



# The advances in IHEP Cloud facility

Tao Cui, Yaodong Cheng, CC-IHEP, CAS

IHEP, 19B Yuquan Road, Beijing, 100049 China

[Cuit@ihep.ac.cn](mailto:Cuit@ihep.ac.cn)



# Outline

---

- Background
- The advances in IHEPCloud
  - private cloud upgrade
  - advance of Virtual Computing Cluster
- future plan
- Conclusion

Virtual Cluster Computing in IHEPCloud  
Li Haibo at Hepix 2016 Spring

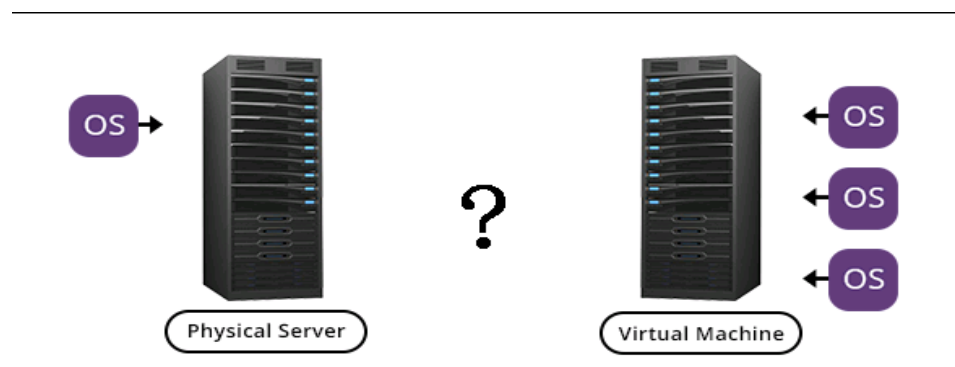
# The group of cloud computing

---

□ founded in May 2014

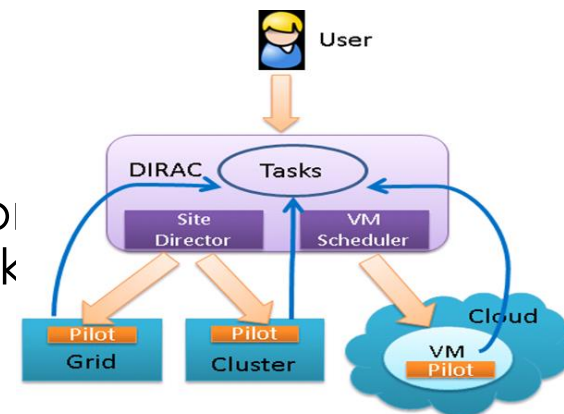
□ The arms

- Computing resource virtualization and flexible virtual resource scheduling
- Hierarchical and heterogeneous virtual system architecture
- Simplified work nodes environment, supported multiple experiments and completely transparent to [users](#)



# IHEPCloud: Three use scenarios

- User self-Service virtual machine platform (IaaS)
  - Launched in Nov 2014
  - User from IHEP unified authentication can launch or destroy VM on-demand
  - Provide UI and testing vm
- A dynamic Virtual Computing Cluster
  - Launched in May 2015
  - Job will be allocated to virtual queue automatically when physical queue is busy
- A distributed computing system
  - Work as a cloud site: Dirac or other application call cloud interface to start or stop virtual work nodes





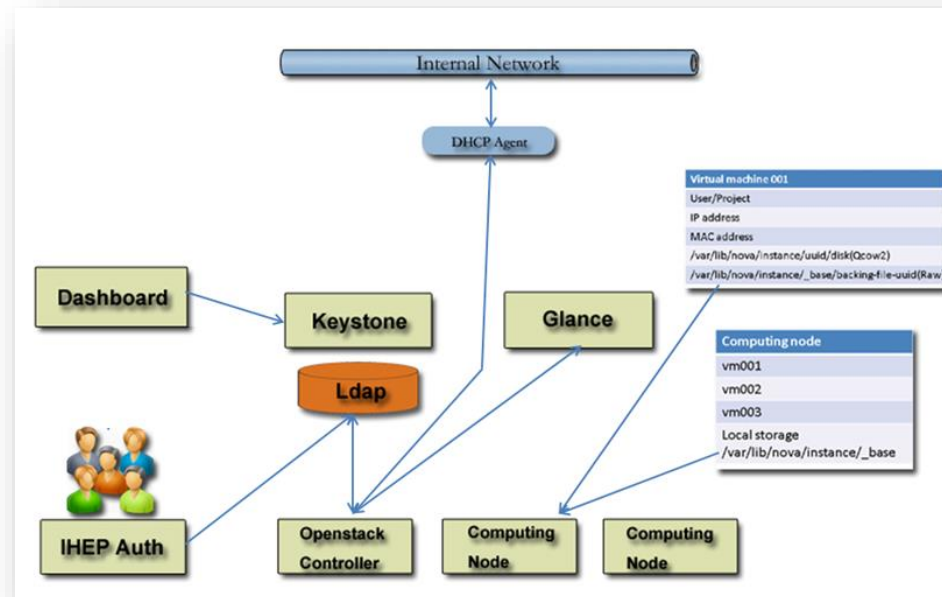
---

# Private Cloud upgrade

# Private Cloud old infrastructure

## Old infrastructure

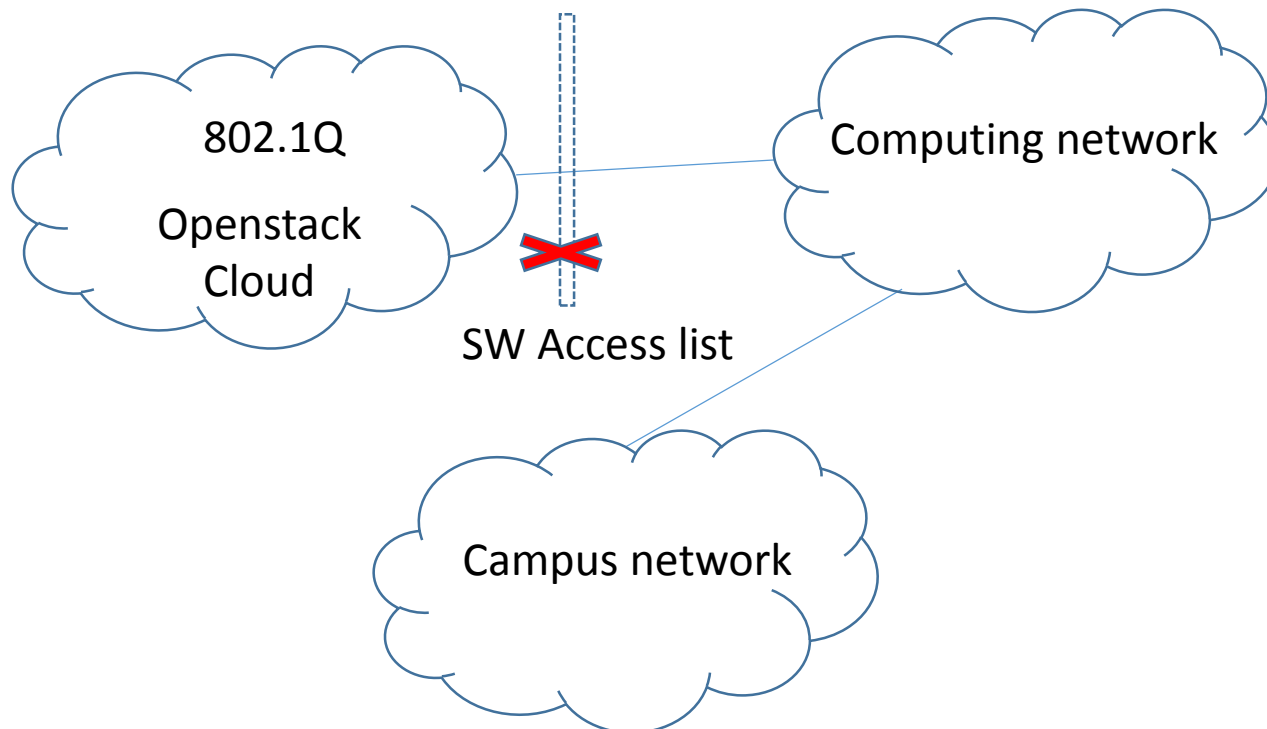
- Version: Openstack Icehouse
- Storage: Local Storage
- Network: Nuetron + OVS + 802.1Q.
- Authentication: user data comes from IHEP unified certification, Users and projects one by one in order to control the quota
- Others: modified dashboard, Limited operating privilege, only vm operation, no volume



# Network of private cloud

---

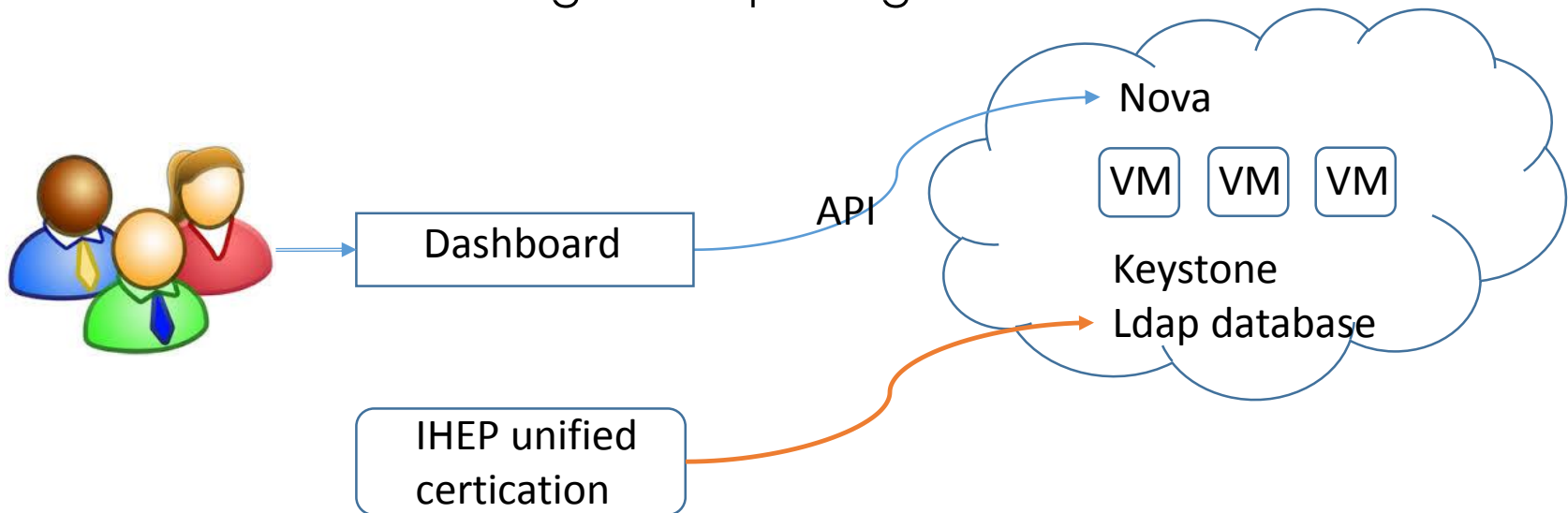
- Independent network connects to computing network with I3 routing and connects to campus network with access list controlled
- To open rules in secgroup and vm ip address is accessible in physical network





# Authentication in private cloud

- ❑ Synchronic Openstack Ldap from IHEP unified certification database
- ❑ The issues
  - ❑ User information depends on IHEP unified certification database
  - ❑ more user info not belong cloud appeared
  - ❑ No local management privilege



# Private Cloud Upgrade

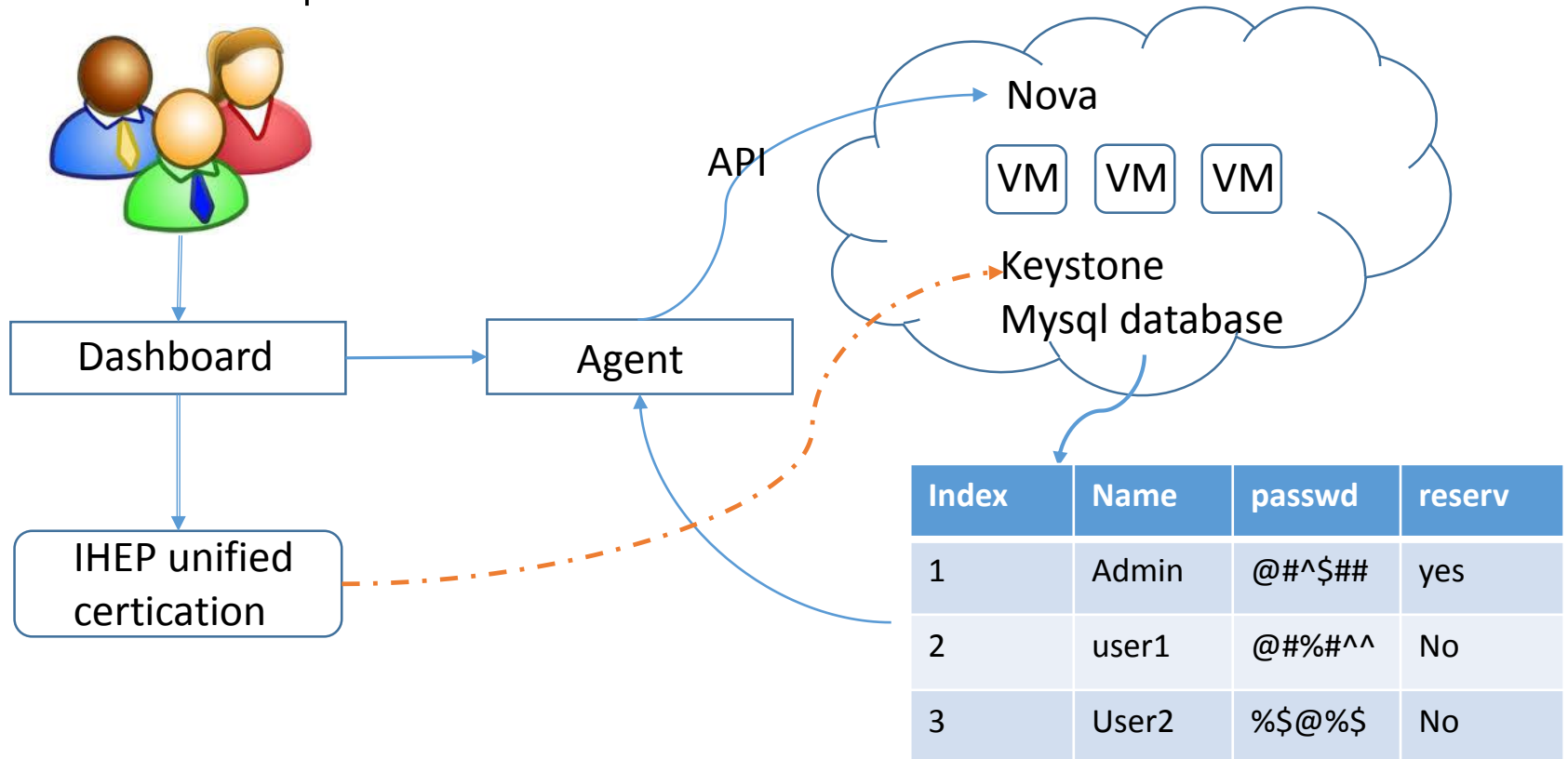
- Upgrade request
  - Maintain virtual machines and users attributes
  - Upgrade from Icehouse to Kilo
  - Improve authentication mechanism
  - Shared storage supported

item	Old Cloud	New Cloud
Version	Openstack Icehouse	Openstack Kilo
Storage	Local Storage for image and instance	Shared storage with GlusterFS
Network	Neutron+OVS+802.1Q	Neutron+OVS+802.1Q
Authentication	Keystone(Ldap)	Agent mechanism Keystone(Mysql) Local Users table and Agent
dashboard	modified and IHEP unified certification supported through Ldap synchronization	Modified and IHEP unified certification through local users table
function	Just virtual machine operation	Just virtual machine operation

Comparing between old cloud and new cloud

# Improvement of authentication in private cloud

- Agent running on the controller
- Agent responsible for replacing password and comparing user name from Mysql database
- white list to protect local user such as admin and services ...



# Openstack pgrade approach

---

## □ The approach


- Deployed a new Openstack Kilo based on SL7
- Configure New environment such as image, flavor
- Same network information between old and new one
- Migrate projects/users and their attributes and dashboard modified for this changes
  - IHEP unified certification
  - Local database and local projects/users
- Migrate virtual machines based on users
- Migrate only increased images but change backing file location

Migration

# The result

---

- ❑ OS from SL6.5 to SL7
- ❑ Openstack from Icehouse to Kilo
- ❑ To migrate 135 vms and ~100 users
- ❑ Shared storage with GlusterFS
- ❑ 45 hours cost



Powered by OpenStack

## Log In

**User Name**

**Password**

[Help](#) [Sign In](#) [IHEP-SSO Sign In](#)

---

# Advances of Virtual Computing Cluster ----VPMManager component

# IHEP Cloud –Virtual Computing Cluster

---

## □ Virtual Computing Cluster

- Launched in May 2015 based on Openstack Kilo
- Support BESIII, JUNO, LHAASO, CEPC,...
- The arms
  - Providing virtual machine as work nodes to computing system
  - Supporting multiple experiments and dynamic virtual resources scheduling.
  - Completely transparent to the computing service and end users

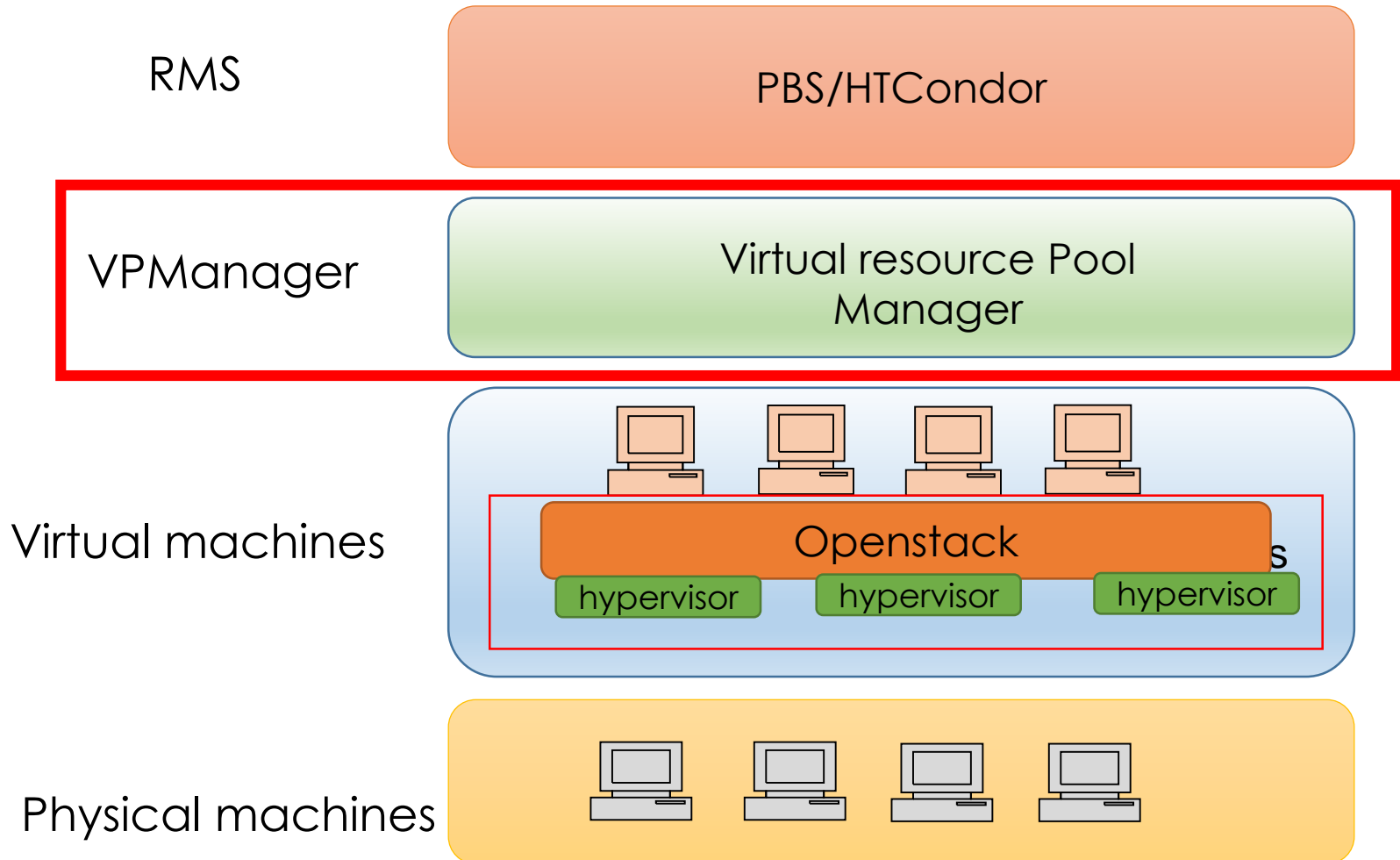
## □ Status

- ~700 CPU cores
- 4 experiments



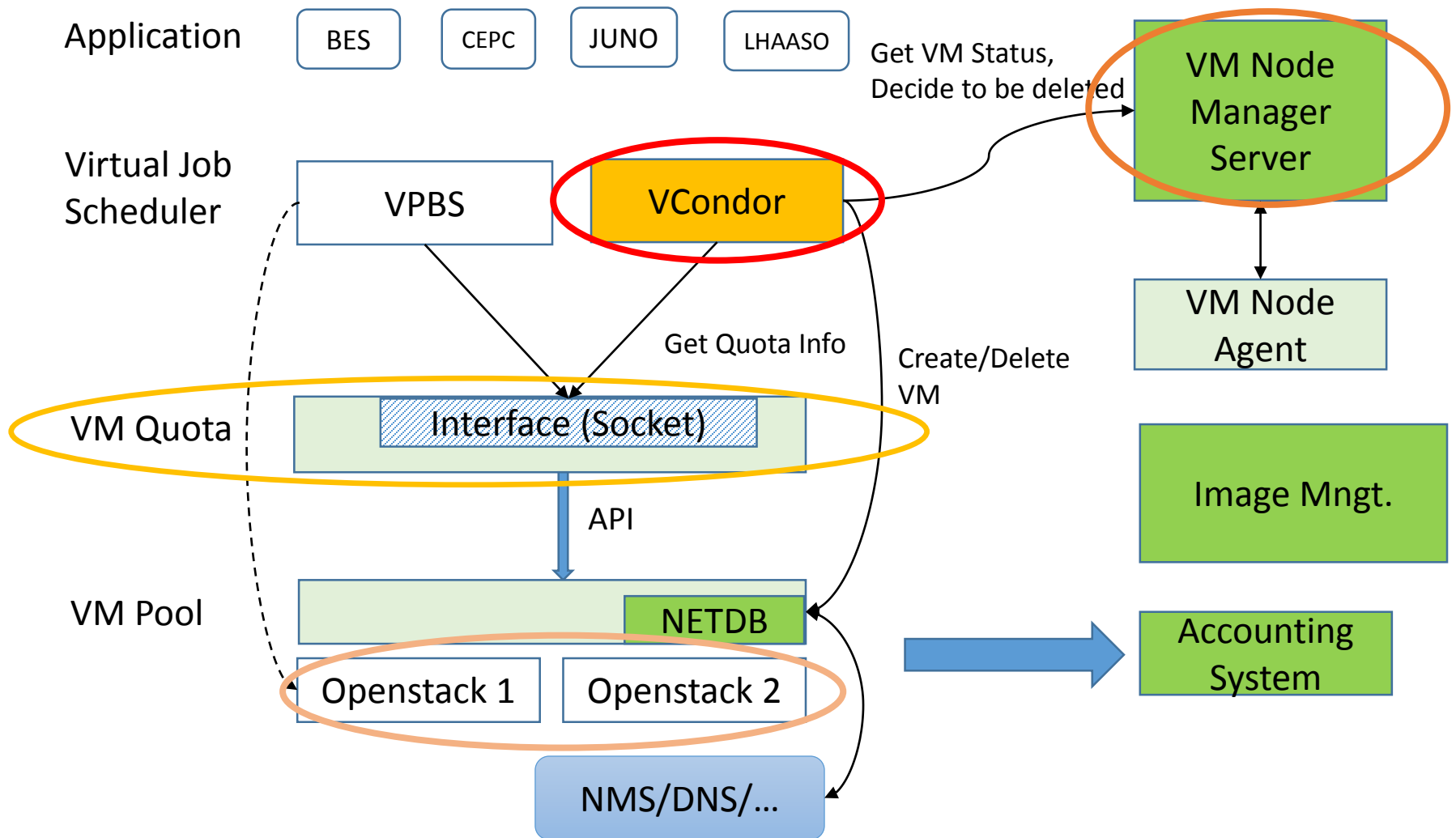
# Infrastructure of Virtual Computing Cluster

- VPManger virtual resource pool manager





# VPManager(Virtual resource Pool Manager)



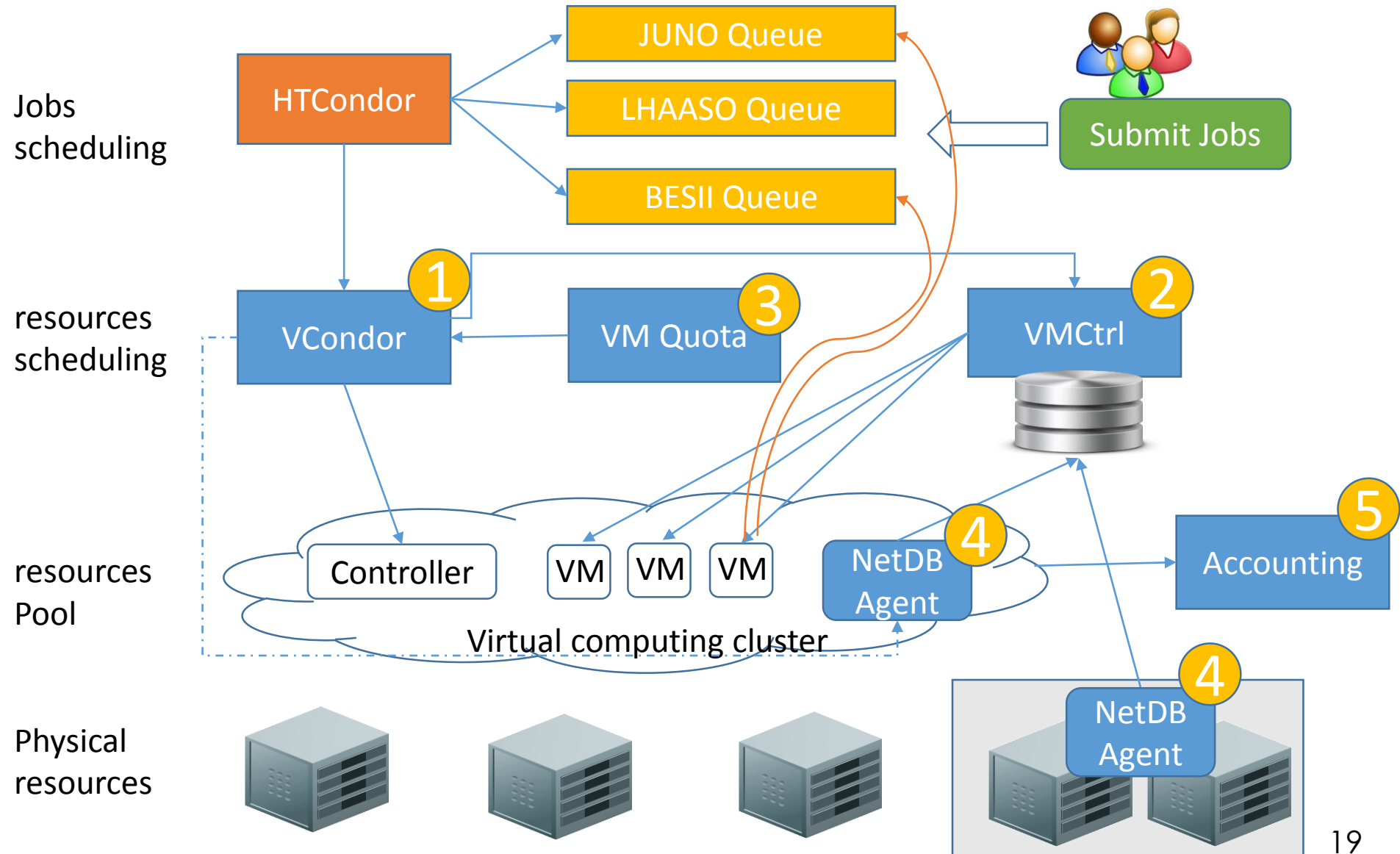
# Scheduler and architecture

## □ Scheduler

- PBS deprecated
- Migrating from PBS to HTCondor
- Virtual Computing Cluster resources scheduler based on HTCondor



# VPManager(Virtual resource Pool Manager)



# VPMManager components

---

## □ VCondor

- A resources scheduler which Increasing or decreasing virtual resources to a queue

## □ VMQuota

- Responsible for resources quota control, vCondor allocates or recycles resources based on maximum or minimum resource limitation.

## □ VMCtrl

- To change virtual machine's queue attribution according to vCondor's request

## □ NetDB Agent

- A Collector which collects info of virtual resources
- A interface which provides resources status to vCondor

## □ Accounting system

- keeps all the usage information of each virtual machine and generate bills to user

# VCondor

## □ VCondor

### □ Resource allocation as demand

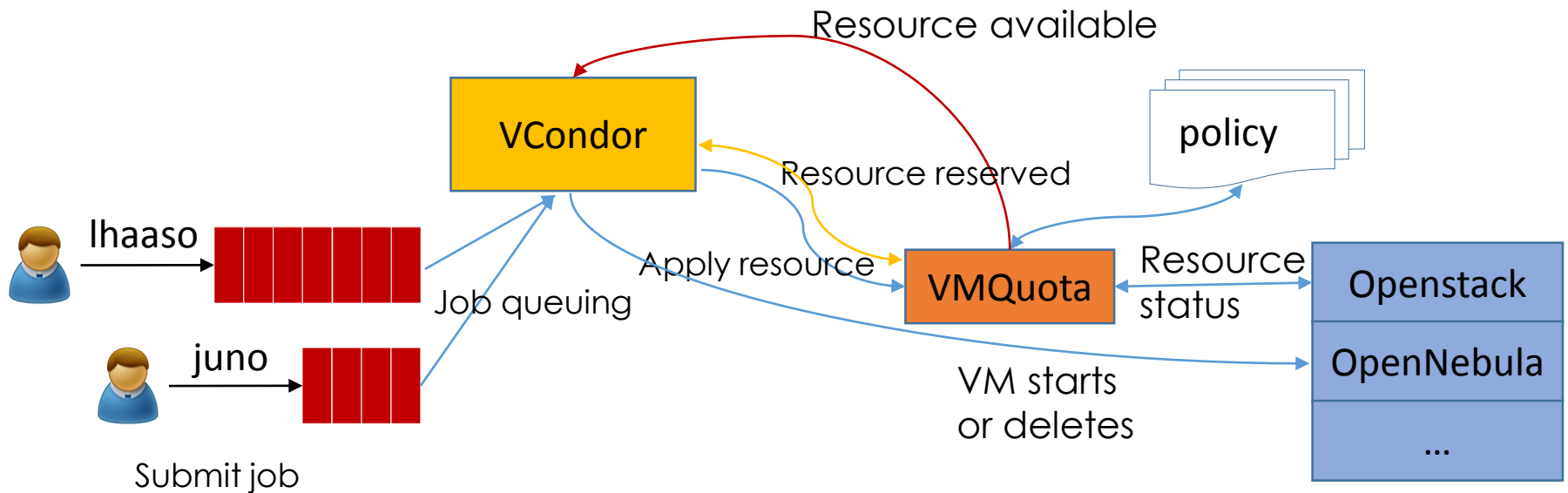
- Queue length
- Resources status
- Resources allocation policy

HTCondor

Openstack Agent

VMQuota

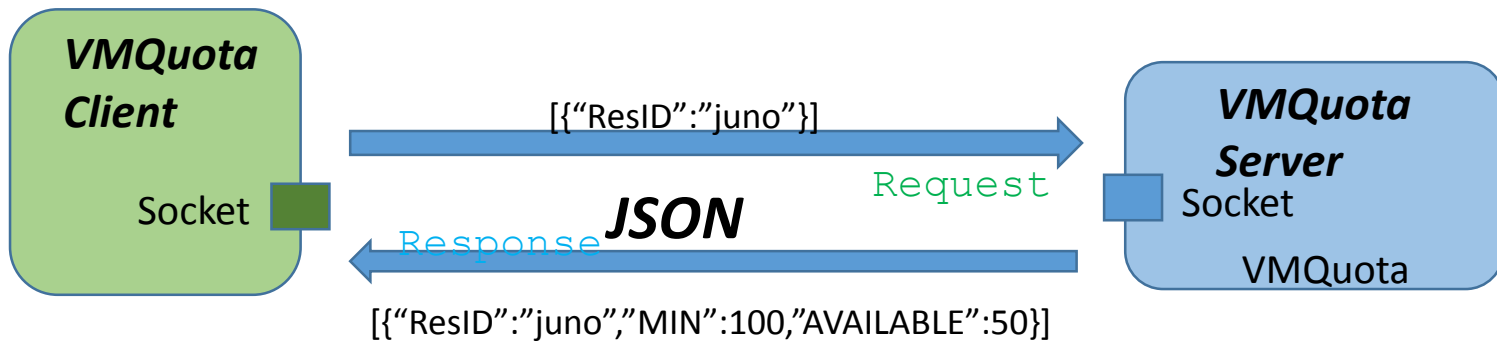
### □ Based on HTCondor



# VMQuota

- VMQuota consists of a web interface and a database

Queue Name	Min	Max	Available	Reserve time(s)
BES	100	400	200	600
JUNO	100	300	200	600



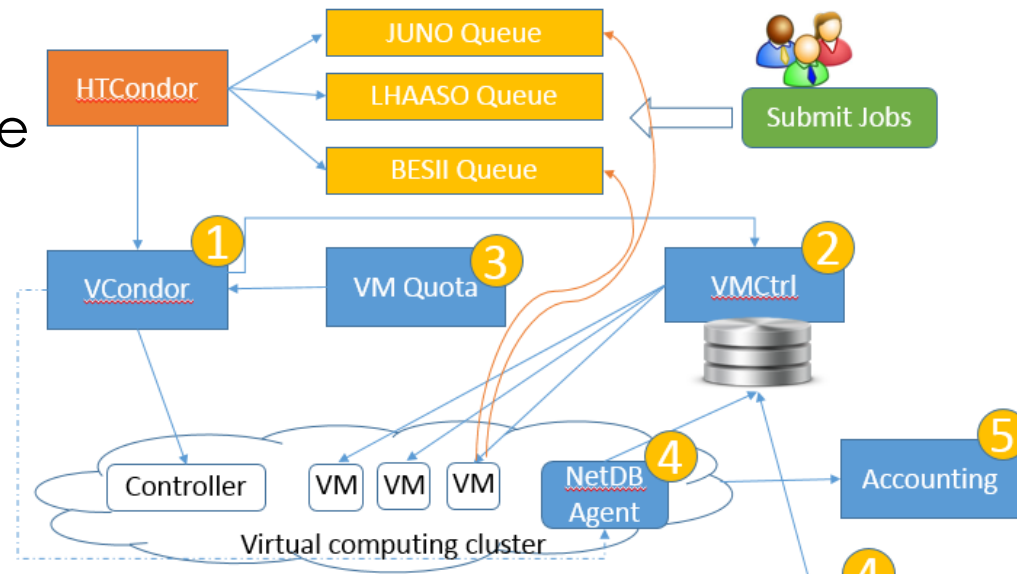
# VMCtrl

## VMCtrl function

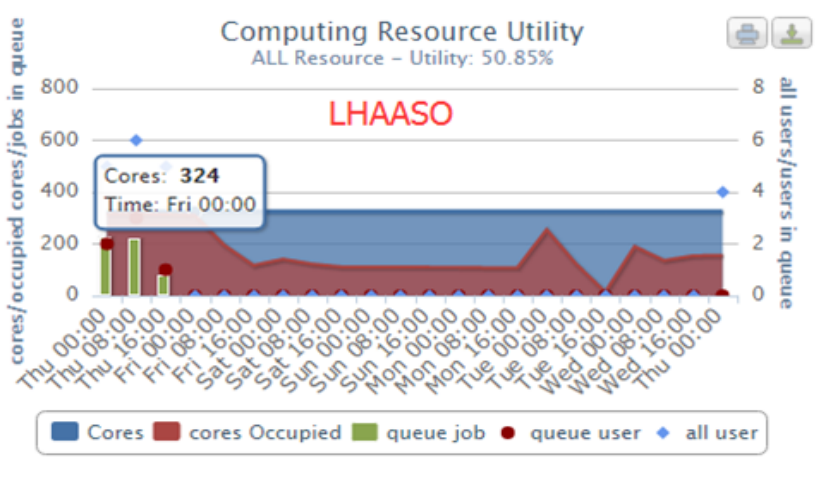
- To increase or decrease HTCondor queue's computing resources
  - To accept virtual machines info from NetDB Agent
  - To accept queue modification request from VCondor
  - To accept query from virtual machines and reconfigure condor client queue name

## component

- A web service interface
- A mysql database
  - Vmctrl table
  - Vms table
- A linux script running on virtual machine

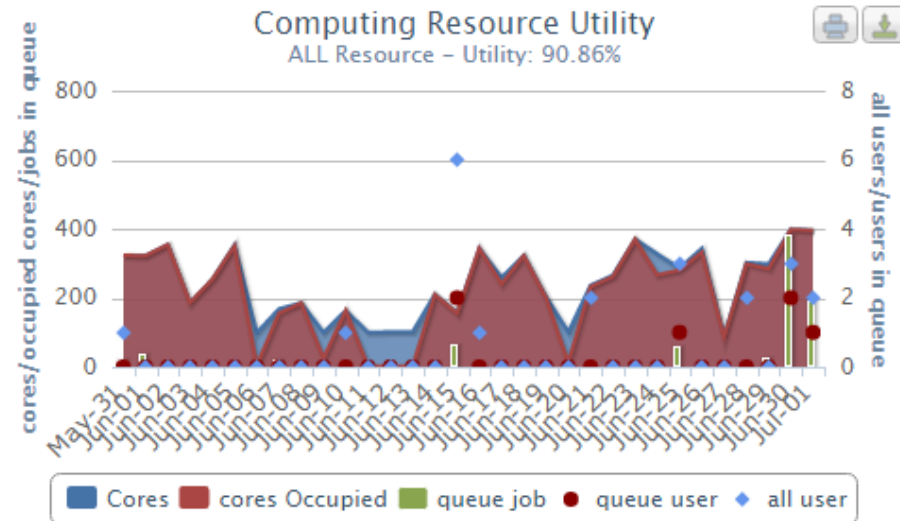
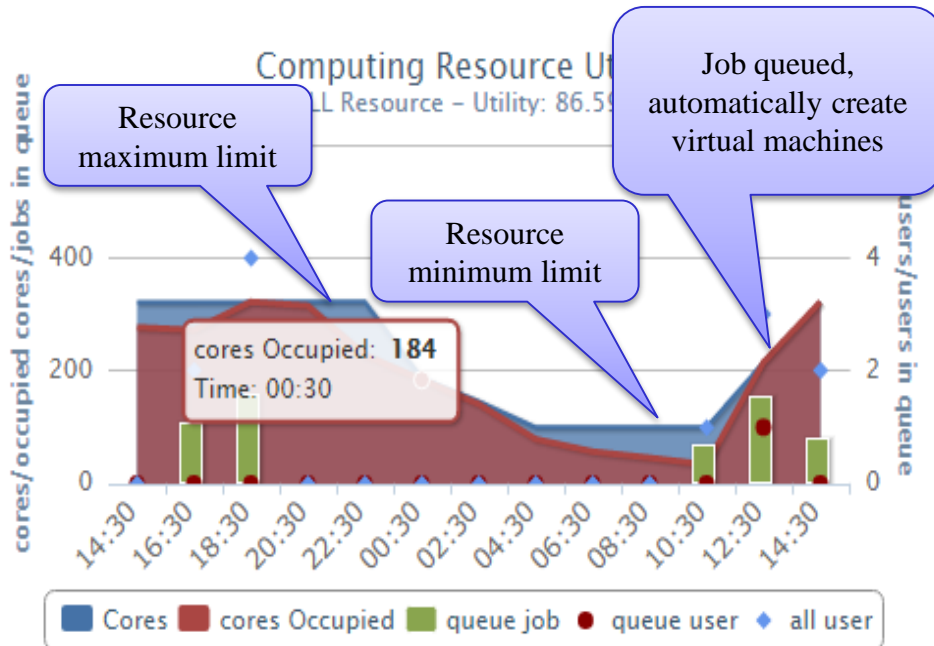


# VCondor scheduling



Computing Resource Utility

~50%





# The advance of virtual computing cloud

---

In the last six months

- Resources dynamic scheduling
  - Experiment resource adjustment
  - Virtual machines amount
- Hierarchical and heterogeneous system architecture
- a stable system maintain experience

# Virtual machine performance testing

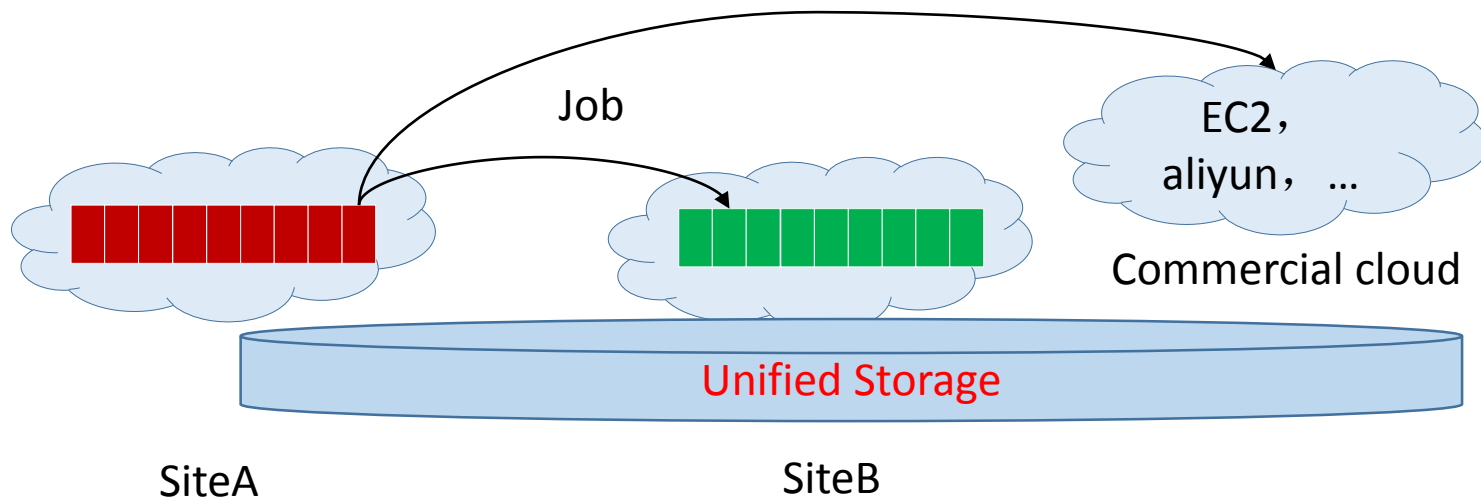
---

- BES simulation job ~2% to ~3% time penalties
- BES reconstruction job ~3% to 6 % time penalties
- CEPC job
  - MPI Parallel Computing job
    - Very small amount of data and network communication
  - Running well but the performance is 50% of the physical machine

# Future plans

---

- ❑ find a solution to scale the virtual computing cloud system
- ❑ China HEP Community Cloud Plan
  - ❑ Sharing resource across different sites
    - ❑ improve resource utilization
  - ❑ Different data management solution
    - ❑ Same storage / file system view across different sites
    - ❑ Data access transparently in every site
    - ❑ Streaming and cache data mechanism
    - ❑ EOS or LEAF (an ongoing system)



# Conclusion

---

- IHEPCloud aims at providing self-service virtual machine platform and virtual computing environment
- VPManger components provides a dynamic and flexible virtual resources scheduling architecture and the system work well under the architecture
- More resources will be added to IHEPCloud this year

# Thank you!

Any Questions?