
Lab 01: A Gentle Introduction to Hadoop

CSC14118 Introduction to Big Data 20KHMT1

your-team-name

2023-02-17

Contents

1	Lab 01: A Gentle Introduction to Hadoop	1
1.1	Setting up Single-node Hadoop Cluster	1
1.2	Introduction to MapReduce	1
1.3	Running a warm-up problem: Word Count	1
1.4	Bonus	1
1.5	References	2

1 Lab 01: A Gentle Introduction to Hadoop

1.1 Setting up Single-node Hadoop Cluster

1.2 Introduction to MapReduce

1.3 Running a warm-up problem: Word Count

1.4 Bonus

Insert table example:

Server IP Address	Ports Open
192.168.1.1	TCP: 21,22,25,80,443
192.168.1.2	TCP: 22,55,90,8080,80
192.168.1.3	TCP: 1433,3389 UDP: 1434,161

Code example:

```
print("Hello")
```

```
cat ~/.bashrc
```

Screenshot example:

Proof of change your shell prompt's name

Figure 1.1: Proof of change your shell prompt's name

ImgPlaceholder

Figure 1.2: ImgPlaceholder

Screenshot example:

Reference examples:

Some text in which I cite an author.¹

More text. Another citation.²

What is this? Yet *another* citation?³

1.5 References

- Three Cloudera version of WordCount problem:
 - https://docs.cloudera.com/documentation/other/tutorial/CDH5/topics/ht_wordcount1.html
 - https://docs.cloudera.com/documentation/other/tutorial/CDH5/topics/ht_wordcount2.html
 - https://docs.cloudera.com/documentation/other/tutorial/CDH5/topics/ht_wordcount3.html
- Book: MapReduce Design Patterns [Donald Miner, Adam Shook, 2012]
- All of StackOverflow link related.

¹So Chris Krycho, "Not Exactly a Millennium," [chriskrycho.com](http://v4.chriskrycho.com/2015/not-exactly-a-millennium.html), July 2015, <http://v4.chriskrycho.com/2015/not-exactly-a-millennium.html> (accessed July 25, 2015)

²Contra Krycho, 15, who has everything *quite* wrong.

³*ibid*