CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

Grid computing is a new paradigm in computing infrastructure and a grid is formed by the contributions of resources from the providers from different locations. Scheduling is an important activity in Grid computing as it determines the effectiveness of the Grid computing environment. The process of scheduling involves, discovering available resources for an application, selecting an appropriate system and submission of an application. The scheduler will aid in doing this process.

To achieve the benefits of tremendous distributed resources, effective and efficient scheduling algorithms are fundamentally important. The need for scheduling is to allocate the appropriate tasks at the right time in the desired environment. Unfortunately, scheduling algorithms in traditional parallel and distributed systems, which usually run on homogeneous and dedicated resources, e.g. computer clusters, cannot work well in the Grid environment. Thus, there is a need for scheduling algorithms that can cater to both the user requirements and optimal utilization of the Grid environment.

The main objective of the current study is to present a family of scheduling algorithms (heuristic based) that cater to various requirements of the user and application. Performances of these algorithms were analyzed based on certain parameters to prove their effectiveness.
The first algorithm proposed was “The three dimensional resource scheduling algorithm” and this algorithm enable jobs with multiple resource requirements to run effectively on Grid Computing Environment. The three dimensions considered were Computation, Data and Deadline. They were chosen to achieve faster computation through proper resource allocation. The various resource requirements of a job are CPU, hard disk space and bandwidth. Inter resource communication was also addressed by the concept of resource potential. The available resources were aggregated and then combined into three major indices namely Computation, Data and Deadline based on the three dimensions. With these indices, a 3D plot was created and this plot described the virtual topology of a job’s resource requirements. Thus, it identified that the best allocation of resources for a job and this algorithm gave minimal waiting and execution time.

The second algorithm proposed was “Economy based scheduling for utilization of space shared resources in a Bag of Tasks Grid”. For Bag of Tasks, Explicit allocation strategy combined with “Deadline and Budget Constraint with Cost Time optimization” algorithm performed effective scheduling of the jobs based on the user’s Quality of Service requirements such as deadline and budget. The cost-time optimization scheduling allocated the cheapest resources to ensure that the deadline could be met and computation cost minimized. In case if there were two resources with the same cost, scheduling was done in affordable resource so that the job gets executed as early as possible. The performance of this scheme was evaluated using cost factor (C\text{factor}) and speed up ratio (T\text{speedup}) and was found to be good.

The third algorithm proposed was “A Penalty Based Heuristic Method Using Continuous Double Auction” which was a real time economics market-like scenario which allowed the entities (user and resource provider) to make decisions autonomously. Consumer aimed at executing jobs within their corresponding deadlines with minimum cost and Provider aimed at
obtaining more profit by trying to sell their resources at higher prices and they competed with each other for accepting more jobs. An auctioneer maintained a list of the current bids and requests and matched the two offers when the highest bid was greater than or equal to the lowest request. The trade occurred at the average of matching requests and bid prices. Further the resource provider provides a penalty amount to the user when deadline is exceeded. The results of simulation illustrated that the proposed method lead to high resource utilization rate, load balancing, minimal average waiting time and queue completion time.

6.2 FUTURE WORK

Future direction for these algorithms may be as follows,

- The three dimensional algorithm can be extended to include the concept of job fragmentation. When a site is unable to satisfy all the resource requirements for a job, by itself, the job needs additional resources from remote sites to execute successfully. This is achieved by the job fragmentation phenomenon. The Three dimensional algorithm may also be extended by considering other parameters like QoS and economic considerations by extending the dimensions in the virtual map.

- The algorithm “Economy based scheduling for utilization of space shared resources in a Bag of Tasks Grid” can incorporate Load Balancing, to further improve the efficiency through effective scheduling. Secondly, “Task Replication” can be incorporated with the system to speed up the processing of jobs.

- The algorithm “Penalty based heuristic method using Continuous double auction” can further incorporate with rewards for the resource providers who complete the jobs prior to the deadline.