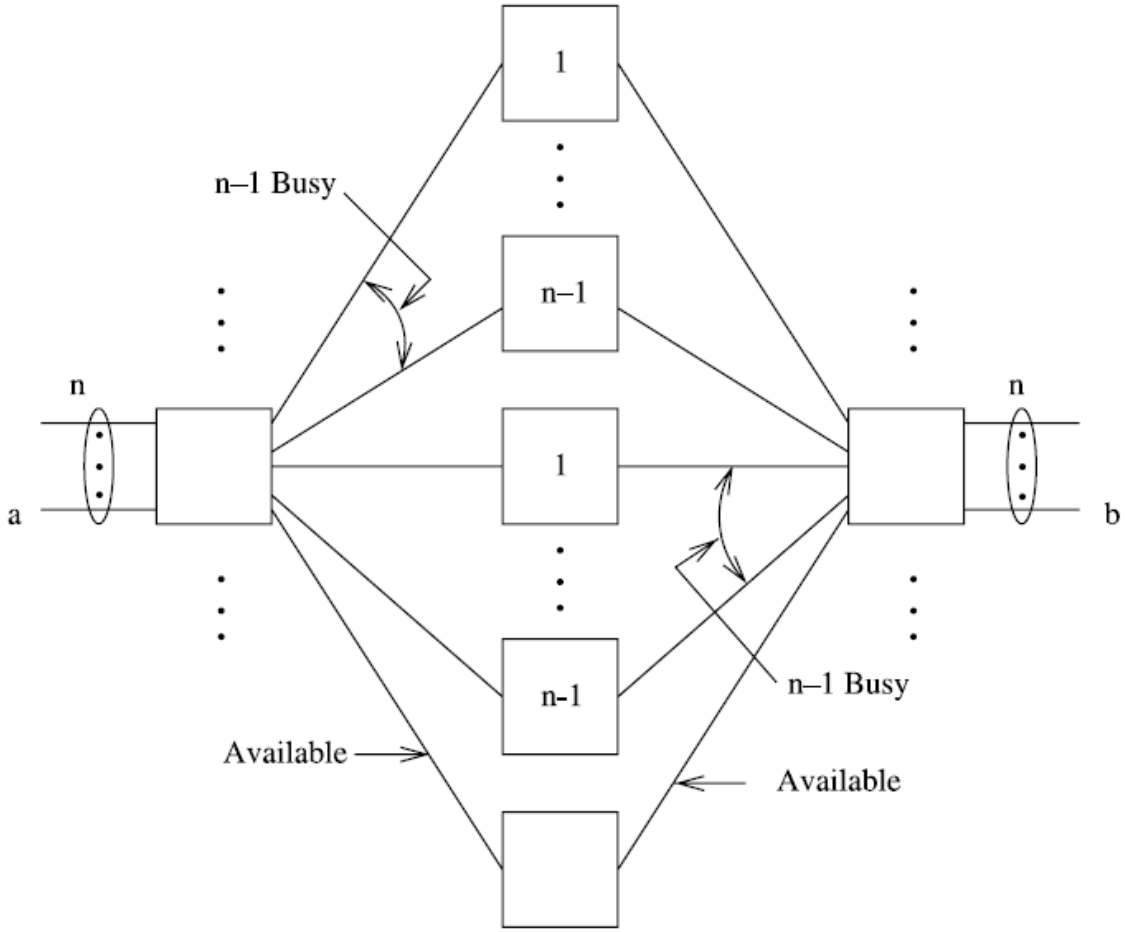
Basics of Packet Switching

CLOS Switching



Basics of Packet Switching

CLOS Switching



Basics of Packet Switching

CLOS Switching (nonblocking condition)

$$m \geq (2n - 2) + 1 = 2n - 1$$

The total number of crosspoints N_x in a three-stage Clos switch when it is symmetric (i.e., $m = k$) is

$$N_x = 2Nm + m\left(\frac{N}{n}\right)^2$$

Substituting $m = 2n - 1$ into N_x, we obtain

$$N_x = 2N(2n - 1) + (2n - 1)\left(\frac{N}{n}\right)^2$$

for a nonblocking switch. For a large switch size, n is large. We can approximate

$$N_x \simeq 2N(2n) + 2n\left(\frac{N}{n}\right)^2 = 4Nn + 2\left(\frac{N^2}{n}\right)$$

To optimize the number of crosspoints, differentiate N_x with respect to n and set the result to 0. The result will be $n \simeq (N/2)^{1/2}$. Substituting into N_x,

$$N_x = 4\sqrt{2}N^{3/2} = O(N^{3/2})$$

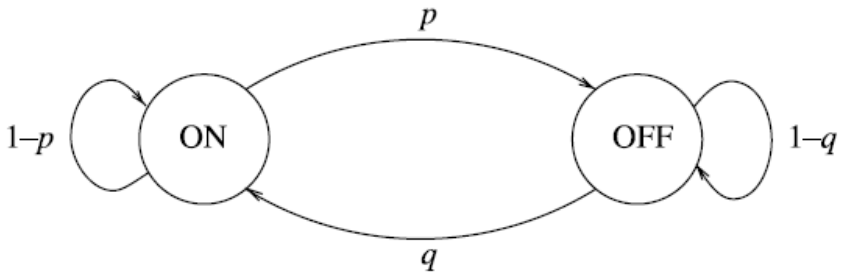
Performance of Basic Switches

- Input-buffered switches,
- Output-buffered switches,
- Completely shared-buffer switches,

Performance of Basic Switches

Traffic Model

- ***Bernoulli Arrival Process and Random Traffic,***
- ***On–Off Model and Bursty Traffic.***



$$\rho = \frac{1/p}{1/p + \sum_{j=0}^{\infty} jq(1-q)^j} = \frac{q}{q+p-pq}.$$

$$\Pr\{X = i\} = p(1-p)^{i-1}, \quad i \geq 1$$

$$\beta = E[X] = \sum_{i=1}^{\infty} i \Pr\{X = i\} = 1/p.$$

$$\Pr\{Y = j\} = q(1-q)^{j-1}, \quad j \geq 0$$

$$\alpha = E[Y] = \sum_{j=0}^{\infty} j \Pr\{Y = j\} = (1-q)/q.$$

Input-buffered switches

B_m^i : the number of remaining cells destined for output i in the mth time slot,

A_m^i : the number of cells moving to the heads of the input queues during the mth time slot and destined for output i,

$$B_m^i = \max\left(0, B_{m-1}^i + A_m^i - 1\right)$$

Input-buffered switches

We assume that each new cell arrival at the head of an input queue has the equal probability $1/N$ of being destined for any given output. As a result, A_m^i has the following binomial distribution:

$$\Pr\left[A_m^i = k\right] = \binom{F_{m-1}}{k}\left(\frac{1}{N}\right)^k\left(1-\frac{1}{N}\right)^{F_{m-1}-k}, \qquad k = 0, 1, \ldots, F_{m-1}, \tag{2.2}$$

where

$$F_{m-1} \triangleq N - \sum_{i=1}^{N} B_{m-1}^i. \tag{2.3}$$

F_{m-1} represents the total number of cells transmitted through the switch during the $(m-1)$st time slot, which in saturation is equal to the total number of input queues that have a new cell moving into the HOL position in the mth time slot. That is,

$$F_{m-1} = \sum_{i=1}^{N} A_m^i. \tag{2.4}$$

Input-buffered switches

When $N \to \infty$, A_m^i has a Poisson distribution with rate $\rho_m^i = F_{m-1}/N$. In steady state, $A_m^i \to A^i$ also has a Poisson distribution. The rate is $\rho_0 = \overline{F}/N$, where $\overline{F}$ is the average of the number of cells passing through the switch, and ρ_0 is the utilization of output lines (i.e., the normalized switch throughput). The state transition of B^i is driven by the same Markov process as the $M/D/1$ queues in the steady state. Using the results for the mean steady-state queue size for an $M/D/1$ queue, for $N \to \infty$ we have

$$\overline{B^i} = \frac{\rho_0^2}{2(1-\rho_0)}. \tag{2.5}$$

In the steady state, (2.3) becomes

$$\overline{F} = N - \sum_{i=1}^{N} \overline{B^i}. \tag{2.6}$$

Input-buffered switches

By symmetricity, $\overline{B^i}$ is equal for all i. In particular,

$$\overline{B^i} = \frac{1}{N}\sum_{i=1}^{N}\overline{B^i} = 1 - \frac{\overline{F}}{N} = 1 - \rho_0. \qquad (2.7)$$

It follows from (2.5) and (2.7) that $\rho_0 = 2 - \sqrt{2} = 0.586$.

The result is based on the following two assumptions:

1. The arrival process to each input queue is a Bernoulli process. That is, the probability that a cell arrives in each time slot is identical and independent of any other slot. We denote this probability as p, and call it the offered load.
2. Each cell is equally likely to be destined for any one output.

Input-buffered switches

The Maximum throughput achievable Using input queuing with FIFO Buffers

N	Throughput
1	1.0000
2	0.7500
3	0.6825
4	0.6553
5	0.6399
6	0.6302
7	0.6234
8	0.6184
∞	0.5858

Input-buffered switches

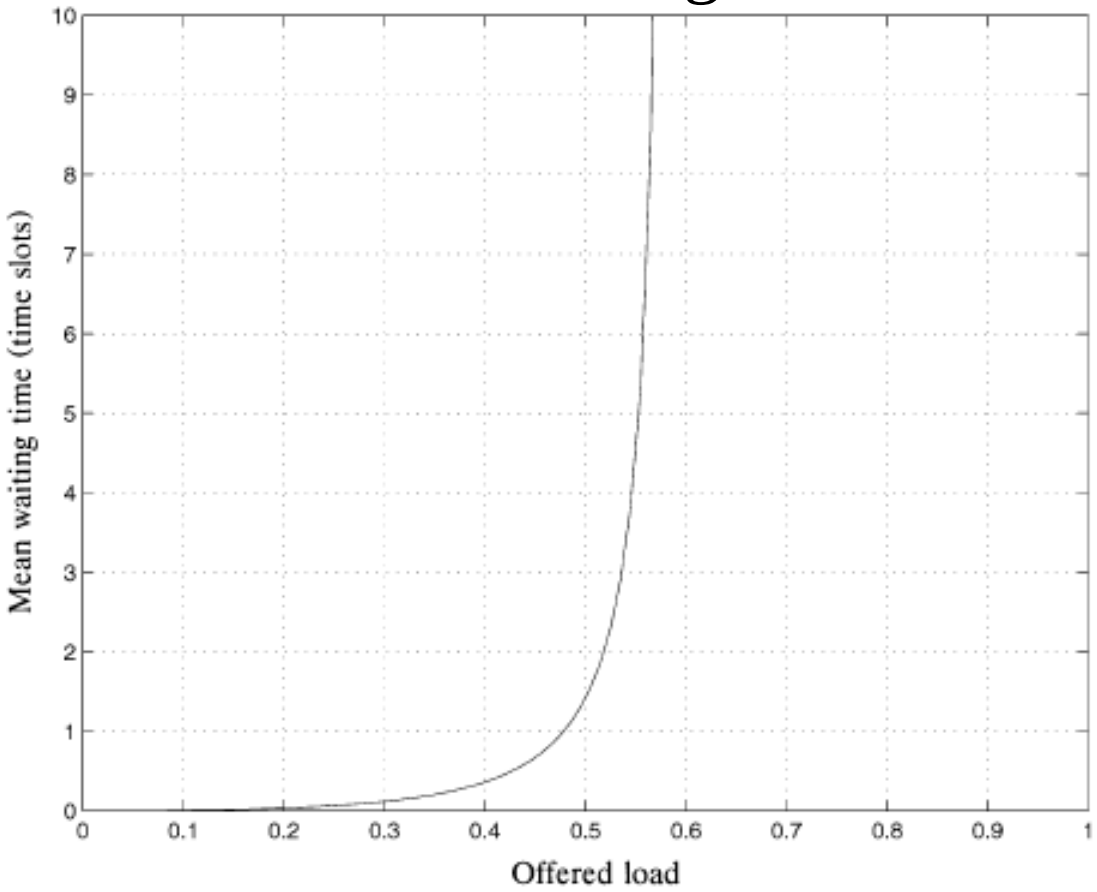
Using standard results for a discrete-time $Geom/G/1$ queue, we obtain the mean cell waiting time for input queuing with FIFO buffers,

$$\overline{W} = \frac{p\overline{S(S-1)}}{2\left(1-p\overline{S}\right)} + \overline{S} - 1, \qquad (2.8)$$

where S is the random variable of the service time obtained from the $M/D/1$ model. This waiting time shown in Figure 2.16 for $N \to \infty$.

Input-buffered switches

The main waiting time for input queuing with FIFO buffers the limiting case for N = ∞



Output-buffered switches

With output queuing, cells are only buffered at outputs, at each of which a separate FIFO is maintained. Consider a particular (i.e., tagged) output queue. Define the random variable A as the number of cell arrivals destined for the tagged output in a given time slot. Based on the same assumptions as in Section 2.3.1 on the arrivals, we have

$$a_k \triangleq \Pr[A = k] = \binom{N}{k}\left(\frac{p}{N}\right)^k\left(1 - \frac{p}{N}\right)^{N-k}, \qquad k = 0, 1, 2, \ldots N \quad (2.9)$$

When $N \to \infty$, (2.9) becomes

$$a_k \triangleq \Pr[A = k] = \frac{p^k e^{-p}}{k!}, \qquad k = 0, 1, 2, \ldots . \quad (2.10)$$

Output-buffered switches

Denote by Q_m the number of cells in the tagged queue at the end of the mth time slot, and by A_m the number of cell arrivals during the mth time slot. We have

$$Q_m = \min\{\max(0, Q_{m-1} + A_m - 1), b\}. \tag{2.11}$$

For finite N and finite b, this can be modeled as a finite-state, discrete-time Markov chain with state transition probabilities $P_{ij} \triangleq \Pr[Q_m = j | Q_{m-1} = i]$ as follows:

$$P_{ij} = \begin{cases} a_0 + a_1, & i = 0, \quad j = 0, \\ a_0, & 1 \le i \le b, \quad j = i - 1, \\ a_{j-i+1}, & 1 \le j \le b - 1, \quad 0 \le i \le j, \\ \sum_{m=j-i+1}^{N} a_m, & j = b, \quad 0 \le i \le j, \\ 0 & \text{otherwise}, \end{cases} \tag{2.12}$$

Output-buffered switches

where a_k is given by (2.9) and (2.10) for a finite N and $N \to \infty$, respectively. The steady-state queue size can be obtained recursively from the following Markov chain balance equations:

$$q_1 \triangleq \Pr[Q = 1] = \frac{1 - a_0 - a_1}{a_0} \cdot q_0$$

$$q_n \triangleq \Pr[Q = n] = \frac{1 - a_1}{a_0} \cdot q_{n-1} - \sum_{k=2}^{n} \frac{a_k}{a_0} \cdot q_{n-k}, \qquad 2 \le n \le b,$$

where

$$q_0 \triangleq \Pr[Q = 0] = \frac{1}{1 + \sum_{n=1}^{b} q_n / q_0} \cdot$$

Output-buffered switches

No cell will be transmitted on the tagged output line during the mth time slot if, and only if, $Q_{m-1} = 0$ and $A_m = 0$. Therefore, the switch throughput ρ_0 is represented as

$$\rho_0 = 1 - q_0 a_0 .$$

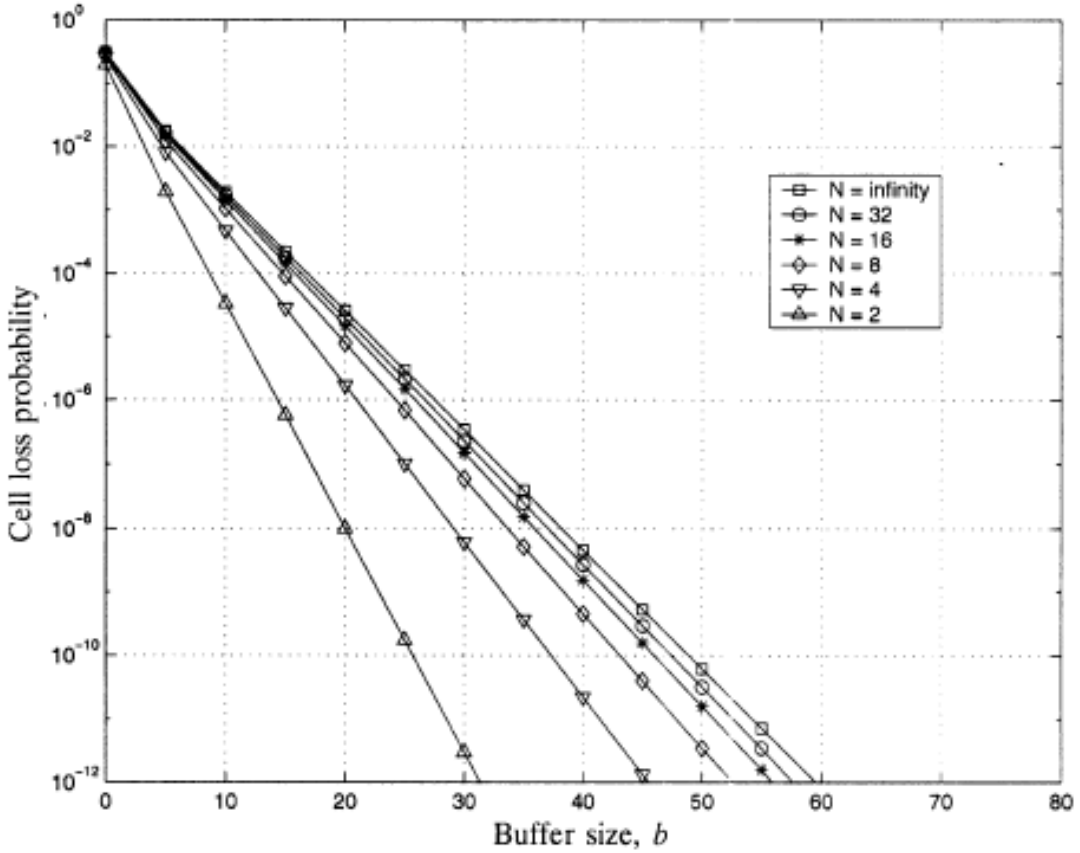
A cell will be lost if, when emerging from the switch fabric, it finds the output buffer already containing b cells. The cell loss probability can be calculated as follows:

$$\Pr[\text{cell loss}] = 1 - \frac{\rho_0}{p},$$

where p is the offered load.

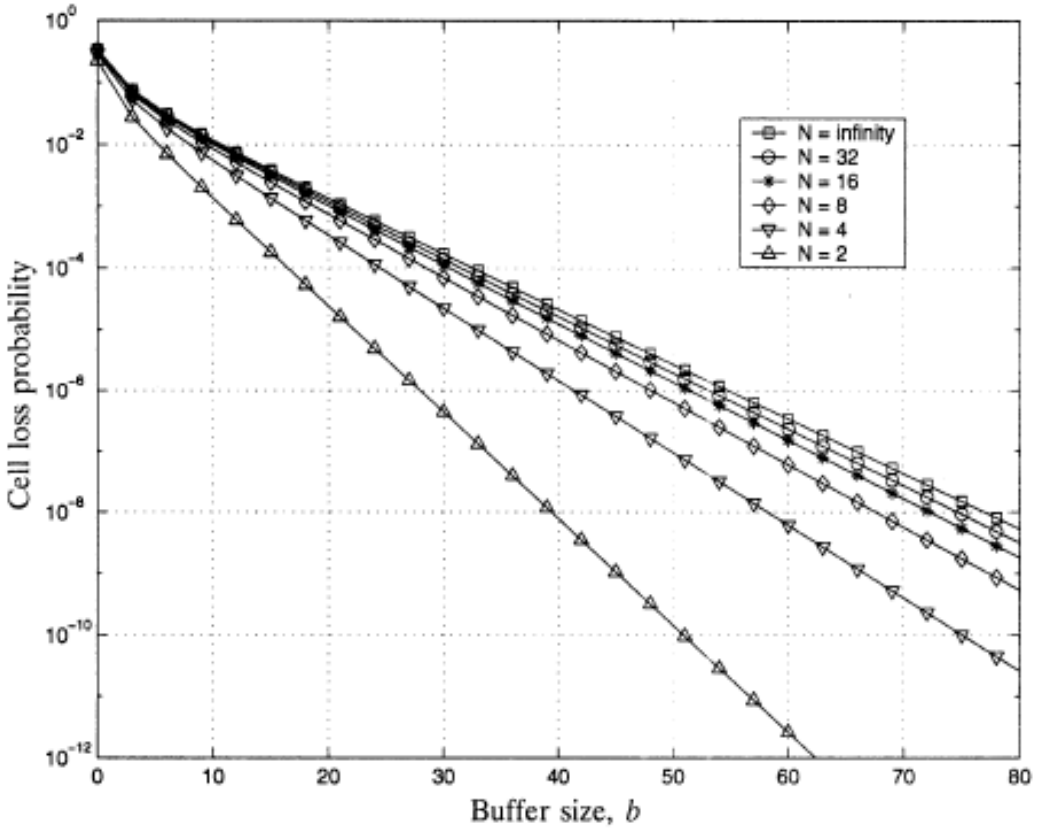
Output-buffered switches

The cell loss probability for output queuing as a function of the buffer size b and the switch size N, for offered loads p=0.8



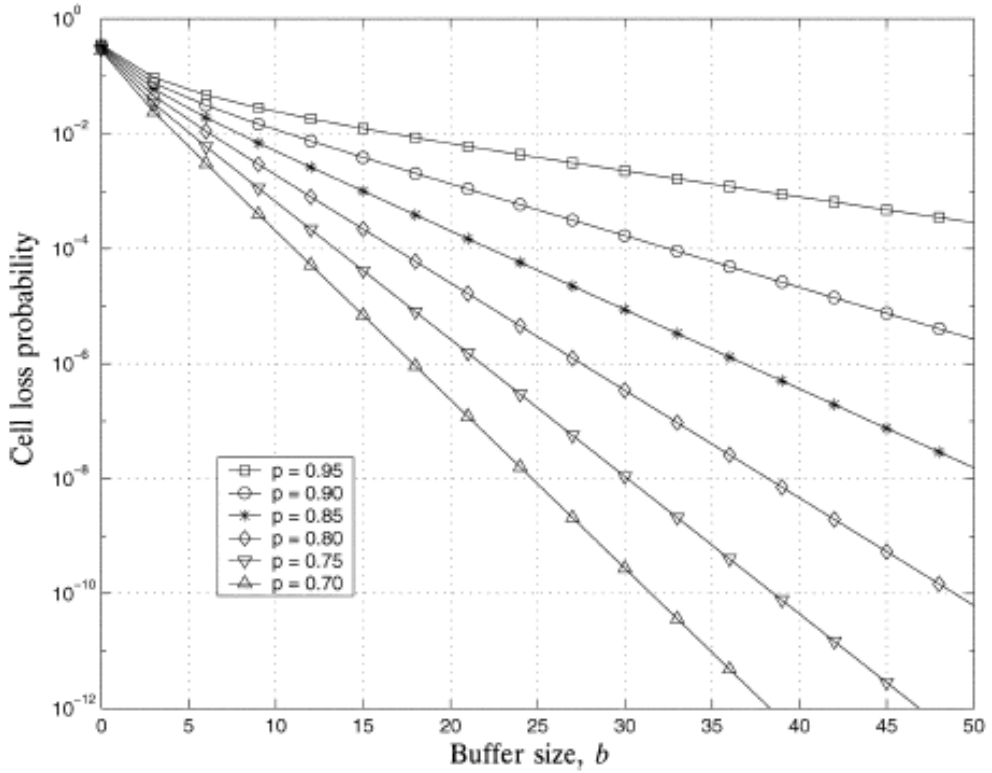
Output-buffered switches

The cell loss probability for output queuing as a function of the buffer size b and the switch size N, for offered loads p=0.9



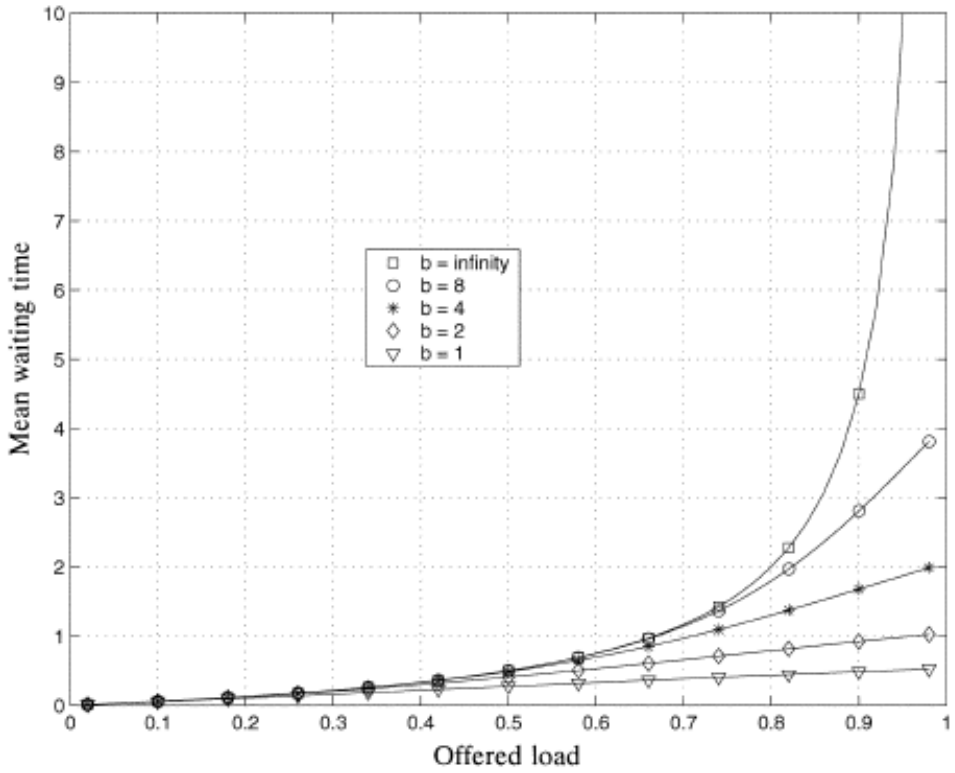
Output-buffered switches

The cell loss probability for output queuing as a function of the buffer size b and offered loads varying from p=0.70 to p=0.95, for the limiting case of $N \rightarrow \infty$



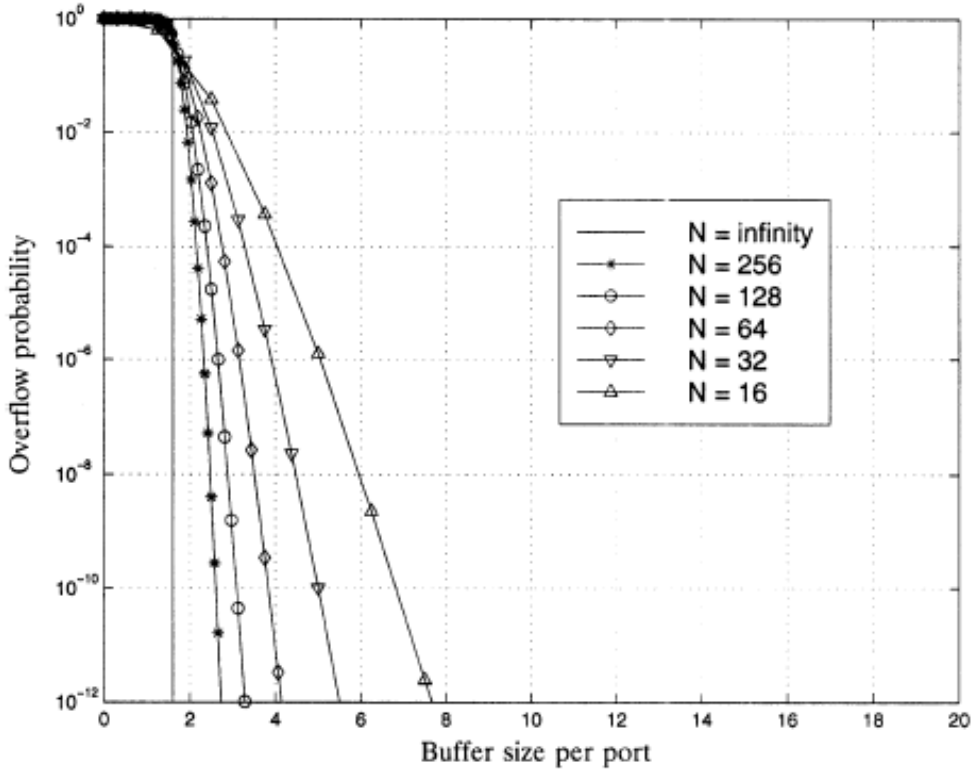
Output-buffered switches

The mean waiting time for output queuing as a function of the offered load p, for $N \rightarrow \infty$ and output FIFO sizes varying from b=1 to ∞



Completely Shared-buffer switches

The cell loss probability for completely shared buffering as a function of the buffer size per output, b, and the switch size N, for offered load p=0.8



Completely Shared-buffer switches

The cell loss probability for completely shared buffering as a function of the buffer size per output, b, and the switch size N, for offered load p=0.9