OPTIMIZATION OF PERFORMANCE AND SCHEDULING OF HPC APPLICATIONS IN CLOUD USING CLOUDSIM AND SCHEDULING APPROACH

D. BoobalaMuralitharan[1],S.ArockiaBabiReebha[2],D. Saravanan[3]

[1]*Assistant Professor, Department of MCA, Saranathan College of Engineering,Trichy.*
[2]*Head Of The Department Of Computer Science And Engineering,PavendarBharathidasan College of Engineering and Technology, Trichy.*
[3]*Principal, PavendarBharathidasan College of Engineering and Technology, Trichy.*

ABSTRACT - Cloud computing is emerging as a promising alternative to supercomputers for some High-Performance Computing (HPC) applications. CloudcomputingisanessentialcomponentofthebackboneoftheInternetofThings(IoT).Cloudsareneededto supporthugenumbersofinteractionswithvaryingqualityrequirements.Hence, Servicequalitywillbea vitaldifferentiatoramongcloudproviders.Inordertodifferentiatethemselvesfromtheircompetitors,cloudproviderss houldofferbestservicesthatmeetcustomers'expectations.Aquality modelcan beusedto represent,measureandcomparethequalityoftheproviders,suchthatamutualunderstandingcanbeestablishedamongc loudstakeholders.With cloud as an additional deployment option, HPC users and providers faces the challenges of dealing with highly heterogeneous resources, where the variability spans across a wide range of processor configurations, interconnects, virtualization environments, and pricing models. HPC applications are increasingly being used in academia and laboratories for scientific research and in industries for business and analytics. Cloud computing offers the benefits of virtualization, elasticity of resources and elimination of cluster setup cost and time to HPC applications users. Effort was taken for holistic viewpoint to answer the questions - why and who should choose cloud for HPC, for what applications and how the cloud can be used for HPC? Comprehensive performance and cost evaluation and analysis of running a set of HPC applications on a range of platforms, varying from supercomputers to clouds was carried out. Further, performance of HPC applications is improved in cloud by optimizing HPC applications' characteristics for cloud and cloud virtualization mechanisms for HPC. In this paper, a novel heuristics for online application-aware job scheduling in multi-platform environments is presented. Experimental results and Simulations using CloudSim show that current clouds cannot substitute supercomputers but can effectively complement them.

Keywords - *Cloud computing, High-Performance Computing (HPC), Job scheduling, CloudSim.*

1. INTRODUCTION

Cloud computing is a computing paradigm, where a large pool of systems are connected in private or public networks, to provide dynamically scalable infrastructure for application, data and file storage. With the advent of this technology, the cost of computation, application hosting, content storage and delivery is reduced significantly. Cloud computing is one of the most explosively expanding technologies in the computing industry today. A Cloud computing implementation typically enables users to migrate their data and computation to a remote location with some varying impact on system performance. This provides a number of benefits which could not otherwise be achieved. Such benefits include:Data Protection, Data Recovery and Availability, Management Capabilities, Scalability, Quality of Service (QoS),Cost Effectiveness and Simplified Access Interfaces.

High Performance Computing (HPC) applications are increasingly being used in academia and laboratories for scientific research and in industries for business and analytics. Clouds can act as a cost-effective and timely solution (e.g. substitute/addition when Supercomputers are heavily loaded such as in case of academic deadlines) to the needs of some academic and commercial HPC users since they do not involve cluster startup and maintenance costs and cluster creation time. In addition, Cloud provides elastic resources which results in elimination of risks caused by under-provisioning and avoidance of wastage of resources (including energy) resulting from underutilization of computing power in case of over-provisioning.

Cloud provides the benefits of virtualization to HPC community. However, traditionally, Clouds have been designed for running business and web applications, whose resource requirements are different from HPC applications. Unlike web applications, HPC applications typically require low latency and high bandwidth inter-processor communication to achieve best performance. In case of Cloud, presences of commodity interconnect and effect of virtualization result in interconnect becoming a bottleneck for HPC applications. A comparison of three platforms using Converse Ping-Pong benchmark and shows that network performance of Cloud is one to two orders of magnitude worse as compared to Infiniband, which is commonly used interconnect in Supercomputers.

Supercomputers have operating systems and I/O subsystems specifically tailored to match HPC application demands. However, recent efforts towards HPC-optimized Clouds, such as Magellan and Nebula are promising signs for research community.

Increasingly, some academic and commercial HPC users are looking at clouds as a cost effective alternative to dedicated HPC clusters. Renting rather than owning a cluster avoids the up-front and operating expenses associated with a dedicated infrastructure. Clouds offer additional advantages of a) elasticity—on-demand provisioning, and b) virtualization-enabled flexibility, customization, and resource control. Despite these advantages, it still remains unclear whether, and when, clouds can become a feasible substitute or complement to supercomputers.

HPC is performance oriented, whereas clouds are cost and resource-utilization oriented. Furthermore, clouds have traditionally been designed to run business and web applications. Previous studies have shown that commodity interconnects and the overhead of virtualization on network and storage performance are major performance barriers to the adoption of cloud for HPC. While the outcome of these studies paints a rather pessimistic view of HPC clouds, recent efforts towards HPC-optimized clouds, such as Magellan and Amazon's EC2 Cluster Compute, point to a promising direction to overcome some of the fundamental inhibitors.

This paper consists of Introduction about the domain and project description is described, System study of the existing system and the proposed system of the project are illustrated, System design architecture diagram of the project is structured, System implementation of the modules are shown and finally Conclusion and future work were discussed.

2. CLOUD COMPUTING

Cloud computing is a practical approach to experience direct cost benefits and it has the potential to transform a data center from a capital-intensive set up to a variable priced environment.

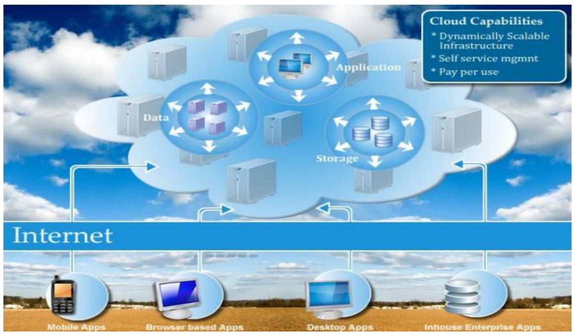

Fig. 1 Conceptual View of Cloud Computing

The idea of cloud computing is based on a very fundamental principle of reusability of IT capabilities. The difference that cloud computing brings compared to traditional concepts of "grid computing", "distributed computing", "utility computing", or "autonomic computing" is to broaden horizons across organizational boundaries.Cloud Providers offer services that can be grouped into three categories as Software as a Service (SaaS), Platform as a Service (Paas), Infrastructure as a Service (Iaas)

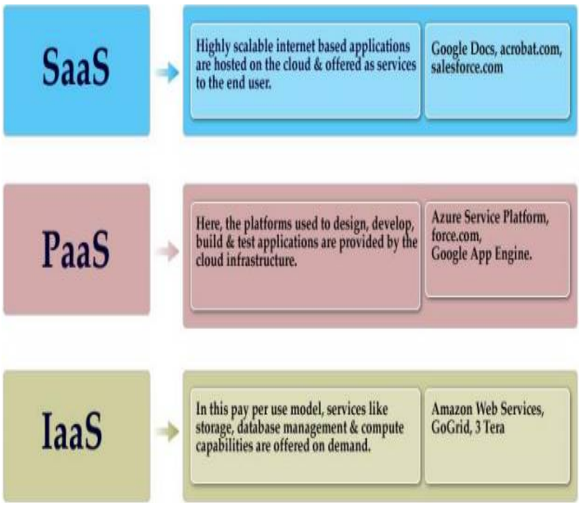

Fig.2 Cloud Models

Enterprises can choose to deploy applications on Public, Private or Hybrid clouds. Cloud Integrators can play a vital part in determining the right cloud path for each organization.Enterprises would need to align their applications, so as to exploit the architecture models that Cloud Computing offers. Some of the typical benefits are listed below: Reduced Cost, Increased Storage and Flexibility.

3. EXISTING SYSTEM

In the existing system**,** HPC clouds rapidly expand the application user base and the available platform choices to run HPC workloads: from in-house dedicated supercomputers, to commodity clusters with and without HPC-optimized interconnects and operating systems, to resources with different degrees of virtualization**.** HPC users and cloud providers are faced with the challenge of choosing the optimal platform based upon a limited knowledge of application characteristics, platform capabilities, and the target metrics such as cost. It results in a potential mismatch between the required and selected resources for HPC application.

Cloudvendors,however,donotyetgivesimilarconsiderationtononfunctionalpropertiesoftheirservices.OftheServiceLevelAgreements(SLAs)typicallyspecifiedwithcloudservices,mostdealwithavailabilityandsomeconsiderreliability.Indeed,nonfunctionalpropertiesmatterbecausetheydetermineservicequality.Forinstance,ifanetworkconnectionbreaksdownorperformancebecomespoor,itmayaffectavailability. Also,ifhardwarefailuresorsoftwarefaultsoccur,theymaydecreasereliability.Still, ifattacksorintrusionshappen,theymayharmsecurity. In short, if nonfunctional properties become problematic,serviceexperiencecanbepoor,whichnegativelyimpactsaprovider'sreputation. There are multiple medium-grained work/data units (objects/tasks) per processors (referred to as over decomposition) and an object needs to wait for a message, control can be asynchronously transferred to another object which has a message to process.

The Disadvantages of Existing System are:

- Increase in Computational granularity.
- Network performance is low.
- Execution time decreases due to increased overlap of computational and communicational.

4. PROPOSED SYSTEM

In the proposed System, a novel heuristics for online application-aware job scheduling in multi-platform environments is proposed in this project to improve performance of HPC applications in cloud by optimizing HPC applications' characteristics for cloud and cloud virtualization mechanisms for HPC. CLOUDQUALis aqualitymodel is to propose forcloudservices.Usability, Availability, Reliability, Responsiveness and Security, Astoitssixqualitydimensions,fiveareobjectiveandnegotiable,whereasusabilityremainssubjectiveandnonnegotiable Cloud-aware HPC execution reduces the penalty caused by the underlying slow physical network in clouds, but it does not address the overhead of network virtualization. Lightweight virtualization reduces the overhead of network virtualization by granting VMs native accesses to physical network interfaces. The mapping of HPC jobs refers to select a platform for a job, and scheduling includes mapping and deciding when to execute the job on the chosen platform.

The Advantages of Proposed System are:

- Better turnaround time for job.
- Assign job to least loaded platform.
- Good performance in cloud.

5. SYSTEM ARCHITECTURE

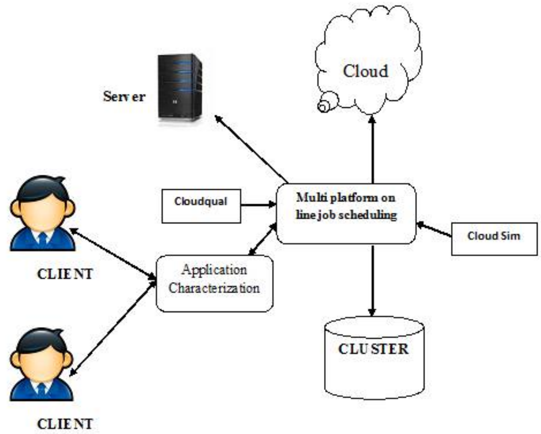

Fig. 3 System Architecture

In Fig.3, the Data Owner as a user upload the content into the cloud server. The Cloudqual analyze the flow of data into the multiplatform based cloud server to check the performance of the cloud server .The CloudSim act as virtualization mechanism to maintain the efficient and availability of the data's into the server. The data's will be stored in the form of clusters into the database server. The Data Consumer will download the data's from the data based cloud server. The CloudSim acts like the proxy to check the valid user and to maintain the scheduling performance in the cloud. The cloud server may be public or private cloud which used in HPC applications.

6. SYSTEM IMPLEMENTATION

The modules designed in the proposed system areCloud Computing Module, Resource Management Module, Virtualization Module, Cloudqual for HPC and Optimizing Cloud for HPC.

6.1 CLOUD COMPUTING MODULE

Cloud computing refers to applications and services offered over the Internet. These services are offered from data centers all over the world, which collectively are referred to as the "cloud."Cloud computing is a movement away from applications needing to be installed on an individual's computer towards the applications being hosted online.

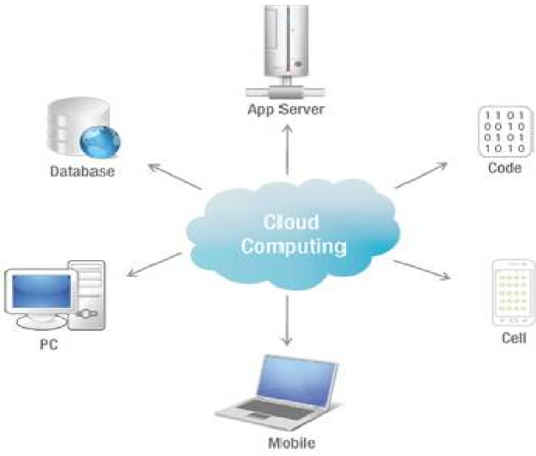

Fig.4 Cloud Computing Model

Cloud resources are usually not only shared by multiple users but as well as dynamically re-allocated as per demand. This can work for allocating resources to users in different time zones.

6.2 RESOURCE MANAGEMENT MODULE

Dynamic resource management has become an active area of research in the Cloud Computing paradigm. Cost of resources varies significantly depending on configuration for using them. Hence efficient management of resources is of prime interest to both Cloud Providers and Cloud Users.

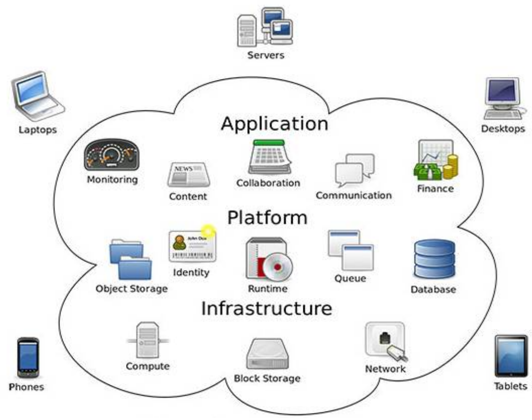

Fig. 5 Resource Management in Cloud

The success of any cloud management software critically de-pends on the flexibility; scale and efficiency with which it can utilize the underlying hardware resources while pro-viding necessary performance isolation. Successful resource management solution for cloud environments, needs to provide a rich set of resource controls for better isolation, while doing initial placement and load balancing for efficient utilization of underlying resources.

6.3 LIGHTWEIGHT VIRTUALIZATION MODULE

Virtualization, in computing, is the creation of a virtual (rather than actual) Version of something, such as a hardware platform, operating system, and a storage device or network resources. Two lightweight virtualization techniques are used, thin VMs configured with PCI pass-through for I/O and containers that is OS-level virtualization.

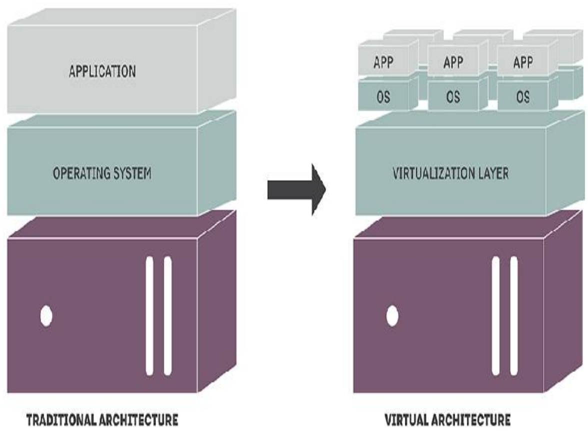

Fig. 6 Traditional and Virtual Machine Architecture

Lightweight virtualization reduces the overhead of network virtualization by granting VMs native accesses to physical network interfaces.VM live migration is a widely used technique for dynamic resource allocation in a virtualized environment. The process of running two or more logical computer system so on one set of physical hardware. Dynamic placement of virtual servers to minimize SLA violations.

6.4 CLOUDQUAL FOR HPC

Usability describes how easy, efficient, and enjoyable the interface to a cloud service is to use, or assesses the ease of invocation if the functionality of a cloud service is exposed as Application Programming Interface (API). Availability is the uptime percentage of cloud services during a time interval, which can be measured by where represents availability; and denote the uptime and the total time of an operational period, respectively. Reliability is the assurance that cloud services are free from hardware failures, software faults, and other defects that could make them break down.Responsiveness is the promptness with which cloud services perform a request during a time interval. Security is the assurance that cloud services are free from viruses, intrusions, spyware, attacks, and other security vulnerabilities that could put them at risk. [1]

6.5 SCHEDULING HPC JOBS IN THE CLOUD

Herein the module mapping refers to selecting a platform for a job, and scheduling includes mapping and deciding when to execute the job on the chosen platform. EJS is defined to capture both, current resource availability and a job's suitability to a particular platform. The Adaptive heuristic optimizes along two dimensions: it balances load across multiple platforms and it matches application characteristics to platforms. CloudSim which is a widely used tool for simulation of scheduling algorithms in a data center or a cloud. In Cloud Sim, a fixed number of VMs are created at the start of simulation, and jobs (cloudlets in CloudSim terminology) can be submitted to these VMs.

7. CONCLUSION

Clouds can successfully complement supercomputers, but using clouds to substitute supercomputers is infeasible. Bursting to cloud is also promising. We have shown that by performing multi-platform dynamic application-aware scheduling, a hybrid cloud-supercomputer platform environment can actually outperform its individual constituents. By using an underutilized resource which is "good enough" to get the job done sooner, it is possible to get better turnaround time for job (user perspective) and improved throughput (provider perspective). Another potential model for HPC in cloud is to use cloud only when there is high demand (cloud burst). My evaluation showed that application-agnostic cloud bursting (e.g. BestFirst heuristic) is unrewarding, but application-aware bursting is a promising research direction. More work is needed to consider other factors in multi-platform scheduling: job quality of service (QoS) contracts, deadlines, priorities, and security.

8.FUTURE WORK

In future work research is required in cloud pricing in multi-platform environments. To evaluate and characterize applications with irregular parallelism and dynamic datasets. The security mechanism is to be implemented in the multiplatform environment.Further development includes the study of a user collaboration based method for deployment. In addition, online optimization is import schedule tasks on IaaS cloud system. Charm mechanism is to be implemented in order to improve the cost efficiency in the cloud. The cost efficient has to be maintained to identify the better cloud.

9. REFERENCES

[1] Avetisyan A.I, R. Campbell, I. Gupta, M. T. Heath, S. Y. Ko, G. R. Ganger, M. A. Kozuch, D. O'Hallaron, M. Kunze, T. T. Kwan, and others, (2010), "Open Cirrus: A Global Cloud Computing Testbed".

[2] Berman .F, A. Chien, K. Cooper, J. Dongarra, I. Foster, D. Gannon, L. Johnsson, K. Kennedy, C. Kesselman, J. Mellor-Crumme, (2001), "The Grads Project: Software Support for High-Level Grid Application Development".

[3] Bhatele .A, S. Kumar, C. Mei, J. C. Phillips, G. Zheng, and L. V. Kale, (2008) "Overcoming Scaling Challenges In Biomolecular Simulations Across Multiple Platforms".

[4] Evangelinos .C and C. N. Hill, (2008) "Cloud Computing for Parallel Scientific HPC Applications: Feasibility of Running Coupled Atmosphere Ocean Climate Modelson Amazon's EC2".

[5]Fan .P, Z. Chen, J. Wang, Z. Zheng, and M. R. Lyu, (2012),"Topology Aware Deployment of Scientific Applications in Cloud Computing".

[6] Gupta .A and D. Milojicic, (2011), "Evaluation of HPC Applications on Cloud".

[7] Gupta .A, L. Kale, D. Milojicic, P. Faraboschi, and S. Balle, (2013), "HPC Aware VM Placement in Infrastructure Clouds".

[8] Gupta .A, L. V. Kal_e, D. S. Milojicic, P.Faraboschi, R. Kaufmann, V. March, F. Gioachin, C. H. Suen, and B.-S. Lee, (2013), "The Who, What, Why, and How of HPC Applications in the Cloud".

[9] Intel Corporation, (2011), Intel(r) Virtualization Technology for Directed I/O. Tech. Rep. D51397-006, (2011) [Online].

[10] Iosup .A, S. Ostermann, M. N. Yigitbasi, R.Prodan, T. Fahringer, and D. H. J. Epema, (2011), "Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing".

[11] Jetley .P, F. Gioachin, C. Mendes, L. V. Kale, and T.R.Quinn, (2008), "Massively Parallel Cosmological Simulations with Changa".

[12] Kale .L and S. Krishnan, (1993), "CHARM++: A Portable Concurrent Object Oriented System Based on C++".

[13] Kivity, A. Y. Kamay, D. Laor, U. Lublin, and A. Liguori, "kvm: the Linux virtual machine monitor".

[14]Mehrotra .M, Djomehri .J, S. Heistand, R. Hood, H.Jin,A. Lazanoff, S. Saini, and R. Biswas, (2012), "Performance Evaluation of Amazon EC2 for NASA HPC Applications".

[15] Napper .J and P.Bientinesi, (2009), "Can Cloud Computing reach the Top500?".

[16] Nurmi .D, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov, (2009), "Cluster Computing and the Grid, 2009.