Achieving cluster reliability & cost effectiveness using IPMI

Richard Libby
Enterprise Products Group, Intel Corporation
Tuesday, May 13, 2003

rml@hpc.intel.com
“It is quite possible by the middle of this decade clusters in their myriad forms will be the dominant High-end computing architecture.”

Dr. Thomas Sterling

JPL/Caltech

http://clusters.top500.org
Overview

- HPC history & motivation
- Achieving cluster reliability & cost effectiveness using IPMI
- Call to Action

*Other names and brands may be claimed as the property of others.*
HPC 20yr Snapshot

10x growth

Performance

Today’s HPC Architecture

Accessible locally

Data Center Model
Accessed by specialists

COST

$4K/GFLOP
All COTS

$5M/GFLOP
No COTS

$250K/GFLOP
Some COTS

Intel® processor based ASCI - Red

Traditional HPC Architecture

Monolithic
Expensive
Proprietary
Exclusive

Applications
Middleware
OS
Hardware

Intel® processor based ASCI - Red

SCore IIIe: An Intel® processor based cluster in Japan.

*COTS - Commercial Off the Shelf

*Other names and brands may be claimed as the property of others.
As your cluster scales out, you need to achieving cluster reliability & cost effectiveness
Cluster Reliability

Problem:
- Even tiny “mean time between failures” can be significant for huge clusters.
  (Example)
  - Cornell*/Intel®/Microsoft* collaboration: the Radar reliability study
  - 64, 4-way IA32 Pentium® III Xeon™ Processor systems over 4 months:
    87 hours lost out of 199,680 hours, i.e. 99.96% uptime
  - In a 1000 node cluster, that means 1 node goes down every 48 hours, i.e. 1 really irritated user every other day.

Solutions:
- Hot swap
- Hardware redundancy
- Monitor/RAS network to anticipate trouble
- Use IPMI compliant hardware
- Use NEBS compliant hardware
- Write applications that are fault tolerant

Example: Cornell Theory Center has documented seven 9's

*Other names and brands may be claimed as the property of others.
Potential Failures

- **Tools and Applications**
  - Minimal Error Checking

- **Middleware**
  - Minimal Error Check

- **OS**
  - Device Driver Failure, Free Space

- **Interconnect**
  - Network Saturation, H/W Failure

- **Compute Node**
  - Unexpected Hardware Failure

- **Environment**
  - Power Failure, Cooling Failure

As your HPC scales out, these problems get worse!
Node Failure

- Need Node Monitoring Capability
- Need Node Control Capability

Solution: IPMI v1.5 Compliant Node Hardware

Goal: Alert User Before Node Fails
IPMI v1.5

- Defines a common, abstracted, message-based interface to intelligent platform management h/w
- Defines common records for describing common platform management devices and their characteristics
- Supports OEM differentiation and value added features
- Promoters: Intel, HP, NEC & Dell
  Adopters: Over 140 and growing*

*Adopters in backup slides*
IPMI v1.5 Functionality

- Monitoring
  - Voltage, Temperature, Fans, etc

- Platform Information
  - SDR, FRU, SEL

- Independent Watchdog Timer
  - Used in Conjunction With BIOS

- Outbound Alerting
  - Platform Event Traps (LAN Connection)
Dealing With the Problem

- Need extensive node hardware control (remote power up, power down, reset)
- Need extensive node hardware configuration control (IP address, boot source, etc…)

Solution: Direct Platform Control

Goal: Over-the-Wire Cluster Node (Re-)Configuration
Direct Platform Control (DPC-IPMI over LAN)

LAN NIC #1 NIC #2

LAN RMCP Port 26Fh LAN

Host Processor

In-band Traffic

Sideband Connection

Management Traffic

In-band Traffic

BMC
Direct Platform Control Implementation

- **Connection**
  - Direct to BMC via the Network Interface
  - RMCP packet format through UDP port 26Fh (NIC Filters)
  - Password protected (can be customized)

- **BIOS Console Redirection**
  - Text mode only
  - Terminates when the operating system boots into protected mode

- **Full Control of Power and Reset**

- **Management Queries of Platform Health**
Example: Simple Architecture

- Hot-spare IPMI
- Compute Nodes
- Master
- Network Access
- Storage
- Data Backup
- Health Master
- Compute Nodes with IPMI 1.5
IPMI Promoter, Contributor, and Adopter News

- Acer Inc.
- Agilent Technologies GmbH
- Alberta Microelectronics
- American Megatrends Inc.
- Arima Computer Corp.
- ASUSTek Computer, Inc.
- Axil Computer, Inc.
- Blue Wave Systems
- Bull S.A.
- Celestica
- Concurrent Technologies, PLC
- CyberGuard Corporation
- Data General Corporation
- Dell Computer Corporation
- Egenera, Inc.
- ElanVital Corporation
- Ericsson UAB
- Evans & Sutherland
- Eversys Corporation
- Exabyte Corporation
- FORCE Computers GmbH
- Fujitsu, Ltd.
- GoAhead Software, Inc.
- HADCO Corporation
- HCL Infosystems Ltd.
- Hewlett-Packard Company
- Hewlett-Packard GmbH
- Hitachi Ltd.
- Hybricon Corporation
- Ibus/Phoenix Corporation
- InnoMediaLogic, Inc.
- Intel Corporation
- Interphase Corporation
- InterWorks Computer Products
- Inventec Corporation
- Ipex ITG
- JMC Products
- L-3 Communications Corp.
- Lynux Works, Inc.
- Macrolink, Inc
- Magnetek, Inc.
- Micro-Star International
- Mirapoint, Inc.
- Mitsubishi Electric Corp.
- Mitsubishi Electric Corp.
- National Semiconductor Corp.
- NEC Corporation
- Nematron Corporation
- Network Appliance, Inc.
- Network Engines, Inc.
- Network Storage Solutions, Inc.
- NOCpulse, Inc.
- Olivetti Computers Worldwide
- Open Source Asia
- PEP Modular Computers
- Phoenix Technologies Ltd.
- Pigeon Point Systems
- Pinnacle Data Systems, Inc.
- Praim, Inc.
- Qlogic Corporation
- Radisys Corporation
- Reliance Computer Corporation
- Samsung Electronics Co., Inc.
- Sanera Systems, Inc.
- SBS Technologies
  (Industrial Computers GmbH)
- Scenix Semiconductor, Inc.
- Siemens AG
- Silicon Graphics, Inc.
- Stratus Computer Systems
  Ireland Ltd.
- Summit Microelectronics, Inc.
- Sun Microsystems
- Super Micro Computer, Inc.
- Symphony Group Intl. Co., Ltd.
- Synergy Microsystems
- Teknor Applicom, Inc.
- T-Netix, Inc.
- Tatung Co.
- Tektronix
- Texas Micro Corporation
- Toshiba Corporation
- Trillogic Systems, LLC
- Trimm Technologies
- Tyan Computer Corporation
- Universal Scientific Industrial Corp.
- USAR Systems, Inc.
- Vitesse Semiconductor Corp.
- Watrin System Design
- Vividon, Inc.
- Vooha, Inc.
- Winbond Electronics Corp.
- Ziatech Corporation

*Other names and brands may be claimed as the property of others.
So what’s it worth to you?

- **Power Control**
  - 8-Port NPS US$500 – 1k Nodes ~US$63K

- **Sensor Model**
  - 1-Port RS232 Sensor US$179 – 1k Nodes ~US$179K

- **Serial Redirection**
  - 40-Port Serial Concentrator US$50K – 1k Nodes ~US$1.2M

Spend more money on additional nodes instead of additional support infrastructure
Summary

- HPC Clusters follow Moore’s law as upgrade path
- Use IPMI-based systems for minimum cluster downtime

*Other names and brands may be claimed as the property of others.*
Call to Action

- Take advantage of affordable COTS hardware to drive large-scale HPC implementations that have minimal downtime using IPMI compliant hardware.
Backup Information
Links & Contact Info

• Links:
  – http://www.intel.com/go/hpc
  – http://www.openpbs.org
  – http://www.openclustergroup.org
  – http://clumon.ncsa.uiuc.edu
  – http://sourceforge.net/projects/clumon

• Richard Libby
  – 503-712-9794 rml@lprime.intel.com

*Other names and brands may be claimed as the property of others.
Enabling HPC with the Intel Ecosystem

Industry Standards
- LCI, GGF, OSDN, Gelatto, IPMI

OSV Enabling
- Microsoft, Red Hat, SuSE, HP/UX

iSV Enabling
- MSC, Accelrys, LSTC, Fluent,
- Schlumberger, Western GECO
- United Devices, Entropia, Platform

Cluster Enabling
- Scyld, Scali, OSCAR, ROCKS

Open Source Support
- OSDN, Source Forge, Linux Bios

Solutions Centers
- Oil, Finance, Life Sciences

Channel Enabling
- WW OEMS
- Regional Leaders

Intel WW Services
- Dedicated HPC Personnel
- End User Developer Enabling
- Infrastructure (WA, NM, VG, OR)
- Premier Technical Support
- HW Design Support

Intel Building Blocks
- CPU
- Chipset
- Interconnect
- Platform/Server boards
- SW Suite

Intel Training
- WW Software College
- WW Face to Face Training

Manufacturing
- .13us, 90nm, 300mm
- Platform Validation

Intel Capital
- SCALI, United Devices

Collaborations
- CTC, CERN, NCSA

System Integrators
- BluePrinting
- Accenture, PWC, CS

*Other names and brands may be claimed as the property of others.
Large Scale HPC Management & Monitoring using CLUMON
In the beginning...

Computing in the dark

bash-2.05# qstat

<table>
<thead>
<tr>
<th>Job id</th>
<th>Name</th>
<th>User</th>
<th>Time Use</th>
<th>S</th>
<th>Queue</th>
</tr>
</thead>
<tbody>
<tr>
<td>6420</td>
<td>NCBI-Blast.nt.SA</td>
<td>yongliu</td>
<td>00:05:20</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6429</td>
<td>dev.nt.f6</td>
<td>essewall</td>
<td>00:52:3</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6430</td>
<td>mpiBlastRun2.be</td>
<td>yongliu</td>
<td>00:00:52</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6431</td>
<td>mpiBlastRun4.be</td>
<td>yongliu</td>
<td>00:00:52</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6432</td>
<td>u1232k123_cc</td>
<td>sandvik</td>
<td>13:02:2</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6433</td>
<td>co.45stg.dsm</td>
<td>subdelva</td>
<td>13:17:0</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6434</td>
<td>CO20_execVMC</td>
<td>prason</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6435</td>
<td>Nav</td>
<td>hornmig</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6436</td>
<td>BSLNP-Rsopt</td>
<td>jwlee</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6437</td>
<td>mpiRun11</td>
<td>spamidig</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6438</td>
<td>sh123p</td>
<td>ricker</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6439</td>
<td>PING0p0.101</td>
<td>bob</td>
<td>19:55:08</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6440</td>
<td>prodsysse108</td>
<td>thumber</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6441</td>
<td>ncmd</td>
<td>rkufrin</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6442</td>
<td>SWDEN</td>
<td>hardner</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6443</td>
<td>mpiBlastRun2.be</td>
<td>yongliu</td>
<td>00:19:03</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6444</td>
<td>dense-apr14</td>
<td>murmurp</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6445</td>
<td>mpiBlastRun16.b</td>
<td>yongliu</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6450</td>
<td>SWDEN</td>
<td>arnoldg</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6451</td>
<td>NCBI-Blast.nt.be</td>
<td>yongliu</td>
<td>00:34:01</td>
<td>R</td>
<td>standard</td>
</tr>
<tr>
<td>6452</td>
<td>tsp128</td>
<td>cmendes</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6453</td>
<td>bezjob</td>
<td>u1403</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
<tr>
<td>6454</td>
<td>job.g3d.1</td>
<td>yongliu</td>
<td>00:00:5</td>
<td>Q</td>
<td>standard</td>
</tr>
</tbody>
</table>

bash-2.05# qstat -B

<table>
<thead>
<tr>
<th>Server</th>
<th>Max Tot Qu Run Hld Wat Trn Ext Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>titan60.ncsa.ui</td>
<td>0 23 15 0 0 0 0 0 0</td>
</tr>
</tbody>
</table>

*Other names and brands may be claimed as the property of others.*
CLUMON Architecture

- PBS Scheduler
- CLUMON Daemon
- Nodes with PCP/IPMI
- MySQL Database
- Apache with PHP
- Hosts, Queue, Jobs, Perf Metrics

*Other names and brands may be claimed as the property of others.*
Current Features

- **Intuitive interface**
  - Point and click web based, others possible

- **Correlation of data**
  - Cross-reference info from different systems

- **Control of resource consumption**
  - Cached data; tunable frequency and quantity

- **Huge set of data available**

- **Portable data**
  - Easily publish data in any format
  - Connect to data from socket, web, mysql

- **Adaptable**
  - Plugins, pcp modules, generic db structure
Upcoming Features

- **V1.3**
  - User and Admin portals
  - Problem Tracking
  - Availability Monitoring

- **V1.4**
  - Myrinet counter PCP module
  - Alert System
  - CluMon Plugin capability
  - Maui Reservation Monitoring Plugin
  - Job Performance Plugin

- **Future**
  - Network Data & Visualization
  - Management via scripts
  - Support for LSF, IPMI, CIM, etc.

*Other names and brands may be claimed as the property of others.*
Main Window (small scale)

![TITAN Cluster Monitor](image)

### Job Table

<table>
<thead>
<tr>
<th>Job</th>
<th>Owner</th>
<th>Job Name</th>
<th>Queue</th>
<th>State</th>
<th>Nodes</th>
<th>Time Used</th>
<th>% Time Allowed</th>
<th>Max Time Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>7687</td>
<td>rkufrn</td>
<td>namd</td>
<td>standard</td>
<td>Queued</td>
<td>128</td>
<td>00:10:00</td>
<td>0</td>
<td>00:10:00</td>
</tr>
<tr>
<td>7958</td>
<td>arnoldg</td>
<td>hpl16x1s</td>
<td>standard</td>
<td>RUNNING</td>
<td>16</td>
<td>02:50:55</td>
<td>48</td>
<td>05:55:00</td>
</tr>
<tr>
<td>7959</td>
<td>arnoldg</td>
<td>hpl16x1s</td>
<td>standard</td>
<td>RUNNING</td>
<td>16</td>
<td>02:50:42</td>
<td>48</td>
<td>05:55:00</td>
</tr>
<tr>
<td>7996</td>
<td>bode</td>
<td>PLAW2p0.027</td>
<td>standard</td>
<td>RUNNING</td>
<td>16</td>
<td>02:51:30</td>
<td>40</td>
<td>07:10:00</td>
</tr>
<tr>
<td>7999</td>
<td>hjyo</td>
<td>testini</td>
<td>standard</td>
<td>RUNNING</td>
<td>4</td>
<td>10:14:37</td>
<td>85</td>
<td>12:00:00</td>
</tr>
<tr>
<td>8046</td>
<td>mduez</td>
<td>merger</td>
<td>standard</td>
<td>RUNNING</td>
<td>4</td>
<td>07:24:18</td>
<td>62</td>
<td>12:00:00</td>
</tr>
<tr>
<td>8134</td>
<td>sandvik</td>
<td>u12132b128_c</td>
<td>standard</td>
<td>RUNNING</td>
<td>8</td>
<td>02:19:03</td>
<td>19</td>
<td>12:00:00</td>
</tr>
</tbody>
</table>

### Queue Table

<table>
<thead>
<tr>
<th>Queue</th>
<th>Type</th>
<th>Max Nodes</th>
<th>Max CPUs</th>
<th>Enabled</th>
<th>Total Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>cap16</td>
<td>Execution</td>
<td>16</td>
<td>32</td>
<td>True</td>
<td>0</td>
</tr>
<tr>
<td>hwtest</td>
<td>Execution</td>
<td>16</td>
<td>32</td>
<td>True</td>
<td>0</td>
</tr>
<tr>
<td>rh72suse</td>
<td>Execution</td>
<td>20</td>
<td>40</td>
<td>True</td>
<td>0</td>
</tr>
<tr>
<td>sandbox</td>
<td>Execution</td>
<td>32</td>
<td>64</td>
<td>True</td>
<td>0</td>
</tr>
</tbody>
</table>
Main Window (large scale)

There were problems encountered while collecting data. The errors were encountered while: Gathering PCP Data

<table>
<thead>
<tr>
<th>Job</th>
<th>Owner</th>
<th>Job Name</th>
<th>Queue</th>
<th>State</th>
<th>Nodes</th>
<th>Time Used</th>
<th>% Time Allowed</th>
<th>Max Time Allowed</th>
</tr>
</thead>
<tbody>
<tr>
<td>62414</td>
<td>ggh</td>
<td>CalcFut_2660</td>
<td>standard</td>
<td>ON HOLD</td>
<td>10</td>
<td></td>
<td>0</td>
<td>24:00:00</td>
</tr>
<tr>
<td>62415</td>
<td>ggh</td>
<td>CalcFut_2480</td>
<td>standard</td>
<td>ON HOLD</td>
<td>10</td>
<td></td>
<td>0</td>
<td>24:00:00</td>
</tr>
<tr>
<td>62417</td>
<td>ggh</td>
<td>CalcFut_2680</td>
<td>standard</td>
<td>ON HOLD</td>
<td>10</td>
<td></td>
<td>0</td>
<td>24:00:00</td>
</tr>
</tbody>
</table>
Core Tenets of Intel’s HPC Strategy

- Building block products based on industry standards and high volume to bring improved economics to HPC.

- Collaborations with thought leadership accounts that drive the evolution of applications, tools, and compute models forward.

- Programs enabling the broadest channels in the world delivering a significant choice of alternatives for the HPC community.

- Ecosystem support and enabling of ISVs, IHVs, and OSV.
Solutions and Case Studies

• Solutions for the Energy Vertical Market
• Significant more computing power than other solutions using IPF
• Short video on how Intel Itanium architecture is enabling new scientific breakthroughs at the National Center for Supercomputing Applications (NCSA), Cornell Theory Center (CTC), and the European Organization for Nuclear Research (CERN)
• Success stories on some of the world's leading research centers that have already turned to Intel-based solutions to meet their high-performance computing needs. Their stories provide a lesson in how to achieve maximum performance at minimal cost using Intel processors.
• Applications and resources that are written to fully take advantage of HPC technology
• HPC enabling software that will make your code run faster on IA (such as compilers, libraries and tools, solution stacks)
• Interconnect products and programs to speed up communication & bandwidth between network nodes
• Processor information

More at www.intel.com/go/hpc/
Cluster Software Stack

**OSCAR**
Open Source Cluster Application Resources

- **Collaboration:** Intel, IBM, ORNL, NCSA, Dell, SGI, Veridian, MSC.software

- **Compared to Beowulf:**
  - Industry support from inception
  - Draws from large community
  - Long-term effort
  - Planned support for multiple OS’s

*Other names and brands may be claimed as the property of others.*
System Control Software
– OSCAR*

• Features
  – Backed by a consortium:
    – Dell, IBM, Intel, NCSA, ORNL, SGI, Veridian, MSC.software
  – Components
    – LUI (Linux Utility for cluster Install uses PXE boot)
    – C3 (Cluster node management)
    – MPI/PVM (Message passing interfaces)
    – PBS (Workload management)
    – OpenSSH (Security)

*Other names and brands may be claimed as the property of others.
Scyld* Computing – Beowulf Cluster Operating System

• Features
  – Second Generation Product
  – Created By Don Becker (part of the original Beowulf development team)
  – Proven Robustness
  – Ease of use
  – GUI based cluster node configuration and control
  – Commercial Package – Fully supported Open Source
  – Now Supports IPMI

*Other names and brands may be claimed as the property of others.
Components

Written in C

clumond is the collection daemon

Uses Open source components

PBS Scheduler

Performance Co-Pilot from SGI

MySQL database

Display windows use PHP on Apache

Renders HTML, XML or any other custom format via a socket interface

*Other names and brands may be claimed as the property of others.
PBS Information

- Clumond gathers information about
  - Queues
  - Jobs
  - Hosts (as it pertains to scheduling)

- Clumond gathers ALL information from the Scheduler in a generic fashion. So if there are changes in PBS, no change should be needed in clumond.

*Other names and brands may be claimed as the property of others.
PCP Information

PCP

Available from:


Provides info on:

system load
memory usage
process information
network information

Customizable and expandable

*Other names and brands may be claimed as the property of others.
PCP Information

• PCP data can include almost any machine metric desired. If not already provided for, PCP is expandable.

• Content is very flexible; Maintainable by a list in the database. Just add a metric to the list and it will be gathered on the next data collection phase.

• No need to change web page either. Data will show up automatically.
Database Configuration

- Database schema is very abstract so that it can be adapted easily to future schedulers and performance APIs.
- All the system’s configuration is done in the database. No need to restart the daemon for changes to take effect.
- The database is very expandable. New areas of data collection include network adapters, switches, and job performance counters.
Links & Contact Info

• Links:
  – http://www.intel.com/go/hpc
  – http://www.openpbs.org
  – http://www.openclustergroup.org
  – http://clumon.ncsa.uiuc.edu
  – http://sourceforge.net/projects/clumon

• Richard Libby
  – 503-712-9794 rml@lprime.intel.com

*Other names and brands may be claimed as the property of others.
For additional information

- General Background
  - http://developer.intel.com/design/servers/ipmi/
  - http://www.intel.com/go/hpc/
  - http://www.beowulf.org
  - http://clusters.top500.org/

- Cluster software platforms
  - http://www.scyld.com/
  - http://www.compaq.com/hpc/
  - http://pdswww.rwcp.or.jp/
  - http://www.openclustergroup.org
For additional information

- URLs
  - http://www.intel.com/go/hpc/
  - http://www.scyld.com
  - http://www.beowulf.org
  - http://www.dell.com/ (search for HPC)
  - http://www.compaq.com/hpc/
  - http://clusters.top500.org
  - http://pds www.rwcp.or.jp
  - http://www.opensce.org
  - http://www.openclusteringgroup.org

*Other names and brands may be claimed as the property of others.*
Glossary

- **Cluster** - a collection of connected independent computers that work in unison to solve a problem.
- **COTS** – Commercial, Off The Shelf. The use of readily available products to create a large machine.
- **FC-PGA** – Flip Chip-Pin Grid Array. A current packaging technique for microprocessors from the Intel Corporation.
- **High Availability** – Another form of clustering to give non-stop computing in the event of partial hardware failures.
- **HPC** – High Performance Computing (supercomputer characteristics).
- **Load Balancing** – Another form of clustering that distributes an incoming workload across a series of computers. Usually in reference to incoming TCP/IP traffic to a web site.
- **MPP** – Massively Parallel Processor. Linking together many (sometimes as much as 9000 processors) in a single machine through a specialized interconnect scheme.
- **Parallel Computing** – Calculating parts of the program at the same time to shorten the overall amount of time needed to solve the problem.
- **SPMD** – Single Program Multiple Data
- **SMP** – Symmetric Multi-Processing. A configuration of processors in a single computer that allows all processors the same access to system resources such as memory or I/O.
Glossary

- **ECC** – error correction circuit. Used to check memory and other bus data for errors. If single bit errors are encountered this circuitry can correct them on the fly.

- **FLOPS** – floating point operations per second. A measure of the capability of a machine or cluster to deal with scientific calculations.

- **FSB** – front side bus. The connection of the microprocessor to the rest of the system (memory, I/O etc).

- **LFS** – log structured file system. A variation of a journaling file system.

- **MPI** – message passing interface. A library specification for message-passing, proposed as a standard by a broadly based committee of vendors, implementers, and users.

- **MPICH** - a portable implementation of the full MPI specification for a wide variety of parallel computing environments, including workstation clusters and massively parallel processors (MPPs) from Argonne National Labs.

- **MPI/Pro** - a commercial implementation of the MPI standard.

- **PVM** – parallel virtual machine. A software package that permits a heterogeneous collection of UNIX*, Linux and/or Windows* NT computers hooked together by a network to be used as a single large parallel computer.

*Other names and brands may be claimed as the property of others.