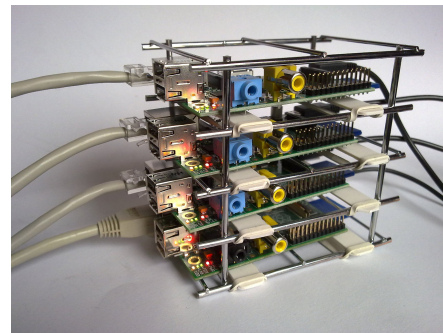# *Tiny Data with Raspberry Pi's*

Kuan-Lin Chen, Jeremy Feinstein, Brian Kutsop

Introduction

This semester we

- Built a little Raspberry Pi cluster
- Wrote an implementation of a distributed file system and MapReduce from scratch in Python
- Measured performance on the Pi's and compared

Motivation

- To see how far we can go with low-cost, commodity computing clusters
- To analyze the advantages and disadvantages of the platform compared to others
- To make physical, cluster computing more accessible for computer science students
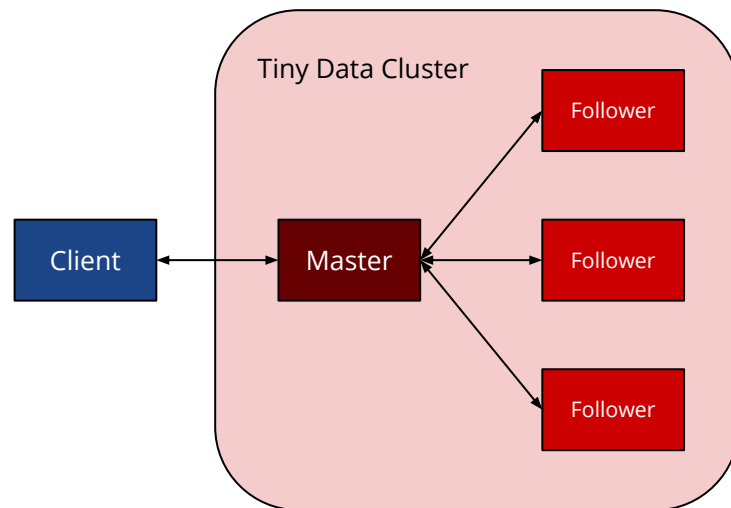
Design

Distributed File System

- Stores data sets, intermediate files, and results
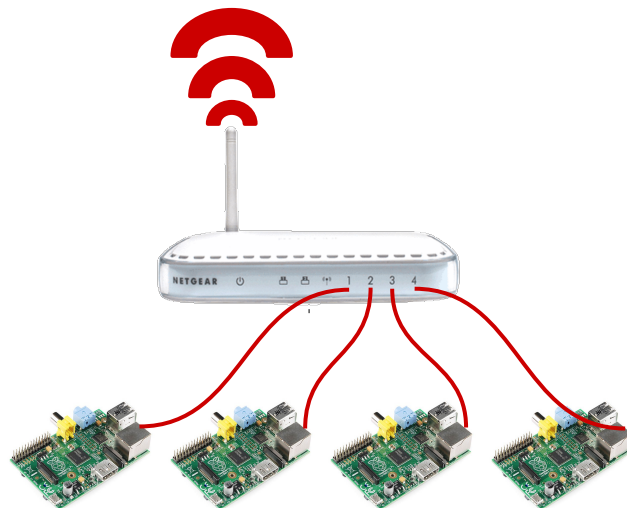- Exposes a simple set of commands : ls, rm, cat, mkdir, mv, upload

MapReduce

- Performs computations on data sets in the DFS
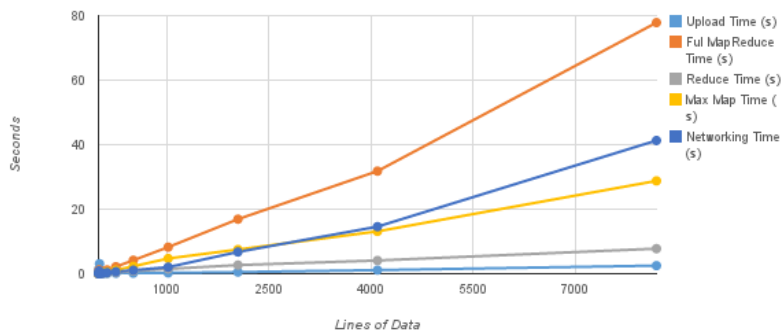- Exposes one command: map_reduce

Setup

- 4 Raspberry Pi's (with power cables, SD cards, etc.): ~$160
- 1 NetGear wireless printer switch: ~$40
- 4 ethernet cables: ~$10
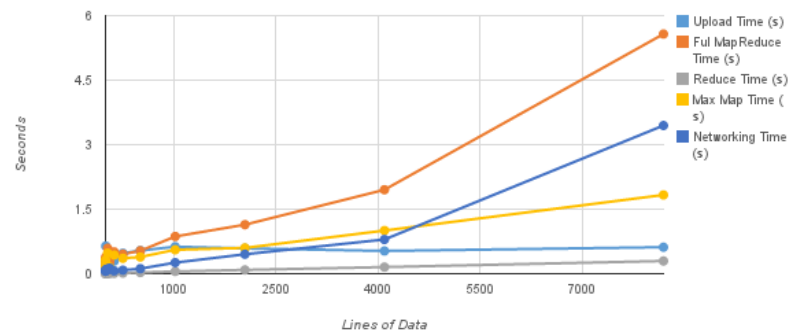- A lot of sweat and elbow grease: $0

Total experimental setup cost: $210
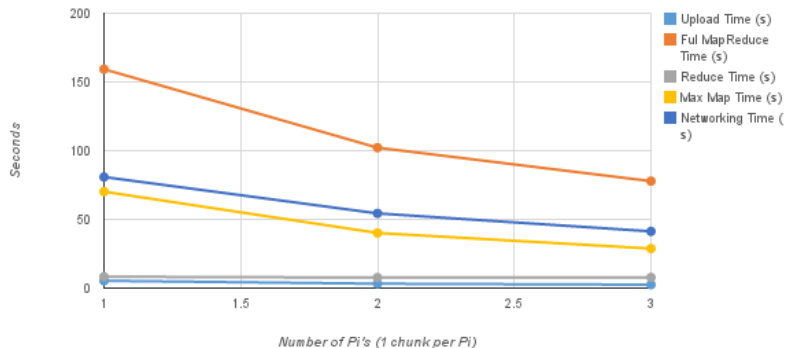
Performance with Three Workers

Raspberry Pi's

CSUG Servers

Performance with Varying Workers
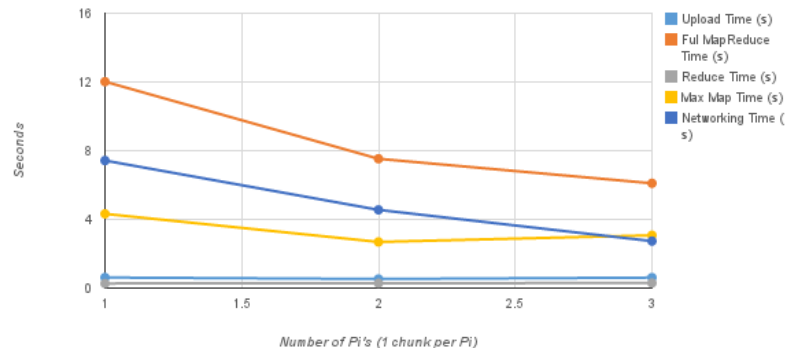
Raspberry Pi's

CSUG Servers

Performance with Varying Chunks
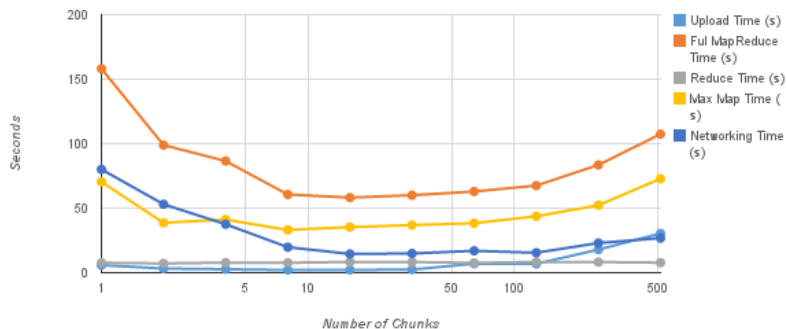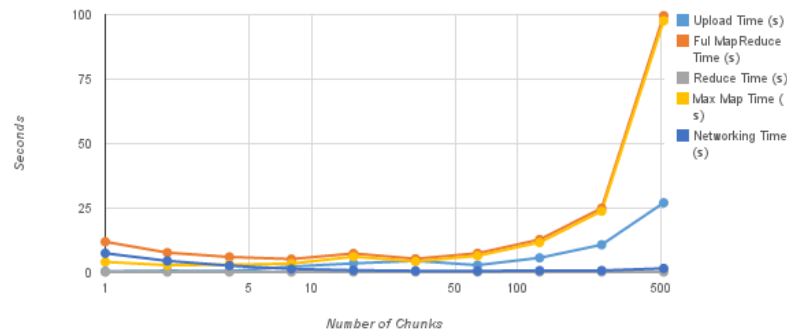
Raspberry Pi's

CSUG Servers

Energy Consumption

Power Usage Estimates During MapReduce Job
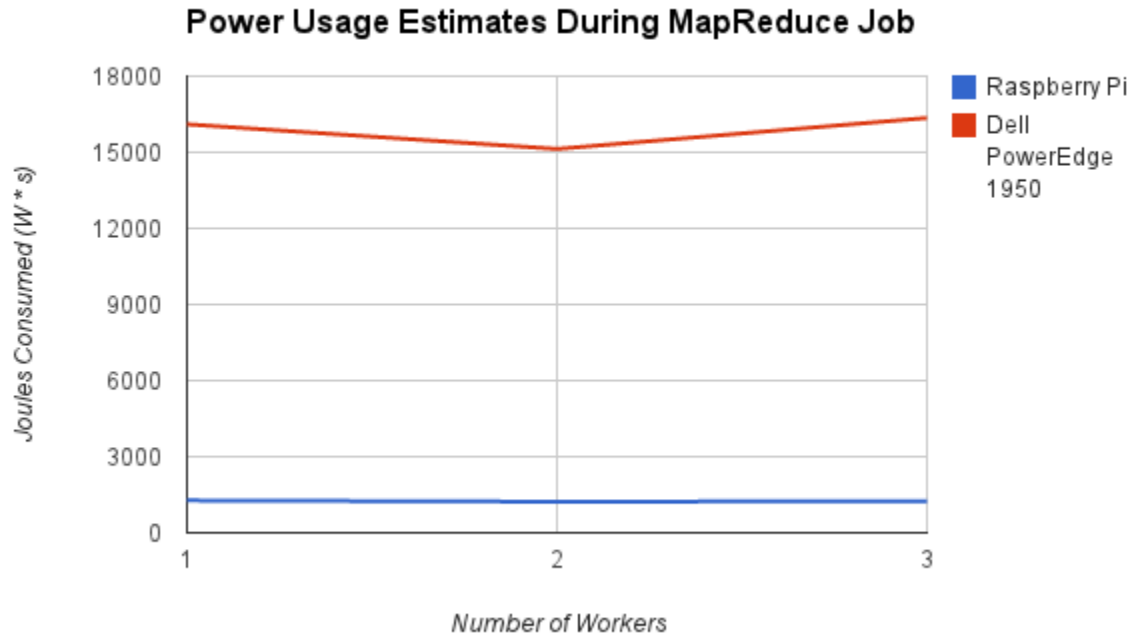
Comparison of Options

Raspberry Pi's

Total Pi Cost: $140

Power: 3W / Pi

Very Unreliable

Amazon EC2 Customer

Total Cost: $456 / year

Power: None for Customer
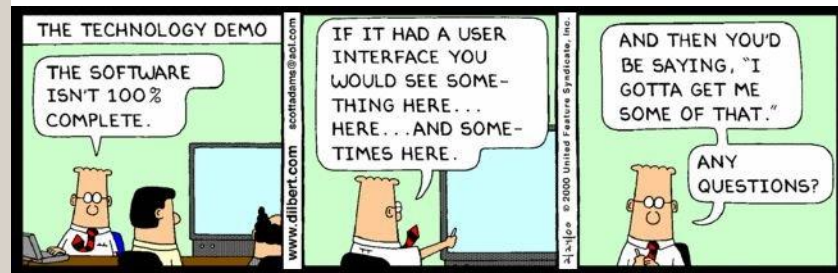
Reliable and 10X Faster

Small Enterprise Cluster

Total Cost: ~$2000

Power: 650W / server

Reliable and 10X Faster

Demo

- Show the client-side interface
- Interact with the distributed file system
- Run some MapReduce examples
  - Word count
  - N-grams

Main Takeaways

- Pi's have same proportion of process time as traditional server
- 10X less power consumption than traditional servers
- 10X less expensive than traditional servers and 3X less expensive than the cloud
- 10X worse Performance than traditional servers
- Larger job size → More bandwidth required
- Increasing number of small jobs → More processing power
- Raspberry Pi's are very "touchy" → Worth investing more

Thank You!

We appreciate the support from:

- Hakim Weatherspoon
- Ki Suh Lee
- Han Wang

Questions?