

## Readme for Data Sync

This Read Me describes changes, updates, and upgrade instructions for Data Sync Version 2.3.

### Topics:

- [Installing Data Sync Version 2.3](#)
- [New Features In Data Sync Version 2.3](#)
- [Known Issues in Data Sync Version 2.3](#)
- [Features Available in Previous Versions](#)

## Installing Data Sync Version 2.3

You install Data Sync differently depending on whether you have an existing Data Sync installation.

### Installing Data Sync For The First Time

Install Data Sync into a new directory using the instructions in the section “Installing and Setting Up Data Sync” in the *Getting Started* guide.

### Upgrading An Existing Data Sync Installation

If you have an earlier version of Data Sync installed, you can upgrade it to the latest version and retain the repository, password, and JDBC libraries that you already use.

1. Install Data Sync into a new directory using the instructions in the section “Download and Install Data Sync” in the *Getting Started* guide.

2. Configure your Java SDK home in the config.bat or config.sh file.

For example, if your Java SDK is installed in C:\Java, configure JAVA\_HOME as:  
set JAVA\_HOME=C:\Java.

3. Copy any JDBC libraries you installed previously for Data Sync to the new Data Sync \lib directory.

By default, Data Sync installs the Oracle JDBC driver only.

If you need to connect to a different database type (for example, Microsoft SQL Server, or DB2) or if you want to use a different Oracle driver from the default, then you must manually copy required files to the lib directory.

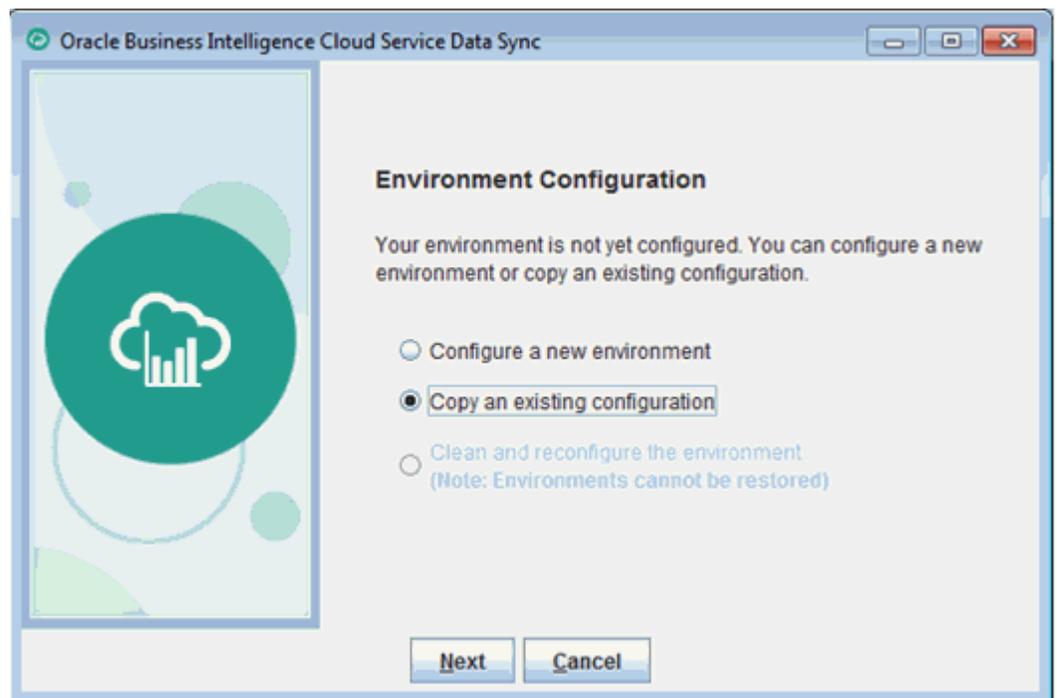
4. When you start Data Sync for the first time, at the Environment Configuration screen, select Copy existing configuration, and specify the location of the previous installation of Data Sync.

The installer copies the repository from the previous installation and prompts you to provide the password originally used for the previous installation.

---

**Note:** If you are loading to Oracle Database as a Service rather than the default Schema As A Service, then you must use the 'Oracle (Thin)' connection type in Data Sync. If you were using the Oracle (BICS) connection type to load to Oracle Database as a Service before upgrading Data Sync, then you must change the connection type to 'Oracle (Thin)'.  

---



We recommend that you change the name of the repository to differentiate your current repository from the new repository.

After you have installed and tested the new version of Data Sync, uninstall the old environment by deleting the old installation folder.

---

**Note:** Always use the **Copy an existing configuration** option to migrate your existing projects and system configuration to a newer Data Sync version. You can't export your existing repository and import it into the newer version.  

---

## New Features in Data Sync Version 2.3

- Transform data in DBaaS or direct database targets, For example, apply default values, calculations, conversions, concatenations, splits, new target data columns, SQL commands.
- Perform pre-load processing.
- Loading data from files:
  - Optionally delete data files after a successful data load.
  - Track where data originates source file metadata (that is, using FILE\_NAME and LINE\_NUMBER).
- Loading data from OTBI:
  - Day Based Partitioned Read from Subject Area.Table.
  - Day Based Partitioned Read from SQL.
- Loading data from Oracle Service Cloud (RightNow):
  - Load data by Named ID.

## Known Issues in Data Sync Version 2.3

- When using Data Sync with Oracle Analytics Cloud:
  - You can't create data sets.
  - You can purge the cache at the end of a job, however the test connection dialog will display the following message:  

```
This connection points to Oracle Analytics Cloud.  
You will not be able to load to it, but can still use it  
to clear BI cache after writing to DBaaS.
```
  - When you click Help, you view the Help topic in the Oracle BI Cloud Service Help Center.

## Features Added in Previous Versions

The following features have been released previously.

## Features Added To Data Sync Version 2.2

- Support for new data sources:
  - Generic JDBC with prepackaged drivers for Greenplum, MongoDB, Salesforce, Redshift, Hive and PostgreSQL
  - Other sources that support JDBC
  - Oracle Financials Cloud
  - Oracle HCM Cloud

- Oracle Procurement Cloud
- Oracle Project Management Cloud
- Oracle Sales Cloud
- Oracle Supply Chain Management Cloud
- Oracle Service Cloud (RightNow)
- Connection enhancements:
  - Time zone - Every time a job successfully completes, all the tables updated by the job are stamped with a timestamp. This timestamp is used as the last refresh date for incremental loads. For known databases (Oracle, SQL Server, DB2, Teradata and Timesten), Data Sync finds the delta timestamp between the local timestamp on the machine where the tool is running and the database's current timestamp and applies the delta before saving the refresh timestamp. For example, a source database is on time zone EST, and Data Sync is running on a machine with time zone PST. The job starts on January 1st 2016 at 10:00 AM (PST). At the end of the run, the refresh dates are persisted as January 1st 2016 1:00 PM (adjusted to EST). If Data Sync cannot determine the delta, you can specify the time zone of the source. When specified, the same logic as above is used to persist the refresh timestamps, with the difference of the time delta calculated based on user specified time zone.
  - Source priority - If more than one data source is to write to the same target table within the same job, Data Sync determines the order of extraction. However, you can override the order by specifying a priority. A priority of 1 takes precedence over a priority of 2.
- Job enhancements:
  - Parallelism:
    - \* For connections to Oracle BI Cloud Service, Data Sync loads a maximum of four tables in parallel. If you attempt to load more than four, an error might be reported if a tenant is overloaded.
    - \* For connections to Database as a Service or on-premises databases, Data Sync defaults to 10 maximum, but you can specify how many tasks can run in parallel. The default is 10. This means that up to 10 replication tasks or 10 post-load processing tasks can run in parallel.
    - \* You specify the number of parallel tasks that you want to allow using the # Parallel Tasks value on the Jobs - Edit tab. When you increase the number of tasks that can run in parallel, the overall throughput will decrease and more resources will be consumed (CPU, memory etc).
  - Purge cache - On Oracle BI Cloud Service connections, data cached on the target Oracle BI Cloud Service purges automatically at the end of every job. When you connect to Database as a Service, you can select which Oracle BI Cloud Service you want to purge from the connection list.

- Pruned extracts - For all tables and objects that are incrementally extracted, the last refresh timestamp is used to filter the data to identify the newly created or modified data. There may be situations when a source system might have transactions that were initiated just before the job was started, and are committed after the job is completed. These records can be missed in the subsequent run as the transaction dates remain as the time before the job started. In this scenario, you can set the refresh timestamp back to a certain period while querying. For example, if the last successful run was at January 1st 2016 10:00 AM, for the subsequent run, the last refresh timestamp will be the timestamp of the last successful run. When the prune time is set to 60 minutes, the table/object will be queried with Jan 1st 2016 9:00 AM as the timestamp to figure out the delta records.
- Execution parameters - you can override a project level parameter using an job level execution parameter with the same name. For example, your project loads five years of data by default, but for a test environment you might want to load one year of data only. In this scenario you have a parameter named NUMBER\_YEARS\_TO\_EXTRACT defined with the value 5 on the Project \Parameter tab. For the default job, TARGET is pointing to a production environment. You create a new job in the same project, and on the Jobs\Job \Execution Parameters tab, you create an execution parameter with the same name NUMBER\_YEARS\_TO\_EXTRACT and set its value to 1.
- Data flow enhancement: Table Size - Use this attribute to characterize source data tables as Extra Large, Large, Medium or Small. During an initial full load, Data Sync uses this information to maximize performance by extracting from larger tables before smaller tables.
- Parameters - Customize data loads dynamically at runtime using parameters.
- The system property named Target Date Time Offset is deprecated. If you are loading data incrementally in a time zone other than UTC by setting Target Date Time Offset, then after you install Data Sync Version 2.2 we recommend that you initialize the full load, by choosing Tools > Re-extract And Load All Data.

## Features Added To Data Sync Version 2.0

- The Oracle (BICS) connection type only loads data into Database Cloud Service Schema Service. To load data into Data Cloud Service Database as a Service (DBaaS), use the connection type Oracle (Thin). If you are currently using the Oracle (BICS) connection to load data into DBaaS, then this will not work after the upgrade. You must change the connection type to Oracle (Thin). This is because the REST API that Data Sync uses will always point to Schema As A Service provisioned with the Oracle BI Cloud Service instance after the Oracle BI Cloud Service upgrade. Therefore, if you were using Oracle (BICS) connection to load a DBaaS instance, then you should make Data Sync connect directly to that database.
- To adjust the memory allocation for the process running a job, you edit the file jvm\_properties.txt in the conf-shared folder.
- Oracle BI data sources. These sources can either be on-premises BI Sources or BI endpoints of Oracle Fusion OTBI. To access either of these, use the connection type Oracle BI Connector under the data sources. Use the Pluggable Sources view

to define your dataflows. The Data from Objects tab enables you to choose a particular Subject Area Table that you would like to replicate. The Manual Entry tab enables you to define one of the following as a data source:

- SQL – Use the logical SQL as the source. You provide for the initial SQL and incremental SQL (recommended). You can get the SQL from BI, from designing a new analysis, or by opening a report and using the advanced tab.
  - Report – Use the report path as the source. You provide for the absolute path to the report (you can obtain this path by decoding the URL used to open the report).
  - Subject Area Table - You provide for the name of the Subject Area Table or Folder. You can specify an additional filter in the format of the BI Logical SQL.
- Upload data to a data set suitable for Visual Analyzer projects (Oracle BI Cloud Service or Data Visualization Cloud Service). Whenever you define a data flow, you change the target type to Data Set, which loads the data in data set format.
  - During data loads, tables are automatically created or altered for direct connection targets (could be a DBaaS or on-premises database). This behavior is in contrast to previous releases where the tables were created/changed automatically for 'Oracle (BICS)' type connections. If you do not want this behavior, then you can turn it off by editing the system property named Automatically Create/Alter Tables (View->System Properties) and set it to false. The default is true.
  - Upload data from Excel spreadsheets and OTBI (Oracle Transactional Business Intelligence) data sources.
  - Data Sync will create and alter target tables based on source definitions for DBaaS and other relational targets.
  - You can import data from CSV files in a number of encoding formats or character sets, for example, ISO-8859-1, US-ASCII, UTF-8, and so on. You specify the encoding type using the Codepage option on the 5 Import Options dialog. For more information, refer to: <https://docs.oracle.com/javase/7/docs/api/java/nio/charset/Charset.html>.
  - Data Loader is deprecated.

## Features Added To Data Sync Version 1.2

- The system columns added in Data Sync Version 1.1 have been optimized to use less space and perform better in the database.
- By default, the data cache for data model objects in Oracle BI Cloud Service is deleted at the end of every completed job run. Set the Delete Data Cache property to false if you do not want the cache deleted. For the cache purge, you need to grant the BI Data Modeler role to the user who is registered to upload data.

## Features Added To Data Sync Version 1.1

- Data Sync automatically retries 10 times to successfully upload data in case of network errors. Three system columns for every target table are used to support

this restart capability. When a failed job is restarted, for those tables where the load strategy is "Insert only" or "Append data", any partially loaded data since the last successful batch is cleaned up and new data is reloaded. This prevents duplicate data in the event of failure restarts.

- For file-based sources, any number that cannot be accommodated by a NUMBER datatype is created with BINARY\_DOUBLE data type.
- Data Sync supports additional delimiters in comma-separated value file sources, including pipe, and custom single character delimiters.
- Data Sync supports table and column names with mixed cases and Oracle reserved words in their names.

---

Oracle® Cloud Data Sync Readme, Release 2.3  
DSRM-230

Copyright © 2015, 2017, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

This document describes changes, updates, and installation instructions for Oracle Business Intelligence Cloud Data Sync Version 2.2.