15640 Project3 Report

Building a Map-Reduce Facility with specialized DFS

Name1: Jian Wang

AndrewID:jianw3

Name2: Yifan Li

AndrewID:yifanl

Part I: Design and Implementation

1. Design on MapReduce
   1. File I/O
   2. Master-Worker Framework
   3. Scheduling

Splitting

Map Tasks dispatching

Reduce Task dispatching

* 1. Management Tool
  2. Failure and Recovery

1. Design for DFS
   1. DFS upload client
   2. DFS Name Node and Data Node framework
   3. DFS abstract file system
   4. DFS file reading and writing
   5. DFS file replication
   6. DFS data node failure and recovery

Part II: System Deployment and Configuration

Build

Deploy

Run

Part III: User Library and Simple Usage

Part IV: Features We haven’t implemented

Part V: Test Case

1. WordCount
2. Maximum