**Guide to Setup and Troubleshoot eCFD DA environment on Hadoop**

1. **Initialize the Kerbros ticket** in the Hadoop environment using below command

kinit -kt /etc/security/keytabs/svchdphdo5ecfd.service.keytab [svchdphdo5ecfd/hdo5252-31.ca.boeing.com@NOS.BOEING.COM](mailto:svchdphdo5ecfd/hdo5252-31.ca.boeing.com@NOS.BOEING.COM)

1. **Test the Knox gateway** using the below command

curl -vvv -k -iv -u svchdphdo5ecfd -X GET 'https://137.136.234.43:8442/gateway/BOEINGHDO5/webhdfs/v1/?op=LISTSTATUS'

or

curl -vvv -k -iv -u svchdptstecfd -X GET 'https://hdedge.ca.boeing.com:8442/gateway/BOEINGHDO1/webhdfs/v1/?op=LISTSTATUS'

1. **Run Hbase utility** from the ecfd-box to check the HBase connectivity from there. Below is the attached jar.



Use the below command to run the utility

sudo -u jboss /boeing/sw/java/default/bin/java -jar ./da\_ping.jar --table.name=ecfd\_prd:test3 --spring.config.location=override\_HDO3.properties

Or this

sudo -u jboss /boeing/sw/java/default/bin/java -jar ./da\_ping.jar  --spring.config.location=override\_HDO3.properties

If the utility fails with port no error then **check the tunnel configuration**

If the utility fails Due to **HBase** **Ranger policy issue** and see this in logs **“ERROR: org.apache.hadoop.hbase.security.AccessDeniedException: Insufficient permissions for user 'svchdpprdecfd/hdo301-31.ca.boeing.com@NOS.BOEING.COM' (action=create)”**

then raise DBSR for the Ranger Policy issue**.**

1. Check if the **OOzie configuration** is done in Ambari UI otherwise raise DBSR for Hadoop IT.

Use this attached document for the steps.

1. Raise a DBSR to **setup the Mount point** for the Hadoop environment and **provide the permission** for the eCFD Service User on the mount point directory (i.e. */mnt/boeing/sw/ecfdda/hpcdata*).Raise a DBSR for the same.
2. Raise a DBSR for the **Read permission on the all Hadoop environment file system directories** for the Service account user. The eCFD application will be using the Hadoop libs etc. internally. Please see the mail attached mail for reference if Point (5) and (6) are not properly set in the Hadoop environment.



1. Raise a DBSR to **copy the Keytab file on all the nodes** in the Hadoop cluster.
2. Raise a DBSR to **get permission to create directory in the HDFS** so that command like *hadoop fs -mkdir -p /user/svchdphdo5ecfd/input/solution/tobeprocessed/* can be run successfully. Devoid of this permission we could not run our setup script “***solutiondpscript.sh***” in the Hadoop environment.
3. Raise a DBSR to create Read/Write **Policy permission** for Oozie etc. on the **Ranger UI.**See the attached error log in case we miss it



1. Cross check the below **checklist** for the new Hadoop and ecfd environment

|  |  |
| --- | --- |
| **Sno** | **Description** |
| 1 | Firewall from app server to hadoop box knox port to be opened |
| 2 | Firewall from app server to hadoop box zoo keeper port to be opened |
| 3 | Firewall from hadoop boxes to Jboss port of DA server should be opened |
| 4 | Kerberos certificate of edge node to be placed in app server |
| 5 | Script/CronJob that renews Kerberos ticket for Jboss user at regular intervals |
| 6 | Script/Cronjob that renews Kerberos ticket for service user at regular intervals |
| 7 | Service user to be added in the **knox proxy user group** (Refer to hadoop.proxyuser.knox.groups property of /etc/hadoop/conf/core-site.xml to get the knox proxy user group) |

If any items in the checklist is missing raise a DBSR.

1. Raise a ticket to get **Knox permission in Ranger** otherwise the 401 unauthorized error while hitting the Knox gateway.
2. **DA Code Setup Procedure** :Unzip the below attached zip file at / data/home/svchdphdo5ecfd/deployment
   1. Run “**solutiondpscript.sh**” script (in Zip) to create the required folder structure in Hadoop and “solution\_table” in **HBase**.

**Commands for running the script**:

chmod 777 solutiondpscript

sed -i 's/\^H//g;s/\^G//g' solutiondpscript

./solutiondpscript

* 1. Copy the latest eCFD application specific .properties file and jar file in the Hadoop as specified in the attached zip file. using below commands

hadoop fs -put -f /data/home/svchdphdo5ecfd/deployment/config/\* /user/svchdphdo5ecfd/input/config

hadoop fs -put /data/home/svchdphdo5ecfd/deployment/solution /user/svchdphdo5ecfd/

* 1. Copy the lib file comprising processor jar at /user/svchdphdo5ecfd/solution/workflow/lib in HDFS.



Note: **Zip/Unzip command**

zip -r lib.zip lib

unzip lib.zip

1. Now **run the backend commands to test** the DA code
   1. Run the Oozie command at the HDOedgenode command prompt
      1. **For testing Forward copy use below command**

oozie job -oozie http://hdo5250-2.ca.boeing.com:11000/oozie -config /data/home/svchdphdo5ecfd/deployment/solution/config/solution-job-processing.properties -run -Dsubdir=FromHPC/9999 -Dhpcplanid=9999

* + 1. **For testing Reverse copy use below command**

oozie job -oozie http://hdo5250-2.ca.boeing.com:11000/oozie -config solution-hdfstolocal-processing.properties -run -DfilesToBeCopied=/user/svchdphdo5ecfd/input/config/SFTPtoHDFSJob.properties:/user/svchdphdo5ecfd/input/config/SolutionJob.properties:/user/svchdphdo5ecfd/input/config/SolutionBulkloadJob.properties:/user/svchdphdo5ecfd/input/config/solutionDelete.properties~081220161631

Note: Use below Oozie commands to check the logs and progress of Oozie steps:

**Command to see Hadoop logs**

hadoop job -logs job\_1494018504007\_0073 attempt\_1494018504007\_0073\_m\_000000\_0 >> ~/joblogs/8may\_3.txt

**Watch the Oozie job steps**

watch oozie job -oozie http://hdo5250-2.ca.boeing.com:11000/oozie -info 0000033-170428161009460-oozie-oozi-W

**Oozie job logs**

oozie job -oozie http://hdp001-oozie:11000/oozie -log 0000026-161212120136531-oozie-oozi-B

**Kill Oozie job**

oozie job -oozie http://hdp001-oozie:11000/oozie -kill 0000095-161021211005891-oozie-oozi-W

* 1. Run the Curl command to test the Oozie from the Knox gateway
     1. **For testing Forward copy use below command**

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"subDir":"FromHPC/2866","planId":"2866"}' http://da-qa.web.boeing.com:8081/rest/api/knox/copyToDA

or

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"subDir":"FromHPC/2866","planId":"2866"}' http://da-dev.web.boeing.com:8081/rest/api/knox/copyToDA

* + 1. **For testing Forward copy use below command**

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"filesToCopy":"/user/svchdphdo5ecfd/input/config/SFTPtoHDFSJob.properties:/user/svchdphdo5ecfd/input/config/SolutionJob.properties:/user/svchdphdo5ecfd/input/config/SolutionBulkloadJob.properties:/user/svchdphdo5ecfd/input/config/solutionDelete.properties~091220161631"}' http://da-qa.web.boeing.com:8081/rest/api/knox/copyFromDA

**New Ones**

Find below the curl commands for the new services (forward and reverse copy). Need to change the URL and input when needed.

Forward Copy (HPC to DA)

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"files": [{"hpcUserId": "zh722e","hpcServerName": "charon.cs.boeing.com","queryJson": "/home/z/zh722e/offshore/20170512/01/al0.solution.query.json","preferredUid": "zh722e","filePath": "/home/z/zh722e/offshore/20170512/01/al0.solution.json"}],"bemsId": "2690084","planId": "3538","jsonType": "Solution","dataSource": "ecfd"}' <http://da-dev.web.boeing.com:8081/web/api/ecfd/data/save>

Reverse Copy – Containers (DA to HPC)

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"planId": "3538", "containers": [{"id": "1494436713991-20170510","containerName": "cname-1","containerId": "cid-1","targetHpcServer": [{"hpcServerName": "charon.cs.boeing.com","refDirectories": ["/home/z/zh722e/offshore/20170511/01"],    "serviceUserId": "zh722e", "preferredUid": "zh722e"}]}]}' <http://da-dev.web.boeing.com:8081/web/api/ecfd/containers/push>

Reverse Copy – Dataset (DA to HPC)

curl -i -H "Accept: application/json" -H "Content-Type: application/json" -X POST -d '{"referenceId" : "refid1","datasetId" : [28, 31], "hpcServerName": "charon.cs.boeing.com", "hpcPath": "/home/z/zh722e/offshore/20170511/01", "serviceUserId": "zh722e", "preferredUid": "zh722e"}'  <http://da-dev.web.boeing.com:8081/web/api/ecfd/data/push>