

ITEC641 Term Project 2

Hadoop and MapReduce

[What to Submit]

1. Technical Report
2. Presentation Slide (presentation length: 50 minutes)

[Topics]

1. What is Hadoop and MapReduce?

Apache Hadoop

<https://hadoop.apache.org/>

<https://hadoop.apache.org/docs/current/>

Introduction to MapReduce and Hadoop

https://people.csail.mit.edu/matei/talks/2010/amp_mapreduce.pdf

MapReduce Tutorial

https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.pdf

2. [Compare] Run Code in the Cloud

(Ex) AWS Lambda <https://aws.amazon.com/blogs/aws/run-code-cloud/>

3. Pros and Cons of building Hadoop in a Cloud Platform vs. a Private Cloud

Why Hadoop in the Cloud? & Reasons to Not Run Hadoop in the Cloud.

<https://www.oreilly.com/library/view/moving-hadoop-to/9781491959626/ch01.html>

4. Install Hadoop

Hadoop: Setting up a Single Node Cluster

<https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/SingleCluster.html>

Hadoop: Setting up a Cluster Setup

<https://hadoop.apache.org/docs/current/hadoop-project-dist/hadoop-common/ClusterSetup.html>

How to Install Hadoop on a Linux Virtual Machine on Windows 10

<https://key2consulting.com/install-hadoop-on-linux-virtual-machine-on-windows-10/>