# Chapter 4 Electronic Space Division Switching

School of Information Science and Engineering, Shandong University Associate Prof., Degiang Wang

#### Outline

- Stored Program Control
- Centralized SPC
- Distributed SPC
- Software Architecture
- Application Software
- Enhanced Services
- Two-Stage Networks
- Three-Stage Networks
- N-Stage Networks

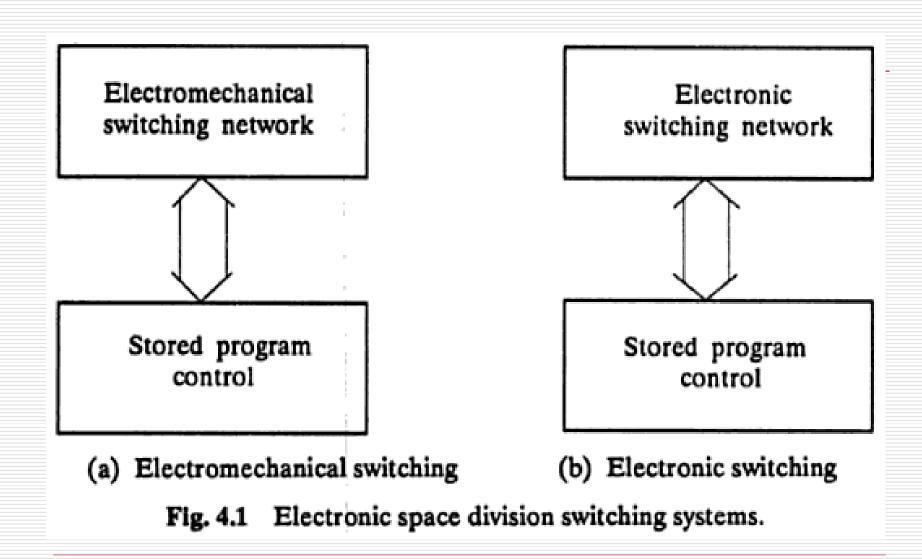
#### Introduction

- Efforts to improve the speed of control and signaling
  - Late 1940s and early 1950s. Use of vacuum tubes, transistors, gas diodes, magnetic drums and cathode ray tubes.
  - Arrival of modern electronic digital computers.
  - The registers and translators of common control systems can be replaced by a single digital computer.

- ☐ Stored Program Control (SPC)
  - Carrying out the exchange control functions through programs stored in the memory of a computer.
  - Consequence
    - Full-scale automation of exchange functions
    - Introduction of a variety of new services

- New features possible for SPC
  - Common Channel Signaling (CCS)
  - Centralized maintenance
  - Automatic fault diagnosis
  - Interactive human-machine interface
- Special Requirements of SPC
  - Operating without interruption
  - Fault tolerant hardware and software

- Two types of SPC switching system
  - Electromechanical Switching SPC+Electromechanical switching network
  - Electronic Switching
    SPC+Electronic switching network



- Organization of SPC
  - Centralized SPC
    Broadly used in early SPC switching systems.
  - Distributed SPC Gaining popularity in modern switching systems.

#### 4.2 Centralized SPC

- Concept
  - All the control equipment is replaced by a single powerful processor.
- □ Configuration of centralized SPC
  - Typical organization
  - Redundant configuration

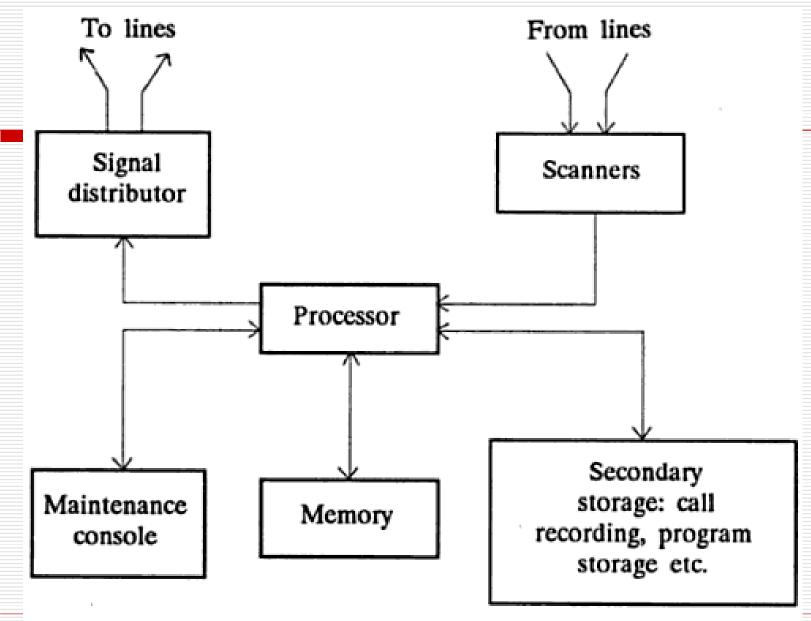
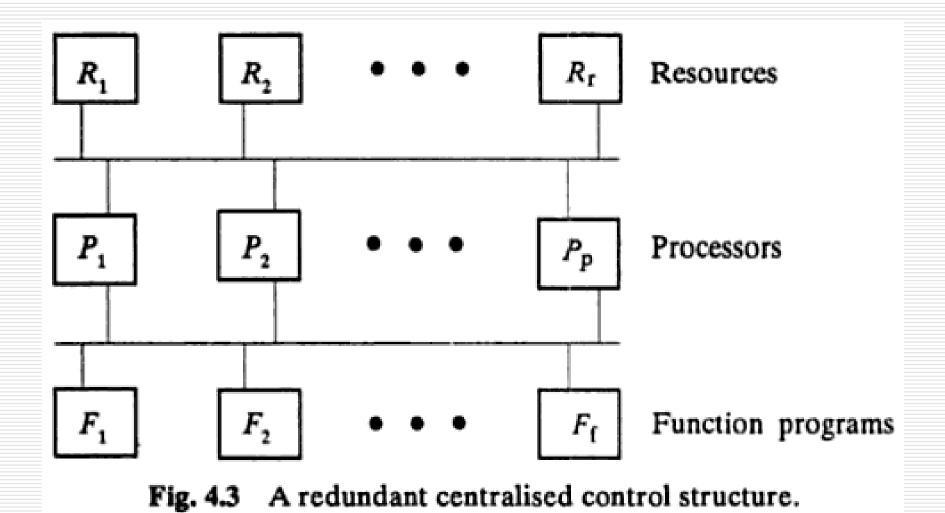


Fig. 4.2 Typical centralised SPC organisation.



#### 4.2 Centralized SPC

- Operation modes in redundant configuration (e.g. dual processor)
  - Standby (备用) mode
  - Synchronous duplex (同步双机) mode
  - Load sharing (负荷分担) mode

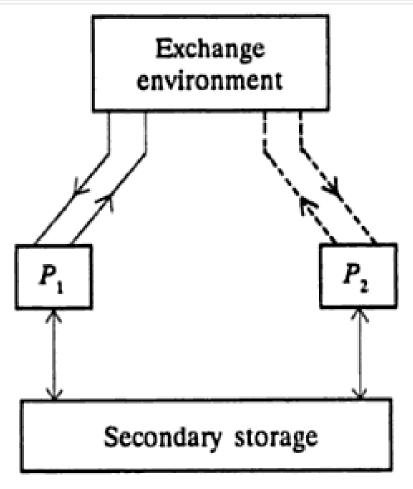
- ☐ How does it work?
  - All processors have the same capability to control the switching procedure.
  - One processor is active and the other is on standby, both hardware and software wise.
  - The standby processor is brought online only when the active processor fails.

- How does the standby processor take over the control properly?
  - State of the exchange system should be clear to the standby processor as its starting point.
    - Which of the subscribers are busy or free?
    - Which of the trunks are busy or free?
    - Which of the paths are connected through the switching network?
    - > So on

- □ Reconstitution of the state
  - Scanning:
    - The standby processor scans all status signals as soon as it is brought into operation.
    - Only the calls which are being established at the time of failure are disturbed.
    - Only feasible for small exchanges.
  - Shared secondary storage: popular.

#### ■ Shared secondary storage:

- The active processor copies system status into a secondary storage periodically, say very 5 seconds.
- As soon as a switchover occurs, the online processor loads the most recent update of the system status from the secondary storage and continues the operations.
- Only the calls which changed status between the last update and the failure are disturbed.
- Feasible for large exchanges.

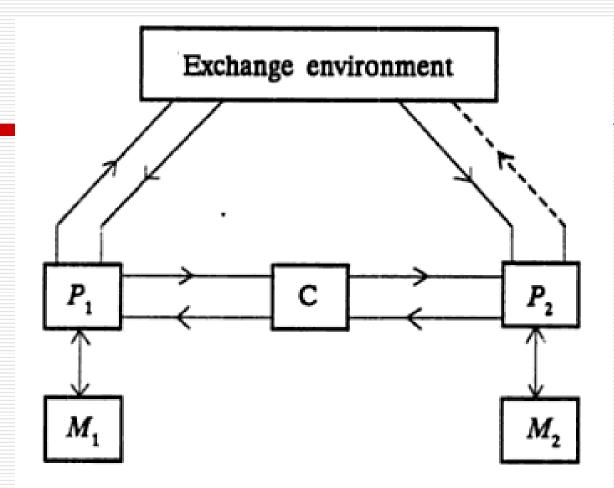


 $P_1$  = active processor  $P_2$  = standby processor

Fig. 4.4 Standby dual processor configuration.

# Synchronous duplex mode

- ☐ How does it work?
  - Both two processor execute the same set of instructions.
  - One of the processor actually controls the exchange.
  - The results from two processors are compared continuously by a comparator.
  - If the results match, the system works normally. Otherwise, a fault occurs, a check-out program is run independently in both two processors to determine which one is faulty.
  - The faulty processor is taken out of service, and the other one works independently.



C = comparator M = memory P = processor

Fig. 4.5 Synchronous duplex operation.

# Synchronous duplex mode

- In case of transient failure of comparator, there are three possibilities exist:
  - Continue with both processors.
  - Take out the active processor and continue with other processor.
  - Continue with the active processor but remove the other processor from service.

# Load sharing mode

- How does it work?
  - Both two processors have access to entire exchange environment. Each of them has independent memories for redundancy purpose.
  - Both two processors are active simultaneously and share the load and the resources dynamically.
  - An incoming call is assigned randomly or in a predefined order to one of the processors which then handles the call right through completion.
  - Inter-processor links are configured for processors to exchange information needed for mutual coordination and verifying the 'state of health' of the other.

# Load sharing mode

- □ If a processor fails, the other processor takes over the entire load including the calls already set up by the failing processor.
- Exclusion mechanism in resource sharing
  - The processors should not seek the same resource at the same time.
  - Implementation: hardware & software

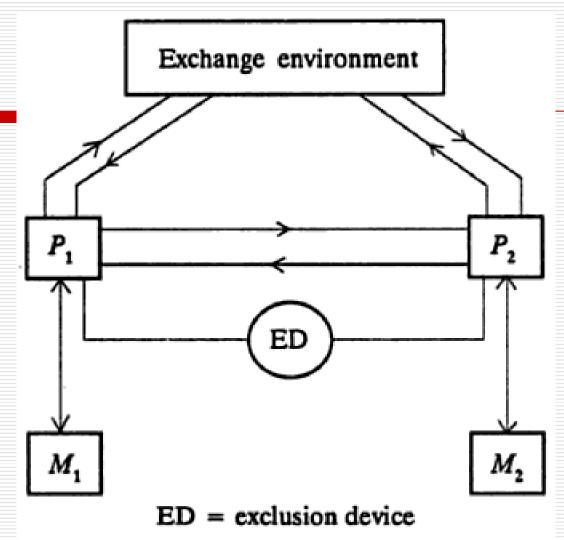


Fig. 4.6 Load sharing configuration.

# Load sharing mode

- Traffic distribution between processors
- □ Load sharing increases the effective traffic capacity by 30 percent compared with synchronous duplex.

#### Availability of system

- A telephone exchange must show more or less a continuous availability over a period of 30 or 40 years.
- □ Availability (可用性,可获得性)
  - Availability of a single processor system
    - ➤ A=MTBF/(MTBF+MTTR)

MTBF: mean time between failures

MTTR: mean time to repair, MTTR < < MTBF

normal normal fault normal fault normal

#### Availability of system

- Unavailability of a single processor system
  - ▶ U=1-A=MTTR/(MTBF+MTTR)≈MTTR/MTBF
- Availability of dual processor system
  - Mean time between failures MTBF<sub>D</sub>=MTBF<sup>2</sup>/(2\*MTTR)
  - $\rightarrow$  AD=MTBFD/(MTBFD+MTTR) =MTBF $^2$ /(MTBF $^2$ +2MTTR $^2$ )
  - $\triangleright$  U<sub>D</sub>=1-A<sub>D</sub> $\approx$  2MTTR<sup>2</sup>/MTBF<sup>2</sup>

normal	norm	al norm	normal		normal	fault	normal
normal	fault	normal	ormal no		normal	fault	normal

# Availability of system

#### ■ Example:

MTBF=2000hours, MTTR=4hours

$$U=?$$
  $U_D=?$ 

#### Solution:

 $U \approx 4/2000 = 2 \times 10^{-3}$ 

i.e. 525 hours in 30 years.

 $U_D \approx 2x16/(2000x2000) = 8x10^{-6}$ 

i.e. 2.1 hours in 30 years.

#### Coordination of functions

- □ Priority of Control functions (控制功能优先级)
  - Level 1: O&M and charging
  - Level 2: Call processing
  - Level 3: Event monitoring and distribution
- Interrupt processing
  - Vectored interrupt (矢量化中断)
    - The interrupting source supplies branch address information to the processor.
  - Nonvectored interrupt (非矢量化中断)
    - A main interrupt service routine scans the interrupt signals and decides on the appropriate routine to service the specific interrupt.

#### Levels of control functions

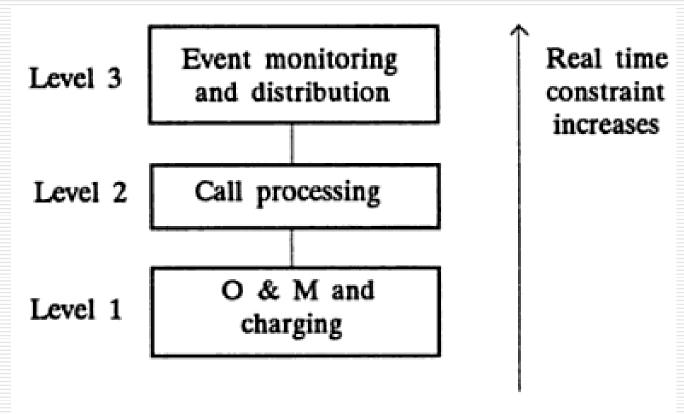


Fig. 4.7 Levels of control functions.

# Interrupt processing

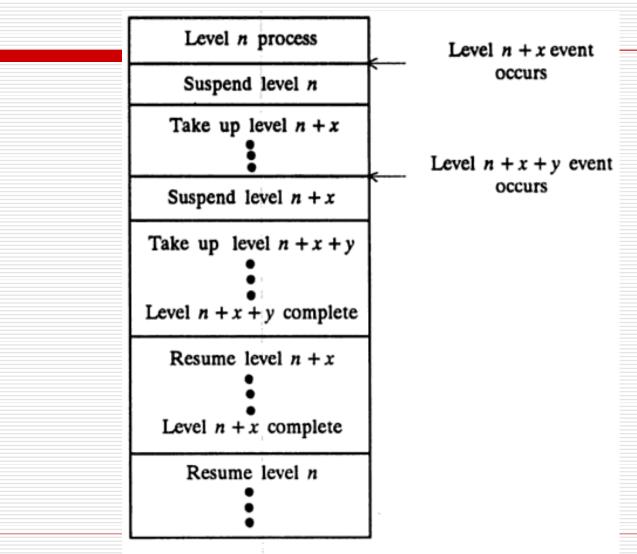
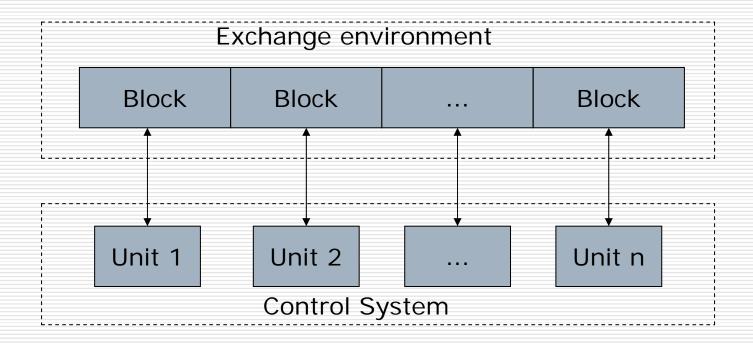


Fig. 4.8 Interrupt processing.

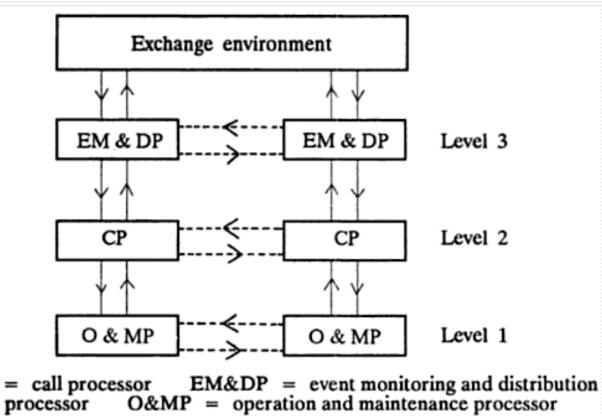
- Concept of distributed SPC
  - The control functions are shared by many processors within the exchanges.
- Background
  - Low cost processors
- Advantages
  - Better Availability (可用性)
  - Better Reliability (可靠性)

- Decomposition of Control Functions
  - Vertical decomposition (垂直分解)
    - The exchange environment is divided into several blocks.
    - Each block is assigned to a processor.
    - A processor performs all control functions related to the corresponding block.
    - The processor in each block may be duplicated for redundancy purposes.
    - Obviously, the control system consists of a number of control units.
    - The modular structure is flexible for system expanding.



Vertical decomposition

- Horizontal decomposition (水平分解)
  - The control functions are divided into groups, e.g. event monitoring, call processing, and O&M functions.
  - Each processor performs only one or some of the exchange control functions.
  - A chain of processors are used to perform the entire control of the exchange.
  - The entire chain may be duplicated to provide redundancy.



processor

Fig. 4.9 Dual chain distributed control.

#### 4.3.1 Level 3 Processing

- □ Tasks
  - Scanning: Subscriber lines/Trunks
  - Distribution: Subscriber lines/Trunks
  - Marking: Registers
- Location
  - Close to switching network, junctors and signalling equipment.

## 4.3.1 Level 3 Processing

- Control schemes
  - Hard-wired control unit
    - Designed as a collection of logic circuits using logic elements
  - Microprogrammed control unit
    - Using microprocessors to implement all control logic

Table 4.1 Characteristics of Electronic Control Schemes

Microprogrammed control	Hard-wired control
Flexible	Not flexible
Slower	Faster
More expensive for moderate processing functions	Less expensive for moderate simple and fixed processing
Easier to implement complex	Difficult to implement complex
processing functions	functions
Introducing new services is easy	Not easily possible
Easier to maintain	Difficult to maintain

## 4.3.2 Level 2 Processing

- □ Task
  - Call processing
  - Level 2 processor = Switching processor
- □ Type of processors used
  - House-developed (自主研发的、定制的)
  - Standard

#### 4.3.2 Level 2 Processing

- Specific characteristics of Switching processors
  - Efficiently designed instructions
  - Special instructions
  - Special architecture for availability, fault tolerance and security
  - Special peripherals for telephones
  - Special communication link for remote control

## 4.3.2 Level 2 Processing

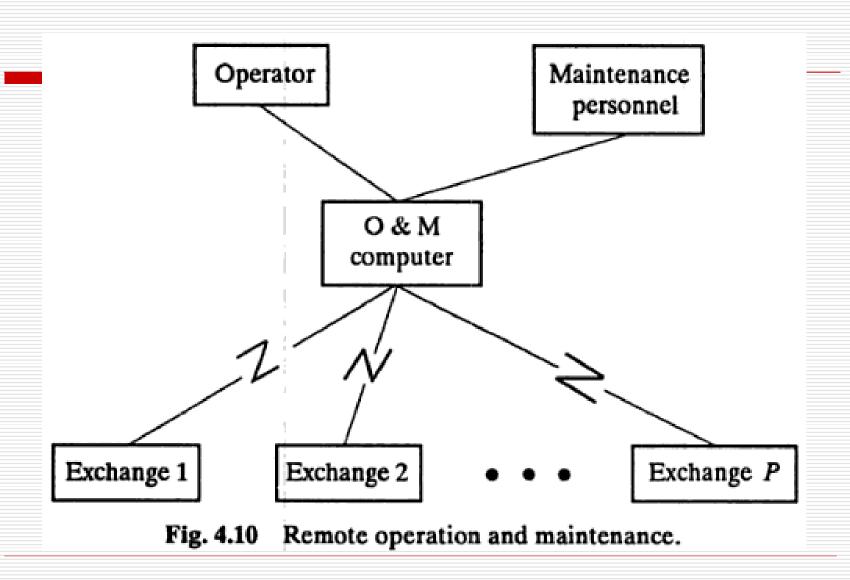
Traffic handling capacity

$$t=a+bN$$
  $N=(t-a)/b$ 

- t: the load of the switching processor measured by its occupancy.
- a: fixed overhead (开销) depending upon the exchange capacity and configuration.
- b: average time to process one call.
- N: number of calls per unit time.

## 4.3.3 Level 1 Processing

- Q&M functions involved
  - Administer the exchange hardware and software.
  - Add, modify or delete information in translation tables.
  - Change subscriber class of services.
  - Put a new line or trunk into operation.
  - Monitor traffic.
  - Detect and locate faults and errors.
  - Run diagnostic and test programs.
  - Man-machine interaction.



#### 4.4 Software Architecture

- Software of SPC systems
  - System software
  - Application software
- We focus on the call processing software systems.

#### Operating system of Call processing

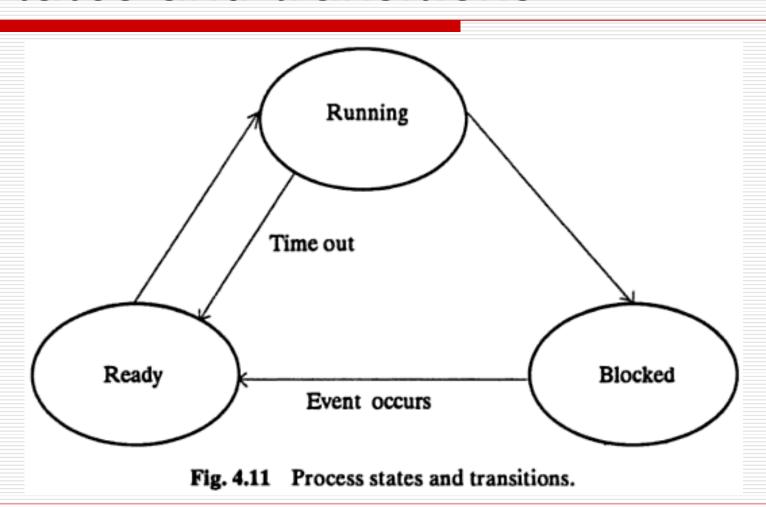
- Multiprogramming (多道程序) Feature
  - Call processing is event oriented.
  - Call setup involves a series of separated processing actions, each action is triggered by an event occurring at a subscriber line or trunk.
  - There are many calls being processed simultaneously, each handled by a process.

Process (进程): a program in execution.

#### Process states and transitions

- States of a process
  - Running state: currently possessing the CPU resources.
    - A CPU can be allocated to only one process at any instant of time.
  - Ready state: waiting for availability of the CPU.
    - All processes in ready state are ordered according to some priority.
    - > The first process in the ordered queue is the next which will receive the CPU.
    - A timer is employed to prevent any process from monopolising the CPU.
  - Blocked state: waiting for some event to occur.

## States and transitions

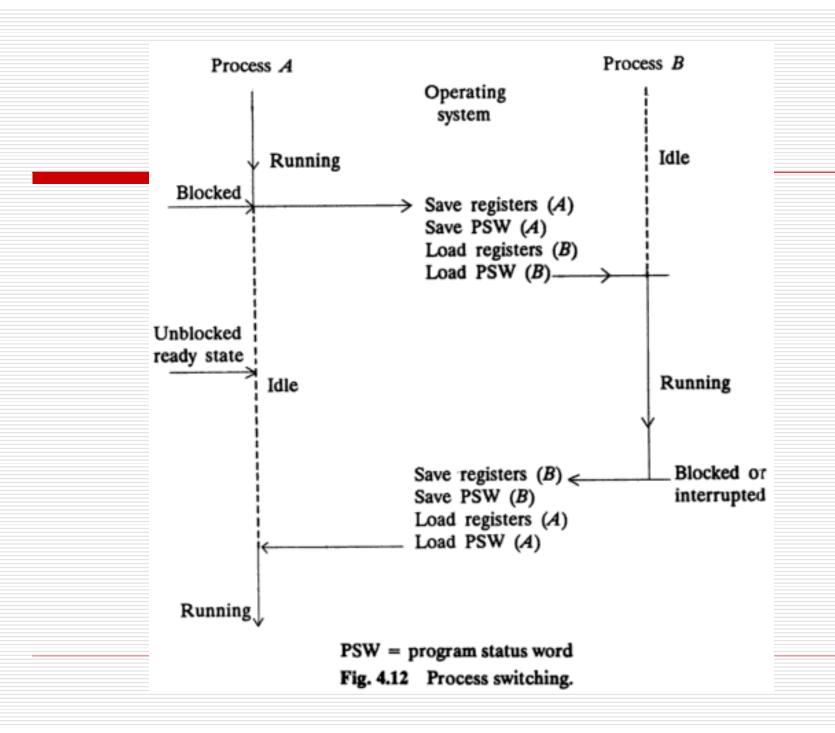


## Representation of a process

- Process Control Block (PCB)
  - PCB is a data structure containing relevant information of a process
    - Current state of the process
    - Process priority and CPU scheduling parameters
    - Register save area to save the contents of the CPU registers when an interrupt occurs
    - Memory allocated to the process
    - Process account like the CPU time limits and usage, process number etc.
    - Status of events and I/O resources associated with the process.

## Process switching (进程切换)

- Process switching occurs in following cases:
  - The current running process becomes blocked;
  - An event or interrupt triggers a high priority process.
- Process switching implies PCB saving and loading.



# Mutual exclusion of processes in accessing shared data

- Data sharing in switching system
  - Processes share common variables, update common tables, write into common files, and so on.
  - Mutual exclusion (互斥): When one process accesses a shared table, all other processes want to access the same table are kept waiting.

## Software production

- Basic factors associated with switching software
  - Complexity and size of the software
  - Long working life required
  - Real time operation
  - Stringent reliability and availability
  - Software portability (可移植性)

#### Four stages in software production

- Functional specification
- Formal description and detailed specification
- Coding and verification
- Testing and debugging

## CCITT standards on language

- □ Z.100
  - specification description language (SDL)
- □ Z.200
  - CCITT high level language (CHILL)
- □ Z.300
  - Man-machine language (MML)

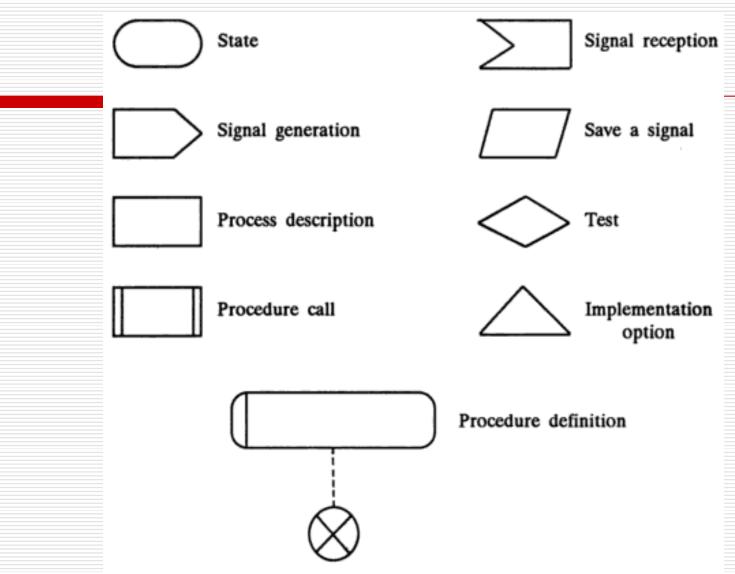
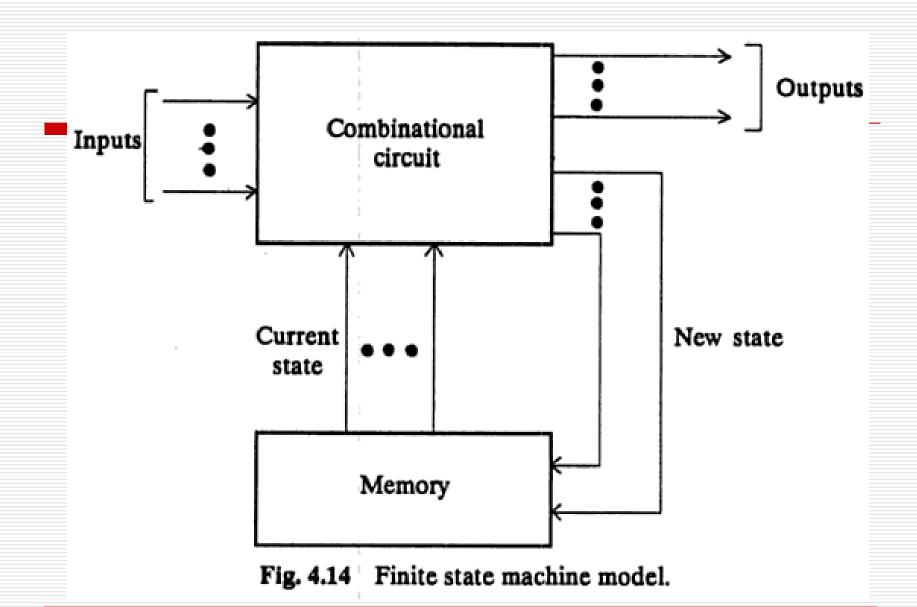
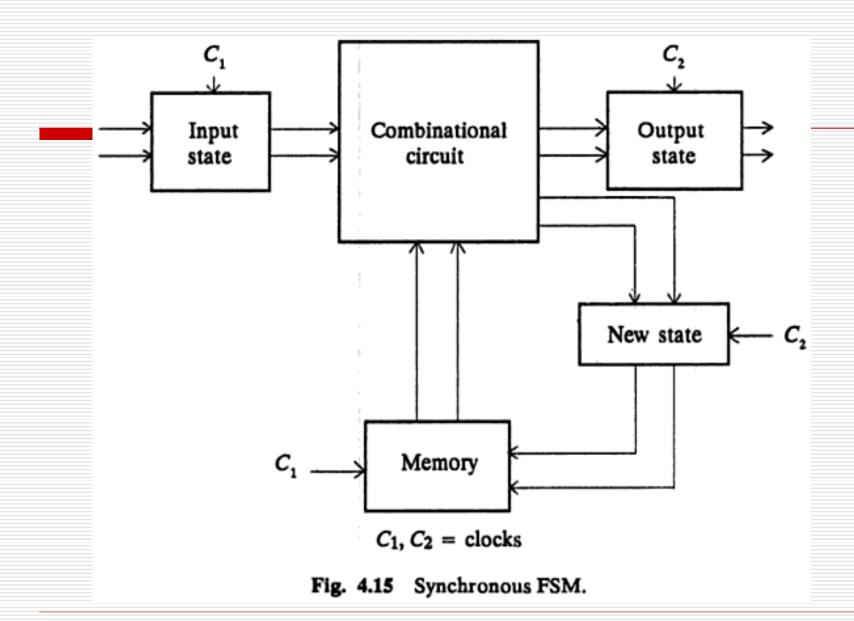


Fig. 4.13 Standard symbols in SDL.





## 4.5 Application software

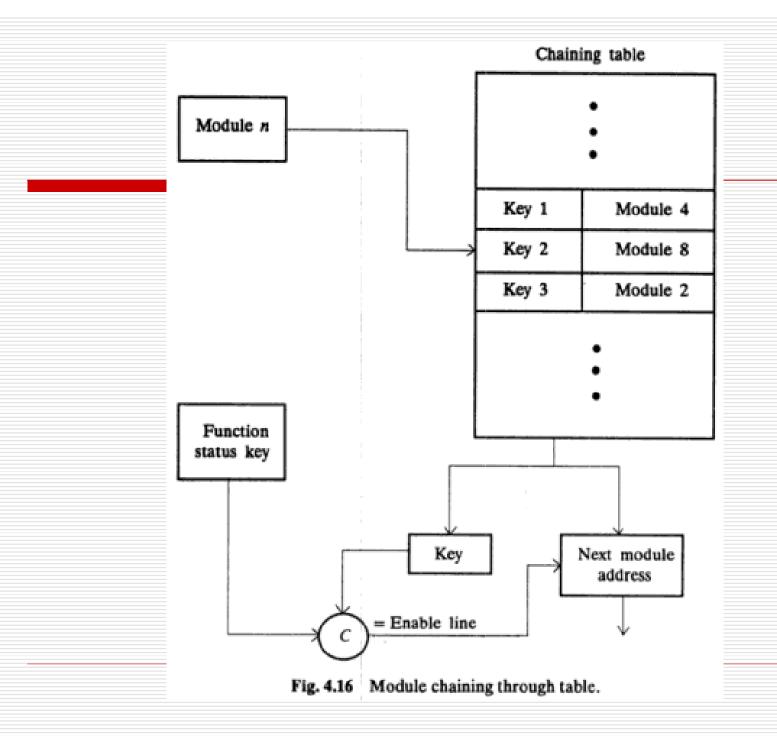
- Classes of application software
  - Call processing software (core resident: 内存驻留)
  - Administrative software (非驻留)
  - Maintenance software (非驻留)
- Description of a software package
  - Organization (组织模式)
  - Data structure (数据结构)
  - Processing functions (处理功能)

#### Organization of switching software

- □ Modular organization(模块化组织模式)
  - The software package is divided into program modules.
  - Each module deals with a specific task.
  - The size of each module varies depending on the task.
  - Modules exchange data through interfaces or data table.
  - Modules are grouped together to constitute independent functional units.
  - A functional unit runs as a separate process.

#### Organization of switching software

- How to string up modules within a process?
  - Special program
  - Chaining table (链表)
    - Each entry consists of a key and a module number
    - When a module completes execution, entries in the chaining table are scanned and the keys are compared with a function status key.
    - The matched module is executed next.
    - Flexibility for modifying a functional unit: adding or removing modules.



- Parameters
  - System parameters
    - Afford flexibility at the overall system level, such as signalling time delays, fault thresholds, etc.
  - Office parameters
    - Define program execution limits at specific exchanges, such as the number of subscribers, the maximum number of simultaneous calls etc.

- □ Data type and storage
  - Semipermanent data:
    - > e.g. hardware parameters
    - Updated rather infrequently
    - Stored as files
  - Temporary data
    - e.g. information specific to a process
    - Have a lifetime equal to the process they pertain to
    - Stored as tables

- Hardware description files
  - Terminal circuit connection file
    - Associated distributor address
    - Associated scanner address
    - Switching network address
    - Physical location address
  - Switching network configuration file
    - Connected subscriber lines
    - Trunks towards other exchanges
    - Rules for digit translation and routing

- Subscriber line information files
  - Correspondence between line equipment number and directory number
  - In-service or out-service status
  - Type of telephone instrument
  - Nominal transmission parameters
  - Call restrictions
  - Subscriber's entitlement
  - Subscriber instructions

- State description tables
  - Subscriber line state table
  - Terminal circuit state table
  - Switching network link state table
  - Working area of a process

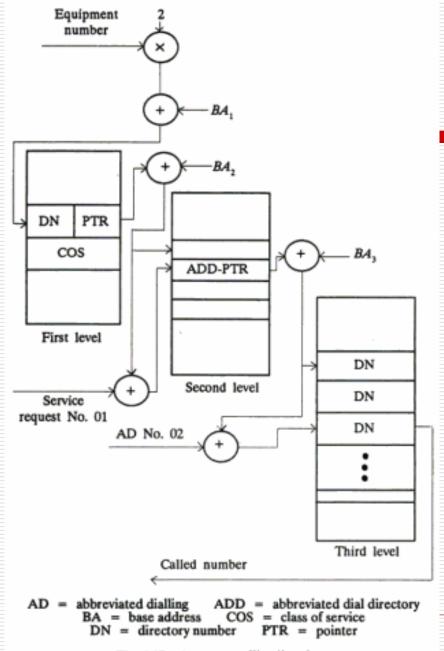
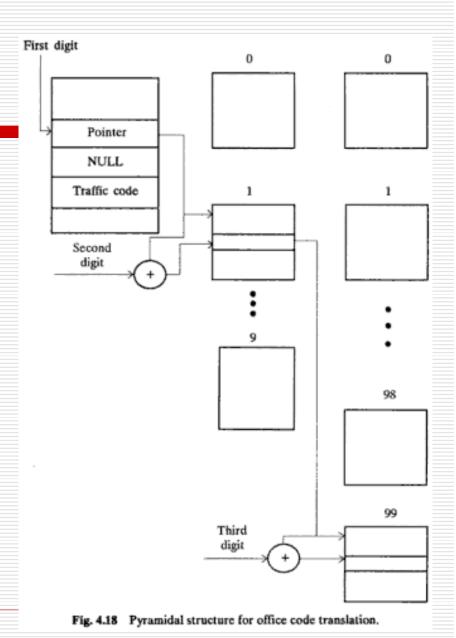


Fig. 4.17 Access to calling line data.

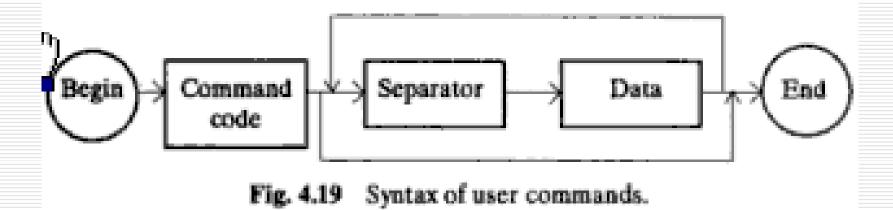


## Functions of switching software

- Call processing
  - Line scanning
  - Calling line related functions
  - Routing and link setup
  - Called line related functions
- Administration
  - Traffic monitoring/control/billing
- Maintenance
  - Diagnostic / preventive maintenances

#### 4.6 Enhanced Services

- Categories of enhanced services
  - Services associated with the calling subscriber and designed to reduce the time spent on dialing and the number of dialing errors.
  - 2. Services associated with the called subscriber and designed to increase the call completion rate.
  - 3. Services involving more than two parties.
  - 4. Miscellaneous services.



## Category 1

- □ Abbreviated dialing (缩位拨号)
- □ Recorded number calls or no dialing calls (无拨号呼叫)
- □ Call back when free (遇忙回呼)

# Category 2

- □ Call forwarding (呼叫前转)
- Operator answer

## Category 3

- □ Calling number record (主叫号码记录)
- □ Call waiting (呼叫等待)
- □ Consultation hold (呼叫保持)
- □ Conference calls (会议电话)

## Category 4

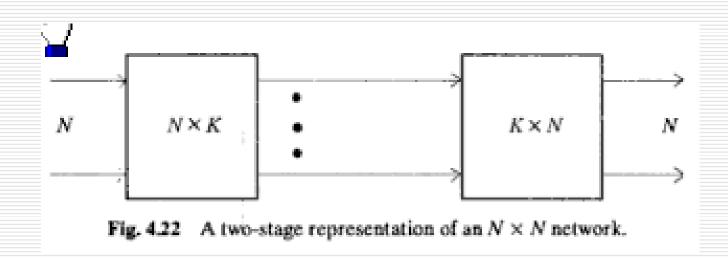
- □ Automatic alarm(自动闹钟)
- □ STD barring (直拨禁用)
- □ Malicious call tracing(恶意呼叫追踪)

STD: subscriber trunk dialing

## 4.7 Two-Stage Networks

- □ Theorem
  - For any single stage network, there exists an equivalent multistage network.
- □ Simple Two-stage NxN network
  - A NxN single stage network with a switching capacity of K connections can be realized by a two-stage network of NxK and KxN.

## Simple Two-stage NxN network



- First Stage: Any of the N inlets can be connected to any of the K outputs. NK switching elements.
- Second Stage: Any of the K inputs can be connected to any of the N outlets. NK switching elements.
- There are K alternative paths for any inlet/outlet pair connection.

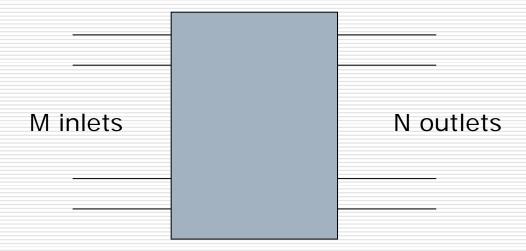
## Simple Two-stage NxN network

- □ Full connectivity/full availability
  - Any of the N inlets can be connected to any of the N outlets.
- Example
  - Assume 10% of the subscribers to be active on average. Set K to be N/16. The number of switching elements is S=N<sup>2</sup>/8.
  - For N=1024, we have K=64, S=131072.
- Note: Feasibility & Flexibility

## 4.7 Two-Stage Networks

- □ Single stage vs. Multistage networks
  - Inlet to outlet connection
  - Quality of link
  - Utility of cross-points
  - Establishment of a specific connection
  - Cross-point & path Redundancy
  - Number of cross-points
  - Capacitive loading problem
  - Blocking feature
  - Call establishing time

- □ Terminology
  - Expanding network: M<N 扩展网络
  - Concentrating network: M>N 汇聚网络
  - Square network: M=N 矩形网络



- □ Architecture of General two-stage networks
  - Multiple small size matrices are used in each stage.
  - Easy to be realized in practice.
  - Flexible in system design.
- MxN two-stage network design
  - Decomposition:  $M = p \times r$ ,  $N = q \times s$
  - Switching matrices: p x s and r x q
  - Full availability: There must be at least one out let from each block in the first stage terminating as inlet on every block of the second stage.

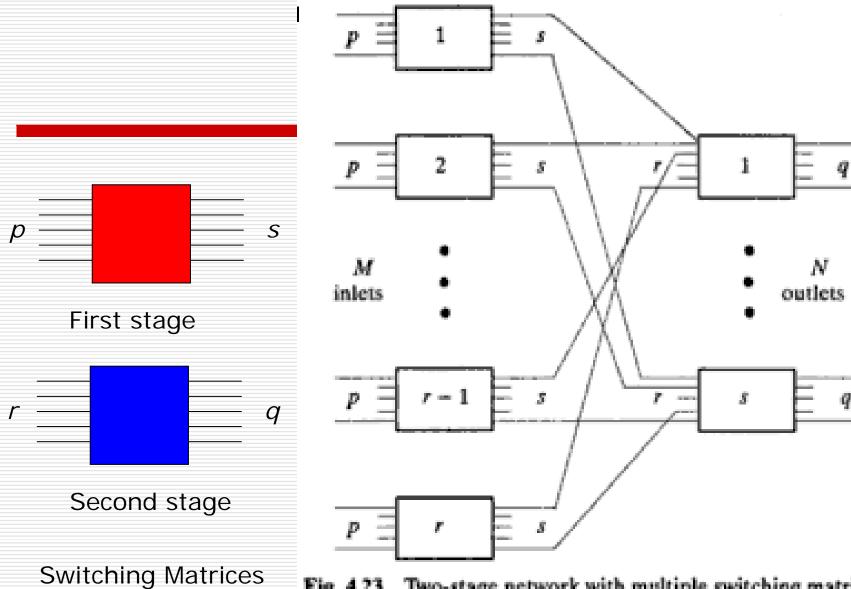


Fig. 4.23 Two-stage network with multiple switching matrices in each stage.

- Parameters
  - Number of switching elements
    - $\triangleright$  S=psr+qrs=Ms+Nr
  - Switching Capacity
    - ▶ i.e., the number of links between the first and the second stages.
    - $\rightarrow$  SC=sr

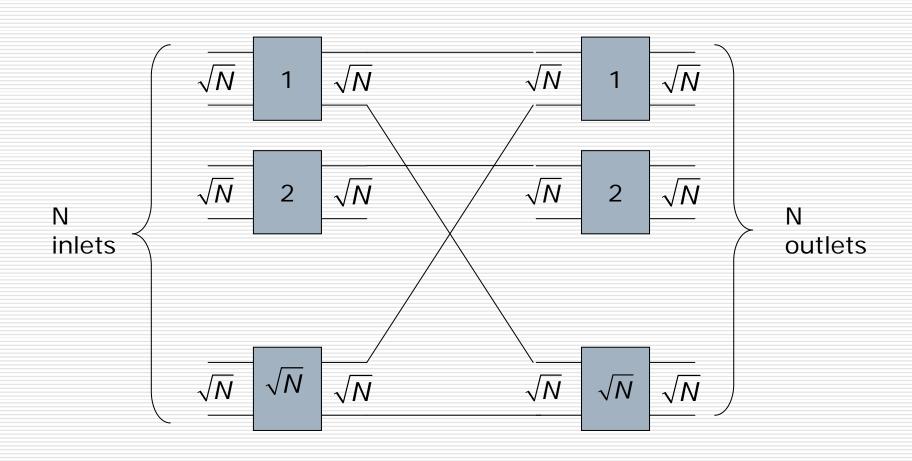
- Parameters
  - Blocking probability
    - Blocking condition 1
      - ✓ There are rxs calls in progress, and the (rs+1)-th call arrives;
      - ✓ The blocking probability P<sub>B</sub> is dependent on the traffic statistics.
    - Blocking condition 2
      - ✓ There is a call in progress from I-th block in the first stage to the J-th block in the second stage, and another call originating in the I-th block destined to the J-th block.
      - Blocking probability  $P_{B} = \frac{M\alpha(s-1)((M/r)-1)\alpha}{rs(s-1)}$

- □ How to choose values of r and s?
  - Both S and SC are proportional to r & s.
  - Blocking probability P<sub>B</sub> is reversely proportional to r & s.
  - Strategy: Tradeoffs should be made between cost and quality of service.
    - The values of r & s should be as small as possible but give sufficient links to provide a reasonable grade of service to subscribers.

## Square two-stage networks

- Baseline networks
  - Square switching matrices are used as building blocks.
  - $p=r=s=q=N^{1/2}$
  - There are N<sup>1/2</sup> blocks, each block is a switching matrix of N<sup>1/2</sup>xN<sup>1/2</sup> inlets and outlets.
  - Switching elements: S=2NxN<sup>1/2</sup>
  - Switching capacity: SC=N
  - Support N simultaneous calls only if the traffic is uniformly distributed.

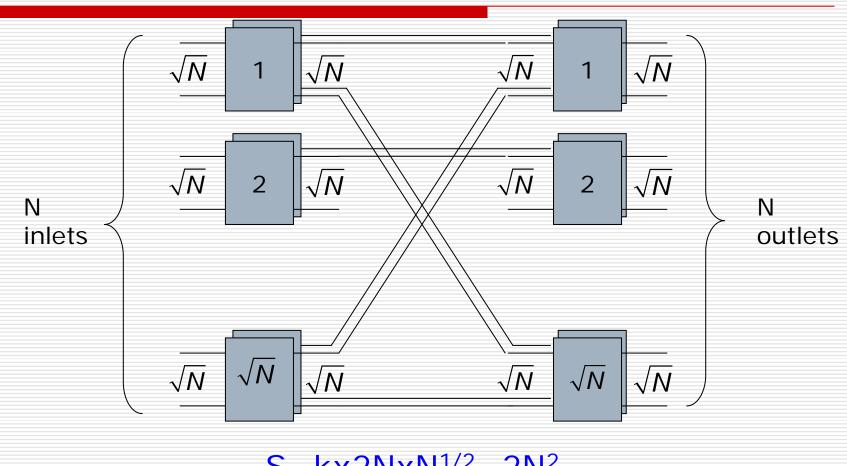
## Baseline networks



## Square two-stage networks

- Nonblocking networks
  - Why does blocking occur?
    - Only one link exists between a pair of first stage and second stage blocks.
  - How to reduce the probability of blocking?
    - Provide more links between the first stage and second stage blocks.
  - How many links should be provided?
    - A group of k=N<sup>1/2</sup> links should be provided for each pair of first stage and second stage blocks. S=2N<sup>2</sup>.
    - In comparison with single stage network, the number of switching elements is doubled.

# Nonblocking networks



 $S = kx2NxN^{1/2} = 2N^2$ 

## Remarks on two-stage networks

- A standard way of designing blocking networks with full availability.
- The expanding and concentrating network structures of small size can be implemented easily.
- Alternative paths for establishing a connection can be provided by two-stage networks, but with more switching elements in nonblocking design.

#### Question

- Is there any way to implement blocking & nonblocking networks with even less switching elements?
  - Solution: three-stage or multi-stage ~

## 4.8 Three-Stage networks

- General structure of an NxN threestage blocking network
  - Stage 1: p x s switching matrices
  - Stage 2: r x r switching matrices
  - Stage 3: s x p switching matrices
  - N=pxr, s is changable
  - Compared with a two-stage network, there are s alternative paths between a pair of inlet and outlet.

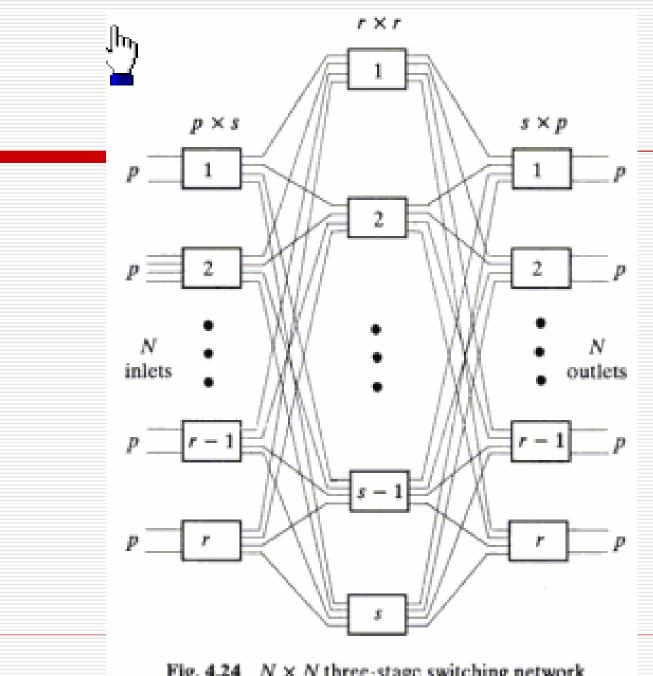


Fig. 4.24  $N \times N$  three-stage switching network.

## NxN three-stage blocking network

- Number of switching elements
  - = S=rps+sr<sup>2</sup>+spr=2Ns+sr<sup>2</sup>=s(2N+r<sup>2</sup>)
  - If square matrices are used in both the first and third stages, then p=s=N/r and S=2N<sup>2</sup>/r+Nr.
  - For a given value of N, there exists an optimal value of r which minimizes the value of S.
  - The optimal value of r is  $r=(2N)^{1/2}$  and the corresponding minimum of S is  $S_{min}=2N$   $(2N)^{1/2}$

$$p=N/r=(N/2)^{1/2}$$

How about the switching capacity?

## NxN three-stage blocking network

- Blocking probability analysis
  - Probability graph
    - Circle: stage
    - Line: link
    - A graph can be broken down into serial and parallel paths

# Inlet Solution Fig. 4.25 Lee's graph for a three-stage network.

- Notation
  - β: probability that a link is busy.
  - $\beta'=1-\beta$ : probability that a link is free.

## Blocking probability analysis

When s serial links complete a connection, the blocking probability is

$$P_B = 1 - (1 - \beta)^s$$



If there are s parallel links, the blocking probability is

$$P_B = \beta$$
 s

## Blocking probability analysis

- The blocking probability of NxN three-stage network
  - If  $\alpha$  denotes the probability that a inlet at the first stage is busy, then  $\beta = \alpha p/s = \alpha /k$ , where k=s/p
  - There are s parallel paths
  - Each path consists of two serial links
- □ Blocking probability of each path is  $1-(1-\beta)^2=1-(1-\alpha/k)^2$
- □ Total Blocking probability is

$$P_{R} = [1-(1-\beta)^{2}]^{s} = [1-(1-\alpha/k)^{2}]^{s}$$

What is the meaning of k?

# Probability graph

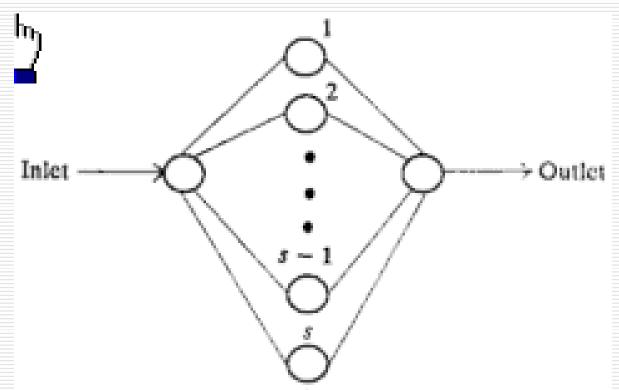


Fig. 4.25 Lee's graph for a three-stage network.

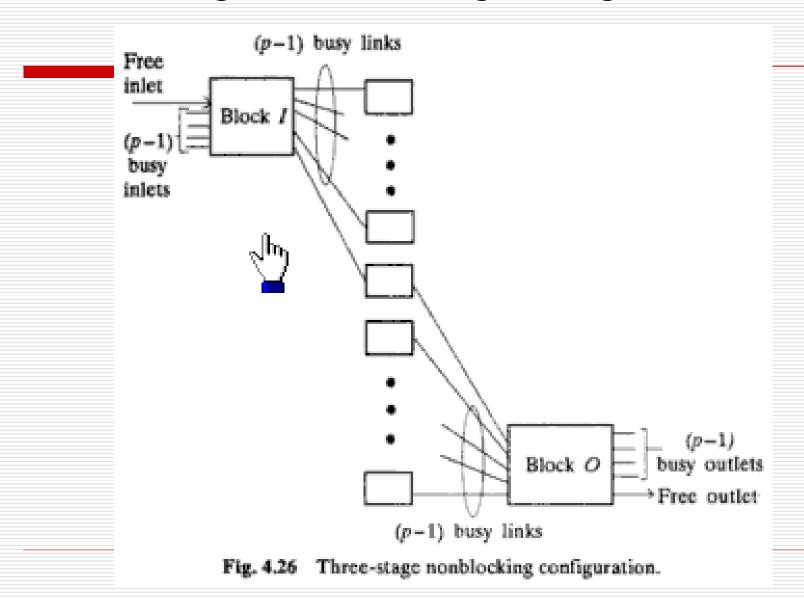
## Remarks on P<sub>B</sub> of three-stage ~

- □ The notation k=s/p represents either space expansion (k>1) or concentration (k<1).</p>
- $\square$  Low  $P_B$  requests small  $\alpha / k$ 
  - In case of small a, e.g. end offices and PBX, k can be small. It means concentration can be used in first stage.
  - In case of large α, e.g. transit exchanges, k should be large. It implies that expansion should be used in first stage.

## Nonblocking three-stage networks

- C. Clos: pioneer researcher of multistage nonblocking networks.
- Clos networks
  - Multistage nonblocking and fully availabile networks.
  - Much less switching elements are used than that in single stage networks.
- Design strategy
  - Providing adequate number of blocks in middle stages. For three-stage networks, the value of s should be large enough.

- Worst situation for blocking
  - p-1 inlets in a block I in the first stage are busy;
  - p-1 outlets in a block O in the third stage are busy;
  - The p-1 second-stage blocks, on which the p-1 outlets from block I are terminated on, are different from the p-1 second-stage blocks from which the links are established to the block O.
  - The free inlet of block I needs to be terminated on the free outlet of block O.
- How big should s be to avoid occurrence of blocking?



- □ The number of blocks required in the second stage for nonblocking operation is s=2(p-1)+1=2p-1.
- The number of switching elements in the nonblocking configuration is given by

$$S=p(2p-1)r+(2p-1)r^2+p(2p-1)r$$

$$=2N(2N/r-1)+r^2(2N/r-1)$$

$$=(4N^2-2Nr)/r+2Nr-r^2$$

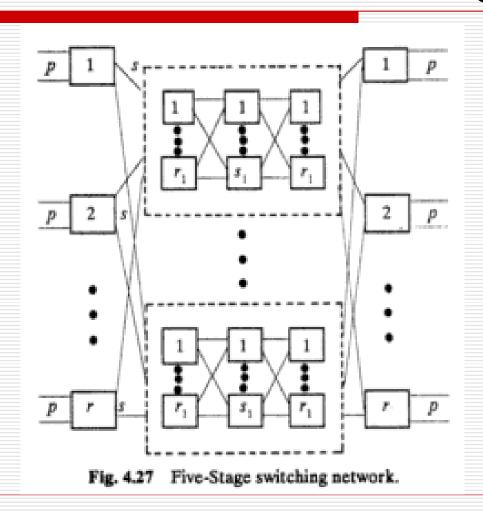
- There exists an optimal value of r for minimizing the value of S.
  - Let dS/dr=0, we have  $r^2(N-r)=2N^2$ .
  - For large values of N, we have  $N-r \approx N$ .
  - Hence,  $r=(2N)^{1/2}$ ,  $p=N/r=(N/2)^{1/2}$
  - The minimum S is  $S_{min} = 4N(2N)^{1/2}$
- Switching elements advantage ratio
- $= \frac{\text{S in nonblocking single-stage network, N}^2}{\text{S in nonblocking three-stage network, 4N(2N)}^{1/2}}$

See table 4.4 on page 95.

## 4.9 n-Stage networks

- ☐ Further reduction in the number of switching elements are possible by using even higher number of stages than three.
- Construction of multi-stage networks
  - By replacing the middle blocks with three-stage network blocks continually, any number of stages can be obtained.

# Construction of five-stage ~



# Assignments

- □ Ex.11
- □ Ex.15