

Modified Scheduling Service Workflow in Hybrid Cloud

Ekta Rai

SVN University, Sagar, Madhya Pradesh, India

Abstract: Workflow is used for representation of variety of application which needs enormous data consumption and storage. To overcome this situation like data consumption and storage cloud computing emerged as best solution for on demand resource provider, but often it's not sufficient to handle this condition, so hybrid cloud is best solution for this problem, where private and public cloud combines. The private cloud is owned by personnel so there is no extra cost in using them while public cloud is owned by organization so we have to pay for that as per our need. While hybrid cloud gives the flexibility for the user but using those two questions overcome first how to break the workflow and second one is which resource we need to borrow from public cloud so that it will fulfill our requirement within range. The modified scheduling workflow gives the less marks span for public cloud and gives us resource to process power to schedule within given deadline.

Keywords: Scheduling Service Workflow, Hybrid Cloud

1. Introduction

The efficient way to use the computer storage is to use the cloud computing where we get enormous storage and it gives the flexibility in the computer environment, we can borrow the borrow environment from other sources like public or private cloud as per their requirement this not only reduces initial investment but also reduces the risk in initial step. By the help of this we can control cost and even reduce operational cost and also increase scalability.

This can be done by creating virtual machine which has benefit over original environment these are network, central processing unit, physical memory and peripherals. This virtual machines has their own operating system, network and application in other words this is similar to original machine because they do have their own server hardware and server like Xen[1]. The user with the help of Xen[1] can use its environment without affecting the other users.

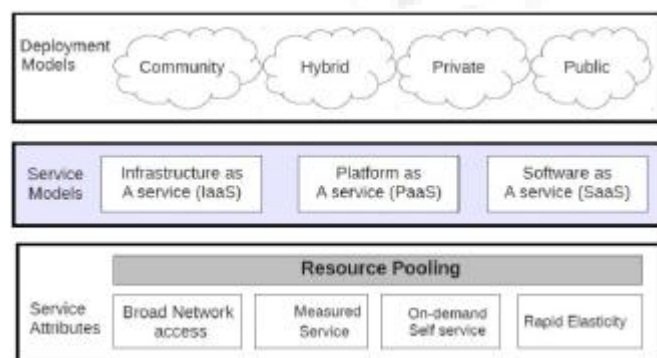


Figure 1: The service models and deployment model

2. Service Models

There are basically three type of this

2.1 Software as a Service (SaaS)

This comprises as software as service where preinstalled application is present with infrastructure, using this one need only browser and device this service is controlled by

centralized architecture, this is very much popular among startup some example are yahoo mail, Google docs, Jio cloud etc. this also don't need update.

2.2 Platform as a Service (PaaS)

Second service named as platform as service in the user get the freedom to represent its application on an available infrastructure, the main advantage is the user don't have to worry about investing in hardware.

2.3 Infrastructure as a Service (IaaS)

The last one where infrastructure is present as a service where the user can customize their entire work with the help of virtual machine, memory and virtual network. The user can customize the virtual software as per their use programmatically.

3. Deployment Model

Based on ownership deployment model [2] comprises of three model [2]

3.1 Private Cloud

Private firm is owned by private firm, sometimes known as company or internal cloud and controlled by individual or firm.

3.2 Public Cloud

It's nothing but controlled or managed by two or more organization and the extension of private cloud.

3.3 Hybrid Cloud

It combination of at least one public and one private cloud

3.4 Problem Concerned

The main problem arise while using the hybrid cloud is how to combine private and public cloud, and how to divide the

workflow between them so that making it more effective in cost cutting and efficiency onwards we will work on this to how reduce this problem.

4. Workflow Scheduling

Scheduling a workflow is one of the major issues in management of workflow execution

4.1 Min-min Algorithm

Min-min algorithm is one of the simplest algorithms used to schedule the task on the available resource. for m task and n resources available we get a matrix of $O(mn)$. For this we find the minimum entry in this matrix and then we find the corresponding task and resource for that entry. Now we schedule that task on that resource. In order to find the next minimum value in second step, we update the available time of this resource and also eliminate the row corresponding to that task. And we repeatedly do this until all the tasks are scheduled. One of the major drawbacks of using the min-min algorithm is that it assigns the smaller tasks to the faster machines.

4.2 Max-min Algorithm

Max-min algorithm was developed so as to overcome the disadvantage of the Min-min algorithm. In this we first find the completion time of each task on each available resource. Then we find out the maximum entry available in the matrix. The row corresponding to this entry gives us the task to be scheduled. Now for the resource we find out the minimum value that is available in this row. Find out the next highest value. And this is repeated until the entire task is scheduled. The disadvantage of this scheduling algorithm is that it may also lead to starvation

4.3 Max-min Algorithm

This algorithm is also based on the completion time of the task. In this algorithm also we first find the completion time of each task on each available resource. Then we find out the maximum entry available in the matrix. The row corresponding to this entry give us the task to be scheduled. Now for the resource we find out the minimum value that is available in this row. Now we schedule the task on this particular resource. For next round we update the available time of the resource and also delete the row corresponding to the task scheduled. And again find out the next highest value. And this is repeated until the entire task is scheduled. The disadvantage of this scheduling algorithm is that it may also lead to starvation. Sometimes it may happen that the task with smaller completion time may not get scheduled. This algorithm also works only for independent tasks.

Algorithm for scheduling the workflow in hybrid cloud was given by L.F. Bittencourt. This scheduling algorithm first schedules the task on the available resources that are present in a private cloud. In this it uses Path clustering heuristics (PCH) [2] algorithm. This algorithm first form the clusters from the workflow. This concept is used so that the time for inter communication between the different virtual machine can be reduced as the there is no need to use the

communication link between as it is the same virtual machine. This concept reduces the time wasted on while communicating between virtual machines.

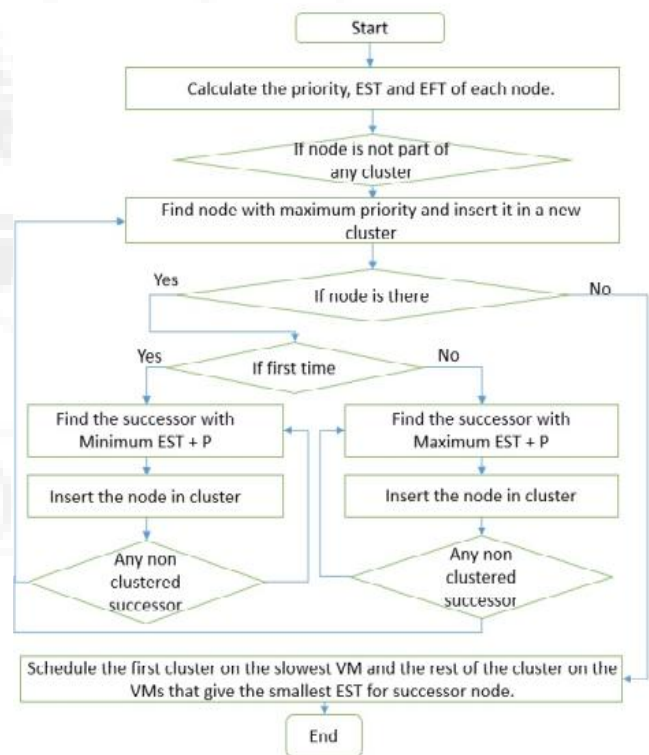
5. Modified Algorithm

Figure below shows the flowchart for the modified algorithm that

5.1 Input

For implementation we take dynamic acyclic graph (DAG) as input. This DAG will have n number of task that is linked together. Each task will have some length which will correspond to number of line of code that task has. This will interconnected and the edge connecting the two task will show the communication cost. The private cloud will have m resources that have the computation power of p_{ri} and these resources are connected to each other with communication links $li;j$. We will also have parameter deadline D , while we be sufficient enough for DAG. If the make span of the DAG is greater than the deadline then the scheduling algorithm will try to get some resources from the public cloud. So that the DAG can be scheduled within the deadline.

5.1.1 Modified Algorithm



References

- [1] ABELS, T., DHAWAN, P., and CHANDRASEKARAN, B., "An overview of xen virtualization," *Dell Power Solutions*, vol. 8, pp. 109–111, 2005.
- [2] MELL, P. and GRANCE, T., "The nist definition of cloud computing," 2011.