Basics of Packet Switching

HIGH PERFORMANCE SWITCHES AND ROUTERS

Wiley

H. JONATHAN CHAO and BIN LIU

Instructor: Mansour Rousta Zadeh

Outlines

- Introduction
- ATM networks
- ATM switch systems
- IP router systems
- switch design criteria and performance requirements
- Basic switching concept
- ATM switching architecture

Introduction

- Internet: Scalable and Distributed System
  - ->Fast Growth in Size and Traffic
  - ->Lead to Great Success
  - ->Exponentially Traffic Volume
  - ->Need to New Infrastructure
- This Problem is solved by Optical Transmission (DWDM, OXC)
- But another Problem risen:

  Slow Growth in Switch/Router Technologies

Switching

- Need to high speed and large capacity switching systems to aggregate lower bit rates
- Need to support QoS specially for Real-time services
- One solution is ATM Network
- Another solution is MPLS
- We will not cover MPLS in this course

Sample High Speed Routers

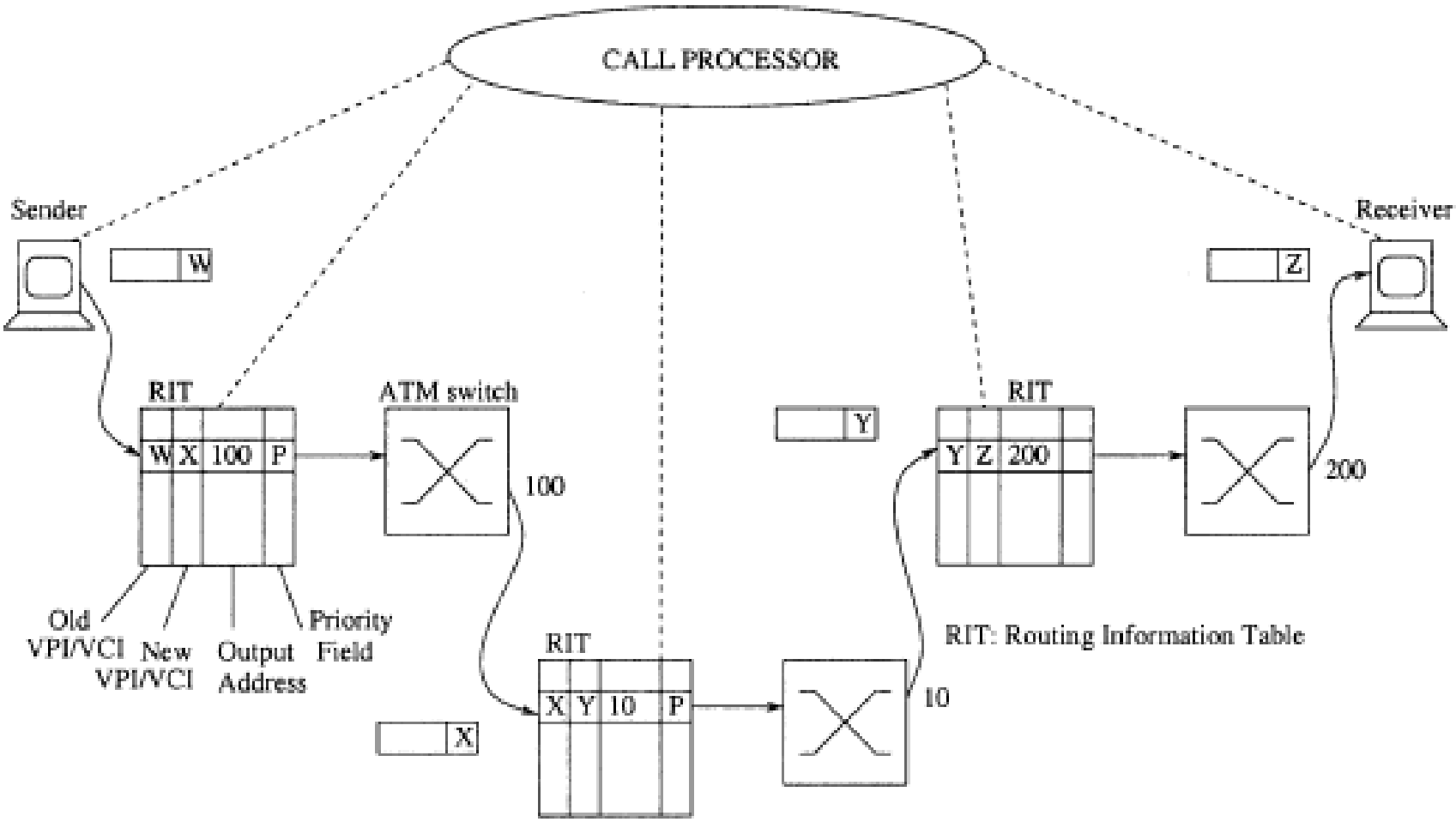
Cisco 12416 Router
up to 160 Gb/s throughput
up to 10 Gb/s ports

Juniper Networks T640 Router
up to 160 Gb/s throughput
up to 10 Gb/s ports

Sample High Capacity Router

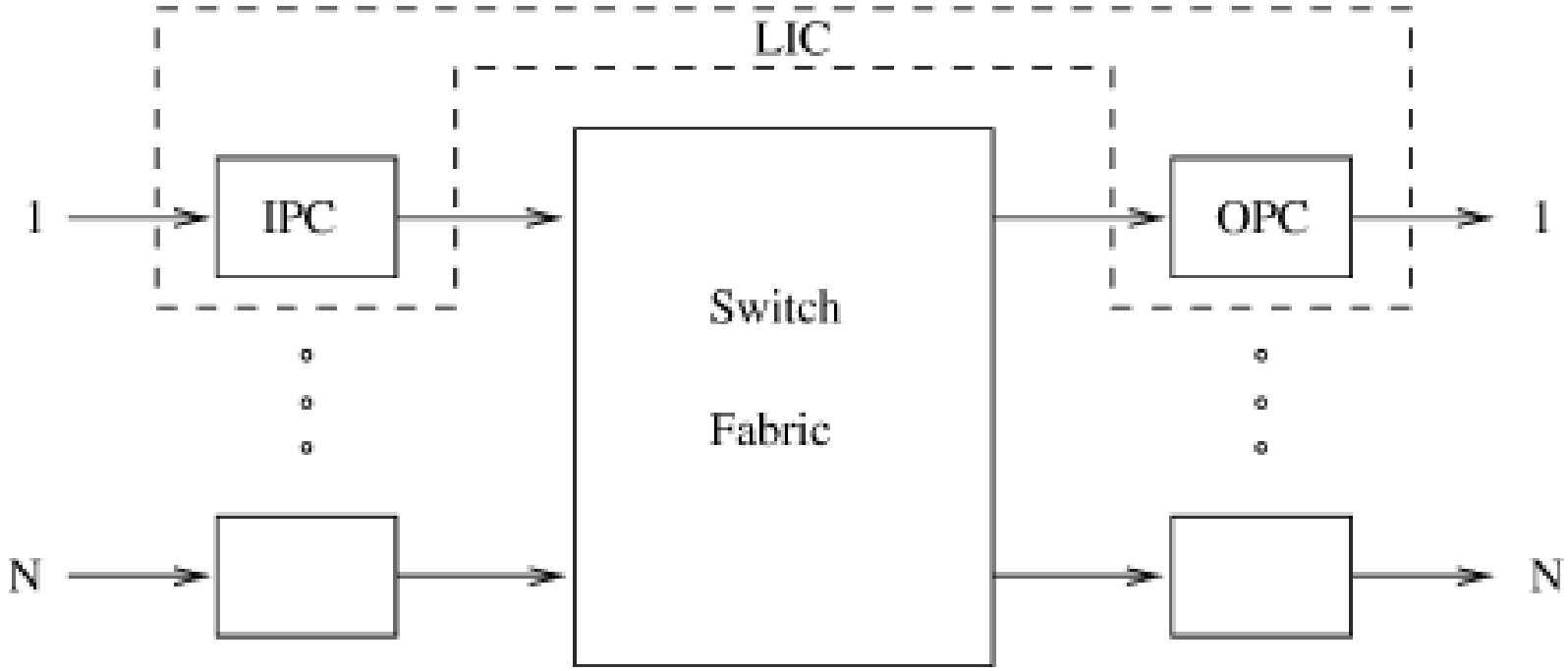
- Cisco CRS-1
  - up to 46 Tb/s throughput
- Two rack types
- Line card rack
  - 640 Gb/s throughput
  - up to 16 line cards
    - up to 40 Gb/s each
  - up to 72 racks
- Switch rack
  - central switch stage
  - up to 8 racks
- In-service scaling.

Basics of ATM Network and Switching

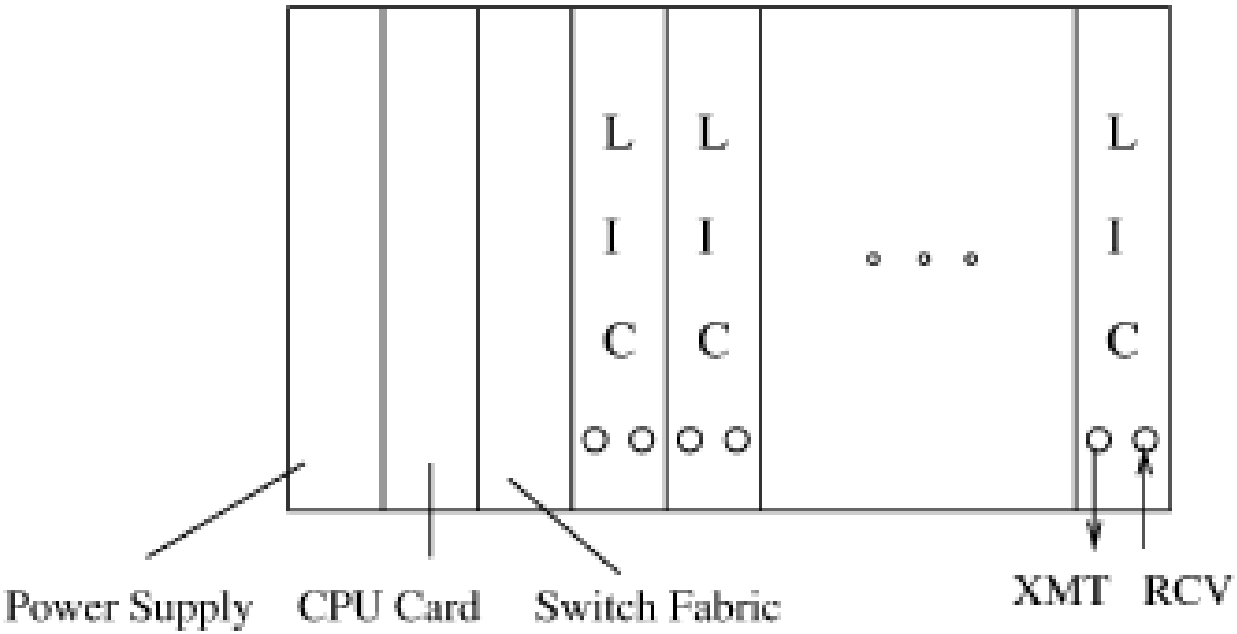


ATM Switch Structure

- Each IPC and OPC pair is located in a single Line Card named LIC.

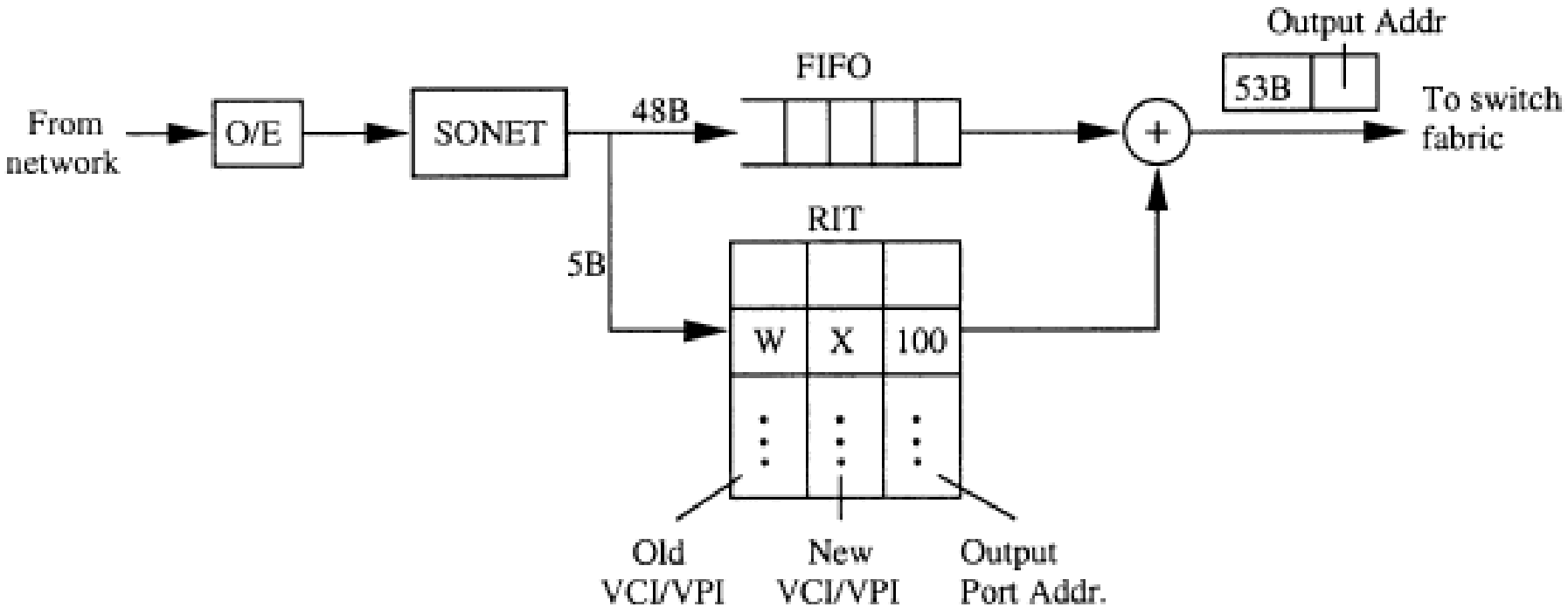


ATM Switch



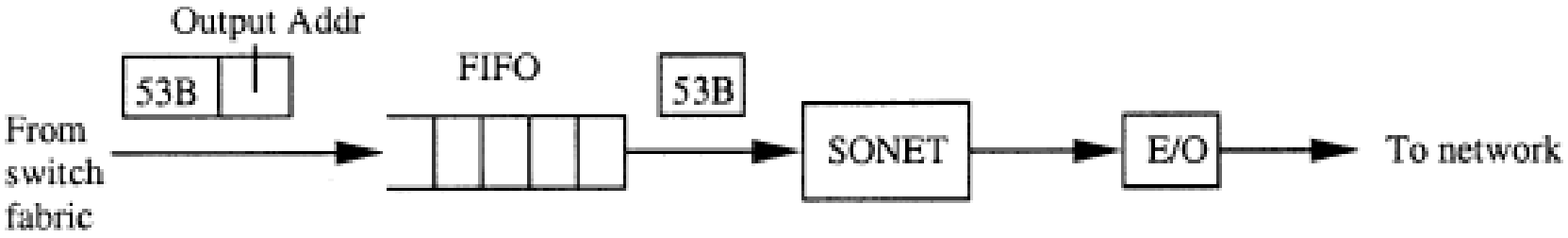
IPC Block Diagram

- This is the structure of IPC block diagram from the network to the fabric:



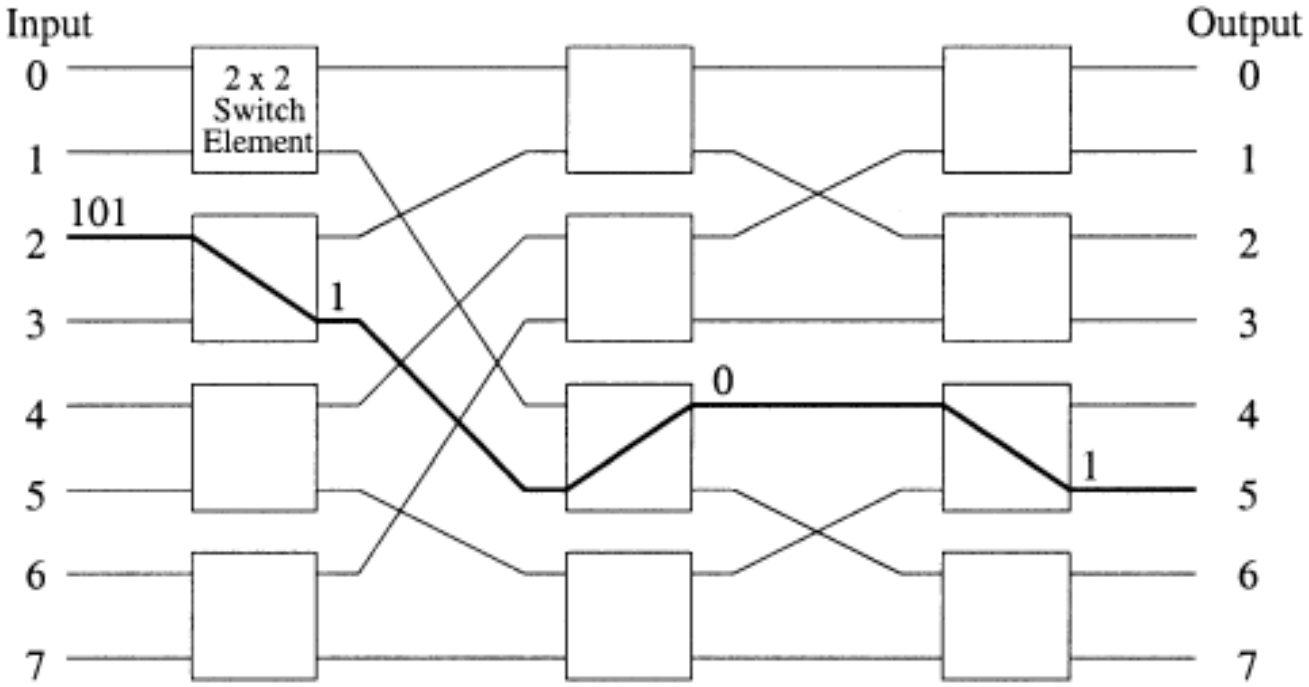
OPC Block Diagram

- This is the structure of OPC block diagram from the switch fabric to the network line:



Routing in ATM Fabric

- Self Routing



- Label Routing

Function of IP Routers

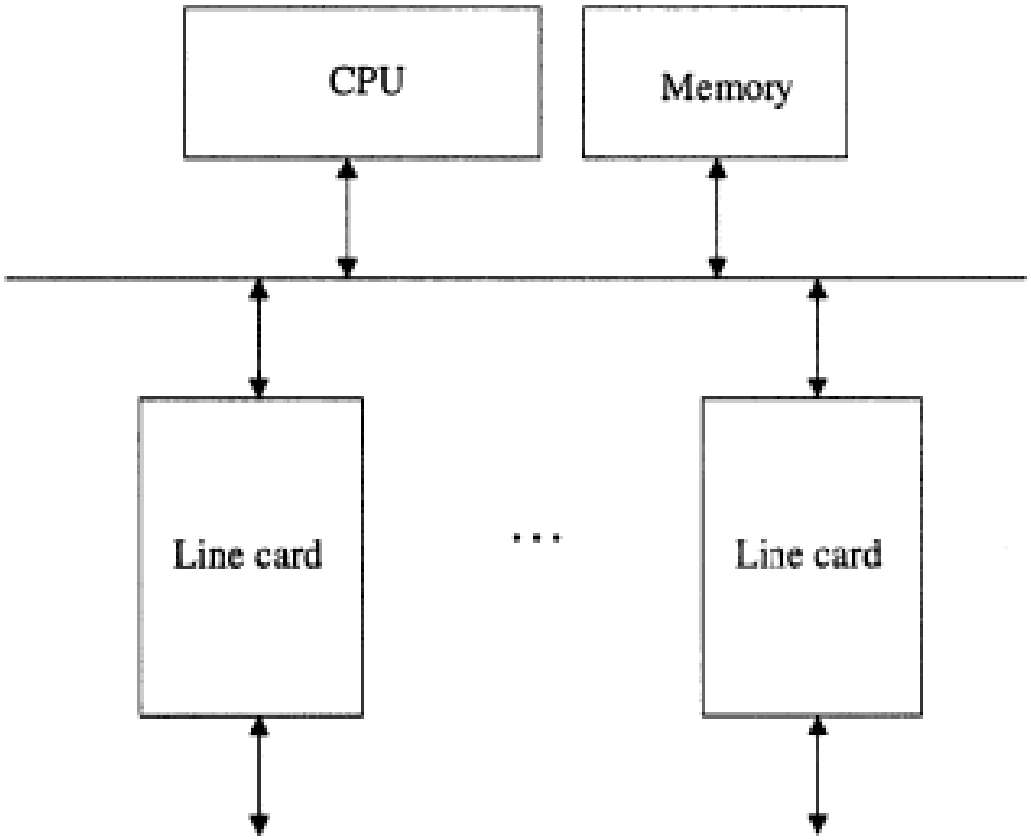
IP Router Functions are classified as:

- **Datapath functions**
- **Control Functions**

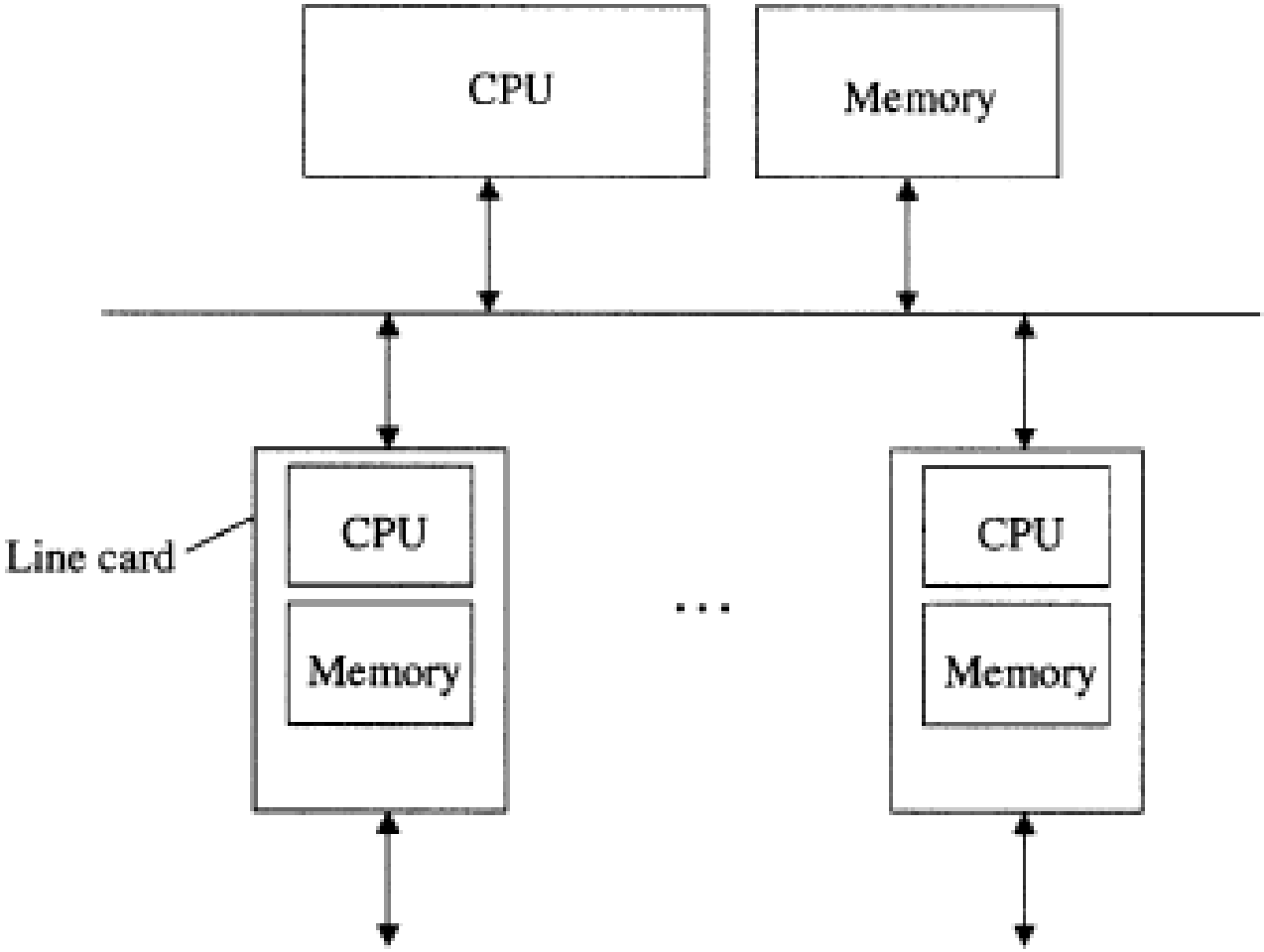
IP Router Systems

- Low-End Routers
- Middle-Size Routers
- High-End Routers

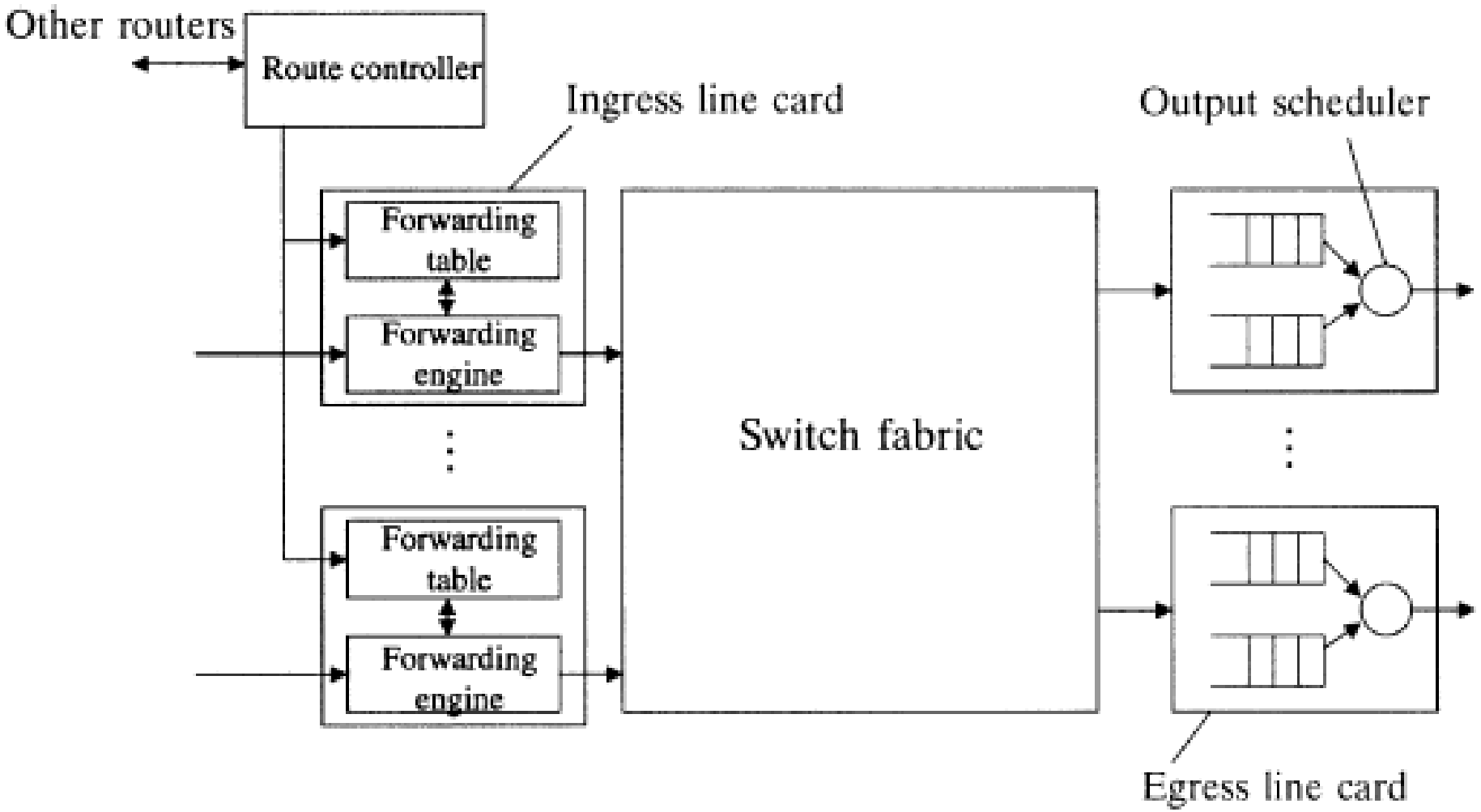
Low-End Routers



Middle-Size Routers



High-End Routers



Design Criteria

- Several Design Criteria
  - Small Delay for multimedia application
  - Small CLP (Cell Loss Probability) in near 100% throughput for loss sensitive Apps
  - High speed and capacity for high performance routers
  - Distributed Control and self routing for scalability

Performance Criteria

Bellcore Performance Criteria

- QoS Class 1
  - Equal to ITU-T **Class A**
  - For cell loss sensitive applications
- QoS Class 3
  - Equal to ITU-T **Class C**
  - For Low latency connection oriented application
- QoS Class 4
  - Equal to ITU-T **Class D**
  - For Low latency connection-less application

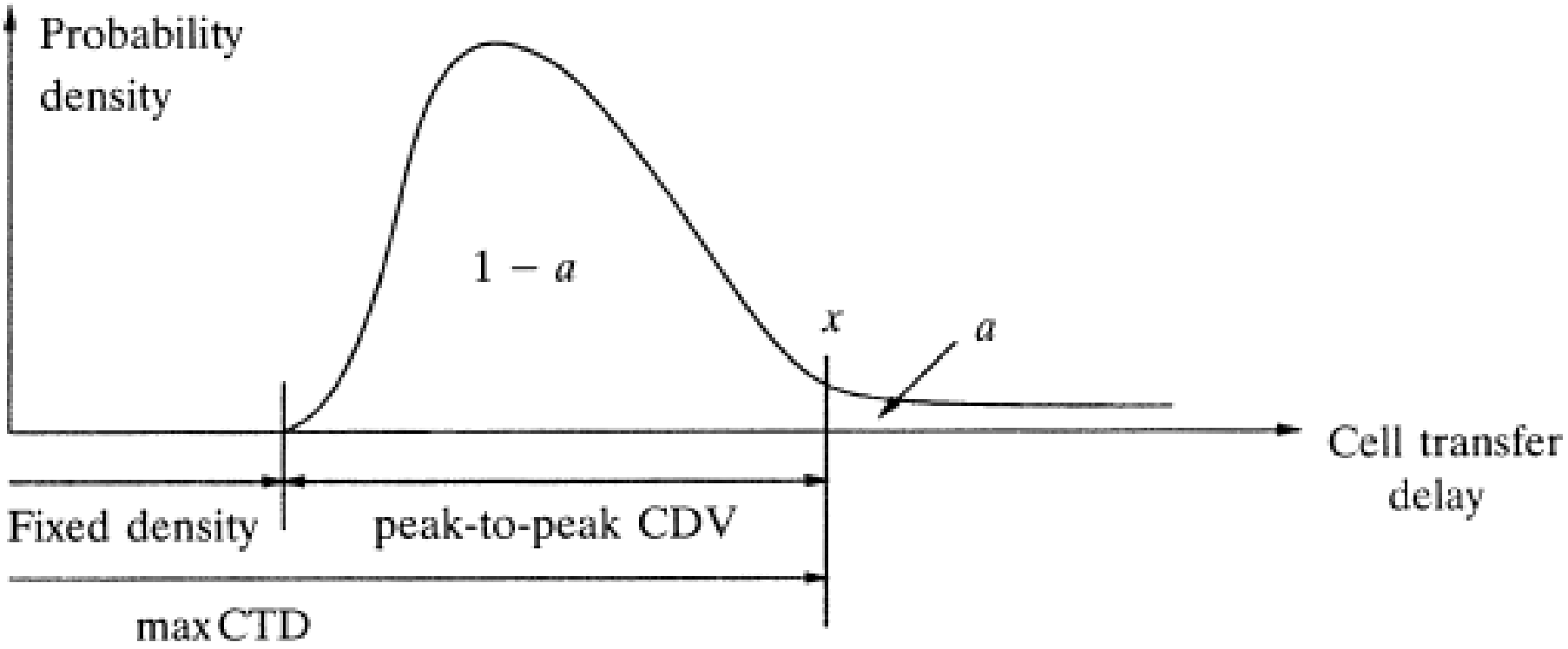
Performance Objective across BSS

Performance Parameter	CLP	QoS 1	QoS 3	QoS 4
Cell loss ratio	0	$< 10^{-10}$	$< 10^{-7}$	$< 10^{-7}$
Cell loss ratio	1	N/S[a]	N/S	N/S
Cell transfer delay (99th percentile)[b]	1/0	150 μs	150 μs	150 μs
Cell delay variation (10^{-10} quantile)	1/0	250 μs	N/S	N/S
Cell delay variation (10^{-7} quantile)	1/0	N/S	250 μs	250 μs

[a] N/S not specified.

[b] Includes nonqueuing related delays, excluding propagation. Does not include delays due to processing above ATM layer.

Distribution of cell transfer delay

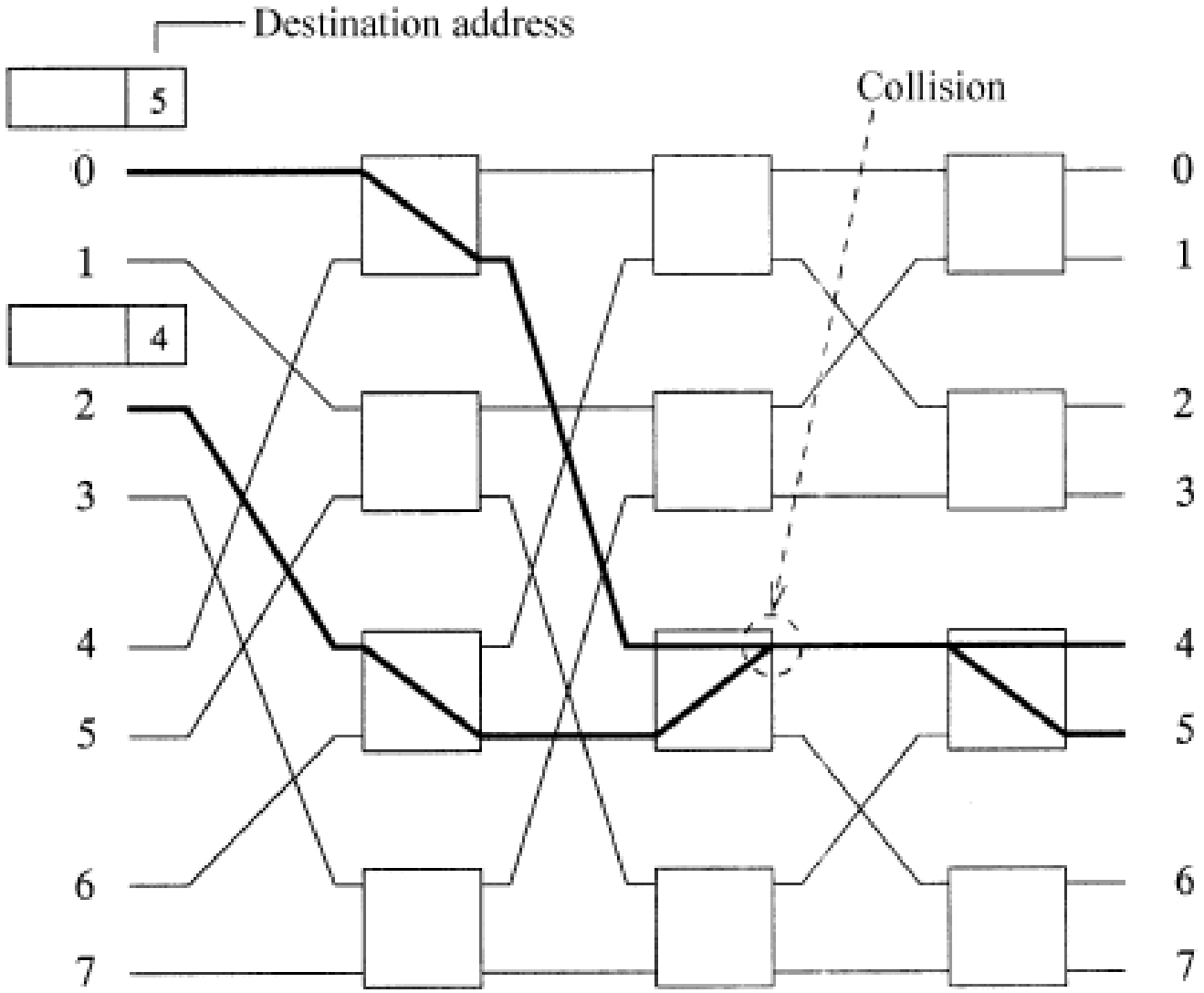


Basics of Packet Switching

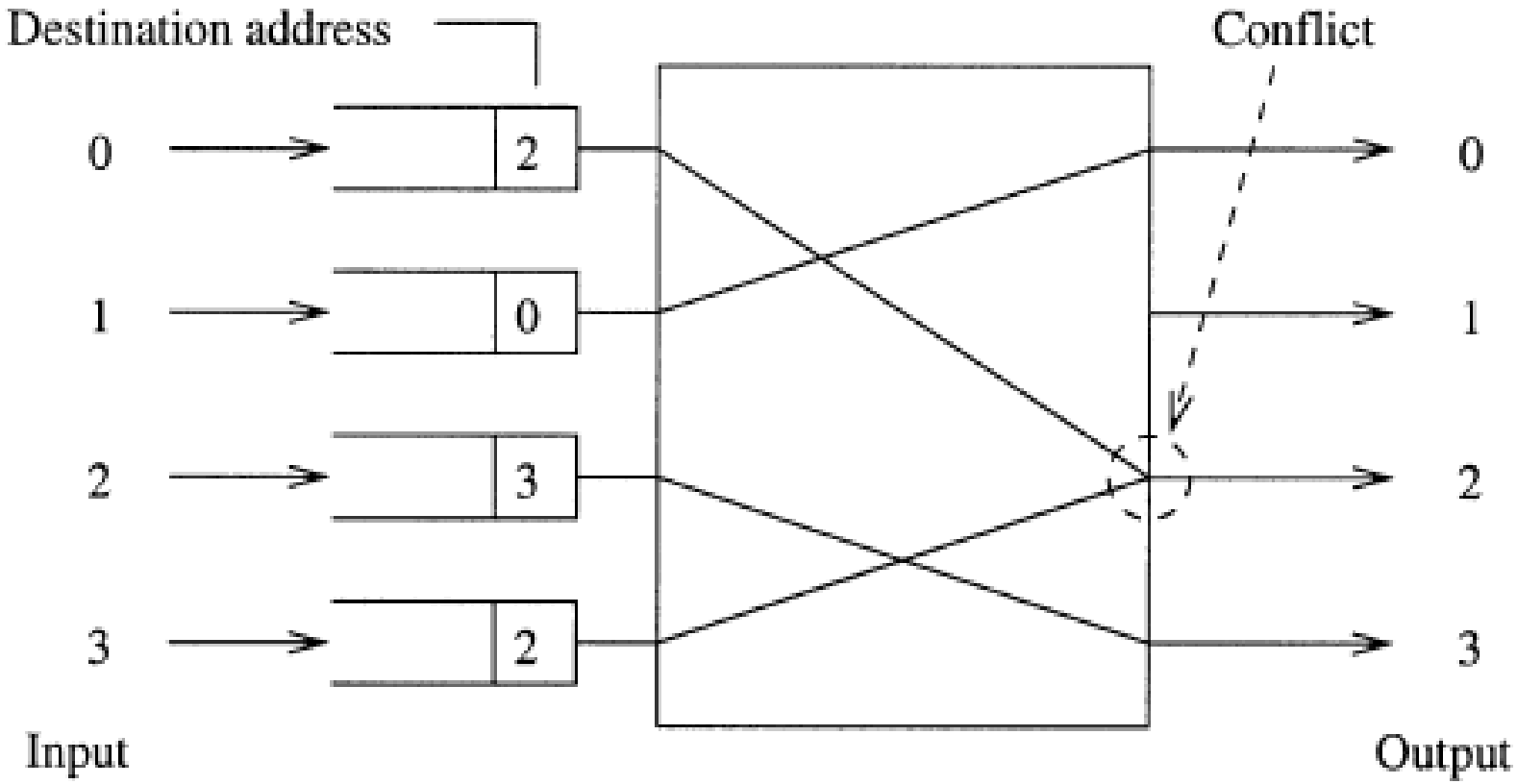
Switching Concepts:

- Internal Link Blocking
- Output Port Contention
- Head-of-Line Blocking
- Multicasting
- Call Splitting

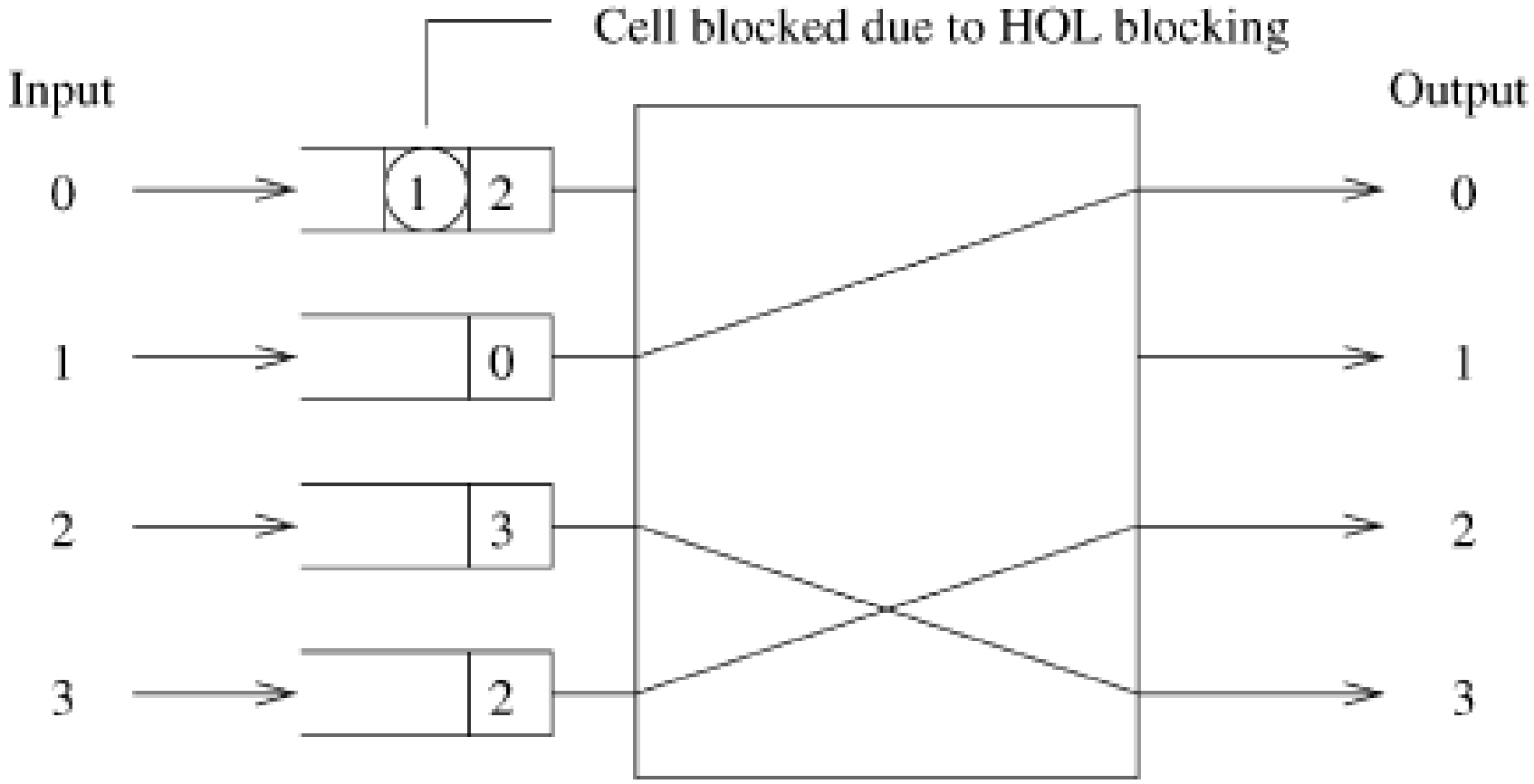
Internal Link Blocking



Output Port Contention



Head of Line (HOL) Blocking



Call Splitting

Input \ Output	1	2	3	4	5
1	1	1	0	0	1
2	0	1	0	0	0
3	0	0	0	1	0
4	0	1	1	0	0
5	0	0	1	1	0

Transmission requets matrix

(1: copy request, 0: no request)

X	X	0	0	X
0	1	0	0	0
0	0	0	X	0
0	1	1	0	0
0	0	1	1	0

(a) One-shot

X	1	0	0	1
0	X	0	0	0
0	0	0	X	0
0	1	X	0	0
0	0	1	1	0

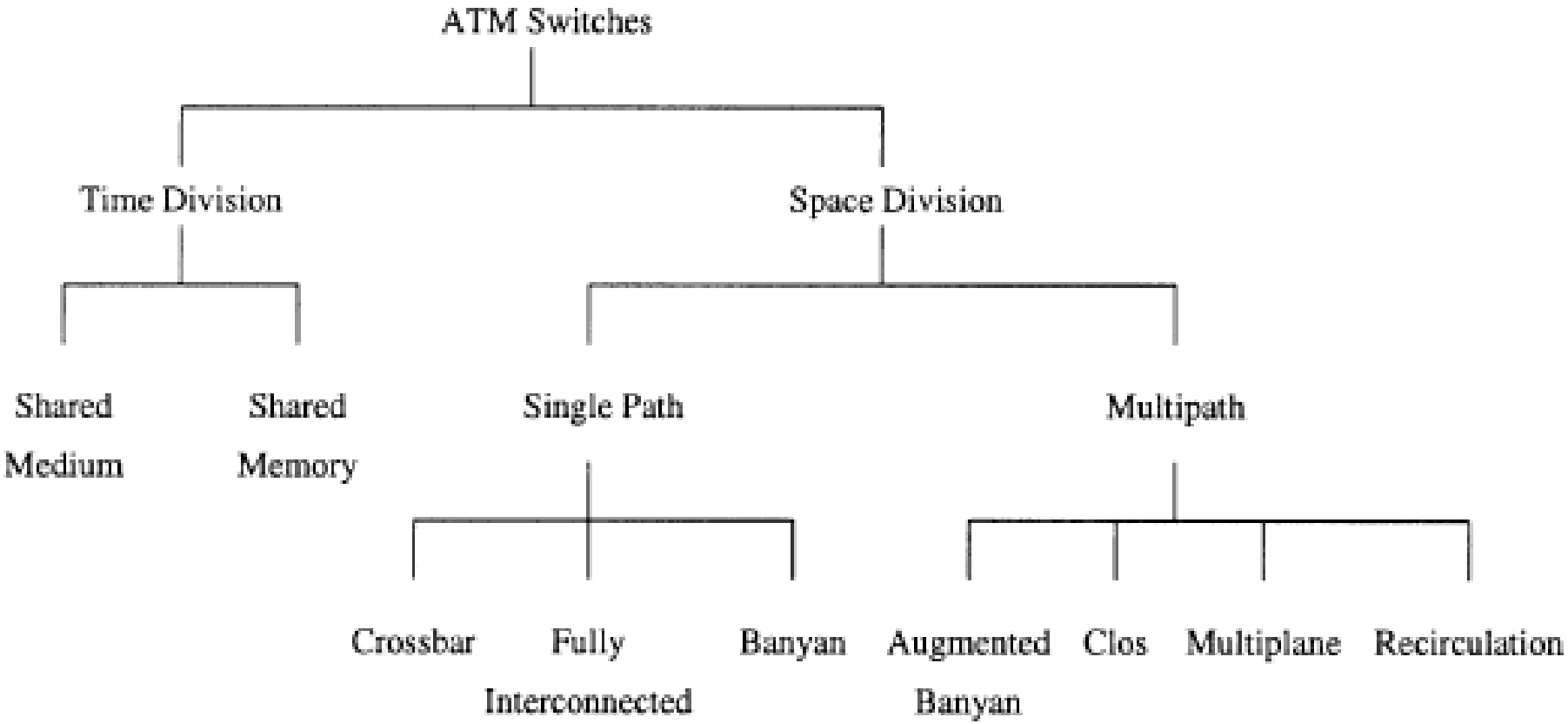
(b) Strict-sense call splitting

X	1	0	0	X
0	X	0	0	0
0	0	0	X	0
0	1	X	0	0
0	0	1	1	0

(c) Wide-sense call splitting

(X: accepted request, 1: rejected request)

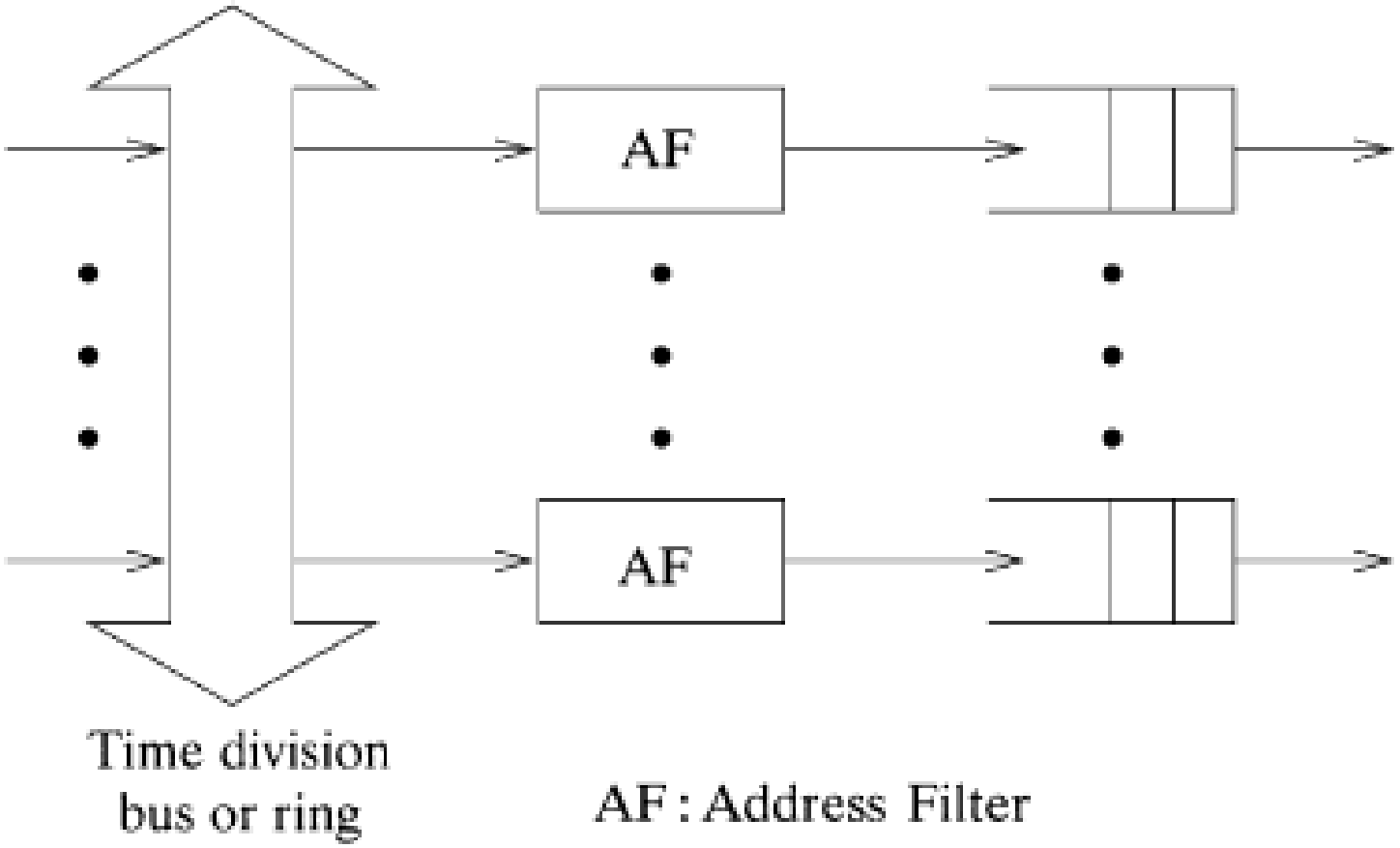
Classification of ATM switching architectures



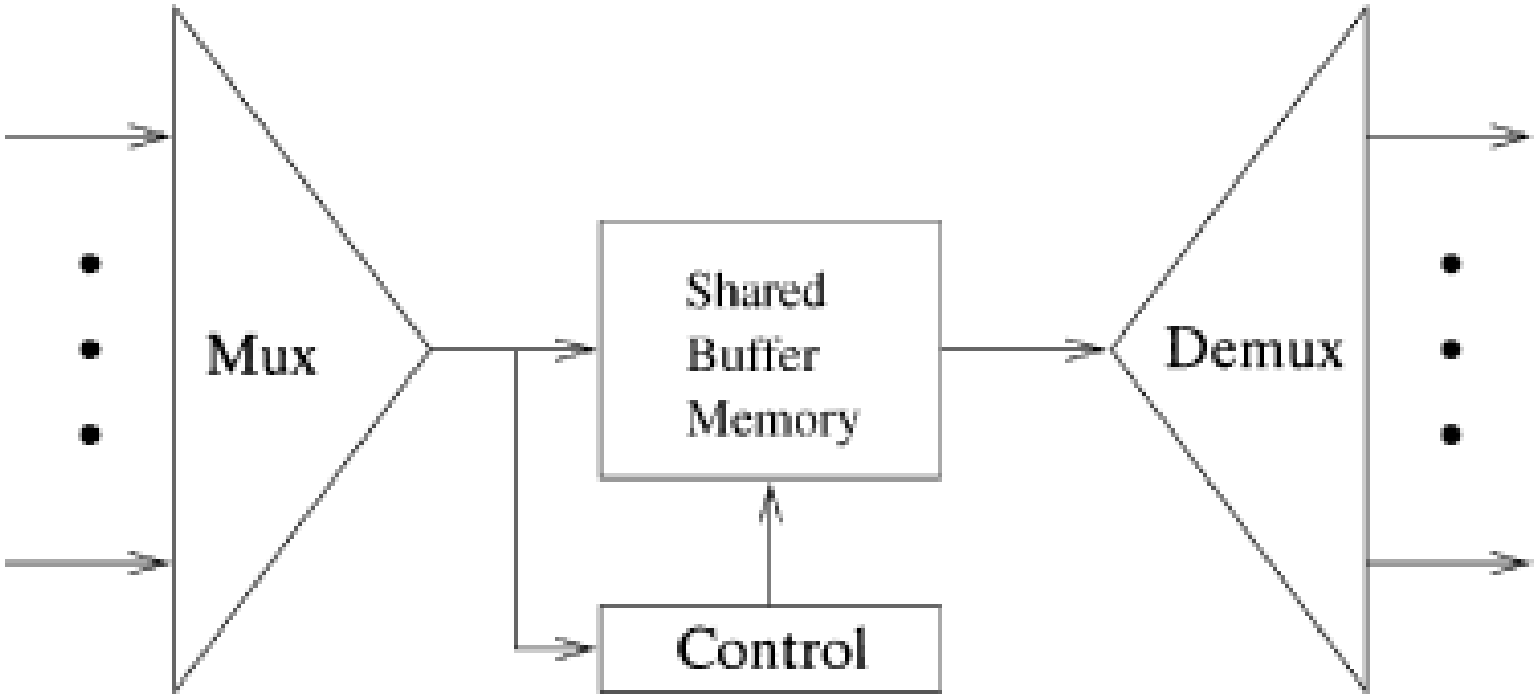
Time Division Switching

- Shared-Medium Switch
- Shared-Memory Switch

Shared-Medium Switch



Shared-Memory Switch



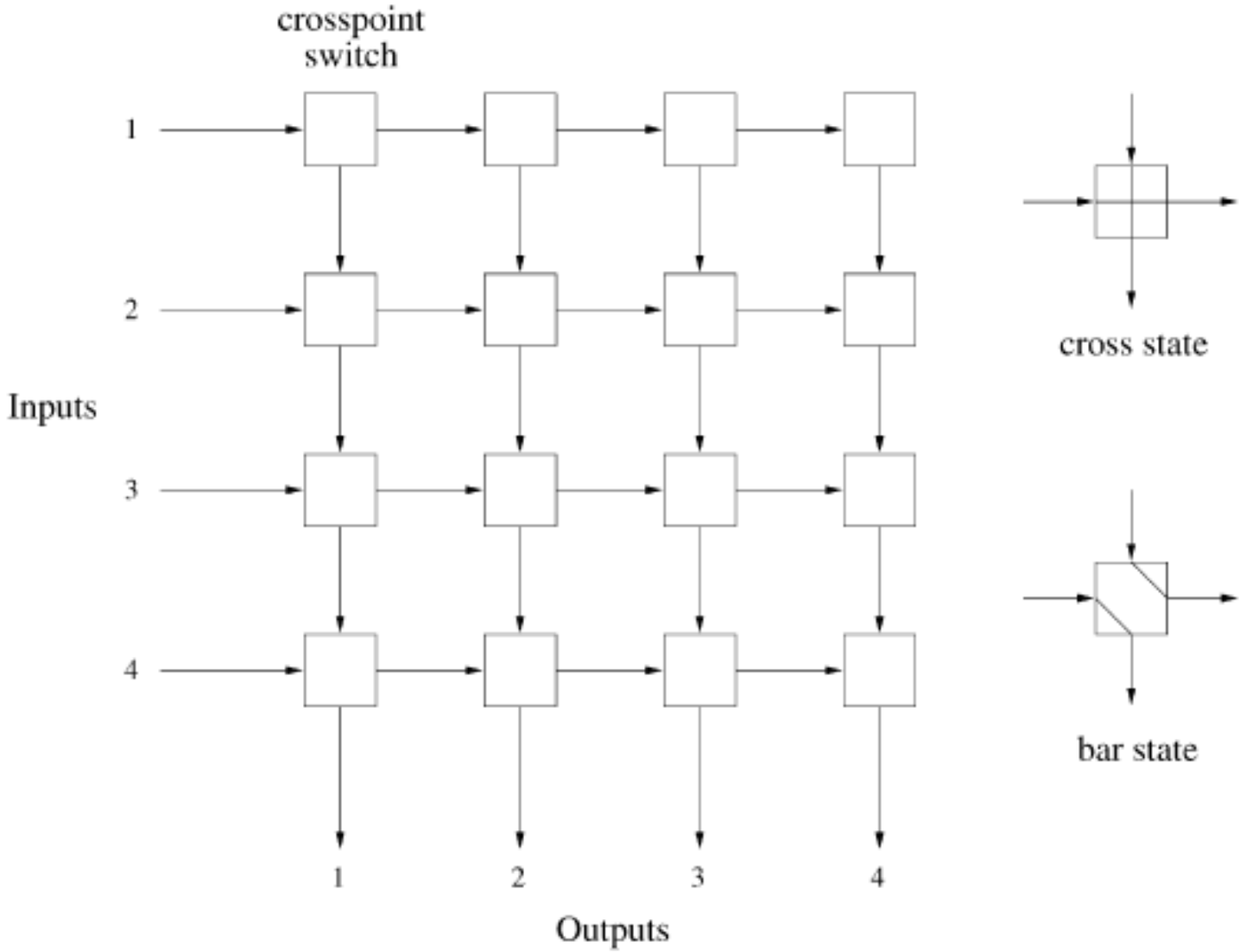
Mux : Multiplexer

Demux : Demultiplier

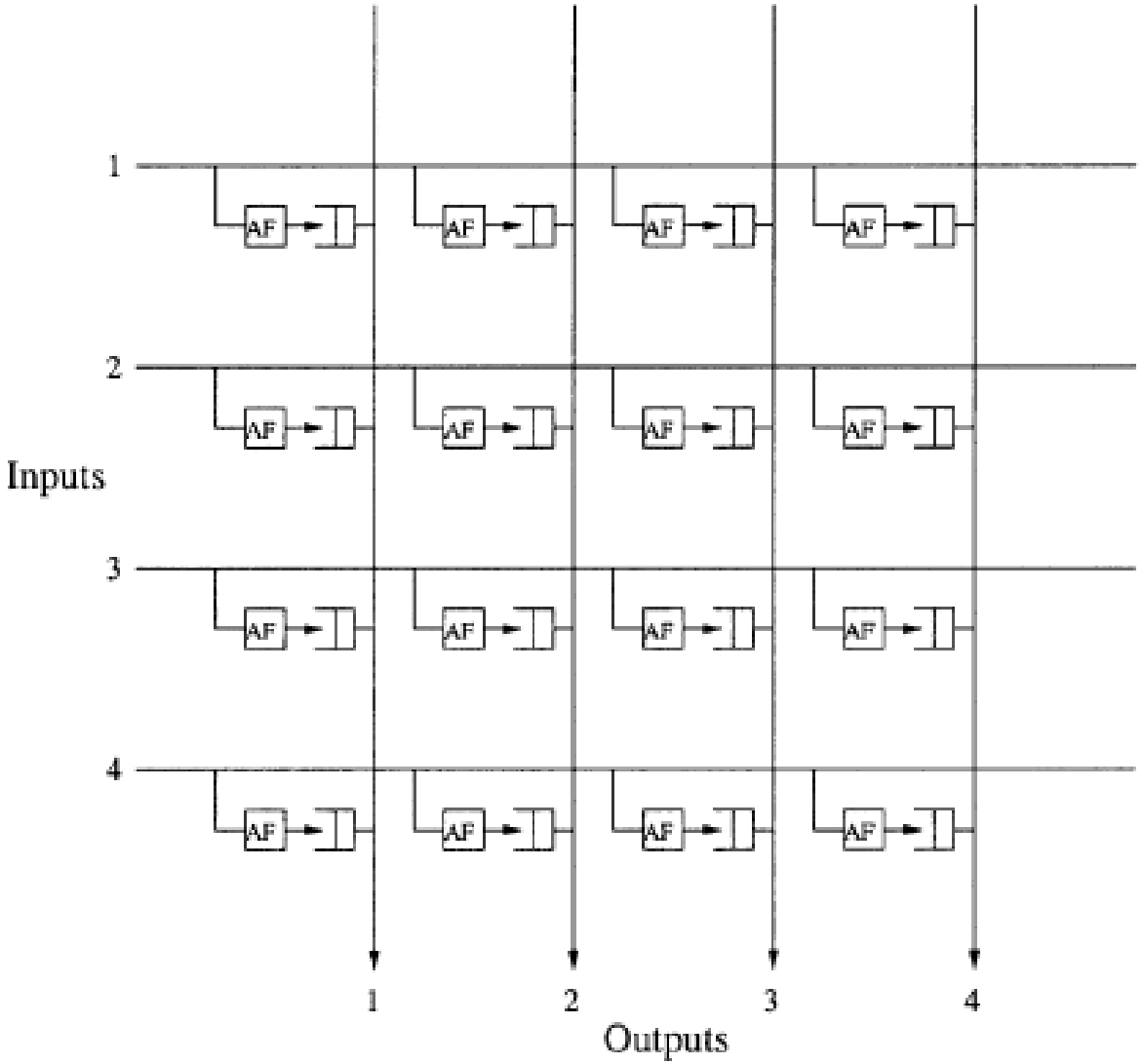
Space division Switching

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

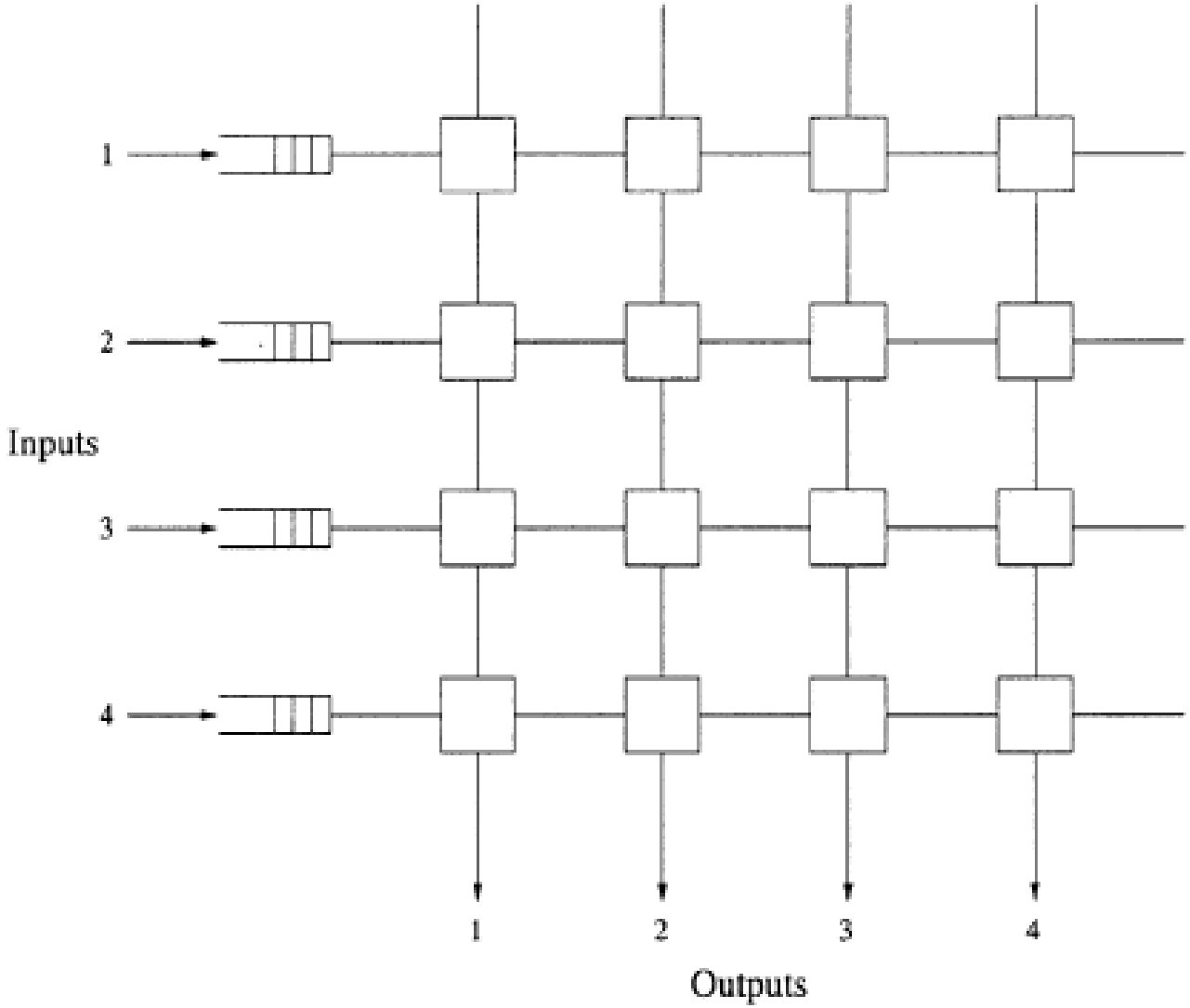
Space Division - Crossbar Switches (1)



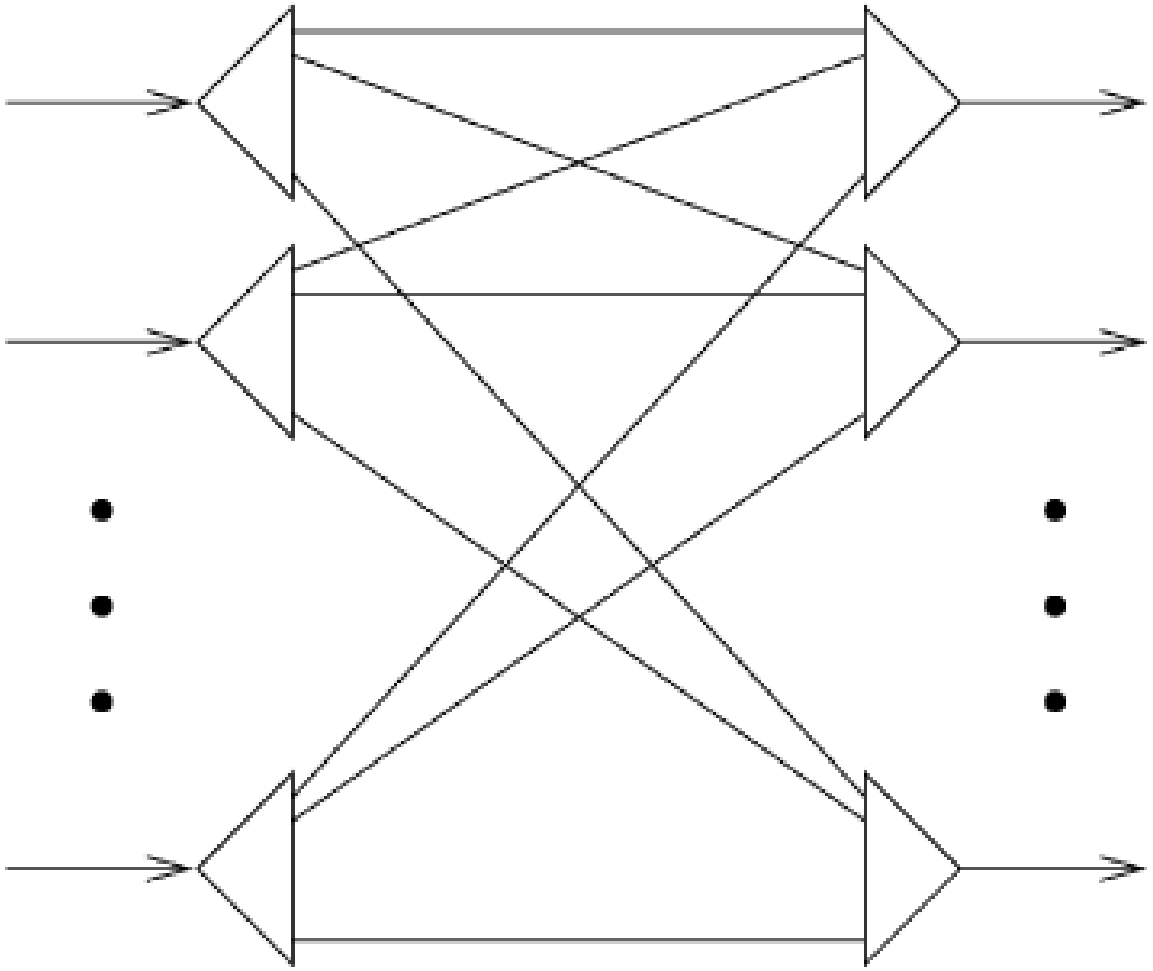
Space Division - Crossbar Switches (2)



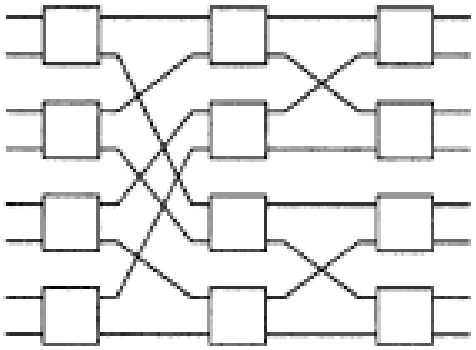
Space Division - Crossbar Switches (3)



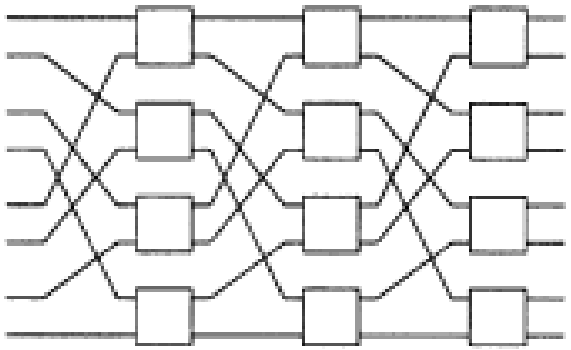
Space Division - Fully Interconnected Switches



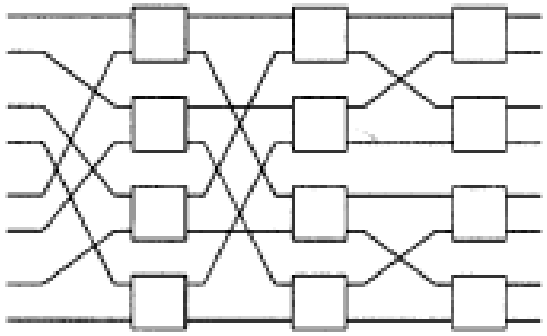
Space Division - Banyan-Based Switches



(a) Delta network



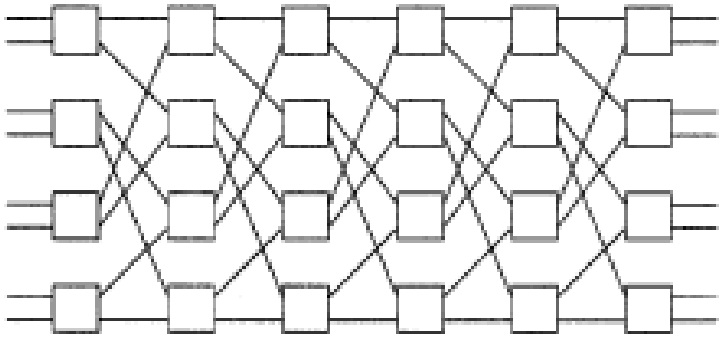
(b) Omega network



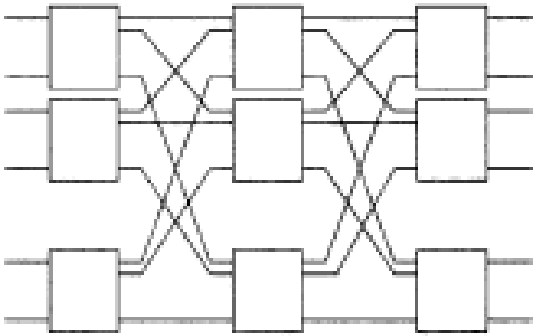
(c) Banyan network

: 2 × 2 Switch Element (SE)

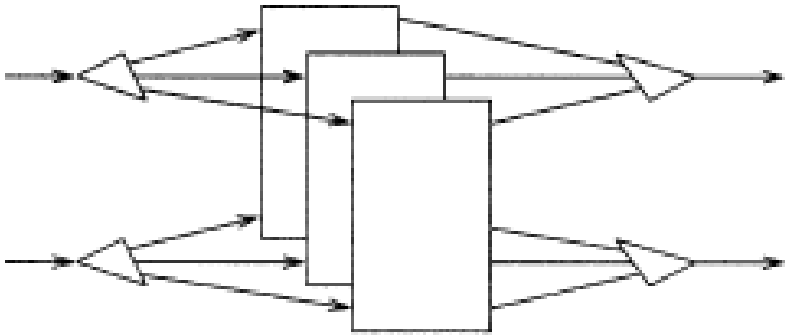
Space Division - Multiple-Path Switches



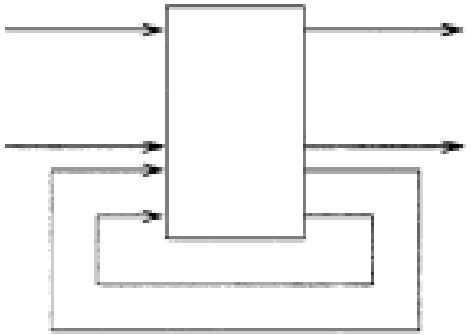
(a) Augmented Banyan



(b) 3-stage Clos

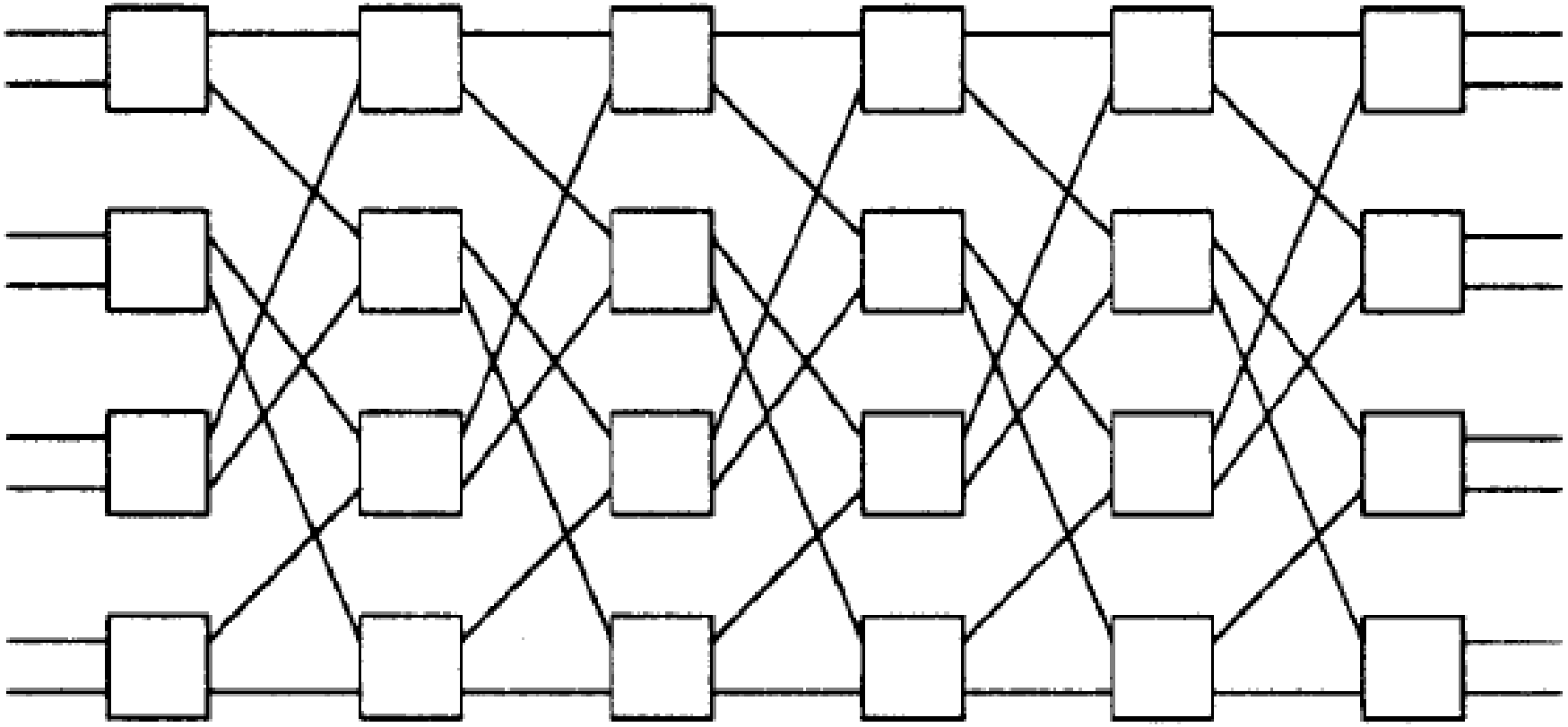


(c) Multiplane

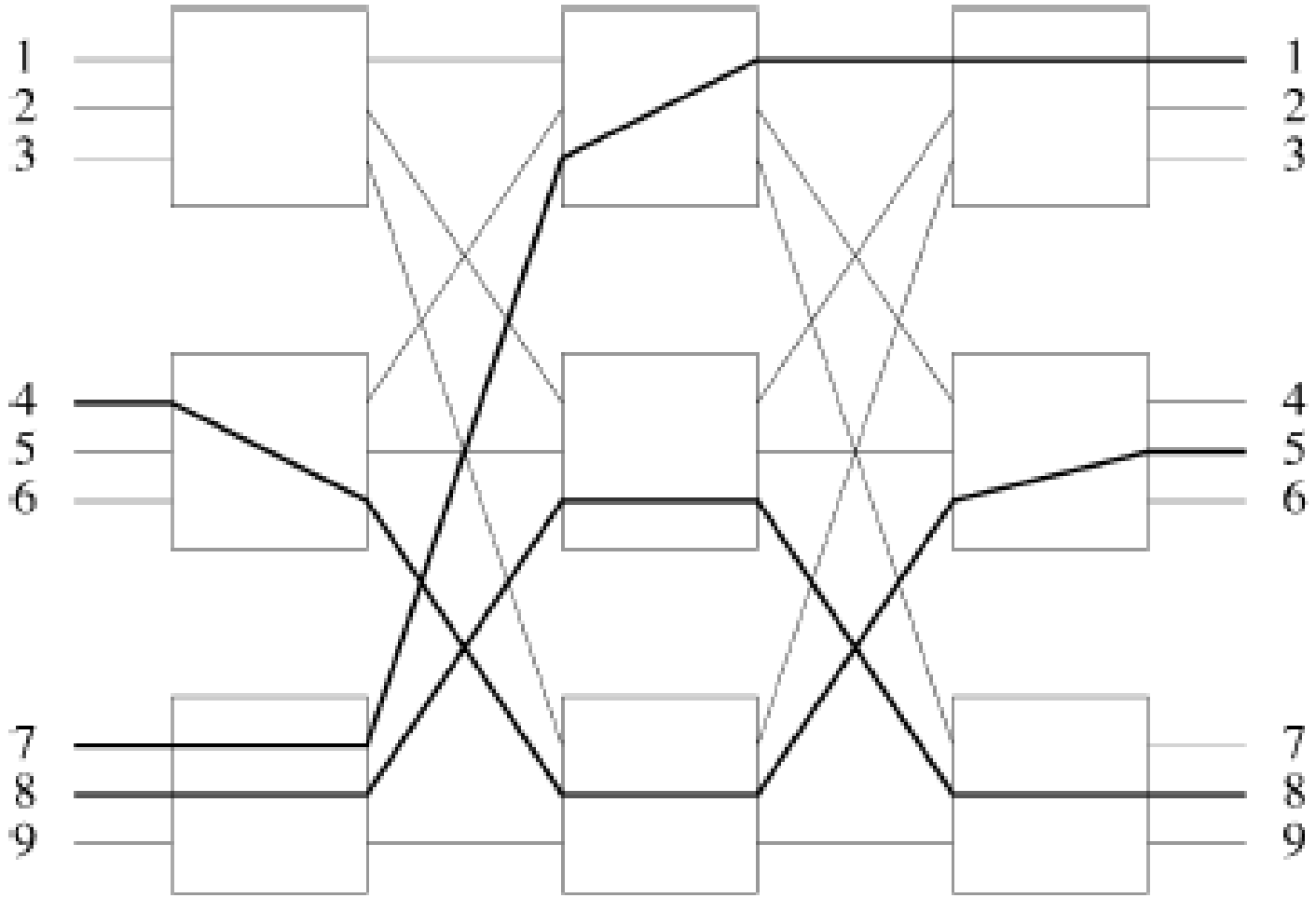


(d) Recirculation

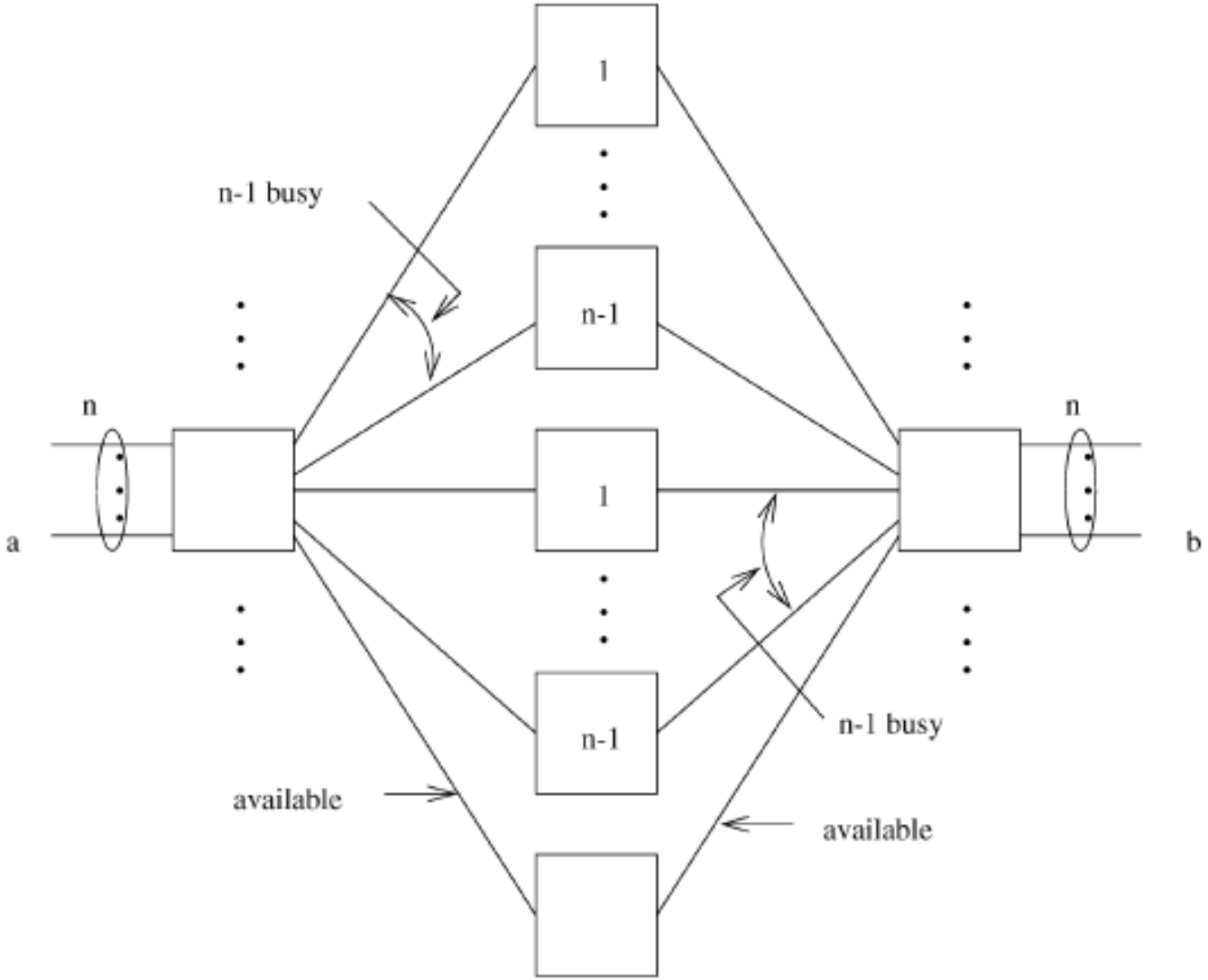
Space Division - Augmented Banyan Switches



Space Division - Three-Stage Clos Switches



Space Division - Clos Switches (1)



Space Division - Clos Switches (2)

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

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

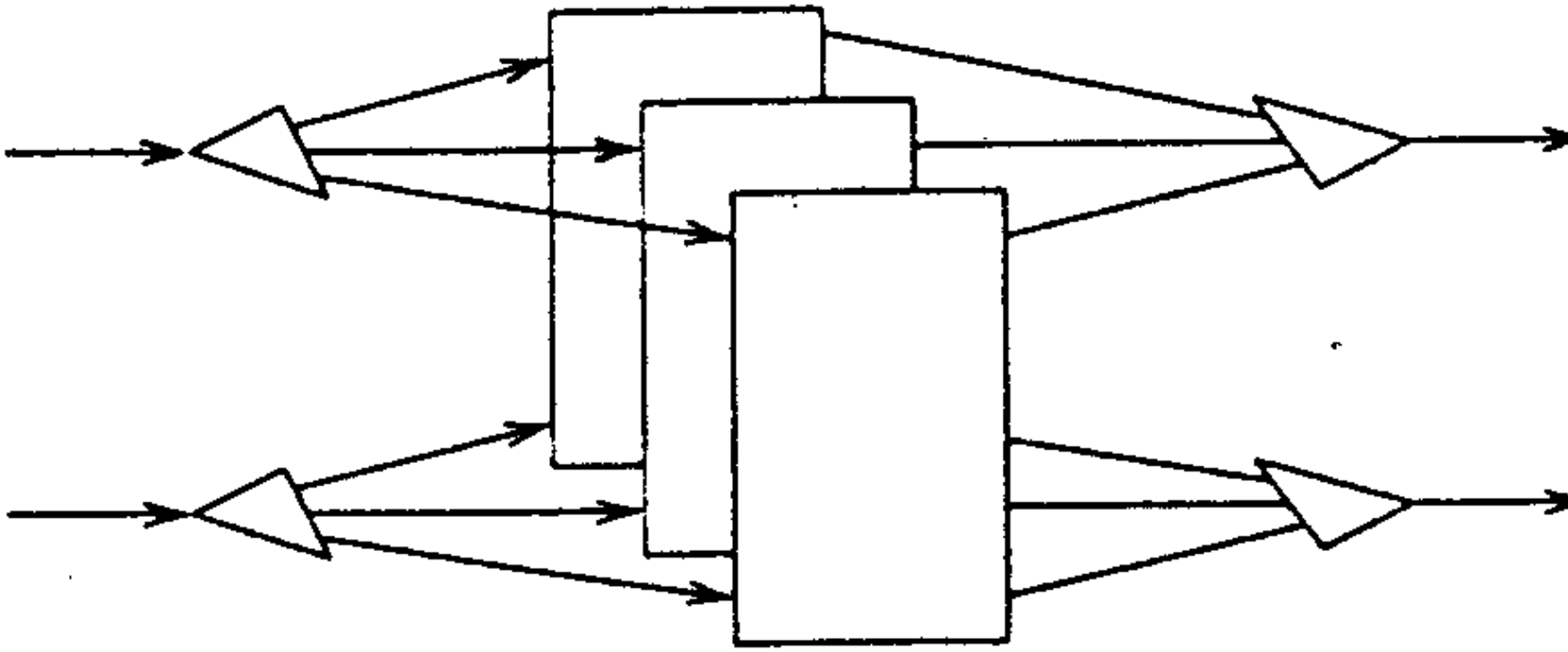
$$N_x = 2N(2n - 1) + (2n - 1)\left(\frac{N}{n}\right)^2.$$

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

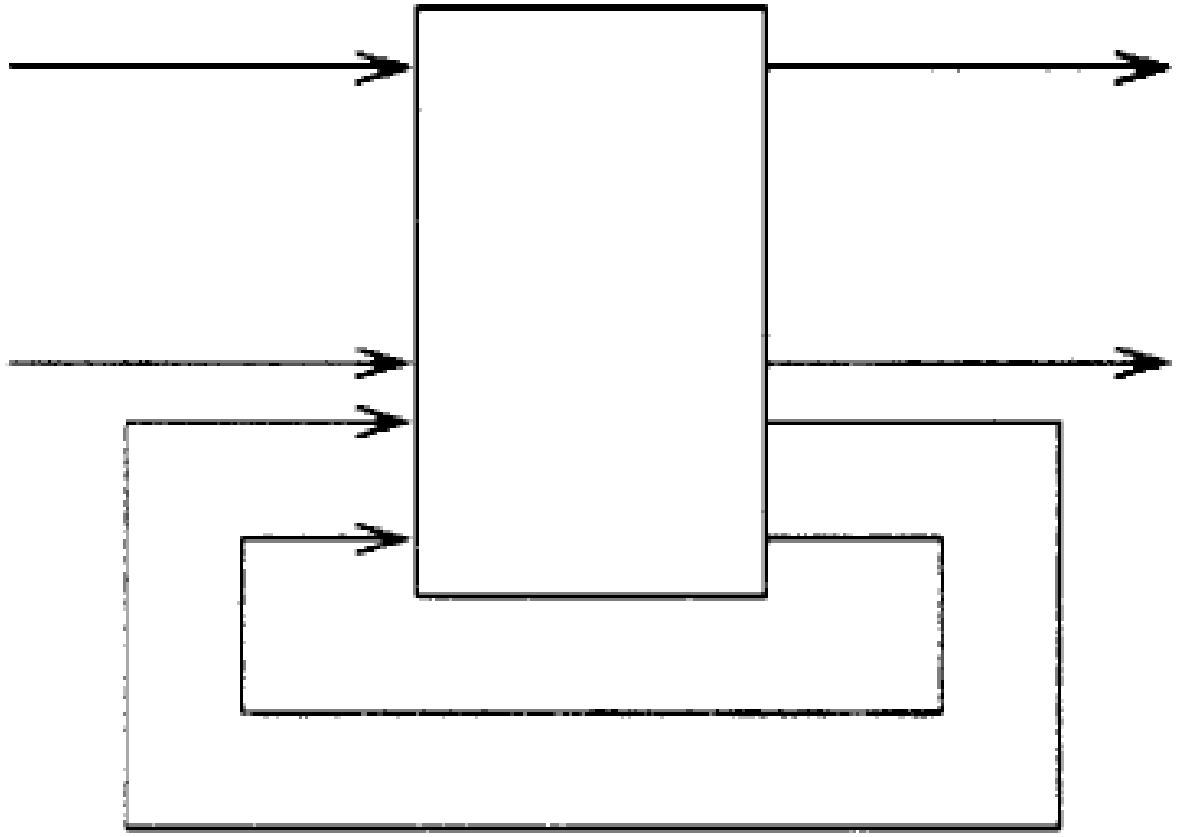
$$n \approx (N/2)^{\frac{1}{2}}$$

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

Space Division - Multiplane Switches



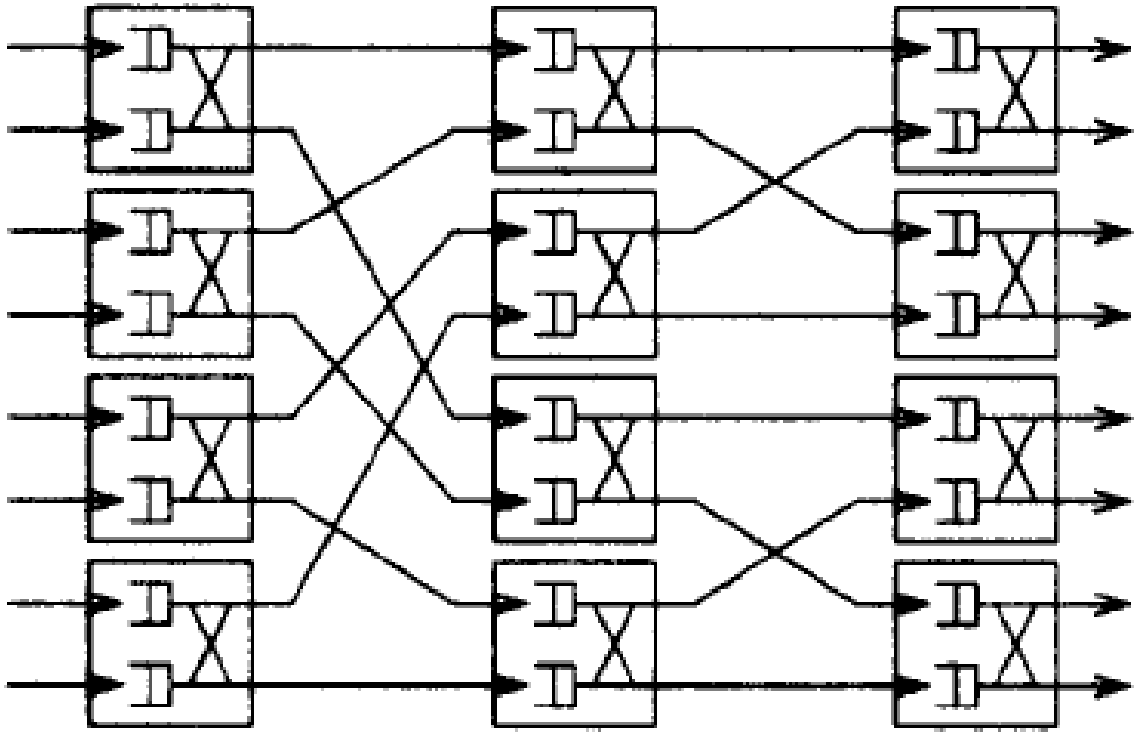
Space Division -Recirculation Switches



Buffering Strategies

- Internally Buffered Switches
- Recirculation Buffered Switches
- Crosspoint-Buffered Switches
- Input-Buffered Switches
- Output-Buffered Switches
- Shared-Buffer Switches
- Multistage Shared-Buffer Switches
- Input- and Output-Buffered Switches
- Virtual-Output-Queueing Switches

Internally Buffered Switches

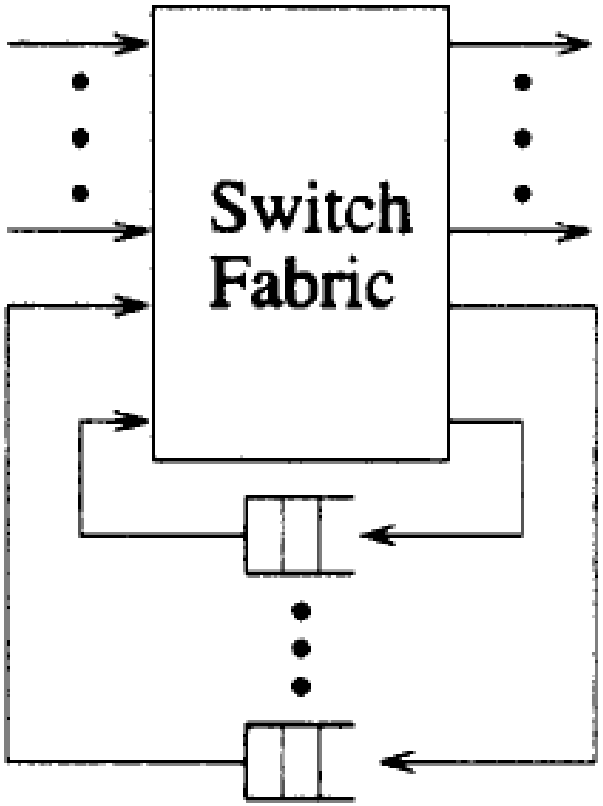


(a) Internal Buffered

Internally Buffered Switches

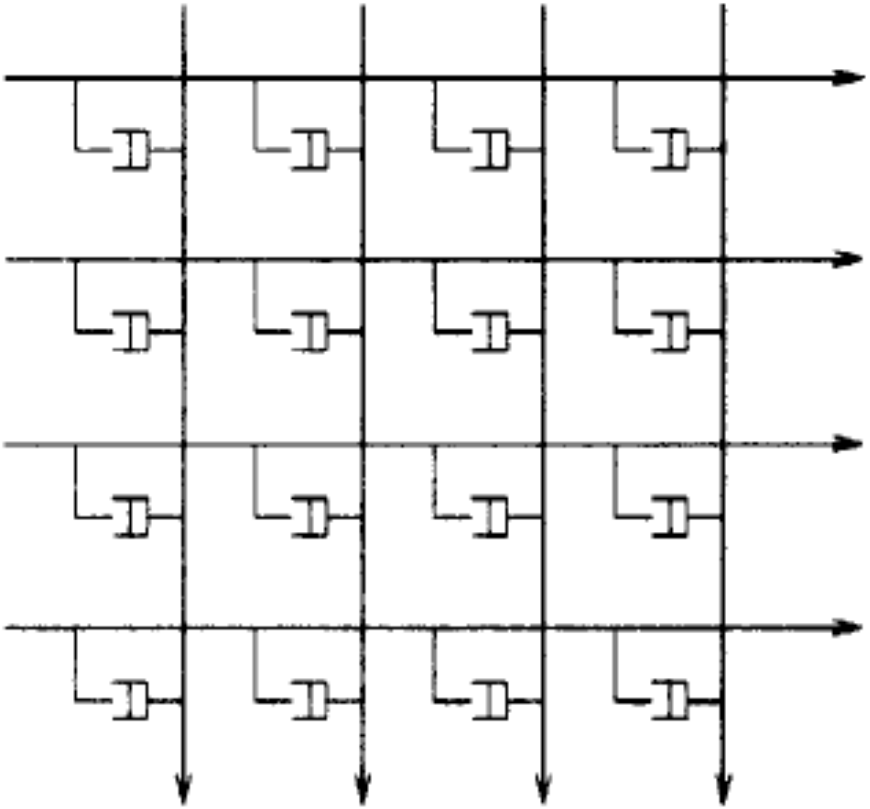
- Advantages:
  - Low cell loss rate
  - Easily scalable
- Disadvantages:
  - Low throughput
  - High transfer delay
  - To meet QoS requirements, some scheduling and buffer management schemes need to be installed at the internal SEs

Recirculation Buffered Switches



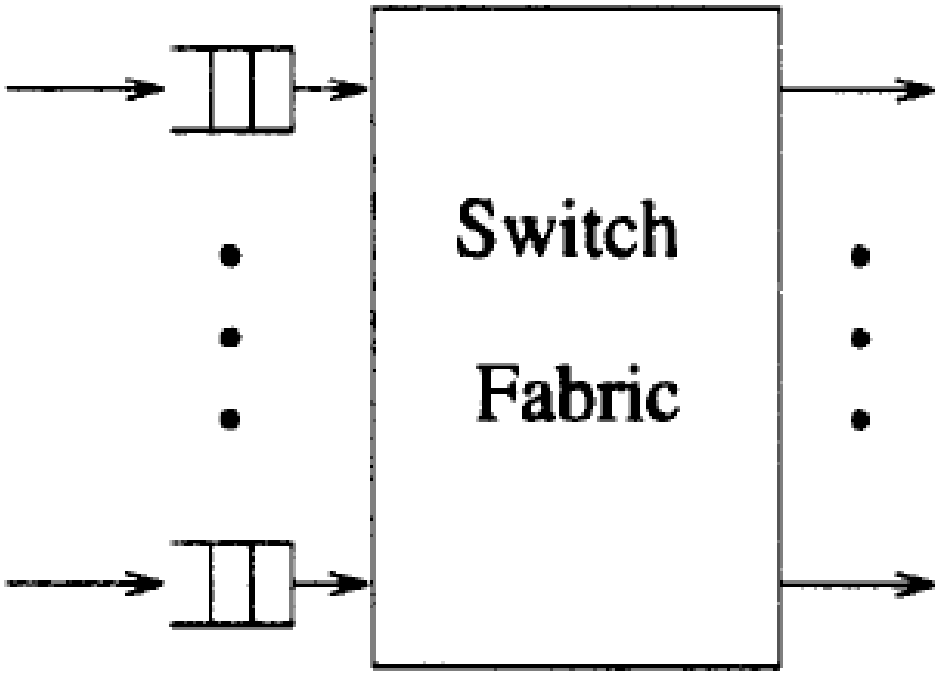
(b) Recirculation Buffered

Crosspoint-Buffered switches



(c) Crosspoint Buffered

Input-Buffered Switches



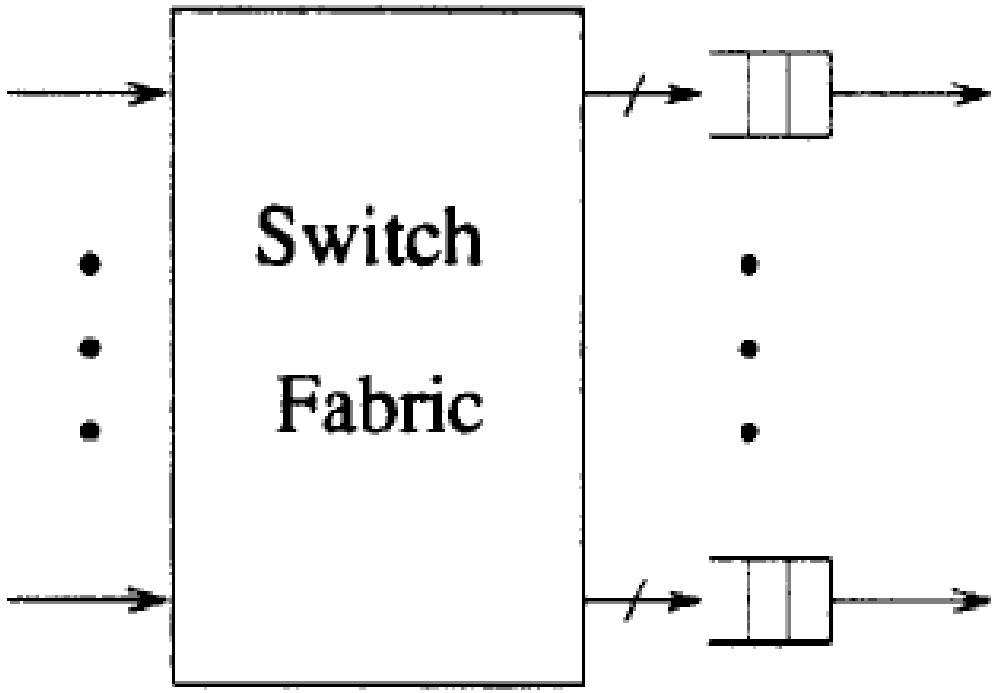
(d) Input Buffered

Input-Buffered Switches

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

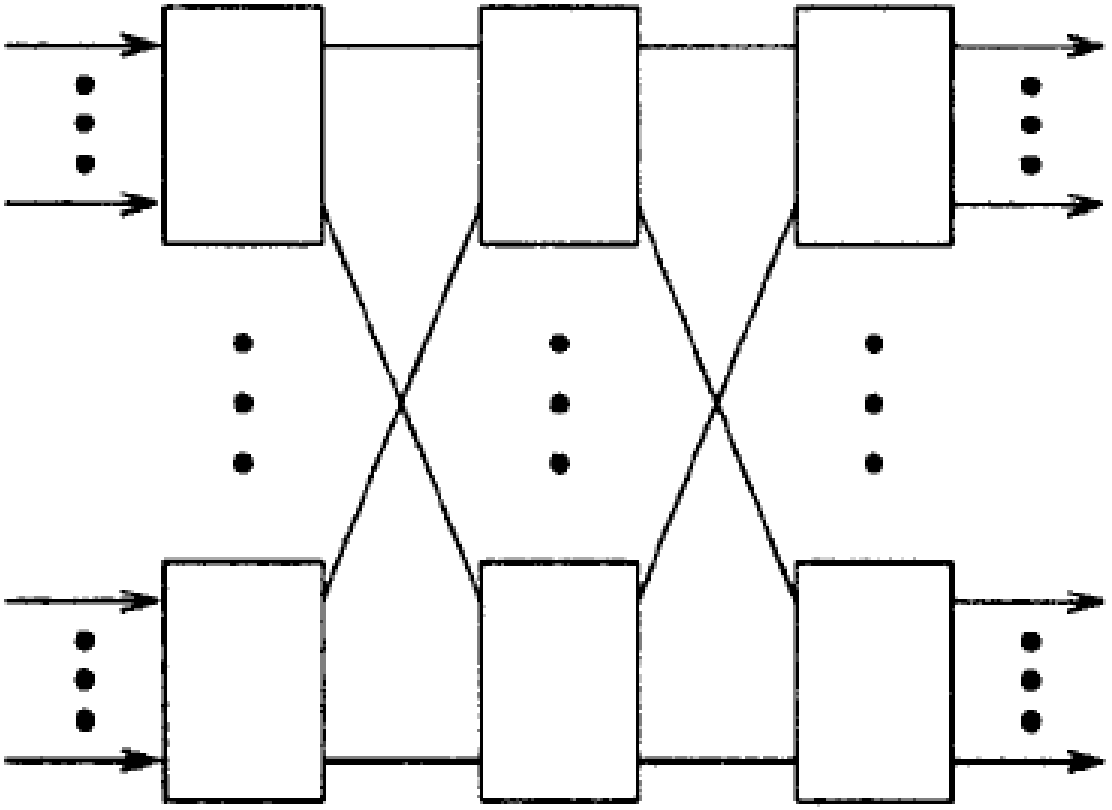
- By using *Windowing* technique, the throughput will be increased. For instance, by increasing the window size to two, the maximum throughput is increased to 70%.

Output Buffered Switches



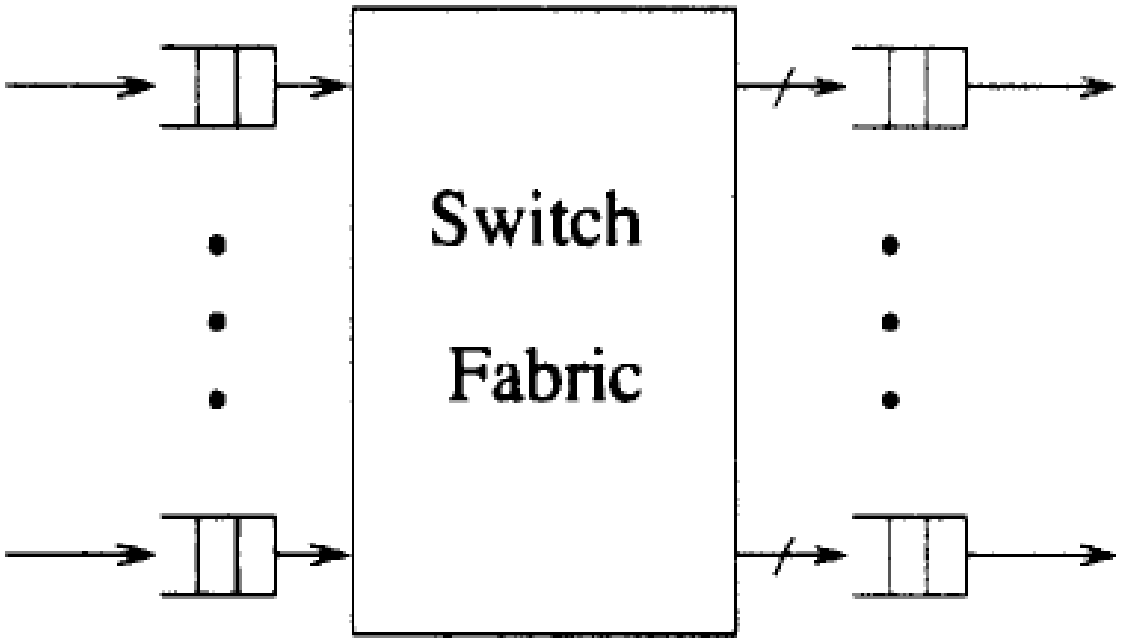
(e) Output Buffered

Multistage Shared-Buffered switches



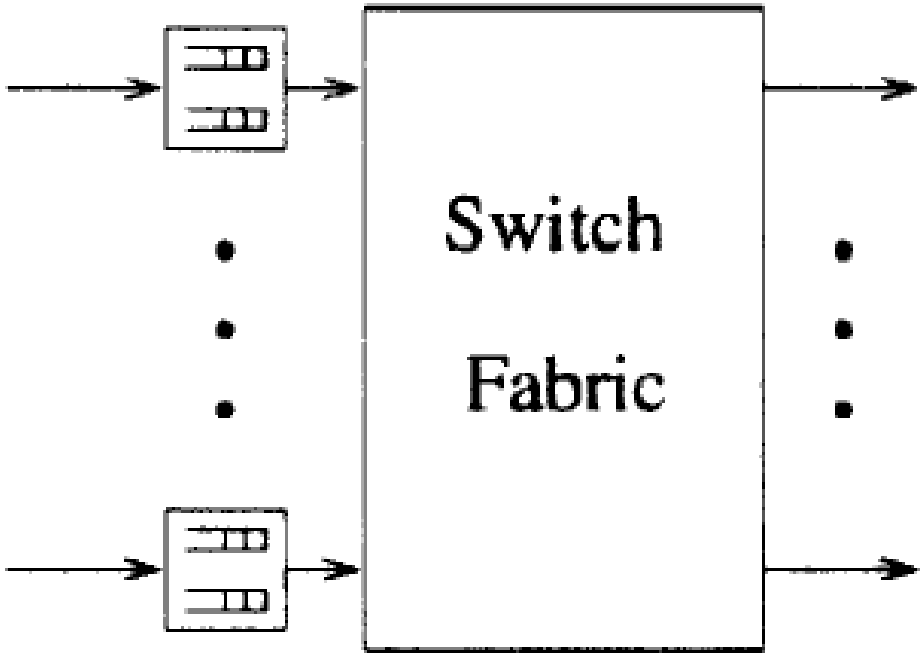
(g) Multistage Shared Buffer

Input- and output-Buffered Switches



(h) Input and Output Buffered

Virtual-Output-Queuing Switches



(i) Virtual Output Queueing

Performance of Basic Switches

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

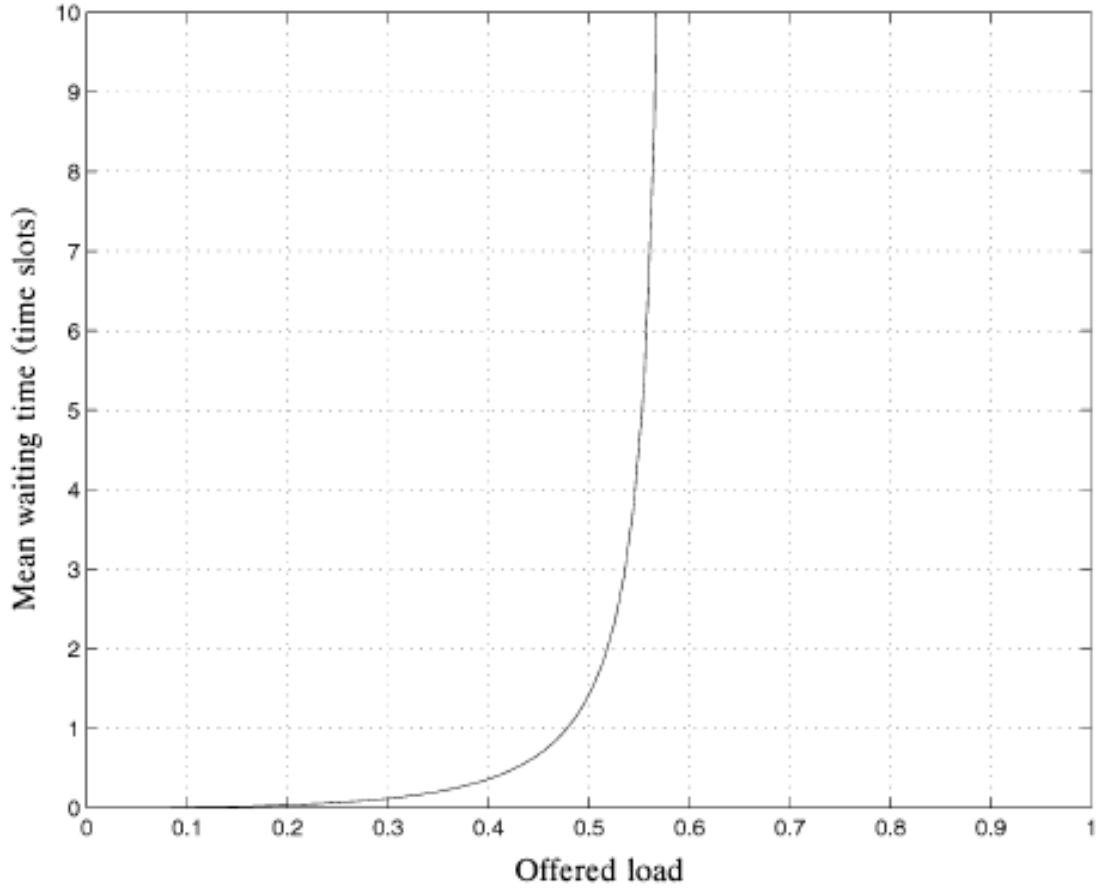
Input-buffered switches

- 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

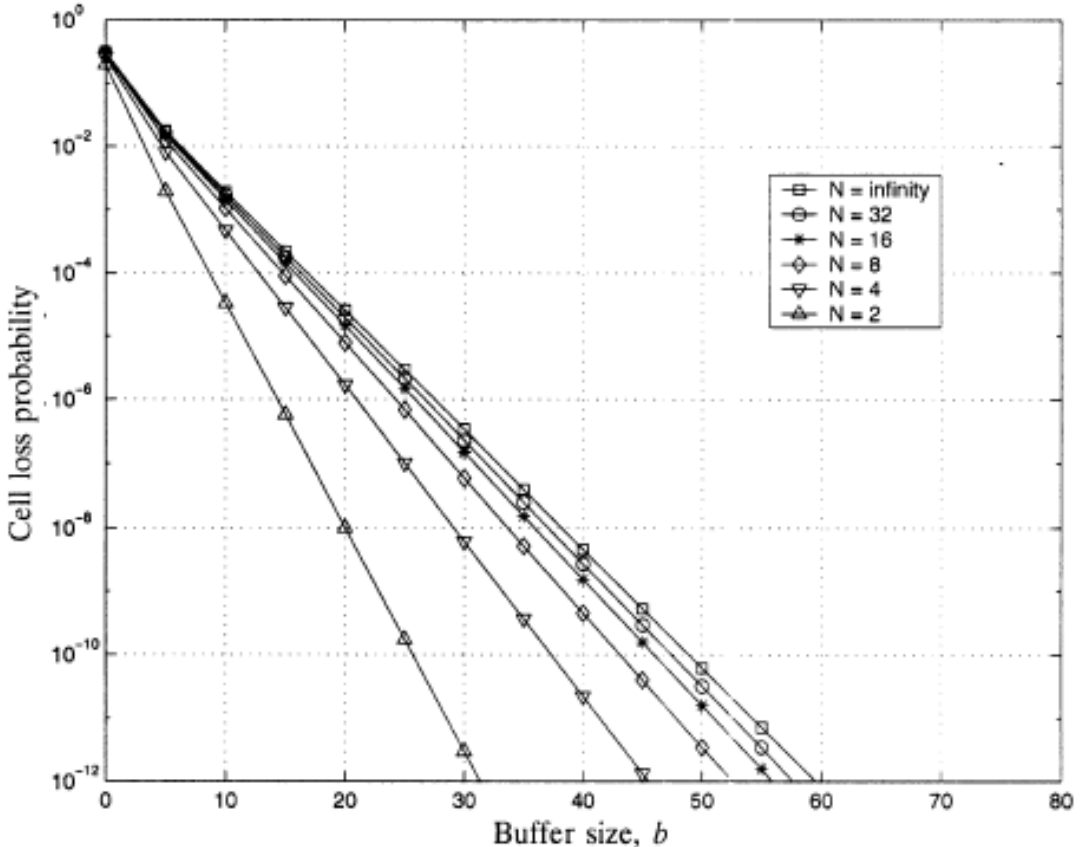
Input-buffered switches

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



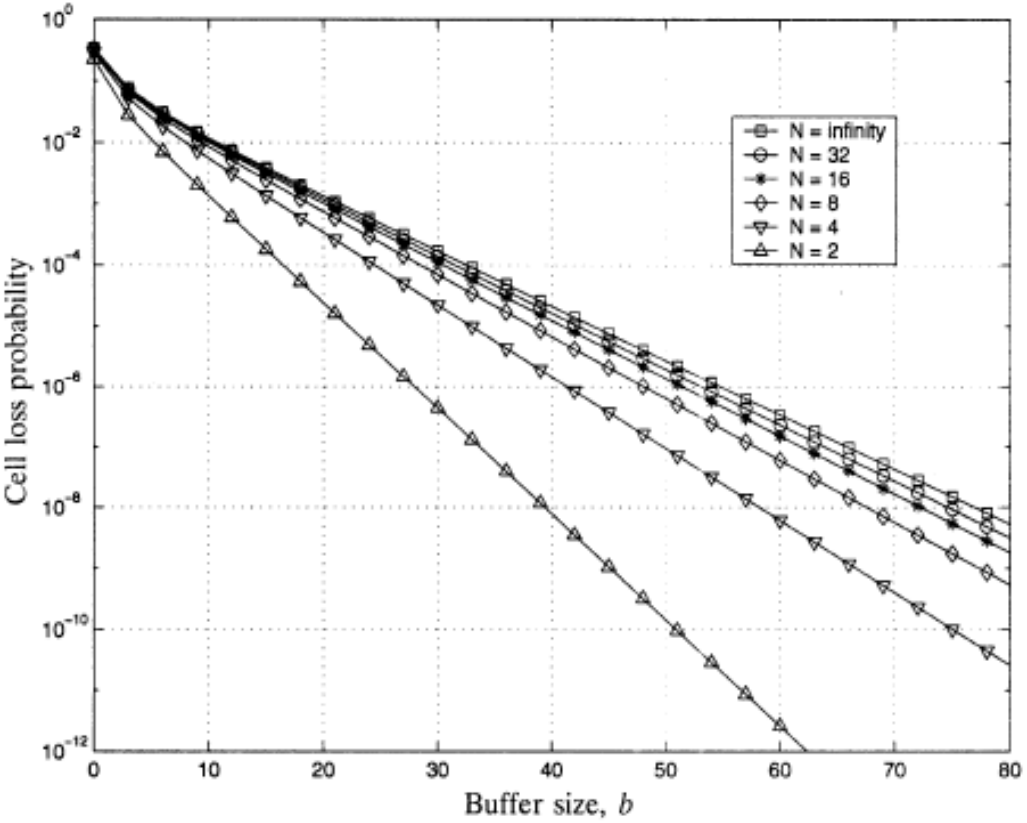
Output-buffered switches

- 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



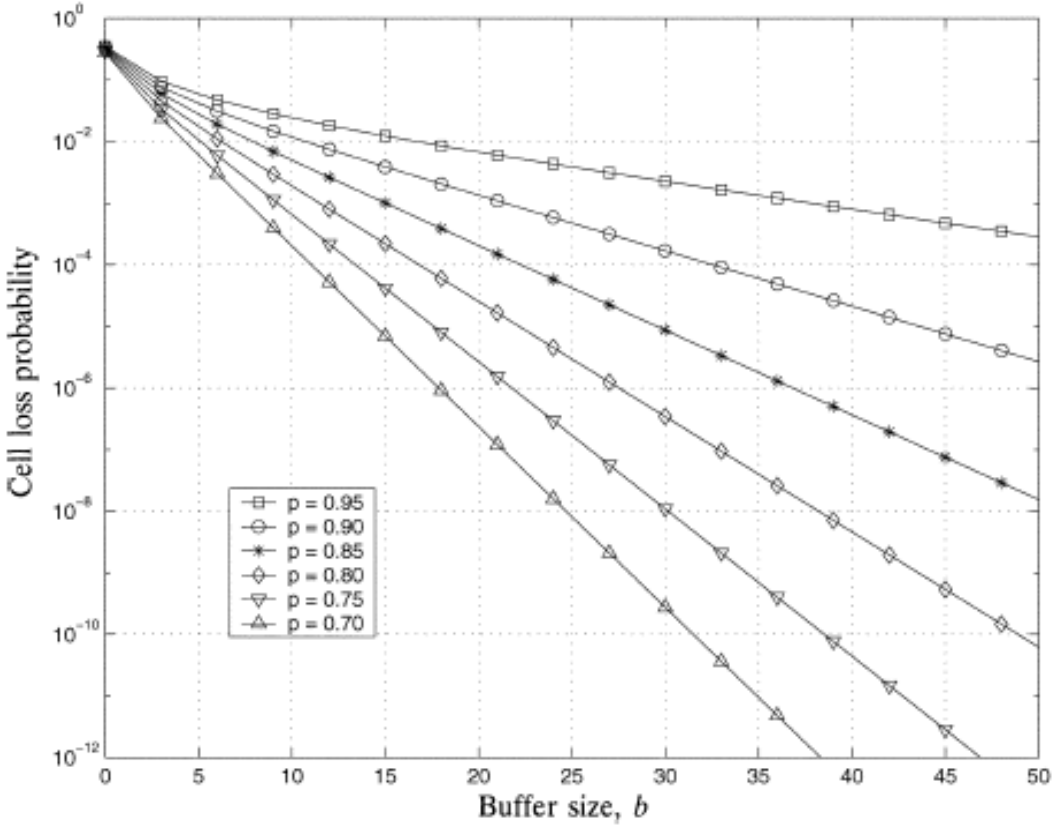
Output-buffered switches

- 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



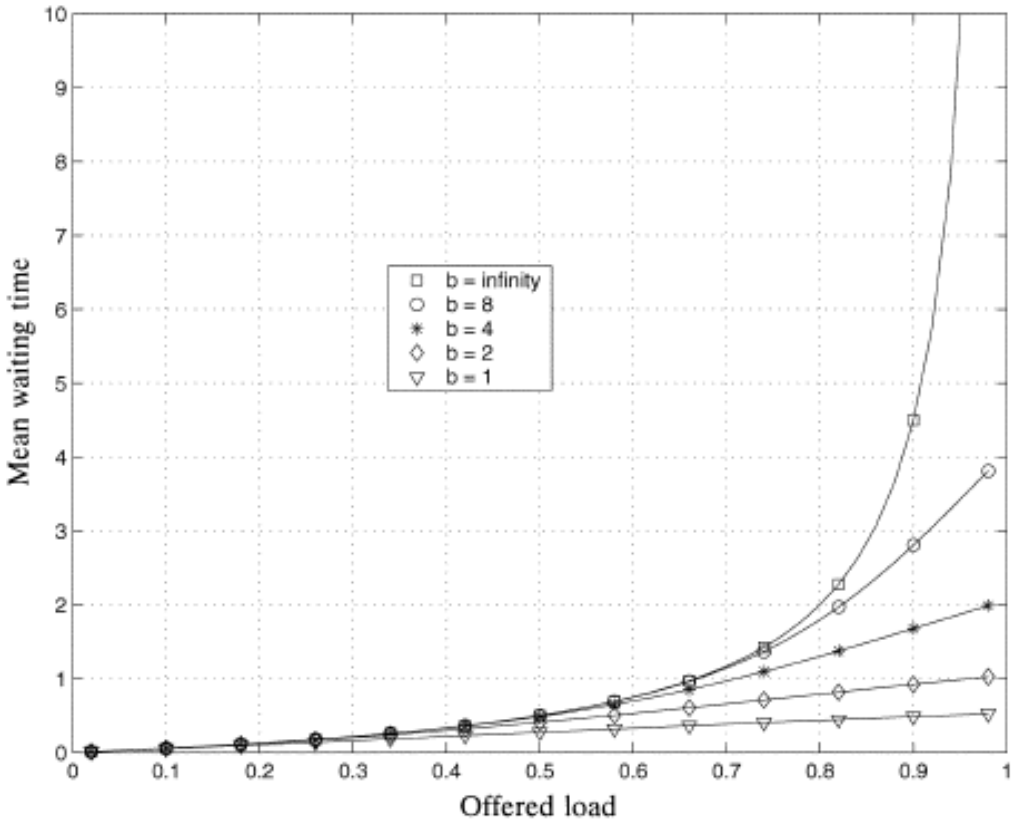
Output-buffered switches

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 $N \rightarrow \infty$



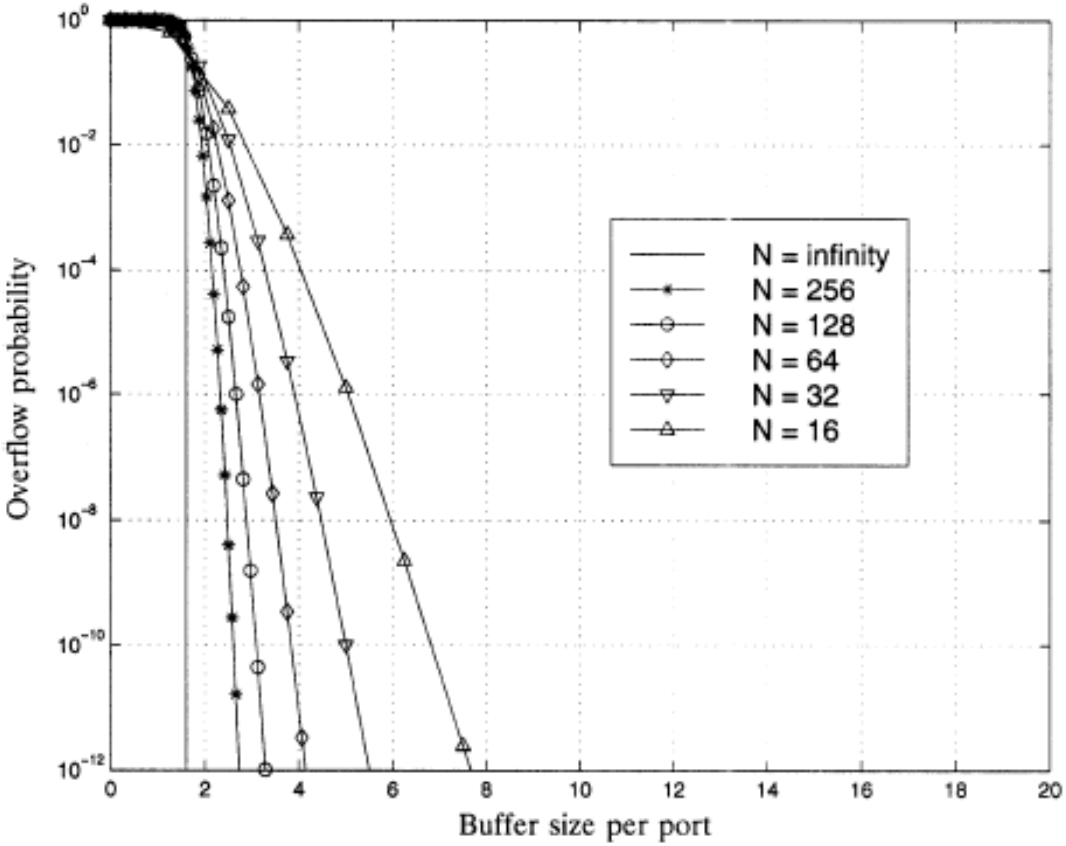
Output-buffered switches

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



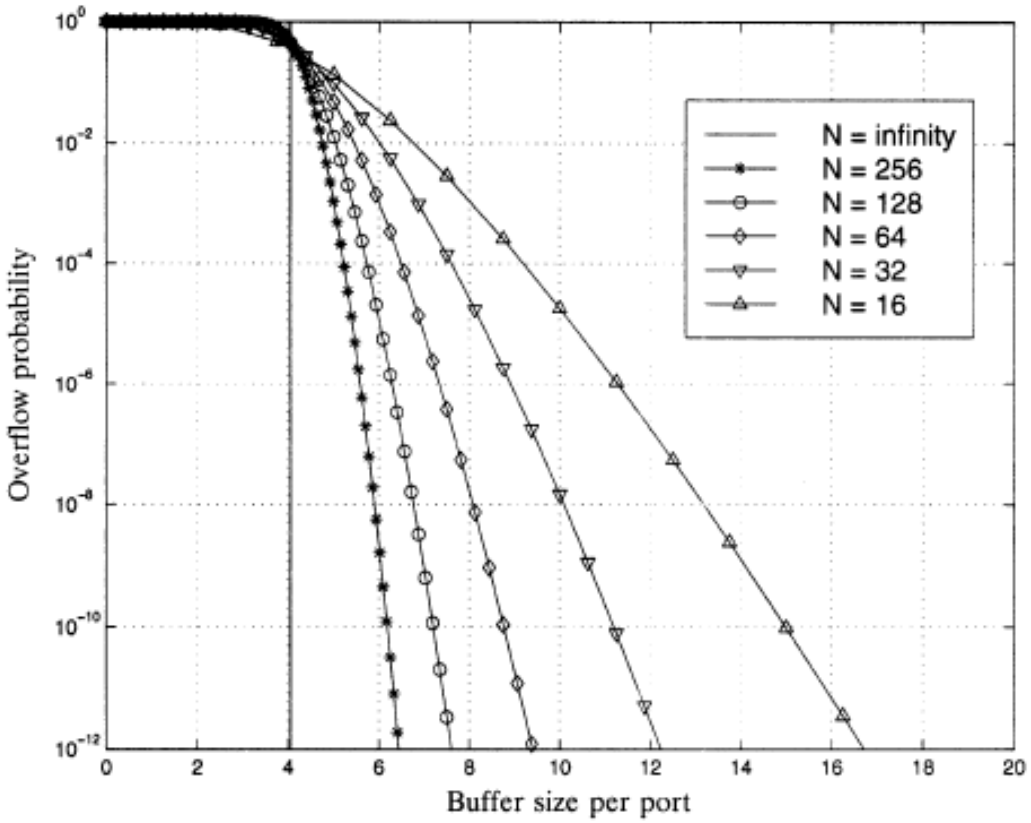
Completely Shared-buffer switches

- 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



Conclusion

- We briefly described major classes of packet switches.
- It is clear that there is no “the best” switch for all situations and applications
- All switches have their own Pros and Cons
- More detail of these switches will be explained in the other sides