Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/13S109

Automated Performance Monitoring and Optimization Method for Cloud Data Centre

A. Poobalan[1,*], *N. Uma Maheswari* [2] *and K. Khaja Mohideen* [3]

[1] Department of Computer Science and Engineering, University College of Engineering, Dindigul, Tamil Nadu, India,
[2] Department of Computer Science and Engineering, PSNA College of Engineering and Technology, Dindigul , Tamilnadu, India
[3] Department of Master of Computer Applications, Aalim Muhamed Salegh College of Engineering, Chennai, India.

Received: 14 Jun. 2018, Revised: 2 Feb. 2019, Accepted: 27 Mar. 2019
Published online: 1 Aug. 2019

Abstract: Data centres are the integral part of the cloud computing which makes the enterprises invest less in their infrastructure. The multi-tenant nature of the cloud also adds value to this. Data is handled more efficiently with the advanced technologies in terms of infrastructure and network communication. With all these facilities, monitoring of the performance of the data centre is an essential task. This is also addressed by cloud computing through various methodologies. Differences in type of servers, user requests and capability of storage are the various factors that have their influence in handling of data and users. A model has been designed based on queuing theory. This model addresses the performance issue and also optimizes it. It follows a multi-tenant queue. The parameters that are considered and optimized are workload, response time, utilization of the resource, capability of serving resource at zero time, reservation of resources in advance and the number of users who are served. The tools provided by oracle such as automated workload reports and ADDT are used for performance measurement and optimization.

Keywords: Data centers, Cloud computing, Multi-tenant, AWR, ADDT and Virtual machine.

1 Introduction

Expensive storage resources in-house and maintenenance can be eliminated with the help of data centres. Data centre provides the entire infrastructure and facilities required to store and retrieve the data remotely. Though cloud service is suitable for various applications due to some issues related to performance, they are not yet a best fit for the high performance computing. The cloud is moving towards multitenant nature in which it serves multiple customers with the same resource. Cloud services are utilized in three flavors as public cloud, private cloud and hybrid cloud. Due to extensive resources and its nature to scale up in distributed environment, the possibility of getting exploited is high in the case of public cloud [1]. Storage service is the one of the main service that has been an important service provided by the cloud. Optimization of storage is a challenging process in the world of on-demand services. Storage support in multi-tenant environment is highly appreciable. The number of servers used for storage varies from hundreds to thousands based on the service provider [2]. The problem that is associated with such service is the queuing service. Studies have shown that there is no state-of-the-art technologies which meet the performance requirements of the multi-tenant environments. When the environment of the single tenant systems satisfies the requirements of a single client, multi-tenant systems should be capable of serving multiple tenants. Multi tenant here refers to the different related clients. For instance, an organization can have its branches in various locations. When such an organization is subscribed to a cloud service, employees of the organization located at various places could be able to access the service.

1.1 Performance evaluation

Performance evaluation is the method that has been followed in every domain. In case of the computers and its related services, the performance evaluation should conform to the service level agreements. For instance, when a user is accessing an service through internet the critical factor to be considered for performance evaluation

* Corresponding author e-mail: apoobal@gmail.com

could be the response time. Various other performance parameters that are widely accepted and followed are the number of resources utilized, power consumed, etc. when performances are measured for every system in nook and corner of the world , it is obvious that cloud with its wide range of services and applications certainly has many challenges in terms of performance [3].

1.2 Data centre management

Sustainability is the main concern of any design and management. The data centre is associated with the partners to produce a system which has utmost efficiency in terms of the energy and environmental conditions. The recent advancements in the network connections make it possible for the users to get services from the cloud remotely [4]. The services provided by the client are platform as a service, software as a service and infrastructure as a service. These services provided by the cloud are reliable and they also meets the quality of service. In recent days, the servers that are used in the data centres are also placed in various locations with the belief on the new data communication technologies.

The service station in the considered queuing model is repairable in the sense that in case of failures, the time required to repair it follows a exponential distribution where the rate $r > 0.5$. Poison process is followed in the arrival of customers. Identical and independent service time is demanded with a common distribution function [5]. The M / M / 1 / K / K queuing model has a single server queue. The arrival of customers follows a Poisson distribution. The input rate is equal to the output rate. The blocking probability is proposed with different approximation based on the one followed by the finite capacity of M/M/m/m + r queues. A model has been proposed to design a multi-tenant queue model which can be utilized in micro data centre. The remaining part of the paper is organized as follows. The second section explains the related work, the third section explains the proposed work and the remaining section explains the experimental results.

2 Related Study

Bandwidth and latency are the major issues that are considered in most of the state-of-the-art works. But with the advancing data centre technologies a more efficient system is required. in the previous studies, the bandwidth used by the cloud resources are only minimal [6] which is not considered. The major performance parameters that are considered in the traditional data centre are as follows.

1. Power Utilization Ratio (PUR)
2. Utilization of CPU
3. Utilization of storage space

These parameters are only used for measuring the performance of the single tenant environments which does not suit for the multi-tenant environment as the one which has been proposed. The queuing model that has been followed is M / M / m / m+r queuing model. The features of the model are, response time distribution, finite buffer size and the correlation between the maximum number of tasks; the minimum number of resources and the highest level of services are identified. Ioan et al. [16] developed a multi cloud system and analyzed the performance of it. The performance parameter that has been considered is the aggregate time to complete and discussed the efficient scheduling model to backup the data [17].

2.1 Problem Statement

– The absence of an efficient method to analyze the performance.
– There is no efficient monitoring mechanism.
– The monitoring mechanisms that are followed is static in nature and that does not meet the requirements of the change in business needs.
– High power consumption.
– Power consumption is increased with the inefficient allocation of virtualized applications which in turn would result in hot spots in server racks. M/G/m/m + r/n/d Queuing Systems are used to solve the above specified issue [7].

2.2 Drawbacks using Existing model

The approximation is accurate with comparatively small number of servers. When the co-efficient of variation reaches 1 or increase above 1 the approximate is inaccurate. Error with approximation occurs only when the traffic intensity is small and when both the co-efficient of variance and number of servers are large. Error also occurs only with the lateral constraint. When the distribution of service time is not known or when the number of server is high, the system becomes useless. When the intensity of the traffic varies in a wide range in small number of servers, the result obtained from M/G/m/m+r queuing model is accurate when $r = 0$ and when $r \neq 0$ the result is reasonable accurate. Here r refers to the buffer rate [2].

3 Proposed Work

In order to solve the problems specified in the previous section a monitoring and management tool designed by oracle is integrated with the system. The main focus of the work is to distribute the work in a data center to various locations through micro data centres. For instance, instead of an enterprise relying on a single huge data centre for its operation, micro data centers are installed in every branch office of the enterprise and they are coordinated to service the requirements raised by the client. This works similar to the distributed systems. Real-world cloud test bed is used for experimentation. The performance monitoring and reporting tool used here is the ADDT and AWT, these are Automated workload reports, which are provided by oracle. Though the utmost objective of the work is to provide a detailed analysis of the performance, other factors that are demonstrated are scalability, efficiency and granularity of the proposed approach. The major advantages of the proposed system are:

–No requirement for a dedicated authentication system
–Data accountability is taken care by the automated workload report.
–Performance of the storage is analyzed and metrics are displayed
–Usage of queuing model with data analytics
–Decentralized approach
–Platform independent
–Effective power utilization.

The ideal value of power calculation for computing task in the data center is 1.0 PUE. The factors that are to be considered in perspective of performance for designing an application that monitors the cloud consumption is

1. Integration of cloud product to another cloud product such that the product to which it is integrated is capable of leveraging its capabilities.
2. Integration of a cloud application to an appliance that already has been deployed in the cloud.
3. Integrating a cloud application to an independent application.

The source and the target domains are linked for improving the data centers performance and this also paves a way to compute the performance of the data centre in cloud environment. In addition to this process, various other processes that have been carried out to compute the performance are as follows. Source and the destination shares the same feature space.

–A common repository is provided to the managers. This enables them to keep track of the performance and utilization of resources
–During normal and peak scenarios, the impact produced by the infrastructure and configuration was measured.
–The classical open network which assumes both inter arrival times and service time as exponential [8] is used for constructing the specified model.

3.1 Proposed model

The objective is to design a model with a cloud application that possesses its own appliances and capable of being integrated with other cloud applications. The model follows SivadonChaisiri et. al [9], which provides computing resources as public resources that could be accessed by the consumers as services. A commonly-prevailing ideology of the environmentalists is that growth in computing is the major contributor of global warming. So, the designed data center is designed in such a way that it resides on a concrete slab and has a weather-tight rack of servers, storage and switching. The proposed architecture is a multitenant in nature and could be adapted to any kind of service such as IaaS, PaaS or SaaS. The proposed model is designed with number of small servers and the performance goals of such a model are energy efficiency, optimal storage, optimization of runtime process, improvement in terms of energy, management of micro-energy, power optimization, and a standard architecture. The proposed model enables the cloud application to evaluate the performance, analyzing the ability of reducing the power consumption per unit. It is proposed to use the muti-tenant queue architecture with multiple clients and multiple servers which is shown in Figure(1).

Figure(1) represents the Multi-tenant Queue Architecture model. The requests that arrive are organized in the queue. The requests in the queue wait for the dispatching unit to allocate the resources that are needed for the execution. Since the IaaS with DaaS cloud model [14] is adopted in the proposed system the dispatching discipline is followed. Virtualization, adjacency, portability of the workload, grouping and storage connectivity are the advancements that are to be achieved. Since the solution of the M/G/c/K model fits only to special cases. There is a requirement for providing a solution that is robust and efficient which also takes into account the optimal buffer size and blocking probabilities [10]. This approach serves better and suits for the proposed model.

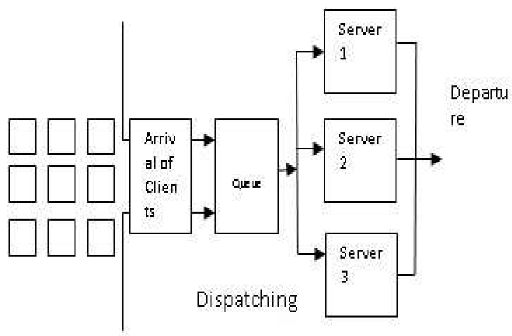

Fig. 1: Multi-tenant Queue Architecture model

The major objectives of the proposed model are:

–Integration of the application to a cloud product with appropriate connections between them.
–Calculating the probability distribution of the response time and number of tasks and also to calculate the probability of service whose waiting time is zero in the input buffer.
–Calculating the probability of a task being blocked for estimating the buffer size.

3.2 Resolved method

–Computation of computing performance in cloud computing environment.
–Improving the performance of the cloud servers by linking source and target domains.
–Intend that when a same feature space is shared by the source and target domains the cloud reveals its maximum performance.
–Optimizing cloud servers to attain better performance.

3.3 Proposed method

The cloud centre follows the queuing system of M/G/m/m+ r/n/d model which has a single task arrival and a task buffer with infinite capacity [11]. This enables the optimization of performance of cloud servers. The same model is followed in the system for queuing.

4 Working Model

Additional modeling and management has been incorporated into the designed model in order to facilitate optimization of system performance dynamically, capacity planning for long termed utilization of asset efficiently. Stochastic approach has been followed to describe the usage pattern of the services by the cloud users. Random request and response are followed. The time period between the requests is not the same but follows random timing. A hypothesis has been proposed assuming that the performance of the service received by the user reduces with the increase in number of users [8].

5 Performance Study

5.1 Experimental Setup

Oracle 10g database server is found to be suitable for the proposed system. Virtualization machine setup is done with VM ware and the cloud servers are introduced. The front end of the system might be any application developed with software such as PHP and Apache. The performance evaluation tool that is used is oracle performance & tuning tool called Automated Workload Report (AWR) [12].

5.2 Snapshot

Performance characteristics at a specified time are called as a snap shot. the rate of change of in the measured statistics can be computed with the data points of the snapshot . Each snapshot is identified with a unique identifier SNAP-ID. The system acitivity and workload was monitored by analyzing the AWR report to Automatic Database Diagnostic Monitor (ADDM) periodically. Root cause of the individual issues is identified and recommendations are provided for resolving the issues. When the value of STATISTICS-LEVEL parameter is set to TYPICAL or ALL, whenever an AWR snapshot is taken, performance analysis is done by ADDM and the findings and the recommendations are written to the database. The Optimized Performance Evaluation of Data Centre (OPEDC) is developed for increasing the performance of the cloud server groups. Queries from the oracle system are used for measuring system performance. Snapshots are used for classifying the results that are in the form of AWR. The classification includes valid users login, performance metric calculation sent by the users, request of optimized performance results by the users and users request to draw a flow graph. The following tables show the performance metrics and the specifications.

Table 2 provides the experimental query evaluation results under assumed 1 CPU (under oracle 10g performance tuninig). Table 3 depicts the experimental query evaluation results under oracle 10g - performance tuning experiment through under assumed1 CPU through an AWR and has recorded online.

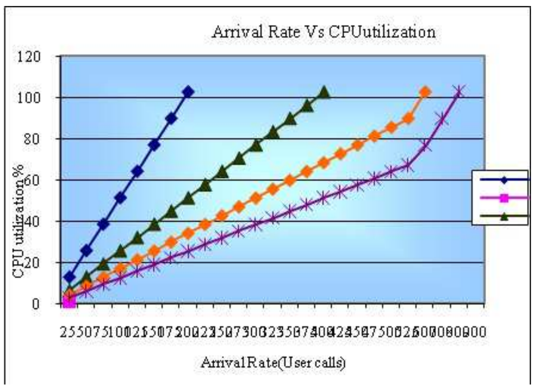

Fig. 2: Comparison of Arrival rate Vs CPU Utilization

5.3 Experimental Results and its Validation

The performance of the cloud servers is enhanced by the results obtained. The queuing model used is M/G/m/m+r

Table 1: Performance metrics formulae

Performance Metrics	Queuing Method
A = The number of arrivals	$U = (S * \lambda)/M$
T = Specific Time period	$\lambda = A/T$
U = CPU Utilization	R=S+Q (or) R=S/(1-U)
S = Service time	Q=R-S trx/s
M =No. of CPUs	
U = CPU Utilization	
R = Response time	
S = Service time	
Q = Wait time in the Queue	

Table 2: Summaries the list of performance metric used for better performance solution to be collected at the receiving end

User Call	Elapsed time	Time Period	Service Time	M (Number of CPU)	Lambda	Utilization	CPU Utilization	Response Time	Queue Waiting Time
3.39	12.29	37.4	0.78	1	0.00459723	0.0035858421	0.358584	0.78280702	0.002807
2.810	72.260	4335.6	0.73	1	0.00064812	0.0004731294	0.047313	0.73034555	0.000346
0.650	124.390	7463.4	0.15	1	0.00008709	0.0000130638	0.001306	0.150002	0.000002
2.740	55.420	3325.2	0.58	1	0.00082401	0.0004779261	0.047793	0.58027733	0.000277
2.680	60.550	3633	0.45	1	0.00073768	0.0003319571	0.033196	0.45014943	0.000149
2.990	59.530	3571.8	0.51	1	0.00083711	0.0004269276	0.042693	0.51021783	0.000218
2.980	60.070	3604.2	0.6	1	0.00082681	0.0004960879	0.049609	0.6002978	0.000298
2.750	60.570	3634.2	0.51	1	0.00075670	0.0003859171	0.038592	0.51019689	0.000197
6	0.15	0.5	1300	113.16	0.45	0.061826	0.388174		

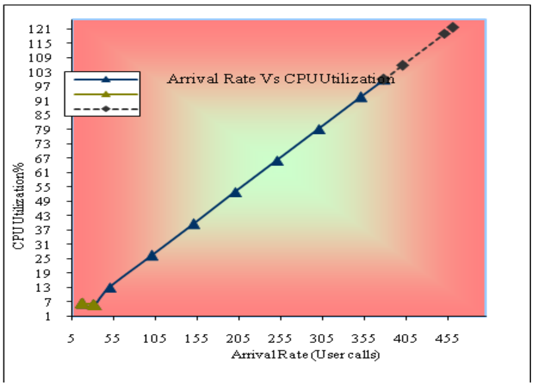

Fig. 3: Comparison of Arrival rate Vs CPU Utilization

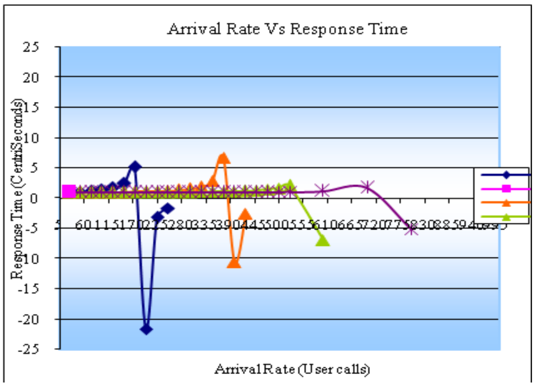

Fig. 4: Comparison of Arrival rate Vs Response Time

and it is powered with First Come First Serve (FCFS) . this combination enables the provision of better service to unlimited clients. Quality of service is accelerated and service level agreements are fulfilled to the clients using the adopted transformation-based model.
By referring to Table 2, the cloud service provider in a real-time cloud environment holds the key go with request(s) and to denial of request(s) from cloud consumers. The real-time data collected and evaluated clearly shows the key solutions of the cloud service providers in the real-time environment depend on the requests and denial of requests from the cloud users. The analysis also discusses the results obtained in terms of the response time, service time and the waiting time. It can be

Table 3: Experimental query evaluation CPU utilization and results with performance tuning under oracle 10g

Assumed 2 CPU

Arrival Rate per Sec	S per sec	*U*	*R*	Q
25	0.010233955	0.127924438	0.010404216	0.000170261
50	0.010233955	0.255848875	0.010950778	0.000716823
75	0.010233955	0.383773313	0.01200157	0.001767615
100	0.010233955	0.51169775	0.01386404	0.003630085
125	0.010233955	0.639622188	0.017319752	0.007085797
150	0.010233955	0.767546625	0.02490788	0.014673925
175	0.010233955	0.895471063	0.051652317	0.041418362
200	0.010233955	1.0233955	-0.216187407	-0.226421362
225	0.010233955	1.151319938	-0.031437091	-0.041671046
250	0.010233955	1.279244375	-0.016079338	-0.026313293

*Unstable System

Assumed 3 CPU

Arrival Rate per Sec	S per sec	*U*	*R*	Q
25	0.010233955	0.085282958	0.010240307	0.00000635184
50	0.010233955	0.170565917	0.010284991	0.00005103650
75	0.010233955	0.255848875	0.010408268	0.000174313
100	0.010233955	0.341131833	0.010657015	0.00042306
125	0.010233955	0.426414792	0.011094137	0.000860182
150	0.010233955	0.51169775	0.011817229	0.001583274
175	0.010233955	0.596980708	0.012999717	0.002765762
200	0.010233955	0.682263667	0.014996621	0.004762666
225	0.010233955	0.767546625	0.018681341	0.008447386
250	0.010233955	0.852829583	0.026951217	0.016717262
275	0.010233955	0.938112542	0.058677827	0.048443872
300	0.010233955	1.0233955	-0.142452146	-0.152686101
325	0.010233955	1.108678458	-0.028211988	-0.038445943
350	0.010233955	1.193961417	-0.014577362	-0.024811317

*Unstable System

noted that the waiting time reduces with the increase in the number of CPUs used .This is inferred from the table 2. The major factors that influence the performance are number of central processing units, response time and waiting time [13,15]. This is also inferred from the data collected and evaluated in real time.

5.4 Flow Graphs

The comparative analysis of the arrival rate and the CPU utilization is illustrated in the figures 2 and 3. Figure 5 depicts the relationship between the response time and the arrival rate. The comparative analysis of the response time and the arrival rate is depicted in the figure 6.

5.5 The Capacity planning calculations

Capacity Planning (S) = N*U = C1 / R1 + + Cn / Rn +/- K. where N is the number of CPUs , U is the average

Assumed 4 CPU

Arrival Rate per Sec	S per sec	*U*	*R*	Q
25	0.010233955	0.063962219	0.010234126	0.000000171
50	0.010233955	0.127924438	0.010236696	0.000002741
75	0.010233955	0.191886656	0.010247849	0.000013894
100	0.010233955	0.255848875	0.010277995	0.000044040
125	0.010233955	0.319811094	0.010342144	0.000108189
150	0.010233955	0.383773313	0.010460872	0.000226917
175	0.010233955	0.447735531	0.010662447	0.000428492
200	0.010233955	0.51169775	0.010987209	0.000753254
225	0.010233955	0.575659969	0.011496444	0.001262489
250	0.010233955	0.639622188	0.012291213	0.002057258
275	0.010233955	0.703584406	0.013555908	0.003321953
300	0.010233955	0.767546625	0.015673931	0.005439976
325	0.010233955	0.831508844	0.019606915	0.00937296
350	0.010233955	0.895471063	0.028665976	0.018432021
375	0.010233955	0.959433281	0.067038774	0.056804819
400	0.010233955	1.0233955	-0.105594372	-0.115828327
425	0.010233955	1.087357719	-0.025717075	-0.03595103

*Unstable System

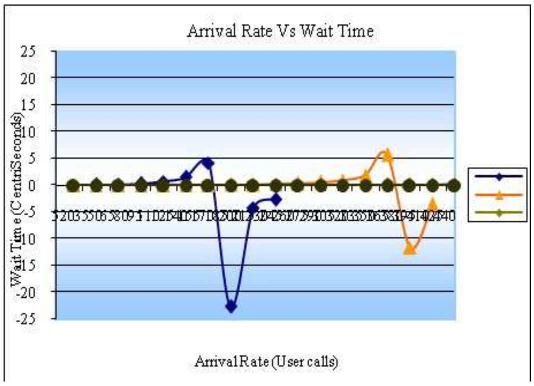

Fig. 5: Comparison of Arrival rate Vs Response Time

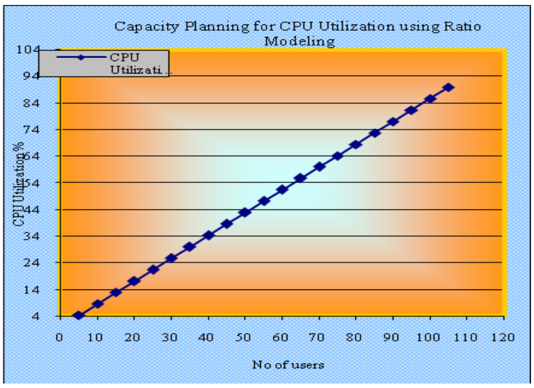

Fig. 6: Comparison of Arrival rate Vs Waiting Time

utilization of the CPU, C refers to the number of users and R is the ratio parameter which defines the ratio between the corresponding parameter to CPU and K is the error. Table 4 illustrates the way that is followed to calculate the CPU capacity with reference to the charts and the tables. The key to hold possessed by the cloud service provider is the requests and denial of requests from the consumers and it is purely based on the considered metrics and model. The major factors that influence the performance are number of central processing units, response time and waiting time. This is observed with the real-time data collected and analyzed.

6 Conclusion and Future Work

The performance measurements of the cloud servers are computed and studied in this work. This performance is the optimized one. The probability distribution of the time for request and response is accurately estimated. The mean number of tasks in the system and the probability

Assumed 6 CPU

Arrival Rate per Sec	S per sec	*U*	*R*	Q
25	0.010233955	0.042641479	0.010233955	0.000000000
50	0.010233955	0.085282958	0.010233959	0.000000004
75	0.010233955	0.127924438	0.010234	0.000000045
100	0.010233955	0.170565917	0.010234207	0.000000252
125	0.010233955	0.213207396	0.010234916	0.000000961
150	0.010233955	0.255848875	0.010236826	0.000002871
175	0.010233955	0.298490354	0.010241198	0.000007243
200	0.010233955	0.341131833	0.010250108	0.000016153
225	0.010233955	0.383773313	0.010266756	0.000032801
250	0.010233955	0.426414792	0.01029585	0.000061895
275	0.010233955	0.469056271	0.01034412	0.000110165
300	0.010233955	0.51169775	0.010421019	0.000187064
325	0.010233955	0.554339229	0.010539789	0.000305834
350	0.010233955	0.596980708	0.010719157	0.000485202
375	0.010233955	0.639622188	0.010986255	0.0007523
400	0.010233955	0.682263667	0.011381921	0.001147966
425	0.010233955	0.724905146	0.01197103	0.001737075
450	0.010233955	0.767546625	0.012864316	0.002630361
475	0.010233955	0.810188104	0.014269782	0.004035827
500	0.010233955	0.852829583	0.016633694	0.006399739
525	0.010233955	0.895471063	0.02112689	0.010892935
600	0.010233955	1.0233955	-0.0687563	-0.078990255
700	0.010233955	1.193961417	-0.005394938	-0.015628893

*Unstable System

that a task will be blocked and probability of a task getting an immediate service are the key performance indicators. The relationship between these parameters along with the input buffer size and the number of servers is studied. We thick that dividing the tasks into sub tasks in the multidirectional networking may be applied further. Virtualized data centres are used for updating the performance measurement. In near future data center management system with self-healing and self-monitoring capabilities will be available if the micro data centers are bridged properly. The performance of the data center will be influenced by the self-renewing energy sources on which the models of the data centers are relied upon. For a designated data center with cloud servers, the accuracy of the performance is measured and optimized in this work.

Acknowledgement

We are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] ImanSadooghi, Sandeep Palur, Ajay Anthony, IshaKapur, KarthikBelagodu, Pankaj Purandare, KiranRamamurty, Ke Wang, IoanRaicuAchieving Efficient Distributed Scheduling with Message Queues in the Cloud for Many-Task Computing and High-Performance Computing. IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (2014).

[2] HamzehKhazaei, JelenaMisic andVojislav B. Misic, Performance Analysis of Cloud Computing Centres Using M/G/m/m + r Queuing Systems, IEEE Trans. parallel and distributed sys. **Vol. 23, No. 5**, (2012).

Assumed 8 CPU

Arrival Rate per Sec	S per sec	*U*	*R*	Q
25	0.010233955	0.031981109	0.010233955	0.000000000
50	0.010233955	0.063962219	0.010233955	0.000000000
75	0.010233955	0.095943328	0.010233955	0.000000001
100	0.010233955	0.127924438	0.010233956	0.000000001
125	0.010233955	0.159905547	0.010233959	0.000000004
150	0.010233955	0.191886656	0.010233974	0.000000019
175	0.010233955	0.223867766	0.01023402	0.000000065
200	0.010233955	0.255848875	0.010234143	0.000000188
225	0.010233955	0.287829984	0.010234437	0.000000482
250	0.010233955	0.319811094	0.010235075	0.000001120
275	0.010233955	0.351792203	0.010236356	0.000002401
300	0.010233955	0.383773313	0.010238773	0.000004818
325	0.010233955	0.415754422	0.010243099	0.000009144
350	0.010233955	0.447735531	0.01025051	0.000016555
375	0.010233955	0.479716641	0.010262738	0.000028783
400	0.010233955	0.51169775	0.010282283	0.000048328
425	0.010233955	0.543678859	0.01031268	0.000078725
450	0.010233955	0.575659969	0.010358878	0.000124923
475	0.010233955	0.607641078	0.010427761	0.000193806
500	0.010233955	0.639622188	0.010528921	0.000294966
525	0.010233955	0.671603297	0.010675835	0.00044188
600	0.010233955	0.767546625	0.011635559	0.001401604
700	0.010233955	0.895471063	0.017447412	0.007213457
800	0.010233955	1.0233955	-0.050356948	-0.060590903
900	0.010233955	1.151319938	-0.004903141	-0.015137096

*Unstable System

[3] AlexandruIosup,SimonOstermann,M.NezihYigitbasi, RaduProdan, Thomas Fahringer, Member, and Dick H.J. Epema Performance Analysis of Cloud ComputingServices for Many-Tasks Scientific Computing IEEE Trans. parallel and distributed sys **Vol.22 No. 6** (2011).

[4] Cong Wang,Qian Wang, Kui Ren, Ning Cao, and WenjingLou,Toward Secure and DependableStorage Services in Cloud Computing IEEE Trans. Services computing **Vol. 5, No. 2**, (2012).

[5] Karim ABBAS and Djamil AISSANI, Approximation in an M.G/1 queuing system with breakdowns and repairs, Laboratory of Modelization and Optimization of Systems University of Bejaia, 06000(Algeria)..

[6] Iman Sadooghi, Jess Hernndez Martin, Tonglin Li, Kevin Brandstatter, KetanMaheshwari, Tiago Pais Pitta de LacerdaRuivo, Gabriele Garzoglio, Steven Timm,Yong Zhao, IoanRaicu, Understanding the Performance and Potential of Cloud Computing for Scientific Applications, IEEE Trans. Cloud computing,(2014).

[7] P. Hokstad, Approximations for the M/G/m Queues, Operations Research, **vol. 26, pp. 510-523,** (1978).

[8] B. Furht, Cloud Computing Fundamentals, Handbook of Cloud Computing, **pp. 3-19, Springer,** (2010).

[9] SivadonChaisiri, Bu-Sung Lee and DusitNiyato, Optimization of Resource Provisioning Cost in Cloud ComputingIEEE Trans. Services computing **Vol. 5, No. 2,** (2012).

[10] J.M. Smith, M/G/c/K Blocking Probability Models and System Performance, Performance Evaluation, **vol. 52, pp. 237-267,** (2003).

[11] SinungSuakanto, Suhono H Supangkat, Suhardi andRoberdSaragih, Performance Measurement of Cloud Computing Services International journal on Cloud computing:services and architecture,**Vol.2, No.2**,(2012).

[12] Kishore Daggubati, Deploying Oracle Database 11g R2Enterprise Edition on Hitachi Converged Platform for

Oracle Database - Reference Architecture Guide by Hitachi Data Systems, (2012).
[13] An Oracle White Paper Oracle Optimized Data Center Reference Architecture Oracle Enterprise Transformation Solution Series ,(2012).
[14] Paul Marshall, Kate Keahey and TiFreeman, Improving Utilization of Infrastructure Clouds IEEE/ACM Cloud Computing,(2011).
[15] Zhiguo Wan, June Liu, and Robert H. DengHASBE: A Hierarchical Attribute-Based Solution for Flexible and Scalable Access Control in Cloud Computing ,IEEE Tran. Information Forensics and Security, **Vol. 7, No. 2,** (2012).
[16] Ioan Petri, Omer F. Rana, Tom Beach, Yacine Rezgui,Performance analysis of multi-institutional data sharing in the Clouds4Coordination system,Computers and Electrical Engineering,**Volume 58,Pages 227-240,**(2017).
[17] Rao kotha et.al., Efficient queuing and scheduling of backups in a multi-tenant cloud computing environment united states patent application publication.

A. Poobalan received his B.E - 2007 -First Class with Distinction from RVS College of Engineering and Technology, Dindigul affiliated to Anna University, Chennai and M.E - 2010 -First Class with Distinction-RVS College of Engineering and Technology, Dindigul affiliated to Anna University, Chennai. Pursuing his Ph.D (Part time) at Anna University, Chennai. Currently working as a Assistant Professor in the Department of Computer Science and Engineering at University College of Engineering, Dindigul, Tamil Nadu India.. His area of interest includes Cloud computing, Network security and Networks and Distributed Computing.

N. Uma Maheswari received her M.E in Computer Science and Engineering from the Madras University, Chennai, India in 2002 and Ph.D. in Information and Communication engineering in 2011 at Anna University, Chennai. Currently, she is working as Professor in the Department of Computer Science and Engineering at the P.S.N.A. College of Engineering and Technology, Dindigul, India. Her current research interests include Image processing, Artificial Intelligence, Speech Processing. She has published 25 papers in International journals, 15 papers in national journals and 20 papers international conferences and 30 papers in national conferences. She is a recognized Ph.D. supervisor under Anna University Chennai in the area of Image processing; Cloud computing, Network security and Networks.

K. Khaja Mohideen received his B.Sc (Computer Science) and M.C.A. (Master of Computer Applications) from Jamala Md. College affiliated to Bharathidasan University, Tiruchirapalli, India in 1991 and 1994 respectively. Completed his M.Phil in Computer Science from Periyar University Salem, India. Completed his M.E. Computer Science and Engineering from University College of Engineering, Tiruchirapalli, India. Currently working as Head and Assistant Professor in the Department of Master of Computer Applications at Aalim Muhamed Salegh College of Engineering, Chennai, India.. His area of interest includes Web Databases, Cloud Computing, and Distributed Computing.