



Performance Comparison of Resource Techniques in Cloud Computing

Sanjeev Dhawan¹, Nitin Kaushik²

¹Assistant Professor of Computer Science & Engineering,

²Research Scholar, M. Tech. (Computer Engineering),

^{1,2}Department of Computer Science & Engineering,

University Institute of Engineering & Technology,

Kurukshetra University, Kurukshetra-136 119, Haryana, INDIA

Abstract: Cloud computing is revolutionizing the IT industry by enabling them to offer access to their infrastructure and application services on a subscription basis. As a result, several enterprises including IBM, Microsoft, Google, and Amazon have started to offer different cloud services to their customers. Due to the vast diversity in the available cloud services from the customer's point of view, it has become difficult to decide whose services they should use and what the basis for their selection is. Currently, there is no framework that can allow customers to evaluate cloud offerings and rank them based on their ability to meet the user's Quality of Service (QoS) requirements. A huge amount of data moves from user to host and hosts to user in the cloud environment. Based on the above two considerations, how to select appropriate host for accessing resources and creating a Virtual Machine (VM) to execute applications so that execution becomes more efficient and access cost becomes low are the challenging tasks. In this paper, a host selection model based on minimum network delay is proposed. The objective is to minimize propagation time of input and output data by selecting nearest host into the network and finally it minimizes the execution time of cloudlet.

Keywords: Cloudlet; Cloud Service; Delay; Resource; Virtual Machine (VM)

I. Introduction

Cloud computing is one of the major technologies predicted to revolutionize the future of computing. The model of delivering IT as a service has several advantages. It enables current businesses to dynamically adapt their computing infrastructure to meet the rapidly changing requirements of the environment. Perhaps more importantly, it greatly reduces the complexities of IT management, enabling more pervasive use of IT. Further, it is an attractive option for small and medium enterprises to reduce upfront investments, enabling them to use sophisticated business intelligence applications that only large enterprises could previously afford. Cloud-hosted services also offer interesting reuse opportunities and design challenges for application developers and platform providers. Cloud computing has, therefore, created considerable excitement among technologists in generation. Cloud computing is basically delivering computing at the Internet scale. Compute, storage, networking infrastructure as well as development and deployment platforms are made available on-demand within minutes. Sophisticated futuristic applications such as those described in the earlier sections are made possible by the abstracted, auto-scaling compute platform provided by cloud computing.

II. Cloud Service Models

Once a cloud is established, how its cloud computing services are deployed in terms of business models can differ depending on requirements. The primary service models being deployed are commonly known as:

A. Software as a Service (SaaS)

Consumers purchase the ability to access and use an application or service that is hosted in the cloud. A benchmark example of this is Salesforce.com, as discussed previously, where necessary information for the interaction between the consumer and the service is hosted as part of the service in the cloud.

Also, Microsoft is expanding its involvement in this area, and as part of the cloud computing option for Microsoft® Office 2010, its Office Web Apps are available to Office in the cloud as it is to have it on the premise. In this context, Communications as a Service (CaaS) could be seen as a subset of SaaS volume licensing customers and Office Web Apps subscriptions through its cloud-based Online Services.

B. Platform as a Service (PaaS)

Consumers purchase access to the platforms, enabling them to deploy their own software and applications in the cloud. The operating systems and network access are not managed by the consumer, and there might be constraints as to which applications can be deployed.

C. Infrastructure as a Service (IaaS)

Consumers control and manage the systems in terms of the operating systems, applications, storage, and network connectivity, but do not themselves control the cloud infrastructure.

Also, known are the various subsets of these models that may be related to a particular industry or market. CaaS is one such subset model used to describe hosted IP telephony services. Along with the move to CaaS is a shift to more IP-centric communications and more SIP trunking deployments. With IP and SIP in place, it can be as easy to have the PBX in the cloud as it is to have it on the premise. In this context, CaaS could be seen as a subset of SaaS.

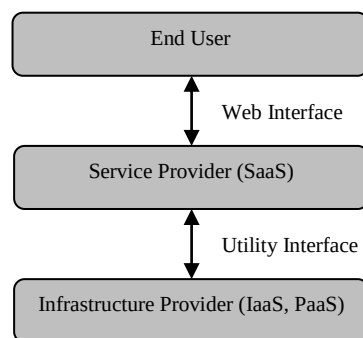


Figure 1: Cloud Computing Business Model

III. Related Work

Paul Martinaitis *et al.* [1] thoroughly investigated the development of a template stream-aware scheduling component. The motivation for developing such a template as the basis for future work to design more sophisticated mechanism than the default Gridbus round robin scheduler, which operates in a queue-oriented manner, iterating over the jobs in the queue. Such an arrangement is reasonable in the usual Gridbus environment, where jobs are independent. However, in a stream computation context, each job corresponds to a component which is almost certainly linked to at least one other component. It is therefore highly desirable that stream-aware schedulers are able to absorb the entire body of information regarding a stream. In the area of resource allocation, Karthik Kumar *et al.* [2] explained the need of IaaS for renting computing resources over the Internet. The user can select from different types of computing resources based on the requirements. For example, Amazon's EC2 Cloud service provides different options for selecting resources as shown in Table I. The user can rent an arbitrarily large number of resources: this is called scalable computing or elasticity, since the number can be scaled dynamically to meet the requirements.

Kun Li *et al.* [3] experienced a rapid development of cloud computing both in academia and industry. They promoted this model by the business rather than academic which determines its focus on user applications. This technology aims to offer distributed, virtualized, and elastic resources as utilities to end users. It has the potential to support full realization of 'computing as a utility' in the near future [1]. With the support of virtualization technology [2, 3, 8], cloud platforms enable enterprises to lease computing power in the form of virtual machines to users. Because these users may use hundreds of thousands of virtual machines (VMs) [4], it is difficult to manually assign tasks to computing resources in clouds [5, 6, 11]. In a similar area, Enda Barrett *et al.* [4] discussed the scheduling of workflow applications to involve the mapping of individual workflow tasks to computational resources based on a range of functional and non-functional quality of service requirements. Workflow applications such as scientific workflows often require extensive computational processing and generate significant amounts of experimental data. The emergence of cloud computing has introduced a utility type market model, where computational resources of varying capacities can be procured on demand, in a pay-per-use fashion.

Boonyarith Saovapakhiran *et al.* [5] used heterogeneous computing platforms such as Grid and Cloud computing for resource management to check their global performance. Under the assumption of jobs comprised of subtasks forming DAG jobs, they focused on how to increase utilization and achieve near-optimal throughput performance on heterogeneous platforms. They analyzed and proposed algorithm which can be analytically derived to aggregate multiple jobs using good scheduling to maximize the throughput. Consequently, its limit is asymptotically converging to a certain value and can be written in the form of the service time of subtasks. Furthermore, they showed how to explicitly compute the optimal throughput of computing system from a complex scheduling problem. In addition, they derived a simple super-job scheduling and showed that its performance in term of throughput is better than the well-known Heterogeneous Earliest-Finish-Time (HEFT) algorithm.

Hsu and Thinn [6] investigated the deployment of Cloud computing on a large set of virtualized computing resources in different infrastructures and various development platforms. One of the significant issues in cloud computing system is the scheduling of virtual resources and virtual machines (VMs). To address this issue, they proposed an efficient approach for virtual machines scheduling in VM management also called EVMSA (Efficient Virtual Machines Scheduling Algorithm) that provides the effective and efficient resource allocation.

The proposed approach is going to test the evaluation on open source private cloud architecture. The major contribution of this paper is to improve the resource utilization such as CPU, memory, disk and to minimize the turnaround time of VMs.

IV. Plan Implementation

A. Problem Formulation

The problem of resource selection in distributed environment has received lots of attention in last few years. In many previous works, resource selection refers to the selection of computational resource in grid environment. In [4], the authors presented a resource selection model using decision theory for selecting the best machine to run a task. In [5], they proposed an algorithm for resource selection problem of computational grids, based on the resource availability prediction using frequent workload patterns. Recently, with the rapid development of data intensive computing, many researchers turned their attention to resource selection of data-intensive environment, such as data grid [6, 8, 9, 10, 11].

In this research, our plan is to derive an algorithm to select data host and compute host for accessing resources and creating a virtual machine (VM) to execute applications to make execution efficiency high which can be efficiently used by one or multiple users. For instance, the Cloud Service Broker [7] developed as part of the Cloudbus project, mediates access to distributed resources by discovering resources, scheduling tasks, monitoring and collating results. A user portal provides a user-friendly environment to the end-users (most often Scientists) to compose, submit and monitor workflow applications. A workflow editor forms the graphical user-interface at the portal to facilitate the composition and visualization of workflows in order to schedule middleware together form a Workflow Management System (WfMS).

B. Plan Work for Algorithm

(i) Resource Selection Technique

In this technique we select an appropriate data center among all other data centers, such that the transfer time to access replicas from other data centers is minimum. Then create a VM on this data center selected finally for the task and submit the task to the VM. Here the selection of the final data center is mainly base on the proximity of the data centers. The following algorithm gives the steps of getting data center which is best for creating VM on it and having shortest path.

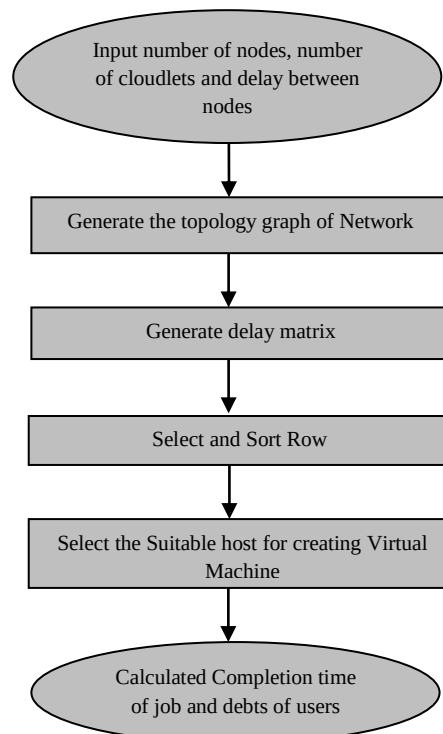


Figure2: Flow steps used for data center for creating VM with shortest path.

V. Conclusion

Cloud computing has become an important paradigm for outsourcing various IT needs of organizations. Currently, there are many Cloud providers who offer different Cloud services with different price and performance attributes. With the growing number of Cloud offerings, even though it opens the chance to leverage the virtually infinite computing resources of the Cloud, it has also becomes challenging for Cloud

customers to find the best Cloud services which can satisfy their QoS requirements in terms of parameters such as performance and security. To choose appropriately between different Cloud services, customers need to have a way to identify and measure key performance criteria that are important to their applications.

VI. References

- [1] Paul Martinaitis, Craig Patten and Andrew Wendelborn "Remote Interaction and Scheduling Aspects of Cloud Based Streams", 2009 IEEE.
- [2] Karthik Kumar, Jing Feng, Yamini Nimmagadda, and Yung-Hsiang Lu "Resource Allocation for Real-Time Tasks using Cloud Computing", IEEE, 2011, 978-1-4577-0638-7 /11
- [3] Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Dan Wang, "Cloud Task scheduling based on Load Balancing Ant Colony Optimization" 2011 Sixth Annual China Grid Conference, IEEE, 2011, 978-0-7695-4472-4/11.
- [4] Enda Barrett, Enda Howley, Jim Duggan, "A Learning Architecture for Scheduling Workflow Applications in the Cloud", 2011 Ninth IEEE European Conference on Web Services, IEEE, 2011, 978-0-7695-4536-3/11.
- [5] Boonyarith Saovapakhiran, George Michailidis, Michael Devetsikiotis, "Aggregated-DAG Scheduling for Job Flow Maximization in Heterogeneous Cloud Computing", IEEE, 2011, 978-1-4244-9268-8/11.
- [6] Hsu Mon Kyi, Thinn Thu Naing, "An Efficient Approach for Virtual Machines Scheduling on a Private Cloud Environment", IEEE, 2011, 978-1-61284-159-5/11.
- [7] V. Nelson, V. Uma, "Semantic based Resource Provisioning and Scheduling in Inter-cloud Environment", IEEE, 2012, 978-1-4673-1601-9/12.
- [8] Jinhua Hu, Jianhua Gu, Guofei Sun Tianhai Zhao, "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment", IEEE, 2010, 978-0-7695-4312-3/10.
- [9] Shu-Ching, Wang, Kuo-Qin, Yan, Shun-Sheng, Wang, Ching-Wei, Chen, "A Three-Phases Scheduling in a Hierarchical Cloud Computing Network", IEEE Third International Conference on Communications and Mobile Computing, 2011, 978-0-7695-4357-4/11.
- [10] Praveen K. Gupta, Nitin Rakesh, "Different Job Scheduling Methodologies for Web Application and Web Server in a Cloud Computing Environment", 2010 IEEE Third International Conference on Emerging Trends in Engineering and Technology.
- [11] Saurabh Kumar Garg, Chee Shin Yeo, Arun Anandasivam, Rajkumar Buyya "Energy-E_client Scheduling of HPC Applications in Cloud Computing Environments", IEEE, 2009.