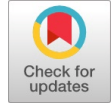


Research Scope and Tools for Workflow Scheduling in Cloud Environment



Menaka M, Sairabanu

Abstract: Workflow is a prototype that executes the behavior of scientific and engineering applications for which the sequence of tasks needs to be automated based on the input parameters specified. Difficulties arise for CSPs primarily during the execution environment due to its direct impact on various QoS parameters. Existing workflow scheduling techniques has research focus with dimensions that include undetermined demands, task failure, delay cost, ambiguous deadlines, bandwidth, cache inclusion, scheduler policies, VM cycles, QoS impact, OS support, fault-tolerant and virtualization level. The comparative analysis made in this paper for workflow scheduling strategies and tools used in the cloud environment towards QoS parameters would have a greater impact on industry task automation which in turn provides the researchers to set their objectives and tools to be used in-order to bring forth new approaches and solutions. This paper can be extended considering the research scope given in the comparative analysis and the data size pertaining to each task in workflow and data migration issues which cannot be done without cost, legal policies and technical issues.

Index Terms: Cloud Computing; Cloud Service Providers(CSP); QoS(Quality of Service); Virtualization; Workflow Scheduling.

I. INTRODUCTION

Cloud Computing [24] pivotal through virtualization technology plays a role in satisfying the on-demand characteristic of the cloud environment. The virtualization layer could be applied to the machine via hypervisor tools which are already available in the market. There exists various open source and licensed tools that gives support for the cloud environment. The major organizations that provide various cloud services [23] are Google, Amazon, IBM, Microsoft etc., Growing number of applications on the mobile platform has brought many challenges, and they are being undertaken to provide seamless services to the consumers. The services were offered either directly through CSP or cloud broker. Each organization develops their SLAs for each service they offer considering the marketing requirements and utilization demand. Workflow scheduling [26] requires a special attention for CSPs in-order to improve their automation process to be effective and efficient, considering all the possibilities supporting dynamic type with QoS met.

The collection of policies to regulate the order of task to be executed by a compute engine is referred to as scheduling process. Scheduling in the cloud environment play a major role for CSP in terms of cost, policies defined, VM type, service type, QoS parameters, load-balancing, energy efficiency, data migration, replication factor, availability, execution, application type, pricing model, constraints enforced. Any CSP has to provide a strong support for managing the demand with respect to backup of audits, data flow process, lock-in issue, security mechanisms, scalability, and validation. So, the need of scheduling strategies and its issues has a continuous focus for the CSPs to move next level of enhancement. Cloud scheduler is responsible for managing the resources and allocating the VM instances based on the consumer needs. The Figure 1 illustrates the generalized process involved in cloud scheduling approaches. The complex task performed by any CSP is the resource management which has an impact over scheduling strategies applied to the execution environment.

The guidelines developed for resource management need to achieve the objectives towards optimization. CSPs have to maintain a global and consistent state information in-order to manage allocation and de-allocation of resources i.e. auto scaling should be done efficiently. The scheduling strategies implemented by researchers so far arrive at optimal or approximate solutions in the cloud environment. The cloud models [21] developed for various needs supported with number of open source tools [25] always pave a way towards the enhancement of resource management and automation in real-time environment.

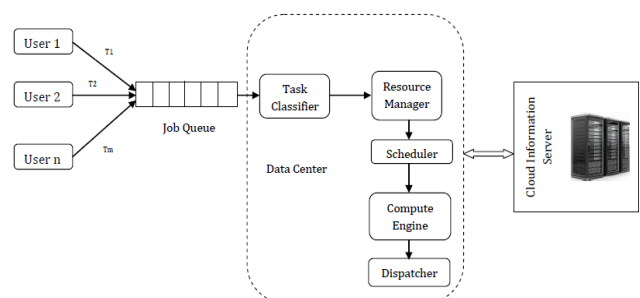


Fig. 1: Scheduling Process in Cloud Environment

II. WORKFLOW SCHEDULING IN CLOUD

Workflow scheduling is the process of resource allocation to single or batch jobs based on the application types and QoS requirements in the cloud environment. Dynamic scheduling support should be achieved by the CSP to meet the real-time needs and also the continuous change in the state of the VM instance strive towards strong approach to meet all the challenges.



Manuscript published on 30 September 2019.

* Correspondence Author (s)

Menaka School of Computer Science and Engineering, Vellore Institute of Technology, Vellore-632014

M, Sairabanu School of Computer Science and Engineering, Vellore Institute of Technology, Vellore-632014

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

The cloud scheduling categorization [22] can be viewed as depicted in the Figure 2 and analysis compared with input and output parameters along with tools are shown in Table I below.

The workflow scheduling hierarchy proceeds with the DAG (Directed Acyclic Graph) or non-DAG input representation that decides the flow of task execution along with the dependency levels. Next, the resources required are mapped to tasks for execution for which the VM instance is needed. All these together provide the hierarchical workflow model that achieves the process automation for the specific application.

III.CONCLUSION & FUTURE WORK

From the literature survey we unearth the following issues with respect to workflow scheduling, they include performance degradation, undefined QoS, low bandwidth, cache inclusion to focused, increased response time, resource failures, availability, cost, security, communication faults, attacks, lack of scheduler policies, and lack of update in number of VM cycles. The research scope can be extended towards data size and data migration issues and a well-defined technique could be imposed for efficient load balanced workflow scheduling process. Further, we would like to focus on automating VM execution environment efficiently considering the parameters like VM performance variation, delays, cost, and cloud characteristics.

REFERENCES

1. Guangshun Yao, Yongsheng Ding, Kuangrong Hao "Using imbalance characteristic for fault-tolerant workflow scheduling in cloud systems", IEEE Transactions on Parallel and Distributed Systems, 2017, 3671 - 3683.
2. Bin Xiang, Bibo Zhang, and Lin Zhang "Greedy-Ant: Ant Colony System-Inspired Workflow Scheduling for Heterogeneous Computing", IEEE Access, 2017, 11404 - 11412.
3. Quanwang Wu, Fuyuki Ishikawa, Qingsheng Zhu, Yunni Xia, Junhao Wen "Deadline-constrained Cost Optimization Approaches for Workflow scheduling in clouds", IEEE Transactions on Parallel and Distributed Systems, 2017, 3401 - 3412.
4. Huangke Chen, Xiaomin Zhu, Dishan Qiu, Ling Liu, Zhihui Du "Scheduling for Workflows with Security-Sensitive Intermediate Data by Selective Tasks Duplication in Clouds", IEEE Transactions on Parallel and Distributed Systems, 2017, 2674 - 2688.
5. Shuibing He, Yang Wang, Xian-He Sun and Chengzhong Xu "Using MinMax-Memory Claims to Improve In-Memory Workflow Computations in the Cloud", IEEE Transactions on Parallel and Distributed Systems, 2017, 1202 - 1214.
6. Bhaskar Prasad Rimal, Martin Maier "Workflow Scheduling in Multi-Tenant Cloud Computing Environments", IEEE Transactions on Parallel and Distributed Systems, 2017, 290 - 304.
7. Seong-Hwan Kim, Dong-Ki Kang, Woo-Joong Kim, Min Chen and Chan-Hyun Youn "A Science Gateway Cloud With Cost-Adaptive VM Management for Computational Science and Applications", IEEE Systems, 2017, 173 - 185.
8. Xiaoping Li, Zhicheng Cai "Elastic Resource Provisioning for Cloud Workflow Applications", IEEE Transactions on Automation Science and Engineering, 2017, 1195 - 1210.
9. Mateusz Zotkiewicz, Mateusz Guzek, Dzmitry Kliazovich, Pascal Bouvry "Minimum Dependencies Energy-Efficient Scheduling in Data Centers", IEEE Transactions on Parallel and Distributed Systems, 2016, 3561-3574.
10. Alexander A. Visheratin, Mikhail Melnik, Denis Nasonov "Workflow scheduling algorithms for hard-deadline constrained cloud environments", Elsevier, 2016, 2098-2106.
11. Mohammad Masdari, Sima ValiKardan, Zahra Shahi, Sonay Imani Azar "Towards workflow scheduling in cloud computing: A comprehensive analysis", Elsevier, Journal of Network and Computer Applications, 2016, 64 - 82.
12. Bing Lin, Wenzhong Guo, Naixue Xiong, Guolong Chen, Athanasios V.Vasilakos, and Hong Zhang "A Pretreatment Workflow Scheduling Approach for Big Data Applications in Multi-cloud Environments", IEEE Transactions on Network and Service Management, 2016, 581 - 594.
13. Xiaomin Zhu, Ji Wang, Hui Guo, Dakai Zhu, Laurence T. Yang, and Ling Liu "Fault-Tolerant Scheduling for Real-Time Scientific Workflows with Elastic Resource Provisioning in Virtualized Clouds", IEEE Transactions on Parallel and Distributed Systems, 2016, 3501 - 3517.
14. Zhaomeng Zhu, Gongxuan Zhang, Miqing Li, Xiaohui Liu "Evolutionary Multi-Objective Workflow Scheduling in Cloud", IEEE Transactions on Parallel and Distributed Systems, 2016, 1344 - 1357.
15. Xiaolong Xu, Wanchun Dou, Xuyun Zhang & Jinjun chun "EnReal: An Energy Aware Resource Allocation Method for Scientific Workflow Executions in cloud environment", IEEE Transactions on Cloud Computing, 2016, 166 - 179.
16. Amelie Chi Zhou, Bingsheng He and Cheng Liu "Monetary Cost Optimizations for Hosting Workflow-as-a-Service in IaaS Clouds", IEEE Transactions on Cloud Computing, 2016, 34 - 48.
17. Zhicheng Cai, Xiaoping Li, Jatinder N. D. Gupta "Heuristics for provisioning services to workflows in XaaS Clouds", IEEE Transactions on Services Computing, 2016, 250 - 263.
18. Yang Wang, Paul Lu, and Kenneth B. Kent "WaFS: A Workflow-Aware File System for Effective Storage Utilization in the Cloud", IEEE Transactions on Computers, 2015, 2716 - 2729.
19. Rodrigo N. Calheiros, Rajkumar Buyya "Meeting Deadlines of Scientific Workflows in Public Clouds with Tasks Replication", IEEE Transactions on Parallel and Distributed Systems, 2014, Vol.25, Issue 07, 1787 - 1796.
20. Maria A. Rodriguez and Rajkumar Buyya "Deadline based Resource Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds", IEEE Transactions on Cloud Computing, 2014, 222 - 235.
21. Ray J Rafaels "Cloud Computing: From Beginning to End", 2015.
22. Arshdeep Bahga and Vijay Madisetti "Cloud Computing: A hands-on approach", 2017.
23. Kai Hwang, Geoffrey C Fox, Jack G Dongarra, "Distributed and Cloud Computing, From Parallel Processing to the Internet of Things", Morgan Kaufmann Publishers, 2012.
24. Toby Velte, Anthony Velte, Robert Elsenpeter, "Cloud Computing, A Practical Approach", TMH, 2009.
25. Rajkumar Buyya, Christian Vecchiola, S.Tamarai Selvi, "Mastering Cloud Computing", TMGH, 2013.
26. Youn, Chan-Hyun, Chen, Min, Dazzi, Patrizio "Cloud Broker and Cloudlet for Workflow Scheduling", Springer Publisher, 2017, ISBN: 978-981-10-5071-8.