

IBM Financial Transactions Repository  
Version 2.0.2

*IBM Financial Transactions Repository  
Guide*



**Note**

Before using this information and the product it supports, read the information in [Notices](#).

**Product Information**

This document applies to Version 2.0.2 and may also apply to subsequent releases.

**Copyright**

Licensed Materials - Property of IBM

© Copyright IBM Corp. 2017.

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "[Copyright and trademark information](http://www.ibm.com/legal/copytrade.shtml)" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

© **Copyright International Business Machines Corporation 2017.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# Contents

- Introduction..... V**
- Chapter 1. IBM Financial Transactions Repository..... 1**
  - Prerequisites..... 1
  - Setting the environment variables.....1
  - Creating a Streams domain and instance.....2
  - Configuring the Kafka server information.....3
  - Installing the Financial Transactions Repository files..... 3
  - Running the ingestion script..... 3
  - Monitoring the data ingestion.....4
  - Verifying the data ingestion..... 5
  - Troubleshooting..... 6
- Appendix A. Accessibility features..... 7**
- Notices.....9**



# Introduction

---

IBM Financial Transactions Repository provides a trade repository for IBM® Surveillance Insight® for Financial Services. IBM Financial Transactions Repository processes daily trading event data from your organization's order management system from variety of formats and models. All of the incoming trading events are stored to meet regulatory auditing requirements.

## **Audience**

This guide is intended for administrators and users of the IBM Financial Transactions Repository solution. It provides information on installation and configuration of the solution, and information about using the solution.

## **Finding information and getting help**

To find product documentation on the web, access [IBM Knowledge Center](http://www.ibm.com/support/knowledgecenter/SSWTQQ) (www.ibm.com/support/knowledgecenter/SSWTQQ).

## **Accessibility features**

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products. Some of the components included in the IBM Financial Transactions Repository have accessibility features. For more information, see [Appendix A, "Accessibility features,"](#) on page 7.

The HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.

## **Forward-looking statements**

This documentation describes the current functionality of the product. References to items that are not currently available may be included. No implication of any future availability should be inferred. Any such references are not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of features or functionality remain at the sole discretion of IBM.

## **Samples disclaimer**

Sample files may contain fictional data manually or machine generated, factual data that is compiled from academic or public sources, or data that is used with permission of the copyright holder, for use as sample data to develop sample applications. Product names that are referenced may be the trademarks of their respective owners. Unauthorized duplication is prohibited.



---

# Chapter 1. IBM Financial Transactions Repository

IBM Financial Transactions Repository is a repository of trade transactions that are ingested from client systems by using analytics components that provide insights to compliance users. Financial Transactions Repository can be configured to process input from various FIX-based order management systems.

Financial Transactions Repository consists of the following components:

1. An ingestion component that is built by using IBM Streams
2. A data archive, or transaction repository, that is built on the Hadoop file system (HDFS)
3. A communication channel for indicating the status of data ingestion that is built by using Apache Kafka

Financial Transactions Repository ingests trade transaction records in FIX protocol versions 4.0, 4.1, and 4.2. The input files of the trading data must be placed in a pre-configured directory for that trading system. This is also called the watch folder for that trading system. Each trading system that ingests data into the Financial Transactions Repository must have its own watch folder.

The IBM Streams component identifies new data files in the watch folder and ingests the data into HDFS. The ingested data is archived and made available to the downstream analytics layers. Messages indicating the status of ingestion at various stages are conveyed by IBM Streams to the Apache Kafka messaging system.

---

## Prerequisites

You must have the following prerequisites before you can install IBM Financial Transactions Repository.

- IBM Surveillance Insight for Financial Services version 2.0.2
- The Hadoop libraries that are required to run the IBM Streams HDFSSink operator. You must set the HADOOP\_HOME environment variable to point to the Hadoop library location.

---

## Setting the environment variables

The IBM Financial Transactions Repository ingestion module receives its configuration by using environment variables, such as the watch directory locations, the HDFS URI, the HDFS file locations, and the Kafka topic name for the ingestion status messages. You must set the variables.

Financial Transactions Repository is installed on the computer where IBM Streams is installed.

### Procedure

1. Log on to the computer as the root user.
2. Open the `.bashrc` file in a text editor.
3. Add the following variables:

```
export HADOOP_HOME='/opt/ibm/ccs/hadoop-2.7.2/'
export HDFS_LOCATION=<SurveillanceInsightAnalyticsServer:port>
export HDFS_RAPTOR_DIRECTORY='ccs/raptor'
export HDFS_NETS_DIRECTORY='ccs/nets'
export LOCAL_RAPTOR_DIRECTORY='/opt/ibm/ccs/data/raptor'
export LOCAL_NETS_DIRECTORY='/opt/ibm/ccs/data/nets'
export TOPIC_NAME='CCSIngest'
```

Use the following table to set the values:

Table 1: Financial Transactions Repository environment variables

Variable name	Variable description	Recommended value
HADOOP_HOME	The Hadoop path	/opt/ibm/ccs/hadoop-2.7.2
HDFS_LOCATION	The server IP and port number where the HDFS web interface is running	webhdfs:// <SurveillanceInsightAnalyticsServer :port>/
HDFS_RAPTOR_DIRECTORY	The path on Hadoop where the Raptor files are stored	ccs/raptor
HDFS_NETS_DIRECTORY	The path on Hadoop where the NETS files are stored	ccs/nets
LOCAL_RAPTOR_DIRECTORY	Path of the Raptor file watch directory for the ingestion module	/opt/ibm/ccs/data/raptor
LOCAL_NETS_DIRECTORY	Path of the NETS file watch directory for the ingestion module	/opt/ibm/ccs/data/nets
TOPIC_NAME	The name of the Kafka topic to populate the status messages	'CCSIgest'

4. Save and close the file.
5. Source the `.bashrc` file for the changes to take effect.

```
source .bashrc
```

## Creating a Streams domain and instance

You must create an IBM Streams domain and instance for IBM Financial Transactions Repository.

### Procedure

1. Log on to the computer where IBM Streams is installed as the `streamsadmin` user.
2. Run the following commands to create and start a domain that is named `ccs`:

```
streamtool mkdomain -d ccs --embeddedzk
```

```
streamtool startdomain --domain-id ccs
```

3. When prompted, enter system's user name and password.
4. Run the following command to check the status of the domain:

```
streamtool getdomainstate --domain-id ccs
```

5. Run the following commands to create and start an instance that is named `ccs_instance`:

```
streamtool mkinstance -i ccs_instance --embeddedzk --domain-id ccs
```

```
streamtool startinstance --instance-id ccs_instance
```

6. Run the following command to check the status of the instance:

```
streamtool getinstancestate --instance-id ccs_instance --embeddedzk --domain-id ccs
```



## Configuring the Kafka server information

---

You must configure the Kafka server properties to enable the ingestion status communication.

### Procedure

1. Create a directory that is named `/opt/ibm/ccs/kafka`.
2. Go to the directory, and create a file that is named `producer.properties`.
3. Open the `producer.properties` file in a text editor, and add the following property:

```
bootstrap.servers=<SurveillanceInsightKafkaNodeIP:Port>
```

Where `<SurveillanceInsightKafkaNodeIP:Port>` is the server name and port number where the Kafka broker is running.

4. Save and close the `producer.properties` file.

## Installing the Financial Transactions Repository files

---

You must download and decompress the IBM Financial Transactions Repository files.

Install Financial Transactions Repository on the computer where IBM Streams is installed.

### Procedure

Decompress the Financial Transactions Repository files:

```
tar -zxvf CNPY6EN.tar.gz
```

The `tar.gz` file contains the following files:

- The Financial Transactions Repository application files: `application.CCSDataIngest.adl` and `application.CCSDataIngest.sab`
- The `IngestionScript.sh` script that creates two IBM Streams applications and runs them in the `ccs` domain. One application is for the Raptor data file ingestion and the other application ingests the NETS data files.
- The license file: `swidtag/ibm.com_IBM_Financial_Transaction_Reporting-2.0.2.swidtag`

## Running the ingestion script

---

Running the `IngestionScript.sh` script starts the IBM Streams application and starts the processes to monitor the watched folders for transaction data.

### Procedure

1. Go to the directory where you decompressed the CNPY6EN file.
2. Use the following command to add execution permissions to the script file:

```
chmod +x IngestionScript.sh
```

3. Run the script:

```
./IngestionScript.sh
```

The following image shows that the `IngestionScript.sh` was successful, and the IBM Streams application is running and monitoring the NETS and Raptor files watch directories to ingest the data into the HDFS.

```

.....
Checking Environment Configurations
.....
NDFS_LOCATION = webndfs://10.173.132.246:50070/
NDFS_RAPTOR_DIRECTORY = ocs/raptor
NDFS_NETS_DIRECTORY = ocs/nets
LOCAL_RAPTOR_DIRECTORY = /home/catuser/data/raptor
LOCAL_NETS_DIRECTORY = /home/catuser/data/nets
TOPIC_NAME = CCSIngest
.....

Starting Ingestion Graph for Raptor
.....
CDISCC079I The following number of applications were submitted to the ocs_instance instance: 1. The instance is in the ocs domain.
CDISCC080I The 136 job was submitted to the ocs_instance instance in the ocs domain.
CDISCC020I Submitted job ID#: 136
CDISCC081I The job IDs that were submitted were written to the following output file: jobidfileForRaptor.txt.
.....

Showing status of Raptor ingestion. Please wait for 25 sec
.....
Instance: ocs_instance
-----
Id State Healthy User Date Name Group
136 Running yes catuser 2017-12-06T03:50:14-0500 application:CCSDataIngest_136 default
.....

Starting Ingestion Graph for Nets
.....
CDISCC079I The following number of applications were submitted to the ocs_instance instance: 1. The instance is in the ocs domain.
CDISCC080I The 137 job was submitted to the ocs_instance instance in the ocs domain.
CDISCC020I Submitted job ID#: 137
CDISCC081I The job IDs that were submitted were written to the following output file: jobidfileForNets.txt.
.....

Showing status of Nets ingestion
.....
Instance: ocs_instance
-----
Id State Healthy User Date Name Group
137 Running yes catuser 2017-12-06T03:50:13-0500 application:CCSDataIngest_137 default
.....

```

Figure 1: Ingestion script screen

## Monitoring the data ingestion

While the ingestion script is running, you can monitor the progress from the CCSIngest Kafka topic by running a Kafka console consumer on the topic.

### Procedure

1. Log in to the Kafka server.
2. Run the following command to list the Docker containers:

```
docker ps
```

3. Open a terminal window for the Kafka docker container:

```
docker exec -it CONTAINERID bash
```

4. Go to the /opt/kafka directory.
5. Run the following command to run the Kafka consumer:

```
bin/kafka-console-consumer.sh --bootstrap-server servername:port --topic
CCSIngest --from-beginning
```

The resulting message format contains the following parameters:

- date: Current date
- protocol: The format of data, such as FIX
- stream\_id: The source from where data is coming, such as NETS or Raptor
- records: The number of records that got ingested
- tRecords: The total number of records to be ingested
- status: The status of the ingestion, such as Ingesting and DataReady

## Verifying the data ingestion

You can verify whether the files are written to the HDFS properly. After the ingestion is completed to HDFS, a status message "DataReady" is sent to the Kafka topic. This message indicates that the data is ready to be used by the other system parts.

### Procedure

1. Enter the following URL to verify that the files were written successfully to HDFS:

```
http://server_name:50070/explorer.html#/user/hadoop/ccs
```

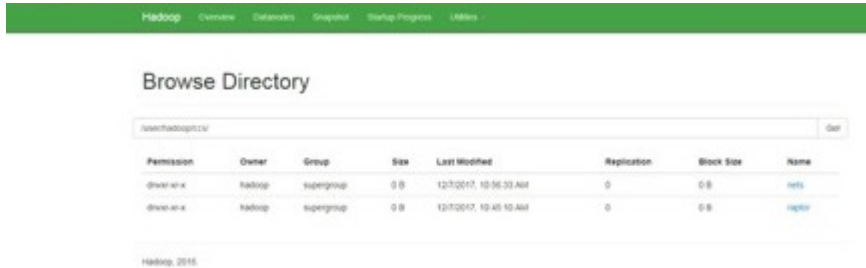


Figure 2: Hadoop file system browse directory

2. Click **raptor**.

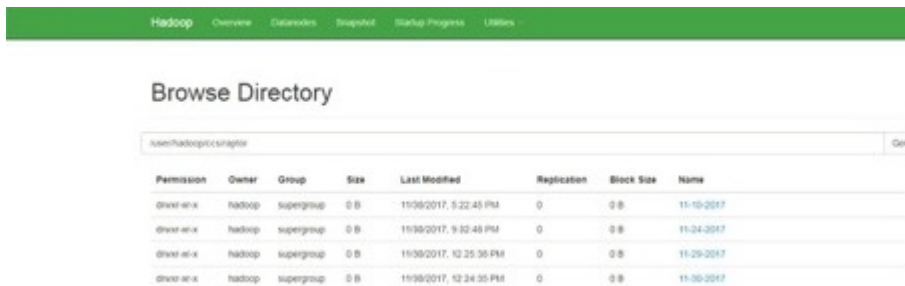


Figure 3: Hadoop Raptor file browse directory

- a) Click the current date, and look for the FIX file.

3. Click **nets**.

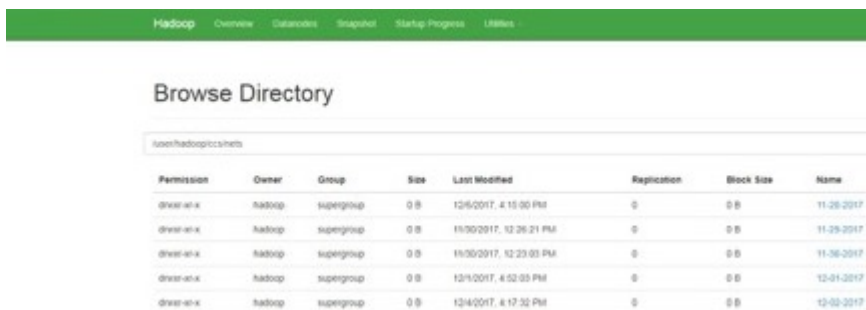


Figure 4: Hadoop NETS file browse directory

- a) Click the current date, and look for the FIX file.

## Troubleshooting

---

The following topic contains some common errors.

### **Files not ingested**

Files are placed in the local watch folder, but they are not being ingested to the HDFS:

- Verify that the HADOOP\_HOME environment variable is pointing to correct Hadoop location.
- Verify that the HADOOP\_LOCATION environment variable is pointing to the correct URL.
- Verify that the Hadoop URL is available.
- Verify that Hadoop is up and running.

### **Kafka consumer is not showing any status messages**

If the Kafka consumer is not showing any status messages after a file is placed in the local folder, verify that Kafka is up and running.

### **Graph status show unhealthy**

If the graph status shows unhealthy, verify that both Kafka and Hadoop are up and running and that all of the environment variables are set correctly.

### **Ingestion script stops and shows environment variable errors**

If the ingestion script stops and displays directory errors, verify that the environment variables are correct and that they point to valid directories.

---

## Appendix A. Accessibility features

Accessibility features help users who have a physical disability, such as restricted mobility or limited vision, to use information technology products.

For information about the commitment that IBM has to accessibility, see the [IBM Accessibility Center](http://www.ibm.com/able) ([www.ibm.com/able](http://www.ibm.com/able)).

HTML documentation has accessibility features. PDF documents are supplemental and, as such, include no added accessibility features.



## Notices

---

This information was developed for products and services offered worldwide.

This material may be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. This document may describe products, services, or features that are not included in the Program or license entitlement that you have purchased.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Software Group  
Attention: Licensing  
3755 Riverside Dr.  
Ottawa, ON  
K1V 1B7  
Canada

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

IBM Surveillance Insight for Financial Services includes Brat (v 1.3) from the following source and licensed under the following agreement:

- [http://weaver.nplab.org/~brat/releases/brat-v1.3\\_Crunchy\\_Frog.tar.gz](http://weaver.nplab.org/~brat/releases/brat-v1.3_Crunchy_Frog.tar.gz)
- <https://creativecommons.org/licenses/by-sa/3.0/legalcode>

IBM Surveillance Insight for Financial Services includes spaCy Models (v 1.2.0) from the following source and licensed under the following agreement:

- [https://github.com/explosion/spacy-models\(en\\_core\\_web\\_sm 1.2.0\)](https://github.com/explosion/spacy-models(en_core_web_sm 1.2.0))
- <https://creativecommons.org/licenses/by-sa/3.0/legalcode>

## Trademarks

---

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "[Copyright and trademark information](http://www.ibm.com/legal/copytrade.shtml)" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).





