IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 2747~2753
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp2747-2753

Performance aware algorithm design for elastic resource workflow management of cluster consolidation to handle enterprise big data

B. J. D. Kalyani[1], Pannala Krishna Murthy[2], Sarabu Neelima[3]
[1]Department of Computer Science and Engineering, Institute of Aeronautical Engineering, Hyderabad, India
[2]Department of Electrical and Electronics Engineering, Sri Chaitanya Institute of Technology and Research, Khammam, India
[2]Department of Computer Science and Engineering, Priyadarshini Institute of Science and Technology for Women, Khammam, India

Article Info

Article history:

Received Mar 21, 2023
Revised Jan 31, 2024
Accepted Mar 9, 2024

Keywords:

Big data
Elasticity
Hetrogeneous workflows
Multi cloud computing
Resource provisioning

ABSTRACT

Integration and deployment of big data and business analytics application with cloud computing are more attractive as a service and are trending practice. This hybrid workflow is rapidly increasing and will trigger a revolution for enterprise data handling, information retrieval and computing. This paper presents hybrid workflow management framework for big data and multi cloud computing systems in a two-step approach. Linear optimization-based resource assessment algorithm is planned in the first step. Cluster oriented elastic resource allocation and workflow management techniques are concentrated in the second step. This paper also focus on performance evaluation parameters includes execution time, through put with multi task work flow optimization model. The proposed framework is efficiently managed the implementation of hybrid workflows by finetuning the evaluation attributes and provides improvement in terms of response time an average of 6%.

Corresponding Author:

B. J. D. Kalyani
Department of Computer Science and Engineering, Institute of Aeronautical Engineering
Dundigal, Hyderabad, Telangana, India
Email: kjd_kalyani@yahoo.co.in

1. INTRODUCTION

Voluminous data processing and pre-processing [1] is required for big data applications has become a major challenge in several incipient domains including scientific, space research, gaming [2], astronomy [3] and healthcare [4]. The need for real data analytics is recognized by the companies like banks are focused on detection of frauds in based on analysing transactional data and smart cities [5] by analysing data from various data sources includes traffic cameras, social media, remote sensing data [6], and global positioning system (GPS) data. For enterprises the cloud based bigdata applications [7] provides business intelligence [8], business strategy adoption and strategies for customer retention. Graphical processing unit (GPUs), tera bytes of storage, datacentres and high speed inter connections are demanded for deployment of hybrid cloud and big data applications. Hence organizations select the cloud computing as fundamental resource provisioning platform [9] to their big data applications. Although each piece of technology has value on its own, many businesses are attempting to integrate them to profit from security and on-demand services. Cloud computing is preferred technology for enterprises to maintain their transactions on demand, reliable deployment of big data in cloud. With the help of cloud computing [10], enterprises can perform better data analysis from the massive amounts of structured and unstructured data [11] in their data processing. This feature of the cloud is origin for the migration of cloud computing across numerous industries and enterprises. Multi cloud computing systems are beneficial for enterprises to implement when integrated to the large-scale big data resources that organizations have used before. Cloud computing also provides platform enables companies to integrate data from numerous different heterogeneous sources with different data formats and can produce better visualization of results with a more consistent performance [12] to facilitate decision making.

In multi cloud environment cooperative virtual machine [13] form as cluster as processing streams with nearby resources and form as middleware layer to backing cloud services. Clusters has a substantial role in dealing out massive data and only uploads processed data to clouds in multi cloud computing systems for improvement in service availability. Hybrid workflow management require the development of well-organized resource provisioning and forecast techniques which coordinate the execution of hybrid workflows [14] on various clusters.

2. RELATED WORK

Saovapakhiran *et al.* [15] focus on coordination and controlling of clusters in multi cloud environment. The authors concentrate on quality of service (QoS) parameters, how to optimize these parameters during integration of clusters to provide cloud services. Latency based stream processing [16] for computational oriented work flow scheduling is demonstrated by Udoh and Kotonya [17]. The author described the procedure for data aggregation, network synchronization and model prediction of clusters in big data applications. Mastroianni *et al.* [18] illustrated significance of elastic state, dynamic virtual machines consolidation and job scheduling in bigdata framework. Shi and Chen [19] illustrates cost time optimization algorithm for deadline and budget distribution among clusters. The scheduling of tasks is carried out with parent and child groups depending on service request.

3. METHODOLOGY

The cluster cloud model is suitable for hybrid task execution paths, because the watercourse tasks with latency sensitivity [20] can benefit from the availability of resources, whereas batch tasks with hefty workloads can be handled at powerful computation nodes in the multi cloud. Generally, hybrid workflow framework includes three layers namely physical layer, cluster layer and application layer. Physical layer contains servers, internet of things (IoT) sensors that provides fundamental resources for multi cloud infrastructure and storage that handles computational intense applications [21] includes business intelligence, complex visualization [22], and data analytics [23]. Cluster layer facilitates data communication between workflow tasks through hybrid resource scheduling algorithm for multi cloud and big data environment. Application layer provides interaction layer for users and is responsible for collecting information and performing operations in order to provide service.

Workflow management is required to estimate resource allocation for workflows based on quality attributes to choose efficient virtual machines for task execution with the help of selected scheduling algorithm [24]. In proposed work hybrid workflow is a combination of stream and batch tasks. The start and end tasks are fake tasks and not considered for hybrid workflow execution. The main aim of hybrid workflow management is to provide best cluster-based task execution framework to provide service with minimum execution time as in Figure 1.

Hybrid workflow scheduling management allows seamless cooperation between clusters to select execution path based on quality parameters. The resource assessment [25] for the cluster is the optimized workflow configuration that is combination of execution time and number of clusters. In the proposed work a cluster can be number of virtual machines as a single core. After resource assessment allocation and scheduling to tasks of each of a cluster is carried out in the multi cloud environment. Each cluster need to consider execution time (T) and cost (C) and need to achieve as (1).

$$Min(T, C) \tag{1}$$

Workflow configuration is carried out with cluster request arrival rate and minimum execution time with a smaller number of resources (section-1), prioritize the cluster based on section-1 attributes then assign cluster to the path with the help of cluster-oriented hybrid workflow management algorithm. This approach can enhance the efficiency and accuracy of data processing and analysis in scenarios where data exhibits natural clusters or groups with different characteristics.

Algorithm: Cluster oriented hybrid workflow management

```
Input: Group of Clusters (G) and Available Computational Resources (D)
Procedure COHWM (G, D)
        Initialize P is set of Paths: P = { }
        Findexecutionpaths(G)
        while clusters do
```

```
                Set cluster to path
                Broadcast message to all clusters
                while cluster proposal arrives do
                        Collect proposal from clusters
                        Prioritize clusters
                end while
                Broadcast message for cluster allocation
                while clusters parameter arrives do
                        Collect parameter message from cluster
                end while
                while cluster disagree proposal do
                        for all cloud users do
                                Build new cluster
                        end for
                end while
        end while
End Procedure
```

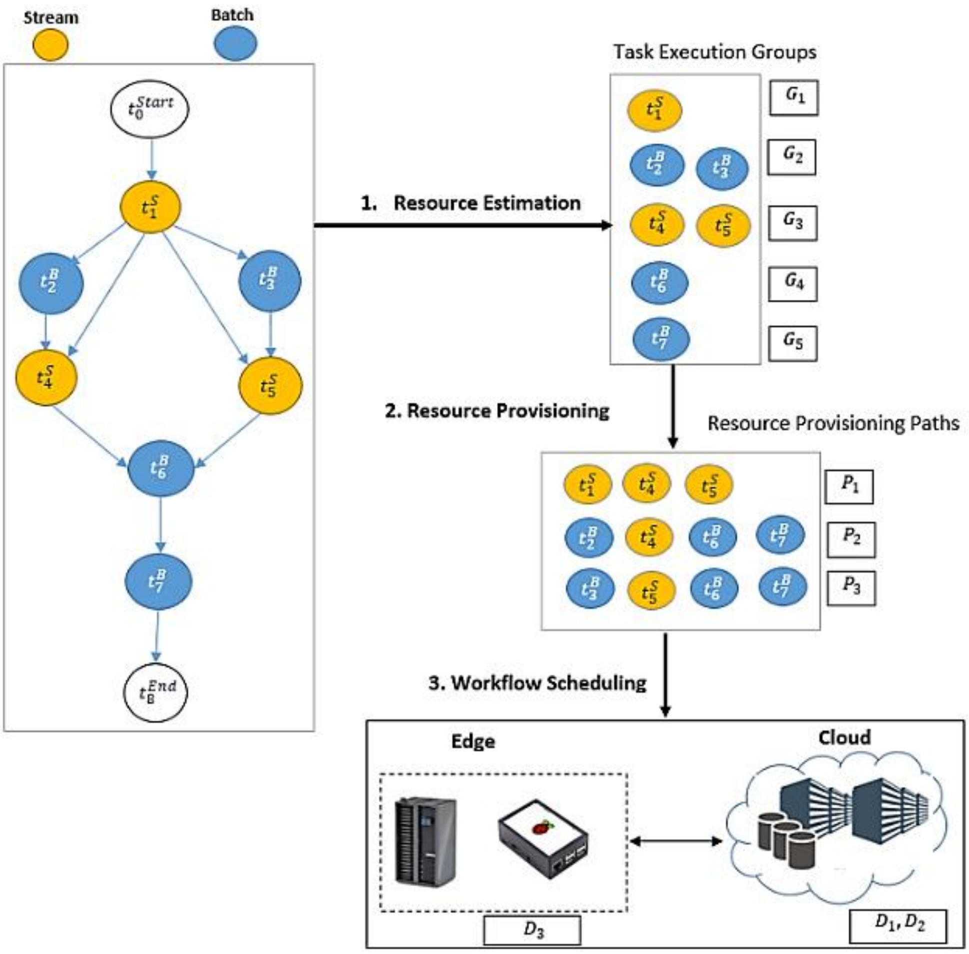

Figure 1. Hybrid workflow scheduling

4. IMPLEMENTATION AND RESULTS

A cluster-oriented hybrid workflow algorithm is a type of algorithm that combines elements from different workflow and clustering techniques to solve specific problems efficiently. This algorithm is often used in data analysis, machine learning, and optimization tasks. The multi-cloud and big data environment with hybrid workflow is deployed with Peacock [26] as a self-governing component with Java [27] Spark [28] add-ins composed with Scala [29]. The proposed work utilised Sparrow [30] combined with code from Eagle [31] to serve the big data enterprise trials as in Figure 2 and the characteristics of the workload are described in Table 1.

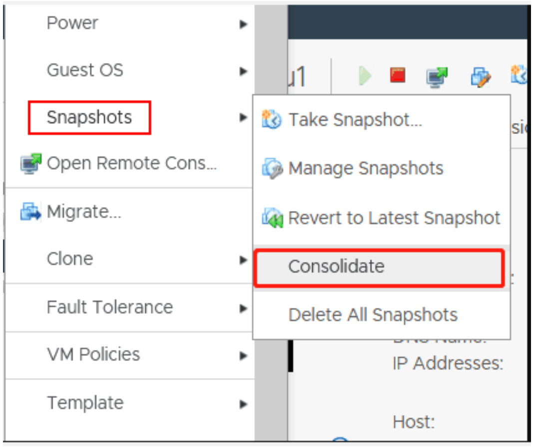

Figure 2. Proposed system environment

Table 1. Workload characteristics

Workloads	Taskset count	Task count	Average task duration
Google	5,04,482	17,80,043	68
Yahoo	29,262	9,92,497	119
Cloudera	21,033	5,76,097	102

To assess average cluster workloads, the task influx time is distributed with a poisson process and a mean task arrival time is estimated based on a predictable average workload percentage, mean task execution time, and mean number of tasks per cluster. Due to heterogeneous tasks, the workload also becomes heterogeneous during the execution of tasks, and the average execution time is 6% and the execution time is illustrated in Figure 3. The proposed work contains 30%, 40%, and 70% light cluster workloads and 100%, 150%, and 200% heavy cluster workloads. The cumulative distribution of task completion for 10,000 clusters is described in Figure 4, (Figure 4(a) Google 300%) illustrates the integrated distribution of tasks termination for 10000 clusters and (Figure 4(b) Google-50%) demonstrates that, with a 50% load, sparrow can only do 2.2% of jobs in less than 100 seconds, compared to 21.6% for Peacock in the same amount of time. As seen in Figure 4(c) Google-300%, when under 300% load, Sparrow completes 0.3% of tasks in less than 100 seconds, compared to 31.8% for Peacock. The Yahoo! trace has longer task durations, so we check for 1000 seconds. At 50% load in Figure 4(d) Yahoo-50%, the percentages for Sparrow and Peacock are in order of 5% and 23.5% but with Cloudera the 300% and 50% comparision is shown in Figure 4(e) Cloudera 300% and Figure 4(f) Cloudera 50% respectively. The workload distribution of a cluster is demonstrated in Figure 5.

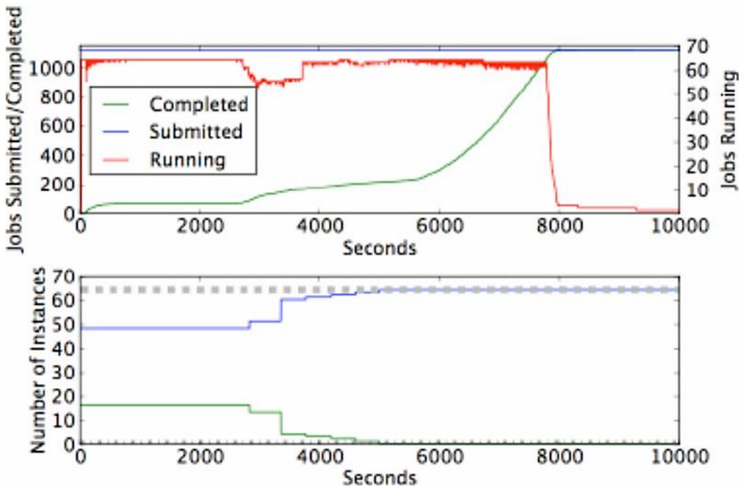

Figure 3. Execution time

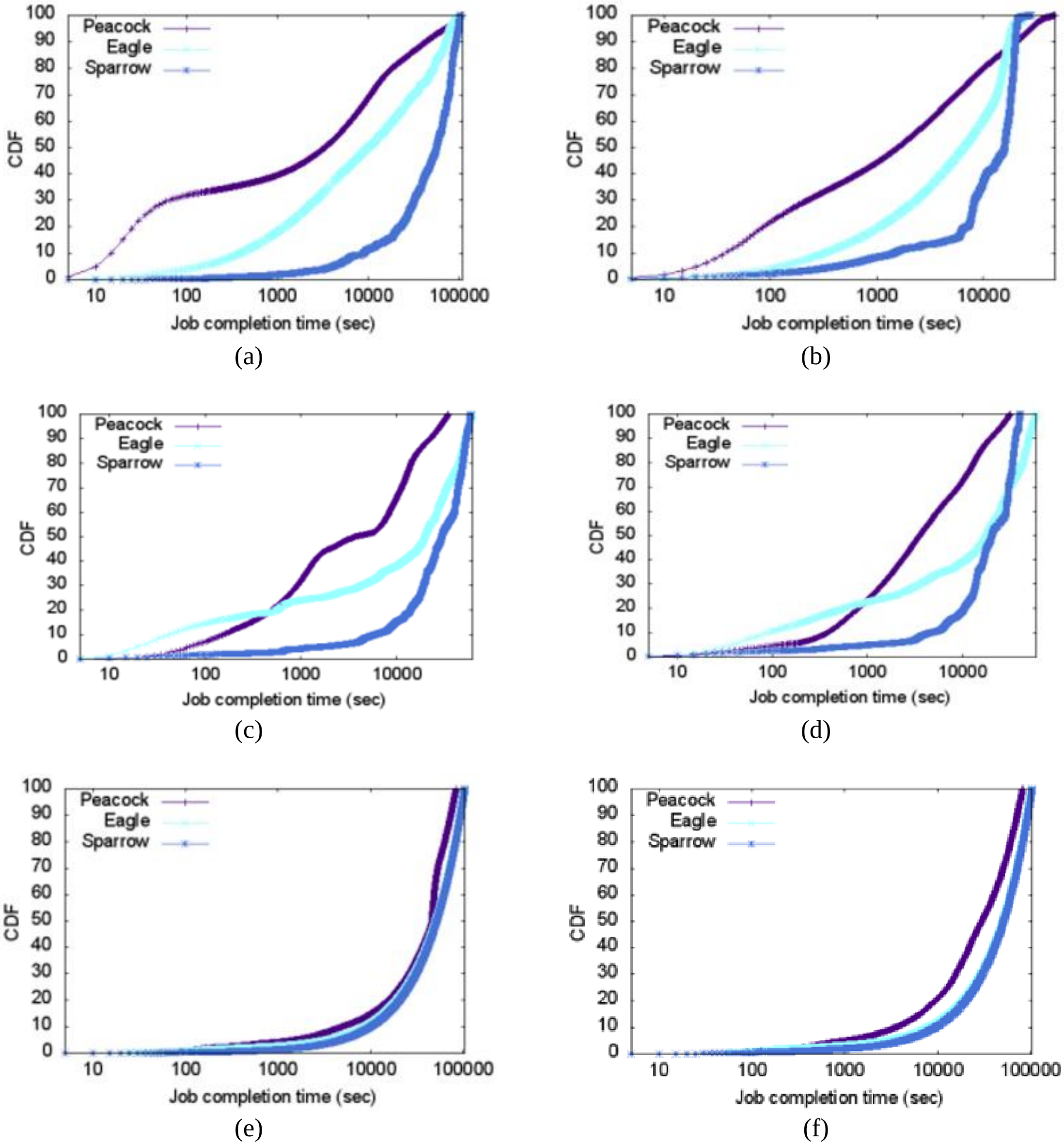

Figure 4. Integrated distribution of tasks termination for 10000 clusters, (a) Google-300%, (b) Google-50%, (c) Yahoo-300%, (d) Yahoo-50%, (e) Cloudera-300%, and (f) Cloudera-50%

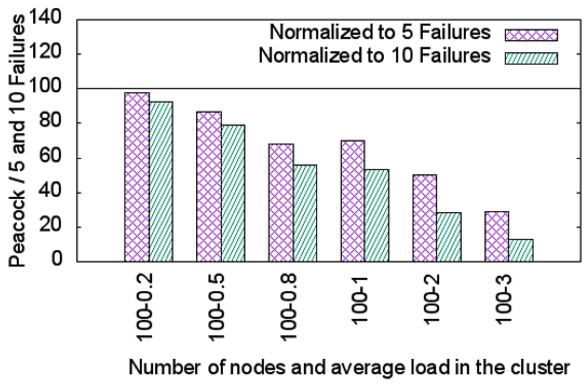

Figure 5. Workload of clusters

5. CONCLUSION

Generally, the big data frameworks split jobs into various parallel processing tasks that are executed with small partition of data with low latency. Such frameworks depend on distributed schedulers to handle the attached overhead. The existing algorithms not efficiently performed during workload variations with heterogeneous jobs. The hybrid workflow management algorithm considers heterogeneous jobs both stream and batch provide improvement in terms of execution time an average of 6%.

REFERENCES

[1] P. Li and J. Cao, "A virtual machine consolidation algorithm based on dynamic load mean and multi-objective optimization in cloud computing," *Sensors*, vol. 22, no. 23, Nov. 2022, doi: 10.3390/s22239154.

[2] N. K. Biswas, S. Banerjee, U. Biswas, and U. Ghosh, "An approach towards development of new linear regression prediction model for reduced energy consumption and SLA violation in the domain of green cloud computing," *Sustainable Energy Technologies and Assessments*, vol. 45, Jun. 2021, doi: 10.1016/j.seta.2021.101087.

[3] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future Generation Computer Systems*, vol. 28, no. 5, pp. 755–768, May 2012, doi: 10.1016/j.future.2011.04.017.

[4] K. Haghshenas, A. Pahlevan, M. Zapater, S. Mohammadi, and D. Atienza, "MAGNETIC: multi-agent machine learning-based approach for energy efficient dynamic consolidation in data centers," *IEEE Transactions on Services Computing*, vol. 15, no. 1, pp. 30–44, Jan. 2022, doi: 10.1109/TSC.2019.2919555.

[5] B. Wang, F. Liu, and W. Lin, "Energy-efficient VM scheduling based on deep reinforcement learning," *Future Generation Computer Systems*, vol. 125, pp. 616–628, Dec. 2021, doi: 10.1016/j.future.2021.07.023.

[6] U. Arshad, M. Aleem, G. Srivastava, and J. C. W. Lin, "Utilizing power consumption and SLA violations using dynamic VM consolidation in cloud data centers," *Renewable and Sustainable Energy Reviews*, vol. 167, Oct. 2022, doi: 10.1016/j.rser.2022.112782.

[7] J. Li, R. Zhang, and Y. Zheng, "QoS-aware and multi-objective virtual machine dynamic scheduling for big data centers in clouds," *Soft Computing*, vol. 26, no. 19, pp. 10239–10252, Oct. 2022, doi: 10.1007/s00500-022-07327-x.

[8] M. H. Sayadnavard, A. T. Haghighat, and A. M. Rahmani, "A multi-objective approach for energy-efficient and reliable dynamic VM consolidation in cloud data centers," *Engineering Science and Technology, an International Journal*, vol. 26, Feb. 2022, doi: 10.1016/j.jestch.2021.04.014.

[9] K. Karmakar, R. K. Das, and S. Khatua, "An ACO-based multi-objective optimization for cooperating VM placement in cloud data center," *Journal of Supercomputing*, vol. 78, no. 3, pp. 3093–3121, Feb. 2022, doi: 10.1007/s11227-021-03978-z.

[10] J. Peake, M. Amos, N. Costen, G. Masala, and H. Lloyd, "PACO-VMP: parallel ant colony optimization for virtual machine placement," *Future Generation Computer Systems*, vol. 129, pp. 174–186, Apr. 2022, doi: 10.1016/j.future.2021.11.019.

[11] Z. Li, X. Yu, L. Yu, S. Guo, and V. Chang, "Energy-efficient and quality-aware VM consolidation method," *Future Generation Computer Systems*, vol. 102, pp. 789–809, Jan. 2020, doi: 10.1016/j.future.2019.08.004.

[12] H. Xiao, Z. Hu, and K. Li, "Multi-objective vm consolidation based on thresholds and ant colony system in cloud computing," *IEEE Access*, vol. 7, pp. 53441–53453, 2019, doi: 10.1109/ACCESS.2019.2912722.

[13] F. F. Moges and S. L. Abebe, "Energy-aware VM placement algorithms for the OpenStack Neat consolidation framework," *Journal of Cloud Computing*, vol. 8, no. 1, Dec. 2019, doi: 10.1186/s13677-019-0126-y.

[14] H. Y. Yun, S. H. Jin, and K. S. Kim, "Workload stability-aware virtual machine consolidation using adaptive harmony search in cloud datacenters," *Applied Sciences*, vol. 11, no. 2, pp. 1–23, Jan. 2021, doi: 10.3390/app11020798.

[15] B. Saovapakhiran, G. Michailidis, and M. Devetsikiotis, "Aggregated-DAG scheduling for job flow maximization in heterogeneous cloud computing," *GLOBECOM - IEEE Global Telecommunications Conference*, 2011, doi: 10.1109/GLOCOM.2011.6133611.

[16] A. Verma and S. Kaushal, "Deadline and budget distribution based cost- time optimization workflow scheduling algorithm for cloud," in *International Conference on Recent Advances and Future Trends in Information Technology*, vol. 4, pp. 1–4, 2012.

[17] I. S. Udoh and G. Kotonya, "Developing IoT applications: challenges and frameworks," *IET Cyber-Physical Systems: Theory & Applications*, vol. 3, no. 2, pp. 65–72, 2018, doi: 10.1049/iet-cps.2017.0068.

[18] C. Mastroianni, M. Meo, and G. Papuzzo, "Probabilistic consolidation of virtual machines in self-organizing cloud data centers," *IEEE Transactions on Cloud Computing*, vol. 1, no. 2, pp. 215–228, 2013, doi: 10.1109/TCC.2013.17.

[19] D. Shi and T. Chen, "Optimal periodic scheduling of sensor networks: A branch and bound approach," *Systems and Control Letters*, vol. 62, no. 9, pp. 732–738, 2013, doi: 10.1016/j.sysconle.2013.04.012.

[20] T. Renugadevi, K. Geetha, K. Muthukumar, and Z. W. Geem, "Optimized energy cost and carbon emission-aware virtual machine allocation in sustainable data centers," *Sustainability*, vol. 12, no. 16, 2020, doi: 10.3390/SU12166383.

[21] M. Zakarya and L. Gillam, "Managing energy, performance and cost in large scale heterogeneous datacenters using migrations," *Future Generation Computer Systems*, vol. 93, pp. 529–547, 2019, doi: 10.1016/j.future.2018.10.044.

[22] S. Jangiti and S. Sriram. V.S., "Scalable and direct vector bin-packing heuristic based on residual resource ratios for virtual machine placement in cloud data centers," *Computers and Electrical Engineering*, vol. 68, pp. 44–61, 2018, doi: 10.1016/j.compeleceng.2018.03.029.

[23] T. Fernando, N. Gureev, M. Matskin, M. Zwick, and T. Natschlager, "WorkflowDSL: scalable workflow execution with provenance for data analysis applications," *International Computer Software and Applications Conference*, vol. 1, pp. 774–779, 2018, doi: 10.1109/COMPSAC.2018.00115.

[24] M. Mezmaz *et al.*, "A parallel bi-objective hybrid metaheuristic for energy-aware scheduling for cloud computing systems," *Journal of Parallel and Distributed Computing*, vol. 71, no. 11, pp. 1497–1508, 2011, doi: 10.1016/j.jpdc.2011.04.007.

[25] F. Quesnel and A. Lèbre, "Cooperative dynamic scheduling of virtual machines in distributed systems," *Euro-Par 2011: Parallel Processing Workshops*, vol. 7156, pp. 457–466, 2012, doi: 10.1007/978-3-642-29740-3_51.

[26] M. Khelghatdoust and V. Gramoli, "Peacock: probe-based scheduling of jobs by rotating between elastic queues," *Euro-Par 2018: Parallel Processing: 24th International Conference on Parallel and Distributed Computing*, pp. 178–191, 2018, doi: 10.1007/978-3-319-96983-1_13.

[27] M. Zaharia *et al.*, "Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing," in *9th USENIX symposium on networked systems design and implementation (NSDI 12)*, pp. 15–28, 2012.

[28] A. Núñez, J. L. V. -Poletti, A. C. Caminero, G. G. Castañé, J. Carretero, and I. M. Llorente, "ICanCloud: a flexible and scalable cloud infrastructure simulator," *Journal of Grid Computing*, vol. 10, no. 1, pp. 185–209, 2012, doi: 10.1007/s10723-012-9208-5.

[29] Y. Oh, J. Choi, E. Song, M. Kim, and Y. Kim, "A SLA-based Spark cluster scaling method in cloud environment," *18th Asia-Pacific Network Operations and Management Symposium, APNOMS 2016: Management of Softwarized Infrastructure - Proceedings*, 2016, doi: 10.1109/APNOMS.2016.7737242.

[30] K. Ousterhout, P. Wendell, M. Zaharia, and I. Stoica, "Sparrow: distributed, low latency scheduling," *SOSP 2013 - Proceedings of the 24th ACM Symposium on Operating Systems Principles*, pp. 69–84, 2013, doi: 10.1145/2517349.2522716.

[31] P. Delgado, D. Didona, F. Dinu, and W. Zwaenepoel, "Job-aware scheduling in eagle: divide and stick to your probes," *Proceedings of the 7th ACM Symposium on Cloud Computing, SoCC 2016*, pp. 497–509, 2016, doi: 10.1145/2987550.2987563.

BIOGRAPHIES OF AUTHORS

B. J. D. Kalyani working as Associate Professor in the Department of Computer Science and Engineering at Institute of Aeronautical Engineering. Awarded doctorate from Acharya Nagarjuna University, Guntur in the area of Cloud computing. She is having 6 years of industry and 12 years of teaching experience. She guided 8 PG projects and 20 UG projects. She published 15 papers in various national/international conferences and journals. Her research areas of interest are cloud computing, data analytics, business intelligence, software engineering and database management systems. She is associated with several committees like R & D, discipline, anti ragging and hospitality. She acted as convener for R&D and actively involved in organizing ICRTEMMS-2018 as registration committee Convener. She can be contacted at email: bjd.kalyani@iare.ac.in.

Pannala Krishna Murthy has a Doctorate Degree in Electrical Engineering from JNTU, Hyderabad. Over 50 International Publications to his credit. One research scholar received Ph.D. from JNTUH, Hyderabad in January 2019 and five Ph.D. scholars registered with various Universities of Telangana, Andhra Pradesh, Madhya Pradesh, and Tamilnadu are working under his supervision. He guided more than 40 UG projects and 28 PG projects. The research areas include, electrical power systems, power electronics, power quality, data mining, optimization algorithms, electrical engineering material characterization, and bio-medical testing and analysis. He is a fellow member IE India, J.Cs. and senior member IEEE, life member, APSMS, ISTE, New Delhi and also member IRC Scientific and Technical Committee & Editorial Review Board on Electrical and Computer Engineering (WASET). He can be contacted at email: krishnamurthy.pannala@gmail.com.

Sarabu Neelima is Dean –Quality Management System and Professor in Computer Science and Engineering, Priyadarshini Institute of Science and Technology for Women (PRIW) is having more than 20 years of academic career in teaching and administration, 7 years of research in Engineering institutions. She obtained her M.Tech. (CSE) from JNTUH in 2011. She has done her Ph.D. research in the field of Data Mining at JNTUH and awarded in 2019. She has guided a number of UG and PG projects. She has 25 publications in journals and conferences at national and international level. Her research work was published in Elsevier, Springer, IEEE, SCOPUS. She can be contacted at email: sarabu.neelima@gmail.com.