International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 12 (2018) pp. 10458-10461

An Optimized Algorithm: Maxchild on Workflow Scheduling

Rakesh Kumar, Kamaljeet Kaur

Department of CET, GNDU, Amritsar, Punjab, India.

Abstract

In Cloud computing environment virtual resources are provided to the user on their requests dynamically. The workflow scheduling suffers from number of issues like security, resource allocation etc. In our research paper we are focusing on workflow scheduling problems by using number of algorithms. In cloud this work of scheduling is done by the broker by selecting the number of services and appropriate start time for each workflow. In this research paper, to schedule the workflow in better way max child algorithm is used. This algorithm provides optimal results in case of cost and workflow is scheduled in better way.

Keywords: workflow, cloud computing, scheduling

INTRODUCTION

Cloud computing has grown in a noticeable quality in later a long time, with its mind blowing on request limit administration demonstration. Numerous cloud suppliers are currently dynamic in the market, giving a rich offering a few sort of administrations to the client. The cloud innovation has moved toward becoming standard in big business server farms, where private, public, hybrid cloud structures [1] are continuously more embraced. These numbers of cloud administrations are oversee, create and convey to the client through virtualization innovation. It will payable by cloud resource clients as it were at the point when client's request is satisfied that mean client pay per utilize premise. The utilization of cloud system is considered by QoS that specifies the levels of dependability, execution, security and accessibility. the attributes that are useful for client provided by supplier can be considered using QoS.

Workflow Scheduling

Work flow scheduling is a way by which appropriate resources are assigned to the task that requires them. The assignments of resources are divided into sub task that are depended upon the resources available to acquire the predefined objectives. The sub assignments of resources are utilized in tackling different issues in bioinformatics, space science and business. [6]. The workflow is process that needs to execute the number of task in a specific form by utilizing parent child relationships. In this the youngest assignment of resource is allocated to the parent and must be executed first. After that child task executed and resources are allocated to them. The Directed acyclic Graph is used to handle workflow. The vertex represented the number of tasks allocated and edges represent the conditions through which resource assignment is done.

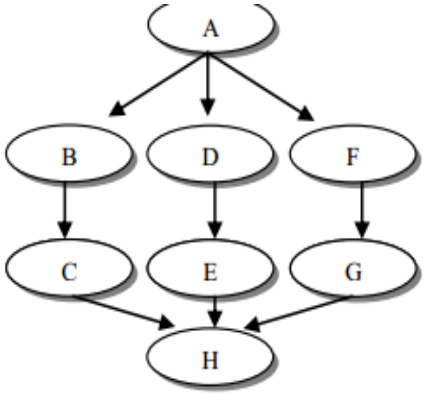

Figure 1: Directed Acyclic Graph

Nature of Workflow Algorithm:

The workflow planning calculations can be:

1. Heuristic in Nature: The heuristic calculations are need based and for the most part issue driven. The creator can utilize his individual learning to designate need for workflow applications and cloud assets.
2. Meta Heuristic in Nature: Meta-heuristic booking calculations are the ones that needn't bother with the human interface and give the general answer for workflow applications.

LITERATURE SURVEY

The resource pool is shared and practically there become infinite resources for use by the federation member and cloud regular user. The study proposed by [1] discussed cloud resource provisioning and availability of resources within hybrid cloud. The study conducted by [2]presented security issues and challenges along with solution within singleton cloud. cost aware scheduling of complex jobs within cloud is proposed by [3]. Advance reservation scheme for scheduling complex jobs within community cloud is proposed by [4] Work is being done towards availability of resources using advance reservation with lack of reliability standards.

Deadline constraint jobs require reliable resources which are ensured by considering reliability metrics. None of the existing work focused on reliability aware advanced reservation scheduling within federation of cloud.

Cloud comes up with resources of logically unlimited resources but in practice, due to large number of users, resources availability is at stake. The resource allocation is on the basis of first come first serve basis. In other words, it becomes impossible for the cloud service provider to satisfy all the needs of the user. To solve this problem [5] suggested a backfilling algorithm with analytical hierarchical process to select best possible solution for lease from best effort queue. Best effort lease allows the selection of resources as soon as they are available. No time constraint is associated with such lease. Hence start time is dynamic. Another category of lease known as advance reservation lease scheduling is used to allow the users to reserve the resources in advance to execute the task. The deadline constraint jobs are executed in such scenario without the starvation problem. Resources are distributed in time critical manner. In the discussed literature performance of the backfilling algorithm is improved by the use of AHP. The backfilling algorithm is required to select the lease from the best effort queue which will provide idle resources to schedule the newly arrived deadline sensitive lease. The slack values are used within backfilling algorithm.

The main goal of the workflow scheduling is to reduce total execution time using limited resource or budget. One of the workflow scheduling algorithms is Heterogeneous Earliest-Finish Time (HEFT) algorithm. It considers critical path to minimize make span which is actual execution time on the distributed computing environment. The critical path means the longest path in the workflow. However, HEFT algorithm is a single objective scheduling considering only make span, not the execution cost. Furthermore, this algorithm is a static scheduling that is hard to cope with unexpected results. Another workflow scheduling algorithm is IaaS Cloud Partial Critical Path (IC-PCP) algorithm. It is multiple objectives scheduling which considers both cost and time, but static scheduling algorithm. Additionally, Time Distribution (TD) Heuristics is dynamic scheduling and multiple objective scheduling. It has a goal to minimize the cost while satisfying the deadline given by user. On task division phase, it classifies simple tasks and synchronization tasks which have multiple parent nodes or child nodes. On planning phase, all workflow tasks are allocated to proper resources based on Markov Decision Process (MDP)

When the workflow applications are executed in Cloud environment, the workflow scheduling and resource provisioning constitutes two main parts. Since the whole procedure of workflow management is considerably complicated for ordinary users, Cloud broker system was suggested as an intermediary. The relationships between workflow service users.

At the beginning, a user submits tasks to be scheduled to Cloud resource broker with SLA constraints. The received workflow is parsed to individual tasks and dependencies between tasks in Workflow Management Module. The parsed workflow is then scheduled by VM allocator using workflow scheduling algorithm. When the workflow scheduling process is executed, the actual allocation of VMs to the given workflow is conducted by Cloud computing resource adaptor. However, the existing workflow scheduling algorithms isolate each request from users and allocate tasks to separate VMs, while each single task usually does not maximally utilize allocated resources such as CPU, memory, and so on. Moreover, the resource pool management could be more efficient if VM requests are predicted and the number of VMs is adjusted, while keeping backup idle VMs to save the startup time.

Flaws of existing approach in terms of downtime or migration time can be further reduced. The fault tolerant strategy which can be optimized for future endeavors is proactive approach.

PROBLEM DEFINITION

The problems and flaws are discussed above. Here the solution and Objectives are discussed. The main objective of this research work is to schedule different workflows in efficient way so that overall completion time would be minimized. In this work, we propose a new heuristic for dynamic allocation of workflow tasks to current available resources. The objective of this scheduling heuristic is to reduce the available resource which is not utilized properly whenever there are no tasks available for scheduling. The objective drawn from the literature is as under

1. To develop a strategy for Energy efficient fault tolerance.
2. If fault occurs then VM's progress will be saved.
3. The down time and migration time is decreased.

PROPOSED METHODOLOGY

MaxChild Algorithm

Input : Cloudlet_List, Dependency_List
Output : Scheduled_List

1. Set Size ← Cloudlet_List.Size

2. for i ← 0 to Size -1

Set Coudlet_List.get (i).Checked to false

end for

3. for i ← 0 to Size -1

a. Set MaxIndex to Zero

b. Set MaxCloudlet to null

c. for j ← 0 to Size -1

if Cloudlet_List.get(j) is not checked

Set MaxCloudlet to Cloudlet_List.get(j)

Set MaxIndex to j

break

end if

end for

d. if MaxCloudlet is null

break

end if

e. for j ← 0 to Size – 1

if Cloudlet_List.get(j) is checked

continue

end if

if Cloudlet_List.get(j).ChildCount > MaxCloudlet.ChildCount Set MaxCloudlet ← Cloudlet_List.get(j)

end if

end for

f. Set MaxCloudlet_Checked to true

g. Set first_IDLE_VM to null

h. for j ← 0 to vmList.Size – 1

if vmList.get(j) status is IDLE

Set MaxCloudlet.VM_ID to j

Set first IDLE_VM to vmList.get(j)

break

end if

end for

if first IDLE_VM is null

break

end for

Fault Injection and Recovery- In this paper we will inject the fault in the system and then by using max child algorithm we will recover the fault that is occurring in the system. Here by fault we mean if eh VM is not working or if the required resources for the subtask can't be fulfilled by the selected VM. The presence of faults in the system causes lots of problem or even can cause disk crash like problems so it is the requirement to recover that type of faults early in the system so that user can not suffer from these problems. This research paper easily handles this problem by using the above proposed methodology.

RESULTS

In terms of Energy consumption

Energy=P1+P2*CPU%age

P1 is the amount of energy consumed by each VM during the execution of job

P2 is the maximum energy which can be consumed

Cpu%age is the Cpu utilization

C	Cloudlets	Energy Consumpion	
		Before fault	After fault
1	50		
2	100		
3	1000	189.6333	215.6312

sno	Cloudlets	Execution Time(us)	
		Before fault	After fault
1	50	389	406
2	100	456	586
3	1000	2987	3293

sno	Cloudlets	Downtime(us)	
		Before fault	After fault
1	50	876	950
2	100	1843	1900
3	1000	18752	19000

Sno	Cloudlets	Migration Time(us)	
		Before fault	After fault
1	50	7833	8550
2	100	16345	17100
3	1000	15234	16245

CONCLUSION

Workflows have been gaining the main focus in scheduling by moving it from Grid computing to Cloud computing. Related work on the existing workflow scheduling algorithms on deadline constraints have been tabulated based on their domain and type of algorithm which aims to satisfy the Quality of Service (QoS) of the user, such as cost minimization within the particular time limit (deadline) specified by the user. But still there are only limited algorithms that are used generally to automate workflows with deadlines as its criterion such as Ant Colony Optimizations, Genetic algorithms, Particle swarm Optimization (PSO), Partial Critical Path (PCP) algorithms and so there is an urge to explore new algorithms in the near future as a deadline constraint is becoming an important aspect.

REFERENCES

[1] A. Juan *et al.*, "OPTIMIS : A holistic approach to cloud service provisioning," *Futur. Gener. Comput. Syst.*, vol. 28, no. 1, pp. 66–77, 2012.

[2] O. Harfoushi, B. Alfawwaz, N. a. Ghatasheh, R. Obiedat, M. M. Abu-Faraj, and H. Faris, "Data Security Issues and Challenges in Cloud Computing: A Conceptual Analysis and Review," *Commun. Netw.*, vol. 6, no. 1, pp. 15–21, 2014.

[3] E. N. Alkhanak, S. P. Lee, S. Ur, and R. Khan, "Cost-aware challenges for workflow scheduling approaches in cloud computing environments : Taxonomy and opportunities," *Futur. Gener. Comput. Syst.*, vol. 50, pp. 3–21, 2015.

[4] A. K. Bhartee, A. Saxena, A. A. Srivastava, and A. K. Patel, "Advance Reservation Policy for Inter-Member Workflow Scheduling in Community Cloud," vol. 4, no. 4, pp. 2891–2894, 2015.

[5] S. C. Nayak and C. Tripathy, "Deadline sensitive lease scheduling in cloud computing environment using AHP," *J. King Saud Univ. - Comput. Inf. Sci.*, 2016.

[6] A. Juan *et al.*, "OPTIMIS : A holistic approach to cloud service provisioning," *Futur. Gener. Comput. Syst.*, vol. 28, no. 1, pp. 66–77, 2012.

[7] O. Harfoushi, B. Alfawwaz, N. a. Ghatasheh, R. Obiedat, M. M. Abu-Faraj, and H. Faris, "Data Security Issues and Challenges in Cloud Computing: A Conceptual Analysis and Review," *Commun. Netw.*, vol. 6, no. 1, pp. 15–21, 2014.

[8] E. N. Alkhanak, S. P. Lee, S. Ur, and R. Khan, "Cost-aware challenges for workflow scheduling approaches in cloud computing environments : Taxonomy and opportunities," *Futur. Gener. Comput. Syst.*, vol. 50, pp. 3–21, 2015.

[9] A. K. Bhartee, A. Saxena, A. A. Srivastava, and A. K. Patel, "Advance Reservation Policy for Inter-Member Workflow Scheduling in Community Cloud," vol. 4, no. 4, pp. 2891–2894, 2015.

[10] S. C. Nayak and C. Tripathy, "Deadline sensitive lease scheduling in cloud computing environment using AHP," *J. King Saud Univ. - Comput. Inf. Sci.*, 2016.