Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 3, March 2014, pg.458 – 461

RESEARCH ARTICLE

QoS Enhanced Architecture for Cloud Computing Environment

ANBUMOZHI ANBUKKARASAN, LIZA M KUNJACHEN

Dept. of Computer Science, SRM University, India

Dept. of Information and Technology, SRM University, India

Anbumozhi27@gmail.com; Lizamk2006@yahoo.co.in

Abstract- Cloud computing is a popular model for enabling network access to shared pool of computing resources that can be provisioned with minimal effort. There are significant issues prevailing with regard to proficient provisioning. Existing works on cloud computing focuses on creation and deletion of static and dynamic VMs and based on the requests the VMs are recycled [1]. But significant amount of time is required for this process which could be applied in serving more user requests. In this paper we introduce provisioning technique that facilitates adaptive management of system offering end users guaranteed Quality of Services (QoS). To improve the efficiency of the system, we use workload analyzer and queuing techniques to achieve high QoS. A loop free path finding algorithm (LPA) is presented to identify the duplicate nodes and replace with the least cost code

I. INTRODUCTION

Cloud computing has reached heights over the last few years in small, medium and large scale industries. It provides services over the internet through several servers' uses virtualization .In cloud computing you can provide service in three types Iaas, Paas, Saas. IaaS is acronym for Infrastructure as a Service, in which service providers offer physical infrastructure for computation [2]. Best examples for IaaS are Amazon EC2, Windows Azure, and Rackspace [3]. PaaS as the name suggests provides computing platform which typically includes operation system, Environment, database, webserver etc. Examples are Force.com and Google App Engine. Saas is Software as a service model, in which access to application software's are provided. Service providers would do the setup and installation. Google

Apps is one such example for software as a service. The services provided by cloud are much more reliable than any other computing model and more scalable in terms of provisioning.

There are various challenges in cloud computing when it comes to QoS (Quality of Service). Some existing works including [4] tried to control VM provisioning in proactive or reactive manner. Some other works for resource provisioning with for adaptive applications in cloud environments has been examined in [5]. To ensure QoS it is essential to meet SLA's agreed by the end users and to provide a secure service [6]. In this paper, we focus on an effective approach to demonstrate the efficiency of the cloud computing in terms of meeting user requirement.

II. RELATED WORKS

Cloud computing has been widely recognized as the next generations computing infrastructure. Allowing users to use infrastructure platforms and software's. Improving battery life time is discussed in [7]. In [8] the authors proposed an energy management system for virtualized data centers where resource sharing is categorized into local and global policies. In [9], the researchers have proposed the framework of adaptive QoS management Process, QoS framework for mobile cloud computing and they have modeled QoS management system based on FCM (Fuzzy Cognitive Map). How many requests will be accepted by the system, in what way the request is handled, what the system will do if it gets congested etc. are not clearly defined in this paper. In our approach, we have described the scenarios clearly. Based on queuing networks the authors of [10] have proposed an architecture form provisioning multitier applications in cloud data centers.

III. PROPOSED MODEL

In this section, we present an overview of our system architecture, assumption and algorithms that derive our design.

A. System Architecture

In this section a brief description of the working environment and network model is explained. The system consists of a set of Virtual machines 'VM' and each of the VM is connected to the VM server. The number of virtual machines required for serving an application depends on the application type and workload. The task which needs to be attended is classified into groups depending on the requirement and assigned to the queue and based on the priority a request is served.

B. System Model

The cloud computing system is owned and maintained by 3rd party provider such as Amazon EC2, Windows Azure, Go Grid and Rackspace. Cloud organizes data centers as networks and allow providers to deploy and access applications from anywhere in the world. All the software and data requirements for processing a request is met by VM. Quality of service is challenging when it comes to managing the service requests over the network. In this paper we implement various techniques to achieve the QoS targets. The QoS enhanced architecture in cloud environment is shown in Figure1.

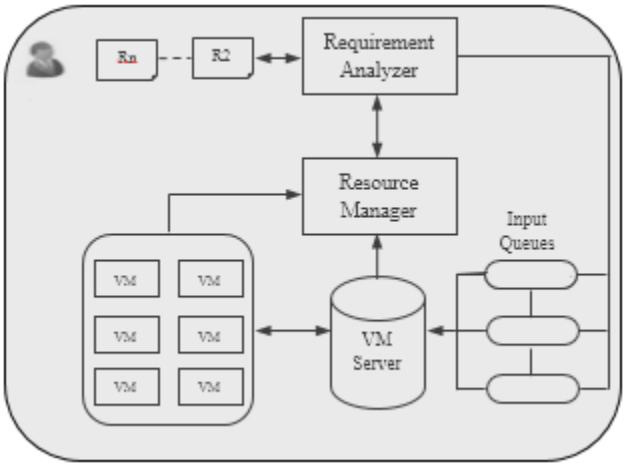

Fig. 1. System Architecture

The proposed model includes the following modules: 1) Requirement analyzer, examines the service requests. When the system reaches the threshold limit or requested data is unavailable the system rejects the service and do not allow the request to be processed further. 2) Resource Manager, periodically check the system usage and resource availability. Based on the recommendation from resource manager the Requirement analyzer decides to process the request or not. 3) VM Server, which determines which of the request time critical is among the requests in all queues and provide virtual machine to that applications, before creating VM wait for those tasks which are waiting for their last request to be finished.

C. Network Model

We set the following model for our approach to continuously ensure Quality of Service at all times. Here, in Fig. 2 the request is controlled by introducing a new method to guarantee service during crunch times. Three protocols are involved in this process, I) Enhanced Service Authentication Protocol (ESA), provides access to VM server directly to existing users. II) Location Request Resource Allocation protocol (LRRA), detects the location of the request data and provides Key based on the node and client ID. The generated key can only be used by the client who requested for the service. III) Genetic Grant Adaptation (GGA) provides access to the requested data by verifying the Key. The purpose of GGA is to check if the key is not misused by any other client.

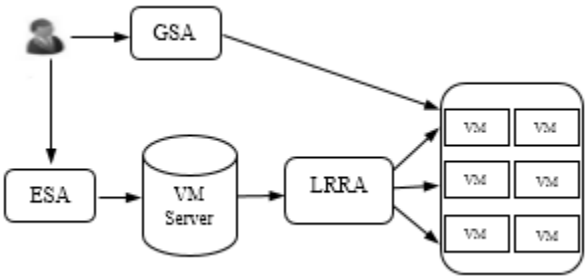

Fig. 2. Network Model

In Fig. 2, the network model is explained. When a client requests for service, the request reaches the Requirement analyzer. If the system is congested the request is not allowed to wait in the Input queue. Rather the client can request the VM server for authentication key for gaining access to the virtual machine directly. The ESA protocol provides access to the server, in term the VM server with the help of index tags identifies the particular VM using the IP address. The LRRA protocol prepares a key based on the location of the data and client details from the ESA protocol. The key is then shared to the VM server which in turn is shared to the client. Now the client uses the key to unlock the VM for the requested data. The GSA protocol verifies the key and grants access to the system. This way waiting time is scrapped for existing users during congestion.

D. Loop- free Path-Finding Algorithm

This paper presents a path finding algorithm which is loop free, LPA clearly checks if the same node is repeated in a single path of source to destination. In addition to this, when a duplicate node is detected the algorithm will do a comparison between the two nodes and selects the least node among the two. In other words, which ever node is nearer is being chosen and the other is removed.

E. Cloud Min Min Scheduling

The biggest drawback in cloud computing is its load imbalanced. CMMS is an improved load balancing algorithm which schedules the service based on the execution time of the request. The task with minimum execution time is given priority over the rest. Hence the minor request need not wait till the completion of major tasks.

F. Priority Based Scheduling

Each process is assigned with a priority and the scheduler always chooses process of higher priority over one of lower priority. The request will low priority is served at the end.

IV. CONCLUSION

Cloud computing is now evolving like never before, with companies of all sizes adapting to this new technology experts believe that this trend will only continue to grow and develop even further in the coming few years. While cloud computing is undoubtedly beneficial for mid-size to large companies, it is not without its downsides, especially when it comes to serving request and security. To counter those complexities over the cloud, this paper presented an adaptive approach for providing service to request during congestion. The goal of the model is to meet the QoS targets related to the service time and rejection rate of requests and utilization of available resources.

Security for confidential data over the cloud is a serious concern. As a future work we are planning to introduce security procedures to third party cloud service provider to ensure company's sensitive information is secure.

REFERENCES

[1] X. Wang, Z. Du, X. Liu, H. Xie, X. Jia, "An adaptive QoS management framework for VoD cloud service centers. 2010 International Conference on Computer Application and System Modeling (ICCASM),Volume: 1, 2010, pp. 527-532.

[2] Y. C. Lee and A. Zomaya, "Rescheduling for reliable job completion with the support of clouds," Future Generation Computer Systems, vol. 26, no. 8, pp. 1192–1199, 2010.

[3] Microsoft Azure Services Platform, "http://www.microsoft.com/azure," accessed date 2nd Feb 2014.

[4] R. N. Calheiros, R. Ranjan and R. Buyya, "Virtual Machine Provisioning Based on Analytical Performance and QoS in Cloud Computing Environments," in Parallel Processing (ICPP), 2011 International Conference, 2011.

[5] Q. Zhu and G. Agrawal, "Resource provisioning with budget constraints for adaptive applications in cloud environments," in Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing (HPDC'10), 2010.

[6] X. Wang, Z. Du, X. Liu, H. Xie, X. Jia, "An adaptive QoS management framework for VoD cloud service centers. 2010 International Conference on Computer Application and System Modeling (ICCASM), Volume: 1, 2010, pp. 527-532.

[7] R. Neugebauer and D. McAuley, "Energy is just another re-source: Energy accounting and energy pricing in the nemesis OS," in Proceedings of the 8th IEEE Workshop on Hot Topics in Operating Systems, 2001, pp. 5964.

[8] A. Beloglazov and R. Buyya, "Energy Efficient Resource Management in Virtualized Cloud Data Centers," in 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing,2010.

[9] P. Zhang and Z. Yan, "A QoS-AWARE SYSTEM FOR MOBILE CLOUD COMPUTING," in Proceedings of IEEE CCIS2011, 2011.

[10] J. Bi, Z. Zhu, R. Tian, and Q. Wang, "Dynamic provisioning modeling for virtualized multi-tier applications in cloud data center," in Proceedings of the 3rd International Conference on Cloud Computing(CLOUD10), 2010.

