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

Performance Evaluation of Cloud Computing Based on Stochastic Reward Nets Analytical Model

M.Vinay Kumar Reddy
M.Tech Student,
Department of CSE,
Vijaya Krishna Institute of Technology and Sciences,
Hyderabad.

A.Mounika
Assistant Professor,
Department of CSE,
Vijaya Krishna Institute of Technology and Sciences,
Hyderabad.

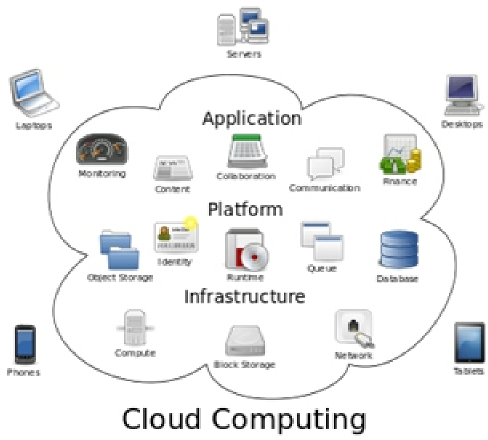
Abstract:

Cloud computing is computing in which large groups of remote servers are networked to allow centralized data storage and online access to computer services or resources. Cloud computing relies on sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over a network. One of the main concerns for enterprises that are considering cloud computing is performance. Achieving high-speed delivery of applications in the cloud is a multifaceted challenge that requires a holistic approach and an end-to-end view of the application request-response path. Performance issues include the geographical proximity of the application and data to the end user, network performance both within the cloud and in-and-out of the cloud and I/O access speed between the compute layer and the multiple tiers of data stores. In this paper we study and implement an analytical model based on stochastic reward net model, that is scalable to model systems composed of several resources like utilization, availability, waiting time is taken into consideration.

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

Introduction:

In a cloud computing system, there's a significant workload shift. Local computers no longer have to do all the heavy lifting when it comes to running applications. The network of computers that make up the cloud handles them instead. Hardware and software demands on the user's side decrease. The only thing the user's computer needs to be able to run is the cloud computing system's interface software, which can be as simple as a Web browser, and the cloud's network takes care of the rest.



The National Institute of Standards and Technology's definition of cloud computing identifies "five essential characteristics":On-demand self-service: A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.Broad network access: Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations). Resource pooling: The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. Rapid elasticity: Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.Measured service: Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Performance Evaluation Criteria:

There is a series of criteria for evaluation of all factors affecting performance of cloud computing some of which will be used in this paper. These criteria are under development.

Some of these criteria have been selected considering the importance and criteria in simulation. It should be mentioned that all of criteria listed in pervious sections cover the factors mentioned in the previous section but some of the factors will be important in special criteria:

• Average response time per unit time, this criterion will cover all factors completely

• Network capacity per second (Mbps)or unit time, the most important factor associated with this criterion is network bandwidth ,availability and scalability.

• The number of I / O commands per second(IOPS)or unit time

• Average waiting time per unit time

• Workload(requests) to be serviced per second(Mbps) or a unit of time

• Throughput (Req / Sec), this criterion will be recovered recovery, buffering capacity and processing power factors

• The average time of processing(exe / sec)

• Percentage of CPU utilization

• The number of requests executed per unit time

• The number of requests per unit time buffer

• The number of rejected requests per unit time

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.

DISADVANTAGES OF EXISTING SYSTEM:

• 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 to conduct 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 to capture the key concepts of an IaaS cloud system.

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

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 to investigate 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:

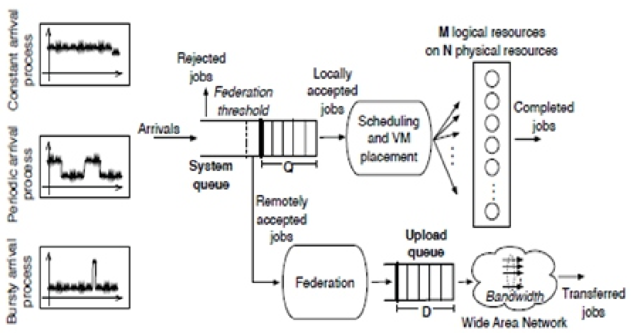


Fig. 1. An IaaS cloud system with federation.

MODULES:

There are five modules in this system:
1. System Queuing
2. Scheduling Module
3. VM Placement Module
4. Federation Module
5. Arrival Process

A. 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.

B. 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 $1/\mu$.

C. 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.

D. 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.

E. 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.

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

CONCLUSION:

A stochastic model was presented here to evaluate the performance of an IaaS cloud system. Several performance metrics have been defined, such as availability, utilization, and responsiveness, allowing investigating theimpact 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-the fly the system configuration in order to react to a change on the working conditions. This can 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.

REFERENCES:

[1] Dario Bruneo, Member, IEEE-" A Stochastic Model to Investigate Data Center Performance and QoS in IaaS Cloud Computing Systems"- IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS 2013.

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

[3] –131 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 Com-puting 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.

[9] Danilo Ardagna, Giuliano Casale, Michele Ciavotta, Juan F Pérez and Weikun Wang, "Quality of service in cloud computing: modeling techniques and applications,"JisaJournal.

[10] 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.