

High Performance Computing

Authors :

Manas Ranjan Sahu, Divyen Patel

[AES Institute of Computer Studies]

26 / FEB / 2005

Topics

- Introduction
- Goals of HPC
- Applications of HPC
- HPC architecture
- HPC Programming concept
- HPC Software

Introduction to HPC

High

Performance

Computing

To achieve more
Than the
normal

Operational ability
in terms of
throughput

Actual Operation
or execution of
An algorithm



In simple term HPC is a computing system that provides more computing performance, power or resources than is generally available.

Why HPC ?

- Traditionally universities, research organizations and large companies use super computer to achieve more computing power.
- Super computers are most expensive machines.
- Small organization can not afford this.
- However most of the organizations need more computing power to solve business problems.
- Right Solution is HPC that bridges this gap.



How HPC ?



Super
Computer

=

Performance

>

Cost



HPC Cluster made up to
cheap workstations

How HPC ?

- The main area of this discipline is developing parallel processing algorithms and software programs that can be divided in to little pieces.
- Each pieces can executed on separate processors.

Goals of HPC

- Increase Processing speed.
- Reduce cost
- Increase efficiency and achieve scalability.

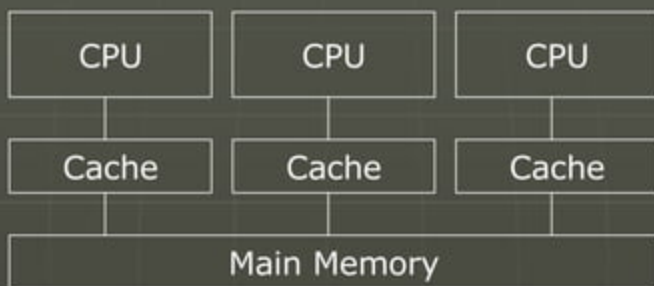
Applications of HPC

- Geographic Information System
- Bio-informatics and drug discovery
- Weather forecasting
- Statistical analysis and mathematical modeling
- Online transaction processing
- Virtual reality
- Search engine
- Other mission critical applications.

HPC Architecture

- HPC Systems can be broadly classified in to
 - Shared Memory Systems
 - Distribute Memory Systems
 - Hybrid Systems

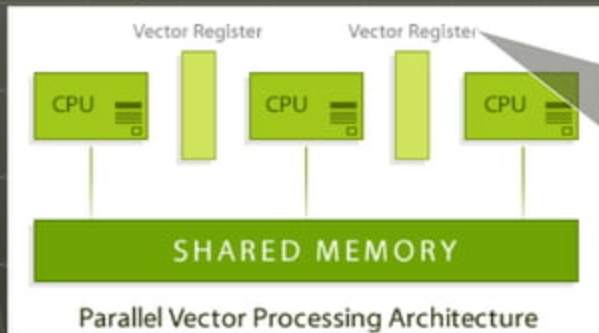
Shared Memory Systems



- Example: Symmetric Multi Processing (SMP) systems like Compaq Alpha Server ES 40
- Operate on single shared bank of memory.
- Operating System manages the allocation of processor time to applications.
- Parallel nature of machine is generally hidden from user.

Shared Memory Systems

- Special case of SMP architecture is Parallel vector processing.



Specially designed memory that is filled and passed along a vector instruction pipeline and returning results much faster than traditional scalar processors.



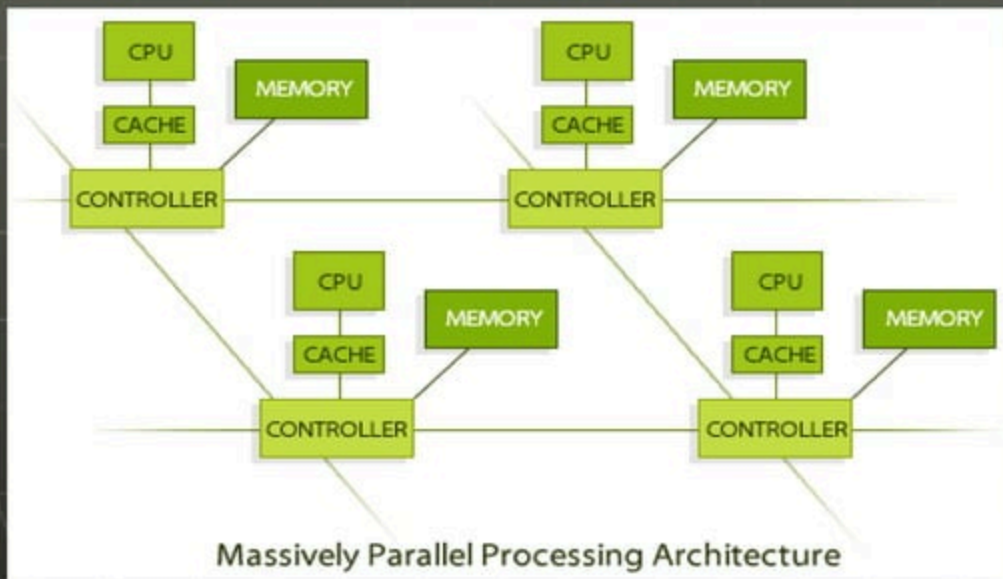
Ability to apply Instruction set to CPU at once rather than single instruction at a time.

Distributed Memory Systems

- It include
 - Massively Parallel Processing (MPP) systems
 - Clustered systems.
- Massively parallel processing architecture
 - It employs several processing nodes consisting of CPU and memory connected together using high speed interconnects.
 - Loosely coupled in nature.
 - Allows to scale up to tens of thousands of CPU.

Distributed Memory Systems

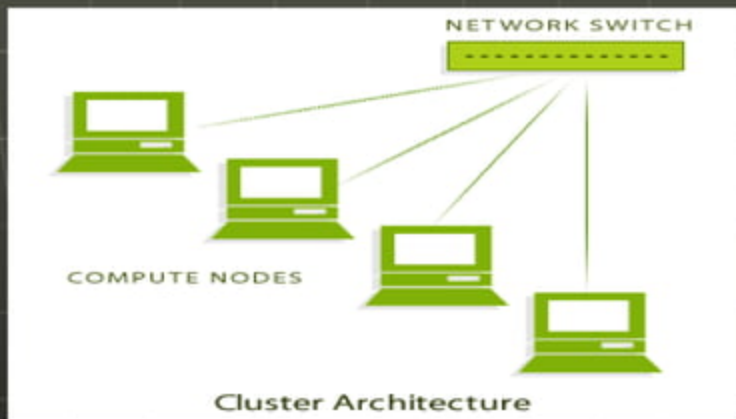
- Cont ...



Distributed Memory Systems

■ Cluster

- Concept is to tie together a lot of computing elements.
- Primary goals are economies of scale and entry level cost.



Distributed Memory Systems

■ Cont ...

- The components used to build HPC clusters include

Commercially off the shelf (COTS) nodes.

Cluster interconnects like Ethernet, Myrinet or ServerNet and

- Operating systems like Linux.

- Cluster management software is an important component to manage the cluster as a single system.

Hybrid Systems

Distributed Memory
Architecture



Shared Memory
Architecture



Non-Uniform Memory Access (NUMA) is an example of Hybrid System

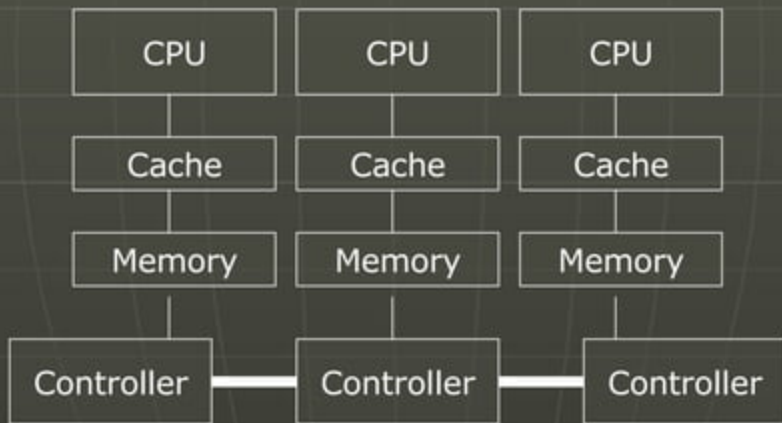
Why Hybrid Systems ?

- Two disadvantage of the SMP shared memory model are
 - Processors must communicate with a single bank of memory, thus gives rise to the danger of memory bottlenecks.
 - Single memory bus to connect the processors with the memory.
 - Adding more processors quickly saturates the bus capacity and kills scalability.

Hybrid Model

- In this Model

- Memory is physically distributed and
- Hardware Controller maintain single shared memory image at the user level.



HPC Programming Concepts

- Vectorisation

The software technique, in which vector arrays are used for partial calculation as an intermediate process.

- Threads and Multithreading

An old concept, but an excellent way to program today's Symmetric Multi-Processors

- Data Decomposition

The divide-and-conquer parallel approach to large data applications.

- Task Based Parallelism

Scalable client-server parallelism for high throughput.

HPC Software

- We take a look at some major software techniques and standards in two key areas of parallel HPC :
 - Data Parallelism
 - Message Passing

HPC Software – Data Parallelism

- Successful when the operations performed on the distributed data subsets are largely independent of each other.
- Easier for the programmer.
- parallelism is applied by the program compiler rather than by the programmer.
- Two important technology fall under this heading
 - HPF – High Performance Fortran
 - Shared memory Programming

HPC Software – Message Passing

- It is an inter-processor communication standards for flexible, efficient applications.
- The programmer has to understand and code by hand the necessary data communications between processors.
- More difficult to write the programs.
- Advantage is the full control over application and detailed knowledge of its ins and outs



Human programmers can almost always write faster and more efficient parallel programs than automatic compilers can generate.

HPC Software – Message Passing

- Cont ...
- Two most important Message Passing systems are

Message Passing Interface (MPI)

Parallel Virtual Machine (PVM)

Conclusion

- Science and technology plays an important role in improving the quality of life.
- The interests in extending capability of workstations to High Performance Computer System will definitely improve quality of life.

Thank You