

Assignment Report for OSG Summer School 2011

Seo-Young Noh
Supercomputing Center
Korea Institute of Science and Technology Information
Daejeon, Korea, 305-806

July 28, 2011

Abstract

This is the report for the assignment of OSG Summer School 2011 which was held at the University of Wisconsin, Madison from June 26 through June 30. In this report, I give my impression on the program of the school and briefly share my conceptual idea on how to apply what I learned at OSG summer school to my current research and work.

1 Overall Comments on OSG Summer School 2011

It was my great opportunity having a chance to attend the OSG Summer School 2011. Before attending the school, I had hard time to understand the concept of OSG and its software stack. Due to the well organized informative program, I am much more comfortable for OSG. Instructors were very knowledgeable and provided me clean and clear answers to my questions, which had been always confusing me. I hope this summer school to be held continuously because its productive and informative educational programs help not only researchers in basic sciences to solve their problems, but also system administrators who have to manage OSG sites in near future.

2 Instructors

Instructors were very helpful and qualified. They were very knowledgeable and exactly pointed out my mistakes with clean explanations. Presentations were very well organized and neat. All topics covered during the school were connected to each other and they were covered from simple to complex, and

from general to advanced items. Time management was also good. They partitioned their lectures into “theory” and “experiment” with very well balanced format. Instructors were very helpful at the evening session too. Sometimes I was in short time to complete some exercises, but I was able to manage them because they opened evening sessions every day. I really appreciate their efforts on lectures and extra services in the evening sessions.

3 Lecture Materials

Exercises used in lectures were very helpful to understand concepts. They were covering not only basic concepts, but also advanced examples which can be directly applied to real world problems. I am sure that valuable examples covered in the lectures will definitely save a lot of time for students to solve their research problems. In my case, DAG examples were very useful because I had trouble to understand its purpose and concept before attending the school.

4 School Organization

I also want to express my sincere thanks to all staffs for their efforts who organized the OSG Summer School 2011. They were very responsive to requests and questions. Lodging was very nice and convenient. Meals served were wonderful. Breaking ice is one of key points to make a good program and good connections among students. We were lucky because we had very friendly instructors - Alain, Igor, Brian, especially Tim! I had really good time with the other students.

5 Applying to My Research and Work

I was thinking about how to apply what I learned from OSG Summer School to my work. Because my main role at KISTI is managing OSG site as a site administrator, it would be good if I can find a flexible way to manage cluster systems, which are robust and easily scalable. After attending the school, I came up with one idea to utilize virtualization technology and Condor. A framework for a virtual cluster which combines independent cloud systems and utilizes their powers will provide great flexibility to handle grid jobs. I named it `vcluster` which is automatic scalable virtual cluster system.

5.1 Concept of vcluster

The basic concept of **vcluster** is utilizing existing private and public cloud systems to make automatically scalable cluster system. For a private cloud system, it can directly control worker nodes running as virtual machines; therefore it can evenly distribute virtual worker nodes in the cloud system, and migrate virtual worker nodes to certain host machines in a certain condition in order to save power consumption. When there are no more rooms available for new virtual worker nodes, **vcluster** can utilize public cloud systems to create virtual worker nodes. Such worker nodes automatically join the cluster pool managed by **vcluster**.

Figure 1 shows a conceptual diagram of **vcluster**. **vcluster** is agnostic on cloud system. Each cloud system is communicated with **vcluster** using a cloud plug-in. It is also agnostic on batch systems. Whenever there is a proper plug-in for a batch system, it can work with **vcluster**.

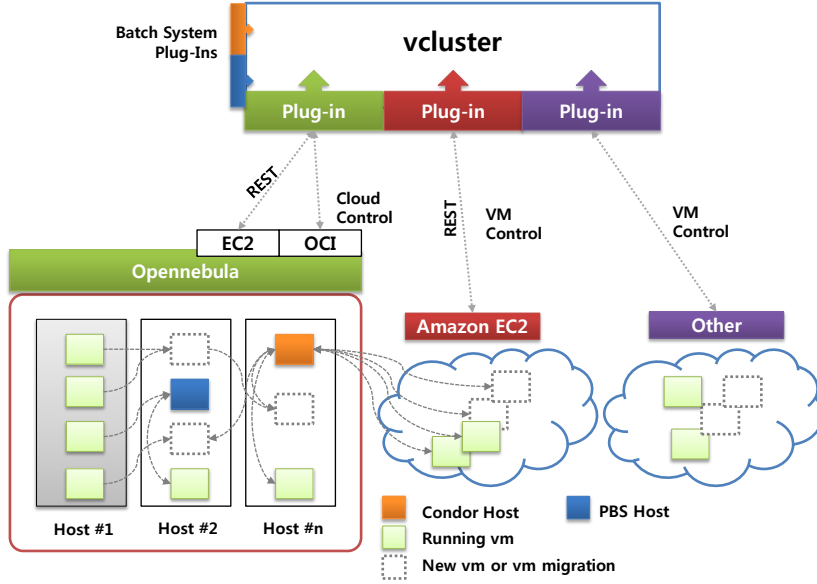


Figure 1: Conceptual Diagram of vcluster

5.2 Components of vcluster

`vcluster` has three main component blocks-**Controller**, **Load Balancer** and **Monitoring** as shown in Figure 2. The underlying design philosophy of `vcluster` is plug-in based and agnostic on underlying batch system and cloud system in order to flexibly increase and decrease its capabilities of handling grid jobs on demands.

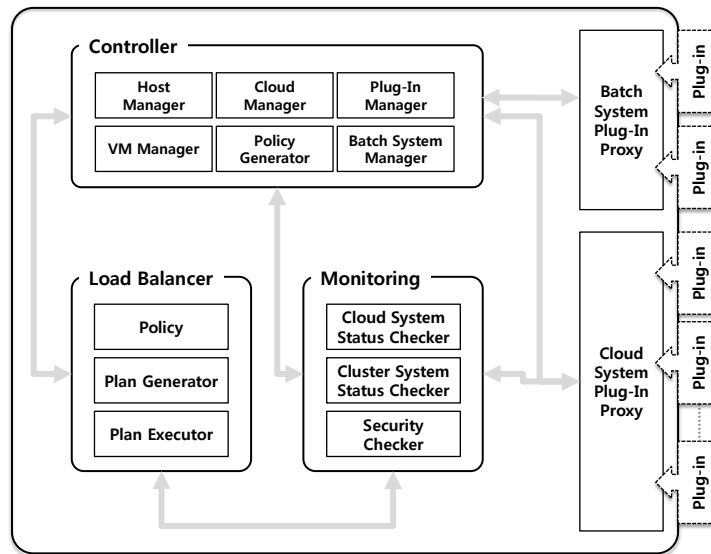


Figure 2: Components of vcluster

5.3 Status of Current Implementation

Current stage of `vcluster` is in prototyping. Below shows some outputs of current implementation. It can communicate with *FermiCloud* system using EC2 REST APIs where FermiCloud is a private cloud system and being managed at Fermi National Accelerator Laboratory.

- Describe Instances:

```
vcluter > dins
```

Inst ID	Status
i-618	running
i-693	running
i-718	running
i-732	running
i-735	running
i-765	running
i-796	running
i-787	running
i-797	running
i-801	running

```
vcluter >
```

- Terminate Instance: terminate instance i-801

```
vcluter > ti i-801
```

```
vcluter > dins
```

Inst ID	Status
i-618	running
i-693	running
i-718	running
i-732	running
i-735	running
i-765	running
i-796	running
i-787	running
i-797	running
i-801	shutting-down

```
vcluter >
```

- Checking instances after shutting-down have been done

```
vcluter > dins
```

Inst ID	Status
i-618	running
i-693	running
i-718	running

```

i-732      running
i-735      running
i-765      running
i-796      running
i-787      running
i-797      running
-----
vcluter >

```

- Describe Images:

```

vcluter > dimg
-----
Img Id      Owner      isPublic?
-----
ami-00000060  rsyoun  true
ami-00000143  rsyoun  false
ami-00000146  rsyoun  false
-----
vcluter >

```

- Run Instance: launch a vm based on ami-00000146

```

vcluter > run ami-00000146
vcluter >
vcluter > dins
-----
Inst ID      Status
-----
i-618        running
i-693        running
i-718        running
i-732        running
i-735        running
i-765        running
i-796        running
i-787        running
i-797        running
i-803        pending
-----
vcluter >
vcluter > dins
-----
Inst ID      Status
-----

```

i-618	running
i-693	running
i-718	running
i-732	running
i-735	running
i-765	running
i-796	running
i-787	running
i-797	running
i-803	running

vcluter >

5.4 Future Plan

The current status of **vcluster** is capable of handling virtual machines using REST API. In a near future, there are two major implementations 1) handling virtual machines over physical machines and 2) launching and terminating virtual machines depending on the number of waiting jobs in a batch system.