RESEARCH PAPER | Computer Science | Volume : 5 | Issue : 6 | June 2015 | ISSN - 2249-555X

A Study on Cloud Computing Models and Resource Management

KEYWORDS: Cloud computing, Distributed computing, Scheduling, Resource Management.

Basant Namdeo

IIPS, DAVV

ABSTRACT *Cloud computing is useful in commercial as well as scientific domain. It encompasses key features of distributed computing, mainframe, virtualization, utility computing and many more computing paradigms. Understanding of proper classification of cloud computing services, deployment models and resource management algorithms and policies are needed to exploit it. This paper focuses on survey based on scheduling and resource management in cloud computing and will be useful for students, researchers, and commercial cloud computing players.*

INTRODUCTION

Cloud computing is a new paradigm of computing. It provides dynamically scalable and often virtualized resources as services over the Internet [1]. Clouds can be classified from deployment point of view as : Public, Private and Hybrid Clouds. Public Cloud offers services to general public and they are open for all, while Private Clouds are used by a single organization. Hybrid Clouds, on the other hand, offer a combination of private and public Clouds. These can be shown as in the figure 1.

There are three Cloud services models: 1) SaaS (Software as a Service), 2) PaaS (Platform as a Service) and 3) IaaS (Infrastructure as a Service). These are shown as in figure 2.

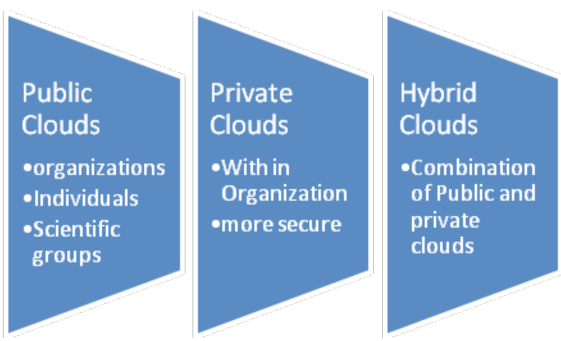

Figure 1: Cloud Depolyment Models

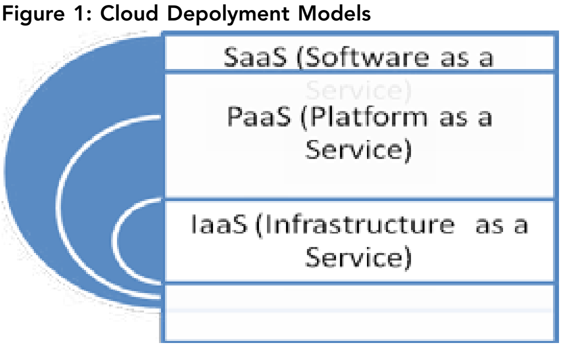

Figure 2: Cloud Service Models

IaaS offers an infrastructure in the form of virtual machines (VMs). These VMs provides resources (storage, databases, among others), as a service. PaaS offers an environment oriented to the development, testing, and deployment and hosting of applications. SaaS offers a set of applications in the form of web services running over Internet. This paper presents a study of PaaS, SaaS and IaaS providers, with the aim of helping users, scientific groups, students to understand the cloud computing and its models that best suits their needs and applications.

Rest of the paper is as follows: section 2 discusses related works of Cloud computing. In Section 3, conclusion and our contribution is presented.

BACKGROUND

Huge amount of literature is available on cloud computing. Various research groups are exploring the ways to use cloud computing as next generations paradigm shift.

Sotomayor et al. presented a model for predicting various runtime overheads involved in using virtual machines, that allows efficient support advance reservations in [2].

Shrivastava et al. proposed two algorithms 1) Starvation-Removal and 2) AR-to-BE Conversion to solve problems related to resource management. They demonstrated experimentally results of the proposed algorithms to stop starvation of BE leases for resources and effectively improve request acceptance rate [3]. Starvation removal algorithm with improvements were also proposed by [4].

In [5] CBUD Micro, a performance evaluation tool was presented that can be used at both cloud host and consumer side for resource management by supporting scheduling decisions. They also described the vision and architecture of CBUD Micro in detail and and the way in which core components were implemented.

In [6] a new leasing policy named CRI (Consumer Rating Index) and an algorithm for prioritizing consumers on the basis of CRI scores was proposed to manage the cloud host resources properly. This policy and algorithm can be used for efficient functioning at cloud hosts side was also presented.

In [7] mEDF (Modified Earliest Deadline First) algorithm and leasing policy for deadline driven resource management was introduced. This algorithm takes care of minimum capacity expenses at cloud host side. mEDF can be used for deadline driven scheduling with minimizing response time and completion time.

COMMA (Cost Oriented, Market and

Volume : 5 | Issue : 6 | June 2015 | ISSN - 2249-555X

Migration Aware) for maintaining a balance between cost, profit and migrations of VM during peak loads was introduced in [8]. This paper showed that COMMA leasing policy and algorithm are effective and can be used for making profit and maintain QoS in a cloud environment.

Cloud service negotiation techniques were presented in [9]. Authors also presented a survey of challenges and current state of resource negotiation. They showed the different agent based methods for cloud negotiation. The focus was on most advanced agent cloud coordinator that will be able to deliver Quality of Service (QoS) for cloud provider so that it allows an increase in performance, reliability and scalability of applications.

Foster et al. presented a study in [10] that strives to compare and contrast Cloud Computing with Grid Computing from various angles and give insights into the essential characteristics of both.

Beloglazov et al. conducted a survey of research in energy-efficient computing and propose: (a) architectural principles for energy-efficient management of Clouds; (b) energy-efficient resource allocation policies and scheduling algorithms considering QoS expectations and power usage characteristics of the devices; and (c) a number of open research challenges, addressing which can bring substantial benefits to both resource providers and consumers. [11]

Buyya presented vision, challenges, and architectural elements for energy-efficient management of Cloud computing environments. they focused on the development of dynamic resource provisioning and allocation algorithms that consider the synergy between various data center infrastructures (i.e., the hardware, power units, cooling and software), and holistically work to boost data center energy efficiency and performance [12].

Pandey et al. presented a particle swarm optimization (PSO) based heuristic to schedule applications to cloud resources that takes into account both computation cost and data transmission cost. They experimented with a workflow application by varying its computation and communication costs. They compared the cost savings when using PSO and existing 'Best Resource Selection' (BRS) algorithm. Their results show that PSO can achieve: a) as much as 3 times cost savings as compared to BRS, and b) good distribution of workload onto resources [13].

Hoffa explored the use of cloud computing for scientific workflows, focusing on a widely used astronomy application-Montage. The approach was to evaluate from the point of view of a scientific workflow the tradeoffs between running in a local environment, if such is available, and running in a virtual environment via remote, wide-area network resource access [14].

Zhang presented a survey of cloud computing, highlighting its key concepts, architectural principles, state-of-the-art implementation as well as research challenges. The aim of his paper was to provide a better understanding of the design challenges of cloud computing and identify important research directions in this increasingly important area [15].

Beloglazov et al. proposed an energy efficient resource management system for virtualized Cloud data centers that reduces operational costs and provides required Quality of Service (QoS). Energy savings were achieved by continuous consolidation of VMs according to current utilization of resources, virtual network topologies established between VMs and thermal state of computing nodes [16].

Conclusion

Cloud computing deployment models and service models affects resource management and leasing policies. Various algorithms and policies for the same are discussed in this paper. New resource management algorithms and leasing policies can also be developed by just improving available features in surveyed literature from this paper.

REFERENCE

[1] Armbrust, Michael, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee et al. "A view of cloud computing." Communications of the ACM 53, no. 4 (2010): 50-58. | [2] Sotomayor, Borja, Rubén S. Montero, I. Martín Llorente, and Ian Foster. "Resource leasing and the art of suspending virtual machines." In High Performance Computing and Communications, 2009. HPCC'09. 11th IEEE International Conference on, pp. 59-68. IEEE, 2009. | [3]Shrivastava, Vivek, and D. S. Bhilare. "Algorithms to Improve Resource Utilization and Request Acceptance Rate in IaaS Cloud Scheduling." International Journal of Advanced Networking and Applications 3, no. 05 (2012): 1367-1374. | [4] Kurdi, Heba, Ebtesam Aloboud, Sarah Alhassan, and Ebtehal T. Alotaibi. "An Algorithm for Handling Starvation and Resource Rejection in Public Clouds." Procedia Computer Science 34 (2014): 242-248. | [5] Shrivastava, Vivek, and D. S. Bhilare. "CBUD Micro: A Micro Benchmark for Performance Measurement and Resource Management in IaaS Clouds." International Journal of Emerging Technolgy and Advanced Engineering 3, no. 11 (2013): 433-437. | [6] Shrivastava, Vivek, and D. S. Bhilare. "CRI: A Novel Rating Based Leasing Policy and Algorithm for Efficient Resource Management in IaaS Clouds." International Journal of Computer Science and Information Technologies 3 (2014): 4226-4230. | [7] Shrivastava, Vivek, and D. S. Bhilare. "mEDF: Deadline Driven Algorithm for Minimizing Response Time and Completion Time in IaaS Clouds." Computing 4: 5. | [8] Shrivastava, Vivek, and D. S. Bhilare. "COMMA: A Cost Oriented, Market and Migration Aware Leasing Policy and Algorithm in IaaS Clouds." In Proceedings of the 2014 International Conference on Information and Communication Technology for Competitive Strategies, p. 52. ACM, 2014. | [9] Ariya, T. K., Christophor Paul, and S. Karthik. "Cloud Service Negotiation Techniques." computer 1, no. 8 (2012). | [10] Foster, Ian, Yong Zhao, Ioan Raicu, and Shiyong Lu. "Cloud computing and grid computing 360-degree compared." In Grid Computing Environments Workshop, 2008. GCE'08, pp. 1-10. Ieee, 2008. | [11] Beloglazov, Anton, Jemal Abawajy, and Rajkumar Buyya. "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing." Future generation computer systems 28, no. 5 (2012): 755-768. | [12] Buyya, Rajkumar, Anton Beloglazov, and Jemal Abawajy. "Energy-efficient management of data center resources for cloud computing: A vision, architectural elements, and open challenges." arXiv preprint arXiv:1006.0308 (2010). | [13] Pandey, Suraj, Linlin Wu, Siddeswara Mayura Guru, and Rajkumar Buyya. "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments." In Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on, pp. 400-407. IEEE, 2010. | [14] Hoffa, Christina, Gaurang Mehta, Timothy Freeman, Ewa Deelman, Kate Keahey, Bruce Berriman, and John Good. "On the use of cloud computing for scientific workflows." In eScience, 2008. eScience'08. IEEE Fourth International Conference on, pp. 640-645. IEEE, 2008. | [15] Zhang, Qi, Lu Cheng, and Raouf Boutaba. "Cloud computing: state-of-the-art and research challenges." Journal of internet services and applications 1, no. 1 (2010): 7-18. | [16] Beloglazov, Anton, and Rajkumar Buyya. "Energy efficient resource management in virtualized cloud data centers." In Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 826-831. IEEE Computer Society, 2010. | |