目录

[**UDH 开启Hadoop和Hbase集群的lzo压缩功能** 1](#_Toc417548659)

[1. 配置修改hdfs服务对应core-site.xml 1](#_Toc417548660)

[2. 配置修改MapReduce2服务对应mapred-site.xml 3](#_Toc417548661)

[3. Hadoop的lzo功能测试 4](#_Toc417548662)

[4. Hbase 的 lzo 功能测试 5](#_Toc417548663)

[5. 总结 6](#_Toc417548664)

**UDH 开启Hadoop和Hbase集群的lzo压缩功能**

在UDH上配置使用Hadoop和Hbase集群的lzo压缩功能，不需要在集群各个节点上去安装Hadoop-lzo和lzo相关的包，因为UDH在管理安装Hadoop集群时默认已将Hadoop-lzo，lzo，lzo-devel等相关包安装在集群各个节点上，而只需要通过修改UDH web（http://manager-server:8080）上hdfs和mapreduce服务的lzo相关配置项，即能同步配置到UDH管理的所有机器节点。

# 配置修改hdfs服务对应core-site.xml

 需要修改的属性io.compression.codecs和io.compression.codec.lzo.class

<property>

<name>io.compression.codecs</name>

<value>org.apache.hadoop.io.compress.DefaultCodec,org.apache.hadoop.io.compress.GzipCodec,org.apache.hadoop.io.compress.BZip2Codec,com.hadoop.compression.lzo.LzoCodec,com.hadoop.compression.lzo.LzopCodec</value>

</property>

<property>

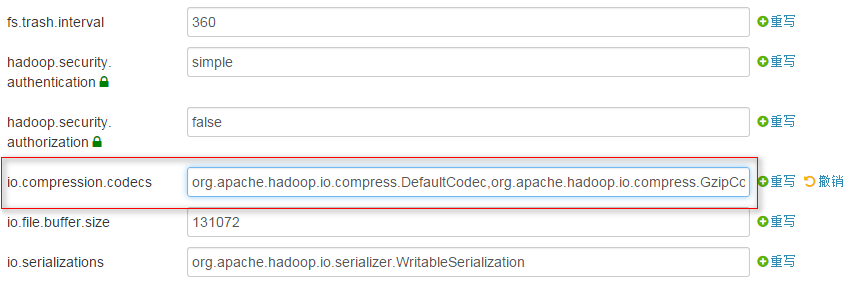
       <name>io.compression.codec.lzo.class</name>

        <value>com.hadoop.compression.lzo.LzoCodec</value>

</property>

  对应 UDH web界面上的修改如下：

io.compression.codecs属性的修改



 io.compression.codec.lzo.class属性的添加分如下3步

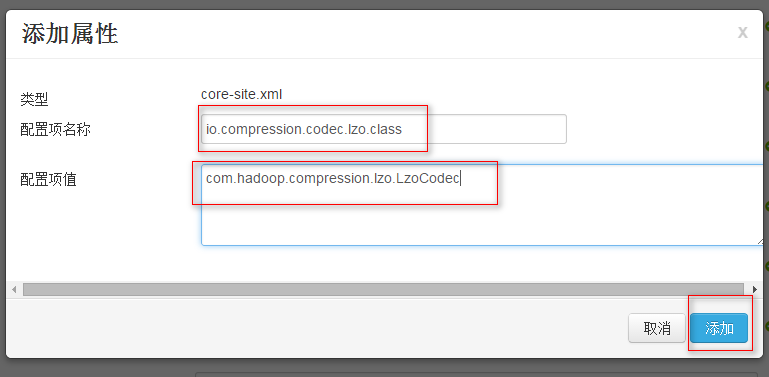
 a，hdfs服务配置界面下，



 b，单击添加属性



  c，添加 io.compression.codec.lzo.class 属性，



注意：lzo.class 的值为com.hadoop而不是org.hadoop ，如果在UDH上配置为org.hadoop在使用lzo压缩功能时会报找不到该类，

保存以上修改并在UDH上重启HDFS服务。

# 配置修改MapReduce2服务对应mapred-site.xml

<property>

        <name>mapred.compress.map.output</name>

        <value>true</value>

</property>

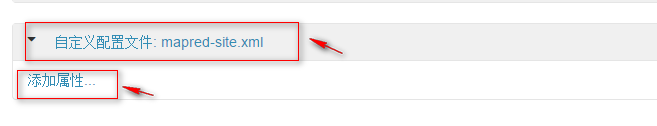
<property>

        <name>mapred.map.output.compression.codec</name>

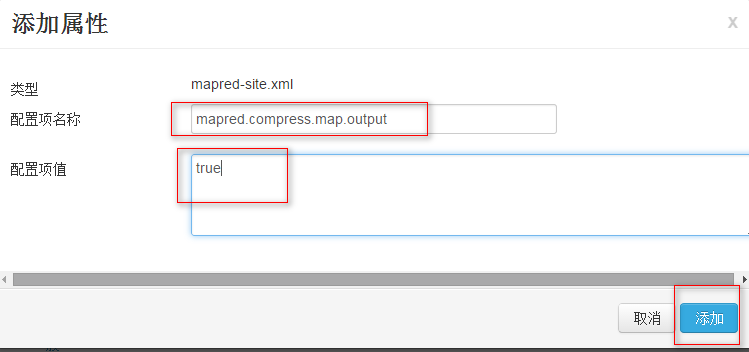
        <value>com.hadoop.compression.lzo.LzoCodec</value>

</property>

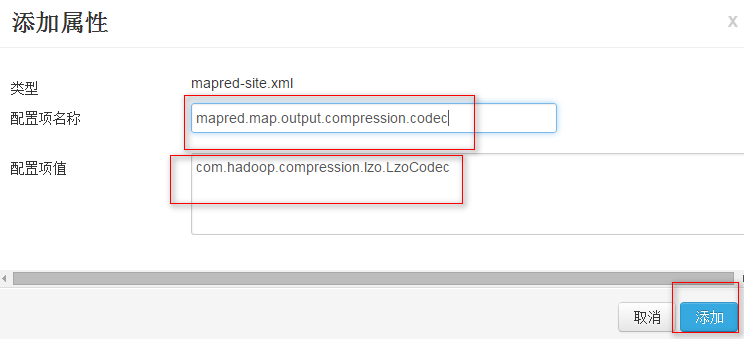
对应 UDH web界面上的修改如下，



a、添加mapred.compress.map.output 属性，



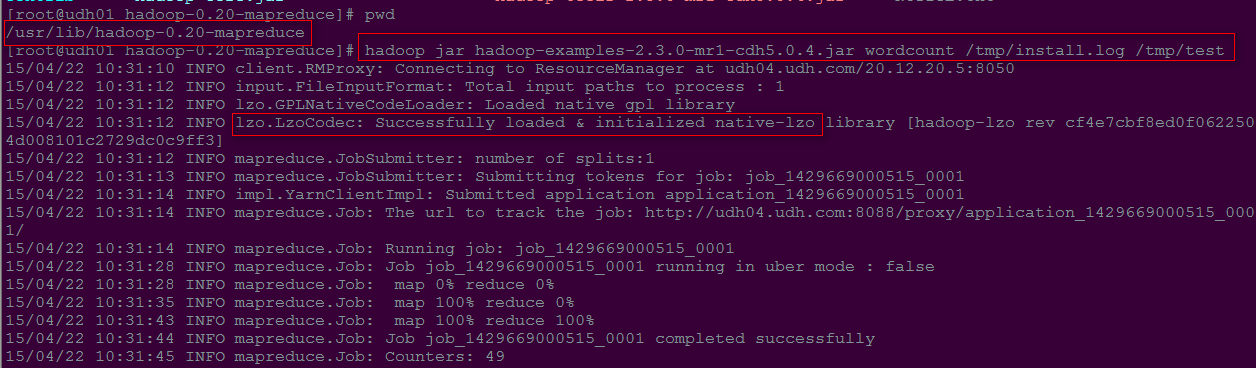
b、添加mapred.map.output.compression.codec属性，



保存以上修改并在UDH上重启MapReduce2服务

# Hadoop的lzo功能测试

在UDH集群的任意一节点上，上传一个文本文件到hdfs上，执行mapreduce wordcount测试，从日志中可以发现lzo加载成功。

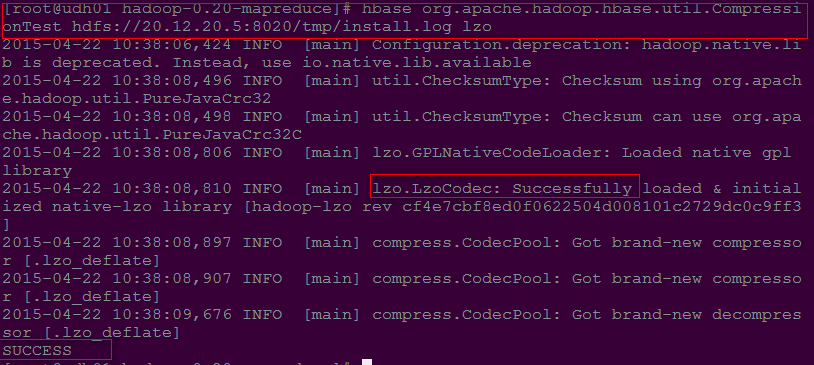


# Hbase 的 lzo 功能测试

执行hbase的lzo example 来验证，执行如下命令：

hbase org.apache.hadoop.hbase.util.CompressionTest

hdfs://20.12.20.5:8020/tmp/install.log lzo



 hbase 应用

LZO对比Hbase默认的Gzip，前者性能比较高，后者压缩比较高，对与想提高Hbase读写性能的用户，采用LZO压缩是比较好的选择，

  create 'test',{NAME=>'f1',COMPRESSION=>'lzo'}   按照列族压缩

# 总结

在UDH上配置开启Hadoop和hbase的lzo压缩功能 只需在UDH server上修改配置，简单易操作，而不需手动修改集群里各个节点的配置。