



Comparative Review of Scheduling and Migration Approaches in Cloud Computing Environment

MS. ALANKRITA AGGARWAL¹, RAJJU²

¹Assistant Professor, ²Research Scholar M.tech,

Department of Computer Science & Engineering, HCTM Technical Campus, Kaithal, India.

Abstract: Cloud computing is one of most essential and popular distributed environment that resides all the services and product at some centralized location. The Task management is the key role in cloud computing systems task scheduling problems are main which relate to the efficiency of the whole cloud computing facilities. Scheduling in cloud means selection of best suitable resources for task execution. An effective scheduling that give the better resource utilization to the process so that user can get better services is also a challenge in cloud computing environment. Several methods have be given to solve the problem of effective scheduling in cloud computing. In this paper we present a study on different scheduling algorithms for effective resource utilization and migration.

I. INTRODUCTION

Cloud computing is an emerging technology that combines a large amount of computer resources in to a virtual place so as to provide an on-demand computing facility to users. The cloud system will provide its computing resources to users according to the user request, in which the amount and capacity of computer resources are highly configurable [5]. Definition of cloud computing provided by the US National institute of standards and technology (NIST). "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, application, and services) that can be rapidly provisioned and released with minimal management effort or services provider interaction." [12]. One of the key promises of cloud computing is an enormous level of flexibility for scaling up or scaling down software and hardware infrastructure without huge upfront investments. Hence, it is expected that any cloud-based infrastructure should have three characteristics: ability to acquire transactional resources on demand, resources publication through a single provider, and mechanisms to bill users on the basis of resource utilization [12]. It includes any subscription based or pay-per-use, where it can obtain network storage space and computer resources. Pay-as-you-go is the means of payment of cloud computing only paying for the actual consumption of resource [4]. Because of its elasticity, cloud computing is suitable for the execution of complicated computational task and scientific simulation, which may require a spike of computational resources, e.g., computing nodes and storages [5]. Cloud computing service models can generally be classified into three categories: Infrastructure as a service (IaaS), platform as a service (PaaS), Software as a Service (SaaS) [12]. Scalability is a prominent quality for all these categories of system. There are mainly three deployment models of cloud computing: private cloud, public cloud and hybrid cloud [9]. Virtualization in cloud refer to multi-layer hardware platforms, operating systems, storage devices, network resources etc., the first prominent feature of virtualization is the ability to hide the technical complexity from users, so it can improve independence of cloud services. Secondly, physical resources can be efficiently configured and utilized, considering that multiple application are run on the same machine. Thirdly, quick recovery and fault tolerance are permitted [4].

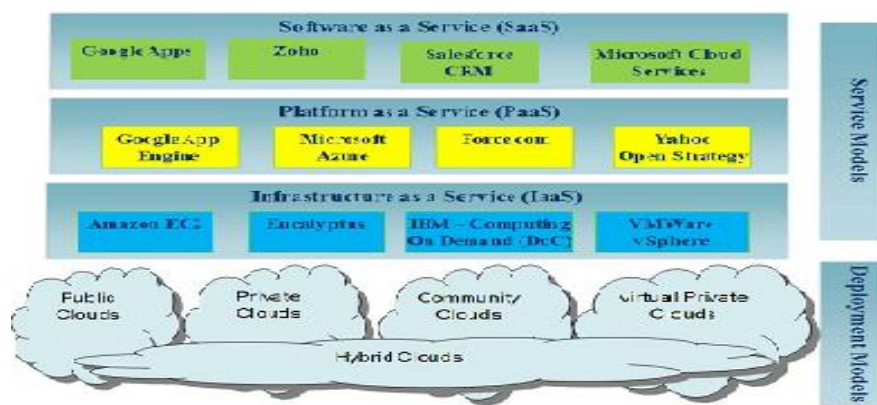


Fig. 1: Service and Deployment model

Task Scheduling is the most effective and important task for a computer system that basically decides the order of the process execution when different processes are kept in a queue. Scheduling is defined as a task that defines the execution of the system under the time constraint. Time is actually the order of the execution of the processes. Co-scheduling defines a process of execution of more than one process simultaneously on different processors. In a cloud based system, this scheduling approach is quite common. Task scheduling involves process of assigning of task to available resources on the basis of task characteristics and requirement. Task scheduling is a necessary process as it schedules task on the basis of the requirement of user. One requirement can be to minimize completion time to complete execution within deadline by reducing the monetary cost [2]. According to a simple classification, job scheduling algorithms in cloud computing can be categorized into two main groups; Batch mode heuristic scheduling algorithms (BMHA) and online mode heuristic algorithms.

II. EXISTING SCHEDULING AND MIGRATION METHODS

A Multiple-objective Workflow Scheduling Framework for Cloud Data Analytics: Orachum Udomkasemsub, Li Xiaorong, Tiranee Achalakul [5] proposed a workflow scheduling framework that can efficiently schedule series workflow with multiple objectives onto a cloud system. In this Artificial Bee Colony (ABC) is used to create an optimized scheduling plan. Conflicts among objectives can also be resolved by using Pareto-based technique. The proposed framework aims to deal with complex workflows for data analytics by handling various workflow structures, short term and long term scheduling, dependency mapping on a heterogeneous computing environment, and most importantly multiple objectives that may contradict one another. Experimental result show that this method is able to reduce 57% cost and 50% scheduling time within a similar makespan of HEFT/LOSS for a typical scientific workflow like Chimera-2.

Deadline and Cost based Workflow Scheduling in Hybrid Cloud: Nitish Chopra, Sarbjeet Singh [2] defined a algorithm for cost optimization by deciding which resources should be taken on lease from public cloud to complete the workflow execution within deadline. The author develop a level based scheduling algorithm which executes task level wise and it uses the concept of sub-deadline which is helpful in finding best resources on public cloud for cost saving and also complete workflow execution within deadlines. In this workflow are represented by DAG. Defined work focuses on scheduling in hybrid cloud. In this deadline and monetary cost are considers as the main factor for scheduling tasks and resources in hybrid cloud environment.

Author define the sub-deadline for each task which finds from the application deadline using a percentage method:

Sub-deadline for the task = percentage of share of task in application * deadline of the application + deadline of task's predecessor.

Author compared it with the min-min algorithm it shows that makespan of proposed level based algorithm is almost double than min-min makespan but the cost incurred is around 3 times lesser then the cost incurred in min-min.

Pre-emptive Scheduling of On-line Real Time Services With Task Migration for Cloud Computing: R. Santhosh, T. Ravichandran [4] defined a new scheduling approach to focus on providing a solution for online scheduling problem of real-time tasks using "Infrastructure as a Service" model offered by cloud computing. This scheduling method sensibly aborts the task when it misses its deadline. In this paper a preemptive online scheduling with task migration algorithm is proposed in order to minimize the response time and to improve the efficiency of the task. Whenever a task misses its deadline, it will be migrated the task to another virtual machine. This will improve the overall system performance and maximizes the total utility. Simulation result shoes that the proposed algorithm can significantly perform the EDF and Non Preemptive scheduling algorithm.

An Efficient Multi Queue Job Scheduling for Cloud Computing: AV. Karthick, Dr. E. Ramaraj, R. Ganapathy Subramanian [1] : defined a multi queue scheduling algorithm to reduce the cost of both reservation and on-demand plans using the global scheduler. This methodology depicts the concept of clustering the jobs based on burst time. This method overcome the problem of fragmentation during scheduling and reduces the starvation with in the process. The proposed MQS method gives more important to select job dynamically in order to achieve the optimum cloud scheduling problem and hence it utilize the unused free space in an economic way. Simulation result show that MQS algorithm gives better and satisfactory result when compared to the traditional algorithm.

An Energy and Deadline Aware Resource Provisioning Scheduling and Optimization Framework for Cloud Systems: Yue Gao, Yanzhi Wang, Sandeep K. Gupta, Massoud Pedram [3] defined the problem of global operation optimization in cloud computing from the perspective of the cloud service provider (CSP). The goal is to provide the CSP with a versatile scheduling and optimization framework that aims to simultaneously maximize energy efficiency and meet all user deadlines, which is also powerful enough to handle multi-user

large scale workloads in large scale cloud platforms. Two types of workload models have been adopted in cloud computing systems: independent batch requests and task graphs with dependencies. In this paper we model the workloads from multiple users as a collection of disjoint task graphs. As for the cloud platform model, it is fully capable of reflecting server resource capacity and energy efficiency heterogeneities. Server communication bottlenecks are also taken into account. This fine-grained treatment of the hardware resources and user workloads provides opportunities for deadline-oriented application acceleration via parallel execution and global energy cost minimization, but also requires additional effort in admission control, resource provisioning, virtual machine placement and task scheduling. In this paper the author propose "Guided Migrate and Pack" (GMaP) as a unified scheduling and optimization framework for the CSP that addresses these issues in a holistic fashion. GMaP is also flexible in search space sizing and algorithm run time control. Experimental results show that when GMaP is deployed for the CSP, global energy consumption costs improves by over 23% when servicing 30 - 50 users, and over 16% when servicing 60 - 100 users.

III. CONCLUSION

With the emerging of cloud computing, cloud workflow systems are designed to facilitate the cloud infrastructure to support large scale distributed collaborative e-business and e-science applications. In this paper we have analyze various scheduling algorithm. Existing scheduling algorithm gives high throughput and cost effective but they do not consider reliability and availability. So we need algorithm that improves availability and reliability in cloud computing environment.

REFERENCES

- [1] A V. Karthick, "An Efficient Multi Queue Job Scheduling for Cloud Computing", 2014 IEEE world congress on computing and communication technology.
- [2] Nitish Chopra, "Deadline and Cost based Workflow Scheduling in Hybrid Cloud" 2013 IEEE, 978-1-4673-6217-7/13.
- [3] Yue Gao, "An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems" 2013 IEEE, 978-1-4799-1417-3/13.
- [4] R. Santhosh, "Pre-emptive Scheduling of On-line Real Time services With Task Migration for Cloud Computing", 2013 IEEE conference on PRIME.
- [5] Orachum udomkasemsub, "A Multiple- Objective Workflow Scheduling Framework for Cloud Data Analytics" JCSSE 2012, IEEE 978-1-4673-1921-8/12.
- [6] Damien Borgetto, "Energy-efficient and SLA-Aware Management of IaaS Clouds", e-Energy 2012, May 9-11 2012, Madrid, Spain. ACM 978-1-4503-1055-0/12/05.
- [7] Balaji Viswanathan, "Rapid Adjustment and Adoption to MaaS Clouds", Middleware 2012 Industry Track, December 3-7, 2012, Montreal, Quebec, Canada. ACM 978-1-4503-1613-2/12/12.
- [8] Shigeru Imai, Thomas Chestna, Carlos A. Varela, "Elastic Scalable Cloud Computing Using Application-Level Migration", 2012 IEEE/ACM Fifth International Conference on Utility and Cloud computing.
- [9] Kejiang Ye, Xiaohong Jiang, "VC-Migration: Live Migration of Virtual Clusters in the Cloud", 2012 ACM/IEEE 13th International Conference on Grid Computing.
- [10] Hadi Goudarzi, Mohammad Ghasemazar, and Massoud Pedram, "SLA-based Optimization of Power and Migration Cost in Cloud Computing", 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing.
- [11] Michael Menzel, "CloudGenius: Decision Support for Web Server Cloud Migration", WWW 2012, April 16-20, 2012, Lyon, France. ACM 978-1-4503-1229-5/12/04.
- [12] Muhammad Ali Babar, "A Tale of Migration to Cloud Computing for Sharing Experiences and Observations", Waikiki, Honolulu, HI, USA. ACM 978-1-4503-0582-2/11/05.
- [13] Sumit Kumar Bose, "CloudSpider: Combining Replication with Scheduling for Optimizing Live Migration of Virtual Machines Across Wide Area Networks", 2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing 978-0-7695-4395-6/11© 2011 IEEE.
- [14] Anton Beloglazov, "Energy Efficient Resource Management in Virtualized Cloud Data Centers", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing 978-0-7695-4039-9/10© 2010 IEEE.
- [15] J. Brandt, "Using Cloud Constructs and Predictive Analysis to Enable Pre-Failure Process Migration in HPC Systems", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing 978-0-7695-4039-9/10© 2010 IEEE.
- [16] Anton Beloglazov, "Energy Efficient Allocation of Virtual Machines in Cloud Data Centers", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing 978-0-7695-4039-9/10 © 2010 IEEE.
- [17] Takahiro Hirofuchi, "Enabling Instantaneous Relocation of Virtual Machines with a Lightweight VMM Extension", 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing 978-0-7695-4039-9/10© 2010 IEEE.
- [18] Kento Sato, "A Model-Based Algorithm for Optimizing I/O Intensive Applications in Clouds using VM-Based Migration", 9th IEEE/ACM International Symposium on Cluster Computing and the Grid 978-0-7695-3622-4/09© 2009 IEEE.