

# Usage of cloud platform for the BY-NCPHEP Tier3 site

V. Mossolov   Dz. Yermak   V. Yermolchyk

Research Institute for Nuclear Problems of Belarusian State University  
Minsk, Belarus

XXV Symposium on Nuclear Electronics and  
Computing  
28 September - 02 October 2015

# Milestones and updates

2008 Site was registered in GocDB as BY-NCPHEP

2009 Site was certified by ROC Russia

2011 Site was registered in CMS SiteDB as  
T3\_BY\_NCPHEP

2013 SE was moved to Supermicro rack mount size  
server

2014 Production instance of PhEDEx was  
commissioned

2015 Grid services were moved to virtualization  
environment

# GRID site usage

- Storage area for analysis output data
- Data postprocessing
- Monte-Carlo sampling
- Other user calculations (Feynman integrals reduction, etc)

# Grid site resources

Small institutions have not enough resources to setup and support large computer cluster:

- limited computational resources (54 cores, raid6 50Tb)
- limited IT personnel (0.5 FTE for site support, 2 admins)



# Why virtualization based IaaS platform?

- Independence from hardware
- Isolation of services
- More efficient resource utilization
- Easy customization and backup
- High availability setup for critically important grid services

# Choice of IaaS platform

Available solutions:

- Proxmox
- OpenNebula
- CloudStack
- OpenStack



# OpenNebula


- Easy provision
- Used as a cloud platform at JINR
- User friendly
- High availability setup is difficult
- Bug? High IO load at frontend node

# OpenNebula

## OpenNebula

 Dashboard

 System

 Virtual Resources

**Virtual Machines**

Templates

Images

Files & Kernels

 Infrastructure

 Marketplace

 OneFlow

## Virtual Machines





| <input type="checkbox"/>            | ID ▾ | Owner    | Group    | Name                  | Status   | Host           | IPs           |
|-------------------------------------|------|----------|----------|-----------------------|----------|----------------|---------------|
| <input type="checkbox"/>            | 94   | testuser | users    | 1                     | UNKNOWN  | node007.hep.by | 195.50.28.238 |
| <input type="checkbox"/>            | 93   | oneadmin | oneadmin | opennebula<br>node 2  | POWEROFF | node007.hep.by | 195.50.28.238 |
| <input type="checkbox"/>            | 92   | oneadmin | oneadmin | opennebula<br>node 1  | POWEROFF | node007.hep.by | 195.50.28.236 |
| <input checked="" type="checkbox"/> | 90   | oneadmin | oneadmin | opennebula -<br>front | POWEROFF | node005.hep.by | 195.50.28.235 |
| <input type="checkbox"/>            | 75   | oneadmin | oneadmin | ceph adm              | RUNNING  | node004.hep.by | 195.50.28.234 |
| <input type="checkbox"/>            | 70   | oneadmin | oneadmin | ceph 3                | POWEROFF | node005.hep.by | 195.50.28.233 |
| <input type="checkbox"/>            | 64   | oneadmin | oneadmin | ceph 1                | POWEROFF | node005.hep.by | 195.50.28.231 |
| <input type="checkbox"/>            | 61   | oneadmin | oneadmin | ceph 2                | POWEROFF | node007.hep.by | 195.50.28.232 |



# Proxmox

- Easy to start and use
- KVM, LXC
- Live Migration
- High Availability Cluster
- Can be used with Lustre shared file system
- Has no single point of failure
- It is not support all features of cloud

# Proxmox

**PROXMOX** Proxmox Virtual Environment  
Version: 3.4-6/102d4547

Server View ▼ **Datacenter**

**Datacenter**

- gridhome1
- gridhome2

**Search** **Summary** **Options** **Storage** **Backup** **Users** **Groups** **Pools** **Permissions** **Roles** **Authentication**

| Type | Description    | Disk usage | Memory usage | CPU usage     | Uptime           |
|------|----------------|------------|--------------|---------------|------------------|
| node | gridhome1      | 26.4%      | 50.4%        | 2.6% of 8CPUs | 47 days 23:00:13 |
| node | gridhome2      | 2.4%       | 7.1%         | 0.1% of 8CPUs | 49 days 21:45:08 |
| qemu | 100 (cfengine) | 0.0%       | 78.4%        | 1.0% of 1CPU  | 47 days 22:58:27 |
| qemu | 101 (grid01)   | 0.0%       | 87.3%        | 0.9% of 1CPU  | 47 days 22:58:23 |
| qemu | 102 (ldap)     | 0.0%       | 81.1%        | 0.4% of 1CPU  | 47 days 22:58:19 |
| qemu | 103 (sbdll)    | 0.0%       | 92.1%        | 4.8% of 1CPU  | 47 days 22:58:15 |
| qemu | 104 (grid02)   | 0.0%       | 41.0%        | 5.2% of 1CPU  | 47 days 22:58:11 |
| qemu | 105 (xp)       | 0.0%       | -            | -             | -                |
| qemu | 106 (ns1)      | 0.0%       | -            | -             | -                |
| qemu | 107 (DC1)      | 0.0%       | -            | -             | -                |
| qemu | 108 (bdc2)     | 0.0%       | 32.0%        | 3.2% of 1CPU  | 02:16:28         |
| qemu | 109 (grid07)   | 0.0%       | 67.7%        | 5.4% of 1CPU  | 47 days 22:58:02 |
| qemu | 110 (node001)  | 0.0%       | 20.6%        | 0.2% of 1CPU  | 40 days 20:49:15 |

# Conclusion

- Grid services moved to virtualization environment.
- High availability setup for critically important grid services is implemented.
- Abilities of Proxmox VE is enough for our current needs. But testing setup of OpenNebula is maintained.
- Replacement of Lustre shared file system is under discussion.