ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

Dynamic Resource Allocation by Exploitation Programming Algorithms in Cloud Computing

RAMYA VANMATHY G[1]
PG Scholar,Shanmuganathan Engineering College, Pudukkottai, Tamilnadu, India,
ramyavanmathy.g@gmail.com

NAGARAJ V[2]
Assistant Professor/CSE,Shanmuganathan Engineering College, Pudukkottai, Tamilnadu, India,
nagaraj.va@gmail.com

MANGAIYARKARASI S[3]
Head of the dept/CSE,Shanmuganathan Engineering college, Pudukkottai, Tamilnadu, India,
mangaiprabhu@gmail.com

***Abstract*—Today, computing needs, as it provides dynamic allocation and flexible, reliable resources with service assurance in the form of payment you use cloud service users. Therefore, there must be a provision of all resources. They can be used effectively requesting user who meets customer needs. This allocation of resources is completed taking into account the service level agreements (SLA) and Support parallel processing. In a different account recent work SLA policy is only one parameter. Thus, considering Assign multiple SLA parameters and resources the best mechanism to achieve high-priority tasks can Improve resource utilization in the cloud. In this article, we proposed to consider the replacement task algorithm SLA enforcement and a plurality of parameters, such as memory, required bandwidth and CPU time. The resulting side the results show that in a situation of resources Containment is intense our algorithm provides better utilization Resources.**

***Keywords*—Cloud computing, dynamic resource allocation, resource provisioning, scheduling, scientific workflow.**

I. INTRODUCTION

With significant advances in information technology, computer science is to meet every day of the most important services Community. Cloud overall computing needs is the kind of model, use computing resources providing information technology services through a network (usually the Internet). Cloud providers use [1] Infrastructure as a Service model that allows consumers to rent computing and storage resources and needs Depending on its use. Cloud infrastructure provider is to consumers through obligations, in order to maximize profits Maximum utilization of the minimum infrastructure and resources.

Cloud providers can achieve the established SLA, through the effective deployment of resources, and in accordance with SLA goals, it must deploy performance-optimized applications suitable VM. Currently, more work in the cloud application programming [1], [2], [3]. These methods are generally considered to be a single goal such as SLA enforcement costs, and other run-time due to multi-objective algorithm programmed workload SLA SLA combination with a plurality of parameters, the nature of optimal mapping of resources was found to be NP-hard [4]. Available solutions are based on the use of heuristics.

When a job is sent to the cloud, it is usually divided into several tasks. 1): After that, when performing these tasks are applied parallel processing of the two issues must be considered. How to allocate resources to the task; 2) tasks in the cloud in what order; 3) how to prepare the virtual machine when the cost of programming, terminate or change tasks. Group scheduling and resource allocation can solve these three problems. In embedded systems [5], [6] and high-performance computing [7], [8], they have studied task scheduling and resource allocation.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
tional Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 20, April 2016

Usually efficient supply requires two steps or different processes: (1) Static initial planning steps: the first group of virtual machine group, then drain and deploy a set of physical hosts; and dynamic configuration (2) Resources: Additional allocation of resources to create and migrate virtual machines dynamically respond to different workloads. Step 2 Run continuous production time from step 1 that is normally performed during the initial system configuration and monthly or semi-annual timetable can only be reproduced on the cleaning and general maintenance.

In this article, we focus on the deployment and step 2 above mentioned dynamic resources in order to achieve the purpose of the SLA agreed to the proposed algorithm now fluctuate in response to dynamic workload future tasks of low priority to high priority task, If preferably it is not possible, because the same priority, then resources to create a new form of virtual machines in the world available.

II. RELATED WORKS

In [9] the author proposed architecture, the use of feedback control theory, adaptive management virtualization Resources, based on the virtual machine. Virtual machine hardware resources based on all of the architecture are brought together in common Hosted applications can access the necessary resources calculated on the basis of shared space cloud infrastructure It is necessary to meet service level objectives (SLO) applications. Adaptive administrator in this framework is Multiple-input multiple-output (MIMO) Explorer, which includes 3: controller CPU, memory controller, and the controller I / S, its purpose is to regulate multi-use virtualized resources, achieved by the application by the CPU entry SLO , memory allocation and VM- control I/O Walsh et al pioneering work. [10] Proposed the use of general utility functions of two-tier architecture, In the dynamic allocation of resources by the autonomous region, local agents and global arbiter. Local actor's responsibility is to calculate the utility, such as range or projected workload and available resources and results are given for each AE shift the global arbiter. When the global computing is almost the best arbiter Results resources based on local agents set. In the global arbiter, the new setting The new application of resources allocation for EA, and calculate any new configuration at the end of the fixed control Intervals or in some way or anticipated event trigger violation SLA.

Capacity in the literature [11], the authors propose a resource allocation algorithm adaptive systems, the task preemption cloud what adaptive algorithm based on the updated actual task execution adjust resource allocation. Listing Adaptation Program and adaptive algorithm, every minute of programming (AMMS) is to use a scheduled task, including static allocation of resources generated by the static scheduling tasks offline. Adaptation Online Process for the re-evaluation of the remaining static allocation of resources many times with a predetermined frequency. In Each re-evaluation process, the programmer recalculates the presentation of each task, rather than the task assigned to the cloud completed.

Dynamically extensible architecture through a description of the size of the Web to provide Web application solutions Applications on a dynamic basis, in a computing environment threshold virtualized cloud. Architecture ago by the load balancer, no. VM Web application. Apache HTTP Web application front-end load balancer routes user requests and balance its Apache HTTP load balancer deployed Linux server virtual machine installed. According to the needs of these virtual machines to start and provided by the subsystem configuration. However, the corresponding threshold-based Web applications and dynamic scaling algorithm configured virtual machine configuration Web server control instances of action.

III. TECHNIQUES USED

In this section we are describing SLA based resource provisioning and online adaptive scheduling for Preempt abletask execution, these two methodologies which are combined in proposed algorithm for effective utilization of cloud resources to meet the SLA objective.

- Adaptive Resource Management:

  Resource management adaptation is also known as adaptive management. Adaptive management is a systematic process of iteration In the face of uncertainty so powerful process. Its main objective is to monitor the system, to reduce the period of uncertainty. In therefore, the decision in response to one or more resource management objectives and passive or bothActive, the information needs to be acquired, terms of future management. Adaptive

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
tional Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 20, April 2016

management is based on the learning process this is why the increase in long-term business results.

- Quality of services

Quality of Service (QoS) refers to the various aspects of the allowed transmission of computer network and telephone, Special traffic such as response time, jitter, latency and throughput.

- Multi-tier applications

It provides a multi-tier application architecture that is used to create reusable and flexible development model application. If developers only need to change or add a specific layer, without having to retype the entire application above or if they decide to change or expand the application only undermine the level of technology.

- Service-Level Agreement

Service Level Agreement (SLA) is essentially a service contract when multiple services formally defined. Sometimes the term refers to the time when the contract SLA for service delivery or performance. As an example, an Internet service Vendors typically include the definition of the terms of their contract of service level agreements with customers Service Level (S) to be sold in plain language use. In this case the SLA will typically have a technical definition in terms of mean time between failures (MTBF),mean time to repair or mean time to recovery (MTTR); various data rates; throughput; jitter; or similar measurable details.

Scheduling Algorithm

In proposed priority based scheduling algorithm we have modified the scheduling heuristic for executing highest priority task with advance reservation by preempting best-effort task as done. Algorithm shows the pseudo codes of priority based scheduling algorithm (PBSA).

Algorithm: Priority Based Scheduling Algorithm (PBSA)

1. **Input:** UserServiceRequest
2. //call Algorithm 2 to form the list of task based on priorities
3. get globalAvailableVMList and gloableUsedVMList and also available ResourceList from each cloud schedular
4. // find the appropriate VMListfromeach cloud scheduler
5. **if** AP(R,AR) != ϕ **then**
6. // call the algorithm 1 load balancer
7. deployableVm=load-balancer(AP(R,AR))
8. Deploy service on deployableVM
9. deploy=true
10. **Else if** R has advance reservation and best-effort task is running on any cloud **then**
11. // Call algorithm 3 CMMS for executing R with advance reservation
12. Deployed=true
13. **Else if** globalResourceAbleToHostExtraVM**then**
14. Start newVMInstance
15. Add VMToAvailbaleVMList
16. Deploy service on newVM
17. Deployed=true
18. **Else**
19. queue serviceReuest until
20. queueTime>waitingTime
21. Deployed=false
22. **End if**
23. If deployed then
24. return successful
25. terminate
26. **Else**
27. return failure
28. terminate

IV. EXPECTED RESULT

A. Setup

We evaluate our performance based on the priority by simulating scheduling algorithm. One by one Simulation of the working group completed in 10 games. In each execution Simulation, a group of 70 different analog service Applications (ie jobs), and each includes a service request up to 18 subtasks. We believe in Simulation of clouds. All 70 will be subject to random cloud service requests any time

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
tional Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Special Issue 20, April 2016

soon. In these requests services 70, 15 Application is in AR mode, while the rest is the best way to work with different SLA objectives. That Table 1 Parameter set randomly in simulation According to their maximum and minimum values. Since we only focus on the planning algorithms, we do simulations locally without implementing in any exiting cloud system or using VM interface API.

Parameter	Minimum	Maximum
No. of virtual machine in cloud	23	120
No. of CPU in a VM	1	7
Disk Space	8000	100000
Memory	16	2048
Speed	100	1000

Table 1: Parameters and its Range

B. Result

In Figure 1 shows the average execution job loose situation. We realized that the algorithm PBSA. The minimum average execution time. Resource Parameters when work occurs AR work best be replaced by. Such as Resource contention at least loosened, it is expected that Target part-time work is nearing completion of the actual time. Therefore Adaptation procedure does not affect the date of execution significant.

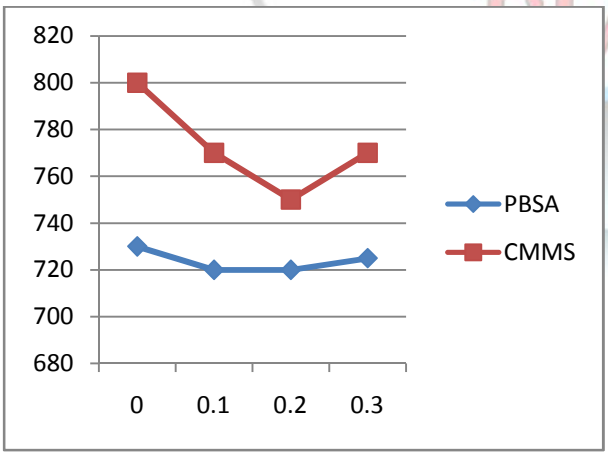

Figure 1. Average job execution time in loose situation

Under prove difficult situation shown in Figure 2 PBSA behavior CMMS better. In stressful situations the scramble for resources more so when the work actually completed it is often later than expected arrival. Because AR preemption works the best, the process of adaptation and upgrade Information more meaningful works in a difficult situation.

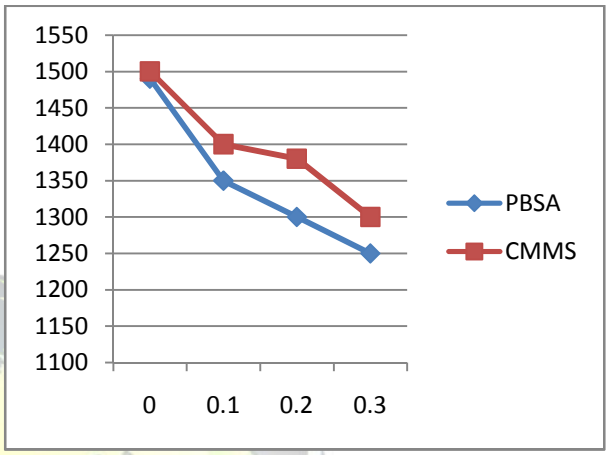

Figure 2 Average job execution time in tight situation

V. CONCLUSION & FUTURE WORK

Cloud computing resources means that in the selection, implementation and management time, management software (e.g., database server, load Balancers, etc.) and hardware resources (for example, CPU Storage, networking, etc.), in order to ensure security application performance. These techniques to improve response time, performance, save Energy, quality of service, SLA. The ultimate goal of resources Configuration is to maximize the benefits of the cloud Prospects for service providers and cloud the user's perspective, in order to reduce costs.

There are many current challenges Strategic resource allocation. A mechanism to overcome the challenges faced by the prior art It must be used. The architecture must be proposed it is suitable for data-intensive applications and high performance computing Also on the actual workload. Mechanism must It recommends that effective use of cloud computing resources to enable QoS and SLA violations in meeting minimization Dynamic Allocation of clouds. Also These mechanisms should also be used to configure SaaS and IaaS users.

REFERENCES

[1] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, "*Cloud computing andemerging IT platforms: vision,*
hype, and reality for delivering computing as the 5th utility",Future Generation Computer Systems 25 (2009) 599–
616.

[2] S. Pandey, L. Wu, S. M. Guru, and R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in
cloud computing environments," in AINA '10: Proceedings of the 2010, *24th IEEE International Conference on Advanced Information Networking and Applications*, pages 400–407, Washington, DC, USA, 2010, IEEE Computer Society.

[3] M. Salehi and R. Buyya, "Adapting market-oriented scheduling policies for cloud computing," In Algorithms and Architectures for
Parallel Processing, volume 6081 of Lecture Notes in Computer Science, pages 351–362. Springer Berlin / Heidelberg, 2010.

[4] J. M. Wilson, "An algorithm for the generalized assignment problemwith special ordered sets," Journal of Heuristics, 11(4):337–350, 2005.

[5] M. Qiu and E. Sha, "Cost minimization while satisfying hard/soft timing constraints for heterogeneous embedded systems," *ACM Transactions on Design Automation of Electronic Systems* (TODAES), vol. 14, no. 2, pp. 1–30, 2009.

[6] M. Qiu, M. Guo, M. Liu, C. J. Xue, and E. H.-M. S. L. T. Yang, "Loop scheduling and bank type assignment for heterogeneousmultibank memory," Journal of Parallel and Distributed Computing(JPDC), vol. 69, no. 6, pp. 546–558, 2009.

[7] A. Dogan and F. Ozguner, "Matching and scheduling algorithms for minimizing execution time and failure probability of applicationsinheterogeneous computing," IEEE Transactions on Parallel and Distributed Systems, pp. 308–323, 2002.

[8] T. Hagras and J. Janecek, "A high performance, low complexity algorithm for compile-time task scheduling in heterogeneoussystems," Parallel Computing, vol. 31, no. 7, pp. 653–670, 2005.

[9] "Adaptive Management of Virtualized Resources in Cloud Computing Using Feedback Control," in First InternationalConferenceonInformation Science and Engineering, April 2010, pp. 99-102.

[10]W. E. Walsh, G. Tesauro, J. O. Kephart, and R. Das, "Utility Functions in Autonomic Systems," in ICAC '04: Proceedings of theFirst International Conference on Autonomic Computing. IEEE Computer Society, pp. 70–77, 2004.

[11] Jiayin Li, MeikangQiu, Jian-Wei Niu, Yu Chen, Zhong Ming, "Adaptive Resource Allocation for Preempt able Jobs in Cloud Systems," in 10th International Conference on Intelligent System Design and Application, Jan. 2011, pp. 31-36.

[12] Yazir Y.O., Matthews C., Farahbod R., Neville S., Guitouni A., Ganti S., Coady Y., "Dynamic resource allocation based ondistributed multiple criteria decisions in computing cloud," in 3rd International Conference on Cloud Computing, Aug. 2010, pp. 91-98.

[13] Goudarzi H., Pedram M., "Multi-dimensional SLA-based Resource Allocation for Multi-tier Cloud Computing Systems,"inIEEEInternational Conference on Cloud Computing, Sep. 2011, pp. 324- 331.

[14] Shi J.Y., Taifi M., KhreishahA.,"Resource Planning for Parallel Processing in the Cloud," in IEEE 13th International ConferenceonHigh Performance and Computing, Nov. 2011, pp. 828-833.

[15] Aoun R., Doumith E.A., Gagnaire M., "Resource Provisioning for Enriched Services in Cloud Environment," IEEESecondInternational Conference on Cloud Computing Technology and Science, Feb. 2011, pp. 296-303.

[16] T. Erl, "Service-oriented Architecture: Concepts, Technology, and Design", Upper Saddle River, Prentice Hall, 2005.

[17] F. Chong, G. Carraro, and R. Wolter,"Multi-Tenant Data Architecture", Microsoft Corporation, 2006.

[18] E. Knorr, "Software as a service: The next big thing", InfoWorld, March 2006.