



Building a Grid System for HPC

Danny Shieh and Hsin Yen Chen

ASGC

ISGC 2008, Taiwan



HPC on Grid

- **High Performance Computing (HPC):** Use of computer system for numerical intense computing. It is commonly associated with the use of computer for scientific research.
- **High Performance Technical Computing:** For engineering applications and computing related to analysis.

Can These Computing Run on Today's Grid System?

- or -

Is Grid System Capable of Support HPC?

Important Issue for the successful of enabled Grid for e-Science



Grid Computing System

- (with a few exceptions) Most of computers on Grid are the cluster of Intel/AMD based microprocessors
- Per CPU, the computing performance of today's microprocessor is closely comparable to special designed 'supercomputer'



Cluster Computer

Cluster of massive Intel/AMD based computer system is fast become the choice of HPC platform. (Thousand of processors)

(Nov, 2007) 406 computing system on Top 500 List are cluster of Intel/AMD based computer.

Does this mean that Grid system can handle all types of HPC requirements?

Also, Cluster based on Blade server?



Nature of Today's HPC Application Programs

- *Large Memory Requirement*
- *Long Running Job*
- *Parallel Processing*
- *Large amount of I/O*



HPC Processes on Grid

- **Workflow Computing: Require system middleware**
- **High Throughput: Suitability - Very High**
- **Parallel Processing: Cluster site dependent**
- **High I/O Jobs I/O system on computing site**
- **Large Memory Job: CPU dependent, 64 bits support**
- **Time Critical Job: Suitability – Low**



Source of HPC Application Program

- **Package Application Software**
 - Mostly, it requires software license
 - Cost of install on every grid site
- **Home Developed Programs**
 - (may-be) Source code modification for every run
 - Static binding job



Porting and Program Installation Issues

- **Capability of Computing System on Grid Site**
- **Compiler and Compiler library**
- **System OS**
- **End User not necessary wants to involve in this**



Parallel Computing Jobs

- **Parallel Computing Models**
 - **Message Passing (MPI Tasks):** Requires interconnect communication
 - **Shared Memory (Threads):** Multiple CPUs shared the common addressable memory
 - **Shared memory computing system on Grid?**
- **Parallelism of Application Program**
 - **Number of CPUs**
 - **Degree of parallelism in a program**
 - **Degree of data sharing among the parallel task**



Parallel Computing Support on Grid (1)

- **Cross-Site parallel: Very, very limited**
 - Inhomogeneous of system across sites
 - Computing performance different from site to site
 - Only a test had been done for specific application
- **Parallel Jobs on a Grid Site**
 - Parallel Computing Environment (at system level)
 - Issue of interconnect communication
 - CPU performance of each CPU on a cluster
 - Number of CPUs on a cluster



Parallel Computing Support on Grid (2)

- Require for enhanced Grid middleware for parallel computing support
- Very, very few sites support parallel computing
- Cost of high performance communication switch
- System support high performance parallel I/O
- Parallelism limited to:
 - ✓ Small to medium parallel (number of CPUs issue)
 - ✓ I/O system that support parallel computing



A Status Summary of Grid for HPC

- **Grid can support HPC applications without major difficult**
 - **Single serial batch jobs**
 - **Job with memory requirement within 2GB**
 - **A perfect solution for high throughput computing project**
- **High Performance Parallel Computing on Grid is not generally available**
- **Porting applications for grid system is an issue**
- **Require for enhancing Grid middleware**
- **Matching Job requirement and Grid resource is a big issue**
- **Need for a better Application User Interface**
- **An improvement for User I/O files support**



ASGC Quanta Blade Server for HPC (1)

- **System Specification**
- **3xQuanta S72A**
 - 10 blades per chassis, each blades 2-way SMP
 - Total 30 nodes (60 CPUs)
- **CPU: Intel Xeon at 3.2 GHz, Cache L1:16KB, L2: 1MB**
- **Memory: 4 GB per node**
- **Internal Disk: 147GB, PCI-X, Ultra 320 SCSI**
- **Default Network: Gigabit Ethernet**
- **High Performance Switch: Mellanox InfiniScale III 2400**
- **System OS: Scientific Linux**



ASGC Quanta Blade Server for HPC (2)

- **Compiler and Library**
- **Intel Fortran and C compiler with MKL**
- **PGI & GNU**
- **MPICH for MPI programming**
- **Other libraries: Mvapich, Atlas, FFTW**



ASGC Quanta Blade Server for HPC (3)

- **Computing Environment and User Support (based on gLite)**
- **Pre-process Procedure**
 - Obtain CA, Join VO, Get UI account, Set Environment
- **Support for Environment Setting on UI: Unix based and Window Users**
- **Job Submission**
 - Grid proxy initialization
 - Submission Methods: Use EDG command or Automatic Job Submission (HPC submit)
- **Parallel Computing Support**
 - Hybrid Parallel model: MPI task per node, then two OpenMP threads in a node
 - Maximum number of CPUs for a job is 48.



Easy of Use for HPC Users on Grid

	<i>Cluster</i>	<i>Grid</i>	<i>ASGC HPC</i>
<i>UI</i>	Front End	Grid UI	Grid UI
<i>Resource</i>	Single Cluster	Cluster	Cluster
<i>Security</i>	Password	Password/CA	Password/CA
<i>Job Submission</i>	PBS Script	JDL Script	Wrapper
<i>Job Maintenance</i>	PBS Job Command	EDG Job Command	EDG Job Command
<i>Share File System</i>	NFS	Storage Element (SE)	NFS
<i>Runtime Input</i>	From NFS	Resource Broker (RB)	From NFS
<i>Output Retrieve</i>	From NFS	RB / SE	From NFS



Quanta Blade Server Status Summary:

1. Quanta Blade Server had been successfully configured and implement for HPC application on Grid (gLite)
2. Performance benchmark indicated the system is of a comparable capability of other dedicated HPC cluster systems.
3. System is on production environment since last year (Note: This system was used in EGEE's Avian Flu Data Challenge in 2006, 2007)
4. Need for a High Performance Share File System
5. Need for an Enhanced UI

Next: Multiple Sites (Grid middleware,.. etc)