

Qlik Compose for Data Lakes Release Notes - April 2020 (v6.6)

This version of Compose for Data Lakes provides enhancements to earlier versions of the product.

In these release notes:

- » [Upgrade](#)
- » [Qlik Product Compatibility](#)
- » [Newly Supported Platform Versions](#)
- » [Enhancements](#)
- » [Resolved Issues](#)
- » [Known Issues](#)

Upgrade

Coexistence with Compose for Data Warehouses

If Compose for Data Lakes is installed on the same machine as Compose for Data Warehouses, before upgrading, first stop the Attunity Compose for Data Warehouses service. Otherwise, the upgrade will be blocked.

Upgrading the Hortonworks JDBC Driver

For optimal performance and connectivity, we always recommend working with the latest driver. However, if you do not encounter any issues, you can continue working with your current driver.

Qlik Product Compatibility

Qlik Compose for Data Lakes 6.6 is compatible with the following Qlik products only:

- » **Qlik Replicate**: Replicate Versions 6.6 and 6.5.
- » **Qlik Enterprise Manager (QEM)**: Version 6.6

Newly Supported Platform Versions

Support was added for the following:

- » Windows Server 2019 (64-bit)
- » Amazon EMR 5.29 (with regular Hive metastore or using AWS Glue Catalog as the metastore)
- » Databricks 6.3. Fully binary compatible versions are also supported.
- » Google Dataproc 1.4

Enhancements

- » Improved metadata performance. You can now choose to generate storage tasks without validation if you are sure that the storage tables are adjusted properly and the mapping is valid. As a result, the generation of storage tasks is much faster.
- » When provisioned data files are transferred to a Google BigQuery target, you can now set the project ID specifically without having to use the Google Dataproc project ID as previously required.

Resolved Issues

The following are the resolved issues in this release.

Component/ Process	Type	Description	Ref #
Server	Issue	Unable to define a provisioning task with Google BigQuery as the target type, as the PROJECT_ID parameter was missing in the UI.	197018

Component/ Process	Type	Description	Ref #
Documentation	Issue	<p>URL redirect configuration in Windows Server High Availability (HA) Cluster environment was incorrect.</p> <p>In a Windows HA Cluster environment, you need to define the cluster network address, and not allow Compose to create it dynamically.</p> <p>For example:</p> <pre><PRODUCT_DIR>\bin>ComposeCtl.exe -d <COMPOSE_DATA_FOLDER> configuration set - -address <CLUSTER_NAME></pre>	195374
Storage	Issue	After resetting a project, the tables counter in the storage zone would not be reset to zero.	194403
Engine	Issue	Compose would not delete log files after the maximum number of days were specified for a log file to exist before being deleted.	191692
Dataproc 1.4	Issue	<p>Unable to install the Compose Agent on Google Dataproc cluster version 1.4.</p> <p>Workaround:</p> <p>Copy the file from the Dataproc 1.4 file location to the Dataproc 1.3 file location, as follows:</p> <pre>cp /usr/local/share/google/dataproc/lib/gcs-connector-hadoop2-1.9.17.jar /usr/lib/hadoop/lib/gcs-connector-hadoop2-1.9.17.jar</pre> <p>Then restart the Compose Agent.</p>	1906177
Server	Issue	When deleting a task definition, the task's context would not be deleted.	1887609
Task Commands	Issue	An error in the SQL logic would cause Compose to read the same partition repeatedly.	1874660

Component/ Process	Type	Description	Ref #
Server	Issue	When testing for characters such as column names with "#", the number of permitted characters was too restricted for column names.	1871843
Task Commands	Issue	In some cases, the <code>header__deleted</code> column would receive a NULL value, when it should have only 0 or 1. If this happens, you should regenerate the tasks which include the mappings for the tables with the incorrect values.	1871196
Engine Issue	Issue	In some cases, adding JDBC parameters would cause corruption of the Compose repository.	1864283
Schema Evolution	Issue	When a new column was added to the model based on the source data type instead of the target data type, an error would be displayed on validation.	CMPS-8295
Upgrade	Issue	Compose Agent tasks and services would still be running during upgrade.	CMPS-8838
Provisioning Tasks	Issue	Provisioning task on Google BigQuery target would fail due to handling of specific date 1-1-1. The issue was resolved by replacing the 1-1-1 date with the default minimum date (1-1-1780).	CMPS-8865
Workflow	Issue	If the workflow contained invalid looping, the monitor would display an empty screen.	CMPS-8893

Known Issues

The following are the known issues in this release.

Component/ Process	Description	Ref #
Provisioning Tasks	If you are using Amazon EMR Hive distribution version 5.20.0 or higher, the value for the <code>spark.sql.parquet.fs.optimized.committer.optimization-enabled</code> parameter is set by default to be True . Before running a Spark provisioning task, you must configure this parameter's value to be False - i.e. <code>spark.sql.parquet.fs.optimized.committer.optimization-enabled=false</code>	191640
Metadata	In an Apache Spark project, when deleting an attribute from the Metadata and then adding back the same attribute to the Metadata, the affected tables need to be dropped and recreated.	CMPS-7862
Provisioning Tasks	In Spark projects, Full Load overwrites previous data, so when running two tasks on the same target, the second task overwrites the first task results. Splitting the Full Load into several tasks also doesn't work for the same reason.	CMPS-7457
Schema Evolution	<p>When a table is added to the source, the new table is added to the task (the mapping is created correctly and associated with the task), and the table is populated with data.</p> <p>However, when opening the mapping, there is an error that the source table does not exist in the database.</p> <p>Workaround:</p> <p>Click Clear Cache in the Manage Data Storage Tasks window (and then open the mapping again).</p>	CMPS-7210

Component/ Process	Description	Ref #
User Interface	When displaying an entity in the Physical Metadata tab and going back to Logical Metadata tab, the entity initially appears without any attributes. Workaround: Click on another entity and then go back to the entity you wanted to view.	CMPS-5421
Schema Change	After running a Compose for Data Lakes Change Processing task and a schema change occurs, if you run the Compose Full Load task again, the task will fail. Workaround: Run the Replicate Full Load again.	CMPS-5394
Replicate Control Tables	If Replicate's attrep_ddl_history and attrep_cdc_partitions Control Tables are not in the same schema, Compose fails at runtime.	CMPS-5346
