Distributed Hybrid Cloud Architecture Based on China Education Research Network for Common Computer Test

Gang Shi[1,2]
[1.] School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China
[2.] Department of Computer Sciences, Tsinghua University, Beijing 100084, China
E-mail: sg11@mails.tsinghua.edu.cn

Chongguo Wang[*]
School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China
E-mail: wcg6808@163.com

***Abstract*—This paper provides a hybrid cloud architecture for Computer Grade Test for Universities in Xinjiang based on China Research Education Network (CERNET), gives the resolving scheme for security, load balancing and disaster backup and develops the prototype system. The hybrid cloud architecture for Xinjiang Common Computer Test (CCT) has the following key characteristics: high concurrency, sudden congestion, scalability and high security.**

Keywords-Hybrid Cloud Architecture; High concurrency; Security; Scalability; Load Balancing; Xinjiang CCT

I. INTRODUCTION

With the development of information technology, non-paper has been a trend for University special examination such as Xinjiang common computer test (CCT), college English test (CET) 4 or 6. Common grade test systems for universities have the similar characteristics facing to the challenges such as high concurrency, sudden congestion, scalability and high security in the practical Implementation. Cloud architecture or security [1-2] is always the hot spot for cloud computing research in recent years.

Information system migrating cloud end [3] is really a trend for technology development. Despite the short-term rental a lot of servers in some data center can meet the requirements of the test system to provide testing services, this simple migration may bring the difficulty of reliable access for a wide range of examination under the condition of current network.

The paper provides the hybrid examination cloud architecture for Xinjiang CCT based on China Research Education Network (CERNET) and it has the characteristics of high concurrency sudden congestion, scalability and high security [4-5]. A distributed test system was deployed on the university private cloud platform and was uniformly managed by the test center which controls the test cloud platform. The hybrid cloud architecture realizes the load balancing service based on the university student examination certificate information, and the backup service based on a district center node in CERNET. It makes test center configure and manage university distributed cloud test system dynamically.

II. CLOUD COMPUTING TECHNOLOGY

Cloud computing [6-7] is a colloquial expression used to describe a variety of different computing concepts in which a large number of computers that are connected through a real-time communication network (typically the Internet).It exhibits the following five key characteristics: dynamical configuration of the resources，self-help demand service, network-centered , measurable services and resource pooling or transparency. Cloud computing provides three level services: Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). (1)IaaS: IT infrastructures provided in the IaaS cloud can be used directly, which include processing, storage, networks, and other fundamental computing resources. IaaS cloud integrates/decomposes physical resources in an ad-hoc manner to meet growing or shrinking resource demand of cloud consumers by the virtualization technology. Setting up independent virtual machines (VM) is the basic strategy of virtualization, that are isolated from both the underlying hardware and other VMs. To be different from the multi-tenancy model, this strategy aims to transform the application software architecture in order to make multiple instances (from multiple cloud consumers) run on a single application (i.e. the same logic machine). Amazon's EC2 is an example of IaaS. (2)PaaS: PaaS is a development platform allowing cloud consumers to develop cloud services and applications (e.g. SaaS) directly on the PaaS cloud and it supports the full "Software Lifecycle". Hence the difference between PaaS and SaaS is in that PaaS offers a development platform that hosts both in-progress and completed cloud applications, whereas SaaS only hosts completed cloud applications. In addition to supporting application hosting environment, PaaS is required to possess development infrastructure including programming environment, tools, configuration management, and so on. An example of PaaS is Google AppEngine. (3) SaaS: Different cloud consumers' applications are organized in a single logical environment on the SaaS cloud to achieve economies of scale and optimization in terms of security, speed, availability, disaster recovery, and maintenance. Examples of SaaS include Google Docs, Google Mail, SalesForce.com, and so forth. Cloud consumers release their applications which can be accessed through networks from various clients (e.g. web browser, PDA, etc.) by application users on a hosting environment. The cloud

infrastructure often employ a multi-tenancy system architecture not controlled over by cloud consumers.

More recently, four cloud deployment models [8] have been defined in the Cloud community: public cloud, private cloud, community cloud and hybrid cloud. The hybrid cloud architecture is adopted for Xinjiang CCT Test. (1) The motivation to setup a private cloud within an organization has several aspects. First, it is to optimize and maximize the utilization of existing in-house resources. Second, security concerns also make Private Cloud an option for many firms, which include data privacy and trust. Third, organizations always require full control over mission-critical activities that reside behind their firewalls. Fourth, the data transfer cost from local IT infrastructure to a Public Cloud is still rather considerable. Last, the private cloud is often built for research and teaching purposes. The cloud infrastructure of Private cloud is operated solely within a single organization, and managed by the organization or a third party regardless of whether it is located premise or off premise. (2)Community cloud is jointly constructed by several organizations. They share the same cloud infrastructure as well as policies, values, requirements, and concerns. The cloud community conforms to a degree of economic scalability and democratic equilibrium. The cloud infrastructure could be hosted by a third-party vendor or one of the organizations in the community. (3) Public cloud is the dominant form of current Cloud computing deployment model. Many popular cloud services are public clouds including Force.com, Amazon EC2, S3 and Google App Engine. The general public cloud consumers usually used the public cloud and the cloud service provider has the full ownership of the public cloud with its own value, policy, costing and profit, and charging model. (4) Hybrid cloud. The cloud infrastructure is a combination of two or more clouds (public, private or community) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds). Hybrid cloud has raised the issues of standardization and cloud interoperability. Organizations adopt the hybrid cloud model to optimize their resources to improve their core competencies by margining out peripheral business functions onto the cloud while controlling core activities on-premise through private cloud.

The paper provides a hybrid cloud architecture based on CERNET for Xinjiang CCT. The bandwidth between Master node of CERNET in Xinjiang area and other universities is usually below 1G, while in oter universities it is far more than 1 G. The large differences in bandwidth make hybrid cloud reasonable.The hybrid cloud architecture of Xinjiang CCT in figure 1 is realized through building the test control cloud and connecting with other university test service clouds.Distributed data centers based on the required data for every university are built in other district university and are deployed in every university private cloud. Every university builds the private cloud platform Iaas with infrasture itself and makes resources pool.When test is required,every university test office can rent private cloud platform in its own university and connect the control cloud of test center.The test center provides test paper distribuding or summary services.Paas is the key of development.SaaS on the top level is realized by every test system.Three level cloud service for Xinjiang CCT sees figure 2.

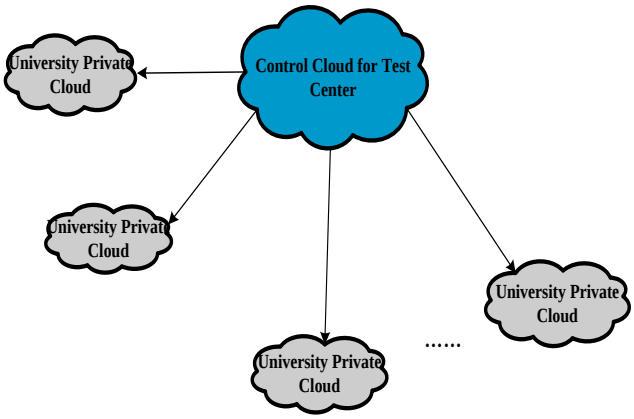

Figure 1.The hybrid cloud architecture for Xinjiang CCT

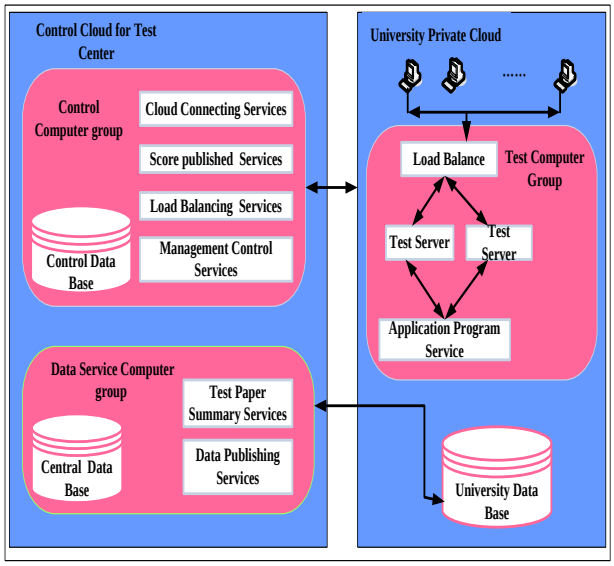

Figure 2.Three level hybrid cloud service for Xinjiang CCT

III. LOAD BALANCING SERVICE BASED ON EXAMINEE INFORMATION

With the growth of WWW application users, improving server performance alone is not enough to solve the problem of the Web server pressure. Systems are often designed to distributed ones to build scalable Web services to evenly distribute required connection to different Web servers, which is known as a Web service load balancing.

Commonly used methods for load balancing are the following: (1) DNS load balancing .To achieve load balancing, this method configures the same name for

multiple addresses in the DNS and assigns different access addresses for multiple user requests in order to access different servers. (2) Proxy server load balancing. This method achieve load balancing by evenly distributing the requests to the internal multiple servers. (3)Address translation gateway load balancing. To achieve load balancing, NAT gateway can map an external IP address for multiple internal IP addresses and only dynamically uses one of the internal addresses for each TCP connection request. (4) NAT (Network Address Translation) load balancing. This method translates an IP address into another IP address. (5) Reverse Proxy load balancing .It refers to the proxy server to accept connection requests on the internet, and then forwards the request to the server on the internal network. At that time, the results obtained from the server is returned back to request connected clients on the internet and the proxy server is considered as a server. (6) Mixed load balancing. You can use the most suitable load balancing methods across multiple servers within a cluster, respectively and load balancing once again is carried on between servers. That can make all the servers provide services to the outside world as a whole so as to achieve the best performance.

From the entire cloud architecture point of view, a distributed examination system uses hybrid load balancing. Reverse Proxy or load balancing is used inside every university examination private cloud based on the specific number of examinees and the hardware environment. NAT load balancing mode is used between the private cloud of colleges and universities and the central control clouds.

However, being different from load balancing scheduling policy of Quality of Service (QoS) and service level agreements (SLA) used by common Web services, it is based on examinee information for load balancing scheduling strategy between university private cloud and the central control and address is translated according to the test center information of examinee . This allows translating the access address of examinee request of different universities to point to its own university private cloud, then the traditional Web server of load balancing within the university makes the requests point to the specific examination server.

This method makes CERNET regional central node be the data disaster recovery center and this node database contains all test papers data of test sites within the region. When a cloud service failures in some university test site within the region, the examination service IP of this test sites is automatically pointed to regional central node so as to improve the reliability of the entire examination system.

IV. THE SOLUTION OF SYSTEM SECURITY

Now let's present physical security strategies and the solution. Security connection is built among every university private cloud through simulating the hardware security of test center control cloud. The securities of access [10-12] include physical security, data security, behavior security and system security. (1) Physical security. The concept of the cloud may be misleading at times, and people forget that everything is somewhere actually tied to a physical location. Security model may need to be reevaluated for customers lose control over physical assets,. The massive investment makes company not build their own physical data centers, but to move to cloud services in the first place. Physical security mainly points to the security of physical devices and basic network communication. Physical devices have the function of disaster recovery. Customers rely heavily on 24/7/365 access to their services and any interruption in access can be catastrophic in the SaaS environment. Virtual server can be copied, backup, and moved just like a file (live migration) by using the virtualization software. Virtual software has the following benefits: Quickly reallocating computing resources without any downtime and ability to deliver on service-level agreements and provide high-quality service. Firstly, the computer group is redundant, which can execute the job of duplexing and disaster recovery. Secondly, the data base server in every university is distributed one and every university only owns its own data. Thirdly, central data base is built in Xinjiang University. If some university cannot provide the service, it may login Xinjiang University and use the backup service. Virtual Private Network (VPN) connects central data base computer group with every university distributed data base before the test. Communication protocol uses the HTTPS which ensures the security of university education network and distributing or uploading data uses safe file transfer protocol (SFTP) based on secure shell (SSH). (2) Data security. Security will need to move to the data level so that enterprises can be sure their data is protected wherever it goes. For example, the enterprise can specify that this data is not allowed to go outside of the European Union with data-level security. It can provide compliance with the Payment Card Industry Data Security Standard (PCI DSS) and also force encryption of certain types of data, and permit only specified users to access the data. Data security includes three parts: first, data is encrypted, which include distributed data base and publishing and summary data transferred. Second, data is prohibited to be tampered, which ensures the security of examinee test papers or answers. Every examinee will own only verification code produced by HashCode function, which is composed of data, verification password, examination certification number and ID number. Third, Access login is added for any data access operation so that it can be audited and traced. (3) Behavior security. Using management strategies ensures the security of related staff. (4)System security is realized by the test system itself.

V. TESTING AND REALIZING THE PROTOTYPE SYSTEM

Publishing or gathering data is realized by below-in-line security channel in order to ensure data security of the test system. That can simplify system model. Every examinee test paper is generated based on entry information

of every university in the test system before the examination starts, then a test paper data file package of encrypted node is built based on examinee data required in every university private cloud and is distributed to every node through the security channel. Encrypted examinee test paper data package is transferred to the test center through security channel after the test is over.

University private cloud platform is constructed by OpenStack. Every university node data base server, test system web server and load balancing server are packaged up machine images which are directly distributed to every university private cloud node.

Test center control cloud is composed of two group of computers: control computer group and data service computer group. On the one hand, Control computer group mainly provide the services such as accessing cloud, publishing scores, management control, load balancing, etc. On the other hand, data service computer group provide the services of publishing data and summary test paper. Every university private cloud provides load balancing service and application program service, which can provide university test service computer group itself when testing. Connection between central data base and every university virtual private cloud is by the way of elastic access. That solves the following key problems: strategies and management mode for providing elastic access service and data distributing or gathering services for distributed data center.

VI. CONCLUSIONS

The paper provides the distributed test system hybrid cloud architecture and develops the prototype system for Xinjiang CCT based on CERNET. The architecture can have these characteristics with high concurrency, sudden congestion, scalability and high security which meet the requirements of test cloud. The resolving scheme is also provided for security, reliability, load balancing and disaster backup. Which can provides as the reference for the application of distributed hybrid cloud architecture based on CERNET. The further research will focus on on-line communication security mechanism among clouds and service quality monitoring of university private cloud node.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China Grant No.61262023.

REFERENCES

[1] Qi Zhang , Lu Cheng, Raouf Boutaba.Cloud computing: state-of-the-art and research challenges[J], J Internet Serv Appl,2010,1:7-18.

[2] S. Subashinin and V.Kavitha, A survey on security issues in service delivery models of cloud computing[J]. Journal of Network and Computer Applications,2011,Vol.34,No 1,pp.:1-11.

[3] Sushil Bhardwaj, Leena Jain, Sandeep Jain. Cloud Computing Research and Development Trend[J]. International Journal of Engineering and Information Technology,2010,Vol.2,No.1,pp:60-63.

[4] Tuncay Ercana,Effective use of cloud computing in educational institutions[J].2010,2:938-942.

[5] Chunye Gong, Jie Liu, Qiang Zhang, Haitao Chen , Zhenghu Gong. The Characteristics of Cloud Computing[C]. 2010 39th International Conference on Parallel Processing Workshops,2010,275-279.

[6] Mariana Carroll, Paula Kotzé, Alta van der Merwe. 2012. Securing Virtual and Cloud Environments. In: Cloud Computing and Services Science, Service Science: Research and Innovations in the Service Economy), edited by I. Ivanov et al., DOI 10.1007/978-1-4614-2326-3 4, © Springer Science+Business Media, LLC 2012.

[7] Sean Marston, Zhi Li, Subhajyoti Bandyopadhyay, Juheng Zhang, Anand Ghalsasi.Cloud computing − The business perspective[J],2011,51:176-189

[8] Tharam Dillon, Chen Wu ,Elizabeth Chang.Cloud Computing: Issues and Challenges[C]. 2010 24th IEEE International Conference on Advanced Information Networking and Applications,2010,pp.27-33.

[9] Mohammed Alhamad, Tharam Dillon, Elizabeth Chang. Conceptual SLA Framework for Cloud Computing [C]. 4th IEEE International Conference on Digital Ecosystems and Technologies (IEEE DEST 2010),2010,pp.606-610..

[10] Flavio Lombardi and RobertoDiPietro, Secure virtualization for cloud computing[J].2011,34:1113-1122.

[11] Dimitrios Zissis and Dimitrios Lekkas.Addressing cloud computing security issues[J].2012,28:583-592.

[12] Krešimir Popović and Željko Hocenski,Cloud computing security issues and challenges[C]. MIPRO 2010, 2010,pp.344-349.