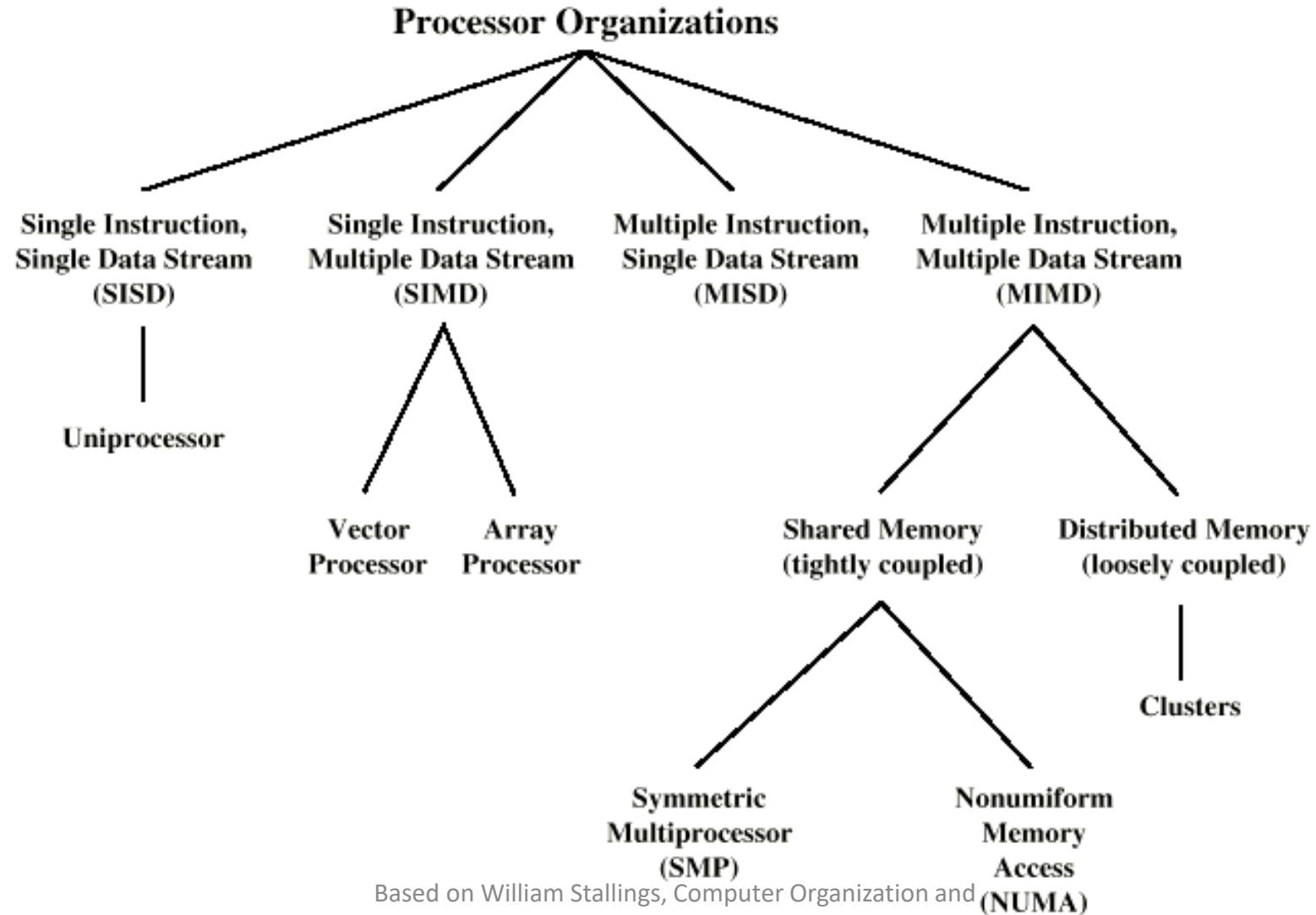


Parallel Processing

ICT 2203 Computer Architecture

Based on William Stallings, Computer Organization and Architecture, 8th Edition

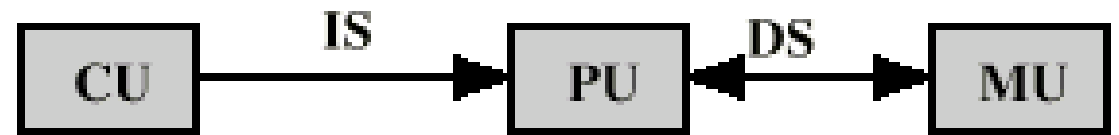
Taxonomy of Parallel Processor Architectures



Based on William Stallings, Computer Organization and Architecture, 8th Edition

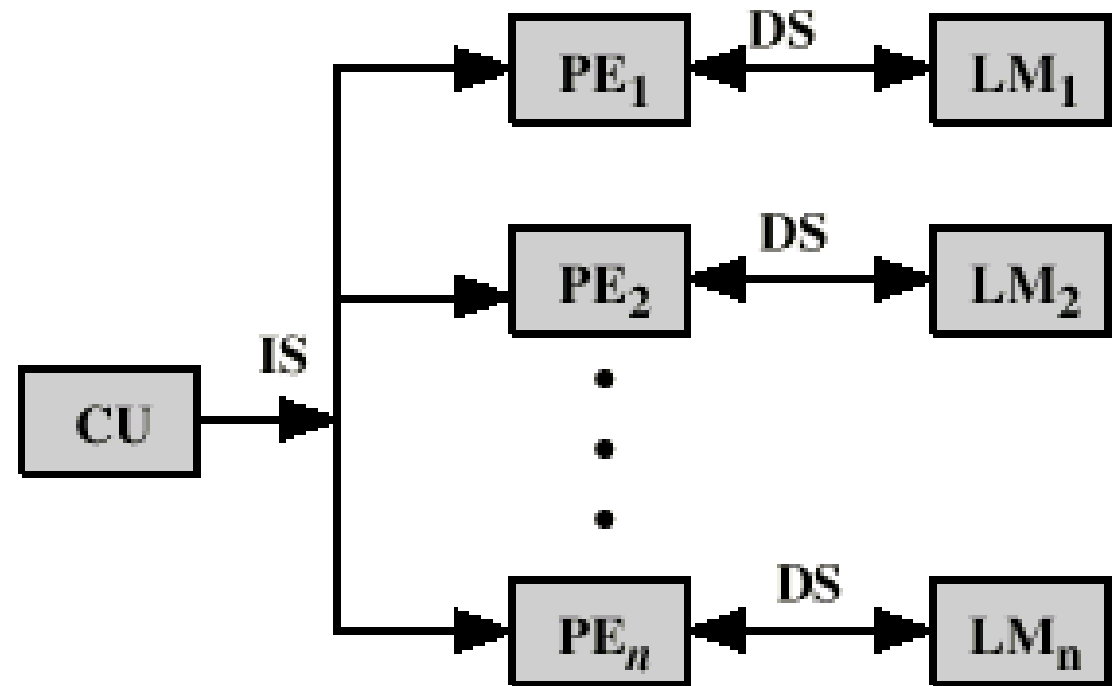
Single Instruction, Single Data Stream – SISD

- Single processor
- Single instruction stream
- Data stored in single memory
- Uni-processor



Single Instruction, Multiple Data Stream – SIMD

- Single machine instruction
- Controls simultaneous execution
- Number of processing elements
- Lockstep basis
- Each processing element has associated data memory
- Each instruction executed on different set of data by different processors



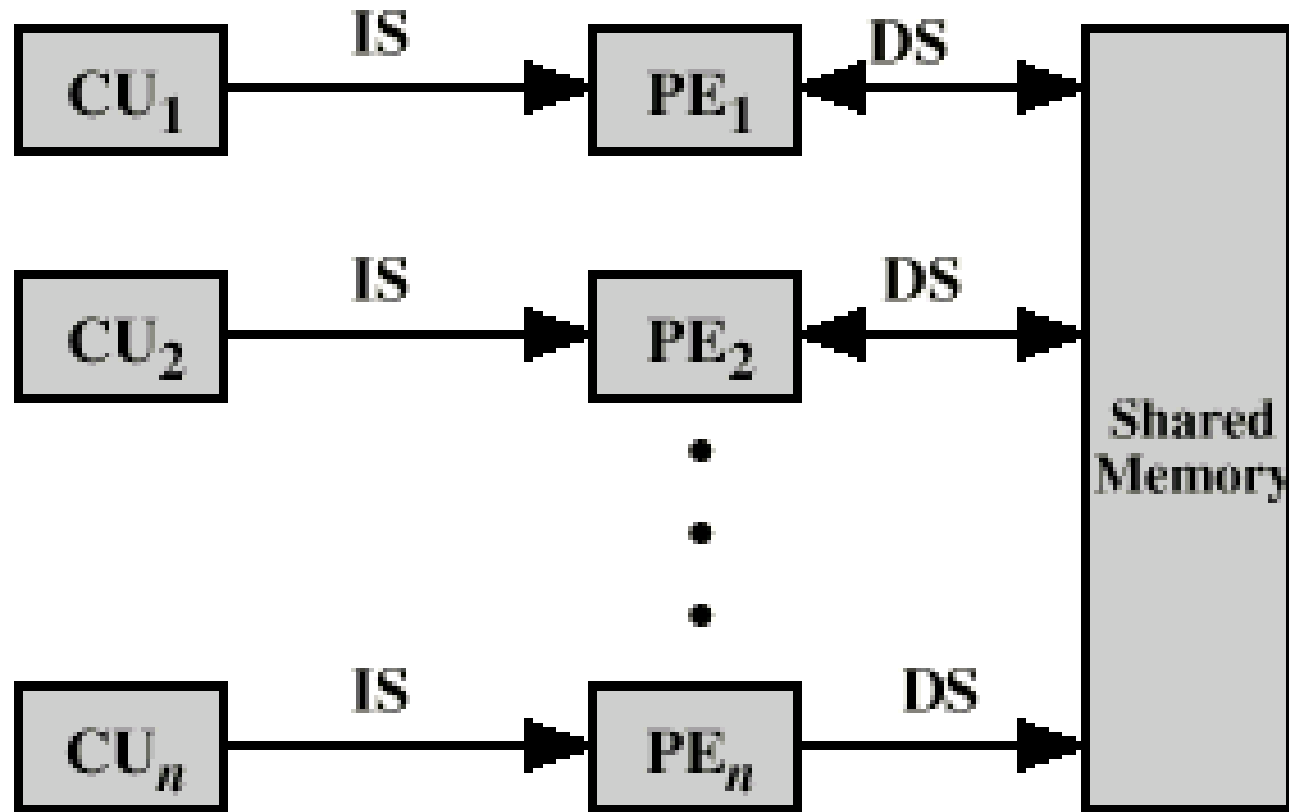
Multiple Instruction, Single Data Stream – MISD

- Sequence of data
- Transmitted to set of processors
- Each processor executes different instruction sequence
- **Never been implemented**

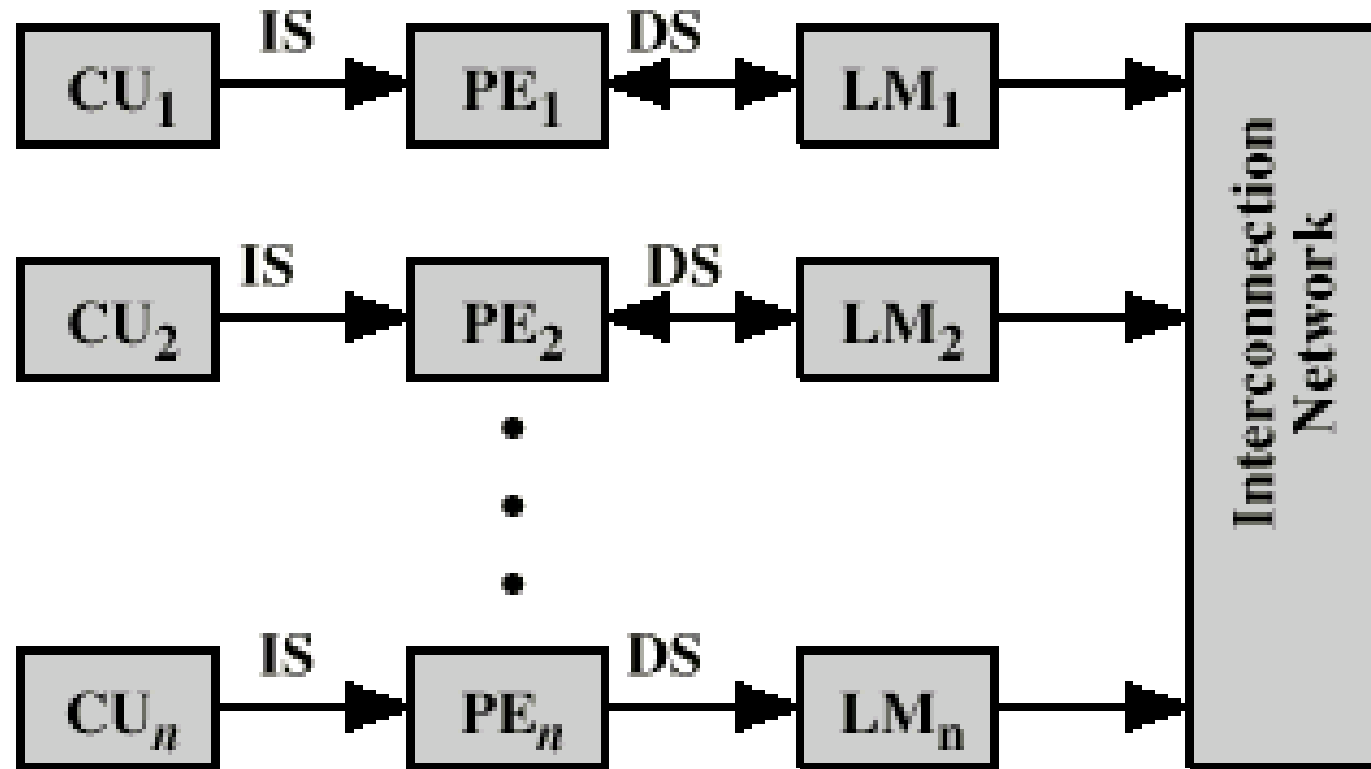
Multiple Instruction, Multiple Data Stream – MIMD

- Set of processors
- Simultaneously execute different instruction sequences
- Different sets of data
- Symmetric Multiprocessors (SMPs), clusters and NUMA systems

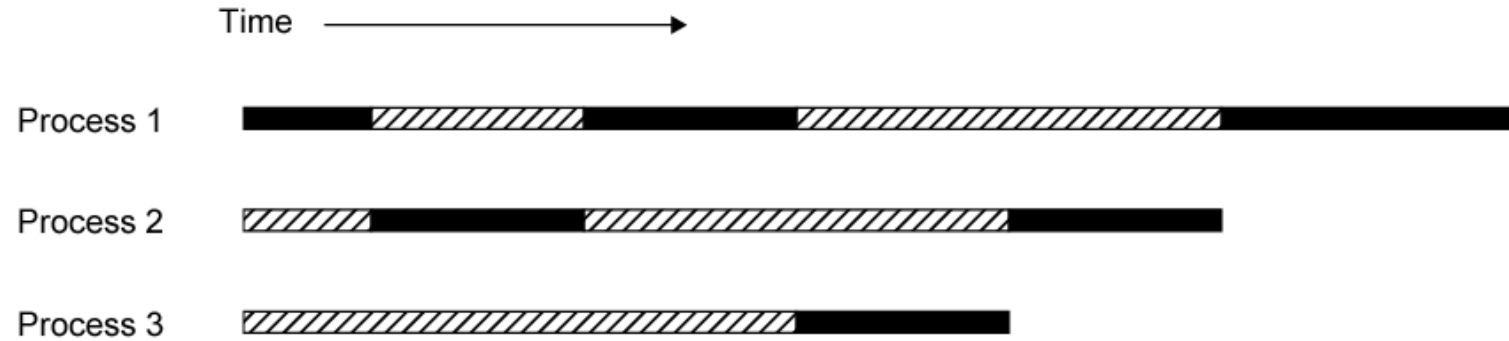
Parallel Organizations – MIMD Shared Memory



Parallel Organizations – MIMD Distributed Memory



Multiprogramming and Multiprocessing



(a) Interleaving (multiprogramming, one processor)



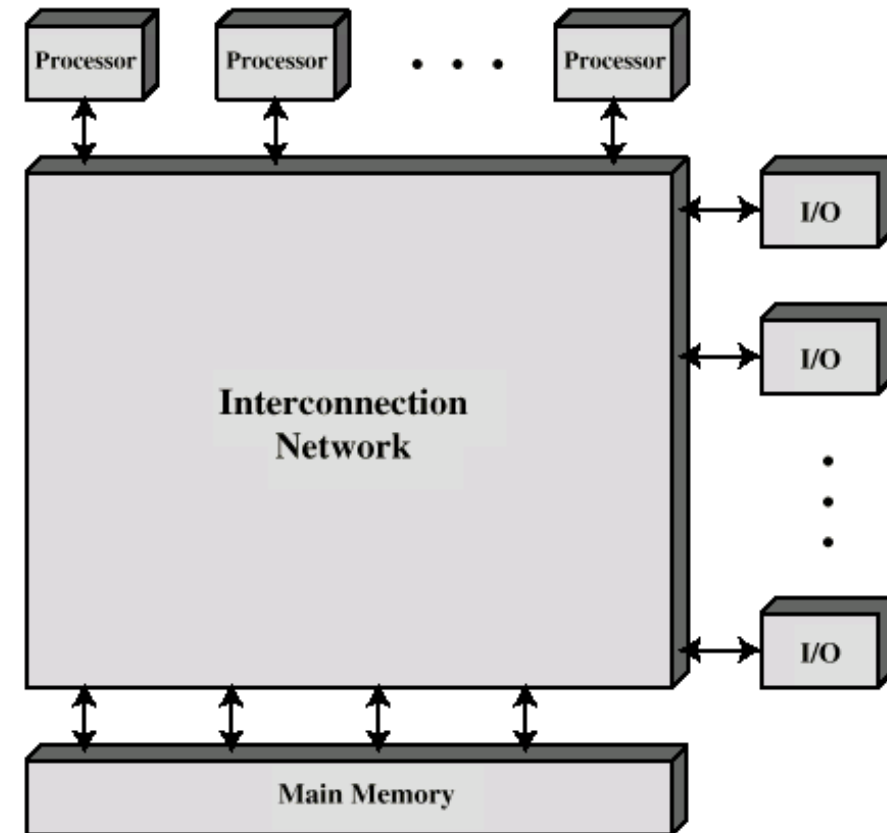
(b) Interleaving and overlapping (multiprocessing; multiple processors)

Operating System Issues

- Simultaneous concurrent processes
- Scheduling
- Synchronization
- Memory management
- Reliability and fault tolerance

Symmetric Multiprocessors

- A stand alone computer with the following characteristics
 - Two or more similar processors of comparable capacity
 - Processors share same memory and I/O
 - Processors are connected by a bus or other internal connection
 - Memory access time is approximately the same for each processor
 - All processors share access to I/O
 - All processors can perform the same functions (hence symmetric)
 - System controlled by integrated operating system



SMP Advantages

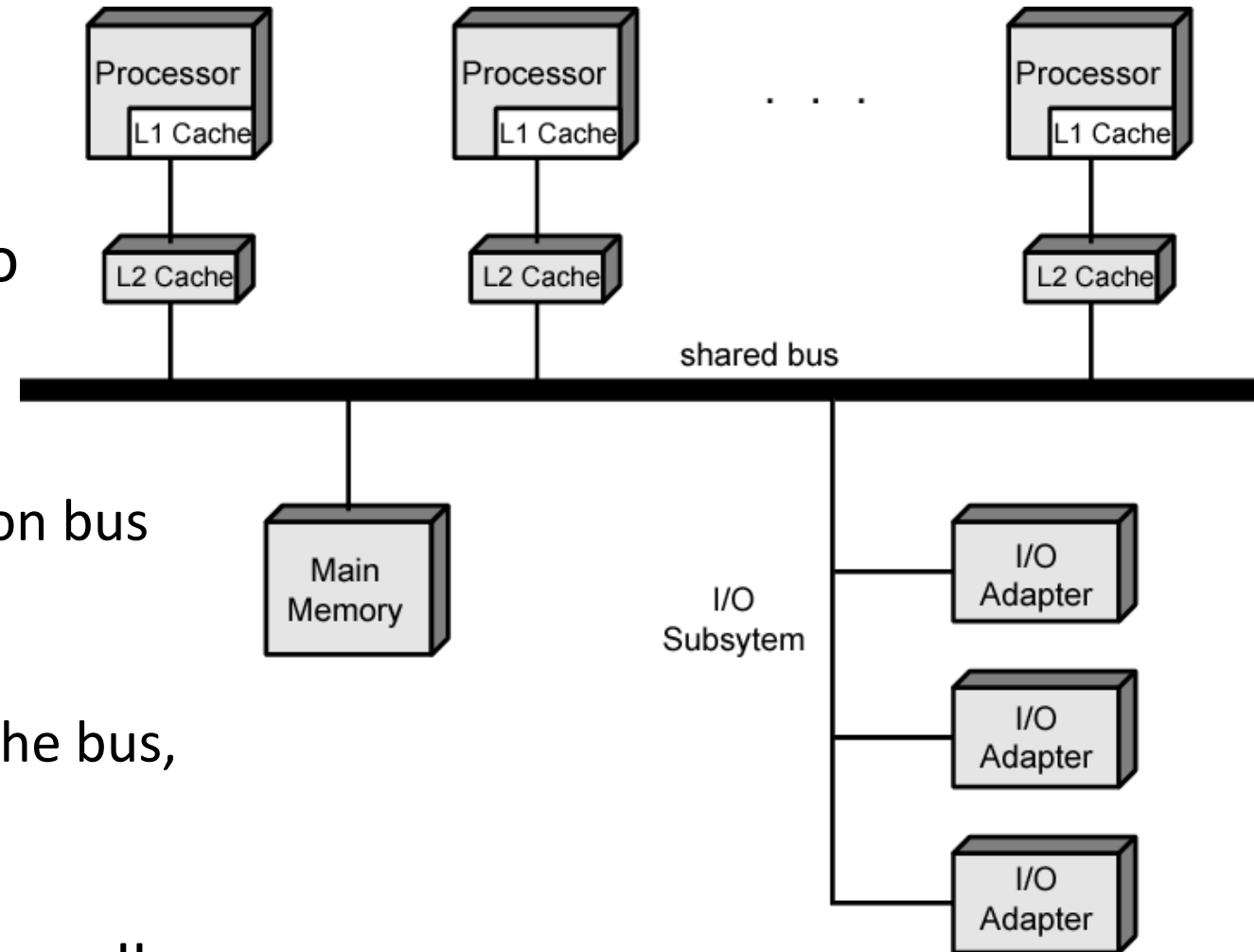
- Performance
 - If some work can be done in parallel
- Availability
 - Since all processors can perform the same functions, failure of a single processor does not halt the system
- Incremental growth
 - User can enhance performance by adding additional processors
- Scaling
 - Vendors can offer range of products based on number of processors

Organization Classification

- Time shared or common bus
- Multiport memory
- Central control unit

Time Shared Bus

- Simplest form
- Structure and interface similar to single processor system
- Following features provided
 - Addressing - distinguish modules on bus
 - Arbitration - any module can be temporary master
 - Time sharing - if one module has the bus, others must wait and may have to suspend
- Now have multiple processors as well as multiple I/O modules



Time Share Bus

- Advantage:

- Simplicity
- Flexibility
- Reliability

- Disadvantage:

- Performance limited by bus cycle time
- Each processor should have local cache
 - Reduce number of bus accesses
- Leads to problems with cache coherence
 - Solved in hardware - see later

Multithreading and Chip Multiprocessors

- Instruction stream divided into smaller streams (threads)
- Executed in parallel
- Wide variety of multithreading designs

Definitions of Threads and Processes

- Thread in multithreaded processors may or may not be same as software threads
- Process:
 - An instance of program running on computer
 - Resource ownership, Scheduling/execution, and Process switching
- Thread: dispatchable unit of work within process
 - Includes processor context (PC and stack pointer) and data area for stack
 - Thread executes sequentially, Interruptible: processor to another thread
- Thread switch
 - Switching processor between threads within same process
 - Typically less costly than process switch

Implicit and Explicit Multithreading

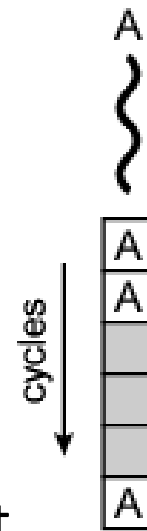
- All commercial processors and most experimental ones use explicit multithreading
 - Concurrently execute instructions from different explicit threads
 - Interleave instructions from different threads on shared pipelines or parallel execution on parallel pipelines
- Implicit multithreading is concurrent execution of multiple threads extracted from single sequential program
 - Implicit threads defined statically by compiler or dynamically by hardware

Approaches to Explicit Multithreading

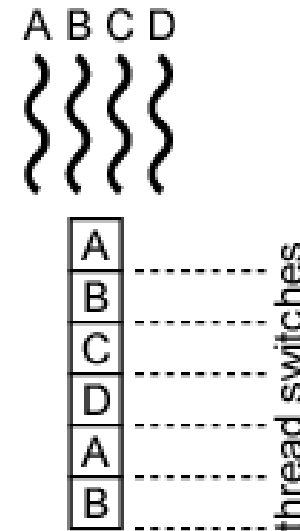
- Interleaved
 - Fine-grained
 - Processor deals with two or more thread contexts at a time
 - Switching thread at each clock cycle
 - If thread is blocked it is skipped
- Blocked
 - Coarse-grained
 - Thread executed until event causes delay
 - E.g. Cache miss
 - Effective on in-order processor
 - Avoids pipeline stall
- Simultaneous (SMT)
 - Instructions simultaneously issued from multiple threads to execution units of superscalar processor
- Chip multiprocessing
 - Processor is replicated on a single chip
 - Each processor handles separate threads

Scalar Processor Approaches

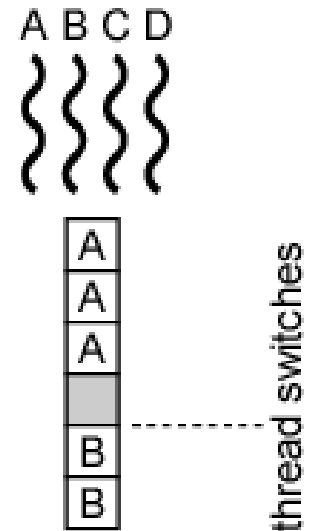
- Single-threaded scalar
 - Simple pipeline
 - No multithreading
- Interleaved multithreaded scalar
 - Easiest multithreading to implement
 - Switch threads at each clock cycle
 - Hardware needs to switch thread context between cycles
- Blocked multithreaded scalar
 - Thread executed until latency event occurs
 - Would stop pipeline
 - Processor switches to another thread



(a) single-threaded scalar



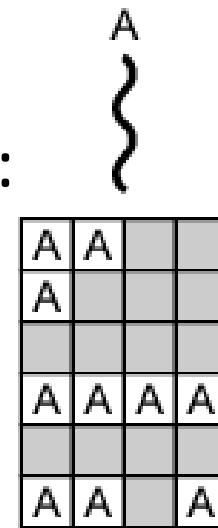
(b) interleaved multithreading scalar



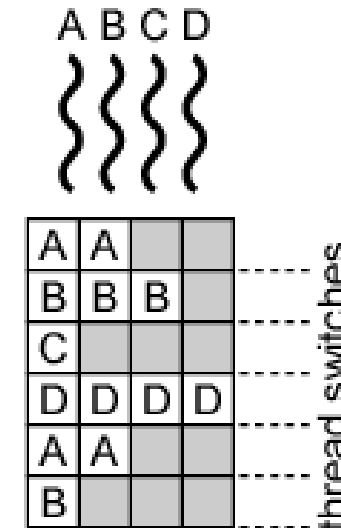
(c) blocked multithreading scalar

Multiple Instruction Issue Processors (1)

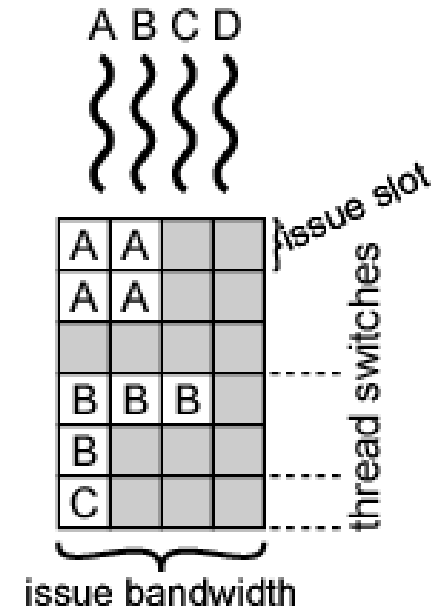
- Superscalar
 - No multithreading
- Interleaved multithreading superscalar:
 - Each cycle, as many instructions as possible issued from single thread
 - Delays due to thread switches eliminated
 - Number of instructions issued in cycle limited by dependencies
- Blocked multithreaded superscalar
 - Instructions from one thread
 - Blocked multithreading used



(d) superscalar



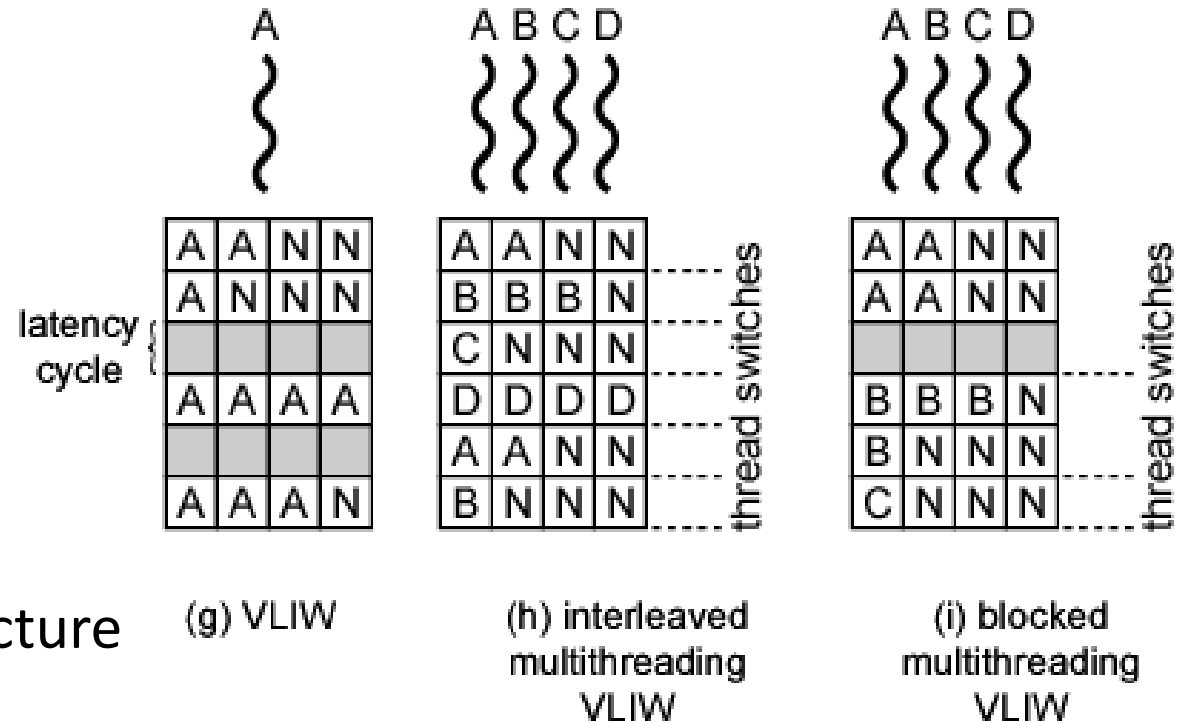
(e) interleaved multithreading superscalar



(f) blocked multithreading superscalar

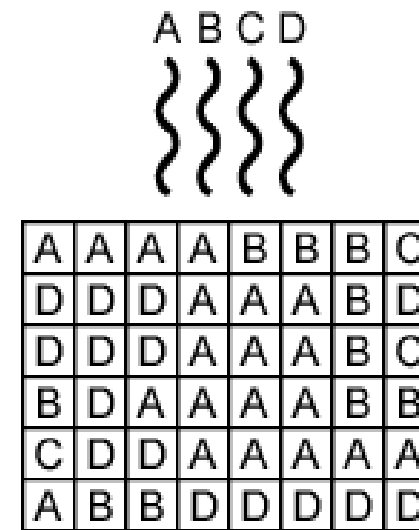
Multiple Instruction Issue Processors (2)

- Very long instruction word (VLIW)
 - Multiple instructions in single word
 - Typically constructed by compiler
 - Operations that may be executed in parallel in same word
 - May pad with no-ops
- Interleaved multithreading VLIW
 - Similar efficiencies to interleaved multithreading on superscalar architecture
- Blocked multithreaded VLIW
 - Similar efficiencies to blocked multithreading on superscalar architecture

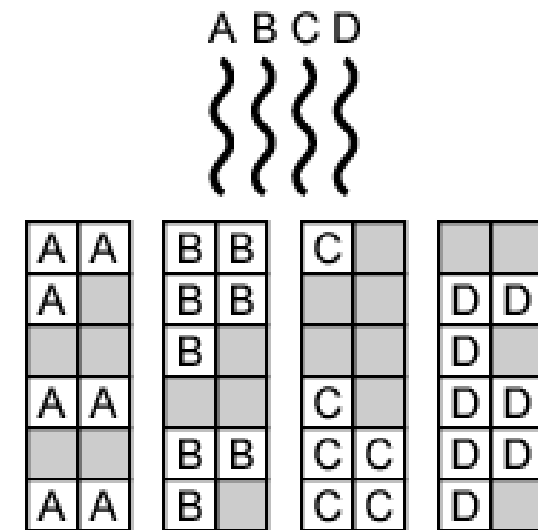


Parallel, Simultaneous Execution of Multiple Threads

- Simultaneous multithreading
 - Issue multiple instructions at a time
 - One thread may fill all horizontal slots
 - Instructions from two or more threads may be issued
 - With enough threads, can issue maximum number of instructions on each cycle
- Chip multiprocessor
 - Multiple processors
 - Each has two-issue superscalar processor
 - Each processor is assigned thread
 - Can issue up to two instructions per cycle per thread



(j) simultaneous multithreading (SMT)



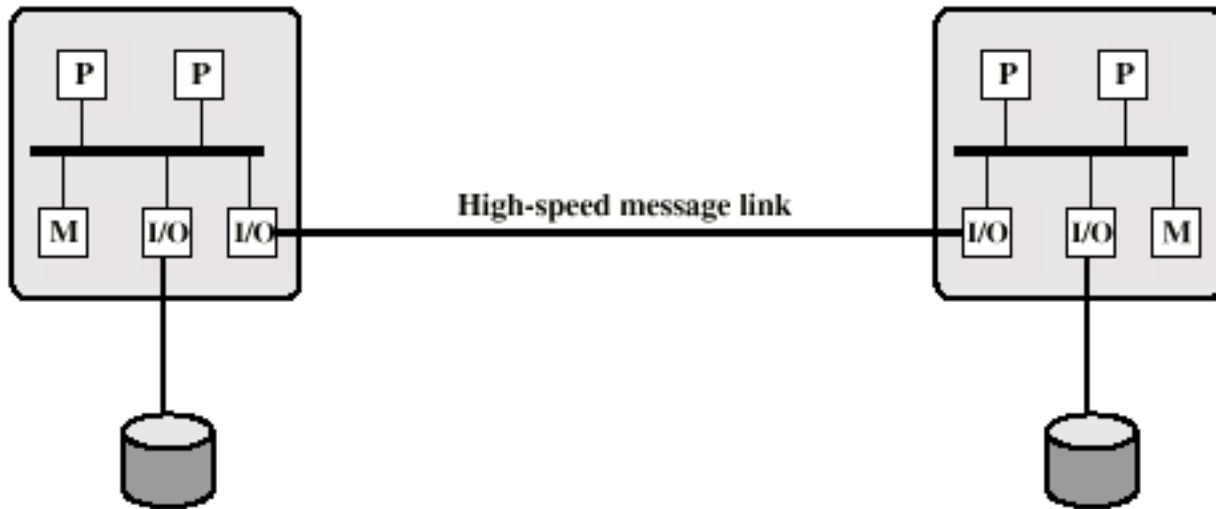
(k) chip multiprocessor

Clusters

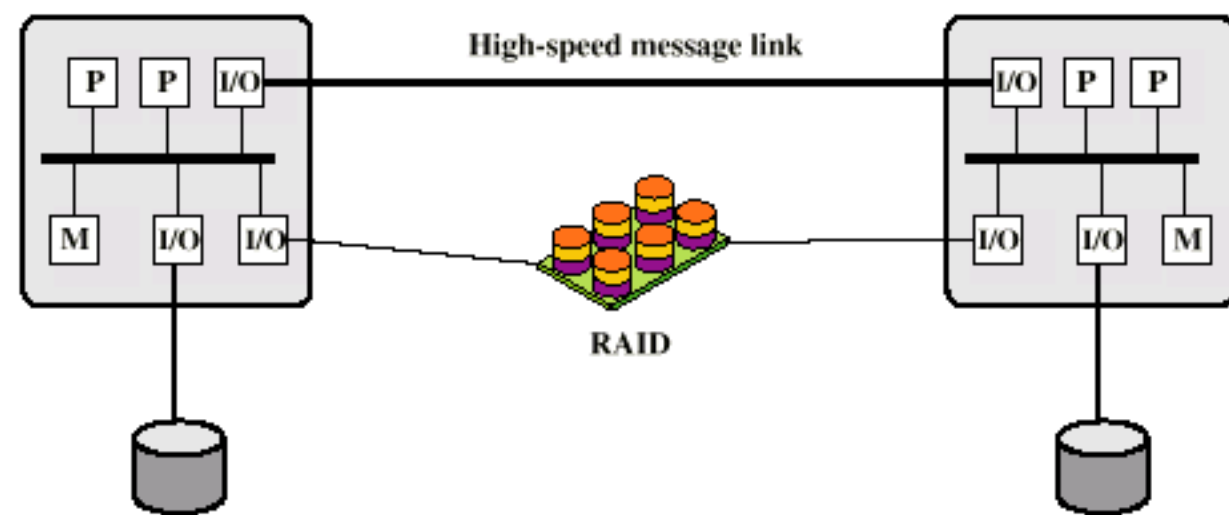
- Alternative to SMP
 - High performance
 - High availability
 - Server applications
 - A group of interconnected computers
 - Working together as unified resource
 - Illusion of being one machine
 - Each computer called a node
- Cluster Benefits:
 - Absolute scalability
 - Incremental scalability
 - High availability
 - Superior price/performance

Cluster Configurations

Standby Server, No Shared Disk



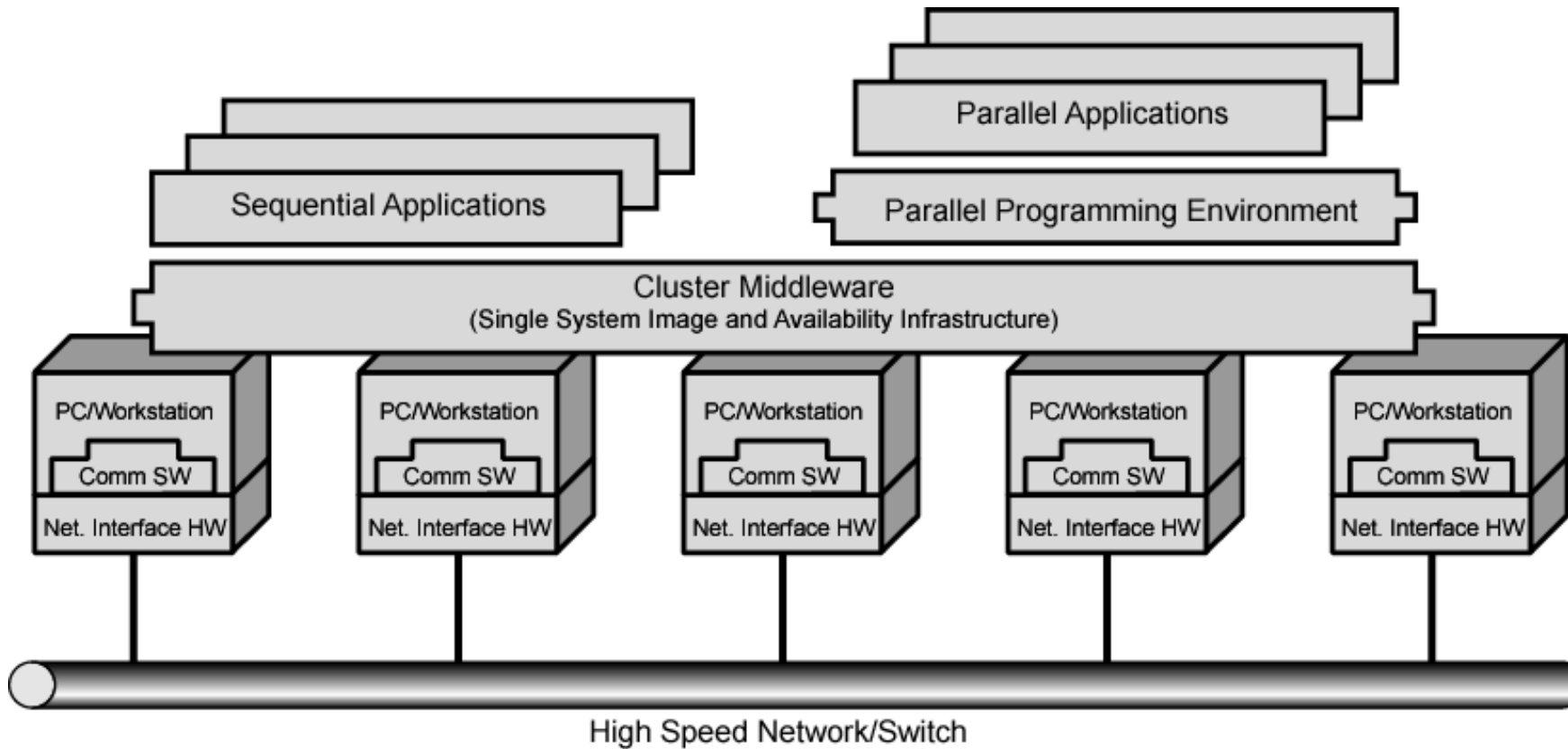
Shared Disk



Operating Systems Design Issues

- Failure Management
 - High availability
 - Fault tolerant
 - Failover
 - Switching applications & data from failed system to alternative within cluster
 - Failback
 - Restoration of applications and data to original system, after problem is fixed
- Load balancing
 - Incremental scalability
 - Automatically include new computers in scheduling
 - Middleware needs to recognise that processes may switch between machines

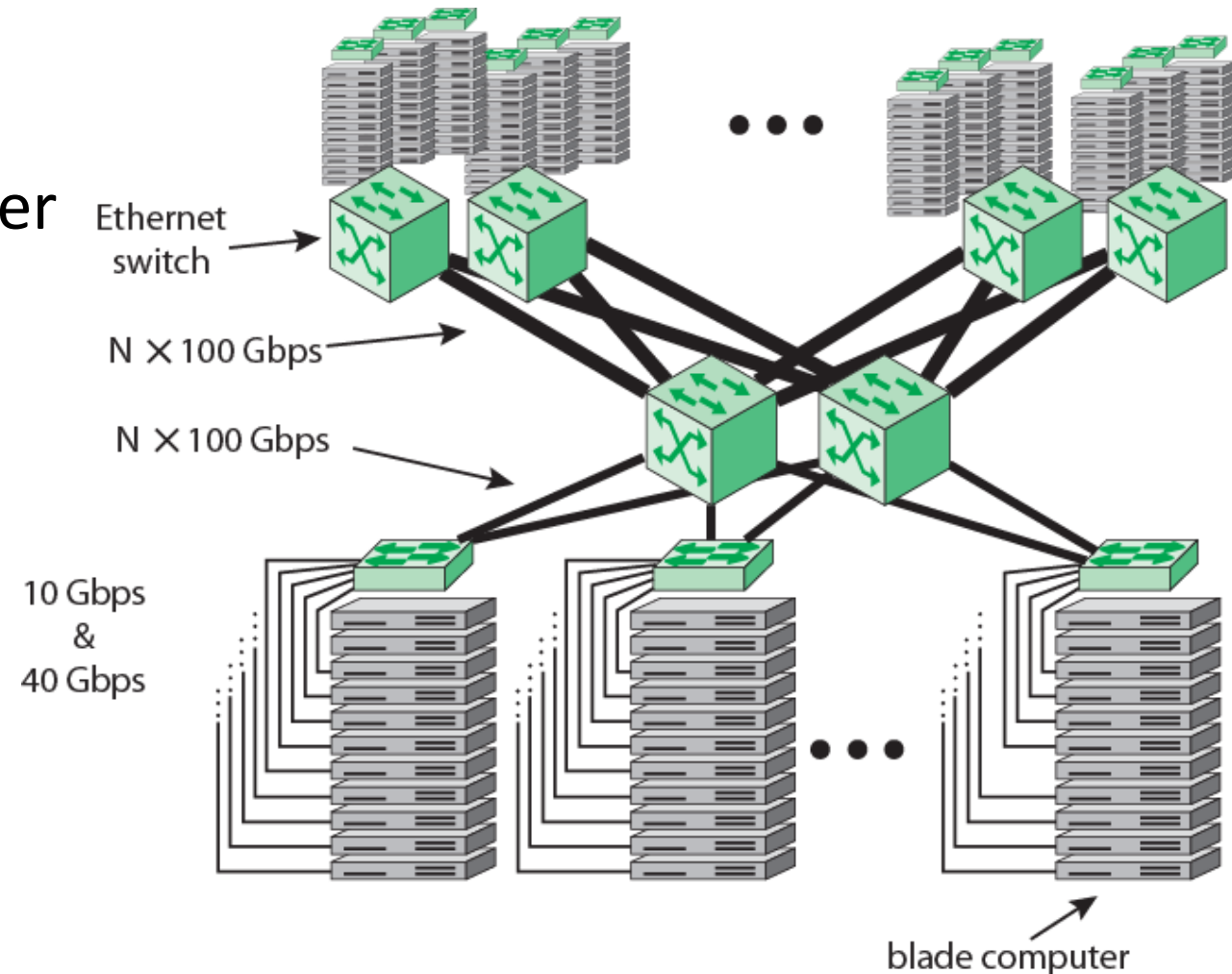
Cluster Computer Architecture



- Single system image
- Single point of entry
- Single file hierarchy
- Single control point
- Single virtual networking
- Single memory space
- Single job management system
- Single user interface
- Single I/O space
- Single process space
- Checkpointing
- Process migration

Blade Servers

- Common implementation of cluster
- Server houses multiple server modules (blades) in single chassis
 - Save space
 - Improve system management
 - Chassis provides power supply
 - Each blade has processor, memory, disk



Example: 100-Gbps Ethernet Configuration for a Massive Blade Server Site

Cluster v. Symmetric Multiprocessors

- Both provide multiprocessor support to high demand applications.
- Both available commercially
 - SMP for longer
- SMP:
 - Easier to manage and control
 - Closer to single processor systems
 - Scheduling is main difference
 - Less physical space
 - Lower power consumption
- Clustering:
 - Superior incremental & absolute scalability
 - Superior availability
 - Redundancy

Next:

Multicore Computers