CHAPTER 7
CONCLUSION AND FUTURE RESEARCH DIRECTIONS

7.1. Conclusion

Resource Management in cloud computing is done effectively with Interference Aware Resource Allocation (IARA) Technique by reducing the interference occurrences on scheduling the task for cloud services. IARA Technique with interference measure provide the support for localization in resource constrained environment, in turn enhances resource scheduling process by improving the processing capability of cloud service provisioning. Moreover, IARA Technique along with multi-tasking based resource scheduler allocates the resources with optimal energy and bandwidth consumption in cloud environment.

The Adaptive Load Balancing (ALB) Approach is presented in modern VM cloud computing data centers, that handle the load demand efficiently for multi-task operation in the cloud environment. ALB supports restriction of information overflow and information losses, and reduces congestion in-between servers with instantaneous gathering of load demand conditions by repetitive query messaging. Transmission of information between the different servers is managed with ALB, to avoid congestion hotspots and data losses caused by load overflow in cloud servers. ALB improves its quality of load balancing service by reducing delays related to communication and data losses due to congestion.

Genetic Clustering with Workload Multi-task (GCWM) Scheduler Scheme organizes the servers in different cluster and manages multiple tasks.

GCWM Scheduler with a fitness function cluster the same type of tasks available in the cloud zone and improve the task scheduling efficiency by assigning optimal resource required. Moreover, it schedules the workload dynamically at continuous time periods without any elapsed time. The scheduler performs multi- tasking on the cloud server efficiently, by reducing the computational cost and minimizing the complexity, with optimal genetic value of the clustered tasks to be processed by the cloud server.

Performance analysis of IARA Technique, ALB Approach and GCWM Scheduler Scheme are conducted with CloudSim Simulator with different cloud environment parameters i.e., number of VM servers, task size, varied load and demand resources such as energy and bandwidth. CloudSim Simulator executes codes developed in java language for the proposed techniques (IARA, ALB and GCWM scheduler) with varied cloud environmental parameters. The proposed techniques are applied on Statlog (Shuttle) Data Set from UCI repository which contains 9 numerical attributes. The proposed work has achieved better results for the performance parameters like interoperability, computational cost, scalability, load balance factor, clustering efficiency, generic based cloud services, multi-task clustering effect, communication cost and computational complexity. More specifically, proposed techniques has improved resource utilization to nearly 5-15%, reduced energy consumption rate to 14-26%, and increased multi-task workload management on the cloud server to 5-25% comparing to existing works.

7.2. Future Research Directions

The present IARA work performs resource allocation of the processes to the tasks required for cloud service in case of under-load and over-load conditions. In case of over-load condition, the migration of the processes is performed from one cloud to other in public cloud service provisioning environment. Future enhancement of proposed work will be extended in the directions of modifying IARA suitable for private and hybrid cloud environment as well.

ALB Approach achieves better performance in terms of energy utilization, load balancing, and evaluation time in handling the load on cloud servers. Further, ALB approach can be extended to investigate the energy drain rate of the cloud servers on managing the over loaded tasks. In future, analysis of ALB approach in maintaining memory requirement, and reliability of multiple type of tasks handling by cloud servers can be evaluated to understand the adaptability in hybrid cloud environment. As a future enhancement, communication constraints for scheduling task in a cloud computing environment can be analyzed along with load demand convergence on varying task requirement.

Further investigations can be carried out on GCWM scheduler in multiple cloud computing platforms to improve the scalability of VM servers. Scalability measure can be proceeded in the direction of deploying GCWM design with Hadoop Platform. The multi-cloud distribution mechanism and algorithms can be used on top of the Hadoop for distributing different applications for highly scalable cloud service provisioning.