How to Provide Public with Pervasive HPC Service?

Hai Jin

Services Computing Technology and System Lab
Cluster and Grid Computing Lab
School of Computer Science and Technology
Huazhong University of Science and Technology
Wuhan, 43074, China
hjin@hust.edu.cn
Outlines

1. Why do we need HPC Cloud, what are the advantages?

2. What types of HPC applications can run on HPC Cloud?

3. Research Challenges

4. Our Practice
Why do we need HPC Cloud?

- HPC market continues to increase
  - CAGR (Compound Annual Growth Rate) is 6.3%
Why do we need HPC Cloud?

- Market has been moving towards multi-polarization
  - Market share of Departmental and Workgroup will hold 49 percent of the total market in 2014 according to IDC’s forecast
Why do we need HPC Cloud?

• The types of HPC applications are diversified and the number of HPC applications is growing

Traditional HPC
Scientific research
Aerospace
Defense
Oil prospecting
Weather forecast, etc.

Edge HPC
E-health
Finance
Government Informatization
Online games
Social network computing, etc.
Why do we need HPC Cloud?

• Traditional HPC patterns are facing the following questions

  - On-demand dynamic scalable computing resources?
  - More flexible service model?
  - Supporting legacy applications?
  - Ability to self-manage computing resources?
  - Customized HPC environment?
Why do we need HPC Cloud?

• Conclusion
  - High-performance computing is no longer just a demand for scientists and engineers

The answer is HPC Cloud!
Why do we need HPC Cloud?

- **HPC Cloud** is a cloud-based high-performance computing resource management and service delivery model

- **Main features**
  - Utilizing *virtualization technology* to manage and organize computing resources
  - *Virtual machine* serves as the basic unit of resource supply
  - Using *IaaS or PaaS* to provide HPC services
Why do we need HPC Cloud?

1. Flexible customization of computing environment
2. Isolation
3. Dynamic scaling resources
4. Transparent fault-tolerant, etc

• These features are just what traditional HPC patterns need!
Why do we need HPC Cloud?

• **Building HPC cloud** with the concept and technologies of cloud computing offers a good choice to address the problems presented above
HPC Cloud Advantages

Fine-grained management of resources
Reallocation of resource depending on changes of resource demand without affecting the normal operation of the job

Providing IaaS
Users can do whatever he wants!

Providing PaaS
Users can customized HPC platform

Advantages

Improving availability and reliability
Achieving transparent fault-tolerant systems
What types of HPC applications can run on HPC Cloud?

- Although HPC cloud is a cost-effective computing model, it introduces virtualization which may lead to performance degradation of HPC applications.

- Therefore, HPC applications which are persistent in pursuit of high computing power are not suitable for running on HPC Cloud.
What types of HPC applications can run on HPC Cloud?

- Researchers from the LBL have proposed a high-level classification of HPC applications.

[Defining Future Platform Requirements for e-Science Clouds, SoCC’10]
What types of HPC applications can run on HPC Cloud?

**Type one**

**Bulk-Synchronous Large-Scale**
- Running on *large-scale supercomputing centers* across the nation
- Using a large number of processors (often in the order of thousands)
- May have long running jobs

- These jobs are served at supercomputing centers through batch queue systems
What types of HPC applications can run on HPC Cloud?

**Type two**

**Bulk-Synchronous Mid-range**
- Need tens to hundreds of processors
- Some of these applications run at supercomputing centers and backfill the queues

- More commonly, users rely on small clusters that are managed by the scientific groups themselves to satisfy these needs
What types of HPC applications can run on HPC Cloud?

Type three

**Asynchronous Massively Independent**

- Some scientific explorations are performed on the desktop or local clusters and have asynchronous, massively independent computations

- The requirements of such applications are similar to those of the internet applications that currently dominate the cloud computing
What types of HPC applications can run on HPC Cloud?

1. Create virtual private clusters (small and medium scale)
2. Guarantee significant cost savings
3. Provide all of the benefits of full control of the software configuration and availability, etc.

Therefore, the applications of type two and three can run on HPC cloud.

They are the main concerns of the current HPC cloud.
What types of HPC applications can run on HPC Cloud?

- For applications of type one, the situation is more complicated, because the virtualization will influence performance.

  - "Minimal-overhead Virtualization of a Large Scale Supercomputer, VEE2011" shows how careful use of hardware and VMM features enables the virtualization of a large-scale HPC system.
What types of HPC applications can run on HPC Cloud?

- Source: "Minimal-overhead Virtualization of a Large Scale Supercomputer, VEE2011"
What types of HPC applications can run on HPC Cloud?

• For this reason, if user can tolerate the 5% performance loss, applications of type one can run in virtualized large-scale supercomputers

• Foundation of widely HPC Cloud
Research Challenges

- Some important issues need to be resolved

  - Improve efficiency of programming in HPC Cloud environment
  - Guarantee QoS in the competitive HPC Cloud environment
  - Rapid scalability of large-scale virtual HPC resources
  - The reliability of large-scale virtual HPC environment
  - Rapid deployment of virtual HPC environment
  - Reduce the performance overhead caused by virtualization
Challenges (1)

- Reduce the performance overhead caused by virtualization
  - Dedicated VMM (Virtual Machine Monitor) for HPC
  - Dedicated OS
  - Symbiotic relationship between VMM and OS

[SymCall: Symbiotic Virtualization Through VMM-to-Guest Upcalls, VEE2011]
Challenges (2)

• Improving the reliability of virtual HPC environment
  - Live migration mechanism of tightly coupled multiple virtual machines (for example, multiple virtual machines which run MPI jobs)
  - Accuracy of failure prediction needs further improvement
Challenges (3)

• **HPC Cloud resource management**
  - Profiling and Modeling of application performance in HPC Cloud
  - SLA-based virtual high performance infrastructure management mechanisms and strategies
Challenges (4)

• Programming model and software environment in HPC Cloud
  - HPC Cloud-oriented programming model and supporting software environment optimized for virtualization
    • Such as parallel compiler, linker, debugger, parallel libraries
Challenges (5)

• To become a commercial platform, HPC Cloud need to address many other issues
  - security of user data
  - measuring software license
  - standards and norms of platform, and so on
Our Practice - CRANE

• **CRANE** - a cloud platform providing special support for HPC services
Basis of CRANE

• Virtualization based: Xen & KVM etc.
• Infrastructure as a Service which could across multiple clusters and sites
• Platform as a Service which contains
  - HPC cloud platform for academia
  - Webapp cloud platform for business
• Efficient resource management and secure user management
HPC towards Functions in CRANE

Two usage patterns

• PaaS based HPC task management
  • Perfect platform for managed tasks

• IaaS based virtualized HPC clusters
  • Used as normal physical clusters

Manage platform

Virtual cluster

Task life-cycle management middleware

HPC job

Job queue

VM

VM

VM

VM
Task Management Platform

• PaaS for HPC Cloud
  – Submit tasks through web interface
    • Choose suitable computing environment
    • Input needed orders and parameters
  – Automatic tasks life-cycle management
    • Transparent construction of computing environment
    • Automatic task execute and result export
  – Convenient
    • Based on prepared environment templates
Virtual Cluster

• IaaS for HPC Cloud
  - Customizable cluster configuration
    • Amount of VM / VCPU / Mem / etc.
    • Type of storage / network / system / etc.
  - Pre-configured MPI cluster environment
    • Contains most used components
  - Scale by apply new VMs through basic IaaS
    • Add new VMs to exist cluster manually
    • Flexible network management
  - Access through SSH with secret key
Practice with CRANE

• Used in CNGI project as a public cloud platform for 4 typical academic applications with ipv6 environment

• Deployed to 4 grid nodes in universities which contain HUST, THU, PKU and SCUT

• Website:
  - http://www.hustcloud.com
Conclusions

- HPC cloud can be **the most powerful competitors** in a small and medium-sized computing cluster market
- HPC cloud may replace the position of user's private physical cluster
- HPC Cloud provides good approaches and ideas to address the requirements of HPC applications (on-demand computing, burst computing, etc.)
Conclusions

• Perceptions transition from purchasing fixed assets to purchasing service is the key to the rise of HPC Cloud

• HPC cloud with its unique characteristics is sure to have a significant influence on the infrastructure management and use of supercomputing centers

• Operating the Supercomputing Center with the concept of cloud computing will become an important trend
Thank You!