Resource Utilization and Cost Reduction in Distributed Systems

Andrew Campbell – 45638292

Introduction

The aim of this project is to create a client for the ds-sim server that takes a job and schedules it to a server based on different factors that will allow it to reduce the overall cost and improve the resource utilization of the servers being used.

This stage of the project focuses on taking one of the pre-defined algorithms (first-fit, best-fit or worst-fit) and improving on it to achieve at least one of the following:

* Minimisation of average turnaround time
* Maximisation of average resource utilisation
* Minimisation of total server rental cost

In this instance, the best-fit algorithm is being improved upon to minimize server rental cost and maximise average resource utilisation.

Problem Definition

The best-fit algorithm makes use of the fitness value calculated using the CPU cores required vs the CPU cores available from a server. Although this method is effective in finding a server that would leave the least amount of CPU cores unused, when choosing a server, it does not take into account the ability to run multiple jobs simultaneously. To improve on the best-fit algorithm, my objective function will look at not only the fitness value of the available servers, but will also look to see if any current jobs are being completed by those servers, giving higher priority to servers that already have jobs running that are capable of running the current job simultaneously.

Algorithm Description

Implementation details

Evaluation

Conclusion

References