International Journal of Computer Trends and Technology (IJCTT) - volume4Issue4 –April 2013

Deployment of Computing Clusters in MultiCloud Environment for MTC Applications

R.Sivakami[1], Dr.R.Saravanan[2]

[1]*Associate Professor, Dept of IT, PSNACET, Dindigul, Tamilnadu, India.*

[2]*Professor and Head, Dept of CSE, PSNACET, Dindigul, Tamilnadu, India.*

Abstract- **Cloud computing is a recent technology and it is the use of computing resources (hardware and software) that are delivered as a service over internet. It has been followed by many IT organizations since it is elastic, reliable, flexible and also it allows companies to avoid upfront infrastructure costs. A single cloud provider scheme is a usual thing, at the same time being everywhere access to cloud resources. But, nowadays, parallel usage of services provided by multiple providers is desired. To address this issue and for solving loosely coupled Many-Task Computing (MTC) applications, we build up a scenario to set up a computing cluster on the top of a multi-cloud infrastructure,. With this model, resources from multiple clouds can be supplied to the cluster nodes to obtain the real advantages of cloud computing.**

The main goal of this work is to analyze the viability with respect to scalability, performance, and cost of deploying large virtual cluster infrastructures distributed over different cloud providers for solving loosely coupled MTC applications. In this work, we have simulated an infrastructure model that includes a larger amount of computing resources and multiple tasks. Performance and cost results obtained from the simulation show that they can be extended to cluster infrastructures and large scale problems.

Keywords-**cloud, multi-cloud infrastructure, many-task computing.**

I. INTRODUCTION

Many-task computing (MTC) aims to bridge the gap between two computing paradigms, high throughput computing (HTC) and high-performance computing (HPC). it differs in the emphasis of using many computing resources over short periods of time to accomplish many computational tasks (i.e. including both dependent and independent tasks). Tasks may be small or large, uniprocessor or multiprocessor, compute-intensive or data-intensive. The set of tasks may be static or dynamic, homogeneous or heterogeneous, loosely coupled or tightly coupled. The aggregate number of tasks, quantity of computing, and volumes of data may be extremely large. MTC includes loosely coupled applications that are generally communication-intensive but not naturally expressed using standard message passing interface commonly found in HPC, drawing attention to the many computations that are heterogeneous but not "happily" parallel. Loosely coupled applications often have dependencies among tasks, and typically use files for inter-process communication. Efficient support for these sorts of applications on existing large scale systems will involve substantial technical challenges. MTC paradigm[1] embraces different types of high-performance applications involving many different tasks, and requiring large number of computational resources over short period of time. These tasks can be of very different nature, with sizes from small to large, loosely coupled or tightly coupled, or compute-intensive or data-intensive.

Cloud computing technologies provide many important benefits for IT organizations and data centers running MTC applications:

- Cost Efficient - Companies can avoid spending large amounts of capital on purchasing and installing their IT infrastructure or applications. Technical support is provided round the clock by reputable providers, reducing administration costs.
- Almost Unlimited Storage - Storing information in the cloud gives almost unlimited storage capacity. Hence, no need to worry about running out of storage space or increasing current storage space availability.
- Backup and Recovery - All data backup and disaster recovery is taken care of as part of the service. Files are stored twice at different remote locations to ensure that there's always a copy available 24 hours a day, 7 days per week.
- Automatic Software Integration - Organizations do not need to take additional efforts to customize and integrate their applications. Software integration is done automatically.
- Easy Access to Information – Cloud -based IT services let customers access their applications and data securely from any location via an internet connection.
- Quick and Easy Deployment - Without the need to purchase hardware, software licenses or implementation services, an organization can get its Cloud-computing arrangement off the ground in minutes.

Our work in this regard is organized as follows:

- Implementation of a real computing cluster testbed on top of multi-cloud infrastructure which is deployed over multiple sites.

- Analysis of the performance of the cluster testbed to solve loosely coupled MTC applications, in order to show the scalability of the multi-cloud solution for the given workloads.
- Analysis of cost and cost-performance ratio of the experimental setup, to compare the different cluster configurations, and to prove the viability of the multi-cloud solution.
- Implementation of a simulated infrastructure model to test larger sized clusters and workloads, proving that results obtained in the real testbed can be extended to large-scale multi-cloud infrastructures.

II. EXISTING MODELS

Existing system concentrates on the deployment of multi-cloud architecture. For having multi-cloud architecture, we need to process more than one task at a time. Multi-processing needs more accuracy, scalability and efficiency. It lacks some performance measures. Efficient management of large-scale cluster infrastructures has been explored for years, and different techniques for on-demand provisioning, dynamic partitioning, or cluster virtualization have been proposed.

Traditional methods for the on-demand provision of computational services consist in overlaying a custom software stack on top of an existing middleware layer. Up port the execution of these two projects on the TeraGrid. The system we describe builds personal clusters for users to submit, manage and monitor computational jobs across sites on the TeraGrid. These personal clusters can be managed by Condor, or Sun Grid Engine (SGE), providing a single uniform job management interface for users of the TeraGrid. Furthermore, the system allows TeraGrid resources to be added to existing departmental Condor or SGE clusters, allowing the expansion of local resources on-demand during busy computation periods.

If Physical clusters are used as computing clusters to solve many-task computing problems, they possess following drawbacks:

- Major capital expenditure and administration costs.
- Possible downtime based on the reliability of Internet connection.
- Lack of support due to overloading and insufficient computational resources during peak demand periods.
- Inflexibility of using applications.

To address these issues, Cloud computing technology has been projected as a feasible solution for deploying elastic computing clusters, or for complementing the in-house data-center infrastructure to accommodate peak workloads. As an example, we consider the BioTeam[2] which has deployed the Univa UD UniCluster Express in a hybrid setup. This setup combines local physical nodes with virtual nodes deployed in the Amazon EC2. This hybrid solution [3] is extended by introducing virtualization in the local site, which provides an elastic, quick and easy management of the whole infrastructure that may include resources from remote providers. Virtualization simplifies the delivery of services by providing a platform for optimizing complex IT resources in a scalable manner which makes cloud computing cost effective. The simultaneous use of different cloud providers to deploy a computing cluster across different clouds can provide several benefits.

III. DEPLOYMENT OF A MULTI-CLOUD VIRTUAL CLUSTER

The main goal of this work is to analyze the viability with respect to scalability, performance, and cost of deploying large virtual cluster infrastructures distributed over different cloud providers for solving loosely coupled MTC applications[6].

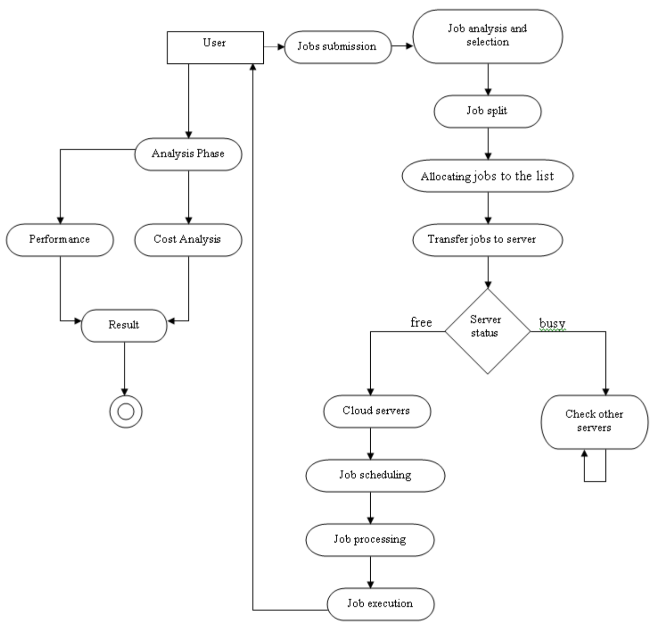

Figure 1. Flowchart

Our work is explained with the above flowchart which is described with the following modules:

A. Job Scheduling
B. Deployment of a Multi-Cloud Infrastructure
C. Infrastructure Model Simulation
D. Performance Analysis
E. Cost Analysis

Before discussing these modules in detail, we explain the architecture briefly. Large numbers of users submit variable numbers of jobs. These jobs are carefully analyzed according to their characteristics and then they are divided into separate groups. Based on this classification, the jobs are assigned to corresponding servers and the jobs are getting done. If no free server is found for processing the job, other idle servers are checked for executing jobs. Once execution is finished, performance and cost measurements are done to identify some better configuration.

A. Job Scheduling

In a large computer, a job queue is established to run a sequence of programs over any period of time. It checks for the free resource and then assigns the task. Job scheduling should not be confused with process scheduling. Clicking on the title of a range in the list of ranges shown in the topmost level of the settings dialog opens the list of sections of the selected range (if any). The (Schema in the range) link, which opens the list of schemas belonging to the selected range, is also to be found here. The first item in this list of schemas is (new entry) which is used to open the dialog used to create a new schema. Under the (new entry) link is a list of entries which have already been made in this schema, and which can be edited using an edit entry dialog. This dialog is opened by clicking on the title of the entry. The jobs are selected and added in database and filtered and then scheduled. The selected jobs in database are laterally scheduled retrieving from the database server.

Sun Grid Engine (SGE) software is used to maintain the queuing system. It consists of an SGE master and a specific amount of virtual worker nodes.

B. Deployment of a Multi-Cloud Infrastructure

In this module, the developed distributed cluster testbed is deployed on top of a multi-cloud infrastructure. This type of multi-cloud deployment introduces numerous difficulties, in the form of cloud interface standard unavailability, service master images distribution and management and the interconnection links between the service components. In this module the scheduled jobs are received here and then stored in a local disk.

The experimental testbed is implemented from a virtual cluster deployed in our local data center, with a queuing system as explained in job scheduling. By providing new virtual worker nodes on remote clouds, this cluster can be scaled-out. ElasticHosts and Amazon EC2 are used as the cloud providers in this work.

The cost of the local resources is calculated based on the model which includes the cost of the computer hardware, the cooling and power expenses, and supporting manpower. In addition to the hardware characteristics of different testbed nodes, this deployment cost is calculated based on the amount charged by the cloud provider for the use of its resources per time unit.

C. Infrastructure Model Simulation

The simulation of infrastructure model to test larger sized clusters and workloads is done in this module, proving that results obtained can be applied extensively to large-scale multi-cloud infrastructures. The scheduled jobs perform their tasks in cloud server.

This simulated infrastructure model includes a larger number of computing resources. Since we have some hardware limitations in our local infrastructure, and also due to the high cost involved in renting many cloud resources for long periods, the tested cluster configurations are limited to a small number of computing resources, running a small number of tasks. We consider up to 16 number of worker nodes in the cluster and up to 128 number of tasks.

From the results obtained, the capability of these multi-cloud cluster solutions for solving MTC problems is realized. This realization can be extended for problems involving large number of tasks with large amount of computing resources.

D. Performance Analysis

Figure 2. Performance Analysis Chart

Performance analysis of the cluster for solving loosely coupled MTC applications proving the scalability of multi-cloud solution for this kind of workloads. For Efficient handling of multiple tasks, we need to have the performance analysis of all tasks. For executing loosely-coupled applications, different configurations of the computing clusters are used by considering variable number of worker nodes with different number of jobs. Their performance are measured and analyzed in this section.

The cluster performance (in jobs completed per second) for executing loosely coupled high throughput computing applications, can be easily assessed using the equation given below:

$$r(n) = (\ r_i)/(1+n_{1/2}/n)$$

where n is the number of jobs completed, r_i is the asymptotic performance which is defined as the maximum rate of performance of the cluster in jobs executed per second, and $n_{1/2}$ is the half-performance length. From the graph shown in Fig.2, we can observe that the performance model provides a good description of the clusters in the execution of the workload taken for study.

The parameter r^i can be used to evaluate the cluster throughput, so that we can compare different cluster configurations. This is due to the accurate approximation of r^i for the maximum performance (in jobs per second) of the cluster in saturation.

E. Cost Analysis

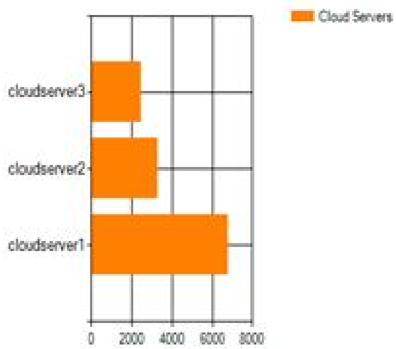

Figure 3. Cost Analysis Chart

Apart from performance measures, the cost of cloud resources used in the solution plays a vital role in analysis. So it is necessary to measure the cost involved in the process and also it is equally important to measure the performance to cost ratio. Since we are executing different number of jobs for every configuration, this helps us in finding the best possible configuration for solving the task. The simulation of different cluster configurations shows that the performance and cost results can be adapted to large-scale problems and cluster infrastructures.

F. Advantages

If different cloud providers are simultaneously used to deploy a computing cluster across different clouds can provide several benefits:

- Almost zero upfront infrastructure investment
- Just-in-time Infrastructure
- More efficient resource utilization
- Usage-based costing
- Reduced time to market
- High availability and Fault tolerance
- Disaster Recovery and Business Continuity
- More Mobility

IV. CONCLUSION

In this paper, a testbed cluster which is based on SGE queuing system is implemented. It comprises computing resources from our in-house infrastructure, and external resources from different clouds. Performance analysis shows that, the cluster throughput is directly proportional to the number of worker nodes for the MTC workload for loosely coupled parameter sweep applications. If we increase the number of nodes from cloud providers, then, in effect, the cluster throughput also increases linearly. This observation proves that the multi-cloud implementation of a computing cluster is having good scalability. Cost analysis shows that, hybrid configurations involving local and cloud nodes reveal better performance-cost ratio than the local setup for the workload considered. This proves that the multi-cloud solution is also attractive from the view of cost.

The simulation of different cluster configurations shows that this realization can be extended for problems involving large number of tasks with large amount of computing resources. But we should note that the results can vary significantly for other MTC applications of different characteristics.

REFERENCES

[1] Raicu, I. Foster, Y. Zhao, Many-Task Computing for Grids and Supercomputers, Workshop on Many-Task Computing on Grids and Supercomputers, 2008, pp. 1–11

[2] BioTeam, Howto: Unicluster and Amazon EC2, Tech. rep., BioTeam Lab Summary (2008).

[3] I. Llorente, R. Moreno-Vozmediano, R. Montero, Cloud Computing for On-Demand Grid Resource Provisioning, Advances in Parallel Computing, Vol. 18 (2009), IOS Press, pp. 177–191.

[4] Amazon Elastic Compute Cloud, http://aws.amazon.com/ec2.

[5] ElasticHosts, http://www.elastichosts.com/.

[6] Rafael Moreno-Vozmediano, Ruben S. Montero, Ignacio M. Llorente, "Multicloud Deployment of Computing Clusters for Loosely Coupled MTC Applications," IEEE Transactions on Parallel and Distributed Systems, vol. 22, no. 6, pp. 924-930, June 2011.