SUN IN GRID AND HPC

PENG Liang, Ph.D
Research Scientist / Technical Consultant
Asia Pacific Science & Technology Center
Sun Microsystems, Inc.
Agenda

HPC@Sun

Sun's Grid Computing &

N1 Grid Engine

Introduction to Sun APSTC

Conclusion
HIGH PERFORMANCE COMPUTING
@ SUN
Sun HPC Vision

To Create the Best Technologies, Products and Services Needed For High Performance Computing
Key Recent Sun HPC Achievements

+ Solaris HPC Development and Deployment Environment
+ University of Southern California (USC) - 13.8 TF (#26)
+ Tokyo Institute of Technology – 38.18 TF (#7)
+ Sun HPC Solutions Center
+ Key Acquisitions and HPC Partnerships

Pain Points Addressed:
+ Scaling to 100+ TFLOPs based on Standards
Innovation at Sun

System

Innovate at the system level with industry-standard chip.

Chip

Innovate at the system AND the chip level
Sun Fire X4600 – Performance & Scalability

- 16 core scaling delivers leading-edge performance
- Up to 128GB of memory
- 25GB/s bi-directional I/O bandwidth

* Pending availability of 4GB DIMM.
Sun Fire X4500 Server
Integrated x86 compute power, massive storage capacity and high data throughput

Compute
- 2 x Dual Core Opteron processors
- 16GB Memory

Storage
- 48 Serial ATA disks (500GB each)
- Up to 24TB raw capacity

I/O
- Very high throughput
- 2x PCI-X slots
- 4 GigE

Availability
- Hot-swap power, fans, disks

Management
- Same management as other x4600 servers

Solaris (TM) ZFS
- Ground-breaking file system performance
Grid Computing Components:
Systems Optimized for All Workloads

64-bit UltraSPARC
64-bit AMD/Opteron x86
64-bit UltraSPARC
Sun HPC Software Stack

• Operating System: Solaris and Linux Choice
• File Systems: Partner with CFS to port Lustre to Solaris and ZFS
  > Integrate better performance and reliability
  > Linux Support is today
• Compilers: Sun Studio 11
• System Management Tool: N1SM
• Resource Management Tool: N1GE
• Open Source
Solaris for HPC and Grid

• Long time support of 64bit computing
• Running on wide range of platforms
  > SPARC, AMD64, Intel x86, ......  
• Support of multi-core architecture
• Outstanding Features & Tools: Zones, DTrace, ...
• Robustness: Self-healing features
Solaris 10: Sun's Innovation Pays Off

- Extreme Performance
- Dynamic Tracing
- Network Entry Systems
- Unparalleled Security
- Process Rights Management
- Crypto Infrastructure
- Platform Choice
- Linux Compatibility
- New UltraSPARC IV, New AMD Opteron
- Optimal Utilization
- Solaris Containers
- Relentless Availability
- Predictive Self Healing
- Next-gen Filesystem
The Right OS for Any Need
Sun Studio 11

Key Benefits

• **C, C++, Fortran Development**
  Compilers support industry & defacto standards to enable portable, maintainable, extensible code

• **World Record Performance**
  Optimized performance for each of the target systems: UltraSPARC, X86 and x64 for maximal system utilization

• **Multi-core Development**
  Powerful debugger, auto-parallelizer, advanced performance analysis tools, and OpenMP 2.5 support

• **Full Featured graphical development environment**
  Based on Award-winning Netbeans open source IDE

Graphical Debugger

Performance Analyzer

30 August 2006 ©

Liang Peng, Research Scientist, Sun APSTC
Sun N1 System Manager

- Enables lifecycle management of groups of Rackmount and Blade Systems
- Provides a single console for managing 1000's of compute nodes
- Free and Open Source
- Functionality includes
  - Complete remote control of all managed systems
  - Discovery, system configuration
  - Provisioning
  - Firmware updating/OS patching
  - Centralized monitoring
  - Advanced GUI interface
  - Full scripting capability
A Success Story

If you attended the Key Note Lecture (yesterday):
“Onto Petascale Grids”
......
Adding to the Galaxy Capabilities ...

More than 45 world records...
38.18 TFlops!!!
Tokyo Institute of Technology

- World’s Largest x64 Cluster
- Over 10,000 x64 processor cores, 20 TB main memory
- Over 1 PegaBytes Sun Storage with Lustre Parallel File System
- World’s Fastest Infiniband Network from Voltaire
- Ranked No.7 in Top500 list and No.1 in APAC Area
- More powerful than the sum of all the rest Japan universities
- Using N1SM and N1GE
**Titech Top Level System Diagram**

**Sun Fire X4600**
- 655 nodes
- 10,480 CPU Cores
- Memory: 21.4TB

**Sun Fire X4500**
- Physical Capacity 1PB

<table>
<thead>
<tr>
<th>Type</th>
<th>Nodes</th>
<th>CPU Clock</th>
<th>MEM/Node</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>639</td>
<td>2.4GHz</td>
<td>32GB</td>
</tr>
<tr>
<td>B</td>
<td>16</td>
<td>2.6GHz</td>
<td>64GB</td>
</tr>
</tbody>
</table>

**InfiniBand Network Voltaire ISR 9288 ×8**

**External Network Devices**

**Storage Server A**
- Sun Fire X4500
- Physical Capacity 1PB

**Storage Server B**
- NEC iStorage S1800AT
- Physical Capacity 96TB RAID6

**External Network Devices**

**FileServer**

**100 Gigabit-class Network Equipment**

**Applications**
- Gaussian03
- GaussView
- TCP Linda for Gaussian
- Molpro
- MOPAC
- AMBER 8
- Materials Explorer
- Materials Studio
- Discovery Studio Modeling
- ABAQUS (Standard, Explicit)
- ABAQUS/CAE
- NASTRAN
- PATRAN
- SAS
- MATLAB
- Mathematica
- AVS/Express PCE
- AVS/Express Developer
- EnSight

**Fields**
- Calculation Chemistry
- Structural Analysis
- Statistical Analysis
- Numerical/Symbolic Computation
- Visualization
Customer's Feedback

“Not only is the performance of the Sun Grid HPC environment extremely impressive today, but the ability of the architecture to scale rapidly is really phenomenal, and will enable us to grow our environment to meet our needs for many years to come—no matter how compute-intensive our projects may be.”

—Professor Satoshi Matsuoka,
Head of Research Infrastructure, Global Scientific Information and Computing Center, Tokyo Institute of Technology
Titech Petascale Grid

- Sun helped build Tsubame cluster, which is now a key building block of the campus Grid in Titech.
- The Tsubame cluster in Titech is one of the good example of Sun's ability of providing solutions of world class supercomputing.
- Sun will continue to contribute to the goal of “everyone supercomputes”.
- Sun's technology for Titech cluster is also applicable to smaller scale clusters with some customizations.
Highlights of Sun's offering at Titech

• Sun designed X4600 16 core SMP
  > Based on commodity AMD Opteron processors

• Sun designed 24 TB storage server
  > Based on commodity enterprise hard disks

• Voltaire High Performance switching network
  > Based on Open standards Infiniband

• ClusterFS High Performance file system
  > Based on the Open Source Lustre file system
Highlights of Sun's offering at Titech

• N1 Grid Engine job scheduler
  > Open source written by Sun

• N1 System manager for cluster management
  > Free for download

• SuSE Linux + Solaris
  > Both are open source

• World Class Partnership
  > Sun + NEC + Voltaire + CFS + Clearspeed
Building the Right Infrastructure

Servers | Storage | Software | Services
A Complete HPC Portfolio From Sun

<table>
<thead>
<tr>
<th>Sun CRS, Support, Architectural, Professional Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applications</td>
</tr>
<tr>
<td>Custom or ISV Applications</td>
</tr>
<tr>
<td>Sun HPC Cluster Tools/Development Tools</td>
</tr>
<tr>
<td>Management</td>
</tr>
<tr>
<td>Workload Management</td>
</tr>
<tr>
<td>Cluster Management</td>
</tr>
<tr>
<td>Sun N1™ Grid Engine Software</td>
</tr>
<tr>
<td>Sun N1™ System Manager Software</td>
</tr>
<tr>
<td>Operating System</td>
</tr>
<tr>
<td>Open, Free</td>
</tr>
<tr>
<td>Node</td>
</tr>
<tr>
<td>Processor</td>
</tr>
<tr>
<td>64 Bit</td>
</tr>
<tr>
<td>Interconnect</td>
</tr>
<tr>
<td>Gigabit Ethernet, Myrinet, Infiniband</td>
</tr>
<tr>
<td>Open, Free</td>
</tr>
</tbody>
</table>
Key Takeaways

HPC components are key building blocks for Grid computing

Sun has a great product portfolio for HPC

Sun is committed to HPC
Grid Computing

• A hardware and software infrastructure that connects distributed resources, such as computers, storage devices, databases and software applications through a network, and is managed by DRM (distributed resource management) software

• A way of using many resources to perform many kinds of tasks, accessible from many places by many people

• A universal computing infrastructure that builds on the power of the Net and enables more efficient computation, collaboration, and communication
Levels of Grid Computing

- **Department Grids**
  - Single user community
  - Single organization

- **Enterprise Grids**
  - Multiple user communities
  - Single organization

- **Global Grids**
  - Multiple user communities
  - Multiple organizations
Grid Enabling Technologies

Global Grid Infrastructure

- Service Discovery
- Authentication/Authorization
- Data Management
- Policy management
- Resource management
- System management
- Data Access

Enterprise Grid Infrastructure

- N1™ Grid Engine (Policy Engine)
- N1™ Grid Containers
- Solaris™ Resource Mgr
- Sun™ Management Center
- Sun™ VTS
- Sun™ QFS/SamFS
- Solaris™ CacheFS

Cluster Grid Infrastructure

OGSA
- Globus Toolkit
- Avaki (3rd party)
Utility Computing: Sun Grid Pay-Per-Use Cycles Service

A virtual compute ranch that provides access to a secure, pay-per-use N1 Grid service, optimized to solve non-transactional workloads at an introductory price of $1/hour/CPU

In the U.S. only, initially; availability to be announced

- Off-load computational workloads that require 10s or 100s or 1,000s of CPUs.
- Use compute power on an hourly basis, without ownership or onerous outsourcing contracts.
- Leverage economies of scale achieved by mapping your workload to a singular service provider with a common infrastructure.
Welcome to Sun Grid

POWER UP YOUR BUSINESS WITH SUN GRID COMPUTE UTILITY
Access the secure compute power you need, when you need it, easily through the internet for just $1/CPU-hr. Pilot US only. Register today »

1 UPLOAD Your Application

Sun Grid Resources
- About Sun Grid
- Sun Grid Developer Community
- Sun Grid ISV Partners
- Utility Computing

Sun Grid Launch
- View Executive Video
- View Customer Testimonials
- Become a Sun Grid Channel Partner

Free Test Drive
Create a spoken-word mp3 audio file from your test input. Try It

Sun Grid Virtual Tour
Take a virtual tour of Sun Grid. Start the tour

Industry Examples
Learn how customers in various industries can take advantage of Sun Grid
- Media and Entertainment
- Electronic Design
- Financial Services and Energy
- Life Sciences

Enter the Grid »
- Register for a Sun Grid Account
- Forgot User Name or Password?
- Sun Grid Frequently Asked Questions
Grid Engine Open Source

- 500,000+ lines of source code
- Binaries for Compaq, HP, IBM, SGI, Linux, 50K+ of downloads
- 50+ development partners
- Contributions to scheduler, broker, parallel, clients, Globus, Grid Engine Portal, ...

http://gridengine.sunsource.net
Grid Landscape and N1GE

Interfaces with

Provides

N1GE

HA
System Mgmt
OS/Svc Prov
Data Grid

Grid Portals
Global Grids

Resource Mgmt
Acct Reporting

Monitoring
Workload Mgmt
Policy Automation
Sun Grid Computing in Asia South National Projects (1)

- Sun's involvement in Singapore's National Grid
  - Actively involved in the workgroups (Dr. Simon See – Middleware and Architecture WG)
  - Contribute resources to NGPP
  - Contribute resource to SG@Home project
Sun Grid Computing in Asia South National Projects (2)

- Sun's involvement in Malaysia's National Grid events (with MIMOS)
  - Conducted Training at MIMOS
    - Grid Service Design and Development
    - Setting up Grid Testbed with Globus Toolkit
    - Sun Grid Engine Administration
    - HPC Benchmarking
    - HPC System and Application Performance Tuning
N1 GRID ENGINE
N1 Grid Engine

- Sun Microsystems licensed product
  - Supported by Sun on all major OS (Windows, all UNIXes, Linux, MAC, etc) with optimized AMD x64 binaries
  - Significant EDU/Research pricing discount
  - GEMM: Interoperation with Sun Control Station (SCS) 2.2
  - ARCo: Advanced accounting, billing & policy allocation
  - Grid Engine Portal (GEP): customized end-user interface for easy job submission and control via secure portal.
    - Available through Sun consulting services
  - N1GE Rocks: Default DRM is Open Source Grid Engine
    - Commercial support for N1GE+ROCKS through Scalable Systems
## Supported Platforms

### HETEROGENEOUS ENVIRONMENTS

<table>
<thead>
<tr>
<th>Master Host</th>
<th>Compute Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solaris 7, 8, 9, 10 on SPARC</td>
<td>Solaris 7, 8, 9, 10 on SPARC</td>
</tr>
<tr>
<td>Solaris 9, 10 on x86/x64</td>
<td>Solaris 9, 10 on x86/x64</td>
</tr>
<tr>
<td>Linux kernel 2.2-2.6 on x86/x64 (any distribution)</td>
<td>Linux kernel 2.2-2.6 on x86/x64 (any distribution)</td>
</tr>
<tr>
<td>Mac OS X</td>
<td></td>
</tr>
<tr>
<td>AIX 4.3, 5.1</td>
<td></td>
</tr>
<tr>
<td>HP-UX 11i</td>
<td></td>
</tr>
<tr>
<td>Irix 6.5</td>
<td></td>
</tr>
</tbody>
</table>
SGE 5.3 -> N1GE 6.0

• N1GE 6.0
  > Accounting & Reporting Console (ARCo)
  > Cluster Queues + Hostgroups
  > More Sophisticated Scheduler Policies
  > Next-gen architecture for even greater scalability

• N1GE 6.0 update 4+
  > Windows Execution Support
  > Grid Engine Management Module (GEMM)
    > Integration with Sun Control Station
N1GE6: Enhanced Performance

- Improved Scalability
  - Berkeley DB spooling
  - Multi-threaded Master Daemon
  - New communication system

- Higher number of hosts/jobs
  - 10,000 unique hosts
  - 500,000 unique jobs

Please Note: N1GE currently supports a higher maximum number of active hosts and jobs than any DRM in the market. Numbers beyond these maxima are supported via multiple Masters config.

An Array Job is considered as one job, allowing for “infinite” scalability!
Campus Grids with N1GE

- **NTU Campus Grid**: Nanyang Techn. University, Singapore
- **UKM/USM**: Collaborative Grid Research, Malaysia
- **White Rose Grid**: Universities of Leads, Sheffield, York, UK
- **HPCVL**: Secure innovative HPC/Grid environment, Canada
- **AIST**: Advanced Industrial Science & Techn. Institute, Tokyo
- **PROGRESS**: Polish Research on Grid Computing, Poland
- **myGrid**: Infrastructure for an e-Biologist Workbench, Manchester
- **NRC-CBR**: BioGrid by National Research Council, Canada
- **... and many more**
References

- N1 Grid Engine 6.0
  > http://www.sun.com/gridware/
- Sun Grid Engine Open Source Community (SGE+GEP)
  > http://gridengine.sunsorce.net/
- Sun in HPC (Newsletters, Blueprints, Whitepapers, etc)
  > http://www.sun.com/hpc/
- Asia Pacific Science & Technology Center
  > http://apstc.sun.com.sg/
- Grid Engine News Tracker
  > http://www.gridengine.info/
APSTC
INTRODUCTION
Asia Pacific Science and Technology Center

• In collaboration with Nanyang Technological University, Singapore
• Network of collaborating researchers across Asia (Singapore, Malaysia, Thailand, India, Japan, China, ...)
• Research and Engineering in Grid and HPC
• URL: http://apstc.sun.com.sg/
APSTC Introduction

• To provide technical expertise for research in areas like bio-computing, engineering, etc.
• Close engagement with universities and research institutes
• To conduct research work in Grid Computing and High Performance Computing
APSTC Introduction

Vision

• APSTC: A well established and recognized center specialized in scientific computing

Objectives

• Work with the scientific community in adopting and developing latest Sun technology
• Monitor computational trend in specific scientific & engineering communities
• Transfer of know-how and exchange of knowledge with academic and research institutions
• Evangelize on specific Sun technology
• Monitor technological trends in Grid and High Performance Computing
Focus Area

- Life Sciences Computing
- Engineering Computing
- Financial Computing
- Mathematical Science and Modeling
- Digital Content

- GRID Computing
- HPTC Environment
- Performance Tuning
APSTF in Thailand

- One of the Asia Pacific Science & Technology Facility on APAC
- Established in 2004
- Focus primarily on building up expertise in HPC and Grid
- Provide support in Grid and HPC for Sun Microsystems Operation in Thailand
- Conducting R&D on Sun products and technology
- [http://apstf.cpe.ku.ac.th/](http://apstf.cpe.ku.ac.th/)
APSTC Related Activities in Asia

**Research**
- Collaborative Treasure Hunt, GriDE Toolkit, Grid Market Framework, Grid Benchmarking, Grid Super Scheduler, NTU Campus Grid, ChinaGrid, Collaborative Engineering, ...

**Engineering**
- SSC World Grid, Bio-ClusterGrid, Math-Grid, MCAE-EduBox, Auto Application Performance Tuning, Grid Simulator, Mobile Multiplayer Game, Bio-DataServer, N1GE in Containers, ...

**Collaboration**
- Nanyang Technological University, National University of Singapore, DSO National Labs, NGO, MIMOS, UKM, USM, General Electric, ST-Micro, AMD, Ericsson, ...

**Integration**
- DSO National Labs, National University of Singapore, NTU, IIT Kanpur, UKM, USM, General Electric, ST-Micro, Chartered Semiconductors, Ericsson, NYP, ...

**Training**
- Knowledge Transfer Implementations

30 August 2006 © Liang Peng, Research Scientist, Sun APSTC
Summary

• Sun Microsystems has shown its capability and commitment in Grid computing
  > Software, hardware, solution

• Titech cluster is a typical success story of build top HPC blocks for Grid computing

• Sun APSTC is doing research in Grid computing and is actively involved in R&D work in Grid computing in APAC universities

• Sun APSTC is providing strong technical and engineering support/consultancy for Sun's solutions
THANK YOU

Liang Peng
liang.peng@sun.com
THANK YOU

Liang Peng
liang.peng@sun.com
What is Grid Computing?

- A hardware and software infrastructure that connects distributed resources, such as computers, storage devices, databases and software applications through a network, and is managed by DRM (distributed resource management) software.
- A way of using many resources to perform many kinds of tasks, accessible from many places by many people.
- A universal computing infrastructure that builds on the power of the Net and enables more efficient computation, collaboration, and communication.
# Application Industries

<table>
<thead>
<tr>
<th>Industry</th>
<th>Computationally Intensive activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Sciences</td>
<td>Genetic sequencing, database queries</td>
</tr>
<tr>
<td>Electronic Design</td>
<td>Simulations, verifications, regression testing</td>
</tr>
<tr>
<td>Financial Services</td>
<td>Risk and portfolio analysis, simulations</td>
</tr>
<tr>
<td>Automotive Manufacturing</td>
<td>Crash test sims, stress testing, aerodynamics modeling</td>
</tr>
<tr>
<td>Scientific Research</td>
<td>Large computational problems, collaboration</td>
</tr>
<tr>
<td>Oil and Gas</td>
<td>Simulations, seismic analysis, visualization</td>
</tr>
<tr>
<td>Digital Content Creation</td>
<td>Frame rendering</td>
</tr>
<tr>
<td>Software Development</td>
<td>Compilations and code builds, automated testing</td>
</tr>
</tbody>
</table>
N1 Grid Engine

- Grid Engine (open source version): http://gridengine.sunsource.net
- Both commercial & open-source versions offer:
  - Easily set up and managed and popular (>10,000 installations) so knowledge of how-to-use is common
  - Highly configurable resource definitions
  - No Linux kernel changes necessary
  - Define jobs by resources needed, not by queues
  - Fine control over resources needed/allocated per job
  - Fine control over resources allocated per project
N1GE 6 Interfaces

3rd PARTY SOFTWARE INTEGRATION

- DRMAA 1.0: Job submission
- XML-based monitoring
  > Status command 'qstat' outputs full status info (jobs, hosts, queues) in fully defined XML
- JDBC/ODBC: usage and accounting information
  > DB scheme fully documented
  > Data unencrypted
- Full functionality via CLI: Datacenter automation
Key Competitive Points

- **Scalability (# of jobs and hosts)**
  - Equal or greater than nearest competitor
- **Performance (throughput rate, responsiveness)**
  - Highest in its class (batch scheduling)
- **Openness**
  - Solaris Enterprise System, open source, large and very active user community with many best practises
  - Standards-based APIs and data formats
- **Completeness**
  - Includes analysis and reporting (not expensive add-on)
Key Competitive Points

• Cost Advantage (>250 cpus)
  > License
  > Annual Support
  > Administration
  > All inclusive pricing model (no additional charges)
  > Sun Grid Engine Open-Source solution
    > Expand the existing code to fit your environment
  > Solaris Enterprise System Program
    > Get the binaries at no-charge, but no license+support included
    > Enterprise-wide “unlimited masters+executions” license
      > Highly discounted and providing ultimate deployment flexibility
Key Competitive Points

• Reduction in number of vendors
  > Deal with only 1 vendor for both hardware and software

• Tight integration with Sun portfolio
  > JES Portal Server
  > Sun N1 System Manager

• Ease of Use
  > Multiple User Interfaces