A Cloudy Future Panel at CCGSC '08

Philip Papadopoulos

15 Sept 2008

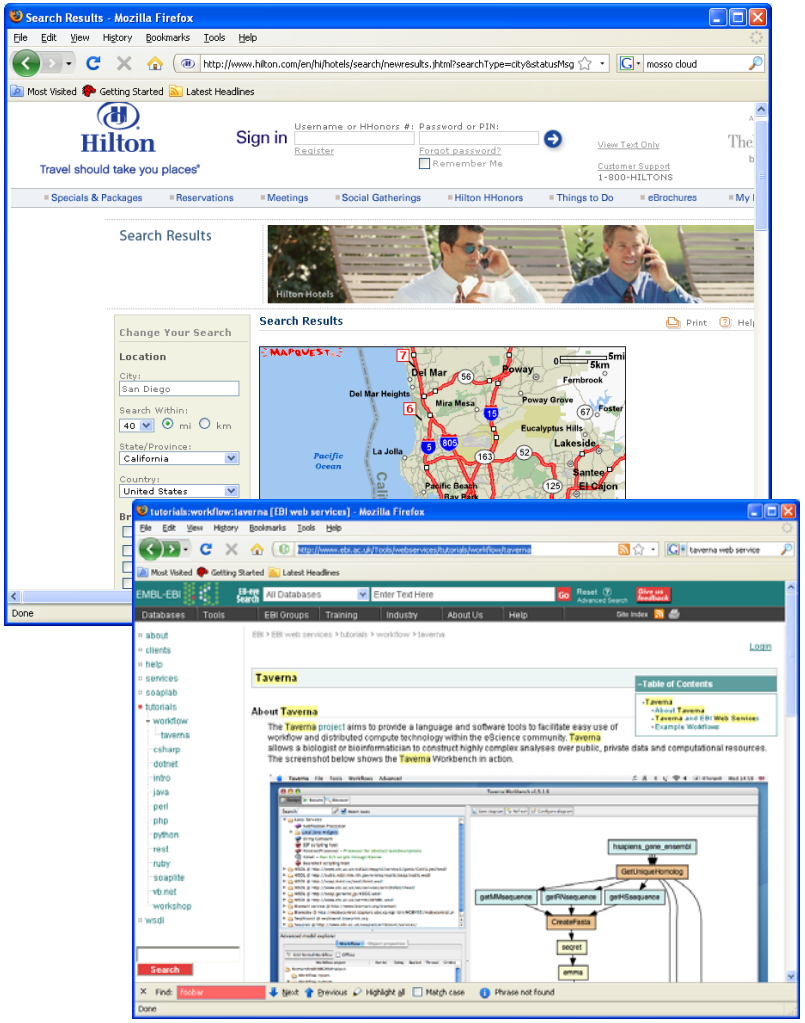
Topics

- Defining a usable subset of Cloud Computing
- Clouds for HPC? A Long Way to Go but not for the reasons you think
  - It's not about the performance
- It's My cluster.
  - Campus cloud provisioning. Extending My Cluster using Rocks.

Possible Definitions of Clouds

- Anything that is a distributed resource
- Another word for Grid or High-Throughput Computing.
- Write your program (workflow) by calling web-services
  - Integrate Google Maps, Weather.com radar, etc
- Software as a Service (SaaS)
  - Mosso Hosting "Cloud"
  - Google Docs (Can I do Computing Here?)
- Infrastructure as a Service (IaaS)
  - Sun's Network.com
  - Amazon EC2 – Run your Software Stack .

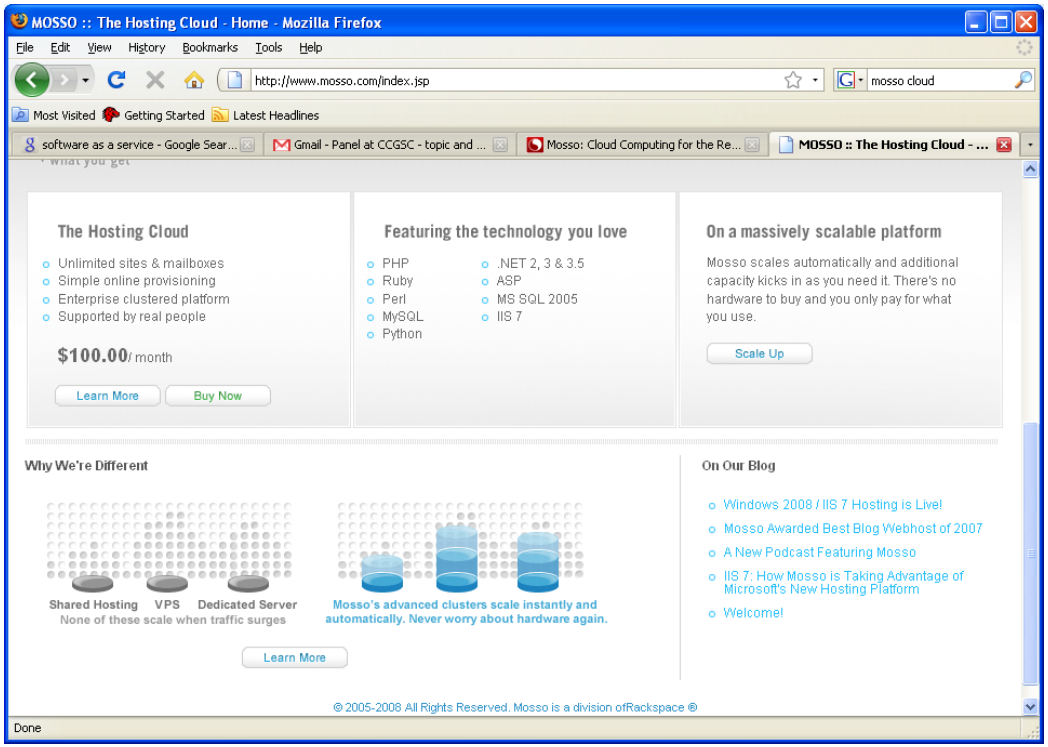
Applications/Workflows using web services.



- They have using web services for 4+ years
- What about Condor, Globus-based, Avaki-based Grids?

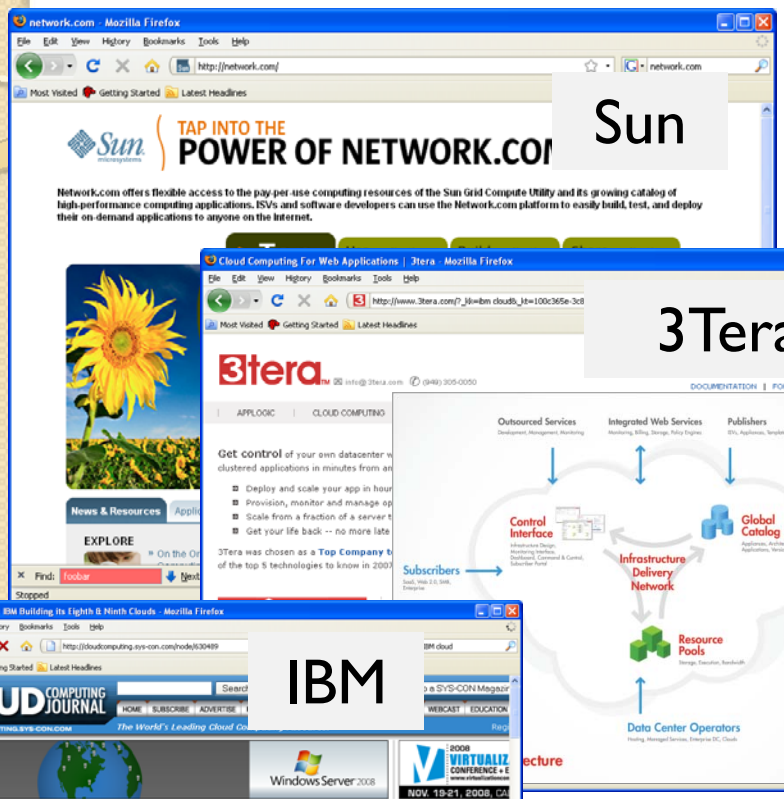
Calling by different name does not a cloud make

SaaS



- If you web server uses (a combination) of 9 software packages, then this works
- Just because you don't know exactly what the hardware is, doesn't make it a cloud.
- Verdict: Mosso (and things like it) are not clouds (opaque to inspection)

IaaS – Cloud with the most promise



Run (virtual) computers to solve your problem, using your software

Cloud Definition

- Remote resources that
  - Run your application in a way that is comfortable to you (interactive, batch-scheduler, web/grid services, …)
  - Complete software stack OS → Application can be inspected
    - Controlled completely, if desired

Clouds for HPC?

- Yes. Inevitable.
- Why? (incomplete list)
  - Bring your software environment with you
  - Performance of Hypervisor:Guest interface improves
  - EC2 monetized CPU hours (it's good business)
    - So cheap that it's "free" for small-scale use.
    - Loads at about 50% are financial break even for cloud provider
    - Not worried about supporting your software.
- Will force HPC centers to support the IaaS paradigm
  - Expensive for traditional HPC providers (eg. Teragrid) to support N x K versions of software packages
  - Machines designed for HPC and IaaS can have much higher I/O rates than EC2

High-level observations

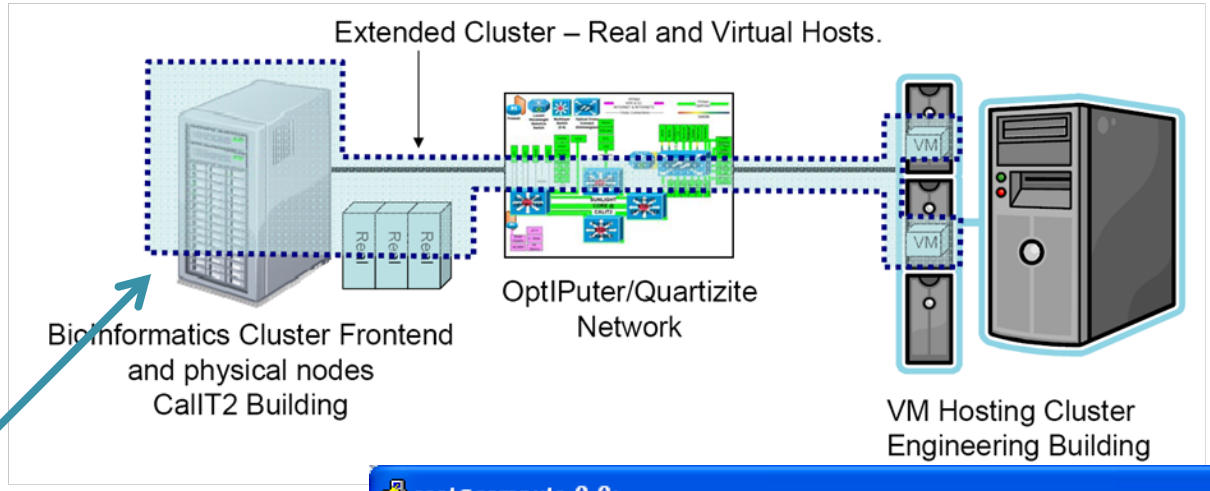
- "Clouds. It's all new, have to learn a whole new vocabulary, software, a whole new way of doing things"
  - This is doomed to failure
- IaaS Clouds will turn all scientists into sysadmins
  - If not addressed, Doomed to failure

- 1992: Vaidy, Bob M. Al, Jack= PVM3. Learn from them . PVM provided a personal "cluster". The personal notion was right on target.

The Neonatal stages of clouds

- (Nearly) everybody focusing on starting virtual machine images
  - E.g., RightScale "dashboard", 3Tera and others
- How are the contents of the VM Defined?
  - Start with "base image," use sysadmin super-skills to make it what you want.
    - Run and scream as we re-invent 1995 Beowulf
  - Use a predefined image on EC2, take what you get.
- What if I want a cluster of VMs with my software?
- What if I want a cluster of 1000 VMs?
- How are the last two different from a real cluster?

Campus Cloud: Cluster Extension



- VMs: Software/OS defined by the frontend:
  - Users, file system mount, queuing system, software versions, etc

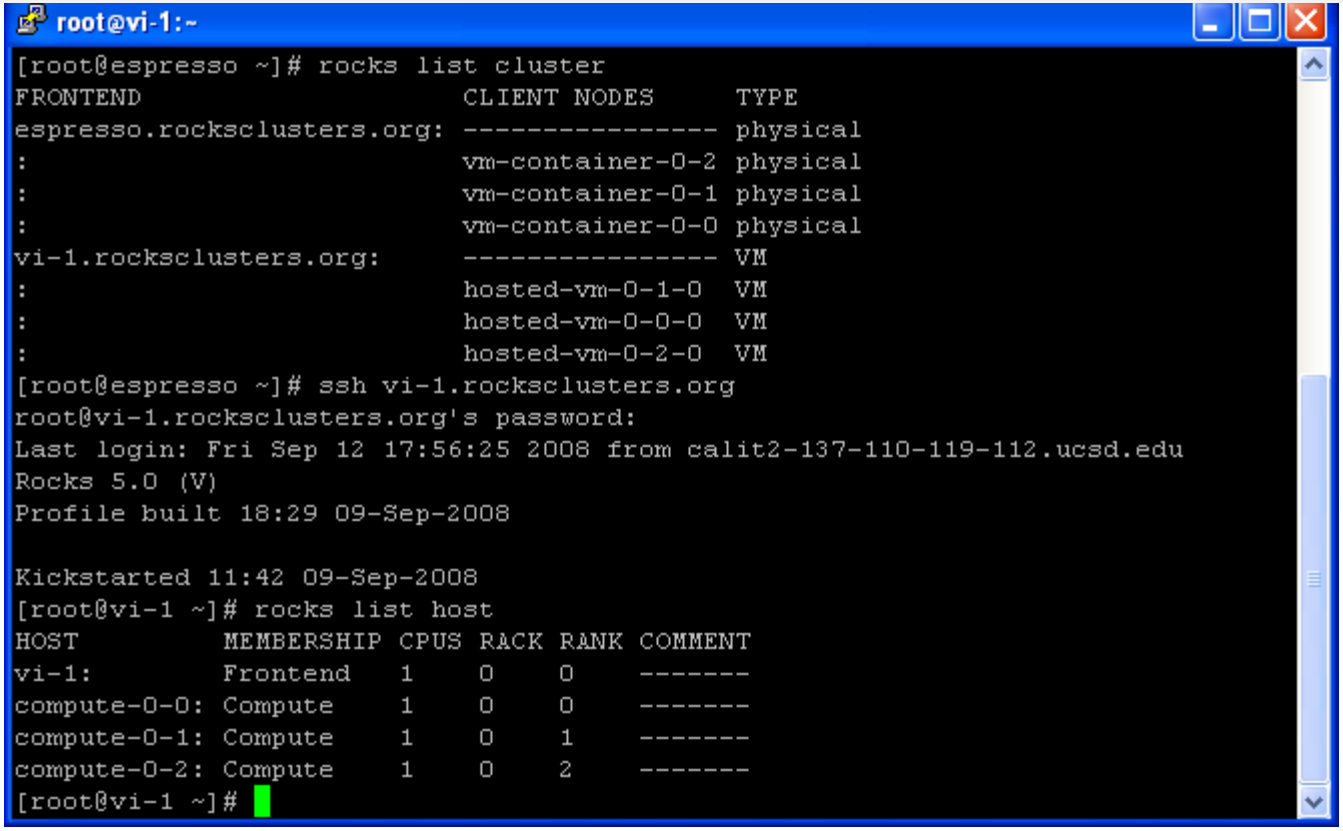
```
root@compute-0-0:~
login as: root
root@ikelite2.rocksclusters.org's password:
Last login: Mon Sep 15 14:05:23 2008 from 12-218-193-156.client.mchsi.com
Rocks 5.0 (V)
Profile built 23:23 01-Jul-2008

Kickstarted 16:44 01-Jul-2008
[root@ikelite2 ~]# ssh compute-0-0
Last login: Mon Sep 15 14:05:39 2008 from ikelite2.local
Rocks Compute Node
Rocks 5.0 (V)
Profile built 12:15 15-Sep-2008

Kickstarted 12:30 15-Sep-2008
[root@compute-0-0 ~]# lspci
[root@compute-0-0 ~]# lsmod | grep xen
xennet                  28617  0 [permanent]
xenblk                  19473  5
[root@compute-0-0 ~]#
```

Clusters in Clusters

```
# rocks create cluster vi-1 137.110.119.137 numnodes=3
```



- Espresso is a VM hosting Cluster
- Vi-1 is virtual – configured to the needs of the owner.