



Red Hat Integration 2019-12

Installing Change Data Capture on RHEL

Install and use Change Data Capture (Debezium) on Red Hat Enterprise Linux (RHEL)

Red Hat Integration 2019-12 Installing Change Data Capture on RHEL

Install and use Change Data Capture (Debezium) on Red Hat Enterprise Linux (RHEL)

Legal Notice

Copyright © 2019 Red Hat, Inc.

The text of and illustrations in this document are licensed by Red Hat under a Creative Commons Attribution–Share Alike 3.0 Unported license ("CC-BY-SA"). An explanation of CC-BY-SA is available at

<http://creativecommons.org/licenses/by-sa/3.0/>

. In accordance with CC-BY-SA, if you distribute this document or an adaptation of it, you must provide the URL for the original version.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, the Red Hat logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux[®] is the registered trademark of Linus Torvalds in the United States and other countries.

Java[®] is a registered trademark of Oracle and/or its affiliates.

XFS[®] is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL[®] is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js[®] is an official trademark of Joyent. Red Hat is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack[®] Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

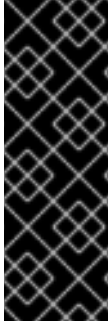
Abstract

This guide contains information about installing Change Data Capture (Debezium) connectors to Red Hat AMQ Streams 1.3 on Red Hat Enterprise Linux (RHEL). Change Data Capture connectors are deployed and managed using Kafka Connect. This Technology Preview release of Change Data Capture supports four connectors: MySQL, PostgreSQL, SQL Server, and MongoDB.

Table of Contents

CHAPTER 1. CHANGE DATA CAPTURE OVERVIEW	3
1.1. DOCUMENT CONVENTIONS	3
CHAPTER 2. INSTALLING CHANGE DATA CAPTURE CONNECTORS	5
2.1. PREREQUISITES	5
2.2. KAFKA TOPIC CREATION RECOMMENDATIONS	5
2.3. DEPLOYING CHANGE DATA CAPTURE WITH AMQ STREAMS ON RHEL	6
APPENDIX A. USING YOUR SUBSCRIPTION	8
Accessing Your Account	8
Activating a Subscription	8
Downloading Zip and Tar Files	8

CHAPTER 1. CHANGE DATA CAPTURE OVERVIEW



IMPORTANT

Technology Preview features are not supported with Red Hat production service-level agreements (SLAs) and might not be functionally complete; therefore, Red Hat does not recommend implementing any Technology Preview features in production environments. This Technology Preview feature provides early access to upcoming product innovations, enabling you to test functionality and provide feedback during the development process. For more information about support scope, see [Technology Preview Features Support Scope](#).

Red Hat Change Data Capture is a distributed platform that monitors databases and creates change event streams. Red Hat Change Data Capture is built on Apache Kafka and is deployed and integrated with AMQ Streams.

Change Data Capture captures row-level changes to a database table and passes corresponding change events to AMQ Streams. Applications can read these *change event streams* and access the change events in the order in which they occurred.

Change Data Capture has multiple uses, including:

- Data replication
- Updating caches and search indexes
- Simplifying monolithic applications
- Data integration
- Enabling streaming queries

Change Data Capture provides connectors (based on Kafka Connect) for the following common databases:

- MySQL
- PostgreSQL
- SQL Server
- MongoDB



NOTE

This guide refers to Debezium documentation. Debezium is the open source project for Change Data Capture.

1.1. DOCUMENT CONVENTIONS

Replaceables

In this document, replaceable text is styled in monospace and italics.

For example, in the following code, you will want to replace ***my-namespace*** with the name of your namespace:

```
sed -i 's/namespace: ./namespace: my-namespace/' install/cluster-operator/*RoleBinding*.yaml
```

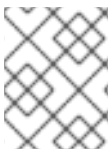
CHAPTER 2. INSTALLING CHANGE DATA CAPTURE CONNECTORS

Install Change Data Capture connectors through AMQ Streams by extending Kafka Connect with connector plugins. Following a deployment of AMQ Streams, you can deploy Change Data Capture as a connector configuration through Kafka Connect.

2.1. PREREQUISITES

A Change Data Capture installation requires the following:

- Red Hat Enterprise Linux version 7.x or 8.x with an x86_64 architecture.
- Administrative privileges (**sudo** access).
- AMQ Streams 1.3 on Red Hat Enterprise Linux is installed on the host machine.
 - AMQ Streams must be running in one of the [supported JVM versions](#).
- Credentials for the **kafka** user that was created when AMQ Streams was installed.
- An AMQ Streams cluster is running.
 - For instructions on running a basic, non-production AMQ Streams cluster containing a single ZooKeeper and a single Kafka node, see [Running a single node AMQ Streams cluster](#).



NOTE

If you have an earlier version of AMQ Streams, you need to upgrade to AMQ Streams 1.3. For upgrade instructions, see [AMQ Streams and Kafka Upgrades](#).

Additional resources

- For more information about how to install AMQ Streams, see [Installing AMQ Streams](#).

2.2. KAFKA TOPIC CREATION RECOMMENDATIONS

Change Data Capture uses multiple Kafka topics for storing data. The topics have to be either created by an administrator, or by Kafka itself by [enabling auto-creation for topics using the `auto.create.topics.enable` broker configuration](#).

The following list describes limitations and recommendations to consider when creating topics:

- Replication, a factor of at least 3 in production
- Single partition
- Infinite (or very long) retention if topic compaction is disabled
- [Log compaction](#) enabled, if you wish to only keep the *last* change event for a given record

**WARNING**

Do not enable topic compaction for the database history topics used by the MySQL and SQL Server connectors.

If you relax the single partition rule, your application must be able to handle out-of-order events for different rows in the database (events for a single row are still fully ordered). If multiple partitions are used, Kafka will determine the partition by hashing the key by default. Other partition strategies require using Simple Message Transforms (SMTs) to set the key for each record.

For log compaction, configure the **min.compaction.lag.ms** and **delete.retention.ms** topic-level settings in Apache Kafka so that consumers have enough time to receive all events and delete markers. Specifically, these values should be larger than the maximum downtime you anticipate for the sink connectors, such as when you are updating them.

2.3. DEPLOYING CHANGE DATA CAPTURE WITH AMQ STREAMS ON RHEL

This procedure describes how to set up connectors for Change Data Capture on Red Hat Enterprise Linux. Connectors are deployed to an AMQ Streams cluster using Kafka Connect, a framework for streaming data between Apache Kafka and external systems. Kafka Connect must be run in distributed mode rather than standalone mode.

This procedure assumes that AMQ Streams is installed and ZooKeeper and Kafka are running.

Procedure

1. Visit [Software Downloads](#) on the Red Hat Customer Portal and download the Change Data Capture connector or connectors that you want to use. For example, download the **Debezium 1.0.0 (Technical Preview) MySQL Connector** to use Change Data Capture with a MySQL database.
2. In **/opt/kafka**, create the **connector-plugins** directory if not already created for other Kafka Connect plugins:

```
$ sudo mkdir /opt/kafka/connector-plugins
```

3. Extract the contents of the Change Data Capture connector archive to the **/opt/kafka/connector-plugins** directory:

```
$ sudo unzip mysql-connector-java-8.0.16.jar -d /opt/kafka/connector-plugins
```

4. Repeat the above step for each connector that you want to install.
5. Switch to the **kafka** user:

```
$ su - kafka
$ Password:
```

6. Check whether Kafka Connect is already running in distributed mode. If it is running, stop the associated process on all Kafka Connect worker nodes. For example:

```
$ jcmd | grep ConnectDistributed
18514 org.apache.kafka.connect.cli.ConnectDistributed /opt/kafka/config/connect-
distributed.properties
$ kill 18514
```

7. Edit the **connect-distributed.properties** file in **/opt/kafka/config/** and specify the location of the Change Data Capture connector:

```
plugin.path=/opt/kafka/connector-plugins
```

8. Run Kafka Connect in distributed mode:

```
$ /opt/kafka/bin/connect-distributed.sh /opt/kafka/config/connect-distributed.properties
```

Kafka Connect runs. During startup, Change Data Capture connectors are loaded from the **connector-plugins** directory.

9. Repeat steps 6–8 for each Kafka Connect worker node.

Additional resources

- [Kafka Connect in distributed mode](#).
- [Adding connector plugins](#).

Updating Kafka Connect

If you need to update your deployment, amend the Change Data Capture connector JAR files in the **/opt/kafka/connector-plugins** directory, and then restart Kafka Connect.

Next Steps

The *Change Data Capture User Guide* describes how to configure each connector and its source database for change data capture. Once configured, a connector will connect to the source database and produce events for each inserted, updated, and deleted row or document.

APPENDIX A. USING YOUR SUBSCRIPTION

Change Data Capture is provided through a software subscription. To manage your subscriptions, access your account at the Red Hat Customer Portal.

Accessing Your Account

1. Go to access.redhat.com.
2. If you do not already have an account, create one.
3. Log in to your account.

Activating a Subscription

1. Go to access.redhat.com.
2. Navigate to **My Subscriptions**.
3. Navigate to **Activate a subscription** and enter your 16-digit activation number.

Downloading Zip and Tar Files

To access zip or tar files, use the customer portal to find the relevant files for download. If you are using RPM packages, this step is not required.

1. Open a browser and log in to the Red Hat Customer Portal **Product Downloads** page at access.redhat.com/downloads.
2. Locate the **Red Hat Change Data Capture** entries in the **JBOSS INTEGRATION AND AUTOMATION** category.
3. Select the desired Change Data Capture product. The **Software Downloads** page opens.
4. Click the **Download** link for your component.

Revised on 2019-12-17 21:40:36 UTC