Research Article

A Survey of QoS aware and Fault Tolerant Workflow Execution in Heterogeneous Cloud Environment

SangethaSangani^aand Dr. Sunil F Rodd ^b

^aAsst.Prof. Dept of CSE. KLS Gogte Institue of Technology, Belgavi ^b Prof.Dept of CSE, KLS Gogte Institute Of Technology,Belgavi

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: One challenging issue in recent trends of distributed environment is workflow scheduling which concentrates on satisfying different constraints of quality of service (QoS). The cloud takes application forms for workflow that contains a set of tasks which are interdependent and are used to solve enterprise and large scale scientific problems. In cloud environment, the workflow scheduling provides broad review of approaches, which are studied comprehensively. Work scheduling article gives the analysis of different workflow scheduling methods and their characteristics. Is also perform the classification based on model of execution and its objectives. Along with this, recent developments in technology like edge-cloud computing are producing new opportunities and requirements for scheduling workflow for background processing tasks like web applications, Internet of-Things (IoT), and event-driven applications in distributed environment. Workflow scheduling article also gives scheduling of workflow in cloud computing trends. At the end, it presents the future research possible direction in creating an effective utilization of resource with good performance in tolerating the faults, execution of workload in the platform of heterogeneous cloud.

Keywords: Cloud computing, edge-cloud, evolutionary computation, resource monitoring, resource provisioning, scientific workflow, scheduling.

1. Introduction

The most promising and an emerging technology in present is cloud computing. Many different researches have been carried out on the cloud computing by several communities (Aarti Singh 2017). It can be used in different organizations and many IT companies for creating custom application services. Health care, Genetic science and other fields are using cloud computing. Cloud computing technologies are driven by the economies of the scale which will give large scale distributed computing. In large scale distributed computing the services and resources like storage, computing power, platform etc. are provided based on their demand through the internet (Yucong Duan, 2015). Cloud technologies provide three types of services: Infrastructure as a service (IaaS) for different hardware computational requirements , Platform as a service (PaaS) for programming languages and frameworks needed for the application /software developments and Software as a service (SaaS) provides application as a service which are ready to use in full-fledge.

Computational resources like Virtual Machines (VM) are provided to the users of cloud on the basis of demand by cloud computing. Optimization of Tasks scheduling is a research area in IaaS Cloud as it is NP-hard problem. However resource heterogeneity and autonomous attribute with in the VM execution and Clouds makes it necessary for various schemes of task scheduling in IaaS Cloud computing. These schemes are tested and used in order to reduce the makespan time, which is directly related to the execution cost of the tasks in environment (Achary R, 2015; Alkhanak EN, 2015; Srikanth GU, 2015).

In past few decades, a large amount of data was generated and transmitted to the cloud for the purpose of execution by Internet-of-Things (IoT) objects. Because of the massive storage capacity and processing cloud computing has become one major paradigms in the distributed environment. On the other hand, sensing data from smart items or objects and sending all the data to cloud environment, which is centralized is not much useful. This will increase the response time and latency time of the data objects. Again this method will increase wastage of resources and energy consuming rate of the server. To avoid these problems in the environment of centralized cloud, an edge-cloud computing pattern is developed by pushing processes to devices of local cloud. The main idea of edge-cloud computing is taking the benefit of the end users' devices which are located locally, and thereby increase IoT resources potential. But the Edge-cloud nodes are having limited capacity of resources for the execution of tasks. Because of the inbuilt weakness and strengths both cloud and edge-cloud devices cannot handle the above-mentioned problems efficiently. But hierarchical model of the edge-cloud along with the cloud architecture satisfies different objectives of real-time applications. Also they build sustainable infrastructure for the IoT. The IoT applications are represented in the form of workflow which can be efficiently executed in the edge-cloud with cloud hierarchical model compared to centralized cloud data center.

The providers of cloud will receive an application in the form of a direct acyclic graph (DAG), which is also known as workflow (Kahina Bessai, 2012). These workflows are characterized as model for various engineering and scientific problems. Each workflow needs enormous number of computing resources to complete the task executions with in limited time. Because of diversity in the applications, specific opportunities and challenges have to be developed for the workflow scheduling. Every workflow task has fixed requirement of resources. Big tasks of the workflow will consume more resources while the small tasks will consume fewer resources for execution in

cloud data center leading to higher make span and improper management of resources. One more significant characteristic of scheduling workflow is to satisfy the multiple QoS constraints such as budget and deadline of the workflow. A main concentration of cloud providers is to finish the execution of workflow within the given deadline and meet the constraint of the budget. The scheduler will select the best-fit resources for the tasks to reduce the time for execution and to boost the utilization of the resources like cache, energy resource.

Scheduling of workflow was extensively studied in the cloud environment (Deepak Poola, 2017; Reihaneh Khorsand, 2017). The review varies from existing strategies in several ways. First, review papers that are existing, discusses strategies of the different workflow scheduling. But the current article discusses both the strategies and the solution that can be provided for the cloud workflow managing system. Second, article discusses different dynamic, static and the metaheuristic-based workflow scheduling strategies. There are no complete surveys or reviews that clearly discuss workflow scheduling strategies in different infrastructures of cloud like edge-cloud model and IaaS. So this review concentrates on the different scheduling strategies of workflow in cloud environment and emerging trends of the distributed environment.

The significance of the work:

- Introduce a cloud workflow management system and summarize the recent advances in workflow scheduling strategy.
- Provide analysis of methods, modalities, strengths, and weaknesses of the existing workflow scheduling strategies.
- Presented possible research direction for utilize resource more efficient with fault tolerance guarantee for executing scientific workloads.

The paper is arranged as follows. In section II, discusses about the workflow management and taxonomy of workflows. In III, study about various existing workload scheduling mechanism on cloud and edge-cloud environment. Section IV highlights the various future issues and challenges that must be addressed for building efficient workflow management technique. In section V, problem statement and possible solution for building efficient workload management model in cloud environment is presented. Lastly, the survey is concluded with future research direction.

2. Workflow Management and Taxonomy of Workflow Model

The workflow management system uses servers (Suraj Pandey, 2011) for execution and management of the workflows. The users submit the task in the form of workflow to the cloud provider. The workflows are produced by workflow generators like Pegasus group (Ewa Deelman, 2005), ASKALON (Thomas Fahringer, 2005), and others. The workflow requests are received by the admission controller of cloud data center. The admission controller admits the requests if resources are available. The admission controller gives workflows to the workflow engine. Resource provisioner, data movement controller, and fault management are the three important components of the workflow engine. Resource provisioner selects the suitable resources from the available resources are monitored by the data movement controller. During the system failure or exception handling the fault manager will provide the mechanism for fault handling. In addition, workflow engine gives feedback to workflow monitor for periodically monitoring the status of the workflow execution. The detailed information regarding currently executing workflows and the availability of the resources in the cloud data center are maintained at the information service center. The system manager will monitor the availability of resources and stores information on resource information service. Finally, the tasks are deployed to the cloud computing resources by workflow engine.

Research Article



Fig. 1. Workflow management system

The workflow scheduling models broadly classified into two categories: Synthetic workflow models and scientific workflow models. Synthetic workflows contain the information that are artificially manufactured rather than generated real-world events. Synthetic workflow is created algorithmically and used for testing various scheduling strategies in a distributed environment. However, scientific workflows are being used today in various large-scale scientific processes. The workflows are generally inspired by nature and our environment, and each has a different structure. The detail of the workflow models is discussed as follows.

The workflow scheduling models are majorly divided into two types: Synthetic workflow models and scientific workflow models. Synthetic workflows consist of the information that are artificially created rather than created by real-world events. These models are algorithmically created and are used in testing different scheduling strategies in the distributed environment. On the other hand scientific workflows are used in different large-scale scientific processes. The workflows are generally induced by our environment and nature, and each has a diverse structure. The details of the workflow models are given below:

2. 1 Synthetic Workflow Models:

Based on the flow of execution of tasks, workflow applications are divided into many forms like Parallel workflow, Sequential workflow, Rules-Driven workflow, State-Machine workflow (Gideon Juve, 2011). In parallel workflow, two or more than two tasks are executed in parallel order. The execution of one set of tasks is independent from the execution of another set of the tasks in workflow. In sequential workflow, every task depends on the immediate predecessor task of it, and the workflow always moves forward without going back. Also, it is not possible for the workflow to trace the path back to a task which is already completed.

2. 2 Scientific Workflow Models:

There are many widely used scientific workflows with various structures and different computational characteristics (Duncan A 2007). Depending on the user application nature, scientific workflows can be CPU, memory or I/O intensive (G. B. Berriman, 2006). CPU intensive workflows will spend maximum time in execution of tasks on processors. But the memory-intensive workflows need large physical memory to save the data on server.

I/O-intensive workflows spend maximum time in doing an input operation and an output operation on server. The scientific workflows are characterized into following categories.

Epigenomics workflow: Based on the concept of the bioinformatics that automates the execution of various genome-sequence operations, the Epigenomics workflow is generated. USC Epigenome Center and the Pegasus Team have developed the Epigenomics. Here almost the all tasks require low utilization of I/O and high utilization of CPU.



Fig. 2. Epigenomics workflow

Cyber-Shake workflow: - It is used to describe earthquake hazards by the Southern California Earthquake Center (SCEC) (Robert Graves, 2011). To provide the automation robustness and reliability required to reach the essential computational scale Cyber Shake relies on scientific workflows. It characterizes earthquake hazards by producing synthetic seismograms. Here the tasks in the workflow use high CPU intensive and large memory for the execution tasks.

Research Article



Fig. 3. Cybershake Workflow

SIPHIT workflow: sRNA Identification Protocol using High-throughput Technology is the expansion of SIPHIT. This workflow is produced by a researcher in Harvard University. It is based on small untranslated RNAs (sRNAs) which control many bacteria processes like virulence and secretion (Jonathan Livny, 2008). The major job of this workflow is encoding and decoding of genes in all bacterial replicas which are stored in National Center for Biotechnology Information (NCBI) database. Almost all the tasks of this workflow have high memory and low CPU utilization.



Fig. 4. SIPHIT workflow.

In the present years, based on the above-mentioned scientific workflows performances strategies of workflow scheduling are evaluated. These scientific workflows are produced by using Workflow Generator tool given by the ASKALON, Cloudbus, Pegasus group and others, that uses real executions traces to create synthetic workflows.

3. Literature Survey

The survey of different workflow scheduling algorithms in cloud computing environment is given in this section. The different workflow scheduling algorithms are mainly divided into heuristic-based, meta-heuristic based, deadline-aware, SLA aware and evolutionary based methodologies.

3. 1 Heuristic based workflow scheduling algorithm

Several existing approaches are considered heuristic methods for task duplication, clustering and scheduling. Some examples are: In (H. Topcuoglu, 2012) proposed an algorithm for task clustering without duplication called CASS-II. This algorithm is compared with DSC in terms of both solution quality and speed. In (H. M. Fard, 2012) proposed scheduling mechanism for heterogeneous network which is based on task duplication. In (D. Shue, 2005) proposed a Heterogeneous Earliest finish time (HEFT) scheduling technique which is used in single work flow. In they have proposed a multi-objective heuristic scheduling for the grid and the cloud environment. Because of the unpredictable performance, these scheduling algorithms are not appropriate for the multi-tenant cloud environment. Unpredictable performance is due to the fact that tenant may opt for performance isolation and some may prefer best effort behaviour. In they showed that, in-time approximated results are preferred in real- time application domains, rather than accurate - but too much delayed - results. Here in this work we are proposing a deployment approach which exploits heterogeneity given by AMP architectures and the tolerance approximation is given by the applications such that quality of the results is increased as much as possible, provided timing and energy constraints. First an optimal approach based on problem linearization and decomposition is proposed. Next, a heuristic approach based on iteration relaxation of the optimal version is developed. The results show that 16.3% of reduction in computation time is achieved in comparison with the conventional optimal approaches.

3.2 Deadline-aware workload scheduling methodologies

In (Q. Zhu 2012), the authors have adopted Q-learning based learning model to meet the user defined deadline for particular application requirement. Deadline constraint weather forecasting system and a heuristic model based on grid scheduling model is proposed in (K. Plankensteiner, 2012) respectively. Scheduling strategies for single workflow instance for IaaS cloud platform is presented in (S. Abrishami, 2015). But these models not considered the multi-tenant cloud environment.

3.3 SLA aware workload scheduling methodologies:

The authors of (Zhou, 2018) presented joint optimization of cost and makespan of scheduling workflows in IaaS clouds. They modeled workflow scheduling scheme. A fuzzy dominance sort based heterogeneous earliest-finish-time (FDHEFT) algorithm is developed. This algorithm integrates closely with the fuzzy dominance sort mechanism with the list scheduling heuristic HEFT.

Efficacy of the schemes developed was demonstrated by the experiments using the synthetic and real-world workflows. Their schemes have achieved better cost-makespan tradeoff than the state-of-art algorithms. In (Zhou, 2019) the authors proposed two workflow scheduling approaches which are efficient for hybrid clouds such that both makespan and monetary cost are considered. In specific, first designed a single-objective workflow scheduling optimization approach called DCOH (deadline-constrained cost optimization for hybrid clouds) is modeled for reducing the monetary cost of scheduling workflows with deadline constraint. Next they modeled a multi-objective workflow scheduling optimization approach called MOH (multi-objective optimization for hybrid clouds) for minimizing monetary cost and makespan of scheduling work- flows simultaneously, which is based on DCOH. To validate the effectiveness of DCOH and MOH extensive simulation experiments were conducted. These simulation experiments indicate that, reduction in the monetary cost for users is achieved by using DCOH approach under the same deadline constraint as compared to other competing algorithms. Also MOH approach can achieve better cost-makespan trade-offs solutions in comparison with the other competing algorithms.

3.4 Energy aware workload scheduling methodologies:

In (G. Xie, 2019), the authors tried to minimize the energy consumption of a deadline constrained parallel application in a heterogeneous cloud computing system by turning off the as many as processors as possible and integrating the tasks on less number of processors. But they come to the conclusion that turning off as many processors as possible does not essentially lead to minimization of total energy consumption. Here they introduced an energy-aware processor merging (EPM) algorithm to select the most effective processor to turn off from the energy saving perspective, and quick EPM (QEPM) algorithm for reducing the complexity in computation of EPM. Experiments are conducted on real and randomly generated parallel applications to validate EPM and QEPM algorithms. The results infer that these algorithms are able to save more energy when compared to the existing methods at different scales, parallelism, and heterogeneity degrees. In (Z. Li, 2018) the authors proposed a cost and energy aware scheduling (CEAS) algorithm for cloud scheduler for minimizing execution cost of workflow and minimize the energy consumption while satisfying deadline constraints. The CEAS algorithm has five sub-

algorithms. First virtual machine (VM) selection algorithm applies cost utility concept to map tasks to their optimal VM types by the constraint of sub-makespan. Then, to reduce consumption of energy by workflow and to minimize the cost of execution two tasks merging methods are employed. Next, VM reuse policy is proposed to reuse the idle VM instances which have been leased. Based on time complexity analysis, they come to the conclusion that each sub-algorithm is having polynomial time complexity. The CEAS algorithm had better performance when compared to the related other well-known approaches.

In (Wen, 2019), the authors proposed an algorithm by name ECIB for scheduling instance-intensive IoT workflows with batch processing in clouds. The aim of this algorithm is to reduce cost of execution and to improve efficiency of energy utilization while satisfying deadline requirements. In specific they gave a prediction based approach for guiding how to manage the resources by historical data and CPU usage prediction results from the physical machines. Next, the two approaches to scale down or scale up the VM resources for optimizing the consumption of energy in the cloud data centers. Further they adopted strategy of batch processing for merging some activity instances of similar type for reducing the cost of execution by cloud users and to improve the utilization of resources for cloud data centers. In (Garg, R., 2019) the authors addressed an issue of reliability for the mission critical applications. In this they proposed an algorithm for scheduling called the reliability and energy efficient workflow scheduling algorithm. This algorithm optimizes application lifetime reliability together with consumption of energy and guarantees the QoS constraint that is specified by the users. The proposed algorithm by the authors has four phases: calculation of priority, task clusterization, target time distribution and assigning a cluster to processing element along with suitable levels of voltage/frequency. Simulation results are produced by using Gaussian Elimination task graphs and randomly generated task graphs. These results show that the presented approach is more effective in optimization of application lifetime reliability together with consumption of energy when compared to other existing algorithms.

3.5 Meta-heuristic and Machine learning algorithm based workload scheduling methodologies:

Meta-heuristic and Machine learning algorithm based workload scheduling methodologies: Here literature review of different existing meta-heuristic methodologies modeled for execution of workload in cloud computing environment is given. The authors of (S. Pandey, 2010; M. A. Rodriguez2014) presented a scheduling technique based on particle swarm optimization (PSO) for reducing the cost of workflow execution in cloud environment. In (Z. Wu, 2013), the authors (H. M. Fard 2015) implemented an optimization of genetic algorithm (GA), Ant colony optimization (ACO) and PSO for reducing the cost of workflow execution in cloud environment. The authors of (G. Juve, 2009) presented a dynamic scheduling and pricing model for single query in multi-cloud platform and it is compared with the traditional model multi-objective evolutionary algorithms, i.e., SPEA2 and NSGA-II. All these models were designed for optimization in grid environment and increase the computation overhead. So these models are not appropriate for using with large workflow application. The authors of (Chunlin, L., 2019) studied public cloud environment for performance and cost involved in computation. Based on their analysis they proved that amazon EC2 is not appropriate for I/O intensive application (NASA HPC cluster) due to the lack of parallel heterogeneous computing platform. In (Li, C., 2019), the authors addressed many challenges related to hybrid cloud environment. Some of the challenges are with minimum cost of monetary how to deploy a new application, heterogeneous jobs, and different cloud providers. They proposed an efficient job scheduling strategy for heterogeneous workloads in private cloud that can ensure the maximum resource utilization. In addition, the task scheduling method which is based on BP neural network in hybrid cloud is presented for ensuring that the tasks are completed within the deadline mentioned by user. From the experimental results it is proved that job response time can be reduced and throughput of cluster can be improved by efficient job scheduling approach. The task scheduling approach can boost the QoS satisfaction rate, reduce the public cloud cost and also minimizes the tasks response time. Presorted locality aware scheduling is presented by the authors of (J. Jin, 2011) for improving performance of the system. But they did not carried out any work about evaluation on dynamic real world workload.

The similar works are presented by the authors of (D. Yuan, 2010) and proposed a data placement approach in the scientific cloud workflows by adopting k-mean clustering. In (Zhu,2015), the authors stressed such complications, and model the workflow scheduling problem which can optimize both cost and makes pan as a Multi-objective Optimization Problem (MOP) for Cloud environments. An Evolutionary Multi-objective Optimization (EMO)-based algorithm is proposed by us to provide solution for the problem of workflow scheduling on an Infrastructure as a Service (IaaS) platform. In this algorithm a unique scheme for fitness evaluation and genetic operators, problem specific encoding and population initialization are presented. Based on the extensive experimental results on real world workflows and randomly generated workflows show that, scheduling done by our proposed evolutionary algorithm is more stable on all most all the workflows with the instance-based IaaS computation of the traditional goals such as budget, makes pan and deadline in workflow scheduling if green aspect of cloud, (i.e., consumption of energy) is considered it will increase the complexity of the problem. Additionally, matching all the interest of stakeholders simultaneously is very difficult as the stakeholder's interests of cloud are

conflicting. Hence they proposed a Multi-Objective Genetic Algorithm (MOGA) in cloud environment for workflow scheduling. This algorithm has considered the interest of cloud stakeholders that are conflicting for optimization and gave a solution. This solution considered the deadline and budget constraints to minimizes the makes pan. It also provides an energy efficient solution by using dynamic voltage frequency scaling and a gap search algorithm for optimizing the utilization of resources of cloud. This model helped to utilize the resources in a more efficient manner (Sardaraz M, 2020).

It is observed in (Abdulhamid SM, 2016), that whenever it is studied with the abundant resources for the execution, it has demonstrated than an ideal correlation among the available resources and the upcoming task for the necessary program for execution. But the programming model under execution with the intermediate results upon the completion of the NP as a nearest substitute. But in case of the data scheduling in the appropriate cloud computing IaaS layer for the NP complete problem. The significant feature of this kind of the design issue does not exist a perfect solution to be discovered. The NP complete design issues can be easily solved by the two various strategies namely non-heuristic and heuristic methods (El-Sisi AB 2014). The obtained results are achieved with the orientation with the scheduler by incorporating the various scheduling algorithms, whereas these algorithms will not directly impact on the kind of the non-centralized control systems and hence related to the corresponding agents. Hence there exists a demand for the exact scheduling policy such as to directly address the issues of the task scheduling in the significant programming environments (Shanahan HP, 2014) for the facilitation of the cloud scheduling. The LCA based approach is a non-heuristic optimization strategic working on the number of the datasets which is proposed in (Kashan HA. 2009). For performing the synthetic based approach for the building of the few simple idealized protocols of execution for the dependent algorithm based on the computer intelligence, which is done based on the few relative results of the round robin resource table for the attainment. A supportive study for the various topics of the LCA is presented in (Abdulhamid SM, 2015). An interesting spike in this domain of LCA has resulted in the significant results for the proposed scheme of non-centric heuristic approach which impress the scheduling issues of the various cloud computing environment. Hence this novel research exhibits the specific required approach for the various scheduling process of the cloud using GBLCA technique of optimal solution for the improvement with continuation of the previously existing reference papers (Abdulhamid SM, 2014) with the enhanced results and performance.

Contemporary models on Reinforcement learning (RL) and game-theoretic and their methodologies are extensively applied to scheduling problems capable of handling multi-constraint scenarios (D. P. Bertsekas, 2018; L. Xue 2018; H. Wang 2017; R. Duan, 2014). The acceptance is widely obtained to use the concept of equilibrium in game theories and also it finds high potential in training methodologies designed for multi-agent scenario, capable of dealing multi-objective and multi-constraint problem optimization techniques. D. Cui, 2016 presents a game algorithm designed using sequential cooperative methodologies for make-span optimization and cost effective and it also fulfills constraints for work-flow schedule in large-scale. (Cui 2018) presents a supplement for learning based method for varieties of work-flow scheduling, having different levels of priorities based on individual necessity available in the cloud. Within present an algorithm for admission-control and distributed load-balancing on the basis of game-theoretic method designed using fuzzy logic, at enormously big data handling in SaaS clouds. In (W. Jiahao, 2018) presents an upgraded algorithm using Q-learning having weighted fitness function of valuation for optimization of load-balancing and time in cloud computing. However, the model is not evaluated considering modern scientific work flows. This work further states research issues and challenges that must be met in future for executing scientific workload.

4. Future Issues and Challenges in Workflow Scheduling in Heterogeneous Cloud Environment

In the present section, we are discussing in detail regarding the various relevant and dependable pathway for the continued work for the present work of scheduling in workflow for the various cloud heterogeneous environment. **4.1 Task Scheduling with resource constraint:**

4.1 Task Scheduling with resource constraint:

The major concern of the cloud computing is to efficiently utilize the data available in the cloud data center for meeting of the different quality of service. The requirements of resources in any cloud computing services will have specific size and contiguous stages for different applications of different users these parameters are to be mentioned in the initial stages of development. Deadline and response time will vary with vary in application. Aim of the assigning task with diminished requirements of resources is to search for suitable scheduling which in turn increases the parallelism of tasks and also maximize the server's resource utilization.

4.2 Dynamic QoS constraint:

Raising demands and availability of variety of resources at cloud computational methods will increase difficulty for algorithms designed using heuristic approach, this results in the difficulty to prediction of QoS requirements. Static QoS requirements are used in many distributed environment to accomplish tasks. The use of static QoS requirement may not suitable for the unstable cloud environment working in an dynamical way. To address this issue of unstable cloud environment a well suitable method is required than static QoS requirements, capable of better performance with available resources provided by cloud service providers and work in a dynamical way. For instance, deadline assigning is to be adjourned, if the task need 21 minutes to complete the workflow, the service

provider will have an option of allocating some time bound to execute the work-flow completely. Effective utilization of resources for computing available in the data center of cloud service provider we need to make use of QoS in an dynamic way this will in turn handle multiple objectives of the application.

4.3 Workload forecasting and makespan modelling:

To handle the QoS parameters effectively we make use of large-scale heterogeneous system in cloud computing systems. Moreover, to know the pattern and characteristics of workload on the servers used in the cloud data center is more difficult task to enhance condition of operations in cloud computing and resource utilization. Workload analysis and projection of target server for execution of the tasks on the basis of real world parameters like memory and CPU usage have to investigate with at most urgency, to know the work load characteristics impact on the cloud. Dynamic nature of cloud is required to put forward prediction of performance of workloads to choose much suitable target server and also to get runtime estimation of tasks requested by the user. Performance model value depends on the work-flow computation time, previous usage of resources like CPU and memory usage, rate of performance of the resources and so on. This method will assist to choose the suitable fit target server for work-flow and maximum utilization of resources of the server.

4.4 Incorporating multiple virtual machine characteristics:

Currently used strategies for scheduling make use of in depth computations and developers make use of VM instances with respect to capacity of the CPU. Moreover, vital cloud computing parameters like disk usage and memory are not considered. For applications with high I/O intensive or memory intensive applications these work-flow properties are valid. For instance, to meet the goal various machine-learning methods are used and the work-flow necessities of storage to fit the data analytically. So, the developers are concentrating on different types of resources of computing and also address work-flow analytically. So the developers must focus on defining the VM instances of various types of computing resources. The performance of the system may have its effect and cost, execution time also could have its effects when designing an work-flow execution application. So there is need for the developers to design different dynamic or heuristic approaches for suitable fit VM instances on the basis of the manifold resource requirements.

4.5 Workflow Fault-Tolerance and Dependability:

To handle the failure mechanisms different techniques of work-flow mechanisms with fault-tolerance are proposed to tackle work-flow execution with relatively reliable and availability, but still dependable on supporting big data work-flow increases complexity. Factors like execution environments and its lengthy execution processes will have its effect on the dynamism of such workflows. Generally, handling the failures that occurred requires first catching the error, identifying its source, then reducing its impact and finally taking the appropriate actions to recover from it. Considering a "Cloud of clouds" environment, achieving those tasks is even harder due not only to the characteristics of big data and big data workflow, but also because of the characteristics of such environments. **4.6 Energy efficiency workflow execution:**

Employing cloud computing needs multiple server's varieties from hundreds to thousands these servers power consumption is much more and power purchase cost is also high, this power consumption will also harm the environment. Moreover, concept of visualization is a vital method used in cloud, this will increase handling of tasks using parallelism in cloud data center. To handle task execution in parallel order, many virtual machines are employed in cloud data center according to their resource capacity. There arises the necessity to design more energy-efficient virtual machine to schedule the task is more challenging in cloud domain for minimization of consumption of energy in cloud data center, much energy efficient strategies are used one such strategy is minimization of consumption of energy by placing the virtual machines such that it makes use of part of the server available in the data cloud center.

4.7 Resource forecasting and workload execution:

Normally while scheduling the work-flow current load on the resources are not considered during the task deployment strategies. Furthermore, choosing of more convenient resources like processing element and cache for execution of the task without affecting the QoS requirements may be the considerable problem because of their dynamic in nature. To avoid and improve the mentioned challenges, work-flow scheduling task is to be done by resource prediction. Based on the resource availability and server loads prediction of resource is done that may help to get the more suitable server at each instance may get improved accuracy and performance of the system. The other vital issue is superior resource prediction may meet the variety of scheduling objectives without affecting the QoS constraints.

In next section this work presents problems statement considering above future open issues and challenges in building effective and fault-tolerant workload scheduling methodology for heterogeneous cloud computational environment.

5. Problem Statement and Possible Solution to Design Resource Efficient and Fault Tolerant Workload Scheduling Algorithm for Multicore Processing Environment

NoC (Network on Chip) has been extensively used in modern multi-core processors (mainly in servers of cloud computing) and System-on-Chips (SoC), as a flexible and scalable solution to the extensive integration of additional components in the chips. furthermore, a main concern is energy-efficiency, performance, fault tolerant and reliability necessity of computing system using big data, we imagine self-aware multi-core architectures will adapt automatically, the behavior accommodate the necessities of users/applications, energy constraints and performance and also availability of the resources. This adaptivity can span all the way from the application (e.g., task mapping/scheduling) to the core (e.g., DVFS/power-gating) to the interconnect fabric (e.g., routing).

Many developers are mainly concentrating on performance and neglecting the energy conservation in environment of cloud computing, so many more efficient power management technologies are implemented by many developers for power dissipation reduction and help to maintain environment, it also improves the performance. Moreover, the interaction cost between the inter-processors will increase for these power efficient technologies. However, the technologies may not provide satisfactory results and consumption of energy in cloud computing devices may not be minimized noticeably due to extensive use of routing, memory utilization etc.

Inter-processor communication cost is very high using existing techniques as well as these algorithms are rely upon classic model which is not realistic. Therefore, existing algorithms are inaccurate and not able to reduce energy consumptions. So, a balanced scheduling with novel approach with DVFS (Dynamic Voltage and Frequency Scaling) is developed for cloud computing needs. This modeled to maintain trade-off between energy consumption and system performance, for cloud multi-core architectures. Further, effective utilization i.e. minimizing last level cache failure will further enhance system performance and reduce energy consumption. Thus, the research work aimed at designing fault-tolerant workload scheduling algorithm.

6. Conclusion

The proposed work is on strategies and various work-flow schedule designed with the in depth analysis on cloud environment. To be more precise, the method is on the scheduling techniques to execute after scheduling the enterprise and scientific work-flow in edge-cloud and cloud model. This work is based on a taxonomic survey based on different work-flow models and strategically scheduling. The work-flow is classified broadly into two categories, scientific work-flow and synthetic work-flow. The problems and applications of scientific in nature are designed on the basis of different scientific work-flows. Furthermore, to test data sheets and design applications of the users synthetic work-flow strategies are used. We are also ponder about different types of scheduling objectives and also their advantages in a cloud environment. The execution model is to be elaborated with in the cloud computing environment. This is done in parallel knowing the disadvantages and advantages of different work-flow anagement technique. Future work would consider developing efficient workload management technique adopting mete-heuristic approach and evaluate its performance considering real-time scientific applications.

References

- 1. Aarti Singh, Dimple Juneja, and Manisha Malhotra. A novel agent based autonomous and service composition framework for cost optimization of resource provisioning in cloud computing. J. King Saud Univ. Comput. Inf. Sci. 29, 1, 19–28, 2017.
- 2. Achary R, Vityanathan V, Raj P, Nagarajan S. "Dynamic Job Scheduling Using Ant Colony Optimization for Mobile Cloud Computing". Intelligent Distributed Computing: Springer. pp. 71–82, 2015.
- 3. Alkhanak EN, Lee SP, Khan SUR "Cost-aware challenges for workflow scheduling approaches in cloud computing environments: Taxonomy and opportunities". Future Generation Computer Systems, 2015.
- Abdulhamid SM, Abd Latiff MS, Abdul-Salaam G, Hussain Madni SH. "Secure Scientific Applications Scheduling Technique for Cloud Computing Environment Using Global League Championship Algorithm". PLoS ONE 11(7): e0158102. <u>https://doi.org/10.1371/journal.pone.0158102</u>, 2016.
- 5. Abdulhamid SM, Latiff MSA, Madni SHH, Oluwafemi O. "A Survey of League Championship Algorithm: Prospects and Challenges". Indian Journal of Science and Technology 8: 101–110, 2015.
- 6. Abdulhamid SM, Abd Latiff MS. League Championship Algorithm Based Job Scheduling Scheme for Infrastructure as a Service Cloud; 2014; Universiti Teknologi Malaysia, Johor Bahru, Malaysia. pp.381– 382.
- 7. Abdulhamid SM, Abd Latiff MS, Ismaila I. "Tasks scheduling technique using league championship algorithm for makespan minimization in IAAS cloud". ARPN Journal of Engineering and Applied Sciences 9: 2528–2533, 2014.

- 8. Arunkarthikeyan, K., Balamurugan, K., Nithya, M. and Jayanthiladevi, A., 2019, December. Study on Deep Cryogenic Treated-Tempered WC-CO insert in turning of AISI 1040 steel. In 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 660-663). IEEE.
- Arunkarthikeyan K., Balamurugan K. & Rao P.M.V (2020) Studies on cryogenically treated WC-Co insert at different soaking conditions, Materials and Manufacturing Processes, 35:5, 545-555, DOI: <u>10.1080/10426914.2020.1726945</u>
- Balamurugan, K., 2020. Compressive Property Examination on Poly Lactic Acid-Copper Composite Filament in Fused Deposition Model–A Green Manufacturing Process. Journal of Green Engineering, 10, pp.843-852.
- 11. Chunlin, L., Jianhang, T. & Youlong, L. Hybrid Cloud Adaptive Scheduling Strategy for Heterogeneous Workloads. J Grid Computing 17, 419–446 (2019).
- 12. D. Yuan, Y. Yang, X. Liu, and J. Chen, i.
- 13. Deepak Poola, Mohsen Amini Salehi, Kotagiri Ramamohanarao, and Rajkumar Buyya. A taxonomy and survey of fault-tolerant workflow management systems in cloud and distributed computing environments. In Software Architecture for Big Data and the Cloud. Elsevier, 285–320, 2017.
- Deepthi, T., Balamurugan, K. and Balamurugan, P., 2020, December. Parametric Studies of Abrasive Waterjet Machining parameters on Al/LaPO4 using Response Surface Method. In IOP Conference Series: Materials Science and Engineering (Vol. 988, No. 1, p. 012018). IOP Publishing.
- 15. Duncan A. Brown, Patrick R. Brady, Alexander Dietz, Junwei Cao, Ben Johnson, and John McNabb. A case study on the use of workflow technologies for scientific analysis: Gravitational wave data analysis. In Workflows for e-Science. Springer, Berlin, 39–59, 2007.
- 16. D. Shue, M. J. Freedman, and A. Shaikh, "Performance isolation and fairness for multi-tenant cloud storage," in Proc., USENIX OSDI, Oct. 2012, pp. 349–362.
- 17. D. Cui, W. Ke, Z. Peng, and J. Zuo, ``Multiple DAGs workflow scheduling algorithm based on reinforcement learning in cloud computing," in Computational Intelligence and Intelligent Systems. Singapore, Springer, pp. 305-311, 2016.
- 18. D. P. Bertsekas, ``Feature-based aggregation and deep reinforcement learning: A survey and some new implementations," IEEE/ACM Trans. Audio, Speech, Language Process., pp. 1-31, 2018.
- 19. El-Sisi AB, Tawfeek MA. "Cloud Task Scheduling for Load Balancing based on Intelligent Strategy". International Journal of Intelligent Systems and Applications (IJISA) 6: 25, 2014.
- 20. E. Iranpour and S. Sharifan, ``A distributed load balancing and admission control algorithm based on Fuzzy type-2 and Game theory for large-scale SaaS cloud architectures," Future Gener. Comput. Syst., vol. 86, pp. 81-98, Sep. 2018.
- 21. Ewa Deelman, Gurmeet Singh, Mei-Hui Su, James Blythe, Yolanda Gil, Carl Kesselman, Gaurang Mehta, Karan Vahi, G. Bruce Berriman, John Good, et al. 2005. *Pegasus: A framework for mapping complex scientific workflows onto distributed systems. Sci. Program. 13, 3 (2005), 219–237.*
- 22. Garikipati, P. and Balamurugan, K., 2021. Abrasive Water Jet Machining Studies on AlSi 7+ 63% SiC Hybrid Composite. In Advances in Industrial Automation and Smart Manufacturing (pp. 743-751). Springer, Singapore.
- 23. Garg, R., Mittal, M. & Son, L. Reliability and energy efficient workflow scheduling in cloud environment. Cluster Comput 22, 1283–1297. <u>https://doi.org/10.1007/s10586-019-02911-7</u>, 2019.
- 24. G. Juve, E. Deelman, K. Vahi, G. Mehta, B. Berriman, B. Berman, and P. Maechling, "Scientific workflow applications on amazon ec2," in Proc., IEEE E-Science Wksp, Dec. 2009, pp. 59–66.
- 25. G. Xie, G. Zeng, R. Li and K. Li, "Energy-Aware Processor Merging Algorithms for Deadline Constrained Parallel Applications in Heterogeneous Cloud Computing," in IEEE Transactions on Sustainable Computing, vol. 2, no. 2, pp. 62-75, doi: 10.1109/TSUSC.2017.2705183, 2017
- 26. Gideon Juve and Ewa Deelman. Scientific workflows in the cloud. In Grids, Clouds and Virtualization. Springer, Berlin, 71–91, 2011.
- 27. G. B. Berriman, A. C. Laity, J. C. Good, J. C. Jacob, D. S. Katz, E. Deelman, G. Singh, M. H. Su, and T. A. Prince. 2006. Montage: The architecture and scientific applications of a national virtual observatory service for computing astronomical image mosaics. In Proceedings of the Earth Sciences Technology Conference, 2006.
- 28. H. M. Fard, R. Prodan, J. J. D. Barrionuevo, and T. Fahringer, "A multi-objective approach for workflow scheduling in heterogeneous environments," in Proc., IEEE/ACM CCGrid, May 2012, pp. 300–309.

- 29. H. Wang, T. Huang, X. Liao, H. Abu-Rub, and G. Chen, "*Reinforcement learning for constrained energy trding games with incomplete information," IEEE Trans. Cybern., vol. 47, no. 10, pp. 3404-3416, Oct. 2017.*
- 30. H. Topcuoglu, S. Hariri, and M. Y.Wu, "Performance-effective and low-complexity task scheduling for heterogeneous computing," IEEE Trans. Parallel and Distrib. Sys., vol. 13, no. 3, pp. 260–274, Mar. 2012.
- 31. J. Jin, J. Luo, A. Song, F. Dong, and R. Xiong, "Bar: An efficient data locality driven task scheduling algorithm for cloud computing," in Proc., IEEE/ACM CCGrid, May 2011, pp. 295–304.
- 32. Jonathan Livny, Hidayat Teonadi, Miron Livny, and Matthew K. Waldor. 2008. *High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PloS One 3, 9, e3197, 2008.*
- 33. Jonathan Livny, Hidayat Teonadi, Miron Livny, and Matthew K. Waldor. 2008. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PloS One 3, 9, e3197, 2008.
- 34. K. Plankensteiner and R. Prodan, "Meeting soft deadlines in scientific workflows using resubmission impact," IEEE Trans. Parallel and Distrib. Sys., vol. 23, no. 5, pp. 890–901, May 2012.
- 35. Kashan HA. League championship algorithm: a new algorithm for numerical function optimization. IEEE. pp. 43–48, 2009.
- 36. Kahina Bessai, Samir Youcef, Ammar Oulamara, Claude Godart, and Selmin Nurcan. Bi-criteria workflow tasks allocation and scheduling in cloud computing environments. In Proceedings of the IEEE 5th International Conference on Cloud Computing (CLOUD'12). IEEE, 638–645, 2012.
- 37. Latchoumi, T.P., Dayanika, J. and Archana, G., 2021. A Comparative Study of Machine Learning Algorithms using Quick-Witted Diabetic Prevention. Annals of the Romanian Society for Cell Biology, pp.4249-4259.
- Latchoumi, T.P., Vasanth, A.V., Bhavya, B., Viswanadapalli, A. and Jayanthiladevi, A., 2020, July. QoS parameters for Comparison and Performance Evaluation of Reactive protocols. In 2020 International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE) (pp. 1-4). IEEE.
- 39. L. Xue, C. Sun, D. Wunsch, Y. Zhou, and F. Yu, ``An adaptive strategy via reinforcement learning for the prisoner's dilemma game," IEEE/CAA J. Autom. Sinica, vol. 5, pp. 301-310, 2018.
- 40. Li, C., Tang, J. & Luo, Y. Cost-aware scheduling for ensuring software performance and reliability under heterogeneous workloads of hybrid cloud. Autom Softw Eng 26, 125–159 (2019).
- 41. National Centre for Biotechnology Information. [n.d.]. Retrieved from http://www.ncbi.nlm.nih.gov.
- 42. S. Pandey, L. Wu, S. Guru, and R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," in Proc., IEEE Advanced Information Networking and Applications, Apr. 2010, pp. 400–407.
- 43. K. Plankensteiner and R. Prodan, "Meeting soft deadlines in scientific workflows using resubmission impact," IEEE Trans. Parallel and Distrib. Sys., vol. 23, no. 5, pp. 890–901, May 2012.
- 44. Q. Zhu and G. Agrawal, "Resource provisioning with budget constraints for adaptive applications in cloud environments," IEEE Trans. Services Comput., vol. 5, no. 4, pp. 497–511, Oct./Dec. 2012.
- 45. M. A. Rodriguez and R. Buyya, "Deadline based resource provisioningand scheduling algorithm for scientific workflows on clouds," IEEE Trans. Cloud Comput., vol. 2, no. 2, pp. 222–235, Apr. 2014.
- 46. R. Duan, R. Prodan, and X. Li, "Multi-objective game theoretic scheduling of bag-of-tasks workflows on hybrid clouds," IEEE Trans. Cloud Comput., vol. 2, no. 1, pp. 29-42, 2014.
 - A. Rehman Syed S. Hussain Zia ur Rehman Seemal Zia Shahaboddin Shamshirband "Multi-objective approach of energy efficient workflow scheduling in cloud environments" available online: https://doi.org/10.1002/cpe.4949, 2018.
- 47. Reihaneh Khorsand, Faramarz Safi-Esfahani, Naser Nematbakhsh, and Mehran Mohsenzade. Taxonomy of workflow partitioning problems and methods in distributed environments. J. Syst. Softw. 132 (2017), 253–271, 2017.
- Robert Graves, Thomas H. Jordan, Scott Callaghan, Ewa Deelman, Edward Field, Gideon Juve, Carl Kesselman, Philip Maechling, Gaurang Mehta, Kevin Milner, et al. 2011. CyberShake: A physics-based seismic hazard model for southern California. Pure Appl. Geophys. 168, 3–4, 367–381, 2011.
- 49. K. Plankensteiner and R. Prodan, "Meeting soft deadlines in scientific workflows using resubmission impact," IEEE Trans. Parallel and Distrib. Sys., vol. 23, no. 5, pp. 890–901, May 2012.

- S. Pandey, L. Wu, S. Guru, and R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," in Proc., IEEE Advanced Information Networking and Applications, Apr. 2010, pp. 400–407.
- Sardaraz M, Tahir M. A parallel multi-objective genetic algorithm for scheduling scientific workflows in cloud computing. International Journal of Distributed Sensor Networks. doi:10.1177/1550147720949142, 2020.
- 52. Shanahan HP, Owen AM, Harrison AP. "Bioinformatics on the cloud computing platform Azure". 2014
- 53. Srikanth GU, Maheswari VU, Shanthi A, Siromoney A "Task Scheduling Model". Indian Journal of Science and Technology 8: 33–42, 2015.
- 54. Suraj Pandey, Dileban Karunamoorthy, and Rajkumar Buyya. 2011. Workflow engine for clouds. Cloud Comput.: Princ. Paradigms 87 (2011), 321–344, 2011.
- 55. D. Shue, M. J. Freedman, and A. Shaikh, "Performance isolation and fairness for multi-tenant cloud storage," in Proc., USENIX OSDI, Oct. 2012, pp. 349–362.
- 56. Tsai C-W, Huang W-C, Chiang M-H, Chiang M-C, Yang C-S "A hyper-heuristic scheduling algorithm for cloud". Cloud Computing, IEEE Transactions on 2: 236–250, 2014.
- 57. USC Epigenome Center. [n.d.]. Retrieved from http://epigenome.usc.edu
- W. Jiahao, P. Zhiping, C. Delong, L. Qirui, and H. Jieguang, ``A multiobject optimization cloud workflow scheduling algorithm based on reinforcement learning," in Intelligent Computing Theories and Application. Cham, Switzerland: Springer, pp. 550-559, 2018.
- Wen, Yiping & Wang, Zhibin & Zhang, Yu & Liu, Jianxun & Cao, Bu-Qing & Chen, Jinjun. Energy and cost aware scheduling with batch processing for instance-intensive IoT workflows in clouds. Future Generation Computer Systems. 101. 10.1016/j.future.2019.05.046, 2019.
- 60. Yarlagaddaa, J., Malkapuram, R. and Balamurugan, K., 2021. Machining Studies on Various Ply Orientations of Glass Fiber Composite. In Advances in Industrial Automation and Smart Manufacturing (pp. 753-769). Springer, Singapore.
- Yookesh, T.L., Boobalan, E.D. and Latchoumi, T.P., 2020, March. Variational Iteration Method to Deal with Time Delay Differential Equations under Uncertainty Conditions. In 2020 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 252-256). IEEE.
- 62. Yucong Duan, Guohua Fu, Nianjun Zhou, Xiaobing Sun, Nanjangud C. Narendra, and Bo Hu. Everything as a service (XaaS) on the cloud: Origins, current and future trends. In Proceedings of the 2015 IEEE 8th International Conference on Cloud Computing. IEEE, 621–628, 2015.
- Chu, Zhaomeng & Zhang, Gongxuan & Li, Miqing & Liu, Xiaohui. (2015). Evolutionary Multi-Objective Workflow Scheduling in Cloud. IEEE Transactions on Parallel and Distribution Systems. 27. 10.1109/TPDS.2015.2446459
- 64. Z. Li, J. Ge, H. Hu, W. Song, H. Hu and B. Luo, "Cost and Energy Aware Scheduling Algorithm for Scientific Workflows with Deadline Constraint in Clouds," in IEEE Transactions on Services Computing, vol. 11, no. 4, pp. 713-726, 1 July-Aug. 2018, doi: 10.1109/TSC.2015.2466545.
- Zhou, Xiumin & Zhang, Gongxuan & Sun, Jin & Zhou, Junlong & Wei, Tongquan & Hu, Shiyan. (2018). Minimizing Cost and Makespan for Workflow Scheduling in Cloud using Fuzzy Dominance Sort based HEFT. Future Generation Computer Systems. 93. 10.1016/j.future.2018.10.046, 2018.
- Zhou, Junlong & Wang, Tian & Cong, Peijin & Lu, Pingping & Wei, Tongquan & Chen, Mingsong. (2019). Cost and Makespan-Aware Workflow Scheduling in Hybrid Clouds. Journal of Systems Architecture. 100. 10.1016/j.sysarc.2019.08.004, 2019.
- 67. Z. Wu, X. Liu, Z. Ni, D. Yuan, and Y. Yang, "A market-oriented hierarchical scheduling strategy incloud workflow systems," J. Supercomputing, vol. 63, no. 1, pp. 256–293, Jan. 2013.