A Novel quality of service in Cloud computing using IAAS by stochastic reward nets

Polapaka Hima Bindu[1]& S.Tirupati Rao[2]

[1]M-Tech Dept. of CSE Geethanjali College of Engineering and Technology Hyderabad, Telangana,India.

[2]ProfessorDept. of CSE Geethanjali College of Engineering and Technology Hyderabad, Telangana, India.

Abstract—

Recent years there is a massive migration ofbusiness applications to cloud. One of the challenges posed are data centre management and providing a survey of QOS modeling approaches and using analytical model in which stochastic reward net model, that is scalable to model systems composed of several resources like utilization, availability, waiting time is taken into consideration. Cloud data center management is a key problem due to the numerous and heterogeneous strategies that can be applied, ranging from the VM placement to the federation with other clouds. Performance evaluation of Cloud Computing infrastructures is required to predict and quantify the cost-benefit of a strategy portfolio and the corresponding Quality of Service (QoS) experienced by users. Such analyses are not feasible by simulation or on-the-field experimentation, due to the great number of parameters that have to be investigated. In this paper, we present an analytical model, based on 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. Several performance metrics are defined and evaluated to analyze the behavior of a Cloud data center: utilization, availability, waiting time, and responsiveness.

Keywords—Cloud computing; quality of service; IaaS; stochastic reward nets(SRN);

1. INTRODUCTION

The cloud system provides services at three different levels IaaS, PaaS, SaaS. IaaS provides the resources in the form of virtual machine deployed in data centre. Quality of service is very important for provisioning service level agreements. Data centre performance and resource provisioning is essential. For performance analysis there are few system models has been used like queuing systems, queuing networks and layered queuing networks(LQN's) has to take into consideration. Several classes of models can be used to model QoS in cloud systems. here we briefly review queuing systems, queuing networks and layered queuing networks, however the other classes exist like stochastic reward nets and models evaluated via probabilistic model checking. The issue that a given method can perform better than other models. 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, storage, servers, 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. First, it is provided on demand, typically by the minute or the hour; second, it is elastic since the user can have as much or as little of a service as they want at any given time; and third, the service is fully managed by the provider. There is no unique definition for cloud computing. The 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." the computing paradigm shift of the last half century .

2. RELATED WORK

Existing system

In order to integrate business requirements and application level needs, in terms of Quality of Service (QoS), cloud service provisioning is regulated by Service Level Agreements (SLAs): contracts between clients and providers that express the price for a service, the QoS levels required during the service provisioning, and the penalties associated with the SLA violations. In such a context, performance evaluation plays a key role allowing system managers to evaluate the effects of different resource management strategies on the data center functioning and to predict the corresponding costs/benefits. Cloud systems differ from traditional distributed systems. First of all, they are characterized by a very large number of resources that can span different administrative domains. Moreover, the high level of resource abstraction allows to implement particular resource management techniques such as VM multiplexing or VM live migration that, even if transparent to final users,

have to be considered in the design of performance models in order to accurately understand the system behavior. Finally, different clouds, belonging to the same or to different organizations, can dynamically join each other to achieve a common goal, usually represented by the optimization of resources utilization. This mechanism, referred to as cloud federation, allows to provide and release resources on demand thus providing elastic capabilities to the whole infrastructure.

Problem definition

On-the-field experiments are mainly focused on the offered QoS, they are based on a black box approach that makes difficult to correlate obtained data to the internal resource management strategies implemented by the system provider. Simulation does not allow conducting comprehensive analyses of the system performance due to the great number of parameters that have to be investigated. PROPOSED SYSTEM: In this paper, we present a stochastic model, based on Stochastic Reward Nets (SRNs), that exhibits the above mentioned features allowing capturing the key concepts of an IaaS cloud system. The 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. With respect to the existing literature, the innovative aspect of the present work is that a generic and comprehensive view of a cloud system is presented. Low level details, such as VM multiplexing, are easily integrated with cloud based actions such as federation, allowing investigating different mixed strategies. An exhaustive set of performance metrics are defined regarding both

the system provider (e.g., utilization) and the final users (e.g., responsiveness).

Advantages of proposed system

To provide a fair comparison among different resource management strategies, also taking into account the system elasticity, a performance evaluation approach is described. Such an approach, based on the concept of system capacity, presents a holistic view of a cloud system and it allows system managers to study the better solution with respect to an established goal and to opportunely set the system parameters.

SYSTEM ARCHITECTURE:
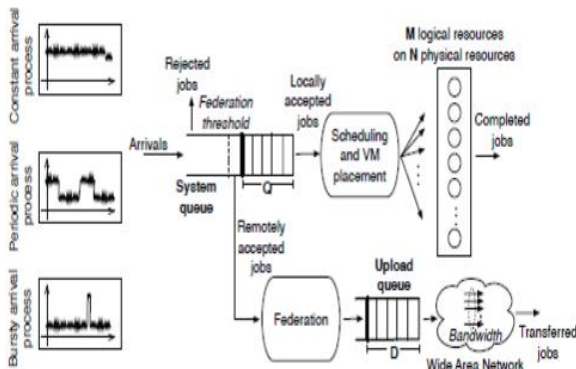
3. IMPLEMENTATION

System Queuing:

Job requests (in terms of VM instantiation requests) are en-queued in the system queue. Such a queue has a finite size Q, once its limit is reached further requests are rejected. The system queue is managed according to a FIFO scheduling policy.

Scheduling Module:

When a resource is available a job is accepted and the corresponding VM is instantiated. We assume that the instantiation time is negligible and that the service time (i.e., the time needed to execute a job) is exponentially distributed with mean $1/\mu$.

VM Placement:

According to the VM multiplexing technique the cloud system can provide a number M of logical resources greater than N. In this case, multiple VMs can be allocated in the same physical machine (PM), e.g., a core in a multicore architecture. Multiple VMs sharing the same PM can incur in a reduction of the performance mainly due to I/O interference between VMs.

Federation Module:

Cloud federation allows the system to use, in particular situations, the resources offered by other public cloud systems through a sharing and paying model. In this way, elastic capabilities can be exploited in order to respond to particular load conditions. Job requests can be redirected to other clouds by transferring the corresponding VM disk images through the network.

Arrival Process:

Finally, we respect to the arrival process we will investigate three different scenarios. In the first one (Constant arrival process) we assume the arrival process be a homogeneous Poisson process with rate λ. However, large scale distributed systems with thousands of users, such as cloud systems, could exhibit self-similarity/long-range dependence with respect to the arrival process. The last scenario (Bursty arrival process) takes into account the presence of a burst whit fixed and short duration and it will be used in order to investigate the system resiliency

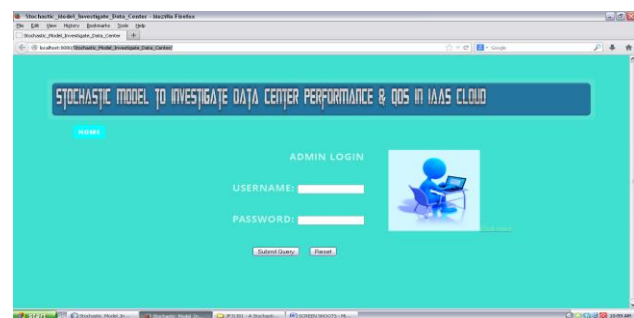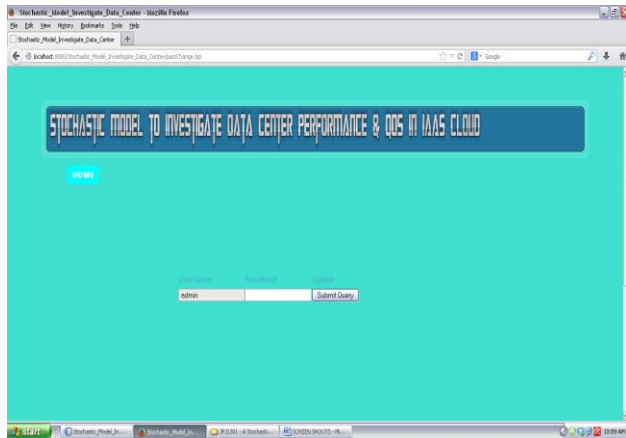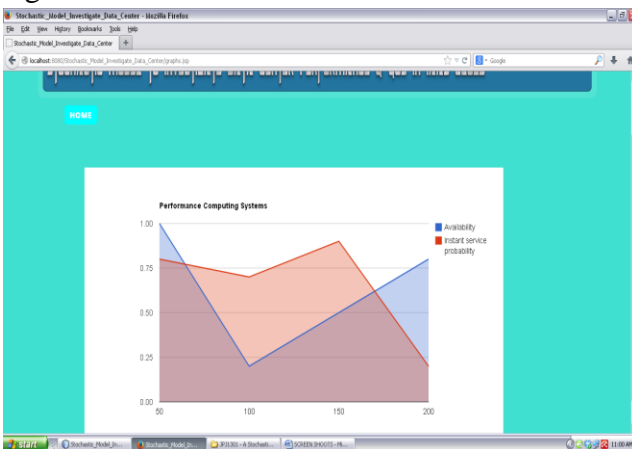### 4. EXPERIMENTAL RESULTS

Fig:-2 Admin Login

Fig:-2 Search

Fig:-2 Results

5. CONCLUSION:

In this paper, we have presented a stochastic model to evaluate the performance of an IaaS cloud system. Several performance metrics have been defined, such as availability, utilization, and responsiveness, allowing to investigate the impact of different strategies on both provider and user point-of-views. In a market-oriented area, such as the Cloud Computing, an accurate evaluation of these parameters is required in order to quantify the offered QoS and opportunely manage SLAs. Future works will include the analysis of autonomic techniques able to change on-thefly the system configuration in order to react to a change on the working conditions. We will also extend the model in order 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.

6. REFERENCES:

[1] R. Buyya et al., "Cloud computing and emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Gener. Comput. Syst., vol. 25, pp. 599–616, June 2009.

[2] X. Meng et al., "Efficient resource provisioning in compute clouds via vm multiplexing," in Proceedings of the 7th international conference on Autonomic computing, ser. ICAC '10. New York, NY, USA: ACM, 2010, pp. 11–20.

[3] H. Liu et al., "Live virtual machine migration via asynchronous replication and state synchronization," Parallel and Distributed Systems, IEEE Transactions on, 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, march 2011.

[5] R. Buyya, R. Ranjan, and R. Calheiros, "Modeling and simulation of scalable cloud computing environments and the cloudsim toolkit: Challenges and opportunities," in High Performance Computing Simulation, 2009. HPCS '09. International Conference on, june 2009, pp. 1 –11.

[6] A. Iosup, N. Yigitbasi, and D. Epema, "On the performance variability of production cloud services," in Cluster, Cloud and Grid Computing

(CCGrid), 2011 11th IEEE/ACM International Symposium on, may 2011, pp. 104 –113.

[7] V. Stantchev, "Performance evaluation of cloud computing offerings," in Advanced Engineering Computing and Applications in Sciences, 2009. ADVCOMP '09.Third International Conference on, oct. 2009, pp. 187 – 192.

[8] S. Ostermann et al., "A Performance Analysis of EC2 Cloud Computing Services for Scientific Computing," in Cloud Computing, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer Berlin Heidelberg, 2010, vol. 34, ch. 9, pp. 115–131.