Informatica PowerExchange for MongoDB
(Version 10.0)

User Guide for PowerCenter

This product includes software licensed under the Academic Free License (http://www.opensource.org/licenses/afl-3-0.php), the Common Development and Distribution License (http://www.opensource.org/licenses/2-clause-fop.php) the Common Public License (http://www.opensource.org/licenses/cpl1.0.php), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (http://www.opensource.org/licenses/bsd-license.php), the new BSD License (http://opensource.org/licenses/BSD-3-Clause), the MIT License (http://www.opensource.org/licenses/mit-license.php), the Artistic License (http://www.opensource.org/licenses/artistic-license-1.0) and the Initial Developer's Public License Version 1.0 (http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at http://xstream.codehaus.org/license.html. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit http://www.extreme.indiana.edu/.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.


DISCLAIMER: Informatica LLC provides this software "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 NONINFRINGEMENT.

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.
# Table of Contents

## Preface ......................................................... 6

Informatica Resources .................................................. 6
- Informatica My Support Portal ................................ 6
- Informatica Documentation ....................................... 6
- Informatica Product Availability Matrixes ..................... 6
- Informatica Web Site .............................................. 7
- Informatica How-To Library ...................................... 7
- Informatica Knowledge Base ..................................... 7
- Informatica Support YouTube Channel ............................ 7
- Informatica Marketplace .......................................... 7
- Informatica Velocity ............................................... 7
- Informatica Global Customer Support ............................ 7

## Chapter 1: Introduction to PowerExchange for MongoDB ............... 9

PowerExchange for MongoDB Overview ................................ 9
Introduction to MongoDB ............................................. 10
PowerExchange for MongoDB Implementation ...................... 10

## Chapter 2: PowerExchange for MongoDB Configuration ............. 13

PowerExchange for MongoDB Configuration Overview ................ 13
Prerequisites .................................................................. 13
- PowerExchange for MongoDB Upgrade ......................... 14
Informatica MongoDB ODBC Driver Configuration .................. 14
  - Configuring the Informatica MongoDB ODBC Driver on Linux ......... 14
Data Source Name Configuration on Windows .......................... 15
  - MongoDB ODBC Connection Properties ......................... 15
  - Advanced Properties ............................................ 16

## Chapter 3: Schema Definition ...................................... 19

Schema Definition Overview ............................................ 19
Schema Editor .......................................................... 19
  - Collection Properties .............................................. 20
  - Column Metadata .................................................... 20
Virtual Tables .......................................................... 21
  - Virtual Table Options ............................................. 22
  - Virtual Tables - An Example ..................................... 23
Metadata Caching ....................................................... 25
  - Defining the Schema for a Collection ............................ 25
    - Updating the Schema File ..................................... 26
# Table of Contents

## Chapter 4: MongoDB Sources

- MongoDB Sources Overview ................................................... 27
- Importing a MongoDB Source Definition ...................................... 27
- MongoDB Reader Sessions ...................................................... 27
- Example: MongoDB Reader Mapping .......................................... 28

## Chapter 5: MongoDB Targets

- MongoDB Targets Overview ..................................................... 31
- Importing MongoDB Target Definitions ........................................ 31
- MongoDB Writer Sessions ....................................................... 31
- Example: MongoDB Target Mapping .......................................... 32

## Appendix A: Datatype Reference

- MongoDB, ODBC, and Transformation Datatypes .......................... 35

## Index

- ............................................................ 36
Preface

The Informatica PowerExchange for MongoDB User Guide describes how to use PowerExchange for MongoDB with PowerCenter to extract data from and load data to MongoDB. The guide is written for database administrators and developers who are responsible for developing mappings and workflows. This guide assumes that you have knowledge of MongoDB and PowerCenter.

Informatica Resources

Informatica My Support Portal

As an Informatica customer, the first step in reaching out to Informatica is through the Informatica My Support Portal at https://mysupport.informatica.com. The My Support Portal is the largest online data integration collaboration platform with over 100,000 Informatica customers and partners worldwide.

As a member, you can:
- Access all of your Informatica resources in one place.
- Review your support cases.
- Search the Knowledge Base, find product documentation, access how-to documents, and watch support videos.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Documentation

The Informatica Documentation team makes every effort to create accurate, usable documentation. If you have questions, comments, or ideas about this documentation, contact the Informatica Documentation team through email at infa_documentation@informatica.com. We will use your feedback to improve our documentation. Let us know if we can contact you regarding your comments.

The Documentation team updates documentation as needed. To get the latest documentation for your product, navigate to Product Documentation from https://mysupport.informatica.com.

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. You can access the PAMs on the Informatica My Support Portal at https://mysupport.informatica.com.
Informatica Web Site

You can access the Informatica corporate web site at https://www.informatica.com. The site contains information about Informatica, its background, upcoming events, and sales offices. You will also find product and partner information. The services area of the site includes important information about technical support, training and education, and implementation services.

Informatica How-To Library

As an Informatica customer, you can access the Informatica How-To Library at https://mysupport.informatica.com. The How-To Library is a collection of resources to help you learn more about Informatica products and features. It includes articles and interactive demonstrations that provide solutions to common problems, compare features and behaviors, and guide you through performing specific real-world tasks.

Informatica Knowledge Base

As an Informatica customer, you can access the Informatica Knowledge Base at https://mysupport.informatica.com. Use the Knowledge Base to search for documented solutions to known technical issues about Informatica products. You can also find answers to frequently asked questions, technical white papers, and technical tips. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team through email at KB.Feedback@informatica.com.

Informatica Support YouTube Channel

You can access the Informatica Support YouTube channel at http://www.youtube.com/user/INFASupport. The Informatica Support YouTube channel includes videos about solutions that guide you through performing specific tasks. If you have questions, comments, or ideas about the Informatica Support YouTube channel, contact the Support YouTube team through email at supportvideos@informatica.com or send a tweet to @INFASupport.

Informatica Marketplace

The Informatica Marketplace is a forum where developers and partners can share solutions that augment, extend, or enhance data integration implementations. By leveraging any of the hundreds of solutions available on the Marketplace, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at http://www.informaticamarketplace.com.

Informatica Velocity

You can access Informatica Velocity at https://mysupport.informatica.com. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Global Customer Support

You can contact a Customer Support Center by telephone or through the Online Support. Online Support requires a user name and password. You can request a user name and password at http://mysupport.informatica.com.
CHAPTER 1

Introduction to PowerExchange for MongoDB

This chapter includes the following topics:

- PowerExchange for MongoDB Overview, 9
- Introduction to MongoDB, 10
- PowerExchange for MongoDB Implementation, 10

PowerExchange for MongoDB Overview

PowerExchange for MongoDB provides connectivity between Informatica and MongoDB. Use PowerExchange for MongoDB to extract and load MongoDB documents through the PowerCenter Integration Service.

You can use PowerExchange for MongoDB to integrate and migrate data from diverse data sources that are incompatible with MongoDB architecture.

You can use PowerExchange for MongoDB for the following data integration scenarios:

- Create a MongoDB data warehouse. You can aggregate data from MongoDB and other source systems, transform the data, and write the data to MongoDB.
- Migrate data from a relational database or other data sources to MongoDB. For example, you want to migrate data from a relational database to MongoDB. You can write data from multiple relational database tables with different schemas to the same MongoDB collection. A MongoDB collection contains the data in a MongoDB database.
- Move data between operational data stores to synchronize data. For example, an online marketplace uses a relational database as the operational data store. You want to use MongoDB instead of the relational database. However, you want to maintain the relational database along with MongoDB for a period of time. You can use PowerExchange for MongoDB to synchronize data between the relational data store and the MongoDB data store.
- Migrate data from MongoDB to a data warehouse for reporting. For example, your organization uses a business intelligence tool that does not support MongoDB. You must migrate the data from MongoDB to a data warehouse so that the business intelligence tool can use the data to generate reports.
Introduction to MongoDB

MongoDB is an open source, document based, NoSQL database that maintains dynamic schema. You can maintain more than one database on a MongoDB server.

A MongoDB database contains a set of collections. A collection is a set of documents and is similar to a table in a relational database. MongoDB stores records as documents that are similar to rows in a relational database. A document contains fields that are similar to columns in a relational database. A document can have a dynamic schema. A document in a collection does not need to have the same set of fields or structure as another document in the same collection. A document can also contain nested documents.

The following schema provides a sample MongoDB document from the collection called Product:

```json
{
  sku: "111445GB3",
  title: "CM Phone",
  description: "The best in the world."
}

manufacture_details: {
  model_number: "CM",
  release_date: new ISODate("2011-07-17T22:14:15.656Z")
},

shipping_details: {
  weight: 350,
  width: 10,
  height: 10,
  depth: 1
},

quantity: 99,

pricing: [
  {
    region: "North America",
    cost_price: 1000,
    sale_price: 1200
  },
  {
    region: "Europe",
    cost_price: 1200,
    sale_price: 1500
  }
]
```

In the example, sku, title, description, quantity, manufacture_details, shipping_details, and pricing are fields. The fields manufacture_details and shipping_details are nested document type fields and pricing is an array type field.

PowerExchange for MongoDB Implementation

To extract and load MongoDB data, create MongoDB source and target definitions in the Designer. You can add a source or target definition to a session and run the session to process the data.

PowerExchange for MongoDB includes the Informatica MongoDB ODBC driver that connects to the MongoDB server. PowerExchange for MongoDB supports the MMAPv1 storage engine in MongoDB. You can create an ODBC connection to extract data from or load data to a MongoDB database. You can also configure the replica sets for the MongoDB server so that the PowerCenter Integration Service can access the secondary servers if the primary server is not available.
The Designer uses the schema of a collection, or you can define the schema for the collection before you import a source or target definition. The Designer flattens the schema if there is any hierarchical element in the collection and retains the original schema of the collection when you import it.

The Designer imports a document based on the schema that you set for the collection. If a document contains hierarchical elements like arrays or nested documents, the Designer imports them as columns at the same level as other columns.

For example, you need to import the collection product_details with the following schema:

```json
{
  sku: "sku_name",
  title: "product_name",
  description: "description",
  manufacture_details: {
    model_number: "model_number",
    release_date: new ISODate("date")
  },
  shipping_details: {
    weight: <value>,
    width: <value>,
    height: <value>,
    depth: <value>
  },
  quantity: <value>,
  pricing: [
    {region: "North America",
      cost_price: 1000,
      sale_price: 1200},
    {region: "Europe",
      cost_price: 1200,
      sale_price: 1500}
  ]
}
```

The Designer imports the collection schema into a tabular format. You can identify arrays and nested documents with the naming convention of the column. The naming convention of a nested document is `<top level element name>.<nested document name>.<nested document element name>`. The naming convention of an array is `<array name>.<element number>`.
The following figure shows the source definition when you import the collection into the Designer if you set the delimiter as a period (.):

<table>
<thead>
<tr>
<th>Name</th>
<th>Datatype</th>
<th>Length/...</th>
</tr>
</thead>
<tbody>
<tr>
<td>description</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>manufacture_details.model_number</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>manufacture_details.release_date</td>
<td>varchar</td>
<td>16</td>
</tr>
<tr>
<td>pricing.0.cost_price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>pricing.0.region</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>pricing.0.sale_price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>pricing.1.cost_price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>pricing.1.region</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>pricing.1.sale_price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>quantity</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>shipping_details.depth</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>shipping_details.height</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>shipping_details.weight</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>shipping_details.width</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>sku</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>title</td>
<td>varchar</td>
<td>255</td>
</tr>
</tbody>
</table>

When you run a session, the PowerCenter Integration Service uses the MongoDB ODBC data source name in the machine that runs the PowerCenter Integration Service to extract data from or load data to a MongoDB database.
CHAPTER 2

PowerExchange for MongoDB Configuration

This chapter includes the following topics:

- PowerExchange for MongoDB Configuration Overview, 13
- Prerequisites, 13
- Informatica MongoDB ODBC Driver Configuration, 14
- Data Source Name Configuration on Windows, 15

PowerExchange for MongoDB Configuration Overview

You can use PowerExchange for MongoDB on Windows or Linux. You must configure PowerExchange for MongoDB before you can extract data from or load data to MongoDB database.

Prerequisites

You must complete the prerequisites before you can use PowerExchange for MongoDB.

Complete the following prerequisites:

- Install or upgrade PowerCenter.
- Ensure that you have the PowerExchange for MongoDB license file. You do not require a separate ODBC license to use PowerExchange for MongoDB.
- On Windows, download and install the Microsoft Visual C++ 2010 Redistributable Package in server and client machines from the Microsoft website. For example, download the vc_redist_x86.exe file.

For more information about product requirements and supported platforms, see the Product Availability Matrix on the Informatica My Support Portal:
https://mysupport.informatica.com/community/my-support/product-availability-matrices
PowerExchange for MongoDB Upgrade

Before you upgrade to Informatica 10.0, back up the odbc.ini file.

After you upgrade to Informatica 10.0, replace the odbc.ini file with the backup copy of the odbc.ini file, and verify if the MongoDB driver name in the odbc.ini file is libinformicamongodbodbc64.so.

Informatica MongoDB ODBC Driver Configuration

The Informatica MongoDB ODBC driver is installed on the machines where you install Informatica services and clients. Configure the Informatica MongoDB ODBC driver on those machines.

The Designer uses the Informatica MongoDB ODBC driver to import MongoDB collections as source or target definitions. The PowerCenter Integration Service uses the driver to extract data from or load data to the MongoDB database. Create ODBC data source names to connect to the MongoDB database.

Configuring the Informatica MongoDB ODBC Driver on Linux

You must configure the Informatica MongoDB ODBC driver with details of the MongoDB database and ODBC driver manager before you can run MongoDB sessions and workflows.

Edit the odbc.ini file to configure the driver in the following location: <INFA_HOME>/tools/mongodb/Setup

1. Enter the correct ODBCInstLib for the ODBC Driver Manager in all the .ini files.
2. Replace <INSTALL_DIR> with the path to the Informatica services installation directory in all the .ini files.
3. Add the following information to the LD_LIBRARY_PATH environment variable:
   - <INFA_HOME>/tools/mongodb/lib
   - 32-bit library directory of the ODBC Driver Manager
4. Add the path of the odbc.ini file to the ODBCINI environment variable.
5. Add entries for all the MongoDB data sources in the odbc.ini file.

The following section shows a sample entry in the odbc.ini file:

```ini
[Sample Informatica MongoDB DSN]
Description=Informatica MongoDB ODBC Driver DSN
Driver=libinformaticamongodbodbc64.so
Host=[Host]
Port=[Port]
Database=[Database]
ReadPreference=primary
ReplicaSetName=""
SecondaryServers=""
UseReplicaSet=0
VirtualTableDetection=0
VTAnyMatchColumnsDetection=0
VTAnyMatchString=ANY
VTAnyMatchTableNameSuffix=any
VTArrayCountPrefix=Number of
VTHideRealTables=0
VTIndexColSuffix=index
VTInsertUpdateSafeMode=0
VTKeyColumnSeparator=. 
VTMainTableNameSeparator=main
VTMainTableNameSeparator=vt_
```
Data Source Name Configuration on Windows

Configure the connection properties, advanced properties, and schema when you configure a data source name.

You must create a data source name in the ODBC datasource administrator to extract data from and load data to a MongoDB database. The connection properties provide information for the MongoDB server and the database. The advanced properties are read and write operations. You can also define a schema after you create a database.

You can find the ODBC datasource administrator in the Control Panel on Windows. Configure the ODBC data source name in the 32-bit ODBC datasource administrator in the client and the machines where you install the Informatica services. You can access the 32-bit ODBC datasource administrator, odbcad32.exe, in 64-bit Windows from the following location: C:\Windows\SysWOW64

MongoDB ODBC Connection Properties

You must configure a MongoDB ODBC data source before you can import MongoDB data sources.

The following table describes the MongoDB ODBC connection properties:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Source Name</td>
<td>Name of the data source name.</td>
</tr>
<tr>
<td>Description</td>
<td>Description to identify a data source name.</td>
</tr>
<tr>
<td>Host</td>
<td>Host name of the MongoDB server.</td>
</tr>
<tr>
<td>Port</td>
<td>Port from which you can access MongoDB.</td>
</tr>
</tbody>
</table>
Database  
Optional. MongoDB user name.

Replica Set Name  
Optional. Name of the replica set of the database.

Additional Servers  
Optional. Host names of the secondary MongoDB servers.

### Advanced Properties

Configure the advanced properties when you create a data source name.

The following table describes the advanced properties in the Informatica MongoDB ODBC driver:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents fetched per block</td>
<td>The maximum number of documents that the PowerCenter Integration Service reads for every call to the MongoDB database. Default is 4096.</td>
</tr>
<tr>
<td>Nested column separator</td>
<td>Separator character for arrays and nested documents. The nested column separator must be consistent across connections used in a mapping. For example, if one connection uses the period (.) as the nested column separator and another connection in the same mapping uses the underscore (<em>) as the separator, then the mapping fails. You can use either the underscore (</em>) or the period (.) as the nested column separator. Default is period (.).</td>
</tr>
<tr>
<td>Maximum number of columns to flatten</td>
<td>The maximum number of array elements that the ODBC driver flattens into multiple nested columns. Default is 5.</td>
</tr>
</tbody>
</table>
| Read preference                 | Server that you prefer to read data from if you configure replica sets. You can select one of the following server options:  
- Primary. The PowerCenter Integration Service reads data from the primary server. If the primary server is offline, the session fails.  
- Primary Preferred. The PowerCenter Integration Service reads data from the primary server if the primary server is available. If the primary server is offline, the PowerCenter Integration Service reads data from the secondary server.  
- Secondary. The PowerCenter Integration Service reads data from the secondary server. If the secondary server is offline, the session fails.  
- Secondary Preferred. The PowerCenter Integration Service reads data from the secondary server if the secondary server is available. If the secondary server is offline, the PowerCenter Integration Service reads data from the primary server.  
- Nearest. The PowerCenter Integration Service reads data from the nearest available server. Default is primary. |
<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
</table>
| Sampling strategy                            | Number of rows to scan in the schema definition. You can select one of the following sampling strategies:  
- Start. Scans the specified number of rows from the start.  
- End. Scans the specified number of rows from the end.  
- Random. Scans the specified number of rows in random order.  
Default is End. |
| Documents to sample (0 to sample all documents) | Number of documents to scan.  
Default is 100. |
| String Columns Lengths                       | The string column length to use for the fields. You can select one of the following string column lengths:  
- Standard. The string column length to use for the standard fields. Default is 255.  
- Container. The string column length to use for the container fields. Default is 511.  
- DocumentAsJSON. The string column length to use for the documentAsJSON fields.  
Default is 1023. |
| Use SQL_WVARCHAR for String datatype         | The PowerCenter Integration Service maps the String datatype to SQL_WVARCHAR ODBC instead of SQL_VARCHAR.  
Default is disabled. |
| Enable reading/writing as JSON document.     | Read or write data as a JSON document. If enabled, the driver reports a special column named documentAsJSON that retrieves or stores whole documents as JSON formatted strings.  
Default is disabled.  
**Note:** For a MongoDB connection, if you toggle between enabling and disabling this option, the metadata cache might lose its integrity. Instead of changing the Enable reading/writing as JSON document property for a MongoDB connection, create separate connections with this property. |
| Show container columns when generating metadata | Show the container columns when the Integration Service generates the metadata.  
Default is disabled. |
| Enable SSL                                   | Establish secure communication to the MongoDB server over SSL. If enabled, you must specify the path where the SSL certificates for the MongoDB server are stored.  
Default is disabled. |
| Check GetLastError on writes                 | Calls the MongoDB CheckGetLastError() function to check for failures after a write operation.  
Default is enabled. |
| Enable Updating Multiple Rows                | The Informatica MongoDB ODBC driver updates multiple rows for each PowerCenter Integration Service write call.  
If enabled, the driver updates all rows that match the filter condition. If disabled, the driver updates only the first row that matches the filter condition.  
Default is disabled. |
| Omit default NULL column on insert           | The PowerCenter Integration Service does not write columns with NULL value to a MongoDB target.  
Default is enabled. |
<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truncate documents larger than 16 MB</td>
<td>Truncate the document size to 16 MB when you load data to MongoDB. Default is disabled.</td>
</tr>
<tr>
<td>Active Metadata Location</td>
<td>Read metadata changes from the MongoDB database or from a local file. Required if you choose to store the metadata in a local file. Default is database.</td>
</tr>
</tbody>
</table>
CHAPTER 3

Schema Definition

This chapter includes the following topics:

- Schema Definition Overview, 19
- Schema Editor, 19
- Virtual Tables, 21
- Metadata Caching, 25
- Defining the Schema for a Collection, 25

Schema Definition Overview

You can define the schema for a MongoDB collection that you want to import as a source or target definition in the Designer. You can define the schema for multiple collections with the same ODBC data source name.

A collection in MongoDB might contain several fields that you do not want to import. When you define the schema you can limit metadata that you import. The driver dynamically detects the collection schema of a MongoDB database. It flattens the MongoDB schema and displays the keys in the a tabular format with each key as a column in the Schema Editor.

You can export the collection to an external schema definition file and edit the schema definition in the Schema Editor. After you modify the collection properties and column metadata, you can save the modifications in the schema definition file. The driver does not modify the schema of the actual MongoDB collection. You can choose to store the modifications in the MongoDB database or as a file.

If you enable virtual table detection in the Informatica MongoDB ODBC driver, the driver creates virtual tables in the schema if the collection contains arrays. You can import the virtual table as a source or target definition in the Designer.

Schema Editor

Use the Schema Editor to view or edit the MongoDB collection schema that you want to import.

You can access the Schema Editor from the ODBC Data Source Administrator when you configure the Informatica MongoDB ODBC Driver DSN. You can also find the Schema Editor in the following location:

$INFA_HOME/clients/tools/mongodb/Tools

When you define a schema in the Informatica MongoDB ODBC driver DSN, you must specify a schema definition file. You can use an existing schema definition file or create a new one. After you specify a schema
definition file, you can import the collections in the database to the schema definition file. You can import all the collections in the database or a particular collection. You can use a JSON filter to filter records on a collection. You can also export those collections that are missing in the schema definition file.

When you open the Schema Editor, all the databases and collections in the schema definition file appear. When you select a collection, the collection properties and document properties appear on the right pane. You can modify the properties and save the schema. You can also save the schema changes to a new schema definition file.

Collection Properties

Before you import a collection, you can view or edit the properties associated with the collection in the Schema Editor.

You can view or edit the following collection properties in the Schema Editor:

**ODBC Table Name**

The name of the collection to use for the schema. Default is the same as source table name. You can modify this value to match the name that you require when you import the ODBC data source in the Designer.

**ODBC Catalog Name**

The name of the catalog that to use for the schema. Default is the same as source catalog name. You can modify this value to match the name that you require when you import the ODBC data source in the Designer.

**Source Table Name**

The name of the collection in the source database. You cannot modify this value.

**Source Catalog Name**

The name of the source database. You cannot modify this value.

**Virtual Type**

Indicates whether the collection is a virtual collection or not. Reserved for future use.

**Permissions**

The permissions assigned to you. Reserved for future use.

Column Metadata

When you select a collection in the Schema Editor, you can view or modify the column metadata of the collection.

The following fields are available in the column metadata:

**ODBC Column Name**

The name of the column that you want to use in the database schema. Default is the source column name. You can modify this value to match the name that you require when you import the ODBC data source in the Designer.

**SQL Type**

The ODBC data type of the column. The PowerCenter Integration Service uses the SQL type when you run the session that uses the ODBC data source. You can modify the datatype based on your requirement.
Source Column Name

The name of the column in the source database. You cannot modify this value.

Source Type

The data type of the column in the source database. The PowerCenter Integration Service uses the SQL type when you run the session that uses the ODBC data source. You can modify the datatype based on your requirement.

Hide Column

You can choose to hide the column so that it does not appear in the schema.

Behavior

The behavior field shows whether the column is scalar or a container. Scalar columns contain a single value like an integer or a string. Container columns have multiple values. Arrays and documents are examples of container columns.

Note: Container columns do not support transformations.

Key Type

The key type field shows whether the column is a key column.

The following values are possible for the key type:

• Primary key
• Foreign key
• Not a key

You cannot modify the key type of a column.

ODBC Type Hint

The ODBC type hint field shows the possible ODBC datatype of the column. You can choose the SQL type of a column based on the hint.

Source Nesting Level

The source nesting level field displays the level at which the column is nested in the document metadata. You can use the MongoDB ODBC driver to read up to five levels of nested columns and write up to three levels of nested columns.

Alternate Source Type

The alternate source type field displays the alternate data type of the column in the source database.

Virtual Tables

You can configure the Informatica MongoDB ODBC driver to create virtual tables in the schema if the collection contains arrays.

Virtual tables depict the normalized view of a MongoDB collection. You can import virtual tables as an ODBC data object and create mappings.

Note: You cannot use the Designer to preview virtual tables.

To configure virtual table creation, open the Informatica MongoDB ODBC Driver DSN. In the Schema Definition dialog box, click Virtual Table Options.

If you enable virtual table creation, the driver creates the following virtual tables:
Main virtual table

The main virtual table contains all the data from the original MongoDB collection except the data in the arrays. The driver replaces the cells that contain arrays with the number of arrays in the cell.

The main virtual table use the following naming convention by default: \(<original\ backquote\ collection\ name>_vt_main\)

The columns that contain arrays use the following naming convention by default: Number of \(<original\ column\ name>\)

Virtual table for array columns

The driver creates a virtual table for each column that contain arrays.

The virtual table for an array column uses the following naming convention by default: \(<original\ backquote\ collection\ name>_vt_<original\ column\ name>\)

Each virtual table has a key column that references back to the primary key column in the original collection. The key column uses the following naming convention by default: \(<original\ backquote\ collection\ name>.<primary\ key\ column\ name>\).

The virtual table has an index column that shows the position of the data within the original array. The index column uses the following naming convention by default: \(<original\ column\ name>.index\)

Other columns in the virtual table represent the elements in the array and are named after the array element. If the array is of scalar type, the data column uses the following naming convention by default: \(<original\ column\ name>.value\)

**Note:** You cannot use a DD_DELETE strategy in an Update Strategy transformation to delete rows from a virtual table. You also cannot use the MongoDB ODBC driver to add an array element to an existing array index because of a limitation from the C API used by the MongoDB driver.

Virtual Table Options

Configure the virtual table options to create virtual tables for a collection that contains arrays.

The following table describes the virtual table options in the Informatica MongoDB ODBC driver:

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enable Virtual Table Detection</td>
<td>The driver creates virtual tables if the collection contains arrays. Default is disabled.</td>
</tr>
<tr>
<td>Virtual Main Table Suffix</td>
<td>The suffix for the main virtual table. Default is main.</td>
</tr>
<tr>
<td>Virtual Key Column Separator</td>
<td>The separator for the key columns in a virtual table. You can use either the underscore (_) or the period (.) as the virtual key column separator. Default is period (.).</td>
</tr>
</tbody>
</table>

Configure the virtual table options to create virtual tables for a collection that contains arrays.
<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Table Name Separator</td>
<td>The separator in the virtual table name. Default is <em>vt</em>. <strong>Note:</strong> If tables in the MongoDB database and virtual tables have the same names, metadata import might be corrupted. To avoid importing corrupted metadata, do not use table names that contain the virtual table separator in the MongoDB database.</td>
</tr>
<tr>
<td>Virtual Table Index Column Suffix</td>
<td>The suffix for the virtual table index column. Default is index.</td>
</tr>
<tr>
<td>Hide Real Table if Virtual Tables Created</td>
<td>Hide the real tables if the corresponding virtual tables are created. Default is disabled.</td>
</tr>
<tr>
<td>Show Array Counts In Virtual Main Table</td>
<td>The virtual tables contain columns that show the array count. Default is disabled.</td>
</tr>
<tr>
<td>Virtual Table Array Count Prefix</td>
<td>The prefix for the virtual table array count column. Default is Number of.</td>
</tr>
<tr>
<td>Enable Any Match Columns Detection</td>
<td>The driver filters the data and selects rows where a value in a top-level array matches a specified expression and then returns the results as columns in a virtual table.</td>
</tr>
<tr>
<td>Any Match Table Name Prefix</td>
<td>The prefix for naming the array column in an any match virtual table.</td>
</tr>
<tr>
<td>Any Match Column Separator</td>
<td>The separator for naming the columns in an any match virtual table.</td>
</tr>
</tbody>
</table>

**Virtual Tables - An Example**

The collection CustomerTable contains arrays. You want to create virtual tables from the arrays and import the virtual tables as data objects in the Designer.

The following table shows the schema of CustomerTable collection:

<table>
<thead>
<tr>
<th>id</th>
<th>Customer Name</th>
<th>Invoices</th>
<th>Service Level</th>
<th>Contacts</th>
<th>Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>John</td>
<td>[invoice_id=123, item=toaster, price=456, discount=0.2], [invoice_id=124, item=oven, price=12345, discount=0.3]</td>
<td>Silver</td>
<td>[{type=primary, name=&quot;John Johnson&quot;}, {type=invoicing, name=&quot;Jane Johnson&quot;}]</td>
<td>[7,8]</td>
</tr>
<tr>
<td>2222</td>
<td>Jane</td>
<td>[invoice_id=125, item=blender, price=7456, discount=0.5]</td>
<td>Gold</td>
<td>[{type=primary, name=&quot;Jane Johnson&quot;}]</td>
<td>[5,6]</td>
</tr>
</tbody>
</table>

If you enable virtual table detection, the driver creates the following virtual tables:
### CustomerTable_vt_main

The following table shows the schema of CustomerTable_vt_main virtual table:

<table>
<thead>
<tr>
<th>id</th>
<th>Customer Name</th>
<th>Number of Invoices</th>
<th>Service Level</th>
<th>Number of Contacts</th>
<th>Number of Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>John</td>
<td>2</td>
<td>Silver</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2222</td>
<td>Jane</td>
<td>1</td>
<td>Gold</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

### CustomerTable_vt_Invoices

The following table shows the schema of CustomerTable_vt_Invoices virtual table:

<table>
<thead>
<tr>
<th>CustomerTable.id</th>
<th>Invoices_index</th>
<th>invoice_id</th>
<th>item</th>
<th>price</th>
<th>discount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>1</td>
<td>123</td>
<td>toaster</td>
<td>456</td>
<td>0.2</td>
</tr>
<tr>
<td>1111</td>
<td>2</td>
<td>124</td>
<td>oven</td>
<td>12345</td>
<td>0.3</td>
</tr>
<tr>
<td>2222</td>
<td>1</td>
<td>125</td>
<td>blender</td>
<td>7456</td>
<td>0.5</td>
</tr>
</tbody>
</table>

### CustomerTable_vt_Contacts

The following table shows the schema of CustomerTable_vt_Contacts virtual table:

<table>
<thead>
<tr>
<th>CustomerTable.id</th>
<th>Contacts_index</th>
<th>type</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>1</td>
<td>primary</td>
<td>John Johnson</td>
</tr>
<tr>
<td>1111</td>
<td>2</td>
<td>invoicing</td>
<td>Jane Johnson</td>
</tr>
<tr>
<td>2222</td>
<td>1</td>
<td>primary</td>
<td>Jane Johnson</td>
</tr>
</tbody>
</table>

### CustomerTable_vt_Ratings

The following table shows the schema of CustomerTable_vt_Ratings virtual table:

<table>
<thead>
<tr>
<th>CustomerTable.id</th>
<th>Ratings_index</th>
<th>Ratings_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1111</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>1111</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>2222</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2222</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
Metadata Caching

The Informatica MongoDB ODBC driver caches the schema in the MongoDB database or a flat file. After you define a schema for the collection, you can store the modifications in the MongoDB database or a file so that the Designer uses the modifications each time you import a definition.

You must modify the schema definition if there are updates to the documents that require a change in the definitions that you created in the Designer.

If you store the schema modification in a file, ensure that the file is available in the location that you configure in the ODBC data source name when you import a source or target definition. If you store the schema modification in the MongoDB database, PowerExchange for MongoDB stores the schema modification in a collection called Mersenne_Collection_Metadata. If you edit Mersenne_Collection_Metadata, you may lose the schema modifications.

**Note:** If you clear the metadata cache, you must re-create or re-import the source and target objects with the same metadata that the existing mapping objects use.

Defining the Schema for a Collection

You can modify and define the schema for a collection that you want to import as source or target definition in the Designer.

1. Open the ODBC Data Source Administrator.
2. Select the Informatica MongoDB ODBC Driver DSN.
3. Click **Configure**.
4. Click **Schema Definition**.
   The **Schema Definition** dialog box appears.
5. Click **Browse** and select a schema definition file.
   You can also enter a file name in the file selection dialog box to create and use a new schema definition file.
6. Choose one of the following collection export options to the schema definition file:
   - Export all the collections in the MongoDB database.
   - Export the tables that are missing from the schema definition file and available in the MongoDB database.
   - Select a particular collection in the MongoDB database. Optionally, you can enter a JSON filter statement to filter records.
7. Click **Launch Schema Editor**.
   The Schema Editor application appears.
8. Select a collection and define the schema in the Schema Editor according to the requirement.
9. Close the schema editor after you save the changes.
   You can also save the schema changes to a different schema definition file.
10. Select whether to store the metadata in the MongoDB database or in a local file.
11. Click **Import File** to store metadata definition from the schema definition file.
   If you read the metadata from a file instead of the MongoDB database, place the schema definition file in the same folder as the metadata file.
Updating the Schema File

You can update the schema file to reflect metadata changes in the MongoDB database or make changes in the imported metadata.

1. Open the schema definition by using the Informatica MongoDB ODBC Driver DSN.
2. Click **Browse** and select a schema definition file.
   You can also enter a file name in the file selection dialog box to create and use a new schema definition file.
3. Export the metadata to the SSD file.
   a. To export the metadata imported by using the MongoDB ODBC driver, click **Export Existing**.
   b. To export metadata sampled from the MongoDB database, click **Generate All**.
   c. To export any missing tables and add metadata, click **Generate Missing**.
4. From the **Database source table** list, select the table to be updated.
5. Click **Generate Table** to update the schema of the table from the database.
6. Click **Edit Schema File** to open the schema file that you exported.
7. In the **Schema Editor**, make the required modifications in the schema file to reflect the metadata changes.
   **Note:** When you update metadata, press **Enter** and then click **Save** to ensure that the changes to the metadata are saved.
8. Save the schema file and close the **Schema Editor** dialog box.
9. In the **Schema Definition** dialog box, click **Update Metadata** to replace the metadata with the metadata from the SSD file.
MongoDB Sources

This chapter includes the following topics:

- **MongoDB Sources Overview, 27**
- **Importing a MongoDB Source Definition, 27**
- **MongoDB Reader Sessions, 27**
- **Example: MongoDB Reader Mapping, 28**

**MongoDB Sources Overview**

You can import a MongoDB collection as an ODBC source definition in the Designer. You can configure advanced read options in the ODBC driver configuration such as the number of rows fetched in every read call.

**Importing a MongoDB Source Definition**

To import a MongoDB source definition, click **Sources > Import from Database** in the Source Analyzer and select a MongoDB ODBC data source. You can select the MongoDB collections and the Designer imports the MongoDB collections that you want to import.

**MongoDB Reader Sessions**

MongoDB reader sessions contain mappings that read data from MongoDB.

When you run a MongoDB reader session, the PowerCenter Integration Service uses the Informatica MongoDB ODBC data source to extract data from MongoDB. The MongoDB reader sessions may fail or produce incorrect results if you enable pushdown optimization in the session properties. Set pushdown optimization as none if the session fails.

You can configure advanced reader properties for the Informatica MongoDB ODBC driver in the ODBC driver properties.

You can configure the following read options in the ODBC driver properties:
Read Preference

MongoDB server that you prefer to read data from if you configure replica sets.

You can select one of the following MongoDB server options:

- **Primary.** The PowerCenter Integration Service reads data from the primary MongoDB server. If the primary MongoDB server is offline, the session fails.
- **Primary Preferred.** The PowerCenter Integration Service reads data from the primary MongoDB server if the primary MongoDB server is available. If the primary MongoDB server is offline, the PowerCenter Integration Service reads data from the secondary MongoDB server.
- **Secondary.** The PowerCenter Integration Service reads data from the secondary MongoDB server. If the secondary MongoDB server is offline, the session fails.
- **Secondary Preferred.** The PowerCenter Integration Service reads data from the secondary MongoDB server if the secondary MongoDB server is available. If the secondary MongoDB server is offline, the PowerCenter Integration Service reads data from the primary MongoDB server.
- **Nearest.** The PowerCenter Integration Service reads data from the nearest available MongoDB server.

Enable Reading/Writing as JSON

Reads the JSON format of the data from the MongoDB document. If you select the option, a column documentAsJSON appears in the collection when you read data from MongoDB through which you can read data as JSON. Default is disabled.

Documents fetched per block

The maximum number of documents fetched from the MongoDB server for every read request. If more documents are available for a query, the PowerCenter Integration Service makes further read requests to the MongoDB server. Default is 4096.

Example: MongoDB Reader Mapping

A large online music store uses MongoDB as a data warehouse to store business inventory details.

The business analysts uses a business intelligence tool that cannot read data from MongoDB. The tool requires the input data to be in a relational database or a flat file.

The data warehouse includes a collection called Music_Contents. The collection Music_Contents contains a catalog of all of the songs in the store. You must move the data in the collection to a flat file to use the data for business analysis. You must also remove those records with zero units to ensure that the data is current.

The following table describes the structure of Music_Contents:

<table>
<thead>
<tr>
<th>Field</th>
<th>Datatype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>String</td>
</tr>
<tr>
<td>Type</td>
<td>Array of strings</td>
</tr>
<tr>
<td>Artist</td>
<td>Array of strings</td>
</tr>
</tbody>
</table>
The following table describes the structure of the nested document, Price:

<table>
<thead>
<tr>
<th>Field</th>
<th>Datatype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost_Price</td>
<td>Int</td>
</tr>
<tr>
<td>Sale_Price</td>
<td>Int</td>
</tr>
</tbody>
</table>

The following document is a sample from the collection, Music_Contents:

```json
{
    "Name": "Happy Birthday",
    "type": ["Folk", "Traditional"],
    "Artist": ["Patty Hill", "Mildred J. Hill", "Derek Underhill"],
    "Units": 1000,
    "Price": {
        "Cost_Price": 1,
        "Sale_Price": 3
    }
}
```

Create a mapping with a MongoDB source definition to read the records from the collection. Include a flat file target definition in the mapping so that the business intelligence tool can consume the data. Use a Filter transformation to remove the documents that have zero units.

The following figure shows the mapping:

Music_Contents (ODBC) ➔ SQ_Music_Contents ➔ filter_mongo ➔ ff_BI_input (Flat File)

The MongoDB reader mapping contains the following components:

**MongoDB ODBC source definition**

Import the collection Store_Catalog as an ODBC source definition.
The following figure shows the source definition created from the collection:

<table>
<thead>
<tr>
<th>Name</th>
<th>Datatype</th>
<th>Length/Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artist.0</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>Artist.1</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>Artist.2</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>Name</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>Price.Cost_Price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>Price.Sale_Price</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>Units</td>
<td>integer</td>
<td>10</td>
</tr>
<tr>
<td>type.0</td>
<td>varchar</td>
<td>255</td>
</tr>
<tr>
<td>type.1</td>
<td>varchar</td>
<td>255</td>
</tr>
</tbody>
</table>

Filter transformation

The filter transformation applies a filter on the Units field and writes those records that have one or more units in the Units field.

The following figure shows the filter transformation:

Flat file target definition

The flat file target definition, ff_BI_input, contains the same columns as in the MongoDB ODBC Source Definition.
CHAPTER 5

MongoDB Targets

This chapter includes the following topics:

- MongoDB Targets Overview, 31
- Importing MongoDB Target Definitions, 31
- MongoDB Writer Sessions, 31
- Example: MongoDB Target Mapping, 32

MongoDB Targets Overview

You can import a MongoDB collection as an ODBC target definition in the Designer. You must configure the ODBC driver and define the MongoDB schema before you import MongoDB collections. You can configure advanced write options in the ODBC driver configuration such as multiple row updates when you write data to MongoDB.

Importing MongoDB Target Definitions

To import a MongoDB target definition, click Targets > Import from Database in the Target Designer and select the MongoDB ODBC data source name that you created. You can select the required MongoDB collections and the Designer imports the MongoDB collections as an ODBC target definition.

MongoDB Writer Sessions

MongoDB writer sessions contain mappings that write data to a MongoDB database.

When you run a MongoDB writer session, the PowerCenter Integration Service uses the Informatica MongoDB ODBC data source to load data to the MongoDB database. The MongoDB writer sessions may fail or produce incorrect results if you enable pushdown optimization in the session properties. Set pushdown optimization as none if the session fails.

You can configure advanced write options for the Informatica MongoDB ODBC Driver in the ODBC driver properties.
You can configure the following write options in the ODBC driver properties:

**Omit default null columns on insert**

Drops columns with null values. Default is enabled.

**Truncate documents larger than 16 MB**

Truncates a document if the size is more than 16 MB in a writer session. MongoDB documents have a size restriction of 16 MB. If enabled, the PowerCenter Integration Service truncates the document that exceeds 16 MB when writing to MongoDB. If you disable the option when you run a write session, the PowerCenter Integration Service rejects the document that exceeds 16 MB. Default is disabled.

**Enable Reading/Writing as JSON**

Writes the JSON format of the data to the MongoDB document. If you select the option, a column with the field documentAsJSON appears in the collection when you write data to MongoDB. You cannot write into individual columns if you select this option. Default is disabled.

**Enable updating multiple rows**

Updates multiple rows in the MongoDB collection for every write operation. If there are multiple documents to update, the PowerCenter Integration Service updates multiple documents in the MongoDB collection for every write operation. If you clear this option and multiple documents require update, the PowerCenter Integration Service initiates write operation for each document update. Default is disabled.

**Check GetLastError on writes**

Calls the MongoDB CheckGetLastError() function to check for failures after each insert or update operation. Select this option to include fault tolerance in write operations. Clear this option to speed up the write operation. Default is enabled.

---

**Example: MongoDB Target Mapping**

A media store uses flat files with comma-separated values to store details of the store inventory with a unique flat file for each type of media. The file FF_Music_Collection stores the details of audio CDs and FF_Movies_Collection stores the details of movie DVDs and Blu-ray disks.

You want to use a MongoDB database to store all inventory details. Create a mapping with two flat file source definitions to read the records from the flat files. Include the MongoDB target definition to write data from the flat files. Use a Joiner transformation with full outer join on the common fields to combine data in the flat file sources before writing the data to MongoDB.

The following figure shows the mapping:

![Mapping Diagram](image)

The mapping contains the following objects:
FF_Music_Data Source Definition

The following table describes the contents of FF_Music_Collection:

<table>
<thead>
<tr>
<th>Field</th>
<th>Datatype</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Artist</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Units</td>
<td>Integer</td>
<td></td>
</tr>
<tr>
<td>Cost Price</td>
<td>Integer</td>
<td></td>
</tr>
<tr>
<td>Sale Price</td>
<td>Integer</td>
<td></td>
</tr>
</tbody>
</table>

FF_Movies_Data Source Definition

The following table describes the contents of FF_Movies_Collection:

<table>
<thead>
<tr>
<th>Field</th>
<th>Datatype</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Director</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Artist1</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Artist2</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>Units</td>
<td>Integer</td>
<td></td>
</tr>
<tr>
<td>Cost Price</td>
<td>Integer</td>
<td></td>
</tr>
<tr>
<td>Sale Price</td>
<td>Integer</td>
<td></td>
</tr>
</tbody>
</table>

MDB_Inventory Target Definition

The collection MDB_Inventory stores audio CD information and movie disks information.

The following sample document shows an audio CD document in the collection:

```json
{
  "Name": "Happy Birthday",
  "Artist": ["Patty Hill", "Mildred J. Hill", "Derek Underhill"],
  "Units": 1000,
  "Price": {
    "Cost_Price": 1,
    "Sale_Price": 3
  }
}
```

The following sample document shows a movie disk document in the collection:

```json
{
  "Name": "City Lights",
  "Type": "Blu-ray",
  "Director": "Charlie Chaplin"
}
"Artist" : ["Charle Chaplin", "Mildred J. Hill", "Derek Underhill"],
"Units" : 1000,
"Price" : {
    "Cost_Price" : 10,
    "Sale_Price" : 15
}

The following figure shows the target definition that you import in the Designer:
Datatype Reference

This appendix includes the following topic:
- MongoDB, ODBC, and Transformation Datatypes, 35

MongoDB, ODBC, and Transformation Datatypes

When you define the schema in the Informatica MongoDB ODBC driver, you can view the ODBC datatypes and edit the datatypes. When you import a MongoDB collection as a source or target definition, the transformation datatypes corresponding to the ODBC datatypes appear in the Designer.

The Informatica MongoDB ODBC driver reads MongoDB data and converts the MongoDB datatypes to ODBC datatypes. The PowerCenter Integration Service converts the ODBC datatypes to transformation datatypes.

The following table lists the MongoDB datatypes and the corresponding ODBC and transformation datatypes:

<table>
<thead>
<tr>
<th>MongoDB Datatypes</th>
<th>ODBC Datatypes</th>
<th>Transformation Datatypes</th>
<th>Range and Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>String</td>
<td>Varchar</td>
<td>String</td>
<td>1 to 104,857,600 characters</td>
</tr>
<tr>
<td>Boolean</td>
<td>Bit</td>
<td>String</td>
<td>Precision of 1</td>
</tr>
<tr>
<td>NumberLong</td>
<td>BigInt</td>
<td>Decimal</td>
<td>Precision 1 to 28 digits, scale 0 to 28</td>
</tr>
<tr>
<td>NumberInt</td>
<td>Int</td>
<td>Integer</td>
<td>Precision 10, scale 0</td>
</tr>
<tr>
<td>NumberDouble</td>
<td>Double</td>
<td>Double</td>
<td>Precision 15</td>
</tr>
<tr>
<td>BinData</td>
<td>Binary</td>
<td>Binary</td>
<td>1 to 104,857,600 bytes</td>
</tr>
<tr>
<td>Date</td>
<td>Timestamp</td>
<td>Date/Time</td>
<td>Jan 1, 0001 A.D. to Dec 31, 9999 A.D. (precision to second)</td>
</tr>
<tr>
<td>jstOID</td>
<td>Varchar</td>
<td>String</td>
<td>1 to 104,857,600 characters</td>
</tr>
</tbody>
</table>
INDEX

I
importing
   targets 31
Introduction
   MongoDB 10
   PowerExchange for MongoDB 10

O
overview
   targets 31

R
read property
   Enable Reading/Writing as JSON 27

read property (continued)
   Read Preference 27
      Rows fetched per block 27

T
targets
   importing 31

W
write options
   Truncate documents larger than 16 MB 31
   Enable readng/writing as JSON 31
   Enable updating multiple row 31
   Omit default null columns on insert 31