

A Meta-Analysis of Cloud Computing Infrastructures Involving Cost regulation and QoS Requirements

Gokul.M, M.Tech^{*}, J.Caroline El Fiorenza^{**}

^{*}Dept of CSE, SRM University, Ramapuram Campus, Chennai-89

^{**}Assistant Professor,Dept of CSE, SRM University, Ramapuram Campus, Chennai-89

Abstract- Cloud computing is a promising technology able to strongly modify the way computing and storage resources will be accessed in the near future. Cloud systems offer services at three different levels: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). In particular, IaaS clouds provide users with computational resources in the form of virtual machine (VM) instances deployed in the provider data center. PaaS and SaaS clouds offer services in terms of specific solution stacks and application software suites. Performance evaluation of Cloud Computing. An infrastructure is required to predict and quantify the cost-benefit of a strategy portfolio and the corresponding Quality of Service (QoS) experienced by users. In a market-oriented area, the cloud computing requires an accurate evaluation of these parameters is required to quantify the offered QoS and opportunely manage SLAs. There are several performance metrics have been defined, like availability, utilization, and responsiveness, allowing us to investigate the impact of different strategies on both provider and user point of views. A resiliency analysis is also provided to take into account load bursts. We can also include the analysis of autonomic techniques able to change on-the-fly the system configuration to react to a change on the working conditions. We can also extend the model to represent PaaS and SaaS cloud systems and to integrate the mechanisms needed to capture VM migration and data center consolidation aspects that cover a crucial role in energy saving policies.

Index Terms- cloud computing, resource allocation and load balancing

I. INTRODUCTION

Cloud computing is a computing paradigm in which different computing resources such as infrastructure, software applications and platform are made accessible over the internet to remote user as services. Infrastructure-as-a-Service (IaaS) clouds are becoming a rich and active branch of commercial services. Uses of IaaS clouds can provision “processing, storage, networks and other fundamental resources” on demand, and paying only for what is actually consumed. However, the increased adaption of clouds and perhaps even the pricing models depend on the ability of (perspective) cloud users to benchmark and compare commercial cloud services. The typical performance evaluation approaches in quality of service such as simulation or on-the-field measurements cannot be easily adapted. Simulation does not allow conducting comprehensive analyses of the system performance due to the great number of parameters that have to

be investigated. To implement particular resource management techniques such as VM multiplexing or VM live migration that is transparent to final user, has to be considered in the design of performance models in order to accurately understand the system behavior. Different clouds, belonging to the same or to different organizations, can join each other to achieve a common goal, represented by the optimization of resources utilization. It allows providing and releasing resources on demand thus providing elastic capabilities to the whole infrastructure. All the parameters do not conform to real time situations or exigencies and are thus are reflective in nature. Field experiments are mainly focused on the QoS, they are based on a black box approach that makes difficult or correlate obtained data to the internal resource management strategies implemented by the system provider. Resource management strategies implemented by the system provider.

Cloud Model And Resources:

Clouds are modeled in the server client mode by a datacenter component for service requests handling. These requests application elements are sandboxed with VMs, which needs to be allocated processing power on data center host components. By processing, it means that a set of operations related to VM life cycle: provisioning of a host for VM creation, destruction, and migration. A Datacenter is composed by set hosts, that are responsible for managing VMs. Host is a component that represents a physical computing node in a cloud: it is assigned a pre-configured processing capability (MIPS expression is million of instruction per second), memory and storage, along with scheduling policy for allocating processing cores to VM. The Host components implement interfaces that support modeling and simulation of both single-core and multi-core nodes.

Client Requests:

Clients are registered with the server for utilizing the resources and application. Needs to access the resources and applications are implemented and customary SLA's are put in place to handle the request handling mechanisms. The Request are considered as jobs that are to be completed. A job is usage of a app or a resource or both and sending back the processed requested data. The job arrival and the request process constitute three different scenarios. In the first one (Constant arrival process) the uniform arrival process be a homogeneous Poisson process. However, in large scale distributed systems with thousands of users, like in cloud systems, could exhibit similarity or long range dependence with respect to the arrival process and in order to take into account the dependencies of the job arrival rate on both the days of a week, the hours of a day, in the second

scenario is the arrival process periodically. This model is also chosen to be the model for the job arrival process as Markov Modulated Poisson Process (MMPP).

Cloud Federation And Monitoring:

This is the analytical algorithm, where parameters like weight; intermediate requests like waiting time, bandwidth calculation completion time are all executed. This is based on the inputs the optimal solution is arrived federation with other clouds is modeled allowing tokens in place pqueue to be moved, through transition tupload, in the upload queue represented by Psend place. In accordance with the assumptions made before, transition tupload is enabled only if the number of tokens in place pqueue is greater than Q and the number of tokens in place Psend is less than D. The federated cloud availability Moreover is order to take into account, concurrent enabled transition tupload and tdrop are managed by setting their weights. For these reasons, typical performance evaluation approaches such black box approach that makes difficult to correlate obtained data to the internal resource management strategies implemented by the system provider. The analytical techniques represent a good candidate, with the help of the limited solution cost of their associated models. In order to accurately represent a cloud system, an analytical model of a cloud has to be:

- **Scalable.** To deal with very large systems composed of hundreds or thousands of resources.
- **Flexible.** Allowing us to easily implement different strategies and policies and toas simulation or on-the-field measurements cannot be easily adopted. The Simulation does not allow us to conduct comprehensive analyses of the system performance due to the great number of parameters that have to be investigated. The On-the-field experiments are mainly focused on the offered QoS; they are based on a represent different working condition.

Our goal is to perform a comprehensive investigation of the long-term variability of performance for production cloud services. Our main contribution towards this end is threefold: We collect performance traces of the largest commercial clouds. The collected traces are analyzed, revealing for each service both summary statistics and the presence or absence of performance time patterns. We evaluate through trace-based simulation the impact of the variability observed in the studied traces on three large-scale applications that are executed today or may be executed in the cloud in the (near) future: scientific computing workloads on clouds, selling virtual goods through cloud payment services and updating the virtual world status of social games through cloud-based database services.

Proposed System:

In the Proposed system we present an analytical model, based on Heuristic algorithms and stochastic reward nets (SRNs), that is both scalable to model systems composed of thousands of resources and flexible to represent different policies and cloud-specific strategies. Data stored in multiple servers when client upload the data in main server. Then we calculate the server performance and user request when client retrieve that data from

cloud. If server performance is high, client communicates with another data center. The performance of the several servers is calculated and the user request is managed periodically. The performance of the system is shown graphically by which we can visualize the performance. Due to the splitting of the workload regarding the performance, the server speed is maximum and will retrieve the files correctly.

Modules:

- Admin
- Display File
- File request
- Performance Check
- Data Download

Modules Description:

Admin

This module upload the file and checks whether the user file request is authenticated otherwise user request will neglect. This module display the user request, online users using grid view control When user login.

Display file

This module used to display all uploaded files and file details to user using grid view control. Then user select file and send request to admin for our usage.

File request

This module used to check the file request to corresponding user. And accept the valid user request. Then that file Display to user and user download that file.

Performance Check

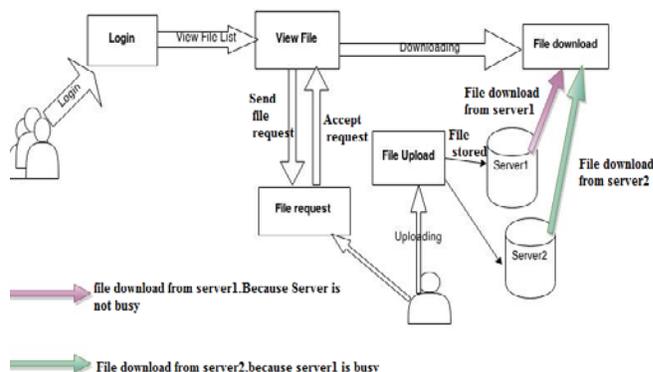
This module used to checking the performance of the system and processor. Computer performance is characterized by the amount of useful work accomplished by a computer system or computer network compared to the time and resources used.

Data Download

This module used to download the file from corresponding server. This module check user Authentication and whether user file request is accepted. Its allowed only valid user and should accept admin. valid user request. Then that file Display to user and user download that file.

Architecture Diagram:

Figure:1.1 Architecture Diagram:



II. CONCLUSION AND RESULT

The viewpoint of this project is to calculate the performance of the admin server. The Admin uploads the several files and stores in the several servers for the various access of the user operation. When the user enters into the system and tries to download the specified file from the server ,they need to get access from the admin and with permission from the admin , according to the higher performance of the server ,it downloads the specified file. SRNs allow us to define reward functions that can be associated to a particular state of the model in order to evaluate the performance level reached by the system during the sojourn in that state performance metrics able to characterize the system behavior from both point-of-views.the stochastic model to evaluate the performance of an IaaS cloud system. Several performance metrics have been calculated and defined, which are utilization, responsiveness and availability, allowing to investigate the impact of different strategies on both user point-of-views and provider's view. In cloud computing, an accurate evaluation of these parameters is required in order to quantify the offered QoS and opportunely manage SLAs and all the analysis are done by autonomic techniques able to change on-the fly the system configuration in order to react to change on the working conditions. This can be extended to models which represent PaaS and SaaS Cloud Systems and to integrate the mechanisms needed to capture VM migration and the data center consolidation aspects that cover a crucial role in energy saving policies. In future the model can be Implemented to include intra and reservoir clouds where other performance parameters like application takes sufficient memory and bandwidth issues are considered. This model may be replicated in web services domain where similar features are used and service agreements are made between the website client and web service provider so the concept can be extended to cover such models and issues as well in the future.

REFERENCES

- [1] R. Buyya et al., "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the Fifth Utility," *Future Generation Computer System*, vol. 25, pp. 599-616, June 2009.
- [2] X. Meng et al., "Efficient Resource Provisioning in Compute Clouds via VM Multiplexing," *Proc. Seventh Int'l Conf. Autonomic Computing (ICAC '10)*, pp. 11-20, 2010.
- [3] H. Liu et al., "Live Virtual Machine Migration via Asynchronous Replication and State Synchronization," *IEEE Trans. Parallel and Distributed Systems*, vol. 22, no. 12, pp. 1986-1999, Dec. 2011.
- [4] B. Rochwerger et al., "Reservoir—When One Cloud Is Not Enough," *Computer*, vol. 44, no. 3, pp. 44-51, Mar. 2011.
- [5] R. Buyya, R. Ranjan, and R. Calheiros, "Modeling and Simulation of Scalable Cloud Computing Environments and the Cloudsim Toolkit: Challenges and Opportunities," *Proc. Int'l Conf. High Performance Computing Simulation (HPCS '09)*, pp. 1-11, June 2009.
- [6] A. Iosup, N. Yigitbasi, and D. Epema, "On the Performance Variability of Production Cloud Services," *Proc. IEEE/ACM 11th Int'l Symp. Cluster, Cloud and Grid Computing (CCGrid)*, pp. 104-113, May 2011.
- [7] V. Stantchev, "Performance Evaluation of Cloud Computing Offerings," *Proc. Third Int'l Conf. Advanced Eng. Computing and Applications in Sciences (ADVCOMP '09)*, pp. 187-192, Oct. 2009.
- [8] S. Ostermann et al., "A Performance Analysis of EC2 Cloud Computing Services for Scientific Computing," *Proc. Int'l Conf. Cloud Computing, LNCS vol. 34*, pp. 115-131, Springer, Heidelberg, 2010.

AUTHORS

First Author – Gokul.M, M.Tech, Dept of CSE, SRM University, Ramapuram campus, Chennai-89, gokulknocks@gmail.com

Second Author – J.Caroline El Fiorenza, Assistant Professor, Dept of CSE, SRM University, Ramapuram campus, Chennai-89, caro.fiorenza@gmail.com