

Hi , I am a newbie into Talend development environment. I have installed **Talend open studio for Big data** in Windows 8 OS and trying to connect to **Cloudera CDH V5.4.3(non commercial)** using VMWare Player. I have followed the Talend getting started manual for Big data HDFS connectivity and able to connect to the CDH environment(tried both NAT and Host Private type of Network Adapters). But when I am trying to put a file from my local to the HDFS I am getting an error

**org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /home/training/states\_demo5mn.txt could only be replicated to 0 nodes instead of minReplication (=1). There are 1 datanode(s) running and 1 node(s) are excluded in this operation.**

As I know its a vary common error with Hadoop ecosystem and there may be many reasons behind this issue, all my possible efforts till now are in vain. I have checked the following

- 1) I am able to ping and telnet the Name Node and Data Node from my Windows host.
- 2) In my VM , the Data Node is up and running and the other statistics are also proper (check screen shot)
- 3) I have given write permission to the folder /var/lib/hadoop-hdfs/cache in my VM
- 4) I have used the IP address of my VM in place of localhost/127.0.0.1 in **core-site.xml(fs.defaultFS)** and **hdfs-site.xml(dfs.datanode.address)** and in **hosts file of my VM.**
- 5) Also , I have used the same IP while setting up the Talend connection with HDFS.

I am not very sure where the actual issue is, is it due some **PortForwarding** issue while connecting through HDFS, or whether there is any set up related issue in my VM's hadoop env or any privilege/ authentication related issue with my user **training**.

NB : I have used this VM in many of my hadoop related works earlier but never faced any such issue before. Also in this case I can see that the file is getting created in HDFS destination but with 0 byte in size!

Here is the log from NameNode log file :

**Starting job write\_to\_hdfs at 18:06 25/08/2016.**

**connecting to socket on port 3742**

**connected**

**: org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable**

**: org.apache.hadoop.util.Shell - Failed to locate the winutils binary in the hadoop binary path**

**java.io.IOException: Could not locate executable null\bin\winutils.exe in the Hadoop binaries.**

**at org.apache.hadoop.util.Shell.getQualifiedBinPath(Shell.java:355)**

**at org.apache.hadoop.util.Shell.getWinUtilsPath(Shell.java:370)**

**at org.apache.hadoop.util.Shell.<clinit>(Shell.java:363)**

**at org.apache.hadoop.util.StringUtils.<clinit>(StringUtils.java:79)**

**at org.apache.hadoop.security.Groups.parseStaticMapping(Groups.java:104)**

**at org.apache.hadoop.security.Groups.<init>(Groups.java:86)**

**at org.apache.hadoop.security.Groups.<init>(Groups.java:66)**

**at org.apache.hadoop.security.Groups.getUserToGroupsMappingService(Groups.java:280)**

**at org.apache.hadoop.security.UserGroupInformation.initialize(UserGroupInformation.java:283)**

**at org.apache.hadoop.security.UserGroupInformation.setConfiguration(UserGroupInformation.java:311)**

**at talend\_demo\_project.write\_to\_hdfs\_0\_1.write\_to\_hdfs.tHDFSput\_1Process(write\_to\_hdfs.java:303)**

**at talend\_demo\_project.write\_to\_hdfs\_0\_1.write\_to\_hdfs.runJobInTOS(write\_to\_hdfs.java:685)**

**at talend\_demo\_project.write\_to\_hdfs\_0\_1.write\_to\_hdfs.main(write\_to\_hdfs.java:542)**

**Exception in component tHDFSput\_1**

**org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /home/training/states\_demo5mn.txt could only be replicated to 0 nodes instead of minReplication (=1). There are 1 datanode(s) running and 1 node(s) are excluded in this operation.**

**at org.apache.hadoop.hdfs.server.blockmanagement.BlockManager.chooseTarget4NewBlock(BlockManager.java:1541)**

**at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:3289)**

**at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.addBlock(NameNodeRpcServer.java:668)**

**at org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.addBlock(AuthorizationProviderProxyClientProtocol.java:212)**

**at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.addBlock(ClientNamenodeProtocolServerSideTranslatorPB.java:483)**

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

**: org.apache.hadoop.hdfs.DFSClient - DataStreamer Exception**

**org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /home/training/states\_demo5mn.txt could only be replicated to 0 nodes instead of minReplication (=1). There are 1 datanode(s) running and 1 node(s) are excluded in this operation.**

**at org.apache.hadoop.hdfs.server.blockmanagement.BlockManager.chooseTarget4NewBlock(BlockManager.java:1541)**

**at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.getAdditionalBlock(FSNamesystem.java:3289)**

**at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.addBlock(NameNodeRpcServer.java:668)**

**at org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.addBlock(AuthorizationProviderProxyClientProtocol.java:212)**

at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.addBlock(ClientNamenodeProtocolServerSideTranslatorPB.java:483)  
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos\$ClientNamenodeProtocol\$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)  
at org.apache.hadoop.ipc.ProtobufRpcEngine\$Server\$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:619)  
at org.apache.hadoop.ipc.RPC\$Server.call(RPC.java:1060)  
at org.apache.hadoop.ipc.Server\$Handler\$1.run(Server.java:2044)  
at org.apache.hadoop.ipc.Server\$Handler\$1.run(Server.java:2040)  
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:1671)  
at org.apache.hadoop.ipc.Server\$Handler.run(Server.java:2038)

at org.apache.hadoop.ipc.Client.call(Client.java:1468)  
at org.apache.hadoop.ipc.Client.call(Client.java:1399)  
at org.apache.hadoop.ipc.ProtobufRpcEngine\$Invoker.invoke(ProtobufRpcEngine.java:232)  
at com.sun.proxy.\$Proxy7.addBlock(Unknown Source)  
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.addBlock(ClientNamenodeProtocolTranslatorPB.java:399)  
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
at sun.reflect.NativeMethodAccessorImpl.invoke(Unknown Source)  
at org.apache.hadoop.ipc.ProtobufRpcEngine\$Server\$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:619)  
at org.apache.hadoop.ipc.RPC\$Server.call(RPC.java:1060)  
at org.apache.hadoop.ipc.Server\$Handler\$1.run(Server.java:2044)  
at org.apache.hadoop.ipc.Server\$Handler\$1.run(Server.java:2040)  
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:1671)  
at org.apache.hadoop.ipc.Server\$Handler.run(Server.java:2038)

at org.apache.hadoop.ipc.Client.call(Client.java:1468)  
at org.apache.hadoop.ipc.Client.call(Client.java:1399)  
at org.apache.hadoop.ipc.ProtobufRpcEngine\$Invoker.invoke(ProtobufRpcEngine.java:232)  
at com.sun.proxy.\$Proxy7.addBlock(Unknown Source)  
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.addBlock(ClientNamenodeProtocolTranslatorPB.java:399)  
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)  
at sun.reflect.NativeMethodAccessorImpl.invoke(Unknown Source)  
at sun.reflect.DelegatingMethodAccessorImpl.invoke(Unknown Source)  
at java.lang.reflect.Method.invoke(Unknown Source)  
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.\$Proxy8.addBlock(Unknown Source)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.locateFollowingBlock(DFSOutputStream.java:1544)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.nextBlockOutputStream(DFSOutputStream.java:1361)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.run(DFSOutputStream.java:600)  
at sun.reflect.DelegatingMethodAccessorImpl.invoke(Unknown Source)  
at java.lang.reflect.Method.invoke(Unknown Source)  
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.\$Proxy8.addBlock(Unknown Source)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.locateFollowingBlock(DFSOutputStream.java:1544)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.nextBlockOutputStream(DFSOutputStream.java:1361)  
at org.apache.hadoop.hdfs.DFSOutputStream\$DataStreamer.run(DFSOutputStream.java:600)

disconnected

Job write\_to\_hdfs ended at 18:06 25/08/2016.

