

# Advanced OS Term Project Proposal

Alexander Benson

February 11, 2021

In this proposal, I present my proposed project for CSE 525 Advanced Operating Systems. I intend to write a simple piece of clustering software that can provide a bare bones implementation of the MapReduce functionality described in the paper "MapReduce: Simplified Data Processing on Large Clusters"[0]. In order to do this, there are two key goals:

- Write a simple clustering specification and program
- Implement MapReduce in this simple cluster

For my simple clustering software, I will use the Go programming language from Google to define two distinct types of nodes: a worker and a scheduler. The scheduler will take in requests from the user, and then allocate these tasks to the worker nodes. The worker nodes will interpret the instructions from the scheduler, perform their tasks, then return the output to the scheduler, who then passes the results back to the original user. This will likely use a simple protocol to communicate. The user will send a packet containing their data and instructions to the scheduler, and the scheduler will then distribute packets to the worker nodes that contain the nodes instructions, and the nodes data it will be processing.

For the MapReduce implementation, I will modify the aforementioned clustering software in order to split it into two classes of worker nodes- Map and Reduce nodes. Then, I will implement an interface for MapReduce in these nodes.

I feel this project will help to deepen my understanding of distributed processing, as well as the MapReduce models.

# 1 Citations

[0] Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." *Communications of the ACM* 51.1 (2008): 107-113.