Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 5.258

IJCSMC, Vol. 5, Issue. 3, March 2016, pg.761 – 765

Scheduling Requirements and Challenges Exploration for Distributed Cloud Environment

Vaishali Balhara[1], Dr. Vandana[2]
[1]Asstt. Prof., PG Dept. of Computer Science, All India Jat Heroes Memorial College, MDU, Rohtak
[2]Asstt. Prof., Comp. Sc. Dept., PT. N.R.S. Govt. College, MDU, Rohtak

[1] vaishali_leo@rediffmail.com; [2] vandanamukeshmalik@gmail.com

Abstract— The cloud system is the current technology which has reduced the load from individual system and provided a centralized control based distributed environment. It provides the sharing of services, data, resource and hardware for public and private environment. This higher scope in the global environment also increased the challenges in the resource distribution and utilization. Cloud system scheduling is one such method applied by middle layer to achieve the controlled distribution of these resources. In this paper, the exploration of various characteristics and challenges of scheduling methods is done. The paper identified different scheduling types and common issues in system execution.

Keywords— Scheduling, Migration, VM Architecture, Distributed Environment

I. INTRODUCTION

Cloud[1][2][3][4][5] system is the distributed system which provides the data and the resource sharing in public and private environment. The services provided by cloud system are connected globally in association to the universal environment. The opportunity driven analysis can be applied to generate the service transition at the lower level. The service cost reduction and the centralized secure information storage has improved the significance of this distributed network. The distribution is provided at hardware and software level. But the criteria specified for service distribution is also integrated to achieve the reliable and robust communication and information sharing in the environment. A cloud system is defined under various strengthen vectors so that the service allocation and distribution can be controlled. The requirement of equalization and uniform distribution is responsible on each factor of cloud system architecture and integrated with each of its layer. To achieve the effective information sharing in load sensitive environment, it is required to connect the system with middle layer specification. The service requirement and the integrated client level estimation is required to provide effective data distribution. One of the major term that enables the equalized service distribution among available clients is cloud scheduling. The scheduling is the process applied with different vectors to set up a process sequence of allocation or service utilization. As the cloud architecture is complex with specification of various hardware, software and virtual components. It is required to classify the variable features, capabilities for cloud system in which way, the allocation in the environment will be done effectively. The basic distribution of this

© 2016, IJCSMC All Rights Reserved

cloud environment with specification of different components is shown in figure 1. The low level of the cloud system is actual the server with some hardware, memory and virtual specification.

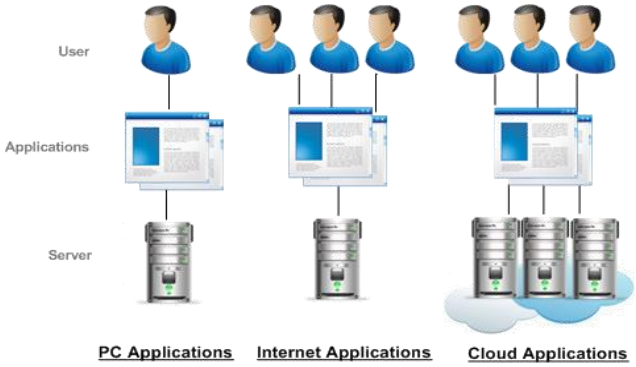

Figure 1 : Cloud System Environment

A cloud system is defined with number of integrated virtual machines. These virtual machines are defined with specific processing power and memory capabilities. Different servers can be formed to control heavy server load. These server formation and categorization can be based on server type, application type and the domain requirement specification. At the middle layer, the application specification is done which requires to map the user profile and the associated control information so that the distribution of resources will be control. The analysis on the requirement and the mapping to availability is done by this layer. The cloud system analysis is done at the static and dynamic parameters. The static parameters here include the physical strengths of the cloud servers and the virtual machines. The dynamic parameters include the response time analysis, availability time, wait time etc. At the top layer, the user itself is placed. The user is here defined with profile specification and the query specification. The middle layer applies analysis on this user query and maps the profile contents to allocate the processes. The process applied for this service distribution among available clients is called scheduling methods. In this paper, the exploration of scheduling process and the relative challenges and problem is defined. The scheduling process is responsible to achieve the effective utilization of cloud resources.

A) Cloud Scheduling

Cloud Scheduling[4][5][6][11][12][13][14][15] is the request driven process applied on various users to define real time processing to achieve the system capabilities unitization in optimum way. The scheduling is the process that actually map the requires to the availability in which way, the overload and underload conditions will be avoided. An effective scheduling algorithm must benefit both the client side and server side. The client side optimization is required to avoid the extra delay and starvation condition whereas for the server side the resource utilization, fail rate must be reduced. There are number of existing scheduling and queuing approaches. But these approaches face some of challenges and issues. In this paper, some such issues are identified.

In this paper, an exploration to the search engine is provided with specification of the impact of ranking algorithms and the associated parameters. The paper has challenges and scope of search engine for real time applications. In this section, the search engine tool, its features and the basic architecture is discussed. In this section, the search impact and associated taxonomy with crawler is discussed. In section II, the work defined by earlier researchers is presented. In section III, the ranking factors and some of ranking algorithms are discussed with feature exploration. In section IV, the conclusion of the work is presented.

II. RELATED WORK

Scheduling in distributed cloud and multi-cloud environment is a challenge. Different researchers provided work to identify and resolve various scheduling problems at different layers of cloud system. Some of earlier work of different researchers is discussed in this section. Author[1] defined work to identify different challenges and problems in computation model for cloud scheduling. Author analyzed the complexities in various scheduling methods and explored the heuristic and meta-heuristic methods. Author identified the concerns associated with real time scheduling problems. Another study work on multi-level scheduling was provided by[4]. The work targeted the role inspection of cloud configuration to extract the service level, task level and machine level scheduling. The feature difference based criteria observation was defined in real time environment under constraint specification. Author identified the comparative observation on various accuracy and scheduling methods. A service request based trust analysis method was proposed by D. Danial[3]. Author applied the global connected analysis on service allocation process to quantify the trust vector. A service level agreement was proposed by the author based on criteria analysis. A grid based fast scheduling method was defined to cover the penalties problem in cloud environment. The trust based request reordering was proposed to

© 2016, IJCSMC All Rights Reserved

improve the service utilization. Multi-cloud[5] scheduling is the key strength of distributed cloud to share the workload among available cloud systems. Author analyzed the available data centers with resource specification and identified the effective distribution using Min-Min and Max Min algorithms. Better resource utilization and effective makespan has proven the successful efforts of author. A work on group[9] specific loud scheduling was defined in multi-cloud environment. Author defined the inter-node analysis to reduce the network latency and provided the algorithmic traffic computation for controlling the scheduling approach. Genez[10] has identified the analysis under performance and cost vectors to define workflow scheduling for hybrid cloud environment. A capacity driven resource analysis is provided in real scenario. Aggregative cost estimation and realistic assumptions are applied to achieve better resource allocation and utilization. Author applied the outsourced monitoring and cost effective resource allocation in distributed environment. Another study work on workflow based scheduling methods was provided by Xin Ye[13]. Author analyzed the associated work taxonomies and identified different aspects of cloud environment. A constraint reconfiguration method is defined to improve the scheduling mechanisms. A security[14] constraint based mining model integration was provided to improve the scheduling method. Author exposed the physical constraints and resource utilization methods for cloud system. A co-scheduling method was defined to apply better characterization of cloud scheduling method. The trust[15] constraint analysis and optimization was defined to improve the customization of clustered scheduling method. Author analyzed the user behavior and improved the workflow modeling for effective cloud scheduling.

Resource adjustment, queue processing are the vectors configured and adjusted by researchers for improving the scheduling mechanisms. Author[2] has applied multi-queue support to each cloud server to improve the job scheduling mechanism. A reservation plan based global scheduler was integrated to provide better resource utilization and sharing. The cloud situation under resource monitoring is observed to handle critical events. Author utilized the available space effectively using balanced spiral method. Another work on queue reformation[7] under size and depth parameter was presented. Author applied the vertical scaled virtualized machine processing at different tiers of cloud architecture. The method is able to provide the scalability in cloud scheduling algorithms. Large request set and multiple cloud systems are processed in parallel for improving the scheduling impact. Tri Queue[8] is the innovative idea to apply resource sharing and improve the space and time complexity with constraint specification. Author applied queuing behavior analysis against various scheduling algorithm. Author provided the effective resource allocation to analyze the scheduling methods.

Researchers also integrated the optimization algorithm with scheduling algorithm to achieve effective and reliable resource allocation. A parallel genetic[6] algorithm was applied to overcome the problem of unbalance assignment problem. Author has investigated the possibility of sequence adjustment for effective cloud scheduling. Mathematical reformed method provided more significant results than traditional genetic approach. A work on workflow[11] analysis based method was defined by the author to improve the scheduling. Author processed the resource provisioning with elasticity vector to compose the massive and heterogeneous resource allocation. Author improved the method using Artificial bee colony algorithm with specification of multiple constraints. Author investigated the performance against multiple feature and algorithmic vectors. A model for dynamic cloud scheduling was proposed using virtual machine migrations[12]. Author analyzed the complexity and pricing criteria for cloud characterization. Author applied the linear integer programming model to provide effective cloud scheduling. The service selection and distribution method is defined at different levels of cloud system.

III.CLOUD SCHEDULING CHARACTERIZATION

Cloud System Scheduling[4][5][6] can be used as a solution for energy efficiency in cloud environment. This solution works at the VMM Level. VM Scheduling can be either single Cloud System Scheduling or dynamic virtual consolidation depending upon the requirement. This section includes the details of Cloud System Scheduling and the methods used for VM Scheduling[5][6][8][9][12].

A. Scheduling and Types

Live Virtual Scheduling is the process of migrating the Cloud System from the host to the destination machine while the machine is in running state. This is one of the most basic operations in virtualised cloud platforms. All such operations are often very mission critical and hence should be completed as fast as possible.. The Scheduling process has two basic steps send and receive. The sending stage copies the memory and disk data into the destination machine and receiving stage maps the guest VM data and the restores its previous stage.

B. Pre Copy Scheduling

When a Scheduling request occurs, pre –copy Scheduling scheme will copy the Cloud System memory state from the source host to the destination host, while it is executing the context of the Cloud System at the source host. When the copy operation for the memory is completed, then the Scheduling manager will identify memory pages which are newly updated on the duration of the copy period, and the time estimate of the latency of

© 2016, IJCSMC All Rights Reserved

transferring those pages. If the latency of transfer is below a particular threshold in that case, the Scheduling manager will stop the Cloud System at source itself. It will then transfer the Cloud System context to the destination machine and then it will resume the execution of the running Cloud System.

C. Post –Copy Scheduling

When a Scheduling request occurs, the post copy Scheduling scheme will transfer the context of the Cloud System, to the destination host. Then resume the Cloud System execution while the state of memory is still present in the source machine. Then a background copying is initiated so that the Cloud System context is sent. The copying process in background is implemented either as linear copy scheme which is transferring the memory pages according to the order of addresses. If the Cloud System context is not sending the requested page of memory to its local memory, then it will trigger a page fault handler remotely in order to obtain the page from the source host. Hence, if the load is not intensive in memory or the copy in Background is not delivering the pages in time, then this policy will not expose the downtime of the machine to the user. Other advantage being, the bounded number of page transfers for memory. This feature makes this scheme rather more silent than pre copy Scheduling.

IV. SCHEDULING CHALLENGES

The scheduling[7][8][9][12][14][15] in the cloud environment is one of the challenging task because of heavy requirement and the limited available resources. The scheduling is not only about to perform the resource allocation but to achieve the balanced distribution of the available services under capabilities observation. The objective of scheduling method is to provide the benefits to both the server side and client side so that the effective resource utilization will be done. The scheduling in the environment is controlled by the middle layer so that the controlled communication will be formed over the environment. Different challenges in the scheduling method are given here under

A. Network bandwidth limitation:

As the cloud system is requested by multiple client at the same time, the parallel communication and communication control is required. To manage these parallel processes and service access, the high bandwidth communication is required. But as the number of requests performed, the available bandwidth get shared and the overall available bandwidth for each process get restricted. In the public domain, the utilization of this bandwidth is a challenge that affects the communication access and sometime affects the response time of cloud servers.

B. Large amount of data:

The cloud system generally manages large volume of data. The requests can be performed by inivididual user to store or access large volume of data. The efficiency and reliability of service allocation also depends on the data size requested by the client. The internal server design is in the form of clusters to manage these large volume data. But to provide the continuous access in the environment, high processing transmission is required over the channel.

C. Synchronous latency:

The latency is the gap between the request performed and the response achieved. The time delay been the arrival time and time of process start is under latency. The latency is estimated based on the expected time of process completion and the actual time of process completion. As some of the scheduling model uses the priority driven resource allocation because of which for some processes or clients the wait time increased. In some cases, this wait time can yield to the starvation. To improve the reliability of scheduling method it is required to avoid the latency of these processes.

D. Intensive Virtual Machine communication:

The virtual machine based communication is applied under application specification. The communication is derived the consideration of its virtual parameters. These parameters are in terms of memory and the processing capabilities. The statistic and dynamic parameters are required to observe for effective scheduling.

E. Number of Migration:

If a scheduled process is not get executed by the allocated virtual machine or cloud server, in such case, it is required to switch the process to some other virtual machine. This process of switching the process between two cloud servers is called migration of processes.

© 2016, IJCSMC All Rights Reserved

V. CONCLUSIONS

This paper has identified the major challenges to the scheduling process applied on distributed cloud system. The paper has identified the scheduling model with specification of different associated components. The paper has categorized the scheduling type in distributed environment. In the final stage, various issues and challenges in cloud system are also been discussed by the author.

REFERENCES

[1] Ashima Mittal," Existing computational models and heuristic techniques for cloud scheduling problems", 978-1-4799-1734-1 @2015.
[2] AV.Karthick," An Efficient Multi Queue Job Scheduling for Cloud Computing", 978-1- 4799-2876-7/13 © 2013 IEEE.
[3] D.Daniel," A Novel Approach for Scheduling Service Request in Cloud with Trust Monitor", 978-1-61284-653-8/11©2011 IEEE.
[4] Yacine Kessaci," Multi-level and Multi-objective Survey on Cloud Scheduling", 2014 IEEE 28th International Parallel & Distributed Processing Symposium Workshops 978-1- 4799-4116-2/14© 2014 IEEE.
[5] Sanjaya K. Panda," An Efficient Task Scheduling Algorithm for Heterogeneous Multi-Cloud Environment", 978-1-4799-3080-7/14@ 2014 IEEE.
[6] Zhongni Zheng," An Approach for Cloud Resource Scheduling Based on Parallel Genetic Algorithm", 978-1-61284-840-2/11©2011 IEEE.
[7] A. Stanislas," MSDoT- Memory Size and Depth of Tier Based Queuing Model for Scheduling a Multi-Tier Cloud", ICETACS 2013, 978-1-4673-5250-5/13 ©2013 IEEE.
[8] AV.Karthick," An Efficient Tri Queue Job Scheduling using Dynamic Quantum Time for Cloud Environment", 978-1-4673-6126-2/13@ 2013 IEEE.
[9] Sheheryar Malik," Latency based Dynamic Grouping aware Cloud Scheduling", 26th International Conference on Advanced Information Networking and Applications Workshops 978-0-7695-4652-0/12 © 2012 IEEE.
[10] Thiago A. L. Genez," On the Performance-cost Tradeoff for Workflow Scheduling in Hybrid Clouds", 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing 978-0-7695-5152-4/13 © 2013 IEEE.
[11] Orachun Udomkasemsub," A Multiple-Objective Workflow Scheduling Framework for Cloud Data Analytics", 978-1-4673-1921-8/12@2012 IEEE.
[12] Wubin Li," Modeling for Dynamic Cloud Scheduling via Migration of Virtual Machines", 2011 Third IEEE International Conference on Cloud Computing Technology and Science 978-0-7695-4622-3/11© 2011 IEEE.
[13] Xin YE," A Survey on Scheduling Workflows in Cloud Environment", 2015 International Conference on Network and Information Systems for Computers 978-1- 4799-1843-0/15© 2015 IEEE.
[14] Christina Delimitrou," Security Implications of Data Mining in Cloud Scheduling", 1556- 6056 (c) 2015 IEEE.
[15] Wenjuan Li," Trust-based and QoS Demand Clustering Analysis Customizable Cloud Workflow Scheduling Strategies", 2012 IEEE International Conference on Cluster Computing Workshops 978-0-7695-4844-9/12 © 2012 IEEE.

© 2016, IJCSMC All Rights Reserved