



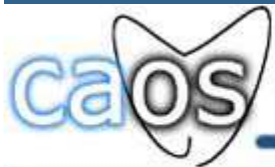
An Efficient Use of Virtualization in Grid/Cloud Environments.

Arindam Choudhury

Supervised by:

Elisa Heymann

Miquel A. Senar



Computer Architecture & Operating Systems

UAB
Universitat Autònoma
de Barcelona



Index

- Introduction
- Motivation
- Objective
- State of Art
- Proposed Solution
- Experimentations and Results
- Conclusions and Future Work





Introduction

Job Requirement:

- Custom Execution Environment:
 - Operating system
 - Specific libraries
- Administrator Privilege





Introduction(contd.)

Cluster Computing

Grid Computing

- Resources are not centrally controlled.
- Resources are from dispersed location.

Cloud Computing

- Easily usable and accessible virtualized resources.
- Illusion of infinite resources.
- Pay-per-use.





Introduction(contd.)

Cluster Computing

Job Requirement:

- Custom Execution Environment:
 - Operating system
 - Specific libraries
- Administrator Privilege





Introduction(contd.)

Grid Computing

Job Requirement:

- Custom Execution Environment:
 - Operating system
 - Specific libraries
- Administrator Privilege





Introduction(contd.)

Cloud Computing

Job Requirement:

- Custom Execution Environment:
 - ✓ • Operating system
 - ✓ • Specific libraries
- ✓ • Administrator Privilege

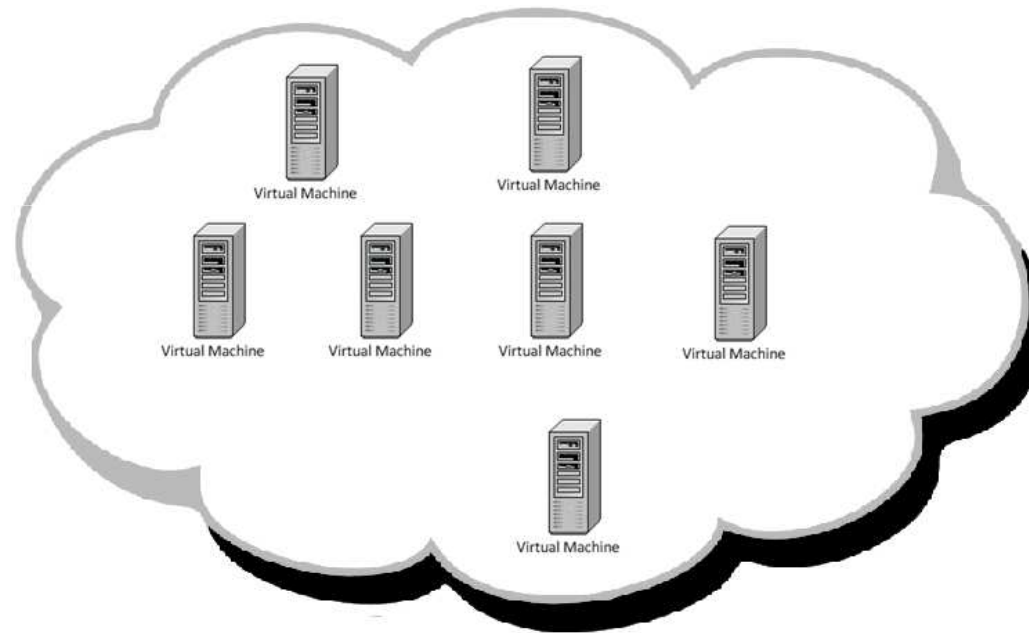




Motivation

How Cloud Computing Provides Custom Execution Environment:

- Virtualization



But Cloud does not support batch job submission



Motivation(contd.)

Virtualization in Grid/Cloud Computing

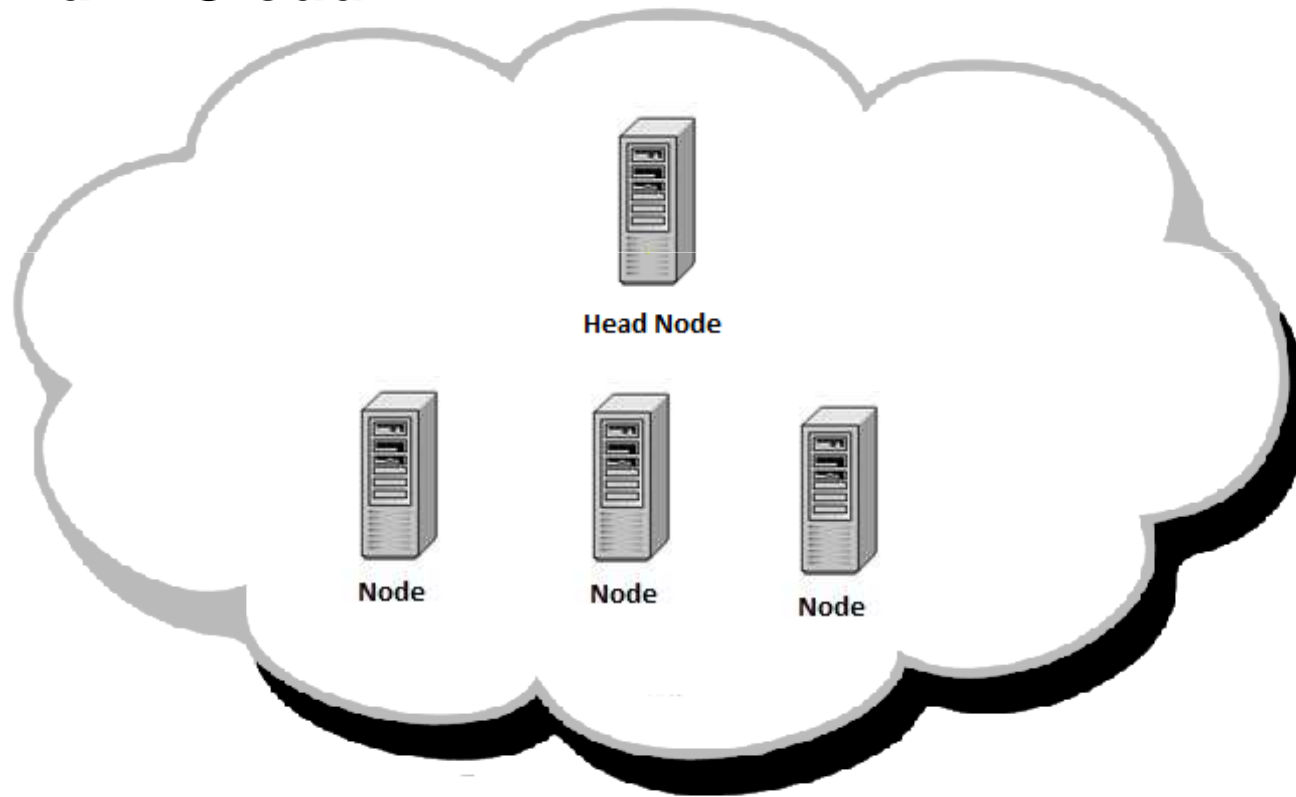
- **Grid in Cloud**
- **Cloud in Grid**
 - Virtual Machine as a Node
 - Virtual Machine as a Job





Motivation(contd.)

Grid in Cloud





Motivation (contd.)

Cloud in Grid: Virtual Machine as a Node

Virtual Organization



Node



Node



Node



Node



Virtual Machine

● ● ● | Motivation (contd.)

Cloud in Grid: Virtual Machine as a Node

Job Requirement:

- Custom Execution Environment:
 - ✗ • Operating system
 - ✗ • Specific libraries
- ✓ • Administrator Privilege





Motivation (contd.)

Cloud in Grid: Virtual Machine as a Job

Submit machine



Virtual Machine

Execute machine

● ● ● | Motivation (contd.)

Cloud in Grid: Virtual Machine as a Node

Job Requirement:

- Custom Execution Environment:
 - ✓ • Operating system
 - ✓ • Specific libraries
- ✓ • Administrator Privilege

But user need to submit the Virtual Machine each time he wants to run job.





Objective

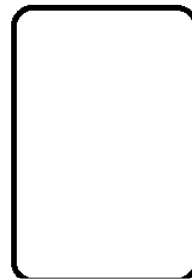
- Provide custom environment to users to run their jobs through virtual machine.
- Users will be able to reuse their virtual machines.
- Only user will be able to use the virtual machine.





State of Art

Globus Virtual Workspaces



Workspace Client

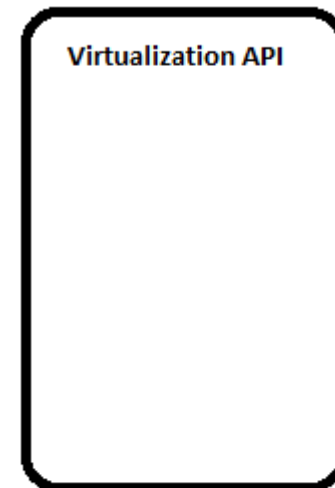
User request to power on virtual machine



VMI Repository



Head Node



Workspace Control





State of Art(contd.)

CernVM

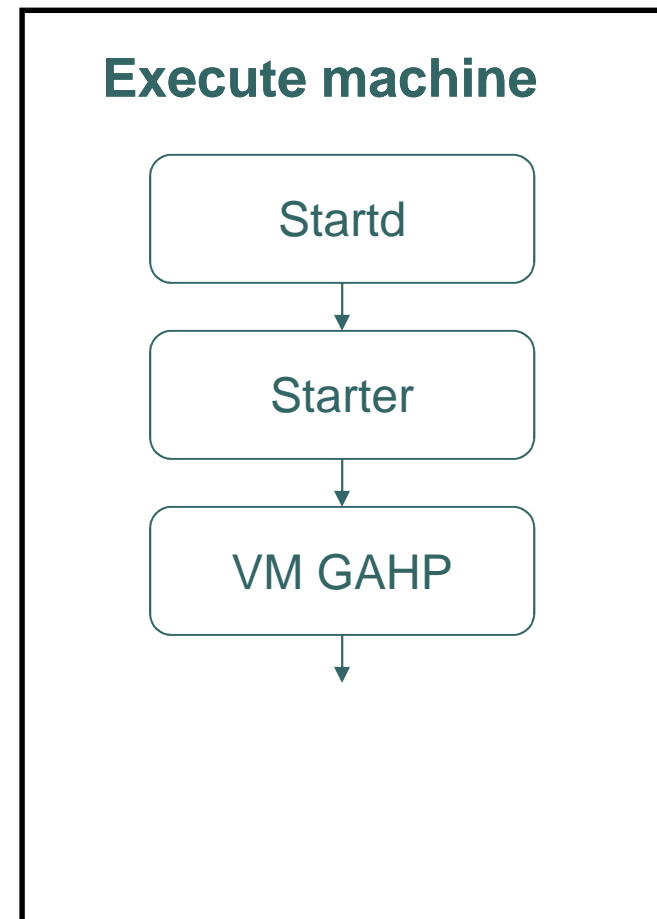
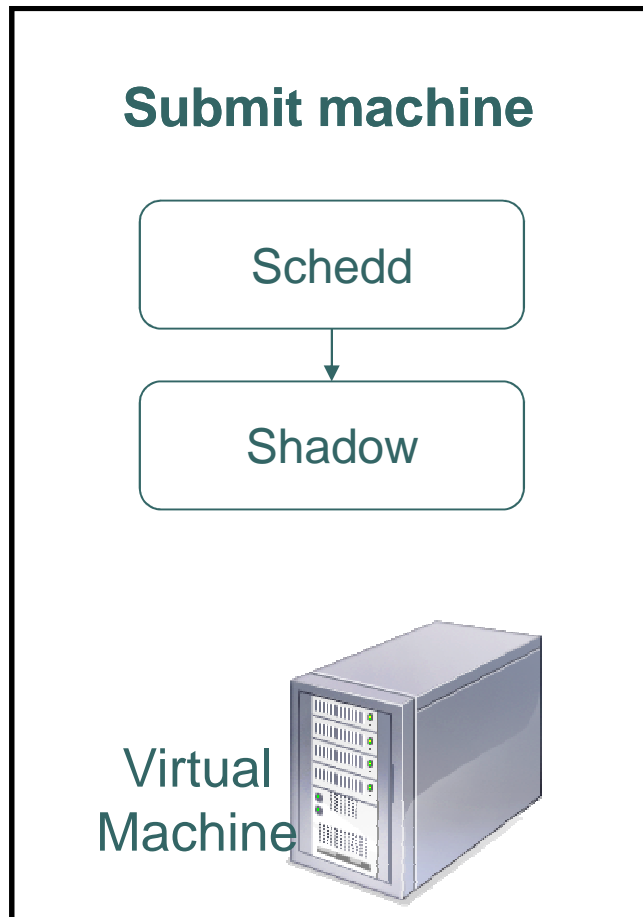
- **Minimal OS:** contains only a minimal operating system required to bootstrap and initiate the experiment software.
- **CernVM-FS:** decouples the operating system from the experiment software life cycle. Pre-built and configured experiment software releases are centrally published. The releases are distributed efficiently on a large scale via a hierarchy of proxy servers or content delivery networks.
- **Configuration and contextualization interfaces:** configures the virtual machines to run correctly on the remote system. It mounts the software repositories before running the job.





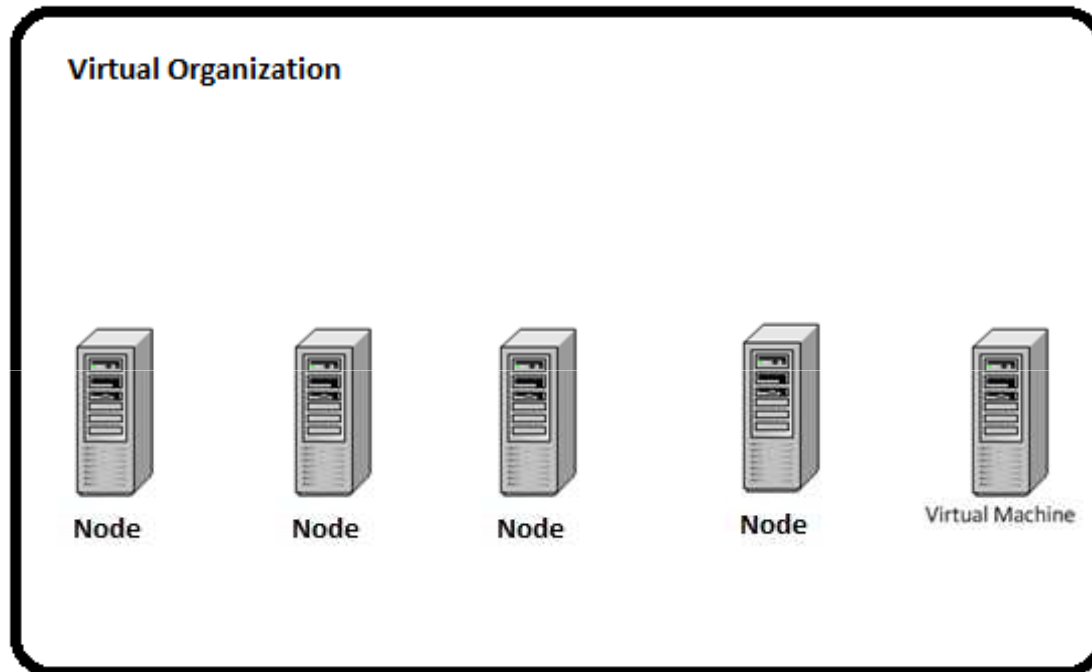
State of Art(contd.)

Condor VM Universe





Proposed Solution



User custom virtual machines will join the pool on user demand.

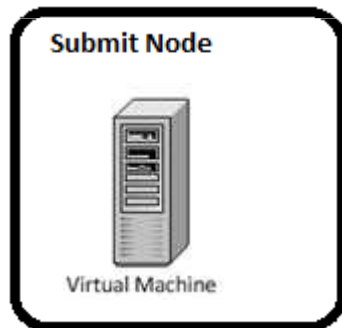
The virtual machine is private to the user.



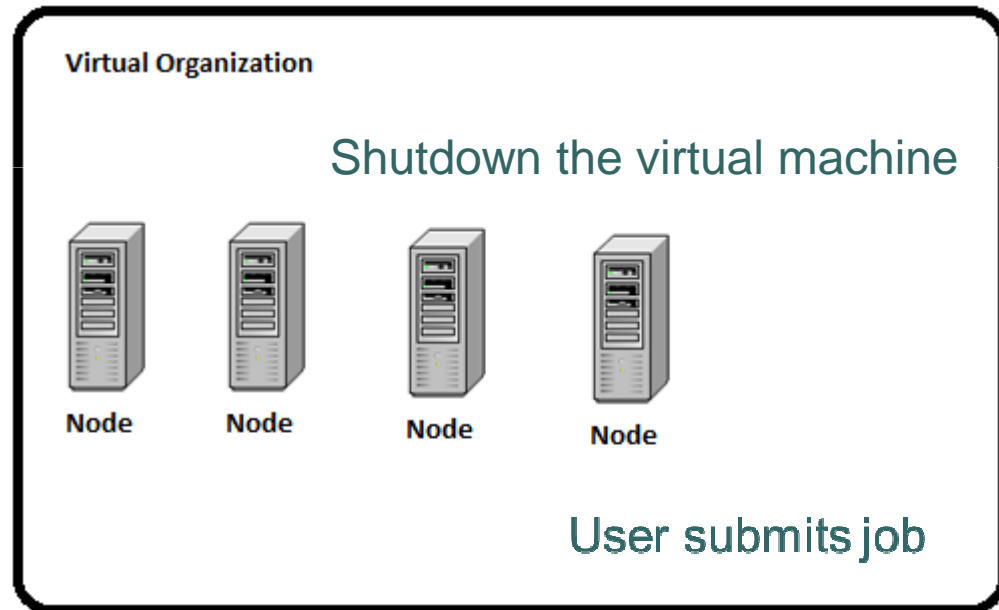
Proposed Solution (contd.)

Scenario 1: Remote Transfer

Sends the virtual machine



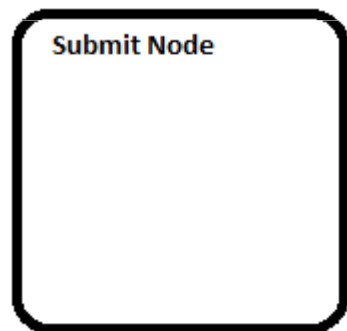
User creates custom virtual machine





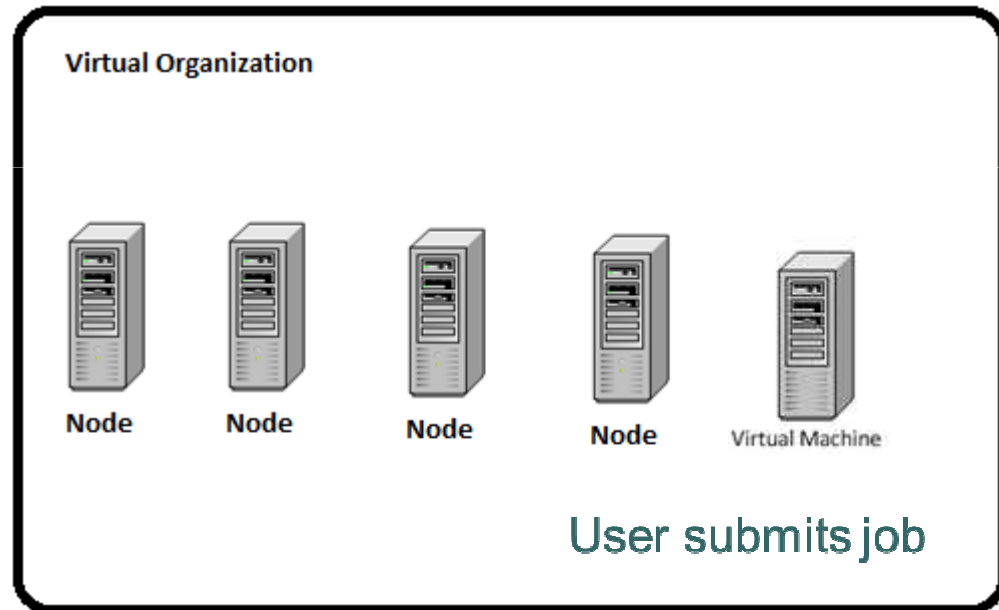
Proposed Solution (contd.)

Scenario 2: Pre- configured



Power On

Shutdown





Used Applications

- Middleware
 - Condor
- Virtualization
 - VMware Server 1
- Virtualization API
 - VMware-VIX





Experimentations & Results

Condor Pool:

```
[condor@aopcach ~]$ condor_status
```

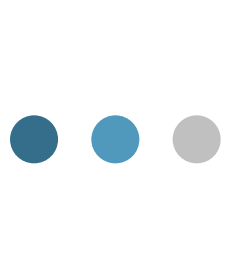
Name	OpSys	Arch	State	Activity	LoadAv	Mem	ActvtyTime
slot1@aopcach.uab.	LINUX	INTEL	Unclaimed Idle	0.020	1762	0+00:25:04	
slot2@aopcach.uab.	LINUX	INTEL	Unclaimed Idle	0.000	1762	0+00:25:05	
aow4grid.uab.es	LINUX	INTEL	Unclaimed Idle	0.550	1518	0+00:55:04	
aow5grid.uab.es	LINUX	INTEL	Unclaimed Idle	0.820	502	0+01:33:26	

aow4grid.uab.es: Condor Host Node, Condor VM Universe is configured on this node.

aow5grid.uab.es: virtual machine image configured on this node.

aopcach.uab.es: condor node with two slots.





Experimentations & Results (contd.)

Condor Pool with Virtual Machine Node:

Every 1.0s: condor_status

Wed Jun 29 17:08:09 2011

Name	OpSys	Arch	State	Activity	LoadAv	Mem	ActvtyTime
slot1@aopcach.uab.	LINUX	INTEL	Unclaimed	Idle	0.010	1762	0+02:45:04
slot2@aopcach.uab.	LINUX	INTEL	Unclaimed	Idle	0.000	1762	0+02:45:05
★ aow12grid.uab.es	LINUX	INTEL	Unclaimed	Idle	2.860	157	0+00:00:04
aow4grid.uab.es	LINUX	INTEL	Unclaimed	Idle	0.010	1518	0+00:31:16
aow5grid.uab.es	LINUX	INTEL	Unclaimed	Idle	1.390	502	0+00:01:26

aow12grid.uab.es: Virtual Machine node





Experimentations & Results (contd.)

Making Virtual Machine Node Private:

- Condor checks START attribute from the configuration to decide when to start running jobs.
- START is modified to enforce that only jobs submitted from user's node can be executed on the virtual machine.





Experimentations & Results (contd.)

Scenario 1: Remote Transfer

- Create virtual machine
- Make the virtual machine private
 - Modify START attribute
- Power On: Submitting Condor VM Universe job
- Virtual machine joins the pool
- Submit jobs
- Power Off: Kill VM Universe job

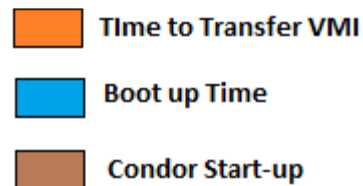




Experimentations & Results (contd.)

Scenario 1: Remote Transfer

Time to be live on the pool:



As in this scenario, virtual machine image need to be transferred to the execution node, it takes 15 minutes on average to be live on the pool.





Experimentations & Results (contd.)

Scenario 1: Remote Transfer Network Usage:

	19.1MB	38.1MB	57.2MB	76.3MB	95.4MB
aow5grid.uab.es => aow4grid.uab.es			85.1MB	80.6MB	83.0MB
			<=	1.50MB	1.44MB
aow5grid.uab.es => aopcach.uab.es			6.99Kb	23.0Kb	13.3Kb
			<=	17.0Kb	8.19Kb
aow5grid.uab.es => smtp-in2.uab.es			0b	2.47Kb	632b
			<=	0b	4.64Kb
aow5grid.uab.es => smtp-in.uab.es			628b	2.37Kb	608b
			<=	988b	4.64Kb
aow5grid.uab.es => dns.uab.es			1.41Kb	1.41Kb	462b
			<=	5.40Kb	5.05Kb
aow5grid.uab.es => xpv-uab-cat.uab.es			2.97Kb	2.42Kb	3.58Kb
			<=	160b	160b
158.109.255.255 => etse-70-67.uab.es			0b	0b	0b
			<=	2.71Kb	555b
255.255.255.255 => ee-68-229.uab.es			0b	0b	0b
			<=	0b	267b
255.255.255.255 => etse-75-70.uab.es			0b	0b	0b
			<=	0b	217b
<hr/>					
TX:	cumm:	5.07GB	peak: 85.9MB	rates: 85.2MB	80.6MB
RX:		91.5MB	1.57MB	1.53MB	1.47MB
TOTAL:		5.16GB	87.3MB	86.7MB	82.1MB

The VMI Transfer uses around 80 Mbps network bandwidth.





Experimentations & Results (contd.)

Scenario 1: Remote Transfer

Fault Tolerance:

The virtual machine runs as a condor job. In case of high load or low memory, condor can power off or migrate the virtual machine. Condor can checkpoint virtual machines.



● ● ● | Experimentations & Results (contd.)

Scenario 1: Remote Transfer

- Pros:
 - Powering on and powering off is simple.
 - User is the owner. Privatizing is easy.
 - Fault-tolerance.
- Cons:
 - Time to join the pool.
 - High network bandwidth.
 - Limited to LAN.





Experimentation & Results

Scenario 2: Pre- configured

- Virtual machine is configured on the execution node
- Power On: execute power on script on execution node
- Virtual machine joins the pool
- Make the virtual machine private
 - Modifying START attribute dynamically
- Submit jobs
- Power Off: execute power off script on execution node





Experimentation & Results

Scenario 2: Pre- configured

- The virtual machine is shared between users. So, high mutual understanding between users is needed.
- The time to become live on pool is small.
- The virtual machine needs to be made private dynamically.
- Network bandwidth usage is normal.





Experimentations & Results (contd.)

Scenario 2: Pre- configured

Fault Tolerance:

The virtual machine runs completely separated from the host operating system. They do not know each others load or usage. So, in case of high load on host system both the virtual machine and the host will crash.





Experimentations & Results (contd.)

Scenario 2: Pre- configured

- Pros:
 - Powering on and powering off.
 - Time join the pool.
 - Network bandwidth usage.
- Cons:
 - Privatizing.
 - Other users can influence user experience.
 - Fault-tolerance.



Experimentations & Results (contd.)

Summary:

	Scenario 1	Scenario 2
Power On	✓	✓
Time to be live on Pool		✓
Privatizing	✓	
Power Off	✓	✓
Fault-Tolerance	✓	
Network Usage		✓
Other user's influence	NONE	YES



● ● ● | Conclusions & Future Work

Conclusions:

- Virtualization can be used in Grid to offer custom job execution environment to the users.
- The virtual machine can be submitted as a job or can be configured previously on execution node.
- Transparent method is required to power on, power off and privatize the virtual machine.



● ● ● | Conclusions & Future Work (contd.)

Future Work:

- Simulating using CloudSim to measure impact on large cluster/Grid.
- Scenario 1
 - Support for static NIC address.
- Scenario 2
 - Information sharing between host and guest system.
 - Portal based management of virtual machines.





THANK YOU



Computer Architecture & Operating Systems

UAB
Universitat Autònoma
de Barcelona



Bibliography

- Ian Foster and Carl Kesselman, *The Grid 2: Blueprint for a New Computing Infrastructure.*: Morgan Kaufmann Publishers Inc., 2003.
- Michael Armbrust et al., "Above the Clouds: A Berkeley View of Cloud Computing," 2009.
- Sean Campbell and Michael Jeronimo, *Applied Virtualization Technology: Usage Models for IT Professionals and Software Developers.*: Intel Press.
- Renato J. O., Peter A. Dinda, and José A. B., "A Case For Grid Computing On Virtual Machine," , 2003, pp. 550-559.
- Rajkumar Buyya, Chee Shin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities," *Computing Research Repository*, vol. abs/0808.3, pp. 5-13, 2008.
- Katarzyna Keahey, Ian T. Foster, Timothy Freeman, and Xuehai Zhang, "Virtual workspaces: Achieving quality of service and quality of life in the Grid," *Scientific Programming*, vol. 13, pp. 265-275, 2005.
- Lizhe Wang, Gregor Von Laszewski, Marcel Kunze, Jie Tao, and Jai Dayal, "Provide Virtual Distributed Environments for Grid computing on demand," *Advances in Engineering Software*, vol. 41, pp. 213-219, 2010.
- A. Agarwal et al., "Deploying HEP applications using Xen and Globus Virtual Workspaces," *Journal of Physics: Conference Series*, vol. 119, 2008.
- P. Buncic, C. Aguado Sánchez, J. Blomer, A. Harutyunyan, and M. Mudrinic, "A practical approach to virtualization in HEP," *The European Physical Journal Plus*, vol. 126, pp. 1-8, 2011, 10.1140/epjp/i2011-11013-1.
- Condor Project Homepage. [Online]. <http://www.cs.wisc.edu/condor/>
- VMware server. [Online]. http://downloads.vmware.com/d/info/datacenter_downloads/vmware_server/1_0
- VIX API. [Online]. <http://www.vmware.com/support/developer/vix-api/>