



TIBCO® Data Virtualization

HBase Adapter Guide

Version 8.7.0 | October 2023

Contents

Contents	2
TDV HBase Adapter	3
Introduction	3
Datasource Configuration	3
Basic Tab	3
Advanced Tab	6
Obtain and Install Drivers for HBase	17
Data Type Mappings	18
HBase to TDV Data Types	18
HBase Function Support	19
HBase Aggregate Function Support	20
HBase Conversion Function Support	20
HBase Date Function Support	21
HBase Numeric Function Support	21
HBase Specific Properties	21
HBase String Function Support	22
References	22
TIBCO Product Documentation and Support Services	24
How to Access TIBCO Documentation	24
How to Contact TIBCO Support	25
Release Version Support	25
How to Join TIBCO Community	26
Legal and Third-Party Notices	27

TDV HBase Adapter

Introduction

This section explains the various connection and configuration options of the datasource HBase as well as the capabilities:

[Datasource Configuration](#)

[Obtain and Install Drivers for HBase](#)

[Data Type Mappings](#)

[HBase Function Support](#)

[References](#)

Datasource Configuration

This section explains the connection properties that are defined while setting up a datasource.

Basic Tab

The following table and the sections below lists and explains the connection properties that are in the Basic Tab of the New Data Source Window.

Datasource Name	The name of the Datasource.
Host	Name of the host machine or the host machine's IP address.

Port	Port number for the data source to connect with the host. The Port number for HBase is 2181.
Database Name	Name or alias of the underlying data source. TDV Server uses this name to find and connect to the data source.
Login/User, Password	User name and password required to access the data source.
Pass-through Login	Flag to indicate whether pass-through login is enabled or not.

Datasource Name

The name of the data source.

Data Type

string

Default Value

""

Host

Name of the host machine or the host machine's IP address.

Data Type

string

Default Value

""

Port

The Port number

Data Type

string

Default Value

""

Database Name

Name or alias of the underlying data source. TDV Server uses this name to find and connect to the data source.

Data Type

string

Default Value

""

Login/User, Password

User name and password required to access the data source.

Data Type

string

Default Value

""

Remarks

When the data source is used as a target for cache tables or data ship, the user must also have permission to create tables, execute DDL, and perform other required tasks. Refer to the individual data source descriptions for details.

Pass-through Login

Flag to indicate whether pass-through login is enabled or not.

Data Type

string

Default Value

""

Remarks

Disabled (default)—This allows automated provisioning of a connection pool. Open connection threads can be used by authorized users after the validation query verifies connection status. If pass-through login is disabled, the Save Password check box is not available.

Enabled—A new connection to the data source uses the credentials supplied by the client when data is requested from that data source for the first time. Subsequent requests by the same user reuse the existing connection. When another user attempts to connect to a data source, a new connection is created.

See “Managing Security for TDV Resources” in the TDV Administration Guide for details.

Advanced Tab

The following table and the sections below lists and explains the connection properties that are in the Advanced Tab of the New Data Source Window.

Connection URL Pattern	A template for generating a URL to connect to the physical data source.
Connection URL String	The URL string generated from the connection URL pattern with the connection information you provide.
JDBC Connection Properties	Lets you specify property-value pairs to pass to the JDBC data source
Connection Pool Minimum Size	Minimum number of connections in the pool even when the pool is inactive.
Connection Pool Maximum Size	Maximum number of connections (both active and idle) allowed for the data source. When the maximum is reached, new requests must wait until a connection is available.
Connection Pool Idle Timeout	Number of seconds that a connection can remain idle without being dropped from the pool when there are more than the minimum number of connections.
Maximum Connection Lifetime	The number of minutes that a connection that was returned to the pool persists if there are more open connections than the minimum pool size.
Connection Validation Query	A data-source-specific query that the TDV query engine sends to see if the data source connection is valid. This query is executed every time a connection is checked out from the pool. Enter a query that returns quickly.
Execution Timeout	The number of seconds an execution query on the data source can run before being canceled.
Execute SELECTs Independently	Lets a SELECT statement be executed using a new connection from the connection pool, and committed immediately after completion. INSERT, UPDATE, and DELETE statements are executed using the same connection as part of the transaction.

Connection Check-out Procedure	A procedure that returns a valid SQL statement that can be used to initialize the connection.
Supports Star Schema	Check only if this data source supports very large predicates and very large cardinalities for star schema semijoins.
Max Source Side of Semi Join To Use OR Syntax	See the documentation for semijoins and the TDV Administration Guide for more information.
Min Target to Source Ratio for Semi Join	Sets the minimum target-to-source ratio of cardinality for semijoins. Refer to the TDV Administration Guide for more information.
Max Source Side Cardinality for Semi Join	See the documentation for semijoins and the TDV Administration Guide for more information.
Enable Native Data Loading	Lets the data source use its proprietary functionality to optimize performance.
Collation Sensitive	TDV does not use the SORT MERGE join algorithm if any data source involved in the join is marked Collation Sensitive.
Concurrent Request Limit	Works with the Massively Parallel Processing engine configuration parameters to control the amount of parallelization for the queries for a particular data source.

Connection URL Pattern

A template for generating a URL to connect to the physical data source.

Data Type

string

Default Value

jdbc:<DATA SOURCE>//<HOST>:<PORT>/<DATABASE_NAME>

Remarks

TDV does not validate modifications at the time of configuration. The data source adapter might not validate changes.

Connection URL String

The URL string generated from the connection URL pattern with the connection information you provide.

Data Type

string

Default Value

“”

Remarks

This string is used by the JDBC adapter to connect to the physical data source. This field cannot be edited. For details, see the section “Connecting through JDBC Adapters” in the *TDV Administration Guide*.

JDBC Connection Properties

Lets you specify property-value pairs to pass to the JDBC data source.

Data Type

string

Default Value

“”

Remarks

Click to add custom connection properties for any JDBC data source. Commonly used properties are populated with default values. Use the Add Argument button to specify other properties and values.

TDV does not validate property names. Some data source adapters ignore invalid property names or values; others return an error.

The driver properties specify connection timeout settings required by specific drivers. To avoid leaving connections open indefinitely, specify properties explicitly for your data source.

Connection Pool Maximum Size

Maximum number of connections (both active and idle) allowed for the data source. When the maximum is reached, new requests must wait until a connection is available.

Data Type

Numeric

Default Value

100

Remarks

If the maximum number of connections is in use when a request comes in (even with pass-through authentication), the new request is blocked and queued until a connection is available or the Connection Pool Idle Timeout is reached.

If no connection was made available within the specified timeout, a check is made for an available connection by the same user. If none is available, the least recently used connection for another user is dropped and a new connection is opened.

Studio reuses pooled connections if they continue to be valid after changes (such as connection name), but JDBC requests are forced to use new connections if any part of the data source connection configuration has changed.

Connection Pool Minimum Size

Minimum number of connections in the pool even when the pool is inactive.

Data Type

Numeric

Default Value

0

Remarks

When a connection has been idle, a validation query is used to verify whether an open connection is still valid just prior to submission of a request. If the connection is invalid, the connection is discarded and another is used.

Connection Pool Idle Timeout

Number of seconds that a connection can remain idle without being dropped from the pool when there are more than the minimum number of connections.

Data Type

Numeric

Default Value

30

Maximum Connection Lifetime

The number of minutes that a connection that was returned to the pool persists if there are more open connections than the minimum pool size.

Data Type

Numeric

Default Value

30

Remarks

The duration is calculated from connection creation. Default value is 60 minutes. Set a smaller value if the pool is likely to run out of connections. Be sure to add a validation query. Set a larger value if you want the connections to be held for a longer period. Set a value of 0 to keep connections alive indefinitely.

Connection Validation Query

A data-source-specific query that the TDV query engine sends to see if the data source connection is valid. This query is executed every time a connection is checked out from the pool. Enter a query that returns quickly.

Data Type

string

Default Value

""

Remarks

If this query returns a non-error result, the data source connection is considered valid. If this query fails, the connection is discarded and a new connection is checked out from the available pool.

No one SELECT statement works with all data sources. To verify that TDV is running and that it can connect to the data source, devise a query against a published table from that data source.

Enable Native Data Loading

Lets the data source use its proprietary functionality to optimize performance.

Data Type

Bool

Default Value

True

Remarks

See the User Guide, Chapter About Data Source Native Load Performance Options” for more details,.

Collation Sensitive

TDV does not use the SORT MERGE join algorithm if any data source involved in the join is marked Collation Sensitive.

Data Type

Bool

Default Value

False

Remarks

None

Concurrent Request Limit

Works with the Massively Parallel Processing engine configuration parameters to control the amount of parallelization for the queries for a particular data source.

Data Type

Numeric

Default Value

0

Remarks

None

Execution Timeout

The number of seconds an execution query on the data source can run before being canceled.

Data Type

Numeric

Default Value

0

Remarks

None

Execute SELECTs Independently

Lets a SELECT statement be executed using a new connection from the connection pool, and committed immediately after completion. INSERT and UPDATE statements are executed using the same connection as part of the transaction.

Data Type

Bool

Default Value

True

Remarks

None

Connection Check-out Procedure

A procedure that returns a valid SQL statement that can be used to initialize the connection.

Data Type

string

Default Value

""

Remarks

The signature of the initialization procedure should be:

```
(IN ds_name VARCHAR, OUT sqlText VARCHAR)
```

Give the full path to the procedure in the Connection Check-out Procedure box.

Max Source Side Cardinality for Semi Join

See the documentation for semijoins and the TDV Administration Guide for more information.

Data Type

Numeric

Default Value

""

Remarks

None

Max Source Side of Semi Join To Use OR Syntax

See the documentation for semijoins and the TDV Administration Guide for more information.

Data Type

Numeric

Default Value

2147483647

Remarks

None

Min Target to Source Ratio for Semi Join

Sets a minimum ratio to trigger use of semi join optimization.

Data Type

Numeric

Default Value

""

Remarks

None

Supports Star Schema

Check only if this data source supports very large predicates and very large cardinalities for star schema semijoins.

Data Type

Bool

Default Value

False

Remarks

Refer to the section Star Schema Semijoin in the User Guide, for more information.

Obtain and Install Drivers for HBase

The Phoenix JDBC driver enables tables or views created through Phoenix to be available to TDV. Tables in HBase do not automatically get mapped to Phoenix. Refer to Apache Phoenix documentation for how to map existing HBase tables to Phoenix views and tables.

To install and configure JDBC drivers for HBase

1. From the Apache Phoenix download page, determine which version of Phoenix is compatible with your version of HBase. For example, Phoenix 4.1.0,.
2. Download the <filename>-bin.tar.gz file.
3. Unpack the zip file and copy the following JAR files:

```
phoenix-<ver>-client-hadoop2.jar
```

4. Paste the necessary client file into the following TDV directory:

```
<TDV_install_dir>\conf\adapters\system\hbase_0_98_apache_phoenix_driver
```

5. Add the Phoenix server JAR file to the HBase classpath on all region servers. See Apache Phoenix documentation for instructions on how to manage the classpath.
6. Restart the TDV Server.

Refer to the Installation guide for more details about the supported versions.

Data Type Mappings

HBase to TDV Data Types

HBase is Apache's Hadoop database. This section provides the data type mappings from HBase to TDV data types.

Note: The maximum CAST function is 2,147,483,647

HBase Data Type	TDV Data Type
BIGINT	BIGINT
BINARY	BINARY
BLOB	BLOB
BOOLEAN	BOOLEAN
CHAR	CHAR
CHARACTER	CHAR
CHARACTER_LARGE_OBJECT	LONGVARCHAR
CHARACTER_VARYING	VARCHAR

HBase Data Type	TDV Data Type
CLOB	CLOB
DATE	DATE
DECIMAL	DECIMAL
DOUBLE	DOUBLE
FLOAT	DOUBLE
INT	INTEGER
LONGVARCHAR	LONGVARCHAR
NUMERIC	DECIMAL
REAL	DOUBLE
SMALLINT	SMALLINT
TIME	TIME
TIMESTAMP	TIMESTAMP
TINYINT	TINYINT
VARBINARY	VARBINARY
VARCHAR	VARCHAR

HBase Function Support

TDV supports the following types of functions for HBase:

- [HBase Aggregate Function Support](#)
- [HBase Conversion Function Support](#)

- [HBase Date Function Support](#)
- [HBase Numeric Function Support](#)
- [HBase String Function Support](#)

HBase Aggregate Function Support

TDV supports the aggregate functions listed in the table below for HBase.

HBase Aggregate Function	Notes
AVG	
COUNT	
MAX	
MIN	
STDDEV_POP	
STDDEV_SAMP	
SUM	

HBase Conversion Function Support

TDV supports the conversion functions listed in the table below for HBase.

HBase Conversion Function	Notes
CAST	
TO_CHAR	

HBase Date Function Support

TDV supports the date functions listed in the table below for HBase.

HBase Date Function	Notes
COALESCE	
CURRENT_DATE	
CURRENT_TIME	

HBase Numeric Function Support

TDV supports the numeric functions listed in the table below for HBase.

HBase Numeric Function	Notes
CEIL	
CEILING	
FLOOR	
ROUND	

HBase Specific Properties

This section describes the connection properties that are specific to the HBase data source.

Port	Port number for the data source to connect with the host. The Port number for HBase is 2181.
------	---

HBase String Function Support

TDV supports the string functions listed in the table below for HBase.

HBase String Function	Notes
CONCAT	
LCASE	
LENGTH	
LOWER	
LTRIM	
RTRIM	
SUBSTR	
UCASE	
UPPER	

References

Refer to the following Guides for further details about the capabilities of the data source:

Capabilities	Section
Query Engine	User Guide, Chapter <i>TDV Query Engine Optimizations</i>
Data ship	User Guide, Chapter <i>Data Ship Performance Optimization</i>

Capabilities	Section
Caching	User Guide, Chapter <i>TDV Caching</i>
Performance Optimization	User Guide, Chapter <i>Performance Tuning</i>
TDV Massively Parallel Processing Engine	User Guide, Chapter <i>Configuring the TDV MPP Engine</i>
Kerberos	Administration Guide Chapter <i>Configuring Kerberos</i>

TIBCO Product Documentation and Support Services

For information about this product, you can read the documentation, contact TIBCO Support, and join the TIBCO Community.

How to Access TIBCO Documentation

Documentation for TIBCO products is available on the [TIBCO Product Documentation](#) website, mainly in HTML and PDF formats.

The [TIBCO Product Documentation](#) website is updated frequently and is more current than any other documentation included with the product.

Product-Specific Documentation

The following documentation for this product is available on the [TIBCO® Data Virtualization](#) page.

- **Users**
 - TDV Getting Started Guide
 - TDV User Guide
 - TDV Web UI User Guide
 - TDV Client Interfaces Guide
 - TDV Tutorial Guide
 - TDV Northbay Example
- **Administration**
 - TDV Installation and Upgrade Guide
 - TDV Administration Guide
 - TDV Active Cluster Guide
 - TDV Security Features Guide
- **Data Sources**

TDV Adapter Guides

TDV Data Source Toolkit Guide (Formerly Extensibility Guide)

- **References**

TDV Reference Guide

TDV Application Programming Interface Guide

- **Other**

TDV Business Directory Guide

TDV Discovery Guide

- *TIBCO TDV and Business Directory Release Notes* Read the release notes for a list of new and changed features. This document also contains lists of known issues and closed issues for this release.

How to Contact TIBCO Support

Get an overview of [TIBCO Support](#). You can contact TIBCO Support in the following ways:

- For accessing the Support Knowledge Base and getting personalized content about products you are interested in, visit the [TIBCO Support](#) website.
- For creating a Support case, you must have a valid maintenance or support contract with TIBCO. You also need a user name and password to log in to [TIBCO Support](#) website. If you do not have a user name, you can request one by clicking **Register** on the website.

Release Version Support

TDV 8.5 is designated as a Long Term Support (LTS) version. Some release versions of TIBCO® Data Virtualization products are selected to be long-term support (LTS) versions. Defect corrections will typically be delivered in a new release version and as hotfixes or service packs to one or more LTS versions. See also

https://docs.tibco.com/pub/tdv/general/LTS/tdv_LTS_releases.htm.

How to Join TIBCO Community

TIBCO Community is the official channel for TIBCO customers, partners, and employee subject matter experts to share and access their collective experience. TIBCO Community offers access to Q&A forums, product wikis, and best practices. It also offers access to extensions, adapters, solution accelerators, and tools that extend and enable customers to gain full value from TIBCO products. In addition, users can submit and vote on feature requests from within the [TIBCO Ideas Portal](#). For a free registration, visit [TIBCO Community](#).

Legal and Third-Party Notices

SOME TIBCO SOFTWARE EMBEDS OR BUNDLES OTHER TIBCO SOFTWARE. USE OF SUCH EMBEDDED OR BUNDLED TIBCO SOFTWARE IS SOLELY TO ENABLE THE FUNCTIONALITY (OR PROVIDE LIMITED ADD-ON FUNCTIONALITY) OF THE LICENSED TIBCO SOFTWARE. THE EMBEDDED OR BUNDLED SOFTWARE IS NOT LICENSED TO BE USED OR ACCESSED BY ANY OTHER TIBCO SOFTWARE OR FOR ANY OTHER PURPOSE.

USE OF TIBCO SOFTWARE AND THIS DOCUMENT IS SUBJECT TO THE TERMS AND CONDITIONS OF A LICENSE AGREEMENT FOUND IN EITHER A SEPARATELY EXECUTED SOFTWARE LICENSE AGREEMENT, OR, IF THERE IS NO SUCH SEPARATE AGREEMENT, THE CLICKWRAP END USER LICENSE AGREEMENT WHICH IS DISPLAYED DURING DOWNLOAD OR INSTALLATION OF THE SOFTWARE (AND WHICH IS DUPLICATED IN THE LICENSE FILE) OR IF THERE IS NO SUCH SOFTWARE LICENSE AGREEMENT OR CLICKWRAP END USER LICENSE AGREEMENT, THE LICENSE(S) LOCATED IN THE “LICENSE” FILE(S) OF THE SOFTWARE. USE OF THIS DOCUMENT IS SUBJECT TO THOSE TERMS AND CONDITIONS, AND YOUR USE HEREOF SHALL CONSTITUTE ACCEPTANCE OF AND AN AGREEMENT TO BE BOUND BY THE SAME.

This document is subject to U.S. and international copyright laws and treaties. No part of this document may be reproduced in any form without the written authorization of TIBCO Software Inc.

TIBCO, TIBCO logo, TIBCO O logo, ActiveSpaces, Enterprise Messaging Service, Spotfire, TERR, S-PLUS, and S+ are either registered trademarks or trademarks of TIBCO Software Inc. in the United States and/or other countries.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle Corporation and/or its affiliates.

All other product and company names and marks mentioned in this document are the property of their respective owners and are mentioned for identification purposes only.

This software may be available on multiple operating systems. However, not all operating system platforms for a specific software version are released at the same time. See the

readme file for the availability of this software version on a specific operating system platform.

THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT.

THIS DOCUMENT COULD INCLUDE TECHNICAL INACCURACIES OR TYPOGRAPHICAL ERRORS. CHANGES ARE PERIODICALLY ADDED TO THE INFORMATION HEREIN; THESE CHANGES WILL BE INCORPORATED IN NEW EDITIONS OF THIS DOCUMENT. TIBCO SOFTWARE INC. MAY MAKE IMPROVEMENTS AND/OR CHANGES IN THE PRODUCT(S) AND/OR THE PROGRAM(S) DESCRIBED IN THIS DOCUMENT AT ANY TIME.

THE CONTENTS OF THIS DOCUMENT MAY BE MODIFIED AND/OR QUALIFIED, DIRECTLY OR INDIRECTLY, BY OTHER DOCUMENTATION WHICH ACCOMPANIES THIS SOFTWARE, INCLUDING BUT NOT LIMITED TO ANY RELEASE NOTES AND "READ ME" FILES.

This and other products of TIBCO Software Inc. may be covered by registered patents. Please refer to TIBCO's Virtual Patent Marking document (<https://www.tibco.com/patents>) for details.

Copyright © 2002-2023 Cloud Software Group, Inc All Rights Reserved.