CHAPTER 7
RESEARCH FINDINGS

Submission of long running jobs to the scheduler and the analysis of non-interactive workloads/job characteristics requires the proper Virtual Machine (VM) monitoring. The utilization of multi-server architecture increases the scalability of processing tasks and the optimal VM selection. The major research findings of this research work are described as follows. The analysis of integrated workloads is the major requirement to select the optimal VM and to reduce the computational complexity. The introduction of load balancing during the job execution efficiently minimizes the memory required for storage. Fuzzy quantifier utilization and the Integer Linear Programming (ILP) application on the decision-making process optimizes the cloud resources, reduces the time required for job execution.

The cloud server utilization requires on-demand sharing and large data processing. The minimum scalability offers less performance in resource allocation methods. The assigning of more jobs to minimum load capacity VM and minimum jobs to more capacity VM (imbalance state) also leads to more time consumption for execution and response provision. The load balancing introduction in this study handles the imbalance problem with the suitable resource allocation mechanism. The provision of well-balancing among the VM by using the integrated Sequence Tweep Processing (STP) and Dynamic Request Selection (DRS) models maximize the performance rate and throughput. The STP-DRS implementation is based on the assumption of the job requests arrived at the machine are in sequence order. But, the job requests to each VM are either in sequence or non-sequence. Hence, the STP-DRS require an enhancement to carry the non-sequence job requests.

The dynamic nature of load balancing improves the response time without considering the previous state information. The major findings for the proposed scheduling are the CPU utilization improvement, throughput enhancement, minimum waiting time and response time. The task priority-based scheduling in proposed research work reduces the number of migrations and idle/waiting time.
Hence, the hybrid such as optimal VM placement, load balancing and the priority-based scheduling in this research work offers better performance than the existing research works regarding the access storage capacity, data transfer rate and throughput with minimum migrations, response, waiting and idle time.

The lack of locational information in optimal VM placement increases the cost performance during the movement from one geographical location to other. Hence, the economy-based preemption policies to be adapted for proposed research work to investigate the job scheduling performance under cloud-environment. The real-time combination of education, income and occupation are termed as the Socio-Economic Status (SES) and their examination reveals that the inequalities during resource access and distribution. The deciding of priority as the QoS factor and the architecture is required to support the resource trading for the trade-off between the budget and QoS satisfaction. With the centralized policies utilization and the diverse distribution of resources according to time zones, the allocation of resources faced the several limitations such as complexities in data aggregation, remote resourcing capabilities.

The provision of multi-threaded facilities enhances the parallelism through multi-threaded programming (US patents). The exploitation of affinity, locality, and the load balancing are the major strategies in job scheduling. The definition and modification of job characteristics according to the usage to be performed dynamically. Hence, the major discovery of the research work is that the enhancement is required with the dynamic provision to improve the execution speed of the multi-threaded programming. The uniform distribution of processes and resources based on the priority levels, placement of jobs in the global run queue and periodical update of the run queue length supports the detection of load balancing problems and reduce the number of migrations in demand situations.