Chapter 7

Conclusions and Future research directions

7.1 Conclusions

Cloud computing is the prominent technology that provides on-demand computing infrastructure for solving large-scale scientific problems. It enables the users to access virtually infinite resources that can be easily configured and scaled according to their application requirements. Because of these advantages, cloud user spending is rapidly increasing. As a consequence CDCs complexity is increasing beyond human ability and consuming a large amount of electrical energy. This over energy utilization impacting both the Total Cost of Ownership (TCO) and the environment by releasing green gases. Moreover, deploying large-scale workflows on clouds is a challenging problem. This thesis has proposed energy-efficient techniques for scheduling workflows and VM consolidation in clouds without compromising the QoS requirements.

Chapter 1 describes the aim of the thesis by defining the problem addressed. It also introduced key concepts, motivation, and a summary of the primary contributions made to enhance the field. Chapter 2 provides a detailed awareness of the energy-aware workflow scheduling algorithms in clouds. It surveyed over 30 energy-efficient scheduling techniques for workflows and presented a detailed review report.

Chapter 3 propose energy and cost-aware scheduling (ECWS) algorithm with the objectives of reducing energy consumption, execution cost, and maximizing resource utilization. The algorithm uses sub-algorithms to identify energy-inefficient VMs by calculating parameter ratios of effectiveness (RE). Then it selects optimal VMs based on RE values by considering workflow deadline. Further, it schedules tasks to selected VMs. Finally, it uses a slack algorithm for further reducing energy consumption. The simulation experimental results with four scientific workflows demonstrated significant energy conservation, maximizing resource utilization, and cost-saving than related three well-known algorithms, HEFT, EES, and EHEFT.

Chapter 4 presents an energy-efficient and reliability-aware workflow scheduling algorithm for the clouds. It is common to use DVFS for energy saving but with the frequency scale down of processors, the probability of transient errors increasing. So, a tradeoff being made between energy-saving and system reliability. The proposed approach consists of five sub-algorithms such as task rank calculation, task clustering, sub-target time distribution, cluster-VM mapping algorithm, and slack algorithm. It improves system reliability and while reducing energy utilization. The performance of the proposed algorithm was evaluated through several numerical experiments. And the experimental results show that the proposed approach can achieve a significant amount of energy-saving while maximizing the application reliability without regard to the diverse workow structures.

Chapter 5 presents metaheuristic-based virtual machine consolidation (WWO-VMC) algorithms. It uses shallow-water waves inspired algorithm called water wave optimization (WWO) for energy-efficient VM consolidation. The proposed algorithms categorize physical machines into underloaded, overloaded, and normally loaded. It migrates VMs from overloaded for load balancing and it migrates all VMs from underloaded to normally loaded servers to switch off idle machines. Performance evaluation of WWO-VMC performed using Cloudsim plus tool with different workloads. The simulation results demonstrated that it surpassed LR, MAD, and IQR algorithms in energy saving.

Finally, Chapter 6 presents a hybrid algorithm for energy-efficient scheduling algorithm for workflows in clouds with virtual machine consolidation using water wave optimization. This algorithm runs in two phases task scheduling and VM consolidation. In the first phase, it schedules tasks to energy-efficient resources without compromising the performance.

The second phase implements VM consolidating which migrates VMs dynamically for load balancing and consolidate VMs. The idle machines are switched off to save more energy. VM consolidation is efficient in minimizing the number of migrations and active running physical machines. The performance of this algorithm was evaluated on WorkflowSim tool with five different scientific workflows. The simulation experiments demonstrated that it is outperformed well-known related algorithms HEFT, EES, EHEFT, ACO-VMC, and WWO-VMC in energy consumption. Further, it reduces the number of VM migrations and improved resource utilization.

7.2 Future Research Directions

Regardless of the contributions of this thesis, there are many open challenges to further advance this area of research. This section explains favorable unexplored ways for future studies in the area of energy-aware workflow task scheduling in a cloud environment.

The contributions of the thesis investigate the dynamic power utilization of the processors. It is worth considering static power consumption such as power utilization of the memory, storage, etc. We investigated the total dynamic energy utilization of the leased cloud servers while executing workflow applications in IaaS clouds. We did not study the energy utilization for the application life cycle. It is required to investigate workflow life cycle for better estimation of energy consumption.

Security is one of the most desired features of cloud services. It is recommended to secure handling of I/O operations of a few scientific applications. Scientific workflows are big data applications that can take a long time to complete execution in large-scale distributed environments like clouds. These computing platforms are not risk-free. Deploying data-intensive workflows in clouds rise challenges in specifying and scheduling. Failing to establish intermediate data security may cause data alternation. Further, some scientific applications demand high-security operations such workflow applications need to deploy on hybrid clouds. Where more sensitive data is stored in the private cloud and rest deployed on the public cloud. It is very important to study the power utilizations of such workflow applications to study the overall carbon footprints of the environment.

Existing algorithms use static scheduling models. Predicting the performance of the resources dynamically helps in better estimation of the energy utilization. These predictions can be used in scheduling and resource management. It can be benefited from the dynamic nature of cloud technology. For example, algorithms can predict energy performance based on historical data, the dynamic monitoring of the performance of computing resources, and time series analysis on future trends. Further, they can use a predicted model or its original scheduling model to get maximum benefits. The resources performance approximations overheads may have a negative impact on the performance of the scheduling algorithm. Hence, it would be better to consider the tradeoff between performance and accuracy.

Container technology has become popular in recent years. Particularly scientific workflow scheduling applications container technology has emerged as an enhancement to VMs. The container-based virtualization helps in modeling energy-aware scheduling algorithms. It provides flexibility to the applications to access the underlying resource as per their needs.