Flynn's Classification of Computer Architectures (Derived from Michael Flynn, 1972)

IS



(a) SISD Uniprocessor Architecture

Captions:

- CU Control Unit ; PU Processing Unit
- MU Memory Unit ; IS Instr
- **DS Date Stream**

IS – Instruction Stream

Flynn's Classification of Computer Architectures (Derived from Michael Flynn, 1972) (contd...)



(b) SIMD Architecture (with Distributed Memory)

Captions:

- **CU Control Unit PU - Processing Unit** ;
- **MU Memory Unit IS - Instruction Stream** ;
- **DS Date Stream** ;
- LM Local Memory

- **PE Processing Element**

Flynn's Classification of Computer Architectures

(Derived from Michael Flynn, 1972) (contd...)



(c) MIMD Architecture (with Shared Memory)

Captions:

- CU Control Unit;PU Processing UnitMU Memory Unit;IS Instruction Stream
- DS Date Stream ; PE Processing Element
- LM Local Memory

Flynn's Classification of Computer Architectures

(Derived from Michael Flynn, 1972) (contd...)



(d) MISD Architecture (the Systolic Array)

Captions:

- CU Control Unit ; PU Processing Unit
- MU Memory Unit ; IS Instruction Stream
- DS Date Stream ; PE Processing Element
- LM Local Memory

Two Approaches to Parallel Programming

a) Implicit Parallelism

Source code written in sequential languages (C, Fortran, Lisp or Pascal)

Parallelizing Compiler produces Parallel Object Code

b) Explicit Parallelism

Source code written in **concurrent** dialects of C, Fortran, Lisp or Pascal

Concurreny preserving compiler produces concurrent Object Code

SIMD and MIMD

SIMDSs appeal more to special purpose applications.SIMDs are not size-scalable.

≻Thinking Machine CM2.

➢MIMDs with distributed memory having globally shared virtual address space is the future trend.

≻CM5 has MIMD architecture

Bell's Taxonomy of MIMD computers



Two Categories of Parallel Computers

- 1. Shared Memory Multiprocessors (tightly coupled systems
- 2. Message Passing Multicomputers

SHARED MEMORY MULTIPROCESSOR MODELS:

- a. Uniform Memory Access (UMA)
- b. Non-Uniform Memory Access (NUMA)
- c. Cache-Only Memory Architecture (COMA)

SHARED MEMORY MULTIPROCESSOR MODELS



The UMA multiprocessor model (e.g., the Sequent Symmetry S-81)

SHARED MEMORY MULTIPROCESSOR MODELS (contd...)



(a) Shared local Memories (e.g., the BBN Butterfly)

NUMA Models for Multiprocessor Systems

SHARED MEMORY MULTIPROCESSOR MODELS (contd...)



(b) A hierarchical cluster model (e.g., the Cedar system at the University of Illinois)

NUMA Models for Multiprocessor Systems

SHARED MEMORY MULTIPROCESSOR MODELS (contd...)



- P: Processor
- C : Cache
- D : Directory

The COMA Model of a multiprocessor (e.g., the KSR-1)

Generic Model of a message-passing multicomputer



e.g., Intel Paragon, nCUBE/2

Important issues: Message Routing Scheme, Network flow control strategies, dead lock avoidance, virtual channels, message-passing primitives, program decomposition techniques.

Theoretical Models for Parallel Computers

- RAM Random Access Machines e.g., conventional uniprocessor computer
- PRAM Parallel Random Access Machines model developed by Fortune & Wyllie(1978) ideal computer with zero synchronization and zero memory access overhead For shared memory machine

PRAM-Variants

depending on how memory read & write are handled



The Architecture of a Vector Supercomputer (contd)

e.g., Convex C3800 8 processors 2G FLOPS peak VAX 9000 125-500 MFLOP CRAY YMP&C90 built with ECL 10K ICS 16 G FLOP

Example for SIMD machines

- MasPar MP-1; 1024 to 16 K RISC processors
- CM-2 from Thinking Machines, bitslice, 65K PE
- DAP 600 from Active memory Tech., bitslice

STATIC Connection Networks



Linear Array



Fully connected Ring



Star



Ring



The Channel width of Fat Tree increases as we ascend from leaves to root. This concept is used in CM5 connection Machine.

Binary Tree







Mesh

Torus

Systolic Array

Degree = t



3-cube



A 4 dimentional cube formed with 3D cubes

Binary Hypercube has been a popular architecture. Binary tree, mesh etc can be embedded in the hypercube.

<u>But:</u> Poor scalability and implementing difficulty for higher dimensional hypercubes.

CM2 – implements hypercube CM5 – Fat tree Intel IPSC/1, IPSC/2 are hypercubes Intel Paragon – 2D mesh

The bottom line for an architecture to survive in future systems is packaging efficiency and scalability to allow modular growth.

New Technologies for Parallel Processing

At present advanced CISC processors are used.

In the next 5 years RISC chips with multiprocessing capabilities will be used for Parallel Computer Design.

Two promising technologies for the next decade : Neural network Optical computing

Neural networks consist of many simple neurons or processors that have densely parallel interconnections.

Journals/Publications of interests in Computer Architecture

- Journal of Parallel & Distributed Computing (Acad. Press, 83-)
- Journal of Parallel Computing (North Holland, 84-)
- IEEE Trans of Parallel & Distributed Systems (90-)
- International Conference Parallel Processing (Penn State Univ, 72-)
- Int. Symp Computer Architecture (IEEE 72-)
- Symp. On Frontiers of Massively Parallel Computation (86-)
- Int Conf Supercomputing (ACM, 87-)
- Symp on Architectural Support for Programming Language and Operating Systems (ACM, 75-)
- Symp. On Parallel Algorithms & Architectures (ACM, 89-)
- Int Parallel Processing Sympo (IEEE Comp. Society 86-)
- IEEE Symp on Parallel & Distributed processing (89-)
- Parallel Processing Technology (?) IEEE Magazine

Digital 21064 Microprocessor - ALPHA

- Full 64 bit Alpha architecture, Advanced RISC optmized for high performance, multiprocessor support, IEEE/VAX floating point
- PAL code Privilieged Architecture Library
 - Optimization for multiple operating system VMS/OSF1
 - Flexible memory management
 - Multi-instruction atomic sequences
 - Dual pipelined architecture
 - 150/180 MHz cycle time
 - 300 MIPS
- 64 or 128 bit data width
 - 75 MHz to 18.75 MHz bus speed
- Pipelined floating point unit
- 8k data cache; 8k instruction cache
- + external cache
- 2 instructions per CPU cycle
- CMOS 4 VLSI, .75 micron, 1.68 million transistors
- 32 floating point registers; 32 integer registers, 32 bit fixed length instruction set
- 300 MIPS & 150 MFLOPS

MIMD BUS



MIMD BUS

- Standards :
 - Intel MULTBUS II
 - Motorola VME
 - Texas Instrument NU BUS
 - IEEE)896 FUTURE BUS
- BUS LATENCY
 - The time for bus and memory to complete a memory access
 - Tiem to acquire BUS + memory read or write time including Parity check, error correction etc.

Hierarchical Caching



Multiprocessor Systems

- 3 types of interconnection between processors:
 - Time shared common bus fig a
 - CROSS-BAR switch network fig b
 - Multiport memory fig c



Fig a – Time shared common bus



Fig b – CROSS BAR switch network



Fig c – Multiport Memory