edubase Cloud: Cloud Platform for Cloud Education

Shigetoshi Yokoyama and Nobukazu Yoshioka
GRACE Center, National Institute of Informatics
Tokyo, Japan
{*yoko, nobukazu*}*@nii.ac.jp*

Takahiro Shida
NTT DATA Intellilink
Tokyo, Japan
shida@intellilink.co.jp

***Abstract*—education of cloud engineers will be crucial for the continued development of cloud technologies. We have developed an open-source software platform called edubase Cloud for education. The platform has multi-cloud architecture. In this paper, we discuss how edubase Cloud provides an alterable cloud platform and how effective it is for educating cloud engineers.**

***Keywords*-Cloud computing; education; alterable platform; customizable platform; Infrastructure as a Service**

I. Introduction

Cloud computing has the potential to dramatically change software engineering. It allows us to manage and use large-scale computing resources efficiently and easily and to develop new software by using these resources for scalability and lowering costs.

Recently, commercial clouds have been used for training cloud users. Such courses are, however, inadequate for cloud engineers because they do not make it possible to learn how to monitor or change the cloud infrastructure's software. Training specifically geared for cloud infrastructure engineers is urgently needed because they are the individuals who design and build reliable and efficient data centers. Educating them will be the key to innovation in cloud technologies.

In response to this need, we have developed edubase Cloud [1], a new cloud platform based on open-source software and using a multi-cloud architecture. An open source platform is needed for training engineers about cloud infrastructure because they need to be able to monitor the cloud's behavior and change it in order to fully understand the mechanism.

We conducted a case study on developing educational content for cloud engineers using edubase Cloud. In this paper, we discuss the case study's findings on the alterability of the cloud platform in detail, which is one of the key features of edubase Cloud.

This paper is organized as follows. Section 2 describes edubase Cloud. Section 3 presents the case study and examines the alterability issue. We summarize our findings from the case study in Section 4 and conclude in Section 5.

II. edubase Cloud

edubase Cloud is an open-source educational cloud that can create an environment in which users can acquire the necessary IT resources as virtual machine, network and storage when needed, and can experiment with their ideas without restraint. Edubase Cloud is expected to be useful for a broad range of training circumstances in the IT field, from basic training on cloud technology to Project Based Learning.

We envision that edubase Cloud will be used in IT education in a number of ways as follows:

Lectures: Users can create several educational settings simultaneously, by copying and allocating an IT environment which has already been built and verified from the archive on demand. Users who want to prepare individual environments for each lecture can quickly establish an environment which is suitable for the lesson content.

Training: By taking advantage of edubase Cloud's exclusiveness and alteration properties, users can undertake practical training, including alteration of the platform, without worrying about impacting other users. It is also easy to set up their IT experiment in their own cloud live and save the midstream results.

Project-based Learning (PBL): Sharing an IT environment enables decentralized cooperative learning among students. Users can save or reuse the results in the archive after completing the project or access external cloud services, if necessary.

It will be a place to learn essential skills to manage and develop cloud platforms. It has the most advanced IT environment and will be the driving force behind the development of new IT specialists.

A cloud platform for educating cloud engineers' must meet the following the key features as described in figure 1:

- Exclusiveness (F1): Users can acquire a physically isolated IT experimental environment for themselves.
- Interoperability (F2): Users can easily work with external cloud services by utilizing its open interface. It must be interoperable with private and public clouds.
- Alteration (F3): The full open-source cloud environment enables customization and tuning of the platform itself. It must be customizable without influencing other cloud lectures or practices.

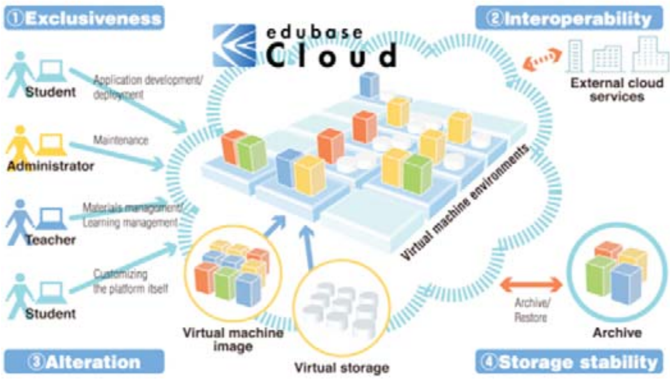


Figure 1. Four key features of edubase Cloud

- Storage Stability (F4): Edubase Cloud has an archive function to save virtual machine images in a reusable form. It must be able to archive its educational content, including machine images, at a location outside of the cloud for the purpose of sharing content with other clouds.

The edubase Cloud consists of 200 nodes, and communication, monitoring, and storage servers as shown in figure 2. The cloud has a multi-cloud architecture to meet (F1) and (F3). Each mini-cloud is a cloud platform that is built up using an enhanced version of Eucalyptus 1.6.2, an open-source cloud platform [2], to meet (F2) and (F3). It consists of a cloud controller, a cluster controller, and 14 node controllers with 8 cores. We can launch more than 100 virtual machines in a mini-cloud. With this architecture, users cannot be accidentally influenced by other users in a different mini-cloud. Communication servers support users' activities in mini-clouds. The project management server allows users to store the results of their projects, including machine images, and resume them in any other cloud; this meets (F4). The learning management server is based on Moodle and provides a portal site for users, and the user management server uses Open-LDAP to unite these servers.

Since edubase Cloud only uses open-source software, we can look at its code and learn. Moreover, it can be modified for experimental learning purposes. It also has a Web API and a machine image format similar to those of AmazonWeb Services. This meets (F2). The monitoring server can monitor each mini-cloud as well as other servers. It is based on open source monitoring software Hinemos [4]. Engineers can learn about cloud monitoring by using Hinemos codes.

III. Alterability

The case study was to develop education content for alternating edubase Cloud infrastructure to improve virtual machine launching performance. End users in the case study used a distributed image processing application on edubase Cloud as a distributed-processing learning material. The main components were Hadoop and openCV. In this application, the time it takes to launch the virtual machine affects usability. Cloud engineering students performed experiments to determine how to improve the virtual machine launching time by modifying the node controller code of eucalyptus [2]. They worked in a mini-cloud and modified the node controller code for the mini-cloud in order to generate a machine image template beforehand.

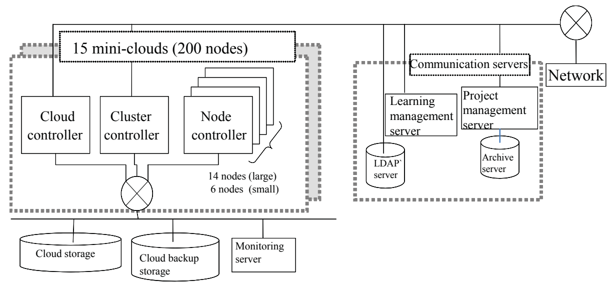


Figure 2. edubase Cloud architecture

The class work proceeded as follows:

1) Bottleneck analysis of virtual machine launch
2) Research on how to improve performance
3) Measurement in a real environment
4) Analysis of measurement results
5) Report and demonstration

Figure 3 shows how to launch virtual machines in eucalyptus.

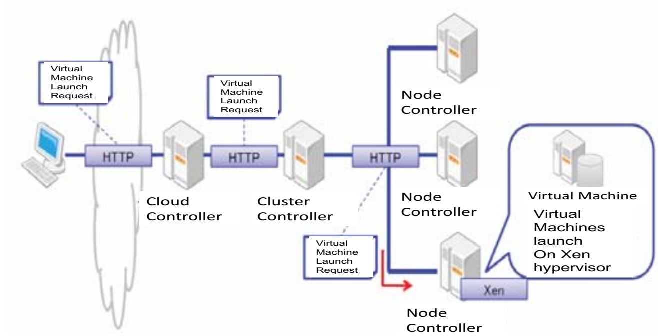


Figure 3. Configuration of the cloud components

A. Bottleneck analysis of virtual machine launch

1) Eucalyptus Source Code Analysis
   The analysis identifies the sub-processes shown in figure 4.
2) Eucalyptus Performance Profiling
   The performance profiling shown in figure 4 is done on the sub-processes. In the measurements of the case study, the machine image size was 4 Gbytes, and had 1 core, 256 Mbytes of memory and 10GBytes of storage.

The sub-processes of creating the ephemeral and swap spaces take up the most time.

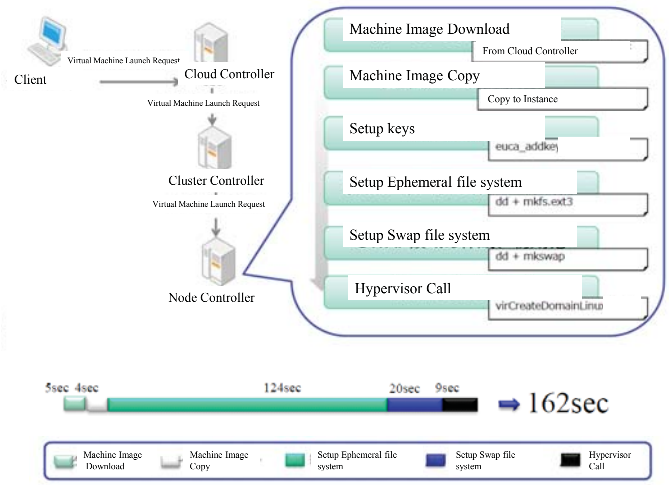


Figure 4. Results of source code analysis and performance measurement

B. Research on how to improve performance

The results of the performance analysis in Step 1 indicate that the pre-execution of the marked procedure in figure 5 is a way to improve performance. The targets of pre-execution are the machine image download from the cloud controller, the machine image copy for launching virtual machines, the ephemeral space and swap space creation.

Because the actual processing is done in the node controller, its code is modified for pre-execution.

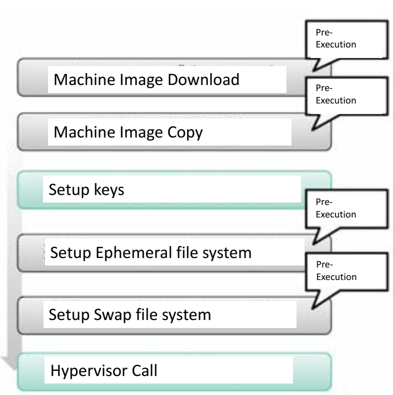


Figure 5. Solution for performance improvement

C. Measurement in a real environment

The students used one mini-cloud exclusively for two months to make modifications to the eucalyptus node controller code and conducted performance measurements.

The results of the measurement are shown in figure 6. Before the improvements were made, the time to launch was proportional to the machine image size and simultaneous launch instance numbers. This is reasonable because of the long time taken to copy a virtual machine image to each virtual machine instance. In contrast, after the improvements were made, the time to launch became independent of the machine image size. In addition, the launch times of simultaneous virtual machines became shorter as well. The modifications made it possible to launch instances about 20 times faster than before.

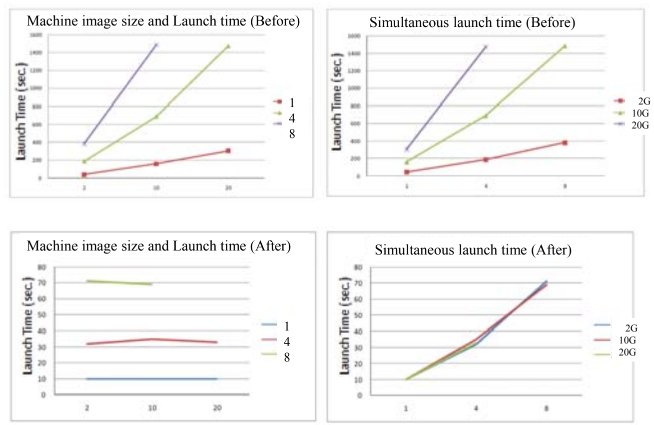
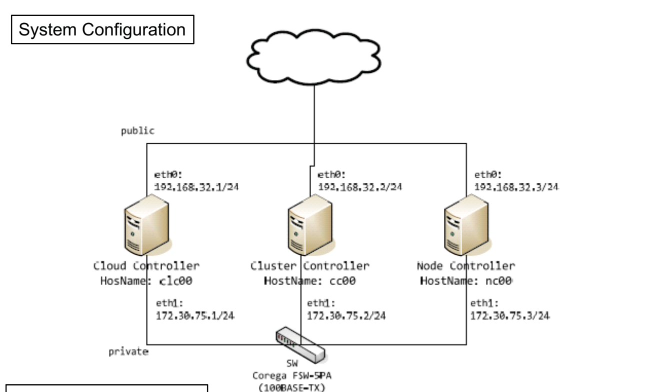


Figure 6. Performance measurement results

Figure 7 shows the system and hardware configuration for the performance measurement.



Hardware Configuration

	Cloud Controller	Cluster Controller	Node Controller
Model	Dell PowerEdge 2950	Dell PowerEdge SC1435	Dell PowerEdge SC1435
CPU	Xeon E5405 2.00GHz (4core x 2cpu)	Opteron 2376 (4core x 2cpu)	Opteron 2376 (4core x 2cpu)
Memory	2006 Mbytes	16121 Mbytes	15732 Mbytes
HDD	2.7 Tbytes	160 Gbytes	160 Gbytes

Figure 7. System and hardware configuration

D. Analysis of measurement results

1) Trade-off between speed and storage space for virtual machine images:

The modifications were effective because they produced a copy in advance when running multiple virtual machines on the same image.

In general, it is not known how many virtual machines will be launched by the same machine image. If the number is large, this pre-execution method is effective. Since virtual machine images are stored on local disk space of the node controller, there is a trade-off between disk space and launch-time speed up.

2) Security considerations:
The code modifications were subject to certain security considerations. Virtual machine image for each instance has to be deleted when users shut it down. This is because references from other users have to be avoided after changes have been made to the virtual machine images. Due to this consideration, the access management to the image copies has to be managed by modified codes.

3) Report and demonstration:
Besides making reports on their work, students participated in a debriefing to share information on the project outcomes. In the debriefing, demonstrations of launching virtual machines were carried out using a virtual machine image of the actual edubase Cloud.
Moreover, we used this use case in the cloud technical course of the TopSE education program [3].

IV. Discussion

The following findings from the case study are noteworthy.

A. The multi-cloud architecture is effective.

The education material developers used a mini-cloud for two months and other users of edubase Cloud could work as usual during this development. Each mini-cloud size is static, but if we could change it dynamically, we would be able to get better performance results for almost the same amount of effort.

B. Archiving outside of mini-clouds works well.

Each time the developers changed the platform code from the original one to the improved one and vice versa, the application machine images were restored to the mini-cloud easily and rapidly from the archive.

C. An open-source cloud platform is essential for cloud engineers' education.

Cloud engineers should not have to develop IaaS from scratch or have to add code to binary IaaS. Instead, they should learn an open-source cloud platform and how to modify it in response to various requests.

D. A mini-cloud initialization tool is important for infrastructure education materials.

For management, an educational cloud should have an automatic mini-cloud initialization tool so that alterations can be easily reset. We have a tool for this purpose but we have yet to study how well it enhances usability.

V. Conclusions

The case study indicates that edubase Cloud architecture is effective for educating cloud engineers. The exclusiveness and alteration features were shown to work well. Without this sort of cloud platform, it would be necessary to provide hardware for each purpose. It would be difficult to provide the necessary hardware in a short period of time. The knowledge we gained in the case study will be applied to the development of other educational materials for cloud platform deployment and cloud-on-cloud systems. These materials will be used in the Top SE education program cloud technical courses [5].

Acknowledgment

We would like to thank al the contributers to edubase Cloud, especially Shin-ichi Honiden, Yoshitaka Kuwata, Masashi Ochida, Satoshi Yazawa, Masaru Nagaku and Prabin Karanjit.

References

[1] Nobukazu Yoshioka, Shigetoshi Yokoyama, Yoshionori Tanabe, and Shinichi Honiden, "edubase Cloud: An Open-source Cloud Platform for Cloud Engineers," SECLOUD '11 Proceedings of the 2nd International Workshop on Software Engineering for Cloud Computing, 2011

[2] Daniel Nurmi, Rich Wolski, Chris Grzegorczyk, Graziano Obertelli, Sunil Soman, Lamia Youseff, and Dmitrii Zagorodnov, "The Eucalyptus Open-source Cloud-computing System," 2009 J. Phys.: Conf. Ser. 180 012051

[3] Shinichi Honiden, Yasuyuki Tahara, Nobukazu Yoshioka, Kenji Taguchi, and Hironori Washizaki, "Top SE: Educating Superarchitects Who Can Apply Software Engineering Tools to Practical Development in Japan," Proceedings of the 29th International Conference on Software Engineering (ICSE 2007), ACM Press, pp.708-717, 2007.

[4] Hinemos: http://en.wikipedia.org/wiki/Hinemos

[5] TOP SE: http://www.topse.jp/en