Chapter VIII

Conclusion and Future Scope
CHAPTER 8

CONCLUSION AND FUTURE SCOPE

This chapter consolidates the design and development of e-governance architecture with the support of advance reservation and scheduling algorithms with the present possible future work.

8.1 Conclusion

Grid technologies represent a significant achievement towards the aggregation of networked resources for solving large-scale data-intensive or compute-intensive applications. Grids has evolved as the next generation computing platform for hosting distributed services for computing, content management and data sharing. Grid environment includes the resources that have varying types and capabilities that are topologically and geographically isolated which is controlled by separate administrative policies. The present work identifies and proposes the design and development of e-governance architecture with the support of advance reservation and scheduling.

A distributed framework called the grid architecture for e-governance framework was developed. The present research proposes the use of hybrid architecture which uses both the advance reservation to ensure whether the specified resources are available for applications and scheduling to sort and speedup the processes in e-governance applications. In addition, this work
recommends a perfect blend of scheduling algorithms with the advance reservation to produce the better results based on cost and time. These studies were carried out through the simulation on GridSim, a discrete-event Grid simulation tool, since different scenarios need to be evaluated and repeated.

The thesis characterises and categorizes several aspects of data grid systems. Data grids have many unique features such as presence of applications with heavy computing requirements and geographically-distributed and heterogeneous resources under different administrative domains.

The present research identifies and proposes the design and development of e-governance architecture with the support of advance reservation and scheduling. As a case study, a scenario of a citizen applying for a new water connection was described. The chain of operations that can be carried out by the citizen applying a new water connection was clearly discussed. To implement this, a distributed framework called the grid framework for e-governance applications was developed. The architecture is generic enough to accommodate different models and maps well onto the architecture of modern distributed systems. Its implementation leverages many existing technologies and provides additional services for scheduling the resources.

Further on, the present work focused on the architecture of the data grids and the requirements of resource allocation and job scheduling. The data grid environment provides the basis for the design of the grid resource broker. The requirements of the broker are to provide a framework that allowes different
types of user objectives. It also supports multiple user interfaces and handles grid characteristics such as job failure and dynamic availability. The architectural separation of interface and core layers enabled support for multiple user interfaces such as command line interfaces and web portals. The separation of core and execution layers allowed the broker to support different implementations of grid services in a standard manner.

The thesis develops the hybrid architecture to schedule task graphs by using advance reservation and scheduling in a grid environment. To implement this architecture a hybrid-core algorithm is developed. The hybrid-core algorithm uses both the advance reservation and the scheduling to improve the resource utilization. By interweaving a set of task graphs, the AR scheduler manages to reduce the overall reservation duration time and when there are many small independent jobs. The advance reservation scheduler is able to fill these jobs into the reservation blocks.

For each task and depending on the objective function, the scheduler is required to select a resource set consisting of one compute resource to execute the task and one data host for each dataset that needs to be accesses for the task. The model also takes into account the costs of using the grid resources. This model is applied to present a greedy approach for cost and time minimization based scheduling of the applications. Empirical results of evaluating this algorithm on a set of grid resources showes that the algorithm is able to reduce either the execution time or the cost of computation depending on the objective chosen.
The GridSim toolkit is used to evaluate the performance of hybrid-core algorithm through a series of simulations. The simulation uses various numbers of users and optimisation strategies. The scheduling simulations with varying deadlines for cost-optimisation strategy shows that as the deadline is increases the cost of computation decreases until it reaches the optimal level. The time-optimisation strategy shows that as the job increases, the completion time decreases. Also, when the number of users competing for the same set of resources increases, there will be proportional impact on others depending on each user’s strategies and constraints. As a result, the hybrid-core algorithm improves the utility of the system significantly compared to the other algorithms.

8.2 Future Scope

The present work reveals areas in data grids where much work is available. The contributions of this work leads to new questions that need addressing for further research. Thus, the future studies can be related to the three key functionalities of grids such as integrating various types of resources, security enhancements and check point policies.

8.2.1 Integrating Various Types of Resources

Common resources that can be reserved are compute nodes, storage elements, network bandwidth or a combination of any of those. However, this thesis is vitally focuses on reserving compute nodes. Therefore, allowing users to
reserve a combination of resource types is highly desirable, since various applications, especially in the area of data grid, can be modeled and studied.

This work leads to another interesting research problem, as it involves coordination and negotiation of multiple resources shared by different organizations. In the case of reserving network bandwidth, a network manager is needed to focus on network management issues, such as establishing a guaranteed end-to-end path, and handling traffic congestion. In addition, implementing the Multi Protocol Label Switching (MPLS) architecture [21] into GridSim may also be required. In the case of data grid applications, a Replica Manager is needed to address various data management issues, such as deletion and replication of data sets. Thus, to reserve a combination of resource types, it needs to collaborate with the Network Manager and Replica Manager.

8.2.2 Security Enhancements

The hybrid-core algorithm proposes in this thesis focuses on time and cost scenarios. However, other parameter such as security levels and fidelity is also required for many applications. For example, confidential applications must be executed on services with high security levels. In addition, the reliability of services needs to be considered during the execution planning stage. These services with high reliability may result in slightly higher cost but reduce the risk of execution failures. So this requires the development of filtering functions to improve the security level.
8.2.3 Check Point Policies

The thesis is assumed that tasks submitted always terminates, but in the case of buggy task that never terminates because it cores dump half way through, the algorithm proposed will never terminate. In order to avoid this situation, it can consider timeout technique: a task will be forced to terminate if it has not completed within the stipulated amount of time. In this way, the algorithm will always finish its execution, and the broker can execute the next task.