Bioinformatics simulations by means of running VMs as a grid jobs in the MolDynGrid Virtual Lab

Andrii Salnikov

National Taras Shevchenko University of Kyiv
Institute of Molecular Biology and Genetics, NAS of Ukraine
VM as a Grid job - Goals

• Use powerful distributed hardware resources (with accelerated virtualization support) of the already deployed Grid infrastructure

• Use this resources as PaaS (interactive access)
  – Ukraine’s Medgrid VO case: interactive medical data (stored on grid storages) analyses using software for OS Windows

• Bring your own software environment with VM (computational jobs)
  – MolDynGrid VO case: bioinformatics simulations with different builds of software
VM as a Grid job - Implementation

- **Client side:** ordinary job submission
  – Let’s start with VM submission live demo!

```bash
&
(jobName="Rainbow Cloud/VM Test")
(runTimeEnvironment="CLOUD/VM")
(executable="/bin/true")
(rsl_substitution="CF_LOCATION" "lfc://lfc.grid.org.ua/grid/testbed.univ.kiev.ua/cloud-test")
("SYSTEM_IMAGE" "vm.img")
("SYSTEM_IMAGE_DESCRIPTION" "$(SYSTEM_IMAGE).description")
)
(inputFiles="$(SYSTEM_IMAGE) $(CF_LOCATION)/$(SYSTEM_IMAGE) "cache=copy" ")
("$(SYSTEM_IMAGE_DESCRIPTION) $(CF_LOCATION)/$(SYSTEM_IMAGE_DESCRIPTION) ")
("nfstest_1MB.log" $(CF_LOCATION)/nfstest_1MB.log)
)
(wallTime="720")
(memory="2048")
(count="2")
(disk="4096")
(environment="CLOUD_VM_SYSIMAGE" "$(SYSTEM_IMAGE))
("CLOUD_VM_DESCRIPTION" "$(SYSTEM_IMAGE_DESCRIPTION)"
("CLOUD_DATA_METHOD" "DISK")
("RUNTIME_PASSWORD" "dynamic")
("NOTIFY_EMAIL" "user@grid.org.ua")
)

Demo: [http://youtu.be/-OgeQkUI2LQ](http://youtu.be/-OgeQkUI2LQ)
VM as a Grid job - Highlights

• **Server side:** developed **Rainbow (“ARC in the Cloud”)** framework to support VM lifecycle
  
  – Project origins is interactive access to platform
  – Runs over Nordugrid ARC in Ukraine (wrappers to run on CREAM CE developed)
  – Extended to transparently support computational jobs
  – Pilot mode added for more flexibility

Rainbow at a Glance: [http://goo.gl/ifC6cl](http://goo.gl/ifC6cl)
VM as a Grid job – How it works

EGI Community Forum 2015, Andrii Salnikov
Rainbow framework components

RunTime Environment
- Pass job info to WN
- VM life cycle handling on WN

Rainbow WN Components
- Helpers for networking interfaces creation
- Helpers for disk image operations
- More helpers for notifications/pilots/staging

Rainbow Gateway Helper
- MAC-addresses leasing for VMs and further MAC-IP bindings
- Network connectivity testing
- Port forwarding and Mailing functions
Rainbow framework install for EL5/EL6/EL7

<table>
<thead>
<tr>
<th>Component</th>
<th>Package Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARC CE</td>
<td>yum install rainbow-RTE-CLOUD-VM</td>
</tr>
<tr>
<td>EMI WN</td>
<td>yum install rainbow-WN-essentials +qemu-kvm +sudo</td>
</tr>
<tr>
<td>Gateway</td>
<td>yum install rainbow-GW-helper +dhcp +iptables +xinetd/systemd</td>
</tr>
<tr>
<td>Wrappers for CREAM CE</td>
<td>yum install rainbow-RTE-CLOUD-VM-CREAM</td>
</tr>
<tr>
<td></td>
<td>yum install rainbow-WN-CLOUD-VM-CREAM</td>
</tr>
</tbody>
</table>

All required packages are available from [Grid Blackjack repo](#)
Rainbow framework configuration

GW Helper: /etc/rainbow-helper.conf:

- **DHCP_LEASES_FILE**=/var/lib/dhcpd/dhcpd.leases
- **DHCP_CONFIG_FILE**=/etc/dhcp/dhcpd.conf

- **RAINBOW_STATEDIR**=/var/spool/rainbow-helper
- **RAINBOW_LIBEXEC**=/usr/libexec/rainbow
- **RAINBOW_MAC_BASE**=52:54:00:c1:0d:00

- **FWD_START_PORT**=30000
- **FWD_END_PORT**=40000
- **FWD_WAN_IP**=194.44.249.94

- **RAINBOW_MAILFROM**=rainbow@grid.org.ua

WN: /etc/rainbow.conf:

- **NETHELPER_SOCKET**="10.75.15.254 10108"
- **VM_NET_TYPE**="direct"
- **VM_NET_INT**="eth0"
- **VM_NET_CHECK_INTERVAL**=15
- **VM_NET_CHECK_COUNT**=20

- **HV_METHOD**="qemu"

+add grid-users allowed to run VMs to “rainbow” group

Simple configuration of cluster infrastructure addresses

No default configuration for CE except default paths.
**Ordinary Job**

- Specify executable
- Request software environment
- Specify data-staging

**Rainbow Job Wrap**

- Automatically pass job environment to VM
- Pass data to/from VM
- `rainbow-VM-jobwrap` service to start job execution
Rainbow JobWrap Data Staging

1. Job Description
2. Stage-in
3. Session Dir
4. Build image with Files
5. Second disk
6. Get results from image
7. ARC Stage-out

ARC

WN

Extract Files

Virtual Machine

One-by-one TAR all

ISO DISK VVFAT

Disk Image
MolDynGrid Virtual Lab Infrastructure

- Own grid-site (UA-IMBG) for computations and development
  - Production ARC CE
  - CREAM CE for development
- 80TB(own)+20TB grid-storage
  - MD trajectories database
  - Jobs data stage-in and stage-out
- MolDynSub CLI for jobs submission
- Web Portal for jobs submission and monitoring
MoldynGrid Virtual Lab Researches

More info on poster #7 from Oleksandr Savytskyi
MolDynSub CLI

Version <= 2.1

- Legacy ARC submission only
- GROMACS MD oriented
- Old portal backend

Version 3.x

- New modular design
- Different software modules for whatsoever (including GPGPU support)
- Different infrastructures support (ARC, CREAM, Rainbow, Gloria)
- New portal backend

MolDynSub CLI submission using Rainbow JobWrap mode live demo!
Old portal
- Submission via old CLI
- Job status pull from ARC Infosys
- Trajectories Database
- UI https://moldyngrid.org

New portal
- Separate backend from UI
- API for jobs submission
- Jobs send push status updates from every infrastructure
- Gloria integration
- UI is under development
Proof of concept (Under development)
More features and Future plans

• **More Rainbow features:**
  – Libguestfs for more image processing options
  – Notify user by e-mail that VM is ready and how to access it
  – Dynamically change VM’s access password (even for windows)
  – Different Stage-in/Stage-out handling methods
  – Gateway helper optional authentication
  – Compressed images support

• **Future Plans**
  – GPU PCI pass-through support
  – LXC/Docker support
• This work is partially supported by NAS of Ukraine
• Acknowledgements to EGI and Dr. Tiziana Ferrari, Dr. Sara Coelho personally for travel grant to EGI CF 2015
Thank you for your attention.

Questions?