

Complications in Infrastructure as a Service Layer of Cloud and its Solution

A.Banushri¹, R.A.Karthika

Abstract: Cloud computing shows a vibrant role in existing scenario and the enactment of infrastructure as a service is perilous because of its discrepancy in the area. The cloud users have increased hastily, and the accessibility of resources for the users are less. Infrastructure as a service (IaaS) mentions the particulars of infrastructure like physical computing resources such as stowage, compute, and networking services. IaaS- cloud providers underwrite these resources based on their necessity from their vast content of resources presents any where all over the universe. Observing of these resources continually is precarious. For monitoring the availability of resources and notifying to the users about the resources is one of the challenges taken in IaaS layer is Service Level Agreement (SLA) and provided with a solution. The foremost objective of the scheduling algorithms in a cloud environment is to exploit the resources proficiently while balancing the load between resources, to get the least possible execution time. Hence, rank based task scheduling algorithm is proposed to utilize the resources efficiently and perform high performance. A simulation result gives the Quality of Service (QoS) Parameters such as length (size), CPU, throughput, and bandwidth.

Keywords: SLA, QoS, CPU, IaaS, Resources, rank-based task scheduling algorithm.

I. INTRODUCTION

Infrastructure as a service (IaaS) discusses the particulars of infrastructure like physical computing resources, for example, compute, storage and networking services servers. This IaaS layer is used to assimilate the hardware which could make cloud computing achievable. Based on the services, the cloud is offering as, Platform as a service(PaaS), Software as a service (SaaS),Infrastructure as a service (IaaS),and Information, Management, Storage, Security, Process, Database, Testing, Application as a service also applicable. The aim of PaaS is to provide the consumer, a virtual platform via the web. The Consumer can utilize own applications and services, deprived of download of software. By using those services the user could get a worthy identify management of strong privilege [1]. SaaS is at the uppermost level of cloud services. In those services the suppliers are responsible for all the features of security. If the Transparency is restricted, then the consumer resources will be gone in this service [2]. IaaS is contributing a virtual computer park which include router, Processors, firewalls etc. This would in cline to eliminate the cost of purchasing material. Essentially, choosing

the service, involves a great eminent network connection and good IT management in consumer organizer [3]. In IaaS services, users could access and make use of the virtualized resources directly. The real-time processing consuming an IaaS service is possible [4]. IaaS service is the foundation layer of other delivery model and an absence of security in this layer which will also affect some deliverable model [5]. In IaaS there are certain security threats which could disturb the infrastructure like external attack, middle attack, flooding attack, data leakage attack, etc. [6]. This service has security issues and solution in both private and public clouds. Data stored in one of the attacks in IaaS which could be examined closely. These types of problems could be resolved by using modern rights management services [7]. Some issues that inevitably be determined before expanding the services. Suitable observing of the resource must tend to be in an operative way to accomplish appropriate quality of services and great performances from the providers

II. RELATED WORK

There are various studies conducted in challenges faced by IaaS layer. In [8], the server surpasses a maximum level it results in decrease in Quality of Service (QoS) and response time. To circumvent overloading trouble, a decentralized two layered load management method is used. The keys used to impede this difficult are load balancing and the system overload protection. For the persistent update of weights in the algorithm, the round robin algorithm is scalable and difficult. The procedure of trusted third party based trusted identifier is used in [9]. The hindrances of using big data were to have reliance of multi-layer outsourced big data scheme in cloud environment. Two policies of enforcement method providing reliable data by encompassing the chain of trust to map-reduce application. Similarly, security disputes in Software-as-a-Service (SaaS) layer are classified as two as attack on development tool and attack on the management tool [10]. Security attacks are on the development tool, on the Application Programming Interface (API) and attack on the publisher. The Security on the management tool is data protection and attack on the interfaces. The security fears are elucidated by mapping to diverse cloud services in order to accomplish veracity and rationality in the cloud system. Dynamic security describes the threat level and category of security needed for every service at diverse cloud layers for a cloud consumer and cloud provider is evaluated. A Cloud Interoperability Broker (CIB) in the IaaS layer was generated to act as intermediate to block the interoperability gap amongst the applications used in IaaS layer. The CIB is instigated for diverse data transfer process [11]. The data is separated into two types as master data and the transaction data. The mapping and the migrating data from one provider to other provider was verified to overcome the vendor lock-in difficult in the IaaS layer. In this layer, a module could be handled using execution time

Revised Manuscript Received on July 22, 2019.

A.Banushri, Department of Computer Science & Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai. banushrics.sc@velsuniv.ac.in

Dr.R.A.Karthika, Department of Computer Science & Engineering, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Chennai.

and constraints from QoS [12]. This module have got online and offline stages. In this proposed system, the number of resources used by applications is diminished, and admission control mechanism, to distribute resources is necessary for applications.

A certification-based adaptation method for cloud services which can provide a stable non-functional behavior is developed in [13]. In this, multi-tenant, general-purpose adaptation practice for the cloud was proposed. This Adaptation practice could play an important role in providing a different application. The performance and quality of this adaptation was enhanced in an extensive range of settings. Virtualization deals with scalability, rationalizes performance, progresses resource management, and stores data proficiently [14]. Problems in the virtual environment poses substantial threats to sensitive information. Difficulties and presenting of cloud-based applications comprises Service Level Agreement (SLA) protection, compliance concerns, legal issues, governance and privacy. Those disputes must be deliberated in the improvement of crucial security measures to safeguard the privacy of user data and the continuousness of cloud-based service operations. Multi-tenancy and isolation is a chief measurement in the cloud security problem that requires a vertical solution from the IaaS layer down to physical infrastructure [15].

III. RESEARCH METHODOLOGY

A. SLA

Service Level Agreement (SLA) is used to identify about the vibrant security and privacy management strategies for the cloud environment. It screens the facilities such as eminence, enactment, precedence, and responsibilities from a service point of view. It contributes a clear sight to Know about the service management necessities in case of cloud service miscarriage [16]. SLA is well-defined as an agreement signed among the Cloud provider and a customer relating functional and non-functional features of service comprising QoS necessities, forfeits in case of abuses, and a set of metrics, which could be used to extent the provisioning of the necessities. To guarantee an established SLA in Clouds, the Cloud provider must be proficient of monitoring its infrastructure (host) resource metrics to impose the agreed service level intents [17]. In a Cloud environment, amenities and applications are accomplished on the physical and virtual resources. Though, the quality of service goal is to regulate the performance of the application is termed as a high-level SLA parameter, for example obtainability, throughput, etc. However, the applications are running on physical or virtual resources, which could be categorized by low-level metrics such as CPU, memory, downtime, uptime etc.

B. Proposed system

In this research, the SLA concentrates mainly on the high-performance parameters such as length, CPU, and bandwidth. The performance analysis is done in the Microsoft Azure and Dynatrace monitoring tool. This tool provide endpoint monitoring and analysis based on the SLA metrics like CPU consumption and response time are calculated to provide more resource availability, also, based on the availability of the resources, users can select high availability resources using Ranking. The foremost objective of the scheduling algorithms in cloud environment is to exploit the resources proficiently while balancing the load among the resources,

in order to get the least possible execution time. Hence, rank based scheduling algorithm is proposed to utilize the resources efficiently and perform high performance.

Resource metrics	SLA parameter	Mapping rule
Uptime, downtime	Availability (A)	$A = \left(1 - \frac{\text{downtime}}{\text{uptime} + \text{downtime}} \right)$
Packet size, bandwidth in, bandwidth out	Response time (R_{total})	$R_{total} = R_{in}(s) + R_{out}(s)$

Table 3.1 Performance metrics

C. Rank-based task scheduling algorithm

In proposed RBTSA algorithm, firstly the rank values are calculated for each resource based on two of the important parameters - the CPU speed and memory requirements of Virtual Machine (VM) resource. Then, a task is selected from all the tasks of workflow whose execution time is greater than or equal to the average execution time of all the tasks of the workflow. The selected task is the first task to be scheduled and is scheduled on the resource which results in the minimum completion time. After scheduling [18, 19], the first task of the workflow the scheduling of rest of the tasks of workflow is done. In every cycle, the task whose completion time is maximum among all the tasks of scientific workflow is selected and is scheduled to the highest rank resource available at that time.

In the proposed RBTSA algorithm for task scheduling of scientific workflows, three scheduling parameters were considered for optimization- Resource utilization, the Execution time of each task. The proposed RBTSA algorithm is shown in above Figure 1.

Scheduling targets to balance the load along with proficient execution of requests leading to value-added QoS. And also, it reduces resource famishment and guarantees fair resource allocation. It prominences on the challenging of decision making about the outstanding requests and which request has to be assigned resources next. Resource allocation is used to consign the resources optimally. Cloud computing also needs a related mechanism to guarantee ideal resource allocation and least possible wait time for the users. Balancing load between accessible VMs lessens the chances of deadlocks and progresses system throughput.

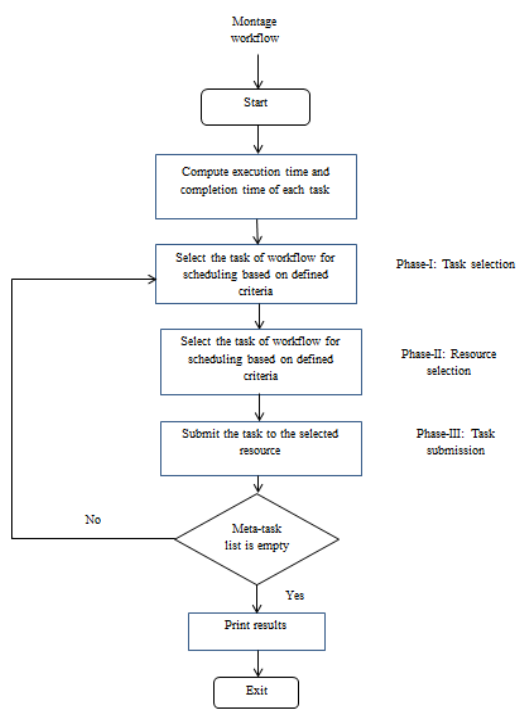


Figure 1. Rank based task scheduling algorithm

The pseudo code for Rank based task scheduling algorithm are as follows:

```

Input: Tasks of scientific workflow
Output: Scheduled list of tasks
Set size ← number of tasks in scientific workflow
Set vmsize ← number of vm resources
Set index ← 0
Set task ← null
Set vm resource ← null
for i ← 0 to size-1
    task ← task whose
    execution_time ≥ avg_execution_time
end for
for j ← 0 to vmsize-1
    vm_resource ← resource which gives minimum
    completion time
    index ← j
end for
set task.vmid ← index
Eliminate task from task list of workflow
Set vm_resource_status ← BUSY
for i ← 0 to size-1
    for k ← 0 to size-1
        (i) if task is already scheduled
            continue
        end if
        (ii) else
            task ← task which has maximum completion
            time
        end if
    end for
end for
for j ← 0 to vmsize-1
    (i) if vm_resource_status ← BUSY
        Continue
    (ii) else
        vm_resource ← highest ranked resource
    (iii) index ← j
end for
i) Set task.vmid ← index
(ii) Eliminate task from task list of workflow
(iii) Set vm_resource_status ← BUSY
end for
Print results
Exit
    
```

```

(ii) else
    vm_resource ← highest ranked resource
(iii) index ← j
end for
i) Set task.vmid ← index
(ii) Eliminate task from task list of workflow
(iii) Set vm_resource_status ← BUSY
end for
    
```

Print results

Exit

D. Resource scheduling factors

The elements that affect resource scheduling are as follows,

CPU Speed (C): It is the measure of a computer's processor speed. It trials the number of operations that might be done per second.

Bandwidth (B): It is the data transfer rate, the amount of data that could able to transfer from one host to another.

Length (L): It is the length of the task. It measures in bytes [20].

IV. FINDINGS AND RESULTS

This section explained the results and performance of the proposed RBSTA for the task scheduling process. The users who had necessity of cloud resources were submitting the tasks as cloudlets. Every cloudlet take explicit parameters such as CPU speed, bandwidth, the firmness of cloudlet, and length termed as Quality of Service (QoS) parameters. In each and every virtual machine (VM), the QoS parameter values are included, and both the cloudlet and VM are sent to data center broker for binding. The simulation results mainly concentrate on task finishing time and resource utilization rate. The performance analysis is done in a Dynatrace monitoring tool to validate the QoS parameters, and the results were shown in this section.

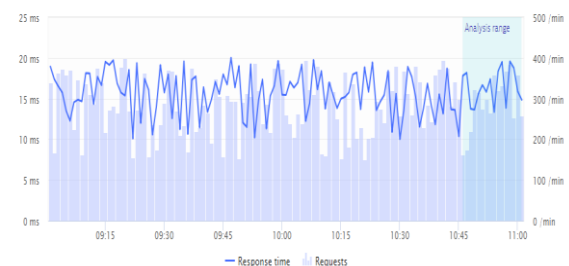


Figure 2. Response time for the allocated requests

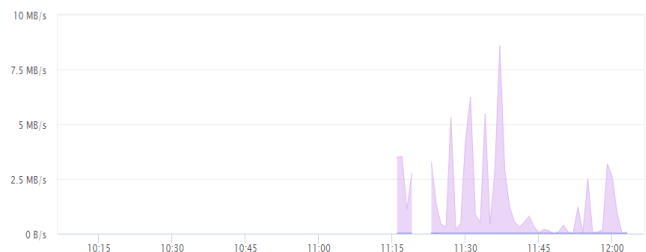


Figure 3. Maximum throughput

V. CONCLUSION

Resource scheduling plays a significant role in resource management. Effective utilization of resources will reduce cost and time. The proposed algorithm improves the resource utility time and minimizes the time needed for allocation. RBTSA scheduling algorithm reduces the makespan of scientific workflow, have proper task distribution on the virtual machines, lessens the execution time of individual workflow jobs and increases the utilization of virtual machine of the IaaS cloud computing environment.

REFERENCES

1. Belbergui, C., Elkamoun, N., & Hilal, R. (2017, October). Cloud computing: Overview and risk identification based on classification by type. In Cloud Computing Technologies and Applications (Cloud Tech), 2017 3rd International Conference of(pp. 1-8). IEEE.
2. Ma, D., & Kauffman, R. J. (2014). Competition between software-as-a-service vendors. IEEE Transactions on Engineering Management, 61(4), 717-729.
3. Calero, J. M. A., & Aguado, J. G. (2015). MonPaaS: an adaptive monitoring platform as a service for cloud computing infrastructures and services. IEEE Transactions on Services Computing, 8(1), 65-78.
4. Luo, F., Zhao, J., Dong, Z. Y., Chen, Y., Xu, Y., Zhang, X., & Wong, K. P. (2016). Cloud-based information infrastructure for next-generation power grid: Conception, architecture, and applications. IEEE Transactions on Smart Grid, 7(4), 1896-1912.
5. Yin, X., Chen, X., Chen, L., Shao, G., Li, H., & Tao, S. (2018). Research of Security as a Service for VMs in IaaS Platform (May 2018). IEEE Access.
6. Joshi, B. K., Shrivastava, M. K., & Joshi, B. (2016). Security threats and their mitigation in infrastructure as a service. Perspectives in Science, 8, 462-464.
7. Jaiswal, P. R., & Rohankar, A. W. (2014). Infrastructure as a service: security issues in cloud computing. International Journal of Computer Science and Mobile Computing, 3(3), 707-711.
8. Iranpour, E., & Sharifian, S. (2018). A distributed load balancing and admission control algorithm based on Fuzzy type-2 and Game theory for large-scale SaaS cloud architectures. Future Generation Computer Systems, 86, 81-98.
9. Zhan, J., Fan, X., Cai, L., Gao, Y., & Zhuang, J. (2018). TPTVer: A trusted third party based trusted verifier for multi-layered outsourced big data system in cloud environment. China Communications, 15(2), 122-137.
10. Hussain, S. A., Fatima, M., Saeed, A., Raza, I., & Shahzad, R. K. (2017). Multilevel classification of security concerns in cloud computing. Applied Computing and Informatics, 13(1), 57-65.
11. Ardagna, C. A., Asal, R., Damiani, E., Dimitrakos, T., El Ioini, N., & Pahl, C. (2018). Certification-Based Cloud Adaptation. IEEE Transactions on Services Computing.
12. Soltanian, A., Belqasmi, F., Yangui, S., Salahuddin, M. A., Glitho, R., & Elbiaze, H. (2018). A Cloud-Based Architecture for Multimedia Conferencing Service Provisioning. IEEE Access, 6, 9792-9806.
13. Ali, H., Moawad, R., & Hosni, A. A. F. (2016). A cloud interoperability broker (CIB) for data migration in SaaS. Future Computing and Informatics Journal, 1(1-2), 27-34.
14. Vázquez-Poletti, J. L., Moreno-Vozmediano, R., Han, R., Wang, W., & Lorente, I. M. (2017). SaaS enabled admission control for MCMC simulation in cloud computing infrastructures. Computer Physics Communications, 211, 88-97.

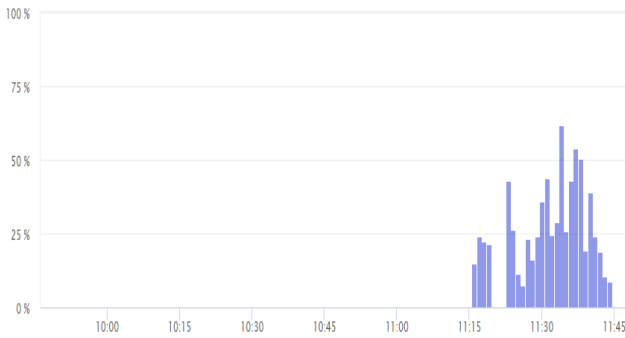


Figure 4. CPU usage

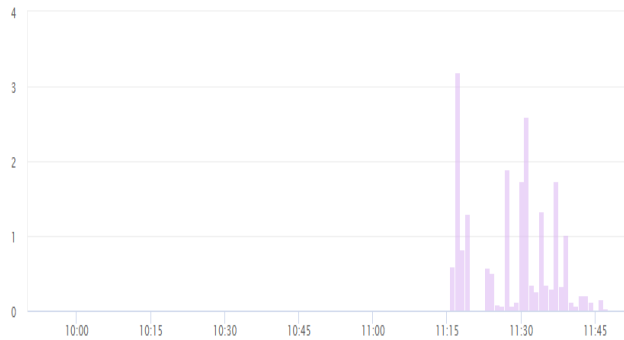


Figure 5. Length of the allocated resource

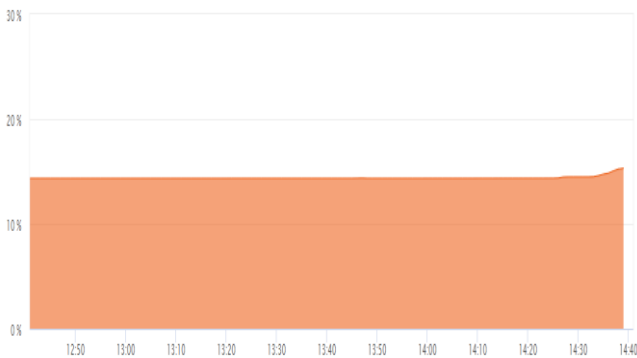


Figure 6. Bandwidth allocation

An efficient scheduling algorithm should increase the utilization of virtual machine resources allocated to it to execute the tasks of scientific workflows. The deployment of virtual machines may be defined as the ratio of the total execution time of virtual machine to the total execution time of the workflow.

QoS Parameters	Simulation results
CPU	50 %
Throughput	1.48 Mb/s
Length	50000
Bandwidth	15000

Table 2. Simulation results

Results of this study show that QoS optimized scheduling algorithms are very efficient with the existing one. This technique lessens the task allocation time.

15. Iqbal, S., Kiah, M. L. M., Anuar, N. B., Daghighi, B., Wahab, A. W. A., & Khan, S. (2016). Service delivery models of cloud computing: security issues and open challenges. *Security and Communication Networks*, 9(17), 4726-4750.
16. S.B.Dash, H.Saini, T.C.Pand, A. Mishra. (2014). Service Level Agreement Assurance in CloudComputing: A Trust Issue. *International Journal of Computer Science and Information Technologies*, 5 (3), 2899-2906.
17. Vincent Chimaobi Emeakaroha. (2012). Managing Cloud Service Provisioning and SLAEnforcement via Holistic Monitoring Techniques. A thesis submitted in Informatics, 1-159.
18. Amit Agarwal, Saloni Jain. (2014). Efficient Optimal Algorithm of Task Scheduling in Cloud Computing Environment. *International Journal of Computer Trends and Technology*. 9(7), 344-349.
19. Rajesh Verma. (2013). Comparative Based Study of Scheduling Algorithms for Resource Management in Cloud Computing Environment. *International Journal of Scientific Research in Computer Science and Engineering*. 2(1), 17-23.
20. Hilda Lawrance and Salaja Silas. (2013). Efficient Qos Based Resource SchedulingUsing PAPRIKA Method for CloudComputing. *International Journal of Engineering Science and Technology*. 5(3), 648-643