



# OpenStack Chances and Practice at IHEP

**Haibo, Li**

**lihaibo@ihep.ac.cn**

**Computing Center, the Institute of High  
Energy Physics, CAS, China**

**2012/10/15**

# Agenda

- OpenStack Overview
- Why choose OpenStack?
- What to do with OpenStack?

# What is OpenStack?

“A collection of open source software for building private and public clouds”

- Founded in July 2010 by NASA and Rackspace.
- Has released six versions.(keeps every 6 months update, “Folsom ” in 2012.10)

# Where to get start?

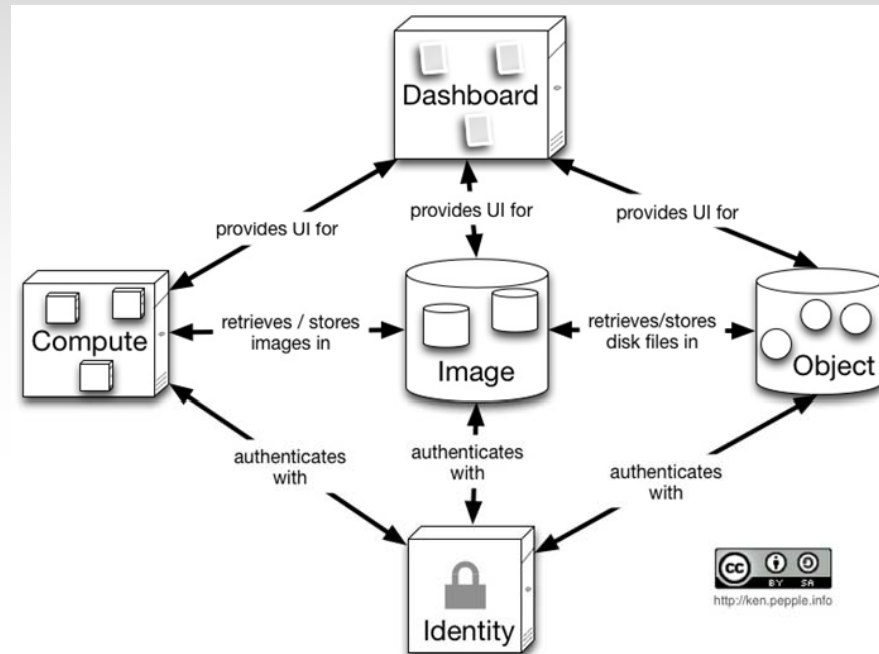


Official documents



Trystack.org Request a test account

# OpenStack Architecture



A stylized and simplified view of the architecture  
(Essex version)

- Each of the constituent services are designed to work together to provide a complete Infrastructure as a Service (IaaS).
- This integration is facilitated through public application programming interfaces (APIs) that each service offers (and in turn can consume).
- While these APIs allow each of the services to use another service, it also allows an implementer to switch out any service as long as they maintain the API.

# Openstack Projects

OpenStack has 5 projects in Essex version.

## Compute – “Nova”

A cloud fabric controller, used to start up virtual instances for either a user or a group. It's also used to configure networking for each instance or project that contains multiple instances for a particular project.

## Storage – “Swift”

A system to store objects in a massively scalable large capacity system with built-in redundancy and failover. Object Storage has a variety of applications, such as backing up or archiving data, serving graphic or videos (streaming data to a user's browser), storing secondary or tertiary static data, developing new applications with data storage integration, storing data when predicting storage capacity is difficult, and creating the elasticity and flexibility of cloud-based storage for your web applications.

# Openstack Projects (cont.)

## Image Service – “Glance”

A lookup and retrieval system for virtual machine images. It can be configured in three ways: using OpenStack Object Store to store images; using Amazon’s Simple Storage Solution (S3) storage directly; or using S3 storage with Object Store as the intermediate for S3 access.

## Identity – “Keystone”

Provides authentication and authorization for all the OpenStack services. It also provides a service catalog of services within a particular deployment.

## Dashboard – “Horizon”

Provides a modular web-based user interface for all the OpenStack service.

# Agenda

- OpenStack Overview
- Why choose OpenStack?
- What to do with OpenStack?



# Why choose OpenStack?

## **Open Source**

Apache 2.0 license, NO 'enterprise' version

## **Open Design**

Open Design Summit, anyone is able to define core architecture

## **Open Development**

Anyone can involve development process via Launchpad & Github

## **Open Community**

OpenStack Foundation in 2012, Now 190+ companies, 3000+ developers

## **100% python**

Python is an interpreted, interactive, object-oriented, extensible programming language

# Agenda

- OpenStack Overview
- Why choose OpenStack?
- **What to do with OpenStack?**

# Infrastructure & Platform

## Virtualization Platform (IaaS)

- VM manager system
- Private cloud with interfaces to customize virtual machines
- Realize auto configuration

## Integrated with our batch system Torque/PBS

- add/remove virtual nodes in a specified job queue dynamically

## Manage resources in remote sites.

- Monitoring and Dashboard development.

# Current Status

Now, we have two openstack environments(based on Essex version):

## 1)Production environment

- Use Dell R510 + ubuntu Juju
- Juju is a tool to deploy openstack, similar to puppet and chef.

Now, almost 100 virtual machines are running.

## 2)Development environment

- Use Dell R510 +scientific linux6.2



<http://juju.ubuntu.com/>

# IaaS

## How to use the IaaS?

- 1) Make a Request (to the administrator an email )
- 2) Administrator creates an account (in dashboard )
- 3) Login and enjoy it!

Now it is merely manual, maybe consider automation later if needed.

## To use it:

- 1) Create image. There are two types of images: Public image and Customized image. Users can choose one according to their needs.
- 2) Launch an instance.

Instances & Volumes

Instances Launch Instance Terminate Instances

<input type="checkbox"/>	Instance Name	IP Address	Size	Status	Task	Power State	Actions
<input type="checkbox"/>	sunzh	192.168.252.12	512MB RAM   1 VCPU   0 Disk	Active	None	Running	Edit Instance
<input type="checkbox"/>	sunzh	192.168.252.7 202.122.39.199	4GB RAM   1 VCPU   40.0GB Disk	Active	None	Running	Edit Instance

Displaying 2 items

- 3) Log on by ssh or vnc.

# Integrated with Torque/PBS

## Problems:

- IHEP computing environment adopts Torque/PBS. Torque process a job according to the task queue, executed sequentially
- Two types of nodes:
  - Exclusive node
  - Shared node
    - can share but cannot be preempted.
- The Checkpoint function in PBS not well supported network file system.

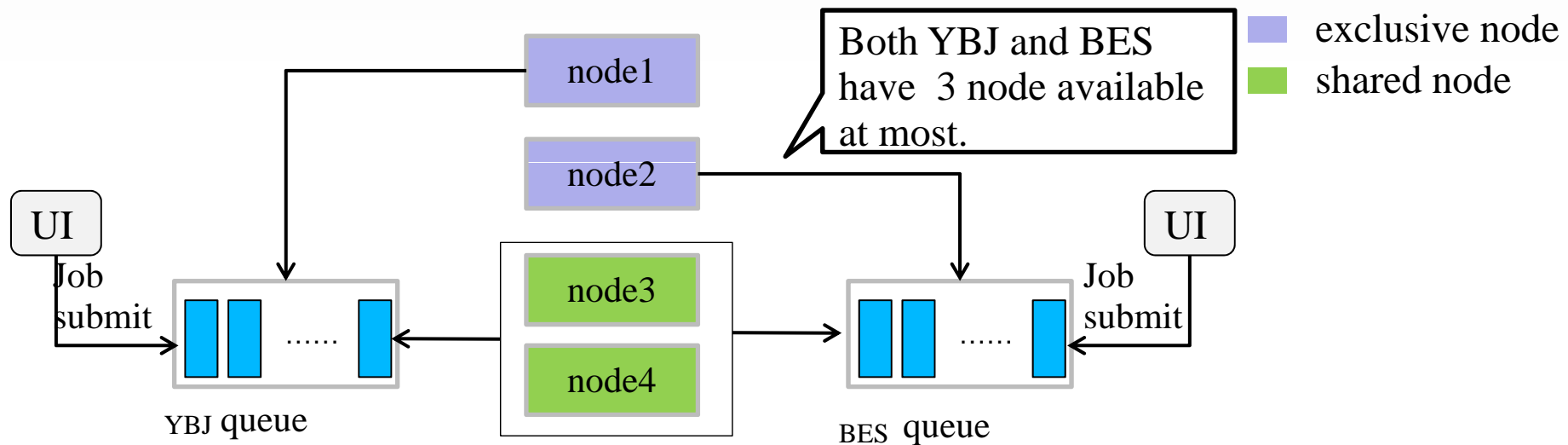
## Result:

- stunt to preemptive scheduling & key resources protection insurance

# Example

Let's use an example to explain it.

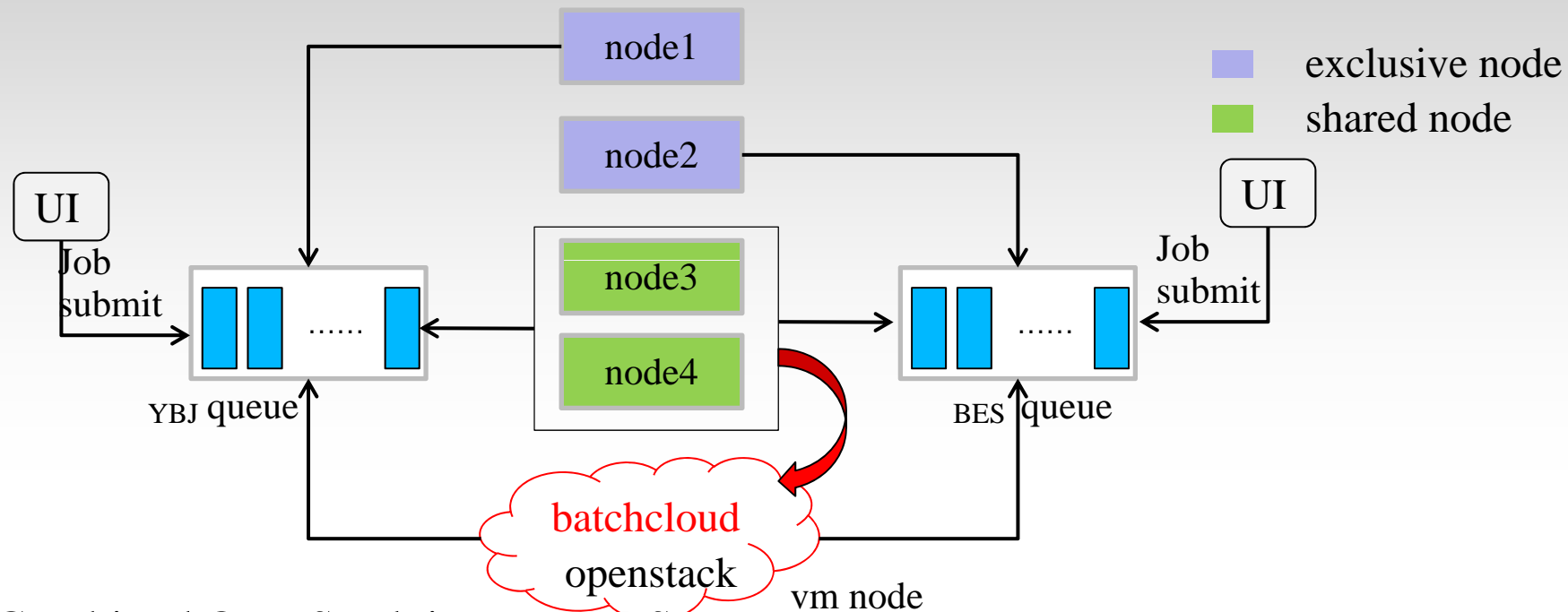
- Assume there are 4 nodes, and two queues, YBJ and BES. YBJ queue owns 1 exclusive node, BES owns 1 exclusive node, and the remaining 2 is sharing. Each node has 8 cores, which has the ability to run 8 **jobs** simultaneously at most.



- If YBJ queue is full, for example needs 10 **nodes** when BES queue is idle, exclusive node in BES queue (eg, node2) can't be used by YBJ queue.

# Solutions

Introducing OpenStack into Torque:



- Combined OpenStack into our PBS.
- Using openstack to manage virtual machines to provide vm nodes. These nodes can be created on-demand in **batchcloud**, and deallocate the virtual machines to take resource from the low priority jobs in accordance with the scheduling policy. When an urgent job arrives, the batchcloud will do as follows:
  - 1) Start an instance.
  - 2) Run Qmgr to add the job to the queue.
  - 3) Wait for the job completion and release the vm.

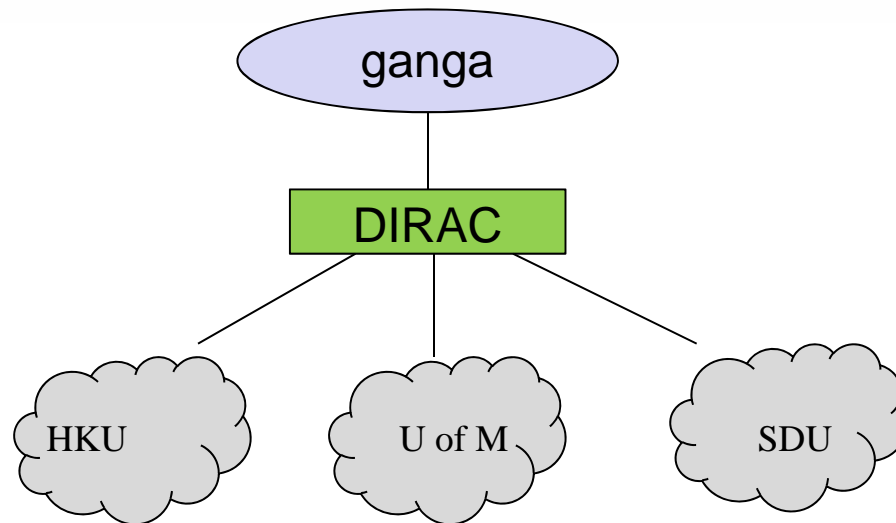
What's more, the shared nodes can be gradually substituted by the vm nodes.



# Manage resources in remote sites

## Problems:

- There are a lot of sites integrated into IHEP BES computing system, especially for some simulation and analysis task.
- Each remote site should run some associated software (Grid software, BES software, Local storage and computing software).
- These software need periodic update.

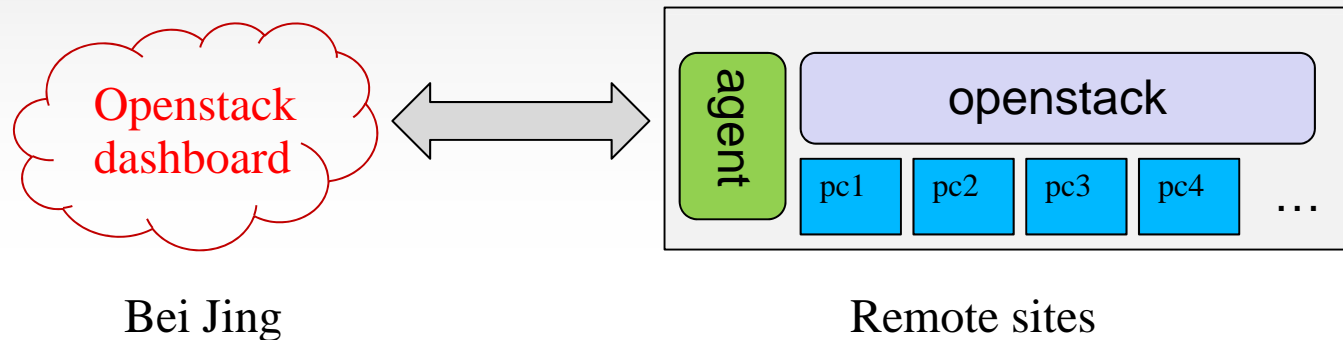


## Result:

- Every remote site needs an dedicated administrator to operate locally, however, some sites is small and it is impractical.

# Solutions

We are going to use OpenStack to improve the usability.



We will install openstack in the remote sites, then create an agent in the remote sites, so the administrator can access and manage using dashboard in Beijing .

## Remote sites:

- Maintenance system (physical machines && OS)

## Beijing local site:

- Image install, software update
- Vms startup and shutdown
- Monitoring all the resource (virtual machine, physical machine, services, CPU, network, storage, etc)

# Conclusion

- Openstack develops very fast, it provides a good chance for IHEP.
- We use openstack to construct our IaaS.
- Integrate openstack with our batch system Torque/PBS to improve resource utilization.
- Use openstack to manage resources in remote sites to improve the usability.

Thanks!  
Q & A