Syniti

Syniti Replicate

Amazon S3 Setup Guide
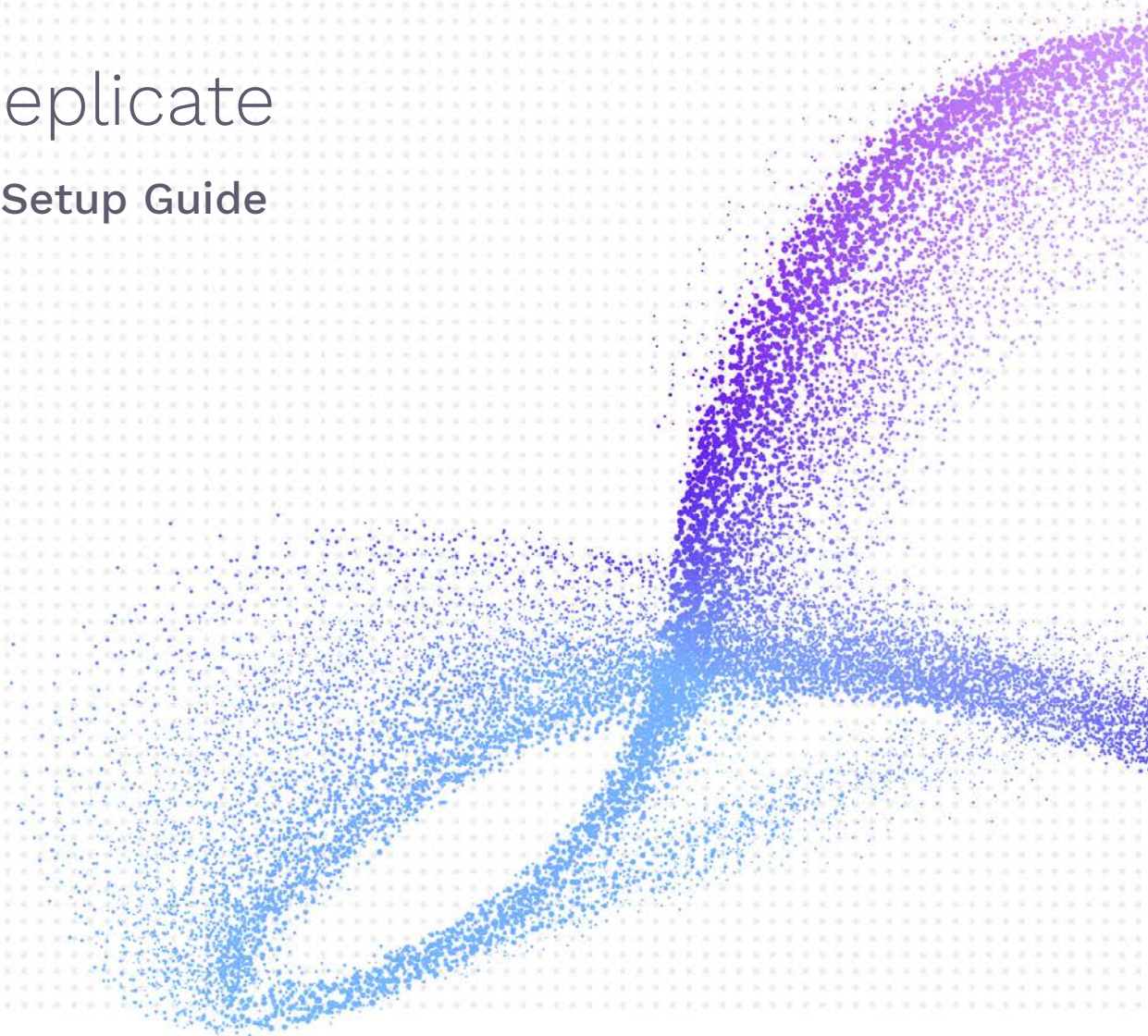
Version 10.3

Syniti

Syniti Replicate

Table of Contents

Syniti Replicate

This guide describes how to set up your Syniti Replicate environment to replicate data to an Amazon S3 bucket. The replicated data is stored in CSV or JSON format in S3. When replicating **from a relational database to Amazon S3**, Syniti Replicate allows you to replicate data in two different ways, or modes:

- **Refresh**
  A one-time complete replication from source table to S3, according to replication settings and scripts. You can control the timing of the replication, identify the columns to be replicated and add scripts to transform data during replication.

- **One-way mirroring**
  A continuous update of a replicated table to S3 based on changes to the source database that have been recorded in the database server log. You can define the replication settings to check the transaction log on the source database at regular intervals. Any changes found in the log would be applied to the target.

Requirements for Amazon S3 Targets

- Install the AWS Toolkit for .NET from:
  https://aws.amazon.com/sdk-for-net/

- Syniti Replicate will later need the pathnames to two of the DLLs in the SDK. Make sure you have the path available when you configure the target connection to Amazon S3.

- Set up your AWS account with either an IAM role or an Access Key and store the access information (either IAM role details or both the Access Key and Secret Access Key) where you will be able to find them when configuring your target connection to Amazon S3 below. The information is needed to access an S3 bucket.

Replicating Data from a Relational Database to Amazon S3

- Make sure you have database connections via a .NET data provider to the supported databases for source database tables.

- Start Syniti Replicate Management Center.
  Syniti Replicate provides a default database (Microsoft SQL Server CE) for your metadata, all the information that Syniti Replicate needs to store about your replication setup.
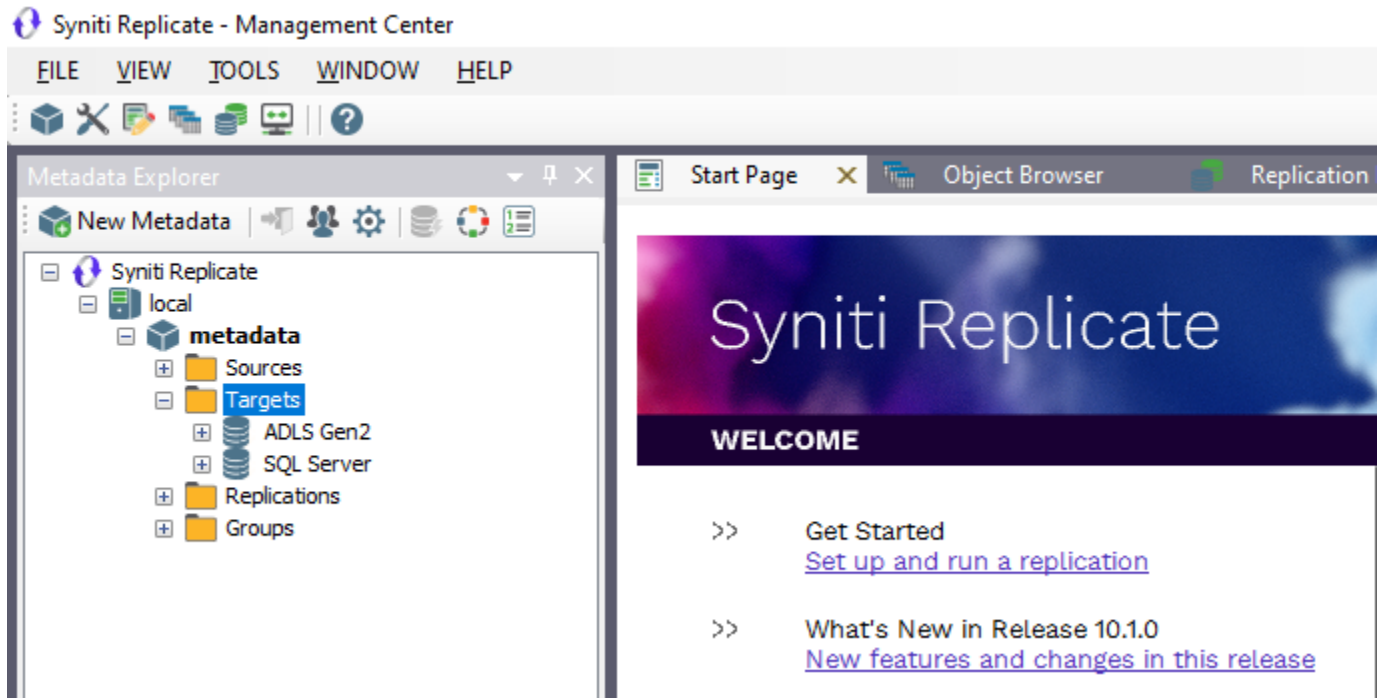
Create a Source Connection

1. In the Metadata Explorer, expand the metadata node to view the **Sources** and **Targets** nodes.

2. Select the **Sources** node.

3. From the right mouse button menu, choose **Add New Connection**.
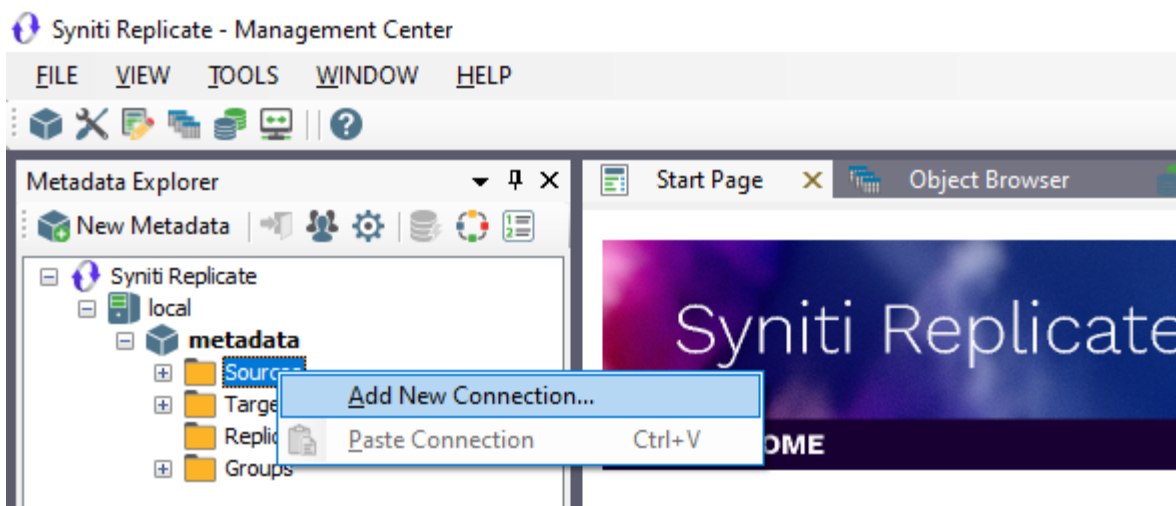
Syniti Replicate

4. In the Add Source Connection Wizard, follow steps to add a connection string and test the connection to the database.
   Check the Supported Provider List in the Help Center before entering a value in the **Assembly** field.

5. In the **Select Tables** screen, choose the tables that you plan to replicate.

6. In the **Actions** screen, check the option **Launch the Enable Transactional Setup Wizard**.

Syniti Replicate
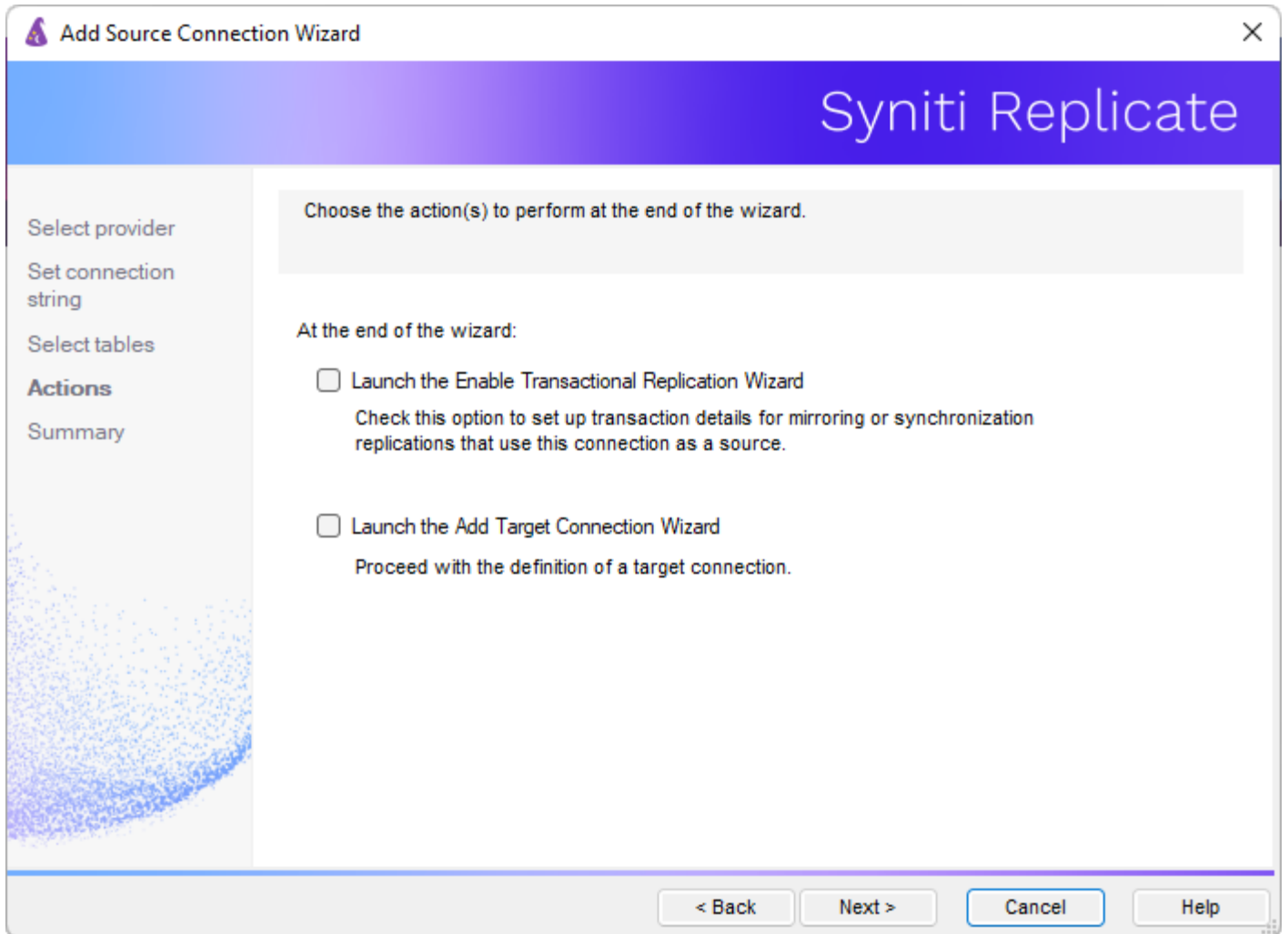
7. Complete the wizard.

Enable Transactional Replication

This section assumes you are planning to replicate data to an Amazon S3 CSV or JSON target using mirroring (rather than Refresh) from a relational database. For Refresh replications, you do not need to complete the Enable Transactional Replication wizard. If you have checked the Source Connection wizard option to launch the Enable Transactional Replication wizard, follow the steps below. To open the wizard from the Management Center, choose the connection in the Metadata Explorer, then right-click to choose Transactional Setup > Enable…

In the **Enable Transactional Replication wizard**:

1. Select the type of transactional replication to use. The options depend on the source database and can include log reader, log server agent, triggers, plus log reader API (for IBM Db2 for i only). Check the help for more information on the best option.

Syniti Replicate
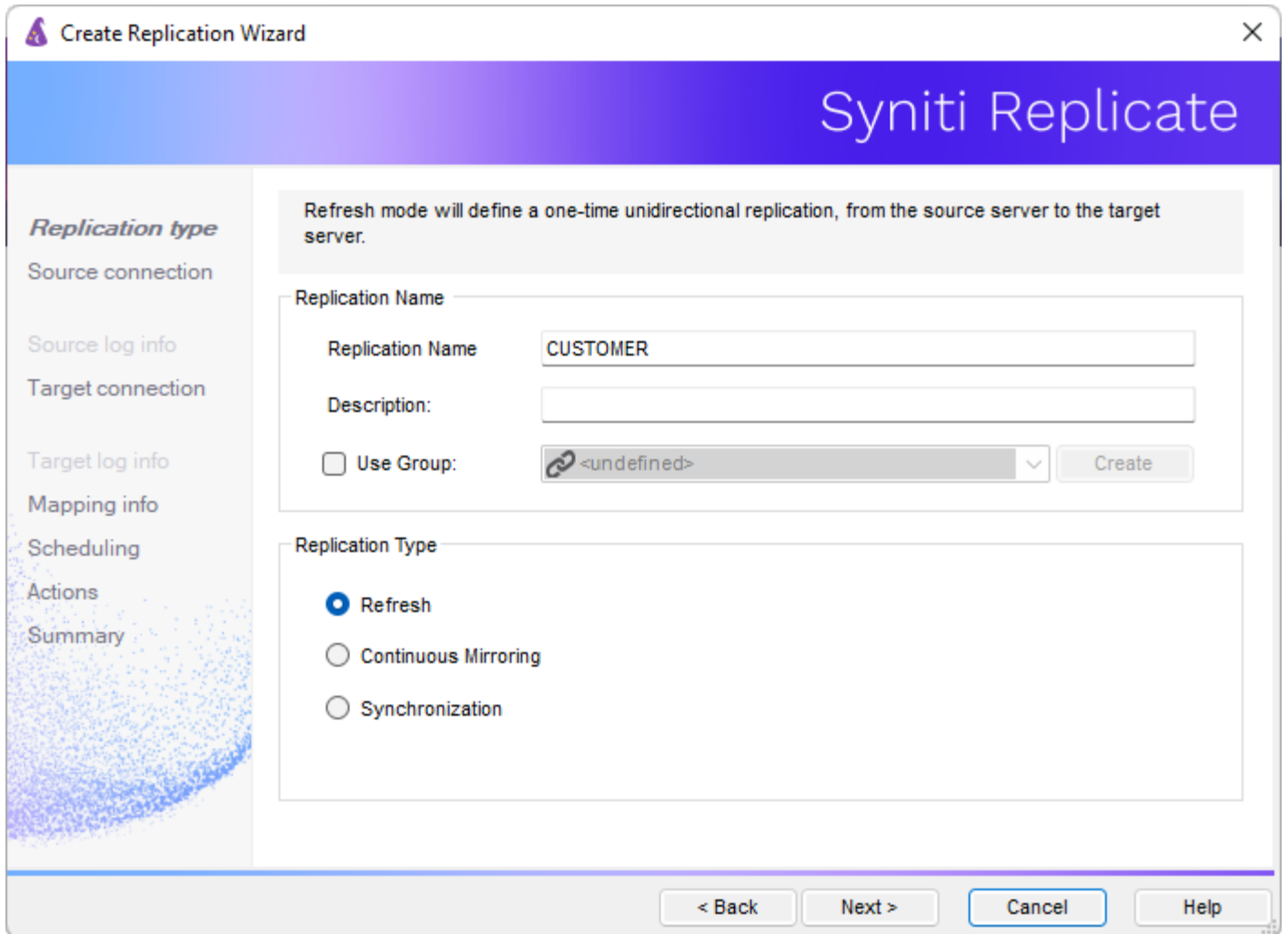
2. Click **Next** to enter the log settings. The fields and appropriate values depend on the database and log type. You can obtain a setup guide for a specific relational database by making a request in the Help Center.

3. Click **Next** to verify your settings against the source connection to the database. If any information is missing, you will not be able to proceed.

4. In the Actions screen, check the option to launch the **Add Target Connection wizard**.

5. Click **Next** to review your changes.

6. Click **Finish** to complete the wizard.

The source connection is now set up for transactional replications.

Set Up Target Connection to Amazon S3

This section assumes you have checked the **Enable Transactional Replication wizard** option to launch the Add Target Connection wizard. If not, to open the wizard from the Management Center, choose **Targets** in

the Metadata Explorer, then right-click to choose **Add New Connection...**
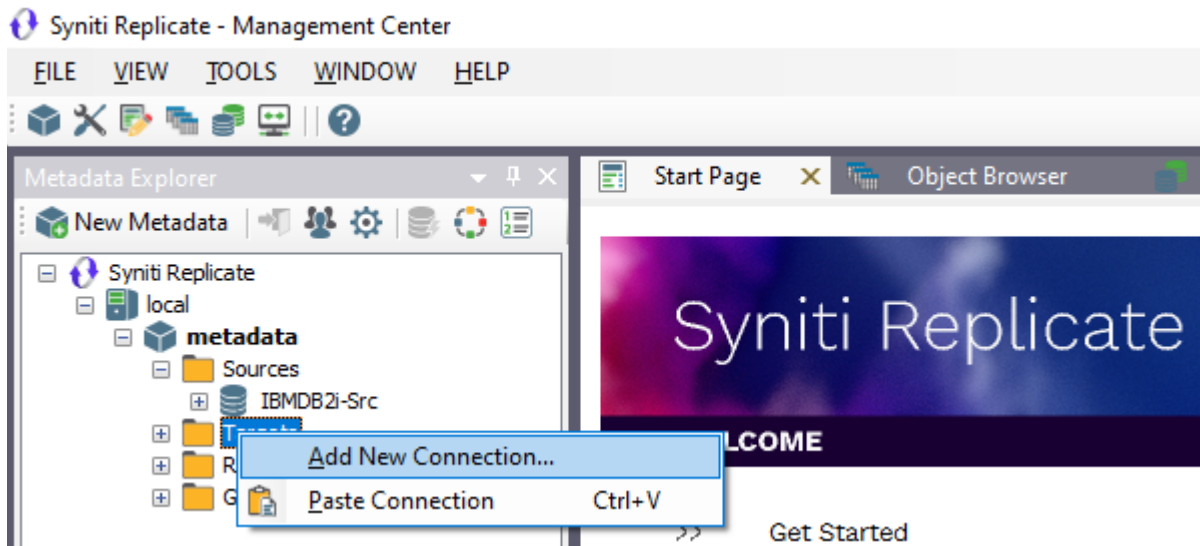
1. In the Target Connection Wizard **Provider** field, select the **Amazon S3** option.

Syniti Replicate
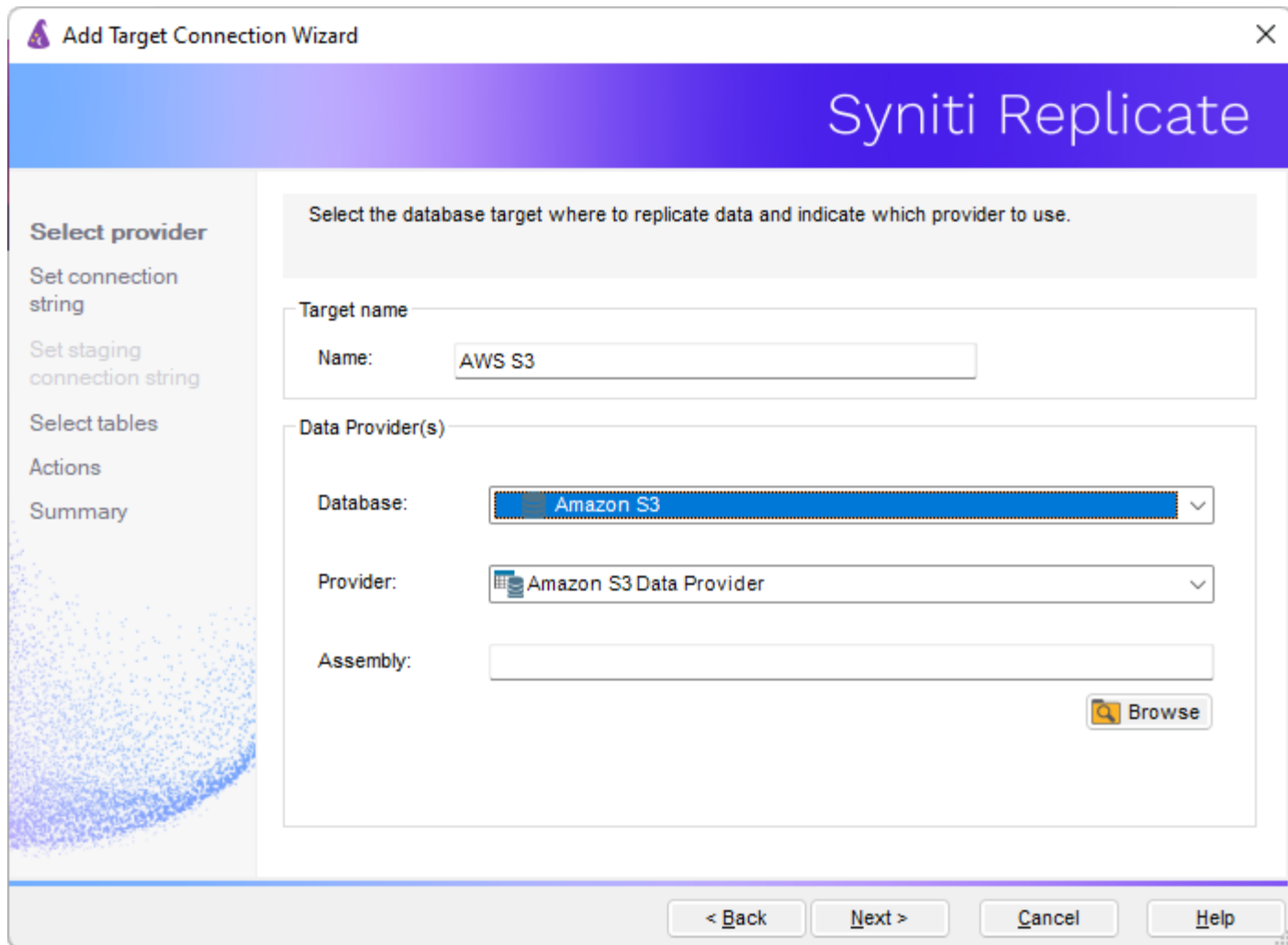
2. Click **Next** to display the **Set connection string** screen.

Syniti Replicate

3. Set values for the following fields:

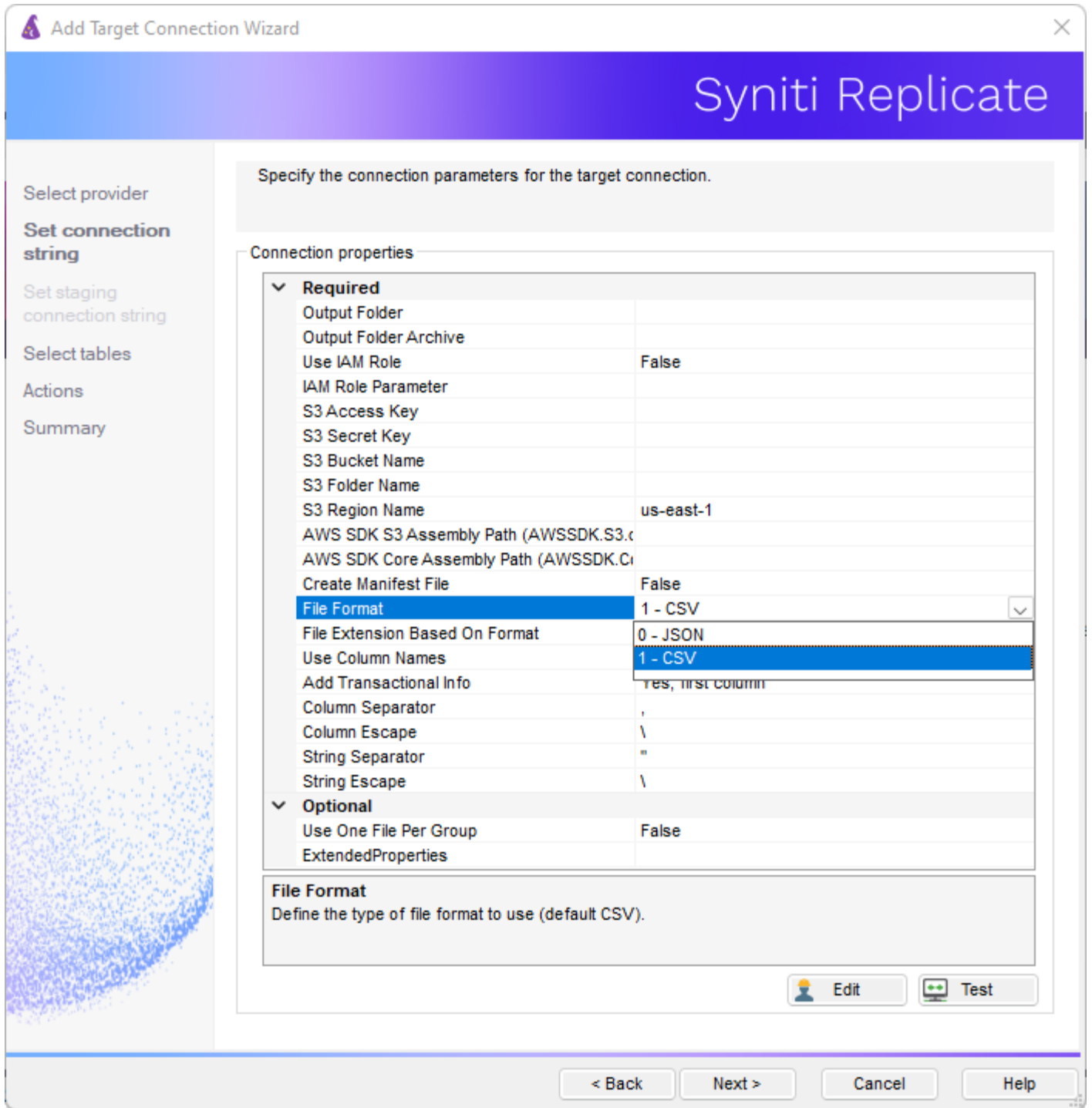|	
Output Folder	Local path to temporarily store replication output files.
Output folder Archive	(Optional) Set a path to archive files that are generated and uploaded.
Use IAM Role	Set to False by default. Set to True to use an IAM role for connections instead of the Access Key/Secret Key approach. For more information about using IAM roles, see https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles.html . Note that IAM roles work only for access between AWS services (such as when using S3 as a staging connection for bulk uploads to Redshift Redshift or when installing Syniti Replicate on a EC2 instance.)
IAM Role Parameter	(Optional – for use with Redshift only) The IAM role parameter should be configured as follows to access S3: arn:aws:iam:<aws-account-id>:role/<role-name>The parameter information is required when using S3 in conjunction with Redshift in the COPY from statement to access files you imported from S3 to Redshift. For more information about using IAM roles with Redshift, see https://docs.aws.amazon.com/redshift/latest/mgmt/authorizing-redshift-service.html
S3 Access Key	The Access Key from your AWS login ID.
S3 Secret Key	The Secret Access Key from your AWS login ID.
S3 Bucket Name	The bucket must already exist. It can either be a bucket that you are already using, or one you have created for this purpose.
AWS SDK S3 Assembly Path	The pathname to AWSSDK.S3.dll installed with the AWS .NET SDK.
AWS SDK Core Assembly Path	The pathname to the AWSSDK.Core.dll installed with the AWS .NET SDK.
Create Manifest	Optional. Default is False. If set to True, a manifest file is created for the data file output.
File Format	Optional. Default is 1-CSV for an output file containing comma-separated values. Choose 0-JSON to generate a JSON format output file.
File Extension Based on Format	Optional. False by default. Output file format will be: <replication_name><*date+time*><serial_number>.<"ref" or "mir"> Set to true to generate output files in the following format: <replication_name><*date+time*><serial_number>.<"ref" or "mir">.<csv or json>
Use Column Names	Optional. Default is True. When True, output file contains a list of column names as the first row.

Add Transactional Info	Optional. Default is Yes, first column. In this case, transaction details (operation type, transaction ID, transaction timestamp etc.) are added to the output file in the first column. Alternative values are None (no transaction details reported) or Yes, last column (transaction details are included in the last column,)
Column Separator	Optional. Default is ",".
Column Escape	Optional. Default is "\". Indicates the excape character used for the column separator.
String Separator	Optional. Default is ".
String Escape	Optional. Default is "\". Indicates the escape character used for the string separator.
Use One File Per Group	Optional. Default is False. One output file per replication is generated. When set to True, one output file per replication group is generated, assuming that one or more Replication Groups have been established.
ExtendedProperties	No additional properties currently supported.

4. Click **Next** to display the Select Tables screen.
   At this point, there is no text output structure available to display. You can add the information after completing the Target Connection wizard.

5. Click **Next** to display the **Summary** page.

6. Click **Finish** to complete the wizard.

Add Table Information to the Target Connection

The target connection is displayed in the Metadata Explorer, but you still need to add the file representation for source table data so that when you create replications below, you can specify a source table and target "file."

1. In the Metadata Explorer, expand the source connection you created above.

2. Select and drag a source table to the target file connection.

3. The Create Target Table wizard is displayed.
   Although you are not actually creating tables, you can use this wizard to create a representation of the file data.

4. In the **Source Connection** screen, you should see the source table you selected above.

5. Click **Next** to display the **Target Connection** screen.

6. Verify that the screen displays the correct target connection name, and table name.

7. Click **Next** to review the table structure.
   At this point, you can modify data types, null values and so on, if you want to modify the data eventually sent to a file.

8. Click **Next** to display the **SQL Script** screen.
   The contents of this screen are inactive because there is no editable SQL script to create a table. Instead, Syniti Replicate outputs the table information to a file.

9. Click **Next** to display the **Summary** screen.

10. Click **Finish** to create the table representation in the Metadata Explorer.

11. Click **Yes** to add the table name to the target connection entry in the Metadata Explorer

12. Repeat steps 2 through 11 for each source table that you want to replicate to a file.

At the end of this process, you should have a list of table representations under the Files target connection in the Metadata Explorer.

Define the Replication

This section assumes you have checked the Create Replication wizard option to launch the Create Replication wizard. If not, to open the wizard from the Management Center, choose the table you want to replicate in the Metadata Explorer, then right-click to choose **Replication > Create New Replication...**

NOTE: This section describes how to create a single replication. However, many situations call for setting up replications on a set of source tables. Use the **Create Multiple Replications** wizard accessible from the source connection in the Metadata Explorer for this purpose.

1. In the **Define Replication Type** screen, type a name to identify the replication.

2. Optionally provide a description of the replication.

3. In the Replication Mode area, choose **Continuous Mirroring**.
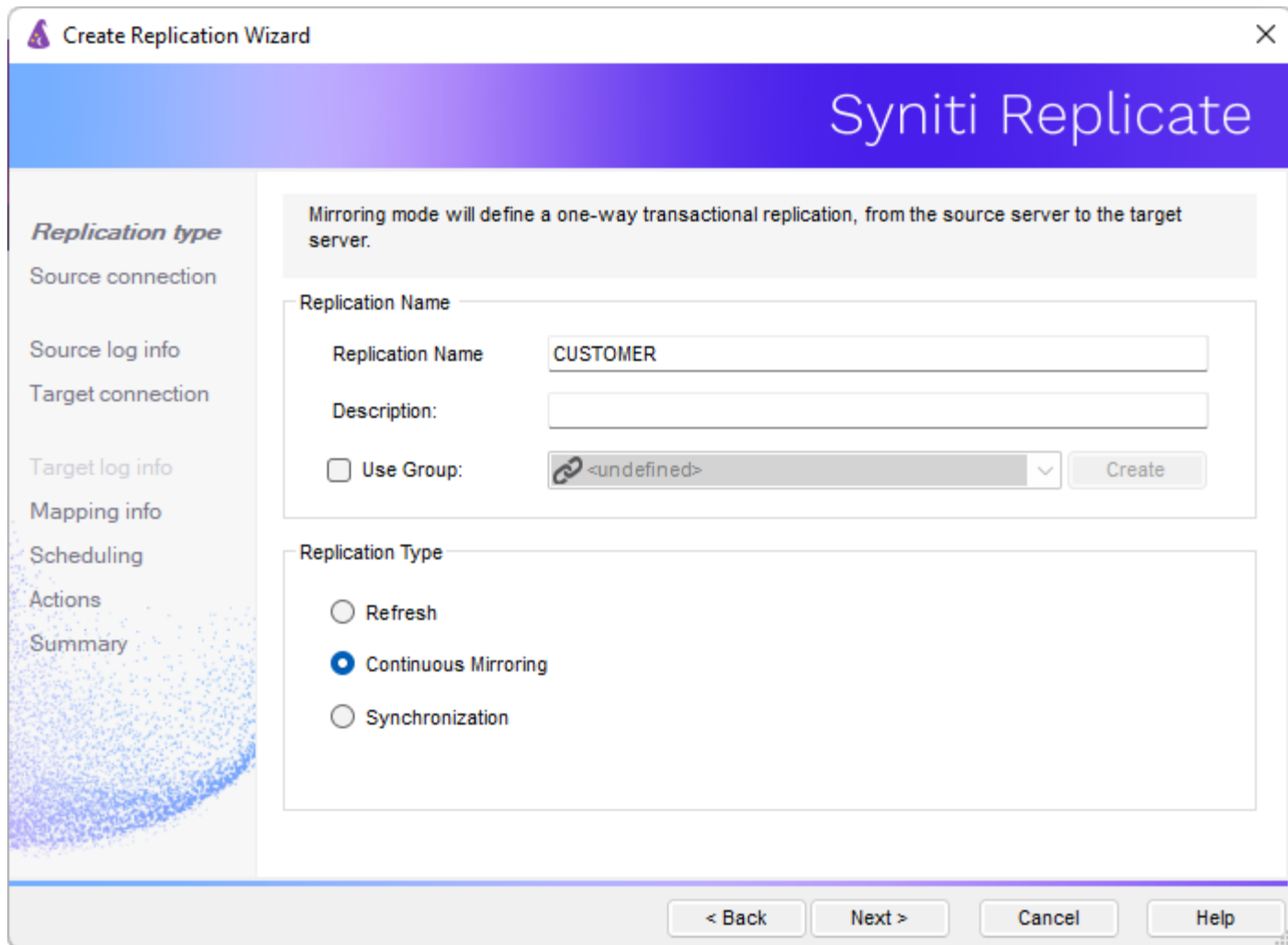
4. Click **Next** to go to the **Select Source Connection** screen.
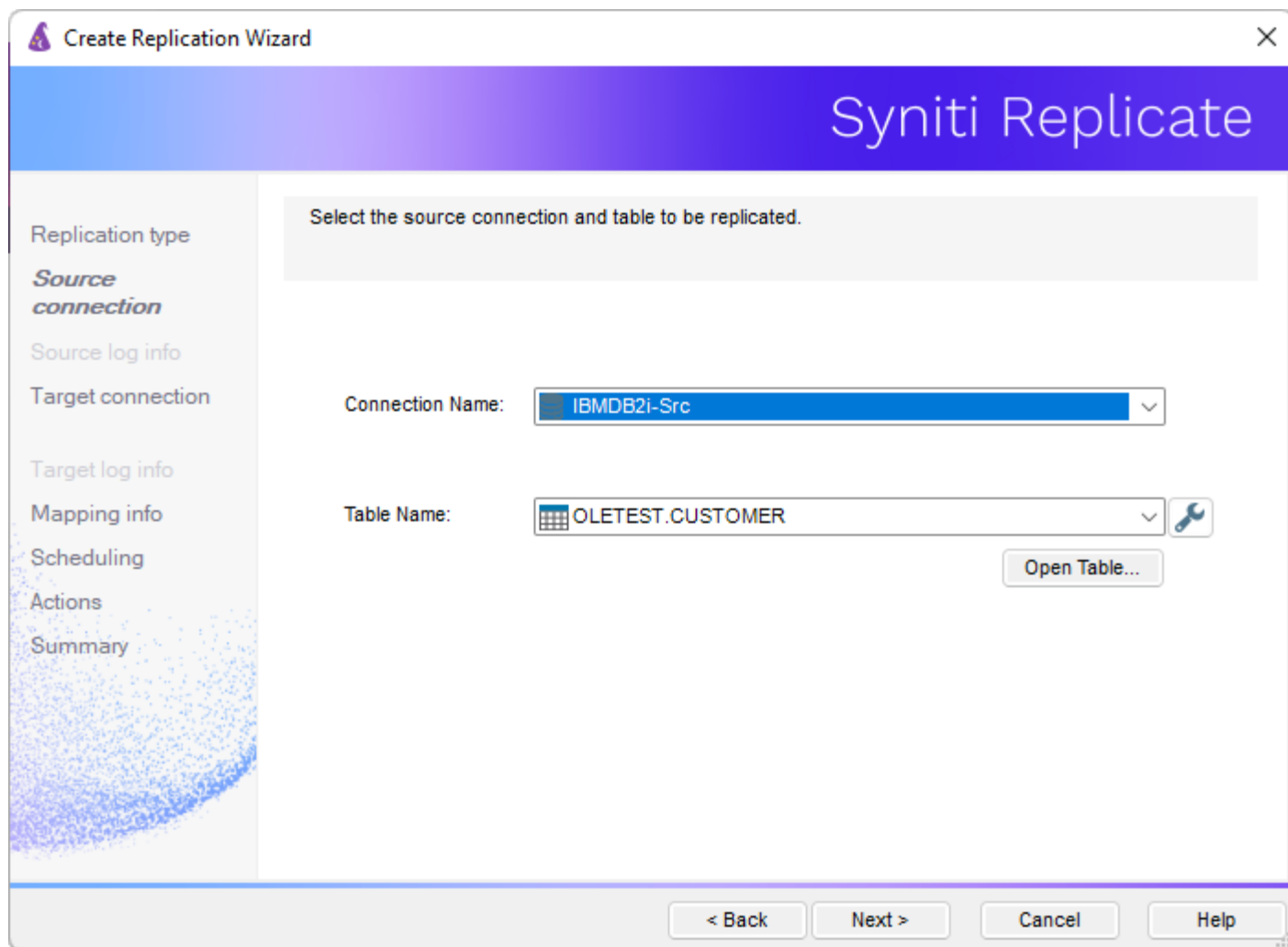
5. Choose the source name from the drop-down list that includes all the source connections you have created in Syniti Replicate.

6. Choose the table that you want to replicate from the drop-down list.

7. If you want more information about the table before proceeding, click **Open Table…**.

8. Click **Next** to go to the **Source Log Info** screen.

Syniti Replicate

The first two fields on this screen depend on the source database you are using. In this explanation, we assume you are using IBM Db2 for i (iSeries or AS/400). Check the help for the Replication Wizard for details on the other databases.

The Journal field is automatically filled in by retrieving the information from Db2 for i. The Receiver field will be automatically filled in after setting the Transaction ID. You do not need to modify these fields. However, if the library that you have specified as a source is not journaled, you will need to ask your system administrator to journal the library.

9. In the Transaction ID field, click **Read** to open the Read Point dialog.

10. Choose either the current transaction or a transaction ID based on a time and date.

11. Click **OK** to add the value to the **Source Log Info** screen.

12. Set the value of the **Read Interval** field to the frequency with which you want Syniti Replicate to check the transaction log for new events to mirror.

13. Click **Next** to go to the **Select Target Connection** screen.

14. Choose the target connection for Amazon S3 from the drop-down list that includes all the target connections you have created in Syniti Replicate.

15. Choose the data set you want to replicate from the drop-down list.
    If the drop-down list is empty, exit the wizard and add or create a target data set.

16. Click **Next** to go to the **Set Mapping Info** screen.
    Source columns and target data with the same name are automatically mapped.

17. Set a start time for the replication. The **Start Time** field indicates the time at which the Replication Agent will begin considering the replication for execution.

18. Check the option to **Execute Initial Refresh**.
    A full replication will be performed from the source table to the data file, prior to starting the mirroring process where only incremental changes will be replicated.

19. Go to the Mirroring Schedule tab.

20. Select how you want to run the replication:

    - Run Continuously: the transaction log will be checked for changes to the table using the frequency that you specified on the Set Log Info tab. Any changes will be replicated to the target table.
    - Schedule Interruptions: the replication process will run as above, except for interruptions specified when you click the Schedule button in the Scheduler dialog.

21. Make sure the **Enable Replication** option is checked. This is required for the replication to run.

22. Click **Next** to go to the **Summary** screen.

23. Click **Finish** to complete the wizard.

Run the Replication

If you installed the Replication Agent as a service during setup, you just need to start the service using the Service Monitor program in the Windows Notification Area.
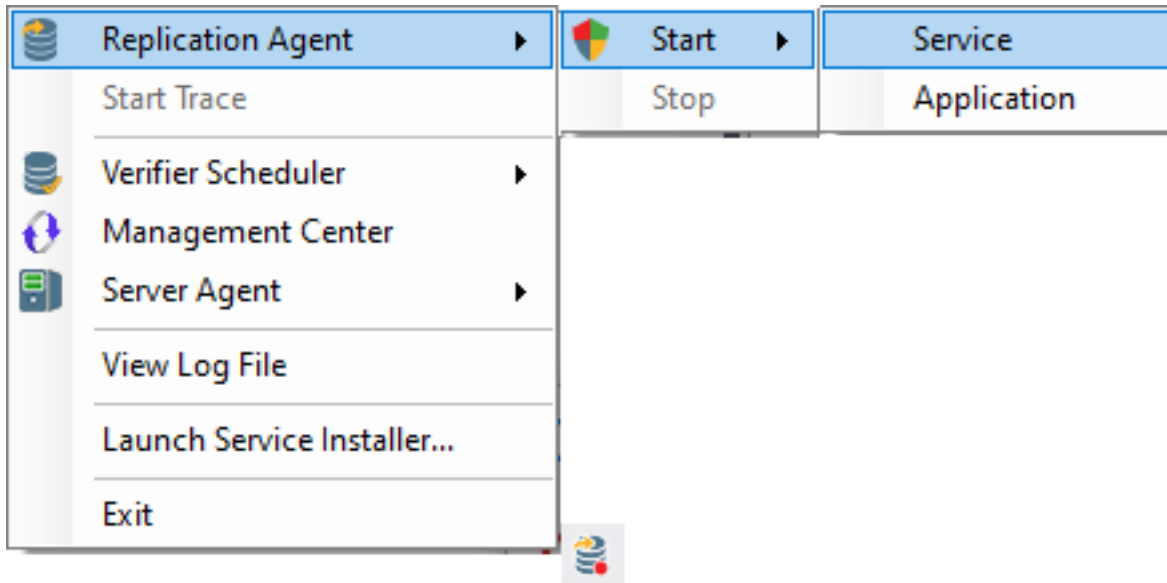
The replication that you have scheduled should start at the specified time. Use the <u>Replication Monitor</u> tab in the Management Center to track the progress of the replication.

If you would like to install the Replication Agent as a service:

● From the Windows Desktop **Start** menu, choose **Syniti Replicate**, then **Service Installer**.

● Manage the service from Service Monitor program (located in the Syniti Replicate install folder or on the Windows **Start** > **Programs** > **Startup** menu).

● Use the Replication Monitor tab in the Management Center to track the progress of the replication.

To run the Replication Agent interactively:

● In the Windows Notification Area, select the Service Monitor icon.

● From the right mouse button menu, choose **Replication Agent**, then **Start** then **Application**.
The replication that you have scheduled should start at the specified time.

● Use the Replication Monitor tab in the Management Center to track the progress of the replication.

CSV File Structure

Files have a .mir or .mir.csv extension when used for mirroring and a .ref or .ref.csv extension when used for snapshot, or refresh replications.

Data files replicated to S3 using mirroring have the following structure if **Yes, first column** is selected.

Column	Value
1	Sequence number
2	Timestamp of transaction
3	User, if provided by the database
4	Table name
5	Operation code – one of I (Insert), U (Update) or D (Delete)
6	Previous primary key value for the database column. This is used for update operations to identify the correct record
7	Next primary key value. This is used for update operations to identify the correct record in cases where the primary key is changed during the update operation
8	Database column value

Example CSV File

Example filename: `corp_20150115T161123155064_1.mir.csv`

```
10,2015-01-22 15:32:39.000000,,"glb","I",,2,"test insert"
11,2015-01-22 15:32:40.000000,," glb ","D",10,,
12,2015-01-22 15:32:42.000000,," glb ","U",104,104,"my updated"
13,2015-01-22 15:32:44.000000,," glb ","U",1,12,"sd"
```

Example JSON File

Example filename: `corp_20150115T161123155064_1.mir.json`

```
{"ID":0,"TransactionID":"00000083000000ec0003-1","TransactionTS":"2021-07-
01T15:30:47.593Z","Table":"DOTNET","Operation":"I","After":{"ID":2,"NAME":"new name
"}}
```