8. CONCLUSION

8.1 INTRODUCTION

This chapter reviews the significance of the proposed failure handling strategies for guaranteed service delivery to grid users, the results obtained during the present work and suggestions for future research. Before proceeding with the highlights of the work done, the objectives of the thesis stated earlier in the introductory chapter are recalled.

The main objective of the research work is to develop an integrated proactive and reactive failure handling framework for a comprehensive resource management scheme and deploy it at the meta-scheduler. It allocates the appropriate resources for client requests, handles resource, process, match-making and time-out failures and successfully completes the client requests within the deadline such that it maximizes the resource utilization & job success rate and minimizes the job completion time. Based on this objective, it is proposed to develop new strategies for resource and process failure handling, match-making failure handling using SLA, time-out failure handling in advance reservation-based scheduling, scheduling with grouping technique, adaptive failure prevention methods, efficient check-pointing management and flexible failure recovery model to provide more reliable and guaranteed service to grid business applications.

8.2 HIGHLIGHTS OF THE WORK DONE

This research has proposed a framework for the integrated proactive and reactive failure handling strategies for reliable and guaranteed services in computational grid. Six proposed strategies are experimented and evaluated using grid simulator GridSim 5.2 and compared with the well acceptable Min-Min scheduling algorithm. The simulation results reveal that the developed strategies are outperformed compared to Min-Min strategy in terms of job success rate, resource utilization and job completion time, which in turn has improved the grid services for reliable and guaranteed completion of the user job.
The introduction of R-FH as a first strategy has provided reliable job execution for grid user and improved job success rate. This algorithm identifies suitable resources for a client application evaluating the reliability and status of the resources in a highly failure prone grid environment. Selection of reliable and available resources at the entry level has reduced the runtime failure during execution and improved the reliability of the grid system services. This strategy can handle resource failure efficiently and has achieved a better performance than the traditional Min-Min scheduling algorithm in terms of increased job success rate and resource utilization. The R-FH strategy minimizes the number of checkpoints by setting the check-pointing interval dynamically using runtime condition of the job and failure information of the resources and it recovers the failed job using latest checkpoints with the same resource or backup resource in minimum check-pointing time and recovery time. The R-FH strategy has provided 25% reduction in job completion time, achieved about 81% of job completion success rate and 95.4% of resource utilization compared with the Min-Min strategy.

The implementation of P-FH as a second strategy has enhanced the R-FH strategy to handle process failure along with the resource failure and has achieved better resource utilization than R-FH and Min-Min strategies. While executing the user job, whenever the process failure occurs due to CPU overloads and memory overloads, P-FH recovers and completes the failed job using a backup process from the recently saved checkpoint status of the job and returns the successfully completed job to the user. Each resource has a monitoring-service to monitor the process failure. The P-FH strategy assures the availability of checkpoints even in failure of the checkpoint server by storing checkpoints in stand-by checkpoint servers. P-FH has improved the resource utilization by 22% compared with the Min-Min strategy.

The M-FH is implemented as the third strategy and T-FH is implemented as the fourth strategy towards the successful implementation of Q-FH model to allocate and execute the user job satisfying the QoS requirements and to guarantee the availability of resources at the required time for the guaranteed job execution with maximum user satisfaction and resource utilization. The M-FH strategy is developed and implemented using SLAs to match user request with the suitable resources by relaxing time resulted in maximizing the job submission rate and hence improved the job success rate. The
T-FH strategy is implemented to determine the required reliable resources in advance for job submission to avoid runtime failures and handled time-out failure by transparently redirecting the reserved job to the suitable resources without abnormally terminating the currently executing job for reducing the number of abnormal job terminations and has improved the job success rate further. The Q-FH model increases the Job success rate about 90% by matching more number of jobs and by redirecting the failed job with the suitable resources. The successful implementation of Q-FH strategies in the P-FH framework to handle the failure of resources, process, match-making and time-out proactively and reactively has provided reliable and better-guaranteed services to grid users.

The GSS-FM strategy is implemented as the fifth strategy for failure management in grouping based scheduling in which the set of jobs are combined into job groups according to the size of jobs and computation power of reliable and available resources and allocated the job groups to the selected resources at the grid scheduler before the deployment to minimize the communication overhead. The suitable resources are selected and allocated to the job groups and the resource failure is handled after it has occurred to minimize the job communication time and maximize the resource utilization. The jobs assigned according to this strategy have reduced the job completion time by 45% compared to Min-Min strategy.

IF-PRAFS is developed as an integrated computational grid framework for proactive, reactive failure handling services with the adaptive failure prevention strategy and is successfully implemented as the sixth strategy in which the resource loads are predicted before scheduling and suitable failure prevention actions are suggested for selection and execution that has considerably reduced the number of reschedulings.
The IF-PRAFS strategy selects the suitable prevention actions such as no action, renegotiating SLA, taking checkpoint and migration prior to the resource allocation itself reduces runtime failures and job migrations. The developed integrated failure handling framework IF-PRAFS has improved the failure handling services with appropriate resource selection based on reliability, availability, QoS parameters, SLAs and resource load and provides guaranteed services to grid users. The performance of the IF-PRAFS strategy has improved the resource utilization by 28% and the job success rate by 25% and reduced 45% job completion time compared with the existing Min-Min strategy which is benefiting both the grid user and the grid provider.

The integrated proactive and reactive failure handling framework developed in this research work has considerably increased the job success rate, the utilization of the resources and decreased the job completion time in the grid. The resource, process, match-making, and time-out failures are handled to provide higher reliable service and the resources are utilized to its maximum level and this benefits the service provider. The service level agreement implemented in this work guarantees the service delivery for the clients and made the grid technology applicable in the commercial arena.

8.3 FURTHER RESEARCH

The future work of this research is to develop a comprehensive failure handling grid framework to assure highly reliable execution environment for the client jobs to avoid runtime failures and job migrations. Reliability has been evaluated between the resources in the same virtual organization and it can be extended to multiple virtual organizations to improve the resource usage. The proposed framework can further be improved to support security features for resource selection and allocation to provide secured grid environment.
The integrated failure handling strategy, IF-PRAFS is presented for grid environments and it should fit into the resource allocation of cloud computing environments and resources can be allocated to different cloud providers. The different components of IF-PRAFS can be converted into appropriate services for cloud resource selection, allocation, scheduling & monitoring to solve the issues of reliability, availability and service agreements.

It is concluded that failure handling over grid computing is a promising but a very challenging issue that needs more efforts to achieve potential success in the future. The integrated failure handling strategies suggested in this research work improves the service of resource selection, allocation, scheduling and monitoring in the high failure prone grid environment to achieve the goal of a commercial application by satisfying both the user and the resource provider and makes grid technology a potential candidate for the business community.