

Migrating Oozie configurations to CDP

Date published: 2020-04-30

Date modified:

CLouDERA

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Migrating Oozie configurations to CDP.....	4
Migrating Oozie Shared Libraries.....	4

Migrating Oozie configurations to CDP

After you migrate the Oozie data to CDP, you must first configure Oozie, and then migrate custom ShareLib JARs to your new cluster.

You must configure Oozie to work with different CDP services including Sqoop actions, Yarn jobs, and HDFS HA. For information on configuring these services, see [Configure Oozie](#).



Note: By default, the Oozie service schedules an internal job to purge all the Oozie workflows older than 30 days from the database. However, actions associated with long-running coordinators do not purge until the coordinators complete. Cloudera recommends that you configure Oozie with an empty database for CDP. To view the old data, you can take a backup and run SQL queries on that data. You must recreate long-running coordinator jobs because the CDP environment is different from that of HDP and CDH.

After you configure Oozie, you must migrate the custom ShareLib JARs to the new cluster. Only the custom JARs and not the service and library JARs as those JARs are used by the components, will be present in Oozie's sharelib.

While upgrading from CDH and HDP, there might have been major third-party library version upgrades. If you have any custom JARs which include third-party libraries, then you have to rebuild those after you sync their versions with the library versions used in CDP. Otherwise, this can lead to runtime binary incompatibility.

If you are upgrading from a stack where the components used Guava version 11 to a CDP version which uses Guava version 28, then you also need to update your custom application JAR to use Guava version 28.

Migrating Oozie Shared Libraries

After you migrate the Oozie data to CDP, you must first configure Oozie, and then migrate custom ShareLib jars to your new cluster.

Procedure

1. Copy the Oozie ShareLib jar from your HDP or CDH cluster:

```
cp /user/oozie/share/lib/lib_{TIMESTAMP}/{COMPONENT}
```

The location of the Oozie ShareLib is the same across the HDP, CDH, and CDP environments.

2. Paste the ShareLib in the new file system of the CDP environment.

```
paste /user/oozie/share/lib/lib_{TIMESTAMP}/{COMPONENT}
```



Note: These files must be present on the storage like HDFS, S3, and so on, and not on the hosts where you installed Oozie.

3. Run a ShareLib update:

```
oozie admin -oozie {OOZIE_URL} -sharelibupdate
```

What to do next

After you migrate the custom ShareLib jars, update the workflow XML files for DFS paths, JDBC URLs (for example, Hive), and so on, to manage the new environment.