ABSTRACT

The growth in network speed, storage size, processing power and Internet usage has changed the way of managing information. To address their processing needs, Distributed Computing Environment (DCE) is extensively used by business organizations and scientific institutions. Clusters, Network of Workstations (NOW), Grids and Clouds are some of the well known examples of Distributed Computing Environment. Due to their simplicity, ease of setup and maintainability, clusters are very popular among the business organizations and scientific institutions.

Dynamic load balancing is a flexible, cost effective and reliable strategy to support distributed scheduling without modifying the systems kernel or underlying hardware and without deploying expensive servers and workstations. In this research work, we have studied various scheduling and load balancing algorithms for clusters, routers and multi-core processors. It is observed that a number of algorithms are available for homogeneous distributed computing. However, very few algorithms have been developed for heterogeneous environment. Heterogeneity may exist in the system due to various factors such as CPU clock speed, cache memory, main memory size, disk capacity and its RPM, front side bus speed etc. Most of the DLB algorithms available for heterogeneous environment have considered processor speed of the participating nodes. Few algorithms have taken into account the other heterogeneity factors but due to their complexity, the algorithms are infeasible. The DLB algorithms proposed in this research incorporate most of the untouched heterogeneity factors highly sensitive to individual nodes. In this research, a general DLB model has been proposed which is applicable to clusters as well as routers. We have also incorporated Linear Programming approach in our load balancer and observed that its application has improved the performance of load balancer in an unstable system. The study has also emphasized the importance of open source software and proposed open source software at front end (load balancer) as well as at backend.
The research proposes an adaptive load balancing model for even workload distribution among the cores of symmetric multi-core servers too.

The work presented in this thesis will have wide applications in the field of distributed computing systems including clusters, grids and clouds. In addition to theoretical interest, algorithmic methodologies suggested in the proposed work have potential application in the practical areas of distributed operating system that are under infancy stage.