



Informatica® Informatica  
10.1.1 Update 2

# Release Notes

Informatica Informatica Release Notes  
10.1.1 Update 2  
March 2017

© Copyright Informatica LLC 1998, 2019

Publication Date: 2019-04-09

# Table of Contents

Abstract..... iv

**Chapter 1: Technical Preview for Big Data Management..... 5**

**Chapter 2: Support Changes..... 6**

**Chapter 3: Installation and Upgrade..... 8**

**Chapter 4: New Products..... 10**

**Chapter 5: New Features and Enhancements..... 11**

**Chapter 6: Changes..... 14**

**Chapter 7: Fixed Limitations..... 16**

**Chapter 8: Known Limitations..... 19**

**Chapter 9: Third-Party Fixed Limitations..... 25**

**Chapter 10: Third-Party Known Limitations..... 26**

**Chapter 11: Informatica Global Customer Support..... 28**

# Abstract

This document contains important information about new products, new features, changes, and limitations for Informatica 10.1.1 Update 2. This release affects Big Data Management, Enterprise Information Catalog, Intelligent Data Lake, and PowerExchange Adapters.

## CHAPTER 1

# Technical Preview for Big Data Management

Big Data Management version 10.1.1 Update 2 includes functionality that is available for technical preview. Technical preview functionality is supported but is unwarranted and is not production-ready. Informatica recommends that you use in non-production environments only. Informatica intends to include the preview functionality in an upcoming GA release for production use, but might choose not to in accordance with changing market or technical circumstances. For more information, contact Informatica Global Customer Support.

Effective in version 10.1.1 Update 2 the following distributions are available for technical preview:

### **Azure HDInsight 3.5**

Version 10.1.1 Update 2 includes the following new functionality:

- You can use Windows Azure Data Storage Blob (WASB) and Azure Data Lake Store (ADLS) as primary storage in an Azure HDInsight cluster.

### **Cloudera CDH 5.10**

Version 10.1.1 Update 2 includes the following new functionality:

- You can use Cloudera Connector Powered by Teradata to run Teradata mappings through Sqoop on the Blaze engine.
- You can use SSL and TLS security modes, including Kerberos authentication, Apache Ranger, Apache Sentry, Name node high availability, and Resource Manager high availability.
- You can read from and write to Hive on Amazon S3 buckets.

**Note:** This functionality is also available for Cloudera CDH versions 5.8 and 5.9.

## CHAPTER 2

# Support Changes

### Distribution support changes for Big Data Management

The following table lists the supported Hadoop distribution versions and changes in 10.1.1 Update 2:

Distribution	Supported Versions	10.1.1 Update 2 Changes
Amazon EMR	5.0.0	No change.
Azure HDInsight	3.5 *	Added support for version 3.5 Dropped support for version 3.4.
Cloudera CDH	5.8, 5.9, 5.10 *	Added support for version 5.10.
Hortonworks HDP	2.3, 2.4, 2.5	Added support for versions 2.3 and 2.4.
IBM BigInsights	4.2	No change.
MapR	5.2	Reinstated support. Added support for version 5.2. Dropped support for version 5.1.
<i>*Azure HDInsight 3.5 and Cloudera CDH 5.10 are available for technical preview. Technical preview functionality is supported but is not production-ready. Informatica recommends that you use in non-production environments only.</i>		

For a complete list of Hadoop support, see the Product Availability Matrix on Informatica Network:

<https://network.informatica.com/community/informatica-network/product-availability-matrices>

### Dropped support for Teradata Connector for Hadoop (TDCH) and Teradata PT objects on the Blaze engine

Effective in version 10.1.1 Update 2, Informatica dropped support for Teradata Connector for Hadoop (TDCH) on the Blaze engine. The configuration for Sqoop connectivity in 10.1.1 Update 2 depends on the Hadoop distribution:

#### IBM BigInsights and MapR

You can configure Sqoop connectivity through the JDBC connection. For information about configuring Sqoop connectivity through JDBC connections, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

### Cloudera CDH

You can configure Sqoop connectivity through the Teradata PT connection and the Cloudera Connector Powered by Teradata.

1. Download the Cloudera Connector Powered by Teradata .jar files and copy them to the node where the Data Integration Service runs. For more information, see the *Informatica 10.1.1 Update 2 PowerExchange for Teradata Parallel Transporter API User Guide*.
2. Move the configuration parameters that you defined in the `InfaTDCHConfig.txt` file to the **Additional Sqoop Arguments** field in the Teradata PT connection. See the Cloudera Connector Powered by Teradata documentation for a list of arguments that you can specify.

### Hortonworks HDP

You can configure Sqoop connectivity through the Teradata PT connection and the Hortonworks Connector for Teradata.

1. Download the Hortonworks Connector for Teradata .jar files and copy them to the node where the Data Integration Service runs. For more information, see the *Informatica 10.1.1 Update 2 PowerExchange for Teradata Parallel Transporter API User Guide*.
2. Move the configuration parameters that you defined in the `InfaTDCHConfig.txt` file to the **Additional Sqoop Arguments** field in the Teradata PT connection. See the Hortonworks Connector for Teradata documentation for a list of arguments that you can specify.

**Note:** You can continue to use TDCH on the Hive engine through Teradata PT connections.

### Deprecated support of Sqoop connectivity through Teradata PT data objects and Teradata PT connections

Effective in version 10.1.1 Update 2, Informatica deprecated Sqoop connectivity through Teradata PT data objects and Teradata PT connections for Cloudera CDH and Hortonworks. Support will be dropped in a future release.

To read data from or write data to Teradata by using TDCH and Sqoop, Informatica recommends that you configure Sqoop connectivity through JDBC connections and relational data objects.

## CHAPTER 3

# Installation and Upgrade

This section describes the installation instructions for 10.1.1 Update 2.

### Upgrade Paths

To apply updates to multiple big data products for Informatica, apply them in the following order:

1. Before you apply version 10.1.1 Update 2, stop the Informatica domain.
2. Apply Update 2 for Big Data Management.
3. Apply Update 2 for Enterprise Information Catalog.
4. Apply Update 2 for Intelligent Data Lake.

### Big Data Management

For information about installing and configuring Big Data Management 10.1.1 Update 2, see the *Informatica Big Data Management 10.1.1 Update 2 Installation and Configuration Guide*.

### Enterprise Information Catalog

To install Enterprise Information Catalog 10.1.1 Update 2, you must first install or upgrade to Enterprise Information Catalog 10.1.1.

For more information about installing Enterprise Information Catalog 10.1.1, see the *Informatica Live Data Map 10.1.1 Installation and Configuration Guide*.

To install Enterprise Information Catalog 10.1.1 Update 2, perform the following tasks:

1. Download and extract one of the following packages to install Enterprise Information Catalog:
  - **Informatica\_1011U2\_Live\_Data\_Map\_linux-x64**. Contains updates for Linux (RHEL 6.5 and RHEL 7.0).
  - **Informatica\_1011U2\_Live\_Data\_Map\_suse11-x64**. Contains updates for SUSE 11SP3 and SUSE 12.
2. In the `Input.properties` file, set the `DEST_DIR` parameter to the destination directory on the machine where Enterprise Information Catalog runs.
3. Type `installEBF.sh` from the command line to install Enterprise Information Catalog.
4. Remove the `<CatalogServiceName>` directory under `$INFA_Home/tomcat/temp/` before you start Informatica domain.

### Intelligent Data Lake

To install Intelligent Data Lake 10.1.1 Update 2, you must first install or upgrade to Intelligent Data Lake 10.1.1.

For more information about installing Intelligent Data Lake 10.1.1, see the *Informatica Intelligent Data Lake 10.1.1 Installation and Configuration Guide*.



To install Intelligent Data Lake 10.1.1 Update 2, perform the following tasks:

1. Download and extract one of the following packages to install Intelligent Data Lake:
  - **Informatica\_1011U2\_Intelligent\_Data\_Lake\_linux-x64**. Contains updates for Linux (RHEL 6.5 and RHEL 7.1).
  - **Informatica\_1011U2\_Intelligent\_Data\_Lake\_suse11-x64**. Contains updates for SUSE 12.
2. In the `Input.properties` file, set the `DEST_DIR` parameter to the destination directory on the machine where Intelligent Data Lake Catalog runs.
3. Type `installEBF.sh` from the command line to install Intelligent Data Lake.

## CHAPTER 4

# New Products

This section describes new products in version 10.1.1 Update 2.

### [PowerExchange for MapR-DB](#)

Effective in version 10.1.1 Update 2, you can use PowerExchange for MapR-DB to read data from and write data to MapR-DB binary tables.

PowerExchange for MapR-DB uses the HBase API to connect to MapR-DB. To connect to a MapR-DB table, you must create an HBase connection in which you must specify the database type as MapR-DB. You must create an HBase data object read or write operation, and add it to a mapping to read or write data.

You can validate and run mappings in the native environment or on the Blaze engine in the Hadoop environment.

For more information, see the *Informatica PowerExchange for MapR-DB 10.1.1 Update 2 User Guide*.

## CHAPTER 5

# New Features and Enhancements

This section describes new features and enhancements in version 10.1.1 Update 2.

### Big Data

#### **Truncate Hive table partitions on mappings that use the Blaze run-time engine**

Effective in version 10.1.1 Update 2, you can truncate Hive table partitions on mappings that use the Blaze run-time engine.

For more information about truncating partitions in a Hive target, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

#### **Filters for partitioned columns on the Blaze engine**

Effective in version 10.1.1 Update 2, the Blaze engine can push filters on partitioned columns down to the Hive source to increase performance.

When a mapping contains a Filter transformation on a partitioned column of a Hive source, the Blaze engine reads only the partitions with data that satisfies the filter condition. To enable the Blaze engine to read specific partitions, the Filter transformation must be the next transformation after the source in the mapping.

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

#### **OraOop support on the Spark engine**

Effective in version 10.1.1 Update 2, you can configure OraOop to run Sqoop mappings on the Spark engine. When you read data from or write data to Oracle, you can configure the direct argument to enable Sqoop to use OraOop.

OraOop is a specialized Sqoop plug-in for Oracle that uses native protocols to connect to the Oracle database. When you configure OraOop, the performance improves.

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

#### **Sqoop support for native Teradata mappings on Cloudera clusters**

Effective in version 10.1.1 Update 2, if you use a Teradata PT connection to run a mapping on a Cloudera cluster and on the Blaze engine, the Data Integration Service invokes the Cloudera Connector Powered by Teradata at run time. The Data Integration Service then runs the mapping through Sqoop.

For more information, see the *Informatica 10.1.1 Update 2 PowerExchange for Teradata Parallel Transporter API User Guide*.

### **Scheduler support on Blaze and Spark engines**

Effective in version 10.1.1 Update 2, the following schedulers are valid for Hadoop distributions on both Blaze and Spark engines:

- Fair Scheduler. Assigns resources to jobs such that all jobs receive, on average, an equal share of resources over time.
- Capacity Scheduler. Designed to run Hadoop applications as a shared, multi-tenant cluster. You can configure Capacity Scheduler with or without node labeling. Node label is a way to group nodes with similar characteristics.

For more information, see the Mappings in the Hadoop Environment chapter of the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

### **Support for YARN queues on Blaze and Spark engines**

Effective in version 10.1.1 Update 2, you can direct Blaze and Spark jobs to a specific YARN scheduler queue. Queues allow multiple tenants to share the cluster. As you submit applications to YARN, the scheduler assigns them to a queue. You configure the YARN queue in the Hadoop connection properties.

For more information, see the Mappings in the Hadoop Environment chapter of the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

### **Hadoop security features on IBM BigInsights 4.2**

Effective in version 10.1.1 Update 2, you can use the following Hadoop security features on the IBM BigInsights 4.2 Hadoop distribution:

- Apache Knox
- Apache Ranger
- HDFS Transparent Encryption

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management Security Guide*.

### **SSL/TLS security modes**

Effective in version 10.1.1 Update 2, you can use the SSL and TLS security modes on the Cloudera and HortonWorks Hadoop distributions, including the following security methods and plugins:

- Kerberos authentication
- Apache Ranger
- Apache Sentry
- Name node high availability
- Resource Manager high availability

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management Installation and Configuration Guide*.

### **Hive sources and targets on Amazon S3**

Effective in version 10.1.1 Update 2, Big Data Management supports reading and writing to Hive on Amazon S3 buckets for clusters configured with the following Hadoop distributions:

- Amazon EMR
- Cloudera
- HortonWorks
- MapR
- BigInsights

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

## Enterprise Information Catalog

### File System resource

Effective in version 10.1.1 Update 2, you can create a **File System** resource to import metadata from files in Windows and Linux file systems.

For more information, see the *Informatica 10.1.1 Update 2 Live Data Map Administrator Guide*.

### Apache Ranger-enabled clusters

Effective in version 10.1.1 Update 2, you can deploy Enterprise Information Catalog on Apache Ranger-enabled clusters. Apache Ranger provides a security framework to manage the security of the clusters.

### Enhanced SSH support for deploying Informatica Cluster Service

Effective in version 10.1.1 Update 2, you can deploy Informatica Cluster Service on hosts where Centrify is enabled. Centrify integrates with an existing Active Directory infrastructure to manage user authentication on remote Linux hosts.

## Intelligent Data Lake

### Hadoop ecosystem

Effective in version 10.1.1 Update 2, you can use following Hadoop distributions as a Hadoop data lake:

- Cloudera CDH 5.9
- Hortonworks HDP 2.3, 2.4, and 2.5
- Azure HDInsight 3.5
- Amazon EMR 5.0
- IBM BigInsights 4.2

### Using MariaDB for the Data Preparation Service

Effective in version 10.1.1 Update 2, you can use MariaDB 10.0.28 for the Data Preparation Service repository.

### Viewing column-level lineage

Effective in version 10.1.1 Update 2, data analysts can view lineage of individual columns in a table corresponding to activities such as data asset copy, import, export, publication, and upload.

### SSL/TLS support

Effective in version 10.1.1 Update 2, you can integrate Intelligent Data Lake with Cloudera 5.9 clusters that are SSL/TLS enabled.

## PowerExchange for Amazon Redshift

Effective in version 10.1.1 Update 2, you can select multiple schemas for Amazon Redshift objects.

For more information, see the *Informatica 10.1.1 Update 2 PowerExchange for Amazon Redshift User Guide*.

## CHAPTER 6

# Changes

This section describes the changes made in version 10.1.1 Update 2.

### Sqoop

Effective in version 10.1.1 Update 2, you can no longer override the user name and password in a Sqoop mapping by using the `--username` and `--password` arguments. Sqoop uses the values that you configure in the **User Name** and **Password** fields of the JDBC connection.

For more information, see the *Informatica 10.1.1 Update 2 Big Data Management User Guide*.

### Enterprise Information Catalog

This section describes the changes to the Enterprise Information Catalog and Live Data Map in version 10.1.1 Update 2.

#### Asset path

Effective in version 10.1.1 Update 2, you can view the path to the asset in the Asset Details view along with other general information about the asset.

For more information, see the *Informatica 10.1.1 Update 2 Enterprise Information Catalog User Guide*.

#### Business terms in the Profile Results section

Effective in version 10.1.1 Update 2, the profile results section for tabular assets also includes business terms. Previously, the profile results section included column names, data types, and data domains.

For more information, see the *Informatica 10.1.1 Update 2 Enterprise Information Catalog User Guide*.

#### URLs as attribute values

Effective in version 10.1.1 Update 2, if you had configured a custom attribute to allow you to enter URLs as the attribute value, you can assign multiple URLs as attribute values to a technical asset.

For more information, see the *Informatica 10.1.1 Update 2 Enterprise Information Catalog User Guide*.

#### Detection of CSV file headers

Effective in version 10.1.1 Update 2, you can configure the following resources to automatically detect headers for CSV files from which you extract metadata:

- Amazon S3
- HDFS
- File System

For more information, see the *Informatica 10.1.1 Update 2 Live Data Map Administrator Guide*.

#### Amazon Redshift resource

Effective in version 10.1.1 Update 2, you can import multiple schemas for an Amazon Redshift resource.

For more information, see the *Informatica 10.1.1 Update 2 Live Data Map Administrator Guide*.

#### **Profiling for Hive resource on Data Integration Service**

Effective in version 10.1.1 Update 2, you can run Hive resources on Data Integration Service for profiling.

For more information, see the *Informatica 10.1.1 Update 2 Live Data Map Administrator Guide*.

#### **PowerExchange for Amazon Redshift**

Effective in version 10.1.1 Update 2, you can select multiple schemas for Amazon Redshift objects. To select multiple schemas, leave the **Schema** field blank in the connection properties. In earlier releases, selecting schema was mandatory and you could select only one schema.

If you upgrade to version 10.1.1 Update 2, the PowerExchange for Redshift mappings created in earlier versions must have the relevant schema name in the connection property. Else, mappings fail when you run them on version 10.1.1 Update 2.

For more information, see the *Informatica 10.1.1 Update 2 PowerExchange for Amazon Redshift User Guide*.

## CHAPTER 7

# Fixed Limitations

This section describes fixed limitations in version 10.1.1 Update 2.

### Big Data Fixed Limitations

Bug	Description
BDM-7174	A mapping that reads from and writes to a Hive database hosted on HDFS fails with an error message about access to the default Hive warehouse directory.
BDM-5328	When you add a filter condition for an Oracle source that has RAPartitioning enabled, the Data Integration Service generates an incorrect WHERE clause and ignores the filter condition. This issue occurs when you run the mapping on the Hive engine.
BDM-5200	A mapping with a DATE_DIFF function returns results that are not valid when the mapping is run on the Hadoop cluster.
BDM-4960	Mappings that use Snappy compression fail on Hadoop clusters if the cluster uses Kerberos authentication.
BDM-4639	Data gets corrupted while reading from or writing to tab-delimited Hive Tables. Reader will read null for columns and writer will corrupt the delimiter.
BDM-4624	When Hive tables are delimited by certain Unicode characters, a mapping that runs with the Blaze engine reads data incorrectly.
BDM-4564	Mappings that use the Blaze engine hang if machines in the cluster are behind a firewall, and a misleading error message was sent.
OCON-7632	When you run a Sqoop mapping with unconnected target ports on the Hive engine, the mapping fails with the following error message: <code>Invalid column reference</code>



## PowerExchange for Amazon Redshift Fixed Limitations

Bug	Description
OCON-7858	If you enable the server-side encryption in the bucket policy when you run an Amazon Redshift mapping, the mapping fails with the following error message:  2017-02-14 21:39:01.761 <DTM-pool-2-thread-5> SEVERE: [LDTM_0072] Access Denied (Service: Amazon S3; Status Code: 403; Error Code: AccessDenied; Request ID: 86EE0E73C9D0B860) 2017-02-14 21:39:01.767 <DTM-pool-2-thread-5> WARNING: [LDTM_0072] Access Denied (Service: Amazon S3; Status Code: 403; Error Code: AccessDenied; Request ID: 86EE0E73C9D0B860)
OCON-7746	When you run an Amazon Redshift mapping to perform an update operation and do not map the primary key, sort key, or distribution key in the Amazon Redshift target, the mapping fails.
OCON-6304	If you specify the UNLOAD and COPY options through a property file and run a mapping on the Blaze engine, the mapping fails.

## PowerExchange for Amazon S3 Fixed Limitations

Bug	Description
OCON-7770	When you read data from an Amazon S3 source in the native environment, the null values in the Bigint and Decimal data type fields are read as zeroes. However, the mapping runs successfully without any error.
OCON-7766	When you write data to an Amazon S3 target in the native environment, the null values in the Bigint data type field are replaced by zeroes. However, the mapping runs successfully without any error.
OCON-7742	Performance is slow when you read data from or write data to the staging target in the native environment.
OCON-7678	When you run a mapping to read data from or write data to an Amazon S3 target and the bucket policy has HTTPS enabled, the mapping fails.
OCON-7677	When you run a mapping to write data to an Amazon S3 target in the native environment and the source table contains null values for the decimal fields, the mapping fails with the following error message:  "CSV parsing issue:null"

## PowerExchange for Hive Fixed Limitations

Bug	Description
BDM-5153	When you add or remove columns in an existing Hive physical data object and run the mapping on the Blaze engine, the mapping fails with the following error:  Plug-in #1010000's target [Write_ot_ver_screendi_protocol] failed in method [init].
OCON-7597	When you use a native DB2 connection and run a mapping on the Hive engine to write data to an IBM DB2 partitioned table that does not exist in the default schema of the user, the mapping fails.

### PowerExchange for Microsoft Azure Blob Storage Fixed Limitations

Bug	Description
OCON-7745	When you run a mapping to read data from or write data to Microsoft Azure BLOB Storage, the mapping takes a long time to execute and results in slow performance.

### PowerExchange for Netezza Fixed Limitations

Bug	Description
OCON-6977	When you select the engine type as <b>Tez</b> and run a mapping on the Hive engine to write data to a Netezza target, the mapping fails.

## CHAPTER 8

# Known Limitations

This section describes known limitations in version 10.1.1 Update 2.

### Big Data Known Limitations

Bug	Description
BDM-7591	Mappings that read from and write to Hive sources and targets on Amazon S3 in a Hortonworks 2.3 cluster fail.
BDM-7348	<p>When the Spark engine runs a mapping that contains a decimal data type on a Hortonworks version 2.3 or 2.4 cluster running on SUSE Linux, the mapping fails.</p> <p>The following error message appears:</p> <pre>java.math.BigDecimal is not a valid external type for schema of int</pre> <p>The issue occurs due to a mismatch between the metadata that Informatica imports and the data type in the Hive table.</p>
BDM-7347	<p>Mappings that use the Snappy compression mode fail.</p> <p>The following error message appears:</p> <pre>java.lang.UnsatisfiedLinkError: org.apache.hadoop.util.NativeCodeLoader.buildSupportsSnappy()Z ...</pre> <p>Workaround for the Blaze, native, and Hive run-time engines:</p> <ol style="list-style-type: none"><li>1. Copy Hadoop libraries from the cluster to the corresponding folders on the machine where the Data Integration Service is installed.</li><li>2. Edit <code>-&lt;Informatica installation directory&gt;/Informatica/services/shared/hadoop/&lt;Hadoop distribution name&gt;_&lt;version&gt;/InfaConf/hadoopEnv.properties</code> to add the following properties: <pre>spark.yarn.appMasterEnv.LD_LIBRARY_PATH=\$INFA_HADOOP_DIST_DIR/lib/native: \$LD_LIBRARY_PATH spark.executorEnv.LD_LIBRARY_PATH=\$INFA_HADOOP_DIST_DIR/lib/ native:\$LD_LIBRARY_PATH</pre></li></ol> <p>There is no workaround if you want to use the Spark run-time engine.</p>
BDM-7126	<p>Mappings that run on the Spark engine fail if you change the operating system profile user or the impersonation user.</p> <p>Workaround: To run the mappings on the Spark engine with a different user, complete any one of the following tasks:</p> <ul style="list-style-type: none"><li>- In the <code>hadoopEnv.properties</code> file, change the value of the <code>infa.osgi.enable.workdir.reuse</code> property to <code>false</code>.</li><li>- Before changing the user, set the value of the <code>infa.osgi.parent.workdir</code> property in the <code>hadoopEnv.properties</code> file to a different working directory.</li></ul>

Bug	Description
BDM-6997	<p>Blaze mappings fail at run time with a validation error if the data types of the Hive table do not match the data types in the physical data object that is used as a Source or Target transformation in the mapping.</p> <p>Workaround: Edit the physical data object to use the same data types as the Hive table and run the mapping again.</p>
BDM-6840	<p>A mapping executed with the Blaze engine writes an inaccurate row count to the target. The row count includes rejected rows.</p>
BDM-6694	<p>When the Blaze engine reads from a compressed Hive table with text format, the mapping fails if the TBLPROPERTIES clause is not set for the Hive table.</p> <p>Workaround: Create or alter the table with the TBLPROPERTIES clause. For example, <code>TBLPROPERTIES ('text.compression'='Snappy')</code>.</p>
BDM-6389	<p>A mapping fails to add statistics to Hive table metadata after loading data to the table on Hortonworks.</p> <p>Workaround: To view statistics for a table, run the following command on the HIVE command line: <code>ANALYZE TABLE &lt;table name&gt; COMPUTE STATISTICS;</code></p>
BDM-6598	<p>When the Blaze engine runs a mapping on an Amazon EMR 5.0 cluster, the Blaze engine does not use the following properties set in yarn-site.xml on the Data Integration Service machine host:</p> <ul style="list-style-type: none"> <li>- fs.s3n.endpoint</li> <li>- fs.s3.awsAccessKeyId</li> <li>- fs.s3.awsSecretAccessKey</li> </ul> <p>The Blaze engine uses the values for these properties from the yarn-site.xml file on the cluster.</p>
BDM-5465	<p>Mappings that read from or write to partitioned or bucketed Hive sources and targets on Amazon S3 take longer to execute than expected.</p>
BDM-5322	<p>Mappings that use the Blaze engine to run a mapping on a Kerberos-enabled IBM BigInsights cluster fail.</p> <p>The following error message appears:</p> <pre>GSS initiate failed [Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)]</pre> <p>Workaround:</p> <ol style="list-style-type: none"> <li>1. Configure the property <code>hadoop.security.token.service.use_ip</code> with a value of <code>FALSE</code> in the Ambari cluster configuration browser.</li> <li>2. Add the following lines to the file <code>&lt;Informatica_installation_home&gt;/services/shared/hadoop/biginsights_&lt;version&gt;/conf/core-site.xml</code>: <pre>&lt;property&gt; &lt;name&gt;hadoop.security.token.service.use_ip&lt;/name&gt; &lt;value&gt;FALSE&lt;/value&gt; &lt;/property&gt;</pre> </li> </ol>
BDM-4276	<p>Using the Spark engine to run a mapping with a Hive source fails when the mappings includes a SQL override with <code>DISTINCT</code> and <code>LIMIT</code> clauses.</p>
OCON-8353	<p>When a mapping reads from or writes to an HDFS on a secure cluster, the mapping log incorrectly states that the Data Integration Service used the HDFS connection user to connect to HDFS. In fact, the Data Integration Service uses the following users, in order of preference, to connect to HDFS:</p> <ol style="list-style-type: none"> <li>1. Operating system profile user, if configured.</li> <li>2. Impersonation user, if configured.</li> <li>3. HDFS connection user, if configured.</li> <li>4. Data Integration Service user.</li> </ol> <p>The mapping log is correct when you read from or write to HDFS on a non-secure cluster.</p>

Bug	Description
OCON-8107	When you read data through Sqoop and the source contains numeric data types with a precision greater than 15, the target data is truncated after 15 digits. This issue occurs when you run the mapping on the Spark engine.
OCON-8036	<p>A mapping that reads data from Netezza fails on the Blaze engine when all of the following conditions are true:</p> <ul style="list-style-type: none"> <li>- You run the mapping on one of the following clusters: <ul style="list-style-type: none"> <li>- Hortonworks HDP 2.5 clusters</li> <li>- Hortonworks HDP 2.3 clusters</li> <li>- MapR Ticket clusters</li> <li>- MapR Kerberos clusters</li> </ul> </li> <li>- You enable SSL and Kerberos authentication for the cluster, or enable only Kerberos authentication for the cluster.</li> </ul>
OCON-7687	When you export data through Sqoop and the columns contain mixed case characters, the mapping fails.
OCON-7669	<p>When you configure Sqoop and OraOop, and export data to an Oracle target that contains mixed case characters in the table name, the mapping fails.</p> <p>Workaround: Use the generic Oracle JDBC driver to export data.</p>
OCON-7620	When you import data from an IBM DB2 source through Sqoop and the table name contains mixed case characters, the mapping fails.
OCON-7521	<p>Column profile run fails when the following conditions are true:</p> <ol style="list-style-type: none"> <li>1. You use the Cloudera Connector Powered by Teradata for Sqoop to read data from Teradata for a Sqoop data source.</li> <li>2. You create a column profile on the Sqoop data source.</li> <li>3. You run the profile using the Blaze engine.</li> </ol> <p>Workaround: Create a separate Sqoop connection for each Sqoop data source with the <b>-split-by</b> option, and run a column profile on the data source.</p>
OCON-7459	<p>When you export data to an IBM DB2 target through Sqoop, the mapping fails if all of the following conditions are true:</p> <ul style="list-style-type: none"> <li>- You create or replace the IBM DB2 target table at run time.</li> <li>- The IBM DB2 target table name or column names contain mixed case characters.</li> <li>- You run the mapping on a Cloudera 5u8 cluster.</li> </ul>
OCON-7431	When you read time data from a Teradata source and write it to a Teradata target, the fractional seconds get corrupted. This issue occurs if you run the Teradata Parallel Transporter mapping on a Hortonworks cluster and on the Blaze engine.
OCON-7429	<p>When you run a Teradata Parallel Transporter mapping on a Hortonworks cluster and on the Blaze engine to write byte and varbyte data to a Teradata target, the data gets corrupted. This issue occurs when you use the <code>hdp-connector-for-teradata-1.5.1.2.5.0.0-1245-distro.tar.gz</code> JAR.</p> <p>Workaround: Use the <code>hdp-connector-for-teradata-1.4.1.2.3.2.0-2950-distro.tar.gz</code> JAR.</p>
OCON-7365	<p>Sqoop mappings fail on MapR 5.2 clusters.</p> <p>Workaround: Add the following property in the <code>mapred-site.xml</code> file on all nodes of the cluster, and restart the Hadoop services and cluster:</p> <pre>&lt;property&gt; &lt;name&gt;mapreduce.jobhistory.address&lt;/name&gt; &lt;value&gt;&lt;Host_Name&gt;:10020&lt;/value&gt; &lt;/property&gt;</pre>

Bug	Description
OCON-7291	<p>Mappings that read data from a Teradata source and contain the != (not equal) operator in the filter override query fail. This issue occurs if you run the Teradata Parallel Transporter mapping on a Hortonworks cluster and on the Blaze engine.</p> <p>Workaround: Use a native expression with the ne operator instead of the != operator.</p>
OCON-7073	<p>When you run a Sqoop mapping on a Cloudera cluster that uses Kerberos authentication, you must manually configure mapreduce properties in the <code>yarn-site.xml</code> file on the Data Integration Service node and restart the Data Integration Service. To run the mapping on the Blaze engine, you must restart the Grid Manager and Blaze Job Monitor.</p>

## Intelligent Data Lake Known Limitations

Bug	Description
IDL-2341	<p>The Intelligent Data Lake Service folder is not cleaned up after installation.</p> <p>Workaround: The administrator needs to clean up the Intelligent Data Lake service folder and its sub-folders in the following location: - - <code>Tomcat/temp/idlservice</code>.</p>
IDL-2336	<p>In a non-secure domain, the Data Preparation Service writes data to durable storage using the system user instead of the user details provided in the HDFS connection.</p> <p>Workaround: Grant the required permissions to the system user to the HDFS location used in the durable storage.</p>
IDL-2333	<p>The Intelligent Data Lake Service tests the Hive connection as an anonymous user instead of with the user details specified in the Hive connection.</p> <p>Workaround: To test the Hive connection, grant the required permissions for the anonymous user in the cluster.</p>

## PowerExchange for Amazon Redshift Known Limitations

Bug	Description
OCON-8301	<p>When the Blaze engine runs a mapping that joins Amazon Redshift sources that use different delimiters, HDFS rejects all target rows.</p>
OCON-8022	<p>If you import an Amazon Redshift table that has a single quote (') in the column name, the mapping fails with the following error message:</p> <pre>[LDTM_0072] [Amazon] (500051) ERROR processing query/statement. Error: Parsing failed, Query: unload ('SELECT "adpqa"."sq_col"."id" FROM "adpqa"."sq_col") TO 's3://infa.qa.bucket/0b0ad503-1c2c-4514-95ac-85a5adb71b3b1489385038407/sq_col' credentials 'aws_access_key_id=*****;aws_secret_access_key=*****' ESCAPE DELIMITER ','</pre>
OCON-8020	<p>If you import an Amazon Redshift table that has a question mark (?) in the column name, the read and write operations fail with the following error message:</p> <pre>[LDTM_0072] [Amazon] (500310) Invalid operation: syntax error at or near "\".\"",</pre>

Bug	Description
OCON-7965	When you run an Amazon Redshift mapping on the Blaze engine to read data from or write data to an Amazon Redshift cluster that requires Version 4 authentication, the mapping fails. This issue occurs if you use the Hortonworks 2.3 distribution.
OCON-7909	When you run an Amazon Redshift mapping on the Blaze engine to read data from or write data to an Amazon Redshift cluster that requires Version 4 authentication, the mapping fails. This issue occurs if you use the MapR 5.2 distribution.
OCON-7322	If you import an Amazon Redshift table that has a single quote (') or a backslash (\) in the table name, the read and write operations fail.

### PowerExchange for Amazon S3 Known Limitations

Bug	Description
OCON-7963	When you run an Amazon S3 mapping on the Blaze engine to read data from or write data to an Amazon S3 bucket that requires Version 4 authentication, the mapping fails. This issue occurs if you use the Hortonworks 2.3 distribution.
OCON-7938	<p>When you run an Amazon S3 mapping in the native environment, the incorrect values in the decimal or bigint data type field are replaced by zeroes. However, the mapping runs successfully and the session log displays the following error message:</p> <pre>2017-03-13 23:46:08.131 &lt;TASK_140116755179264-READER_1_1_1&gt; SEVERE: [APPSDK_Msg_1762] Data for column [age] of type [bigint] should be a of type [java.lang.Number] or its sub-types. 2017-03-13 23:46:08.132 &lt;TASK_140116755179264-READER_1_1_1&gt; SEVERE: [APPSDK_Msg_1762] Row containing invalid data = [abc, abc, abc, abc]</pre>
OCON-7930	If you set a code page other than UTF-8 encoding for the flat file parser, PowerExchange for Amazon S3 still uses UTF-8 encoding to read or write data and the mapping runs successfully.
OCON-7911	When you run an Amazon S3 mapping on the Blaze engine to read data from or write data to an Amazon S3 bucket that requires Version 4 authentication, the mapping fails. This issue occurs if you use the MapR 5.2 distribution.
OCON-7763	When you write data to an Amazon S3 target in the native environment, the null values in the String data type field are replaced by double quote characters (" "). However, the mapping runs successfully without any error.

### PowerExchange for Hive Known Limitations

Bug	Description
OCON-8015	<p>The Data Integration Service converts Hive 1.1.0 float data to double with incorrect values.</p> <p><b>Note:</b> The Data Integration Service correctly converts the float data type in Hive 1.2.1.</p>

## PowerExchange for MapR-DB Known Limitations

Bug	Description
OCON-8092	If you specify a filter condition with a negative value for numeric data types, the Data Integration Service does not apply the filter condition when it reads data from a MapR-DB table.
OCON-7815	When you run a mapping on the Hive or Spark engine to read data from or write data to MapR-DB, the Data Integration Service does not validate the mapping.
OCON-7599	When you use PowerExchange for MapR-DB to import MapR-DB JSON tables, the import operation fails.

## PowerCenter Known Limitations

Bug	Description
PLAT-16265	<p>When you run the <code>infacmd ipc genReuseReportFromPC</code> command for a large repository, the request hangs.</p> <p>Workaround: Execute the report against partial set of folders instead of the entire repository.</p>



## CHAPTER 9

# Third-Party Fixed Limitations

This section describes third-party fixed limitations in version 10.1.1 Update 2.

### PowerExchange for Amazon Redshift Third-Party Fixed Limitations

Bug	Description
OCON-7760	<p>When you run an Amazon Redshift mapping on the Blaze engine using a MapR 5.2 secure cluster, the mapping fails with the following error message:</p> <p>Unable to execute HTTP request: Peer not authenticated.</p> <p>This issue occurs if you do not run the following command on all the nodes of the MapR cluster:</p> <pre>keytool -importkeystore -srckeystore &lt;JDK_HOME&gt;/jre/lib/security/cacerts - destkeystore &lt;mapR_Conf&gt;/ssl_truststore -deststorepass mapr123 -v (00045573)</pre>

### PowerExchange for Amazon S3 Third-Party Fixed Limitations

Bug	Description
OCON-7759	<p>When you run an Amazon S3 mapping on the Blaze engine using a MapR 5.2 secure cluster, the mapping fails with the following error message:</p> <p>Unable to execute HTTP request: Peer not authenticated.</p> <p>This issue occurs if you do not run the following command on all the nodes of the MapR cluster:</p> <pre>keytool -importkeystore -srckeystore &lt;JDK_HOME&gt;/jre/lib/security/cacerts - destkeystore &lt;mapR_Conf&gt;/ssl_truststore -deststorepass mapr123 -v (00045573)</pre>

## CHAPTER 10

# Third-Party Known Limitations

This section describes third-party known limitations in version 10.1.1 Update 2.

### Big Data Third-Party Known Limitations

Bug	Description
BDM-7339	<p>In a Hortonworks HDP 2.3 or 2.4 environment, when the Data Integration Service is idle for more than 10 hours, mappings fail if the Hadoop cluster uses Hadoop Key Management Server.</p> <p>Workaround: Recycle the Data Integration Service.</p> <p>Hadoop issue numbers: HADOOP-13155, HADOOP-13251, and HADOOP-13255.</p>
BDM-7321	<p>When the Blaze engine runs a mapping on a Hortonworks cluster, and use an s3a URI to access Hive tables on Amazon S3 as sources and targets, the following error message appears :</p> <pre>WARNING: [LDTM_3001] Partitioning is disabled for the mapping. Exception Class: [java.lang.RuntimeException] Exception Message: [java.nio.file.AccessDeniedException: /HWS3a_testing: getFileStatus on / HWS3a_testing: com.amazonaws.services.s3.model.AmazonS3Exception: Forbidden (Service: Amazon S3; Status Code: 403; Error Code: 403 Forbidden; ...</pre> <p>Workaround: To use the s3a URI to access Hive tables, copy the hive-exec .jar file from the cluster to the directory defined in the Hadoop Distribution Directory property for the Data Integration Service.</p> <p><b>Note:</b> The hive-exec .jar file enables use of the s3a URI to access Hive tables, but does not enable other Hive-related operations. To fully read from and write to Hive tables on a Hortonworks cluster, use an s3n URI instead.</p>
BDM-7283	<p>You cannot delete the INFORMATICA BDM Service in the Ambari Configuration Manager version 2.2.</p> <p>Workaround: To delete the INFORMATICA BDM Service, run the following command on the cluster name node:</p> <pre>curl -u admin:admin -H "X-Requested-By: ambari" -X DELETE http://&lt;hostname&gt;: 8080/api/v1/clusters/&lt;CLUSTERNAME&gt;/services/INFORMATICA BDM</pre>
BDM-6813	<p>When you use an S3a URI to access Hive source and target tables on a Hortonworks cluster on Amazon S3, mappings might fail due to a table creation issue. The following error message appears:</p> <pre>ERROR [main]: exec.DDLTask (DDLTask.java:failed(533)) - org.apache.hadoop.hive.ql.metadata.HiveException: MetaException(message:java.nio.file.AccessDeniedException: s3a://&lt;Bucket_Name&gt;/ &lt;Table_name&gt;: getFileStatus on s3a://&lt;Bucket_Name&gt;/&lt;Table_name&gt;: com.amazonaws.services.s3.model.AmazonS3Exception: Forbidden (Service: Amazon S3; Status Code: 403; Error Code: 403 Forbidden; Request ID: ....), S3 Extended Request ID....)?</pre> <p>Workaround: Follow the steps in <a href="#">Knowledge Base article 509701</a> to patch the Hortonworks instance.</p> <p>Hortonworks issue number HADOOP-3733.</p>

Bug	Description
BDM-6812	<p>When you use an S3a URI to access Hive source and target tables on an Amazon S3 MapR cluster, mappings might fail due to a table creation issue. The following error message appears:</p> <pre>hive&gt; create table &lt;table_name&gt; (c1 int) location 's3a://&lt;Bucketname&gt;/&lt;table_name&gt;'; FAILED: Execution Error, return code 1 from org.apache.hadoop.hive.ql.exec.DDLTask. MetaException(message:Got exception: javax.net.ssl.SSLPeerUnverifiedException peer not authenticated)</pre> <p>Workaround: Follow the steps in <a href="#">Knowledge Base article 509702</a> to import security certificates to the SSL truststore.</p> <p>MapR case number 00046563.</p>
BDM-5431	<p>When the Blaze engine runs mappings as the impersonation user on a MapR cluster, the following error message appears:</p> <pre>[org.apache.hadoop.security.AccessControlException: User impUser3(user id 2002) does not have access</pre> <p>Workaround: Disable user impersonation. In the Hadoop connection properties, in the Common Properties area, set the impersonation user to the user ID of the Data Integration Service user.</p> <p>MapR case number 00045006.</p>
BDM-5022	<p>When you run a mapping with the Spark engine on a MapR cluster, the execution engine uses the default MapR staging directory and not the staging directory that you configured in the "SPARK HDFS Staging Directory" property in the Hadoop connection.</p> <p>For example, the engine uses the directory /user/&lt;Data Integration Service username&gt;/. If you configured the impersonation user, the engine uses the directory /user/&lt;impersonation user&gt;. If you configured the operating system profile (OSP) user, the engine uses the directory /user/&lt;OSP user&gt;.</p> <p>MapR case number 00045736.</p>
BDM-1363	<p>Big Data Management on the MapR Hadoop distribution does not support the use of operating system profiles when you run mappings using the Blaze engine.</p> <p>MapR case number 00045006.</p>
OCON-7974	<p>Sqoop mappings fail when a column name contains spaces.</p> <p>Apache ticket reference number: SQOOP-2737</p>
OCON-7505	<p>Sqoop mappings that read byte or varbyte data from a Teradata source and write it to a Teradata target fail on the Blaze engine. This issue occurs if you use Cloudera Connector Powered by Teradata.</p> <p>Cloudera ticket reference number: 124305</p>
OCON-7504	<p>When you use Sqoop to read time data from a Teradata source and write it to a Teradata target, only milliseconds are written to the target. This issue occurs if you run the Teradata Parallel Transporter mapping on a Cloudera cluster and on the Blaze engine.</p> <p>Cloudera ticket reference number: 124302</p>
OCON-7503	<p>When you use Sqoop to read time data from a Teradata source and write it to a Teradata target, the fractional seconds get corrupted. This issue occurs if you use Cloudera Connector Powered by Teradata or Hortonworks Connector for Teradata, and run the mapping on the Blaze engine.</p> <p>Cloudera ticket reference number: 124306</p>

## CHAPTER 11

# Informatica Global Customer Support

You can contact a Global Support Center by telephone or through Online Support on Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

If you are an Informatica Network member, you can use Online Support at <http://network.informatica.com>.