

# **Oracle Financial Services Crime and Compliance Studio**

**Deployment and Configuration Guide**

**Release 8.0.7.3.0**

**April 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"><li>• 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 <a href="#">“Configuring the Elastic Search Component” on page 17</a> and <a href="#">“Appendix - Configuring the Elastic Search component” on page 68</a>.</li><li>• A new script is introduced for the preparation of ICIJ to resolve data quality issues before graph loading. For more information, see <a href="#">“Cleaning the ICIJ Data” on page 31</a>.</li></ul>
8.0.7.2.0	Updated: February 2020	<ul style="list-style-type: none"><li>• Deploy Studio application with or without OFSAA on the Kubernetes cluster. For more information, see <a href="#">“Deploying FCC Studio with OFSAA” on page 17</a> and <a href="#">“Deploying FCC Studio with Non-OFSAA on the Kubernetes Cluster” on page 44</a>.</li><li>• The support for the Data Forward service has been deprecated.</li><li>• Configuration for the newly introduced interpreters such as Spark and PySpark interpreters. For more information, see <a href="#">“Interpreter Settings” on page 14</a>.</li></ul>
8.0.7.1.0	Updated: October 2019	Created the first version of Oracle Financial Services Crime and Compliance Studio Deployment Guide for v8.0.7.1.0 Release.

---

## Table of Contents

<b>1</b>	<b>Preface .....</b>	<b>4</b>
1.1	Summary.....	4
1.2	Audience .....	4
1.3	Related Documents.....	4
1.4	Abbreviations .....	4
<b>2</b>	<b>Deployment Overview .....</b>	<b>6</b>
2.1	Introduction.....	6
2.2	Quick Start Steps to Deploy FCC Studio with OFSAA on the Kubernetes Cluster.....	7
2.3	Quick Start Steps to Deploy FCC Studio with Non-OFSAA on the Kubernetes Cluster .....	7
<b>3</b>	<b>Preparing for Deployment .....</b>	<b>9</b>
3.1	Prerequisites.....	9
3.2	Hardware and Software Requirements .....	9
3.3	Prerequisite Environmental Settings .....	10
3.4	Performing Common Pre-installation Tasks.....	12
3.4.1	<i>Obtaining the Software .....</i>	<i>12</i>
3.4.2	<i>Extracting the Software .....</i>	<i>12</i>
3.5	Required File Structure .....	13
3.6	Interpreter Settings.....	14
3.6.1	<i>Configuring Spark Interpreter .....</i>	<i>14</i>
3.6.2	<i>Configuring PySpark Interpreter .....</i>	<i>15</i>
<b>4</b>	<b>Deploying FCC Studio with OFSAA.....</b>	<b>17</b>
4.1	Configuring the Elastic Search Component .....	17
4.2	Configuring Wallet .....	17
4.3	Configuring the studio-env.yml File .....	17
4.4	Configuring ETL Services .....	25
4.5	Verifying Resource Allocation for FCC Studio Services .....	27
4.6	Deploying FCC Studio on the Kubernetes Cluster .....	27
4.7	Verifying the FCC Studio Deployment.....	27
4.8	Verifying the Entity Resolution .....	27
4.9	Accessing the FCC Studio Application.....	28

---

<b>5</b>	<b>Post-deployment Configuration for FCC Studio Installed with OFSAA .....</b>	<b>29</b>
5.1	Configuring Interpreters.....	29
5.2	Performing OFSAA Configuration for Batch Execution.....	29
5.3	Performing Hive Data Movement.....	29
5.3.1	Configuring Schema Creation .....	29
5.3.2	Creating Credential Keystore .....	30
5.3.3	Configuring ETL .....	31
5.4	Performing Configuration to Run Published Notebooks.....	33
<b>6</b>	<b>Upgrading FCC Studio with OFSAA on the Kubernetes Cluster .....</b>	<b>34</b>
6.1	Upgrade Overview.....	34
6.2	Additional Cleanup for Upgrade .....	35
6.3	Configuring the studio-env.yml File for Upgrade .....	35
<b>7</b>	<b>Deploying FCC Studio with Non-OFSAA on the Kubernetes Cluster .....</b>	<b>44</b>
7.1	Configuring the Elastic Search Component .....	44
7.2	Configuring Wallet .....	44
7.3	Configuring the studio-env.yml File .....	44
7.4	Configuring ETL Services .....	54
7.5	Verifying Resource Allocation for FCC Studio Services .....	55
7.6	Deploying FCC Studio on Kubernetes Cluster .....	55
7.7	Verifying the FCC Studio Deployment.....	56
7.8	Verifying the Entity Resolution .....	56
7.9	Accessing the FCC Studio Application.....	56
<b>8</b>	<b>Post-deployment Configuration for FCC Studio with Non-OFSAA .....</b>	<b>57</b>
8.1	Configuring Interpreters.....	57
8.2	Configuring ICIJ.....	57
8.2.1	Cleaning the ICIJ Data .....	57
8.2.2	Configuring the FILEPATH for ICIJ .....	57
<b>9</b>	<b>Redeploying FCC Studio .....</b>	<b>59</b>
9.0.1	Cleanup for Studio Schema .....	59
9.0.2	Cleanup for BD Atomic Schema .....	60
9.0.3	Cleanup for BD Config Schema .....	60

---

<b>10 Appendix - Setting Up Password Stores with Oracle Wallet .....</b>	<b>61</b>
10.1 Overview.....	61
10.2 Setting Up Password Stores for Database User Accounts .....	61
10.3 Verifying the Connectivity of the Wallet.....	62
<b>11 Appendix - Installing Oracle R Distribution .....</b>	<b>64</b>
<b>12 Appendix - Starting/Stopping FCC Studio Services.....</b>	<b>65</b>
12.1 Starting a Service.....	65
12.2 Stopping a Service .....	65
12.3 Starting the FCC Studio Application.....	65
<b>13 Appendix - Checking Logs of FCC Studio Services .....</b>	<b>66</b>
<b>14 Appendix - Configuring the Elastic Search component.....</b>	<b>68</b>
<b>15 Appendix - Accessing the FCC Studio Application .....</b>	<b>70</b>
<b>16 Appendix - Tables and Sequences .....</b>	<b>71</b>
16.1 Studio Schema Tables.....	71
16.2 Studio Schema Sequences .....	71
<b>17 Appendix - Resource Allocation for FCC Studio Services .....</b>	<b>73</b>
17.1 Resource Limits .....	73
17.2 Resource Types .....	73
17.3 Resource Parameters in FCC Studio .....	73
<b>18 Appendix - Providing Spark Libraries.....</b>	<b>78</b>
<b>19 Appendix - Setting Spark Master .....</b>	<b>79</b>
<b>20 Appendix - Using Python Virtual Environments with PySpark.....</b>	<b>80</b>
20.1 Creating a Virtual Environment with Conda.....	80
20.2 Including Virtual Environment in the Init Container .....	80
20.3 Updating Interpreter Properties.....	80
<b>21 Appendix - Adding Packages to Python Interpreter .....</b>	<b>82</b>
<b>22 Appendix - Uninstalling FCC Studio.....</b>	<b>83</b>
<b>23 Appendix - Executing Graph_Alive Notebook .....</b>	<b>84</b>

# 1 Preface

This section provides supporting information for the Oracle Financial Services (OFS) Crime and Compliance Studio (FCC Studio) application Installation Guide and includes the following topics:

- [Summary](#)
- [Audience](#)
- [Related Documents](#)

## 1.1 Summary

Before you begin the deployment, 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 deploying and configuring the OFS Crime and Compliance Studio application.

The document assumes that you have experience in installing Enterprise components. The basic knowledge of the following is recommended:

- UNIX commands
- Database concepts
- Big Data
- Kubernetes
- Docker

## 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 Installation Guide*
- *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

**Table 1: Abbreviations**

<b>Abbreviation</b>	<b>Meaning</b>
FCC Studio	Financial Crime and Compliance Studio
OFSAA	Oracle Financial Services Analytical Application
BD	Behavior Detection
FCDM	Financial Crime Data Model
ICIJ	International Consortium of Investigative Journalists



## **2 Deployment Overview**

This chapter provides the information required to understand the deployment of the Oracle Financial Services (OFS) Crime and Compliance Studio (FCC Studio) application on the Kubernetes cluster.

### **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:

- Deploy a new instance FCC Studio (v8.0.7.3.0) with or without OFSAA on the Kubernetes cluster.
- Upgrade an existing instance of FCC Studio (v8.0.7.x) to v8.0.7.3.0 version with OFSAA.

## 2.2 Quick Start Steps to Deploy FCC Studio with OFSAA on the Kubernetes Cluster

**Table 1: Quick Start Steps to Deploy FCC Studio with OFSAA on the Kubernetes Cluster**

Sl. No.	Steps	Reference Links
1.	Prepare for Deployment	<ol style="list-style-type: none"> <li>1. <a href="#">Prerequisites</a></li> <li>2. <a href="#">Hardware and Software Requirements</a></li> <li>3. <a href="#">Prerequisite Environmental Settings</a></li> <li>4. <a href="#">Performing Common Pre-installation Tasks</a> <ol style="list-style-type: none"> <li>a. <a href="#">Obtaining the Software</a></li> <li>b. <a href="#">Extracting the Software</a></li> </ol> </li> <li>5. <a href="#">Required File Structure</a></li> <li>6. <a href="#">Interpreter Settings</a></li> </ol>
2.	Deploy FCC Studio with OFSAA on Kubernetes Cluster	<ol style="list-style-type: none"> <li>1. <a href="#">Configuring the Elastic Search Component</a></li> <li>2. <a href="#">Configuring Wallet</a></li> <li>3. <a href="#">Configuring the studio-env.yml File</a></li> <li>4. <a href="#">Configuring ETL Services</a></li> <li>5. <a href="#">Verifying Resource Allocation for FCC Studio Services</a></li> <li>6. <a href="#">Deploying FCC Studio on the Kubernetes Cluster</a></li> <li>7. <a href="#">Verifying the FCC Studio Deployment</a></li> <li>8. <a href="#">Verifying the Entity Resolution</a></li> <li>9. <a href="#">Accessing the FCC Studio Application</a></li> </ol>
3.	Post-deployment configuration	<ul style="list-style-type: none"> <li>• <a href="#">Configuring Interpreters</a></li> <li>• <a href="#">Performing OFSAA Configuration for Batch Execution</a></li> <li>• <a href="#">Performing Hive Data Movement</a></li> <li>• <a href="#">Performing Configuration to Run Published Notebooks</a></li> </ul>

## 2.3 Quick Start Steps to Deploy FCC Studio with Non-OFSAA

## on the Kubernetes Cluster

**Table 2: Quick Start Steps to Deploy FCC Studio with Non-OFSAA on the Kubernetes Cluster**

Sl. No.	Steps	Reference Links
1.	Prepare for Deployment	<ol style="list-style-type: none"> <li>1. <a href="#">Prerequisites</a></li> <li>2. <a href="#">Hardware and Software Requirements</a></li> <li>3. <a href="#">Prerequisite Environmental Settings</a></li> <li>4. <a href="#">Performing Common Pre-installation Tasks</a> <ol style="list-style-type: none"> <li>a. <a href="#">Obtaining the Software</a></li> <li>b. <a href="#">Extracting the Software</a></li> </ol> </li> <li>5. <a href="#">Required File Structure</a></li> <li>6. <a href="#">Interpreter Settings</a></li> </ol>
2.	Deploy FCC Studio with non-OFSAA on Kubernetes Cluster	<ol style="list-style-type: none"> <li>1. <a href="#">Configuring the Elastic Search Component</a></li> <li>2. <a href="#">Configuring Wallet</a></li> <li>3. <a href="#">Configuring the studio-env.yml File</a></li> <li>4. <a href="#">Configuring ETL Services</a></li> <li>5. <a href="#">Verifying Resource Allocation for FCC Studio Services</a></li> <li>6. <a href="#">Deploying FCC Studio on Kubernetes Cluster</a></li> <li>7. <a href="#">Verifying the FCC Studio Deployment</a></li> <li>8. <a href="#">Verifying the Entity Resolution</a></li> <li>9. <a href="#">Accessing the FCC Studio Application</a></li> </ol>
3.	Post-deployment configuration	<ul style="list-style-type: none"> <li>• <a href="#">Configuring Interpreters</a></li> <li>• <a href="#">Configuring ICIJ</a></li> </ul>

## 3 Preparing for Deployment

This chapter provides the necessary information to review before deploying FCC Studio. 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

The Linux machine must satisfy the following conditions:

- FCC Studio can be installed with or without OFSAA. To install FCC Studio with OFSAA, ensure the BD (Behavior Detection) application pack is installed.
- Kubernetes (k8s) cluster must be installed to include the following:
  - Registry to store docker images.
  - Minimum of 8 GB memory (inclusive of all nodes) available for the installation. The actual memory requirement depends on the workload/container size configuration.
  - Must not contain a namespace called, fccs. If it already exists, delete the namespace before running the deployment script.
- Docker and kubectl must be installed.
- kubectl is configured (that is, connected to cluster where you want to install FCC Studio).
- Docker has push access to a private registry.
- 12GB free space is available to store the FCC Studio Installer zip file in some directory.
- 45GB free space is available in the docker root directory. Run `docker info` command to find the docker root directory.

### 3.2 Hardware and Software Requirements

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

**Table 1: Hardware and Software Requirements**

Hardware/Software Category	Component Version
Browser	<ul style="list-style-type: none"> <li>• Chrome 57.x</li> <li>• Firefox 52.x</li> </ul>
Java Version	Java 8
Docker Registry	<ul style="list-style-type: none"> <li>• Docker registry must be present to store docker images</li> <li>• Min of 45GBspace is required to save docker images</li> </ul>

**Table 1: Hardware and Software Requirements**

Hardware/Software Category	Component Version
Database Server	<ul style="list-style-type: none"> <li>Oracle Database Server 12c Release 2 (12.2.0.1+) Enterprise Edition</li> <li>Oracle R Enterprise 1.5.1 with Oracle R Distribution 3.3.0</li> </ul>
Elastic Search	elasticsearch-7.3.2
<b>Kubernetes Cluster</b>	
Processing Server	<ul style="list-style-type: none"> <li>RHEL 7.4+</li> <li>SFTP</li> <li>Oracle JRE Standard Edition 1.8.x(with JCE)</li> <li>Kubernetes(K8s) cluster. For more information, see <a href="#">Prerequisites</a>.</li> </ul>
PGX (Graph) Server	<ul style="list-style-type: none"> <li>RHEL 7.4+</li> <li>Kubernetes(K8s) cluster. For more information, see <a href="#">Prerequisites</a>.</li> </ul>
<b>Big Data</b>	
Cloudera Distribution Hadoop 5.12	<ul style="list-style-type: none"> <li>CDH Version 5.12</li> <li>Hadoop-2.5.0+cdh5.3.3+844</li> <li>Hive-0.13.1+cdh5.3.3+350</li> <li>Sqoop1 V 1.4.4+cdh5.3.3+67</li> <li>The <code>.profile</code> file must be present with the <code>SPARK_HOME</code> and <code>PYTHON_HOME</code> parameters already set.</li> <li>Set <code>spark2-shell</code> alias in the <code>.profile</code> file as follows: <code>alias spark2-shell=spark-shell</code></li> </ul>
Cloudera Hive Connectors	Hive JDBC Connectors V 2.5.15
Hadoop Security Protocol	<ul style="list-style-type: none"> <li>Kerberos R release 1.6.1</li> <li>Sentry-1.4.0</li> </ul>

### 3.3 Prerequisite Environmental Settings

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

**Table 2: Prerequisite Information**

Category	Expected Value
Java Settings	<ul style="list-style-type: none"> <li>PATH in the <code>.profile</code> file must be set to include <code>kubectl</code> and the Java Runtime Environment (Java 8) absolute path.</li> </ul> <p><b>NOTE:</b></p> <ul style="list-style-type: none"> <li>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></li> <li>Ensure no <code>SYMBOLIC</code> links to Java installation are set in the <code>PATH</code> variable.</li> </ul>
Oracle Database Settings	<p><b>NOTE:</b> This setting is required only if the Wallet has to be created on the same server as that of the Studio server.</p> <p><b>Oracle Processing Server</b></p> <ul style="list-style-type: none"> <li><code>ORACLE_HOME</code> must be set in the <code>.profile</code> file pointing to the appropriate Oracle DB Client installation.</li> <li><code>PATH</code> in the <code>.profile</code> file must be set to include the appropriate <code>\$ORACLE_HOME/bin</code> path.</li> </ul>
Download Directory	Indicates the directory where the product installer zip file is downloaded/copied. The user permission must be set to 755 for this Download directory.
Installation Directory	Indicates the directory where the product installer zip file is extracted and the installation files are placed. The user permission must be set to 755 for this installation directory.  <b>NOTE:</b> The Installation and the Download Directory can be the same if the product installer zip file is not copied separately to another directory.
OS Locale	Linux: <code>en_US.utf8</code> Execute the following command to check the locale installed: <code>locale -a   grep -i 'en_US.utf'</code> The installed locale is displayed.
Studio Schema	<ol style="list-style-type: none"> <li>Create a new Oracle Database schema user using the following script: <code>CREATE USER &lt;Studio Schema User Name&gt; IDENTIFIED BY &lt;Password&gt;;</code> A new oracle Database schema is created.</li> <li>Grant the permissions that are given in the next row. This newly created schema is referred to as Studio Schema.</li> </ol>

**Table 2: Prerequisite Information**

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

## 3.4 Performing Common Pre-installation Tasks

The common pre-installation tasks that you must perform before installing FCC Studio are:

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

### 3.4.1 Obtaining the Software

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

1. Login to [My Oracle Support](#) with a valid Oracle account and search for the Bug ID **30995978** under the *Patches & Updates* tab.
2. Download the following installer archive files to the download directory (in Binary Mode) on the setup identified for the FCC Studio installation:
  - OFS\_FCCM\_STUDIO\_8.0.7.3.0\_LINUX\_part1of2.zip
  - OFS\_FCCM\_STUDIO\_8.0.7.3.0\_LINUX\_part2of2.zip

### 3.4.2 Extracting the Software

1. Extract the contents of the downloaded zip files using the following command:

```
unzip OFS_FCCM_STUDIO_8.0.7.3.0_LINUX_1of2.zip
unzip OFS_FCCM_STUDIO_8.0.7.3.0_LINUX_2of2.zip
```

Both the installer zip files are extracted to the same folder and the `OFS_FCCM_STUDIO` directory is obtained and is referred to as `<Studio_Installation_Path>`.

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

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

```
chmod 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 which 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 existing jars of Cloudera installation.
- The HiveJDBC4.jar file is not available in the Cloudera 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.  <b>NOTE:</b> 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 <a href="#">Appendix 11, "Appendix - Installing Oracle R Distribution"</a> .
fcc-pyspark	<ul style="list-style-type: none"> <li>Install the py4j package in the Spark cluster.</li> <li>Install the Livy server (0.5.0) on the master node of the Big Data cluster.</li> </ul>
fcc-python.	<ul style="list-style-type: none"> <li>Install the py4j package.</li> <li>If you want to add packages to the Python interpreter, see <a href="#">"Appendix - Adding Packages to Python Interpreter" on page 82</a>.</li> </ul>
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 <a href="#">Configuring PySpark Interpreter</a> .
spark:	For the required configuration, see <a href="#">Configuring Spark Interpreter</a> .

### 3.6.1 Configuring Spark Interpreter

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

### 3.6.1.1 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 [Appendix 19, "Appendix - Setting Spark Master"](#).

#### 3.6.1.1.1 Local Mode

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

#### 3.6.1.1.2 Yarn Mode

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

- Provide custom Spark libraries. For more information, see [Appendix 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.

### 3.6.1.2 Configuration

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 through the **Interpreters** page of the FCC Studio application UI.

## 3.6.2 Configuring PySpark Interpreter

- [Prerequisites](#)

- [Configuration](#)

### 3.6.2.1 Prerequisites

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

### 3.6.2.2 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, which 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 the case where your components run on different machines, you can use [Appendix 20, "Appendix - Using Python Virtual Environments with PySpark"](#).

This step is not needed if different machines have matching python versions available.

## 4 Deploying FCC Studio with OFSAA

This chapter provides the necessary information to deploy a new instance of the Crime and Compliance Studio application on the Kubernetes cluster. It includes the following sections:

1. [Configuring the Elastic Search Component](#)
2. [Configuring Wallet](#)
3. [Configuring the studio-env.yml File](#)
4. [Configuring ETL Services](#)
5. [Verifying Resource Allocation for FCC Studio Services](#)
6. [Deploying FCC Studio on the Kubernetes Cluster](#)
7. [Verifying the FCC Studio Deployment](#)
8. [Verifying the Entity Resolution](#)
9. [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 [Appendix 14, "Appendix - Configuring the Elastic Search component"](#).

### 4.2 Configuring Wallet

To configure wallets, follow these steps:

1. Create a wallet. For information on creating wallets, see [Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"](#).
2. Copy the wallet files, `cwallet.sso`, `ewallet.p12` and `tnsnames.ora`, and place in the `<Studio_Installation_Path>/configmaps/wallet` directory.

### 4.3 Configuring the studio-env.yml File

To configure the `studio-env.yml` file, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/secrets/studio-env.yml` file.
3. Configure the `studio-env.yml` file as mentioned in [Table 1](#).

#### NOTE

- You must manually set the Interaction Variable parameter values. If a value is not applicable, enter NA and ensure that the value is not entered as NULL.
- Do not alter the parameter values that are already set in the `studio-env.yml` file

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
apiVersion	For example: v1	Yes
kind	For example: Secret	Yes
<b>metadata</b>		
name	For example: studio-env	Yes
<b>stringData</b>		
NON_OFSAA	Enter "false" to deploy FCC Studio with OFSAA on the Kubernetes cluster.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.FCCMRealm</code>	Yes
OFSAA_SERVICE_URL	Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL.  <b>NOTE:</b> For OFSAAAI, the value must be in the following format: <code>https://&lt;HostName&gt;:&lt;PortNo&gt;/&lt;ContextName&gt;/rest-api</code>	Yes
LIVY_HOST_URL	Indicates the URL of the Livy application. The format for the URL is as follows: <code>http://&lt;HostName&gt;:&lt;PortNo&gt;</code>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
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 for the newly created schema.	Yes

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
STUDIO_ALIAS_NAME	<p>Indicates the Studio alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> Enter the alias name that was created during wallet creation.</p>	Yes
STUDIO_WALLET_LOCATION	<p>Indicates the Studio wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
STUDIO_TNS_ADMIN_PATH	<p>Indicates the path of the tnsnames.ora file where an entry for the STUDIO_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
BD_CONFIG_HOSTNAME	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_ALIAS_NAME	<p>Indicates the BD config alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> Enter the alias name that was created during wallet creation.</p>	Yes

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
BD_CONFIG_WALLET_LOCATION	Indicates the BD config wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_CONFIG_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BD_ATOMIC_HOSTNAME	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_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Enter the alias name that was created during wallet creation.	Yes
BD_ATOMIC_WALLET_LOCATION	Indicates the BD atomic wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BD_ATOMIC_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_ATOMIC_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
FSINFODOM	Indicates the name of the OFSAA or BD Infodom.	Yes
FSSEGMENT	Indicates the name of the OFSAA or BD segment.	Yes
DATAMOVE- MENT_LINK_NAME	<ul style="list-style-type: none"> <li>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. Alternatively, you can provide the source schema name.</li> <li>If no DB link is present, provide the &lt;SCHEMA_NAME&gt; in this parameter.</li> <li>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.</li> </ul>	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
HADOOP_CREDEN- TIAL_PROVIDER_- PATH	Indicates the path where Hadoop credential is stored.	Yes
HADOOP_PASS- WORD_ALIAS	<p>Indicates the Hadoop alias given when creating the Hadoop credentials.</p> <p><b>NOTE:</b> Enter the alias name that was created during wallet creation. For information on how to create credential keystore, see <a href="#">Creating Credential Keystore</a></p>	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 the new Hive schema name.	Yes
JAAS_CONF_FILE_- PATH	Created for future use.	No
Krb_Host_FQD- N_Name	Indicates the Kerberos host FQDN name.	Yes
Krb_Realm_Name	Indicates the Kerberos realm name.	Yes
Krb_Service_Name	Indicates the Kerberos service name. Example: Hive	Yes
KRB5_CONF_FILE_- PATH	Created for future use.	No



**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
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.	Yes
server_kerberos_principal	Created for future use.	Yes
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Big Data server where SQOOP will run.	Yes
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> <li>1. Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry to the same: <code>oracle.jdbc.mapDateToTimestamp=false</code></li> <li>2. Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsaas/</code></li> </ol> <p><b>NOTE:</b> 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_HOSTNAME	Indicates the hostname of the Big Data server where SQOOP will run. Example: <code>&lt;HostName&gt;</code>	Yes
SQOOP_TRG_PASSWORD	Indicates the password of the user of the Big Data server where SQOOP will run.	Yes
SQOOP_WORKDIR_HDFS	Indicates the SQOOP working directory in HDFS. Example: <code>/user/ofsaas</code>	Yes
AUTH_SERVICE_URL	Indicates the AUTH service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the Auth service URL is as follows: <code>http://&lt;HostName&gt;:7041/authservice</code>  <b>NOTE:</b> The value is already set. Do not change the value.	Yes

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
BATCH_SERVICE_URL	<p>Indicates the Batch service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the Batch service URL is as follows: http://&lt;HostName&gt;:7043/batchservice</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
META_SERVICE_URL	<p>Indicates the META service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the META service URL is as follows: http://&lt;HostName&gt;:7045/metaservice</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
SESSION_SERVICE_URL	<p>Indicates the Session service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the Session service URL is as follows: http://&lt;HostName&gt;:7047/sessionservice</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
PGX_SERVER_URL	<p>Indicates the URL of the PGX server.</p> <p>http://&lt;HostName&gt;:&lt;PortNo&gt;</p> <p>The value for PortNo must be 7007.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
RSERVE_USERNAME	<p>Indicates the RServe username.</p> <p>Value: oml</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
RSERVE_PASSWORD	<p>Indicates the RServe password.</p> <p>Value: password</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
HTTP_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No
HTTPS_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
REPO_CRAN_URL	Indicates the URL from where the R packages are obtained. The format for the REPO_CRAN_URL is as follows: <a href="https://cran.r-project.org/">https://cran.r-project.org/</a>	No
USERS_LIB_PATH	Indicates the path where the R packages are installed. Default value: /usr/lib64/R/library  <b>NOTE:</b> The value is already set. Do not change the value.	Yes, only if the ORE interpreter is to be used.
RSERVE_CONF_PATH	Indicates the path where the Rserve.conf file is present. Default value: /var/ore-interpreter/rserve  <b>NOTE:</b> The value is already set. Do not change the value.	Yes, only if the ORE interpreter is to be used.
ELASTIC_SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
ELASTIC_SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	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
ELASTICRESPONSE_BUFFERLIMIT_BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
MATCHING_SERVICE_HOSTNAME	Indicates the hostname of the database where matching service is installed.  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
MATCHING_SERVICE_PORT	Indicates the port number where matching service is installed.  <b>NOTE:</b> The value is already set. Do not change the value.	Yes

**Table 1: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
ER_SERVICE_URL	Indicates the URL of the entity resolution service. For example: http://<HostName>:7051  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default value: 7051  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
HDFS_GRAPH_FILES_PATH	Indicates the filepath in the HDFS where the graph.json 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 Configuring ETL Services

To configure the ETL services, follow these steps:

1. Place the Hadoop Cluster files in the <Studio\_Installation\_Path>/configmaps/spark directory. For more information on the file structure, see [Required File Structure](#).
2. Place the Kerberos files in the <Studio\_Installation\_Path>/configmaps/batchservice/user/conf/ directory. For more information on the file structure, see [Required File Structure](#).
3. Place the following jars in the <Studio\_Installation\_Path>/docker/user/batchservice/lib/ directory:
  - 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 are client/user-specific. These jars can be obtained from existing jars of Cloudera installation.
- The HiveJDBC4.jar file is not available in the Cloudera setup. You must download the same from the Cloudera website.

- Configure the `config.sh` file in `<Studio_Installation_Path>/bin` directory to replace the placeholder values as described in the following table:

**NOTE** Do not alter the parameter values that are already set in the `config.sh` file

**Table 2: Configuring config.sh File**

Parameter	Description
URL_GLOBAL_GRAPH_CONFIG_JSON	Indicates the HDFS URL where the PGX graph configuration .json file is stored at the end of the ETL. The location can be either local or hdfs path. For example: <code>hdfs:///user/ofsaas/STUDIO_ETL/graph.json</code>
PGX_SERVER_NUM_REPLICAS	Indicates the number of replicas of the PGX server. For example: 1
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: <code>GlobalGraphH</code>

- Grant Execute permission to the `<Studio_Installation_Path>/bin` directory by executing the following command:

```
chmod 755 install.sh config.sh
```

- Run the following command:

```
./install.sh
```

**NOTE**

- Execution of the `install.sh` command does not generate any log file.
- The values for the `<URL_GLOBAL_GRAPH_CONFIG_JSON>` and `<PGX_GLOBAL_GRAPH_NAME>` parameters in the `<Studio_Installation_Path>/configmaps/pgx-server/pgx.conf` file are auto-populated with the values that are configured in the `<Studio_Installation_Path>/bin/config.sh` file.

- Navigate to the `<Studio_Installation_Path>/configmaps/pgx-server/` directory and modify the `pgx.conf` file as follows:

Comment the following preload graph section:

```
<!--


```
"preload_graphs": [
  {
    "path": "<URL_GLOBAL_GRAPH_CONFIG_JSON>",
    "name": "<PGX_GLOBAL_GRAPH_NAME>"
  }
]
```


```

-->

## 4.5 Verifying Resource Allocation for FCC Studio Services

The required resources must be allocated to the FCC Studio services as per the architecture. For information on resource allocation, see [Appendix 17, "Appendix - Resource Allocation for FCC Studio Services"](#).

## 4.6 Deploying FCC Studio on the Kubernetes Cluster

To deploy FCC Studio on the Kubernetes cluster, follow these steps:

1. Navigate to the `<Studio_Installation_Path>` directory.
2. Execute the following command:

```
./fcc-studio.sh --registry <registry URL>:<registry port>
```

**NOTE** Refer to `./fcc-studio.sh -h` for usage instructions.

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

3. Verify the FCC Studio deployment. See [Verifying the FCC Studio Deployment](#).
4. Verify the Entity Resolution. [Verifying the Entity Resolution](#).

## 4.7 Verifying the FCC Studio Deployment

To verify the FCC Studio deployment, follow these steps:

1. Wait for a minimum of 10 minutes, after completing the execution of the `./fcc-studio.sh` command, and run the following command:

```
kubectl get pods -n fccs
```

The pod details are displayed to indicate the status of the services. You can also check the logs of the FCC Studio services from the Kubernetes Dashboard. For more information, see [Appendix 13, "Appendix - Checking Logs of FCC Studio Services"](#).

**NOTE** Ensure all the pods are ready before accessing the FCC Studio application.

## 4.8 Verifying the Entity Resolution

To verify the entity resolution service, follow these steps:

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

```
http://<Master_Node>:30751/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](#).

## 4.9 Accessing the FCC Studio Application

Access the FCC Studio application. For more information, see [“Appendix - Accessing the FCC Studio Application”](#) on page 70.

## 5 Post-deployment Configuration for FCC Studio Installed with OFSAA

On the successful deployment of FCC Studio, perform the following post-deployment configuration:

- [Configuring Interpreters](#)
- [Performing OFSAA Configuration for Batch Execution](#)
- [Performing Hive Data Movement](#)
- [Performing Configuration to Run Published Notebooks](#)

### 5.1 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 [OFS Crime and Compliance Studio Administration Guide](#).

### 5.2 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>/out/ficdb/bin` directory to the server where the BD pack is installed and to the `$FIC_DB_HOME/bin` directory of the OFSAA setup.
2. Execute the following command to grant Execute permission to the files:  

```
chmod +x <filenames>
```
3. Copy all the files in the `<Studio_Installation_Path>/out/ficdb/lib` directory to the `$FIC_DB_HOME/lib` directory.

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

### 5.3 Performing Hive Data Movement

To perform Hive data movement, follow these steps:

- [Configuring Schema Creation](#)
- [Creating Credential Keystore](#)
- [Configuring ETL](#)

#### 5.3.1 Configuring Schema Creation

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

##### 5.3.1.1 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>/out/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 `studio-env.yml` file.

For information on `studio-env.yml` file, see [Configuring the studio-env.yml File](#).

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

```
FCCM_Studio_SchemaCreation.sh HIVE
```

This creates tables in the Hive Schema.

4. Check Batch Service logs for more information.

### 5.3.1.2 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>/out/ficdb/lib` directory to the `<OFSAA_FIC_HOME_PATH>/ficdb/lib` directory.
2. Copy all the .sh files located in the `<Studio_Installation_Path>/out/ficdb/bin` directory to the `<OFSAA_FIC_HOME_PATH>/ficdb/bin` directory.
3. Create a Hive Schema with the name mentioned in the `HIVE_SCHEMA` parameter in the `studio-env.yml` file.

For information on `studio-env.yml` file, see [Configuring the studio-env.yml File](#).

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

```
FCCM_Studio_SchemaCreation.sh HIVE
```

This creates tables in the Hive Schema.

5. Check Batch Service logs for more information.

### 5.3.2 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.password.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 correct values of the credential keystore file path and the alias are provided in the `studio-env.yml` file.

### 5.3.3 Configuring ETL

- [Configuring Data Movement and Graph Load](#)
- [Cleaning the ICIJ Data](#)
- [Configuring the FILEPATH for ICIJ](#)
- [Configuring the Pre-load Global Graph for PGX Server](#)

#### 5.3.3.1 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>/out/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 [Crime and Compliance Studio Administration Guide](#).

#### 5.3.3.2 Cleaning the ICIJ Data

To clean the ICIJ (International Consortium of Investigative Journalists) 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.

### 5.3.3.3 Configuring the 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 file in HDFS, which 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

### 5.3.3.4 Configuring the Pre-load Global Graph for PGX Server

**NOTE**

- Ignore this section if you do not want to pre-load global graph.
- 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](#).

To configure pre-load global graph for PGX server, follow these steps:

1. Navigate to the `<Studio_Installation_Path>/configmaps/pgx-server/` directory and modify the `pgx.conf` file as follows:

Uncomment the following preload graph section in the `pgx.conf` file:

```
"preload_graphs": [
  {
    "path": "<URL_GLOBAL_GRAPH_CONFIG_JSON>",
    "name": "<PGX_GLOBAL_GRAPH_NAME>"
  }
]
```

**NOTE** The values for the `<URL_GLOBAL_GRAPH_CONFIG_JSON>` and `<PGX_GLOBAL_GRAPH_NAME>` parameters in the `<Studio_Installation_Path>/configmaps/pgx-server/pgx.conf` file are auto-populated with the values that are configured in the `<Studio_Installation_Path>/bin/config.sh` file.

2. Execute the following command:

```
kubectl -n fccs delete configmap pgx-config
kubectl -n fccs delete -f deployments/pgx-server.yml
```

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

```
kubectl -n fccs create configmap pgx-config --from-file=configmaps/pgx-server
```

```
kubectl -n fccs apply -f deployments/pgx-server.yml
```

4. Check the Kubernetes Dashboard, [Appendix 13, "Appendix - Checking Logs of FCC Studio Services"](#), to ensure that the `pgx-server` service is up and running.
5. Execute the `Graph_Alive` notebook after each time you start/restart the PGX service. For more information, see [Chapter 23, "Appendix - Executing Graph\\_Alive Notebook"](#).

## 5.4 Performing Configuration to Run 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>/out/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 on the Kubernetes Cluster

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

### 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 to FCC Studio with OFSAA on the Kubernetes Cluster**

Sl. No.	Steps	Reference Links
1.	Prepare for Deployment	<p><b>NOTE:</b> Ensure that the <a href="#">Prerequisites</a> and the <a href="#">Prerequisite Environmental Settings</a> are met before proceeding with the upgrade.</p> <ol style="list-style-type: none"> <li>1. <a href="#">Additional Cleanup for Upgrade</a></li> <li>2. <a href="#">Performing Common Pre-installation Tasks</a> <ol style="list-style-type: none"> <li>a. <a href="#">Obtaining the Software</a></li> <li>b. <a href="#">Extracting the Software</a></li> </ol> </li> <li>3. <a href="#">Required File Structure</a></li> <li>4. <a href="#">Interpreter Settings</a></li> </ol>
2.	Deploy FCC Studio with OFSAA on Kubernetes Cluster	<ol style="list-style-type: none"> <li>1. <a href="#">Configuring the Elastic Search Component</a></li> <li>2. <a href="#">Configuring Wallet</a></li> <li>3. <a href="#">Configuring the studio-env.yml File</a></li> <li>4. <a href="#">Configuring ETL Services</a></li> <li>5. <a href="#">Verifying Resource Allocation for FCC Studio Services</a></li> <li>6. <a href="#">Deploying FCC Studio on the Kubernetes Cluster</a></li> <li>7. <a href="#">Verifying the FCC Studio Deployment</a></li> <li>8. <a href="#">Verifying the Entity Resolution</a></li> <li>9. <a href="#">Accessing the FCC Studio Application</a></li> </ol>

**Table 1: Quick Start Steps to Upgrade to FCC Studio with OFSAA on the Kubernetes Cluster**

Sl. No.	Steps	Reference Links
3.	Post-deployment configuration	<p><b>NOTE:</b> 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"> <li>• <a href="#">Configuring Interpreters</a></li> <li>• <a href="#">Performing OFSAA Configuration for Batch Execution</a></li> <li>• <a href="#">Performing Hive Data Movement</a></li> <li>• <a href="#">Performing Configuration to Run Published Notebooks</a></li> </ul>

## 6.2 Additional Cleanup for Upgrade

To perform the cleanup required for the upgrade, follow these steps:

1. Navigate to the DATABASECHANGELOG table in the Studio Schema.
2. Delete the entry with the following ID:

```
FCC_DATASTUDIO_CONFIG_8.0.7
```

## 6.3 Configuring the studio-env.yml File for Upgrade

To configure the `studio-env.yml` file for an upgrade, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/secrets/studio-env.yml` file.
3. Configure the `studio-env.yml` file as follows:

**NOTE**

- You must manually set the Interaction Variable parameter values. If a value is not applicable, enter NA and ensure that the value is not entered as NULL.
- Do not alter the parameter values that are already set in the `studio-env.yml` file
- During the upgrade, ensure to provide the same BD database, Studio schema, Hive schema, and wallet related information in the `studio-env.yml`, that you used during the installation of the existing instance of the FCC Studio application.

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
apiVersion	For example: v1	Yes
kind	For example: Secret	Yes

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
<b>metadata</b>		
name	For example: studio-env	Yes
<b>stringData</b>		
NON_OFSAA	Enter "false" to deploy FCC Studio with OFSAA on the Kubernetes cluster.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.FCCMRealm</code>	Yes
OFSAA_SERVICE_URL	Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL.  <b>NOTE:</b> For OFSAAAI, the value must be in the following format: <code>https://&lt;HostName&gt;:&lt;PortNo&gt;/&lt;ContextName&gt;/rest-api</code>	Yes
LIVY_HOST_URL	Indicates the URL of the Livy application. The format for the URL is as follows: <code>http://&lt;HostName&gt;:&lt;PortNo&gt;</code>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
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 for the newly created schema.	Yes
STUDIO_ALIAS_NAME	Indicates the Studio alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Enter the alias name that was created during wallet creation.	Yes

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
STUDIO_WALLET_LOCATION	<p>Indicates the Studio wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
STUDIO_TNS_ADMIN_PATH	<p>Indicates the path of the tnsnames.ora file where an entry for the STUDIO_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
BD_CONFIG_HOSTNAME	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_ALIAS_NAME	<p>Indicates the BD config alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> Enter the alias name that was created during wallet creation.</p>	Yes
BD_CONFIG_WALLET_LOCATION	<p>Indicates the BD config wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes



**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_CONFIG_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BD_ATOMIC_HOSTNAME	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_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Enter the alias name that was created during wallet creation.	Yes
BD_ATOMIC_WALLET_LOCATION	Indicates the BD atomic wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BD_ATOMIC_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_ATOMIC_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
FSINFODOM	Indicates the name of the OFSAA or BD Infodom.	Yes
FSSEGMENT	Indicates the name of the OFSAA or BD segment.	Yes

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
DATAMOVE- MENT_LINK_NAME	<ul style="list-style-type: none"> <li>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. Alternatively, you can provide the source schema name.</li> <li>If no DB link is present, provide the &lt;SCHEMA_NAME&gt; in this parameter.</li> <li>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.</li> </ul>	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
HADOOP_CREDEN- TIAL_PROVIDER_- PATH	Indicates the path where Hadoop credential is stored.	Yes
HADOOP_PASS- WORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials.  <b>NOTE:</b> Enter the alias name that was created during wallet creation. For information on how to create credential keystore, see <a href="#">Creating Credential Keystore</a>	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 the new Hive schema name.	Yes
JAAS_CONF_FILE_- PATH	Created for future use.	No
Krb_Host_FQD- N_Name	Indicates the Kerberos host FQDN name.	Yes
Krb_Realm_Name	Indicates the Kerberos realm name.	Yes
Krb_Service_Name	Indicates the Kerberos 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

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
server_kerberos_keytab_file	Created for future use.	Yes
server_kerberos_principal	Created for future use.	Yes
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Big Data server where SQOOP will run.	Yes
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> <li>Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry to the same: <code>oracle.jdbc.mapDateToTimestamp=false</code></li> <li>Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsaas/</code></li> </ol> <p><b>NOTE:</b> 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 Big Data server where SQOOP will run. Example: <code>&lt;HostName&gt;</code>	Yes
SQOOP_TRG_PASSWORD	Indicates the password of the user of the Big Data server where SQOOP will run.	Yes
SQOOP_WORK_DIR_HDFS	Indicates the SQOOP working directory in HDFS. Example: <code>/user/ofsaas</code>	Yes
AUTH_SERVICE_URL	Indicates the AUTH service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the Auth service URL is as follows: <code>http://&lt;HostName&gt;:7041/authservice</code>  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
BATCH_SERVICE_URL	Indicates the Batch service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the Batch service URL is as follows: <code>http://&lt;HostName&gt;:7043/batchservice</code>  <b>NOTE:</b> The value is already set. Do not change the value.	Yes

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
META_SERVICE_URL	<p>Indicates the META service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the META service URL is as follows:                      http://&lt;HostName&gt;:7045/metaservice</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes
SESSION_SERVICE_URL	<p>Indicates the Session service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the Session service URL is as follows:                      http://&lt;HostName&gt;:7047/sessionservice</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes
PGX_SERVER_URL	<p>Indicates the URL of the PGX server.</p> <p>http://&lt;HostName&gt;:&lt;PortNo&gt;</p> <p>The value for PortNo must be 7007.</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes
RSERVE_USERNAME	<p>Indicates the RServe username.</p> <p>Value: oml</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
RSERVE_PASSWORD	<p>Indicates the RServe password.</p> <p>Value: password</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
HTTP_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No
HTTPS_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No
REPO_CRAN_URL	<p>Indicates the URL from where the R packages are obtained.</p> <p>The format for the REPO_CRAN_URL is as follows:  <a href="https://cran.r-project.org/">https://cran.r-project.org/</a></p>	No
USERS_LIB_PATH	<p>Indicates the path where the R packages are installed.</p> <p>Default value: /usr/lib64/R/library</p> <p><b>NOTE:</b>                      The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
RSERVE_CONF_PATH	Indicates the path where the <code>Rserve.conf</code> file is present. Default value: <code>/var/ore-interpreter/rserve</code>  <b>NOTE:</b> The value is already set. Do not change the value.	Yes, only if the ORE interpreter is to be used.
ELASTIC_SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
ELASTIC_SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	Yes
SCROLL_TIME	Indicates the duration for which the <code>scroll_size</code> 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
ELASTICRESPONSE_BUFFERLIMIT_BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
MATCHING_SERVICE_HOSTNAME	Indicates the hostname of the database where matching service is installed.  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
MATCHING_SERVICE_PORT	Indicates the port number where matching service is installed.  <b>NOTE:</b> The 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://&lt;HostName&gt;:7051</code>  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default value: 7051  <b>NOTE:</b> The value is already set. Do not change the value.	Yes

**Table 2: studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
HDFS_GRAPH_FILES_PATH	Indicates the filepath in the HDFS where the graph.json 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 Deploying FCC Studio with Non-OFSAA on the Kubernetes Cluster

This chapter provides the necessary information to deploy a new instance of FCC Studio with non-OFSAA on the Kubernetes cluster. It includes the following sections:

1. [Configuring the Elastic Search Component](#)
2. [Configuring Wallet](#)
3. [Configuring the studio-env.yml File](#)
4. [Configuring ETL Services](#)
5. [Verifying Resource Allocation for FCC Studio Services](#)
6. [Deploying FCC Studio on Kubernetes Cluster](#)
7. [Verifying the FCC Studio Deployment](#)
8. [Verifying the Entity Resolution](#)
9. [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 [Appendix 14, "Appendix - Configuring the Elastic Search component"](#).

### 7.2 Configuring Wallet

To configure wallets, follow these steps:

1. Create a wallet. For information on creating wallets, see [Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"](#).
2. Copy the wallet files, `cwallet.sso`, `ewallet.p12` and `tnsnames.ora`, and place in the `<Studio_Installation_Path>/configmaps/wallet` directory.

### 7.3 Configuring the studio-env.yml File

To configure the `studio-env.yml` file, follow these steps:

1. Login to the server as a non-root user.
2. Navigate to the `<Studio_Installation_Path>/secrets/studio-env.yml` file.
3. Configure the `studio-env.yml` file as mentioned in [Table C-1](#).

#### NOTE

- You must manually set the Interaction Variable parameter values. If a value is not applicable, enter NA and ensure that the value is not entered as NULL.
- Do not alter the parameter values that are already set in the `studio-env.yml` file

**Table C-1 studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
apiVersion	For example: v1	Yes
kind	For example: Secret	Yes
<b>metadata</b>		
name	For example: studio-env	Yes
<b>stringData</b>		
NON_OFSAA	Enter "true" to deploy FCC Studio with non-OFSAA on the Kubernetes cluster.	Yes
REALM	For example: <code>com.oracle.ofss.fccm.studio.datastudio.auth.DemoRealm</code>	Yes
OFSAA_SERVICE_URL	Indicates the URL of the OFSAA instance. Do not enter '/' at the end of the URL.  <b>NOTE:</b> For OFSAAAI, the value must be in the following format: <code>https://&lt;HostName&gt;:&lt;PortNo&gt;/&lt;ContextName&gt;/rest-api</code>	Yes
LIVY_HOST_URL	Indicates the URL of the Livy application. The format for the URL is as follows: <code>http://&lt;HostName&gt;:&lt;PortNo&gt;</code>	Yes, only if the fcc-spark-sql, fcc-spark-scala and/or fcc-pyspark interpreters are to be used.
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 for the newly created Studio schema.	Yes



**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
STUDIO_ALIAS_NAME	<p>Indicates the Studio alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> Enter the alias name that was created during wallet creation.</p>	Yes
STUDIO_WALLET_LOCATION	<p>Indicates the location of the wallet created for the Studio Schema. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
STUDIO_TNS_ADMIN_PATH	<p>Indicates the Studio TNS admin path. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a>.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
BD_CONFIG_HOSTNAME	<p>Indicates the hostname of the database where BD config schema is installed.</p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
BD_CONFIG_PORT	<p>Indicates the port of the database where BD config schema is installed.</p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
BD_CONFIG_SERVICE_NAME	<p>Indicates the service name of the database where BD config schema is installed.</p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
BD_CONFIG_SID	<p>Indicates the SID of the database where BD config schema is installed.</p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
BD_CONFIG_USERNAME	<p>Indicates the username for the BD config schema.</p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
BD_CONFIG_PASS-WORD	Indicates the password for the BD config schema.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_CONFIG_ALIAS_NAME	Indicates the BD config alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_CONFIG_WALLET_LOCATION	Indicates the BD config wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_CONFIG_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_CONFIG_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_HOST-NAME	Indicates the BD atomic schema hostname.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_PORT	Indicates the BD atomic schema port number.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_SERVICE_NAME	Indicates the BD atomic schema service name.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_SID	Indicates the BD atomic schema SID.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_USER-NAME	Indicates the username of the BD atomic schema.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

Interaction Variable Name	Significance	Required
BD_ATOMIC_PASSWORD	Indicates the password of the BD atomic schema.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_ALIAS_NAME	Indicates the BD atomic alias name. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_WALLET_LOCATION	Indicates the BD atomic wallet location. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
BD_ATOMIC_TNS_ADMIN_PATH	Indicates the path of the tnsnames.ora file where an entry for the BD_ATOMIC_ALIAS_NAME is present. For more information, see <a href="#">Appendix 10, "Appendix - Setting Up Password Stores with Oracle Wallet"</a> .  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
FSINFODOM	Indicates the name of the OFSAA or BD Infodomain.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
FSSEGMENT	Indicates the name of the OFSAA or BD segment.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
DATAMOVEMENT_LINK_NAME	<ul style="list-style-type: none"> <li>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. Alternatively, you can provide the source schema name. If no DB link is present, provide the &lt;SCHEMA_NAME&gt; in this parameter.</li> <li>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.</li> </ul> <b>NOTE:</b> Ensure to provide the value as NA.	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
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.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
HADOOP_CREDEN- TIAL_PROVIDER_- PATH	Indicates the path where Hadoop credential is stored.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
HADOOP_PASS- WORD_ALIAS	Indicates the Hadoop alias given when creating the Hadoop credentials.  <b>NOTE:</b> <ul style="list-style-type: none"> <li>Enter the alias name that was created during wallet creation.</li> </ul> For information on how to create credential keystore, see <a href="#">Creating Credential Keystore</a> . <ul style="list-style-type: none"> <li>Ensure to provide the value as NA.</li> </ul>	Yes
Hive_Host_Name	Indicates the Hive hostname.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
Hive_Port_number	Indicates the Hive port number. Contact System Administrator to obtain the port number.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
HIVE_PRINCIPAL	Indicates the Hive Principal. Contact System Administrator to obtain the HIVE_PRINCIPAL.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
HIVE_SCHEMA	Indicates the new Hive schema name.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
JAAS_CONF_FILE_- PATH	Created for future use.	No
Krb_Host_FQD- N_Name	Indicates the Kerberos host FQDN name.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
Krb_Realm_Name	Indicates the Kerberos realm name.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
Krb_Service_Name	Indicates the Kerberos service name. Example: Hive  <b>NOTE:</b> Ensure to provide the value as NA.	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.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
server_kerberos_principal	Created for future use.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
SQOOP_HOSTMACHINE_USER_NAME	Indicates the user name of the Big Data server where SQOOP will run.  <b>NOTE:</b> Ensure to provide the value as NA.	Yes
SQOOP_PARAMFILE_PATH	<ol style="list-style-type: none"> <li>Create a file with the name <code>sqoop.properties</code> in the Big Data server and add the following entry to the same:  <code>oracle.jdbc.mapDateToTimestamp=false</code></li> <li>Enter the location of the <code>sqoop.properties</code> file in the <code>SQOOP_PARAMFILE_PATH</code> parameter. Example: <code>/scratch/ofsa/</code></li> </ol> <b>NOTE:</b> <ul style="list-style-type: none"> <li>Ensure to provide the value as NA.</li> <li>Ensure that the location name ends with a <code>'/'</code>.</li> </ul>	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

Interaction Variable Name	Significance	Required
SQOOP_PARTITION_COL	<p>Indicates the column in which the HIVE table is partitioned. The value must be <code>SNAPSHOT_DT</code></p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
SQOOP_TRG_HOST_NAME	<p>Indicates the hostname of the Big Data server where SQOOP will run. Example: <code>&lt;HostName&gt;</code></p> <p><b>NOTE:</b> Ensure to provide the value as NA.</p>	Yes
SQOOP_TRG_PASSWORD	<p>Indicates the password of the user of the Big Data server where SQOOP will run.</p>	Yes
SQOOP_WORK_DIR_HDFS	<p>Indicates the SQOOP working directory in HDFS. Example: <code>/user/ofsa</code></p>	Yes
AUTH_SERVICE_URL	<p>Indicates the AUTH service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the Auth service URL is as follows: <code>http://&lt;HostName&gt;:7041/authservice</code></p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
BATCH_SERVICE_URL	<p>Indicates the Batch service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the Batch service URL is as follows: <code>http://&lt;HostName&gt;:7043/batchservice</code></p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
META_SERVICE_URL	<p>Indicates the META service URL that gets activated after the <code>fccstudio.sh</code> file runs. The format for the META service URL is as follows: <code>http://&lt;HostName&gt;:7045/metaservice</code></p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes

**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
SESSION_SERVICE_URL	<p>Indicates the Session service URL that gets activated after the fccstudio.sh file runs.</p> <p>The format for the Session service URL is as follows: http://&lt;HostName&gt;:7047/sessionservice</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
PGX_SERVER_URL	<p>Indicates the URL of the PGX server.</p> <p>http://&lt;HostName&gt;:&lt;PortNo&gt; The value for PortNo must be 7007.</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes
RSERVE_USERNAME	<p>Indicates the RServe username.</p> <p>Value: oml</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
RSERVE_PASSWORD	<p>Indicates the RServe password.</p> <p>Value: password</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
HTTP_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No
HTTPS_PROXY	Indicates the proxy for the host where FCC Studio is deployed.	No
REPO_CRAN_URL	<p>Indicates the URL from where the R packages are obtained.</p> <p>The format for the REPO_CRAN_URL is as follows: <a href="https://cran.r-project.org/">https://cran.r-project.org/</a></p>	No
USERS_LIB_PATH	<p>Indicates the path where R packages are installed.</p> <p>Value: /usr/lib64/R/library</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.
RSERVE_CONF_PATH	<p>Indicates the path where the Rserve.conf file is present.</p> <p>Value: /var/ore-interpreter/rserve</p> <p><b>NOTE:</b> The value is already set. Do not change the value.</p>	Yes, only if the ORE interpreter is to be used.

**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
ELASTIC_-SEARCH_HOSTNAME	Indicates the hostname of the database where the elastic search service is installed.	Yes
ELASTIC_-SEARCH_PORT	Indicates the port number where the elastic search service is installed.	Yes
EXECUTOR_THREADS	Indicates the number of threads to run in parallel during one scroll. For example: 10	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
ELASTICRESPONSE_BUFFERLIMIT_BYTE	Indicates the buffer size of the response obtained from the elastic search service. For example: 200	Yes
MATCHING_SERVICE_HOSTNAME	Indicates the hostname of the database where matching service is installed.  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
MATCHING_SERVICE_PORT	Indicates the port number where matching service is installed.  <b>NOTE:</b> The 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  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
ER_SERVICE_PORT	Indicates the port number where the entity resolution service is installed. Default value: 7051  <b>NOTE:</b> The value is already set. Do not change the value.	Yes
HDFS_GRAPH_FILES_PATH	Indicates the filepath in the HDFS where the graph.json is formed.	Yes
GRAPH_FILES_PATH	Indicates the directory in the Big Data server for graph files.	Yes



**Table C-1 (Continued) studio-env.yml Parameters**

InteractionVariable Name	Significance	Required
GRAPH_NAME	Indicates the name you want to assign to the global graph at the end of ETL.	Yes

## 7.4 Configuring ETL Services

To configure ETL Services, follow these steps:

1. Place the Hadoop Cluster files in the `<Studio_Installation_Path>/configmaps/spark` directory. For more information on the file structure, see [Required File Structure](#).
2. Place the Kerberos files in the `<Studio_Installation_Path>/configmaps/batchservice/user/conf/` directory. For more information on the file structure, see [Required File Structure](#).
3. Place the following jars in the `<Studio_Installation_Path>/docker/user/batchservice/lib/` directory:
  - `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 are client/user-specific. These jars can be obtained from existing jars of Cloudera installation.
- The `HiveJDBC4.jar` file is not available in the Cloudera setup. You must download the same from the Cloudera website.

4. Configure the `config.sh` file in `<Studio_Installation_Path>/bin` directory to replace the placeholder values as described in the following table:

**NOTE**

Do not alter the parameter values that are already set in the `config.sh` file

**Table 1: Configuring config.sh File**

Parameter	Description
URL_GLOBAL_GRAPH_CONFIG_JSON	Indicates the HDFS URL where the PGX graph configuration .json file is stored at the end of the ETL. The location can be either local or hdfs path. For example: <code>hdfs:///user/ofsaas/STUDIO_ETL/graph.json</code>
PGX_SERVER_NUM_REPLICAS	Indicates the number of replicas of the PGX server. For example: 1
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: <code>GlobalGraphIH</code>

5. Grant Execute permission to the `<Studio_Installation_Path>/bin` directory by executing the following command:

```
chmod 755 install.sh config.sh
```

6. Run the following command:

```
./install.sh
```

**NOTE**

- Execution of the `install.sh` command does not generate any log file.
- The values for the `<URL_GLOBAL_GRAPH_CONFIG_JSON>` and `<PGX_GLOBAL_GRAPH_NAME>` parameters in the `<Studio_Installation_Path>/configmaps/pgx-server/pgx.conf` are auto-populated with the values configured in the `<Studio_Installation_Path>/bin/config.sh` file.

7. Navigate to the `<Studio_Installation_Path>/configmaps/pgx-server/` directory and modify the `pgx.conf` file as follows:

Comment the following preload graph section:

```
<!--  
"preload_graphs": [  
  {  
    "path": "<URL_GLOBAL_GRAPH_CONFIG_JSON>",  
    "name": "<PGX_GLOBAL_GRAPH_NAME>"  
  }  
]  
-->
```

## 7.5 Verifying Resource Allocation for FCC Studio Services

The required resources must be allocated to the FCC Studio services as per the architecture. For information on resource allocation, see [Appendix 17, "Appendix - Resource Allocation for FCC Studio Services"](#)

## 7.6 Deploying FCC Studio on Kubernetes Cluster

To deploy FCC Studio on the Kubernetes cluster, follow these steps:

1. Navigate to the `<Studio_Installation_Path>`.
2. Execute the following command:

```
./fcc-studio.sh --registry <registry URL>:<registry port>
```

**NOTE**

Refer to `./fcc-studio.sh -h` for usage instructions.

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

3. Verify the deployment. See [Verifying the FCC Studio Deployment](#).
4. If you have added new data sources at this stage, you must redeploy FCC Studio. For more information, see [Chapter 9, "Redeploying FCC Studio"](#).

## 7.7 Verifying the FCC Studio Deployment

To verify the FCC Studio deployment, follow these steps:

1. Wait for a minimum of 10 minutes, after completing the execution of the `./fcc-studio.sh` command, and run the following command:

```
kubectl get pods -n fccs
```

The pod details are displayed to indicate the status of the services. You can also check the logs of the FCC Studio services from the Kubernetes Dashboard. For more information, see [Appendix 13, "Appendix - Checking Logs of FCC Studio Services"](#).

### NOTE

Ensure the metaservice is up and running before accessing the FCC Studio application.

## 7.8 Verifying the Entity Resolution

To verify the entity resolution service, follow these steps:

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

```
http://<Master_Node>:30751/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](#).

## 7.9 Accessing the FCC Studio Application

Access the FCC Studio application. For more information, see ["Appendix - Accessing the FCC Studio Application" on page 70](#).

## 8 Post-deployment Configuration for FCC Studio with Non-OFSAA

On the successful deployment of FCC Studio, perform the following post-deployment configuration:

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

### 8.1 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](#).

### 8.2 Configuring ICIJ

- [Cleaning the ICIJ Data](#)
- [Configuring the FILEPATH for ICIJ](#)

#### 8.2.1 Cleaning the ICIJ Data

To clean the ICIJ (International Consortium of Investigative Journalists) 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.

#### 8.2.2 Configuring the 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 file in HDFS, which 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 Redeploying FCC Studio

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

To redeploy FCC Studio, follow these steps:

1. Execute the following command to delete namespace of FCC Studio:
 

```
kubectl delete namespace fccs
```
2. Navigate to the `<Studio_Installation_Path>` path and correct the parameters or files as suggested by the error.
3. Navigate to `<Studio_Installation_Path>/secrets` directory and re-enter the values of sensitive information in the `studio-env.yml` file. For example, `STUDIO_DB_PASSWORD`, `HADOOP_PASSWORD-ALIAS` and so on.
4. Perform database cleanup by performing the following:

**Table 1:**

Schema	Applicable for OFSAA	Applicable for non-OFSAA
Cleanup for Studio Schema	Yes	Yes
Cleanup for BD Atomic Schema	Yes	No
Cleanup for BD Config Schema	Yes	No

5. Redeploy FCC Studio. For more information, see [Deploying FCC Studio on the Kubernetes Cluster](#).

### 9.0.1 Cleanup for Studio Schema

To cleanup 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_ols 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 cleanup 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 cleanup the BD Config schema:

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>` directory 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 -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 1:**

Schema	Applicable for OFSAA	Applicable for 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. Then 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 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](https://docs.oracle.com/cd/E83411_01/OREAD/installing-R-for-ORE.htm#ORE-AD129).

## 12 Appendix - Starting/Stopping FCC Studio Services

This section describes how to start and stop the services needed for FCC Studio and includes the following topics:

- [Starting a Service](#)
- [Stopping a Service](#)
- [Starting the FCC Studio Application](#)

### 12.1 Starting a Service

1. Navigate to the `<Studio_Installation_Path>/out/deploy.sh` file  
The `deploy.sh` file contains details of the commands to start the FCC Studio services.
2. Navigate to the `<Studio_Installation_Path>/out` directory.
3. Run the start command in the console to start a service:

For example:

```
kubectl -n fccs apply -f deployments/authservice.yml
```

You can check the logs of FCC Studio from the Kubernetes Dashboard. For more information, see [Appendix 13, "Appendix - Checking Logs of FCC Studio Services"](#).

### 12.2 Stopping a Service

1. Navigate to the `<Studio_Installation_Path>/out` directory.
2. Run the stop command in the console to stop a service:

For example:

```
kubectl -n fccs delete -f deployments/authservice.yml
```

### 12.3 Starting the FCC Studio Application

To start the FCC Studio application:

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

```
./fcc-studio.sh --registry <registry URL>:<registry port>
```

The FCC Studio application is restarted.

Check the logs of the FCC Studio services from the Kubernetes Dashboard. For more information, see [Appendix 13, "Appendix - Checking Logs of FCC Studio Services"](#).

Once all the services are up and running, the FCC Studio application can be accessed using the following URL:

```
http://<HostName>:30078
```

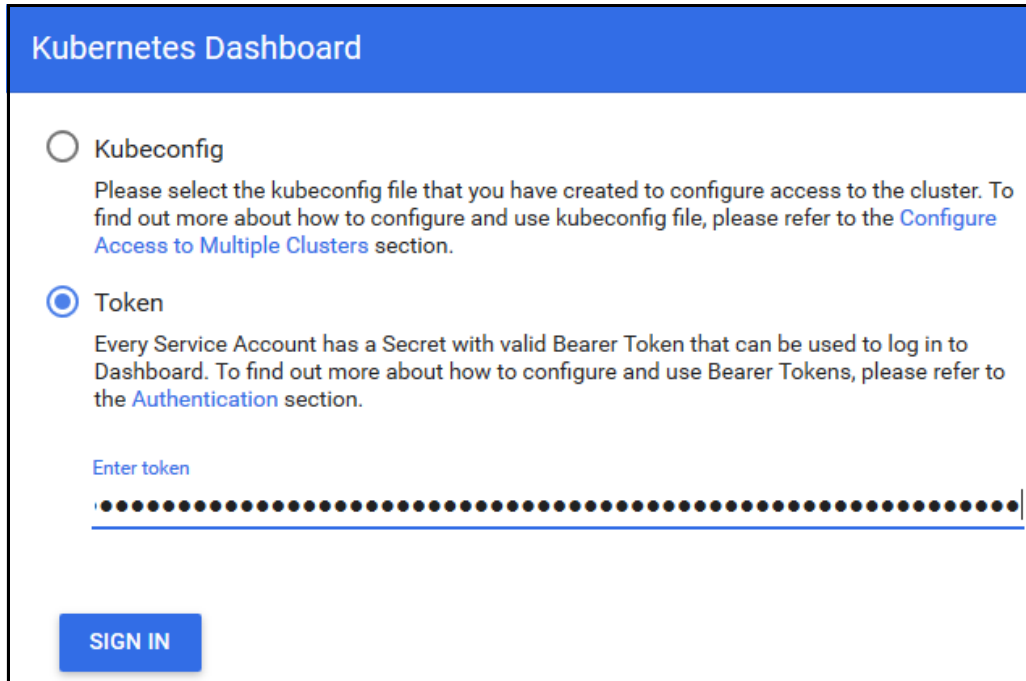
## 13 Appendix - Checking Logs of FCC Studio Services

The Kubernetes UI enables you to view the logs of the services installed as part of FCC Studio.

To check the logs of the FCC Studio services from the Kubernetes Dashboard, follow these steps:

1. Access the Kubernetes Dashboard.

The **Kubernetes Dashboard** login page is displayed.

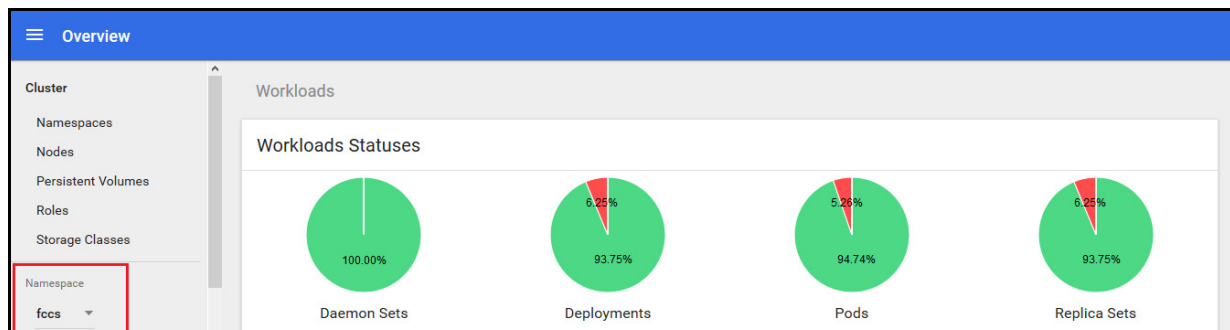


2. Select **Token** and enter the Admin user secret token in the **Enter Token** field.

3. Click **SIGN IN**.

The Kubernetes Dashboard page is displayed.

4. Select **fccs** from the **Namespace** drop-down list on the menu items displayed on the LHS.



5. Navigate to **Workloads > Pods** from the menu items displayed on the LHS.

The **Pods** page is displayed with the details of all the services installed as part of the FCC Studio installation.

| Name                      | Node     | Status                    | Restarts | Age    |
|---------------------------|----------|---------------------------|----------|--------|
| <b>fcc-studio-server</b>  | whf00bfk | Waiting: CrashLoopBackOff | 1009     | 3 days |
| fcc-pgql-t                | whf00bfk | Running                   | 0        | 3 days |
| pgx-server-l              | whf00bfk | Running                   | 0        | 3 days |
| fcc-pgx-java-l            | whf00bfk | Running                   | 0        | 3 days |
| fcc-pgx-algorithm         | whf00bfk | Running                   | 0        | 3 days |
| svclb-pgx-server-l        | whf00bfk | Running                   | 0        | 3 days |
| etl                       | whf00bfk | Running                   | 0        | 3 days |
| svclb-crosslangnamematch- | whf00bfk | Running                   | 0        | 3 days |

6. Click the service name from the **Name** column.

The service details are displayed.

7. Click the **Logs** tab.

**Details**

Name: fcc-pgql-  
 Namespace: fccs  
 Labels: app: fcc-pgql pod-template-hash:   
 Annotations: version:   
 Creation Time: 2020-01-17T11:41 UTC  
 Status: Running  
 QoS Class: Burstable

Network  
 Node: wl  
 IP:

The service logs are displayed and you can download the logs by clicking **Download Logs** icon.

Logs from python-interpreter in fcc-python-

```
11:41:23.362 [main] INFO pelin.interpreter.remote.RemoteInterpreterServer - Starting remote interpreter server on port 7777
11:41:23.374 [Thread-0] INFO pelin.interpreter.remote.RemoteInterpreterServer - Starting remote interpreter server on port 7777
```

# 14 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` file as follows:

**Table B-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   |
| <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 elastic search. |
| <code>http.port</code>                    | Indicates the port number where elastic search 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 1:** *elasticsearch.yml* File

| InteractionVariable Name | Significance  |
|--------------------------|---|
| <code>-Xms1g</code>      | <ul style="list-style-type: none"> <li>• Indicates the maximum and minimum heap memory size (mainly used for storing graphs' string properties) for the Java process of PGX.</li> <li>• Set the value for these parameters.</li> <li>• The maximum value set can be upto 50% of the ram size of the machine.</li> <li>• Recommended Value: Less than 32GB.</li> </ul> |
| <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 Elastic Search.

```
{
  "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"
}
```



## 15 Appendix - Accessing the FCC Studio Application

To access the FCC Studio application, follow these steps:

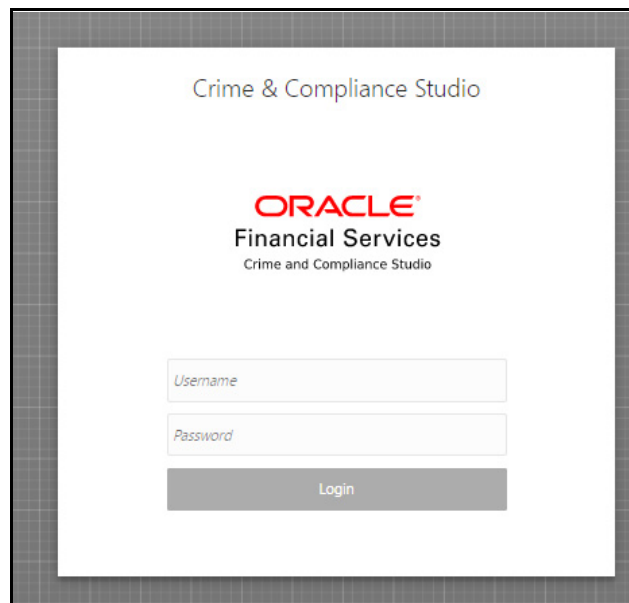
1. Enter the URL in the following format in the web browser:

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

Here <Port\_Number> is 30078 for the FCC Studio application deployed on the Kubernetes cluster.

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.

You can check the logs of the FCC Studio services from the Kubernetes Dashboard. For more information, see [Chapter 13, "Appendix - Checking Logs of FCC Studio Services"](#).

## 16 Appendix - Tables and Sequences

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

- Studio Schema Tables
- Studio Schema Sequences

### 16.1 Studio Schema Tables

The following table includes the details of the Studio Schema tables that must be dropped during the redeployment of the 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     |                           |                       |

### 16.2 Studio Schema Sequences

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

**Table 2: Studio Schema Sequences**

|             |                        |                         |
|-------------|------------------------|-------------------------|
| SEQ_COMMENT | SEQ_ENTITY_PERMISSIONS | SEQ_GRAPH               |
| SEQ_GROUP   | SEQ_INTERPRETER_RESULT | SEQ_INTERPRETER_VARIANT |

**Table 2: Studio Schema Sequences**

|                            |                        |                    |
|----------------------------|------------------------|--------------------|
| SEQ_JOB                    | SEQ_NOTEBOOK           | SEQ_PARAGRAPH      |
| SEQ_PERMISSION             | SEQ_PERMISSION_MAPPING | SEQ_RESULT_MESSAGE |
| SEQ_ROLE                   | SEQ_TASK               | SEQ_USER           |
| SEQ_VISUALIZATION_TEMPLATE |                        |                    |

# 17 Appendix - Resource Allocation for FCC Studio Services

## 17.1 Resource Limits

For FCC Studio to run reliably, the available resources of the Kubernetes cluster have to be allocated accordingly. The components are mainly memory intensive and therefore we recommend setting memory constraints for each component.

## 17.2 Resource Types

Each container requires a memory request and memory limit size as defined by the Kubernetes API. In short, containers specify a request, which is the amount of that resource that the system will guarantee to the container and a limit which is the maximum amount that the system will allow the container to use. For more information on troubleshooting tips, see [Managing Compute Resources for Containers](#).

Some components require additional resource limits which are set as environment variables.

## 17.3 Resource Parameters in FCC Studio

After extracting the FCC Studio application installer software, the resource limits have to be adjusted for each component. The configuration files can be found in the `<Studio_Installation_Path>` directory.

**NOTE**

- The sizing recommendations are preliminary. In the case of deployment failures, a manual configuration of the sizing parameters is needed.
- Depending on the use case, the recommended value changes.
- The default value in the following table is the value that is already set in the file.

**Table 1: Resource Parameters in FCC Studio**

| Configuration File/Container | Parameter type | Parameter Name                              | Description  | Recommendation |
|------------------------------|----------------|---|--|----------------|
| server.yml / server          | k8             | spec.containers[].resources.requests.memory | Memory request size for the FCC server (web application) component     | default        |
|                              | k8             | spec.containers[].resources.requests.memory | Memory limit size for the FCC server (web application) component       | default        |
| agent.yml / agent            | k8             | spec.containers[].resources.requests.memory | Memory request size for the Agent (manages all interpreters) component | default        |

**Table 1: Resource Parameters in FCC Studio**

| Configuration File/Container  | Parameter type      | Parameter Name                              | Description  | Recommendation  |
|-------------------------------|---------------------|---|--|---|
|                               | k8                  | spec.containers[].resources.limits.memory   | Memory limit size for the Agent (manages all interpreters) component   | default   |
| pgx-server.yml / pgx-server   | k8                  | spec.containers[].resources.requests.memory | Memory request size for the PGX server (manages graph processing) component  | Slightly less than the memory of the PGX server as calculated in the sizing guide.  |
|                               | k8                  | spec.containers[].resources.requests.memory | Memory limit size for the PGX server (manages graph processing) component  | The same as the request size above.   |
|                               | ENV VAR (JAVA_OPTS) | -Xmx<br>-Xms                                | The maximum and minimum heap memory size (mainly used for storing graphs' string properties) for the Java process of PGX.                    | 58% of the container's memory limit size above.<br><br>For better understanding of this sizing parameter, please consult the PGX 20.0.0-RC1 Memory Consumption documentation. |
|                               | ENV VAR (JAVA_OPTS) | -Dpgx.max_off_heap_size                     | The maximum off-heap memory size in megabytes (mainly used for storing graphs except for their string properties) that PGX tries to respect. | 42% of the container's memory limit size above.<br><br>For better understanding of this sizing parameter, please consult the PGX 20.0.0-RC1 Memory Consumption documentation. |
| fcc-pgx.yml / pgx-interpreter | k8                  | spec.containers[].resources.requests.memory | Memory request size for the PGX interpreter  | 4Gi   |
|                               | k8                  | spec.containers[].resources.limits.memory   | Memory limit size for the PGX interpreter  | 16Gi<br><br>Sizing should depend on the number and behavior (memory requirements of sessions) of concurrent users   |
| authservice.yml / authservice | k8                  | spec.containers[].resources.requests.memory | Memory request size for the authservice (used for getting roles of a user from DB) component   | default   |

**Table 1: Resource Parameters in FCC Studio**

| Configuration File/Container            | Parameter type | Parameter Name                              | Description   | Recommendation |
|---|----------------|---|---|----------------|
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the authservice (used for getting roles of a user from DB) component  | default        |
| metaservice.yml / metaservice           | k8             | spec.containers[].resources.requests.memory | Memory request size for the metaservice (used for custom interpreter api's like loaddataset, listdataset in scala interpreter etc.) component | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the metaservice (used for custom interpreter api's like loaddataset, listdataset in scala interpreter etc.) component   | default        |
| sessionservice.yml / session-service    | k8             | spec.containers[].resources.requests.memory | Memory request size for the session-service (used for managing session between pgx and scala interpreter) component                           | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the session-service (used for managing session between pgx and scala interpreter) component                             | default        |
| batchservice.yml / batchservice         | k8             | spec.containers[].resources.requests.memory | Memory request size for the batchservice (used for managing batches like sqoopjob, graph load, notebook execution etc) component              | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the batchservice (used for managing batches like sqoopjob, graph load, notebook execution etc) component                | default        |
| entity-resolution.yml/entity resolution | k8             | spec.containers[].resources.requests.memory | Memory request size for the Entity Resolution component.  | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the Entity Resolution component.  | default        |

**Table 1: Resource Parameters in FCC Studio**

| Configuration File/Container  | Parameter type | Parameter Name                              | Description  | Recommendation |
|---|----------------|---|--|----------------|
| matching-service.yml/<br>matching service   | k8             | spec.containers[].resources.requests.memory | Memory request size for the Matching Service component                 | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the Matching Service component                   | default        |
| spark.yml/spark<br>and pyspark<br>Interpreter   | k8             | spec.containers[].resources.requests.memory | Memory request size for the Spark interpreter.                         | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the Spark interpreter.                           | default        |
| fcc-jdbc.yml /<br>fcc-jdbc  | k8             | spec.containers[].resources.requests.memory | Memory request size for the jdbc connection                            | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the jdbc connection                              | default        |
| fcc-livy.yml /<br>fcc-spark-scala,<br>fcc-spark-sql,<br>and fcc-pyspark<br>interpreters | k8             | spec.containers[].resources.requests.memory | Memory request size for the livy connection to big data Spark cluster. | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the livy connection to big data Spark cluster.   | default        |
| fcc-markdown.yml /<br>markdown-<br>interpreter  | k8             | spec.containers[].resources.requests.memory | Memory request size for the Markdown interpreter                       | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the Markdown interpreter                         | default        |
| fcc-ore.yml /<br>ore-interpreter  | k8             | spec.containers[].resources.requests.memory | Memory request size for the ore connection                             | default        |
|   | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the ore connection                               | default        |

**Table 1: Resource Parameters in FCC Studio**

| Configuration File/Container        | Parameter type | Parameter Name                              | Description                                    | Recommendation        |
|-------------------------------------|----------------|---|--|-----------------------|
| fcc-python.yml / python-interpreter | k8             | spec.containers[].resources.requests.memory | Memory request size for the Python interpreter | depending on use case |
|                                     | k8             | spec.containers[].resources.limits.memory   | Memory limit size for the Python interpreter   | depending on use case |



## 18 Appendix - Providing Spark Libraries

To provide your own Spark libraries and/or Hadoop client-configuration files to connect to a Yarn cluster, follow these steps to create a new Init Container/Configmap that copies your Spark libraries directory and/or Hadoop client-configuration directory into the right location, where the Spark interpreter can access it.

To provide Spark libraries, follow these steps:

1. Download the desired Spark libraries from the [Spark Official Website](#).
2. Prepare Spark libraries.
  - a. Place your libraries in the `<Studio_Installation_Path>/docker/user/spark-interpreter-libraries/` directory.

**NOTE** Ensure that the directory name of the Spark library is prefixed with the term 'spark'

- b. To use separate Hadoop libraries, download them as well and place the Hadoop libraries directory in the same directory.
  - c. Be cautious when linking the two libraries, since the path where they are located is in the K8s pod, and the location is as follows:  

```
/var/olds-spark-interpreter/interpreter/spark/libs/
```
3. Change the image of the Spark interpreter Init Container in the `<Studio_Installation_Path>/deployments/spark.yml` file to `{{imageRepository}}/fcc-studio/3rd-party:init`.
4. (Optional) Place your Hadoop Client Configuration files in the `<Studio_Installation_Path>/configmaps/spark-interpreter-conf/` directory.

## 19 Appendix - Setting Spark Master

You can update the Spark Master by changing the default value of the `spark.master` property in the Interpreter Settings of the Spark interpreter (`spark.json`) or the **Interpreters** page of the FCC Studio application UI after startup.

- The default value of `spark.master` property is `local[*]`, which means that the interpreter will run in local mode.
- For a Yarn cluster, you must change the default value of `spark.master` property to `yarn-client`.

---

**NOTE**

The Hadoop client-configuration files are required to connect to a Yarn cluster.

## 20 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. [Including Virtual Environment in the Init Container](#)
3. [Updating Interpreter Properties](#)

### 20.1 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 using the following command:

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

**NOTE** The `<environment-name>` can be chosen freely and subsequently must be used in further commands.

3. Activate your virtual environment using:
4. Execute the following command to obtain the path to your virtual environment:

```
which python
```

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

5. Compress your virtual environment using the following command:

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

### 20.2 Including Virtual Environment in the Init Container

To include the virtual environment in the Init container, you must place the Virtual Environment in the same path as the Spark libraries. For more information, see [Appendix 18, "Appendix - Providing Spark Libraries"](#).

### 20.3 Updating Interpreter Properties

All the 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` file), change the following:

- Change the value of the `spark.yarn.dist.archives` property to `/var/olds-spark-  
interpreter/interpreter/spark/libs/<environment-name>/<environment-  
name>.tar.gz#<environment-name>`
- Change the value of the `spark.pyspark.python` property to `./<environment-  
name>/bin/python`
- In the **PySpark Interpreter Settings** page of the FCC Studio application UI (or `pyspark.json` file), change the value of the `zeppelin.pyspark.python` property to `/var/olds-spark-  
interpreter/interpreter/spark/libs/<environment-name>/bin/python`.

## 21 **Appendix - Adding Packages to Python Interpreter**

For information on adding packages to python interpreter, see *Adding Packages to Python Interpreter* chapter in the [OFS Crime and Compliance Studio Administration Guide](#).

## 22 Appendix - Uninstalling FCC Studio

To uninstall FCC Studio, follow these steps:

**NOTE**

Uninstalling the FCC Studio application deletes all the data from FCC Studio namespace.

1. Delete the FCC Studio namespace using the following command:  

```
kubectl delete namespace fccs
```
2. Manually delete the FCC studio images for each Kubernetes node using the following command:  


```
docker rmi <Image ID>
```

You can get the list of image IDs by running the docker images.

## 23 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.

