ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

A Comparative Study of Heuristics for Different Cloud Computing Scheduling Algorithm

R. Ganesan [1], Dr. S. Ravichandran Ph.D[2]

Research Scholar, H.H. The Rajah's College (Autonomous), Pudukkottai, India [1]

Head, Department of Computer Applications, H.H The Rajah's College (Autonomous), Pudukkottai, India [2]

Abstract: Cloud computing has gained popularity in recent times. Cloud computing is internet based computing, whereby shared resources, software and information are provided to computers and other devices on demand, like a public utility. Cloud computing is technology that uses the internet and central remote servers to maintain data and applications. This technology allows consumers and businesses to use application without installation and access their personal files at any computer with internet access. The main aim of my work is to study various problems, issues and types of scheduling algorithms for cloud workflows as well as on designing new workflow algorithms for cloud Workflow management system. The algorithms are compared with each other on the basis of parameters like Total execution time, Execution time for algorithm, Estimated execution time. Experimental results generated via simulation shown that Algorithm 2 is much better than Algorithm 1, as it reduced makes pan time.

Keywords: Cloud Computing, Min Min Algorithm, Max min Algorithm, FCFS Algorithm, Scheduling Algorithm.

I. INTRODUCTION

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services), which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in either privately owned, or third-party data centers that may be located far from the user–ranging in distance from across a city to across the world. Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network

Cloud computing is one of the latest technology that is very popular now a days in IT industries as well as in R&D. This cloud computing technology is a model of development that comes after the introduction of distributed computing. As compare the cloud computing with the distributed computing in this there is a multilevel virtualization. The whole work that is related to cloud computing works in a virtual environment. To get the advantages of cloud user needs to only connect to the internet and after that user can easily use the powerful computing and capacity of storage. Cloud computing services provided by CSP (cloud service provider) as per user requirements. In order to fulfil the demand of different users, they provide different quality of services. In order to conclude the term cloud is an executable environment having dynamic behaviour of resources as well as users providing multiple services. Scheduling is the one of the most prominent activities that executes in the cloud computing environment. To increase the efficiency of the work load of cloud computing, scheduling is one of the tasks performed to get maximum profit. The main objective of the scheduling algorithms in cloud environment is to utilize the resources

There are so many algorithms for scheduling in cloud computing. The main advantage of scheduling algorithm is to obtain a high performance. The main examples of scheduling algorithms are FCFS, Round-Robin, Min-Min algorithm, Max-Min algorithm and meta- heuristic algorithms (ACO, GA, Simulated annealing, PSO, Tabu search and many more). FCFS: First come First serve basis means that task that come first will be execute first. Round-Robin algorithm (RRA): In this Scheduling algorithm time is to be given to resources in a time slice manner. Min-Min Algorithm: Min-Min algorithm selects the smaller tasks to be executed first. Max-Min algorithm: Max-Min algorithm selects the bigger tasks to be executed first. In this paper we will discuss other scheduling algorithms i.e, Heuristic algorithms

II. LITERATURE REVIEW

Deepa.K et al [1]. They descript the improved Hyper- Heuristic Scheduling Approach to schedule resources, by taking account of computation time and makespan with two detection operators. Operators are used to select the low-level

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

heuristics automatically. Conditional Revealing Algorithm (CRA) idea is applied for finding the job failures while allocating the resources. We believe that proposed hyper-heuristic achieve better results than other individual heuristics.

Manel Femmam et al [2]. They demonstrate Depending on the precision of these applications or the business service orientation, three scheduling levels were identified (service level scheduling, task level scheduling, virtual machine level scheduling). The availability and the load on the resources increase with time, so scheduling in the Cloud is a complicated problem especially the need to ensure different (QoS) quality requirements. This article presents a bibliographic research about works dealing with the various scheduling levels in the Cloud Computing.

Shimpy Er et al [3]. This paper descripted by a scheduling with meta-heuristic algorithms is one of the active research area in cloud computing. The goal of cloud task scheduling is to achieve high system throughput and to allocate various computing resources to applications. The Complexness of scheduling problem increases with the size of the task and becomes highly difficult to solve effectively. Many different techniques have been proposed to solve this problem. Some of the methods are based on heuristic techniques that provide an optimal or near optimal solution for tasks i.e, large in size. In this paper we study different scheduling algorithms in different environments with their respective parameters

C T Lin et al [4].They demonstrated scheduling algorithm is needed to manage the access to the different resources. There are different types of resource scheduling technologies in CC environment. These are implemented at different levels based on different parameters like cost, performance, resource utilization, time, priority, physical distances, throughput, bandwidth, resource availability etc. In this research paper various types of resource allocation scheduling algorithms that provide efficient cloud services have been surveyed and analysed

M. Maheswaran et a [5].Optimization problems are in Class NP-hard. These problems can be solved by enumeration method, heuristic method or approximation method. In enumeration method, an optimal solution can be selected if all the possible solutions are enumerated and compared one by one. When number of instances is large, exhaustive enumeration is not feasible for scheduling problems. In that case heuristic is a suboptimal algorithm to find reasonably good solutions reasonably fast. Approximation algorithms are used to find approximate solutions to optimized solution.

III.PROBLEM DESCRIPTION

In recent years ad-hoc parallel data processing has emerged to be one of the killer applications for Infrastructure-as-a-Service (IaaS) clouds.

A growing number of companies have to process huge amounts of data in a cost-efficient manner.

Once a user has fit his program into the required map and reduce pattern, the execution framework takes care of splitting the job into subtasks, distributing and executing them. A single Map Reduce job always consists of a distinct map and reduce program.

Typical Cloud environments do not present regular performance in terms of execution and data transfer times. This is caused by technological and strategic factors and can cause performance variation of up to 30 percent for execution times and 65 percent for data transfer time. Fluctuations in performance of Cloud resources delay tasks execution, what also delays such tasks' successors. If the delayed tasks were part of the critical path of the workflow, it will delay its completion time, and may cause its deadline to be missed. Workflow execution is also subject to delays if one or more of the virtual machines fail during task execution.

To reduce the impact of performance variation of public Cloud resources in the deadlines of workflows, we proposed a new algorithm, called EIPR, which takes into consideration the behavior of Cloud resources during the scheduling process and applies replication of tasks to increase the chance of meeting application deadlines.A.EIPR Algorithm.

The existing algorithm performs following steps
1. Combined Provisioning and Scheduling
2. Data-Transfer Aware Provisioning Adjust
3. Task Replication

A. Combined Provisioning and Scheduling

The first step of the EIPR algorithm consists in the determination of the number and type of VMs to be used for workflow execution as well as start and finish time of each VM (provisioning) and the determination of ordering and placement of tasks on such allocated resources (scheduling).

B. Data-Transfer Aware Provisioning Adjust

In the second step of the EIPR algorithm, the provisioning decision performed in Step 1 is adjusted to account for the aforementioned factors. the algorithm meets the required communication time by setting the start time of

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

the machine earlier than st of the first task, and/or setting the end time of the machine later than the finish time of the last task.

Demerits of Existing Work

- Expensive
- Complex
- Increases data base organization

C. Proposed work

The proposed priority algorithm helps cloud admin to decide priority among the users and allocate resources efficiently according to priority.Once the user's request will be received at the cloud end, after that according to the user's requirement, the resources will be checked for assigning to the user. Batches of the user's requirement will be created according to the type of task, the amount of processor required by the user, and time for the execution of the user.If the resources are not available then the user needs to wait for the resources to be available. The user's waiting request will be compared with all the waiting resources and priority will be assigned accordingly. Algorithms assign resources to workflow application based on the priority of hardware resources or tasks. We felt that capability of hardware resources and performance requirement should be considered while assigning resources to workflow application. We are motivated by the score concept to measure hardware capability and to calculate minimum performance requirement for workflow tasks execution. Microsoft is using the WINSAT (Window System Assessment Tool) module in Windows to measure the hardware capabilities it is running in terms of WEI (Windows Experience Index) score.

IV. METHODOLOGY

A. Scheduling In Cloud

The efficiency of task scheduling has a direct impact on the performance of the entire cloud environment; many heuristic scheduling algorithms were used to optimize it. Scheduling in cloud computing environment is performed at various levels like workflow, VM level, task level etc. The scheduling algorithms are also divided according to scheduling policies i.e. preemptive or non-preemptive. In a non-preemptive (FCFS, SJF) pure multiprogramming system the short-term scheduler lets the current process run until the task is not finished.

B. First-Come-First-Serve (FCFS)

FCFS for parallel processing and is aiming at the resource with the smallest waiting queue time and is selected for the incoming task. The CloudSim toolkit supports First Come First Serve (FCFS) scheduling strategy for internal scheduling of jobs. Allocation of application-specific VMs to Hosts in a Cloud-based data center is the responsibility of the virtual machine provisioner component. Its turnaround and response is quite low. Round robinRound Robin (RR) algorithm focuses on the fairness. The advantage of RR algorithm is that each job will be executed in turn and they don't have to be waited for the previous one to get completed.

C. Max-Min Algorithm

Max-min algorithm allocates task Ti on the resource Rj where large tasks have highest priority rather than smaller tasks. It would be similar to the Min-min make span. For these cases, original Max-min algorithm losses some of its major advantages as load balance between available resources in small distributed system configuration and small total completion time for all submitted tasks in large scale distributed environment. We can't use the Max-min and wait submitted tasks to decide what would be the allocation map, make span, load balance, etc. We try to minimize waiting time of short jobs through assigning large tasks to be executed by slower resources. The algorithm calculates the expected completion time of the submitted tasks on each resource. Then thetask with the overall maximum expected execution time is assigned toa resource that has the minimum overall completion time.

Max min algorithm Steps

Step 1: For all submitted tasks in meta-task Ti
Step 2: For all resource Rj
Step 3: Compute Cij = Eij + rj
Step 4: While meta-task is not empty
Step 5: Find the task Tm consumes maximum completion time.
Step 6: Assign task Tm to the resource Rj with minimum execution time.
Step 7: Remove the task Tm from meta-tasks set
Step 8: Update rj for selected Rj
Step 9:Update Cij for all Ti

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

D. Min-Min Algorithm

Minimum completion time for each task in min-min is computed for all machines. The task with overall minimum completion time is chosen and assigned to corresponding machine. The newly mapped task is removed and the process is repeated till all tasks are mapped. Min-min is a simple and fast algorithm capable of good performance. Min-mean algorithm is a Min-Min algorithm based on resource execution time average value meanCT. The algorithm balances system load with secondary scheduling based on a prime scheduling of Min-Min algorithm .compared to the Min-Min algorithm, the Min-mean algorithm has a certain improvement on load balance; its scheduling performance is not that well as expected due to the heterogeneous nature of grid resources and tasks.

Algorithm Steps

1. For all tasks Ti in meta-task Mv
2. for all resources Rj
3. Cij=Eij+rj
4. Do until all tasks in Mv are mapped
5. For each task in Mv find the earliest completion time and the resource that obtains it
6. Find the task Tk with the minimum earliest completion time
7. Assign task Tk to the resource Rl that gives the earliest completion time
8. Delete task Tk from Mv
9. Update rl 10. Update Cil for all i 11. End do

V. EXPERIMENTAL RESULTS

The proposed work uses simulation to test and verify the efficiency and correctness of the scheduling algorithm presented in this research

A. Simulation Setup

The proposed algorithm is simulated in a simulation toolkit workflowsim. We have created a datacenter with the following properties and created 20 virtual machines.

Arch	OS	Vmm	Time Zone	Cost	Cost per Mem	Cost per Storage	Cost per BW
X86	**Linux**	**Xen**	**10.0**	**0.3**	**0.05**	**0.1**	**0.1**

The CloudSim toolkit is used to simulate heterogeneous resource environment and the communication environment. CloudSim simulator is used to verify the results. The experiments are performed with Sequential assignment which is default in CloudSim and the proposed algorithm. The jobs arrival is Uniformly Randomly Distributed to get generalized scenario. The scheduler submits these jobs on available resources according to these algorithms. All parameters are varied in a similar fashion for judging the performance of the algorithms

B. Analysis of results

The figure shows workflow of task dependency and shows simulation results when MINMIN simulation algorithm is used.

Here the total cost is calculated using the formula

cost= (communication cost + computation cost) = (data (both input and output) * (unit cost of data + runtime * cpu cost) for each task

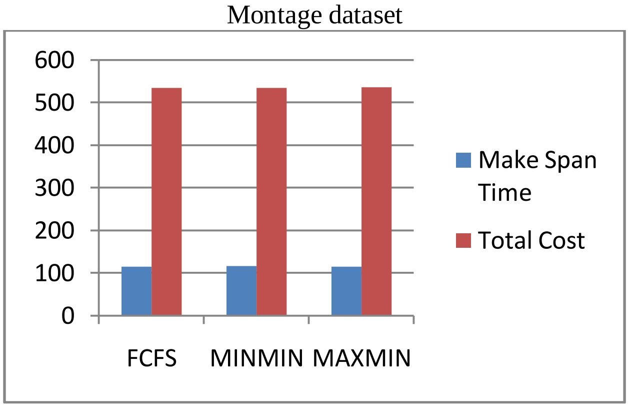
The table shows the makespan and total cost of scheduling of tasks using the algorithms FCFS, MAXMIN and MINMIN with Montage_100 and CyberShake_50 datasets

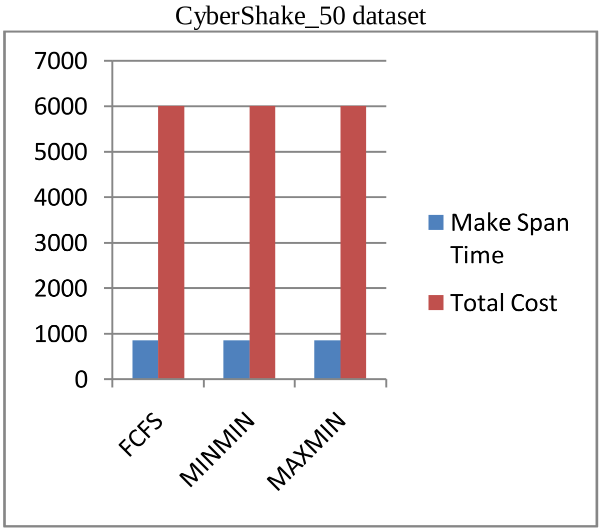
Table 1: Results of various scheduling algorithms with different datasets

Scheduling Algorithms	Dataset	Make Span Time	Total Cost
FCFS	Montage_100	114.2	533.58
MINMIIN	Montage_100	115.19	534.56
MAXMIN	Montage_100	113.55	534.7
FCFS	CyberShake_50	851.2	5996.73
MINMIIN	CyberShake_50	851.2	5996.73
MAXMIN	CyberShake_50	846.95	5996.69

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

Our experiment shows that MAXMIN is the most efficient with respect to the makespan and total cost in comparison to the other algorithms FCFS, MINMIN





VI. CONCLUSION

Minimizing Makspan is one of the objective in scheduling algorithm. The Experiments shown MAXMIN is the most efficient with respect to the makespan and total cost in comparison to the other algorithms FCFS, MINMIN. These algorithms still can be improved by considering multiple objective functions, fault tolerance. Because after a job is submitted to the resource, if the resource becomes unavailable it may affect the makespan and cost of execution.

REFERENCES

[1] Deepa.K, Prabhu.S , Dr.N.Sengottaiyan, “A Hyper-Heuristic Method For Scheduling Thejobs In Cloud Environment”, Informatics Engineering, an International Journal (IEIJ), 2016
[2] Manel Femmam , Okba Kazar , Sohaib Baarir , and Mohamed El-kabir Fareh, “A Comparative Study of Levels Scheduling Algorithms in the Cloud Computing”, nt'l Journal of Computing, Communications & Instrumentation Engg. (IJCCIE), 2015
[3] Shimpy Er., and Mr. Jagandeep Sidhu, “Different Scheduling Algorithms In Different Cloud Environment” , International Journal of Advanced Research in Computer and Communication Engineering, 2014
[4] Lin C T, “Comparative Based Analysis of Scheduling Algorithms for Resource Management in Cloud Computing Environment”, International Journal of Computer Science and Engin and Engineerin, 2013.
[5] J. D. Ullman, “NP-Complete Scheduling Problems,” Journal of Computer System Sciences, Vol. 10, No. 3, 1975, pp. 384-393.
[6] Jiahui Jin, Junzhou Luo, Aibo Song, Fang Dong, Runqun Xiong, “BAR: An Efficient Data Locality Driven Task Scheduling Algorithm for Cloud Computing”, IEEE, 2011
[7] Xin Lu, Zilong Gu, “A load-adapative cloud resource scheduling model based on ant colony algorithm”, IEEE, 2011
[8] Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Wang, D., "Cloud Task Scheduling Based on Load Balancing Ant Colony Optimization", IEEE, 2011.
[9] Xiangqian Song, Lin Gao, Jieping Wang, "Job scheduling based on ant colony optimization in cloud computing", IEEE, 2011
[10] Zhang Bo, Gao Ji, Ai Jieqing, "Cloud Loading Balance algorithm", IEEE, 2011