ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A Stochatistic Model to Investigate Data Center Performance and QoS in Laas Cloud Computing System

K.Srikanth
M.Tech Scholar,
Christu Jyoti Institute of Technology And Science
Colombonagar, Yeshwanthapur, Jangaon, Telangana

Pranay Kumar BV
Associate Professor
Christu Jyoti Institute of Technology And Science
Colombonagar, Yeshwanthapur, Jangaon, Telangana

ABSTRACT: ***The future research topics to be investigated as a continuation of this and We plan to extend the heterogeneity of our performance model to other parts of IaaS cloud computing centers. And we would like to support full heterogeneity in VMs, PMs and user tasks; more specifically, and VMs are required to be varied in terms of memory, CPU, disk, vCPU and core; The PMs may differ in computation networking, capacity, disk, memory, and Graphics Processing Unit (GPU) and hypervisor (VMM) capabilities; The user tasks may request various number of resources for different amount of time (i.e., task service time different in probability distributions). For Assuming heterogeneous environment brings another challenge which is known as decision making of task assignment policy. While this issue has been extensively studied in the area of cluster and grid computing, proposed techniques are not quite applicable/useful to cloud computing paradigm due to different nature of cloud architecture and user requests. At one direction of future research can be proposing an efficient task assignment strategy whereby the utilization of cloud resources as well as the performance indicators (e.g., response time, waiting time, probability of immediate service and blocking probability) are improved simultaneously.***

I. INTRODUCTION

We first describe Cloud Computing as a new paradigm of computing. In this main challenge and obstacle of cloud computing, Quality of Service (QoS), is defined and the general aspects of QoS are explored. Later on the performance evaluation in cloud computing, the related work and the challenges are highlighted. Here We wrap up the Chapter with outline and organization of the thesis proposal. The development of human society, the basic essential services are commonly provided such that everyone can easily obtain access to them. Nowadays, utility services such as electricity, water, gas, telephony are deemed necessary for fulfilling daily life routines. Frequently these utility services are accessed so that they need to be available at any time. The Consumers are then able to pay service providers based on their usage of these utility services.

In the year1969, Leonard Kleinrock said: "Since of now, computer networks are still in their infancy, however as they grow up and become sophisticated, and then we will probably see the spread of 'computer utilities' which, like present telephone and electric utilities, resolve service individual homes and offices across the country." In this vision of the computing utility based on a service provisioning model anticipates the massive transformation of the entire computing industry in the 21st century, where by computing services will be readily available on demand, like other utility services. For as a result, there will be no need for the consumers to invest heavily in building and maintaining complex IT infrastructure; instead, they will pay only when they access computing services.

Nowadays, significant innovations in virtualization and distributed computing, and to better access to high-speed Internet, and they have accelerated interest in cloud computing. This is quickly gaining acceptance: For IDC, they spent 17 billion dollars on cloud-related technologies, for software and Hardware in 2009, and spending 45 billion by 2013 is expected to grow. The Cloud computing has a service oriented architecture in which services are broadly divided into three categories:

Infrastructure-as-a-Service (IaaS), where equipment such as hardware, servers, storage, and networking components are made accessible over the Internet); Platform-asa- Service (PaaS), which includes computing platforms—hardware with operating systems, virtualized servers, and the like; and Software-as-a-Service (SaaS), which includes software applications and other hosted services.

A cloud service differs from traditional hosting in three principal aspects. At first, it is provided on demand; and second, typically by the minute or the hour,it is elastic since the user can have as much or as little of a service as they want at any given time; third, the provider service is fully managed. Here is no unique definition for cloud computing. In these common definitions we present and also state our definition of cloud computing as follows: "Cloud computing is a new computing paradigm, whereby shared resources such as, infrastructure, hardware platform, and software applications are provided to users on-demand over the Internet (Cloud) as services." Computing paradigm shift of the last half century.

Specifically, the figure identifies six distinct phases. At Phase 1, people used terminals to connect to powerful mainframes shared by many users. Back then, the terminals were mostly little more than monitors and keyboards. At Phase 2, stand-alone personal computers (PCs) became powerful enough to satisfy users' daily work; One didn't have to share a mainframe with anyone else.

In Phase 3 ushered in computer networks that allowed multiple computers to connect to each other. PC and connect to other computers through local networks to share resources work on a one could. At Phase 4 saw the advent of local networks that could connect to other local networks to establish a more global network users could now connect to the Internet to utilize remote applications and resources.

At Phase 5 brought us the concept of an electronic grid to facilitate shared computing power and storage resources (distributed computing). Here People used PCs to access a grid of computers in a transparent manner. Now, At Phase 6, cloud computing lets us exploit all available resources on the Internet in a scalable and simple way. And Authors et al. We have listed 10 top obstacles and opportunities for growth of Cloud Computing. First three concern adoptions, the next five affect growth, and the last two are policy and business obstacles. At each obstacle is paired with an opportunity, which can overcome that obstacle Although benefits and opportunities that cloud computing has been bringing about are tremendous, ranging from product development to research projects, its challenges and problems that require huge amount of effort to be addressed by researchers. Since 2007 that gotten high attention until now the cloud computing , the trend of search volume index has been increasing sharply.The trend of search volume index for Cloud Computing versus Grid Computing using Google trend service during past years.

II. RELATED WORK

In Existing system we present a stochastic model, based on Stochastic Reward Nets (SRNs), that exhibits the above mentioned features allowing to capture the key concepts of an IaaS cloud system. In proposed model is scalable enough to represent systems composed of thousands of resources and it makes possible to represent both physical and virtual resources exploiting cloud specific concepts such as the infrastructure elasticity. And with respect to the existing literature, the generic and comprehensive view of a cloud system is presented innovative aspect on the present work. The Low level details, such as VM multiplexing they are easily integrated with cloud based actions such as federation, and to investigate different mixed strategies is allowing. In an exhaustive set of performance metrics are defined regarding both the system provider (e.g., utilization) and the final users (e.g., responsiveness). In proposed system interacting analytical sub-models to capture other important aspects including power consumption, resource assigning process and pool management, and virtual machine deployment of nowadays cloud centers. At finally, a performance model suitable for cloud computing centers with heterogeneous requests and resources using interacting stochastic models is proposed and evaluated.

III. SYSTEM PREMELIRIES

A. POOL MANAGEMENT SUB-MODEL

The Pool Management Sub-Model (PMSM) captures the details of pools management in the cloud center. Here we model PMSM as a three dimensional CTMC. System starts with state (i; j; k) which indicates that i, j and k PMs are in the hot, warm and cold pool respectively. And if next search in hot pool was successful (probability of Ph), system will remain in the starting state, state (i; j; k), otherwise PMSM will bring one of the warm PMs into the hot pool (requires some time).

B. POWER CONSUMPTION

We have calculated the performance metrics for a number of different pool configurations. For this can be seen 'stronger' configurations easily outperform the 'weaker' ones. While power consumption is also taken into account, and the results reveal that any given pool arrangement has a unique pool check rate that results in optimum balance between availability and throughput, and on one side, and, power consumption on the other side. At from the results it can be seen that for [15,10,10] arrangement the optimum pool check rate is 30/hour, and while for [18,12,11] the best rate is 20/hr. For other configurations, a pool check rate of 10/hour can be considered as a good choice.

C. RESOURCE ASSIGNING PROCESS

Resource allocation process is described in the resource allocation sub-model (RASM). Resource allocation process(RASM)is a two dimensional CTMC in which we take care of number of super-task in the queue as well as the current pool on which provisioning is taking place.

Since the number of potential users is high, but each user submits super-tasks one at a time with low probability, and the super-task arrival can be modeled as a Poisson process. The Super-tasks are lined up in the global finite queue to be processed in a first-in, first-out (FIFO) basis. For each state of Markov chain is labeled as (i; j), where i indicates the number of super-tasks in queue and j denotes the pool on which the leading super-task is under provisioning. State (0; 0) indicates that system is empty which means there is no request under provisioning or in the queue. In index j can be c, h or w that indicates current super-task is undergoing provisioning on cold, hot or worm pool respectively.

D. VIRTUAL MACHINE DEPLOYMENT

The virtual Machine Provisioning Sub-Model (VMPSM) captures the instantiation, the deployment and provisioning of VMs on a PM. The VMPSM also incorporate the actual servicing of each task (VM) on a PM. This note that each task within super-task is provisioned with an individual VM.Though RASM makes sure that all tasks within a super-task get provisioned at the same PM. In this model, we assume homogeneous PMs, VMs and tasks.

IV. CONCLUSION

In this paper we have extended our proposed interacting analytical model to capture important aspects including pool management, resource assigning process and power consumption, and virtual machine deployment of nowadays at cloud centers. Performance model can assist cloud providers to predict the expected task rejection probability, servicing delay, and steady state arrangement of server pools and power consumption. In this we carried out extensive numerical experiments to study the effects of various parameters such as arrival rate of super tasks, virtualization degree, task service time, and response time and normalized power consumption super-task size and pool check rate on the task rejection probability. For the activities of cloud center for given configurations has been characterized in order to facilitate the capacity planning, cloud economic analysis and tradeoffs by cloud service providers and SLA analysis. By using the proposed pool management model, and the most appropriate arrangement of server pools and the amount of required electricity power can be identified in advance for anticipated arrival process and super-task characteristics. In this paper we have also presented a performance model suitable for cloud computing centers with heterogeneous requests and resources using

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

interacting stochastic models. Additional specifically, and a user task may request different types of VMs; VMs can be varied in terms of CPU cores per virtual CPU. Unlike previous performance models that have been presented our final performance model can support heterogeneous PMs.

REFERENCES

[1] S. Alam, R. Barrett, M. Bast, M. R. Fahey, J. Kuehn, C. McCurdy, J. Rogers, P. Roth, R. Sankaran, and J. S. Vetter. Early evaluation of IBM BlueGene. 2008 SC International Conference for High Performance Computing Networking Storage and Analysis, pages 1–12, May 2008.

[2] Amazon Elastic Compute Cloud. User Guide. Amazon Web Service LLC or its affiliate, Aug. 2012.

[3] An amazon.com company. Amazon Elastic Compute Cloud, Amazon EC2. Website, Last Access Oct. 2012. http://aws.amazon.com/ec2.

[4] T. M. Apostol. Mathematical analysis. Addison Wesley, 2nd edition, 1974.

[5] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A view of cloud computing. Commun. ACM, 53:50–58, Apr. 2010.

[6] A. Baig. UniCloud Virtualization Benchmark Report. white paper, March 2010. http://www.oracle.com/us/technologies/linux/intel-univa-virtualization- 00164.pdf.

[7] J. Baker, C. Bond, J. Corbett, J. J. Furman, A. Khorlin, J. Larsonand, J. M. Leon, Y. Li, A. Lloyd, and V. Yushprakh. Megastore: Providing Scalable, Highly Available Storage for Interactive Services. In Conference on Innovative Data Systems Research (CIDR), pages 223–234, Jan. 2011.

[8] F. Bonomi and A. Kumar. Adaptive optimal load balancing in a nonhomogeneous multiserver system with a central job scheduler. IEEE Trans. on Comput., 39:1232– 1250, Oct. 1990.

[9] S. C. Borst. Optimal probabilistic allocation of customer types to servers. SIGMETRICS Perform. Eval. Rev., 23:116–125, May. 1995.

[10] O. J. Boxma, J.W. Cohen, and N. Huffel. Approximations of the mean waiting time in an M=G=s queueing system. Operations Research, 27:1115–1127, 1979.

[11] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. In IEEE International Conference on High Performance Computing and Communications, pages 5 –13, Sep. 2008.

[12] J. Cao, W. Zhang, and W. Tan. Dynamic control of data streaming and processing in a virtualized environment. IEEE Transactions on Automation Science and Engineering, 9(2):365–376, April 2012.

[13] A.L.E. Corral-Ruiz, F.A. Cruz-Perez, and G. Hernandez-Valdez. Teletraffic model for the performance evaluation of cellular networks with hyper-erlang distributed cell dwell time. In 71st IEEE Vehicular Technology Conference (VTC 2010-Spring), pages 1–6, May. 2010.

[14] G. P. Cosmetatos. Some practical considerations on multi-server queues with multiple poisson arrivals. Omega, 6(5):443–448, 1978.

[15] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good. The cost of doing

science on the cloud: The montage example. 2008 SC International Conference for High Performance Computing Networking Storage and Analysis, C(Nov.):1–12, 2008.