Informatica®

Informatica® Cloud Data Integration

Google Bigtable Connector

Informatica Cloud Data Integration Google Bigtable Connector
April 2024

© Copyright Informatica LLC 2020, 2024

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, Informatica Cloud, and PowerCenter are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at https://www.informatica.com/trademarks.html. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

See patents at https://www.informatica.com/legal/patents.html.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2024-04-08

Table of Contents

Preface 4

Informatica Resources 4

- Informatica Documentation 4
- Informatica Intelligent Cloud Services web site 4
- Informatica Intelligent Cloud Services Communities 4
- Informatica Intelligent Cloud Services Marketplace 4
- Data Integration connector documentation 5
- Informatica Knowledge Base 5
- Informatica Intelligent Cloud Services Trust Center 5
- Informatica Global Customer Support 5

Chapter 1: Introduction to Google Bigtable Connector 6

Data Integration Hosted Agent 6

Google Bigtable Connector task and objects types 6

Administration of Google Bigtable Connector 7

Introduction to Google Cloud Bigtable 7

Chapter 2: Google Bigtable Connections 8

Google Bigtable connection properties 8

Chapter 3: Mappings and Mapping Tasks with Google Bigtable Connector 9

Google Cloud Bigtable sources in mappings 9

Google Cloud Bigtable targets in mappings 10

Appendix A: Data Type Reference 11

Google Cloud Bigtable and transformation data types 11

Index 12

Preface

Use *Google Bigtable Connector* to learn how to read from or write to Google Cloud Bigtable by using Cloud Data Integration. Learn to create a connection, develop mappings, and run mapping tasks in Cloud Data Integration.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Intelligent Cloud Services web site

You can access the Informatica Intelligent Cloud Services web site at http://www.informatica.com/cloud. This site contains information about Informatica Cloud integration services.

Informatica Intelligent Cloud Services Communities

Use the Informatica Intelligent Cloud Services Community to discuss and resolve technical issues. You can also find technical tips, documentation updates, and answers to frequently asked questions.

Access the Informatica Intelligent Cloud Services Community at:

https://network.informatica.com/community/informatica-network/products/cloud-integration

Developers can learn more and share tips at the Cloud Developer community:

https://network.informatica.com/community/informatica-network/products/cloud-integration/cloud-developers

Informatica Intelligent Cloud Services Marketplace

Visit the Informatica Marketplace to try and buy Data Integration Connectors, templates, and mapplets:

https://marketplace.informatica.com/

Data Integration connector documentation

You can access documentation for Data Integration Connectors at the Documentation Portal. To explore the Documentation Portal, visit https://docs.informatica.com.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Intelligent Cloud Services Trust Center

The Informatica Intelligent Cloud Services Trust Center provides information about Informatica security policies and real-time system availability.

You can access the trust center at https://www.informatica.com/trust-center.html.

Subscribe to the Informatica Intelligent Cloud Services Trust Center to receive upgrade, maintenance, and incident notifications. The Informatica Intelligent Cloud Services Status page displays the production status of all the Informatica cloud products. All maintenance updates are posted to this page, and during an outage, it will have the most current information. To ensure you are notified of updates and outages, you can subscribe to receive updates for a single component or all Informatica Intelligent Cloud Services components. Subscribing to all components is the best way to be certain you never miss an update.

To subscribe, on the Informatica Intelligent Cloud Services Status page, click **SUBSCRIBE TO UPDATES**. You can choose to receive notifications sent as emails, SMS text messages, webhooks, RSS feeds, or any combination of the four.

Informatica Global Customer Support

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

To find online support resources on the Informatica Network, click **Contact Support** in the Informatica Intelligent Cloud Services Help menu to go to the **Cloud Support** page. The **Cloud Support** page includes system status information and community discussions. Log in to Informatica Network and click **Need Help** to find additional resources and to contact Informatica Global Customer Support through email.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at https://www.informatica.com/services-and-training/support-services/contact-us.html.

CHAPTER 1

Introduction to Google Bigtable Connector

You can use Google Bigtable Connector to connect to Google Cloud Bigtable from Data Integration. Use Google Bigtable Connector to read data from or write data to Google Cloud Bigtable. You can use Google Cloud Bigtable objects as sources and targets in mapping tasks and mappings.

When you run a task or mapping, the Secure Agent uses the Google Cloud Bigtable and Google Cloud BigTable Admin APIs to perform the specified operation and reads data from or writes data to Google Cloud Bigtable tables.

Note: This connector is developed by Informatica partners, third-party system integrators, or customers. You can look up the Product Availability Matrix or Informatica Marketplace for more information.

Data Integration Hosted Agent

If you have the Cloud Runtime license, you can use the Data Integration Hosted Agent (Hosted Agent) as a runtime environment for a Google Bigtable connection.

Data Integration Secure Agents are installed locally. As an alternative to installing a Secure Agent, you can use a Hosted Agent. Hosted Agents are hosted at a Data Integration hosting facility. The Data Integration hosting facility manages the Hosted Agent runtime environment and the Agents that run in it. You cannot add, delete, or configure a Hosted Agent runtime environment. Because you do not install a Hosted Agent, you do not have access to files normally stored in the Secure Agent directory, such as configuration, success, and reject files.

Google Bigtable Connector task and objects types

You can use Google Bigtable Connector to read data from a Google Cloud Bigtable source. You can insert or update data in a Google Cloud Bigtable target and delete data from a Google Cloud Bigtable target. You cannot perform upsert operations on a Google Cloud Bigtable target.

The following table lists the Google Cloud Bigtable object types that you can include in Data Integration tasks:

Task Type	Source	Target
Mapping	Yes	Yes
Synchronization	No	No

Administration of Google Bigtable Connector

Google Cloud Bigtable is a sparsely populated table that can scale to billions of rows and thousands of columns. Google Cloud Bigtable enables you to store terabytes or even petabytes of data.

Before you use Google Bigtable Connector, you must complete the following prerequisite tasks:

- Create a Google account to access Google Cloud Bigtable.
- Ensure that you have the client_email, project_id, and private_key values for the service account. You will need to enter these details when you create a Google Bigtable connection in Data Integration.
- Enable billing for the project that you created.
- Ensure that you have enabled the Cloud Bigtable and Cloud Bigtable Admin APIs for your service account. Google Bigtable Connector uses the Google Cloud Bigtable and Cloud Bigtable Admin APIs to integrate with Google Cloud Bigtable.
- Ensure that you have created a Google Cloud Bigtable instance with the same name as the user name in the service account ID.
  For more information about creating a Google Cloud Bigtable instance, click the following URL:

  https://cloud.google.com/bigtable/docs/creating-instance#creating-instance

Introduction to Google Cloud Bigtable

Google Cloud Bigtable is a web service that allows global storage and retrieval of large volumes of data at any time.

Google Cloud Bigtable is a sparsely populated table that can scale to billions of rows and thousands of columns and enables you to store terabytes or even petabytes of data. A Bigtable is a sparse, distributed, persistent multidimensional sorted map.

Cloud Bigtable stores data in massively scalable tables, each of which is a sorted key-value map. The table is composed of rows and columns. Each row typically describes a single entity and the columns contain individual values for each row. Each row is indexed by a single row key and columns that are related to one another are typically grouped together into a column family. Each column is identified by a combination of the column family and a column qualifier, which is a unique name within the column family.

Each row-column intersection can contain multiple cells, or versions, at different timestamp, providing a record of how the stored data has been altered over time. If a cell in Google Cloud Bigtable table does not contain any data, it does not take up any space.

CHAPTER 2

Google Bigtable Connections

Create a Google Bigtable connection to read data from a Google Cloud Bigtable source and write data to a Google Cloud Bigtable target. You can use Google Bigtable connections in mapping tasks and mappings.

Google Bigtable connection properties

When you create a Google Bigtable connection, you must configure the connection properties.

The following table describes the Google Cloud Bigtable connection properties:

Property	Description
Connection Name	The name of the connection. The name is not case sensitive and must be unique within the domain. You can change this property after you create the connection. The name cannot exceed 128 characters, contain spaces, or contain the following special characters:~ ` ! $ % ^ & * ( ) - + = { [}] | \ : ; " ' < , > . ? /
Description	Optional. The description of the connection. The description cannot exceed 4,000 characters.
Type	The **googleBigTable** connection type.
Runtime Environment	Runtime environment that contains the Secure Agent used to access Google Cloud Bigtable.
Project ID	Specifies the project_id value present in the JSON file that you download after you create a service account.
Service Account ID	Specifies the client_email value present in the JSON file that you download after you create a service account.
Service Account Key	Specifies the private_key value present in the JSON file that you download after you create a service account.

CHAPTER 3

Mappings and Mapping Tasks with Google Bigtable Connector

Use a mapping to define the data flow logic between the source and the target. You can also add other transformations in the mapping to transform the data.

Create the mapping from the Data Integration Mapping Designer. In advanced mode, the Mapping Designer updates the mapping canvas to include transformations and functions that enable advanced functionality.

Google Cloud Bigtable sources in mappings

To read data from a Google Cloud Bigtable bucket, configure a Google Cloud Bigtable object as the Source transformation in a mapping.

Specify the name and description of the Google Cloud Bigtable source. Configure the source and advanced properties for the source object.

The following table describes the properties that you can configure for a Google Cloud Bigtable source:

Property	Description
Connection	Name of the Google Bigtable source connection.
Source Type	Type of the Google Cloud Bigtable source object available. You can read data to a single Google Cloud Bigtable source object.
Object	Name of the source object for the mapping.

You can set the tracing level in the advanced properties session to determine the amount of details that logs contain.

The following table describes the tracing levels that you can configure:

Tracing Level	Description
Terse	The Secure Agent logs initialization information, error messages, and notification of rejected data.
Normal	The Secure Agent logs initialization and status information, errors encountered, and skipped rows due to transformation row errors. Summarizes session results, but not at the level of individual rows.
Verbose Initialization	In addition to normal tracing, the Secure Agent logs additional initialization details, names of index and data files used, and detailed transformation statistics.
Verbose Data	In addition to verbose initialization tracing, the Secure Agent logs each row that passes into the mapping. Also notes where the Secure Agent truncates string data to fit the precision of a column and provides detailed transformation statistics. When you configure the tracing level to verbose data, the Secure Agent writes row data for all rows in a block when it processes a transformation.

Google Cloud Bigtable targets in mappings

To write data to a Google Cloud Bigtable target, configure a Google Cloud Bigtable object as the Target transformation in a mapping.

Specify the name and description of the Google Cloud Bigtable target. You can configure the target and advanced properties for the target object.

The following table describes the properties that you can configure for a Google Cloud Bigtable target:

Property	Description
Connection	Name of the Google Bigtable target connection.
Target Type	Type of the Google Cloud Bigtable target objects available. You can write data to a single or parameterize Google Cloud Bigtable target object. You cannot write data to multiple objects.
Object	Name of the Google Cloud Bigtable target object.
Operation	Select insert, update, or delete as the target operation. You cannot perform upsert operations on a Google Cloud Bigtable object.

APPENDIX A

Data Type Reference

Data Integration uses the following data types in mappings, synchronization tasks, and mapping tasks with Google Cloud Bigtable:

Google Cloud Bigtable native data types

Google Cloud Bigtable data types appear in the **Fields** tab for Source and Target transformations when you choose to edit metadata for the fields.

Transformation data types

Set of data types that appear in the remaining transformations. They are internal data types based on ANSI SQL-92 generic data types, which Data Integration uses to move data across platforms. Transformation data types appear in all remaining transformations in a mapping, synchronization task, or mapping task.

When Data Integration reads source data, it converts the native data types to the comparable transformation data types before transforming the data. When Data Integration writes to a target, it converts the transformation data types to the comparable native data types.

Google Cloud Bigtable and transformation data types

The following table describes the data types that Data Integration supports for Google Cloud Bigtable sources and targets:

Google Cloud Bigtable Data Type	Transformation Data Type	Range and Description for the Transformation Data Type
STRING	String	1 to 104,857,600 characters. Default precision is 256. You can increase the value up to 104857600 characters.

INDEX

C

Cloud Application Integration community
  URL 4
Cloud Developer community
  URL 4
connections
  Google Bigtable 8

D

Data Integration community
  URL 4
data type reference
  overview 11

G

Google Bigtable
  connection properties 8
Google Bigtable connections
  overview 8
Google Bigtable Connector
  administration 7
  overview 6
  supported task and object types 6
Google Cloud Bigtable
  data types 11
  features 7
  introduction 7
  native data types 11

H

Hosted Agent 6

I

Informatica Global Customer Support
  contact information 5
Informatica Intelligent Cloud Services
  web site 4

M

maintenance outages 5
mapping
  Google Cloud Bigtable sources 9
  Google Cloud Bigtable targets 10
mapping advanced properties
  Google Cloud Bigtable source 9
  Google Cloud Bigtable target 10
mapping properties
  Google Cloud Bigtable source 9
  Google Cloud Bigtable targets 10

S

status
  Informatica Intelligent Cloud Services 5
system status 5

T

transformation
  data types 11
trust site
  description 5

U

upgrade notifications 5

W

web site 4