Grenze International Journal of Engineering and Technology, Jan Issue

Dynamic Multi-Workflow Scheduling Analysis of Real-time Healthcare Ecosystem: An Emerging Research Area

Sugandha Rathi[1], Renuka Nagpal[2], Deepti Mehrotra[3] and Gautam Srivastava[4]
[1-3]Department of Computer Science and Engineering, Amity University, Noida, India
[4]Department of Math and Computer Science, Brandon University, Canada.
E-mail: sugandharathi@gmail.com, mehdeepti@gmail.com, rnagpal1@amity.edu

***Abstract*— In today's high technology environment, most of the workflows are migrated to the cloud and related resourcing technologies like fog, mist and edge computing. Here, in this research work, we have given a detailed idea about dynamic multi-workflow scheduling (DMWS) in fog environment. The fog computing parameters that are affected are also mentioned. Various standard workflows are discussed mentioning their usage in various fields. DMWS has two major branches which are the most researched in past few years. One is multi-workflow scheduling and the other one is dynamic workflow scheduling. Both workflow scheduling have its own pros and cons. A detailed literature review of the papers published since 2019 on DMWS is discussed in detail. Also, a case study regarding DMWS in the healthcare field is conducted and discussed. A section is included that explains the various challenges and issues in DMWS. Lastly future scope and areas of implementation are discussed for DMWS.**

***Index Terms*— Workflow scheduling, Dynamic-workflow scheduling, Multi-workflow scheduling, Fog computing, Standard workflows.**

I. INTRODUCTION

Cloud computing is the latest technology which allows the dynamic allocation of resources using a variety of techniques of distributed resource allocation, high-speed computing and platform virtualization. The services like IaaS, PaaS and SaaS have helped and enabled the cloud to work efficiently.
This popularity has proved to be both boon and a bane for cloud computing. As the number of users is increasing exponentially, a drastic increase in traffic on the cloud is seen. A few major issues of delay, security, and cost have risen. To address them, fog computing came into light.
Fog computing is an extension of the cloud and its services. All the properties and features that the cloud provides are also provided by fog but in a better way like performing correlated functions are performed in the local resource pool leading to less usage of time. Fog computing plays an important part in time-related platform-based applications which are part of IoT services like smart car parking, smart city, etc.
All the prominent areas which follow a particular path or process for task completion are said to be known as workflows. Workflow scheduling is a process that allocates the execution of tasks to distributed resources. Most real-world applications work on workflow structures. With the rise of fog computing, scheduling workflows in fog becomes an essential trend [1]. However, many factors affect the performance of the workflows in a fog computing environment. A few of the factors that affect performance include workflow patterns, size, fog resources, etc.

Grenze ID: 01.GIJET.10.1.552

The problem of workflow scheduling is known as NP-complete [1]. Most existing works focus on workflow scheduling in a static scenario. A static scenario means a pre-defined path is set for the workflows in advance. It is assumed that they follow the same path mentioned only. However, in a real-life scenario, the workflow paths keep changing and the pattern is not known in advance. Task scheduling techniques are utilized for scheduling tasks on computing processors to minimize total makespan without violating the specified restrictions [2]. Various techniques are being developed to address the problems in computation offloading. Furthermore, optimization of offloading decisions and resource allocation, such as distribution of bandwidth, transmit power, and computation resources, are employed for attaining system performance gain, energy efficiency improvement, and reduction of energy consumption and delay [3,4].
There are mainly two concepts of workflows used in the real world.

- In the real world, fog computing structures are showing dynamic variations that are needed by the users. Fog computing paradigms are capable of satisfying the need for time-sensitive applications which need dynamic alternations of the user necessities. Dynamic workflow scheduling is an integral part of today's world. It aims at allocating dynamically allocating the tasks to the fog resources with optimal cost, time, makespan, reliability, etc.
- Second is the use of multi-workflow in a single problem. The scenario uses multiple workflows from the source to the destination. Fig 1. Shows a multi-workflow structure.

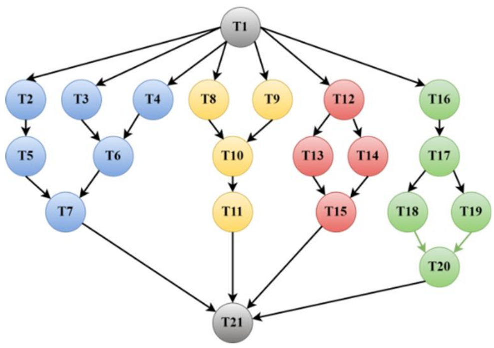

Fig 1.Multi-Workflow Structure

Dynamic and multi-workflows can together be used in the user necessity of scenario creation. Both workflows can be described with the help of the Venn diagram in Fig. 2.

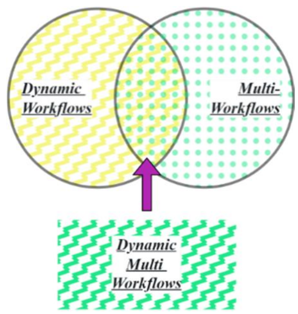

Fig 2. Dynamic Multi Workflows

Both the techniques i.e. multi-WS and dynamic WS have their positive side. A few of them discussed here are the pros of both types of WS techniques.

A. Pros:

- On-the-fly process modifications are available in realistic scenarios.
- To continue to work on the algorithm, check the execution status after equal intervals.
- Analyze the algorithm based on WS type and rebuild
- Both WS gives scope for flexibility

Both support interactive workflows - workflows are involved with human interactions and mainly consider the factors such as response time, stability and security, etc.

Task scheduling techniques are utilized for scheduling tasks on computing processors to minimize total makespan without violating the specified restrictions [5, 6]. Furthermore, optimization of offloading decisions and resource allocation, such as distribution of bandwidth, transmit power, and computation resources, are employed for attaining system performance gain, energy efficiency improvement, and reduction of energy consumption and delay [3,4].

Fog computing provides a good environment for dynamic requests that can be easily and efficiently deployed on hardware and also the proper use of heterogeneous workflows to improve the throughput, cost and performance. The multiple workflow applications are connected by the VMs and tasks in a particular manner and are priority-based. These workflow applications have interconnected data and tasks that are to be performed in a particular way only.

Dynamic Multi workflows are better handled by heuristic algorithms. As the tasks keep coming dynamically, so whenever a new task arrives, an optimized rescheduling is done with the help of a heuristic approach by calculating the best options. The focus is on globally optimizing the performance of the workflows and giving reliable and optimized results. Therefore, there is a strict need to design algorithms which can easily be implemented on multi-workflows and can optimize the dynamic workflow environment by calculating and changing the approach accordingly.

The WS depends on a few parameters to optimize the results. These factors are the drawing force of WS. The parameters that are taken into consideration are cost, time, make span, security, energy consumption, reliability and a few more..

In practice, cloud/fog users want low cost, high reliability, low energy consumption, high security, small make span and less execution time at the same time. The objectives here don't allow us to make a clear boundary between the two definitions, so, we can define dynamic workflow scheduling as a multi and many-objective problem depending upon the number of parameters taken to optimize to jointly work on all the parameters.

B. Types of Standard Workflows

There are numerous scientific Workflows such as Montage, LIGO, Cybershake, SIPHT and epigenomics. These workflows were initially designed for complex-sized data and data which needed very high processing power.

1. *Montage*- It was designed by NASA to be applied in the field of astronomy. It is an open-source toolkit to construct image mosaics of sky images.
2. *Cybershake*- It has its application in earthquakes and other seismic hazard prediction in the southernCalifornian region.
3. *Epigenomics*- It is used to generate gene sequences.
4. *Inspiral*- Used to generate waveforms from collected data in a binary system.
5. *SIPHT*- used in bioinformatics

II. LITERATURE SURVEY

An integrated Information Technology - Operational Technology solution is proposed in this paper [5] by the authors to support dynamic scheduling in the production environment. It shows the implementation in the real-world environment showing implementation of quality control measures in the pharmaceutical industry. This paper focuses on supporting the implementation and rescheduling triggers.

In [6], a workload prediction method and resource provisioning technique are presented utilizing feed-forward ANN and the results are applied for resource provisioning. On the results, a marine predator is applied. Opposition-based learning is applied for training multi-layer ANN and multi-workflow scheduling. Makespan is the main factor to be optimized. NASA and Saskatchewan datasets are used.

A dynamic scheduling algorithm called Mayfly Taylor Optimisation is proposed by the authors [7]. It is based on a deep-q-network which showed better performance in terms of energy consumption, SLA and computation cost. A better convergence is shown with the help of the proposed algorithm.

In [8], a genetic programming hyper-heuristic (GPHH) approach is proposed to generate multiple heuristics for multi-objective DWS.

In [9], DWS genetic programming algorithm is proposed to generate scheduling heuristics for workflow scheduling. It was implemented for a single-objective function i.e. makespan. The makespan is minimized for long sequence workflows that are arriving dynamically.

In [10], High-Level Heuristic Dynamic Workflow Scheduling Genetic Programming (HLH-DSGP) algorithm was proposed to evolve high-level heuristics. HLH-DSGP considered multiple aspects of scheduling and balanced them by jointly applying heuristics.

The authors [11] proposed a stochastic multi-WDS algorithm called SMWDSA. It was designed for deadline constraints for optimizing the cost factor. This algorithm has three main stages pre-processing, scheduling, and feedback. Task sub-deadlines are designed for each task in a multi-workflow. This process of scheduling is based on minimal time slot availability to minimize the scheduling cost while meeting the constraints and deadlines. The experiments are conducted on both static and real-world data.

A DWS technique that incorporates dynamism on resource failures is proposed in [12] by the authors. An evolutionary algorithm NSGA-II and its variants DNSGA-II-HM, DNSGAII-A, DNSGA-II-B, DNSGA-II-RI and dynamic extension of the multi-objective particle swarm optimization algorithm (DMOPSO) are implemented and compared to show the improvements. Five main objectives were considered which are makespan, the cost and the energy, and maximization of reliability and resource utilization.

A competent DWS[13] is proposed for resource utilization of the DC using the VM reusability technique and completing the execution of the workflows of the multi-workflow application within their deadline correspondingly.

III. MOTIVATION

As the literature study shows work done in the field of workflow scheduling, but, a very little work has been done considering dynamic multi-workflow scheduling scenarios. In past 5 years, researches have been taking place in dynamic scheduling so as to consider real-world scenarios. Also, the multi-workflows are one of the growing topics of discussion in workflow scheduling. A scenario can't always follow a single workflow, it can be a combination of two or more workflows. So, keeping this in mind, this concept of multi-workflow was designed. All this research and the ideas of being more realistic in drawing the aspects of real-world scenarios in a workflow environment have drawn our interest in this research field.

IV. PROBLEM DOMAIN

The above discussion clearly explains workflow scheduling which is a part of resource scheduling done in cloud and fog computing. As cloud computing is the main domain, a lot of research has been done in order to optimize the computing environment. The cloud services are extended by fog to the edge of the network. Resource scheduling is an aspect and an important factor which schedules the resources in such a manner that they work together and gives the most optimized results. Workflow scheduling is a type of scheduling that is addressed to a particular type of scheduling problems. Thus, the problem domain is to optimize workflow scheduling in fog environment.

V. STATEMENT

The solution towards getting an optimized result is required for various metrics that affect workflow scheduling. The dynamicity is required to test the results on a real-time environment. The workflow scheduling is one such aspect that is needed to be optimized and varies according to the multiple workflows included and how the tasks flow in the system. This is a new research area that focuses on dynamic environment for multi-workflow scheduling.

VI. INNOVATIVE CONTENT

This study as mentioned in the research paper above is focused on explaining the new field that is being less explored till now and is now attracting the researchers to work towards multi-workflow scheduling in dynamic environment. A lot of work has been done workflow scheduling and in last 5 years, many researchers have started their work focused on dynamic scheduling, but, focusing multi-workflow with dynamicity of the present environment is still very less explored field. This field is seeking attention as it is more relatable to the real-time environment. Here, we have case studied healthcare field as a subject of dynamic multi-workflow environment. A thorough study is performed on this field. Pros and cons are discussed along with the challenges and issues.

VII. CASE STUDY

Workflow scheduling is a constraint classification problem (CSP) where a set of constraints are set for CSP. Many algorithms are used to solve CSPs like backtracking, etc. these algorithms have limitations and their properties are enhanced by heuristic techniques. For diversified and scattered constraints, special CSPs are used which fir the dynamic behavior of the CSP algorithms.

Workflow scheduling in the healthcare department is difficult because of the different scenarios which need coordination of multiple resources and procedures' timings. No industry involves more workflows than healthcare. The workflows that can be mentioned are:

a) Care coordination workflow-

- Outreach
- Scheduling
- Insurance verification
- Appointment reminders

b) Registration workflows-

- Check-in
- Demographic verification
- Co-pay collection
- Informed consent collection

c) Intake workflows-

- Medical history review
- Family history review
- Allergy review

d) Documentation workflows-

- Clinical note creation
- Order Placement

e) Revenue Integrity workflows-

- Charge capture
- Charge entry
- Coding
- Claim editing
- Claim attachments

f) Billing workflows-

- Denial management
- Payment posting
- Claim status updating
- Utilization review
- Clinical documentation improvement
- Fax management and routing

Working on these workflows simultaneously is a tedious task. Also, with the regular workflows, the emergencies and urgencies have a great impact on the workflow of the tasks too. For example, a doctor needs certain machines to start the procedure like an X-ray machine before a root canal and also a lab technician to perform the x-ray, at the same time an emergency patient arrives with a broken elbow, so, the time slot, resources and the required personnel have to be allotted to the emergency patients which will require multiple changes in the pre-defined schedules.

Every task in the healthcare field is associated with deadlines to ensure the quality of service. However, maintaining an appropriate schedule for every workflow on the respective machines is not a possibility thus, compromising with low energy consumption, minimum makespan feature, etc. Rough allocation can lead to deadline violations, ignorance of patients, disruption of emergency services, high energy utilization, wastage of resources, etc.

In contrast, if all tasks of healthcare workflow applications run on slower computing machines, the task execution time will increase, and the deadline will be violated. Therefore, researchers are focusing on developing methods to find a good trade-off between makespan and energy consumption when scheduling workflow tasks in a cloud, especially in the healthcare field.

In a dynamic environment, a lack of control over scheduling can be chaotic and highly inefficient. Current methods of scheduling are not efficient enough to be used dynamically and can handle dynamic situations. Additionally, the current methods fail to adequately address the issue of uncertainty. Therefore, alternative systems and methods are required to integrate scheduling and provide timely, efficient, and centralized scheduling of workflow.
The dynamic multi-workflow scheduling techniques provide improved methods and apparatus for dynamic workflow scheduling that is going on in parallel within a dynamic environment. The methods of scheduling use a computer system to receive one or more work orders, determine scheduling constraints for fulfilling the orders, determine the availability of one or more resources for fulfilling the orders, and schedule one or more orders based on the scheduling constraints and the availability of the one or more resources. The knowledge base has at least one temporal constraint list, a resource constraint list, a medical constraint list, a timing list, and a workflow template. The scheduling system uses CSP algorithms to determine the schedules that can run parallelly and keep the rest in a workflow structure. After the schedule is determined, the resources are reserved and schedules are generated for each of the orders and the resources. In one or more embodiments, multiple possible schedules are determined and one or more parameters of the schedules are ranked.
In healthcare, the term “dynamic” is everything around which tasks revolve. The multi-workflows as mentioned here are processed according to their data flows and are thus, tried to obtain the Pareto optimal solutions by finding their optimal solutions considering various factors.

VIII. Application Areas

The workflow can be applied to any domain to enhance the automation process of the workflows. The factors affecting workflows are workflow types, VM types, number of tasks, costing, delays, constraints, etc.
Any workflow systems [14] includes the processing sequence from problem identification, collection of datasets, load & summarize datasets, data segregation, model evaluation, feature scaling, algorithm or technique to be applied for the model, training the model, model validation and prediction.
The applications areas [15] where dynamic workflow scheduling can be applied are listed below:

- Healthcare services
- Object detection systems
- Plant-based disease detection system
- Gaming
- Speech Emotion Analysis
- Recommendation Systems
- Cyber-physical systems
- Traffic prediction systems

IX. Current Challenges And Issues

There are tonnes of challenges faced by new technologies that are coming these days. The scheduling of the work along with its proper flow is a huge task. A few of the challenges faced in workflow scheduling in fog environments are mentioned below:

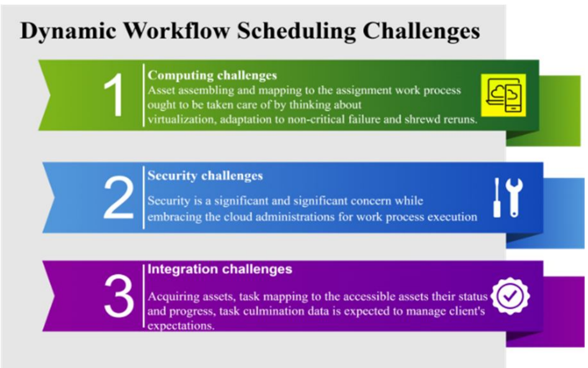

Fig 5. Dynamic Workflow Scheduling Challenges

A. Issues with workflow scheduling

A few of the issues that can be discussed in both dynamic workflow and multi-workflow are also discussed in table 1 and table 2.

TABLE I. ISSUES IN DYNAMIC WORKFLOW SCHEDULING

Issue	Description
Complexity	It increases complexity
Complicate concurrency	The complicated re-engineering process to increase concurrency
Completion time	Variable duration of completing the individual tasks
Resource utilization rate	Very high resource utilization rate
Makespan	Makespan could increase in a few cases
Variability	deal with cloud infrastructure variability

TABLE II. ISSUES IN MULTI-WORKFLOW SCHEDULING

Issue	Description
Execution time	Increases the overall execution time
Overhead	Increased overhead
Power consumption	More power consumption than a single workflow
Makespan	Makespan could increase in a few cases

X. CONCLUSION AND FUTURE WORK

In this paper, a brief discussion of dynamic multi-workflow scheduling is presented. It is made of two techniques i.e. multi-workflow scheduling and dynamic workflow scheduling. A literature survey of the research papers about this technique that have been published since 2019 is presented. The various factors that are affected by workflow scheduling are mentioned. We present a study for dynamic workflow scheduling in fog computing along with a case study on a real-world application. Lastly, various challenges and issues are discussed for DWS in fog environments. This paper introduces a concise overview of fog environments, how they can as an extension of the cloud, workflow scheduling and the concept of DMWS. There is a possibility to improve the work for real-life scenarios on various parameters such as makespan, execution time, total cost, resource utilization rate, quality of service parameters, etc. In the future, we would conduct research on dynamic workflow scheduling using the concept of optimization techniques by including more relevant studies and trying to implement them in DWS applications.

REFERENCES

[1] Z. Zhu, G. Zhang, M. Li, and X. Liu, “Evolutionary multi-objective workflow scheduling in cloud,” IEEE Transactions on parallel and distributed Systems, vol. 27, no. 5, pp. 1344–1357, 2015.

[2] Y. Yu, H. Ma, and G. Chen, “Achieving multi-objective scheduling of heterogeneous workflows in cloud through a genetic programming-based approach,” in 2021 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2021, pp. 1880–1887.

[3] K.-R. Escott, H. Ma, and G. Chen, “Genetic programming based hyper heuristic approach for dynamic workflow scheduling in the cloud,” in International Conference on Database and Expert Systems Applications, Springer, 2020, pp. 76–90.

[4] Y. Yang, G. Chen, H. Ma, M. Zhang, and V. Huang, “Budget and sla aware dynamic workflow scheduling in cloud computing with heterogeneous resources,” in 2021 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2021, pp. 2141–2148.

[5] Coito, Tiago, et al. "Integration of industrial IoT architectures for dynamic scheduling." Computers & Industrial Engineering 171 (2022): 108387.

[6] Kuang, Fangjun, Zhangze Xu, and Mohammad Masdari. "Multi-workflow scheduling and resource provisioning in Mobile Edge Computing using opposition-based Marine-Predator Algorithm." Pervasive and Mobile Computing 87 (2022): 101715.

[7] Shruthi, G., et al. "Mayfly taylor optimisation-based scheduling algorithm with deep reinforcement learning for dynamic scheduling in fog-cloud computing." Applied Computational Intelligence and Soft Computing 2022 (2022).

[8] Yu, Yongbo, et al. "A Genetic Programming-Based Hyper-Heuristic Approach for Multi-Objective Dynamic Workflow Scheduling in Cloud Environment." 2022 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2022.

[9] K.-R. Escott, H. Ma, and G. Chen, “Genetic programming based hyper heuristic approach for dynamic workflow scheduling in the cloud,” in International Conference on Database and Expert Systems Applications.Springer, 2020, pp. 76–90.

[10] K.-R. Escott, H. Ma, and G. Chen, “A genetic programming hyperheuristic approach to design high-level heuristics for dynamic workflow scheduling in cloud,” in 2020 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2020, pp. 3141–3148.

[11] Ye, Lingjuan, et al. "Dynamic Scheduling Stochastic Multiworkflows With Deadline Constraints in Clouds." IEEE Transactions on Automation Science and Engineering (2022).

[12] Ismayilov, Goshgar, and Haluk Rahmi Topcuoglu. "Dynamic multi-objective workflow scheduling for cloud computing based on evolutionary algorithms." 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion). IEEE, 2018.
[13] Adhikari, Mainak, and Santanu Koley. "Cloud computing: a multi-workflow scheduling algorithm with dynamic reusability." Arabian Journal for Science and Engineering 43.2 (2018): 645-660.
[14] Menaka, M., and KS Sendhil Kumar. "Workflow scheduling in cloud environment–Challenges, tools, limitations & methodologies: A review." Measurement: Sensors (2022): 100436.
[15] D. Singh and R. Garg, "A robust multi-objective optimization to workflow scheduling for dynamic grid," in Proceedings of the International Conference on Advances in Computing and Artificial Intelligence, 2011, pp. 183–188.