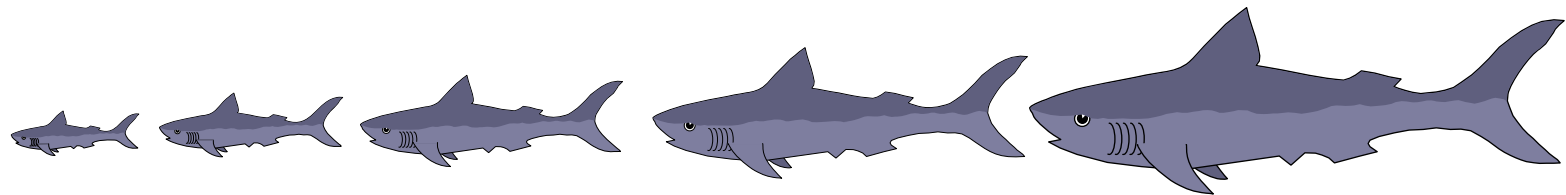


High Performance Computing Course Notes

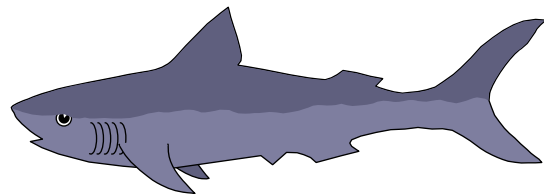
Cluster Technologies



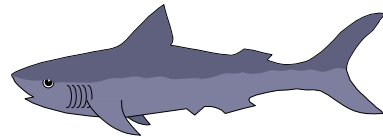
Food Chain Picture



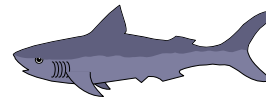
1980s Computer Food Chain



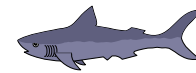
Mainframe



Vector Supercomputer



Mini Computer

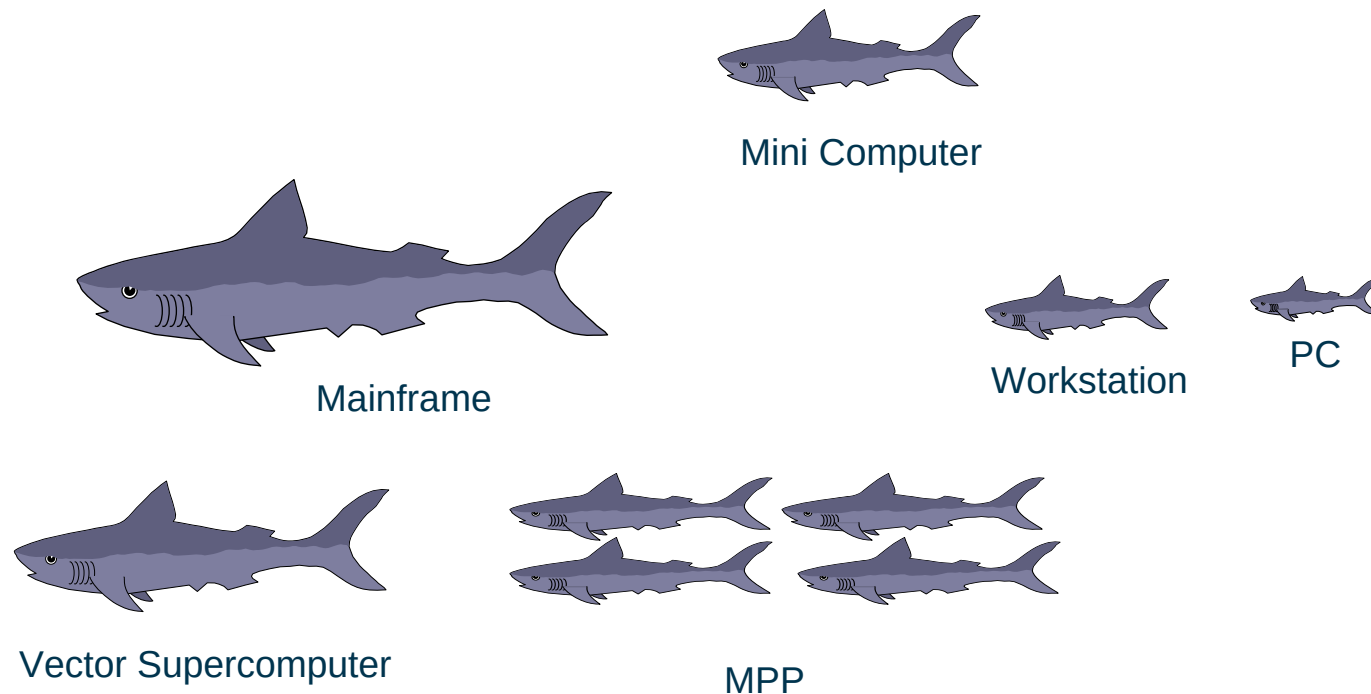


Workstation

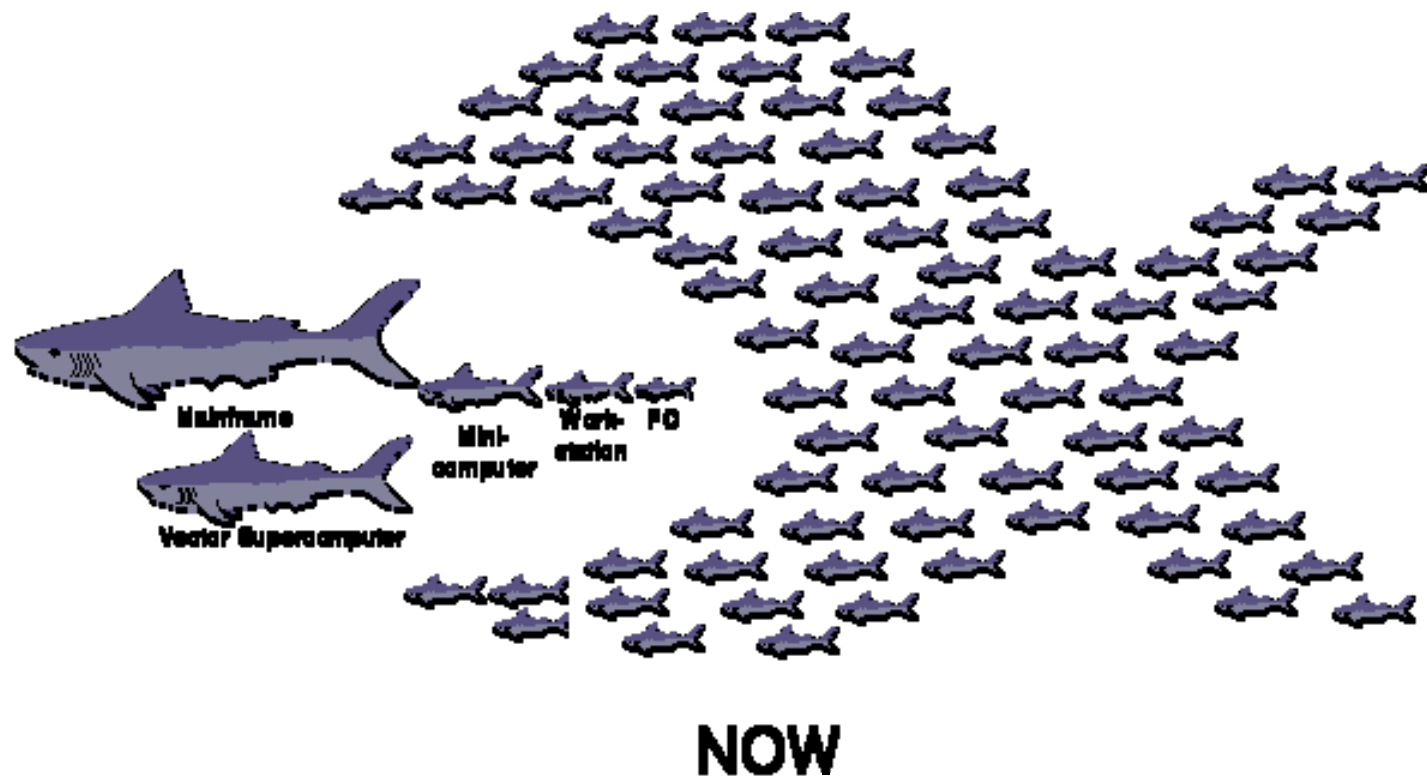


PC

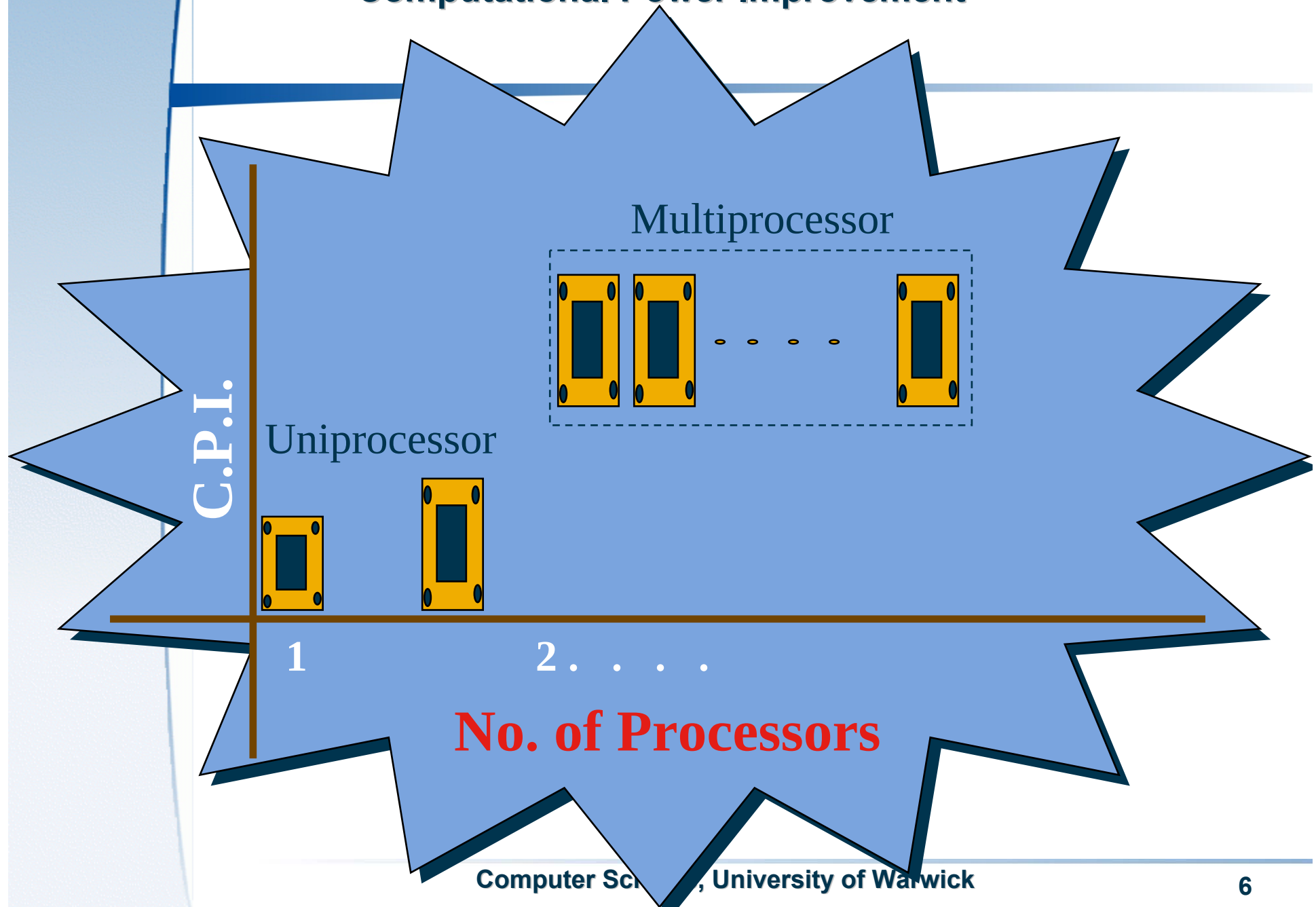
Early 1990s Computer Food Chain



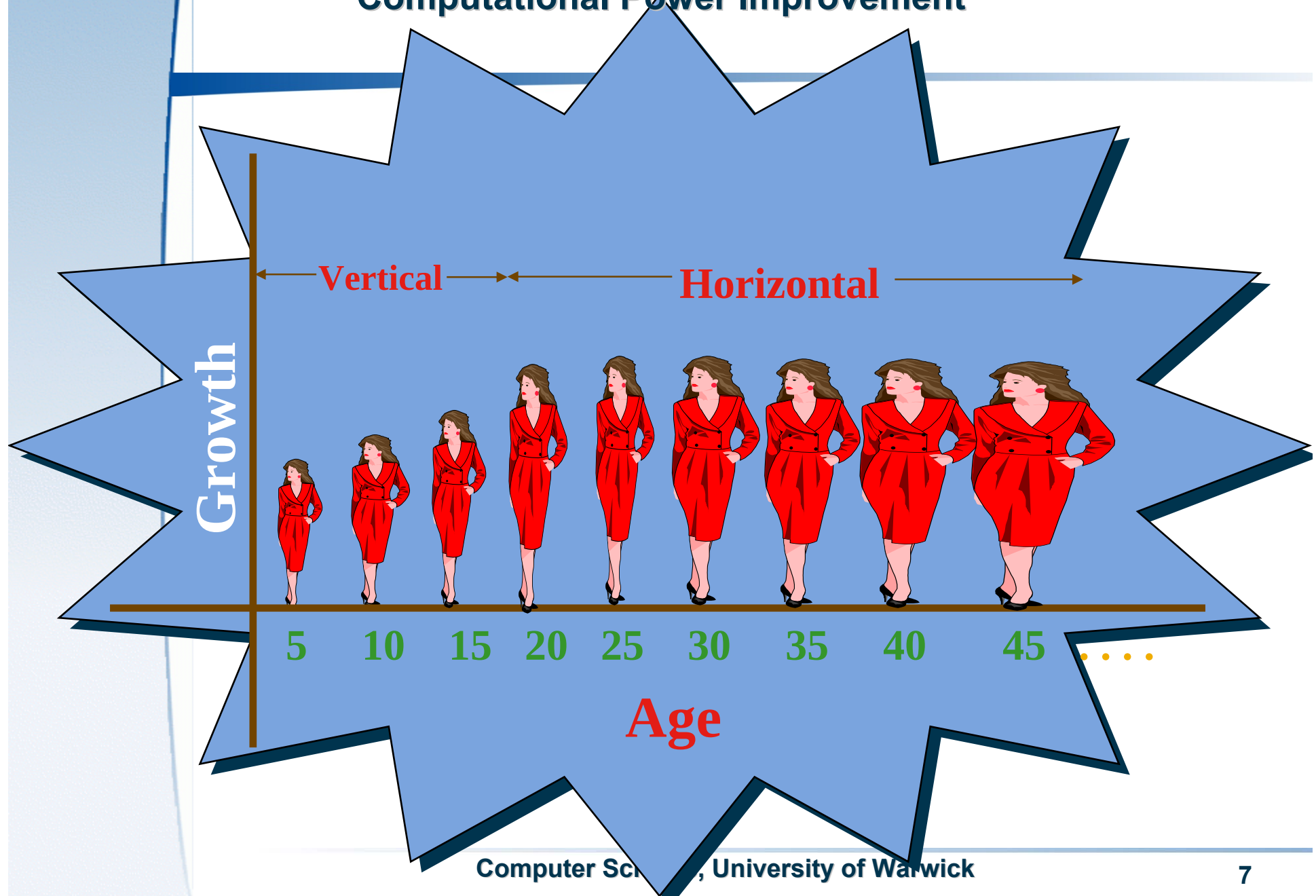
Computer Food Chain (since late 1990s)



Computational Power Improvement



Human Physical Growth Analogy: Computational Power Improvement



Why Clusters now?

Clustering gained new wave of interests when 3 technologies converged:

- ❑ **1. Very high performance Microprocessors**
 - **workstation performance = yesterday supercomputers**
- ❑ **2. High speed communication**
- ❑ **3. Standard tools for parallel/ distributed programming**

Cluster Components...1

Nodes

Multiple High Performance Components:

- ❑ PCs
- ❑ Workstations
- ❑ SMPs (CLUMPS)

They can be based on different architectures and running different OS

**But usually, the nodes in a cluster are homogenous
- they have same architecture and performance
and are installed with the same os**

Cluster Components ... 2

OS

OS used in various cluster systems :

- ❑ Linux (Beowulf)
- ❑ Microsoft NT (Illinois HPVM)
- ❑ SUN Solaris (Berkeley NOW)
- ❑ IBM AIX (IBM SP2)
- ❑ HP UX (Illinois - PANDA)
- ❑ Mach (Microkernel based OS) (CMU)

Cluster Components ... 3

High Performance Networks

Ethernet (10Mbps),

FDDI (100Mbps): Fibre Distributed Data Interface

Fast Ethernet (100Mbps),

Gigabit Ethernet (1Gbps)

Myrinet (10Gbps)

Infiniband (80Gbps)

Cluster Components ... 4

Network Interfaces

Network Interface Card

- ❑ Myrinet NIC
- ❑ Ethernet NIC
- ❑ ...

Cluster Components ... 5 Communication Software

Traditional OS supported facilities (heavy weight due to protocol processing)..

- ❑ Pipes, Sockets (TCP/IP), etc.

Light weight protocols (User Level)

- ❑ Active Messages (Berkeley): used in NOW system
- ❑ Fast Messages (Illinois): used in HPVM
- ❑ U-net (Cornell)
- ❑ XTP (Virginia)

System can be built on top of the above protocols

Cluster Components ... 6

Cluster Middleware

Resides Between OS and Applications and support:

- ❑ **Single System Image (SSI)**
- ❑ **System Availability (SA)**

**SSI makes collection appear as single machine (globalised view of system resources).
ssh cluster.myinstitute.edu**

SA - Check pointing and process migration..

PBS, Condor

Cluster Components ... 7

Programming environments

Threads (PCs, SMPs, NOW..)

- ❑ **POSIX Threads**
- ❑ **Java Threads**

MPI

- ❑ **Linux, Windows**

DSMs

PVM

Cluster Components ... 8

Development Tools

Compilers

- ❑ C/C++/Java;
- ❑ MPICC

Debuggers: e.g. Buster (based on sequential debuggers like gdb and dbx)

Performance Analysis Tools and Visualization Tools: e.g. Vampir Trace

Cluster Components ... 9

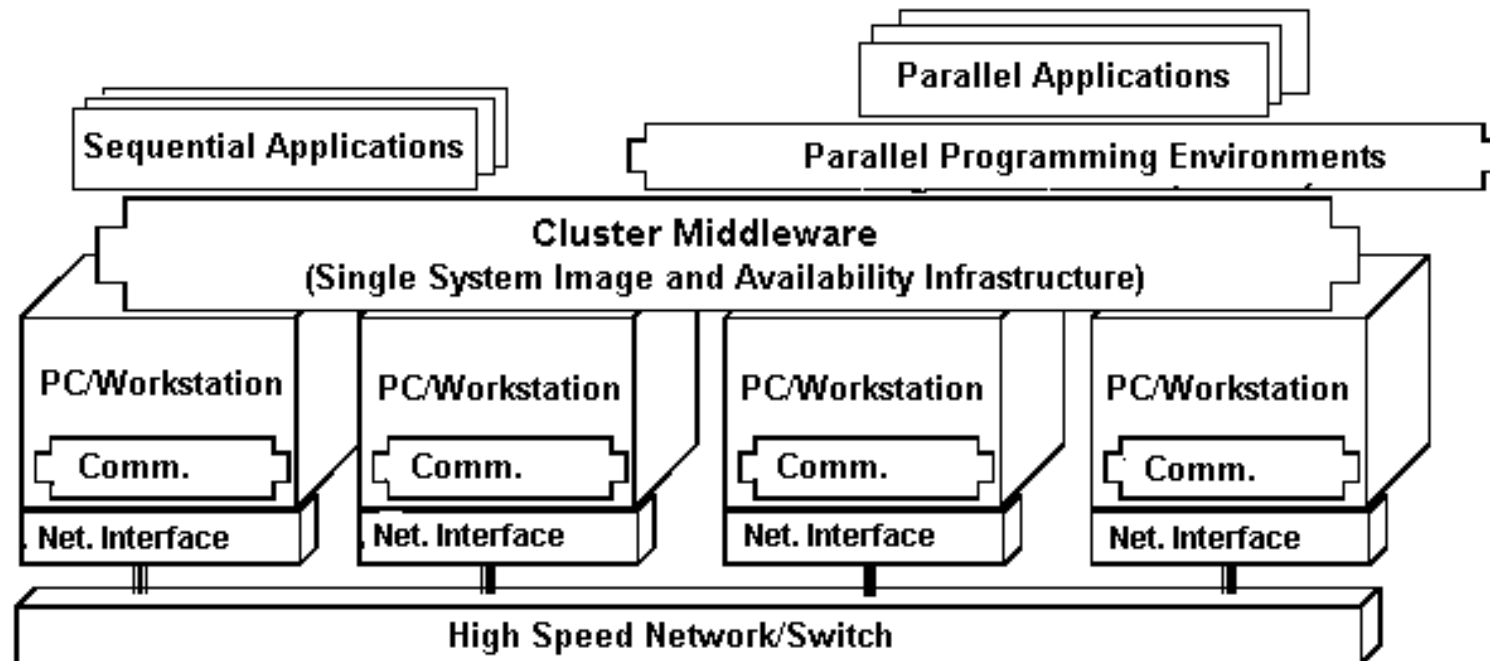
Applications

Sequential

Parallel / Distributed

- ❑ **Grand Challenging applications**
 - **Weather Forecasting**
 - **Computational Fluid Dynamics**
 - **Molecular Biology Modeling**
 - **Engineering Analysis (CAD/CAM)**
 -
- ❑ **Web servers applications**
 - **Google**
 - **Ebay**
 - **amazon**

Cluster Architectures



Cluster system architecture

Clusters Classification..1

Based on Focus (in Market)

- ❑ High Performance (HP) Clusters
 - Grand Challenging Applications
- ❑ High Availability (HA) Clusters
 - Mission Critical applications

Clusters Classification..2

Based on Workstation/PC Ownership

- ❑ **Dedicated Clusters**
- ❑ **Non-dedicated clusters**

Clusters Classification..3

Based on Node Architecture..

- ❑ **Clusters of PCs (CoPs)**
- ❑ **Clusters of Workstations (COWs)**
- ❑ **Clusters of SMPs (CLUMPs)**

Clusters Classification..4

Based on Node OS Type..

- ❑ Linux Clusters (Beowulf)
- ❑ Solaris Clusters (Berkeley NOW)
- ❑ NT Clusters (HPVM)
- ❑ AIX Clusters (IBM SP2)

Clusters Classification..5

**Based on node components
architecture & configuration
(Processor Arch, Node Type:
PC/Workstation.. & OS: Linux/NT..):**

- ❑ **Homogeneous Clusters**

- **All nodes will have similar configuration**

- ❑ **Heterogeneous Clusters**

- **Nodes based on different processors and running different OSes.**

What is Single System Image (SSI) ?

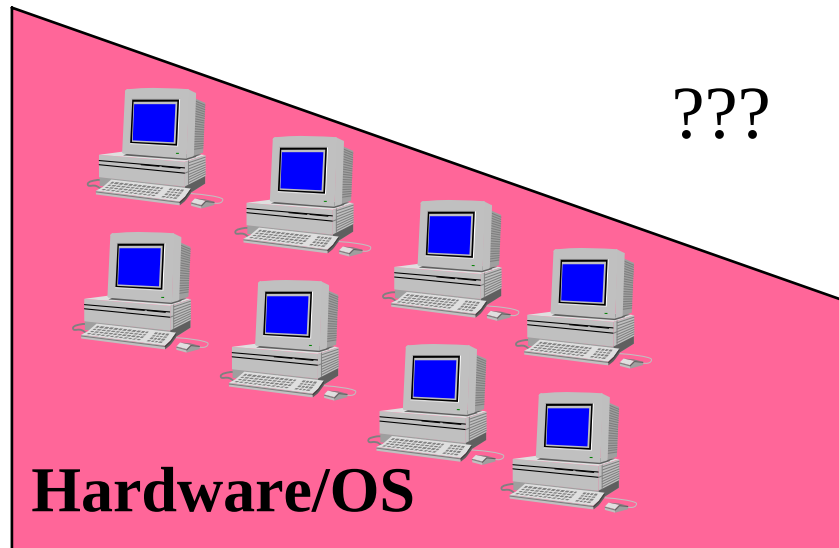
A single system image is the illusion, created by software or hardware, that presents a collection of resources as one, more powerful resource.

SSI makes the cluster appear like a single machine to the user, to applications, and to the network.

A typical Cluster Computing Environment

Application

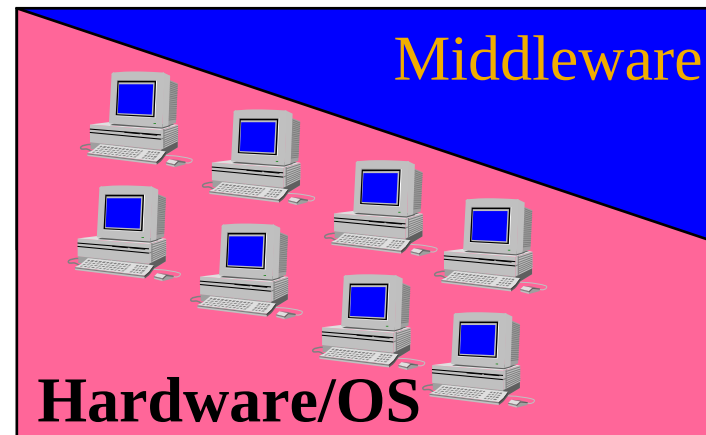
PVM / MPI/ SSH



The missing link is provide by cluster
middleware/underware

Application

PVM / MPI/ SSH



Single Entry Point

`ssh cluster.my_institute.edu`

`ssh node1.cluster.institute.edu`



Benefits of Single System Image

- Usage of system resources transparently
- Transparent process migration and load balancing across nodes.
- Improved reliability and higher availability
- Improved system response time and performance
- Simplified system management
- User need not be aware of the underlying system architecture to use these machines effectively