

实验报告

实验名称	实验二 Hadoop 伪分布运行模式		
实验教室	丹青 922	实验日期	2021 年 5 月 31 日
学 号	2019210173	姓 名	刘思远
专业班级	计算机科学与技术 04 班		
指导教师	卢洋		

东北林业大学
信息与计算机科学技术实验中心

一、实验目的

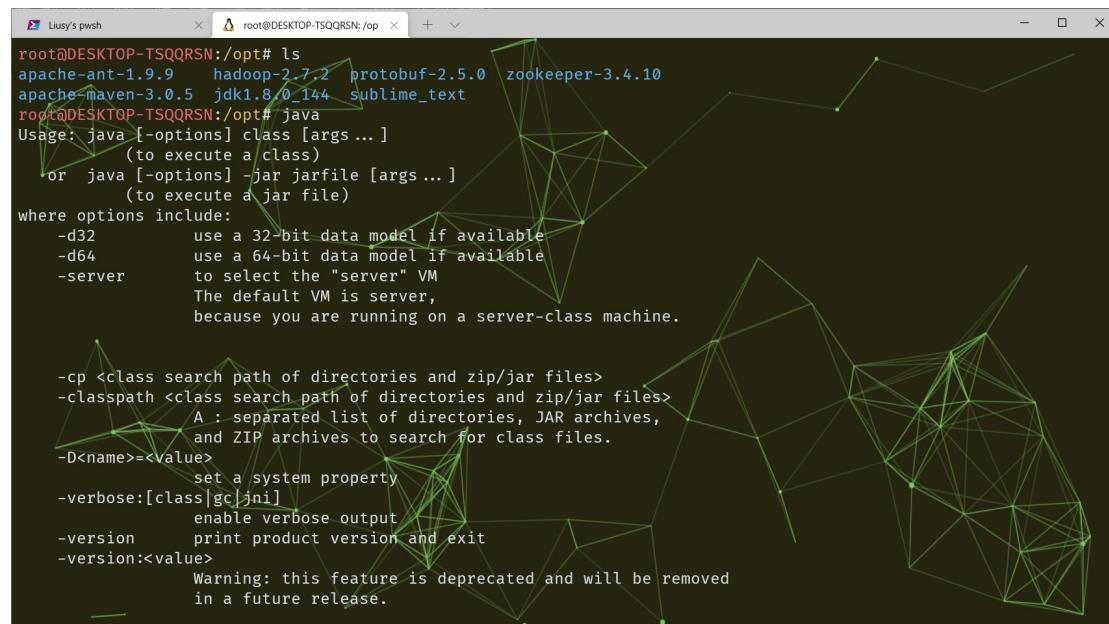
1. 掌握 JDK 的安装和配置；
2. 掌握 Hadoop 伪分布模式的安装和配置。

二、实验环境

- (1) 计算机的硬件配置 PC 系列微机。
- (2)计算机的软件配置 VMware 虚拟机软件及 Ubuntu 虚拟机。

三、实验内容及结果

1. 下载安装超星在线课提供的 JDK，版本为 1.8.0_144；



The screenshot shows a terminal window with a dark background and light-colored text. It displays the usage information for the Java command. The text includes:

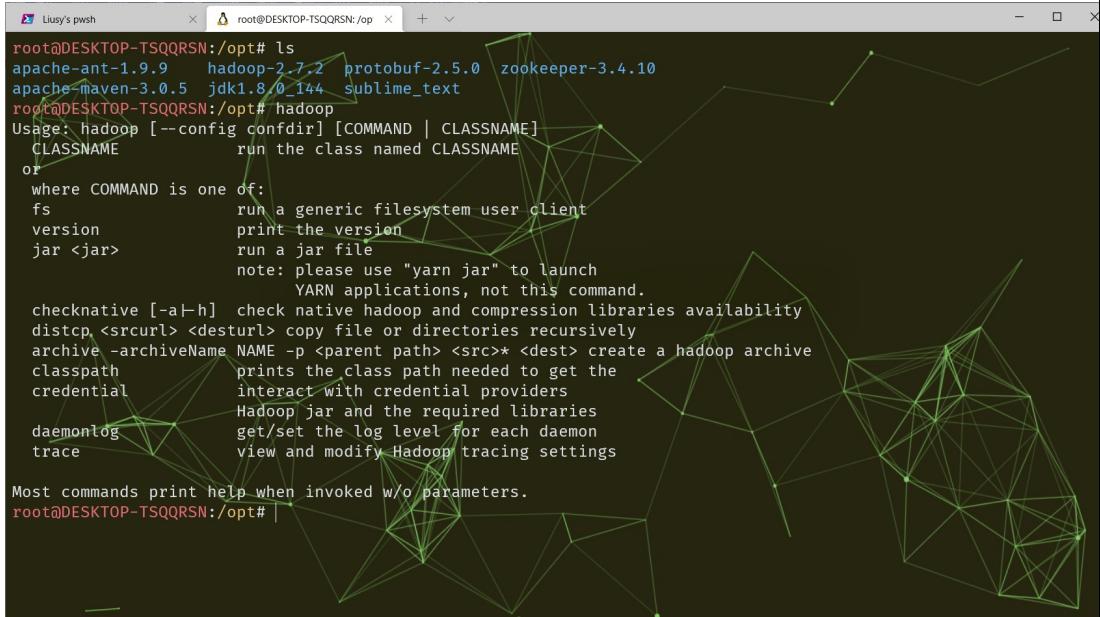
```
root@DESKTOP-TSQQRSN:/opt# ls
apache-ant-1.9.9    hadoop-2.7.2  protobuf-2.5.0  zookeeper-3.4.10
apache-maven-3.0.5   jdk1.8.0_144  sublime_text
root@DESKTOP-TSQQRSN:/opt# java
Usage: java [-options] class [args ...]
            (to execute a class)
        or  java [-options] -jar jarfile [args ...]
            (to execute a jar file)
where options include:
  -d32      use a 32-bit data model if available
  -d64      use a 64-bit data model if available
  -server   to select the "server" VM
            The default VM is server,
            because you are running on a server-class machine.

  -cp <class search path of directories and zip/jar files>
  -classpath <class search path of directories and zip/jar files>
            A : separated list of directories, JAR archives,
            and ZIP archives to search for class files.
  -D<name>=<value>
            set a system property
  -verbose:[class|gc|jni]
            enable verbose output
  -version   print product version and exit
  -version:<value>
            Warning: this feature is deprecated and will be removed
            in a future release.
```

2. 验证所安装 Java 的版本；

```
root@DESKTOP-TSQQRSPN:/opt# java -version
java version "1.8.0_144"
Java(TM) SE Runtime Environment (build 1.8.0_144-b01)
Java HotSpot(TM) 64-Bit Server VM (build 25.144-b01, mixed mode)
root@DESKTOP-TSQQRSPN:/opt# |
```

3. 下载安装超星在线课提供的 Hadoop , 版本为 2.7.2 ;



```
Liusy's pwsh      root@DESKTOP-TSQQRSPN:/op |
root@DESKTOP-TSQQRSPN:/opt# ls
apache-ant-1.9.9  hadoop-2.7.2  protobuf-2.5.0  zookeeper-3.4.10
apache-maven-3.0.5  jdk1.8.0_144  sublime_text
root@DESKTOP-TSQQRSPN:/opt# hadoop
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]
  CLASSNAME          run the class named CLASSNAME
or
  where COMMAND is one of:
    fs                run a generic filesystem user client
    version           print the version
    jar <jar>          run a jar file
                      note: please use "yarn jar" to launch
                           YARN applications, not this command.
    checknative [-a|-h]  check native hadoop and compression libraries availability
    distcp <srcurl> <desturl> copy file or directories recursively
    archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archive
    classpath          prints the class path needed to get the
    credential         interact with credential providers
    Hadoop jar and the required libraries
    daemonlog          get/set the log level for each daemon
    trace              view and modify Hadoop tracing settings

Most commands print help when invoked w/o parameters.
root@DESKTOP-TSQQRSPN:/opt# |
```

4. 验证所安装 Hadoop 的版本;

```
root@DESKTOP-TSQQRSPN:/opt# hadoop version
Hadoop 2.7.2
Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r b165c4fe8a74265c792ce23f546c64604acf0e41
Compiled by jenkins on 2016-01-26T00:08Z
Compiled with protoc 2.5.0
From source with checksum d0fd2a6633fa762bff87ec759ebe689c
This command was run using /opt/hadoop-2.7.2/share/hadoop/common/hadoop-common-2.7.2.jar
```

5. 使用 jps 工具 , 验证 NameNode 、 DataNode 、 ResourceManager 、 NodeManager 和 JobHistoryServer 的运行状态;

```
Liusy's pwsh      root@DESKTOP-TSQQRSPN:/op |
root@DESKTOP-TSQQRSPN:/opt/hadoop-2.7.2# jps
929 ResourceManager
1142 NodeManager
566 DataNode
1319 JobHistoryServer
393 NameNode
748 SecondaryNameNode
4270 Jps
```

6. 访问 HDFS 的 Web 页面;

Overview 'DESKTOP-TSQQRSN:9000' (active)

Started:	Mon May 31 19:00:15 CST 2021
Version:	2.7.2, rb165c4fe8a74265c792ce23f546c64604acf0e41
Compiled:	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
Cluster ID:	CID-f52f4b91-6613-4b07-a646-fa592e48f0b7
Block Pool ID:	BP-1542253369-10.191.53.85-1622264819009

Summary

Security is off.

Safemode is off.

97 files and directories, 60 blocks = 157 total filesystem object(s).

Heap Memory used 135.65 MB of 207.5 MB Heap Memory. Max Heap Memory is 889 MB.

Non Heap Memory used 55.56 MB of 56.72 MB Committed Non Heap Memory. Max Non Heap Memory is -1 B.

Configured Capacity:	200 GB
DFS Used:	2.33 MB (0%)

7. 访问 Yarn 的 Web 页面;

All Applications

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
application_1622458832718_0005	root	word count	MAPREDUCE	default	Mon May 31 19:15:37 +0800 2021	Mon May 31 19:15:37 +0800 2021	FINISHED	SUCCEEDED		History	N/A
application_1622458832718_0004	root	grep-sort	MAPREDUCE	default	Mon May 31 19:07:58 +0800 2021	Mon May 31 19:08:21 +0800 2021	FINISHED	SUCCEEDED		History	N/A
application_1622458832718_0003	root	grep-search	MAPREDUCE	default	Mon May 31 19:07:24 +0800 2021	Mon May 31 19:07:57 +0800 2021	FINISHED	SUCCEEDED		History	N/A
application_1622458832718_0001	root	grep-search	MAPREDUCE	default	Mon May 31 19:02:30 +0800 2021	Mon May 31 19:02:47 +0800 2021	FINISHED	SUCCEEDED		History	N/A

8. 访问 HistoryServer 的 Web 页面;

JobHistory

Submit Time	Start Time	Finish Time	Job ID	Name	User	Queue	State	Maps Total	Maps Completed	Reduces Total	Reduces Completed
2021-05-31 19:15:21 CST	2021-05-31 19:15:26 CST	2021-05-31 19:15:37 CST	job_1622458832718_0005	word count	root	default	SUCCEEDED	1	1	1	1
2021-05-31 19:07:58 CST	2021-05-31 19:08:10 CST	2021-05-31 19:08:20 CST	job_1622458832718_0004	grep-sort	root	default	SUCCEEDED	1	1	1	1
2021-05-31 19:07:24 CST	2021-05-31 19:07:31 CST	2021-05-31 19:07:57 CST	job_1622458832718_0003	grep-search	root	default	SUCCEEDED	9	9	1	1
2021-05-31 19:02:30 CST	2021-05-31 19:02:36 CST	2021-05-31 19:02:46 CST	job_1622458832718_0001	grep-search	root	default	SUCCEEDED	1	1	1	1
2021-05-29 14:43:04 CST	2021-05-29 14:43:11 CST	2021-05-29 14:44:07 CST	job_1622270333365_0001	word count	root	default	SUCCEEDED	29	29	1	1
2021-05-29 13:09:16 CST	2021-05-29 13:09:25 CST	2021-05-29 13:09:37 CST	job_1622264873807_0001	word count	root	default	SUCCEEDED	1	1	1	1

9. 运行 Apache Hadoop 官方提供的 grep 案例;

```

root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# ls
LICENSE.txt README.txt data include lib logs sbin wcinput
NOTICE.txt bin etc input libexec output share wcoutput
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -put input /user/root
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root
Found 5 items
drwxr-xr-x - root supergroup 0 2021-05-31 19:28 /user/root/input
drwxr-xr-x - root supergroup 0 2021-05-29 14:41 /user/root/input0x00
drwxr-xr-x - root supergroup 0 2021-05-29 14:44 /user/root/output0x00
drwxr-xr-x - root supergroup 0 2021-05-31 19:13 /user/root/wcinput
drwxr-xr-x - root supergroup 0 2021-05-31 19:15 /user/root/wcoutput
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar
grep input output 'dfs[a-z]+'
21/05/31 19:29:51 INFO client.RMProxy: Connecting to ResourceManager at DESKTOP-TSQQRSN/10.191.53.85:8032
21/05/31 19:29:51 INFO input.FileInputFormat: Total input paths to process : 9
21/05/31 19:29:51 INFO mapreduce.JobSubmitter: number of splits:9
21/05/31 19:29:51 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1622458832718_0006
21/05/31 19:29:52 INFO impl.YarnClientImpl: Submitted application application_1622458832718_0006
21/05/31 19:29:52 INFO mapreduce.Job: The url to track the job: http://DESKTOP-TSQQRSN:8088/proxy/application_1622458832718_0006/
21/05/31 19:29:52 INFO mapreduce.Job: Running job: job_1622458832718_0006
21/05/31 19:29:59 INFO mapreduce.Job: Job job_1622458832718_0006 running in uber mode : false
21/05/31 19:29:59 INFO mapreduce.Job: map 0% reduce 0%
21/05/31 19:30:09 INFO mapreduce.Job: map 67% reduce 0%
21/05/31 19:30:17 INFO mapreduce.Job: map 100% reduce 0%
21/05/31 19:30:18 INFO mapreduce.Job: map 100% reduce 100%
21/05/31 19:30:19 INFO mapreduce.Job: Job job_1622458832718_0006 completed successfully
21/05/31 19:30:19 INFO mapreduce.Job: Counters: 49

```

```

Liusy's pwsh      root@DESKTOP-TSQQRSN:/op
Combine input records=0
Combine output records=0
Reduce input groups=1
Reduce shuffle bytes=78
Reduce input records=3
Reduce output records=3
Spilled Records=6
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=78
CPU time spent (ms)=1060
Physical memory (bytes) snapshot=433438720
Virtual memory (bytes) snapshot=2542836764672
Total committed heap usage (bytes)=320339968

Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
Bytes Read=176
File Output Format Counters
Bytes Written=48
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# |
```

```

root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root
Found 6 items
drwxr-xr-x  - root supergroup          0 2021-05-31 19:28 /user/root/input
drwxr-xr-x  - root supergroup          0 2021-05-29 14:41 /user/root/input0x00
drwxr-xr-x  - root supergroup          0 2021-05-31 19:30 /user/root/output
drwxr-xr-x  - root supergroup          0 2021-05-29 14:44 /user/root/output0x00
drwxr-xr-x  - root supergroup          0 2021-05-31 19:13 /user/root/wcinput
drwxr-xr-x  - root supergroup          0 2021-05-31 19:15 /user/root/wcoutput
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root/output
Found 2 items
-rw-r--r--  1 root supergroup          0 2021-05-31 19:30 /user/root/output/_SUCCESS
-rw-r--r--  1 root supergroup          48 2021-05-31 19:30 /user/root/output/part-r-00000
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -cat /user/root/output/p*
1  dfsadmin
1  dfs.replication
1  dfs.http.address
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# |
```

10. 运行 Apache Hadoop 官方提供的 WordCount 案例;

```

Liusy's pwsh      root@DESKTOP-TSQQRSN:/op
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -put wcinput /user/root/
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root
Found 5 items
drwxr-xr-x  - root supergroup          0 2021-05-31 19:28 /user/root/input
drwxr-xr-x  - root supergroup          0 2021-05-29 14:41 /user/root/input0x00
drwxr-xr-x  - root supergroup          0 2021-05-31 19:30 /user/root/output
drwxr-xr-x  - root supergroup          0 2021-05-29 14:44 /user/root/output0x00
drwxr-xr-x  - root supergroup          0 2021-05-31 19:33 /user/root/wcinput
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar
wordcount wcinput
21/05/31 19:34:02 INFO client.RMProxy: Connecting to ResourceManager at DESKTOP-TSQQRSN/10.191.53.85:8032
21/05/31 19:34:02 INFO input.FileInputFormat: Total input paths to process : 1
21/05/31 19:34:02 INFO mapreduce.JobSubmitter: number of splits:1
21/05/31 19:34:03 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1622458832718_0008
21/05/31 19:34:03 INFO impl.YarnClientImpl: Submitted application application_1622458832718_0008
21/05/31 19:34:03 INFO mapreduce.Job: The url to track the job: http://DESKTOP-TSQQRSN:8088/proxy/application_1622458832718_0008/
21/05/31 19:34:03 INFO mapreduce.Job: Running job: job_1622458832718_0008
21/05/31 19:34:09 INFO mapreduce.Job: Job job_1622458832718_0008 running in uber mode : false
21/05/31 19:34:09 INFO mapreduce.Job: map 0% reduce 0%
21/05/31 19:34:14 INFO mapreduce.Job: map 100% reduce 0%
21/05/31 19:34:19 INFO mapreduce.Job: map 100% reduce 100%
21/05/31 19:34:19 INFO mapreduce.Job: Job job_1622458832718_0008 completed successfully
21/05/31 19:34:19 INFO mapreduce.Job: Counters: 49
      File System Counters
          FILE: Number of bytes read=84
          FILE: Number of bytes written=235043
          FILE: Number of read operations=0

```

```

Liusy's pwsh      root@DESKTOP-TSQQRSN:/op
Input split bytes=119
Combine input records=10
Combine output records=7
Reduce input groups=7
Reduce shuffle bytes=84
Reduce input records=7
Reduce output records=7
Spilled Records=14
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=82
CPU time spent (ms)=1280
Physical memory (bytes) snapshot=436436992
Virtual memory (bytes) snapshot=1770440224768
Total committed heap usage (bytes)=320864256
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
Bytes Read=53
File Output Format Counters
Bytes Written=50
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# 

root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root
Found 6 items
drwxr-xr-x  - root supergroup 0 2021-05-31 19:28 /user/root/input
drwxr-xr-x  - root supergroup 0 2021-05-29 14:41 /user/root/input0x00
drwxr-xr-x  - root supergroup 0 2021-05-31 19:30 /user/root/output
drwxr-xr-x  - root supergroup 0 2021-05-29 14:44 /user/root/output0x00
drwxr-xr-x  - root supergroup 0 2021-05-31 19:33 /user/root/wcinput
drwxr-xr-x  - root supergroup 0 2021-05-31 19:34 /user/root/wcoutput
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -ls /user/root/wcoutput
Found 2 items
-rw-r--r-- 1 root supergroup 0 2021-05-31 19:34 /user/root/wcoutput/_SUCCESS
-rw-r--r-- 1 root supergroup 50 2021-05-31 19:34 /user/root/wcoutput/part-r-00000
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2# hdfs dfs -cat /user/root/wcoutput/p*
4ever 1
hadoop 2
lby 1
love 1
lsy 2
root 2
yarn 1
root@DESKTOP-TSQQRSN:/opt/hadoop-2.7.2#

```

11. 在 Hadoop 的 Web 界面上查看案例的运行情况及结果

JobHistory

Submit Time	Start Time	Finish Time	Job ID	Name	User	Queue	State	Maps Total	Maps Completed	Reduces Total	Reduces Completed
2021.05.31 19:34:03 CST	2021.05.31 19:34:08 CST	2021.05.31 19:34:18 CST	job_1622458832718_0008	word count	root	default	SUCCEEDED	1	1	1	1
2021.05.31 19:30:19 CST	2021.05.31 19:30:28 CST	2021.05.31 19:30:39 CST	job_1622458832718_0007	grep-sort	root	default	SUCCEEDED	1	1	1	1

Grep

The screenshot shows the Hadoop MapReduce Job interface. The top navigation bar includes links for '学生学习页面' (Student Learning Page), 'MapReduce Job job_1622458832718_0007', and '不安全 | desktop-tsqrn:19888/jobhistory/job/job_1622458832718_0007'. The left sidebar has sections for Application, Job (Overview, Counters, Configuration, Map tasks, Reduce tasks), and Tools. The main content area displays job details: Job Name: grep-sort, User Name: root, Queue: default, State: SUCCEEDED, Uberized: false, Submitted: Mon May 31 19:30:19 CST 2021, Started: Mon May 31 19:30:28 CST 2021, Finished: Mon May 31 19:30:39 CST 2021, Elapsed: 10sec. Diagnostics show Average Map Time: 2sec, Average Shuffle Time: 2sec, Average Merge Time: 0sec, and Average Reduce Time: 0sec. Below this is the ApplicationMaster table:

Attempt Number	Start Time	Node	Logs
1	Mon May 31 19:30:25 CST 2021	DESKTOP-TSQRN:8042	logs
Task Type	Total	Complete	
Map	1	1	
Reduce	1	1	
Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

The bottom section shows the log upload time as Mon May 31 19:30:46 +0800 2021.

The screenshot shows the Hadoop MapReduce Job interface with the logs tab selected. The logs are displayed in a terminal-like window. The logs start with:

```
Log Type: stderr
Log Upload Time: Mon May 31 19:30:46 +0800 2021
Log Length: 1703
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.GuiceComponentProviderFactory register
INFO: Registering org.apache.hadoop.mapreduce.v2.app.job.JobHistoryEventHandler as a provider class
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.GuiceComponentProviderFactory register
INFO: Registering org.apache.hadoop.yarn.webapp.GenericExceptionHandler as a provider class
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.GuiceComponentProviderFactory register
INFO: Registering org.apache.hadoop.mapreduce.v2.app.webapp.MasterService as a root resource class
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.GuiceComponentProviderFactory register
INFO: Initiating Jersey application, version 'Jersey: 1.9.0@2002/2011 11:17 AM'
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.ContainerFactory getGuiceComponentProvider
INFO: GuiceComponentProviderFactory registered with the scope "Singleton"
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.GuiceComponentProviderFactory getGuiceComponentProvider
INFO: Binding org.apache.hadoop.yarn.webapp.GenericExceptionHandler to GuiceManagedComponentProvider with the scope "Singleton"
May 31, 2021 7:30:28 PM com.sun.jersey.spi.container.ContainerFactory getGuiceComponentProvider
INFO: Binding org.apache.hadoop.yarn.webapp.MasterService to GuiceManagedComponentProvider with the scope "PerRequest"
1log4j:WARN No appenders could be found for logger (org.apache.hadoop.ipc.Server).
1log4j:WARN Please initialize the log4j system properly.
1log4j:WARN See http://logging.apache.org/log4j/1.2/manual/appenders.html#config for more info.
```

Below this, the log type is set to 'stdout' and the log length is 35156 bytes. The logs show the execution of a WordCount application:

```
Log Type: stdout
Log Upload Time: Mon May 31 19:30:46 +0800 2021
Log Length: 0
Log Type: syslog
Log Upload Time: Mon May 31 19:30:46 +0800 2021
Log Length: 35156
Showing 0-26 bytes of 35156 total. Click here for the full log.
1log4j:WARN No appenders could be found for logger [job_1622458832718_0007].1.Hist to /DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,116 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Copied to done location: hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,118 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Copying hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/root/.staging/job_1622458832718_0007/j
2021-05-31 19:30:39,137 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Copied to done location: hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,140 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Moved tap to done hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,143 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Moved tap to done hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,146 INFO [eventHandlingThread] org.apache.hadoop.mapreduce.JobHistoryEventHandler: Moved tap to done hdfs://DESKTOP-TSQRN:9000/tap/hadoop-yarn/staging/history/done_intermediate/root/job_1622458832718_0007
2021-05-31 19:30:39,149 INFO [Thread-67] org.apache.hadoop.mapreduce.v2.app.rm.RMComunicator: Setting job diagnostics to
2021-05-31 19:30:39,149 INFO [Thread-67] org.apache.hadoop.mapreduce.v2.app.rm.RMComunicator: History url is http://DESKTOP-TSQRN:9000/jobhistory/job/job_1622458832718_0007
2021-05-31 19:30:39,155 INFO [Thread-67] org.apache.hadoop.mapreduce.v2.app.rm.RMComunicator: Waiting for application to be successfully unregistered.
```

WordCount

The screenshot shows the Hadoop MapReduce Job interface. The main title is "MapReduce Job job_1622458832718_0008". On the left, there's a sidebar with "Application" and "Job" sections, and a "Tools" button. The "Job" section is expanded, showing details like Job Name: word count, User Name: root, Queue: default, State: SUCCEEDED, and Uberized: false. It also lists submission and start times, and completion metrics: Average Map Time 2sec, Average Shuffle Time 2sec, Average Merge Time 0sec, and Average Reduce Time 0sec. Below this is the "ApplicationMaster" table with one row for Attempt Number 1. The last column has a "Logs" link. At the bottom, there are two tables: "Task Type" and "Attempt Type". The "Task Type" table shows 1 Map task and 1 Reduce task, both completed. The "Attempt Type" table shows 0 Failed Maps and Reduces, with 1 Killed Map and 1 Successful Reduce.

Task Type	Total	Complete
Map	1	1
Reduce	1	1

Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	1

四、实验过程分析与讨论

通过配置七个文件，core-site.xml、hadoop-env.sh、hdfs-site.xml、yarn-env.sh、yarn-site.xml、mapred-env.sh、mapred-site.xml，配置 ssh 免密登录，慢慢的调试，才将伪分布式搭建好。期间通过阅读官方文档获取资料并解决问题，还帮助别的同学搭建，过程艰辛，但收获满满。

五、指导教师意见

指导教师签字：卢洋