

Hive ODBC Connector

This page contains details about setting up and using the ODBC Connector for Hive. This page contains the following topics:

- [Before You Begin](#)
 - [The SQL Connector](#)
- [Hive ODBC Connector on Linux](#)
 - [System Requirements](#)
 - [Install the Hive ODBC Connector on Linux](#)
 - [Configure the Hive ODBC Connector Driver on Linux](#)
 - [Configure ODBC Connections on Linux](#)
- [Hive ODBC Connector on Windows](#)
 - [Software and Hardware Requirements](#)
 - [Install the Hive ODBC Connector on Windows](#)
 - [Configure Hive ODBC Connections on Windows](#)
- [Hive ODBC Connector on Mac OS X](#)
 - [System Requirements](#)
 - [Install the Hive ODBC Connector on Mac OS X](#)
 - [Configure the Hive ODBC Connector Driver on Mac OS X](#)
 - [Configure Hive ODBC Connections on Mac OS X](#)

Before You Begin

The MapR Hive ODBC Connector is an ODBC driver for Apache Hive that complies with the ODBC 3.52 specification.

To use the ODBC driver, configure a *Data Source Name* (DSN), a definition that specifies how to connect to Hive. DSNs are typically managed by the operating system and may be used by multiple applications. Some applications do not use DSNs. You will need to refer to your particular application's documentation to understand how it connects using ODBC.

The standard query language for ODBC is SQL. HiveQL, the standard query language for Hive, includes a subset of ANSI SQL-92. Applications that connect to Hive using ODBC may need queries altered if the queries use SQL features that are not present in Hive. Applications that use SQL will recognize HiveQL, but might not provide access to HiveQL-specific features such as multi-table insert. Please refer to the [Hive Language Manual](#) for up-to-date information on HiveQL.

The SQL Connector

The SQL Connector feature translates standard SQL-92 queries into equivalent HiveQL queries. The SQL Connector performs syntactical translations and structural transformations. For example:

- **Quoted Identifiers:** When quoting identifiers, HiveQL uses back quotes (` `), while SQL uses double quotes (" "). Even when a driver reports the back quote as the quote character, some applications still generate double-quoted identifiers.
- **Table Aliases:** HiveQL does not support the AS keyword between a table reference and its alias.
- The JOIN, INNER JOIN, and CROSS JOIN SQL syntaxes are translated to the HiveQL JOIN syntax.
- SQL TOP N queries are transformed to HiveQL LIMIT queries.

Hive ODBC Connector on Linux

System Requirements

- The 32-bit and 64-bit version of the following operating systems:
 - Red Hat® Enterprise Linux® (RHEL) 5.0 or 6.0
 - CentOS 5.0 or 6.0
 - SUSE Linux EnterpriseServer (SLES) 11
- 45 MB of available disk space.
- An installed ODBC driver manager:
 - iODBC 3.52.7 or above
 - OR
 - unixODBC 2.2.12 or above

The MapR ODBC Driver with SQL Connector for Apache Hive requires a Hadoop cluster with the Hive service installed and running. The MapR ODBC Driver with SQL Connector for Apache Hive is suitable for use with all versions of Hive. Download the ODBC connector from the following locations:

- Red Hat 5 32-bit and 64-bit:

- http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/Linux/EL5/
- Red Hat 6 32-bit and 64-bit:
http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/Linux/EL6/
- SuSE 11 32-bit and 64-bit:
http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/Linux/SLES11/

Note: You must update the version placeholders in the provided download locations.



For Hive 1.2.1, you must install version 2.1.8 or above. The 2.1.8 version also supports Hive 0.13 and Hive 1.0.

Install the Hive ODBC Connector on Linux

The MapR ODBC Driver with SQL Connector for Apache Hive driver files are installed in the following directories:

- `/opt/mapr/hiveodbc/ErrorMessage` – Error messages files directory
- `/opt/mapr/hiveodbc/Setup` – Sample configuration files directory
- `/opt/mapr/hiveodbc/lib/32` – 32-bit shared libraries directory
- `/opt/mapr/hiveodbc/lib/64` – 64-bit shared libraries directory

To install the MapR ODBC Driver with SQL Connector for Apache Hive:

1. Log in as the root user.
2. Navigate to the folder containing the packages for installation, and type the command corresponding to your Linux distribution:
 - Red Hat Enterprise Linux 5.0 or CentOS 5.0: `yum --nogpgcheck localinstall <RPMFileName>`
 - SUSE Linux Enterprise Server 1: `zypper install <RPMFileName>`

The MapR ODBC Driver with SQL Connector for Apache Hive depends on the following resources:

- `cyrus-sasl-2.1.22-7` or above
- `cyrus-sasl-gssapi-2.1.22-7` or above
- `cyrus-sasl-plain-2.1.22-7` or above

If the package manager in your Linux distribution cannot resolve the dependencies automatically when installing the driver, then download and manually install the packages required by the version of the driver that you want to install.

Configure the Hive ODBC Connector Driver on Linux

Setting the LD_LIBRARY_PATH Environment Variable

The LD_LIBRARY_PATH environment variable must include the paths to:

- Installed ODBC driver manager libraries
- Installed The MapR ODBC Driver with SQL Connector for Apache Hive shared libraries

Important: The Linux version of the driver bundles together functionality for both 32-bit and 64-bit environments. Do not include the paths to both 32- and 64-bit shared libraries in LD_LIBRARY_PATH at the same time. Only include the path to the shared libraries corresponding to the driver matching the bitness of the client application used. For example, if you are using a 64-bit client application and ODBC driver manager libraries are installed in `/usr/local/lib`, then set LD_LIBRARY_PATH as follows:

```
export LD_LIBRARY_PATH=/usr/local/lib/opt/mapr/hiveodbc/lib/64
```

Refer to your Linux shell documentation for details on how to set environment variables permanently.

Configure ODBC Connections on Linux

Files

ODBC driver managers use configuration files to define and configure ODBC data sources and drivers. By default, the following configuration files residing in the user's home directory are used:

- `.odbc.ini` – The file used to define ODBC data sources (required)
- `.odbcinst.ini` – The file used to define ODBC drivers (optional)
- `.mapr.hiveodbc.ini` – The file used to configure the MapR ODBC Driver with SQL Connector for Apache Hive (required)

Sample Files

The driver installation contains the following sample configuration files in the Setup directory:

- odbc.ini
- odbcinst.ini
- mapr.hiveodbc.ini

The names of the sample configuration files do not begin with a period (.) so that they will appear in directory listings by default. A filename beginning with a period (.) is hidden. For odbc.ini and odbcinst.ini, if the default location is used, then the filenames must begin with a period (.). For mapr.hiveodbc.ini, the filename must begin with a period (.) and must reside in the user's home directory.

If the configuration files do not already exist in the user's home directory, then the sample configuration files can be copied to that directory and renamed. If the configuration files already exist in the user's home directory, then the sample configuration files should be used as a guide for modifying the existing configuration files.

Configuring the Environment

By default, the configuration files reside in the user's home directory. However, two environment variables, ODBCINI and ODBCYSINI, can be used to specify different locations for the odbc.ini and odbcinst.ini configuration files. Set ODBCINI to point to your odbc.ini file. Set ODBCYSINI to point to the directory containing the odbcinst.ini file. For example, if your odbc.ini file is located in /etc and your odbcinst.ini file is located in /usr/local/odbc, then set the environment variables as follows:

```
export ODBCINI=/etc/odbc.ini
export ODBCYSINI=/usr/local/odbc
```

For version 2.1.8 and above, you must also set MAPRHIVEINI to point to the mapr.hiveodbc.ini file in the user's home directory:

```
export MAPRHIVEINI=/etc/.mapr.hiveodbc.ini
```

For version 2.1.5 and below, you set MAPRINI to point to the mapr.hiveodbc.ini file in the user's home directory:

```
export MAPRINI=/<user_home>/mapr.hiveodbc.ini
```

Configuring the odbc.ini File

ODBC Data Sources are defined in the odbc.ini configuration file. The file is divided into several sections:

- [ODBC] is optional and used to control global ODBC configuration, such as ODBC tracing.
- [ODBC Data Sources] is required, listing DSNs and associating DSNs with a driver.
- A section having the same name as the data source specified in the [ODBC Data Sources] section is required to configure the data source.

Here is an example odbc.ini configuration file for Linux:

```
[ODBC Data Sources]
Sample MapR Hive DSN 32=MapR Hive ODBC Driver 32-bit
[Sample MapR Hive DSN 32]
Driver=/opt/mapr/hiveodbc/lib/32/libmaprhiveodbc32.so
HOST=MyHiveServer
PORT=10000
```

To create a data source:

1. Open the .odbc.ini configuration file in a text editor.
2. Add a new entry to the [ODBC Data Sources] section. Type the data source name (DSN) and the driver name.
3. To set configuration options, add a new section having a name matching the data source name (DSN) you specified in step 2. Specify configuration options as keyvalue pairs.
4. Save the .odbc.ini configuration file.

Note: You can set configuration options in your odbc.ini and .mapr.hiveodbc.ini files. Configuration options set in a .mapr.hiveodbc.ini file apply to all connections, whereas configuration options set in an odbc.ini file are specific to a connection. Configuration options set in odbc.ini take precedence over configuration options set in .mapr.hiveodbc.ini

Configuring the odbcinst.ini File

ODBC Drivers are defined in the odbcinst.ini configuration file. The configuration file is optional because drivers can be specified directly in the odbc.ini configuration file. The odbcinst.ini file is divided into the following sections:

- [ODBC Drivers] lists the names of all the installed ODBC drivers.

- section having the same name as the driver name specified in the [ODBC Drivers] section lists driver attributes and values.

Here is an example odbcinst.ini file for Linux:

```
[ODBC Drivers]
Mapr Hive ODBC Driver 32-bit=Installed
Mapr Hive ODBC Driver 64-bit=Installed
[Mapr Hive ODBC Driver 32-bit]
Description=Mapr Hive ODBC Driver (32-bit)
Driver=/opt/mapr/hiveodbc/lib/32/libmaprhiveodbc32.so
[Mapr Hive ODBC Driver 64-bit]
Description=Mapr Hive ODBC Driver (64-bit)
Driver=/opt/mapr/hiveodbc/lib/64/libmaprhiveodbc64.so
```

To define a driver:

1. Open the .odbcinst.ini configuration file in a text editor.
2. Add a new entry to the [ODBC Drivers] section. Type the driver name, and then type **=Installed**
Note: Assign the driver name as the value of the Driver attribute in the data source definition instead of the driver shared library name.
3. In .odbcinst.ini, add a new section having a name matching the driver name you typed in step 2, and then add configuration options to the section based on the sample odbcinst.ini file provided with the MapR ODBC Driver with SQL Connector for Apache Hive in the Setup directory. Specify configuration options as keyvalue pairs.
4. Save the .odbcinst.ini configuration file.

Configuring the mapr.hiveodbc.ini File

To configure the MapR ODBC Driver with SQL Connector for Apache Hive to work with your ODBC driver manager:

1. Open the .mapr.hiveodbc.ini configuration file in a text editor.
2. Edit the DriverManagerEncoding setting. The value usually must be UTF-16 or UTF-32, depending on the ODBC driver manager you use. iODBC uses UTF-32 and unixODBC uses UTF-16. Consult your ODBC Driver Manager documentation for the correct setting to use.
3. Edit the ODBCInstLib setting. The value is the name of the ODBCInst shared library for the ODBC driver manager you use. The configuration file defaults to the shared library for iODBC. In Linux, the shared library name for iODBC is `libiodbcinst.so`. Note: Consult your ODBC driver manager documentation for the correct library to specify. You can specify an absolute or relative filename for the library. If you intend to use the relative filename, then the path to the library must be included in the library path environment variable. In Linux, the library path environment variable is named `LD_LIBRARY_PATH`.
4. Save the .mapr.hiveodbc.ini configuration file.

Configure Authentication for ODBC Connections on Linux

You can configure the following types of authentication:

- No authentication
- User name
- User name and password
- Kerberos

When `hive.server2.authentication` is set to `KERBEROS`, then you must configure your connection to use Kerberos.

To find out the authentication setting your Hive Server 2 is set to use, review the following properties in the `hive-site.xml` file:

- `hive.server2.authentication`
- `hive.server2.enable.doAs`

Using No Authentication

To use no authentication:

- Set the `AuthMech` configuration key for the DSN to 0

Using User Name

To configure User Name authentication:

1. Set the AuthMech configuration key for the DSN to 2
2. Set the UID key to the appropriate user name recognized by the Hive server.

Using User Name and Password

To configure User Name and Password authentication:

1. Set the AuthMech configuration key for the DSN to 3
2. Set the UID key to the appropriate user name recognized by the Hive server.
3. Set the PWD key to the password corresponding to the user name you provided in step 2.

Using Kerberos

To configure Kerberos authentication:

1. Set the H2SAuthMech configuration key for the DSN to 1.
2. If your Kerberos setup does not define a default realm or if the realm of your Hive server is not the default, then set the appropriate realm using the HS2KrbRealm key.
3. Set the HS2HostFQDN key to the fully qualified domain name of the Hive Server 2 host.
4. Set the HS2KrbServiceName key to the service name of the Hive Server 2 host.

Hive ODBC Connector on Windows

There are versions of the connector for 32-bit and 64-bit applications. The 64-bit version of the connector works only with 64-bit DSNs; the 32-bit connector works only with 32-bit DSNs. Because 64-bit Windows machines can run both 64-bit and 32-bit applications, install both versions of the connector in order to set up DSNs to work with both types of applications. If both the 32-bit connector and the 64-bit connector are installed, you must configure DSNs for each independently, in their separate Data Source Administrators.

Software and Hardware Requirements

To use MapR Hive ODBC Connector on Windows requires:

- Windows® 7 Professional or Windows® 2008 R2. Both 32 and 64-bit editions are supported.
- The Microsoft Visual C++ 2010 Redistributable Package (runtimes required to run applications developed with Visual C++ on a computer that does not have Visual C++ 2010 installed.)
- A Hadoop cluster with the Hive service installed and running. You should find out from the cluster administrator the hostname or IP address for the Hive service and the port that the service is running on. (The default port for Hive is 10000.)

Install the Hive ODBC Connector on Windows

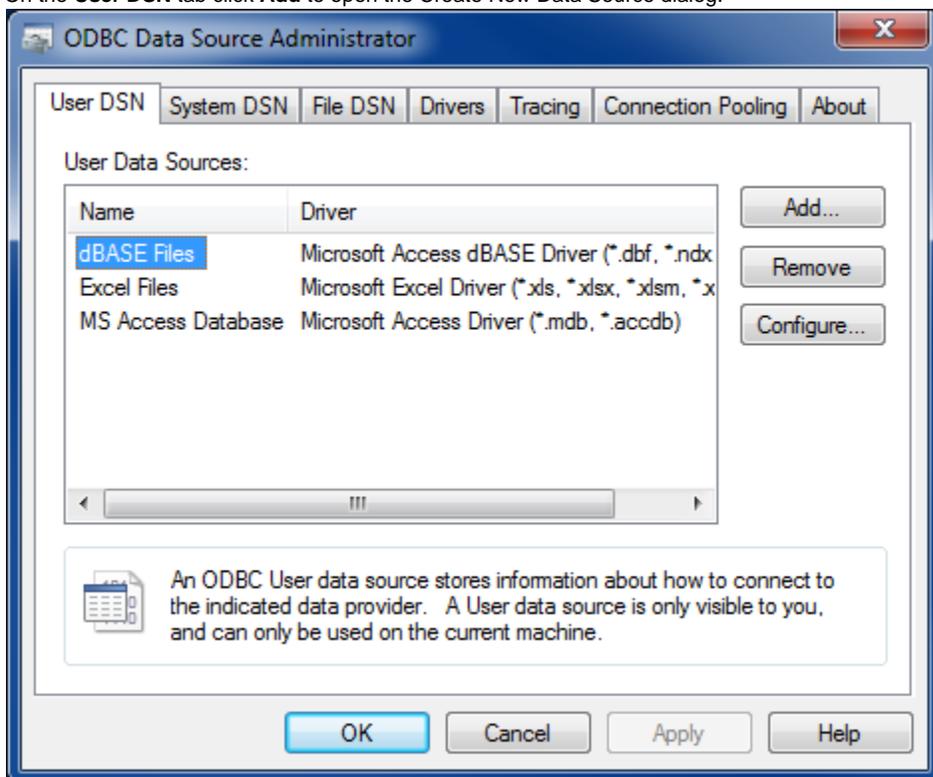
To install the Hive ODBC Connector:

1. Run the installer to get started:
 - To install the 64-bit connector, download and run the following package:
http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/Windows/MapRHiveODBC64.msi (Note: You must update the version placeholder in this link with one of the available versions of the connector [here](#). Here, in the directory named in the format MapRHive_odbc_<X.X.X.YYYY>, <X.X.X.YYYY> is the connector version.)
 - To install the 32-bit connector, download and run the following package:
http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/Windows/MapRHiveODBC32.msi
(Note: You must update the version placeholder in this link with one of the available versions of the connector [here](#). Here, in the directory named in the format MapRHive_odbc_<X.X.X.YYYY>, <X.X.X.YYYY> is the connector version.)
2. Accept the license agreement.
3. Select an installation folder.
4. On the Information window, click **Next**.
5. On the Completing... window, click **Finish**.
6. Install a DSN corresponding to your Hive server.

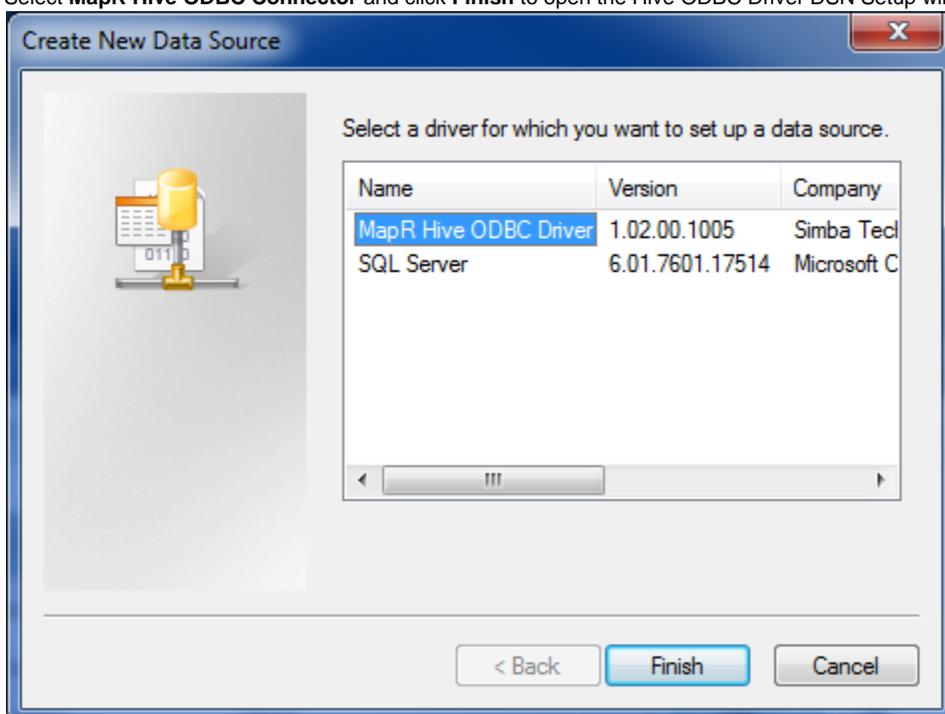
Configure Hive ODBC Connections on Windows

To create a Data Source Name (DSN)

1. Open the Data Source Administrator from the Start menu. Example:
Start > MapR Hive ODBC Driver 2.1 > 64-Bit ODBC Administrator
2. On the **User DSN** tab click **Add** to open the Create New Data Source dialog.



3. Select **MapR Hive ODBC Connector** and click **Finish** to open the Hive ODBC Driver DSN Setup window.



4. Enter the connection information for the Hive instance:

Simba Hive ODBC Driver DSN Setup

Data Source Name: HiveServer2 on MapR

Description:

Host: hivenode

Port: 10000

Database: default

Hive Server Type: Hive Server 2

Authentication

Mechanism: User Name

Realm:

Host FQDN:

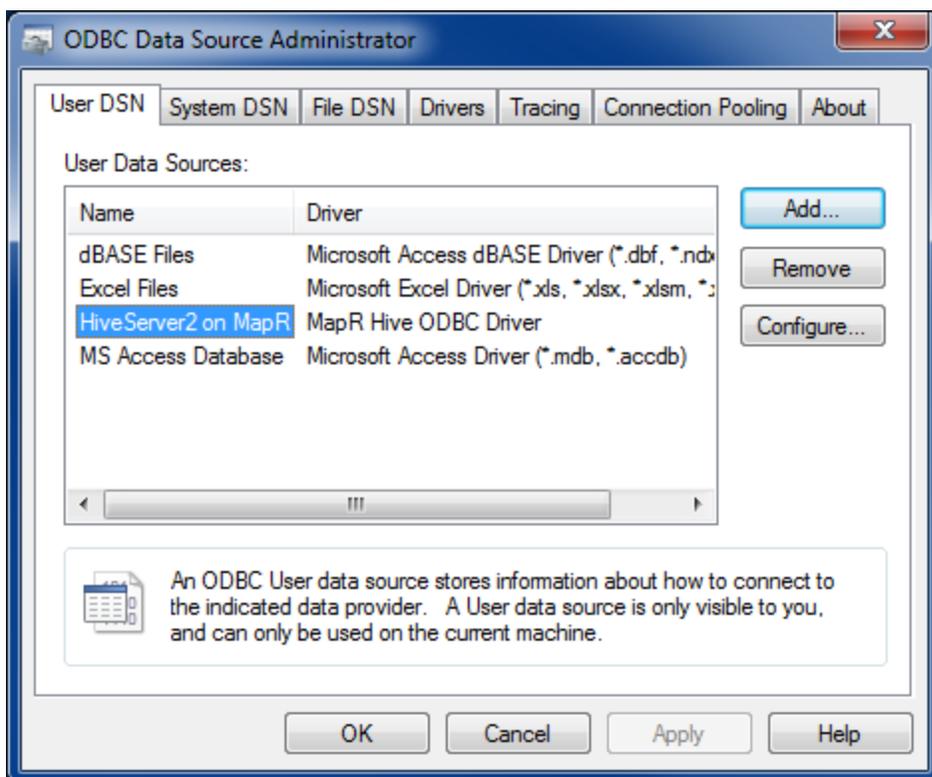
Service Name:

User Name: mapr

Advanced Options...

v1.2.0.1005 (64 bit) Test OK Cancel

- **Data Source Name** — Specify a name for the DSN.
 - **Description** — Enter an optional description for the DSN.
 - **Host** — Enter the hostname or IP of the server running HiveServer1 or HiveServer2.
 - **Port** — Enter the listening port for the Hive service.
 - **Database** — Leave as `default` to connect to the default Hive database, or enter a specific database name.
 - **Hive Server Type:** — Set to HiveServer1 or HiveServer2.
 - **Authentication** — If you are using HiveServer2, set the following.
 - **Mechanism:** — Set to the authentication mechanism you're using. The MapR ODBC driver supports user name, user name and password, username and password over SSL authentication, and Kerberos.
 - **User Name:** — Set the user to run queries as.
 - **Password:** — The user's password, if your selected authentication mechanism requires one.
5. Optionally, click **Test** to test the connection.
6. Click **OK**.
- Your new connector will appear in the User Data Sources list.



Hive ODBC Connector on Mac OS X

System Requirements

- Mac OS X version 10.6.8 or later
- 100 MB of available disk space.
- iODBC 3.52.7 or above

The MapR ODBC Driver with SQL Connector for Apache Hive requires a Hadoop cluster with the Hive service installed and running. The MapR ODBC Driver with SQL Connector for Apache Hive is suitable for use with all versions of Hive. The driver supports both 32- and 64-bit client applications.

Download the MacOS Hive ODBC connector from http://package.mapr.com/tools/MapR-ODBC/MapR_Hive/MapRHive_odbc_<version>/MacOSX/.

Note: You must update the version placeholder in the provided download location.

Install the Hive ODBC Connector on Mac OS X

The MapR ODBC Driver with SQL Connector for Apache Hive driver files are installed in the following directories:

- `/opt/mapr/hiveodbc/ErrorMessage` – Error messages files directory
- `/opt/mapr/hiveodbc/Setup` – Sample configuration files directory
- `/opt/mapr/hiveodbc/lib/universal` – Binaries directory

To install the MapR ODBC Driver with SQL Connector for Apache Hive:

1. Double-click to mount the `MapRHiveODBC.dmg` disk image.
2. Double-click `MapRHiveODBC.pkg` to run the Installer.
3. Follow the instructions in the Installer to complete the installation process.
4. When the installation completes, click **Close**.

Configure the Hive ODBC Connector Driver on Mac OS X

Setting the DYLD_LIBRARY_PATH Environment Variable

The DYLD_LIBRARY_PATH environment variable must include the paths to:

- Installed ODBC driver manager libraries
- Installed MapR ODBC Driver with SQL Connector for Apache Hive shared libraries

For example, if ODBC driver manager libraries are installed in `/usr/local/lib`, then set DYLD_LIBRARY_PATH as follows:

```
export DYLD_LIBRARY_PATH=/usr/local/lib/opt/mapr/hiveodbc/lib/universal
```

Refer to your Mac OS X shell documentation for details on how to set environment variables permanently.

Configure Hive ODBC Connections on Mac OS X

See [Configuring ODBC Connections for Linux](#) for details on creating ODBC connections.