**Zeppelin User Impersonation Per User and Isolated Per User**

This document completely focuses how to set up Zeppelin Interpreter **Per User**, **Isolated Per User** with **User Impersonation** using zeppelin 0.7.0/0.7.2/0.7.3 or HDP 2.6

How to set up user impersonation at Zeppelin Shell Interpreter level with shiro authentication can found from [here](https://zeppelin.apache.org/docs/0.7.0/manual/userimpersonation.html)

What the different kinds of ***Interpreter binding modes*** is well discussed in the upcoming version 0.8.0 of Zeppelin [documentation](https://zeppelin.apache.org/docs/0.8.0-SNAPSHOT/usage/interpreter/interpreter_binding_mode.html)

Before setting up the Zeppelin User Impersonation make sure you have integrated your zeppelin and your cluster with AD/LDAP, below is the *shiro.ini* for AD authentication:

*[users]*

*# List of users with their password allowed to access Zeppelin.*

*# To use a different strategy (LDAP / Database / ...) check the shiro doc at http://shiro.apache.org/configuration.html#Configuration-INISections*

*admin = admin, admin*

*user1 = user1, role1, role2*

*user2 = user2, role3*

*user3 = user3, role2*

*# Sample LDAP configuration, for user Authentication, currently tested for single Realm*

*[main]*

*### A sample for configuring Active Directory Realm*

*activeDirectoryRealm = org.apache.zeppelin.realm.ActiveDirectoryGroupRealm*

*activeDirectoryRealm.systemUsername = CN=Administrator,CN=Users,DC=ADSERVER,DC=com*

*activeDirectoryRealm.systemPassword = loginpassword*

*activeDirectoryRealm.searchBase = CN=Users,DC=ADSERVER,DC=com*

*activeDirectoryRealm.url = ldap://ldap.ad.ADSERVER.com:389*

*activeDirectoryRealm.authorizationCachingEnabled = true*

*sessionManager = org.apache.shiro.web.session.mgt.DefaultWebSessionManager*

*securityManager.sessionManager = $sessionManager*

*### If caching of user is required then uncomment below lines*

*cacheManager = org.apache.shiro.cache.MemoryConstrainedCacheManager*

*securityManager.cacheManager = $cacheManager*

*#securityManager.sessionManager = $sessionManager*

*# 86,400,000 milliseconds = 24 hour*

*securityManager.sessionManager.globalSessionTimeout = 86400000*

*shiro.loginUrl = /api/login*

*[roles]*

*role1 = \**

*role2 = \**

*role3 = \**

*admin = \**

*[urls]*

*# This section is used for url-based security.*

*# You can secure interpreter, configuration and credential information by urls. Comment or uncomment the below urls that you want to hide.*

*# anon means the access is anonymous.*

*# authc means Form based Auth Security*

*# To enfore security, comment the line below and uncomment the next one*

*/api/version = anon*

*#/api/interpreter/\*\* = authc, roles[admin]*

*#/api/configurations/\*\* = authc, roles[admin]*

*#/api/credential/\*\* = authc, roles[admin]*

*#/\*\* = anon*

*/\*\* = authc*

**Below are the configurations required from the environment:**

* On the Zeppelin server node make sure to add the below to */etc/sudoers:*

With the root user:

*$visudo*

*zeppelin ALL=(ALL) NOPASSWD: ALL*

If this entry is missing in the /etc/sudoers then the interpreter fails with the below error:

*org.apache.zeppelin.interpreter.InterpreterException: Permission denied, please try again.*

*Permission denied, please try again.*

*Permission denied (publickey,gssapi-keyex,gssapi-with-mic,password).*

* Log folder (usually /var/log/zeppelin or /var/log/hadoop/zeppelin) of zeppelin should be of 777 instead of 755 if not the Spark interpreter keep failing with below error as the impersonated user won’t able to write to the logs folder:

*java.lang.NullPointerException*

*at org.apache.zeppelin.spark.Utils.invokeMethod(Utils.java:38)*

*at org.apache.zeppelin.spark.Utils.invokeMethod(Utils.java:33)*

* If you are on a HDP platform this can’t be achieved from the CLI which needs a Ambari Zeppelin code level change

*$ vi /var/lib/ambari-server/resources/common-services/ZEPPELIN/0.6.0.2.5/package/scripts/master.py*

Find the below:

*def create\_zeppelin\_log\_dir(self, env):*

*import params*

*env.set\_params(params)*

*Directory([params.zeppelin\_log\_dir],*

*owner=params.zeppelin\_user,*

*group=params.zeppelin\_group,*

*cd\_access="a",*

*create\_parents=True,*

***mode=0755***

*)*

**Change this to:**

*def create\_zeppelin\_log\_dir(self, env):*

*import params*

*env.set\_params(params)*

*Directory([params.zeppelin\_log\_dir],*

*owner=params.zeppelin\_user,*

*group=params.zeppelin\_group,*

*cd\_access="a",*

*create\_parents=True,*

***mode=0777***

*)*

Please note this requires the stop and start of AMBARI-SERVER.

**Configurations required for Zeppelin User Impersonation:**

* Add the proxy settings for ***zeppelin*** user under core-site.xml and restart the required services (HDFS, Mapreduce2 and YARN)

*hadoop.proxyuser.zeppelin.hosts=\**

*hadoop.procyuser.zeppelin.goups=\**

This configuration is required as we will be giving the control to zeppelin user to impersonate other users (users using the zeppelin UI)

* Add the below lines to ***zeppelin-env.sh*** file (or zeppelin-env from *Ambari 🡪 Zeppelin 🡪 Configs 🡪 Advanced zeppelin-env*)

*# --- FOR IMPERSONATION -- START ----*

*ZEPPELIN\_IMPERSONATE\_USER=`echo ${ZEPPELIN\_IMPERSONATE\_USER} | cut -d "@" -f1` # we are trimming the user to get short name i.e. username from username@domain.com*

*export ZEPPELIN\_IMPERSONATE\_CMD='sudo -H -u zeppelin bash -c' #zeppelin user in command to do the impersonation*

*export SPARK\_HOME=/usr/hdp/current/spark2-client/*

*export PYTHONPATH=$SPARK\_HOME/python/:$PYTHONPATH*

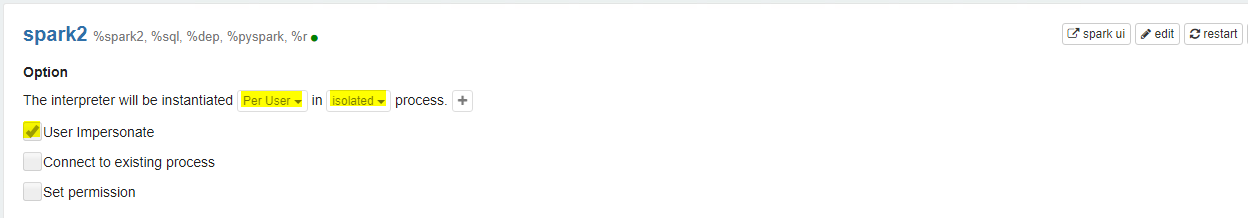
*export PYTHONPATH=$SPARK\_HOME/python/lib/py4j-0.10.4-src.zip:$PYTHONPATH # This is version based config and needs to be changed based on the spark version*

*export SPARK\_YARN\_USER\_ENV="PYTHONPATH=${PYTHONPATH}"*

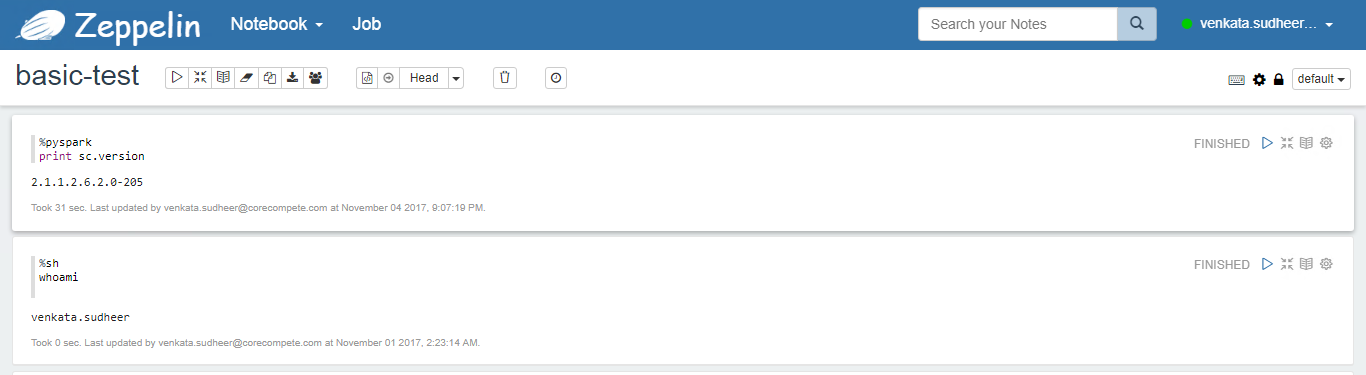
*# --- FOR IMPERSONATION -- END ----*

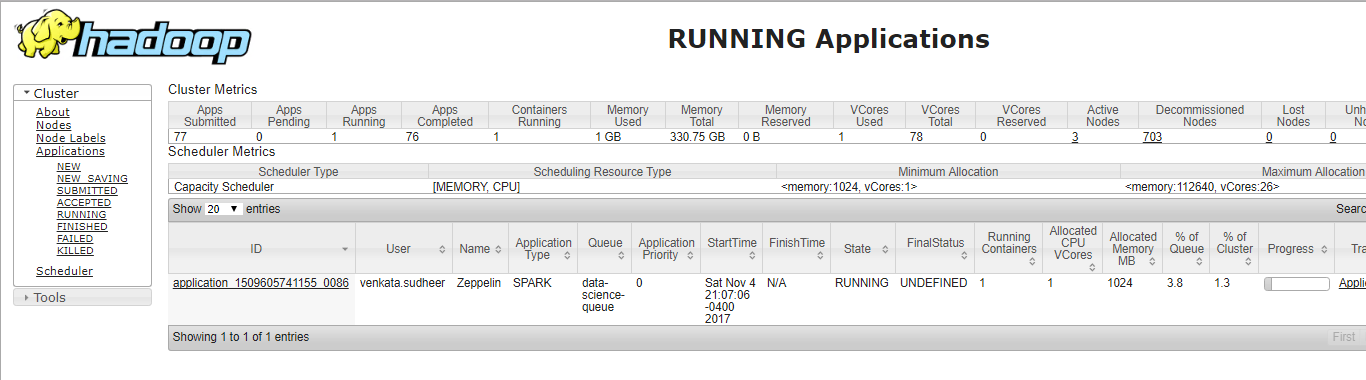
And restart the zeppelin service.

Once all the above configurations are completed and restarted the required services, now make sure to do the required configurations (Per User, Isolated Per User and User Impersonation as shown below) from ***Zeppelin 🡪 Interpreter 🡪 Spark/Spark2***



Once any command/job run from Spark Interpreter with YARN will launched with the login user and from the zeppelin Shell interpreter we can see the user id:





Hope this helps someone who is struggling with Zeppelin User Impersonation while running the ***Zeppelin 🡪 Spark 🡪 YARN***

More details about the issue and the errors can be found from:

<https://issues.apache.org/jira/browse/ZEPPELIN-3016>