

Shong,  
Followed your suggestion, we did the read test. Got error too. Any suggestions?  
*Starting job HDFSInput at 13:59 25/05/2016.*

connecting to socket on port 3744  
connected

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

: org.apache.hadoop.hdfs.BlockReaderFactory - I/O error constructing remote block reader.

java.net.ConnectException: Connection refused: no further information

at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)

at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)

at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)

at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)

at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)

at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)

at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)

at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)

at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)

at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)

at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)

at java.io.DataInputStream.read(DataInputStream.java:149)

at java.io.FilterInputStream.read(FilterInputStream.java:133)

at java.io.PushbackInputStream.read(PushbackInputStream.java:186)

at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)

at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)

at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)

at local\_project.hdfsinput\_0\_1.HDFSInput.tHDFSInput\_1Process(HDFSInput.java:541)

at local\_project.hdfsinput\_0\_1.HDFSInput.runJobInTOS(HDFSInput.java:903)

at local\_project.hdfsinput\_0\_1.HDFSInput.main(HDFSInput.java:760)

: org.apache.hadoop.hdfs.DFSCClient - Failed to connect to sandbox.hortonworks.com/10.0.0.4:50010 for block, add to deadNodes and continue. java.net.ConnectException: Connection refused: no further information

java.net.ConnectException: Connection refused: no further information

at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)

at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)

at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)

at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)

at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)

at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)

at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)

at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)

at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)

at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)

at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)

at java.io.DataInputStream.read(DataInputStream.java:149)

at java.io.FilterInputStream.read(FilterInputStream.java:133)

at java.io.PushbackInputStream.read(PushbackInputStream.java:186)

at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)

at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)

at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)

at local\_project.hdfsinput\_0\_1.HDFSInput.tHDFSInput\_1Process(HDFSInput.java:541)

at local\_project.hdfsinput\_0\_1.HDFSInput.runJobInTOS(HDFSInput.java:903)

at local\_project.hdfsinput\_0\_1.HDFSInput.main(HDFSInput.java:760)

: org.apache.hadoop.hdfs.DFSCClient - DFS chooseDataNode: got # 1 IOException, will wait for 1439.7276636017177 msec.

```
: org.apache.hadoop.hdfs.BlockReaderFactory - I/O error constructing remote block reader.
java.net.ConnectException: Connection refused: no further information
    at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)
    at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
    at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
    at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
    at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
    at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
    at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
    at java.io.DataInputStream.read(DataInputStream.java:149)
    at java.io.FilterInputStream.read(FilterInputStream.java:133)
    at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
    at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
    at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
    at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
    at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
    at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
    at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSCClient - Failed to connect to sandbox.hortonworks.com/10.0.0.4:50010 for block, add to deadNodes and continue. java.net.ConnectException: Connection
refused: no further information
java.net.ConnectException: Connection refused: no further information
    at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)
    at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
    at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
    at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
    at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
    at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
    at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
    at java.io.DataInputStream.read(DataInputStream.java:149)
    at java.io.FilterInputStream.read(FilterInputStream.java:133)
    at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
    at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
    at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
    at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
    at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
    at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
    at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSCClient - DFS chooseDataNode: got # 2 IOException, will wait for 4769.762006394212 msec.
: org.apache.hadoop.hdfs.BlockReaderFactory - I/O error constructing remote block reader.
java.net.ConnectException: Connection refused: no further information
    at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
    at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
    at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
    at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
    at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)
```

```
at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSClient - Failed to connect to sandbox.hortonworks.com/10.0.0.4:50010 for block, add to deadNodes and continue. java.net.ConnectException: Connection
refused: no further information
java.net.ConnectException: Connection refused: no further information
at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
at org.apache.hadoop.hdfs.DFSClient.newConnectedPeer(DFSClient.java:3444)
at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSClient - DFS chooseDataNode: got # 3 IOException, will wait for 11059.781468103138 msec.
: org.apache.hadoop.hdfs.BlockReaderFactory - I/O error constructing remote block reader.
java.net.ConnectException: Connection refused: no further information
at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
at org.apache.hadoop.hdfs.DFSClient.newConnectedPeer(DFSClient.java:3444)
at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
```

```
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSCClient - Failed to connect to sandbox.hortonworks.com/10.0.0.4:50010 for block, add to deadNodes and continue. java.net.ConnectException: Connection
refused: no further information
java.net.ConnectException: Connection refused: no further information
at sun.nio.ch.SocketChannellImpl.checkConnect(Native Method)
at sun.nio.ch.SocketChannellImpl.finishConnect(SocketChannellImpl.java:735)
at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)
at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:531)
at org.apache.hadoop.hdfs.DFSCClient.newConnectedPeer(DFSCClient.java:3444)
at org.apache.hadoop.hdfs.BlockReaderFactory.nextTcpPeer(BlockReaderFactory.java:777)
at org.apache.hadoop.hdfs.BlockReaderFactory.getRemoteBlockReaderFromTcp(BlockReaderFactory.java:694)
at org.apache.hadoop.hdfs.BlockReaderFactory.build(BlockReaderFactory.java:355)
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:618)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
: org.apache.hadoop.hdfs.DFSCClient - Could not obtain block: BP-595454498-172.16.137.143-1456768655900:blk_1073742615_1799 file=/tmp/maria_dev/data/geolocation.csv No live nodes
contain current block Block locations: DatanodeInfoWithStorage Dead nodes: DatanodeInfoWithStorage. Throwing a BlockMissingException
: org.apache.hadoop.hdfs.DFSCClient - Could not obtain block: BP-595454498-172.16.137.143-1456768655900:blk_1073742615_1799 file=/tmp/maria_dev/data/geolocation.csv No live nodes
contain current block Block locations: DatanodeInfoWithStorage Dead nodes: DatanodeInfoWithStorage. Throwing a BlockMissingException
: org.apache.hadoop.hdfs.DFSCClient - DFS Read
org.apache.hadoop.hdfs.BlockMissingException: Could not obtain block: BP-595454498-172.16.137.143-1456768655900:blk_1073742615_1799 file=/tmp/maria_dev/data/geolocation.csv
at org.apache.hadoop.hdfs.DFSInputStream.chooseDataNode(DFSInputStream.java:945)
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:604)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
Exception in component tHDFSInput_1
org.apache.hadoop.hdfs.BlockMissingException: Could not obtain block: BP-595454498-172.16.137.143-1456768655900:blk_1073742615_1799 file=/tmp/maria_dev/data/geolocation.csv
at org.apache.hadoop.hdfs.DFSInputStream.chooseDataNode(DFSInputStream.java:945)
```

```
at org.apache.hadoop.hdfs.DFSInputStream.blockSeekTo(DFSInputStream.java:604)
at org.apache.hadoop.hdfs.DFSInputStream.readWithStrategy(DFSInputStream.java:844)
at org.apache.hadoop.hdfs.DFSInputStream.read(DFSInputStream.java:896)
at java.io.DataInputStream.read(DataInputStream.java:149)
at java.io.FilterInputStream.read(FilterInputStream.java:133)
at java.io.PushbackInputStream.read(PushbackInputStream.java:186)
at org.talend.fileprocess.UnicodeReader.<init>(UnicodeReader.java:25)
at org.talend.fileprocess.TOSDelimitedReader.<init>(TOSDelimitedReader.java:77)
at org.talend.fileprocess.FileInputDelimited.<init>(FileInputDelimited.java:93)
at local_project.hdfsinput_0_1.HDFSInput.tHDFSInput_1Process(HDFSInput.java:541)
disconnected
  at local_project.hdfsinput_0_1.HDFSInput.runJobInTOS(HDFSInput.java:903)
  at local_project.hdfsinput_0_1.HDFSInput.main(HDFSInput.java:760)
Job HDFSInput ended at 13:59 25/05/2016.
```