

Lab 7

Introduction to Cloud Computing -- Hadoop and MapReduce

Prof. Zichen Xu

Lab 7: MapReduce

Dear Pretty Little Pumpkins

- You are now invited to make experiments on Hadoop, with the programming model for MapReduce
- Still, I ain't ask you to build one from raw
- Pull a Hadoop-docker image will do

Goal for Lab 7

- Objectives:
 - Understanding the concept of MapReduce Model
 - Building a file system in a distributed way that store large data trunks
 - Merging all things together into a docker image
 - Writing some sample demo on running MapReduce models

To Start

- Read the Hadoop wiki page and learn it
 - <https://wiki.apache.org/hadoop>
- Install a Hadoop Docker image, you know where to find it
- Finish the report on your experience of realizing two classic demos
 - Word Count: <https://hadoop.apache.org/docs/stable/hadoop-mapreduce-client/hadoop-mapreduce-client-core/MapReduceTutorial.html#Example: WordCount v1.0>
 - Quantity Analysis: <https://docs.mongodb.com/manual/tutorial/map-reduce-examples/>

Lab 7: Submission Requirement

- Use the provided basic Markdown template for your lab report online
- Fill the report with the Markdown template, including all your test procedure
- Compile it and submit it to ncuhomework@outlook.com
- Subject shall be Course Z6110X0035 CC Lab 7 #name #ID
- The deadline for this lab is **May. 19th, 11:59PM (Hard Deadline)**