

Oracle Financial Services Crime and Compliance Studio

Installation and Configuration Guide

Release 8.0.7.3.0

March 2020

E91246.01

ORACLE®
Financial Services

Oracle Financial Services Crime and Compliance Studio

Copyright © 2020 Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are “commercial computer software” pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

For information on third party licenses, click [here](#).

Document Control

Version Number	Revision Date	Changes Done
8.0.7.3.0	Updated: March 2020	<ul style="list-style-type: none"> • A new component called the Entity Resolution is introduced that enables Entity linking in graphs and Entity Searching in notebooks based on multi-attribute name matching using the Elastic Search service. For more information, see "Configuring the Elastic Search Component" on page 16 and "Appendix - Configuring the Elastic Search Component" on page 64. • A new script is introduced for the preparation of ICIJ to resolve data quality issues before graph loading. For more information, see "Cleaning the ICIJ Data" on page 31.
8.0.7.2.0	Updated: February 2020	<ul style="list-style-type: none"> • Install FCC Studio with or without OFSAA. For more information, see "Installing FCC Studio With OFSAA" on page 16 and "Installing FCC Studio With Non-OFSAA" on page 41. • The support for the Data Forward service is deprecated. • Upgrade an existing instance of FCC Studio (v8.0.7.x) to v8.0.7.2.0 version. For more information, see "Upgrading FCC Studio With OFSAA" on page 32. • Configure newly introduced interpreters such as Spark and PySpark interpreters. For more information, see "Interpreter Settings" on page 13.
8.0.7.1.0	Updated: October 2019	<ul style="list-style-type: none"> • Updated the guide for the following newly introduced FCC Studio services: <ul style="list-style-type: none"> ■ ETL ■ Data Forwarding ■ Cross Language Name Matching
8.0.7.0.0	Created: November 2019	Created the first version of the Oracle Financial Services Crime and Compliance Studio Installation Guide for v8.0.7.0.0 Release.

Table of Contents

1	Preface	5
1.1	Summary.....	5
1.2	Audience	5
1.3	Related Documents.....	5
1.4	Abbreviations	5
2	Installation Overview	6
2.1	Introduction.....	6
2.2	Quick Start Steps to Install Studio with OFSAA	6
2.3	Quick Start Steps to Install Studio with Non-OFSAA.....	7
3	Preparing for Installation	8
3.1	Prerequisites.....	8
3.2	Hardware and Software Requirements	8
3.3	Prerequisite Environmental Settings	9
3.4	Performing Common Pre-installation Tasks.....	11
3.4.1	<i>Obtaining the Software</i>	11
3.4.2	<i>Extracting the Software</i>	11
3.5	Required File Structure	12
3.6	Interpreter Settings.....	13
3.6.1	<i>Configuring PySpark Interpreter</i>	13
3.6.2	<i>Configuring Spark Interpreter</i>	14
4	Installing FCC Studio With OFSAA	16
4.1	Configuring the Elastic Search Component	16
4.2	Placing the Required Files	16
4.3	Configuring the config.sh File for FCC Studio	17
4.4	Running the FCC Studio Installer	24
4.5	Verifying the FCC Studio Installation	24
4.6	Verifying the Entity Resolution	25
4.7	Installing PGX Service.....	25
4.8	Accessing the FCC Studio Application.....	27

5	Post-installation Configuration for FCC Studio Installed With OFSAA	28
5.1	Configuring SSH Connection.....	28
5.2	Configuring Interpreters.....	28
5.3	Performing OFSAA Configuration for Batch Execution.....	28
5.4	Performing Hive Data Movement.....	29
5.4.1	<i>Configuring Schema Creation</i>	29
5.4.2	<i>Creating Credential Keystore</i>	30
5.4.3	<i>Configuring Data Movement and Graph Load</i>	30
5.5	Configuring ICIJ.....	30
5.5.1	<i>Cleaning the ICIJ Data</i>	31
5.5.2	<i>Configuring FILEPATH for ICIJ</i>	31
5.6	Configuration for Running Published Notebooks.....	31
6	Upgrading FCC Studio With OFSAA	32
6.1	Upgrade Overview.....	32
6.1.1	<i>Configuring the config.sh File for Upgrade</i>	33
7	Installing FCC Studio With Non-OFSAA.....	41
7.1	Configuring the Elastic Search Component	41
7.2	Placing the Required Files	41
7.3	Configuring the config.sh File for FCC Studio	42
7.4	Running the FCC Studio Installer	49
7.5	Verifying the FCC Studio Installation	50
7.6	Verifying the Entity Resolution	50
7.7	Installing PGX Service.....	50
7.8	Accessing the FCC Studio Application.....	52
8	Post-installation Configuration for Studio Installed With Non-OFSAA.....	53
8.1	Configuring SSH Connection.....	53
8.2	Configuring Interpreters.....	53
8.3	Configuring ICIJ.....	53
8.3.1	<i>Cleaning the ICIJ Data</i>	53
8.3.2	<i>Configuring FILEPATH for ICIJ</i>	54
9	Reinstalling FCC Studio Application	55

9.0.1	<i>Clean up for Studio Schema</i>	55
9.0.2	<i>Cleanup for BD Atomic Schema</i>	56
9.0.3	<i>Cleanup for BD Config Schema</i>	56
10	Appendix - Setting Up Password Stores with Oracle Wallet	57
10.1	Overview.....	57
10.2	Setting Up Password Stores for Database User Accounts	57
10.3	Verifying the Connectivity of the Wallet.....	58
11	Appendix - Installing RServe Manually	60
11.1	Overview.....	60
11.2	Prerequisites.....	60
11.2.1	<i>Installing Oracle R Distribution</i>	60
11.2.2	<i>Installing Dependencies</i>	61
11.2.3	<i>Installing ORE Client</i>	61
11.3	Installing Rserve	61
11.4	Configuring Rserve	61
11.5	Starting Rserve.....	62
11.6	Adding the Certificate to the Keystore	62
11.7	Installing Additional Libraries	63
12	Appendix - Configuring the Elastic Search Component	64
13	Appendix - Starting/Stopping FCC Studio Services	66
13.1	Starting/Stopping PGX Service.....	66
13.2	Starting/Stopping FCC Studio.....	66
14	Appendix - Tables and Sequences	67
14.1	Studio Schema Tables.....	67
14.2	Studio Schema Sequences	67
15	Appendix - Accessing the FCC Studio Application	68
16	Appendix - Using Python Virtual Environments with PySpark	69
16.1	Creating a Virtual Environment with Conda.....	69
16.2	Updating Interpreter Properties.....	69
17	Appendix - Enabling a Second Spark/PySpark Interpreter	70
18	Appendix - Providing Spark Libraries	73

19 Appendix - Executing Graph_Alive Notebook	74
---	-----------

1 Preface

1.1 Summary

Before you begin the installation, ensure that you have access to the Oracle Support Portal with valid login credentials to quickly notify us of any issues at any stage. You can obtain the login credentials by contacting the Oracle Support.

1.2 Audience

This document is intended for System Engineers who are responsible for installing and configuring or upgrading the Oracle Financial Services (OFS) Crime and Compliance Studio (FCC Studio) application.

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

- UNIX commands
- Database concepts
- Big Data

1.3 Related Documents

You can access the following additional documents related to the OFS Crime and Compliance Studio application from the [Oracle Help Center \(OHC\)](#) Documentation Library:

- *Oracle Financial Services Crime and Compliance Studio Deployment Guide (Using Kubernetes)*
- *Oracle Financial Services Crime and Compliance Studio Administration Guide*
- *Oracle Financial Services Crime and Compliance Studio User Guide*
- *Oracle Financial Services Crime and Compliance Studio Data Model Guides*
- *Oracle Financial Services Crime and Compliance Studio Release Notes and Readme*

1.4 Abbreviations

The following table lists the abbreviations used in this document:

Table 1: Abbreviations

Abbreviation	Meaning
OFS	Oracle Financial Services
FCC Studio	Financial Crime and Compliance Studio
OFSAA	Oracle Financial Services Analytical Application
BD	Behavior Detection
ICIJ	International Consortium of Investigative Journalists

2 Installation Overview

This chapter provides the information required to understand the installation of the Oracle Financial Services (OFS) Crime and Compliance Studio (FCC Studio) application with or without OFSAA.

2.1 Introduction

This release (v8.0.7.3.0) of the OFS Crime and Compliance Studio application pack can be used for the following:

- Install a new instance of FCC Studio (v8.0.7.3.0) with or without OFSAA.
- Upgrade an existing instance of FCC Studio (v8.0.7.x) to v8.0.7.3.0 version.

2.2 Quick Start Steps to Install Studio with OFSAA

This section provides the quick start steps to install a new instance of FCC Studio with OFSAA.

Table 1: Quick Start Steps to Install Studio with OFSAA

Sl. No.	Steps	Reference Links
1.	Prepare for installation	<ol style="list-style-type: none"> 1. Prerequisites 2. Hardware and Software Requirements 3. Prerequisite Environmental Settings 4. Performing Common Pre-installation Tasks <ol style="list-style-type: none"> a. Obtaining the Software b. Extracting the Software 5. Required File Structure 6. Interpreter Settings
2.	Install FCC Studio with OFSAA	<ol style="list-style-type: none"> 1. Configuring the Elastic Search Component 2. Placing the Required Files 3. Configuring the config.sh File for FCC Studio 4. Running the FCC Studio Installer 5. Verifying the FCC Studio Installation 6. Verifying the Entity Resolution 7. Installing PGX Service 8. Accessing the FCC Studio Application
3	Post-installation configuration	<ul style="list-style-type: none"> • Configuring SSH Connection • Configuring Interpreters • Performing OFSAA Configuration for Batch Execution • Performing Hive Data Movement • Configuring ICIJ • Configuration for Running Published Notebooks

2.3 Quick Start Steps to Install Studio with Non-OFSAA

This section provides the quick start steps to install a new instance of FCC Studio with Non-OFSAA.

Table 2: Quick Start Steps to Install Studio with Non-OFSAA

Sl. No.	Steps	Reference Links
1.	Prepare for installation	<ol style="list-style-type: none"> 1. Prerequisites 2. Hardware and Software Requirements 3. Prerequisite Environmental Settings 4. Performing Common Pre-installation Tasks <ol style="list-style-type: none"> a. Obtaining the Software b. Extracting the Software 5. Required File Structure 6. Interpreter Settings
2.	Install FCC Studio with non-OFSAA	<ol style="list-style-type: none"> 1. Configuring the Elastic Search Component 2. Placing the Required Files 3. Configuring the config.sh File for FCC Studio 4. Running the FCC Studio Installer 5. Verifying the FCC Studio Installation 6. Verifying the Entity Resolution 7. Installing PGX Service 8. Accessing the FCC Studio Application
3	Post-installation configuration	<ul style="list-style-type: none"> • Configuring SSH Connection • Configuring Interpreters • Configuring ICIJ

3 Preparing for Installation

This chapter provides the information to review before installing the OFS Crime and Compliance Studio (FCC Studio) application. It includes the following sections:

- [Prerequisites](#)
- [Hardware and Software Requirements](#)
- [Prerequisite Environmental Settings](#)
- [Performing Common Pre-installation Tasks](#)
- [Required File Structure](#)
- [Interpreter Settings](#)

3.1 Prerequisites

A new instance of FCC Studio can be installed with or without OFSAA. To install FCC Studio with OFSAA, ensure the BD (Behavior Detection) application pack is installed.

3.2 Hardware and Software Requirements

The hardware and software required to install FCC Studio are as follows:

Table 1: Hardware and Software Requirements

Hardware/Software Category	Component Version
Browser	<ul style="list-style-type: none"> • Chrome 57.x • Firefox 52.x
Java Version	Java 8
Processing Server	<ul style="list-style-type: none"> • RHEL 7.4+ • SFTP • Oracle JRE Standard Edition 1.8.x(with JCE)
Database Server	<ul style="list-style-type: none"> • Oracle Database Server 12c Release 2 (12.2.0.1+) Enterprise Edition • Oracle R Enterprise 1.5.1 with Oracle R Distribution 3.3.0
PGX (Graph) Server	RHEL 7.4+
Elastic Search	elasticsearch-7.3.2
Big Data	
Cloudera Distribution Hadoop 5.12	<ul style="list-style-type: none"> • CDH Version 5.12 • Hadoop-2.5.0+cdh5.3.3+844 • Hive-0.13.1+cdh5.3.3+350 • Sqoop1 V 1.4.4+cdh5.3.3+67 • The <code>.profile</code> file must be present with the <code>SPARK_HOME</code> and <code>PYTHON_HOME</code> parameters already set. • Set <code>spark2-shell</code> alias in the <code>.profile</code> file as follows: <code>alias spark2-shell=spark-shell</code>

Table 1: Hardware and Software Requirements

Hardware/Software Category	Component Version
Cloudera Hive Connectors	Hive JDBC Connectors V 2.5.15
Hadoop Security Protocol	<ul style="list-style-type: none"> • Kerberos R release 1.6.1 • Sentry-1.4.0

3.3 Prerequisite Environmental Settings

The prerequisite environmental settings to be performed before beginning the installation of FCC Studio are as follows:

Table 2: Prerequisite Environmental Settings

Category	Expected Value
PGX Settings	<p>PGX version 20.0.0-RC1</p> <p>Set the following paths in the environment variables in the <code>.profile</code> file:</p> <ul style="list-style-type: none"> • <code>PGX_HOME</code>: Indicates the path of the server where PGX client is installed. For more information, see Installing PGX Service. • <code>PGX_TMP_DIR</code>: Indicates the path of the PGX temporary directory. • <code>SPARK_HOME</code>: Indicates the path where <code>SPARK_HOME</code> is installed by the client.
Java Settings	<ul style="list-style-type: none"> • <code>PATH</code> in the <code>.profile</code> file must be set to include the Java Runtime Environment (Java 8) absolute path. <p>NOTE:</p> <ul style="list-style-type: none"> • Ensure the absolute path to <code>JRE/bin</code> is set at the beginning of the <code>PATH</code> variable. For example: <code>PATH=/usr/java/jre1.8/bin:\$PATH</code> • Ensure no <code>SYMBOLIC</code> links to Java installation are set in the <code>PATH</code> variable.
PGX Server	<ul style="list-style-type: none"> • The following packages must be installed or present in the server where PGX service is installed: <ul style="list-style-type: none"> ■ <code>krb5-libs</code> ■ <code>krb5-workstation</code> ■ <code>procps-ng</code> ■ <code>nc</code> • Execute the following command to install the above-mentioned packages: <pre>yum install -y krb5-libs krb5-workstation procps-ng nc</pre>

Table 2: Prerequisite Environmental Settings

Category	Expected Value
Oracle Data-base Settings	<p>NOTE: This setting is required only if the Wallet has to be created on the same server as that of the FCC Studio server.</p> <p>Oracle Processing Server</p> <ul style="list-style-type: none"> • 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.
Download Directory	Indicates the directory where the product installer zip file is downloaded/copied. The user permission must be set to 755 for this directory.
Installation Directory	<p>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.</p> <p>NOTE: The Installation and the Download Directory can be the same if the product installer zip file is not copied separately to another directory.</p>
OS Locale	<p>Linux: en_US.utf8</p> <p>Execute the following command to check the locale:</p> <pre>locale -a grep -i 'en_US.utf'</pre> <p>The locale is displayed.</p>
Studio Schema	<ol style="list-style-type: none"> 1. Create a new Oracle Database schema user using the following script: <pre>CREATE USER <Studio Schema User Name> IDENTIFIED BY <Password>;</pre> <p>A new Oracle Database schema is created.</p> 2. Grant the permissions given in the next row. <p>This newly created schema is referred to as Studio Schema.</p>

Table 2: Prerequisite Environmental Settings

Category	Expected Value
Oracle Database Schema Settings	<p>Grant the following permissions to the newly created Oracle Database Schema:</p> <pre> GRANT create session TO <Studio Schema User>; GRANT create table TO <Studio Schema User>; GRANT create view TO <Studio Schema User>; GRANT create any trigger TO <Studio Schema User>; GRANT create any procedure TO <Studio Schema User>; GRANT create sequence TO <Studio Schema User>; GRANT execute on dbms_ols TO <Studio Schema User>; GRANT execute on sys.dbms_session TO <Studio Schema User>; ALTER USER <Studio Schema User> QUOTA 100M ON users; GRANT create sequence TO <Studio Schema User>; GRANT create SYNONYM TO <Studio Schema User>; GRANT create any context TO <BD Schema User>; GRANT execute on dbms_ols TO <BD Schema User>; GRANT ALL privileges TO <Studio Schema User>; </pre>
Wallet Settings	<p>Set a password store with Oracle Wallet. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet".</p>

3.4 Performing Common Pre-installation Tasks

The common pre-installation tasks to perform before installing FCC Studio are as follows:

1. [Obtaining the Software](#)
2. [Extracting the Software](#)

3.4.1 Obtaining the Software

To download the FCC Studio application installer software, follow these steps:

1. Login to [My Oracle Support](#) with a valid Oracle account and search for the Patch ID **30995557** under the *Patches & Updates* tab.
2. Download the installer archive `OFS_FCCM_STUDIO_8.0.7.3.0_LINUX.zip` file to the download directory (in Binary Mode) on the setup identified for FCC Studio installation.

3.4.2 Extracting the Software

1. Extract the contents of the `OFS_FCCM_STUDIO_8.0.7.3.0_Linux.zip` installer archive file in the download directory using the following command:

```
unzip -a OFS_FCCM_STUDIO_8.0.7.3.0_LINUX.zip
```

The FCC Studio installer file is extracted and the `OFS_FCCM_STUDIO` folder is obtained and is referred to as `<Studio_Installation_Path>`.

NOTE Do not rename the application installer folder 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 0755 OFS_FCCM_STUDIO -R
```

3.5 Required File Structure

Obtain the following configuration files from the Cloudera installation setup:

NOTE These files must be kept ready and provided in the following file structure that is used during FCC Studio installation.

- Hadoop Cluster
 - `core-site.xml`
 - `hadoop-env.sh`
 - `hdfs-site.xml`
 - `log4j.properties`
 - `ssl-client.xml`
 - `topology.map`
 - `topology.py`
- Kerberos Files
 - `krb5.conf`
 - `ofsaa.keytab`

NOTE Ensure to rename your `.keytab` file to `ofsaa.keytab`.

- Additional Jars
 - `hive-exec-1.1.0-cdh5.13.0.jar`
 - `HiveJDBC4.jar`
 - `hive-metastore-1.1.0-cdh5.13.0.jar`
 - `hive-service-1.1.0-cdh5.13.0.jar`

NOTE

- The version of the jars is client/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.

3.6 Interpreter Settings

NOTE

Perform the following pre-requisite settings only for the interpreters that you need.

Table 3: Interpreter Settings

Interpreter	Prerequisite Settings
fcc-jdbc	No additional configuration is required. NOTE: The FCC Studio application installed with non-OFSAA can use the Vanilla jdbc interpreter instead of the fcc-jdbc interpreter to connect to the Studio schema.
fcc-ore	For the required configuration, see Chapter 11, "Appendix - Installing RServe Manually" .
fcc-pyspark	<ul style="list-style-type: none"> • Install the py4j package in the Spark cluster. • Install the Livy server (0.5.0) on the master node of the Big Data cluster.
fcc-python	Install the py4j package.
fcc-spark-scala	Install the Livy server (0.5.0) on the master node of the Big Data cluster.
fcc-spark-sql	Install the Livy server (0.5.0) on the master node of the Big Data cluster.
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.
pyspark	For the required configuration, see Configuring PySpark Interpreter .
spark	For the required configuration, see Configuring Spark Interpreter

3.6.1 Configuring PySpark Interpreter

- [Prerequisites](#)

- [Configuration](#)

3.6.1.1 Prerequisites

The PySpark interpreter has the same prerequisites as that as the Spark Interpreter. For more information, see [Configuring Spark Interpreter](#). In addition, all Spark components must be configured to use the same Python version.

3.6.1.2 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 FCC Studio application UI during runtime by changing the following properties:

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

To ensure that the two Python versions match, in case your components run on different machines, see [Chapter 16, "Appendix - Using Python Virtual Environments with PySpark"](#).

3.6.2 Configuring Spark Interpreter

- [Prerequisites](#)
- [Configuration](#)

3.6.2.1 Prerequisites

To operate the Spark interpreter in local mode or Yarn mode, the `spark.master` property must be set accordingly. For information on setting the Spark Master properties, see [Chapter 18, "Appendix - Providing Spark Libraries"](#).

- Local Mode
- Yarn Mode

No additional configuration is required to operate the Spark interpreter in local mode.

To operate the Spark interpreter in yarn mode, follow these steps:

- Provide custom Spark libraries. For more information, see [Chapter 18, "Appendix - Providing Spark Libraries"](#).
- The cluster's Hadoop client-side configuration files that include XML files such as `yarn-site.xml` are required and must be supplied with the Spark libraries. These files are available in the Hadoop configuration directory, `HADOOP_CONF_DIR` of the cluster or can be downloaded from the cluster manager's UI if you are using a Cloudera cluster.

3.6.2.2 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 FCC Studio application UI. This includes properties such as `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 can be found in the [Spark's Official Documentation](#). All the properties prefixed with the term "zeppelin", that are listed in the [Zeppelin Spark Configuration Document](#) can also be set from the **Interpreters** page of the FCC Studio application UI.

4 Installing FCC Studio With OFSAA

To install FCC Studio with OFSAA, follow these steps:

1. [Configuring the Elastic Search Component](#)
2. [Placing the Required Files](#)
3. [Configuring the config.sh File for FCC Studio](#)
4. [Running the FCC Studio Installer](#)
5. [Verifying the FCC Studio Installation](#)
6. [Verifying the Entity Resolution](#)
7. [Installing PGX Service](#)
8. [Accessing the FCC Studio Application](#)

4.1 Configuring the Elastic Search Component

Configure the Elastic Search component as per FCC Studio requirements. For more information, see [Chapter 12, "Appendix - Configuring the Elastic Search Component"](#).

4.2 Placing the Required Files

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

1. To place the additional jar files, follow these steps:
 - a. Navigate to the `<Studio_Installation_Path>/batchservice/user/lib` directory.
 - b. Place the following additional jar files:
 - `hive-exec-1.1.0-cdh5.13.0.jar`
 - `HiveJDBC4.jar`
 - `hive-metastore-1.1.0-cdh5.13.0.jar`
 - `hive-service-1.1.0-cdh5.13.0.jar`

NOTE

- The version of the jars is client/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 `<Studio_Installation_Path>/batchservice/user/conf` directory.
 - b. Place the following Kerberos files:
 - `krb5.conf`
 - `ofsa.keytab`
3. To enable the Synonym and Stopword service with the Elastic Search service, follow these steps:

- a. Navigate to the `<Elastic_Search_Installed_Path>/config` directory.
- b. Create a directory named `analysis` using the following command:

```
mkdir analysis
```
- c. Place all the Stopword and Synonym files in the newly created `analysis` directory.

NOTE The filenames must be the same as the column names of the index where you want to enable the Synonym and Stopword services.

For example:

In the index, if the column name is `Country` and you want to enable the Stopword and Synonym service on that column, the file name must be `Country.txt`.

4.3 Configuring the config.sh File for FCC Studio

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

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/bin/` directory.
3. Configure the `config.sh` file as follows:

- NOTE**
- You must manually set the `InteractionVariable` parameter value. If a value is not applicable, enter `NA` and ensure that the value is not entered as `NULL`.
 - Depending on the installation architecture, ensure to provide the correct hostname for URL of PGX service in the `config.sh` file.
 - Do not alter the parameter values that are already set in the `config.sh` file

Table C-1 *config.sh File for FCC Studio*

InteractionVariable Name	Significance	Mandatory
FCC_STUDIO_INSTALLATION_PATH	Indicates the path where the FCC Studio installer file is extracted.	Yes
NON_OFSSAA	Enter "false" to install FCC Studio with OFSAA.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.FCCMRealm</code>	Yes
External Services		

Table C-1 (Continued) config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
OFSAA_SERVICE_URL	Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL. NOTE: For OFSAAAI, the value must be set in the following format: <code>https://<HostName>:<PortNo>/<ContextName>/rest-api</code>	Yes
PGX_SERVER_URL	Indicates the URL of the PGX server. Example: <code>http://<HostName>:<PortNo>/</code> The value for PortNo must be 7007.	Yes
LIVY_HOST_URL	Indicates the URL of the Livy application. The format for the URL is as follows: <code>http://<HostName>:<PortNo></code>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
Internal Services		
AUTH_SERVICE_URL	Indicates the AUTH service URL that gets activated after the <code>fcc-studio.sh</code> file runs. The format for the Auth service URL is as follows: <code>http://<HostName>:7041/authservice</code>	Yes
BATCH_SERVICE_URL	Indicates the batch service URL that gets activated after the <code>fcc-studio.sh</code> file runs. The format for the batch service URL is as follows: <code>http://<HostName>:7043/batchservice</code>	Yes
META_SERVICE_URL	Indicates the metaservice URL that gets activated after the <code>fcc-studio.sh</code> file runs. The format for the metaservice URL is as follows: <code>http://<HostName>:7045/metaservice</code>	Yes
SESSION_SERVICE_URL	Indicates the session service URL that gets activated after the <code>fcc-studio.sh</code> file runs. The format for the session service URL is as follows: <code>http://<HostName>:7047/session-service</code>	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
FCC_STUDIO_URL	Indicates the FCC Studio URL. The format for the FCC Studio URL is as follows: <code>http://<HostName>:7008</code>	Yes
ORE Interpreter Settings		
RSERVE_USERNAME	Indicates the RServe username.	No
RSERVE_PASSWORD	Indicates the RServe password.	No
HTTP_PROXY	Indicates the HTTP proxy.	No
HTTPS_PROXY	Indicates the HTTPS proxy.	No
REPO_CRAN_URL	Indicates the URL from where the R packages are obtained. The format for the REPO_CRAN_URL is as follows: https://cran.r-project.org/ .	No
USERS_LIB_PATH	Indicates the path where the R packages are installed.	No
RSERVE_CONF_PATH	Indicates the path where the <code>Rserve.conf</code> file is present.	No
DB Details for Studio Schema		
STUDIO_DB_HOST-NAME	Indicates the hostname of the database where Studio schema is created.	Yes
STUDIO_DB_PORT	Indicates the port number where Studio schema is created.	Yes
STUDIO_DB_SERVICE_NAME	Indicates the service name of the database where Studio schema is created.	Yes
STUDIO_DB_SID	Indicates the SID of the database where Studio schema is created.	Yes
STUDIO_DB_USERNAME	Indicates the username of the Studio Schema (newly created Oracle Schema).	Yes
STUDIO_DB_PASSWORD	Indicates the password of the Studio schema.	Yes
Studio DB Wallet Details		
STUDIO_ALIAS_NAME	Indicates the Studio alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
STUDIO_WALLET_LOCATION	Indicates the Studio wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
STUDIO_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>STUDIO_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
DB Details for BD Config Schema		

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
BD_CONFIG_HOST-NAME	Indicates the hostname of the database where BD config schema is installed.	Yes
BD_CONFIG_PORT	Indicates the port of the database where BD config schema is installed.	Yes
BD_CONFIG_SERVICE_NAME	Indicates the service name of the database where BD config schema is installed.	Yes
BD_CONFIG_SID	Indicates the SID of the database where BD config schema is installed.	Yes
BD_CONFIG_USERNAME	Indicates the username for the BD config schema.	Yes
BD_CONFIG_PASSWORD	Indicates the password for the BD config schema.	Yes
BD Config Wallet Details		
BD_CONFIG_ALIAS_NAME	Indicates the BD config alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_CONFIG_WALLET_LOCATION	Indicates the BD config wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_CONFIG_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
DB Details for BD Atomic Schema		
BD_ATOMIC_HOST-NAME	Indicates the BD atomic schema hostname.	Yes
BD_ATOMIC_PORT	Indicates the BD atomic schema port number.	Yes
BD_ATOMIC_SERVICE_NAME	Indicates the BD atomic schema service name.	Yes
BD_ATOMIC_SID	Indicates the BD atomic schema SID.	Yes
BD_ATOMIC_USERNAME	Indicates the username of the BD atomic schema.	Yes
BD_ATOMIC_PASSWORD	Indicates the password of the BD atomic schema.	Yes
BD Atomic Wallet Details		
BD-ATOMIC_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
BD_ATOMIC_WALLET_LOCATION	Indicates the BD atomic wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_ATOMIC_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_ATOMIC_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
SQL Scripts		
FSINFODOM	Indicates the name of the BD Infodom.	Yes
FSSEGMENT	Indicates the name of the BD segment.	Yes
DATAMOVEMENT_LINK_NAME	<ul style="list-style-type: none"> If the newly created schema is in a different database host, you must create a DB link and provide the same link in this parameter. If no DB link is present, provide the <code><SCHEMA_NAME></code> in this parameter. If the newly created schema is in the same database host, the value for this parameter is the user name of the BD atomic schema. 	Yes
DATAMOVEMENT_LINK_TYPE	If the DB link is used, enter <code>DBLINK</code> in this field. If the DB link is not used, enter <code>SCHEMA</code> in this field.	Yes
PGX Setup Details		
PGX_INSTALLATION_PATH	Indicates the installation path of the PGX server. Example: <code><STUDIO_INSTALLATION_PATH>/pgx/server/</code>	Yes
PGX_PGB_PATH	Indicates the path where you want to obtain the output graph PGB file. Example for Hive Installation: <code>hdfs:/user/ofsaa</code> Example for Oracle DB Installation: <code>/scratch/ofsaa</code>	Yes
Cloudera Setup Details		
HADOOP_CREDENTIAL_PROVIDER_PATH	Indicates the path where Hadoop credential is stored. Note: Indicates the Hadoop alias given when creating the Hadoop credentials. For information on how to create a credential keystore, see Creating Credential Keystore .	Yes
HADOOP_PASSWORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials.	Yes
Hive_Host_Name	Indicates the Hive hostname.	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
Hive_Port_number	Indicates the Hive port number. Contact System Administrator to obtain the port number.	Yes
HIVE_PRINCIPAL	Indicates the Hive Principal. Contact System Administrator to obtain HIVE_PRINCIPAL.	Yes
HIVE_SCHEMA	Indicates to create a schema in HIVE.	Yes
JAAS_CONF_FILE_PATH	Created for future use.	No
Krb_Host_FQDN_Name	Indicates the Krb host FQDN name.	Yes
Krb_Realm_Name	Indicates the Krb realm name.	Yes
Krb_Service_Name	Indicates the Krb service name. Example: Hive	Yes
KRB5_CONF_FILE_PATH	Created for future use.	No
security_krb5_kdc_server	Created for future use.	No
security_krb5_realm	Created for future use.	No
server_kerberos_keytab_file	Created for future use.	No
server_kerberos_principal	Created for future use.	No
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Host machine where sqoop will run.	Yes
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> 1. Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry: <code>oracle.jdbc.mapDateToTimestamp=false</code> 2. Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsa/</code> <p>NOTE: Ensure that the location name ends with a <code>'/'</code>.</p>	Yes
SQOOP_PARTITION_COL	Indicates the column in which the HIVE table is partitioned. The value must be <code>SNAPSHOT_DT</code> .	Yes
SQOOP_TRG_HOST_NAME	Indicates the hostname of the SQOOP server. Example: <code><HostName></code>	Yes

Table C-1 (Continued) config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
SQOOP_WORK- DIR_HDFS	Indicates the SQOOP working directory in HDFS. Example: /user/ofsaa	Yes
Elastic Search Cluster details		
ELASTIC_- SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
ELASTIC_- SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
Matching Service		
MATCHING_SER- VICE_HOSTNAME	Indicates the hostname where the matching service is installed.	Yes
MATCHING_SER- VICE_PORT	Indicates the port number where the matching service is installed. NOTE: This parameter is same as the SERVER_START_PORT parameter given below.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	Yes
ELASTICRESPON- SE_BUFFERLIMIT_- BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
SCROLL_TIME	Indicates the duration for which the scroll_size output is active. For example: 5	Yes
SCROLL_SIZE	Indicates the amount of data that must be obtained in one attempt when a query is fired on an index in the elastic search service. For example: 1000	Yes
SERVER_START_POR- T	Indicates the port number where matching service is installed. Default Value: 7049 NOTE: This value is already set. Do not change the value.	Yes
Entity Resolution		
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default Value: 7051 NOTE: This value is already set. Do not change the value.	Yes
ER_SERVICE_URL	Indicates the URL of the entity resolution service. For example: http://<HostName>:7051	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
Graphs		
HDFS_GRAPH_FILES_PATH	Indicates the filepath in the HDFS where the <code>graph.json</code> is formed.	Yes
GRAPH_FILES_PATH	Indicates the directory in the Big Data server for graph files.	Yes
GRAPH_NAME	Indicates the name you want to assign to the global graph at the end of ETL.	Yes

4.4 Running the FCC Studio Installer

To run the FCC Studio installer, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/bin/` directory.
2. Execute the following command in the console:

```
./install.sh
```

NOTE

Execution of `install.sh` command does not generate any log file.

3. Execute the following command in the console:

```
./fcc_studio.sh
```

The OFS Crime and Compliance Studio application is installed with OFSAA. The FCC Studio and all the interpreters are started.

After the successful completion of the FCC Studio installation, the script displays a URL that can be used to access the FCC Studio Application. For more information, see [Accessing the FCC Studio Application](#).

4. Verify the FCC Studio installation. See [Verifying the FCC Studio Installation](#)
5. Verify the Entity Resolution. For more information, see [Verifying the Entity Resolution](#).

4.5 Verifying the FCC Studio Installation

To verify the FCC Studio installation, follow these steps:

1. Check the log files in the `<STUDIO_INSTALLATION_PATH>/logs` directory.

If all the servers are up and running, it indicates that the installation is complete.

NOTE

Any error encountered in the process is displayed with an appropriate error code in the log file. Do not proceed with further installation and contact Oracle Support with relevant log files.

If the installation of FCC Studio is unsuccessful, you must reinstall the application after performing the cleanup tasks. For more information, see [Chapter 9, "Reinstalling FCC Studio Application"](#).

4.6 Verifying the Entity Resolution

To verify the installation of the entity resolution service, follow these steps:

1. Enter the URL in the following format into the browser:
`http://<Entity Resolution Hostname>:<Entity Resolution Port>/er/Rule.jsp`
The Entity Resolution page is displayed.
2. Verify if all the out of the box rulesets are displayed in the Entity Resolution page. For more information, see [Financial Crime Graph Model Matching Guide](#).

4.7 Installing PGX Service

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

To install PGX service, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/pgx/server/` directory.
2. Perform the following:
 - If PGX service is to be installed on the same server where FCC Studio is installed, unzip the `pgx-distribution-20.0.0-RC1-server.zip` file.
 - If PGX service is to be installed on a different server, follow these steps:
 - i. Copy the `pgx-distribution-20.0.0-RC1-server.zip` file to the PGX server.
 - ii. Unzip the `pgx-distribution-20.0.0-RC1-server.zip` file.

NOTE The path where the `pgx-distribution-20.0.0-RC1-server.zip` file is unzipped is referred to as `<PGX_Installation_Path>`.

3. Navigate to the `<PGX_Installation_Path>/pgx/server/conf` directory.
4. Configure the following properties as per the requirement:
 - Set the following in the `server.conf` file:

```
enable_tls: false,  
enable_client_authentication: false
```
 - Set the following in the `pgx.conf` file:

```
allow_local_filesystem: true
```

Here, true indicates to enable, and false indicates to disable.
5. Copy and paste the following Kerberos Files into the `<PGX_Installation_Path>/pgx/server/conf/kerberos` directory:
 - `krb5.conf`
 - `ofsaa.keytab`
6. Copy and paste the following Hadoop configuration files into 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

7. Navigate to the <PGX_Installation_Path>/pgx/server/bin and configure the config.sh file as described in the following table:

Table G-1 config.sh Parameters

InteractionVariable Name	Significance
KERBEROS_TICKET_RENEWAL_PERIOD	For example: 7200 would mean every 2 hours
KERBEROS_PRINCIPAL	For example: ofsaa@WHFAYK.ORACLE.COM
KERBEROS_KEYTAB_FILENAME	For example: ofsaa.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 container's memory limit size above. For example: 10240
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 container's memory limit size above. For example: 10240
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/ofsaa/hive_etl_out_80702_pgx/graph.json
PGX_GLOBAL_GRAPH_NAME	Indicates the name that the pre-loaded global graph is published with and the FCC Studio users can use to reference the global graph. For example: GlobalGraphIH

8. Navigate to the <PGX_Installation_Path>/pgx/server/bin directory and execute the following commands:

```
./install.sh
```

9. Start the PGX service. For more information, see [Starting/Stopping PGX Service](#).

NOTE

- You must run at least one successful ETL batch to start the PGX service with the PGX graph configuration.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 Crime and Compliance Studio Administration Guide](#).
- Ensure to execute the Graph_Alive notebook after each time you start/restart the PGX service. For more information, see [Chapter 19, "Appendix - Executing Graph_Alive Notebook"](#).

4.8 Accessing the FCC Studio Application

Access the FCC Studio application. For more information, see [Chapter 15, "Appendix - Accessing the FCC Studio Application"](#).

5 Post-installation Configuration for FCC Studio Installed With OFSAA

On successful installation of FCC Studio with OFSAA, perform the following post-installation configuration:

- [Configuring SSH Connection](#)
- [Configuring Interpreters](#)
- [Performing OFSAA Configuration for Batch Execution](#)
- [Performing Hive Data Movement](#)
- [Configuring ICIJ](#)
- [Configuration for Running Published Notebooks](#)

NOTE

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

5.1 Configuring SSH Connection

To configure SSH connection, follow these steps:

1. Run the following command in the console:
 - a. Run `ssh-keygen`
Generating public/private rsa key pair
 - b. Enter file in which to save the key (<Linux_Home>/`.ssh/id_rsa`):[Press Enter]
 - c. Enter passphrase (empty for no passphrase) : [Press Enter]
 - d. Enter same passphrase again:[Press Enter]
 - e. `ssh-copy-id -i ~/.ssh/id_rsa.pub <BigData_Server>`
 - f. `ssh <BigData_Server>`

5.2 Configuring Interpreters

After starting the FCC Studio application, the configuration for the interpreters can be performed from the user interface (UI). For information on configuring interpreters, see the *Configuring Interpreters* chapter in the [Oracle Financial Services Crime and Compliance Studio Administration Guide](#).

5.3 Performing OFSAA Configuration for Batch Execution

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

1. Copy the files in the `<Studio_Installation_Path>/ficdb/bin` directory to the server where the BD pack is installed and to the `$FIC_DB_HOME/bin` path of the OFSAA setup.
2. Execute the following command to grant Execute permission to the files:
`chmod +x <filenames>`

3. Copy all the files in the `<Studio_Installation_Path>/ficdb/lib` directory to the `$FIC_DB_HOME/lib` directory.

For information on running FCC Studio Batches, see *Managing Studio Batches* chapter in the [Oracle Financial Services Crime and Compliance Studio Administration Guide](#).

5.4 Performing Hive Data Movement

To perform Hive data movement, follow these steps:

- [Configuring Schema Creation](#)
- [Creating Credential Keystore](#)
- [Configuring Data Movement and Graph Load](#)

5.4.1 Configuring Schema Creation

- [Configuring Schema Creation from FCC Studio Server](#)
- [Configuring Schema Creation from OFSAA Server](#)

5.4.1.1 Configuring Schema Creation from FCC Studio Server

To configure Schema Creation from FCC Studio server, follow these steps:

1. Set `FIC_DB_HOME` path to `<Studio_Installation_Path>/ficdb`.

NOTE The `$FIC_DB_HOME` path can be set from the `.profile` file as well.

2. Create a Hive Schema with the name mentioned in the `HIVE_SCHEMA` parameter in the `config.sh` file.

For information on `config.sh` file, see [Configuring the config.sh File for FCC Studio](#).

3. Execute the following shell script in the `<Studio_Installation_Path>/ficdb/bin/` path to create tables in Hive Schema:

```
FCCM_Studio_SchemaCreation.sh HIVE
```

This creates tables in the Hive Schema.

4. Check the `<Studio_Installation_Path>/logs/batchservice.logs` for more information.

5.4.1.2 Configuring Schema Creation from OFSAA Server

To configure Schema Creation from OFSAA server, follow these steps:

1. Copy all the jar files located in the `<Studio_Installation_Path>/ficdb/lib` path to the `<OFSAA_FIC_HOME_PATH>/ficdb/lib` path.
2. Copy all the .sh files located in the `<Studio_Installation_Path>/ficdb/bin` path to the `<OFSAA_FIC_HOME_PATH>/ficdb/bin` path.
3. Create a Hive Schema with the name mentioned in the `HIVE_SCHEMA` parameter in the `config.sh` file.

For information on `config.sh` file, see [Configuring the config.sh File for FCC Studio](#).

4. Execute the following shell script in the `<OFSAA_FIC_HOME_PATH>/ficdb/bin` path to create tables in Hive Schema:

```
FCCM_Studio_SchemaCreation.sh HIVE
```

This creates tables in the Hive Schema.

5. Check the `<Studio_Installation_Path>/logs/batchservice.logs` for more information.

5.4.2 Creating Credential Keystore

To create a credential keystore, follow these steps:

1. Login as HDFS SuperUser.
2. Create a credential keystore on HDFS by executing the following command:

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

3. Verify the credential keystore file by executing the following command:

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

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

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

NOTE

Ensure the credential keystore file path and the alias are correctly provided in the `config.sh` file.

5.4.3 Configuring Data Movement and Graph Load

NOTE

The Big Data System Administrator must place the `batchservice-8.0.7.3.0.jar` file in all nodes of the Spark cluster. Ensure that the path of the jar file is present in the Spark `classpath` in the `spark-defaults.conf` file.

To configure the Data Movement and Graph Load, follow these steps:

1. Copy the required `FCCM_Studio_SqoopJob.sh` files from the `<Studio_Installation_Path>/ficdb/bin` directory to the `<FIC_HOME of OFSAA_Installed_Path>/ficdb/bin` directory.

For information on performing Data Movement and Graph Load, see the *Data Movement and Graph Loading for Big Data Environment* section in the [OFS Crime and Compliance Studio Administration Guide](#).

5.5 Configuring ICIJ

To configure ICIJ (International Consortium of Investigative Journalists), perform the following:

1. [Cleaning the ICIJ Data](#)
2. [Configuring FILEPATH for ICIJ](#)

5.5.1 Cleaning the ICIJ Data

To clean the ICIJ data, follow these steps:

1. Download the four dataset directories from <https://offshoreleaks.icij.org/pages/database>.
2. Unzip the four dataset directories and place the unzipped directories in the <Studio_Installation_Path>/`icij_data_cleaning` directory.
3. Navigate to the <Studio_Installation_Path>/`icij_data_cleaning/bin` directory and execute the following command:

```
./clean.sh
```

NOTE Ensure that Python 3 is available in the machine before executing this command.

After successful execution of the command:

- The cleaned data is available for the sqoop job to load it in Hive and HDFS.
- A directory named "clean" is created inside each of the dataset where a clean version of each CSV file is created.

5.5.2 Configuring FILEPATH for ICIJ

NOTE The FCC Studio graph model is configured to include ICIJ watch list files.

To configure FILEPATH for ICIJ, follow these steps:

1. Place the watch list files in HDFS, that is accessible by the user.
2. Update the FILEPATH of the watch list files in the `fcc_studio_etl_files` table.

Figure 1: `fcc_studio_etl_files` Table

DF_NAME	FILEPATH	DF_SEQ_NO	FILE_ORDER
1 Offshore edges is related to	12	1
2 Bahama External Address	13	1

5.6 Configuration for Running Published Notebooks

To perform the configuration required to run published notebooks, follow these steps:

1. Copy the required `FCCM_Studio_NotebookExecution.sh` file from the <Studio_Installation_Path>/`ficdb/bin` directory to the <FIC_HOME of OFSAA_Installed_Path>/`ficdb/bin` directory.

For information on performing Data Movement and Graph Load, see the *Executing Published Notebook* section in the [OFS Crime and Compliance Studio Administration Guide](#).

6 Upgrading FCC Studio With OFSAA

This chapter provides the information required to understand the upgrade of the FCC Studio application with OFSAA.

6.1 Upgrade Overview

This release (v8.0.7.3.0) of the OFS Crime and Compliance Studio application pack can be used for the following:

- Install a new instance of the FCC Studio application (v8.0.7.3.0).
- Upgrade an existing instance of the FCC Studio application (v8.0.7.x) to the v8.0.7.3.0 version.

NOTE During the upgrade, ensure to provide the same BD database, Studio schema, Hive schema, wallet related information that you used during the installation of the existing instance of the FCC Studio application.

This section provides quick-start steps to upgrade an existing instance of the FCC Studio application (v8.0.7.x) with OFSAA to v8.0.7.3.0 version.

Table 1: Quick-Start Steps to Upgrade FCC Studio with OFSAA

Sl. No.	Steps	Reference Links
1.	Prepare for installation	<p>Note: Ensure that the Prerequisite Environmental Settings are met before proceeding with the upgrade.</p> <ol style="list-style-type: none"> 1. Performing Common Pre-installation Tasks <ol style="list-style-type: none"> a. Obtaining the Software b. Extracting the Software 2. Required File Structure 3. Interpreter Settings
2.	Install FCC Studio with OFSAA	<ol style="list-style-type: none"> 1. Configuring the Elastic Search Component 2. Placing the Required Files 3. Configuring the config.sh File for FCC Studio 4. Running the FCC Studio Installer 5. Verifying the FCC Studio Installation 6. Verifying the Entity Resolution 7. Installing PGX Service 8. Accessing the FCC Studio Application

Table 1: Quick-Start Steps to Upgrade FCC Studio with OFSAA

Sl. No.	Steps	Reference Links
3	Post-installation configuration	<p>NOTE: The post installation configuration that are not performed as part of the existing installation of FCC Studio must be performed during upgrade.</p> <ul style="list-style-type: none"> • Configuring SSH Connection • Configuring Interpreters • Performing OFSAA Configuration for Batch Execution • Performing Hive Data Movement • Configuring ICIJ • Configuration for Running Published Notebooks

6.1.1 Configuring the config.sh File for Upgrade

To configure the `config.sh` file for an upgrade, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/bin/config.sh` file.
3. Configure the `config.sh` file as follows:

NOTE

- You must manually set the `InteractionVariable` parameter value. If a value is not applicable, enter NA and ensure that the value is not entered as NULL.
- Depending on the installation architecture, ensure to enter the correct hostname for URL of PGX service in the `config.sh` file.
- Do not alter the parameter values that are already set in the `config.sh` file
- During the upgrade, ensure to provide the same BD database, Studio schema, Hive schema, and wallet related information in the `config.sh` file, that you used during the installation of the existing instance of the FCC Studio application.

Table C-1 config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
FCC_STUDIO_INSTALLATION_PATH	Indicates the path where the FCC Studio installer file is extracted.	Yes
NON_OFSAA	Enter "false" to install FCC Studio with OFSAA.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.FCCMRealm</code>	Yes

Table C-1 (Continued) config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
External Services		
OFSAA_SERVICE_URL	<p>Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL.</p> <p>NOTE: For OFSAAAI, the value must be set in the following format: <code>https://<HostName>:<PortNo>/<ContextName>/rest-api</code></p>	Yes
PGX_SERVER_URL	<p>Indicates the URL of the PGX server.</p> <p>Example: <code>http://<HostName>:<PortNo>/</code> The value for PortNo must be 7007.</p>	Yes
LIVY_HOST_URL	<p>Indicates the URL of the Livy application.</p> <p>The format for the URL is as follows: <code>http://<HostName>:<PortNo></code></p>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
Internal Services		
AUTH_SERVICE_URL	<p>Indicates the AUTH service URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the Auth service URL is as follows: <code>http://<HostName>:7041/authservice</code></p>	Yes
BATCH_SERVICE_URL	<p>Indicates the batch service URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the batch service URL is as follows: <code>http://<HostName>:7043/batchservice</code></p>	Yes
META_SERVICE_URL	<p>Indicates the metaservice URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the metaservice URL is as follows: <code>http://<HostName>:7045/metaservice</code></p>	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
SESSION_SERVICE_URL	Indicates the session service URL that gets activated after the <code>fcc-studio.sh</code> file runs. The format for the session service URL is as follows: <code>http://<HostName>:7047/session-service</code>	Yes
FCC_STUDIO_URL	Indicates the FCC Studio URL. The format for the FCC Studio URL is as follows: <code>http://<HostName>:7008</code>	Yes
ORE Interpreter Settings		
RSERVE_USERNAME	Indicates the RServe username.	No
RSERVE_PASSWORD	Indicates the RServe password.	No
HTTP_PROXY	Indicates the HTTP proxy.	No
HTTPS_PROXY	Indicates the HTTPS proxy.	No
REPO_CRAN_URL	Indicates the URL from where the R packages are obtained. The format for the REPO_CRAN_URL is as follows: https://cran.r-project.org/ .	No
USERS_LIB_PATH	Indicates the path where the R packages are installed.	No
RSERVE_CONF_PATH	Indicates the path where the <code>Rserve.conf</code> file is present.	No
DB Details for Studio Schema		
STUDIO_DB_HOSTNAME	Indicates the hostname of the database where Studio schema is created.	Yes
STUDIO_DB_PORT	Indicates the port number where Studio schema is created.	Yes
STUDIO_DB_SERVICE_NAME	Indicates the service name of the database where Studio schema is created.	Yes
STUDIO_DB_SID	Indicates the SID of the database where Studio schema is created.	Yes
STUDIO_DB_USERNAME	Indicates the username of the Studio Schema (newly created Oracle Schema).	Yes
STUDIO_DB_PASSWORD	Indicates the password of the Studio schema.	Yes
Studio DB Wallet Details		
STUDIO_ALIAS_NAME	Indicates the Studio alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
STUDIO_WALLET_LOCATION	Indicates the Studio wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
STUDIO_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>STUDIO_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
DB Details for BD Config Schema		
BD_CONFIG_HOST_NAME	Indicates the hostname of the database where BD config schema is installed.	Yes
BD_CONFIG_PORT	Indicates the port of the database where BD config schema is installed.	Yes
BD_CONFIG_SERVICE_NAME	Indicates the service name of the database where BD config schema is installed.	Yes
BD_CONFIG_SID	Indicates the SID of the database where BD config schema is installed.	Yes
BD_CONFIG_USERNAME	Indicates the username for the BD config schema.	Yes
BD_CONFIG_PASSWORD	Indicates the password for the BD config schema.	Yes
BD Config Wallet Details		
BD_CONFIG_ALIAS_NAME	Indicates the BD config alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_CONFIG_WALLET_LOCATION	Indicates the BD config wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_CONFIG_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
DB Details for BD Atomic Schema		
BD_ATOMIC_HOST_NAME	Indicates the BD atomic schema hostname.	Yes
BD_ATOMIC_PORT	Indicates the BD atomic schema port number.	Yes
BD_ATOMIC_SERVICE_NAME	Indicates the BD atomic schema service name.	Yes
BD_ATOMIC_SID	Indicates the BD atomic schema SID.	Yes
BD_ATOMIC_USERNAME	Indicates the username of the BD atomic schema.	Yes
BD_ATOMIC_PASSWORD	Indicates the password of the BD atomic schema.	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
BD Atomic Wallet Details		
BD_- ATOM- IC_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_ATOMIC_WAL- LET_LOCATION	Indicates the BD atomic wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
BD_ATOMIC_TNS_AD- MIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_ATOMIC_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
SQL Scripts		
FSINFODOM	Indicates the name of the BD Infodom.	Yes
FSSEGMENT	Indicates the name of the BD segment.	Yes
DATAMOVE- MENT_LINK_NAME	<ul style="list-style-type: none"> If the newly created schema is in a different database host, you must create a DB link and provide the same link in this parameter. If no DB link is present, provide the <SCHEMA_NAME> in this parameter. If the newly created schema is in the same database host, the value for this parameter is the user name of the BD atomic schema. 	Yes
DATAMOVE- MENT_LINK_TYPE	If the DB link is used, enter DBLINK in this field. If the DB link is not used, enter SCHEMA in this field.	Yes
PGX Setup Details		
PGX_INSTALLATION_- PATH	Indicates the installation path of the PGX server. Example: <STUDIO_INSTALLATION_PATH>/pgx/server/	Yes
PGX_PGB_PATH	Indicates the path where you want to obtain the output graph PGB file. Example for Hive Installation: <code>hdfs:/user/ofsa</code> Example for Oracle DB Installation: <code>/scratch/ofsa</code>	Yes
Cloudera Setup Details		
HADOOP_CREDEN- TIAL_PROVIDER_- PATH	Indicates the path where Hadoop credential is stored. Note: Indicates the Hadoop alias given when creating the Hadoop credentials. For information on how to create a credential keystore, see Creating Credential Keystore .	Yes

Table C-1 (Continued) config.sh File for FCC Studio

Interaction Variable Name	Significance	Mandatory
HADOOP_PASSWORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials.	Yes
Hive_Host_Name	Indicates the Hive hostname.	Yes
Hive_Port_number	Indicates the Hive port number. Contact System Administrator to obtain the port number.	Yes
HIVE_PRINCIPAL	Indicates the Hive Principal. Contact System Administrator to obtain HIVE_PRINCIPAL.	Yes
HIVE_SCHEMA	Indicates to create a schema in HIVE.	Yes
JAAS_CONF_FILE_PATH	Created for future use.	No
Krb_Host_FQDN_Name	Indicates the Krb host FQDN name.	Yes
Krb_Realm_Name	Indicates the Krb realm name.	Yes
Krb_Service_Name	Indicates the Krb service name. Example: Hive	Yes
KRB5_CONF_FILE_PATH	Created for future use.	No
security_krb5_kdc_server	Created for future use.	No
security_krb5_realm	Created for future use.	No
server_kerberos_keytab_file	Created for future use.	No
server_kerberos_principal	Created for future use.	No
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Host machine where sqoop will run.	Yes
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> 1. Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry: <code>oracle.jdbc.mapDateToTimestamp=false</code> 2. Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsaas/</code> <p>NOTE: Ensure that the location name ends with a <code>'/'</code>.</p>	Yes
SQOOP_PARTITION_COL	Indicates the column in which the HIVE table is partitioned. The value must be <code>SNAPSHOT_DT</code> .	Yes

Table C-1 (Continued) config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
SQOOP_TRG_HOST-NAME	Indicates the hostname of the SQOOP server. Example: <HostName>	Yes
SQOOP_WORK- DIR_HDFS	Indicates the SQOOP working directory in HDFS. Example: /user/ofsaa	Yes
Elastic Search Cluster details		
ELASTIC_- SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
ELASTIC_- SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
Matching Service		
MATCHING_SER- VICE_HOSTNAME	Indicates the hostname where the matching service is installed.	Yes
MATCHING_SER- VICE_PORT	Indicates the port number where the matching service is installed. NOTE: This parameter is same as the SERVER_START_PORT parameter given below.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	Yes
ELASTICRESPON- SE_BUFFERLIMIT_- BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
SCROLL_TIME	Indicates the duration for which the scroll_size output is active. For example: 5	Yes
SCROLL_SIZE	Indicates the amount of data that must be obtained in one attempt when a query is fired on an index in the elastic search service. For example: 1000	Yes
SERVER_START_POR T	Indicates the port number where matching service is installed. Default Value: 7049 NOTE: This value is already set. Do not change the value.	Yes
Entity Resolution		
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default Value: 7051 NOTE: This value is already set. Do not change the value.	Yes

Table C-1 (Continued) config.sh File for FCC Studio

InteractionVariable Name	Significance	Mandatory
ER_SERVICE_URL	Indicates the URL of the entity resolution service. For example: <code>http://<HostName>:7051</code>	Yes
Graphs		
HDFS_GRAPH_FILES_PATH	Indicates the filepath in HDFS where the <code>graph.json</code> is formed.	Yes
GRAPH_FILES_PATH	Indicates the directory in the Big Data server for graph files.	Yes
GRAPH_NAME	Indicates the name you want to assign to the global graph at the end of ETL.	Yes

7 Installing FCC Studio With Non-OFSAA

This chapter provides the necessary information to install FCC Studio with non-OFSAA. It includes the following topics:

1. [Configuring the Elastic Search Component](#)
2. [Placing the Required Files](#)
3. [Configuring the config.sh File for FCC Studio](#)
4. [Running the FCC Studio Installer](#)
5. [Verifying the FCC Studio Installation](#)
6. [Verifying the Entity Resolution](#)
7. [Installing PGX Service](#)
8. [Accessing the FCC Studio Application](#)

7.1 Configuring the Elastic Search Component

Configure the Elastic Search component as per FCC Studio requirements. For more information, see [Chapter 12, "Appendix - Configuring the Elastic Search Component"](#).

7.2 Placing the Required Files

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

1. To place the additional jar files, follow these steps:
 - a. Navigate to the `<Studio_Installation_Path>/batchservice/user/lib` directory.
 - b. Place the following additional jar files:
 - `hive-exec-1.1.0-cdh5.13.0.jar`
 - `HiveJDBC4.jar`
 - `hive-metastore-1.1.0-cdh5.13.0.jar`
 - `hive-service-1.1.0-cdh5.13.0.jar`

NOTE

- The version of the jars is client/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 `<Studio_Installation_Path>/batchservice/user/conf` directory.
 - b. Place the following Kerberos files:
 - `krb5.conf`
 - `ofsaa.keytab`

3. To enable the Synonym and Stopword service with the Elastic Search service, follow these steps:
 - a. Navigate to the `<Elastic_Search_Installed_Path>/config` directory.
 - b. Create a directory named `analysis` using the following command:


```
mkdir analysis
```
 - c. Place all the Stopword and Synonym files in the newly created `analysis` directory.

NOTE The filenames must be the same as the column names of the index where you want to enable the Synonym and Stopword services.

For example:

In the index, if the column name is `Country` and you want to enable the Stopword and Synonym service on that column, the file name must be `Country.txt`.

7.3 Configuring the config.sh File for FCC Studio

To configure the `config.sh` file, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/bin/config.sh` directory.
3. Configure the `config.sh` file as follows:

- NOTE**
- You must manually set the `InteractionVariable` parameter value. If a value is not applicable, enter `NA` and ensure that the value is not entered as `NULL`.
 - Depending on the installation architecture, ensure to enter the correct hostname for URL of PGX service in the `config.sh` file.
 - Do not alter the parameter values that are already set in the `config.sh` file
 - Retain the existing placeholder values for the non-mandatory parameters.

Table C-1 *config.sh Parameters for FCC Studio*

InteractionVariable Name	Significance	Mandatory
FCC_STUDIO_INSTALLATION_PATH	Indicates the path where the FCC Studio installer file is extracted.	Yes
NON_OFSAA	Enter "true" to install FCC Studio with non-OFSAA.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.DemoRealm</code>	Yes

Table C-1 (Continued) config.sh Parameters for FCC Studio

InteractionVariable Name	Significance	Mandatory
External Services		
OFSAA_SERVICE_URL	<p>Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL.</p> <p>NOTE: For OFSAAAI, the value must be set in the following format: <code>https://<HostName>:<PortNo>/<ContextName>/rest-api</code></p>	No
PGX_SERVER_URL	<p>Indicates the URL of the PGX server.</p> <p>Example: <code>http://<HostName>:<PortNo>/</code> The value for PortNo must be 7007.</p>	Yes
LIVY_HOST_URL	<p>Indicates the URL of the Livy application.</p> <p>The format for the URL is as follows: <code>http://<HostName>:<PortNo></code></p>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
Internal Services		
AUTH_SERVICE_URL	<p>Indicates the AUTH service URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the Auth service URL is as follows: <code>http://<HostName>:7041/authservice</code></p>	Yes
BATCH_SERVICE_URL	<p>Indicates the batch service URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the batch service URL is as follows: <code>http://<HostName>:7043/batchservice</code></p>	Yes
META_SERVICE_URL	<p>Indicates the metaservice URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the metaservice URL is as follows: <code>http://<HostName>:7045/metaservice</code></p>	Yes
SESSION_SERVICE_URL	<p>Indicates the session service URL that gets activated after the <code>fcc-studio.sh</code> file runs.</p> <p>The format for the session service URL is as follows: <code>http://<HostName>:7047/session-service</code></p>	Yes

Table C-1 (Continued) config.sh Parameters for FCC Studio

InteractionVariable Name	Significance	Mandatory
FCC_STUDIO_URL	Indicates the FCC Studio URL. The format for the FCC Studio URL is as follows: <code>http://<HostName>:7008</code>	Yes
ORE Interpreter Settings		
RSERVE_USERNAME	Indicates the RServe username.	No
RSERVE_PASSWORD	Indicates the RServe password.	No
HTTP_PROXY	Indicates the HTTP proxy.	No
HTTPS_PROXY	Indicates the HTTPS proxy.	No
REPO_CRAN_URL	Indicates the URL from where the R packages are obtained. The format for the REPO_CRAN_URL is as follows: https://cran.r-project.org/	No
USERS_LIB_PATH	Indicates the path where the R packages are installed.	No
RSERVE_CONF_PATH	Indicates the path where the <code>Rserve.conf</code> file is present.	No
DB Details for Studio Schema		
STUDIO_DB_HOST-NAME	Indicates the hostname of the database where the Studio schema is created.	Yes
STUDIO_DB_PORT	Indicates the port number where the Studio schema is created.	Yes
STUDIO_DB_SERVICE_NAME	Indicates the service name of the database where the Studio schema is created.	Yes
STUDIO_DB_SID	Indicates the SID of the database where the Studio schema is created.	Yes
STUDIO_DB_USERNAME	Indicates the username of the Studio schema (newly created Oracle Schema).	Yes
STUDIO_DB_PASSWORD	Indicates the password for the Studio schema.	Yes
Studio DB Wallet Details		
STUDIO_ALIAS_NAME	Indicates the Studio alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
STUDIO_WALLET_LOCATION	Indicates the Studio wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes
STUDIO_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>STUDIO_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	Yes

Table C-1 (Continued) config.sh Parameters for FCC Studio

Interaction Variable Name	Significance	Mandatory
DB Details for BD Config Schema		
BD_CONFIG_HOST-NAME	Indicates the hostname of the database where BD config schema is installed.	No
BD_CONFIG_PORT	Indicates the port of the database where BD config schema is installed.	No
BD_CONFIG_SERVICE_NAME	Indicates the service name of the database where BD config schema is installed.	No
BD_CONFIG_SID	Indicates the SID of the database where BD config schema is installed.	No
BD_CONFIG_USERNAME	Indicates the username for the BD config schema. NOTE: Ensure to enter the value as NA.	Yes
BD_CONFIG_PASSWORD	Indicates the password for the BD config schema.	No
BD Config Wallet Details		
BD_CONFIG_ALIAS_NAME	Indicates the BD config alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
BD_CONFIG_WALLET_LOCATION	Indicates the BD config wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_CONFIG_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
DB Details for BD Atomic Schema		
BD_ATOMIC_HOST-NAME	Indicates the BD atomic schema hostname.	No
BD_ATOMIC_PORT	Indicates the BD atomic schema port number.	No
BD_ATOMIC_SERVICE_NAME	Indicates the BD atomic schema service name.	No
BD_ATOMIC_SID	Indicates the BD atomic schema SID.	No
BD_ATOMIC_USERNAME	Indicates the username of the BD atomic schema.	No
BD_ATOMIC_PASSWORD	Indicates the password of the BD atomic schema.	No
BD Atomic Wallet Details		

Table C-1 (Continued) config.sh Parameters for FCC Studio

Interaction Variable Name	Significance	Mandatory
BD_- ATOM- IC_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
BD_ATOMIC_WAL- LET_LOCATION	Indicates the BD atomic wallet location. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
BD_ATOMIC_TNS_AD- MIN_PATH	Indicates the path of the <code>tnsnames.ora</code> file where an entry for the <code>BD_ATOMIC_ALIAS_NAME</code> is present. For more information, see Chapter 10, "Appendix - Setting Up Password Stores with Oracle Wallet" .	No
SQL Scripts		
FSINFODOM	Indicates the name of the BD Infodomain.	No
FSSEGMENT	Indicates the name of the BD segment.	No
DATAMOVE- MENT_LINK_NAME	<ul style="list-style-type: none"> If the newly created schema is in a different database host, you must create a DB link and provide the same link in this parameter. If no DB link is present, provide the <code><SCHEMA_NAME></code> in this parameter. If the newly created schema is in the same database host, the value for this parameter is the user name of the BD atomic schema. <p>NOTE: Ensure to enter the value as NA</p>	Yes
DATAMOVE- MENT_LINK_TYPE	If the DB link is used, enter DBLINK in this field. If the DB link is not used, enter SCHEMA in this field. NOTE: Ensure to enter the value as NA	Yes
PGX Setup Details		
PGX_INSTALLATION_- PATH	Indicates the installation path of the PGX server. Example: <code><STUDIO_INSTALLATION_PATH>/pgx/ server/</code>	Yes
PGX_PGB_PATH	Indicates the path where you want to obtain the output graph PGB file. Example for Hive Installation: <code>hdfs:/user/ofsa</code> Example for Oracle DB Installation: <code>/scratch/ofsa</code>	Yes
Cloudera Setup Details		

Table C-1 (Continued) config.sh Parameters for FCC Studio

InteractionVariable Name	Significance	Mandatory
HADOOP_CREDENTIAL_PROVIDER_PATH	Indicates the path where Hadoop credential is stored. NOTE: Indicates the Hadoop alias given when creating the Hadoop credentials.	No
HADOOP_PASSWORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials.	No
Hive_Host_Name	Indicates the Hive hostname.	No
Hive_Port_number	Indicates the Hive port number. Contact System Administrator to obtain the port number.	No
HIVE_PRINCIPAL	Indicates the Hive Principal. Contact System Administrator to obtain HIVE_PRINCIPAL.	No
HIVE_SCHEMA	Indicates to create a schema in HIVE.	No
JAAS_CONF_FILE_PATH	Created for future use.	No
Krb_Host_FQDN_Name	Indicates the Krb host FQDN name.	No
Krb_Realm_Name	Indicates the Krb realm name.	No
Krb_Service_Name	Indicates the Krb service name. Example: Hive	No
KRB5_CONF_FILE_PATH	Created for future use.	No
security_krb5_kdc_server	Created for future use.	No
security_krb5_realm	Created for future use.	No
server_kerberos_keytab_file	Created for future use.	No
server_kerberos_principal	Created for future use.	No
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Host machine where sqoop will run.	No

Table C-1 (Continued) config.sh Parameters for FCC Studio

InteractionVariable Name	Significance	Mandatory
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> 1. Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry: <code>oracle.jdbc.mapDateToTimestamp=false</code> 2. Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsaas/</code> <p>NOTE: Ensure that the location name ends with a <code>'/'</code>.</p>	No
SQOOP_PARTITION_COL	Indicates the column in which the HIVE table is partitioned. The value must be <code>SNAPSHOT_DT</code> .	No
SQOOP_TRG_HOST_NAME	Indicates the hostname of the SQOOP server. Example: <code><HostName></code>	No
SQOOP_WORK_DIR_HDFS	Indicates the SQOOP working directory in HDFS. Example: <code>/user/ofsaas</code>	No
Elastic Search Cluster details		
ELASTIC_SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
ELASTIC_SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
Matching Service		
MATCHING_SERVICE_HOSTNAME	Indicates the hostname where the matching service is installed.	Yes
MATCHING_SERVICE_PORT	Indicates the port number where the matching service is installed. NOTE: This parameter is same as the <code>SERVER_START_PORT</code> parameter given below.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	Yes
ELASTICRESPONSE_BUFFERLIMIT_BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
SCROLL_TIME	Indicates the duration for which the <code>scroll_size</code> output is active. For example: 5	Yes

Table C-1 (Continued) config.sh Parameters for FCC Studio

InteractionVariable Name	Significance	Mandatory
SCROLL_SIZE	Indicates the amount of data that must be obtained in one attempt when a query is fired on an index in the elastic search service. For example: 1000	Yes
SERVER_START_PORT	Indicates the port number where the matching service is installed Default Value: 7049 NOTE: This value is already set. Do not change the value.	Yes
Entity Resolution		
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default Value: 7051 NOTE: This value is already set. Do not change the value.	Yes
ER_SERVICE_URL	Indicates the URL of the entity resolution service. For example: <code>http://<HostName>:7051</code>	Yes
Graphs		
HDFS_GRAPH_FILES_PATH	Indicates the filepath in HDFS where the <code>graph.json</code> is formed.	Yes
GRAPH_FILES_PATH	Indicates the directory in the Big Data server for graph files.	Yes
GRAPH_NAME	Indicates the name you want to assign to the global graph at the end of ETL.	Yes

7.4 Running the FCC Studio Installer

To run the installer, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/bin/` directory.
2. Execute the following command in the console:

```
./install.sh
```

NOTE Execution of `install.sh` command does not generate any log file.

3. Execute the following command in the console:

```
./fcc_studio.sh
```

The OFS Crime and Compliance Studio application is installed with non-OFSAA. The FCC Studio and all the interpreters are started.

After the successful completion of the FCC Studio installation, the script displays a URL that can be used to access the FCC Studio Application. For more information, see [Accessing the FCC Studio Application](#).

4. Verify the Installation. See [Verifying the FCC Studio Installation](#)
5. Verify the Entity Resolution. See [Verifying the Entity Resolution](#).

7.5 Verifying the FCC Studio Installation

To verify the installation, follow these steps:

- Check the log files located in the `<STUDIO_INSTALLATION_PATH>/logs` directory.
 - Ensure that the metaservice is up and running.
 - The authservice and batchservice will not be up.
- In the FCC Studio application UI, navigate to the **Interpreters** page to ensure all the interpreters are displayed and the JDBC interpreter is working.

NOTE Any errors encountered in the process is displayed with an appropriate error code in the log file. Do not proceed with the further installation and contact Oracle Support with relevant log files.

If the installation of the FCC Studio application is unsuccessful, you must reinstall the application after performing the cleanup tasks. For more information, see [Chapter 9, "Reinstalling FCC Studio Application"](#).

7.6 Verifying the Entity Resolution

To verify the installation of the entity resolution service, follow these steps:

1. Enter the URL in the following format into the browser:
`http://<Entity Resolution Hostname>:<Entity Resolution Port>/er/Rule.jsp`
The Entity Resolution page is displayed.
2. Verify if all the out of the box rulesets are displayed in the Entity Resolution page. For more information, see [Financial Crime Graph Model Matching Guide](#).

7.7 Installing PGX Service

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

To install PGX service, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/pgx/server/` directory.
2. Perform the following:
 - If PGX service is to be installed on the same server where FCC Studio is installed, unzip the `pgx-distribution-20.0.0-RC1-server.zip` file.
 - If PGX service is to be installed on a different server, follow these steps:

- i. Copy the `pgx-distribution-20.0.0-RC1-server.zip` file to the PGX server.
- ii. Unzip the `pgx-distribution-20.0.0-RC1-server.zip` file.

NOTE The path where the `pgx-distribution-20.0.0-RC1-server.zip` file is unzipped is referred to as `<PGX_Installation_Path>`.

- 3. Navigate to the `<PGX_Installation_Path>/pgx/server/conf` directory.
- 4. Configure the following properties as per the requirement:
 - Set the following in the `server.conf` file:


```
enable_tls: false,
enable_client_authentication: false
```
 - Set the following in the `pgx.conf` file:


```
allow_local_filesystem: true
```

Here, true indicates to enable, and false indicates to disable.
- 5. Copy and paste the following Kerberos Files into the `<PGX_Installation_Path>/pgx/server/conf/kerberos` directory:
 - `krb5.conf`
 - `ofsa.keytab`
- 6. Copy and paste the following Hadoop configuration files into 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`
- 7. Navigate to the `<PGX_Installation_Path>/pgx/server/bin` and configure the `config.sh` file as described in the following table:

Table G-1 config.sh Parameters

InteractionVariable Name	Significance
KERBEROS_TICK-ET_RENEWAL_PERIOD	For example: 7200 would mean every 2 hours
KERBEROS_PRINCIPAL	For example: ofsaa@WHFAYK.ORACLE.COM

Table G-1 (Continued) config.sh Parameters

InteractionVariable Name	Significance
KERBEROS_KEYTAB_FILENAME	For example: ofsaa.keytab
KRB5_CONFIG_FILE-NAME	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 container's memory limit size above. For example: 10240
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 container's memory limit size above. For example: 10240
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/ofsaa/hive_etl_out_80702_pgx/graph.json
PGX_GLOBAL_GRAPH_NAME	Indicates the name that the pre-loaded global graph is published with and the FCC Studio users can use to reference the global graph. For example: GlobalGraphIH

8. Navigate to the <PGX_Installation_Path>/pgx/server/bin directory and execute the following commands:

```
./install.sh
```

9. Start the PGX service. For more information, see [Starting/Stopping PGX Service](#).

NOTE

- You must run at least one successful ETL batch to start the PGX service with the PGX graph configuration.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 Crime and Compliance Studio Administration Guide](#).
- Ensure to execute the Graph_Alive notebook after each time you start/restart the PGX service. For more information, see [Chapter 19, "Appendix - Executing Graph_Alive Notebook"](#).

7.8 Accessing the FCC Studio Application

Access the FCC Studio application. For more information, see [Chapter 15, "Appendix - Accessing the FCC Studio Application"](#).

8 Post-installation Configuration for Studio Installed With Non-OFSAA

On successful installation of FCC Studio with non-OFSAA, perform the following post-installation configuration:

- [Configuring SSH Connection](#)
- [Configuring Interpreters](#)
- [Configuring ICIJ](#)

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

8.1 Configuring SSH Connection

To configure SSH connection, follow these steps:

1. Run the following command in the console:
 - a. Run `ssh-keygen`
Generating public/private rsa key pair
 - b. Enter file in which to save the key (<Linux_Home>/`.ssh/id_rsa`):[Press Enter]
 - c. Enter passphrase (empty for no passphrase): [Press Enter]
 - d. Enter same passphrase again:[Press Enter]
 - e. `ssh-copy-id -i ~/.ssh/id_rsa.pub <BigData_Server>`
 - f. `ssh <BigData_Server>`

8.2 Configuring Interpreters

After starting the FCC Studio application, the configuration for the interpreters can be performed from the user interface (UI). For information on configuring interpreters, see the *Configuring Interpreters* chapter in the [Oracle Financial Services Crime and Compliance Studio Administration Guide](#).

8.3 Configuring ICIJ

To configure ICIJ (International Consortium of Investigative Journalists), perform the following:

1. [Cleaning the ICIJ Data](#)
2. [Configuring FILEPATH for ICIJ](#)

8.3.1 Cleaning the ICIJ Data

To clean the ICIJ data, follow these steps:

1. Download the four dataset directories from <https://offshoreleaks.icij.org/pages/database>.
2. Unzip the four dataset directories and place the unzipped directories in the <Studio_Installation_Path>/`icij_data_cleaning` directory.

3. Navigate to the `<Studio_Installation_Path>/icij_data_cleaning/bin` directory and execute the following command:

```
./clean.sh
```

NOTE Ensure that Python 3 is available in the machine before executing this command.

After successful execution of the command:

- The cleaned data is available for the sqoop job to load it in Hive and HDFS.
- A directory named "clean" is created inside each of the dataset where a clean version of each CSV file is created.

8.3.2 Configuring FILEPATH for ICIJ

NOTE The FCC Studio graph model is configured to include ICIJ watch list files.

To configure FILEPATH for ICIJ, follow these steps:

1. Place the watch list files in HDFS, that is accessible by the user.
2. Update the FILEPATH of the watch list files in the `fcc_studio_etl_files` table.

Figure 1: fcc_studio_etl_files Table

	DF_NAME	FILEPATH	DF_SEQ_NO	FILE_ORDER
1	Offshore_edges_is_related_to	12	1
2	Bahama_External_Address	13	1

9 Reinstalling FCC Studio Application

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

To reinstall the Studio application, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/bin` directory.
2. Create a backup for the existing `config.sh` file.
3. Delete the `<Studio_Installation_Path>` folder.
4. Unzip the installer archive file, `OFS_FCCM_STUDIO_8.0.7.3.0_Linux.zip` and extract the software. For more information, see [Extracting the Software](#).
5. Perform database cleanup by performing the following:

Table 1: Database Cleanup

Schema	Applicable for FCC Studio with OFSAA	Applicable for FCC Studio with Non-OFSAA
Clean up for Studio Schema	Yes	Yes
Cleanup for BD Atomic Schema	Yes	No
Cleanup for BD Config Schema	Yes	No

6. Reinstall the FCC Studio application. For more information, see [Chapter 4, "Installing FCC Studio With OFSAA"](#) and [Chapter 7, "Installing FCC Studio With Non-OFSAA"](#).

9.0.1 Clean up for Studio Schema

To clean up the Studio schema, follow these steps:

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

NOTE The username and password credentials of the 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_qls TO <Schema User>;`
- `GRANT execute on sys.dbms_session TO <Schema User>;`
- `ALTER USER <Schema User> QUOTA 100M ON users;`
- `GRANT create sequence TO <Schema User>;`

- GRANT create SYNONYM TO <Schema User>;
- GRANT ALL privileges TO <Studio Schema User>;

NOTE

If dropping the schema is not an option, drop the tables and sequences as mentioned in the [Studio Schema Tables](#) section.

9.0.2 Cleanup for BD Atomic Schema

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

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

```
TRUNCATE TABLE DATABASECHANGELOGLOCK;  
TRUNCATE TABLE DATABASECHANGELOG;
```

9.0.3 Cleanup for BD Config Schema

To clean up the BD Config schema, follow these steps:

1. Login to the BD Config Schema.
2. Truncate the DATABASECHANGELOG and DATABASECHANGELOGLOCK tables using the following command:

```
TRUNCATE TABLE DATABASECHANGELOGLOCK;  
TRUNCATE TABLE DATABASECHANGELOG;
```

10 Appendix - Setting Up Password Stores with Oracle Wallet

This section includes the following topics:

- [Overview](#)
- [Setting Up Password Stores for Database User Accounts](#)
- [Verifying the Connectivity of the Wallet](#)

10.1 Overview

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.

10.2 Setting Up 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 set up a wallet and the aliases for the database user accounts. For more information on configuring authentication and password stores, refer to the Oracle Database Security Guide.

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

To set up a password store for the database user accounts, follow these steps:

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

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

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

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

The wallet is created with the auto-login feature enabled. This feature enables the database client to access the wallet contents without using the password. For more information, refer to the Oracle Database Advanced Security Administrator's Guide.

3. Create the database connection credentials in the wallet using the following command:

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

Run the above command for the following <alias-name>:

Table C-1

Schema	Applicable for FCC Studio with OFSAA	Applicable for FCC Studio with Non-OFSAA
BD_Config_Schema	Yes	No
BD_Atomic_Schema	Yes	No
Studio_Schema	Yes	Yes

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.

- Repeat step 2 for all the database user accounts.
- Update the `tnsnames.ora` file to include the following entry for each alias name to be set up.

```
<alias-name> =
(DESCRIPTION =
(AADDRESS_LIST =
(AADDRESS = (PROTOCOL = TCP) (HOST = <host>) (PORT = <port>))
)
(CONNECT_DATA =
(SERVICE_NAME = <service>)
)
)
```

NOTE

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

10.3 Verifying the Connectivity of the Wallet

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

- Create a `sqlnet.ora` 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 what the wallet was created with.

```
$ export WALLET_LOCATION=<wallet_location>
```

```
$ export TNS_ADMIN=<tnsnames.ora_location>, Here ensure to use the wallet to point to the alternate tnsnames.ora as created above.
```

```
$ 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>"
```

11 Appendix - Installing RServe Manually

11.1 Overview

You must install Rserve manually on a host to expose the local R installation on that host to the network, so that remote Rserve clients such as the R interpreter can use the local R installation.

The R interpreter always connects to an Rserve instance and runs the R code remotely. The interpreter needs to be configured with the hostname or IP and the port of the remote instance (where Rserve is running). When the interpreter is initialized, it connects to the remote instance.

This section includes the following topics:

- [Prerequisites](#)
- [Installing Rserve](#)
- [Configuring Rserve](#)
- [Starting Rserve](#)
- [Adding the Certificate to the Keystore](#)
- [Installing Additional Libraries](#)

11.2 Prerequisites

The following is a list of prerequisites required before beginning the installation of RServe:

- Ensure that Oracle Linux 7.x and Oracle JDK 8 are validated against Oracle Linux 7.4 and Oracle JDK 8u161.
- The user must be a root user.
- 800 MB disk space is required for package installation.

The following subsections provide more details for prerequisites:

- [Installing Oracle R Distribution](#)
- [Installing Dependencies](#)
- [Installing ORE Client](#)

11.2.1 Installing Oracle R Distribution

To install the Oracle R Distribution (ORD), enable the **addons** and **optional_latest** channels in yum as shown in the following:

```
```bash
(root)# yum-config-manager --enable ol7_addons
(root)# yum-config-manager --enable ol7_optional_latest
```
```

After completing the previous step, pull ORD from the yum repository using the following command:

```
```bash
(root)# yum install R.x86_64 R-core-extra
```
```

To install ORD, see https://docs.oracle.com/cd/E83411_01/OREAD/installing-R-for-ORE.htm#ORE-AD129.

11.2.2 Installing Dependencies

Rserve has certain dependencies to run correctly. The **openssl-devel** is required for SSL support. The dependencies change based on the libraries you have installed. For example, to let **knitr** send plots as **base64 encoded** strings, you require **pango-devel**.

The following dependent packages must be installed for Rserve to support SSL:

```
```bash
(root)# yum install openssl openssl-devel pango-devel
```
```

11.2.3 Installing ORE Client

To connect to ORE through Rserve, follow these steps:

- Install the corresponding client libraries. For more information, see https://docs.oracle.com/cd/E83411_01/OREAD/installing-ORE-client.htm#OREAD167 from the ORE project.
- Install the knitr and PrintR packages.

11.3 Installing Rserve

To install Rserve, call the following code in your R shell:

```
```R
> install.packages('Rserve', repos='https://www.rforge.net/')
```
```

If you are behind a proxy, ensure that R is communicated about it when you start the R shell.

For example, you can start R shell as shown in the following before installing any package:

```
```bash
$ http_proxy=http://your-proxy:80 R
```
```

11.4 Configuring Rserve

You can base your config on the following example configuration, that you should store in an `Rserve.conf` file. You will require `Rserve.conf` as a reference when you start Rserve.

```
```
auth required
plaintext disabled
pwdfile /path/to/Rserve.pwd
remote enable
switch.gap.tls enable
```
```



```
tls.port 6311
qap disable
interactive no
rsa.key /path/to/server.key
tls.key /path/to/server.key
tls.cert /path/to/server.crt
```

This configuration tells Rserve to encrypt the communication with TLS and listen for incoming connections on port **6311**. The `Rserve.pwd` file appears as shown in the following example:

```
...
user $5baa61e4c9b93f3f0682250b6cf8331b7ee68fd8
...
```

The file contains one line per user, where the first part is the username and the second part is the password. The password can either be plain text or a **MD5/SHA1** hash. In this example, the password `password` is hashed with SHA1. If you use hashed passwords, the password must start with a ``$`` sign.

The `rsa.key`, `tls.key`, and `tls.cert` settings point to the private key files you require for TLS. These keys can be generated using the **openssl** command-line tool as shown in the following example:

```
```bash
$ openssl genrsa -out server.key 2048
$ openssl req -new -key server.key -out server.csr
$ openssl x509 -req -days 265 -in server.csr -signkey server.key -out
server.crt
...`
```

The preceding sample is an example and for a production deployment, you should use relevant certificates. You can find more information about configuration options on the Rserve homepage - <https://www.rforge.net/Rserve/doc.html>.

## 11.5 Starting Rserve

After installing Rserve and placing configuration files in the correct location, start the Rserve as given here:

```
```bash
$ R CMD Rserve --no-save --RS-conf /path/to/Rserve.conf
```

Rserve starts in the background. After it starts, the R Interpreter can connect to it. The Rserve process is running before you start the R Interpreter.

11.6 Adding the Certificate to the Keystore

The certificates that were generated in the previous step to configure Rserve to encrypt the communication must be added to the Java keystore to be used by the R interpreter. The add procedure depends on your setup.

Add the certificate to a keystore as given here:

```
```bash
```

```
$ $JAVA_HOME/bin/keytool -import -alias rserve -file /path/to/server.crt -keystore /path/to/key-
store-storepass storepassword -noprompt
```
```

The certificate must be imported correctly and the correct keystore is used by the Java process you use to start the R interpreter. Else you will get SSL related exceptions when the interpreter attempts to connect to Rserve.

You can specify the keystore when starting the R interpreter as shown in the following example:

```
```bash  
$ $JAVA_HOME/bin/java -Djavax.net.ssl.trustStore=/path/to/keystore -
Djavax.net.ssl.trustStorePassword=storepassword <additional paramters>
```
```

11.7 Installing Additional Libraries

Depending on your use case, you must install other R libraries. For example, you can install **knitr** or **ggplot2**, in the same manner, that you installed Rserve previously. You can use the *package.install* within your R shell to perform the installation.

For example:

```
```R  
> install.packages('knitr')
```
```

12 Appendix - Configuring the Elastic Search Component

To configure the Elastic Search component, follow these steps:

1. Navigate to the `elasticsearch-7.3.2/config` directory.
2. Configure the `elasticsearch.yml` with the following variables:

Table 1: elasticsearch.yml File

| InteractionVariable Name | Significance |
|---|---|
| <code>cluster.name</code> | Indicates the name of the cluster. |
| <code>node.name</code> | Indicates the name given for the node. |
| <code>node.master</code> | Indicates whether the node is a master. |
| <code>node.data</code> | Indicates the node data. |
| <code>path.data</code> | Indicates the directory where you want to store the data. |
| <code>path.logs</code> | Indicates the directory where you want to store the logs. |
| <code>network.host</code> | Indicates the hostname of the machine where you want to install the elastic search service. |
| <code>http.port</code> | Indicates the port number where the elastic search service is installed. |
| <code>discovery.seed_hosts</code> | (Optional) Indicates the hostnames of the nodes of the cluster. |
| <code>cluster.initial_master_nodes</code> | (Optional) Indicates the number given to the nodes of the cluster. |

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

Table 2: elasticsearch.yml File

| InteractionVariable Name | Significance |
|--------------------------|---|
| <code>-Xms1g</code> | <ul style="list-style-type: none"> • 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. |
| <code>-Xmx1g</code> | |

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

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

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

```
{
  "name" : "node-1",
  "cluster_name" : "my-application",
  "cluster_uuid" : "_D-2qEGGSgKQPd3i-UtgWQ",
```

```
"version" : {
  "number" : "7.3.2",
  "build_flavor" : "oss",
  "build_type" : "tar",
  "build_hash" : "1c1faf1",
  "build_date" : "2019-09-06T14:40:30.409026Z",
  "build_snapshot" : false,
  "lucene_version" : "8.1.0",
  "minimum_wire_compatibility_version" : "6.8.0",
  "minimum_index_compatibility_version" : "6.0.0-beta1"
},
"tagline" : "You Know, for Search"
}
```

13 Appendix - Starting/Stopping FCC Studio Services

This section describes how to start and stop the services needed for the FCC Studio application. It includes the following sections:

Table 0-1

| Start/Stop FCC Studio Services | Applicable for FCC Studio with OFSAA | Applicable for FCC Studio with Non-OFSAA |
|--------------------------------|--------------------------------------|--|
| Starting/Stopping PGX Service | Yes | No |
| Starting/Stopping FCC Studio | Yes | Yes |

13.1 Starting/Stopping PGX Service

- To start the PGX service:
 - a. Navigate to the path where 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:
`./start-pgx.sh`

NOTE Ensure to execute the Graph_Alive notebook after each time you start/restart the PGX service. For more information, [Chapter 19, "Appendix - Executing Graph_Alive Notebook"](#).

- To stop the PGX service, kill the process.

13.2 Starting/Stopping FCC Studio

- To start FCC Studio:
 - a. Navigate to the `<Studio_Installation_Path>/bin/` directory.
 - b. Run the following command:
`./fcc-studio.sh`
 - To stop FCC Studio, kill all the FCC Studio processes.
1. Once all the services are up and running, you can access the FCC Studio application. For more information, see [Chapter 15, "Appendix - Accessing the FCC Studio Application"](#).

14 Appendix - Tables and Sequences

The list of tables and sequences that are to be dropped during reinstallation of the FCC Studio application are as follows:

- Studio Schema Tables
- Studio Schema Sequences

14.1 Studio Schema Tables

The following table includes the details of the Studio Schema tables that must be dropped during the reinstallation of the FCC Studio application.

Table 1: Studio Schema Tables

| | | |
|----------------------------|---------------------------|-----------------------|
| DS_PARAGRAPH | DS_NOTEBOOK_TAGS | DS_TASK_RESULTS |
| DS_ENTITY_PERMISSIONS | DS_ROLE | DS_PERMISSION_ACTIONS |
| DS_GROUP | DS_IS_PERMITTED | DS_PERMISSION_MAPPING |
| DS_USER_PERMS_MAP | DS_USER_ROLES | DS_NOTEBOOK |
| DS_INTERPRETER_RESULT_MSGS | DS_USER | DS_PERMS_MAP_ACTIONS |
| DS_ENTITY_PERMS_MAP | DS_TASK | DS_GRAPH |
| DS_INTERPRETER_RESULT | DS_GROUP_PERMS_MAP | DS_NOTEBOOK_RELATIONS |
| DS_INTERPRETER_PROPS | DS_JOB | DS_PERMISSION |
| DS_ROLE_PERMS_MAP | DS_VISUALIZATION_TEMPLATE | DS_RESULT_MESSAGE |
| DS_INTERPRETER_ABILITIES | DATABASECHANGELOG | DATABASECHANGELOGLOCK |
| DS_USER_GROUPS | DS_INTERPRETER_VARIANT | DS_COMMENT |
| DS_PARAGRAPH_RELATIONS | | |

14.2 Studio Schema Sequences

The following table includes the details of the Studio Schema sequences that must be dropped during the reinstallation of the FCC Studio application:

Table 2: Studio Schema Sequences

| | | |
|----------------------------|------------------------|-------------------------|
| SEQ_COMMENT | SEQ_ENTITY_PERMISSIONS | SEQ_GRAPH |
| SEQ_GROUP | SEQ_INTERPRETER_RESULT | SEQ_INTERPRETER_VARIANT |
| SEQ_JOB | SEQ_NOTEBOOK | SEQ_PARAGRAPH |
| SEQ_PERMISSION | SEQ_PERMISSION_MAPPING | SEQ_RESULT_MESSAGE |
| SEQ_ROLE | SEQ_TASK | SEQ_USER |
| SEQ_VISUALIZATION_TEMPLATE | | |

15 Appendix - Accessing the FCC Studio Application

To access the FCC Studio application, follow these steps:

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

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

Here <Port_Number> is 7008 for the FCC Studio application installed on-premise.

The FCC Studio application login page is displayed.

Figure 1: Crime and Compliance Studio Login Page



2. Log in with the **Username** and **Password**.
3. Click **Login**.

The Crime and Compliance Studio application's landing page is displayed with the list of all the out-of-the-box notebooks packaged with FCC Studio.

16 Appendix - Using Python Virtual Environments with PySpark

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

1. [Creating a Virtual Environment with Conda](#)
2. [Updating Interpreter Properties](#)

16.1 Creating a Virtual Environment with Conda

NOTE You can also use `virtualenv` to create your virtual environment instead of Conda.

1. Ensure that you have Conda and Conda-Pack installed.
2. Create your virtual environment by executing 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 by executing 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 by executing the following command:

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

16.2 Updating Interpreter Properties

All these properties can either be configured in the interpreter JSON files or from the *Interpreters* page of the FCC Studio application UI after starting the FCC Studio application.

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

17 Appendix - Enabling a Second Spark/PySpark Interpreter

To set up a second Spark/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 The PySpark interpreter does not have or require its own start-script.

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

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

- b. Edit the `start-spark2-interpreter.sh` file in the `<Studio_Installation_Path>/interpreters/bin/` directory to update the port number to a new port number that is not in use (for example, 7030) and rename the log file (for example, `spark2.sh`)

In the `start-spark2-interpreter.sh` file,

— Line 13 becomes:

```
{SPARK_SUBMIT} --class oracle.datastudio.interpreterserver.Zeppe-
linRemoteInterpreterServer --driver-class-path ${CLASSPATH} --
driver-java-options "${JAVA_OPTS} ${SPARK_INTERPRETER_OPTS} " --
files ${py4j[0]},${SPARK_HOME}/python/lib/pyspark.zip ${SPARK_
CONF} ${SPARK_APP_JAR} ${1:-7030} > $DIR/../../logs/spark2.log
```

— Line 24 becomes:

```
java -DlogFileName=spark -Dfile.encoding=UTF-8 ${JAVA_OPTS}
${SPARK_INTERPRETER_OPTS} oracle.datastudio.interpreterserver.Zep-
pelinRemoteInterpreterServer ${1:-7030} > $DIR/../../logs/
spark2.log
```

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

Insert the following code:

```
sh "$DIR"/start-spark2-interpreter.sh &
```

Below the following code:

```
sh "$DIR"/start-spark-interpreter.sh &
```

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

- a. Navigate to the `<Studio_Installation_Path>/interpreters/conf` directory and execute the following command to create the new interpreter JSON called `spark2.json`:

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

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

Update the following parameter values:

```
group, name, groupSettings.initialCodeCapability, port, capabilities.name, capabilities.button.label
```

to the following:

```
<new-spark-interpreter-name> ,<new-spark-interpreter-name>, <new-spark-interpreter-name>, 7030 (the port chosen in the first step), <new-spark-interpreter-name>, <new-spark-interpreter-name>
```

The first 23 lines of the new JSON file must be similar to the following code, where Spark2 is used for <new-spark-interpreter-name>:

```
[
  {
    "group": "spark2",
    "name": "spark2",
    "className": "org.apache.zepplin.spark.SparkInterpreter",
    "groupSettings": {
      "initialCode": "True\n",
      "initialCodeCapability": "spark2"
    },
    "host": "localhost",
    "port": 7030,
    "capabilities": [
      {
        "name": "spark2",
        "highlightLanguage": "scala",
        "formEscapeCharacter": "@",
        "button": {
          "defaultCode": "println(\"Hello, world\")",
          "icon": "fa fa-fw fa-building-o",
          "label": "Spark2"
        }
      }
    ]
  },
]
```

3. Create the interpreter JSON for the second PySpark interpreter.
 - a. Navigate to the <Studio_Installation_Path>/interpreters/conf directory and execute the following command to create the new interpreter JSON called pyspark2.json:

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

- b. Edit the pyspark2.json file in the <Studio_Installation_Path>/interpreters/conf/ directory as follows:

Update the following parameter values:

group, name, port, capabilities.name, capabilities.button.label

to the following:

```
<new-spark-interpreter-name> ,<new-pyspark-interpreter-name>, 7030
(the port chosen in the first step), <new-pyspark-interpreter-name>,
<new-pyspark-interpreter-name>
```

The first 19 lines of the new JSON file must be similar to the following code, where spark2 is used for <new-spark-interpreter-name> and pyspark2 is used for <new-pyspark-interpreter-name>:

```
[
{
"group": "spark2",
"name": "pyspark2",
"className": "org.apache.zepppelin.spark.PySparkInterpreter",
"host": "localhost",
"port": 7030,
"capabilities": [
{
"name": "pyspark2",
"highlightLanguage": "python",
"button": {
"defaultCode": "print('Hello World')",
"icon": "icon-python",
"label": "pyspark2"
},
"formEscapeCharacter": "$"
}
],
```

NOTE

If you try to connect the two interpreters to different external clusters when setting the environment variables, SPARK_HOME and HADOOP_CONF_DIR as part of [Chapter 18, "Appendix - Providing Spark Libraries"](#), To prepend the declaration of those environment variables to the respective Spark interpreter start-scripts, so that they don't interfere, rather than setting them system wide.

- Restart the FCC Studio application. For more information, see [Starting/Stopping FCC Studio](#).

18 Appendix - Providing Spark Libraries

To provide your own Spark libraries and/or Hadoop client-configuration files for the Spark interpreter to connect to a Spark Standalone or Yarn cluster, follow these steps:

1. Place your unarchived Spark libraries in a directory accessible by FCC Studio.

You can download the Spark libraries from the [Spark's Official Release Archive](#).

2. Set the environment variable `SPARK_HOME` to the absolute path of your Spark libraries.
3. Place your Hadoop client-configuration files to a directory accessible by FCC Studio.

You can obtain the Hadoop client-configuration files from the Yarn cluster by performing one of the following:

- Copy the Hadoop configuration directory, `$HADOOP_HOME/etc/hadoop` or `$HADOOP_HOME/conf` that contains the `core-site.xml` and `hdfs-site.xml` files.

Where,


`HADOOP_HOME` is either `/etc/hadoop` or `/usr/lib/hadoop`.

- Download the configuration files directly from the cluster manager's UI if you are using a Cloudera cluster.
4. Set the system environment variable `HADOOP_CONF_DIR` to the absolute path of the directory that contains the configuration files.

19 Appendix - Executing Graph_Alive Notebook

In an Investigation Hub notebook, the graph is lost whenever a session is reset and this occurs as part of the session clean-up. You must execute the Graph_Alive notebook to retain the link to the graph even when a session is reset.

To execute the Graph_Alive notebook, follow these steps:

1. Access the FCC Studio application. For more information, see [Chapter 15, "Appendix - Accessing the FCC Studio Application"](#).
2. Click the Graph_Alive notebook.
The Graph_Alive notebook is displayed.
3. Click the **Run Paragraphs** icon  to execute all the paragraphs in the notebook in a sequential order.

