**Accessing Hadoop Ecosystem Components**

Version -V1.0

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version | Description | Author |
| 03/26/2018 | 1.0 | Initial Version | Yagnaraman Madhavan |
|  |  |  |  |
|  |  |  |  |

**Version History**

Contents

[1.0 Accessing Spark 4](#_Toc510016509)

[1.1 Logging into the REPL 4](#_Toc510016510)

[1.2 Interactive Data Processing: REPL 5](#_Toc510016511)

[1.3 Executing hive queries using Spark SQL & JDBC connection in Beeline: 6](#_Toc510016512)

[2.0 Accessing Hive: 6](#_Toc510016513)

[2.1 Dev On Prem: 6](#_Toc510016514)

[2.2 Dev Azure: 6](#_Toc510016515)

[2.3 Test: 7](#_Toc510016516)

[2.4 Prod 7](#_Toc510016517)

[3.0 Accessing Pig Grunt shell 7](#_Toc510016518)

[*3.1* Local Mode: 7](#_Toc510016519)

[3.2 Mapreduce Mode 8](#_Toc510016520)

[4.0 Accessing Talend 8](#_Toc510016521)

[4.1 Create a new Project 8](#_Toc510016522)

[4.2 Selecting the project 9](#_Toc510016523)

[4.3 Develop the components in the workspace 9](#_Toc510016524)

[4.4 Building the Talend Job 9](#_Toc510016525)

[5.0 Accessing Hbase Shell & Pheonix 11](#_Toc510016526)

[5.1 Hbase Shell 11](#_Toc510016527)

[5.2 Pheonix 12](#_Toc510016528)

[5.2.1 Dev Azure 12](#_Toc510016529)

[5.2.2 Dev On Prem 12](#_Toc510016530)

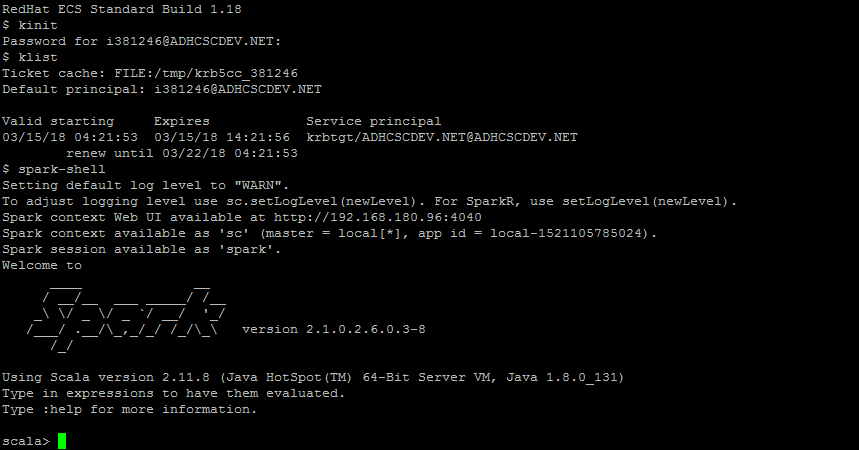
[5.2.3 Test 12](#_Toc510016531)

[5.2.4 Prod 12](#_Toc510016532)

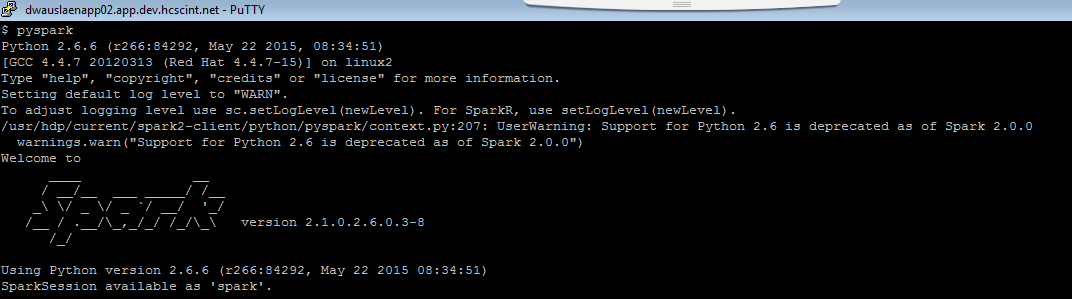
# Accessing Spark

# Logging into the REPL

* **Scala shell**

****

* Type :q or sys.exit to exit from the shell
* By default, spark context & spark session will be available.
* **Python shell**

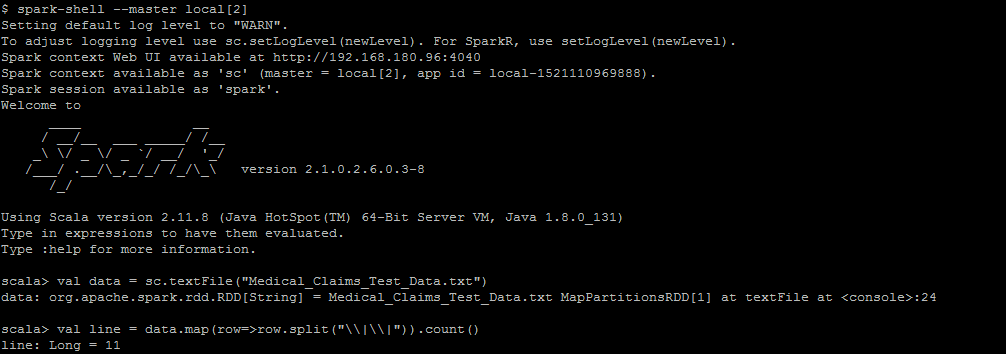
****

* By default, spark context & spark session will be available.
* Type exit() to exit from REPL.

# **Interactive Data Processing: REPL**

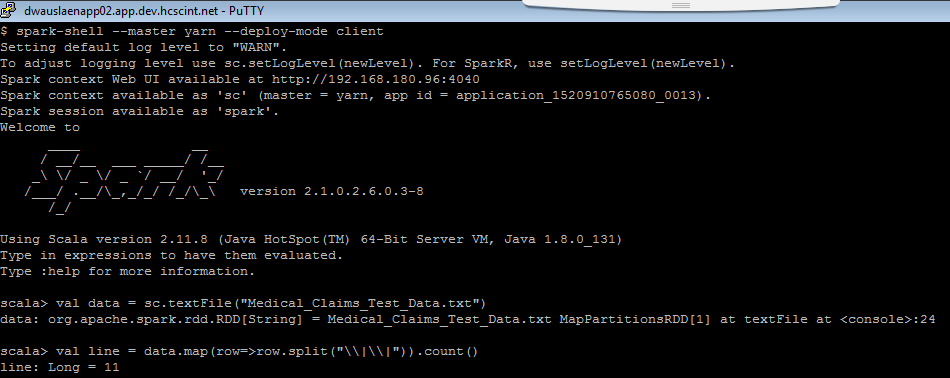
Spark shell in Local Mode

* Spark shell with deployment mode => Local 2 threads (cores)

******

Spark shell with Yarn client mode

* Spark shell with deployment mode => Yarn client

******

# **Executing hive queries using Spark SQL & JDBC connection in Beeline:**

Thriftserver details: Dev On-prem

dwauslenapp02.app.dev.hcscint.net:10016

Hive Queries can be executed using JDBC connection & sparksql & thriftserver

There are 2 methods to execute hive queries using spark:

1. HTTP method
2. Non-HTTP method.

beeline -u "jdbc:hive2://dwauslenapp02.app.dev.hcscint.net:10016/;transportmode=http;principal=hive/\_HOST@ADHCSCDEV.NET;httpPath=cliservice"?

**The following connection string accesses Spark SQL through JDBC on a Kerberos-enabled cluster:**

beeline -u "jdbc:hive2://dwauslenapp02.app.dev.hcscint.net:10016/;transportmode=http;principal=hive/\_HOST@ADHCSCDEV.NET;httpPath=/"?

**The following connection string accesses Spark SQL through JDBC over HTTP transport on a Kerberos-enabled cluster:**

beeline -u "jdbc:hive2://dwauslenapp02.app.dev.hcscint.net:10016/;principal=hive/\_HOST@ADHCSCDEV.NET;httpPath=/"?

# **Accessing Hive:**

Beeline:

# **Dev On Prem:**

beeline -u "jdbc:hive2://dwauslmnapp02.app.dev.hcscint.net:10001/;principal=hive/\_HOST@ADHCSCDEV.NET;transportMode=http;httpPath=cliservice"?

# **Dev Azure:**

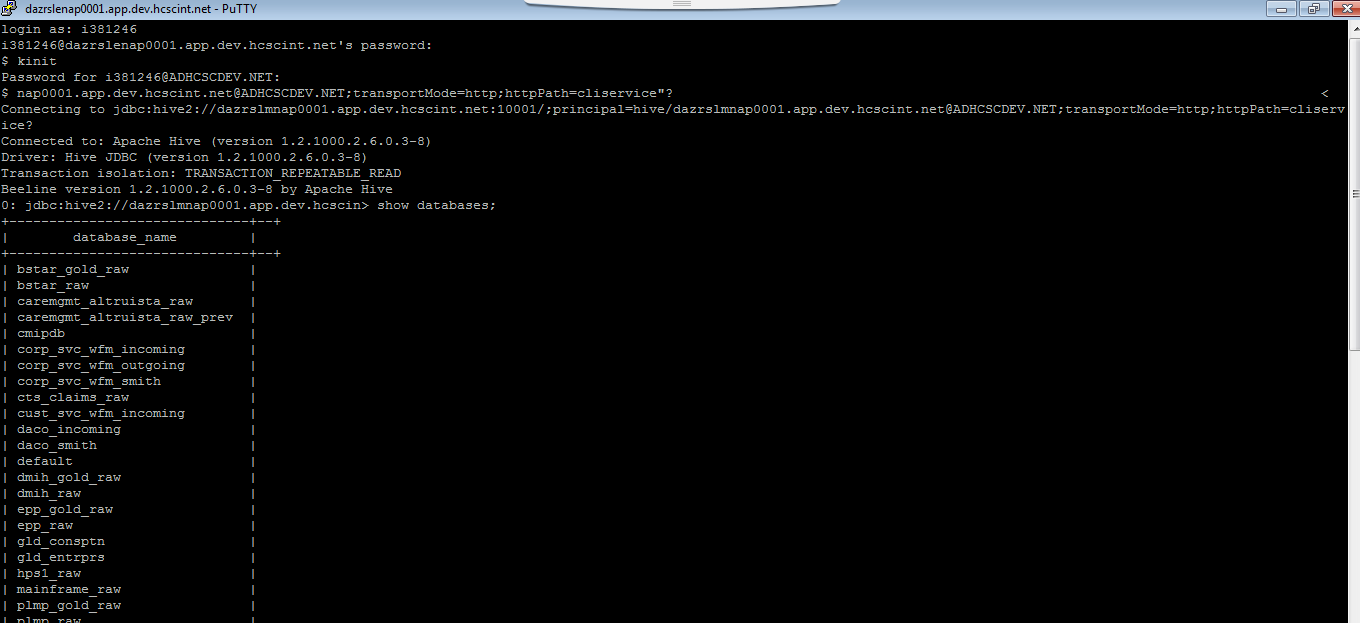
beeline -u "jdbc:hive2://dazrslmnap0001.app.dev.hcscint.net:10001/;principal=hive/dazrslmnap0001.app.dev.hcscint.net@ADHCSCDEV.NET;transportMode=http;httpPath=cliservice"?

# **Test:**

beeline -u "jdbc:hive2://twauslmnapp01.app.test.hcscint.net:10001/;principal=hive/\_HOST@ADHCSCTST.NET;transportMode=http;httpPath=cliservice"

# **Prod**

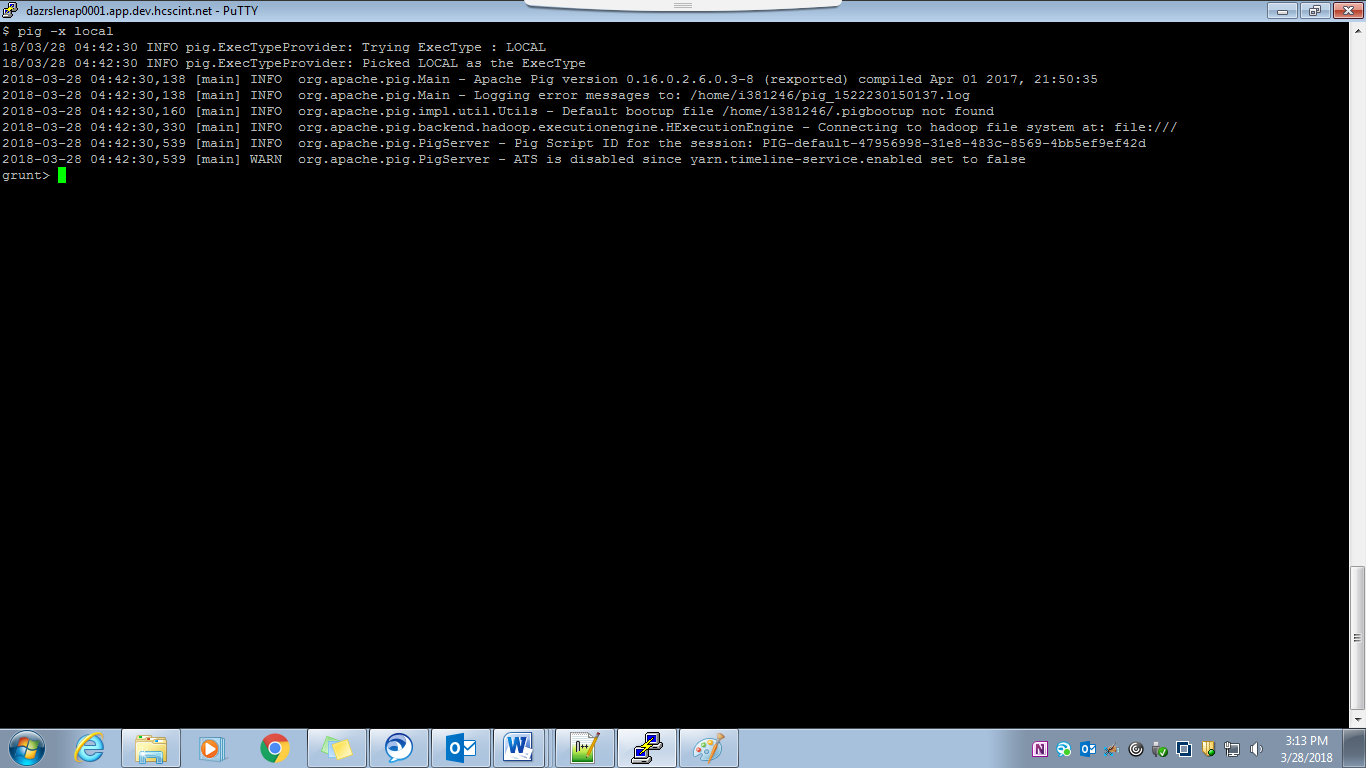
beeline -u ?"jdbc:hive2://pwauslmnapp03.app.hcscint.net:10001/;transportMode=http;httpPath=cliservice;principal=hive/pwauslmnapp03.app.hcscint.net@ADHCSCINT.NET;":



# **Accessing Pig Grunt shell**

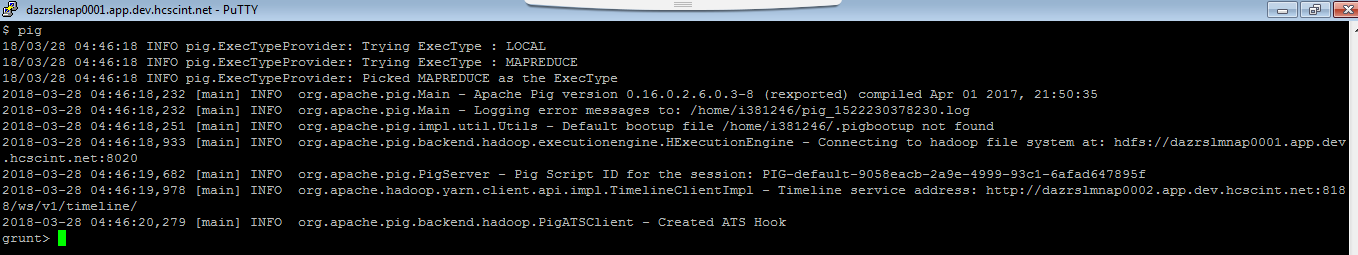
# **Local Mode:**

Type pig –x local in the command line



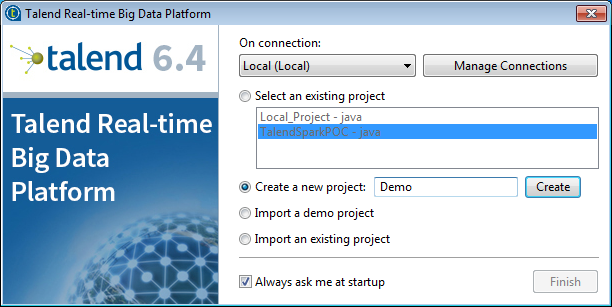
# Mapreduce Mode

Type pig in the command line

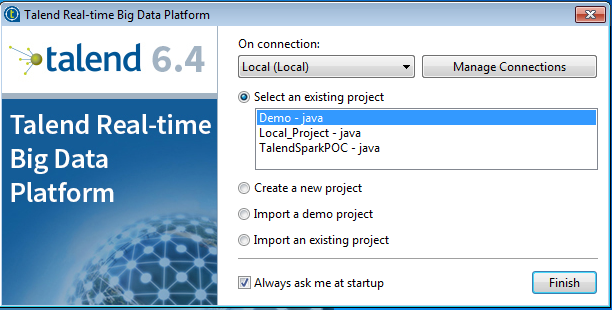


# Accessing Talend

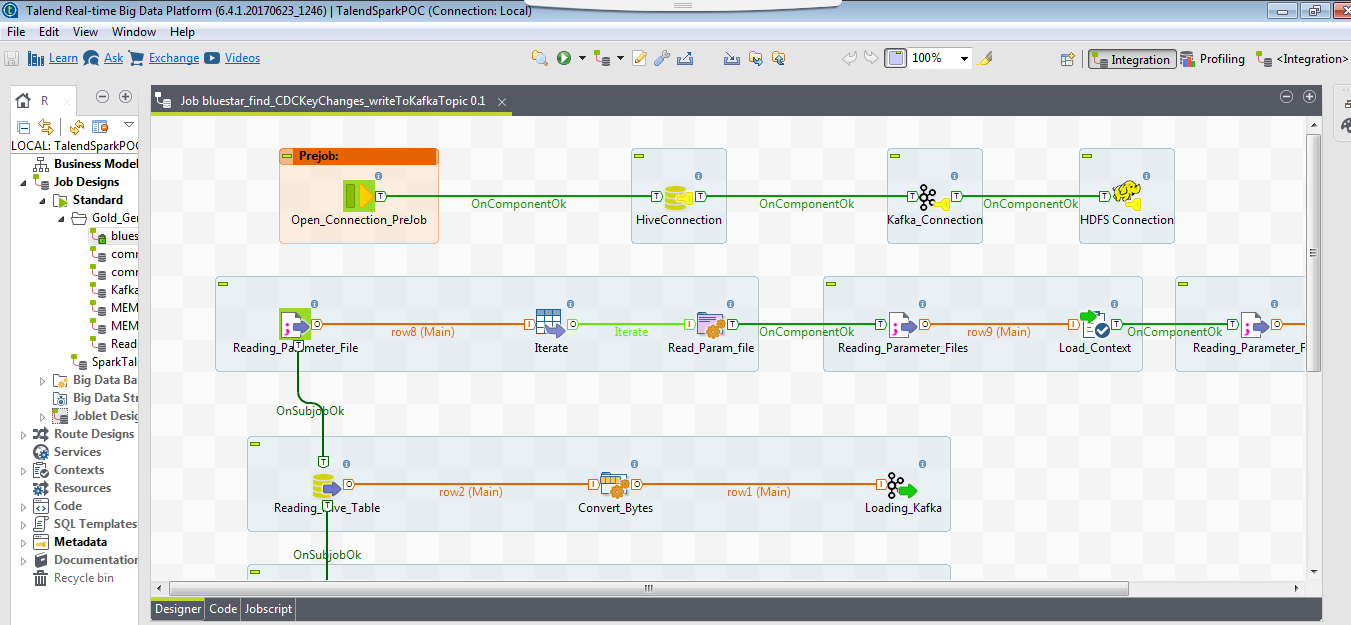
# Create a new Project



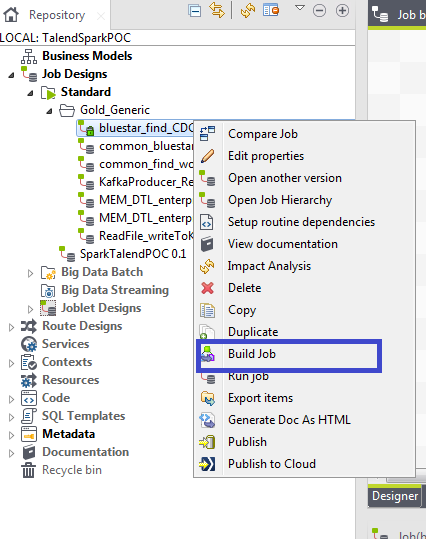
# Selecting the project



# Develop the components in the workspace



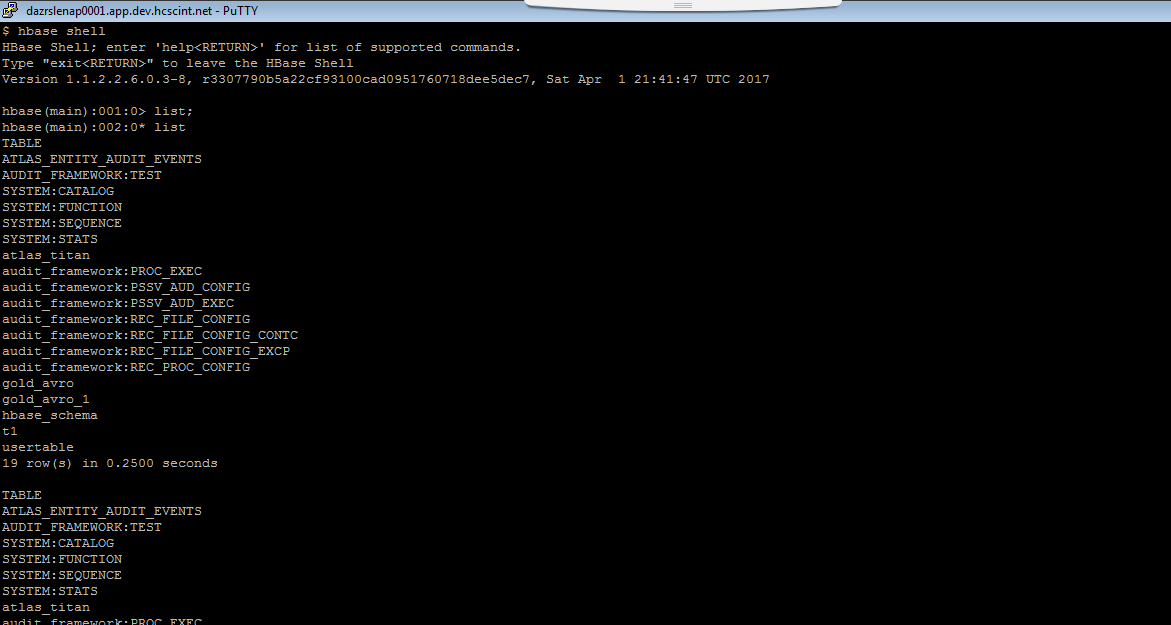
# Building the Talend Job

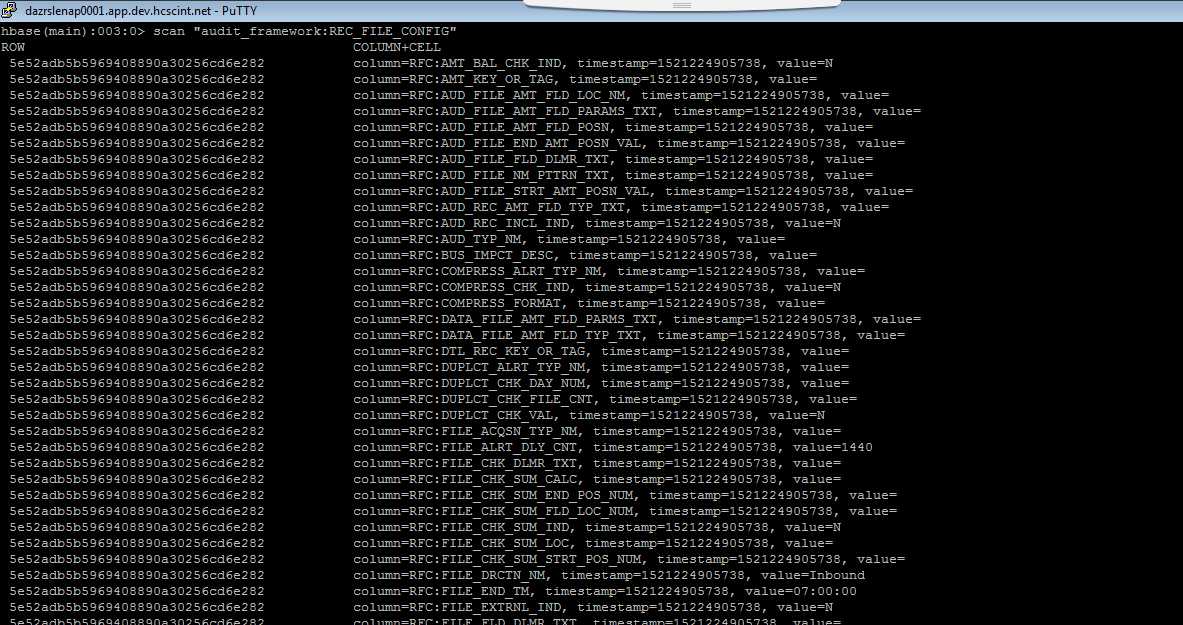


# Accessing Hbase Shell & Pheonix

# Hbase Shell

Type hbase shell in the command line

****

****

Type quit or exit in the command line to exit from the hbase shell.

# Pheonix

# Dev Azure

/usr/hdp/current/phoenix-client/bin/sqlline.py dazrslmnap0002.app.dev.hcscint.net,dazrslmnap0003.app.dev.hcscint.net,dazrslsrap0001.app.dev.hcscint.net:2181/hbase-secure

# Dev On Prem

/usr/hdp/current/phoenix-client/bin/sqlline.py dwauslmnapp02.app.dev.hcscint.net,dwauslmnapp03.app.dev.hcscint.net,dwausldnapp04.app.dev.hcscint.net:2181:/hbase-secure

# Test

/usr/hdp/current/phoenix-client/bin/sqlline.py twauslmnapp04.app.test.hcscint.net,twauslmnapp05.app.test.hcscint.net,twauslmnapp06.app.test.hcscint.net:2181:/hbase-secure

# Prod

/usr/hdp/current/phoenix-client/bin/sqlline.py pwauslmnapp01.app.hcscint.net,pwauslmnapp02.app.hcscint.net,pwauslmnapp03.app.hcscint.net:2181:/hbase-secure

