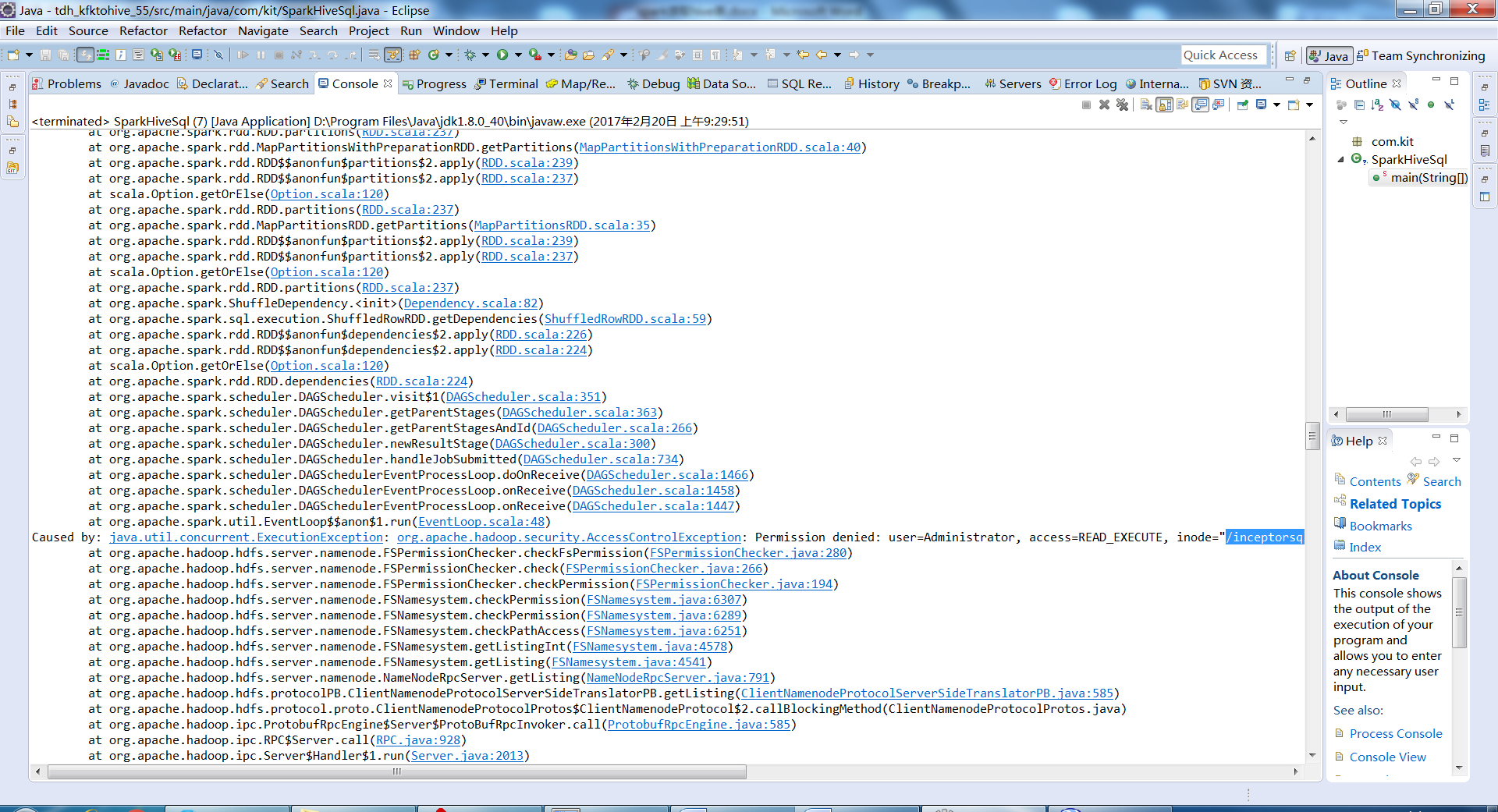
# Spark 读取hive表

Tips: 注意jar版本问题

修改权限：hdfs文件

sudo -u hdfs hadoop fs -chmod -R 777 /inceptorsql1/user/hive/warehouse/default.db/hive/dim\_kpi\_organization



<http://blog.csdn.net/seven_zhao/article/details/44178185>



[**Hive**](http://lib.csdn.net/base/hive)启动时报错，报错信息如下：

Logging initialized using configuration in jar:file:/usr/local/opt/hive-0.14.0/lib/hive-common-0.14.0.jar!/hive-log4j.properties  
Exception in thread "main" [**Java**](http://lib.csdn.net/base/javase).lang.NoSuchMethodError: org.apache.[**Hadoop**](http://lib.csdn.net/base/hadoop).conf.Configuration.unset(Ljava/lang/String;)V  
    at org.apache.hadoop.hive.ql.exec.Utilities.resetUmaskInConf(Utilities.java:3561)  
    at org.apache.hadoop.hive.ql.exec.Utilities.createDirsWithPermission(Utilities.java:3546)  
    at org.apache.hadoop.hive.ql.session.SessionState.createRootHDFSDir(SessionState.java:521)  
    at org.apache.hadoop.hive.ql.session.SessionState.createSessionDirs(SessionState.java:478)  
    at org.apache.hadoop.hive.ql.session.SessionState.start(SessionState.java:430)  
    at org.apache.hadoop.hive.cli.CliDriver.run(CliDriver.java:672)  
    at org.apache.hadoop.hive.cli.CliDriver.main(CliDriver.java:616)  
    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.RunJar.main(RunJar.java:156)  
**第一种解决方案：**  
这个问题是由于hive版本较高，hadoop版本较低，不相匹配造成的，解决方法时降低hive的版本，比如降低到hive-0.9.0。

**第二种解决方案：**

1、在hive-site.xml中添加如下配置：

<property>  
    <name>fs.permissions.umask-mode</name>  
    <value>022</value>  
    <description> Setting a value for fs.permissions.umask-mode to work around issue in HIVE-6962.  
 It has no impact in hadoop 1.x line on hdfs operations.  
    </description>  
</property>

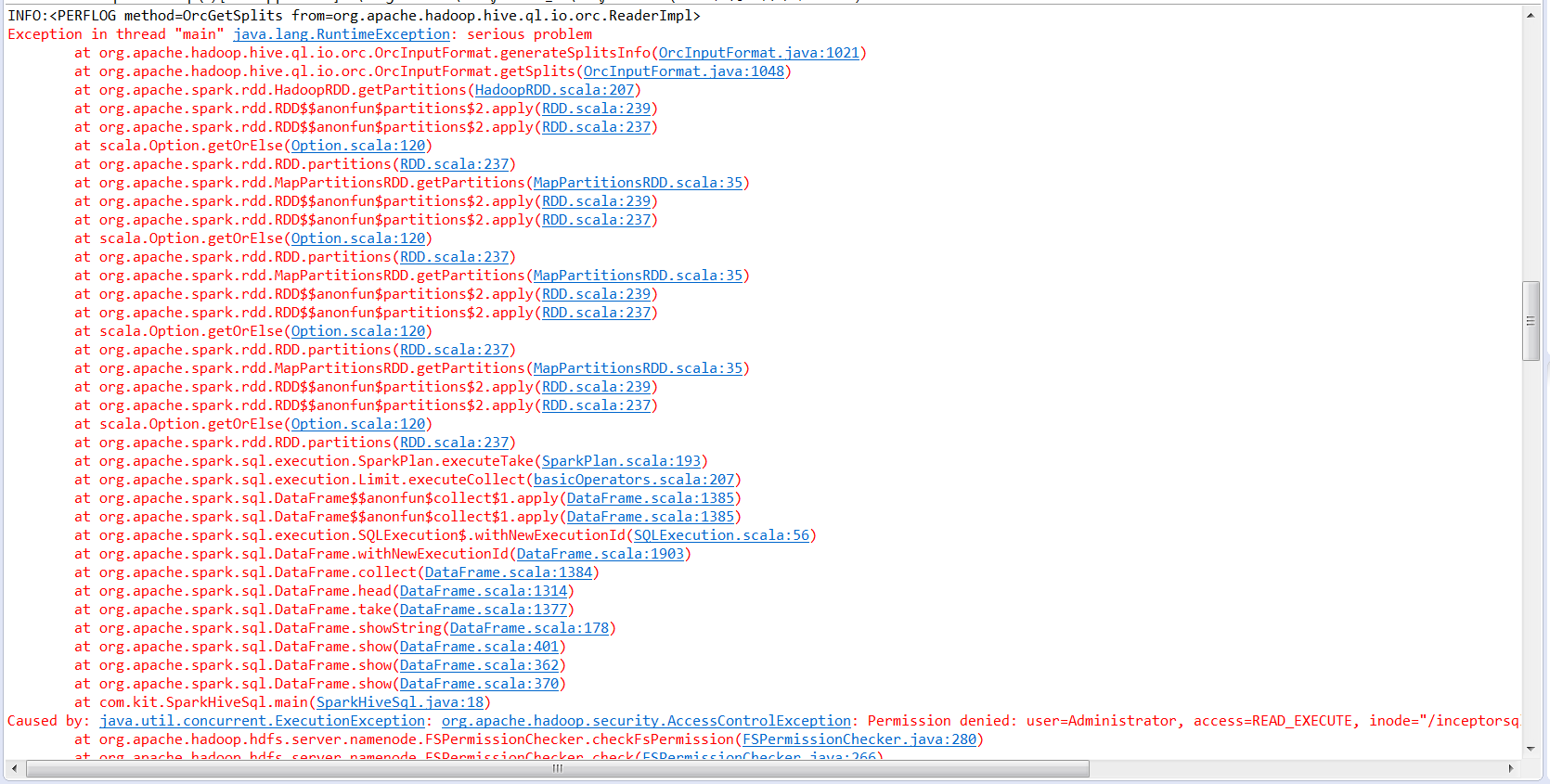
2、修改hadoop hdfs tmp文件夹的权限

hadoop fs -chmod -R 777 /tmp

3、将hive-site.xml中所有的配置值含有"system:java.io.tmpdir"的配置项替换为绝对路径

  在hive-0.14.0下创建临时文件目录hive\_tmp：mkdir hive\_tmp

将system开始的路径替换为/usr/local/opt/hive-0.14.0/hive\_tmp



Exception in thread "main" java.lang.RuntimeException: serious problem

at org.apache.hadoop.hive.ql.io.orc.OrcInputFormat.generateSplitsInfo(OrcInputFormat.java:1021)

at org.apache.hadoop.hive.ql.io.orc.OrcInputFormat.getSplits(OrcInputFormat.java:1048)

at org.apache.spark.rdd.HadoopRDD.getPartitions(HadoopRDD.scala:207)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:239)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:237)

at scala.Option.getOrElse(Option.scala:120)

at org.apache.spark.rdd.RDD.partitions(RDD.scala:237)

at org.apache.spark.rdd.MapPartitionsRDD.getPartitions(MapPartitionsRDD.scala:35)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:239)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:237)

at scala.Option.getOrElse(Option.scala:120)

at org.apache.spark.rdd.RDD.partitions(RDD.scala:237)

at org.apache.spark.rdd.MapPartitionsRDD.getPartitions(MapPartitionsRDD.scala:35)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:239)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:237)

at scala.Option.getOrElse(Option.scala:120)

at org.apache.spark.rdd.RDD.partitions(RDD.scala:237)

at org.apache.spark.rdd.MapPartitionsRDD.getPartitions(MapPartitionsRDD.scala:35)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:239)

at org.apache.spark.rdd.RDD$$anonfun$partitions$2.apply(RDD.scala:237)

at scala.Option.getOrElse(Option.scala:120)

at org.apache.spark.rdd.RDD.partitions(RDD.scala:237)

at org.apache.spark.sql.execution.SparkPlan.executeTake(SparkPlan.scala:193)

at org.apache.spark.sql.execution.Limit.executeCollect(basicOperators.scala:207)

at org.apache.spark.sql.DataFrame$$anonfun$collect$1.apply(DataFrame.scala:1385)

at org.apache.spark.sql.DataFrame$$anonfun$collect$1.apply(DataFrame.scala:1385)

at org.apache.spark.sql.execution.SQLExecution$.withNewExecutionId(SQLExecution.scala:56)

at org.apache.spark.sql.DataFrame.withNewExecutionId(DataFrame.scala:1903)

at org.apache.spark.sql.DataFrame.collect(DataFrame.scala:1384)

at org.apache.spark.sql.DataFrame.head(DataFrame.scala:1314)

at org.apache.spark.sql.DataFrame.take(DataFrame.scala:1377)

at org.apache.spark.sql.DataFrame.showString(DataFrame.scala:178)

at org.apache.spark.sql.DataFrame.show(DataFrame.scala:401)

at org.apache.spark.sql.DataFrame.show(DataFrame.scala:362)

at org.apache.spark.sql.DataFrame.show(DataFrame.scala:370)

at com.kit.SparkHiveSql.main(SparkHiveSql.java:18)

Caused by: java.util.concurrent.ExecutionException: org.apache.hadoop.security.AccessControlException: Permission denied: user=Administrator, access=READ\_EXECUTE, inode="/inceptorsql1/user/hive/warehouse/kitdev.db/hive/test\_orc":hive:hadoop:drwx--x--x

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkFsPermission(FSPermissionChecker.java:280)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:266)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:194)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6307)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6289)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPathAccess(FSNamesystem.java:6251)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListingInt(FSNamesystem.java:4578)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListing(FSNamesystem.java:4541)

at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getListing(NameNodeRpcServer.java:791)

at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getListing(ClientNamenodeProtocolServerSideTranslatorPB.java:585)

at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)

at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:585)

at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:928)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2013)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2009)

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:1854)

at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2007)

at java.util.concurrent.FutureTask.report(FutureTask.java:122)

at java.util.concurrent.FutureTask.get(FutureTask.java:192)

at org.apache.hadoop.hive.ql.io.orc.OrcInputFormat.generateSplitsInfo(OrcInputFormat.java:998)

... 35 more

Caused by: org.apache.hadoop.security.AccessControlException: Permission denied: user=Administrator, access=READ\_EXECUTE, inode="/inceptorsql1/user/hive/warehouse/kitdev.db/hive/test\_orc":hive:hadoop:drwx--x--x

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkFsPermission(FSPermissionChecker.java:280)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:266)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:194)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6307)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6289)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPathAccess(FSNamesystem.java:6251)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListingInt(FSNamesystem.java:4578)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListing(FSNamesystem.java:4541)

at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getListing(NameNodeRpcServer.java:791)

at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getListing(ClientNamenodeProtocolServerSideTranslatorPB.java:585)

at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)

at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:585)

at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:928)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2013)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2009)

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:1854)

at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2007)

at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)

at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)

at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)

at java.lang.reflect.Constructor.newInstance(Constructor.java:422)

at org.apache.hadoop.ipc.RemoteException.instantiateException(RemoteException.java:106)

at org.apache.hadoop.ipc.RemoteException.unwrapRemoteException(RemoteException.java:73)

at org.apache.hadoop.hdfs.DFSClient.listPaths(DFSClient.java:1728)

at org.apache.hadoop.hdfs.DistributedFileSystem$15.<init>(DistributedFileSystem.java:731)

at org.apache.hadoop.hdfs.DistributedFileSystem.listLocatedStatus(DistributedFileSystem.java:720)

at org.apache.hadoop.fs.FileSystem.listLocatedStatus(FileSystem.java:1662)

at org.apache.hadoop.hive.shims.Hadoop23Shims.listLocatedStatus(Hadoop23Shims.java:667)

at org.apache.hadoop.hive.ql.io.AcidUtils.getAcidState(AcidUtils.java:361)

at org.apache.hadoop.hive.ql.io.orc.OrcInputFormat$FileGenerator.call(OrcInputFormat.java:634)

at org.apache.hadoop.hive.ql.io.orc.OrcInputFormat$FileGenerator.call(OrcInputFormat.java:620)

at java.util.concurrent.FutureTask.run(FutureTask.java:266)

at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1142)

at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:617)

at java.lang.Thread.run(Thread.java:745)

Caused by: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessControlException): Permission denied: user=Administrator, access=READ\_EXECUTE, inode="/inceptorsql1/user/hive/warehouse/kitdev.db/hive/test\_orc":hive:hadoop:drwx--x--x

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkFsPermission(FSPermissionChecker.java:280)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:266)

at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:194)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6307)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPermission(FSNamesystem.java:6289)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkPathAccess(FSNamesystem.java:6251)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListingInt(FSNamesystem.java:4578)

at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getListing(FSNamesystem.java:4541)

at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getListing(NameNodeRpcServer.java:791)

at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getListing(ClientNamenodeProtocolServerSideTranslatorPB.java:585)

at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)

at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:585)

at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:928)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2013)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2009)

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:1854)

at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2007)

at org.apache.hadoop.ipc.Client.call(Client.java:1420)

at org.apache.hadoop.ipc.Client.call(Client.java:1373)

at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:206)

at com.sun.proxy.$Proxy21.getListing(Unknown Source)

at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.getListing(ClientNamenodeProtocolTranslatorPB.java:523)

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

at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)

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

at java.lang.reflect.Method.invoke(Method.java:497)

at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:187)

at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:102)

at com.sun.proxy.$Proxy22.getListing(Unknown Source)

at org.apache.hadoop.hdfs.DFSClient.listPaths(DFSClient.java:1726)

... 11 more