Science in the Cloud
Exploring Cloud Computing for Science
Shane Canon
Moab Con
May 11, 2011
Outline

• Definitions
• The Magellan Project
• Experience and Lessons Learned
• Cloud Misconceptions
• Closing remarks
What is a Cloud?

Definition

According to the National Institute of Standards & Technology (NIST)...

• **Resource pooling.** Computing resources are pooled to serve multiple consumers.

• **Broad network access.** Capabilities are available over the network.

• **Measured Service.** Resource usage is monitored and reported for transparency.

• **Rapid elasticity.** Capabilities can be rapidly scaled out and in (pay-as-you-go)

• **On-demand self-service.** Consumers can provision capabilities automatically.
# What is a cloud?

## Cloud Models

<table>
<thead>
<tr>
<th>Hardware focus</th>
<th>Application focus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infrastructure as a Service (IaaS)</strong></td>
<td><strong>Platform as a Service (PaaS)</strong></td>
</tr>
<tr>
<td>Provisions processing, storage, networks, and other fundamental computing resources. Consumer can deploy and run arbitrary software, including OS.</td>
<td>Provides programming languages and tools. Consumer applications created with provider’s tools.</td>
</tr>
<tr>
<td>• Amazon EC2</td>
<td>• Microsoft Azure</td>
</tr>
<tr>
<td>• RackSpace</td>
<td>• Google AppEngine</td>
</tr>
<tr>
<td><strong>Software as a Service (SaaS)</strong></td>
<td></td>
</tr>
<tr>
<td>Provides applications on a cloud infrastructure. Consumer provides data.</td>
<td></td>
</tr>
<tr>
<td>• Salesforce.com</td>
<td>• Google Docs</td>
</tr>
<tr>
<td></td>
<td>• Application Portals</td>
</tr>
</tbody>
</table>
Co-located at two DOE-SC Facilities

- Argonne Leadership Computing Facility (ALCF)
- National Energy Research Scientific Computing Center (NERSC)
- Funded by DOE under the American Recovery and Reinvestment Act (ARRA)
Magellan Scope

• Mission
  – Determine the appropriate role for private cloud computing for DOE/SC midrange workloads

• Approach
  – Deploy a test bed to investigate the use of cloud computing for mid-range scientific computing
  – Evaluate the effectiveness of cloud computing models for a wide spectrum of DOE/SC applications
Magellan Test Bed at NERSC
Purpose-built for Science Applications

720 nodes, 5760 cores in 9 Scalable Units (SUs) ➔ 61.9 Teraflops
SU = IBM iDataplex rack with 640 Intel Nehalem cores

QDR IB Fabric

18 Login/network nodes
14 I/O nodes (shared)
10G Ethernet
8G FC
ANL

Load Balancer
Network
100-G Router
HPSS (15PB)
<table>
<thead>
<tr>
<th>Purpose</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLUSTER RESOURCES INC.</td>
<td>Mix of node types and queues. Future: Dynamic provisioning, VMs, and virtual private clusters</td>
</tr>
<tr>
<td>Eucalyptus Systems</td>
<td>Can expand based on demand. Supports: VMs, block storage</td>
</tr>
<tr>
<td>Hadoop</td>
<td>MapReduce. Both configured with HDFS</td>
</tr>
</tbody>
</table>
Magellan Research Agenda and Lines of Inquiry

• Are the open source cloud software stacks ready for DOE HPC science?
• Can DOE cyber security requirements be met within a cloud?
• Are the new cloud programming models useful for scientific computing?
• Can DOE HPC applications run efficiently in the cloud? What applications are suitable for clouds?
• How usable are cloud environments for scientific applications?
• When is it cost effective to run DOE HPC science in a cloud?
• What are the ramifications for data intensive computing?
What aspects of Cloud computing are attractive to Scientists? Why are they looking at Clouds?

- Survey users who expressed an interest in clouds.
- Follow-up with deep-dive discussions
Cloud Computing Interest Survey

- User interfaces/Science Gateways: Use of clouds to host science gateways and/or access to cloud
- Hadoop File System
- MapReduce Programming Model/Hadoop
- Cost associativity? (i.e., I can get 10 cpus for 1 hr now or 2 cpus for 5 hrs at the same cost)
- Easier to acquire/operate than a local cluster
- Exclusive access to the computing resources/ability to schedule independently of other groups/
- Ability to control groups/users
- Ability to share setup of software or experiments with collaborators
- Ability to control software environments specific to my application
- Access to on-demand (commercial) paid resources closer to deadlines
- Access to additional resources
Attractive Features of the Cloud

- On-demand access to compute resources
  - Cycles from a credit card! Avoid lengthy procurements.
- Overflow capacity to supplement existing systems
  - Berkeley Water Center has analysis that far exceeds the capacity of desktops
- Customized and controlled environments
  - Supernova Factory codes have sensitivity to OS/compiler version
- Parallel programming models for data intensive science
  - Hadoop (data parallel, parametric runs)
- Science Gateways (Software as a Service)
  - Deep Sky provides an Astrophysics community data base
Can DOE HPC applications run efficiently in the cloud? What applications are suitable for clouds?

- Can parallel applications run effectively in virtualized environments?
- How critical are high-performance interconnects that are available in current HPC systems?
- Are some applications better suited than others?
Application Performance Application Benchmark

Runtime Relative to Magellan (non-VM)

- Carver
- Franklin
- Lawrencium
- Amazon CC
- Amazon EC2

Applications:
- GAMESS
- GTC
- IMPACT
- fvCAM
- MAESTRO256
Application Performance
Application Benchmarks

![Application Performance Graph](image-url)

- MILC
- PARATEC

Runtime relative to Carver:
- Carver
- Franklin
- Lawrencium
- Amazon CC
- Amazon EC2
How usable are cloud environments for scientific applications?

• How difficult is it to port applications to Cloud environments?
• How should users manage their data and workflow?
- **Magellan resources made available to JGI to facilitate disaster recovery efforts**
  - Used up to 120 nodes
  - Linked sites over layer-2 bridge across ESnet SDN link
  - Manual provisioning took ~1 week including learning curve
  - Operation was transparent to JGI users
- **Practical demonstration of HaaS**
  - Reserve capacity can be quickly provisioned (but automation is highly desirable)
  - Magellan + ESnet were able to support remote departmental mission computing
Early Science - STAR

Details

- STAR performed Real-time analysis of data coming from RHIC at BNL
- First time data was analyzed in real-time to such a high degree
- Leveraged existing OS image from NERSC system
- Used 20 8-core instances to keep pace with data from the detector
- STAR is pleased with the results
• Clouds are simple to use and don’t require system administrators.
• My job will run immediately in the cloud.
• Clouds are more efficient.
• Clouds allow you to *ride* Moore’s Law without additional investment.
• Commercial Clouds are much cheaper than operating your own system.
Are Clouds Easy to Use?

From Experience with Magellan we have Learned
• IaaS Clouds can require significant amounts of system administration expertise
• Images must be customized for the application
• No batch environment. No global file system.
• Users must properly secure and protect their images and instances.
• Do we want to turn scientists into system administrators?
Cloud Misconceptions

• Clouds are simple to use and don’t require system administrators.
• My job will run immediately in the cloud.
• Clouds are more efficient.
• Clouds allow you to ride Moore’s Law without additional investment.
• Commercial Clouds are much cheaper than operating your own system.
Is the Cloud Elastic enough for HPC?

Number of cores required to run a job immediately upon submission to Franklin
Cloud Misconceptions

• Clouds are simple to use and don’t require system administrators.
• My job will run immediately in the cloud.
• **Clouds are more efficient.**
• Clouds allow you to *ride* Moore’s Law without additional investment.
• Commercial Clouds are much cheaper than operating your own system.
Assumptions for cost saving from Clouds aren’t true for HPC Centers.

<table>
<thead>
<tr>
<th>EFFICIENCY</th>
<th>Current Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Benefits</td>
<td></td>
</tr>
<tr>
<td>Improved asset utilization (server utilization &gt; 60-70%)</td>
<td>Low asset utilization (server utilization &lt; 30% typical)</td>
</tr>
<tr>
<td>Aggregated demand and accelerated system consolidation (e.g., Federal Data Center Consolidation Initiative)</td>
<td>Fragmented demand and duplicative systems</td>
</tr>
<tr>
<td>Improved productivity in application development, application management, network, and end-user</td>
<td>Difficult-to-manage systems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AGILITY</th>
<th>Current Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Benefits</td>
<td></td>
</tr>
<tr>
<td>Purchase “as-a-service” from trusted cloud providers</td>
<td>Years required to build data centers for new services</td>
</tr>
<tr>
<td>Near-instantaneous increases and reductions in capacity</td>
<td>Months required to increase capacity of existing services</td>
</tr>
<tr>
<td>More responsive to urgent agency needs</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INNOVATION</th>
<th>Current Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Benefits</td>
<td></td>
</tr>
<tr>
<td>Shift focus from asset ownership to service management</td>
<td>Burdened by asset management</td>
</tr>
<tr>
<td>Tap into private sector innovation</td>
<td>De-coupled from private sector innovation engines</td>
</tr>
<tr>
<td>Encourages entrepreneurial culture</td>
<td>Risk-adverse culture</td>
</tr>
<tr>
<td>Better linked to emerging technologies (e.g., devices)</td>
<td></td>
</tr>
</tbody>
</table>

- HPC Centers run at >90% CPU utilization and >90% scheduled utilization.
- HPC Centers partner with Vendors to field cutting edge systems.
- HPC more aggressive with technical risks.

From the Federal Cloud Computing Strategy
# Enterprise IT versus HPC

<table>
<thead>
<tr>
<th></th>
<th>Traditional Enterprise IT</th>
<th>HPC Centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical Load Average</td>
<td>30% *</td>
<td>90%</td>
</tr>
<tr>
<td>Computational Needs</td>
<td>Bounded computing requirements – Sufficient to meet customer demand or transaction rates. (i.e. If you gave a typical business free computing, would they suddenly be able to take advantage of it?)</td>
<td>Virtually unbounded requirements – Scientist always have larger, more complicated problems to simulate or analyze.</td>
</tr>
</tbody>
</table>
Cloud Misconceptions

- Clouds are simple to use and don’t require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to ride Moore’s Law without additional investment.
- Commercial clouds are much cheaper than operating your own system.
The cost of a standard cloud instance has dropped 18% over 5 years. Meanwhile, cores per socket have increased 2x-5x per socket in the same time-frame at roughly constant cost.
Cloud Misconceptions

- Clouds are simple to use and don’t require system administrators.
- My job will run immediately in the cloud.
- Clouds are more efficient.
- Clouds allow you to *ride* Moore’s Law without additional investment.
- Commercial Clouds are much cheaper than operating your own system.
Cost of NERSC in the Cloud

<table>
<thead>
<tr>
<th>Component</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Systems (1.38B hours)</td>
<td>$180,900,000</td>
</tr>
<tr>
<td>HPSS (17 PB)</td>
<td>$12,200,000</td>
</tr>
<tr>
<td>File Systems (2 PB)</td>
<td>$2,500,000</td>
</tr>
<tr>
<td>Total (Annual Cost)</td>
<td>$195,600,000</td>
</tr>
</tbody>
</table>

Assumes 85% utilization and zero growth in HPSS and File System data. Doesn’t include the 2x-10x performance impact that has been measured. This still only captures about 65% of NERSC’s $55M annual budget. **No consulting staff, no administration, no support.**
Where are (commercial) clouds effective?

• Individual projects with high-burst needs.
  – Avoid paying for idle hardware
  – Access to larger scale (elasticity)
  – Alternative: Pool with other users (condo model)

• High-Throughput Applications with modest data needs
  – Bioinformatics
  – Monte-Carlo simulations

• Infrastructure Challenged Sites
  – Facilities cost >> IT costs
  – Consider the long-term costs

• Undetermined or Volatile Needs
  – Use Clouds to baseline requirements and build in-house
What HPC Can Learn from Clouds

• Need to support surge computing
  – Predictable: monthly processing of genome data; nightly processing of telescope data
  – Unpredictable: computing for disaster recovery; response to facility outage

• Support for tailored software stack

• Different levels of service
  – Virtual private cluster: guaranteed service
  – Regular: low average wait time
  – Scavenger mode, including preemption
Potential Role of Resource Managers for the Cloud

- Dynamically provision cloud resources based on demand.
- Apply a cost model to utilize the most cost effective resource.
- Dynamically provision local resources to provide custom images.
- Data locality-aware scheduling (i.e. MapReduce style)
What should an HPC Cloud Solution Look Like?

• High-performance interconnect (high bandwidth, low latency) with fast access from the application
• Fast access to a high-performance file system
• No penalty to gather resources
• Non-Virtualized/bare-metal?
Is an HPC Center a Cloud?

- Resource pooling.
- Broad network access.
- Measured Service.
- Rapid elasticity.
  - Usage can grow/shrink; pay-as-you-go.
- On-demand self-service.
  - Users cannot demand (or pay for) more service than their allocation allows
  - Jobs often wait for hours or days in queues

HPC Centers ?

- ✓
- ✓
- ✓
- ✓
- ✗
• Cloud computing is a business model
• It can be applied to HPC systems as well as traditional clouds (ethernet clusters)
• Can get on-demand elasticity through:
  – Idle hardware (at ownership cost)
  – Sharing cores/nodes (at performance cost)
  – Scheduling policies (pre-emption)
Closing Remarks

• Cloud Computing is changing the face of computing
• HPC Centers can learn new tricks from the Cloud Computing space
• Cloud Computing is not a panacea: they can be more difficult to use, slower and more expensive than in-house solutions.
Acknowledgements

• Lavanya Ramakrishnan
• Iwona Sakrejda
• Tina Declerck
• Others
  – Keith Jackson
  – Nick Wright
  – John Shalf
  – Krishna Muriki
  (not picture)
Thank you!

Contact Info:
Shane Canon
Scanon@lbl.gov
magellan.nersc.gov

Kathy Yelick
KAYelick@lbl.gov