CHAPTER 8

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Request scheduling has a major impact on the performance of the service delivery model in a large-scale distributed computing environment. Adopting a suitable request scheduling technique helps to achieve high performance. The overall objective of this research is to design a scheduling principle that assigns the requests based on the values of the preferred attribute of the servers satisfying multiple constraints. Appropriate methods were identified and employed to achieve these objectives. The conclusions are summarized below.

8.1 Conclusions

This research work developed a request scheduling technique based on a quantification principle that decides the number of requests a server is required to serve. The conclusions of the research are summarized as follows.

i) The first objective of this research was achieved by introducing a method to identify the most preferred attribute of a server. This research used a mathematical statistical method called Conjoint Analysis to enumerate the level of influence of each attribute among the set of attributes of a server. The attribute with the highest level of influence is then chosen as the most influencing attribute. The capacity of the server is weighed based on this identified attribute.

ii) The second objective of this research was achieved by identifying a method that determines each server’s allocation share. A statistical method called Z-Score was used to do this. Z-Score takes the value set of the attribute which is identified by conjoint analysis and calculates the allocation share of each server.

iii) The third objective of this research was accomplished by designing a scheduling principle that prioritizes the requests based on the services they seek and assigns
requests to each server based on its allocation share satisfying the capacity constraints.

iv) The next objective of this research was accomplished by proposing a method to solve the request scheduling problem of any distributed system with suitable parameters. The proposed method was amply verified.

v) The last objective of this research was accomplished by developing a simulator. This simulator is used to compare the quantification scheduling principle developed in this research work with a few commonly adopted scheduling techniques. This simulator can be used to ascertain the best possible resource allocation technique by facilitating the designer to apply and test different scheduling principles. This simulator is comparable to portion of several Cloud simulators already available.

8.2 Future Research Directions

The request scheduling based on quantification principle is generic and it can be adapted to other distributed computing architectures like cluster, Peer-to-Peer, grid and Big data processing systems. Extending the problem scenario, parameters, objective policies and constraints can further this research.

The quantification principle is in some sense befits Infrastructure as a Service Cloud model. Extending the designed solution for other Cloud models is a desirable extension of this research.

Moreover, modeling the chosen problem as a linear programming problem will further enhance the effectiveness of the system.