




micmiu — 软件开发+生活点滴

hadoop 2.2.0 集群模式安装配置和测试

 作者: Michael 日期: 2014 年 1 月 22 日

 查看评论  发表评论 (0)

本文详细记录Hadoop 2.2.0 集群安装配置的步骤,并运行演示一个简单的job。基本目录结构如下:

- 环境准备
- Hadoop安装配置
- 启动及演示

[一]、环境准备

本文所有集群节点的操作系统均为:CentOS 6.0 64位,不管是实体机还是虚拟机都可以,在这里统一叫做 “实例”吧,以4 台主机实例作为集群配置的演示,具体的划分如下:

| hostname | IP | 用途 |
|---------------|---------------|-------------------------|
| Master.Hadoop | 192.168.6.77 | NameNode/ResouceManager |
| Slave5.Hadoop | 192.168.8.205 | DataNode/NodeManager |
| Slave6.Hadoop | 192.168.8.206 | DataNode/NodeManager |
| Slave7.Hadoop | 192.168.8.207 | DataNode/NodeManager |

ps:如果是虚拟机可以把环境配置好后,copy多个实例即可,需要注意修改hostname

1、`vi /etc/hosts` 添加如下内容:

| | | |
|---|---------------|---------------|
| 1 | 192.168.6.77 | Master.Hadoop |
| 2 | 192.168.8.205 | Slave5.Hadoop |
| 3 | 192.168.8.206 | Slave6.Hadoop |
| 4 | 192.168.8.207 | Slave7.Hadoop |

2、JDK

到Java 的官网下载jdk6 64位的版本,安装最基础的安装即可,当然由于CentOS6 自带了OpenJDK,本文直接用OpenJDK来演示(ps: OpenJDK的目录一般在`/usr/lib/jvm/` 路径下),该系统的JAVA_HOME 配置如下:`export JAVA_HOME = /usr/lib/jvm/java-1.6.0-openjdk.x86_64`

3、SSHD服务

确保系统已经安装了SSHD相关服务,并启动(CentOS默认已经安装好)。

4、创建用户

创建一个专用的账户:hadoop

| | |
|---|--------------------------------|
| 1 | <code>\$ useradd hadoop</code> |
|---|--------------------------------|

5、配置SSH无密码登录

需要实现 Master到所有的Slave以及所有Slave 到Master的SSH无密码登录



 微博

 **孙吴空的爸爸** 上海 浦东新区

+ 加关注

@吴柚柚 // @电商行业: 【手机QQ也能发红包了！来一发？】7月15日-20日，拍拍网联合手机QQ开展抢红包活动，上亿全站通用无门槛代金券火热发放。京东接手运营半年时间，拍拍网开始发力吸引客户了，期待好体验！

拍拍网  ：7月15日-20日，拍拍网将联合手机QQ开展抢红包活动。上亿全站通用无门槛代金券火热发放，小伙伴们速速来抢吧。今天，小编就来告诉大家具体的攻略，明天上午准时登陆手机QQ哦，不

近期文章

-  Spring AOP拦截Struts1记录操作日志
-  Baby
-  Hadoop2.x在Ubuntu系统中编译源码
-  Ubuntu设置默认jdk
-  Ubuntu12.04中Eclipse安装svn插件问题
-  eclipse中开发Hadoop2.x的Map/Reduce项目
-  Hadoop2.x eclipse plugin插件编译安装配置
-  maven编译Spark源码
-  java.net.ConnectException: to 0.0.0.0:10020 failed on connection exception
-  ERROR tool.ImportTool: Imported Failed: Attempted to generate class with no columns!

浏览

-  图文介绍openLDAP在windows上的安装配置 - 12,237 views
-  Thrift入门及Java实例演示 - 11,664 views
-  SSO之CAS单点登录实例演示 - 11,648 views
-  java keytool证书工具使用小结 - 9,516 views
-  java自带的MD5、SHA1算法演示 - 9,217 views
-  VirtualBox安装Ubuntu图文介绍 - 8,792 views
-  Openfire服务端源代码开发配置指南 - 8,719 views
-  Eclipse+Maven构建web项目及部署时Maven lib依赖问题的解决 - 8,147 views
-  Mac OSX 10.8安装SVN的方法 - 7,050 views

有关SSH无密码登录的详细介绍可以参见:[Linux\(Centos\)配置OpenSSH无密码登陆](#)

6、配置时钟同步

```
1 $ cp /usr/share/zoneinfo/Asia/Shanghai /etc/localtime
2 $ ntpdate us.pool.ntp.org
3 $ crontab -e
4 0-59/10 * * * * /usr/sbin/ntpdate us.pool.ntp.org | logger -t NTP
```

ps: 如果是实体机以上所有步骤需要在每个实例里都操作一遍;如果是虚拟机只需要一个实例中完成,其他实例复制即可。

[二]、Hadoop安装配置

1、下载源码编译本地库

由于官方的发布包中的本地库是32位的,不符合我们的要求,需要自己编译本地库,编译本地库的过程可以参考:[Hadoop 2.x build native library on Mac os x](#) ,大同小异,编译完成后,替换<HADOOP_HOME>/lib/native/ 下的文件即可,注意lib文件名。

ps:这步只需要做一次即可,因为集群中的4个实例的环境一样。

2、下载发布包

打开官方下载链接 <http://hadoop.apache.org/releases.html#Download> ,选择2.2.0版本的发布包下载后解压到指定路径下:

```
1 $ tar -zxf hadoop-2.2.0.tar.gz -C /usr/local/
2 $ cd /usr/local
3 $ ln -s hadoop-2.2.0 hadoop
```

那么本文中 `HADOOP_HOME = /usr/local/hadoop/` .

3、配置hadoop用户的环境变量 `vi ~/.bash_profile` ,添加如下内容:

```
1 # set java environment
2 export JAVA_HOME=/usr/lib/jvm/java-1.6.0-openjdk.x86_64
3 export CLASSPATH=.:$CLASSPATH:$JAVA_HOME/lib:$JAVA_HOME/jre/lib
4 export PATH=$PATH:$JAVA_HOME/bin:$JAVA_HOME/jre/bin
5
6 # Michael@micmiu.com
7 # Hadoop
8 export HADOOP_PREFIX="/usr/local/hadoop"
9 export PATH=$PATH:$HADOOP_PREFIX/bin:$HADOOP_PREFIX/sbin
10 export HADOOP_COMMON_HOME=${HADOOP_PREFIX}
11 export HADOOP_HDFS_HOME=${HADOOP_PREFIX}
12 export HADOOP_MAPRED_HOME=${HADOOP_PREFIX}
13 export HADOOP_YARN_HOME=${HADOOP_PREFIX}
```

4、编辑 `<HADOOP_HOME>/etc/hadoop/hadoop-env.sh`

修改JAVA_HOME的配置:

```
1 export JAVA_HOME=/usr/lib/jvm/java-1.6.0-openjdk.x86_64
```

5、编辑 `<HADOOP_HOME>/etc/hadoop/yarn-env.sh`

修改JAVA_HOME的配置:

```
1 export JAVA_HOME=/usr/lib/jvm/java-1.6.0-openjdk.x86_64
```

6、编辑 `<HADOOP_HOME>/etc/hadoop/core-site.xml`

java.lang.OutOfMemoryError: GC overhead limit exceeded - 6,505 views

分类目录

- J2EE (41)
 - Hibernate (12)
 - Struts (5)
 - Spring (11)
 - JDBC (6)
 - JTA (3)
 - EJB (4)
- 海量数据 (28)
 - Hadoop (12)
 - Hive (8)
 - HBase (4)
 - Spark (2)
 - Sqoop (2)
- SOA (14)
 - webservice (12)
 - RPC (2)
- OpenSource (27)
 - openflashchart (2)
 - 生成文档 (5)
 - Openfire (4)
 - Security (1)
 - CORBA (4)
- 架构设计 (7)
 - 集群 (6)
 - 缓存 (1)
- 企业应用 (19)
 - SNMP (8)
 - SSO (4)
 - 服务器软件 (4)
- 编程语言 (34)
 - Java (24)
 - javascript (5)
 - Groovy (4)
 - nodejs (1)
- NoSQL (12)
 - Berkeley (6)
- 数据库 (10)
 - Oracle (7)
 - mysql (2)
- Android (2)
- web UI (1)
- 软件工具 (32)
 - 常用软件 (1)
 - 项目构建 (15)
 - 测试工具 (2)
- 操作系统 (21)
 - linux (15)
 - windows (1)
 - Mac (5)
- 小小 (12)
 - 心底软绵绵 (4)
 - 时间静悄悄 (5)
 - 肚皮圆滚滚 (3)

在<configuration>节点下添加或者更新下面的配置信息:

```
1 <!-- 新变量f:s.defaultFS 代替旧的:fs.default.name |micmiu.com-->
2 <property>
3   <name>fs.defaultFS</name>
4   <value>hdfs://Master.Hadoop:9000</value>
5   <description>The name of the default file system.</description>
6 </property>
7 <property>
8   <name>hadoop.tmp.dir</name>
9   <!-- 注意创建相关的目录结构 -->
10  <value>/usr/local/hadoop/temp</value>
11  <description>A base for other temporary directories.</description>
12 </property>
```

7、编辑<HADOOP_HOME>/etc/hadoop/hdfs-site.xml

在<configuration>节点下添加或者更新下面的配置信息:

```
1 <property>
2   <name>dfs.replication</name>
3   <!-- 值需要与实际的DataNode节点数要一致,本文为3 -->
4   <value>3</value>
5 </property>
6 <property>
7   <name>dfs.namenode.name.dir</name>
8   <!-- 注意创建相关的目录结构 -->
9   <value>file:/usr/local/hadoop/dfs/name</value>
10  <final>true</final>
11 </property>
12 <property>
13   <name>dfs.datanode.data.dir</name>
14   <!-- 注意创建相关的目录结构 -->
15   <value>file:/usr/local/hadoop/dfs/data</value>
16 </property>
```

8、编辑<HADOOP_HOME>/etc/hadoop/yarn-site.xml

在<configuration>节点下添加或者更新下面的配置信息:

```
1 <!-- micmiu.com -->
2 <property>
3   <name>yarn.nodemanager.aux-services</name>
4   <value>mapreduce_shuffle</value>
5 </property>
6 <property>
7   <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
8   <value>org.apache.hadoop.mapred.ShuffleHandler</value>
9 </property>
10
11 <!-- resourcemanager hostname或ip地址-->
12 <property>
13   <name>yarn.resourcemanager.hostname</name>
14   <value>Master.Hadoop</value>
15 </property>
```

9、编辑<HADOOP_HOME>/etc/hadoop/mapred-site.xml

默认没有mapred-site.xml文件,copy mapred-site.xml.template 一份为 mapred-site.xml即可

在<configuration>节点下添加或者更新下面的配置信息:

```
1 <!-- micmiu.com -->
2 <property>
3   <name>mapreduce.framework.name</name>
4   <value>yarn</value>
5   <final>true</final>
6 </property>
```

[三]、启动和测试

1、启动Hadoop

异常处理 (19)

杂谈 (2)

文章归档

- 2014年七月 (1)
- 2014年五月 (1)
- 2014年四月 (12)
- 2014年三月 (4)
- 2014年二月 (6)
- 2014年一月 (8)
- 2013年十二月 (1)
- 2013年十月 (9)
- 2013年九月 (7)
- 2013年八月 (8)
- 2013年六月 (4)
- 2013年五月 (4)
- 2013年四月 (3)
- 2013年三月 (10)
- 2013年二月 (1)
- 2013年一月 (2)
- 2012年十二月 (1)
- 2012年十一月 (7)
- 2012年十月 (2)
- 2012年九月 (5)
- 2012年八月 (13)
- 2012年七月 (14)
- 2012年六月 (12)
- 2012年五月 (16)
- 2012年四月 (20)
- 2012年三月 (22)
- 2012年二月 (88)

标签云(3D)

技术链接

- 10→me@iteye
- 21→nosqlfan
- 22→mysqllops
- 23→OpenCloudDB

1.1、第一次启动需要在Master.Hadoop 执行format `hdfs namenode -format`:

```
1 [hadoop@Master ~]$ hdfs namenode -format
2 14/01/22 15:43:10 INFO namenode.NameNode: STARTUP_MSG:
3 /*****
4 STARTUP_MSG: Starting NameNode
5 STARTUP_MSG: host = Master.Hadoop/192.168.6.77
6 STARTUP_MSG: args = [-format]
7 STARTUP_MSG: version = 2.2.0
8 STARTUP_MSG: classpath =
9 .....
10 .....micmiu.com.....
11 .....
12 STARTUP_MSG: java = 1.6.0_20
13 *****/
14 14/01/22 15:43:10 INFO namenode.NameNode: registered UNIX signal handlers for
15 [TERM, HUP, INT]
16 Formatting using clusterid: CID-645f2ed2-6f02-4c24-8cbc-82b09eca963d
17 14/01/22 15:43:11 INFO namenode.HostFileManager: read includes:
18 HostSet(
19 )
20 14/01/22 15:43:11 INFO namenode.HostFileManager: read excludes:
21 HostSet(
22 )
23 14/01/22 15:43:11 INFO blockmanagement.DatanodeManager:
24 dfs.block.invalidate.limit=1000
25 14/01/22 15:43:11 INFO util.GSet: Computing capacity for map BlocksMap
26 14/01/22 15:43:11 INFO util.GSet: VM type = 64-bit
27 14/01/22 15:43:11 INFO util.GSet: 2.0% max memory = 888.9 MB
28 14/01/22 15:43:11 INFO util.GSet: capacity = 2^21 = 2097152 entries
29 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
30 dfs.block.access.token.enable=false
31 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
32 defaultReplication = 3
33 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
34 maxReplication = 512
35 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
36 minReplication = 1
37 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
38 maxReplicationStreams = 2
39 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
40 shouldCheckForEnoughRacks = false
41 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
42 replicationRecheckInterval = 3000
43 14/01/22 15:43:11 INFO blockmanagement.BlockManager:
44 encryptDataTransfer = false
45 14/01/22 15:43:11 INFO namenode.FSNamesystem: fsOwner = hadoop
46 (auth:SIMPLE)
47 14/01/22 15:43:11 INFO namenode.FSNamesystem: supergroup = supergroup
48 14/01/22 15:43:11 INFO namenode.FSNamesystem: isPermissionEnabled = true
49 14/01/22 15:43:11 INFO namenode.FSNamesystem: HA Enabled: false
50 14/01/22 15:43:11 INFO namenode.FSNamesystem: Append Enabled: true
51 14/01/22 15:43:11 INFO util.GSet: Computing capacity for map INodeMap
52 14/01/22 15:43:11 INFO util.GSet: VM type = 64-bit
53 14/01/22 15:43:11 INFO util.GSet: 1.0% max memory = 888.9 MB
54 14/01/22 15:43:11 INFO util.GSet: capacity = 2^20 = 1048576 entries
55 14/01/22 15:43:11 INFO namenode.NameNode: Caching file names occurring more
56 than 10 times
57 14/01/22 15:43:11 INFO namenode.FSNamesystem: dfs.namenode.safemode.threshold-
58 pct = 0.9990000128746033
59 14/01/22 15:43:11 INFO namenode.FSNamesystem:
60 dfs.namenode.safemode.min.datanodes = 0
61 14/01/22 15:43:11 INFO namenode.FSNamesystem:
62 dfs.namenode.safemode.extension = 30000
63 14/01/22 15:43:11 INFO namenode.FSNamesystem: Retry cache on namenode is
64 enabled
65 14/01/22 15:43:11 INFO namenode.FSNamesystem: Retry cache will use 0.03 of
66 total heap and retry cache entry expiry time is 600000 millis
67 14/01/22 15:43:11 INFO util.GSet: Computing capacity for map Namenode Retry
68 Cache
69 14/01/22 15:43:11 INFO util.GSet: VM type = 64-bit
70 14/01/22 15:43:11 INFO util.GSet: 0.029999999329447746% max memory = 888.9 MB
71 14/01/22 15:43:11 INFO util.GSet: capacity = 2^15 = 32768 entries
72 14/01/22 15:43:11 INFO common.Storage: Storage directory
73 /usr/local/hadoop/dfs/name has been successfully formatted.
74 14/01/22 15:43:11 INFO namenode.FSImage: Saving image file
75 /usr/local/hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 using no
76 compression
77 14/01/22 15:43:11 INFO namenode.FSImage: Image file
78 /usr/local/hadoop/dfs/name/current/fsimage.ckpt_000000000000000000 of size
79 198 bytes saved in 0 seconds.
```

- 31→CrazyJvm
- 32→封神无度
- 33→董的博客

功能

- 登录
- 文章RSS
- 评论RSS
- WordPress.org

```
57 14/01/22 15:43:11 INFO namenode.NNStorageRetentionManager: Going to retain 1
images with txid >= 0
58 14/01/22 15:43:11 INFO util.ExitUtil: Exiting with status 0
59 14/01/22 15:43:11 INFO namenode.NameNode: SHUTDOWN_MSG:
60 /*****
61 SHUTDOWN_MSG: Shutting down NameNode at Master.Hadoop/192.168.6.77
62 *****/
```

1.2、在Master.Hadoop执行 start-dfs.sh :

```
1 [hadoop@Master ~]$ start-dfs.sh
2 Starting namenodes on [Master.Hadoop]
3 Master.Hadoop: starting namenode, logging to /usr/local/hadoop-
2.2.0/logs/hadoop-hadoop-namenode-Master.Hadoop.out
4 Slave7.Hadoop: starting datanode, logging to /usr/local/hadoop-
2.2.0/logs/hadoop-hadoop-datanode-Slave7.Hadoop.out
5 Slave5.Hadoop: starting datanode, logging to /usr/local/hadoop-
2.2.0/logs/hadoop-hadoop-datanode-Slave5.Hadoop.out
6 Slave6.Hadoop: starting datanode, logging to /usr/local/hadoop-
2.2.0/logs/hadoop-hadoop-datanode-Slave6.Hadoop.out
7 Starting secondary namenodes [0.0.0.0]
8 0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop-
2.2.0/logs/hadoop-hadoop-secondarynamenode-Master.Hadoop.out
```

在Master.Hadoop 验证启动进程:

```
1 [hadoop@Master ~]$ jps
2 7695 Jps
3 7589 SecondaryNameNode
4 7403 NameNode
```

在SlaveX.Hadop 验证启动进程如下:

```
1 [hadoop@Slave5 ~]$ jps
2 8724 DataNode
3 8815 Jps
```

1.3、在Master.Hadoop 执行 start-yarn.sh :

```
1 [hadoop@Master ~]$ start-yarn.sh
2 starting yarn daemons
3 starting resourcemanager, logging to /usr/local/hadoop-2.2.0/logs/yarn-hadoop-
resourcemanager-Master.Hadoop.out
4 Slave7.Hadoop: starting nodemanager, logging to /usr/local/hadoop-
2.2.0/logs/yarn-hadoop-nodemanager-Slave7.Hadoop.out
5 Slave5.Hadoop: starting nodemanager, logging to /usr/local/hadoop-
2.2.0/logs/yarn-hadoop-nodemanager-Slave5.Hadoop.out
6 Slave6.Hadoop: starting nodemanager, logging to /usr/local/hadoop-
2.2.0/logs/yarn-hadoop-nodemanager-Slave6.Hadoop.out
```

在Master.Hadoop 验证启动进程:

```
1 [hadoop@Master ~]$ jps
2 8071 Jps
3 7589 SecondaryNameNode
4 7821 ResourceManager
5 7403 NameNode
```

在SlaveX.Hadop 验证启动进程如下:

```
1 [hadoop@Slave5 ~]$ jps
2 9013 Jps
3 8724 DataNode
4 8882 NodeManager
```

2、演示

2.1、演示hdfs 一些常用命令,为wordcount演示做准备:

```
1 [hadoop@Master ~]$ hdfs dfs -ls /
2 [hadoop@Master ~]$ hdfs dfs -mkdir /user
3 [hadoop@Master ~]$ hdfs dfs -mkdir -p /user/micmiu/wordcount/in
4 [hadoop@Master ~]$ hdfs dfs -ls /user/micmiu/wordcount
5 Found 1 items
```

```
6 drwxr-xr-x - hadoop supergroup 0 2014-01-22 16:01
  /user/micmiu/wordcount/in
```

2.2、本地创建三个文件 micmiu-01.txt、micmiu-02.txt、micmiu-03.txt, 分别写入如下内容:

micmiu-01.txt:

```
Hi Michael welcome to Hadoop
more see micmiu.com
```

micmiu-02.txt:

```
Hi Michael welcome to BigData
more see micmiu.com
```

micmiu-03.txt:

```
Hi Michael welcome to Spark
more see micmiu.com
```

把 micmiu 打头的三个文件上传到hdfs:

```
1 [hadoop@Master ~]$ hdfs dfs -put micmiu*.txt /user/micmiu/wordcount/in
2 [hadoop@Master ~]$ hdfs dfs -ls /user/micmiu/wordcount/in
3 Found 3 items
4 -rw-r--r-- 3 hadoop supergroup 50 2014-01-22 16:06
  /user/micmiu/wordcount/in/micmiu-01.txt
5 -rw-r--r-- 3 hadoop supergroup 50 2014-01-22 16:06
  /user/micmiu/wordcount/in/micmiu-02.txt
6 -rw-r--r-- 3 hadoop supergroup 49 2014-01-22 16:06
  /user/micmiu/wordcount/in/micmiu-03.txt
```

2.3、然后cd 切换到Hadoop的根目录下执行:

```
hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.2.0.jar
wordcount /user/micmiu/wordcount/in /user/micmiu/wordcount/out
```

ps: hdfs 中 /user/micmiu/wordcount/out 目录不能存在 否则运行报错。

看到类似如下的日志信息:

```
1 [hadoop@Master hadoop]$ hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-
  examples-2.2.0.jar wordcount /user/micmiu/wordcount/in
  /user/micmiu/wordcount/out
2 14/01/22 16:36:28 INFO client.RMProxy: Connecting to ResourceManager at
  Master.Hadoop/192.168.6.77:8032
3 14/01/22 16:36:29 INFO input.FileInputFormat: Total input paths to process : 3
4 14/01/22 16:36:29 INFO mapreduce.JobSubmitter: number of splits:3
5 .....
6 .....micmiu.com.....
7 .....
8 File System Counters
9 FILE: Number of bytes read=297
10 FILE: Number of bytes written=317359
11 FILE: Number of read operations=0
12 FILE: Number of large read operations=0
13 FILE: Number of write operations=0
14 HDFS: Number of bytes read=536
15 HDFS: Number of bytes written=83
16 HDFS: Number of read operations=12
17 HDFS: Number of large read operations=0
18 HDFS: Number of write operations=2
19 Job Counters
20 Launched map tasks=3
21 Launched reduce tasks=1
22 Data-local map tasks=3
23 Total time spent by all maps in occupied slots (ms)=55742
```

| | |
|----|---|
| 24 | Total time spent by all reduces in occupied slots (ms)=3933 |
| 25 | Map-Reduce Framework |
| 26 | Map input records=6 |
| 27 | Map output records=24 |
| 28 | Map output bytes=243 |
| 29 | Map output materialized bytes=309 |
| 30 | Input split bytes=387 |
| 31 | Combine input records=24 |
| 32 | Combine output records=24 |
| 33 | Reduce input groups=10 |
| 34 | Reduce shuffle bytes=309 |
| 35 | Reduce input records=24 |
| 36 | Reduce output records=10 |
| 37 | Spilled Records=48 |
| 38 | Shuffled Maps =3 |
| 39 | Failed Shuffles=0 |
| 40 | Merged Map outputs=3 |
| 41 | GC time elapsed (ms)=1069 |
| 42 | CPU time spent (ms)=12390 |
| 43 | Physical memory (bytes) snapshot=846753792 |
| 44 | Virtual memory (bytes) snapshot=5155561472 |
| 45 | Total committed heap usage (bytes)=499580928 |
| 46 | Shuffle Errors |
| 47 | BAD_ID=0 |
| 48 | CONNECTION=0 |
| 49 | IO_ERROR=0 |
| 50 | WRONG_LENGTH=0 |
| 51 | WRONG_MAP=0 |
| 52 | WRONG_REDUCE=0 |
| 53 | File Input Format Counters |
| 54 | Bytes Read=149 |
| 55 | File Output Format Counters |
| 56 | Bytes Written=83 |

到此 wordcount的job已经执行完成,执行如下命令可以查看刚才job的执行结果:

| | |
|----|--|
| 1 | [hadoop@Master hadoop]\$ hdfs dfs -ls /user/micmiu/wordcount/out |
| 2 | Found 2 items |
| 3 | -rw-r--r-- 3 hadoop supergroup 0 2014-01-22 16:38 |
| 4 | /user/micmiu/wordcount/out/_SUCCESS |
| 5 | -rw-r--r-- 3 hadoop supergroup 83 2014-01-22 16:38 |
| 6 | /user/micmiu/wordcount/out/part-r-00000 |
| 7 | [hadoop@Master hadoop]\$ hdfs dfs -cat /user/micmiu/wordcount/out/part-r-00000 |
| 8 | BigData 1 |
| 9 | Hadoop 1 |
| 10 | Hi 3 |
| 11 | Michael 3 |
| 12 | Spark 1 |
| 13 | micmiu.com 3 |
| 14 | more 3 |
| 15 | see 3 |
| 16 | to 3 |
| 17 | welcome 3 |

打开浏览器输入:http://192.168.6.77(Master.Hadoop):8088 可查看相关的应用运行情况。

----- EOF @Michael Sun -----


原创文章,转载请注明: 转载自 micmiu – 软件开发+生活点滴 [<http://www.micmiu.com/>]

本文链接地址: <http://www.micmiu.com/bigdata/hadoop/hadoop2x-cluster-setup/>

 Hadoop  Hadoop

[← hadoop 2.2.0 单节点安装和测试](#)

[Hive安装配置详解 →](#)

 发表评论? 0 条评论。

发表评论

昵称

邮箱

网址



注意 - 你可以用以下 HTML tags and attributes:
 <abbr title=""> <acronym title=""> <blockquote
cite=""> <code> <del datetime=""> <i> <q cite=""> <strike>
