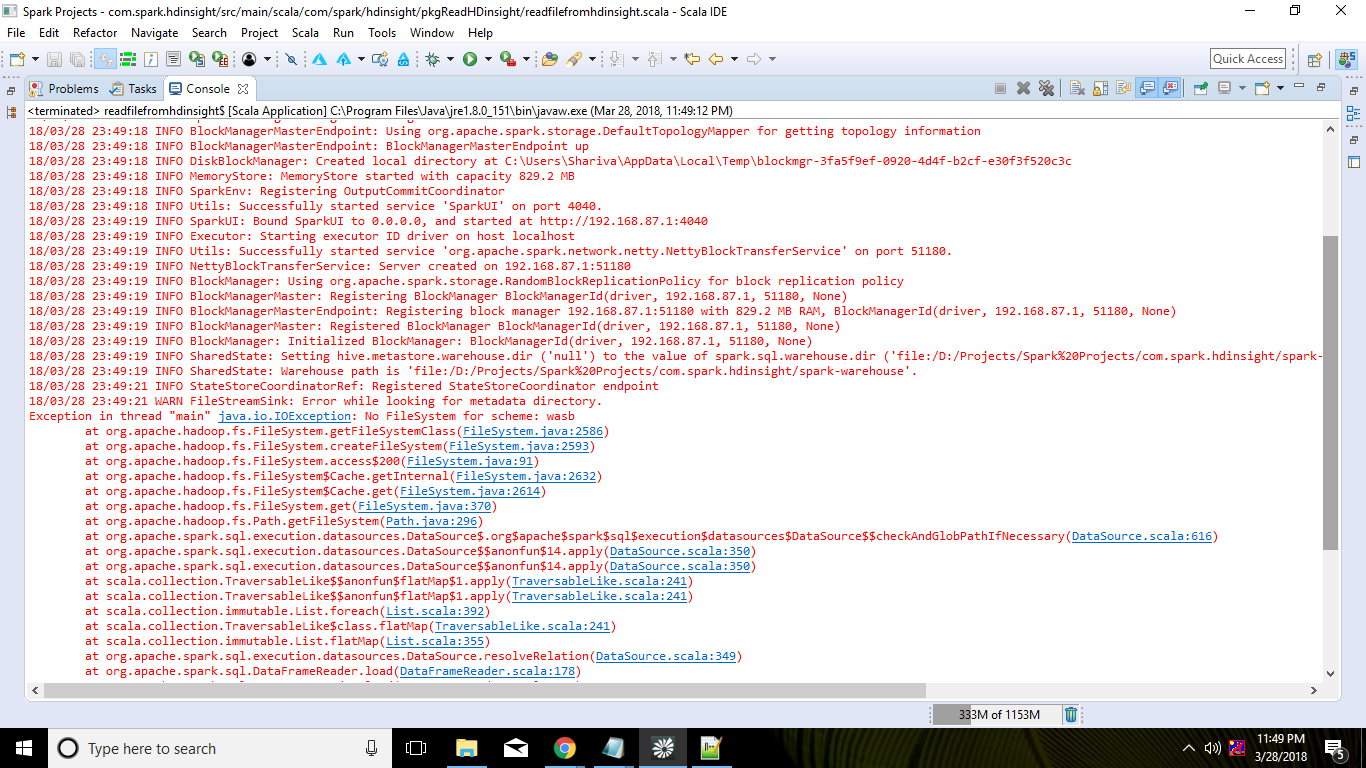
**How do I connect to HDInsight cluster from eclipse and access the file**

1. Create HDInsight Cluster for spark
2. Create Storage
3. In our case below is cluster and storage details
   1. Storage : myhdstorage1
   2. Cluster: hdcluster
4. Upload some input file to blob storage under **Example/Data** folder.
5. You will get the connection path as below

|  |
| --- |
| wasb:[//hdcluster-2018-03-28t17-42-47-461z@myhdstorage1.blob.core.windows.net](about:blank)/  Example/data/customer.csv |

Where wasb is the azure file system, followed by container name, followed by storage name.

1. Write a spark program to read a file from the cluster using above file path.
2. Run the program. You should see below error



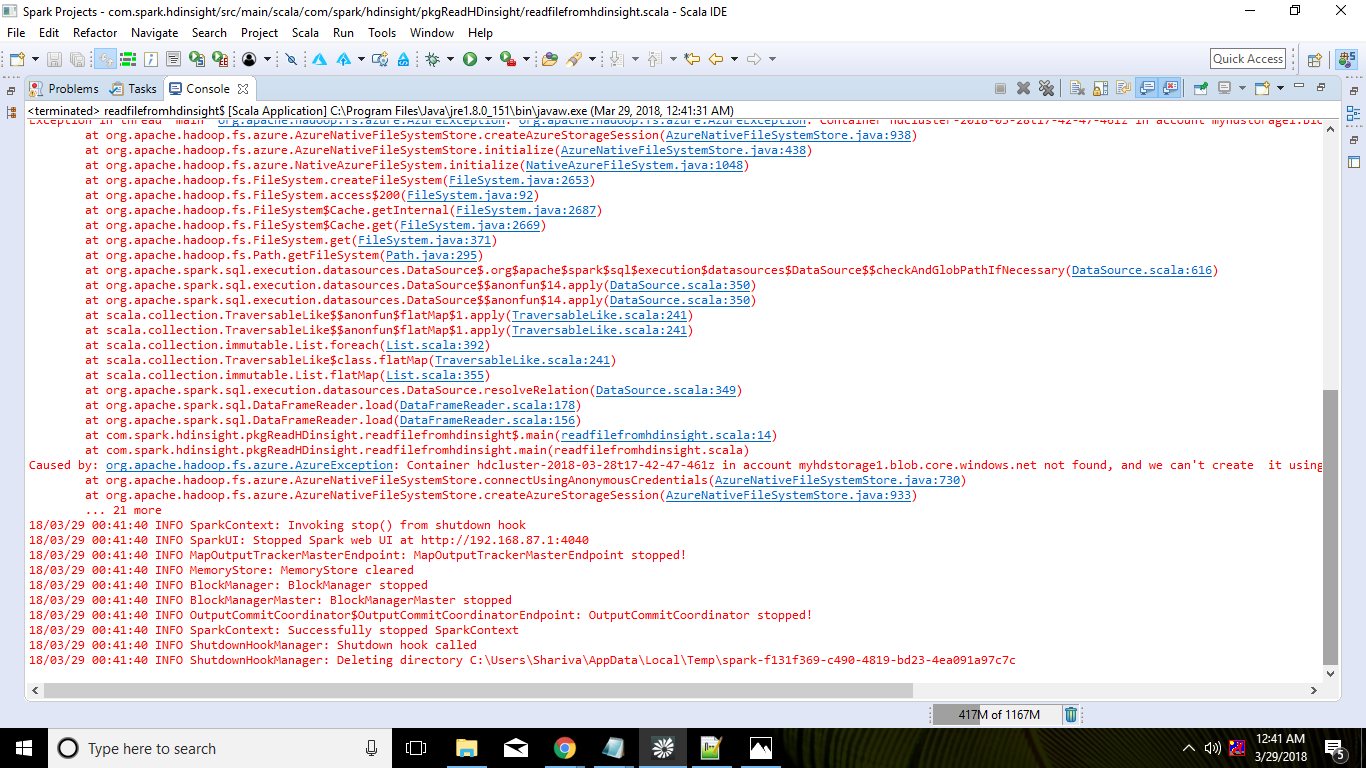
1. Above error is related to file system i.e. unable to recognize the wasb file system.
2. In order to fix this error we need to add related dependencies in pom.xml file

|  |
| --- |
| <dependency>  <groupId>org.apache.hadoop</groupId>  <artifactId>hadoop-hdfs</artifactId>  <version>2.7.5</version>  </dependency> |

|  |
| --- |
| <dependency>  <groupId>org.apache.hadoop</groupId>  <artifactId>hadoop-azure</artifactId>  <version>2.7.0</version>  </dependency> |

|  |
| --- |
| <dependency>  <groupId>com.fasterxml.jackson.core</groupId>  <artifactId>jackson-core</artifactId>  <version>2.9.0</version>  </dependency> |

1. After running the application at this point we will encounter below error which is related to container access policy.



1. To fix above error we need to grant the read access to the container. By default it is private.

