

A Novel Approach for Workflow Scheduling in Hybrid Cloud with Dynamic Datacenter

Anuradha Gupta¹, Abhigya Saxena², Brij Kishore³

¹M.Tech. Scholar C.S.E, AIET, Jaipur, India

^{2,3}Assistant Professor, CSE Dept, Jaipur, India

Abstract— Work process is used to speak to variety of use which requires huge data figuring and limit. To vanquish this need of data computation and limit dispersed processing has created as one of the best responses for on asks for resource provider. Nevertheless, on occasion the benefits open to us may not be satisfactory, so the need develops to gather more sources from various fogs. This is done by using the Hybrid cloud. Half and half cloud is mix of open and private cloud. The private cloud is guaranteed by the customer consequently there are no extra charges for using the benefits available in it, however open cloud is controlled by others so we have to pay for the using the advantage as indicated by the businesses. The usage of the cream cloud offers adaptability to the customer. While using the half breed cloud, two most basic inquiries rises. The first is the best approach to segment the work procedure. In addition, the second one is the thing that benefit we need to get from individuals as a rule cloud so it can meet our essential inside the foreordained due date. The changed booking organization work handle for hybrid cloud give the less make span for the DAG than the main figuring and give us the best resources that we need to secure from open cloud to have enough planning vitality to arrange the work procedure inside given due date. We have gone through with two different approaches after scheduling has been performed to check the successful transaction between dynamic and static data center. We contemplated a hybrid approach which will reduce the complexity of the network but at the same time it also perform successful transaction between dynamic datacenters .Each time the transaction has been performed between different centers or clusters the consequences are successful.

Keywords—DAG, Virtulation, Cloud, Datacenter.

I. INTRODUCTION

The most effective route for information calculation and capacity is the utilization of cloud. The fundamental advantage of utilizing the cloud is that it gives flexibility to figuring condition. It can adjust to the client current need as cloud condition gives the obtaining of the assets by the client according to their need. What's more, it likewise decreases the underlying speculation that we make to

purchase the assets that are required by the client to play out the assignment. Not just that it likewise spares the cost of upkeep and furthermore brings down the operation cost and it additionally gives versatility.

The conceptual method for speaking to the processing assets in a subset or an intelligent gathering is called virtualization. This reflection has advantage over the first configure-proportion. Through virtualization we are making an interface for virtual machine which contains the virtual assets like system association, physical memory, focal master ceasing units and peripherals. The virtual machines made have their own working framework, arrange administrations and applications. At the end of the day they are comparable, they are in-ward. As a result of the virtualization same machine can have application detachment, server solidification and additionally equipment standardization. The most utilized virtualization innovation is Xen [1]. Cloud client can have diverse programming and equipment configuration which are given by the cloud supplier. As a result of Xen [1], one client of the cloud can oversee and arrange the machines without influencing alternate clients.

Hybrid Cloud Infrastructure

Today there is a need of on-request processing. Also, it requires adaptability, accessibility and versatility. To have these necessities we need to either include new assets or we have refreshed the assets while remembering that the executing procedures are not influenced. The hybrid cloud framework requires the blend of open clouds and administration situated lattice which are executed utilizing a dynamic administration deployer, which makes the matrix fill in as a private cloud.

The hybrid framework is a mix of private cloud and open cloud [3], of which we can get to and utilize a couple of assets. While utilizing the hybrid cloud, the workflow administration ought to give the office of presenting the errand on the private assets first. The workflow director ought to have the capacity to use the assets accessible in private cloud to their best usage. Furthermore, it ought to likewise has some dynamic organization facility. Not just this, it ought to have the capacity to speak with general society cloud assets so they can be gotten as and when required. The general population assets can contract when

the neighborhood assets are not ready to finish the workflow inside a few necessities. Along these lines, we can expand the computational energy of private cloud without including the new assets so that the on request figuring prerequisites can be met. By giving these prerequisites, the hybrid cloud foundation underpins the basic leadership done by the planning calculation.

The Cloud Interconnection

When we have to acquire the assets from open cloud we utilize the DDVR [3]. It has two gathering administrations. The first is the DDVR. It deals with the functionalities and furthermore speaks with the foundation. The second one is Cloud Interface Service (CIS). It is utilized to interface with general society cloud. It enables the straightforward access to the assets in broad daylight cloud which are same with respect to the assets accessible in the private cloud. For the straightforward get to, the CIS epitomizes the particular specifics of each cloud type (Nimbus, Amazon, Eucalyptus, etc.). One instance of couple DDVR/CIS is required for binding the between workflow manager and the cloud resources for each resource available. To use the cloud resources and request a resource a series of communication takes place. First the GPO communicates with the DDVR, and then DDVR communicates with the cloud through CIS. Then the cloud returns the local values of the VM hired. And now the DDVR can start using it by starting the GT4 and it also creates a DDVR instance in the cloud resource.

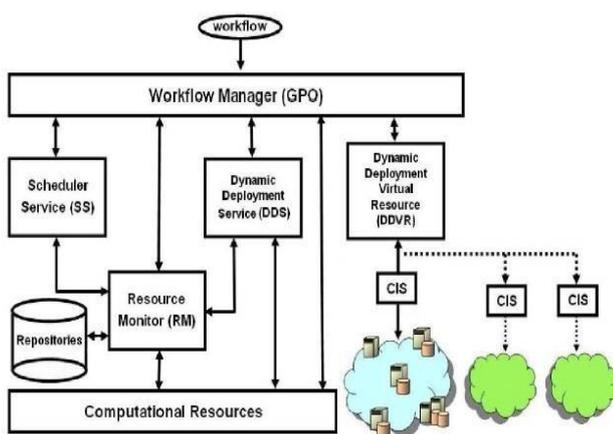


Fig.1.1: Hybrid Architecture

II. ISSUES IN HYBRID CLOUD

Hybrid Cloud requires careful determination of so as to have the best split between public and private cloud. Here the problem arises because the workflow consists of dependent tasks. The problem also includes the dividing of work on heterogeneous resources with heterogeneous link and money charged for using the public cloud. While using the hybrid cloud, we have to carefully split the workflow so that the task can be scheduled on the public and private cloud components. This splitting of the workflow is one of

the major problems that arise in hybrid cloud because the task present in the workflow is dependent not only that, the resources available are heterogeneous in nature and the links connecting them are also heterogeneous in nature.

The resources that are borrowed from public cloud have some cost. So we have to keep them in mind also and try to minimize the cost for using them. So in this thesis, we are presenting a way to divide the workflow of dependent tasks on private and public resources so that the workflow can be completed within a deadline D. And this work also tries to minimize the cost for using the public resources.

Motivation

The problem of scheduling the undertaking for computation is not new. It is one of the fundamental hassle nevertheless exists. Now and again the resources available to us are not enough for scheduling the undertaking inside a constraint. This constraint can close date or budget or any other person precise QOS, but in maximum of cases it's far closing date. So for scheduling the workflow inside a cut-off date we should either upload new resources or we will borrow the resources available to other customers but installing new resources is a high priced enterprise and it is able to time to put in them. So the alternative option to be had is to borrow the assets from others in pay in step with use foundation that is where the idea of hybrid cloud is available in play. In hybrid cloud the scheduling problem receives a chunk complex. Because now we must reflect on consideration on how to divide the workflow in order that it can scheduled on private and public assets. If we schedule most of the assignment on public assets then we can also turn out to be paying a lot more than we've got predicted and if we give less challenge to the general public assets than we leave out deadline. so we need to transfer minimal project to the general public resources so that the undertaking can be finished within the cut-off date and we ought to pay minimum for using the general public sources.

III. PROBLEM FORMULATION AND PROPOSED SOLUTION

INTRODUCTION

The present computation for arranging the work technique in hybrid cloud is given by means of I.f. bittencourt et.al. This arranging figuring before everything designs the errand at the open sources which are available in a non-public cloud. on this essential arranging matter it makes use of route batching heuristics (pch) [2] estimation. This computation at first part the clusters from the work system. The notion for confining the % is that the middle points in same gathering may be moved toward the same digital system. This idea is used so the appropriate open door for intercommunication among the various digital gadget may be decreased as the there's no convincing motivation to

apply the correspondence interface between as its miles the equal virtual gadget. This concept decreases the time wasted on while passing on among virtual machines. After this hidden timetable the make span is checked in opposition to the due date. in case it misses the due date then we will enterprise to transport some p.c. from the non-public to open assets. For doing this we are able to discover the advantage that we can do our work faster and it will likely be much less immoderate. Recalling these conditions we will pick which aid we have to get from people by and large cloud. Thus it will lessen the cost and the paintings technique is moreover completed within the due date.

The main problem that arises while using the hybrid cloud is that how we will divide the workflow between the public and private resources. Our main concern here is to divide the workflow so that cost for using the public resources can be minimized and we can complete the task within a deadline and also the communication between the resources can be minimized.

Computational Power

$$W_{i,r} = \frac{\text{instructions}}{P_r}$$

Where $w_{i,r}$ is the computational cost of task I on resource r.

- $Suc(n_i)$: is the immediate successor of node i.
- $Pre(n_i)$: is the immediate predecessor of node i.

$$\text{Priority: } P_i = \begin{cases} W_{i,r} \\ W_{i,r} + \max \forall suc(n_i)(c_{i,j} + P_j) \end{cases} \text{ otherwise}$$

Where P_i gives the priority level of node i

➤ Earliest Start Time:

$$\text{EST}(n_i, r_k) = \begin{cases} Time_{r_k} & \text{if } i = 1 \\ \max\{Time_{r_k}, ST\} & \text{otherwise} \end{cases}$$

Where $ST_i = \max\{Time_{r_k}, ST(\text{EST}(n_h, r_k) + w_{h,k} + c_{h,i})\}$.

$EST(n_i, r_k)$ is the earliest start time of node i in resource r. $Time(r_k)$ is time of resource r when it will be available.

Simulation and Implementation

There are some activities shown below in Process model. We show a dynamic output of each activity and it also shows the performance of each activity. We have seen ideal condition though blue color as shown below and we also show the operation of activity (i.e green color).

As we seen in model there are two capacity activities, one is single capacity and other is multiple capacities. We have work on the single capacity in which at a time single

data can flow which is efficient and easy to identify the comparison. Therefore our main emphases on single capacity activity model.

Our main task is to reduce the ideal condition of each activity, so we can gain the maximum throughput of overall process. And then we flow the data through Activity.

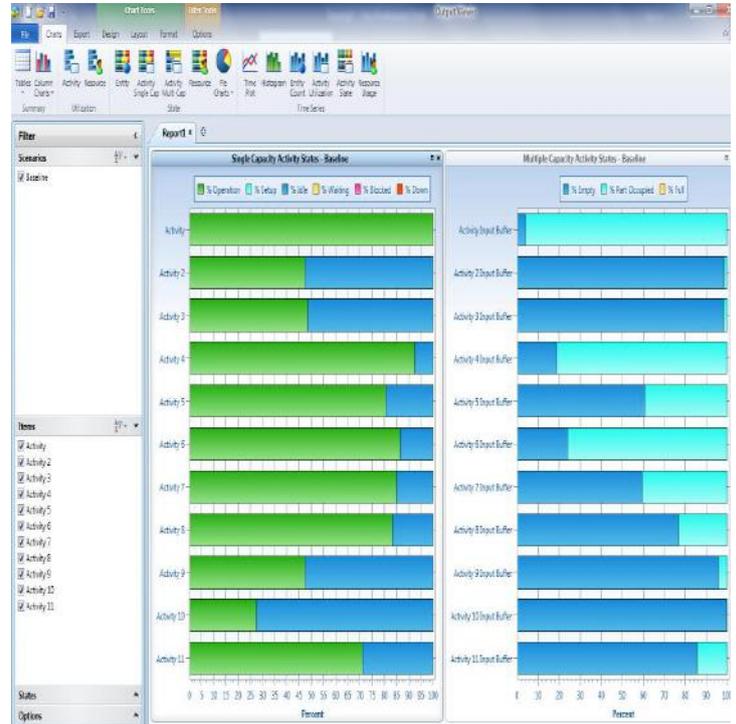


Fig.1.2: Dynamic response of activity

Comparison between existing and propose result

As we seen below the ideal time of activity in existing approach is more (i.e free from data process) but in the purposed work we have reduced the idle time of data process and increase the operational timing of data flow for each activity, so we can gain the maximum throughput of process model (i.e busy in scheduling). This thing we have explained through the process model of hybrid cloud below. As green color shows the operation on data and blue color show the ideal situation of activity.

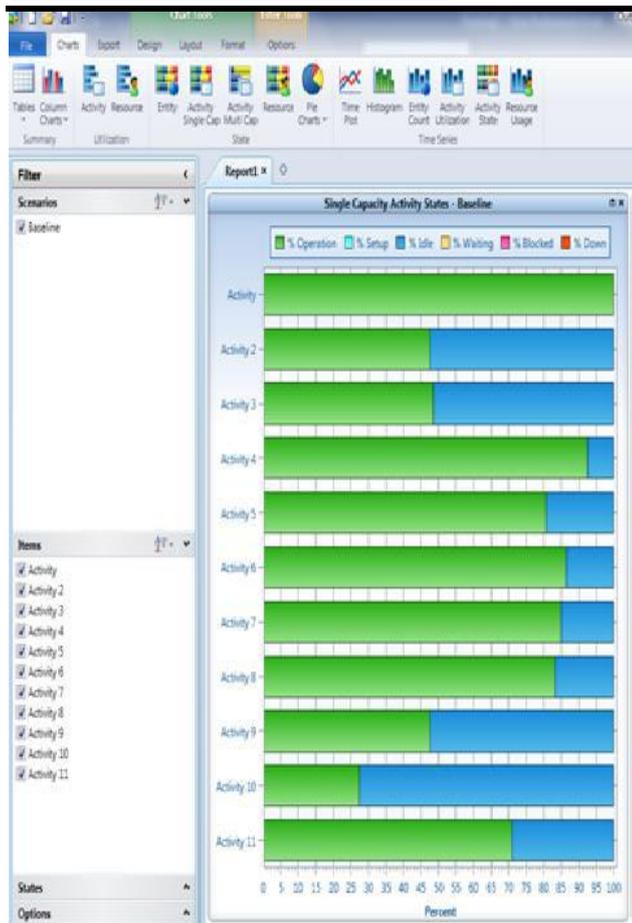


Fig.1.3: Existence output of hybrid cloud

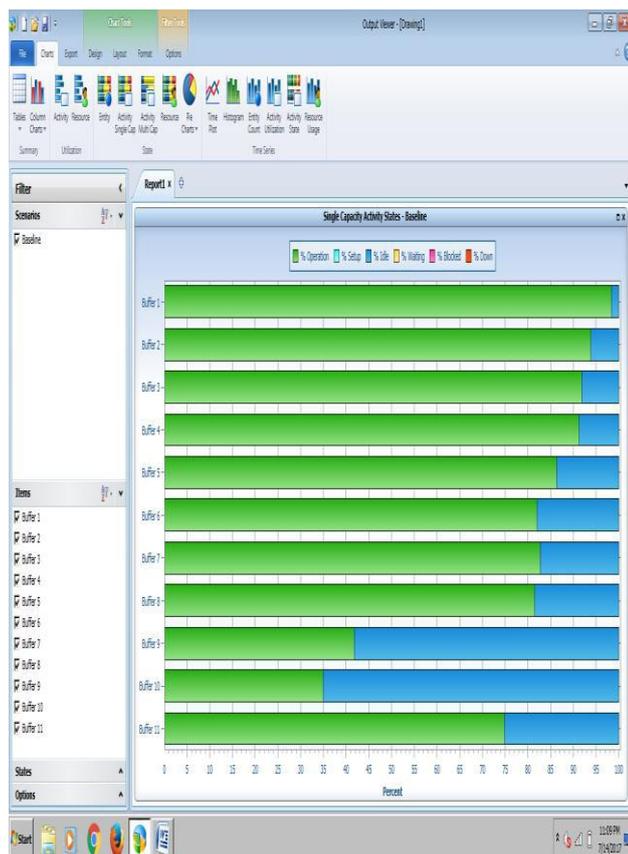


Fig.1.4: Final output of hybrid cloud

This is a pie chart representation for transaction considered for different buffers at the same time in context of work flow scheduling which is shown below.

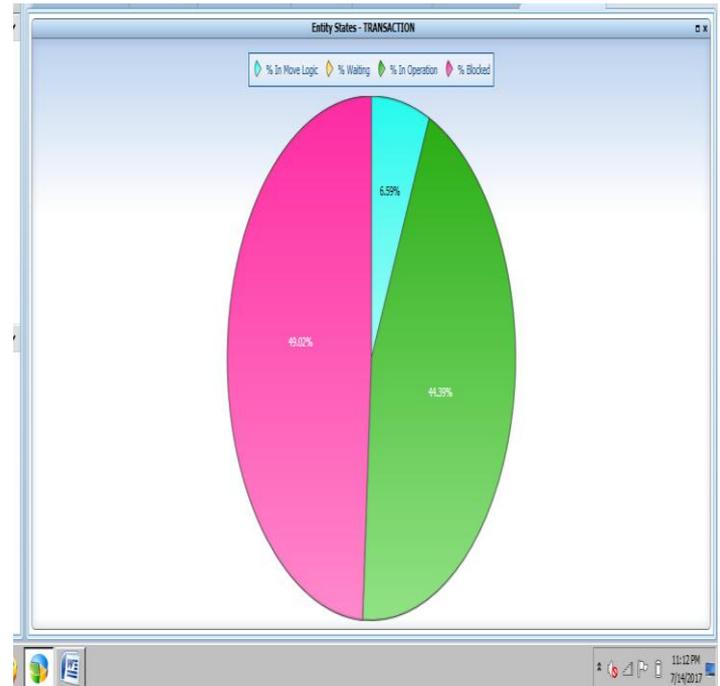


Fig.1.5: Final output of hybrid cloud

IV. CONCLUSION AND FUTURE WORK

In this we've proposed technique work which allows us in lowering the perfect time. This is accomplished by means of lowering the waiting time of the final node in the dag. For lowering the ready time, we are locating the shortest direction in the dag from beginning to ending node rather than the longest route inside the first round and storing the node gift within the direction in a cluster from the second spherical we are doing the same, i.e., find next longest path from the nodes which aren't selected inside the previous iteration and once more node present within the direction are stored in cluster.

REFERENCES

- [1] A BELS, T., DHAWAN, P., and CHANDRASEKARAN, B., "An overview of xen virtualization," Dell Power Solutions, vol. 8, pp. 109–111, 2005.
- [2] B ITTENCOURT, L. F., MADEIRA, E. R., CICERRE, F., and BUZATO, L., "A path clustering heuristic for scheduling task graphs onto a grid," in 3rd International Workshop on Middleware for Grid Computing (MGC05), 2005.
- [3] BITTENCOURT, L. F., SENNA, C. R., and MADEIRA, E. R., "Scheduling ser-vicе workflows for cost optimization in hybrid clouds," in Network and Service Management (CNSM), 2010 International Conference on, pp. 394–397, IEEE, 2010.

- [4] C HAWLA, Y. and BHONSLE, M., "A study on scheduling methods in cloud computing," *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 1, no. 3, pp. 12–17, 2012.
- [5] M. Jaiswal, "Cloud computing and Infrastructure", *International Journal of Research and Analytical Reviews*, vol. 4, no. 2, pp. 742-746, 2017.
- [6] E LZEKI, O., RESHAD, M., and ELSOUD, M., "Improved max-min algorithm in cloud computing," *International Journal of Computer Applications*, vol. 50, no. 12, pp. 22–27, 2012.
- [7] G RAY, M., "Cloud computing: Demystifying iaas, paas and saas," Retrieved July, vol. 17, p. 2011, 2010.
- [8] K OKILAVANI, T. and AMALARETHINAM, D. D. G., "Load balanced min-min algorithm for static meta-task scheduling in grid computing," *International Journal of Computer Applications*, vol. 20, no. 2, pp. 43–49, 2011.
- [9] L IU, G., LI, J., and XU, J., "An improved min-min algorithm in cloud computing, in *Proceedings of the 2012 International Conference of Modern Computer Science and Applications*, pp. 47–52, Springer, 2013.
- [10] M ELL, P. and GRANCE, T., "The nist definition of cloud computing," 2011.
- [11] P ARSA, S. and ENTEZARI-MALEKI, R., "Rasa: A new task scheduling algorithm in grid environment," *World Applied sciences journal*, vol. 7, pp. 152–160, 2009.
- [12] J. Qiu, J. Ekanayake, T. Gunarathne, J. Y. Choi, S.-H. Bae, H. Li, B. Zhang, T.-L. Wu, Y. Ruan, S. Ekanayake, *et al.*, "Hybrid cloud and cluster computing paradigms for life science applications," *BMC bioinformatics*, vol. 11, no. Suppl 12, p. S3, 2010.
- [13] M. A Vouk, "Cloud computing—issues, research and implementations," *CIT. Journal of Computing and Information Technology*, vol. 16, no. 4, pp. 235–246, 2008.
- [14] T. Binz, G. Breiter, F. Leyman, and T. Spatzier, "Portable cloud services using toasca," *IEEE Internet Computing*, vol. 16, no. 3, 2012.
- [15] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya, "Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Software: Practice and Experience*, vol. 41, no. 1, pp. 23–50, 2011.
- [16] K. A. Nuaimi, N. Mohamed, M. A. Nuaimi, and J. Al-Jaroodi, "A survey of load balancing in cloud computing: Challenges and algorithms," in *Network Cloud Computing and Applications (NCCA), 2012 Second Symposium on*, pp. 137–142, IEEE, 2012.
- [17] T. Velte, A. Velte, and R. Elsenpeter, *Cloud computing, a practical approach*. McGraw-Hill, Inc., 2009.
- [18] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in *Computer Security—ESORICS 2009*, pp. 355–370, Springer, 2009.
- [19] M. Hogan, F. Liu, A. Sokol, and J. Tong, "Nist cloud computing standards roadmap," *NIST Special Publication*, vol. 35, 2011.
- [20] S. G. Grivas, T. U. Kumar, and H. Wache, "Cloud broker: Bringing intelligence into the cloud," in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pp. 544–545, IEEE, 2010.
- [21] P. Pawluk, B. Simmons, M. Smit, M. Litoiu, and S. Mankovski, "Introducing stratos: A cloud broker service," in *IEEE CLOUD*, pp. 891–898, 2012