

Shark 0.9.1安装遇到的问题及解决办法

这几天在集群上部署了Shark 0.9.1，我下载的是已经编译好的，Hadoop版本是2.2.0，下面就总结一下我在安装Shark的过程中遇到的问题及其解决方案。

一、YARN mode not available ?

```
Exception in thread "main" org.apache.spark.SparkException: YARN mode not available ?
  at org.apache.spark.SparkContext$.org$apache$spark$SparkContext$$createTaskScheduler(S
parkContext.scala:1275)
  at org.apache.spark.SparkContext.<init>(SparkContext.scala:201)
  at shark.SharkContext.</init><init>(SharkContext.scala:42)
  at shark.SharkContext.</init><init>(SharkContext.scala:61)
  at shark.SharkEnv$.initWithSharkContext(SharkEnv.scala:78)
  at shark.SharkEnv$.init(SharkEnv.scala:38)
  at shark.SharkCliDriver.</init><init>(SharkCliDriver.scala:278)
  at shark.SharkCliDriver$.main(SharkCliDriver.scala:162)
  at shark.SharkCliDriver.main(SharkCliDriver.scala)
Caused by: java.lang.ClassNotFoundException: org.apache.spark.scheduler.cluster.YarnClientC
lusterScheduler
  at java.net.URLClassLoader$1.run(URLClassLoader.java:366)
  at java.net.URLClassLoader$1.run(URLClassLoader.java:355)
  at java.security.AccessController.doPrivileged(Native Method)
  at java.net.URLClassLoader.findClass(URLClassLoader.java:354)
  at java.lang.ClassLoader.loadClass(ClassLoader.java:425)
  at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:308)
  at java.lang.ClassLoader.loadClass(ClassLoader.java:358)
  at java.lang.Class.forName0(Native Method)
  at java.lang.Class.forName(Class.java:190)
  at org.apache.spark.SparkContext$.org$apache$spark$SparkContext$$createTaskScheduler(S
parkContext.scala:1269)
  ... 8 more
14/04/29 11:23:37 WARN conf.Configuration: org.apache.hadoop.hive.conf.LoopingByteArrayI
nputStream@1da98852:an attempt to override final parameter: mapreduce.job.end-
notification.max.retry.interval; Ignoring.
```

解决方法是在SHARK_HOME中的run文件加入spark-assembly-0.9.1-hadoop2.2.0.jar文件到SPARK_CLASSPATH中：

```
SPARK_CLASSPATH+=:$SHARK_HOME"/jars/spark-assembly-0.9.1-hadoop2.2.0.jar"
```

二、overrides final method getUnknownFields

这个错误经常遇到

```
14/04/29 11:26:35 INFO service.AbstractService: Service org.apache.hadoop.yarn.client.api.impl.YarnClientImpl failed in state STARTED; cause: org.apache.hadoop.yarn.exceptions.YarnRuntimeException: java.lang.reflect.InvocationTargetException
org.apache.hadoop.yarn.exceptions.YarnRuntimeException: java.lang.reflect.InvocationTargetException
    at org.apache.hadoop.yarn.factories.impl.pb.RpcClientFactoryPBImpl.getClient(RpcClientFactoryPBImpl.java:79)
    at org.apache.hadoop.yarn.ipc.HadoopYarnProtoRPC.getProxy(HadoopYarnProtoRPC.java:48)
    at org.apache.hadoop.yarn.client.RMProxy$1.run(RMProxy.java:134)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:356)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1528)
    at org.apache.hadoop.yarn.client.RMProxy.getProxy(RMProxy.java:130)
    at org.apache.hadoop.yarn.client.RMProxy.createRMProxy(RMProxy.java:93)
    at org.apache.hadoop.yarn.client.ClientRMProxy.createRMProxy(ClientRMProxy.java:70)
    at org.apache.hadoop.yarn.client.api.impl.YarnClientImpl.serviceStart(YarnClientImpl.java:114)
    at org.apache.hadoop.service.AbstractService.start(AbstractService.java:193)
    at org.apache.spark.deploy.yarn.Client.runApp(Client.scala:76)
    at org.apache.spark.scheduler.cluster.YarnClientSchedulerBackend.start(YarnClientSchedulerBackend.scala:78)
    at org.apache.spark.scheduler.TaskSchedulerImpl.start(TaskSchedulerImpl.scala:126)
    at org.apache.spark.SparkContext.</init><init>(SparkContext.scala:202)
    at shark.SharkContext.</init><init>(SharkContext.scala:42)
    at shark.SharkContext.</init><init>(SharkContext.scala:61)
    at shark.SharkEnv$.initWithSharkContext(SharkEnv.scala:78)
    at shark.SharkEnv$.init(SharkEnv.scala:38)
    at shark.SharkCliDriver.</init><init>(SharkCliDriver.scala:278)
    at shark.SharkCliDriver$.main(SharkCliDriver.scala:162)
    at shark.SharkCliDriver.main(SharkCliDriver.scala)
Caused by: java.lang.reflect.InvocationTargetException
    at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
    at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:57)
    at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
    at java.lang.reflect.Constructor.newInstance(Constructor.java:526)
```

at org.apache.hadoop.yarn.factories.impl.pb.RpcClientFactoryPBImpImpl.getClient(RpcClientFacto
ryPBImpImpl.java:76)

... 21 more

Caused by: java.lang.VerifyError: class org.apache.hadoop.security.proto.SecurityProtos\$GetD
elegationTokenRequestProto overrides final method getUnknownFields().()Lcom/google/protob
uf/UnknownFieldSet;

at java.lang.ClassLoader.defineClass1(Native Method)
at java.lang.ClassLoader.defineClass(ClassLoader.java:800)
at java.security.SecureClassLoader.defineClass(SecureClassLoader.java:142)
at java.net.URLClassLoader.defineClass(URLClassLoader.java:449)
at java.net.URLClassLoader.access\$100(URLClassLoader.java:71)
at java.net.URLClassLoader\$1.run(URLClassLoader.java:361)
at java.net.URLClassLoader\$1.run(URLClassLoader.java:355)
at java.security.AccessController.doPrivileged(Native Method)
at java.net.URLClassLoader.findClass(URLClassLoader.java:354)
at java.lang.ClassLoader.loadClass(ClassLoader.java:425)
at sun.misc.Launcher\$AppClassLoader.loadClass(Launcher.java:308)
at java.lang.ClassLoader.loadClass(ClassLoader.java:358)
at java.lang.Class.getDeclaredMethods0(Native Method)
at java.lang.Class.privateGetDeclaredMethods(Class.java:2531)
at java.lang.Class.privateGetPublicMethods(Class.java:2651)
at java.lang.Class.privateGetPublicMethods(Class.java:2661)
at java.lang.Class.getMethods(Class.java:1467)
at sun.misc.ProxyGenerator.generateClassFile(ProxyGenerator.java:426)
at sun.misc.ProxyGenerator.generateProxyClass(ProxyGenerator.java:323)
at java.lang.reflect.Proxy.getProxyClass0(Proxy.java:636)
at java.lang.reflect.Proxy.newProxyInstance(Proxy.java:722)
at org.apache.hadoop.ipc.ProtobufRpcEngine.getProxy(ProtobufRpcEngine.java:92)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:537)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:482)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:447)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:600)
at org.apache.hadoop.ipc.RPC.getProxy(RPC.java:557)
at org.apache.hadoop.yarn.api.impl.pb.client.ApplicationClientProtocolPBClientImpl.</init><i
nit>(ApplicationClientProtocolPBClientImpl.java:111)

... 26 more

Exception in thread "main" org.apache.hadoop.yarn.exceptions.YarnRuntimeException: java.la
ng.reflect.InvocationTargetException

at org.apache.hadoop.yarn.factories.impl.pb.RpcClientFactoryPBImpImpl.getClient(RpcClientFacto
ryPBImpImpl.java:79)
at org.apache.hadoop.yarn.ipc.HadoopYarnProtoRPC.getProxy(HadoopYarnProtoRPC.java:48)
at org.apache.hadoop.yarn.client.RMProxy\$1.run(RMProxy.java:134)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:356)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1528
)

```
at org.apache.hadoop.yarn.client.RMProxy.getProxy(RMProxy.java:130)
at org.apache.hadoop.yarn.client.RMProxy.createRMProxy(RMProxy.java:93)
at org.apache.hadoop.yarn.client.ClientRMProxy.createRMProxy(ClientRMProxy.java:70)
at org.apache.hadoop.yarn.client.api.impl.YarnClientImpl.serviceStart(YarnClientImpl.java:114)
at org.apache.hadoop.service.AbstractService.start(AbstractService.java:193)
at org.apache.spark.deploy.yarn.Client.runApp(Client.scala:76)
at org.apache.spark.scheduler.cluster.YarnClientSchedulerBackend.start(YarnClientSchedulerBackend.scala:78)
at org.apache.spark.scheduler.TaskSchedulerImpl.start(TaskSchedulerImpl.scala:126)
at org.apache.spark.SparkContext.</init><init>(SparkContext.scala:202)
at shark.SharkContext.</init><init>(SharkContext.scala:42)
at shark.SharkContext.</init><init>(SharkContext.scala:61)
at shark.SharkEnv$.initWithSharkContext(SharkEnv.scala:78)
at shark.SharkEnv$.init(SharkEnv.scala:38)
at shark.SharkCliDriver.</init><init>(SharkCliDriver.scala:278)
at shark.SharkCliDriver$.main(SharkCliDriver.scala:162)
at shark.SharkCliDriver.main(SharkCliDriver.scala)
Caused by: java.lang.reflect.InvocationTargetException
at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:57)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:526)
at org.apache.hadoop.yarn.factories.impl.pb.RpcClientFactoryPBImpl.getClient(RpcClientFactoryPBImpl.java:76)
... 21 more
Caused by: java.lang.VerifyError: class org.apache.hadoop.security.proto.SecurityProtos$GetDelegationTokenRequestProto overrides final method getUnknownFields()Lcom/google/protobuf/UnknownFieldSet;
at java.lang.ClassLoader.defineClass1(Native Method)
at java.lang.ClassLoader.defineClass(ClassLoader.java:800)
at java.security.SecureClassLoader.defineClass(SecureClassLoader.java:142)
at java.net.URLClassLoader.defineClass(URLClassLoader.java:449)
at java.net.URLClassLoader.access$100(URLClassLoader.java:71)
at java.net.URLClassLoader$1.run(URLClassLoader.java:361)
at java.net.URLClassLoader$1.run(URLClassLoader.java:355)
at java.security.AccessController.doPrivileged(Native Method)
at java.net.URLClassLoader.findClass(URLClassLoader.java:354)
at java.lang.ClassLoader.loadClass(ClassLoader.java:425)
at sun.misc.Launcher$AppClassLoader.loadClass(Launcher.java:308)
at java.lang.ClassLoader.loadClass(ClassLoader.java:358)
at java.lang.Class.getDeclaredMethods0(Native Method)
at java.lang.Class.privateGetDeclaredMethods(Class.java:2531)
at java.lang.Class.privateGetPublicMethods(Class.java:2651)
```

```
at java.lang.Class.privateGetPublicMethods(Class.java:2661)
at java.lang.Class.getMethods(Class.java:1467)
at sun.misc.ProxyGenerator.generateClassFile(ProxyGenerator.java:426)
at sun.misc.ProxyGenerator.generateProxyClass(ProxyGenerator.java:323)
at java.lang.reflect.Proxy.getProxyClass0(Proxy.java:636)
at java.lang.reflect.Proxy.newProxyInstance(Proxy.java:722)
at org.apache.hadoop.ipc.ProtobufRpcEngine.getProxy(ProtobufRpcEngine.java:92)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:537)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:482)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:447)
at org.apache.hadoop.ipc.RPC.getProtocolProxy(RPC.java:600)
at org.apache.hadoop.ipc.RPC.getProxy(RPC.java:557)
at org.apache.hadoop.yarn.api.impl.pb.client.ApplicationClientProtocolPBClientImpl.</init><i
nit>(ApplicationClientProtocolPBClientImpl.java:111)
... 26 more
```

解决办法：

到\$SHARK_HOME/lib_managed/jars/org.spark-project.protobuf/protobuf-java目录下可以看到有个protobuf-java-2.4.1-shaded.jar文件，将这个文件删掉，并拷贝protobuf-java-2.5.0.jar到该目录即可。

三、java.lang.NullPointerException

```
Exception in thread "main" java.lang.NullPointerException
at scala.collection.mutable.ArrayOps$ofRef$.length$extension(ArrayOps.scala:114)
at scala.collection.mutable.ArrayOps$ofRef.length(ArrayOps.scala:114)
at scala.collection.IndexedSeqOptimized$class.foreach(IndexedSeqOptimized.scala:32)
at scala.collection.mutable.ArrayOps$ofRef.foreach(ArrayOps.scala:108)
at org.apache.spark.deploy.yarn.Client$.populateHadoopClasspath(Client.scala:498)
at org.apache.spark.deploy.yarn.Client$.populateClasspath(Client.scala:519)
at org.apache.spark.deploy.yarn.Client.setupLaunchEnv(Client.scala:333)
at org.apache.spark.deploy.yarn.Client.runApp(Client.scala:94)
at org.apache.spark.scheduler.cluster.YarnClientSchedulerBackend.start(YarnClientSchedulerBackend.scala:78)
at org.apache.spark.scheduler.TaskSchedulerImpl.start(TaskSchedulerImpl.scala:126)
at org.apache.spark.SparkContext.</init><init>(SparkContext.scala:202)
at shark.SharkContext.</init><init>(SharkContext.scala:42)
at shark.SharkContext.</init><init>(SharkContext.scala:61)
at shark.SharkEnv$.initWithSharkContext(SharkEnv.scala:78)
at shark.SharkEnv$.init(SharkEnv.scala:38)
at shark.SharkCliDriver.</init><init>(SharkCliDriver.scala:278)
at shark.SharkCliDriver$.main(SharkCliDriver.scala:162)
at shark.SharkCliDriver.main(SharkCliDriver.scala)</init>
```

这是因为spark启动时需要读取yarn-site.xml中的yarn.application.classpath，如果此参数没有显示配置，则默认的值是空，这时会抛出异常，将yarn.application.classpath设置一下即可。

四、Unable to instantiate org.apache.hadoop.hive.metastore.HiveMetaStoreClient

```
Exception in thread "main" org.apache.hadoop.hive.ql.metadata.HiveException: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.metastore.HiveMetaStoreClient
    at org.apache.hadoop.hive.ql.metadata.Hive.getAllDatabases(Hive.java:1072)
    at shark.memstore2.TableRecovery$.reloadRdds(TableRecovery.scala:49)
    at shark.SharkCliDriver.<init>(SharkCliDriver.scala:283)
    at shark.SharkCliDriver$.main(SharkCliDriver.scala:162)
    at shark.SharkCliDriver.main(SharkCliDriver.scala)
Caused by: java.lang.RuntimeException: Unable to instantiate org.apache.hadoop.hive.metastore.HiveMetaStoreClient
    at org.apache.hadoop.hive.metastore.MetaStoreUtils.newInstance(MetaStoreUtils.java:1139)
    at org.apache.hadoop.hive.metastore.RetryingMetaStoreClient.<init>(RetryingMetaStoreClient.java:51)
    at org.apache.hadoop.hive.metastore.RetryingMetaStoreClient.getProxy(RetryingMetaStoreClient.java:61)
    at org.apache.hadoop.hive.ql.metadata.Hive.createMetaStoreClient(Hive.java:2288)
    at org.apache.hadoop.hive.ql.metadata.Hive.getMSC(Hive.java:2299)
    at org.apache.hadoop.hive.ql.metadata.Hive.getAllDatabases(Hive.java:1070)
    ... 4 more
Caused by: java.lang.reflect.InvocationTargetException
    at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
    at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:57)
    at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
    at java.lang.reflect.Constructor.newInstance(Constructor.java:526)
    at org.apache.hadoop.hive.metastore.MetaStoreUtils.newInstance(MetaStoreUtils.java:1137)
    ... 9 more
Caused by: javax.jdo.JDOFatalInternalException: Error creating transactional connection factory
NestedThrowables:
java.lang.reflect.InvocationTargetException
    at org.datanucleus.api.jdo.NucleusJDOHelper.getJDOExceptionForNucleusException(NucleusJDOHelper.java:587)
    at org.datanucleus.api.jdo.JDOPersistenceManagerFactory.freezeConfiguration(JDOPersistenceManagerFactory.java:781)
    at org.datanucleus.api.jdo.JDOPersistenceManagerFactory.createPersistenceManagerFactory(JDOPersistenceManagerFactory.java:326)
```



```
at org.datanucleus.api.jdo.JDOPersistenceManagerFactory.getPersistenceManagerFactory(JDOPersistenceManagerFactory.java:195)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at javax.jdo.JDOHelper$16.run(JDOHelper.java:1958)
at java.security.AccessController.doPrivileged(Native Method)
at javax.jdo.JDOHelper.invoke(JDOHelper.java:1953)
at javax.jdo.JDOHelper.invokeGetPersistenceManagerFactoryOnImplementation(JDOHelper.java:1159)
at javax.jdo.JDOHelper.getPersistenceManagerFactory(JDOHelper.java:803)
at javax.jdo.JDOHelper.getPersistenceManagerFactory(JDOHelper.java:698)
at org.apache.hadoop.hive.metastore.ObjectStore.getPMF(ObjectStore.java:270)
at org.apache.hadoop.hive.metastore.ObjectStore.getPersistenceManager(ObjectStore.java:299)
at org.apache.hadoop.hive.metastore.ObjectStore.initialize(ObjectStore.java:229)
at org.apache.hadoop.hive.metastore.ObjectStore.setConf(ObjectStore.java:204)
at org.apache.hadoop.util.ReflectionUtils.setConf(ReflectionUtils.java:73)
at org.apache.hadoop.util.ReflectionUtils.newInstance(ReflectionUtils.java:133)
at org.apache.hadoop.hive.metastore.RetryingRawStore.</init><init>(RetryingRawStore.java:62)
at org.apache.hadoop.hive.metastore.RetryingRawStore.getProxy(RetryingRawStore.java:71)
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.newRawStore(HiveMetaStore.java:413)
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.getMS(HiveMetaStore.java:401)
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.createDefaultDB(HiveMetaStore.java:439)
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.init(HiveMetaStore.java:325)
at org.apache.hadoop.hive.metastore.HiveMetaStore$HMSHandler.</init><init>(HiveMetaStore.java:285)
at org.apache.hadoop.hive.metastore.RetryingHMSHandler.</init><init>(RetryingHMSHandler.java:54)
at org.apache.hadoop.hive.metastore.RetryingHMSHandler.getProxy(RetryingHMSHandler.java:59)
at org.apache.hadoop.hive.metastore.HiveMetaStore.newHMSHandler(HiveMetaStore.java:4102)
at org.apache.hadoop.hive.metastore.HiveMetaStoreClient.</init><init>(HiveMetaStoreClient.java:121)
... 14 more
Caused by: java.lang.reflect.InvocationTargetException
at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:57)
```

```
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccesso
rImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:526)
at org.datanucleus.plugin.NonManagedPluginRegistry.createExecutableExtension(NonManage
dPluginRegistry.java:631)
at org.datanucleus.plugin.PluginManager.createExecutableExtension(PluginManager.java:325)
at org.datanucleus.store.AbstractStoreManager.registerConnectionFactory(AbstractStoreMan
ager.java:281)
at org.datanucleus.store.AbstractStoreManager.</init><init>(AbstractStoreManager.java:239)
at org.datanucleus.store.rdbms.RDBMSStoreManager.</init><init>(RDBMSStoreManager.java:
292)
at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.ja
va:57)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccesso
rImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:526)
at org.datanucleus.plugin.NonManagedPluginRegistry.createExecutableExtension(NonManage
dPluginRegistry.java:631)
at org.datanucleus.plugin.PluginManager.createExecutableExtension(PluginManager.java:301)
at org.datanucleus.NucleusContext.createStoreManagerForProperties(NucleusContext.java:10
69)
at org.datanucleus.NucleusContext.initialise(NucleusContext.java:359)
at org.datanucleus.api.jdo.JDOPersistenceManagerFactory.freezeConfiguration(JDOPersistenc
eManagerFactory.java:768)
... 43 more
Caused by: org.datanucleus.exceptions.NucleusException: Attempt to invoke the "DBCP" plugi
n to create a ConnectionPool gave an error : The specified datastore driver ("com.mysql.jdbc.D
river") was not found in the CLASSPATH. Please check your CLASSPATH specification, and the n
ame of the driver.
at org.datanucleus.store.rdbms.ConnectionFactoryImpl.generateDataSources(ConnectionFacto
ryImpl.java:237)
at org.datanucleus.store.rdbms.ConnectionFactoryImpl.initialiseDataSources(ConnectionFact
oryImpl.java:110)
at org.datanucleus.store.rdbms.ConnectionFactoryImpl.</init><init>(ConnectionFactoryImpl.j
ava:82)
... 61 more
Caused by: org.datanucleus.store.rdbms.datasource.DatastoreDriverNotFoundException: The
specified datastore driver ("com.mysql.jdbc.Driver") was not found in the CLASSPATH. Please c
heck your CLASSPATH specification, and the name of the driver.
at org.datanucleus.store.rdbms.datasource.AbstractDataSourceFactory.loadDriver(AbstractDa
taSourceFactory.java:58)
at org.datanucleus.store.rdbms.datasource.DBCPDataSourceFactory.makePooledDataSource(DBC
PDataSourceFactory.java:55)
at org.datanucleus.store.rdbms.ConnectionFactoryImpl.generateDataSources(ConnectionFacto
ryImpl.java:217)
```


... 63 more

很明显是缺少了mysql驱动。解决办法是：到一个地方拷贝mysql-connector-java-5.1.22-bin.jar到\$SHARK_HOME/lib_managed/jars/edu.berkeley.cs.shark/hive-jdbc路径下。

五、Cannot create an instance of InputFormat class org.apache.hadoop.mapred.TextInputFormat as specified in mapredWork!

14/04/29 11:57:40 ERROR CliDriver: Failed with exception java.io.IOException:java.io.IOException: Cannot create an instance of InputFormat class org.apache.hadoop.mapred.TextInputFormat as specified in mapredWork!

java.io.IOException: java.io.IOException: Cannot create an instance of InputFormat class org.apache.hadoop.mapred.TextInputFormat as specified in mapredWork!

at org.apache.hadoop.hive ql.exec.FetchOperator.getNextRow(FetchOperator.java:544)

at org.apache.hadoop.hive ql.exec.FetchOperator.pushRow(FetchOperator.java:488)

at org.apache.hadoop.hive ql.exec.FetchTask.fetch(FetchTask.java:136)

at org.apache.hadoop.hive ql.Driver.getResults(Driver.java:1419)

at shark.SharkCliDriver.processCmd(SharkCliDriver.scala:355)

at org.apache.hadoop.hive.cli.CliDriver.processLine(CliDriver.java:413)

at shark.SharkCliDriver\$.main(SharkCliDriver.scala:235)

at shark.SharkCliDriver.main(SharkCliDriver.scala)

Caused by: java.io.IOException: Cannot create an instance of InputFormat class org.apache.hadoop.mapred.TextInputFormat as specified in mapredWork!

at org.apache.hadoop.hive ql.exec.FetchOperator.getInputFormatFromCache(FetchOperator.java:222)

at org.apache.hadoop.hive ql.exec.FetchOperator.getRecordReader(FetchOperator.java:378)

at org.apache.hadoop.hive ql.exec.FetchOperator.getNextRow(FetchOperator.java:508)

... 7 more

Caused by: java.lang.RuntimeException: Error in configuring object

at org.apache.hadoop.util.ReflectionUtils.setJobConf(ReflectionUtils.java:109)

at org.apache.hadoop.util.ReflectionUtils.setConf(ReflectionUtils.java:75)

at org.apache.hadoop.util.ReflectionUtils.newInstance(ReflectionUtils.java:133)

at org.apache.hadoop.hive ql.exec.FetchOperator.getInputFormatFromCache(FetchOperator.java:218)

... 9 more

Caused by: java.lang.reflect.InvocationTargetException

at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)

at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)

at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)

at java.lang.reflect.Method.invoke(Method.java:606)

at org.apache.hadoop.util.ReflectionUtils.setJobConf(ReflectionUtils.java:106)

... 12 more

Caused by: java.lang.IllegalArgumentException: Compression codec com.hadoop.compression

```
.lzo.LzoCodec not found.  
at org.apache.hadoop.io.compress.CompressionCodecFactory.getCodecClasses(Compression  
CodecFactory.java:135)  
at org.apache.hadoop.io.compress.CompressionCodecFactory.<init>(CompressionCodecFacto  
ry.java:175)  
at org.apache.hadoop.mapred.TextInputFormat.configure(TextInputFormat.java:45)  
... 17 more  
Caused by: java.lang.ClassNotFoundException: Class com.hadoop.compression.lzo.LzoCodec n  
ot found  
at org.apache.hadoop.conf.Configuration.getClassByName(Configuration.java:1801)  
at org.apache.hadoop.io.compress.CompressionCodecFactory.getCodecClasses(Compression  
CodecFactory.java:128)  
... 19 more
```

解决方法是去拷贝一个hadoop-lzo-0.4.20-SNAPSHOT.jar到\$SHARK_HOME/lib目录下即可。

六、readObject can't find class org.apache.hadoop.hive.conf.HiveConf

```
java.lang.RuntimeException: readObject can't find class org.apache.hadoop.hive.conf.HiveConf  
at org.apache.hadoop.io.ObjectWritable.loadClass(ObjectWritable.java:377)  
at org.apache.hadoop.io.ObjectWritable.readObject(ObjectWritable.java:228)  
at org.apache.hadoop.io.ObjectWritable.readFields(ObjectWritable.java:77)  
at org.apache.spark.SerializableWritable.readObject(SerializableWritable.scala:39)  
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)  
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)  
at java.lang.reflect.Method.invoke(Method.java:606)  
at java.io.ObjectStreamClass.invokeReadObject(ObjectStreamClass.java:1017)  
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1893)  
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)  
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)  
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)  
at org.apache.spark.serializer.JavaDeserializationStream.readObject(JavaSerializer.scala:40)  
at org.apache.spark.broadcast.HttpBroadcast$.read(HttpBroadcast.scala:165)  
at org.apache.spark.broadcast.HttpBroadcast.readObject(HttpBroadcast.scala:56)  
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)  
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)  
at java.lang.reflect.Method.invoke(Method.java:606)  
at java.io.ObjectStreamClass.invokeReadObject(ObjectStreamClass.java:1017)  
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1893)  
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)  
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
```

```
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)
at scala.collection.immutable.$colon$colon.readObject(List.scala:362)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at java.io.ObjectStreamClass.invokeReadObject(ObjectStreamClass.java:1017)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1893)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)
at scala.collection.immutable.$colon$colon.readObject(List.scala:362)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at java.io.ObjectStreamClass.invokeReadObject(ObjectStreamClass.java:1017)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1893)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)
at scala.collection.immutable.$colon$colon.readObject(List.scala:362)
```

```
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:606)
at java.io.ObjectStreamClass.invokeReadObject(ObjectStreamClass.java:1017)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1893)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.defaultReadFields(ObjectInputStream.java:1990)
at java.io.ObjectInputStream.readSerialData(ObjectInputStream.java:1915)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1798)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)
at org.apache.spark.serializer.JavaDeserializationStream.readObject(JavaSerializer.scala:40)
at org.apache.spark.scheduler.ResultTask$.deserializeInfo(ResultTask.scala:63)
at org.apache.spark.scheduler.ResultTask.readExternal(ResultTask.scala:139)
at java.io.ObjectInputStream.readExternalData(ObjectInputStream.java:1837)
at java.io.ObjectInputStream.readOrdinaryObject(ObjectInputStream.java:1796)
at java.io.ObjectInputStream.readObject0(ObjectInputStream.java:1350)
at java.io.ObjectInputStream.readObject(ObjectInputStream.java:370)
at org.apache.spark.serializer.JavaDeserializationStream.readObject(JavaSerializer.scala:40)
at org.apache.spark.serializer.JavaSerializerInstance.deserialize(JavaSerializer.scala:62)
at org.apache.spark.executor.Executor$TaskRunner$$anonfun$run$1.apply$mcV$sp(Executor.scala:193)
at org.apache.spark.deploy.SparkHadoopUtil$$anon$1.run(SparkHadoopUtil.scala:42)
at org.apache.spark.deploy.SparkHadoopUtil$$anon$1.run(SparkHadoopUtil.scala:41)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:415)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1491)
)
at org.apache.spark.deploy.SparkHadoopUtil.runAsUser(SparkHadoopUtil.scala:41)
at org.apache.spark.executor.Executor$TaskRunner.run(Executor.scala:176)
at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1145)
at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:615)
at java.lang.Thread.run(Thread.java:744)</init>
```

解决办法是在每个子节点上面都部署好相应的shark、spark、hive等，并设置好环境变量。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接：[【】（）](#)