iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** IV **Month of publication:** April 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Analysis of Multicloud Priority Based Job Scheduling For Loosely Coupled Tasks

S.Ammathul Fareedha[1], S.Adhikesavan[2]
Assistant Professor, Department of Information Technology, Ganadipathy Tulsi's Jain Engineering College, Vellore

Abstract: Cloud computing technologies can offer important benefits for IT organizations and data centers running MTC applications: elasticity and rapid provisioning, reduced capital costs, access to potentially unlimited resources and flexibility to deploy their service platforms using outsourced resources. In this project, we are going to explore this scenario to deploy a cluster of resources from multicloud infrastructure for solving loosely coupled MTC applications. The proposed work of system is that the cluster of nodes can be provisioned with resources from different clouds which are organized by priority before deploying the job. Then, the jobs are submitted based on priority scheduling algorithm. This way we prove the viability by evaluating the scalability, performance and cost of different configurations, deployed on a multicloud infrastructure. Then, showing the result of multicloud provider compared with single provider.

Keywords: Cloud Computing, Priority based job scheduling, Multicloud infrastructure, Loosely Coupled Application

I. INTRODUCTION

Cloud computing is a marketing term for technologies that provide computation, software, data access, and storage services that do not require end-user knowledge of the physical location and configuration of the system that delivers the service. Cloud computing describes a new supplement, consumption, and delivery model for IT services based on Internet protocols, and it typically involves provisioning of dynamically scalable and often virtualized resources. It is a byproduct and consequence of the ease-of-access to remote computing sites provided by the Internet. This may take the form of web-based tools or applications that users can access and use through a web browser as if the programs were installed locally on their own computers. At the foundation of cloud computing is the broader concept of infrastructure convergence (or Converged Infrastructure) and shared services. This type of data center environment allows enterprises to get their applications up and running faster, with easier manageability and less maintenance, and enables IT to more rapidly adjust IT resources (such as servers, storage, and networking) to meet fluctuating and unpredictable business demand. Most cloud computing infrastructures consist of services delivered through shared data-centers and appearing as a single point of access for consumers' computing needs. Commercial offerings may be required to meet service-level agreements (SLAs), but specific terms are less often negotiated by smaller companies. Many Task Computing(MTC) paradigm embraces different types of high-performance applications involving many different tasks and requiring large number of computational resources over short periods of time. These tasks can be very different nature, with sizes from small to large, loosely coupled or tightly coupled, or compute intensive or data intensive. Cloud computing technologies can offer important benefits for IT organizations and data centers running MTC applications: elasticity and rapid provisioning, enabling the organization to increase or decrease its infrastructure capacity within minutes, according to the computing necessities; *pay-as-you-go model,* allowing organization to purchase and pay for the exact amount of infrastructure they require at any specific time*; reduced capital* cost, since organization can reduce or even eliminate their in-house infrastructures, resulting on a reduction in capital investment and the personnel costs*; access to potentially "unlimited" resources,*, as most cloud providers allow to deploy hundreds or even thousand of server instances simultaneously; *flexibility*, because the user can deploy cloud instances with different hardware configurations, operating systems, and software packages. The proposed work of this system, once got a services from multicloud infrastructure has to allocate job based on priority. Here, have to set the priority based on performance and cost of cloud provider such as Amazon ec2, windows azure etc., After priority has been given to cloud provider, jobs are scheduled to that provider. Finally, analyzing the performance and cost of each cloud provider, how much time required to execute the job is calculated. When comparing with single cloud provider, multicloud structure gives the performance and cost in more optimized way.

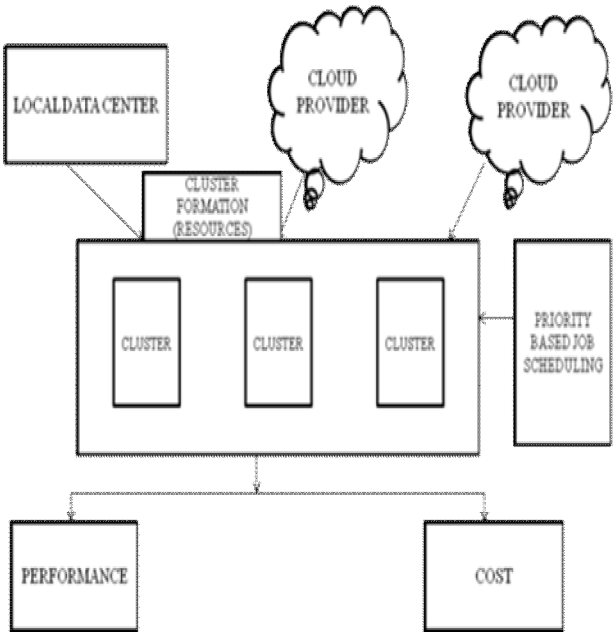


FIG-I SYSTEM DESIGN

II. RELATED WORK

Rafael Moreno-Vozmediano, Ruben S. Montero and Ignacio M. Llorente, "Multicloud Deployment of Computing Clusters for Loosely Coupled MTC Applications"IEEE Transactions on Parallel and Distributed Systems, vol 22, N0.6, June 2011[1]
The paper deals with the different cluster configurations considered in this work have been selected manually, without considering any scheduling policy or optimization criteria.The main goal of analyzing the viability of the multicloud solution from the points of view of performance and
cost.

R.S. Montero, R. Moreno-Vozmediano and I. Llorente, "An Elasticity Model for High Throughput Computing Clusters", to be published in J. Parallel and Distributed Computing, 2010.[2]
The paper deals with the combination of virtual machines and cloud computing, dynamically deliver heterogeneous computational environments. Moreover, the introduction of new virtualization layer between computational environment(cluster VM worker nodes) and physical infrastructure(local node) makes it possible to adjust the capacity allocated to each environment and to supply resources from cloud provider.

M. Murphy, B. Kagey, M. Fenn, and S. Goasguen ,"Dynamic Provisioning of Virtual Organization Clusters", Proc, Ninth IEEE Int'l Symp. Cluster Computing and the Grid, 2009.[3]
An approach to solve the provisioning of workspace environment to meet specific resource constraints and matching job owned by a particular virtual organization (VO) to VO specific virtual cluster running on private networks has not been addressed in the context of virtual clusters. Both of these problems are solved by dynamically instantiating, resizing and removing virtual clusters, while ensuring that grid computing jobs submitted by users affiliated with a specific VO execute only on the VOC owned by that specific VO.

Aboulanga, K. Salem, A. Soror, U. Minhas, P. Kokosielis ,and S. Kamath ,"Deploying Database Appliances in the cloud", Bull of the IEEE Computer Soc. Technical committee on Data Eng.,vol.32,no.1,pp.13-20,2009.[4]
The paper deals with the partitioning of CPU capacity among database appliances as an example end-end solution for virtualized environments. This can be achieved by deploying software in VM environments is the virtual appliance model. By using this, VM image is pre-installed pre-configured application. Deploying this application requires copying this VM image to a physical machine, starting the VM and performing any required configuration tasks.

III. PROBLEM STATEMENT

The objective of this project is to deploy a cluster of resources from multicloud infrastructure for solving loosely coupled MTC applications which is used to embrace different types of high performance applications involving many different tasks and requiring large number of computational resources over short period of time. The proposed work is that the cluster of nodes can be provisioned with resources from different clouds which are organized by priority before deploying the job. The highest priority is given to the system that has the lowest performance ranking of system. The jobs are scheduled to the nodes based on the priority which is done by the scheduling algorithm. This way, we prove the viability by evaluating the scalability, performance and cost of different configurations, deployed on a multicloud infrastructure. When comparing with single cloud provider, multicloud structure gives the performance and cost in more optimized way.

IV. DEPLOYMENT OF VIRTUAL MACHINE

Two cloud provider has been selected. They are windows azure and amazon ec2.Windows azure Remote Desktop Connection:

A. Login windows azure portal using windows live id.
B. In windows azure management portal, create a hosted services and certificate to access it.
C. To create a new instance count, go to windows azure project in visual studio 2010.
From that, select roles (rt. Click) ->Debug->step into new instance
->configuration->instance count (change the count how much VM you want).
Additionally, to get a remote desktop connection, step into new instance->virtual network->Specify subscription ID.
D. Now, go to instance count of windows azure portal, click into it then connect remote access. Then, will get a VM.

Amazon ec2 Remote Desktop Connection:

A. Create Amazon management console account, select ec2
->instances->launch instance
->Quick start instance->select which platform you want->create Key pair id and download.pem file (save it in local system for future use)->retrieve password by attaching that .pem file
->gives public DNS, username and password.
B. Using that go to all programs in our local system, remote desktop connection->Enter computer name, user name and password. Verify it and give virtual machine.

1) Cluster Formation (Based on Priority): After selection of cloud provider, have to allocate job based on priority.The cluster of nodes are provisioned with resources from different clouds. Before deploying a job, measure the performance rate of each cloud provider based on Pass Mark Performance Test Benchmark. Then, set priority to cloud provider based on which process has high performance rate and low cost rate. The parameters considered here for performance are CPU mark, Disc mark, Memory mark, CD mark and Graphics mark.

Algorithm:

1. Read pRate1,cRate1,pRate2,cRate2.

2. Intialise provider1priorityflag=0

provider2priorityflag=0

3. if((pRate1<pRate2) &&(cRate1>cRate2))

{

Provider1priorityflag=1;

Provider2priorityflag=2;

}

else if((pRate2)<(pRate1)&&(cRate2)>(cRate1))

{\

Provider2priorityflag=1;

Provider1priorityflag=2;

}

else if((pRate2)<(pRate1))

{

Provider1priorityflag=1;

Provider2priorityflag=2:

}

Performance monitor	Windows azure	Amazon ec2
Cpu mark	746.0	935.8
Memory mark	365.3	469.2
Disk mark	3722.9	454.3
2Dgraphics mark	-	35.0
3D graphics mark	-	135.2
Cd mark	-	398.6
Pass mark rating	541.1	271.3

TABLE-I
Pass Mark Rating Performance Test Benchmark

2) *Priority Based Job Scheduling:* After priority has been given to cloud provider, jobs are scheduled on virtual machine (cloud provider) based on priority. In which, the highest performance applications are scheduled to the highest priority provider and the lowest performance applications are scheduled to other providers.

3) *Analysis of Performance And Cost:* After priority has been set and jobs are scheduled to particular cloud provider, jobs are measured while allocating it in two cloud providers separately how much the processing capabilities of system differ from jobs are allocated in each provider parallel. It can be measured by Pass Mark Performance Test Benchmark.

When comparing with single cloud provider, multicloud structure gives the performance and cost in more optimized way.

V. IMPLEMENTATION

The Service provider selection module has been implemented by checking processing capability of each node. The two service providers (cloud providers) are windows azure and amazon ec2.Then using two virtual machines clusters are formed based on priority which provider has high performance and low cost. After priority has been given to cloud provider, jobs are scheduled on virtual machine (cloud provider) based on priority. In which, the highest performance applications are scheduled to the highest priority provider and the lowest performance applications are scheduled to other providers. Here, have considered two different kinds of applications. By running these two applications in each machine calculating execution time how long it will take to execute which is differ from one another. The execution time can be expressed as,

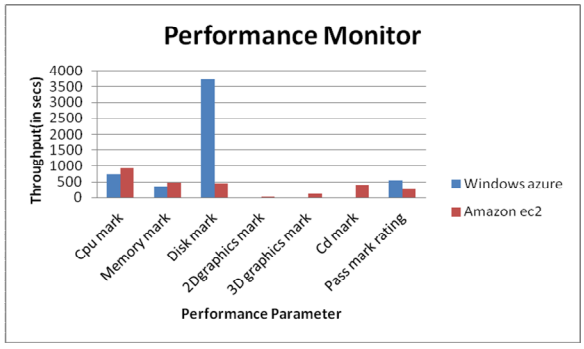
Execution time =end time –start time.

Finally, the performance and cost of each machine is measured and represented by a graph.

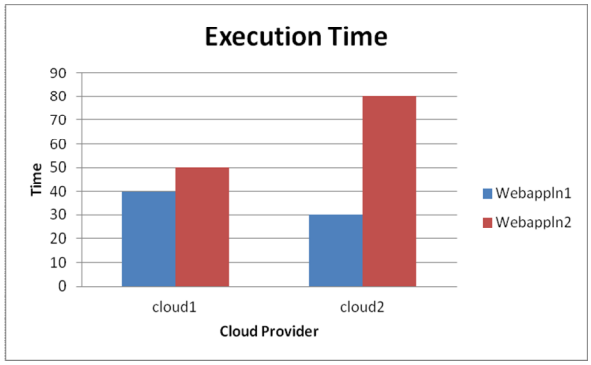
When comparing with single cloud provider, multicloud structure gives the performance and cost in more optimized way.

VI. RESULTS

A new method to solve loosely coupled MTC applications is created using scheduling policy. In proposed system, we can achieve better and effective performance and cost analysis in multi-cloud infrastructure when compared with single cloud structure.
Finally, the comparison of the multicloud infrastructure with single cloud provider shall be represented by a graph.



GRAPH-I
Pass Mark Rating Performance Test Benchmark



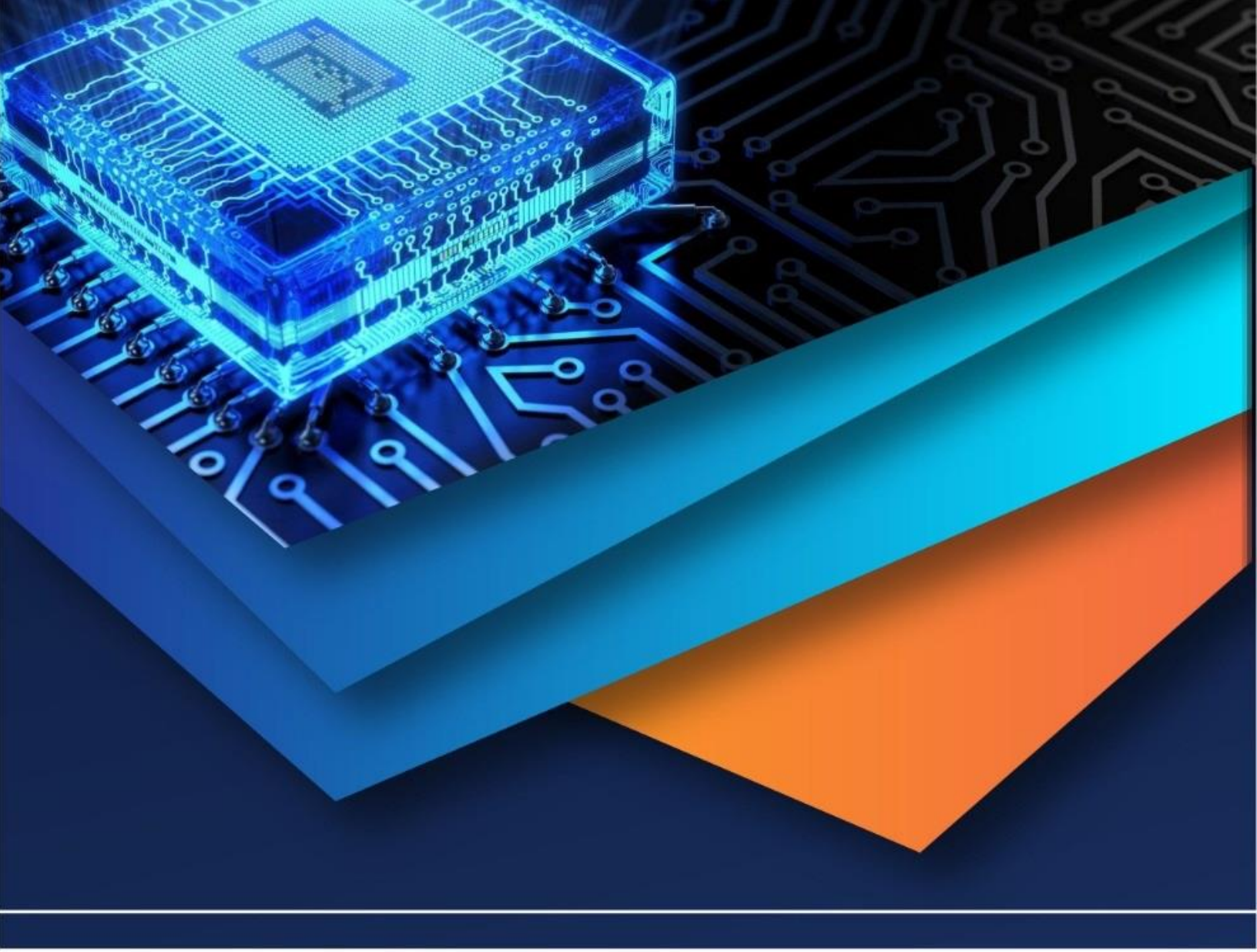
GRAPH-II
Execution Time Calculation

VII. CONCLUSION AND FUTURE WORK

In this paper, I have analyzed the challenges and viability of deploying clusters for solving loosely coupled MTC applications. Here, I have implemented in real time that two different kinds of applications running on two different cloud providers. From this, I have measured the performance and cost of cloud providers based on priority scheduling and the results are represented by graph. The future work is to analyze the infrastructure cost reduction, high availability and fault tolerance capabilities.

REFERENCES

[1] Rafael Moreno-Vozmediano, Ruben S. Montero and Ignacio M. Llorente, "Multicloud Deployment of Computing Clusters for Loosely Coupled MTC Applications"IEEE Transactions on Parallel and Distributed Systems, vol 22, N0.6, June 2011.
[2] R.S. Montero, R. Moreno-Vozmediano and I. Llorente, "An Elasticity Model for High Throughput Computing Clusters", to be published in J. Parallel and Distributed Computing, 2010.
[3] M. Murphy, B. Kagey, M. Fenn, and S. Goasguen ,"Dynamic Provisioning of Virtual Organization Clusters", Proc, Ninth IEEE Int'l Symp. Cluster Computing and the Grid, 2009.
[4] A. Aboulanga, K. Salem, A. Soror, U. Minhas, P. Kokosielis ,and S. Kamath ,"Deploying Database Appliances in the cloud", Bull of the IEEE Computer Soc. Technical committee on Data Eng.,vol.32,no.1,pp.13-20,2009.
[5] I. Llorente, R. Moreno-Vozmediano, and R. Monetero, "Cloud computing for On-Demand Grid Resource Provisioning," Advances in parallel computing,vol.18,pp.177-191,IOS Press,2009.
[6] I.Raicu,I.Foster, and Y.Zhao,"Many Task Computing for Grids and Supercomputers",Proc.Workshop Many-Task Computing on Grids and Supercomputers.pp.1-11,2008.
[7] BioTeam "How to Unicluster and Amazon EC2," technical report,BioTeam Lab Summary,2008.
[8] I.Raicu, Y.Zhao, ,C.Dumitrescu, I.Foster, and M.Wilde, "Falkon:A.Fast and Light-Weight Task Execution Framework," Proc.IEEE/ACM conf. Super Computing,2007.
[9] E.Walker, J.Gardener, V.Litvin and E.Turner, "Creating Personal Adaptive Clusters for Managing Scientific Jobs in a Distributed Computing Environment," Proc.IEEE Second Int'l Workshop Challenges of Large Applications in Distributed Environments(CLADE'06).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY