Oracle Financial Services Compliance Studio

Installation Guide

Release 8.1.2.0.0

July 2023

F48800-01

ORACLE

Financial Services

OFS Compliance Studio Installation Guide

Document Control

Table 1 lists the document control of this guide:

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.2.5.0	July 2023	Updated steps in the Generate truststore File for Elasticsearch section.
8.1.2.3.0	January 2023	Added the following sections: ● Create the Tablespace ● Create the Sandbox Schema ● Assign Grants for the Sandbox Schema Java supported version is added in the Prerequisite Environmental Settings and Frequently Asked Questions in Compliance Studio sections.
8.1.2.1.0	November 2022	Added a note and updated the value of **maxTotal** in the Configure the resources.xml for Multiple ER Schemas section.
8.1.2.1.0	October 2022	Added a new sub-step (19.d) in the Frequently Asked Questions in Compliance Studio section. Added FAQ on interpreter settings and upgrade the python virtual environment for the fcc-python interpreter in the Frequently Asked Questions in Compliance Studio section.
8.1.2.1.0	September 2022	Updated with note information for CDH in the following sections: ● Hardware and Software Requirements (Big Data) ● Download the Big Data Files (Additional Jars) ● Appendix C – Additional Jars – PGX ● Appendix D – Additional Jars – Batch Service Updated with correct reference topics in Configure the Extract Transfer and Load (ETL) Process section. Updated Installing Analytics ICU Plugin section. Updated Generate API token for CS API User section. Updated SQL statement in the Create the Studio Schema section. Added FAQ on retaining logs after restart in the Frequently Asked Questions in Compliance Studio section. Added FAQ on system's JDK 8 and bundled JDK in the Frequently Asked Questions in Compliance Studio section. Added Configure the PGX Interpreter section. Added Generate Signed Certificate section. Added FAQ on java memory error in the Frequently Asked Questions in Compliance Studio section.

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.2.0.1	May 2022	As part of this release, the following sections are updated: ● Updated the upgrade version, steps in Installation Checklist table with OFSAA and without OFSAA in the Introduction section. ● Updated bug number in Download the Installer Kit section. ● Updated the notes in STUDIO_DB_SID and AUTOMIC_DB_SID in the Configure the config.sh File section. ● Updated the Place Files in Wallet section. ● Updated steps in Stop the PGX Service and Upgrade Steps without OFSAA sections. ● Added Upgrade from 8.1.2.0.0 to 8.1.2.0.1 section. ● Added Perform Cleanup for Entity Resolution section. ● Added Appendix F – Create Users, Groups, and Mappings section.
8.1.2.0.0	April 2022	Removed the following: ● Configure the ore Interpreter section. ● Configure the fcc-python interpreter section. ● ORE Interpreter settings from Configure the config.sh File section. ● Generate an Encrypted Password for the Elastic Search section. ● One permission from Clean up for Compliance Studio Schema section. ● FAQ 16 in the Frequently Asked Questions in Compliance Studio section. Updated the following: ● Modified the component versions in the Hardware and Software Requirements table for Elastic Search, Logstash, and ES Hadoop Jars. ● Updated the note in Configure the Extract Transfer and Load (ETL) Process section. ● Updated Loading sample graph without running ETL section. ● Updated the description in STUDIO_DB_ENCRYPTED_PASSWORD, ELASTIC_SEARCH_ENCRYPTED_PASSWORD, ENCRYPTED_QUANTIFIND_TOKEN parameters and modified the note in Configure the config.sh File section.

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.2.0.0	April 2022	- Updated significance for parameters in the table in Install the PGX Service table. Added the following: - Configure Logstash section. - Added a note in Create the Studio Schema section. - Added a note in Assign Grants for the Sandbox Schema section. - Added a note in the Clean up for Compliance Studio Schema section. - Added a note in Loading sample graph without running ETL section. - FAQ 18 in the Frequently Asked Questions in Compliance Studio section. - Note in Appendix C – Additional Jars – PGX chapter.
8.1.2.0.0	March 2022	Updated the following sections: - Updated Hardware and Software Requirements table. - Added pgx-python in the Configure the Interpreter Settings - Configure the Spark Interpreter - Download the Installer Kit - Extract the Installer Kit - Generate the Public and Private Keys - Updated UI screenshots in the Configure Python Interpreter Setting - Updated API_USERS and SSO_TOKEN parameter in the Configure the config.sh File - Added from 13 to 18 FAQs in the Frequently Asked Questions in Compliance Studio - Updated aopalliance-1.0.jar in Appendix C – Additional Jars – PGX Added the following sections: - Upgrade from 8.0.8.2.0 to 8.1.2.0.0 - Upgrade from 8.1.1.1.0 to 8.1.2.0.0 - Generate API token for CS API User - Perform Cleanup for Templates - Perform Cleanup for Interpreters - Sample spark-default.conf Configuration File
8.1.1.1.0	December 2021	The Appendix E – Apache Log4j Security Alert CVE-2021-44228 Patch Details section is added for the Patch 33684394 release.

Table 1: Document Control

Version Number	Revision Date	Change Log
8.1.1.1.0	November 2021	This is created for the v8.1.1.1.0 release.
8.1.1.0.0	October 2021	This is created for the v8.1.1.0.0 release.

Table of Contents

1 Preface

This section provides the Oracle Financial Services (OFS) Compliance Studio Installation Guide information.

Topics:

- Audience
- Related Documents
- Conventions
- Abbreviations

1.1 Audience

OFS Compliance Studio Installation Guide is intended for System Engineers who are responsible for installing and maintaining the application.

This document assumes that you have experience in installing Enterprise components and basic knowledge about the following:

- UNIX commands
- Database concepts
- Big Data concepts

1.2 Related Documents

You can strive to keep this and all other related documents updated regularly; visit the OHC Documentation Library to download the latest version available there. The list of related documents is provided here.

- Oracle Financial Services Compliance Studio Administration and Configuration Guide
- Oracle Financial Services Compliance Studio User Guide
- Oracle Financial Services Compliance Studio Matching Guide
- Oracle Financial Services Compliance Studio Data Model Guide
- Oracle Financial Services Compliance Studio Release Notes
- Oracle Financial Services Compliance Studio Use Case Guide

1.3 Conventions

Table 2 lists text conventions are used in this document.

Table 2: Document Conventions

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.

Table 2: Document Conventions

Convention	Meaning
monospace	Monospace type indicates commands within a paragraph, URLs, code in examples, file names, text on the screen, or text you enter.
Hyperlink	Hyperlink type indicates the links to external websites and internal document links to sections.

1.4 Abbreviations

Table 3 lists the abbreviations used in this document.

Table 3: Abbreviations

Abbreviation	Meaning
OFS	Oracle Financial Services
Compliance Studio	Oracle Financial Services Compliance Studio
OFSAA	Oracle Financial Services Analytical Application
BD	Behavior Detection
FCDM	Financial Crime Data Model
ICIJ	International Consortium of Investigative Journalists
IDCS	Oracle Identity Cloud Service
ECM	Enterprise Case Management
SSO	Single Sign-On
SSH	Secure Shell

2 Introduction

OFS Compliance Studio is an advanced analytics application that supercharges anti-financial crime programs for better customer due diligence, transaction monitoring, and investigations by leveraging the latest innovations in artificial intelligence, open-source technologies, and data management. It combines Oracle's Parallel Graph Analytics (PGX), Machine Learning for AML, Entity Resolution, and notebook-based code development and enables Contextual Investigations in one platform with complete and robust model management and governance functionality.

This chapter provides the information required to understand the installation of the OFS Compliance Studio application.

This release (v8.1.2.0.0) of Compliance Studio can be used for the following:

- To install a new instance of Compliance Studio as follows:
- Compliance Studio with OFSAA (Oracle Financial Services Analytical Application). Here, OFSAA is with Behavior Detection (BD) or Enterprise Case Management (ECM).
- Compliance Studio without OFSAA

> **NOTE** You can install the v8.1.2.0.1 directly. The process of installation is similar to 8.1.2.0.0 Installation.

To upgrade an existing instance of Compliance Studio as follows:

- Upgrade Compliance Studio from v8.1.1.1.0 onwards to Compliance Studio v8.1.2.0.0.

  OR
- Upgrade FCC Studio from v8.0.8.2.0 onwards to Compliance Studio v8.1.2.0.

Then you can upgrade Compliance Studio from v8.1.2.0.0 onwards to Compliance Studio v8.1.2.0.1.

Topics:

- Installation Check List when Studio is installed with OFSAA
- Installation Check List when Studio is installed without OFSAA

2.1 Installation Check List when Studio is installed with OFSAA

To complete the installation process, you must perform the steps listed in the Table 4 Checklist. Use this checklist to verify whether these steps are completed or not. Click the reference link to go to the topic.

Table 4: Installation Check List

Sl. No.	Activity	Mandatory	Description
	Pre-installation Steps		-
1	Install all the prerequisite Hardware and Software Requirements.	Yes	-

Table 4: Installation Check List

2	Setup the environmental settings (System Configuration).	Yes	-
3	Download the Big Data Files	No	It is required for graph analytics and leverages fragmented data or as a datasource for models.
4	Configure the Elastic Search Component	No	It is required for graph analytics and leverage fragmented data or for matching service and Entity Resolution
5	Configure the Interpreter Settings	Yes	-
6	`GRANT DROP ANY TRIGGER TO <SANDBOX SCHEMA USER>;`	Yes	-
7	See the Configure the resources.xml for Multiple ER Schemas section for more details.	Yes	-
8	Setup Password Stores with Oracle Wallet	Yes	-
9	Create the Credential Keystore	No	It is required for graph analytics and leverages fragmented data or as a datasource for models
10	Download the Installer Kit	Yes	-
	Installation Steps		-
1	Extract the Installer Kit	Yes	-
2	Place Files in the Installation Directories	Yes	-
3	Generate an Encrypted Password	Yes	-
4	Generate the Public and Private Keys	Yes	-
5	Generate API token for CS API User	Yes	-
6	Generate the Key Store File for Secure Batch Service	Yes	-
7	Add the Batch Service (SSL) to PGX Configuration	Yes	-
8	Configure the Extract Transfer and Load (ETL) Process	No	It is required for graph analytics and leveraging fragmented data
9	Configure the config.sh File	Yes	-
10	Run the Compliance Studio Installer	Yes	-
11	Install the PGX Service	Yes	-
	Post-Installation Steps		-
1	Verify the Installation	Yes	-

Table 4: Installation Check List

2	Start the PGX Service	Yes	-
3	Access the Compliance Studio Application	Yes	-
4	Perform the OFSAA Configuration for Batch Execution	No	It is required if leverage OFSAA's scheduling and executing capability.
5	Configure and Run Published Notebooks	No	It is required if leveraging OFSAA's batch execution.

2.2 Installation Check List when Studio is installed without OFSAA

To complete the installation process, you must perform the steps listed in the Table 5 Checklist. Use this checklist to verify whether these steps are completed or not. Click the reference link to go to the topic.

Table 5: Installation Check List

Sl. No.	Activity	Mandatory	Details
	Pre-installation Steps		
1	Install all the prerequisite Hardware and Software Requirements.	Yes	-
2	Setup the environmental settings (System Configuration).	Yes	-
3	Configure the Interpreter Settings	Yes	-
4	`GRANT DROP ANY TRIGGER TO <SANDBOX SCHEMA USER>;`	Yes	-
5	See the Configure the resources.xml for Multiple ER Schemas section for more details.	Yes	-
6	Setup Password Stores with Oracle Wallet	Yes	-
7	Create the Credential Keystore	Yes	-
8	Download the Installer Kit	Yes	-
	Installation Steps		
1	Extract the Installer Kit	Yes	-
2	Place Files in the Installation Directories	Yes	-
3	Generate an Encrypted Password	Yes	-
4	Generate the Public and Private Keys	Yes	-
5	Generate API token for CS API User	Yes	-
6	Generate the Key Store File for Secure Batch Service	Yes	-

Table 5: Installation Check List

7	Configure the config.sh File	Yes	-
8	Run the Compliance Studio Installer	Yes	-
	Post-Installation Steps		
1	Verify the Installation	Yes	-
2	Access the Compliance Studio Application	Yes	-

3 Pre-installation

This chapter provides information about the tasks that must be performed before installing Compliance Studio. To install Compliance Studio with OFSAA, ensure the Behavior Detection (BD) or the Enterprise Case Management (ECM) application pack is installed.

The following patches are required only when integrating with old versions for ECM:

- On top of ECM 8.0.8.0.0, apply the following ECM patch for ML-ECM integrations.

  8.0.8.0.28 (BUG: **31497997**)
- On top of ECM 8.0.8.1.0, apply the following ECM patch for ML-ECM integrations.

  8.0.8.1.4 (BUG: **33395125**)

> **NOTE** From ECM 8.1.1.0.0 and later versions, the above patches are not required for ML-ECM integrations.

Topics:

- Hardware and Software Requirements
- Setup Password Stores with Oracle Wallet

3.1 Hardware and Software Requirements

The following hardware and software are required for this version of Compliance Studio. The installation environment or setup must have these requirements for an application to run smoothly and efficiently.

Topics:

- System Configuration
- Prerequisite Environmental Settings
- Download the Big Data Files
- Validation Checklist
- Configure the Elastic Search Component
- Configure Logstash
- Installing Analytics ICU Plugin
- Configure the Interpreter Settings
- Create the Hive Schema
- Create the Tablespace
- Create the Studio Schema
- Assign Grants for the Studio Schema
- Create the Sandbox Schema
- Assign Grants for the Sandbox Schema
- Entity Resolution

Table 6 lists the Hardware and Software Requirements:

Table 6: Hardware and Software Requirements

Hardware or Software Category	Component Version
Browser	`Chrome`
Java Version	`Java 8`
Processing Server	● `RHEL 7.6+` ● `Oracle JRE Standard Edition 1.8.x(with JCE)`
Database Server	● `Oracle Database Release 19c (19.3+)` ● `Oracle Machine Learning for R (OML4R) (formerly ORE) 1.5.1 with Open source R or Oracle R Distribution 3.6.1` Click here to get the supported DB versions.
PGX (Graph) Server	● `RHEL 7.4+` ● `Minimum gcc library v4.8.2`
Elastic Search	`Elastic Search 7.13.4 and 7.14 versions` **NOTE:** Compliance Studio certified with 7.13.4 and 7.14 versions.
Logstash	`7.13.4 and 7.14 versions` **NOTE** : ● `Compliance Studio is certified with 7.13.4 and 7.14 versions.` ● `Logstash version should be the same as Elastic Search` For example, if the ES version is 7.14.0, the Logstash version should also be 7.14.0.
Elastic Search Hadoop Jars	`Elastic search 7.13.4 and 7.14` versions are also supported. Elastic Search can be downloaded from Elastic Search.
Oracle Instant Client	`instantclient-basic-linux.x64-19.8.0.0.0` **NOTE**: The version should be the same as the Database version, and this should be present in the processing server.
Big Data **NOTE:** You can use either **Cloudera** or open-source **Apache** for a big data cluster.	

Table 6: Hardware and Software Requirements

Hadoop and Spark	**NOTE**: Kerberos authentication must be enabled for Big Data. ● Apache Hadoop Version 3.0.0 ● Apache Hive Version 2.1.1 ● Apache Spark Version 2.4.0 ● Apache Sqoop Version 1.4.7 ● The **.profile** file must be present with the SPARK_HOME and PYTHON_HOME parameters already set. **NOTE**: The product is certified for Apache-Hadoop, and any vendor-specific Hadoop distributions have to confirm compliance with Apache-Hadoop standards, and if not, the vendor the customer chooses to work with for Hadoop should ensure compliance to Apache-Hadoop standards. Any issue raised on vendor-specific distributions will be fixed only if the issue is reproducible on Apache-Hadoop, Apache-Hive, and Apache-Spark.
Hive Connectors	`Hive JDBC Connectors V 2.5.15`
Apache	● `Kerberos 1.19.1` ● `Hadoop Version 3.0.0` ● `Hive Version 3.1.2` ● `Spark Version 2.4.8 (with Hadoop)` ● `Sqoop Version 1.4.7` **NOTE**: ● The **.profile** file must be present with the **SPARK_HOME and PYTHON_HOME** parameters already set. ● Kerberos authentication must be enabled for the above services and ensure these services are Apache standards. ● The product is certified for Apache-Hadoop, and any vendor-specific Hadoop distributions must confirm compliance with Apache-Hadoop standards. If not, the vendor, the customer, who chooses to work with Hadoop should comply with the Apache-Hadoop standards. Any issue raised on vendor-specific distributions will be fixed only if the issue is reproducible on Apache-Hadoop, Apache-Hive, and Apache-Spark.
Hadoop Security Protocol	● `Kerberos 5` ● `Apache Sentry-2.1.0`

3.1.1 System Configuration

1. Log in to the server as a root user.
2. Navigate to UNIX file path `/etc/security/limits.conf` to edit the file.
3. Add the following values at the end of the file for Compliance Studio:

   ```
   <Username> hard nproc 65536
   <Username> soft nproc 65536
   ```

   For example,

   ```
   compliancestudio hard nproc 65536
   compliancestudio soft nproc 65536
   ```

3.1.2 Prerequisite Environmental Settings

The following prerequisite environmental settings must be set before beginning the installation of Compliance Studio. These settings are the configuration that a system must have for an application to run smoothly and efficiently.

Table 7 lists the Prerequisite Environmental Settings:

Table 7: Prerequisite Environmental Settings

Category	Expected Value
Java Settings	`PATH` in the `.profile` file must be set to include the Java Runtime Environment (Java 8) absolute path. **Supported version:** jdk 1.8.0 **NOTE**: Ensure the absolute path to `JRE/bin` is set at the beginning of the PATH variable. For example: `PATH=/usr/java/jre1.8/bin:$PATH` Ensure no SYMBOLIC links to Java installation are set in the PATH variable.

PGX Server	The following packages must be installed or present in the server where the PGX service is installed:
	`krb5-libs`
	`krb5-workstation`
	`procps-ng`
	`nc`
	Execute the following command to install the packages as mentioned above:
	`yum install -y krb5-libs krb5-workstation procps-ng nc`

Oracle Database Settings	**Oracle Processing Server** `ORACLE_HOME` must be set in the .profile file pointing to the appropriate Oracle DB Client installation. `PATH` in the `.profile` file must be set to include the appropriate `$ORACLE_HOME/bin` directory.

Table 7: Prerequisite Environmental Settings

Category	Expected Value
Download Directory	Indicates the directory where the product installer zip file is downloaded or copied. The user permission must be set to 755 for this directory.
Installation Directory	Indicates the directory where the product installer zip file is extracted, and the installation files are placed. The user permission must be set to 755 for this directory. **NOTE:** The Installation and the Download Directory can be the same if the product installer zip file is not copied separately to another directory.
OS Locale	Linux: `en_US.utf8` Execute the following command to check the locale: `locale -a | grep -i 'en_US.utf'` The locale is displayed.
Oracle Instant client	Install oracle instant client in the server where compliance Studio is installed and provide the configuration `LD_LIBRARY_PATH` in `config.sh`

3.1.3 Download the Big Data Files

Download the following configuration files from the Big Data server or contact the Big Data Administrator:

> **NOTE** These files must be kept ready and provided in the following file structure used during Compliance Studio installation.

Table 8 lists the required file structure:

Table 8: Required File Structure

File Category	File Names
Hadoop Cluster	• `core-site.xml` • `hive-env.sh` • `hive-site.xml` • `hadoop-env.sh` • `hdfs-site.xml` • `mapred-site.xml` • `yarn-site.xml` • `redaction-rules.json` • `log4j.properties` • `ssl-client.xml` • `topology.map` • `topology.py`

Table 8: Required File Structure

Kerberos Files	- `krb5.conf` - keytab file name as mentioned in the `config.sh` file.
Additional Jars	- `hive-exec-*.jar.` - `HiveJDBC4.jar.` - `hive-metastore-*.jar.` - `hive-service-*.jar.` **NOTE**: - The version of the jars is client or user-specific. These jars can be obtained from the existing jars of the Cloudera installation. - The `HiveJDBC4.jar` file is not available in the Cloudera installation setup. You must download the same from the Cloudera website. This is applicable only for Cloudera Cluster. - For additional jars, see the Appendix C – Additional Jars – PGX and Appendix D – Additional Jars – Batch Service.
ES-Hadoop Jars	`elasticsearch-spark-20_2.11-7.14.jar` To download the `elasticsearch-spark-20_2.11-7.14.jar` file, follow these steps: 1. Download the ZIP file from Elasticsearch 7.14 2. Extract the downloaded file. 3. Navigate to the dist directory and download the `elasticsearch-spark-20_2.11-7.14.jar` **NOTE:** The version should be the same as the Elastic Search version.

3.1.4 Validation Checklist

The Validation Checklist section provides you with the parameters that you can validate to avoid installation issues. This section explains the validation and actions that can be taken for some of the common parameters that are used in the `config.sh` file for the installation. The parameters that can be validated are as follows:

Table 9 lists the required file structure:

Table 9: Required File Structure

Parameters	Validation
External Service (OFSAA_SERVICE URL)	The OFSAA_Service URL can be validated by clicking the URL for verification.
DB Details for Studio Schema	You can log in to SQL developer and verify the DB Details for Studio Schema.
Compliance Studio Schema Wallet Details	You can verify the Wallet details by reviewing the steps in Verify the Connectivity of the Wallet.

Table 9: Required File Structure

Atomic Wallet Detail	You can verify the Wallet details by reviewing the steps in Setup Password Stores with Oracle Wallet.
SQL Scripts	You can log in to Compliance Studio using SQL developer and validate the **Studio_DBLINK_BD**. If the link type is DBLINK, if Schema is not DBLINK, there is no validation required.
Cloudera	You can verify the Cloudera details and validate them by reviewing the steps in Create the Credential Keystore.
Cloudera (SSH Connection)	Run the command `ssh <hostname of the Cloudera machine>`. You must run this command from the host where the Studio is installed.
Cloudera (Keytab)	Run the command `kinit -V <KERBEROS_PRINCIPAL> -k -t <KEYTAB_FILEPATH>` to verify the keytab.

3.1.5 Configure the Elastic Search Component

To configure the Elastic Search component, follow these steps:

> **NOTE**
> - Ensure that a minimum of 4GB free RAM space is available for elastic search. If RAM is low, the shards of the elastic search fail, and the correct result is not fetched.
> - You must manually clean the cache if facing a performance issue.
> - As a prerequisite, download the `analysis-icu-<Elastic Search Version>.zip` from Elastic Search official website.

1. Navigate to the `<Elastic search installed path>/config` directory.
2. Configure the `elasticsearch.yml` with the following variables:

Table 10 lists the parameters of `elasticsearch.yml` file:

Table 10: Elasticsearch.yml File

Interaction Variable Name	Significance
cluster.name	Indicates the name of the cluster.
node.name	Indicates the name given for the node.

Table 10: Elasticsearch.yml File

node.master	Indicates whether the node is a master.
node.data	Indicates the node data.
path.data	Indicates the directory where you want to store the data.
path.logs	Indicates the directory where you want to store the logs.
network.host	Indicates the hostname of the machine where you want to install the elastic search service.
http.port	Indicates the port number where the elastic search service is installed.
discovery.seed_hosts	(Optional) Indicates the hostnames of the nodes of the cluster.
cluster.initial_master_nodes	(Optional) Indicates the number given to the nodes of the cluster.
indices.breaker.total.use_real_memory	● Indicates the static setting to determine whether the parent breaker must consider the real memory usage or only consider the amount reserved by the child circuit breakers. ● This setting is used to prevent the OutOfMemory error.

3. Configure the `jvm.options` file as follows:

   The following table lists Interaction variable names for Configure jvm.options File

Table 11: Configure jvm.options File

Interaction Variable Name	Significance
-Xms1g	● Set the value for these parameters. ● The maximum value set can be up to 50% of the RAM size of the machine. ● Recommended value: Less than 32GB.
-Xmx1g	

4. Unzip the `analysis-icu-<Elastic Search Version>.zip` and copy to `<Elastic Search Home>/plugins`.

5. Enter the URL in the following format into the browser:

   ```
   http://<network.host>:<http.port>
   ```

   The following output is displayed to indicate the successful installation of the Elastic Search service.

   ```
   {
     "name" : "node-1",
     "cluster_name" : "my-application",
     "cluster_uuid" : "<Cluster UUID>",
     "version" : {
       "number" : "7.13.4",
       "build_flavor" : "default",
       "build_type" : "tar",
       "build_hash" : "c5f60e894ca0c61cdbae4f5a686d9f08bcefc942",
       "build_date" : "2021-07-14T18:33:36.673943207Z",
       "build_snapshot" : false,
       "lucene_version" : "8.8.2",
       "minimum_wire_compatibility_version" : "6.8.0",
       "minimum_index_compatibility_version" : "6.0.0-beta1"
     },
     "tagline" : "You Know, for Search"
   }
   ```

3.1.5.1 Enable or Disable HTTPS and Authentication for Elastic Search

To enable the HTTPS and Authentication, ensure the below codes lines are not commented (remove **#** symbol at the starting of line) in `elasticsearch.yml`.

To disable the HTTPS and Authentication, ensure the below codes lines are commented (add # symbol at the starting of line) in `elasticsearch.yml`.

3.1.5.1.1 Enable HTTPS and Authentication

1. Navigate to `<Elastic Search Installation Path>/config/elasticsearch.yml`.
2. Verify the below code lines if anything commented, if yes, remove it.

   ```
   xpack.security.enabled: true
   xpack.security.http.ssl.enabled: true
   xpack.security.transport.ssl.enabled: true
   xpack.security.http.ssl.key: certs/node-1.key
   xpack.security.http.ssl.certificate: certs/node-1.crt
   xpack.security.http.ssl.certificate_authorities: certs/ca.crt
   xpack.security.transport.ssl.key: certs/node-1.key
   ```

```
xpack.security.transport.ssl.certificate: certs/node-1.crt
xpack.security.transport.ssl.certificate_authorities: certs/ca.crt
```

3.1.5.1.2 Disable HTTPS and Authentication

1. Navigate to `<Elastic Search Installation Path>/config/elasticsearch.yml`.
2. Verify the below code lines and comment all as shown below:

```
#xpack.security.enabled: true
#xpack.security.http.ssl.enabled: true
#xpack.security.transport.ssl.enabled: true
#xpack.security.http.ssl.key: certs/node-1.key
#xpack.security.http.ssl.certificate: certs/node-1.crt
#xpack.security.http.ssl.certificate_authorities: certs/ca.crt
#xpack.security.transport.ssl.key: certs/node-1.key
#xpack.security.transport.ssl.certificate: certs/node-1.crt
#xpack.security.transport.ssl.certificate_authorities: certs/ca.crt
```

3.1.5.1.3 Enable HTTPS and Disable Authentication

1. Navigate to `<Elastic Search Installation Path>/config/elasticsearch.yml`.
2. Verify the below code lines, and add comment (#) to the required code to disable authentication as shown below:

```
#xpack.security.enabled: true
xpack.security.http.ssl.enabled: true
#xpack.security.transport.ssl.enabled: true
xpack.security.http.ssl.key: certs/node-1.key
xpack.security.http.ssl.certificate: certs/node-1.crt
xpack.security.http.ssl.certificate_authorities: certs/ca.crt
#xpack.security.transport.ssl.key: certs/node-1.key
#xpack.security.transport.ssl.certificate: certs/node-1.crt
#xpack.security.transport.ssl.certificate_authorities: certs/ca.crt
```

3.1.5.1.4 Disable HTTPS and Enable Authentication

1. Navigate to `<Elastic Search Installation Path>/config/elasticsearch.yml`.
2. Verify the below code lines, and add comment (#) to the required code to disable HTTPS as shown below:

```
xpack.security.enabled: true
#xpack.security.http.ssl.enabled: true
xpack.security.transport.ssl.enabled: true
#xpack.security.http.ssl.key: certs/node-1.key
```

```
#xpack.security.http.ssl.certificate: certs/node-1.crt
#xpack.security.http.ssl.certificate_authorities: certs/ca.crt
xpack.security.transport.ssl.key: certs/node-1.key
xpack.security.transport.ssl.certificate: certs/node-1.crt
xpack.security.transport.ssl.certificate_authorities: certs/ca.crt
```

3.1.5.2 Cleanup of Elastic Search Indexes

To clean up the Elastic Search indexes, run the following command:

```
curl -XDELETE http://<FULLY QUALIFIED HOSTNAME>:<PORT of Load To Elastic
Search Service>/load-to-elastic-search/idx/deleteIndex/<INDEX NAME>
```

For example,

```
curl -XDELETE http://testserver.in.oracle.com:7053/load-to-elastic-
search/idx/deleteIndex/test_index
```

3.1.5.3 Generate truststore File for Elasticsearch

NOTE This section is applicable only when https and authentication are enabled.

To generate file for Elasticsearch, follow these steps:

1. Run the following jks command in the Studio Server:

   ```
   keytool -import -alias elasticCA -file
   <path_to_elasticsearch_ca_crt_file> -keystore
   <path_to_save_elastic.jks_file>
   ```

   For example,

   ```
   keytool -import -alias elasticCA -file /scratch/elastic/Elasticsearch/
   elasticsearch/config/certs/ca/ca.crt -keystore /scratch/elastic/
   Elasticsearch/elastic.jks
   ```

2. Specify the keystore password.
3. Execute the following command in the studio server to generate the `.crt` certificate:

   ```
   keytool -importcert -keystore <path_to_elastic.jks_file> -alias
   <alias_name> -file <path_to_node_crt>
   ```

   For Example,

   ```
   keytool -importcert -keystore /scratch/elastic/Elasticsearch/elastic.jks
   -alias myEsNode -file /scratch/elastic/Elasticsearch/elasticsearch/
   config/certs/node-1/node-1.crt
   ```

4. Specify the keystore password.
5. When generating the keytool ensure to provide the hostname in the first name. For example:

   Question: What is your first and last name?

   Answer: Provide the fully qualified studio server hostname.

   For example, <hostname>.<domain name>

6. Specify any name for the other questions.

7. Specify the keystore password. The `jks` file is created in the Studio Server.

> **NOTE** You must use the same password and alias that is provided in the `config.sh` file.

3.1.6 Configure Logstash

1. Download the Logstash tar file compatible with your Elastic Search version. For example, if the Elastic Search version is 7.14.0, the Logstash version should also be 7.14.0.

   You can download logstash from the official website:
2. Untar the tar file in one of the Server locations where you are installing Compliance Studio.
3. Provide this path as Logstash_Home in `config.sh` file.
4. Create a folder "Logstash" under CS install path.
5. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/Logstash`
6. Untar the contents of the tar file.
7. Provide this folder path for the parameter "Logstash_Home" in `config.sh` file. The Compliance Studio installer will automatically configure the Logstash properties where necessary.

> **NOTE** The ca.crt file should be copied from the elastic search server into the `Logstash_Home/config` path when https is enabled in elastic search.

3.1.7 Installing Analytics ICU Plugin

To install the Analytics ICU plugin, perform the following.

1. To Obtain the ICU plugin, follow these steps:
   - Run the following command to download:

   `wget` https://artifacts.elastic.co/downloads/elasticsearch-plugins/analysis-icu/analysis-icu-<version>.zip

   `Example:`

   `wget` https://artifacts.elastic.co/downloads/elasticsearch-plugins/analysis-icu/analysis-icu-7.9.2.zip

   - You can also download the required version from the browser.
2. Navigate to `<Elastic Search Installation Path>`.

   For example,

   ```
   elasticsearch-<version>
   ```
3. Run the following command to install the plugins:

   ```
   elasticsearch-<version>/bin/elasticsearch-plugin install file:///
   <ElasticSearch Installation Path>/analysis-icu-<version>.zip
   ```

   ```
   Example:
   ```

   ```
   elasticsearch-7.14/bin/elasticsearch-plugin install file:///
   <ElasticSearch Installation Path>/analysis-icu-7.14.zip
   ```

3.1.8 Configure the Interpreter Settings

Before installing Compliance Studio, you must configure the interpreter settings for the following interpreters.

> **NOTE** Ensure to configure the settings only for the interpreters that you require.

Table 12 lists the Pre-installation Interpreter Settings:

Table 12: Pre-installation Interpreter Settings

Interpreter	Prerequisite Settings
jdbc	No additional configuration is required.
md	No additional configuration is required.
pgql	No additional configuration is required.
pgx-algorithm	No additional configuration is required.
pgx-java	No additional configuration is required.
pgx-python	No additional configuration is required. You can point to any other python virtual environment.
pyspark	For the required configuration, see Configure the PySpark Interpreter.
spark	For the required configuration, see Configure the Spark Interpreter.
fcc-python	No additional configuration is required.
ore	The ore Interpreter has been deprecated. It is recommend using this interpreter since it will be removed in future versions of OFS Compliance Studio. It will be introducing “R” Interpreter instead of ore Interpreter.

3.1.8.1 Configure the PGX Interpreter

- To update the bundled JDK, see the How to update the bundled JDK version? in the Frequently Asked Questions in Compliance Studio.
- To use system's JDK instead of bundled JDK, see the How to use the system's JDK 8 instead of bundled JDK? in the Frequently Asked Questions in Compliance Studio.

3.1.8.2 Configure the jdbc Interpreter

To create the context for the jdbc interpreter, follow these steps:

1. Log in to Oracle Database as an SYSDBA user.
2. Grant Execute permission to the user using the following command:

   ```
   GRANT execute dbms_rls to <Compliance Studio_DB_Username>;
   ```

   The Execute permission is granted to the user.
3. Grant Create permission to the context using the following command:

   ```
   GRANT create any context to <Compliance Studio_DB_Username>;
   ```

   The Create permission is granted to context.

3.1.8.3 Configure the Spark Interpreter

3.1.8.3.1 Prerequisites for using the Spark Interpreter

To configure Spark Interpreter, you must download the desired spark distribution from Spark's official website.

For example, spark-2.4.0-bin-hadoop2.7.tgz from the website.

Configure the Spark Interpreter can be used in several situations as follows:

- Connecting to remote spark cluster
  - With/without Kerberos
  - Custom Hadoop client configuration
  - Custom libraries
- Spark in local mode.

In case you want to connect to a remote spark cluster, then obtain the following files:

- Hadoop or Hive client configuration as per your use case
- Kerberos files (if applicable)
  - `krb5.conf`
  - `keytab file`

3.1.8.3.2 Setting up spark-interpreter

The spark interpreter requires spark distribution to start. If you do not intend to use a spark-interpreter, disable the interpreter by performing the following steps:

> **NOTE** If you do not intend to use a spark interpreter, disable the interpreter. You can perform the following steps.

1. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory.
2. Open the `config.sh` file and set `export SPARK_ENABLED=`**`false`**.

The default configuration of Spark is configured for yarn-client deployment mode with Kerberos enabled remote spark cluster. For local mode, skip this section and follow below.

3.1.8.3.3 Spark Interpreter with remote spark cluster

The Spark Interpreter with remote spark cluster can be performed for the following:

- Configuration with Kerberos enabled remote spark cluster:
  - **krb5.conf**
  - ***.keytab** (For example, **fccstudio.keytab**)
- Configuration with Kerberos disabled remote spark cluster

3.1.8.3.3.1 Configuration with Kerberos enabled remote spark cluster

1. Move the obtained Kerberos files to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/user/conf` directory.

   NOTE These are the same Kerberos files used for ETL.

2. Place the `spark-<version>-bin-hadoop<version>` files to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/interpreter/spark/extralibs` directory.

   For example, `spark-2.4.0-bin-hadoop2.7`
3. Create a **conf** folder in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/interpreter/spark/extralibs.`
4. Place the Hadoop or Hive client configuration files to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/interpreter/spark/extralibs/conf` directory.

   NOTE Do not remove the `spark-env.sh` file.

5. Create spark-default.conf and update the spark configurations accordingly. See the Sample spark-default.conf Configuration File section for more information.
6. Update `spark.yarn.dist.files` and `spark.executorEnv.PYTHONPATH`.

   NOTE The path must be the same as the path given for the downloaded spark distribution. For example: path for `spark-2.4.0-bin-hadoop2.7spark` distribution.

7. Update the `spark.driver.host.`
8. Update the `spark.yarn.keytab.`
9. Update the `spark.yarn.principal.`
10. If required, you can add an additional spark configuration.

3.1.8.3.3.2 *Configuration with Kerberos disabled remote spark cluster:*

1. Place the Hadoop or Hive client configuration files to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/interpreter/spark/extralibs/conf` directory.

   > **NOTE** Do not remove the `spark-env.sh` file.

2. Create `spark-default.conf` and update the spark configurations accordingly. See the Sample spark-default.conf Configuration File section for more information.
3. Update `spark.yarn.dist.files` and `spark.executorEnv.PYTHONPATH`.

   > **NOTE** The path must be the same as the path given for the downloaded spark distribution. For example: path for spark-2.4.0-bin-hadoop2.7 spark distribution.

4. Update the `spark.driver.host.`
5. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ interpreters/interpreter/spark/extralibs/conf` directory.
6. Open `spark-default.conf` file and update spark.driver.defaultJavaOptions by removing:

```
"-Dsun.security.krb5.debug=false -
Djavax.security.auth.useSubjectCredsOnly=false -
Djava.security.krb5.conf=<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/
batchservice/user/conf/krb5.conf"
```

3.1.8.3.4 Spark Interpreter in local mode

1. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ interpreters/interpreter/spark/extralibs/conf` directory.
2. Create `spark-default.conf` and update the spark configurations accordingly. See the Sample spark-default.conf Configuration File section for more information.
3. Open `spark-default.conf` file and update spark.driver.defaultJavaOptions by removing:

```
"-Dsun.security.krb5.debug=false -
Djavax.security.auth.useSubjectCredsOnly=false -
Djava.security.krb5.conf=<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/
batchservice/user/conf/krb5.conf"
```

4. Set `spark.master as local[*]` in interpreter configuration file.

3.1.8.3.5 Configuration

The Spark interpreter configuration can be divided into the following categories:

- Configuration related to deployment

  These properties can be set either in the Spark libraries, for example, the `spark-defaults.conf` file, or through the system environment variable, **SPARK_CONF**.

For example, `SPARK_CONF="--conf spark.driver.memory=2g"`.

> **NOTE** These properties cannot be changed when the Spark interpreter is running.

- Configuration related to Spark runtime control

  These properties can be set from the Interpreters page of the Compliance Studio application UI. This includes properties such as a `spark.executor.memory`.

> **NOTE** The properties related to the driver cannot be set during runtime and are considered deployment configuration. The properties related to the executors can be set during runtime. Hence, the latter option of runtime control configuration is preferred.

A list of possible properties is available in the Spark Official Documentation. All the properties prefixed with the term "zeppelin" listed in the Zeppelin Spark Configuration Document can also be set from the Interpreters page of the Compliance Studio application UI.

3.1.8.4 Configure the PySpark Interpreter

Compliance Studio uses PySpark 2.4.0. Before you begin the configurations, check the prerequisites depending on your operation mode.

3.1.8.4.1 Prerequisites

The PySpark interpreter has the same prerequisites as that as the Spark Interpreter. For more information, see Configure the Spark Interpreter. Also, all Spark components must be configured to use the same Python version.

3.1.8.4.2 Configuration

The PySpark interpreter can be configured through the Spark interpreter, with the only exception being the Python version used. By default, the Python version is set to 3 that can be changed either in the interpreter JSON files before the startup or from the **Interpreters** page of the Compliance Studio application UI during runtime by changing the following properties:

- In the **Spark Interpreter Settings** page of the Compliance Studio application UI (or `spark.json` file), change the value of the `spark.pyspark.python` property to the Python executable path that is to be used by the Spark executors.
- In the **PySpark Interpreter Settings** page of the Compliance Studio application UI (or `pyspark.json` file), change the value of the `zeppelin.pyspark.python` property to the Python executable path that is to be used by the Spark driver.

3.1.8.4.3 Use the Python Virtual Environments with PySpark

To ensure that the two Python versions match, in case your components run on different machines, you must use the Python virtual environments with PySpark.

To use Python Virtual Environments with PySpark, follow these steps:

1. Create a Virtual Environment with Conda
2. Update the Interpreter Properties

3.1.8.4.3.1 Create a Virtual Environment with Conda

> **NOTE** You can also use **virtualenv** to create your virtual environment instead of **conda**.

To create a virtual environment with Conda, follow these steps:

1. Ensure that you have conda and conda-Pack installed.
2. Create your virtual environment using the following command:

   ```
   conda create -y -n <environment-name> python=<python-version>
   <additional-packages>
   ```

   > **NOTE** The <environment-name> can be chosen freely and subsequently has to be substituted in further commands.

3. Activate your virtual environment using the following command:

   ```
   conda activate <environment-name>
   ```

4. Execute the following to obtain the path to your virtual environment:

   ```
   which python
   ```

   The obtained result is referred to as `<environment-abs-path>.`

5. Compress your virtual environment using the following command:

   ```
   conda pack -n <environment-name> -o <environment-abs-path>/<environment-
   name>.tar.gz
   ```

3.1.8.4.3.2 Update the Interpreter Properties

The interpreter properties can either be configured in the interpreter JSON files or from the Interpreters page of the Compliance Studio application UI after starting the Compliance Studio application.

- In the **Spark Interpreter Settings** page of the Compliance Studio application UI (or `spark.json`), change the following:
  - Change the value of the `spark.yarn.dist.archives` property to `<environment-abs-path>/<environment-name>.tar.gz#<environment-name>`
  - Change the value of the `spark.pyspark.python property to ./<environment-name>/bin/python`
- In the **PySpark Interpreter Settings** page of the Compliance Studio application UI (or `pyspark.json`), change the value of the `zeppelin.pyspark.python` parameter to `<environment-abs-path>/bin/python`.

3.1.9 Create the Hive Schema

To create a hive schema, perform the following steps:

1. Login to the server where **cloudera/hive** is installed.
2. Open a hive session in the command prompt.

   ```
   hive
   ```

3. Create a new hive schema using the following script:

   ```
   create database <hive schema name>;
   ```

4. Use the hive schema that is created by the following command:

   ```
   Use <hive schema name>
   ```

A new hive schema is created.

3.1.10 Create the Tablespace

To create a tablespace in the Oracle Database using the script as described in the Table 13.

Table 13: Create Tablespace

User	Script
AIF_USER_TEMP_TS	CREATE TABLESPACE AIF_USER_TEMP_TS DATAFILE '<Datafile Path>' SIZE <size in byte> REUSE AUTOEXTEND ON NEXT <size in megabyte> MAXSIZE UNLIMITED;
AIF_USER_TS	CREATE TABLESPACE AIF_USER_TS DATAFILE '<Datafile Path>' SIZE <size in byte> REUSE AUTOEXTEND ON NEXT <size in megabyte> MAXSIZE UNLIMITED;
<CS_USER_TS>	CREATE TABLESPACE <CS_USER_TS> DATAFILE '<Datafile Path>' SIZE <size in byte> REUSE AUTOEXTEND ON NEXT <size in megabyte> MAXSIZE UNLIMITED;

NOTE The tablespace size should be defined based on the size of the data.

3.1.11 Create the Studio Schema

To create a studio schema, create a new Oracle Database schema user using the following script:

```
CREATE USER <Compliance Studio Schema User Name> IDENTIFIED BY <Password>
DEFAULT TABLESPACE <Studio Tablespace>;

ALTER USER <SCHEMA USER> QUOTA 2000M ON <STUDIO TABLESPACE>;

ALTER USER <SCHEMA USER> QUOTA <size in megabyte> ON AIF_USER_TS;
```

For example;

```
ALTER USER CS812_USER QUOTA 500M ON AIF_USER_TS;
```

NOTE The tablespace and quota sizes should be defined based on the size of the data.

A new Oracle Database schema (Studio schema) is created.

3.1.12 Assign Grants for the Studio Schema

Grant the following permissions to the newly created Oracle Database studio schema:

- `GRANT CREATE SESSION TO <FSDF Schema>;`
- `GRANT CREATE TABLE TO <FSDF SCHEMA>;`
- `GRANT CREATE VIEW TO <FSDF SCHEMA>;`
- `GRANT CREATE ANY TRIGGER TO <FSDF SCHEMA>;`
- `GRANT CREATE ANY PROCEDURE TO <FSDF SCHEMA>;`
- `GRANT CREATE SEQUENCE TO <FSDF SCHEMA>;`
- `GRANT CREATE SYNONYM TO <FSDF SCHEMA>;`
- `GRANT CREATE RULE TO <FSDF SCHEMA>;`
- `GRANT CREATE JOB TO <FSDF SCHEMA>;`
- `GRANT CREATE MATERIALIZED VIEW TO <FSDF SCHEMA>;`
- `GRANT DROP ANY TRIGGER TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON DBMS_LOCK TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON DBMS_STATS TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON DBMS_RLS TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON SYS.DBMS_SESSION TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON DBMS_REDEFINITION TO <FSDF SCHEMA>;`
- `GRANT REDEFINE ANY TABLE TO <FSDF SCHEMA>;`
- `GRANT SELECT ON SYS.V_$PARAMETER TO <FSDF SCHEMA>;`
- `GRANT SELECT ON SYS.DBA_FREE_SPACE TO <FSDF SCHEMA>;`
- `GRANT SELECT ON SYS.DBA_TABLES TO <FSDF SCHEMA>;`
- `GRANT SELECT ON SYS.DBA_TAB_COLUMNS TO <FSDF SCHEMA>;`
- `GRANT SELECT ON SYS.DBA_RECYCLEBIN TO <FSDF SCHEMA>;`
- `GRANT EXECUTE ON CTXSYS.CTX_DDL TO <FSDF Schema>;`

3.1.13 Create the Sandbox Schema

To create a sandbox schema, create a new Oracle Database sandbox schema user using the following script:

```
create user <USER_NAME>

IDENTIFIED BY <password>

default tablespace AIF_USER_TS

temporary tablespace TEMP

profile DEFAULT

quota unlimited on AIF_USER_TS
quota <size in megabyte> on <USER_NAME>;
```

> **NOTE**
>
> - The sandbox will always be on a different database other than the production schema.
> - The tablespace and quota sizes should be defined based on the size of the data.

A new Oracle Database schema (Sandbox schema) is created.

3.1.14 Assign Grants for the Sandbox Schema

Grant the following permissions to the newly created Oracle Database sandbox schema:

- `GRANT CONNECT, RESOURCE, DBA TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE SESSION TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE PROCEDURE TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE SEQUENCE TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE TABLE TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE TRIGGER TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE VIEW TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE MATERIALIZED VIEW TO <SANDBOX SCHEMA USER>;`
- `GRANT SELECT ON SYS.V_$PARAMETER TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE SYNONYM TO <SANDBOX SCHEMA USER>;`
- `GRANT SELECT ON SYS.V_$PARAMETER TO <SANDBOX SCHEMA USER>;`
- `GRANT SELECT ON SYS.DBA_FREE_SPACE TO <SANDBOX SCHEMA USER>;`
- `GRANT SELECT ON SYS.DBA_TABLES TO <SANDBOX SCHEMA USER>;`
- `GRANT SELECT ON SYS.DBA_TAB_COLUMNS TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE RULE TO <SANDBOX SCHEMA USER>;`
- `GRANT CREATE ANY TRIGGER TO <SANDBOX SCHEMA USER>;`
- `GRANT DROP ANY TRIGGER TO <SANDBOX SCHEMA USER>;`

3.1.15 Entity Resolution

3.1.15.1 Create Entity Resolution Schema and Grant Permission

To create ER schema, create a new Oracle Database schema user using the following script:

```
CREATE USER <ER SCHEMA USERNAME> IDENTIFIED BY <PASSWORD>;
```

A new Oracle Database schema (ER schema) will be created.

To assign grants, see the Assign Grants for the Studio Schema section.

3.1.15.2 Create a wallet for ER schema

See **step 4** in the Setup the Password Stores for Database User Accounts section.

NOTE

- ER schema can be in the same database where CS is installed or a different database.
- You can create multiple ER schemas.

3.1.15.3 Configure Resource XML

See the Configure the resources.xml for Multiple ER Schemas section for more details.

3.1.15.4 Configure ER schema Profile

Set the SESSIONS_PER_USER limit to UNLIMITED for ER Schema by executing the below steps:

1. Get the ER schema profile by executing the below query:

```
select profile from dba_users where username ='<ER Schema User>’;
```

2. Change the profile which is obtained from the step 1 by executing the below query:

```
ALTER PROFILE <profile> LIMIT SESSIONS_PER_USER UNLIMITED;
```

3.2 Setup Password Stores with Oracle Wallet

As part of an application installation, administrators must set up password stores for database user accounts using Oracle Wallet. These password stores must be installed on the application database side. The installer handles much of this process. The administrators must perform some additional steps.

A password store for the application and application server user accounts must also be installed; however, the installer takes care of this entire process.

Topics:

- Setup the Password Stores for Database User Accounts
- Verify the Connectivity of the Wallet
- Create the Credential Keystore
- Download the Installer Kit

3.2.1 Setup the Password Stores for Database User Accounts

After the database is installed and the default database user accounts are set up, administrators must set up a password store using the Oracle Wallet. This involves assigning an alias for the username and associated password for each database user account. The alias is used later during the application installation. This password store must be created on the system where the application server and database client are installed.

This section describes the steps to create a wallet and the aliases for the database user accounts. For more information on configuring authentication and password stores, see Oracle Database Security Guide.

> **NOTE** In this section, `<wallet_location>` is a placeholder text for illustration purpose. Before running the command, ensure that you have already created the <wallet_location> directory where you want to create and store the wallet.

To create a wallet, follow these steps:

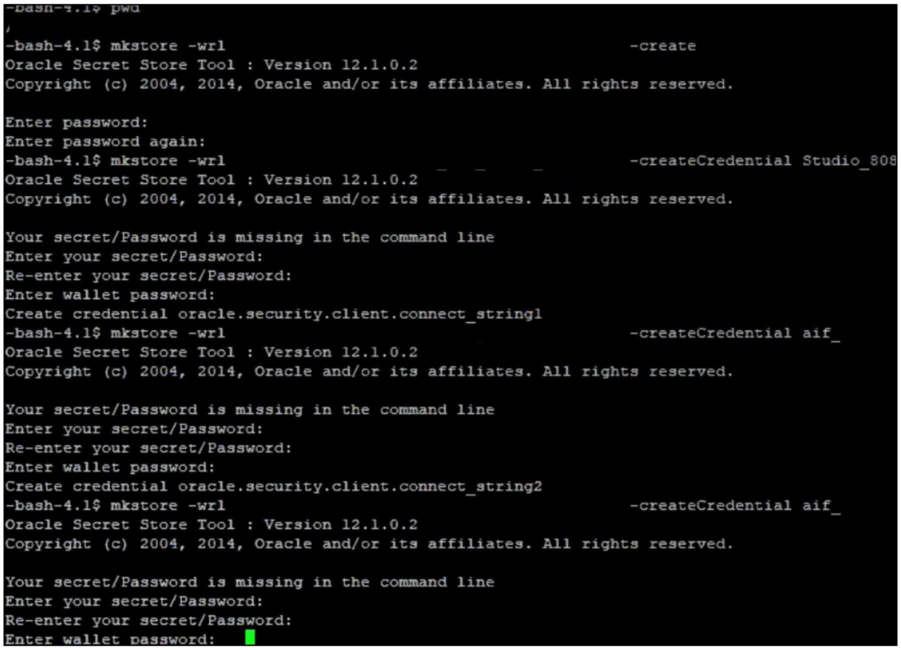
1. Log in to the server as a Linux user.
2. Create a wallet in the <wallet_location> using the following command:

   ```
   mkstore -wrl <wallet_location> -create
   ```

> **NOTE** The mkstore utility is included in the Oracle Database Client installation.

3. After you run the command, a prompt appears. Enter a password for the Oracle Wallet in the prompt.

   Figure 1: Wallet Creation



```
-bash-4.1$ pwd
/
-bash-4.1$ mkstore -wrl                                        -create
Oracle Secret Store Tool : Version 12.1.0.2
Copyright (c) 2004, 2014, Oracle and/or its affiliates. All rights reserved.

Enter password:
Enter password again:
-bash-4.1$ mkstore -wrl                                        -createCredential Studio_808
Oracle Secret Store Tool : Version 12.1.0.2
Copyright (c) 2004, 2014, Oracle and/or its affiliates. All rights reserved.

Your secret/Password is missing in the command line
Enter your secret/Password:
Re-enter your secret/Password:
Enter wallet password:
Create credential oracle.security.client.connect_string1
-bash-4.1$ mkstore -wrl                                        -createCredential aif_
Oracle Secret Store Tool : Version 12.1.0.2
Copyright (c) 2004, 2014, Oracle and/or its affiliates. All rights reserved.

Your secret/Password is missing in the command line
Enter your secret/Password:
Re-enter your secret/Password:
Enter wallet password:
Create credential oracle.security.client.connect_string2
-bash-4.1$ mkstore -wrl                                        -createCredential aif_
Oracle Secret Store Tool : Version 12.1.0.2
Copyright (c) 2004, 2014, Oracle and/or its affiliates. All rights reserved.

Your secret/Password is missing in the command line
Enter your secret/Password:
Re-enter your secret/Password:
Enter wallet password:
```

4. Create the database connection credentials for the studio schema/ER Schema alias using the following command:

   ```
   mkstore -wrl <wallet_location> -createCredential <alias-name> <database-user-name>
   ```

   After you run the command, a prompt appears. Enter the password associated with the database user account in the prompt. You are prompted to re-enter the password. You are prompted for the wallet password used in Step 1.

5. Create the database connection credentials for the atomic schema alias using the following command:

```
mkstore -wrl <wallet_location> -createCredential <alias-name> <database-
user-name>
```

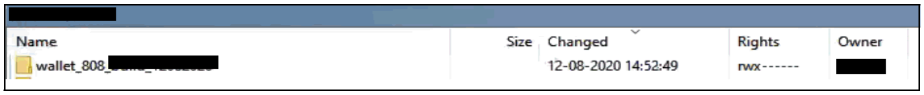
NOTE Creating an atomic schema is not required when installing Compliance Studio without OFSAA.

In this manner, create a wallet and associated database connection credentials for all the database user accounts.

The wallet is created in the <wallet_location> directory with the auto-login feature enabled. This feature enables the database client to access the wallet contents without using the password. For more information, see Oracle Database Security Guide.

After the wallet is created, go to the <wallet_location> directory and click **Refresh** to view the created wallet folder.

Figure 2: Location of the Created Wallet Folder



[redacted]

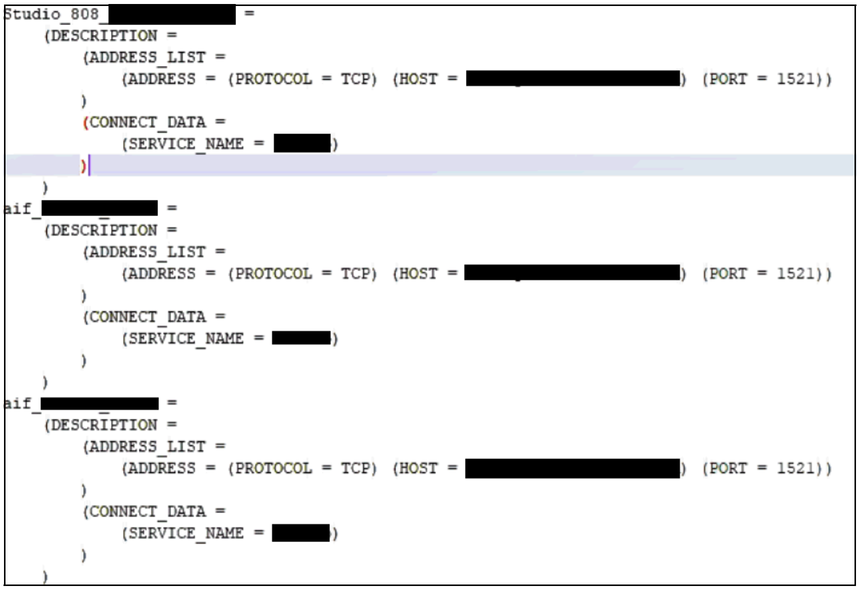
Name	Size	Changed	Rights	Owner
wallet_808_[redacted]		12-08-2020 14:52:49	rwx------	[redacted]

The wallet folder contains two files: **ewallet.p12** and **cwallet.sso**.

6. Move the wallet folder to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/<alias-name>` directory.

7. In the `<wallet_location>` directory, configure the tnsnames.ora file to include the entry for each alias name to be set up.

Figure 3: Location of the Created Wallet Folder



```
Studio_808_[redacted] =
    (DESCRIPTION =
        (ADDRESS_LIST =
            (ADDRESS = (PROTOCOL = TCP) (HOST = [redacted]) (PORT = 1521))
        )
        (CONNECT_DATA =
            (SERVICE_NAME = [redacted])
        )
    )
aif_[redacted] =
    (DESCRIPTION =
        (ADDRESS_LIST =
            (ADDRESS = (PROTOCOL = TCP) (HOST = [redacted]) (PORT = 1521))
        )
        (CONNECT_DATA =
            (SERVICE_NAME = [redacted])
        )
    )
aif_[redacted] =
    (DESCRIPTION =
        (ADDRESS_LIST =
            (ADDRESS = (PROTOCOL = TCP) (HOST = [redacted]) (PORT = 1521))
        )
        (CONNECT_DATA =
            (SERVICE_NAME = [redacted])
        )
    )
```

> **NOTE**
>
> - You can either update the existing tnsnames.ora file with the above details or create new tnsnames.ora file and add the required entries.
> - <alias-name> is a user-defined value.

8. Create a sqlnet.ora file in the wallet directory using the following content:

```
WALLET_LOCATION = (SOURCE = (METHOD = FILE) (METHOD_DATA = (DIRECTORY =
<Wallet_Location>)) )

SQLNET.WALLET_OVERRIDE=TRUE

SSL_CLIENT_AUTHENTICATION=FALSE
```

3.2.2 Verify the Connectivity of the Wallet

To verify the connectivity of the wallet, follow these steps:

1. Create a sqlnet.ora file in the wallet directory using the following content:

```
WALLET_LOCATION = (SOURCE = (METHOD = FILE) (METHOD_DATA = (DIRECTORY =
<Wallet_Location>)) )

SQLNET.WALLET_OVERRIDE=TRUE

SSL_CLIENT_AUTHENTICATION=FALSE
```

2. Test the connectivity using the following command:

> **NOTE** The ORACLE_HOME used with the wallet must be the same version or higher than the wallet created.

```
$ export WALLET_LOCATION=<wallet_location>

$ export TNS_ADMIN=<tnsnames.ora_location>. If you have created a new
tnsnames.ora file, provide the location of the new file.

$ sqlplus /@<alias_name>
```

The output is similar to:

```
SQL*Plus: Release 11

Connected to:

Oracle Database 12c

To verify if you are connected to the correct user:

SQL> show user

The output is similar to:

USER is "<database-user-name>"
```

3.2.3 Create the Credential Keystore

Credential keystore must be created for the Behavior Detection (BD) or Enterprise Case Management (ECM) **Atomic schema** and **Compliance Studio Schema**. To create a credential keystore, follow these steps:

1. Login as HDFS Superuser.
2. Create a credential keystore on HDFS using the following command:

   ```
   hadoop credential create mydb.password.alias -provider jceks://hdfs/
   user/root/oracle.password.jceks
   ```

3. Verify the credential keystore file using the following command:

   ```
   hadoop credential list -provider jceks://hdfs/user/root/
   oracle.password.jceks
   ```

4. Grant Read permission to the keystore file using the following command:

   ```
   hadoop fs -chmod 744 /user/root/oracle.password.jceks
   ```

> **NOTE** Ensure the credential keystore file path and the alias are correctly mentioned in the `config.sh` file.

3.2.3.1 Copying and Adding Files

To copy the jar files, follow these steps:

1. Create the folder in the `GRAPH_FILES_PATH` parameter in a node of the big data cluster.
2. Create a folder called jars inside the folder that is created in the previous step.
3. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/etlJars directory` and copy all the jars in this directory into the big data node inside the folder jars.
4. Perform this step if https is enabled for Elastic Search:
   a. Create a new folder with the name as **conf** in the **GRAPH_FILES_PATH** parameter in a node of the big data cluster.
   b. Place the **es_truststore.jks** file in the **conf** folder.

> **NOTE** To use the ES-Hadoop connector, download the `commons-httpclient-3.0.1.jar` and `elasticsearch-spark-20_2.11-7.14.jar` (depending on which Elastic version is used) files and place them in the jars folder.
>
> This is applicable only in the case of ETL for Graph.

3.2.3.2 Create Credential Keystore for Elastic Search

Credential keystore must be created for the Elastic Search if https is enabled for Elastic Search.

To create a credential keystore, follow these steps:

1. Login as HDFS Superuser.
2. Create a credential keystore on HDFS using the following command:

   ```
   hadoop credential create elastic.password.alias -value <Elastic search
   password> \
   ```

```
-provider jceks://hdfs/user/fccstudio/elastic/elastic.password.jceks

hadoop credential create elastic.keystore.password.alias -value password
\

-provider jceks://hdfs/user/fccstudio/elastic/elastic.password.jceks

Where,
```

- `elastic.password.alias` is the elastic search password alias name
- `elastic.keystore.password.alias` is the elastic search keystore password alias name
- `<Elastic search password>` is elastic search password
- `password` is elastic search keystore password
- `hdfs/user/fccstudio/elastic/elastic.password.jceks` is the file path of the credential keystore

3. Verify the credential keystore file using the following command:

   ```
   hadoop credential list -provider jceks:// hdfs/user/fccstudio/elastic/
   elastic.password.jceks
   ```

4. Grant Read permission to the keystore file using the following command:

   ```
   hadoop fs -chmod 744 /user/fccstudio/elastic/elastic.password.jceks
   ```

> **NOTE**
> - Ensure the credential keystore file path and the alias are correctly mentioned in the `config.sh` file.
> - The version of the elastic search jar should be the same as the version of Elastic Search installed.

3.2.4 Download the Installer Kit

To download the software as a .zip folder, download the latest installer **33874169** for the **v8.1.2.0.0** release from My Oracle Support (MOS).

To download the software as a .zip folder, download the latest installer **34094831** for the **v8.1.2.0.1** release from My Oracle Support (MOS).

4 Installation

Perform the following steps to complete the installation:

- Extract the Installer Kit
- Place Files in the Installation Directories
- Add Synonyms and Stopword files in Elastic Search
- Place Files in Wallet
- Generate an Encrypted Password
- Generate the Public and Private Keys
- Generate API token for CS API User
- Generate the Key Store File for Secure Batch Service
- Generate Compliance Studio Server SSL Configuration Mandatory File
- Add the Batch Service (SSL) to PGX Configuration
- Configure the Extract Transfer and Load (ETL) Process
- Apply Fine-Grained access control and Redaction Changes for Compliance Studio
- Configure the config.sh File
- Configure the resources.xml for Multiple ER Schemas
- Run the Compliance Studio Installer
- Install the PGX Service

4.1 Extract the Installer Kit

After downloading the .zip folder, follow these steps to extract the folder contents:

1. Extract the contents of the installer archive file in the download directory using the following command:

   ```
   unzip -a <Compliance_Studio_Installer_Archive_File>.zip
   ```

   The Compliance Studio installer file is extracted, and the `OFS_COMPLIANCE_STUDIO` directory is obtained and is referred to as `<COMPLIANCE_STUDIO_INSTALLATION_PATH>`.

   WARNING Do not rename the application installer directory name after extraction from the archive.

2. Navigate to the download directory where the installer archive is extracted, and assign execute permission to the installer directory using the following command:

   ```
   chmod -R 0755 OFS_COMPLIANCE_STUDIO
   ```

4.2 Place Files in the Installation Directories

To place the required jars and Kerberos files in the required locations, follow these steps:

1. To place the additional jar files, follow these steps:

c. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/user/lib` directory.

d. Place the following additional jar files:

- `hive-exec-*.jar. For example, hive-exec-1.1.0.jar.`
- `HiveJDBC4.jar`
- `hive-metastore-*.jar. For example, hive-metastore-1.1.0.jar.`
- `hive-service-*.jar. For example, hive-service-1.1.0.jar.`

For additional jars, see the Appendix C – Additional Jars – PGX and Appendix D – Additional Jars – Batch Service sections.

> **NOTE**
> - The version of the jars is client or user-specific. These jars can be obtained from the existing jars of the Cloudera installation.
> - The `HiveJDBC4.jar` file is not available in the Cloudera setup. You must download the same from the Cloudera website.

2. To place the Kerberos files, follow these steps:

a. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/user/conf` directory.

b. Place the following Kerberos files:

- `krb5.conf`
- keytab file name as mentioned in the `config.sh` file.

3. Perform this step if https is enabled for Elastic Search:

a. Copy `es_truststore.jks` file from `<Elastic_Search_Installation_Path>`.

b. Place the `es_truststore.jk`s file in `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/matching-service/conf` directory.

> **NOTE** Generate the `es_truststore.jks` file in the `<Elastic_Search_Installation_Path>` before performing this step. This file contains Keystore certificates.

4.3 Add Synonyms and Stopword files in Elastic Search

To consider the similarity when performing the elastic search, you can add the synonyms and keyword files in the Elastic search.

To add synonyms and keyword files in Elastic search, perform the following steps:

1. Create a folder in the name of “analysis” in the `<Elastic Search Installation path>/config` directory.
2. You can add your synonyms and stop words to these files and place the files in the analysis folder:
   - `Country.txt`
   - `Gender.txt`
   - `Organisation_strip.txt`

- Organisation_suffix.txt
- Name_synonym.txt
- Title.txt
- Namestop.txt
- Cardinal_ordinal.txt
- Organisational_level2.txt
- Organisational_stopwords.txt
- Oraganisational_businesswords.txt

NOTE

- User can decide to provide any data in the Stopword or Synonym files.
- Each Stopword must be provided in a separate line.
- All related synonyms must be provided in the same line, separated by a comma.
- All the synonyms must be provided in the same line and ensure that there are no repetitions of the synonym. For Example, rob, robi, robie, roby, robbi.

4.4 Place Files in Wallet

To place the files in the wallet in the required locations, follow these steps:

1. To place the files in the wallet, follow these steps:
   a. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>.`
   b. Create a folder 'wallet' and place the following files.
   c. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/wallet.`
   d. Place the following files:
      - tnsnames.ora
      - ewallet.p12
      - cwallet.sso
      - ewallet.p12.lck
      - cwallet.sso.lck

NOTE This folder path will be referred to as "`WALLET_LOCATION`" and "`TNS_ADMIN_PATH`" in `config.sh` while configuring Compliance Studio. If you want to maintain `tnsname.ora` in a different folder, then "`TNS_ADMIN_PATH`" will be that folder location.

2. Place the **sqlnet.ora** file into the wallet and update the path for the current wallet location.

```
WALLET_LOCATION =

     (SOURCE =

          (METHOD = FILE)
```

```
                (METHOD_DATA =
                    (DIRECTORY = <wallet location>)
                )
            )
        SQLNET.WALLET_OVERRIDE = TRUE
```

4.5 Generate an Encrypted Password

To generate encrypted passwords required during configuration, i.e., while configuring encrypted passwords, for example. STUDIO_DB_ENCRYPTED_PASSWORD, follow the below steps.

1. Set the export `FIC_DB_HOME` path in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb` directory.
2. Run the echo `$FIC_DB_HOME` command.
3. Go to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/bin` directory and run the `./FCCM_Studio_Base64Encoder.sh <password to be encrypted>` command.

4.6 Generate the Public and Private Keys

The Public and Private keys are JSON Web Tokens (JWT) that are generated for Authentication from Compliance Studio.

To generate the keys, follow these steps:

> **NOTE** The following steps are mandatory for the first time Compliance Studio installation.

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/bin` directory.
2. Run the Shell Script `FCCM_Studio_JWT_Keygen.sh` from the directory.

   The Public and Private Keys are generated and available in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/conf` directory.
3. Copy the `private.key` and `public.key` files to the following paths:
   - `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/datastudio/server/conf directory`
   - `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/conf directory`
   - `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/pgx/server/conf directory`

4.7 Generate API token for CS API User

To generate the API token, follow these steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/bin` directory.
2. Run the following command:

   ```
   export FIC_DB_HOME=<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb
   ```
3. Run the following shell script:

```
./FCCM_Studio_Generate_APIToken.sh <FCC_API_USER>
```

This will generate the API token on the terminal.

4. Save the generated token that is required while configuring `config.sh` file in the path `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin`.

4.8 Generate the Key Store File for Secure Batch Service

Generating the Key Store file for Secure Batch Service generates the key store parameters and changes the key store parameters from HTTP to HTTPS protocol.

> **NOTE**
> - The following steps are only applicable if the user wants to create a self-signed certificate.
> - It is recommend strongly that obtaining a signed certificate from your IT admin team for this host.

To configure the Key Store file for Secure Batch Service, follow these steps:

1. Run the `keytool -genkey -alias batchservice -keyalg RSA -keysize 2048 -keystore <COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/conf/<Keystore file name>.jks` command in the Studio Server.

   When generating the keytool ensure to provide the hostname in the first name.

   Question: What is your first and last name?

   Answer: Provide the fully qualified studio server hostname.

   For example, <hostname>.<domain name>

2. Specify the keystore password. The `<Keystore file name>.jks` file is created in the path `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/conf` directory.

3. Specify the following parameters in the `config.sh` file.

   - `export KEYSTORE_FILE_NAME=<Keystore file name>.jks`
   - `export KEYSTORE_PASS=”your password”`

4.9 Generate Compliance Studio Server SSL Configuration Mandatory File

Topics:

- Generate Self-signed Certificate
- Generate Signed Certificate

4.9.1 Generate Self-signed Certificate

To generate the self-signed certificate, perform the following steps:

1. Run the following jks command in the Studio Server:

```
keytool -genkey -alias <alias> -keyalg RSA -keystore <alias>.jks
```

> **NOTE** You must use the same password and alias that is provided in the `config.sh` file.

2. Specify the keystore password.
3. When generating the keytool ensure to provide the hostname in the first name. For example:

   Question: What is your first and last name?

   Answer: Provide the fully qualified studio server hostname.

   For example, <hostname>.<domain name>
4. Specify any name for the other questions.
5. Specify the keystore password. The `jks` file is created in the Studio Server.

> **NOTE** You must use the same password and alias that is provided in the `config.sh` file.

6. Run the following jks command in the Studio Server to generate the `.p12` file using the `.jks` file.;

```
keytool -importkeystore -srckeystore <alias>.jks -destkeystore
<alias_name>.p12 -srcalias <alias> -srcstoretype jks -deststoretype
pkcs12
```

7. Specify the keystore password. The .`p12` file is created in the Studio Server.
8. Copy the .`p12` files and place in the `<Studio Installation path>/datastudio/server/conf` directory.

4.9.2 Generate Signed Certificate

To generate the signed certificate, perform the following steps:

1. Log in to the server as a Linux user.
2. Generate the CSR file that describes the certificate requested and needed by the signing authority.
3. Openssl default configuration does not include subject alternative names by default.
4. SANs should be updated in `cert.conf` file. Additional SANs or IPs can be added through properties such as DNS.2, DNS.3, IP.1, and IP.2 in the [alt_names] section.
5. Once the configuration file is placed, generate the CSR file and associated private key by running the following command:

```
openssl req -new -newkey rsa:2048 -nodes -keyout server.key -out
server.csr -config cert.conf
```

6. Provide the requested entries, and some entries can be left blank.

> **NOTE** You can check the CSR contains SANs by running the following command:
>
> `openssl req -text -noout -verify -in server.csr`
>
> This step is optional only.

7. Request certificate from the signing authority. Once the certificate is received, convert the `server.cer` into PEM format if required by running the following command:

```
openssl x509 -in server.cer -out server.pem -outform PEM
```

> **NOTE** You can check the contents of the certificate to make sure that the SANs are included by running the following command:
>
> `openssl x509 -in server.pem -text`
>
> This step is optional only.

8. Create `.p12` keystore.

> **NOTE**
>
> - The -name parameter must match the value of the **STUDIO_SERVER_SSL_ALIAS** variable from the path `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin/config.sh`
> - To store the password, run the following command:
>
>   `openssl pkcs12 -export -out studio_server.p12 -inkey server.key -in server.pem -name studio_alias`
> - The password must match the value of the **STUDIO_SERVER_SSL_PASSWORD** variable from `<COMPLIANCE_STUDIO_INSTALLATION_PATH >/bin/config.sh`
> - To check the keystore, run the following command:
>
>   `openssl pkcs12 -export -out studio_server.p12 -inkey server.key -in server.pem -name studio_alias`
>
>   This step is optional only.

9. Copy the `cp studio_server.p12` file and place in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/datastudio/server/conf/studio_server.p12` and `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/datastudio/server/conf/studio_server.p12` directories.

10. Restart Compliance Studio. To do this, navigate to the

    `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory and run the `./compliance-studio.sh restart`

4.10 Add the Batch Service (SSL) to PGX Configuration

Adding the Batch Service (SSL) to PGX Trust Store facilitates you to apply redaction on the graph batch service and connect with PGX.

To add the Batch Service to PGX Trust Store, copy the `<Keystore file name>.jks` file to the `<PGX Server path>/server/conf` directory. To create a `.jks` file, see Generate the Key Store File for Secure Batch Service.

After generating the key store file and adding the batch service to the PGX trust store, you must configure the user mapping for the changes made in the database. For more information about configuring user mapping, see the OFS Compliance Studio Administration and Configuration Guide.

4.11 Configure the Extract Transfer and Load (ETL) Process

Extract Transfer and Load (ETL) is the procedure of copying data from one or more sources into a destination system that represents the data differently from the source or in a different context. Data movement and graph loading is performed using ETL.

To configure the Data Movement and Graph Load, copy the applicable `FCCM_Studio_SqoopJob.sh` files from the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/bin` directory and add in the `<FIC_HOME of OFSAA_Installed_Path>/ficdb/bin` directory.

For more information, see the Configure ETL and Execute ETL sections in the OFS Compliance Studio Administration and Configuration Guide.

> **NOTE** Before running the sqoop job, ensure that the correct values are the `server-config.properties` file from the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/batchservice/conf` directory.

Topics:

- Loading Graphs

4.11.1 Loading Graphs

Loading graphs to Compliance Studio can be based on the following scenarios:

4.11.1.1 Loading sample graph without running ETL

To load the sample graph without running ETL, perform the following steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/sample-graph` folder.
2. Unzip the contents of the `sample-graph-8.*.zip` file in the same folder.
3. Copy the entire path of the folder `sample-graph`.
4. Open the `sample-graph-8.*.json` file and paste the copied `<sample-graph folder path>` into the placeholder `<SERVER_PATH>` under the parameter `"uris"`.

> **NOTE** Ensure to replace all the placeholders with the copied path of the folder `sample-graph`.

5. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/pgx/server/conf` directory and modify the `pgx.conf` file for the highlighted parameters:

```
"preload_graphs": [
    {
      "path": "<sample-graph folder path>/sample-graph.json",
      "name": "GlobalGraphIH",
      "publish": false,
      "publish_with_snapshots": true
    }
  ],
  "pgx_realm": {
    "implementation": "com.oracle.ofss.fccm.studio.pgx.FCCMPgxRealm"
  },
  "file_locations": [
    {
      "name": "hdfs_storage",
      "location": "<sample-graph folder path>"
    }
  ]
```

6. Restart the PGX server.

4.11.1.2 Loading the graph generated from ETL

You can load a graph generated from ETL based on the following:

PGX fails until you have a graph generated from ETL on the HDFS. Once the graph is generated, and then as soon as the PGX server pods restart, the graph is pre-loaded to the Compliance Studio.

- Create a backup of `pgx.conf`. The backup can be used when the graph is generated from ETL.
- At the time of deployment, you must delete the following lines from the `pgx.conf` file.

```
"preload_graphs": [
  {
    "path": "##URL_GLOBAL_GRAPH_CONFIG_JSON##",
    "name": "##PGX_GLOBAL_GRAPH_NAME##",
    "publish": false,
    "publish_with_snapshots": true
  }
],
```

  - The following lines must be delete multiple times.

```
,
        {
          "preloaded_graph": "##PGX_GLOBAL_GRAPH_NAME##",
            "grant": "manage"
        }
```

- Proceed with the Compliance Studio deployment.
- Once the graph is generated, perform the following:
    - Replace the existing `pgx.conf` file with the backed up pgx.conf file
- Restart Compliance Studio.

4.12 Apply Fine-Grained access control and Redaction Changes for Compliance Studio

After generating the key store file and adding the batch service to the PGX trust store, you must configure the user mapping for the changes made in the database. For more information about configuring user mapping, see the OFS Compliance Studio Administration and Configuration Guide.

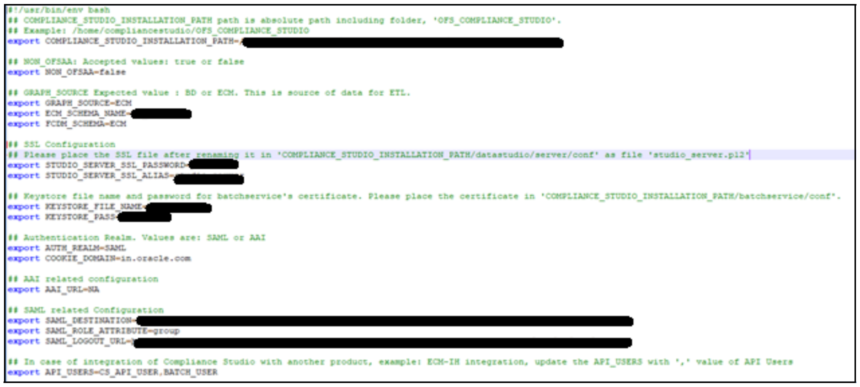
4.13 Configure the config.sh File

To configure the `config.sh` file for installing Compliance Studio, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory.
3. Configure the applicable `config.sh` attributes are shown in the following table.

    A sample `config.sh` file is shown:

Figure 4: Sample Config.sh File



```
#!/usr/bin/env bash
## COMPLIANCE_STUDIO_INSTALLATION_PATH path is absolute path including folder, 'OFS_COMPLIANCE_STUDIO'.
## Example: /home/compliancestudio/OFS_COMPLIANCE_STUDIO
export COMPLIANCE_STUDIO_INSTALLATION_PATH=[redacted]

## NON_OFSAA: Accepted values: true or false
export NON_OFSAA=false

## GRAPH_SOURCE Expected value : BD or ECM. This is source of data for ETL.
export GRAPH_SOURCE=ECM
export ECM_SCHEMA_NAME=[redacted]
export FCDM_SCHEMA=ECM

## SSL Configuration
## Please place the SSL file after renaming it in 'COMPLIANCE_STUDIO_INSTALLATION_PATH/datastudio/server/conf' as file 'studio_server.p12'
export STUDIO_SERVER_SSL_PASSWORD=[redacted]
export STUDIO_SERVER_SSL_ALIAS=[redacted]

## Keystore file name and password for batchservice's certificate. Please place the certificate in 'COMPLIANCE_STUDIO_INSTALLATION_PATH/batchservice/conf'.
export KEYSTORE_FILE_NAME=[redacted]
export KEYSTORE_PASS=[redacted]

## Authentication Realm. Values are: SAML or AAI
export AUTH_REALM=SAML
export COOKIE_DOMAIN=in.oracle.com

## AAI related configuration
export AAI_URL=NA

## SAML related Configuration
export SAML_DESTINATION=[redacted]
export SAML_ROLE_ATTRIBUTE=group
export SAML_LOGOUT_URL=[redacted]

## In case of integration of Compliance Studio with another product, example: ECM-IH integration, update the API_USERS with ',' value of API Users
export API_USERS=CS_API_USER,BATCH_USER
```

NOTE

- You must manually set the parameter value in the `config.sh` file. If a value is not applicable, enter NA and ensure that the value is not entered as **NULL**.
- If the parameter STUDIO_DB_SERVICE_NAME has been filled, the parameter STUDIO_DB_SID should be left **blank**, and vice versa.
- If the parameter ATOMIC_DB_SERVICE_NAME has been filled, the parameter ATOMIC_DB_SID should be left **blank**, and vice versa.

Table 14 lists configuration parameters of the `config.sh` file:

Table 14: config.sh file

Parameter	Significance	Installing with OFSAA (Mandatory)	Upgrading with OFSAA (Mandatory)	Installing without OFSAA (Mandatory)	Upgrading without OFSAA (Mandatory)
COMPLIANCE_STUDIO_INSTALLATION_PATH	Indicates the path where the Compliance Studio installer file is extracted.	Yes	Yes	Yes	Yes
NON_OFSAA	To install Compliance Studio with OFSAA, enter "false". To install Compliance Studio without OFSAA, enter "true".	Enter false	Enter false	Enter true	Enter true
GRAPH_SOURCE					
GRAPH_SOURCE	Indicates the source database for Compliance Studio. The available options are ECM and BD. **NOTE:** ● Compliance Studio can use either the BD or ECM schema as the source of FCDM data for the graph. ● Ensure to enter the value as ECM whenever ECM integration is required with Investigation Hub. Here, the ECM schema is used as the source of the FCDM data to load the case information into the graph.	Enter BD or ECM	Enter BD or ECM	Enter NA	Enter NA

Table 14: config.sh file

FCDM_SCHEMA	This indicated the datasource for the Production workspace. The available options are ECM and BD.	Enter BD or ECM	Enter BD or ECM	Enter NA	Enter NA
ECM_SCHEMA_NAME	ECM Schema name	ECM Schema name	ECM Schema name	Enter NA	Enter NA
SSL file					
STUDIO_SERVER_SSL_PASSWORD	Indicates the password for Studio Server P12 that is required for HTTPS configuration.	Yes	Yes	Yes	Yes
STUDIO_SERVER_SSL_ALIAS	Indicates the alias name for P12 for the Studio Server	Yes	Yes	Yes	Yes
Keystore file and pass details for batch service					
KEYSTORE_FILE_NAME	Indicates the keystore file name that is used for secure batch service.	Yes	Yes	Yes	Yes
KEYSTORE_PASS	Indicates the keystore password details for the secure batch service.	Yes	Yes	Yes	Yes
Authentication Realm					

Table 14: config.sh file

AUTH_REALM	Realm indicates the functional grouping of database schemas and roles that must be secured for an application. Realms protect data from access through system privileges; realms do not give its owner or participant's additional privileges. Compliance Studio uses realm-based authorization and authentication for its users. For more information, see the Realm Based Authorization for Compliance Studio section in the OFS Compliance Studio Administration and Configuration Guide. The Compliance Studio application can be accessed using the following realms: **FCCMRealm** Value=AAI **FCCSamlRealm** Value=SAML	Yes	Yes	Yes	Yes
COOKIE_DOMAIN	The domain of the server. Example: in.oracle.com	Yes	Yes	Yes	Yes
AAI related configuration					
AAI_URL	OFSAA URL.	Yes	Yes	No	No

Table 14: config.sh file

SAML **The SAML-related parameters are applicable only if SAMLRealm is used in the Realm parameter.**	1. In the case of SAML Realm, the certificate from IDP (**key.cert** file) is required. 2. The certificate that is obtained from the IDP must be renamed to key.cert and placed in the < `COMPLIANCE_STUDIO_INSTALLATION_PATH` >/ `datastudio/server/conf` directory. 3. This certificate is used to identify the trust of the SAML response from the Identity Provider. 4. Specify the Role Attribute name from IDP, in which the User Roles are present in the SAML response.				
SAML_DESTINATION	Indicates the SAML IDP URL that the Identity Provider provides after creating the SAML Application.	Yes	Yes	Yes	Yes
SAML_ROLE_ATTRIBUTE	Indicates the SAML client identifier provided by the SAML Administrator for the Role and Attributes information while creating the SAML application for Compliance Studio.	Yes	Yes	Yes	Yes
SAML_LOGOUT_URL	Indicates the SAML client identifier provided by the SAML Administrator for the Logout URL information while creating the SAML application for Compliance Studio.	Yes	Yes	Yes	Yes

Table 14: config.sh file

Integrate with other products	In case of integration of Compliance Studio with another product, for example, ECM-IH integration, update the API_USERS with ',' value of API Users				
API_USERS	Indicates the API users. Comma-separated API Users, which accesses datastudio using API token. Example: ECM_USER,BATCH_USER,MMG_USER	Yes	Yes	Yes	Yes
MMG Service Configurations					
SESSION_TOKEN_CREDENTIALS	Set password to generate Authorization Header Token to communicate with mmg-services	Yes	Yes	Yes	Yes
FCC_API_USER	API User for Compliance Studio.	Yes	Yes	Yes	Yes
SSO_TOKEN	This is the API token for FCC_API_USER. See the Generate API token for CS API User for token value.	Yes	Yes	Yes	Yes
MMG_DATASOURCE_MAX_POOL_SIZE	Maximum connection pool size allowed for Config Data Source. 50	Yes	Yes	Yes	Yes
MMG_DATASOURCE_IDLE_TIMEOUT	Idle timeout for Config Data Source in a millisecond. 30000	Yes	Yes	Yes	Yes
MMG_DATASOURCE_CONN_TIMEOUT	Connection timeout for Config Data Source in milliseconds. 30000	Yes	Yes	Yes	Yes
EXT_DATASOURCE_MAX_POOL_SIZE	Maximum connection pool size allowed for Meta/Data Schemas. 50	Yes	Yes	Yes	Yes

Table 14: config.sh file

EXT_DATASOURCE_IDLE_TIMEOUT	Idle timeout for Meta/ Data Schemas in milliseconds. 30000	Yes	Yes	Yes	Yes
EXT_DATASOURCE_CONN_TIMEOUT	Connection timeout for Meta/Data Schemas in milliseconds. 30000	Yes	Yes	Yes	Yes
SERVER_COOKIE_TIMEOUT	Connection timeout for server cookie in milliseconds. 86400	Yes	Yes	Yes	Yes
DB Details for Studio Schema **You must be logged in as SYSDBA to perform these configurations.**					
STUDIO_DB_HOSTNAME	Indicates the hostname of the database where the Compliance Studio schema is created.	Yes	Yes	Yes	Yes
STUDIO_DB_PORT	Indicates the port number where the Compliance Studio schema is created.	Yes	Yes	Yes	Yes
STUDIO_DB_SERVICE_NAME	Indicates the service name of the database where the Studio schema is created.	Yes	Yes	Yes	Yes
STUDIO_DB_SID	Indicates the SID of the database where the Studio schema is created. **NOTE:** Set this field as blank if there is no SID for Database.	Yes	Yes	Yes	Yes
STUDIO_DB_USERNAME	Indicates the username of the Compliance Studio Schema (newly created Oracle Schema).	Yes	Yes	Yes	Yes

Table 14: config.sh file

STUDIO_DB_PASSWORD	Indicates the password of the Studio schema.	Yes	Yes	Yes	Yes
STUDIO_DB_ENCRYPTED_PASSWORD	Indicates the encrypted password that is provided for the Studio schema. For example, cGFzc3dvcmQ. **NOTE:** See Generate an Encrypted Password section to generate this encrypted password.	Yes	Yes	Yes	Yes
DB Details of Atomic Schema					
ATOMIC_DB_HOSTNAME	The hostname of the database where Atomic schema is present. (BD/ECM config)	Yes	Yes	Yes	Yes
ATOMIC_DB_PORT	Port number of database where Atomic schema is present.	Yes	Yes	Yes	Yes
ATOMIC_DB_SERVICE_NAME	The service name of the database where Atomic schema is present.	Yes	Yes	Yes	Yes
ATOMIC_DB_SID	Service id of database where Atomic schema is present. **NOTE:** Set this field as blank if there is no SID for Database.	Yes	Yes	Yes	Yes
ATOMIC_DB_USERNAME	Username of Atomic schema	Yes	Yes	Yes	Yes
ATOMIC_DB_PASSWORD	The password of the Atomic schema	Yes	Yes	Yes	Yes
Studio DB Wallet Details **For information on creating a wallet, see Setup Password Stores with Oracle Wallet.**					

Table 14: config.sh file

STUDIO_ALIAS_NAME	Indicates the Studio alias name. **NOTE:** Enter the alias name that was created during wallet creation.	Yes	Yes	Yes	Yes
WALLET_LOCATION	Indicates the Compliance Studio wallet location.	Yes	Yes	Yes	Yes
TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the STUDIO_ALIAS_NAME is present.	Yes	Yes	Yes	Yes
ATOMIC_ALIAS_NAME	Indicates alias name of FCDM source atomic schema given in wallet	Yes	Yes	Yes	Yes
Cloudera Setup Details **Contact your System Administrator to obtain the required details for these parameters.**					
STUDIO_HADOOP_CREDENTIAL_ALIAS	Indicated the alias password saved in Hadoop. For example, studio.password.alias	Yes	Yes	Yes	Yes
STUDIO_HADOOP_CREDENTIAL_PATH	Indicates the credentials path. For example, <Compliance Studio Installed Path>oracle.password.jceks	Yes	Yes	Yes	Yes
HADOOP_CREDENTIAL_PROVIDER_PATH	Indicates the path where the Hadoop credential is stored.	Yes	Yes	Enter NA	Enter NA

Table 14: config.sh file

HADOOP_PASSWORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials. For information on creating a credential keystore, see Create the Credential Keystore.	Yes	Yes	Enter NA	Enter NA
Hive_Host_Name	Indicates the Hive hostname.	Yes	Yes	Enter NA	Enter NA
Hive_Port_number	Indicates the Hive port number. Contact your Studio Administrator to obtain the port number.	Yes	Yes	Enter NA	Enter NA
HIVE_PRINCIPAL	Indicates the Hive Principal. Contact your Studio Administrator to obtain the HIVE_PRINCIPAL value.	Yes	Yes	Enter NA	Enter NA
HIVE_SCHEMA	Indicates to create a schema in HIVE.	Yes	Yes	Enter NA	Enter NA
Krb_Host_FQDN_Name	Indicates the Kerberos host FQDN name.	Yes	Yes	Enter NA	Enter NA
Krb_Realm_Name	Indicates the Kerberos realm name.	Yes	Yes	Enter NA	Enter NA
Krb_Service_Name	Indicates the Kerberos service name. Example: Hive	Yes	Yes	Enter NA	Enter NA
server_kerberos_keytab_file	Indicates the Kerberos keytab file.	Yes	Yes	Enter NA	Enter NA
server_kerberos_principal	Indicates the Kerberos Principal.	Yes	Yes	Enter NA	Enter NA
server_kerberos_krb5_conf_file	Indicates the krb5.conf file name.	Yes	Yes	Enter NA	Enter NA
SQOOP_HOSTMACHINE_USER_NAME	Indicates the username of the Host machine where sqoop will run.	Yes	Yes	Enter NA	Enter NA

Table 14: config.sh file

SQOOP_PARAMFILE_PATH	1. Create a file with the name `sqoop.properties` in the Big Data server and add the following entry: oracle.jdbc.mapDateToTimestamp=false 2. Enter the location of the `sqoop.properties` file as the value for this parameter. Example: /scratch/ofsaa/ **NOTE:** Ensure that the location name ends with a'/'.	Yes	Yes	Enter NA	Enter NA
SQOOP_PARTITION_COL	Indicates the column in which the HIVE table is partitioned. The value must be SNAPSHOT_DT.	Yes	Yes	Enter NA	Enter NA
SQOOP_TRG_HOSTNAME	Indicates the hostname of the Big Data server where SQOOP will run. Example: <HostName>	Yes	Yes	Enter NA	Enter NA
SQOOP_WORKDIR_HDFS	Indicates the Sqoop working directory in HDFS. Example: /user/ofsaa	Yes	Yes	Enter NA	Enter NA
ETL					
HDFS_GRAPH_FILES_PATH	Indicates the file path in the HDFS where the graph.json is formed.	Yes	Yes	No	No
GRAPH_FILES_PATH	Indicates the directory in the Big Data server for graph files.	Yes	Yes	No	No
GRAPH_NAME	Indicates the name you want to assign to the global graph at the end of ETL.	Yes	Yes	No	No

Table 14: config.sh file

ETL_PROCESSING_RANGE	Indicates the duration for which the data would be moved from Oracle to Hive. For example: If the ETL_PROCESSING_RANGE = 2Y, 3M, 10D, that is, 2 years, 3 months, and 10 days, and the current date is 20200814, then the data movement occurs for the range 20180504 to 20200814.	Yes	Yes	No	No
OLD_GRAPH_SESSION_DURATION	Indicates that the session older than this specified duration will be removed from the PGX server. If unsure, you can set this value for a week (7D).	Yes	Yes	No	No
REMOVE_TRNXS_EDGE_AFTER_DURATION	Indicates the date range for which transaction edges will be maintained in the graph. For example, 6Y, 3M, 10D, which means 6 years, 3 months, and 15 days.	Yes	Yes	No	No
CONNECTOR_CHANGESET_SIZE	Indicates the number of nodes or edges you want to process during an update of the graph. If unsure, you can set it to 10000.	Yes	Yes	No	No
CB_CONFIGURED	Indicates the setting of the graph edges. When the corresponding edges of the graph are needed, set the value to true.	Enter true or false	Enter true or false	Enter NA	Enter NA
Elastic Search Cluster details					

Table 14: config.sh file

ELASTIC_SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_USERNAME	Elastic Search Username (Not Applicable, if https enabled is false and authentication is not supported).	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_ENCRYPTED_PASSWORD	Encrypted password (Not Applicable, if https enabled is false and authentication is not supported). **NOTE:** See Generate an Encrypted Passwordsection to generate this encrypted password.	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_HTTPS_ENABLED	True (If ES is https enabled, else false)	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_TRUSTSTORE_FILE_NAME	The filename of the ElasticSearch keystore that contains the certificates of ES host to trust (Not Applicable, if https enabled is false)	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_TRUSTSTORE_PASSWORD	The password of the Elasticsearch keystore file. (Not Applicable, if https enabled is false).	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_HADOOP_PASSWORD_ALIAS	Indicates the password alias for Elastic Search (Not applicable if ES ELASTIC_SEARCH_HTTPS_ENABLED is false).	Yes	Yes	Yes	Yes

Table 14: config.sh file

ELASTIC_SEARCH_KEYSTORE_HADOOP_CREDENTIAL_ALIAS	Indicates the password alias for Elastic Search (Not applicable if ES ELASTIC_SEARCH_HTTPS_ENABLED is false).	Yes	Yes	Yes	Yes
ELASTIC_SEARCH_HADOOP_CREDENTIAL_PATH	Indicates the elastic search hadoop credential path.	Yes	Yes	Yes	Yes
Logstash					
LOGSTASH_HOME	Logstash home Example: "/ `<COMPLIANCE_STUDIO_INSTALLATION_PATH>`/Logstash/ logstash-7.14.0" **NOTE:** See the section Configure Logstash for more details.	Yes	Yes	Yes	Yes
Service URLs					
PGX_SERVER_URL	Indicates the comma ',' separated values of PGX URLs. If you have only one PGX URL, the value is http://<server1>:7007. NOTE: Ensure to provide the correct hostname for the URL of the PGX service.	Yes	Yes	No	No
PGX server configuration, i.e., Interpreter, data memory limits					
NUM_CACHED_RESULTSET	Indicates the ached result set. For example, 0	Yes	Yes	No	No
RESULTSET_EXPIRATION_TIME_SECS	Indicates the Result set expiration time. For example, 3600.	Yes	Yes	No	No

Table 14: config.sh file

MAX_TOTAL_SHARED_DATA_MEMORY_SIZE	The absolute memory limit of shared data (includes published graphs and pinned non-referenced graphs). For example: 20G	Yes	Yes	No	No
MAX_TOTAL_PRIVATE_DATA_MEMORY_SIZE	The memory limit of private data (includes non-published graphs and PGQL results) relative to the total PGX engine memory limit. For example, 8G	Yes	Yes	No	No
MAX_PER_SESSION_DATA_MEMORY_SIZE	Absolute memory limit for any one session of the PGX engine. For example: 700M	Yes	Yes	No	No
MAX_DATA_MEMORY_SIZE_DSADMIN	Absolute memory limit for any user of the PGX engine whose role is DSADMIN. For example: 2G	Yes	Yes	No	No
MAX_DATA_MEMORY_SIZE_DSBATCH	Absolute memory limit for any user of the PGX engine whose role is DSBATCH. For example: 10G	Yes	Yes	No	No
MAX_DATA_MEMORY_SIZE_DSINTER	Absolute memory limit for any user of the PGX engine whose role is DSINTER. For example: 5G	Yes	Yes	No	No
MAX_DATA_MEMORY_SIZE_DSAPPROVER	Absolute memory limit for any user of the PGX engine whose role is DSAPPROVER. For example: 5G	Yes	Yes	No	No
MAX_DATA_MEMORY_SIZE_DSUSER	Absolute memory limit for any user of the PGX engine whose role is DSUSER. For example, 5G	Yes	Yes	No	No

Table 14: config.sh file

Quantifind Details In the case of Quantifind, the generated Quantifind token must be encoded. Use the `<Fic_DB_path>/FCCM_Studio_Base64Encoder.sh` file for encoding Quantifind token.					
QUANTIFIND_URL	Indicates the URL of the Quantifind. For example, https://api-test.quantifind.com	Yes	Yes	Yes	Yes
ENCRYPTED_QUANTIFIND_TOKEN	Indicates the token that is generated when integrating with Quantifind. For example, c2FtcGxlX2VuY3J5cHRlZF9xdWFudGlmaW5kX3Rva2Vu **NOTE:** See Generate an Encrypted Password section to generate this encrypted password.	Yes	Yes	Yes	Yes
QUANTIFIND_APPNAME	Indicates the Quantifind App Name. For example, OracleIntegrationTest	Yes	Yes	Yes	Yes
QUANTIFIND_ENABLED	Indicates that Quantifind is enabled. Options are True or False.	Yes	Yes	Yes	Yes
HTTPS_PROXY_HOST	Indicates the proxy host that is used. For example, www-proxy-idc.in.oracle.com	Yes	Yes	Yes	Yes
HTTPS_PROXY_PORT	Indicates the proxy port that is used. For example, 80	Yes	Yes	Yes	Yes

Table 14: config.sh file

HTTP_PROXY_USERNAME	Indicates the proxy username used, if there is any. For example, ##HTTP_PROXY_USERNAME##	Yes	Yes	Yes	Yes
HTTP_PROXY_PASSWORD	Indicates the proxy password used if there is any. For example, ##HTTP_PROXY_PASSWORD##	Yes	Yes	Yes	Yes
Additional Environment variables					
LD_LIBRARY_PATH	Oracle Instant client path For example: /opt/oracle/instantclient_19_8/:$LD_LIBRARY_PATH				
All Services	Set the value of the parameter, DEPLOY_ALL_SERVICE, as : ● **true** for starting all services ● **false** for starting selected services Examples: ● Compliance Studio independent of OFSAA: set "false" for service(s): entity-resolution, matching-service, and load-to-elastic ● Compliance Studio lite: set "false" for service(s): fcc-pgql, fcc-pgx-algorithm, fcc-pgx-java and pgx-server.				

Table 14: config.sh file

DEPLOY_ALL_SERVICE	True: Indicates that all services are deployed.	Yes	Yes	Yes	Yes
Services					
SERVER_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
AUTHSERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
METASERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
BATCHSERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
SESSIONSERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
FCC_JDBC_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
JDBC_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
FCC_MARKDOWN_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
ORE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
FCC_PYTHON_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
SPARK_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
PGX_SERVER_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
FCC_PGX_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
ENTITY_RESOLUTION_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
MATCHING_SERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
LOAD_TO_ELASTIC_SEARCH_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
MMG_UI_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes
MMG_SERVICE_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes

Table 14: config.sh file

MMG_SCHEMA_CREATOR_ENABLED	True: Indicates that the service is enabled.	Yes	Yes	Yes	Yes

4.14 Configure the resources.xml for Multiple ER Schemas

NOTE

- **ER_Schema ID** should always be unique.
- For ease of execution, it is recommended to have the same **Er_Data_Schema_Alias_Name** as the **ER_Schema_ID**.
- **Er_Data_Schema_Alias_Name** and **ER_Schema_ID** are case sensitive, so it is recommended to use the same case for both of them.

1. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/ficdb/conf`

NOTE If the user wants to add additional ER schemas post-installation, the path will change to:

`<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/conf`

The remaining steps will remain the same.

2. Open `resources.xml` file.
3. Provide the **id** as `ER_Schema_ID` and **ER_DATA_SCHEMA_ALIAS_NAME** as `ER_Schema_Alias.`

   The sample resource tag will look like the following, users can change the values as applicable:

```
<Resource
            id="##ER_DATA_SCHEMA_ALIAS_NAME##"
            name="jdbc/erdataschema"
            auth="Container"
            type="javax.sql.DataSource"
            driverClassName="oracle.jdbc.OracleDriver"
            url="jdbc:oracle:thin:@##ER_DATA_SCHEMA_ALIAS_NAME##"
            connectionProperties="oracle.net.wallet_location
=##STUDIO_WALLET_LOCATION##;
oracle.net.tns_admin=##STUDIO_TNS_ADMIN_PATH##;"
            maxTotal="20"
            maxIdle="0"
            maxWaitMillis="-1" >
    </Resource>
```

```
Example resource.xml tag with single ER Schema:
<Resource
            id="ER1"
            name="jdbc/erdataschema"
            auth="Container"
            type="javax.sql.DataSource"
            driverClassName="oracle.jdbc.OracleDriver"
            url="jdbc:oracle:thin:@ER1"
            connectionProperties="oracle.net.wallet_location
=##STUDIO_WALLET_LOCATION##;
oracle.net.tns_admin=##STUDIO_TNS_ADMIN_PATH##;"
            maxTotal="20"
            maxIdle="0"
            maxWaitMillis="-1" >
    </Resource>
```

4. The sample can be repeated for multiple ER Schemas with a unique id and ER_Schema_Alias.

   Example **`resource.xml`** tag with multiple ER Schemas:

```
<Resource
            id="ER1"
            name="jdbc/erdataschema"
            auth="Container"
            type="javax.sql.DataSource"
            driverClassName="oracle.jdbc.OracleDriver"
            url="jdbc:oracle:thin:@ER1"
            connectionProperties="oracle.net.wallet_location
=##STUDIO_WALLET_LOCATION##;
oracle.net.tns_admin=##STUDIO_TNS_ADMIN_PATH##;"
            maxTotal="20"
            maxIdle="0"
            maxWaitMillis="-1" >
      </Resource>
<Resource
            id="ER2"
            name="jdbc/erdataschema"
     auth="Container"
```

```
                type="javax.sql.DataSource"
                driverClassName="oracle.jdbc.OracleDriver"
                url="jdbc:oracle:thin:@ER2"
                connectionProperties="oracle.net.wallet_location
=##STUDIO_WALLET_LOCATION##;
oracle.net.tns_admin=##STUDIO_TNS_ADMIN_PATH##;"
                maxTotal="20"
                maxIdle="0"
                maxWaitMillis="-1" >
    </Resource>
```

> **NOTE** Make sure that the following parameters are updated with the values:
>
> `maxTotal="20"`
>
> `maxIdle="0"`

4.15 Run the Compliance Studio Installer

This section provides the install, reinstall, start and stop of the services.

Topics:

- Installing for the first time
- Starting Compliance Studio
- Stopping Compliance Studio
- Restarting Compliance Studio
- Reinstalling Compliance Studio

The Compliance Studio application is installed with or without OFSAA, depending on the configuration provided in the `config.sh` file. The Compliance Studio application and all the interpreters are started.

After completing the Compliance Studio installation, the script displays a URL that can be used to access the Compliance Studio Application.

4.15.1 Installing for the first time

For first-time installation, you can pass argument '-i' or '--install'.

To run the Compliance Studio installer, follow these steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory.
2. Run the following command:

   ```
   ./compliance-studio.sh -i
   Or
   ./compliance-studio.sh --install
   ```

This will copy the whole compliance studio into the folder 'deployed' and then replaces the placeholders. Now, you can start Compliance Studio.

NOTE

- Run these commands only from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin.`
- It should not be run from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/bin.`
- Upon executing `./compliance-studio.sh -i` command. A deployed folder is created that copies all the folders. And replaces placeholders inside the deployed folder.

Congratulations! Your installation is complete.

NOTE For any help on installation commands, Run

```
./compliance-studio.sh --help
```

4.15.2 Starting Compliance Studio

To start the application, you can run pass argument '-s' or'--start'. Example:

```
./compliance-studio.sh --start
```

This will start the application and, on successful installation, will make the sensitive details blank in `config.sh`

4.15.3 Stopping Compliance Studio

To stop the application, you can run pass argument '-k' or '--stop'. Example:

```
./compliance-studio.sh --stop
```

4.15.4 Restarting Compliance Studio

To restart the application, you can run pass argument '-r' or '--restart'. Example:

```
./compliance-studio.sh --restart
```

4.15.5 Reinstalling Compliance Studio

In case if you need to reinstall compliance Studio due to the wrong configuration or need to update configuration details. Then:

- Stop the Compliance Studio
- Update the `config.sh` file. Do not forget to reconfigure the sensitive details which were removed earlier.

To restart the application, you can run pass argument '-R' or '--reinstall'. Example:

```
./compliance-studio.sh --reinstall
```

Once reinstallation is done, you can start the application.

4.16 Install the PGX Service

> **NOTE** PGX service can be installed on the same server where Compliance Studio is installed or on a different server.

To install PGX service, follow these steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/pgx/server/` directory.
2. Perform the following:
   - If PGX service is to be installed on the same server where Compliance Studio is installed, extract the `pgx-distribution-*-server.zip` file.
   - If PGX service is to be installed on a different server, follow these steps:
     - Copy the `pgx-distribution-*-server.zip` file to the PGX server.
     - Extract the `pgx-distribution-*-server.zip` file.

> **NOTE** The path where the `pgx-distribution-*-server.zip` file is unzipped is referred to as `<PGX_Installation_Path>.`

3. Navigate to the `<PGX_Installation_Path>/pgx/server/conf` directory.

> **NOTE** Configure the following properties if applicable:
>
> In the `server.conf` file, configure the following properties:
>
> - **enable_tls: false,**
> - enable_client_authentication: false
>
> The property value is true by default, which means that the SSL certificate is enabled and recommended. Change to false only if you do not have the SSL certificate enabled.

4. Replace the following Kerberos Files in the `<PGX_Installation_Path>/pgx/server/conf/kerberos` directory:

   ```
   krb5.conf
   keytab file name as mentioned in the config.sh file.
   ```

5. Replace the following Hadoop configuration files in the `<PGX_Installation_Path>/pgx/server/conf/hadoop_cluster` directory:
   - `core-site.xml`
   - `hadoop-env.sh`
   - `hdfs-site.xml`
   - `log4j.properties`
   - `ssl-client.xml`
   - `topology.map`
   - `topology.py`
   - `hive-site.xml`

- `yarn-site.xml`
- `redaction-rules.json`
- `hive-env.sh`
- `mapred-site.xml`
- For additional jars, see Appendix C – Additional Jars – PGX section. Contact your administrator to get the files.

6. Copy all the obtained jars into `<PGX_Installation_path>/server/conf/hdfs_libs` directory.

7. Navigate to the `<PGX_Installation_Path>/pgx/server/bin` directory and configure the `config.sh` file as described in the Table 15:

Table 15: config.sh Parameters

Interaction Variable Name	Significance
`KERBEROS_TICKET_RENEWAL_PERIOD`	For example, 7200 would mean every 2 hours
`KERBEROS_PRINCIPAL`	For example: USER@PRINCIPAL
`KERBEROS_KEYTAB_FILENAME`	For example: fccstudio.keytab
`KRB5_CONFIG_FILENAME`	For example: `krb5.conf`
`PGX_SERVER_OFF_HEAP_MB`	Indicates the maximum off-heap memory size in megabytes (mainly used for storing graphs except for their string properties) that PGX tries to respect. Recommended Value: 42% of the PGX server memory limit size above.
`PGX_SERVER_ON_HEAP_MB`	Indicates the maximum and minimum heap memory size (mainly used for storing graphs' string properties) for the Java process of PGX. Recommended Value: 58% of the PGX server memory limit size above.
`PGX_SERVER_YOUNG_SPACE_MB`	Indicates the amount of young space (new space) configured for the java heap.
`URL_GLOBAL_GRAPH_CONFIG_JSON`	Indicates the URL of the global graph to be pre-loaded. The value can be on HDFS. For example, `hdfs:///user/fccstudio/graph.json`

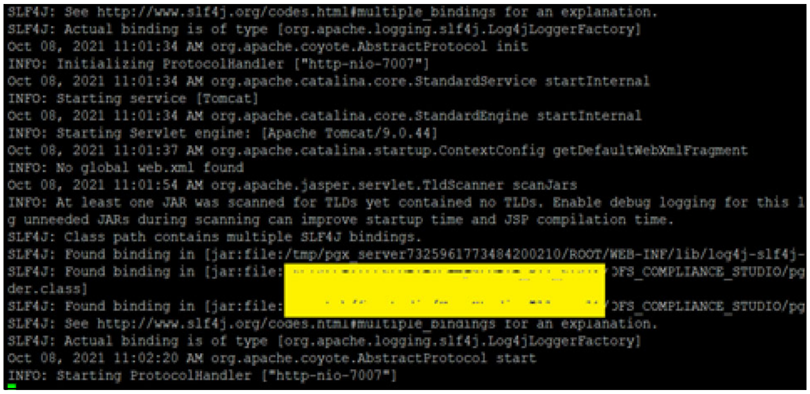
Table 15: config.sh Parameters

PGX_GLOBAL_GRAPH_NAME	Indicates the name that the pre-loaded global graph is published with, and the Compliance Studio users can use it to reference the global graph. For example, GlobalGraphIH
HDFS_GRAPH_FILES_PATH	Indicates the path of the graph files.
MAX_TOTAL_SHARED_DATA_MEMORY_SIZE	The absolute memory limit of shared data (includes published graphs and pinned non-referenced graphs). For example, 20G
MAX_TOTAL_PRIVATE_DATA_MEMORY_SIZE	The memory limit of private data (includes non-published graphs and PGQL results) relative to the total PGX engine memory limit. For example, 8G
MAX_PER_SESSION_DATA_MEMORY_SIZE	Absolute memory limit for any one session of the PGX engine. For example, 700M
MAX_DATA_MEMORY_SIZE_DSADMIN	Absolute memory limit for any user of the PGX engine whose role is DSADMIN. For example, 2G
MAX_DATA_MEMORY_SIZE_DSBATCH	Absolute memory limit for any user of the PGX engine whose role is DSBATCH. For example, 10G
MAX_DATA_MEMORY_SIZE_DSINTER	Absolute memory limit for any user of the PGX engine whose role is DSINTER. For example, 5G
MAX_DATA_MEMORY_SIZE_DSAPPROVER	Absolute memory limit for any user of the PGX engine whose role is DSAPPROVER. For example, 5G
MAX_DATA_MEMORY_SIZE_DSUSER	Absolute memory limit for any user of the PGX engine whose role is DSUSER. For example, 5G
KEYSTORE_FILE_NAME	Indicates keystore file name of Batchservice's certificates.
KEYSTORE_PASS	Indicates keystore password of Batchservice's certificates.

8. Navigate to the `<PGX_Installation_Path>/pgx/server/bin` directory and run the following command:

```
./install.sh
```

Figure 5: PGX start service



```
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Oct 08, 2021 11:01:34 AM org.apache.coyote.AbstractProtocol init
INFO: Initializing ProtocolHandler ["http-nio-7007"]
Oct 08, 2021 11:01:34 AM org.apache.catalina.core.StandardService startInternal
INFO: Starting service [Tomcat]
Oct 08, 2021 11:01:34 AM org.apache.catalina.core.StandardEngine startInternal
INFO: Starting Servlet engine: [Apache Tomcat/9.0.44]
Oct 08, 2021 11:01:37 AM org.apache.catalina.startup.ContextConfig getDefaultWebXmlFragment
INFO: No global web.xml found
Oct 08, 2021 11:01:54 AM org.apache.jasper.servlet.TldScanner scanJars
INFO: At least one JAR was scanned for TLDs yet contained no TLDs. Enable debug logging for this l
g unneeded JARs during scanning can improve startup time and JSP compilation time.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/tmp/pgx_server7325961773484200210/ROOT/WEB-INF/lib/log4j-slf4j-
SLF4J: Found binding in [jar:file:[redacted]OFS_COMPLIANCE_STUDIO/pg
der.class]
SLF4J: Found binding in [jar:file:/OFS_COMPLIANCE_STUDIO/pg
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Oct 08, 2021 11:02:20 AM org.apache.coyote.AbstractProtocol start
INFO: Starting ProtocolHandler ["http-nio-7007"]
```

9. Start the PGX service.

   To start the PGX service, follow these steps:

   a. Navigate to the path where the PGX service is installed.

   b. Navigate to the following directory where the start service for PGX is located:

      `<PGX_Installation_Path>/pgx/server/bin`

   c. Run the following command:

      `“nohup ./start-pgx.sh &”`

10. Stop the PGX service.

    To stop the PGX service, run the following command:

    `./stop-script.sh`

> **NOTE** You must run at least one successful ETL batch to start the PGX service with the `graph.json` file located in the `URL_GLOBAL_GRAPH_CONFIG_JSON` path is present. For more information, see the *Data Movement and Graph Loading for Big Data Environment* section in the OFS Compliance Studio Administration and Configuration Guide.

4.17 Run ER in different workspaces

1. The ER Data Schema and Compliance Studio Schema should be in the same wallet. For more information on how to create a wallet, see Create a wallet for ER schema section.

2. Update the following details for ER schema in the `resources.xml` file. The file can be found in: `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/conf`

   Example:

```
<Resource

            id="ER2_CSA_ABCD"

            name="jdbc/erdataschema"

            auth="Container"

            type="javax.sql.DataSource"
            driverClassName="oracle.jdbc.OracleDriver"
            url="jdbc:oracle:thin:@ER2_CSA_ABCD"
   connectionProperties= "oracle.net.wallet_location
   =<WALLET_PATH/ABCD>;
oracle.net.tns_admin=<WALLET_PATH/ABCD>;"
            maxTotal="5"
            maxIdle="0"
            maxWaitMillis="-1" >
</Resource>
```

NOTE Log in as either an SYS user or DBA user and grant these permissions to the ER schema created.

3. Ensure that the pre-staging and output tables are present in the given ER Data Schema.
   a. The following are the pre-staging table names by version:
      i. FSDF 808:
      — STG_PARTY_MASTER_PRE
      — STG_PARTY_DETAILS_PRE
      — STG_PARTY_EMAIL_ADDRESS_PRE
      — STG_PARTY_ADDRESS_PRE
      — STG_PARTY_PHONE_PRE
      — STG_CUSTOMER_IDENTIFCTN_DOC_PRE
      ii. FSDF 811:
      — STG_PARTY_MASTER_PRE
      — STG_PARTY_DETAILS_PRE
      — STG_PARTY_EMAIL_MAP_PRE
      — STG_ADDRESS_MASTER_PRE
      — STG_PARTY_ADDRESS_MAP_PRE
      — STG_PARTY_PHONE_MAP_PRE
      — STG_CUSTOMER_IDENTIFCTN_DOC_PRE
      iii. FSDF 812:
      — STG_PARTY_MASTER_PRE
      — STG_PARTY_DETAILS_PRE
      — STG_CUSTOMER_IDENTIFCTN_DOC_PRE
      — STG_ADDRESS_MASTER_PRE
      — STG_PARTY_ADDRESS_MAP_PRE
      — STG_PARTY_PHONE_MAP_PRE

- STG_PARTY_EMAIL_MAP_PRE
- FCC_ER_MAPPING
- FCC_ER_MANUAL_MAPPING

b. The following are the output table names by version:

i. **FSDF 808:**

- STG_PARTY_MASTER
- STG_PARTY_DETAILS
- STG_PARTY_EMAIL_ADDRESS
- STG_PARTY_ADDRESS
- STG_PARTY_PHONE
- STG_CUSTOMER_IDENTIFCTN_DOC
- FCC_ER_MAPPING
- FCC_ER_OUTPUT

ii. FSDF 811:

- STG_PARTY_MASTER
- STG_PARTY_DETAILS
- STG_PARTY_EMAIL_MAP
- STG_ADDRESS_MASTER
- STG_PARTY_ADDRESS_MAP
- STG_PARTY_PHONE_MAP
- STG_CUSTOMER_IDENTIFCTN_DOC
- FCC_ER_MAPPING
- FCC_ER_OUTPUT

iii. FSDF 812:

- STG_PARTY_MASTER
- STG_PARTY_DETAILS
- STG_PARTY_EMAIL_MAP
- STG_ADDRESS_MASTER
- STG_PARTY_ADDRESS_MAP
- STG_PARTY_PHONE_MAP
- STG_CUSTOMER_IDENTIFCTN_DOC
- FCC_ER_MAPPING
- FCC_ER_OUTPUT

5 Post-installation Steps when OFSAA is installed

On successful installation of Compliance Studio, you must perform the following post-installation configurations.

ATTENTION For the utility shell script and patch for Security Alert CVE-2021-44228, see Appendix E – Apache Log4j Security Alert CVE-2021-44228 Patch Details section.

Topics:

- Verify the Installation
- Start the PGX Service
- Access the Compliance Studio Application
- Perform the OFSAA Configuration for Batch Execution
- Configure and Run Published Notebooks

NOTE Before running the post-installation steps, an SSH connection to the Big Data server must be configured.

5.1 Verify the Installation

To verify the Compliance Studio installation with OFSAA, check the log files in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory. If all the servers are up and running, it indicates that the installation is complete.

WARNING If you notice any errors in the log files, do not proceed further. Contact My Oracle Support (MOS) provide the applicable error code and log files.

If the installation of Compliance Studio is unsuccessful, you must reinstall the application after performing the cleanup tasks. For more information, see Reinstalling Compliance Studio.

5.2 Start the PGX Service

To start the PGX service, follow these steps:

1. Navigate to the path where the PGX service is installed.
2. Navigate to the following directory where the start service for PGX is located:

   ```
   <PGX_Installation_Path>/pgx/server/bin
   ```

3. Run the following command:

   ```
   "nohup./start-pgx.sh &"
   ```

NOTE Make sure to update the correct location of `graph.json` and `csv files` in `config.sh` inside `<PGX Installation Path>/bin` directory before starting the PGX server.

For more information, see the OFS Compliance Studio Administration and Configuration Guide.

5.3 Access the Compliance Studio Application

To access Compliance Studio, follow these steps:

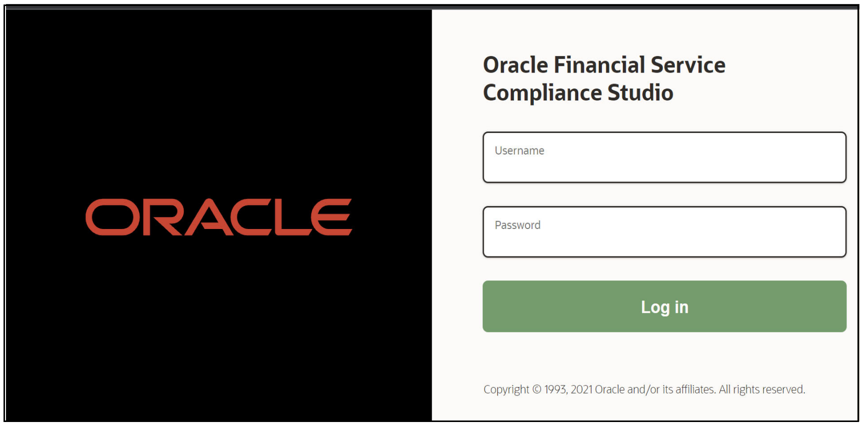
1. Enter the URL in the following format in the web browser:

   `https://<Host_Name>:<Port_Number>/cs/home`

   Here `<Port_Number>` is 7001 for the Compliance Studio application installed on-premise.

   The Compliance Studio application login page is displayed.

   Figure 6: Compliance Studio Application Login Page


ORACLE
Oracle Financial Service
Compliance Studio
Username
Password
Log in
Copyright © 1993, 2021 Oracle and/or its affiliates. All rights reserved.

2. Enter the Username and Password.

   For Creating Users, Groups, and Mappings in AAI. See Appendix F – Create Users, Groups, and Mappings section.

3. Click **Login**.

   After you access the application, you can view the ready-to-use notebooks. To check if you have been assigned any roles, create a notebook. If you cannot create a notebook, contact My Oracle Support (MOS).

5.4 Perform the OFSAA Configuration for Batch Execution

> **NOTE** This configuration is not applicable for Compliance Studio installed without OFSAA.

To perform OFSAA configuration for batch execution, follow these steps:

1. Copy the files in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/bin` directory to the server where the BD or ECM pack is installed and to the `$FIC_DB_HOME/bin` directory of the OFSAA setup.

2. Execute the following command to grant Execute permission to the files:

   `chmod +x <filenames>`

3. Copy all the files from the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/lib` directory into the `$FIC_DB_HOME/lib` directory.

   See the OFS Compliance Studio Administration and Configuration Guide for running Compliance Studio Batches.

5.5 Configure and Run Published Notebooks

NOTE This configuration is not applicable for Compliance Studio installed without OFSAA.

To perform the configuration required to run published notebooks, copy the required `FCCM_Studio_NotebookExecution.sh` file from the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/bin` directory into the `<FIC_HOME of OFSAA_Installed_Path>/deployed/ficdb/bin` directory.

For information on running published notebooks, see the Executing Published Notebook section in the OFS Compliance Studio Administration and Configuration Guide.

6 Post-installation Steps when OFSAA is Not Installed

On successful installation of Compliance Studio, you must perform the following post-installation configurations.

ATTENTION For the utility shell script and patch for Security Alert CVE-2021-44228, see Appendix E – Apache Log4j Security Alert CVE-2021-44228 Patch Details section.

Topics:

- Verify the Installation
- Start the PGX Service
- Access the Compliance Studio Application

NOTE Before running the post-installation steps, an SSH connection to the Big Data server must be configured.

6.1 Verify the Installation

To verify the Compliance Studio installation without OFSAA, check the log files in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory. If all the servers are up and running, it indicates that the installation is complete. Also, ensure all the interpreters are displayed and the JDBC interpreter is working on the Compliance Studio application home page.

WARNING If you notice any errors in the log files, do not proceed further. For additional information, see the Frequently Asked Questions in Compliance Studio section first and Contact My Oracle Support (MOS) provide the applicable error code and log files.

If the installation of Compliance Studio is unsuccessful, you must reinstall the application after performing the cleanup tasks. For more information, see Reinstalling Compliance Studio.

6.2 Start the PGX Service

To start the PGX service, follow these steps:

1. Navigate to the path where the PGX service is installed.
2. Navigate to the following directory where the start service for PGX is located:

   ```
   <PGX_Installation_Path>/pgx/server/bin
   ```

3. Run the following command:

   ```
   "nohup./start-pgx.sh &"
   ```

For more information, see the OFS Compliance Studio Administration and Configuration Guide.

6.3 Access the Compliance Studio Application

To access Compliance Studio, follow these steps:

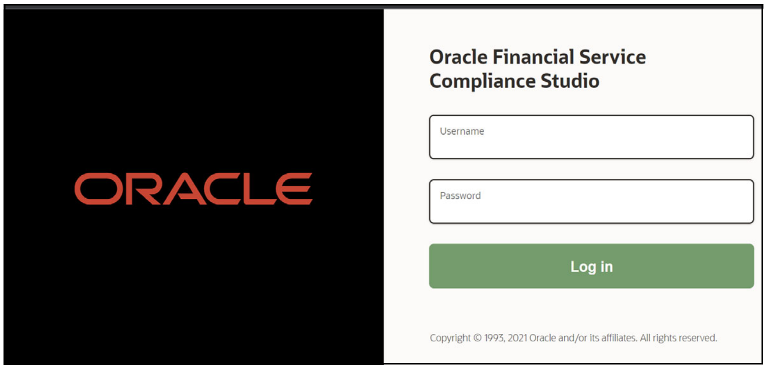
1. Enter the URL in the following format in the web browser:

   `https://<Host_Name>:<Port_Number>/cs/home`

   Here `<Port_Number>` is **7001** for the Compliance Studio application installed on-premise.

   The Compliance Studio application login page is displayed.

   Figure 7: Compliance Studio Application Login Page

   Oracle Financial Service Compliance Studio

   Copyright © 1993, 2021 Oracle and/or its affiliates. All rights reserved.

2. Enter the Username and Password.
3. Click Login.

   After you access the application, you can view the ready-to-use notebooks. To check if you have been assigned any roles, create a notebook. If you cannot create a notebook, contact My Oracle Support (MOS).

7 Upgrade

Follow these steps to upgrade an existing instance of Compliance Studio:

Topics:

- Upgrade Steps with OFSAA
- Pre-Upgrade Steps
- Additional Upgrade Steps
- Cleanup for Upgrade
- Stop the PGX Service
- Stop the Compliance Studio Installer
- Upgrade Steps without OFSAA
- Configure Python Interpreter Setting

You can upgrade an existing instance of Compliance Studio as follows:

Upgrade FCC Studio from v8.0.8.2.0 onwards to Compliance Studio v8.1.2.0.0.

> **NOTE** Here, ensure to provide the same BD database, Studio schema, Hive schema, and wallet-related information you used while installing the existing instance Compliance Studio.

Upgrade FCC Studio from v8.1.1.1.0 onwards to Compliance Studio v8.1.2.0.0.

> **NOTE** Here, ensure to provide the same Compliance Studio schema and wallet-related information you used while installing the existing instance of Compliance Studio.

7.1 Upgrade Steps with OFSAA

This section describes generic steps for the upgrade. For specific upgrades, see Additional Upgrade Steps section.

Table 16 provides the steps to upgrade Compliance Studio with OFSAA.

Table 16: Upgrade Steps with OFSAA

Sl. No.	Activity
Pre-installation Steps	
1	Download the Installer Kit
Installation Steps	
1	Extract the Installer Kit
2	Configure the Elastic Search Component
3	Add Synonyms and Stopword files in Elastic Search
4	Place Files in the Installation Directories

Table 16: Upgrade Steps with OFSAA

5	Generate an Encrypted Password
6	Generate the Public and Private Keys
7	Generate API token for CS API User
8	Generate the Key Store File for Secure Batch Service
9	Configure the Extract Transfer and Load (ETL) Process
10	Configure the config.sh File
11	Run the Compliance Studio Installer
12	Install the PGX Service
Post-Installation Steps	
1	Verify the Installation
2	Stop the PGX Service
3	Stop the Compliance Studio Installer
4	**Configure the SSH Connection** – See OFS Compliance Studio Administration and Configuration Guide
5	**Add the Python Packages to Compliance Studio** - See OFS Compliance Studio Administration and Configuration Guide
6	**Configure the Schema Creation** – See OFS Compliance Studio Administration and Configuration Guide
7	**Configure the ICIJ Data** – See OFS Compliance Studio Administration and Configuration Guide
8	Start the PGX Service
9	Starting Compliance Studio
10	Access the Compliance Studio Application

7.2 Pre-Upgrade Steps

To do pre-upgrade, follow these steps:

1. Stop studio using `./stop-studio.sh` from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin`
2. Stop pgx server. To stop, see Stop the PGX Service.
3. Configure `config.sh` in `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path`
4. Use the same HIVE name and Compliance Studio Schema name during configuration.
5. Install the PGX Service. For more details, see Install the PGX Service.

7.3 Additional Upgrade Steps

This section provides additional steps for upgrade and post-upgrade.

7.3.1 Upgrade from 8.0.8.2.0 to 8.1.2.0.0

In case the user is going to use Graph ETL, below are the steps user needs to follow:

1. Drop the tables starting with FCDM, and ICIJ as the prefix in the HIVE schema.
2. Truncate below tables in studio schema:
   - `fcc_studio_graph_entity_provider;`
   - `fcc_studio_etl_connector_log;`
   - `fcc_studio_etl_graph_log;`
   - `fcc_studio_graph_plug_edge_status;`
3. Remove the jars from `<GRAPH_FILES_PATH >/jars` except `elasticsearch-spark-20_2.11-<Version Number> jar.`
4. Copy all the jars from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/etlJars` to `<GRAPH_FILES_PATH >/jars`.
5. To remove `<HDFS_GRAPH_FILES_PATH>,` run the following command:

   ```
   hadoop fs -rm -r <HDFS_GRAPH_FILES_PATH>
   ```

> **NOTE** You can use **http** or **https** in the command depending upon Elastic search configuration.
>
> If existing indices are not replaced in Elastic Search of **80820 ETL Batch** with new indices, then run the following command to delete existing indices:
>
> ```
> curl -XDELETE http://<Elastic Search hostname>:<port>/
> load-to-elastic-search/idx/deleteIndex/<INDEX NAME>
> ```

7.3.1.1 Upgrade Steps

1. Run below command from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` path to install new compliance studio:

```
2. ./compliance-studio.sh –i
```

3. Run below command from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path to deploy new compliance studio:`

```
4. ./compliance-studio.sh -s
```

7.3.1.2 Post Upgrade Steps

In case the user is going to use Graph ETL, follow the below steps:

1. Run `FCCM_Studio_SchemaCreation.sh` from `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/bin`
2. Run `FCCM_Studio_SchemaCreation.sh` from `<compliance studio installation-path>/deployed/ficdb/bin ONLY ONCE.`
3. WARNING: Do not modify the following tables;

- fcc_datastudio_schemaobjects table in the Studio schema
- fcc_orahive_datatypemapping table in the Atomic Schema

4. Run the Sqoop, ETL Batches, and Graph job.
5. Start PGX server.

7.3.2 Upgrade from 8.1.1.1.0 to 8.1.2.0.0

7.3.2.1 Upgrade Steps

1. Update all the jars in <GRAPH_FILES_PATH>/jars from new compliance studio/deployed/ficdb/etlJars.
2. Run below command from <COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path to install new compliance studio:

   ```
   ./compliance-studio.sh –i
   ```

3. Run below command from <COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path to deploy new compliance studio:

   ```
   ./compliance-studio.sh -s
   ```

4. Pgx can be brought up using <compliance studio installation path>/pgx/server/bin

7.3.3 Upgrade from 8.1.2.0.0 to 8.1.2.0.1

7.3.3.1 Upgrade Steps

1. Update all the jars in <GRAPH_FILES_PATH>/jars from new compliance studio/deployed/ficdb/etlJars.
2. Run below command from <COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path to install new compliance studio:

   ```
   ./compliance-studio.sh –i
   ```

3. Run below command from <COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin path to deploy new compliance studio:

   ```
   ./compliance-studio.sh -s
   ```

4. Pgx can be brought up using <compliance studio installation path>/pgx/server/bin

7.3.3.2 Post-Upgrade Steps

1. Run FCCM_Studio_SchemaCreation.sh from <COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb/bin ONLY ONCE.

   WARNING: Do not modify the following tables;

   c. fcc_datastudio_schemaobjects table in the Studio schema

   d. fcc_orahive_datatypemapping table in the Atomic Schema

2. Run the Sqoop and ETL Batches.
3. Start PGX server.

NOTE You can use http or https in the command depending upon Elastic search configuration.

If existing indices are not replaced in Elastic Search of <Previous version of Compliance Studio> ETL Batch with new indices, then run the following command to delete existing indices:

```
curl -XDELETE http://<Elastic Search hostname>:<port>/
load-to-elastic-search/idx/deleteIndex/<INDEX NAME>
```

7.4 Cleanup for Upgrade

This section provides cleanup steps for the upgrade.

7.4.1 Perform Extract Transfer and Load (ETL) Cleanup

To perform the ETL cleanup, follow these steps:

- Extract the contents of the installer archive file in the download directory using the `unzip -a <Compliance_Studio_Installer_Archive_File>.zip.` The Compliance Studio installer file is extracted in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>` directory.
- Configure the applicable parameters in the `config.sh` file. For more information, see Configure the config.sh File.
- Generate the keystore file. For more information, see Generate the Key Store File for Secure Batch Service.
- Generate an encrypted password. For more information, see Generate an Encrypted Password.

7.4.2 Perform Cleanup for Templates

NOTE This is applicable only if you want to use the new FCGM Default Template. Otherwise, the template will not be updated.

To delete the templates, perform the following:

1. Log in to the Compliance Studio application.
2. Launch the **CS Production** Workspace.
3. Hover the mouse over the **Data Studio Options** widget and Click **Templates.**

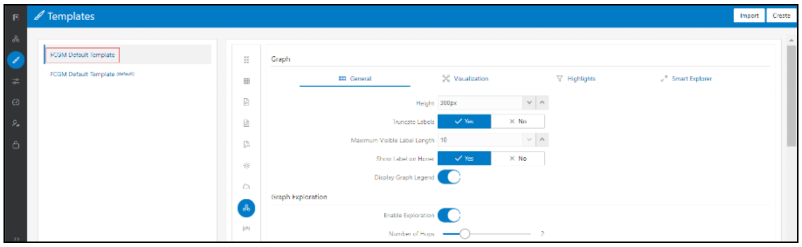
   By default, the Templates page lists all the available templates.

   You can see the following templates among all the templates:

   - FCGM Default Template (default)
   - FCGM Default Template

   You should delete the **FCGM Default Template** that is without **(default).**
4. Click the **FCGM Default Template** on the LHS. The default details are displayed on the RHS:

Figure 8: Template screens



Height: 300px
Truncate Labels: ✓ Yes ✕ No
Maximum Visible Label Length: 10
Show Label on Hover: ✓ Yes ✕ No
Display Graph Legend
Enable Exploration
Number of Hops: 2

5. Click **Delete** on the RHS. A confirmation message is displayed for deletion.
6. Click **Delete**. The template will be deleted.

7.4.3 Perform Cleanup for Interpreters

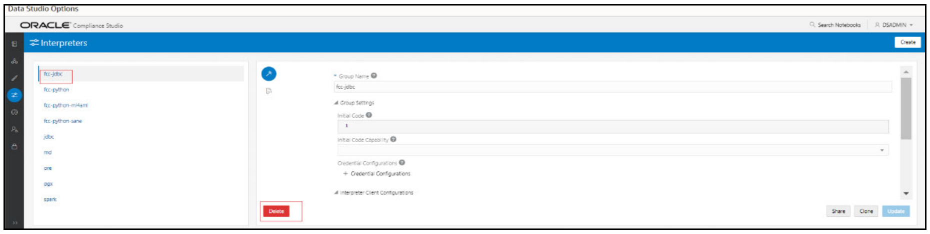
NOTE

- Ensure that the following interpreters are deleted:
  - fcc-jdbc
  - fcc-ore
  - fcc-pyspark
  - fcc-spark-scala
  - fcc-spark-sql
- For 8.1.2.0.0 and later versions, you can rename the fcc interpreter variants in all cases except for the different Python Virtual Environments, so simpler interpreter’s names will be used.
- The steps in this section explain removing the fcc versions before installing the generic versions.
- See **Create an Interpreter Variant** in the OFS Compliance Studio Administration and Configuration Guide on creating new interpreter variants if you want to use the notebooks that use the deleted interpreter name.
- For example, if the notebook has an **fcc-jdbc** paragraph, and these paragraphs' interpreter cannot be replaced with **jdbc**, you can create/clone an interpreter variant of jdbc with the name **fcc-jdbc**.

To delete the interpreter, perform the following:

1. Log in to the Compliance Studio application.
2. Launch the **CS Production** Workspace.
3. Hover the mouse over the **Data Studio Options** widget and Click **Interpreters**.
4. By default, the Interpreters page lists all the available interpreters.
5. Click the **fcc-jdbc** interpreter on the LHS. The default configured interpreter variant is displayed on the RHS:

Figure 9: fcc-jdbc interpreter screens



* Group Name: fcc-jdbc

Group Settings

Initial Code: 1

Initial Code Capability

Credential Configurations

+ Credential Configurations

Interpreter Client Configurations

6. Click **Delete** on the RHS. A confirmation message is displayed for deletion.
7. Click **Delete**. The template will be deleted.
8. Repeat the steps **4, 5,** and **6** for the following interpreters:
    - fcc-ore,
    - fcc-pyspark,
    - fcc-spark-scala
    - fcc-spark-sql

7.4.4 Perform Cleanup for Entity Resolution

You can follow the approach based on the following scenario:

In case of resetting Entity Resolution completely, see the **Resetting Entity Resolution Back to Day 0** section in the OFS Compliance Studio Administration and Configuration Guide.

In case of detailed cleanup steps to continue with Entity Resolution, you can contact My Oracle Support (MOS).

7.5 Stop the PGX Service

To stop the PGX service, follow these steps:

1. Navigate to the path where the PGX service is installed.
2. Navigate to the following directory where the start service for PGX is located:

    ```
    <PGX_Installation_Path>/pgx/server/bin
    ```

3. Run `./stop-script.sh.`

7.6 Stop the Compliance Studio Installer

To stop the Compliance Studio installer, follow these steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin/`directory.
2. Run `./compliance-studio.sh -k`

7.7 Upgrade Steps without OFSAA

Table 17 lists the steps to upgrade Compliance Studio without OFSAA.

Table 17: Upgrade Steps without OFSAA

Sl. No.	Activity
Pre-installation Steps	
1	Download the Installer Kit
Installation Steps	
1	Extract the Installer Kit
2	Place Files in the Installation Directories
3	Generate an Encrypted Password
4	Generate API token for CS API User
5	Generate the Public and Private Keys
6	Generate the Key Store File for Secure Batch Service
7	Configure the config.sh File
8	Run the Compliance Studio Installer
Post-Installation Steps	
1	Stop the Compliance Studio Installer
2	**Add the Python Packages to Compliance Studio** - See OFS Compliance Studio Administration and Configuration Guide
3	Starting Compliance Studio
4	Access the Compliance Studio Application

7.8 Configure Python Interpreter Setting

To use a python interpreter in an upgraded environment, you need to configure the following:

Zeppelin.python:

```
<COMPLIANCE STUDIO INSTALLATION PATH>/deployed/python-packages/
defaultVirtualEnv/bin/python3
```

Initialization:

```
import os; os.environ['TNS_ADMIN'] = '<WALLET_LOCATION>';

from ds_interpreter_client.context.ds_context import PyDataStudioContext

ds = PyDataStudioContext()
```

To configure, perform the following:

1. Login to the Compliance Studio application.

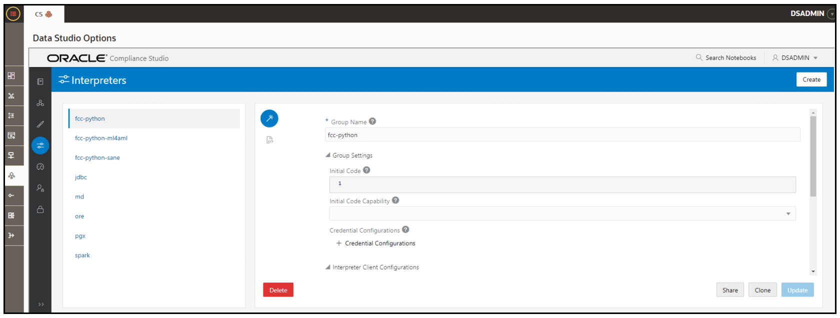
2. Launch the **CS Production** Workspace.

3. Hover the mouse over the Data Studio Options widget and Click **Interpreters**.

4. By default, the Interpreters page lists all the available interpreters.

5. Click the **fcc-python** interpreter on the LHS. The default configured interpreter variant is displayed on the RHS:

Figure 10: fcc-python interpreter screens



* Group Name
fcc-python

Group Settings

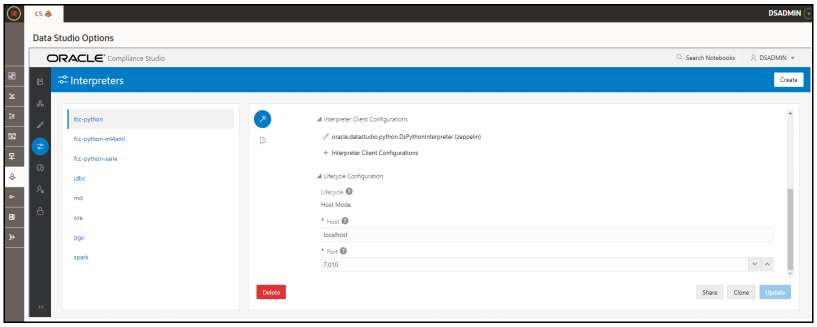
Initial Code
1

Initial Code Capability

Credential Configurations
+ Credential Configurations

Interpreter Client Configurations

Figure 11: Interpreters



Interpreter Client Configurations
oracle.datastudio.python.DsPythonInterpreter (zeppelin)
+ Interpreter Client Configurations

Lifecycle Configuration
Lifecycle
Host Mode

* Host
localhost

* Port
7,010

8 Reinstall Compliance Studio

If the installation of Compliance Studio is unsuccessful, you must reinstall the application after performing the required cleanup tasks.

To reinstall Compliance Studio, follow these steps:

1. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory to update `config.sh` file.
2. Run the following command:

   ```
   ./compliance-studio.sh -k and ./compliance-studio.sh -R
   ```

3. Download and extract the Compliance Studio installer archive file. For more information, see Download the Installer Kit.
4. Perform the database cleanup for the following schemas:

   The following table lists Schemas applicable for cleanup

   Table 18: Schemas applicable for cleanup

Schema	Applicable for Compliance Studio with OFSAA	Applicable for Compliance Studio without OFSAA
Clean up for Compliance Studio Schema	Yes	Yes
Cleanup for BD or ECM Atomic Schema	Yes	No

5. Reinstall Compliance Studio.

Topics:

- Clean up for Compliance Studio Schema
- Cleanup for BD or ECM Atomic Schema

8.1 Clean up for Compliance Studio Schema

To clean up the Studio schema, follow these steps:

1. Drop the existing Compliance Studio schema and create a new Studio schema.

   > **NOTE** The username and password credentials of the Compliance Studio Schema in the wallet files must be updated accordingly. (If applicable)

2. Grant the following permissions to the newly created Oracle Database Schema:
   - `GRANT create session to <schema user>;`
   - `GRANT create table to <schema user>;`
   - `GRANT create view to <schema user>;`
   - `GRANT create any trigger to <schema user>;`
   - `GRANT create any procedure to <schema user>;`
   - `GRANT create sequence to <schema user>;`

- `GRANT execute on dbms_rls to <schema user>;`
- `GRANT execute on sys.dbms_session to <schema user>;`
- `ALTER user <schema user> quota 2000m on <studio tablespace>;`

NOTE Note that the tablespace size can be as per the user's requirement.

- `GRANT create sequence to <schema user>;`
- `GRANT create synonym to <schema user>;`
- `GRANT execute on dbms_redefinition to <schema user>;`
- `GRANT redefine any table to <schema user>;`
- `GRANT create materialized view to <schema user>;`
- `GRANT select on sys.v_$parameter to <schema user>;`
- `GRANT select on sys.dba_free_space to <schema user>;`
- `GRANT select on sys.dba_tables to <schema user>;`
- `GRANT select on sys.dba_tab_columns to <schema user>;`
- `GRANT create rule to <schema user>;`
- `GRANT drop any trigger to <schema user>;`
- `GRANT select on sys.dba_recyclebin to <schema user>;`
- `GRANT create job to <schema user>;`

NOTE

- The **AIF_USER_TS** tablespace will not exist in the BD/ECM in case of the new installation. You can create it manually.

  For example,

  Run the following command to create the tablespace, AIF_USER_TS:
- `CREATE TABLESPACE AIF_USER_TS DATAFILE '<Path of dbf files from table dba_data_files>/aiftestuser.dbf' size 500M;`
- Note that the tablespace size can be as per the user's requirement.

8.2 Cleanup for BD or ECM Atomic Schema

To clean up the BD or ECM Atomic schema, follow these steps:

1. Login to the BD or ECM Atomic Schema.
2. Truncate the `DATABASECHANGELOG` and `DATABASECHANGELOGLOCK` tables using the following command:

```
TRUNCATE TABLE DATABASECHANGELOGLOCK;

TRUNCATE TABLE DATABASECHANGELOG;
```

9 Appendix A - Change Port Numbers for the Applicable Services

Change the number in the applicable files as shown in the following sections to change the port number.

> **NOTE** Only follow this if you want to update the port number of all the service(s).

Topics:

- Server
- Authservice, Batchservice, Metaservice, and Sessionservice
- Interpreter Service
- PGX Service
- Matching Service
- Entity Resolution Service

9.1 Server

To change the port number for the server, go to the **application.yml** file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/datastudio/server/conf/` directory and edit the following values with the new port, for example, 7008:

- `authserviceUrl: "http://<hostname>:<port>/authservice"`
- `metaserviceUrl: "http://<hostname>:<port>/metaservice"`
- `erserviceUrl: "http://<hostname>:<port>"`
- `batchserviceUrl: "https://<hostname>:<port>/batchservice"`
- `mmgServiceUrl: "https://<hostname>:<port>/cs"`

9.2 Authservice, Batchservice, Metaservice, and Sessionservice

To change the port number for the Authservice server, go to the `server-config.properties` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/authservice/conf/` directory and edit the following values with the new port.

- `server.http.port:7041`
- `server.shutdownPort:7042`

Follow this step to make the same changes to the Batchservice, Metaservice, and Sessionservice server.

9.3 Interpreter Service

To change the port number for the Interpreter service, follow these steps:

1. Navigate to the `start-jdbc-interpreter.sh` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/interpreters/bin/` directory and edit `java -DlogFileName=jdbc -Dfile.encoding=UTF-8 ${JAVA_OPTS}`

`${FCC_JDBC_INTERPRETER_OPTS}`
`oracle.datastudio.interpreterserver.ZeppelinRemoteInterpreterServer`
`${1:-7010} > $DIR/../../logs/jdbc.log` with the new port, for example, **7008**.

2. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/interpreters/conf/` directory and update the JSON files with the modified port number.

9.4 PGX Service

To change the port number for the PGX service, go to the `server.conf` file in the `<PGX installation Path>/server/conf/` directory and update the new port number as **7007**.

9.5 Matching Service

To change the port number for the matching service, go to the `application.yml` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/matching-service/conf` directory and update the new port number as **7049**.

9.6 Entity Resolution Service

To change the port number for the entity resolution service, go to the `application.yml` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/entity-resolution/conf` directory and update the new port number **7051**.

10 Appendix B – Spark or PySpark Interpreter

This section provides additional details for Spark or PySpark Interpreter.

Topics:

- Spark Interpreter User Impersonation
- Sample spark-default.conf Configuration File

To set up an additional Spark or PySpark interpreter, for example, to connect to two different external clusters at the same time, follow these steps:

1. Create a start-script for the second Spark interpreter.

> **NOTE** This is an optional step.

e. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/bin` directory and create a new start-script called `start-spark2-interpreter.sh` using the following command:

```
cp start-spark-interpreter.sh start-spark2-interpreter.sh
```

f. Edit the start-spark2-interpreter.sh file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/bin/` directory to update:

iv. Port number to a new port number that is not in use (for example, 7030)

v. Rename the log file, search for the text, .log and give a new name to the log (for example, from spark.log to spark2.log).

g. Edit the `start-all-interpreters.sh` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/interpreters/bin/` directory as follows:

i. Search for the `text sh "$DEPLOY_APP_HOME"/interpreters/bin/start-spark-interpreter.sh &`

ii. Add an additional entry with `sh "$DEPLOY_APP_HOME"/interpreters/bin/start-spark2-interpreter.sh &`

> **NOTE** For the **2nd Spark** interpreter variant, use `start-spark2-interpreter.sh`, when configuring for a 3rd variant, use as `start-spark3-interpreter.sh` etc.

2. Create the interpreter JSON for the additional Spark interpreter.

a. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/conf` directory and create the new interpreter JSON called `spark2.json` using the following command:

```
cp spark.json spark2.json
```

b. Edit the `spark2.json` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/conf/` directory as follows:

i. Update the following parameter values:

```
group: <new-spark-interpreter-name>,
name: <new-spark-interpreter-name>,
groupSettings.initialCodeCapability: <new-spark-interpreter-name>,
port: 7030 (the port chosen in the step 1),
capabilities.name: <new-spark-interpreter-name>,
capabilities.button.label: <new-spark-interpreter-name>,
```

3. After the update, the file will look like the following:

```
[
  {
    "group": "spark",
    "name": "spark",
    "className": "org.apache.zeppelin.spark.SparkInterpreter",
    "groupSettings": {
      "initialCode": "1+1",
      "initialCodeCapability": "spark"
    },
    "host": "localhost",
    "port": 7017,
    "capabilities": [
      {
        "name": "spark",
        "highlightLanguage": "scala",
        "formEscapeCharacter": "@",
        "button": {
          "defaultCode": "println(\"Hello, world\")",
          "icon": "fa fa-fw fa-building-o",
          "label": "Spark"
        }
      }
    ],
    "defaultInterpreter": true,
    "properties": {
      "spark.executor.memory": {
        "envName": null,
        "propertyName": "spark.executor.memory",
        "defaultValue": "",
        "description": "Executor memory per worker instance. ex) 512m,
32g",
        "type": "string"
      },
      "args": {
        "envName": null,
        "propertyName": null,
        "defaultValue": "",
        "description": "spark commandline args",
        "type": "textarea"
      },
      "zeppelin.spark.useHiveContext": {
        "envName": "ZEPPELIN_SPARK_USEHIVECONTEXT",
        "propertyName": "zeppelin.spark.useHiveContext",
        "defaultValue": true,
        "description": "Use HiveContext instead of SQLContext if it is
true.",
        "type": "checkbox"
      },
      "spark.app.name": {
        "envName": "SPARK_APP_NAME",
        "propertyName": "spark.app.name",
        "defaultValue": "Zeppelin",
        "description": "The name of spark application.",
        "type": "string"
      },
      "spark.pyspark.python": {
        "envName": null,
        "propertyName": "spark.pyspark.python",
        "defaultValue": "python3",
        "description": "Python command to run pyspark workers with",
        "type": "string"
      },
      "zeppelin.spark.printREPLOutput": {
        "envName": null,
        "propertyName": "zeppelin.spark.printREPLOutput",
        "defaultValue": true,
        "description": "Print REPL output",
        "type": "checkbox"
      },
      "spark.cores.max": {
        "envName": null,
        "propertyName": "spark.cores.max",
        "defaultValue": "",
        "description": "Total number of cores to use. Empty value uses
all available core.",
        "type": "number"
      },
      "zeppelin.spark.maxResult": {
        "envName": "ZEPPELIN_SPARK_MAXRESULT",
        "propertyName": "zeppelin.spark.maxResult",
        "defaultValue": "1000",
        "description": "Max number of Spark SQL result to display.",
        "type": "number"
      },
      "spark.master": {
        "envName": "MASTER",
        "propertyName": "spark.master",
        "defaultValue": "yarn",
        "description": "Spark master uri. ex) spark://masterhost:7077",
        "type": "string"
      },
      "spark.yarn.archive": {
        "envName": null,
        "propertyName": "spark.yarn.archive",
        "defaultValue": "",
        "description": "An archive containing needed Spark jars for
distribution to the YARN cache",
        "type": "string"
      },
      "spark.driver.bindAddress": {
        "envName": "DRIVER_BIND_ADDRESS",
        "propertyName": "spark.driver.bindAddress",
        "defaultValue": "0.0.0.0",
        "description": "Hostname or IP address where to bind listening
sockets.",
        "type": "string"
      },
      "zeppelin.spark.enableSupportedVersionCheck": {
        "envName": null,
        "propertyName": "zeppelin.spark.enableSupportedVersionCheck",
        "defaultValue": true,
        "description": "Do not change - developer only setting, not for
production use",
        "type": "checkbox"
      },
      "zeppelin.spark.uiWebUrl": {
        "envName": null,
        "propertyName": "zeppelin.spark.uiWebUrl",
        "defaultValue": "",
        "description": "Override Spark UI default URL",
        "type": "string"
      },
      "zeppelin.spark.useNew": {
        "envName": null,
        "propertyName": "zeppelin.spark.useNew",
        "defaultValue": true,
        "description": "Whether use new spark interpreter
implementation",
        "type": "checkbox"
      },
      "zeppelin.spark.ui.hidden": {
        "envName": null,
        "propertyName": "zeppelin.spark.ui.hidden",
        "defaultValue": false,
        "description": "Whether to hide spark ui in zeppelin ui",
        "type": "checkbox"
      },
  "zeppelin.interpreter.output.limit": {
        "envName": null,
        "propertyName": "zeppelin.interpreter.output.limit",
        "defaultValue": "102400",
        "description": "Output message from interpreter exceeding the
limit will be truncated",
        "type": "number"
      }
    },
    "initialCode": [],
    "editor": {
      "language": "scala",
      "editOnDblClick": false
    }
  }
]
```

4. Create the interpreter JSON for the second PySpark interpreter.

   a. Navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/conf` directory and create the new interpreter JSON called `pyspark2.json` using the following command:

      ```
      cp pyspark.json pyspark2.json
      ```

   b. Edit the `pyspark2.json` file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/conf/` directory as follows:

      i. Update the following parameter values:

         ```
         group: <new-spark-interpreter-name>,
         name: <new-spark-interpreter-name>,
         groupSettings.initialCodeCapability: <new-spark-interpreter-name>,
         port: 7030 (the port chosen in the step 1),
         capabilities.name: <new-spark-interpreter-name>,
         capabilities.button.label: <new-spark-interpreter-name>,
         ```

5. After the update, the file will look like the following:

```
[
  {
    "group": "spark",
    "name": "pyspark",
    "className": "org.apache.zeppelin.spark.PySparkInterpreter",
    "host": "localhost",
    "port": 7017,
    "capabilities": [
      {
        "name": "pyspark",
```

```
        "highlightLanguage": "python",
        "button": {
          "defaultCode": "print('Hello World')",
          "icon": "icon-python",
          "label": "PySpark"
        },
        "formEscapeCharacter": "$"
      }
    ],
    "properties": {
      "zeppelin.pyspark.python": {
        "envName": "PYSPARK_PYTHON",
        "propertyName": null,
        "defaultValue": "python3",
        "description": "Python executable to run pyspark with",
        "type": "string"
      },
      "zeppelin.pyspark.useIPython": {
        "envName": null,
        "propertyName": "zeppelin.pyspark.useIPython",
        "defaultValue": false,
        "description": "whether use IPython when it is available",
        "type": "checkbox"
      },
  "zeppelin.interpreter.output.limit": {
        "envName": null,
        "propertyName": "zeppelin.interpreter.output.limit",
        "defaultValue": "102400",
        "description": "Output message from interpreter exceeding the
limit will be truncated",
        "type": "number"
      }
    },
    "initialCode": []
  }
]
```

> **NOTE** If you try to connect two interpreters to different external clusters when setting the environment variables, SPARK_HOME and HADOOP_CONF_DIR, as part of providing custom Spark libraries in Yarn Mode, ensure that you append the environment variables to the respective Spark interpreter start-scripts.

6. Restart Compliance Studio. To do this, navigate to the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin/` directory and run the `./compliance-studio.sh -restart or ./compliance-studio.sh -r script`

10.1 Spark Interpreter User Impersonation

Configure the Spark cluster and Studio to allow proxy users.

Add the below properties and values in `core-site.xml` in the Spark cluster as well as Studio and restart the Spark cluster and Studio:

```
<property>
  <name>hadoop.proxyuser.zeppelin.groups</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.proxyuser.zeppelin.hosts</name>
  <value>*</value>
</property>
```

Configure the Spark interpreter to run the spark-submit job as the currently logged-in user.

Add the below property in `spark.json`:

```
"zeppelin.spark.run.asLoginUser": {
    "envName": null,
    "propertyName": "zeppelin.spark.run.asLoginUser",
    "defaultValue": true,
    "description": "Whether run spark job as the zeppelin login user, it is only applied when running spark job in hadoop yarn cluster and shiro is enabled",
    "type": "checkbox"
}
```

> **NOTE** There will be only a single keytab used by all Spark interpreter runs.

10.2 Sample spark-default.conf Configuration File

Here is the sample code block for creating `spark-default.conf` file:

```
spark.driver.port 30303

spark.blockManager.port 31313

spark.driver.bindAddress 0.0.0.0

spark.yarn.dist.files <COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/
interpreters/interpreter/spark/extralibs/spark-<version>-bin-
hadoop<version>/python/lib/
pyspark.zip,<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/
interpreter/spark/extralibs/spark-<version>-bin-hadoop<version>/python/lib/
py4j-0.10.7-src.zip

spark.executorEnv.PYTHONPATH pyspark.zip:py4j-0.10.7-src.zip

spark.driver.defaultJavaOptions "-Dsun.security.krb5.debug=false -
Djavax.security.auth.useSubjectCredsOnly=false -
Djava.security.krb5.conf=<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/
batchservice/user/conf/krb5.conf"

spark.driver.host <FQDN_HOSTNAME>

spark.yarn.keytab <COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/
batchservice/user/conf/fccstudio.keytab

spark.yarn.principal <KRBS_PRINCIPAL>

spark.yarn.kerberos.relogin.period 1m
```

NOTE

- **FQDN_HOSTNAME** stands for compliance Studio Fully Qualified hostname, and **KRBS_PRINCIPAL** stands for Kerberos principal.
- For example, the Spark version is spark-2.4.0-bin-hadoop2.7.

11 Frequently Asked Questions (FAQs) and Error Dictionary

This section consists of resolutions to the frequently asked questions and error codes noticed during the Compliance Studio installation.

Topics:

- Frequently Asked Questions in Compliance Studio

The Compliance Studio installer performs all the pre-requisite validation checks during installation. Any error encountered in the process is displayed with an appropriate Error Code. You can refer to the Error Dictionary to find the exact cause and resolution to rectify the error.

11.1 Frequently Asked Questions in Compliance Studio

You can refer to the Frequently Asked Questions, which are developed with interest to help you resolve some of the Compliance Studio Installation and configuration issues. This intends to share problem resolution knowledge to a few of the known issues. This is not an official support document and just attempts to share problem resolution knowledge to a few known issues.

1. Why does my console show an unsuccessful message during wallet creation?

   You can check if you have run the following commands correctly. For more information on wallet creation, see Setup Password Stores with Oracle Wallet.

   a. `mkstore -wrl <wallet_location> -create` //creates a wallet in the specified location

   b. `mkstore -wrl <wallet_location> -createCredential <alias-name> <database-user-name>` //creates an alias in the studio schema

   c. `mkstore -wrl <wallet_location> -createCredential <alias-name> <database-user-name>` //creates an alias in the atomic schema

   d. `mkstore -wrl <wallet_location> -createCredential <alias-name> <database-user-name>` //creates an alias in the config schema

   If your issue is still not resolved, contact My Oracle Support (MOS).

2. Where can I find my created wallet?

   Your wallet will be in the directory you have set as your wallet location.

   If your issue is still not resolved, contact My Oracle Support (MOS).

3. When should I create a Database link, and if yes, how do I do it?

   Create a Database link to connect the Atomic and Config database schemas to the Studio database schema if the databases are different. You must create the link in the Studio database.

   In the following example, a link has been created from the config schema to the atomic schema by running the following script:

   ```
   create public database link <studio database link>

   connect to <Config Schema>

   identified by password

   using ' (DESCRIPTION = ADDRESS_LIST = (ADDRESS = (PROTOCOL = TCP) (HOST
   =<host name> (PORT = <port number>)) (CONNECT_DATA = (SERVICE_NAME =
   <service name>))) ';
   ```

```
Config schema : <Config Schema>/password
' (DESCRIPTION = ADDRESS_LIST = (ADDRESS = (PROTOCOL = TCP) (HOST =<host
name> (PORT = <port number>)) (CONNECT_DATA = (SERVICE_NAME = <service
name>))) ';
```

After running the script, run the FCDM connector and ICIJ connector jobs.

4. Why does my installed studio setup not have any notebooks?

   Some default notebooks are ready to use when you install Compliance Studio. If you do not see any notebooks when you log in to the application, you may not be assigned any roles. Check the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory to see if you have been assigned any roles, and if not, contact your administrator.

   If your issue is still not resolved, contact My Oracle Support (MOS).

5. What can I do if the schema creation fails?

   If the Atomic schema creation fails, login to the BD and ECM Atomic schemas and run the following query:

   ```
   select * from fcc_orahive_datatypemapping;
   ```

   The `fcc_orahive_datatypemapping` table must not have duplicate data types.

   If the Compliance Studio schema creation fails, login as a Studio user and run the following query:

   ```
   select * from fcc_datastudio_schemaobjects
   ```

   Run the following query to replace all `Y` values with '':

   ```
     update fcc_datastudio_schemaobjects set SCHEMA_OBJ_GENERATED=''
   ```

   After the schema creation is successful, the value of the `SCHEMA_OBJ_GENERATED` attribute changes to Y.

   You can also check for errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

   If your issue is still not resolved, contact My Oracle Support (MOS).

6. What can I do if the Import_training_model batch execution fails?

   Batch execution status always displays success in case of success or failure.

   You can also check for errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

   According to the log details, you can fix the failure and rerun the same batch.

7. Why is the sqoop job not successful?

   The Sqoop job may fail if some of the applicable values are null or if the service name or SID value is not provided. Do one of the following:

   - Check if there are any null values for the applicable configurations in the `config.sh` and `FCC_DATASTUDIO_CONFIG` tables. If there are any null values, add the required value.
   - Check for any errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

   If your issue is still not resolved, contact My Oracle Support (MOS).

8. Why am I getting the following error when I run the sqoop job?

   ```
   Error: Could not find or load main class
   com.oracle.ofss.fccm.studio.batchclient.client.BatchExecute
   ```

   Set the `FIC_DB_HOME` path in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/ficdb` directory.

   You can also check for any errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

9. Why is the PGX Server not starting?

   The PGX server starts only after the FCDM tables are created after the FCDM connector job is run. Check if all FCDM tables are created, and start the PGX server. You can also check for any errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

   If your issue is still not resolved, contact My Oracle Support (MOS).

10. Why is the ICIJ connector job failing?

    This can happen because of a missing `csv` file path in the FCC_STUDIO_ETL_FILES table. Add the `CSV` file path. You can also check for any errors in the application log file in the `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/logs` directory.

    If your issue is still not resolved, contact My Oracle Support (MOS).

11. What should I do if there is a below Error while selecting edges in manual Decision UI?

    ```
    java.lang.IllegalStateException: Unable to create
    PgxSessionWrapperjava.lang.IllegalStateException: Unable to create
    PgxSessionWrapper at
    oracle.datastudio.interpreter.pgx.CombinedPgxDriver.getOrCreateSession(C
    ombinedPgxDriver.java:147) at
    oracle.pgx.graphviz.driver.PgxDriver.getGraph(PgxDriver.java:334) at
    oracle.pgx.graphviz.library.QueryEnhancer.createEnhancer(QueryEnhancer.j
    ava:223) at
    oracle.pgx.graphviz.library.QueryEnhancer.createEnhancer(QueryEnhancer.j
    ava:209) at
    oracle.pgx.graphviz.library.QueryEnhancer.query(QueryEnhancer.java:150)
    at
    oracle.pgx.graphviz.library.QueryEnhancer.execute(QueryEnhancer.java:136
    ) at
    oracle.pgx.graphviz.interpreter.PgqlInterpreter.interpret(PgqlInterprete
    r.java:131) at
    oracle.datastudio.interpreter.pgx.PgxInterpreter.interpret(PgxInterprete
    r.java:120) at
    org.apache.zeppelin.interpreter.LazyOpenInterpreter.interpret(LazyOpenIn
    terpreter.java:103) at
    org.apache.zeppelin.interpreter.remote.RemoteInterpreterServer$Interpret
    Job.jobRun(RemoteInterpreterServer.java:632) at
    org.apache.zeppelin.scheduler.Job.run(Job.java:188) at
    org.apache.zeppelin.scheduler.FIFOScheduler$1.run(FIFOScheduler.java:140
    ) at java.base/
    java.util.concurrent.Executors$RunnableAdapter.call(Executors.java:515)
    at java.base/java.util.concurrent.FutureTask.run(FutureTask.java:264) at
    java.base/
    java.util.concurrent.ScheduledThreadPoolExecutor$ScheduledFutureTask.run
    ```

```
(ScheduledThreadPoolExecutor.java:304) at java.base/
java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.jav
a:1128) at java.base/
java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.ja
va:628) at java.base/java.lang.Thread.run(Thread.java:834)Caused by:
java.util.concurrent.ExecutionException:
oracle.pgx.common.auth.AuthorizationException: PgxUser(FCCMDSADMIN) does
not own session 6007f00a-8305-4576-9a56-9fa0f061586f or the session does
not exist code: PGX-ERROR-CQAZPV67UM4H at java.base/
java.util.concurrent.CompletableFuture.reportGet(CompletableFuture.java:
395) at java.base/
java.util.concurrent.CompletableFuture.get(CompletableFuture.java:1999)
at oracle.pgx.api.PgxFuture.get(PgxFuture.java:99) at
oracle.pgx.api.ServerInstance.getSession(ServerInstance.java:670) at
oracle.datastudio.interpreter.pgx.CombinedPgxDriver.getOrCreateSession(C
ombinedPgxDriver.java:145) ... 17 moreCaused by:
oracle.pgx.common.auth.AuthorizationException: PgxUser(FCCMDSADMIN) does
not own session 6007f00a-8305-4576-9a56-9fa0f061586f or the session does
not exist code: PGX-ERROR-CQAZPV67UM4H at
oracle.pgx.common.marshalers.ExceptionMarshaler.toUnserializedException(
ExceptionMarshaler.java:107) at
oracle.pgx.common.marshalers.ExceptionMarshaler.unmarshal(ExceptionMarsh
aler.java:123) at
oracle.pgx.client.RemoteUtils.parseExceptionalResponse(RemoteUtils.java:
130) at
oracle.pgx.client.HttpRequestExecutor.executeRequest(HttpRequestExecutor
.java:198) at
oracle.pgx.client.HttpRequestExecutor.get(HttpRequestExecutor.java:165)
at
oracle.pgx.client.RemoteControlImpl$10.request(RemoteControlImpl.java:31
3) at
oracle.pgx.client.RemoteControlImpl$ControlRequest.request(RemoteControl
Impl.java:119) at
oracle.pgx.client.RemoteControlImpl$ControlRequest.request(RemoteControl
Impl.java:110) at
oracle.pgx.client.AbstractAsyncRequest.execute(AbstractAsyncRequest.java
:47) at
oracle.pgx.client.RemoteControlImpl.request(RemoteControlImpl.java:107)
at
oracle.pgx.client.RemoteControlImpl.getSessionInfo(RemoteControlImpl.jav
a:296) at
oracle.pgx.api.ServerInstance.lambda$getSessionInfoAsync$14(ServerInstan
ce.java:490) at java.base/
java.util.concurrent.CompletableFuture.uniComposeStage(CompletableFuture
.java:1106) at java.base/
java.util.concurrent.CompletableFuture.thenCompose(CompletableFuture.jav
a:2235) at oracle.pgx.api.PgxFuture.thenCompose(PgxFuture.java:158)
```

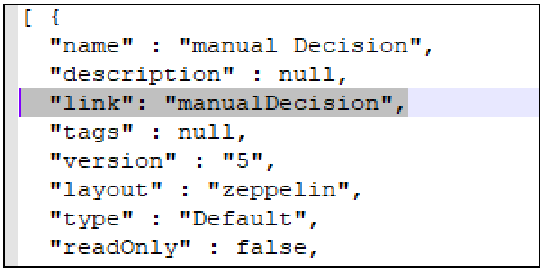
You can perform the following steps as a workaround -

c. Export the "Manual Decision" Notebook

d. Add the link parameter just below Description

   For example, "link": "manual Decision",

Figure 12: link parameter



```
[ {
  "name" : "manual Decision",
  "description" : null,
  "link": "manualDecision",
  "tags" : null,
  "version" : "5",
  "layout" : "zeppelin",
  "type" : "Default",
  "readOnly" : false,
```

e. Truncate the table “fcc_er_paragraph_manual" in Studio Schema.

f. Import the modified notebook again.

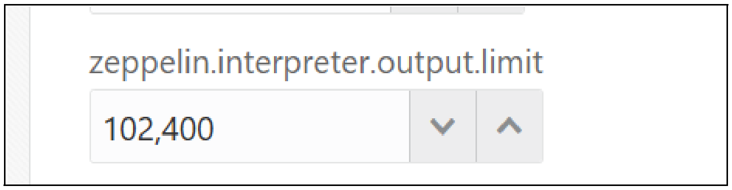
12. What should I do when the result set is truncated if the size goes above '102400' bytes?

a. Perform the following steps:

b. Login to Compliance Studio.

c. Navigate to interpreter zeppelin.interpreter.output.limit.

Figure 13: Interpreter zeppelin parameter



zeppelin.interpreter.output.limit

102,400

d. Set the value to the required size.

e. Restart the Studio Application.

13. What should I do when the spark interpreter is not working?

a. Log in to the server where Compliance Studio is installed.

b. Navigate to $SPARK_HOME directory. If the path is not set, then navigate to <`Compliance Studio_HOME>/deployed/interpreters/interpreter/spark/extralibs/` `spark*`directory.

c. Export the following environment variables:

```
export HADOOP_CONF_DIR=<HADOOP Configuration Directory path>
export SPARK_HOME=<SPARK CLIENT DIRECTORY path>
export SPARK_CONF_DIR=<spark-defaults.conf directory path >
export SPARK_SUBMIT_OPTS="-Djava.security.krb5.conf=<kerberos
directory path>/krb5.conf"
```

d. Run the following commands for specific cases:

— The result of the following command should be Pie value. (It ensures that the client is configured successfully.

```
./bin/run-example --master yarn SparkPi 10
```

— The result of the following command is displayed as a Pie value. (It ensures that the client can successfully connect to the remote cluster

```
./bin/spark-submit  --class org.apache.spark.examples.SparkPi --
master yarn <SPARK_HOME/examples/jars/>/spark-examples_<Ver-
sion>.jar 10
```

For example, in case of spark 2.11-2.4.0, the command is as follows:

```
./bin/spark-submit  --class org.apache.spark.examples.SparkPi --
master yarn <SPARK_HOME/examples/jars/>/spark-examples_2.11-
2.4.0.jar 10
```

— The result of the following command displays the list of databases that exist in HIVE.

```
./bin/spark-submit --class org.apache.spark.sql.hive.thrifts-
erver.SparkSQLCLIDriver --master yarn -e "Show databases"
```

— The result of the following command ensures that the client can query from the HIVE schema.

```
./bin/spark-submit --class org.apache.spark.sql.hive.thrifts-
erver.SparkSQLCLIDriver --master yarn -e "select * from
<hiveSchema>.<tableName> limit 10"
```

14. What should I do when you see the following error in the `spark.log` file?

```
Could not find or load main class
org.apache.spark.deploy.yarn.ExecutorLauncher
```

    a. Log in to the Compliance Studio.
    b. Navigate to Interpreter configurations.
    c. Click on Spark Interpreter.
    d. The `spark.yarn.dist.archives` field value must be empty.

15. What should I do when you see the following error in the `spark.log` file?

```
INFO client.TransportClientFactory: Successfully created connection to
after 105 ms (0 ms spent in bootstraps)
```

```
Exception in thread "main"
java.lang.reflect.UndeclaredThrowableException
at
org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformatio
n.java:1713)
at
org.apache.spark.deploy.SparkHadoopUtil.runAsSparkUser(SparkHadoopUtil.s
cala:64)
at
org.apache.spark.executor.CoarseGrainedExecutorBackend$.run(CoarseGraine
dExecutorBackend.scala:188)
at
org.apache.spark.executor.CoarseGrainedExecutorBackend$.main(CoarseGrain
edExecutorBackend.scala:281)
at
org.apache.spark.executor.CoarseGrainedExecutorBackend.main(CoarseGraine
dExecutorBackend.scala)
Caused by: org.apache.spark.rpc.RpcTimeoutException: Futures timed out
after [120 seconds]. This timeout is controlled by spark.rpc.askTimeout
at
org.apache.spark.rpc.RpcTimeout.org$apache$spark$rpc$RpcTimeout$$createR
```

```
pcTimeoutException(RpcTimeout.scala:47)
at
org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrEl
se(RpcTimeout.scala:62)
at
org.apache.spark.rpc.RpcTimeout$$anonfun$addMessageIfTimeout$1.applyOrEl
se(RpcTimeout.scala:58)
at
scala.runtime.AbstractPartialFunction.apply(AbstractPartialFunction.scal
a:36)
atorg.apache.spark.rpc.RpcTimeout.awaitResult(RpcTimeout.scala:76)
atorg.apache.spark.rpc.RpcEndpointRef.askSync(RpcEndpointRef.scala:92)
atorg.apache.spark.rpc.RpcEndpointRef.askSync(RpcEndpointRef.scala:76)
at
org.apache.spark.executor.CoarseGrainedExecutorBackend$$anonfun$run$1.ap
ply$mcV$sp(CoarseGrainedExecutorBackend.scala:202)
at
org.apache.spark.deploy.SparkHadoopUtil$$anon$2.run(SparkHadoopUtil.scal
a:65)
at
org.apache.spark.deploy.SparkHadoopUtil$$anon$2.run(SparkHadoopUtil.scal
a:64)
atjava.security.AccessController.doPrivileged(NativeMethod)
atjavax.security.auth.Subject.doAs(Subject.java:422)
at
org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformatio
n.java:1698)
```

a. Log in to the Compliance Studio.

b. Navigate to Interpreter configurations.

c. Click on **Spark Interpreter**.

d. The `spark.master` field value must be configured as `yarn`.

e. The `spark.master` should not be set in the `spark-default.conf` file.

16. How can I increase the memory of entity resolution and matching services?

    For more information on increasing memory of entity resolution and matching services, see the **Appendix - Setting Memory of Entity Resolution and Matching Services** in the OFS Compliance Studio Administration and Configuration Guide.

17. What should I do when a runtime error occurs while executing a paragraph in Compliance Studio?

    When Compliance Studio is just started (restart/upgrade/fresh installation), every interpreter gives a runtime error for the first time. Re-run the paragraph to get a result.

    In addition, a user with admin privileges has to run a dummy notebook with a simple paragraph of all the used interpreters once.

18. What should I do if I encounter an error on the login?

    If you log in to Compliance Studio for the first time, log out and log back in to resolve the error.

19. How can I retain the logs after restarting the Compliance Studio?

    a. Log in to the Compliance Studio.

b. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory .

c. Open the `compliance-studio.sh` file and modify the following for service(s) as per your requirement:

   Search with **"$LOGS_FOLDER"** text for each service and add > (Greater than) special character and space before the text as specified below:

```
"$DEPLOY_APP_HOME"/<service name>/bin/<service name> >>
"$LOGS_FOLDER"/<service name>.log
```

   For example, batchservice, entity-resolution

```
function start_services() {
  service=$1
  case $service in
  batchservice)
    export JAVA_OPTS="-Djavax.net.ssl.trustStore=$DEPLOY_APP_HOME/
mmg-home/mmg-studio/conf/<studio server>
    -Djavax.net.ssl.trustStorePassword=$STUDIO_SERVER_SSL_PASSWORD"
    sh "$DEPLOY_APP_HOME"/batchservice/bin/batchservice >>
"$LOGS_FOLDER"/batchservice.log 2>&1 &
    unset JAVA_OPTS
    ;;
  entity-resolution)
    export JAVA_OPTS=<JAVA Options>
    export ER_LOG_PATH="$COMPLIANCE_STUDIO_INSTALLATION_PATH/
deployed"
    export ER_LOG_LEVEL=INFO
    export LD_LIBRARY_PATH="$COMPLIANCE_STUDIO_INSTALLATION_PATH/
deployed/python-packages/saneVirtualEnv/lib/python<version>/site-
packages/jep:$COMPLIANCE_STUDIO_INSTALLATION_PATH/deployed/python-
packages/saneVirtualEnv/lib/":$LD_LIBRARY_PATH
    export PATH_ORG=$PATH
    export PATH=$DEPLOY_APP_HOME/python-packages/saneVirtualEnv/
bin:$PATH
    export TNS_ADMIN=$TNS_ADMIN_PATH
    export PYTHONPATH_ORG=$PYTHONPATH
    export PYTHONPATH="$DEPLOY_APP_HOME"/python-packages/
saneVirtualEnv/lib/python<version>/site-packages:$PYTHONPATH_ORG
    sh "$DEPLOY_APP_HOME"/entity-resolution/bin/entity-resolution >>
"$LOGS_FOLDER"/entity-resolution.log &
    unset JAVA_OPTS
    export PATH=$PATH_ORG
    ;;
```

d. For load to elastic search, you need to add one more > (Greater than) special character as specified below:

```
sh "$DEPLOY_APP_HOME"/load-to-elastic-search/bin/load-to-elastic-search
```

```
>>"$DEPLOY_APP_HOME"/logs/load-to-elastic-search.log &
```

e. Restart Compliance Studio. To do this, run the following command:

```
./compliance-studio.sh -restart
```

Or

```
./compliance-studio.sh -r script
```

20. How to use the system's JDK 8 instead of bundled JDK?

    To use the system's JDK 8 instead of bundled JRE in the Compliance Studio, perform the following.

    a. Set Java home as **JAVA8_HOME** in `.profile` or `.bash_profile`.

    b. Restart Compliance Studio.

> **NOTE** jdk 1.8.0 is the supported version and anything above is not supported.

21. How to update the bundled JDK version?

    Ensure that the Oracle JDK8 should be available in the environment.

    Oracle JDK8 versions details, see Oracle JDK8.

    a. Navigate to `<Compliance Studio Installation Path>/mmg-home/mmg-studio/interpreter-server/pgx-interpreter-bundledJRE-<version>/`

    b. Run the following shell-script, **update-jdk.sh**, with **jdk8_home** and **output_dir** path:

    ```
     ./update-jdk.sh [-j JDK8_HOME ] [-o OUTPUT_DIR]
    ```

    — <JDK8_HOME> specifies the path to the downloaded JDK8

    — <OUTPUT_DIR> where the updated interpreter is saved.

    c. Back up **pgx-interpreter-bundledJRE-<version>** folder.

    d. Copy the **pgx-interpreter** generated inside **<OUTPUT_DIR>** and place it at `<Compliance Studio Installation Path>/mmg-home/mmg-studio/interpreter-server/`

    e. Rename **pgx-interpreter** to p**gx-interpreter-bundledJRE-<version>**.

    f. Install/Re-install Compliance Studio.

> **NOTE** jdk 1.8.0 is the supported version and anything above is not supported.

22. What should I do if the following error message is displayed while starting Compliance Studio services?

    ```
    Java Memory error: unable to create new native thread
    ```

    The user should perform the following steps:

    a. Login to the Linux server as a root user where Compliance Studio is installed.

    b. Open `/etc/security/limits.conf` file.

c. Add the following parameters in the file:

soft nofile 65536

hard nofile 65536

<linux username> soft nproc 10240

@svrtech soft memlock 500000

@svrtech hard memlock 500000

d. Save the file.

e. Restart the Compliance Studio.

23. What should I do if interpreter settings are changed after restarting the Compliance Studio?

To retain the interpreter settings, follow these steps:

a. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/datastudio/server/conf` directory.

b. Open the `application.yml` file and change the value of **overwrite-builtin** to **false** in the interpreter parameter.

> **NOTE** While upgrading Compliance Studio, you should change the value to **true**.

c. Restart Compliance Studio.

24. How to upgrade the python virtual environment for the fcc-python interpreter?

To upgrade, follow these steps:

a. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/bin` directory.

b. Open the `compliance-studio.sh` file and modify the **PYTHONPATH** variable in the default fcc-python interpreter as per your requirement:

For example,

```
export PYTHONPATH=<absolute/path/to/virtual-environment-or-python-
installation-folder/lib/python<version>>/site-
packages:$PYTHONPATH_ORG
```

c. Navigate to `<COMPLIANCE_STUDIO_INSTALLATION_PATH>/deployed/interpreters/bin` directory.

d. Open the `start-fcc-python-interpreter.sh` file and modify the **CLASSPATH** variable as specified below:

```
export CLASSPATH="$DIR/../interpreter/fcc-python/python-interpreter-
21.4.9.jar:$DIR/../interpreter/fcc-python/*:$DIR/../lib/*:$DIR/../
conf"
```

12 Appendix C – Additional Jars – PGX

PGX-Server does not include Hadoop-client for reading graphs from HDFS.

> **NOTE** This section can be skipped if the deployer intends to use only ready to use sample-graph or PGX server without ETL.

When deploying Studio, you must obtain the following libraries. These libraries can be obtained from your existing big data cluster or the internet. The following list of jars is for **Hadoop-client 3.0.0-cdh6.3.0**. These libraries are referred to as '**hdfs-libs**'.

> **NOTE** The following Jar files for your reference. you can use the similar **hdfs-libs** jars based on your Big Data cluster.

Table 19 lists required libraries:

Table 19: List of libraries

accessors-smart-1.2.jar	jaxb-api-2.2.11.jar
aopalliance-1.0.jar	jaxb-impl-2.2.3-1.jar
asm-5.0.4.jar	jcip-annotations-1.0-1.jar
avro-1.8.2-cdh6.3.1.jar	jersey-client-1.19.jar
commons-beanutils-1.9.4.jar	jersey-core-1.19.jar
commons-cli-1.2.jar	jersey-guice-1.19.jar
commons-codec-1.11.jar	jersey-json-1.19.jar
commons-collections-3.2.2.jar	jersey-server-1.19.jar
commons-compress-1.18.jar	jersey-servlet-1.19.jar
commons-configuration2-2.1.1.jar	jettison-1.1.jar
commons-io-2.6.jar	jetty-security-9.3.25.v20180904.jar
commons-lang-2.6.jar	jetty-servlet-9.3.25.v20180904.jar
commons-lang3-3.7.jar	jetty-util-9.3.25.v20180904.jar
commons-logging-1.2.jar	jetty-webapp-9.3.25.v20180904.jar
commons-math3-3.1.1.jar	jetty-xml-9.3.25.v20180904.jar
commons-net-3.1.jar	jline-0.9.94.jar
curator-client-2.12.0.jar	json-smart-2.3.jar
curator-framework-2.12.0.jar	jsp-api-2.1.jar
curator-recipes-2.12.0.jar	jsr305-3.0.0.jar

Table 19: List of libraries

gson-2.2.4.jar	jsr311-api-1.1.1.jar
guava-16.0.1.jar	kerb-admin-1.0.0.jar
guice-4.0.jar	kerb-client-1.0.0.jar
hadoop-annotations-3.0.0-cdh6.3.1.jar	kerb-common-1.0.0.jar
hadoop-auth-3.0.0-cdh6.3.1.jar	kerb-core-1.0.0.jar
hadoop-client-3.0.0-cdh6.3.1.jar	kerb-crypto-1.0.0.jar
hadoop-common-3.0.0-cdh6.3.1.jar	kerb-identity-1.0.0.jar
hadoop-hdfs-client-3.0.0-cdh6.3.1.jar	kerb-server-1.0.0.jar
hadoop-mapreduce-client-common-3.0.0-cdh6.3.1.jar	kerb-simplekdc-1.0.0.jar
hadoop-mapreduce-client-core-3.0.0-cdh6.3.1.jar	kerb-util-1.0.0.jar
hadoop-mapreduce-client-jobclient-3.0.0-cdh6.3.1.jar	kerby-asn1-1.0.0.jar
hadoop-yarn-api-3.0.0-cdh6.3.1.jar	kerby-config-1.0.0.jar
hadoop-yarn-client-3.0.0-cdh6.3.1.jar	kerby-pkix-1.0.0.jar
hadoop-yarn-common-3.0.0-cdh6.3.1.jar	kerby-util-1.0.0.jar
htrace-core4-4.1.0-incubating.jar	kerby-xdr-1.0.0.jar
httpclient-4.5.3.jar	log4j-1.2.17.jar
httpcore-4.4.6.jar	netty-3.7.0.Final.jar
jackson-annotations-2.9.9.jar	nimbus-jose-jwt-4.41.1.jar
jackson-core-2.9.9.jar	okhttp-2.7.5.jar
jackson-core-asl-1.9.13.jar	okio-1.6.0.jar
jackson-databind-2.9.9.3.jar	paranamer-2.8.jar
jackson-jaxrs-1.9.2.jar	protobuf-java-2.5.0.jar
jackson-jaxrs-base-2.9.9.jar	re2j-1.1.jar
jackson-jaxrs-json-provider-2.9.9.jar	slf4j-api-1.7.25.jar
jackson-mapper-asl-1.9.13-cloudera.1.jar	slf4j-log4j12-1.7.25.jar
jackson-module-jaxb-annotations-2.9.9.jar	snappy-java-1.1.4.jar
jackson-xc-1.9.2.jar	stax2-api-3.1.4.jar
javax.activation-api-1.2.0.jar	woodstox-core-5.0.3.jar
javax.inject-1.jar	xz-1.6.jar
javax.servlet-api-3.1.0.jar	zookeeper-3.4.8.jar

13 Appendix D – Additional Jars – Batch Service

When deploying Studio, you must obtain the following files for Batch Service.

> **NOTE** The following Jar files for your reference. you can use the similar **hdfs-libs** jars based on your Big Data cluster.

Table 20 lists the required files:

Table 20: List of Files

accessors-smart-1.2.jar	jersey-server-1.19.jar
activation-1.1.jar	jersey-servlet-1.19.jar
asm-5.0.4.jar	jettison-1.1.jar
avro-1.8.2-cdh6.3.1.jar	jetty-http-9.3.25.v20180904.jar
commons-beanutils-1.9.4.jar	jetty-io-9.3.25.v20180904.jar
commons-cli-1.2.jar	jetty-security-9.3.25.v20180904.jar
commons-codec-1.11.jar	jetty-server-9.3.25.v20180904.jar
commons-collections-3.2.2.jar	jetty-servlet-9.3.25.v20180904.jar
commons-compress-1.18.jar	jetty-util-9.3.25.v20180904.jar
commons-configuration2-2.1.1.jar	jetty-webapp-9.3.25.v20180904.jar
commons-io-2.6.jar	jetty-xml-9.3.25.v20180904.jar
commons-lang-2.6.jar	jline-0.9.94.jar
commons-lang3-3.7.jar	jsch-0.1.54.jar
commons-logging-1.2.jar	json-smart-2.3.jar
commons-math3-3.1.1.jar	jsp-api-2.1.jar
commons-net-3.1.jar	jsr305-3.0.0.jar
curator-client-2.12.0.jar	jsr311-api-1.1.1.jar
curator-framework-2.12.0.jar	kerb-admin-1.0.0.jar
curator-recipes-2.12.0.jar	kerb-client-1.0.0.jar
gson-2.2.4.jar	kerb-common-1.0.0.jar
guava-16.0.1.jar	kerb-core-1.0.0.jar
hadoop-annotations-3.0.0-cdh6.3.1.jar	kerb-crypto-1.0.0.jar
hadoop-auth-3.0.0-cdh6.3.1.jar	kerb-identity-1.0.0.jar
hadoop-common-3.0.0-cdh6.3.1.jar	kerb-server-1.0.0.jar
hive-exec-1.1.0-cdh5.13.0.jar	kerb-simplekdc-1.0.0.jar
HiveJDBC4.jar	kerb-util-1.0.0.jar
hive-metastore-1.1.0-cdh5.13.0.jar	kerby-asn1-1.0.0.jar

Table 20: List of Files

hive-service-1.1.0-cdh5.13.0.jar	kerby-config-1.0.0.jar
htrace-core4-4.1.0-incubating.jar	kerby-pkix-1.0.0.jar
httpclient-4.5.3.jar	kerby-util-1.0.0.jar
httpcore-4.4.6.jar	kerby-xdr-1.0.0.jar
jackson-annotations-2.9.0.jar	log4j-1.2.17.jar
jackson-core-2.9.9.jar	netty-3.7.0.Final.jar
jackson-core-asl-1.9.13.jar	nimbus-jose-jwt-4.41.1.jar
jackson-databind-2.9.9.3.jar	paranamer-2.8.jar
jackson-jaxrs-1.9.2.jar	protobuf-java-2.5.0.jar
jackson-mapper-asl-1.9.13-cloudera.1.jar	re2j-1.1.jar
jackson-xc-1.9.2.jar	slf4j-api-1.7.25.jar
javax.activation-api-1.2.0.jar	slf4j-log4j12-1.7.25.jar
javax.servlet-api-3.1.0.jar	snappy-java-1.1.4.jar
jaxb-api-2.2.2.jar	stax2-api-3.1.4.jar
jaxb-impl-2.2.3-1.jar	stax-api-1.0-2.jar
jcip-annotations-1.0-1.jar	woodstox-core-5.0.3.jar
jersey-core-1.19.jar	xz-1.6.jar
jersey-json-1.19.jar	zookeeper-3.4.8.jar

14 Appendix E – Apache Log4j Security Alert CVE-2021-44228 Patch Details

To address the vulnerability on Apache Log4J v2, Patch 33684394 is released as remediation for a new or upgraded installation of any Compliance Studio Instance.

The patch is based on removing JndiLookup class from the log4j2 jars. To remove this class from the jars in Compliance Studio, perform the following steps:

> **NOTE** The following utilities are required to execute the `studio-patch.sh` script.
>
> - bash
> - tar
> - zip
> - unzip
> - jar

1. Download the shell-script `studio-patch.sh` from Patch 33684394.
2. Place this shell-script in the Compliance Studio Home directory.
3. Grant execute permission by using the command: `chmod +x studio-patch.sh`.
4. Stop Compliance Studio services (including PGX server).
5. Set `STUDIO_HOME` and execute the shell script, where `STUDIO_HOME` is the path where Studio is installed.
   For example:

   ```
   /user/studio/OFS_COMPLIANCE_STUDIO
   ```

6. Set the `STUDIO_HOME` by either of the below options:
   e. Edit the shell-script to update the path as shown below (as applicable):

   ```
   export STUDIO_HOME=/user/studio/OFS_COMPLIANCE_STUDIO
   ```

   f. While execution (use `./studio-patch.sh`) it will ask for Studio Home. The message will be like this:

   ```
   STUDIO_HOME path is not set. Please set it.
   Enter the STUDIO_HOME:
   ```

7. Run `./studio-patch.sh` to execute this shell-script. This will patch the application.
8. Restart Compliance Studio and the PGX server (if applicable).
9. Post-patch Steps:
   a. Refresh the jars in Big data environments for ETL from `STUDIO_HOME/ficdb/etlJars`.
   b. If your PGX server is deployed on another server, refresh it with the PGX server from Studio Home and restart.

15 Appendix F – Create Users, Groups, and Mappings

This section describes how to create users and groups and map groups to the User.

1. Log in to the OFSAAI application as **SYSADMN** user. The landing page is displayed after successful login. See the **Accessing OFSAA Applications** section in OFSAAI User Guide.
2. Navigate to **Identity Management** > **User Maintenance**. The Identity Management window is displayed.

   For more information on adding, updating, and deleting Users, see the **System Configuration and Identity Management** section in the OFSAAI User Guide.

   You can create a new user with the following parameters and select the **EnableUser** and **Login on Holidays** checkboxes:

   - User ID
   - UserName
   - Start Date
   - End Date
   - Password
3. Save the changes and then log out.
4. Log in to the OFSAA application as an **SYSAUTH** user to the Authorize.
5. Log in to the OFSAA application as an **SYSADMN** user.
6. Navigate to **Identity Management** > **User Group Maintenance**.
7. Create Groups using the following names:
   - SANDBOXADM
   - IDNTYADMN
   - IDNTYAUTH
   - MDLUSR
   - MDLREV
   - MDLAPPR
   - WKSPADMIN
   - MDLBATCHUSR
   - DSREDACTGRP

   See the OFS Compliance Studio Administration and Configuration Guide for pre-configured Groups in Compliance Studio.
8. Click **User Group Role Map** and map any AAI available role(s) to the above-created groups.
9. Click **User Group Domain Map** and map the groups to any available Domain(s) in AAI to the above-created groups.
10. Save the changes and then log out.

11. Log in to the OFSAAI application as **SYSAUTH** user to authorize Groups that are created and log out.

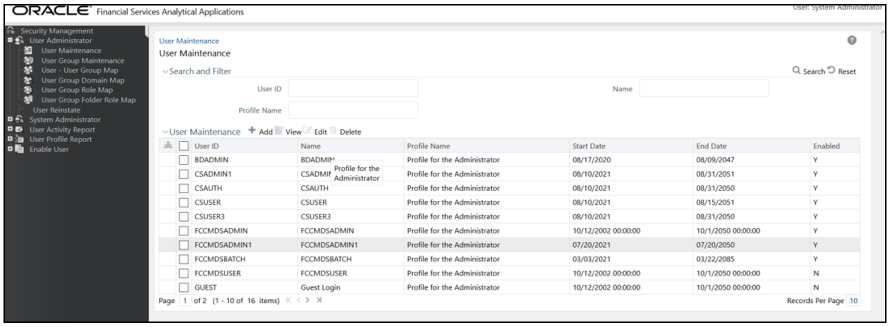
> **NOTE** Roles and Domain mapping are required to authorize Groups only in AAI. These mappings are not significant in the Compliance Studio.

12. Log in to the OFSAAI application as **SYSADMN** user.
13. Navigate to **Identity Management** > **User-User Group Map**.
14. Click on the **User** that is newly created and map the following Groups:
    - SANDBOXADM
    - IDNTYADMN
    - IDNTYAUTH
    - MDLUSR
    - MDLREV
    - MDLAPPR
    - WKSPADMIN
15. Save the changes and then log out.
16. Login to the OFSAAI application as **SYSAUTH** user to authorize the groups and log out.
17. Login to the OFSAAI application as **SYSADMN** user.
18. Navigate to **Identity Management** > **User-User Group** Map to see the Groups mapped to the User.

    For example,

    The following figure illustrates the Creating of User in AAI

Figure 14: Creating of User in AAI



User Maintenance

User ID

Name

Profile Name

	User ID	Name	Profile Name	Start Date	End Date	Enabled
□	BDADMIN	BDADMIN	Profile for the Administrator	08/17/2020	08/09/2047	Y
□	CSADMIN1	CSADMI	Profile for the Administrator	08/10/2021	08/31/2051	Y
□	CSAUTH	CSAUTH	Profile for the Administrator	08/10/2021	08/31/2050	Y
□	CSUSER	CSUSER	Profile for the Administrator	08/10/2021	08/15/2051	Y
□	CSUSER3	CSUSER3	Profile for the Administrator	08/10/2021	08/31/2050	Y
□	FCCMDSADMIN	FCCMDSADMIN	Profile for the Administrator	10/12/2002 00:00:00	10/1/2050 00:00:00	Y
□	FCCMDSADMIN1	FCCMDSADMIN1	Profile for the Administrator	07/20/2021	07/20/2050	Y
□	FCCMDSBATCH	FCCMDSBATCH	Profile for the Administrator	03/03/2021	03/22/2085	Y
□	FCCMDSUSER	FCCMDSUSER	Profile for the Administrator	10/12/2002 00:00:00	10/1/2050 00:00:00	N
□	GUEST	Guest Login	Profile for the Administrator	10/12/2002 00:00:00	10/1/2050 00:00:00	N

Page 1 of 2 (1 - 10 of 16 items)

Records Per Page 10

OFSAA Support

Raise a Service Request (SR) in My Oracle Support (MOS) for queries related to OFSAA applications.

Send Us Your Comments

Oracle welcomes your comments and suggestions on the quality and usefulness of this publication. Your input is an important part of the information used for revision.

- Did you find any errors?
- Is the information clearly presented?
- Do you need more information? If so, where?
- Are the examples correct? Do you need more examples?
- What features did you like most about this manual?

If you find any errors or have any other suggestions for improvement, indicate the title and part number of the documentation along with the chapter/section/page number (if available) and contact the Oracle Support.

Before sending us your comments, you might like to ensure that you have the latest version of the document wherein any of your concerns have already been addressed. You can access My Oracle Support site which has all the revised/recently released documents.

ORACLE®