World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:10, 2012

Ezilla Cloud Service with Cassandra Database for Sensor Observation System

Kuo-Yang Cheng, Yi-Lun Pan, Chang-Hsing Wu, His-En Yu, Hui-Shan Chen, Weicheng Huang

***Abstract*—**The main mission of Ezilla is to provide a friendly interface to access the virtual machine and quickly deploy the high performance computing environment. Ezilla has been developed by Pervasive Computing Team at National Center for High-performance Computing (NCHC). Ezilla integrates the Cloud middleware, virtualization technology, and Web-based Operating System (WebOS) to form a virtual computer in distributed computing environment. In order to upgrade the dataset and speedup, we proposed the sensor observation system to deal with a huge amount of data in the Cassandra database. The sensor observation system is based on the Ezilla to store sensor raw data into distributed database. We adopt the Ezilla Cloud service to create virtual machines and login into virtual machine to deploy the sensor observation system. Integrating the sensor observation system with Ezilla is to quickly deploy experiment environment and access a huge amount of data with distributed database that support the replication mechanism to protect the data security.

***Keywords*—**Cloud, Virtualization, Cassandra, WebOS

I. Introduction

THE main mission of Ezilla is to provide a friendly interface to access the virtual machine and quickly deploy the high performance computing environment. The Ezilla [1], [17]a private Cloud deployment toolkit, has been developed by the Pervasive Computing Team at the NCHC. Through the Ezilla, Cloud users can create and customize their virtual computing environment for the applications by sending their specified requirements. Therefore, users can install the software on demand and configure variables according to their requirements and users do not worry that their own virtual cluster would be interfered by other users. Actually, users need to execute scientific or engineering programs and simulation that probably spend lots of time to set their specified environment. The various simulations will need the different software and configurations on demand. If all of Cloud users utilize the same physical cluster simultaneously, it is possible to affect the other Cloud users. To avoid this problem, the virtual clusters can make users that have the independent computing environment to execute specific programs on demand. The main goal of this project is to simplify the process to utilize Cloud service, and thus to provide scientists as well as general users a friendly Cloud environment.

Kuo-Yang Cheng, Yi-Lun Pan, Chang-Hsing Wu, His-En Yu, Hui-Shan Chen and Weicheng Huang are with the National Center for High-Performance Computing, Taiwan; email:{kycheng, serenapan, hsing, yun, chwhs, whuang}@nchc.narl.org.tw.

It is a lightweight approach helping users to access virtual computing resources and storage service.

Cloud computing is an important service of providing access to remote resources anytime and anywhere. Actually, the virtual computing resource management is one of the most important issues in Cloud Computing environment. However, specific Cloud users manually build specified virtual cluster with the console mode in order to create or manage virtual resources. To improve this condition, an Ezilla toolkit has been developed and built based on the "Carry-On-Cloud" concept [1], [17]. Cloud users can conveniently and directly utilize the virtual machines on Ezilla instead of using local physical resources. Cloud users can build on-demand private Cloud resources just with one click in Ezilla environment. Ezilla provides a user-friendly interface that is based on WebOS [2] for Cloud users. The major mission of Ezilla is to reduce the barrier for using Cloud computing environment. Therefore, the sensor observation system is based on the Ezilla to store sensor raw data into distributed database. We adopt the Ezilla Cloud service to create virtual machines and login into virtual machine to deploy the sensor observation system. There are three major components that consist of Cassandra database, real time information and management interface in sensor observation system. If the sensor observation sensor system receives sensor data, we will store the data into Cassandra database. Therefore, this research not only focuses on virtual resources management with an interactive graphical user interface, but also integrates the Cassandra database technology to provide the high quality of storage in the virtualized environment.

The remainder sections of the paper are organized as follows: Section 2 gives a briefly review on related studies. In Section 3, we proposed the architecture of Ezilla. And the following Section 4, the design and implementation of sensor observation system on Ezilla are elaborately described. Finally we conclude this paper and picture future plans in Section 5.

II. Related Works

A. Existing Web-based Operating System (WebOS) Projects

The WebOS - Chrome OS is one of the famous Web-based Operating System that is based on the AJAX technique [3] and implement a web application to communicate with a server in the background.

The Cloud WebOS platform provides a practicable service via AJAX technique. The Chrome OS is practicable, but Chrome OS does not support on-demand applications and computing services to users in Clouds.

There are various Web-based Operating System projects that consist of Network of Workstations [4], FlyakiteOSX [5], Glide OS [6], Xindesktop [7], etc.

B. Virtualization Technologies

The virtualization is one of the important technology, but the traditional processor is not integrated the virtualization technology on processor. Because the traditional virtualization technology is based on software to simulate each component of guest OS, the performance is limit to simulate and communication. Therefore, both the Intel and the AMD enhance the original processor to support the virtualization technology in 2006. Intel and AMD also implemented the I/O virtualization technology that consists of memory, disk and network on the chipset. The virtualization technology was called hardware-assisted virtualization that improves the performance by reducing a layer of complex software techniques between the hardware and guest operation system. There are various platforms that consist of Linux KVM, VMware Workstation and Microsoft Hyper-V can work on the hardware with virtualization technology.

The hypervisor is one of the major technologies for virtualization. Hypervisor can centrally manage multiple virtual machines on a single computer system. Hypervisor is between the operating system and hardware that handle the computing hardware and virtual machines. The implementations of hypervisor are divided into two categories that consist of Host-based and Bare-metal approaches. The host-based approach adopts modified operating systems to provide virtual machine monitor particular information , for instance Linux-VServer [8], Solaris Zones [9], and Kernel-based Virtual Machine (KVM) [10]. Otherwise, the bare-metal approach adopts small-dedicated hypervisors to directly execute on physical machines. There are various famous examples of the bare-metal approach that consists of The VMware ESX server [11] and the XenServer [12]. We integrate the success of the virtual technologies that consist of KVM and WebOS to build the Ezilla environment. Ezilla provide a new and lightweight approach to access virtual computing services in Cloud.

C. Distributed Database

Apache Cassandra is an open source distributed database management system and the data model is from Google's BigTable. Apache Cassandra supports a highly scalable, eventually consistent, distributed, structured key-value store and provides a replication method to protect the data security [13]. Cassandra provides a ColumnFamily-based data model richer than typical key/value systems [14]. The advantage of distributed database is reliability and availability. The distributed database is scalable service that increases dynamically database scale to provide a mass storage service. In order to upgrade the reliability, a replication technology is implemented on distributed database.

In addition to support reliability, the replication technology can improve the performance.

Because data is stored near the site of greatest demand, parallel access database systems and load on the databases to be balanced among servers. If one of the database servers is failure, system will start the replication technology to provide other replicas to access for users.

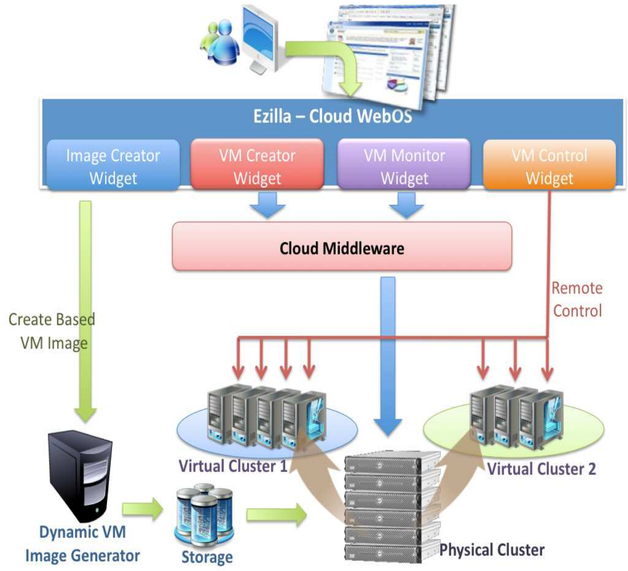

Fig. 1 The architecture of Ezilla

III. The EZILLA Cloud Service

A. The Architecture of Ezilla

Ezilla can integrate computing, networking and storage resources to provide a user-friendly web-based interface for Cloud users. If Cloud users adopt Ezilla toolkit, users can easily and quickly build a private Cloud by themselves.

Cloud users need an independent computing environment to execute particular programes and jobs. Therefore, a private Cloud is aservice-oriented application strategy that implemented by virtualization technology and developed an automatic management mechanism.

Therefore, Ezilla provides an open Cloud architecture and a fully web-based solution that enables Cloud users to simply utilize virtual and real computing resources via a user-friendly interface. There are three key components in the Ezilla toolkit that consists of Ezilla Server, Ezilla Client, and intuitive Ezilla WebOS Interface [17] as shown in the Fig. 1.

The Ezilla Server must manage the all computing resources, and monitor the all status that consist of physical and virtual machines via Cloud middleware.

In addition, the virtual machine images are stored in the file system and Ezilla is depended on the image file server to manage each virtual machine images. If Cloud users send a request to Ezilla server, the server will create a new virtual machine with specific image for users.

A Cloud user can own his personal WebOS interface to create a Cloud environment with storage, networking, and computing resources simply and quickly. The Cloud users can generate dynamically a virtual cluster that includes a couple of virtual machines via Ezilla WebOS Interface.

Users can determine the scale of resources and the life cycle of the virtual cluster on demand. Once the job is done via virtual cluster and the user is finished the job, the physical computing resource can be released for other users to exercise. Because the adoption of the XML-RPC based API for web-based WebOS, users can manipulate the whole operation via a web browser, such as Chrome.

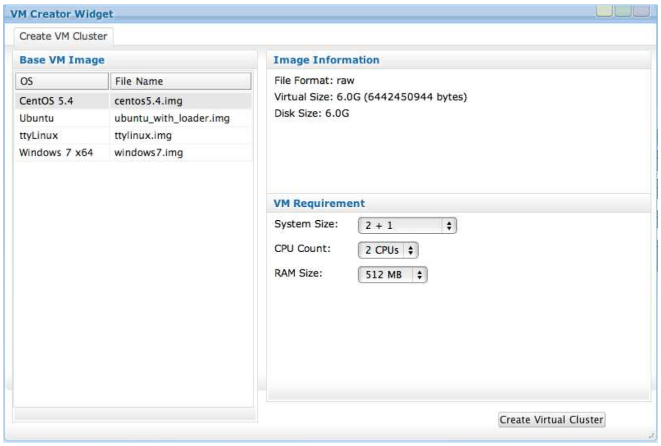

Fig. 2 VM creator widget

There are various middlewares embedded in Ezilla Toolkit that consists of the basic Operation System and packages and OpenNebula [15]. The Ezilla toolkit supports the Debian OS and basic packages that consist of KVM hypervisor, and Libvirt [16]. Ezilla adopt OpenNebula to handle the Cloud management. OpenNebula is responsible for allocating available computing resources, creating VMs based on user selected VM image, and deploying the image onto the physical computing resources.

As network configuration, OpenNebula manages unique IP address, MAC address, and virtual network (vNet) ID. Therefore, each virtual cluster executes on dedicated vNet. Once the Ezilla Server is working, Ezilla clients can be setup automatically. However, Ezilla can regulate dynamically computing resources, thus to enhance the flexibility of Cloud resource allocation.

B. The Ezilla Cloud Service

Ezilla is based on WebOS to develop basic Widgets and advance Cloud Computing Widgets. Ezilla WebOS Interface is one of key techniques that develop many Cloud Widgets with a friendly graphical user interface in Ezilla [17].

There are four key Widgets that consist of Image Creator Widget, VM Creator Widget, VM Monitor Widget, and VM Control Widget in Ezilla. Duo to users have a graphical user interface, users can simply and intuitively manage all of these widgets in Ezilla. According to user's requirements, user can adopt the particular Widgets to arrange their complicated computing tasks via Ezilla Toolkit.

The Image Creator Widget is to generate the specific base image that consists of specified applications for user. The Image Creator Widget provides a complete software stack that consists of operating system, management tools, monitor resources and commercial package.

According to user's requirements, VM Creator Widget will generate a profile that consists of image, system size, CPU count and Memory size shown in Fig. 2. This profile is parsed by the VM Creator engine to create an on-demand virtual cluster on the physical computing resources. The main mission of VM Monitor Widget is to monitor the all the status of virtual machines and the physical hardware that are shown in Fig. 3. Ezilla toolkit provides two methods to login and control VMs that consist of VNC and SSH on Web.

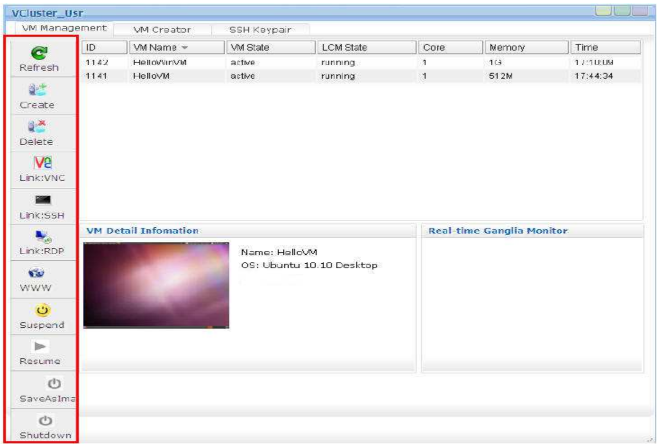

Fig. 3 VM monitor widget

IV. Design and Implementation of Sensor Observation System

A. The Architecture of Sensor Observation System

The sensor observation system is based on the Ezilla to store sensor raw data into distributed database. We adopt the Ezilla Cloud service to create virtual machines and login into virtual machine to deploy the sensor observation system.

The architecture of sensor observation system is shown in Fig. 4. There three key components that consist of Cassandra database, real time information and management interface in sensor observation system. If the sensor observation sensor system receives sensor data, we will store the data into Cassandra database.

As shown in Fig. 5, the sensor observation system provides the particular website for user to query the sensor data from Cassandra database. The user can select one of the multi sensor modules that consist of various sensors. If user selects specific multi sensor module, user can choose one of the sensor in particular multi sensor module.

As shown in Fig. 6 user selects the humidity sensor to query the historical data during the particular time. According to these query results, the sensor observation system can transfer the data to a statistic chart to provide the visualization service.

As the real time information, the sensor observation system received a large amount of sensor data and stored into Cassandra database. The sensor observation system supports a real time information service that can show the newest data on website dynamically. However, user can observe the newest data of specific sensor on demand.

As the management interface, there are three key components that consist of login authentication, multi sensor module and sensor information in the sensor observation system. The administrator inputs username and password can login to management interface to manage multi sensor module and sensor information. The administrator can create, edit or delete any multi sensor module and sensor information. If you modify any data, the change will be updated to Cassandra database to support the newest sensor data and information to query by normal users.

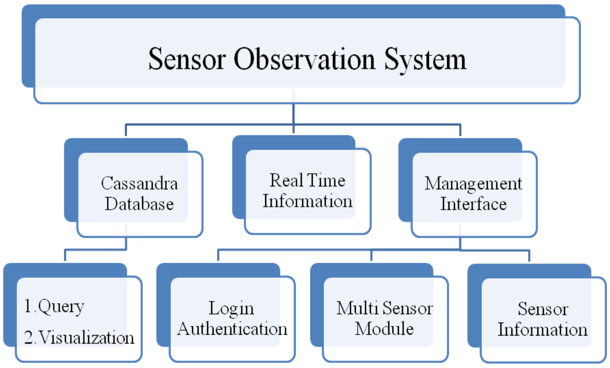

Fig. 4 The architecture of sensor observation system

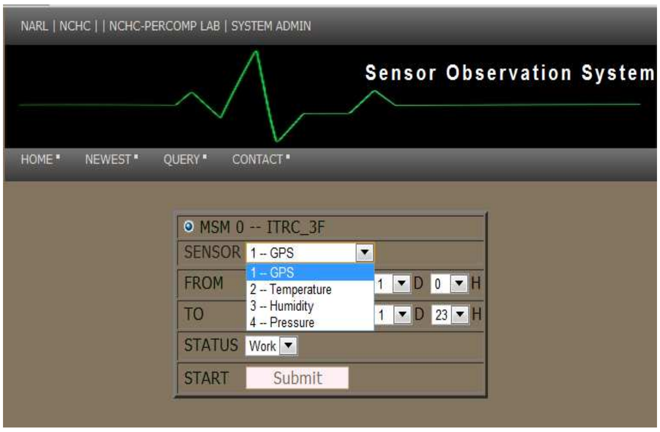

Fig. 5 The form of data query

V. Conclusions and Future Work

Ezilla has been developed by Pervasive Computing Team at National Center for High-performance Computing (NCHC). Ezilla integrates the Cloud middleware, virtualization technology, and Web-based Operating System (WebOS) to form a virtual computer in distributed computing environment.

The main mission of Ezilla is to provide a friendly interface to access the virtual machine and quickly deploy the high performance computing environment. In order to upgrade the dataset and speedup, we proposed the sensor observation system to deal with a huge amount of data in the Cassandra database.

Integrating the sensor observation system with Ezilla is to quickly deploy experiment environment and access a huge amount of data with distributed database that support the replication mechanism to protect the data security.

The NCHC's Ezilla Development Team will develop additional Cloud service to improve speedup and deal with a huge amount of data in the future.

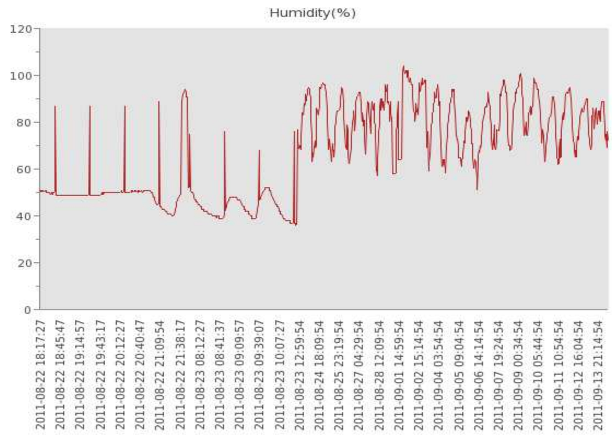

Fig. 6 The query results of humidity

References

[1] Ezilla , http://ezilla.info

[2] G. Gu, and X. Lu, "Simple Web OS System Based on Ext Framework and Cloud Computing," IEEE International Forum on Information Technology and Applications, pp. 448-450, 2010.

[3] http://www.chromium.org/chromium-os/

[4] Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, David E. Culler, Joseph M. Hellerstein, and David A, Patterson, "Searching for the Sorting Record: Experiences in Tuning NOW-Sort," The 1998 Symposium on Parallel and Distributed Tools , Welches, Oregon, August 3-4, 1998.

[5] http://osx.portraitofakite.com/logon.htm

[6] http://www.glidedigital.com/

[7] http://www.xindesk.com/

[8] S. Soltesz, H. P¨otzl, M. E. Fiuczynski, A. Bavier, and L. Peterson, "Container-based Operating System Virtualization: A Scalable, High-Performance Alternative to Hypervisors, " Proceedings of ACM SIGOPS/Eurosys European Conf. on Computer Systems, pp. 275-287, Mar, 2007.

[9] D. Price and A. Tucker, "Solaris Zones: Operating Systems Support for Consolidating Commercial Workloads, " Proceedings of 18th Large Installation System Administration Conf., pp. 241-254, Nov, 2004.

[10] B. Zhang, X. Wang, R. Lai, Y. Liang, Z. Wang, Y. Luo, and X. Li, "Evaluating and Optimizing I/O Virtualization in Kernel-based Virtual Machine (KVM)," International Conference on Network and Parallel Computing, pp. 220-231, Zhengzhou, China, September 13-15, 2010.

[11] John Paul, "VMWare ESX Server Workload Analysis: How to Determine Good Candi-dates for Virtualization," 33rd International Computer Measurement Group Conference, pp. 483-484, San Diego, CA, USA, December 2-7, 2007.

[12] X. Ge, H. Jin; S. Wu, X. Shi, W. Gao, "A method of multi-VM automatic network configuration," Intelligent Computing and Intelligent Systems, pp. 309-313, 2009.

[13] http://cassandra.apache.org/

[14] http://wiki.apache.org/cassandra/

[15] Milojičić, D., Llorente, I.M., Montero, R.S, 'OpenNebula: A Cloud Management Tool', Internet Computing, IEEE, pp. 11-14, 2011.

[16] Bolte, M., Sievers M., Birkenheuer, G., Niehorster O., and Brinkmann A. 'Non-intrusive Virtualization Management using libvirt', Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, Germany, pp. 574-579, 2010.

[17] Chang-Hsing Wu, Yi-Lun Pan, Hsi-En Yu, Hui-Shan Chen and Weicheng Huang, "Power Saving for Fast Deployment Private Cloud Toolkit – Ezilla with Infrastructure Services", The 2012 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA'12), Las Vegas, Nevada, USA, July 16-19, 2012.