7

Conclusions & Future Scope

7.1 Conclusions

Grid computing has emerged as an important discipline through an increased focus on resource sharing, co-ordination, manageability and high performance. Computational Grid is a distributed and heterogeneous computing system that combines open, shared, geographically distributed and heterogeneous resources to achieve high computational performance. Since a World Wide Grid can span over million of computers connected over the internet, hence resource management becomes quite challenging and complex. To find a set of resource that match the job

requirements and then actually scheduling the jobs on to the resources is a compute-intensive problem and the complexity grows as the size of grid increases. This pushes a strong need for meta-heuristic based algorithms to be devised for grid resource management systems that can explore the large search space and can quickly converge to an optimal or near optimal solution.

In this thesis three new meta-heuristics based algorithms for grid resource scheduling have been proposed, implemented and validated. These are: SexualGA based Resource Scheduling Algorithms (SGASchedule), ACO based Resource Scheduling Algorithm (ACOSchedule) and Hybrid Genetic Simulated Annealing based Grid Scheduling Algorithm (HybridGSA). It is clear from the experimental results that SGASchedule performs best among all three algorithms and also from existing algorithms. Genetic Algorithms (GA) based scheduling algorithm is the best known heuristic till now. SGASchedule further improves over GA by using SexualGA based selection mechanism, which uses two different selection mechanisms to select different parents. This truly mimics the selection in the human genetics by correctly simulating the concept of male vigor and female choice. This helps to maintain the genetic diversity of the population and enables better exploration of search space. ACOSchedule comes second best in the list. It performs a little better over GA based algorithm in some cases and is almost approaching the existing GA based algorithms in the other. ACOSchedule uses the ACO based algorithm in the first place to find the schedule. This result is further enhanced by applying genetic operators: crossover and mutation. HybridGSA uses SGASchedule as the base algorithm where the probability of acceptance or rejection of new offsprings is based on the system temperature that decreases after every iteration. Even though it is assumed the Hybrid approach based on GA and SA can harness the parallel

exploration of search space and better convergence properties of GA and SA respectively but still the experimental results show that HybridGSA performs poor as compared to proposed algorithms SGASchedule and ACOSchedule and also to existing GA based algorithm.

A new ACO guided Mobile agents based resource discovery algorithm (ACORD) has also been proposed. This algorithm uses ACO based learning mechanism for resource discovery. Mobile Agents are used to simulate the ants in the actual resource discovery process. Hence this approach reduces the network traffic by encapsulating and moving the processing to the data than to have data transferred to a central machine and processed there. ACO guided search enables the Mobile Agents to quickly reach to the fruitful nodes. It has been established from the experiments that for finding the exact match that this approach gives approximately 85% query hits, which could increase if we go for approximate match instead of exact match. Moreover, it has been observed that with the adjustment of few parameters in this novel approach it is possible to extend the scalability of the grid size effectively.

7.2 Future Scope of work

This thesis demonstrated the usability of meta-heuristic techniques for Grid resource management. This work has laid down the foundation for using meta-heuristic techniques for grid environment but also opens up new avenues for future work in this direction.

7.2.1 Supporting Parallelization of Grid Scheduling Algorithms

While the meta-heuristics based algorithms perform well for grid scheduling, a further enhancement that can be done to the proposed algorithms is the evaluation of parallel version of these algorithms. Since all scheduling algorithms are population

based algorithms, hence the actual process of match-making and scheduling itself can be parallelized and distributed to more than one processor. Research can also be focused on designing the scheduler / broker itself as a grid application and distribute the task of schedule generation to the systems registered with the grid with in a virtual organization. ACO Guided Mobile agent resource discovery algorithm can also be further enhanced by integrating with Dynamic Hash Tables to speed up the resource discovery process.

7.2.2 Data intensive Programming and Scheduling Framework

Further research can also be focused on developing the programming and scheduling frameworks for data intensive applications from commerce, science and medicine that require need large processing power as well as data analysis / processing on large distributed data bases. Mobile Agents is another area of interest which can be further explored. Mobile agent based grid application framework can be developed for application that have little computation part but huge data that needs to be analyzed / processed. Mobile agents can travel to the data store with computation code encapsulated with them and can move the processing part rather than moving the huge data to a processing node.

7.2.3 Scheduling Framework for Real Time applications

Another area that needs attention is to develop grid based framework for real time applications where a quick response to an event is required rather than only optimally utilizing the grid resources. An example is a social community web site where one user needs to search for some specific person in the whole user database, where as the other user may query for other set of people who shared the same school with him/her. These types of requests generally require a soft real time response.

7.2.4 Grid based Generic Mathematical / Software libraries

Mathematical libraries for matrix operation are a good choice to be developed on a grid based application frameworks. Similarly, Linear Programming (LP) or Mixed Integer Linear Programming (MILP) solvers can be a good choice to be developed as grid enabled software libraries. Future research can be focused on designing grid based architectures for these libraries.

7.2.5 Grid based DWDM Network Design and Planning Tool

Designing Dense Wave length Division Multiplexing (DWDM) based network design and planning has many problems like wavelength assignment, Amplifier placement, link budget design etc. that are computationally hard and require lot of optimization. A good DWDM design tool tries to optimize the hardware placement so as to reduce the overall cost of network equipment, thus helps sales personnel to quickly place cost effective bids to the customers. Designing large and mesh DWDM networks generally take 2-3 days or more on single desktop computer with some sub-optimal designs that may result in lose of business. A grid enabled distributed architecture for DWDM design tool may solve all these problems.