Pentaho Data Integration Installation Guide

Help and Support Resources

If you have questions that are not covered in this guide, or if you would like to report errors in the documentation, please contact your Pentaho technical support representative.

Support-related questions should be submitted through the Pentaho Customer Support Portal at http://support.pentaho.com.

For information about how to purchase support or enable an additional named support contact, please contact your sales representative, or send an email to sales@pentaho.com.

For information about instructor-led training on the topics covered in this guide, visit http://www.pentaho.com/training.

Limits of Liability and Disclaimer of Warranty

The author(s) of this document have used their best efforts in preparing the content and the programs contained in it. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, express or implied, with regard to these programs or the documentation contained in this book.

The author(s) and Pentaho shall not be liable in the event of incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of the programs, associated instructions, and/or claims.

Trademarks

Pentaho (TM) and the Pentaho logo are registered trademarks of Pentaho Corporation. All other trademarks are the property of their respective owners. Trademarked names may appear throughout this document. Rather than list the names and entities that own the trademarks or insert a trademark symbol with each mention of the trademarked name, Pentaho states that it is using the names for editorial purposes only and to the benefit of the trademark owner, with no intention of infringing upon that trademark.

Company Information

Pentaho Corporation
Citadel International, Suite 340
5950 Hazeltine National Drive
Orlando, FL 32822
Phone: +1 407 812-OPEN (6736)
Fax: +1 407 517-4575
http://www.pentaho.com

E-mail: communityconnection@pentaho.com

Sales Inquiries: sales@pentaho.com

Documentation Suggestions: documentation@pentaho.com

Sign-up for our newsletter: http://community.pentaho.com/newsletter/

Contents

Installation Overview

This guide explains how to install the Pentaho Data Integration Enterprise Edition version 4.4 on both servers and workstations using either the standalone PDI graphical installer, or the equivalent archive packages. Typically, the Data Integration Server and Pentaho Enterprise Console will go on the server, and each workstation will have its own copy of the data integration design tools.

If you need more information about using Pentaho Data Integration with big data technology, see the *Big Data Guide*.

To ensure that your installation goes smoothly, refer to the *Compatability Matrix: Supported Components* for details about the versions of databases, application servers, software, and browsers that are compatible with Pentaho Data Integration.

Use the archive-based installation if you are a Windows, Linux, OS X, or Solaris user and you

- Have your own or are creating your own deployment scripts
- Are using a command line interface or are performing a headless installation
- Are performing a remote installation
- Plan to connect to a database that is not the default installation database in the installation wizard
- Are independently upgrading or migrating to a new version of a particular product or plugin
- Need more fine-grained control over configuration and deployment options for servers

Use the graphical installation if you are a Windows or Linux user and prefer to

- Install using a graphical installer
- Install quickly and/or for evaluation purposes

The PDI Graphical Installer Provides	You Must Supply
Data Integration server **Data Integration tools:** • Spoon (graphical interface) • Kitchen (CLI job interface) • Pan (CLI transformation interface) • Carte (CLI execution engine for PDI content) **Pentaho Enterprise Console** **A Java Runtime Environment** (only if you use the PDI graphical installer)	**A supported operating system:** • Linux • Windows • Solaris • Mac OS **A Java Runtime Environment** (only if you use the PDI archive packages) **One or more data sources:** • Any JDBC-compliant database • A spreadsheet • A flat file containing comma-separated values **A supported Big Data structure:** (for Big Data deployments only) • Hadoop • Hive • MongoDb • HBase • Cassandra If you are using a Hadoop distribution other than the default Apache Hadoop distribution you need to configure PDI for your specific version of Hadoop. See the *Setting the Active Hadoop Configuration* section of the *Pentaho Data Integration Admin Guide* for more guidance on changing which version of Hadoop works with PDI.

Installation Methods

There are two ways to install Pentaho Data Integration: through the graphical installer and through the archive packages.

Graphical Installation

The installer provides a Java Runtime Environment, a master control script for starting and stopping the PDI servers, a **Start** menu entry in Windows, automatic service configuration in Windows, and it auto-checks for available port numbers and reassigns them if there are port collisions.

Archive Package Installation

The archive packages are more useful for **headless**, **unattended** (if you have your own deployment scripts), and **remote** installations. Installing from archives is the easiest and quickest way to deploy to one server and many workstations. If you perform an archive-based installation, you will have to manually configure your operating system to start the PDI servers as boot services. This guide provides basic instructions for these processes, but you will need to modify them to accommodate your operating environment.

The Graphical Installer

Pentaho provides self-contained graphical installation utilities for the Windows, Linux, and Mac OS platforms. These utilities will install the DI Server, Pentaho Enterprise Console, Java Runtime Environment, and Pentaho Data Integration client tools on a single machine. The Report Designer, the BA Server, and other Business Analytics software are not included in this package.

Consult the Welcome Kit email that was sent to you after completing the sales process. This email contains login credentials for the Customer Support Portal, where you can download all Pentaho Enterprise Edition software. The specific file names of each package are:

- **Windows 32-bit installer:** `PDI-4.4.0-GA-i386.exe`
- **Windows 64-bit installer:**`PDI-4.4.0-GA-x64.exe`
- **Linux 32-bit installer:** `PDI-4.4.0-GA-i386.bin`
- **Linux 64-bit installer:** `PDI-4.4.0-GA-x64.bin`
- **OS X 64-bit installer:** `pdi-4.4.0-GA-x64.app`

Note: It is also possible to use the much larger Pentaho Business Analytics graphical installer to install PDI as a server or client. The process will take a little longer.

Next Steps

After downloading the installer, follow the installation prompts and complete the installation procedure.

You may find it helpful to review these sections after running the graphical installation utility.

- *Starting the DI Server at Boot Time on Linux*
- *Switching LDAP Authentication*
- *Connecting to an Enterprise Repository*
- *Adding PDI Enterprise Repository Content Support to the BA Server*
- *Adding a JDBC Driver*

The Archive-Based Installation Process

If you prefer to install Pentaho Data Integration using the archive packages, follow the remaining content in this guide.

Because the server and workstation will be on different machines, it is helpful to reference the server and workstation checklists to ensure that you complete all the necessary steps.

Installation Checklist: Server

The Installation Checklist is a concise list of instructions intended to show a high-level overview of the installation and configuration process. It also serves as a quick reference for administrators and developers who have performed several installations in the past and only need a brief rehash of the necessary steps. **If you need more details than are provided in this checklist**, consult the appropriate section in the verbose instruction set that comprises the rest of this guide.

Step	Procedure	Done
Step 1	If you are a Unix operating systems user, create a pentaho system user.	
Step 2	Download the DI Server archive package from the Pentaho Customer Support Portal.	
Step 3	Create a `/pentaho/server/` directory in an appropriate location on your file system and unpack the DI Server and Pentaho Enterprise Console archive packages to `/pentaho/server/`.	
Step 4	Start the DI Server and Pentaho Enterprise Console.	
Step 5	Log into the Pentaho Enterprise Console, which by default is located at `http://localhost:8088`.	
Step 6	Install an Enterprise Edition key for Pentaho Data Integration.	
Step 7	Continue on to the workstation installation procedures.	
Step 8	As needed, perform post-install configuration.	

Installation Checklist: Workstation

The Installation Checklist is a concise list of instructions intended to show a high-level overview of the installation and configuration process. It also serves as a quick reference for administrators and developers who have performed several installations in the past and only need a brief rehash of the necessary steps. **If you need more details than are provided in this checklist**, consult the appropriate section in the verbose instruction set that comprises the rest of this guide.

Step	Procedure	Done
Step 1	Download the client tool archive package from the Pentaho Customer Support Portal.	
Step 2	Create a `/pentaho/design-tools/` directory in an appropriate location on your file system.	
Step 3	Unpack the PDI client archive package to `/pentaho/design-tools/`.	
Step 4	Install a PDI Enterprise Edition license key using the command line tool in the `/pentaho/design-tools/data-integration/license/` directory.	
Step 5	Start the PDI client tool (Spoon).	
Step 6	Create or connect to an enterprise repository. (This requires installing the DI Server.)	
Step 7	Test the new PDI installation by creating, sharing, and scheduling new content.	
Step 8	Remove any temporary files, such as the original archive packages you installed from.	
Step 9	Optionally, perform post-installation configuration.	

Prerequisites for Archive Installation

In order to install the Pentaho Data Integration archive packages, you must be very familiar with technical concepts.

- You must be familiar with system administration operations pertaining to network services, including modifying your firewall to open specific ports, and adding services to the system startup and shutdown scripts.
- You must feel comfortable using the operating system command line interface and/or graphical system administration tools.
- You or a system administrator must have the ability to install software, open firewall ports, and start and stop system services on the machine you are installing on.

Installing the Pentaho Data Integration client tools onto user workstations is a much simpler process that requires little more than being able to unpack a file archive or run the graphical installation utility as a regular user.

Refer to the *Compatibility Matrix: Supported Components* in this guide for a list of supported operating systems, browsers, and databases.

How to Check Your Java Version

Pentaho Business Analytics requires a Java Runtime Environment (JRE) or Java Development Kit (JDK). Follow this procedure to see which version of Java is installed on your system and configured to be the default Java executable. There may be multiple JREs or JDKs on your system, but only one can be set as the global default. If a particular JRE or JDK is not specified by an application on startup, the default is used. Pentaho establishes a specific system variable named **PENTAHO_JAVA_HOME** to declare which Java instance it will use.

1. Open a terminal or command prompt window.
2. Type this command in: **java -version** and press **Enter**.

   Along with the Java version, the bit-ness, 32-bit or 64-bit, and patch level also shows in the output. For example ...

```
java version "1.6.0_21"
Java(TM) SE Runtime Environment (build 1.6.0_21-b06)
Java HotSpot(TM) 64-Bit Server VM (build 17.0-b16, mixed mode)
```

Local User Accounts on Linux

If you create a new local user account called **pentaho** as described in *Creating a Pentaho System User (Linux/Solaris)* , you must put the pentaho-solutions directory and store license information recorded by the Pentaho Enterprise Console in that directory. You must also adjust your `init` scripts to start your Web application server (which will run the BA Server) and the Pentaho Enterprise Console as this user. If you are unable to do this, you must modify the instructions in the rest of this guide to support your own custom user and directory configuration.

If you do not run the Pentaho Enterprise Console and the application server that runs the BA Server as the same system user, the BA Server will not be able to find the appropriate license information and will not operate with full functionality. If you need to launch the Pentaho Enterprise Console and the BA Server as separate users or services, you will have to set a **-D** parameter in your web application server's service configuration or startup scripts to specify a static location for your Pentaho license file. The Java parameter is **-Dpentaho.installed.licenses.file=/pentaho/.installedLicenses.xml**, though you will need to change this example to match your configuration.

Creating a Pentaho System User on Linux

You may be able to use a different local user account, however you will have to modify all of the instructions to match that configuration.

Pentaho licenses are installed to an XML file in the home directory of the user that starts the Pentaho Enterprise Console. While you can use any system user to install licenses, it is easier to create a new user to start and stop the Pentaho Enterprise Console, and to install and update licenses with.

1. Open a local terminal on, or an OpenSSH session to the server you are hosting the BA Server on.

```
ssh pgibbons@192.168.1.133
```

2. With root permissions, create a new user account called **pentaho**.

   Bash is not a requirement, but it is typically the shell that Linux users want to standardize on. On many Linux distributions, the default new user shell is /bin/sh or some equivalent (such as Dash) that may not use the ~/.bashrc configuration file that you will work with later. If you don't have or want to use Bash, adjust the instructions throughout this guide accordingly.

   ```
   sudo useradd -s /bin/bash -m pentaho
   ```

3. Set a password for the pentaho user (note that when using the **sudo** command, you must first supply the root password).

   ```
   sudo passwd pentaho
   ```

4. Verify that you can log in using the credentials you specified.

   ```
   su pentaho -
   ```

You now have a user account created specifically for running the BA Server and controlling Pentaho Enterprise Console start and stop scripts. You should stay logged into this new account to create the pentaho directory and perform all other installation tasks that do not explicitly require root access.

You must use this new user account for starting and stopping the Pentaho Enterprise Console, and for installing and updating licenses if you use the command line tool to manage them instead of the graphical interface in Pentaho Enterprise Console. If you create any RC or init scripts to start Pentaho Enterprise Console automatically at boot time, then you will have to write those scripts such that they start the service with the **pentaho** user credentials.

Setting the PENTAHO_JAVA_HOME Variable on Linux

To ensure that Business Analytics will always use the correct Sun Java Runtime Environment, especially in software environments that contain multiple JREs, you must create a **PENTAHO_JAVA_HOME** system variable for your **pentaho** user account and point it to a supported JRE or JDK. If you do not set this variable, Business Analytics will attempt to use the JRE that the JAVA_HOME variable points to. We do not recommend that your BA Server installation rely on the **JAVA_HOME** variable, as this is the global default JRE for the entire system, and changing it could adversely affect other Java-based applications in the system. Therefore, Pentaho recommends that you set PENTAHO_JAVA_HOME as described in these steps.

1. Edit your **/etc/environment** file with a text editor.

   If you're using Solaris, you will have to set this environment variable through whatever means are available to you.

2. Add this line in a convenient place (replacing the path with the location of the JRE on your system): **export PENTAHO_JAVA_HOME=/usr/lib/jvm/java-6-sun**.
3. You must log out and log back into the operating system for the change to take effect.
4. Verify that the variable is properly set.

   ```
   env | grep PENTAHO_JAVA_HOME
   ```

Setting the PENTAHO_JAVA_HOME Variable on Windows

To ensure that Business Analytics will always use the correct Sun Java Runtime Environment, especially in software environments that contain multiple JREs, you must create a **PENTAHO_JAVA_HOME** system variable for your **pentaho** user account and point it to a supported JRE or JDK. If you do not set this variable, Business Analytics will attempt to use the JRE that the JAVA_HOME variable points to. We do not recommend that your BA Server installation rely on the **JAVA_HOME** variable, as this is the global default JRE for the entire system, and changing it could adversely affect other Java-based applications in the system. Therefore, Pentaho recommends that you set PENTAHO_JAVA_HOME as described in these steps.

1. In **Windows**, right-click on **Computer**, then select **Properties** from context menu, then click **Advanced System Settings**.

   The **System Properties** window will come up.

2. In the System Properties window, click the **Advanced** tab, then click **Environment Variables**.

3. In the System Variable section, click New.
4. A popup dialog will ask for a variable name and value. Type **PENTAHO_JAVA_HOME** into the name field.
5. In the value field, enter the directory for the JRE. An example of a JRE directory is `C:\Program Files\Java \jre6`. Click **OK**.
6. In the parent window, click **Apply Changes**.
7. You must restart your computer for the change to take effect.
8. Verify that the variable is properly set.

```
echo %PENTAHO_JAVA_HOME%
```

Obtaining the Archive Packages

Consult the Welcome Kit email that was sent to you after completing the sales process. This email contains user credentials for the Pentaho Customer Support Portal, where you can download individual archive packages for the DI Server and Data Integration client tools. Here are the packages you need for each platform and distribution:

- **DI Server for Windows:** `pdi-ee-server-4.4.0-GA.zip`
- **DI Server for Linux/Solaris/OS X:** `pdi-ee-server-4.4.0.0-GA.tar.gz`
- **Data Integration client tool Windows package:** `pdi-ee-client-4.4.0-GA.zip`
- **Data Integration client tool Linux/Solaris/OS X package:** `pdi-ee-client-4.4.0-GA.tar.gz`

If you download the **pdi-ee-server** package, you must also download the Pentaho Enterprise Console package:

- **Pentaho Enterprise Console for Linux/Solaris/OS X:** `pec-4.4.0-GA.tar.gz`
- **Pentaho Enterprise Console for Windows:** `pec-4.4.0-GA.zip`

Server Installation Procedure

To install and configure the Data Integration Server and Pentaho Enterprise Console from archive packages, follow the below procedures in the order they are presented.

Server Archive Package Deployment

Follow the below instructions to install the Data Integration Server and Pentaho Enterprise Console on a dedicated server. If you intend to deploy the client tools and servers on one machine, you can combine this with *Workstation Installation Procedures* on page 15, or use the PDI graphical installer.

Note: The example commands in this and other sections are specific to Linux. You will have to adjust or ignore them on other operating systems.

1. Create a **/pentaho/server/** directory in an appropriate place in your hierarchy.

   This directory should be accessible to the system users who will be controlling services. Typically only root or the users in the wheel or administrator group will need to do this.

   

   Note: If you are using the graphical installer, it will create this directory structure for you, so you can skip this step.

   ```
   mkdir -p /home/pentaho/pentaho/server/
   ```

2. Unpack the **pdi-ee-server-4.4.0-GA** archive to `/pentaho/server/`.

   ```
   tar zxvf pdi-ee-server-4.4.0-GA.tar.gz -C /home/pentaho/pentaho/server/
   ```

3. Unpack the **pec-4.8.0-GA** archive to `/pentaho/server/`.

   ```
   tar zxvf pec-4.8.0-GA.tar.gz -C /home/pentaho/pentaho/server/
   ```

4. Switch to the `/pentaho/server/data-integration-server/` directory and run the **start-pentaho** script to start the DI Server.

   ```
   cd /home/pentaho/pentaho/server/data-integration-server/ && ./start-pentaho.sh
   ```

5. Switch to the `/pentaho/server/enterprise-console-server/` directory and run the **start-pec** script to start the Pentaho Enterprise Console.

   ```
   cd /home/pentaho/pentaho/server/enterprise-console-server/ && ./start-pec.sh
   ```

The DI Server and Enterprise Console are now installed, and should be operational. The DI Server will not be accessible from workstations until a license key is installed.

Installing or Updating an Enterprise Edition Key

You must install Pentaho Enterprise Edition keys associated with products for which you have purchased licenses. The keys you install determine the layout and capabilities of the Pentaho Enterprise Console, and the functionality of the BA Server and DI Server. Follow the instructions below to install an Enterprise Edition key through the Pentaho Enterprise Console for the first time, or to update an expired or expiring key. If you would prefer to use a command line tool instead, see *Working From the Command Line Interface* on page 12.

Note: If your Pentaho Enterprise Console server is running on a different machine than your BA or DI Server, you must use the command line tool to install and update license files; you will not be able to use the Pentaho Enterprise Console for this task.

Note: License installation is a user-specific operation. You must install licenses from the user accounts that will start all affected Pentaho software. If your BA or DI Server starts automatically at boot time, you must install licenses under the user account that is responsible for system services. There is no harm in installing the licenses under multiple local user accounts, if necessary.

1. If you have not done so already, log into the Pentaho Enterprise Console by opening a Web browser and navigating to `http://server-hostname:8088`, changing **server-hostname** to the hostname or IP address of your BA or DI server.
2. Click the **+** (plus) button in the upper right corner of the Subscriptions section.

   An **Install License** dialog box will appear.
3. Click **Browse**, then navigate to the location you saved your LIC files to, then click **Open**.

   LIC files for each of your supported Pentaho products were emailed to you along with your Pentaho Welcome Kit. If you did not receive this email, or if you have lost these files, contact your Pentaho support representative. If you do not yet have a support representative, contact the Pentaho salesperson you were working with.

   

   Note: Do not open your LIC files with a text editor; they are binary files, and will become corrupt if they are saved as ASCII.
4. Click **OK**.

   The Setup page changes according to the LIC file you installed.

You can now configure your licensed products through the Pentaho Enterprise Console.

Working From the Command Line Interface

Though the Pentaho Enterprise Console is the quickest, easiest, and most comprehensive way to manage PDI and/or the BA Server, some Pentaho customers may be in environments where it is difficult or impossible to deploy or use the console. See the alternative instructions for command line interface (CLI) license registration for step-by-step instructions.

Installing an Enterprise Edition Key on Windows (CLI)

To install a Pentaho Enterprise Edition Key from the command line interface, follow the below instructions.

Note: Do not open your LIC files with a text editor; they are binary files, and will become corrupt if they are saved as ASCII.

1. Navigate to the `\pentaho\server\enterprise-console\license-installer\` directory, or the `\license-installer\` directory that was part of the archive package you downloaded.
2. Run the **install_license.bat** script with the **install** switch and the location and name of your license file as a parameter.

```
install_license.bat install "C:\Users\pgibbons\Downloads\Pentaho BA Platform
 Enterprise Edition.lic"
```

Upon completing this task, you should see a message that says, "The license has been successfully processed. Thank you."

Installing an Enterprise Edition Key on Linux (CLI)

To install a Pentaho Enterprise Edition Key from the command line interface, follow the below instructions.

Note: Do not open your LIC files with a text editor; they are binary files, and will become corrupt if they are saved as ASCII.

1. Navigate to the `/pentaho/server/enterprise-console/license-installer/` directory, or the `/license-installer/` directory that was part of the archive package you downloaded.
2. Run the **install_license.sh** script with the **install** switch and the location and name of your license file as a parameter. You can specify multiple files, separated by spaces, if you have more than one license key to install.

   

   Note: Be sure to use backslashes to escape any spaces in the path or file name.

```
install_license.sh install /home/pgibbons/downloads/Pentaho\ BI\ Platform\
 Enterprise\ Edition.lic
```

Upon completing this task, you should see a message that says, "The license has been successfully processed. Thank you."

Post-Install Configuration

After you've installed PDI software to your server, you must perform some extra tasks to register license keys, connect your workstations to the server, and configure the server to start at boot time. Follow the sections below that apply to your situation.

Starting the DI Server At Boot Time On Linux

This procedure assumes that you will be running your DI Server and Pentaho Enterprise Console server under the **pentaho** local user account. If you are using a different account to start these services, substitute it in the script below.

You can start and stop the DI Server manually at any time by running the **start-pentaho.sh** and **stop-pentaho.sh** scripts. To start the DI Server automatically at boot time, and stop automatically during shutdown, follow the below procedure.

1. With root permissions, create a file in `/etc/init.d/` called **pdi**.
2. Using a text editor, copy the following content into the new pdi script. Adjust the paths to the DI Server and Pentaho Enterprise Console scripts to match your situation.

```
#!/bin/sh -e
### BEGIN INIT INFO
# Provides: pdi
# Required-Start: networking
# Required-Stop:
# Default-Start: 2 3 4 5
# Default-Stop: 0 1 6
# Description: Pentaho DI Server
### END INIT INFO

case "$1" in
"start")
su - pentaho -c "/home/pentaho/pentaho/server/data-integration-server/start-
pentaho.sh"
su - pentaho -c "cd /home/pentaho/pentaho/server/enterprise-console && ./start-
pec.sh"
;;
"stop")
su - pentaho -c "/home/pentaho/pentaho/server/data-integration-server/stop-
pentaho.sh"
su - pentaho -c "cd /home/pentaho/pentaho/server/enterprise-console && ./stop-
pec.sh"
;;
*)
echo "Usage: $0 { start | stop }"
;;
esac
exit 0
```

3. Save the file and close the text editor.
4. Change the file permissions to make the init script executable.

```
chmod +x /etc/init.d/pdi
```

5. Add the pdi init script to the standard runlevels so that it will run when the system starts, and stop when the system is shut down or rebooted, by using the update-rc.d command.

   This command may not exist on your computer if it is not Debian-based. If that is the case, consult your distribution documentation or contact your distribution's support department to determine how to add init scripts to the default runlevels.

```
update-rc.d pdi defaults
```

The Pentaho DI Server will now start at boot time, and shut down when the system stops or restarts.

Switching LDAP Authentication

If you intend to use your own LDAP or Microsoft Active Directory server for DI Server authentication, you should now proceed to the *Pentaho Data Integration Administrator's Guide* before starting the Data Integration client tool. If you connect to a PDI enterprise repository before switching the security backend to LDAP, you will have to take extra steps later to delete and regenerate the default user and role configuration files.

Workstation Installation Procedures

To install and configure the Data Integration client tools from the archive package, follow the below procedures in the order they are presented.

Workstation Archive Package Deployment

Follow the below instructions to install the Data Integration client tools on your workstations. If you intend to deploy the client tools and servers on one machine, you can combine this with *Server Installation Procedure* on page 11, or use the PDI graphical installer.

Note: The example commands in this and other sections are specific to Linux. You will have to adjust or ignore them on other operating systems. If you need instructions for installing a license from a Windows command line, see *Installing an Enterprise Edition Key on Windows (CLI)* on page 12.

1. If necessary, download the Archive packages as described in *Obtaining the Archive Packages*.
2. Create a **/pentaho/design-tools/** directory in an appropriate place in your hierarchy.

```
mkdir -p /home/pgibbons/pentaho/design-tools/
```

3. Unpack the **pdi-ee-client-4.4.0-GA** archive to `/pentaho/design-tools/`.

```
tar zxvf pdi-ee-client-4.4.0-GA.tar.gz -C /home/pgibbons/pentaho/design-tools/
```

4. Navigate to the `/pentaho/design-tools/license-installer/` directory.
5. Run the **install_license.sh** script with the **install** switch and as a parameter, the location and name of your license file.

```
./install_license.sh install /home/rwilco/downloads/Pentaho\ PDI\ Enterprise\
 Edition.lic
```

The Data Integration client tools are now installed.

Connecting to an Enterprise Repository

Follow the below instructions to create a new enterprise repository connection from a PDI workstation. By default, the DI Server comes with a preconfigured enterprise repository, so there is no need to create one, but you must still connect to it.

1. Start Spoon by running the `/pentaho/design-tools/data-integration/spoon` script.

   The **Repository Connection** dialogue will appear.
2. Click the round green **+** icon in the upper right corner of the window.

   The **Repository Type** dialogue will appear.
3. Select **Enterprise Repository** in the list, then click **OK**.

   The **Repository Configuration** dialogue will appear.
4. Ensure that the **URL** field corresponds to your DI Server address and port number. Type in a system-identifiable value (a unique internal name for this repository instance) in the **ID** field, and a friendly name or description in the **Name** field.

   If you only intend to have one repository for all users, you can un-check the **Show this dialogue at startup** option before clicking OK. This will prevent the dialogue from appearing every time you start Spoon. If you need to make repository connection changes later, you can still get to this screen through the **Tools** menu.
5. Use the default credentials of **admin** and **secret** for this repository, and click **OK** to complete repository configuration.

   This account is part of the default PDI configuration. Refer to the *PDI Administrator's Guide* to learn more about setting up users and roles in PDI.

You are now connected to an enterprise repository, and are enabled to begin creating users and roles for your organization.

Testing and Cleanup

You should now have a complete PDI environment, from the DI Server to individual client workstations. Before you go into production, you should perform the following tests:

- Ensure that each PDI workstation has an enterprise repository connection.
- Create a new job, transformation, and/or Agile BI analysis schema.
- Schedule a job or transformation and ensure that the schedule executes properly.
- From one PDI user workstation, share a job or transformation with another user and verify that this other user can access it.
- Physically restart the server and ensure that the DI Server and Enterprise Console are automatically started as services.

Once you're certain that your PDI environment is ready for production, you can remove any installation artifacts, such as ZIP or tar.gz archives and installers. See the *PDI Administrator's Guide* for further guidance on system administration, configuration, and maintenance.

Adding PDI Enterprise Repository Content Support to the BA Server

If you are using a Pentaho Data Integration (PDI) enterprise repository (through a Data Integration Server) to store PDI jobs and transformations, and you plan on using those jobs and transformations in action sequences that will be run on the BA Server, you must install some BA Server plugins from the PDI client tool package. This is not a typical scenario, but there is no harm in performing it if you aren't sure of the details.

1. Download a PDI Enterprise Edition 4.4 client tool archive package from the Pentaho Customer Support Portal.
   The package name (available in both tar.gz and zip formats) is: **pdi-ee-client-4.4.0-GA**
2. Unpack the archive to a temporary location.
3. Edit the `/pentaho/server/biserver-ee/pentaho-solutions/system/kettle/settings.xml` file.
4. Change the value of the **<repository.type>** node from **files** to **rdbms**.
5. Enter your enterprise repository connection information in the proper nodes.
6. Enter the location of your local **repositories.xml** file in the **<repositories.xml.file>** node.

   

   Note: This file is created on your PDI client workstation when you establish a connection to an enterprise repository. Once you have made all of your repository connections on a workstation, copy the **repositories.xml** file to the `~/.kettle/` directory on the BA Server and DI Server machines. If the client tool and servers are all on the same machine, you do not have to copy the file. If you have not yet established any repositories, you will have to revisit this procedure later when your PDI environment is fully configured.
7. Copy the contents of `/data-integration/plugins/` to the `/pentaho/server/biserver-ee/pentaho-solutions/system/kettle/plugins/` directory.

   ```
   cp -r /tmp/data-integration/plugins/* /home/pentaho/pentaho/server/biserver-ee/
   pentaho-solutions/system/kettle/plugins/
   ```
8. Remove the unpacked archive.

   ```
   rm -rf /tmp/data-integration/
   ```

Your BA Server is now configured to run content stored in the DI Server.

Adding a JDBC Driver

Before you can connect to a data source in any Pentaho server or client tool, you must first install the appropriate database driver. Your database administrator, CIO, or IT manager should be able to provide you with the proper driver JAR. If not, you can download a JDBC driver JAR file from your database vendor or driver developer's web site. Once you have the JAR, follow the instructions below to copy it to the driver directories for all of the Data Integration components that need to connect to this data source.

Note: Microsoft SQL Server users frequently use an alternative, non-vendor-supported driver called JTDS. If you are adding an MSSQL data source, ensure that you are installing the correct driver.

Backing up old drivers

You must also ensure that there are no other versions of the same vendor's JDBC driver installed in these directories. If there are, you may have to back them up and remove them to avoid confusion and potential class loading problems. This is of particular concern when you are installing a driver JAR for a data source that is the same database type as your Pentaho solution repository. If you have any doubts as to how to proceed, contact your Pentaho support representative for guidance.

Installing JDBC drivers

Copy the driver JAR file to the following directories, depending on which servers and client tools you are using (Dashboard Designer, Pentaho Interactive Reporting, and Analyzer are all part of the BA Server):

Note: For the DI Server: before copying a new JDBC driver, ensure that there is not a different version of the same JAR in the destination directory. If there is, you must remove the old JAR to avoid version conflicts.

- **BA Server:** `/pentaho/server/biserver-ee/tomcat/lib/`
- **Enterprise Console:** `/pentaho/server/enterprise-console/jdbc/`
- **Data Integration Server:** `/pentaho/server/data-integration-server/tomcat/webapps/pentaho-di/WEB-INF/lib/`
- **Data Integration client:** `/pentaho/design-tools/data-integration/libext/JDBC/`
- **Report Designer:** `/pentaho/design-tools/report-designer/lib/jdbc/`
- **Schema Workbench:** `/pentaho/design-tools/schema-workbench/drivers/`
- **Aggregation Designer:** `/pentaho/design-tools/agg-designer/drivers/`
- **Metadata Editor:** `/pentaho/design-tools/metadata-editor/libext/JDBC/`

Note: To establish a data source in the Pentaho Enterprise Console, you must install the driver in both the Enterprise Console and the BA Server or Data Integration Server. If you are just adding a data source through the Pentaho User Console, you do not need to install the driver to Enterprise Console.

Restarting

Once the driver JAR is in place, you must restart the server or client tool that you added it to.

Connecting to a Microsoft SQL Server using Integrated or Windows Authentication

The Microsoft SQL Server JDBC driver supports Type 2 integrated authentication on Windows operating systems through the **integratedSecurity** connection string property. To use integrated authentication, copy the **sqljdbc_auth.dll** file to all the directories to which you copied the JDBC files.

The **sqljdbc_auth.dll** files are installed in the following location:

```
<installation directory>\sqljdbc_<version>\<language>\auth\
```

Note: Use the **sqljdbc_auth.dll** file, in the x86 folder, if you are running a 32-bit Java Virtual Machine (JVM) even if the operating system is version x64. Use the **sqljdbc_auth.dll** file in the x64 folder, if you are running a 64-bit JVM on a x64 processor. Use the **sqljdbc_auth.dll** file in the IA64 folder, you are running a 64-bit JVM on an Itanium processor.

Compatibility Matrix: Supported Components

Pentaho aims to accommodate our clients' diverse computing environments. This list provides details about the environment components and versions we support. If you have questions about your particular computing environment, please contact Pentaho support.

Client

Pentaho client software is hardware-independent and runs on client-class computers that comply with these specifications for minimum hardware and required operation systems.

Pentaho Software
Pentaho Aggregation Designer
Pentaho Data Integration
Pentaho Design Studio
Pentaho Metadata Editor
Pentaho Report Designer
Pentaho Schema Workbench

Hardware—32 or 64 bit	Operating System—32 or 64 bit
Processors: • Apple Macintosh Dual-Core • Intel EM64T or AMD64 Dual-Core **RAM:** 2 GB RAM **Disk Space:** 2 GB free after installation	• Apple Macintosh OS 10.7 & 10.8 • Microsoft Windows 7 • Ubuntu Server 10.X and 12.X

Server

Pentaho server software is hardware-independent and runs on server-class computers that comply with these specifications for minimum hardware and required operation systems.

Pentaho Software
Pentaho Business Analysis Server
Pentaho Data Integration Server
Pentaho Enterprise Console

Hardware—64 bit	Operating System—64 bit
• Apple Macintosh Pro Quad-Core or Macintosh Mini Quad-Core • Intel EM64T or AMD64 Dual-Core **RAM:** 8 GB with 4 GB dedicated to Pentaho servers, 1 GB to Pentaho Enterprise Console **Disk Space:** 20 GB free after installation	• Apple Macintosh OS X Server 10.6 & 10.7 • CentOS Linux 5 & 6 • Microsoft Windows 2008 Server R1 & R2 • Red Hat Enterprise Linux 5 & 6 • Solaris 10 • Ubuntu Server 10.X & 12.X

Embedded Software

When embedding Pentaho software into other applications, the computing environment should comply with these specifications for minimum hardware and required operation systems.

Pentaho Software
Embedded Pentaho Reporting
Embedded Pentaho Analysis
Embedded Pentaho Data Integration

Hardware—32 or 64 bit	Operating System—32 or 64 bit
Processors: • Apple Macintosh Pro Quad-Core or Macintosh Mini Quad-Core • Intel EM64T or AMD64 Dual-Core **RAM:** 8 GB with 4 GB dedicated to Pentaho servers **Disk Space:** 20 GB free after installation	• Apple Macintosh OS X Server 10.6 & 10.7 • CentOS Linux 5 & 6 • Microsoft Windows 2008 Server R1 & R2 • Microsoft Windows 7 • Red Hat Enterprise Linux 5 & 6 • Solaris 10 • Ubuntu Server 10.X & 12.X

Application Servers

Servers to which you deploy Pentaho software must run one of these application servers.

Pentaho Software	Application Server
Pentaho Business Analysis Server	• Jboss 5.1.x • Tomcat 6.0.x
Pentaho Data Integration Server	Tomcat 6.0.x

Solution Database Repositories

Pentaho software stores processing artifacts in these solution database repositories.

Pentaho Software	Database Repository
Pentaho Business Analysis Server	• MySQL 5.x • Oracle 10g/11i • PostgreSQL 8.x & 9.1.x*
Pentaho Data Integration Server	Integrated Pentaho-specific, H2 1.2.131

*Default installed solution database

Data Sources

Pentaho software connects to these relational and non-relational data sources.

Pentaho Software	Data Source
Pentaho Reporting	• JDBC 3** • ODBC • OLAP4J • XML • Pentaho Analysis • Pentaho Data Integration • Pentaho Metadata
Pentaho Business Analysis Server, Action Sequences	• Relational (JDBC) • Hibernate • Javascript

Pentaho Software	Data Source
	• Metadata (MQL) • Mondrian (MDX) • XML (XQuery) • Security User/Role List Provider • Data Integration Steps (PDI) • Other Action Sequences • Web Services • XMLA

Pentaho Data Integration	• JDBC 3**
	• OLAP4J
	• Salesforce
	• XML
	• CSV
	• Microsoft Excel
	• Pentaho Analysis
	• Apache Hadoop 0.20.2 & 0.20.203.0***
	• Cloudera CDH3u4***
	• CDH4***
	• MapR 1.1.3 & 1.2.0
	• Cassandra distributions
	  • Apache 1.1.2
	  • DataStax 1.1.2
	• MongoDB 2.0.4

**Use a JDBC 3.x compliant driver that is compatible with SQL-92 standards when communicating with relational data sources. For your convenience, we provide a list of drivers used to get data from relational JDBC databases.

***From one of these distributions: HBase 0.90.5 and Hive 0.7.1

SQL Dialect-Specific

Pentaho software generates dialect-specific SQL when communicating with these data sources.

Pentaho Software	Data Source
Pentaho Analysis	• Access • DB2 • Derby • Firebird • Greenplum • Hive • Hsqldb • Infobright • Informix • Ingres • Interbase • LucidDb • MicrosoftSqlServer • MySql • Neoview • Netezza • Oracle • PostgreSQL

Pentaho Software	Data Source
	• SqlStream • Sybase • Teradata • Vectorwise • Vertica • Other SQL-92 compliant****
Pentaho Metadata	• DB2 • Firebird • H2 • Hypersonic • Ingres • MS Access • ASSQL • MSSQLNative • MySQL • Netezza • Oracle • PostgresSQL • Sybase • Other SQL-92 compliant****
Pentaho Data Integration	• Apache Derby • AS/400 • InfiniDB • Exasol 4 • Firebird SQL • Greenplum • H2 • Hypersonic • IBM DB2 • Infobright • Informix • Ingres • Ingres VectorWise • LucidDB • MaxDB (SAP DB) • MonetDB • MySQL • MS SQL Server • Neoview • Netezza • Oracle • Oracle RDB • PostgreSQL • SQLite • Teradata • UniVerse database • Vertica • Other SQL-92 compliant****

****If your data source is not in this list and is compatible with SQL-92, Pentaho software uses a generic SQL dialect.

Security

Pentaho software integrates with these third-party security authentication systems.

Pentaho Software	Authentication System
Pentaho Business Analysis Server Pentaho Enterprise Console	• Active Directory • CAS • Integrated Microsoft Windows Authentication • LDAP • RDBMS
Pentaho Data Integration Server	• Active Directory • LDAP • RDBMS

Java Virtual Machine

All Pentaho software, except the Pentaho Mobile App, requires the Sun/Oracle version 1.6 (6.0) distribution of the Java Runtime Environment (JRE) or Java Development Kit (JDK).

Web Browsers

Pentaho supports these major versions of Web browsers that are publicly available six weeks prior to when Pentaho begins to finalize a release. We also support the preceding major version.

Pentaho Software	Web Browser
Pentaho User Console Pentaho Enterprise Console Pentaho Report Designer*****	• Apple Safari 5.x • Google Chrome 19 • Microsoft Internet Explorer 8 & 9 • Mozilla Firefox 13 & 14

*****Requires a web browser to preview the exported HTML reports.

Mobile Apps

Pentaho mobile apps run on the Apple iPad 2 and 3 using iOS 5.x and 6.

JDBC Drivers

JDBC Drivers

This reference is a continuous work in progress. If you are viewing it in the Pentaho InfoCenter and see something that is not correct, know of a driver that is not listed here, or have a tip you want to share, please let us know by using the comments fields found in the bottom right corner.

Database	Vendor	URL
Apache Derby	IBM	*http://db.apache.org/derby/derby_downloads.html*
Cache'	InterSystems	*http://www.cachemonitor.de/intersystems-documentation/cache-jdbc-driver*
CUBRID	CUBRID	*http://www.cubrid.org/?mid=downloadsitem=jdbc_driver*
Daffodil DB	Daffodil Software	*http://sourceforge.net/projects/daffodildb/*

Database	Vendor	URL
DB2 AS/400	IBM	*http://www-03.ibm.com/systems/i/software/toolbox/*
DB2 Universal Database	IBM	*http://www-306.ibm.com/software/data/db2/java*
Firebird	Firebird Foundation	*http://www.firebirdsql.org/en/jdbc-driver/*
FrontBase	FrontBase	*http://www.frontbase.com/cgi-bin/WebObjects/FBWebSite*
Greenplum	EMC2	*http://jdbc.postgresql.org/download.html*
H2 Database	H2	*http://www.h2database.com*
Hive	Apache	*http://hive.apache.org/*
HSQLDB	HyperSQL	*http://sourceforge.net/projects/hsqldb/*
Informix	IBM	*http://www-01.ibm.com/software/*
Ingres	Actian	*http://esd.actian.com/product/drivers/JDBC/java*
InterBase	Embarcadero	*http://edn.embarcadero.com*
jTDS Free MS Sybase	jTDS	*http://jtds.sourceforge.net/*
LucidDB	DynamoDB	*http://www.dynamobi.com/c/downloads/stable/*
MaxDB	SAP	*http://maxdb.sap.com*
Mckoi	Mckoi SQL Database	*http://www.mckoi.com/originalmckoisql/index.html*
Mimer	Mimer Information Technology	*http://www.mimer.com*
MonetDB	MonetDB	*http://www.monetdb.org/*
MySQL	Oracle	*http://dev.mysql.com/downloads/connector/j/*
Neoview	HP	*https://h20392.www2.hp.com/portal/swdepot/displayProductInfo.do?productNumber=NEO10*
Netezza	IBM	*http://www.netezza.com*
OpenBase SQL	OpenBase International	*http://www.openbase.com/index.php/products/downloads*
Oracle	Oracle	*http://www.oracle.com/technetwork/database/features/jdbc/index.html*
Pervasive	Pervasive	*http://www.pervasivedb.com/download/Pages/PDBDownloads.aspx*
PostgreSQL	PostgreSQL Global Development Group	*http://jdbc.postgresql.org/*
SAP DB	SAP DB	*http://www.sapdb.org/sap_db_jdbc.htm*
SQLite	Xerial	*http://www.xerial.org/trac/Xerial/wiki/SQLiteJDBC*

Database	Vendor	URL
SQL Server	Microsoft	*http://msdn.microsoft.com/en-us/sqlserver/aa937724.aspx*
Sybase ASE	SAP	*http://www.sybase.com/products/allproductsa-z/softwaredeveloperkit/jconnect*
Sybase SQL Anywhere	SAP	*http://www.sybase.com/products/allproductsa-z/softwaredeveloperkit/jconnect*
SmallSQL	SmallSQL	*http://www.smallsql.de/download.html*
Teradata	Teradata	*http://downloads.teradata.com/download/connectivity/jdbc-driver*
Vertica	HP	*http://www.vertica.com*

Apache Derby

Vendor Name	Details	
Recommended Native Driver		
IBM	**Company URL** *http://www.ibm.com*	
	Driver URL *http://db.apache.org/derby/derby_downloads.html*	
	JDBC URL Syntax by Type Server—jdbc:derby://<server>[:<port>]/<databaseName>[;<URL attribute>=>value>] Embedded—jdbc:derby:<databaseName>[;create=true]	**Default Port** 1527
	JDBC Class org.apache.derby.jdbc.ClientDriver org.apache.derby.jdbc.EmbeddedDriver	**JDBC JAR File Name** derby.jar
	Shipped with Pentaho Products Pentaho Data Integration	
	Comments Open source database	

Caché

Vendor Name	Details
Recommended Native Driver	
InterSystems	**Company URL** *http://www.cachemonitor.de*
	Driver URL *http://www.cachemonitor.de/intersystems-documentation/cache-jdbc-driver*

Vendor Name	Details	
	JDBC URL Syntax by Type Server—jdbc:Cache://<server>[:<port>]/<namespace>	**Default Port** 1972
	JDBC Class com.intersys.jdbc.CacheDriver	**JDBC JAR File Name** cachedb.jar

CUBRID

Vendor Name	Details	
Recommended Native Driver		
CUBRID	**Company URL** ***http://www.cubrid.org***	
	Driver URL ***http://www.cubrid.org/?mid=downloads&item=jdbc_driver***	
	JDBC URL Syntax by Type Server—jdbc:cubrid:<server>:<port>:<databaseName>:<username>:<password> :[?<URL attribute>=<value>[&<URL attribute>=<value>] …]	**Default Port** 33000
	JDBC Class cubrid.jdbc.driver.CUBRIDDriver	**JDBC JAR File Name** N/A
	Comments Open source database highly optimized for Web applications.	

Daffodil DB

Vendor Name	Details	
Recommended Native Driver		
Daffodil Software	**Company URL** *http://db.daffodilsw.com*	
	Driver URL *http://sourceforge.net/projects/daffodildb/*	
	JDBC URL Syntax by Type Server—jdbc:daffodilDB://<server>[:<port>]/<databaseName> Embedded—jdbc:daffodilDB_embedded:<databaseName>	**Default Port** 3456 N/A
	JDBC Class in.co.daffodil.db.rmi.RmiDaffodilDBDriver	**JDBC JAR File Name** DaffodilDB_client.jar DaffodilDB_Embedded.jar,

Vendor Name	Details	
	in.co.daffodil.db.jdbc.DaffodilDBDriver	DaffodilDB_Common.jar
	Comments Open source database	

DB2 AS/400

Vendor Name	Details	
Recommended Native Driver		
IBM	**Company URL** *http://www.ibm.com*	
	Driver URL *http://www-03.ibm.com/systems/i/software/toolbox/*	
	JDBC URL Syntax by Type Server—jdbc:as400:// <server>naming=sql;errors=full	**Default Port** N/A
	JDBC Class com.ibm.as400.access.AS400JDBCDriver	**JDBC JAR File Name** jt400.jar
	Shipped with Pentaho Products Pentaho Data Integration	

DB2 Universal Database

Vendor Name	Details	
Recommended Native Driver		
IBM	**Company URL** *http://www.ibm.com*	
	Driver URL *http://www-306.ibm.com/software/data/db2/java*	
	JDBC URL Syntax by Type Server—jdbc:db2://<server>[:<port>]/ <databaseName>[;<URL attribute>=<value>]	**Default Port** 50000
	JDBC Class com.frontbase.jdbc.FBJDriver	**JDBC JAR File Name** frontbasejdbc.jar

Firebird

Vendor Name	Details
Recommended Native Driver	
Firebird Foundation	**Company URL**

Vendor Name	Details	
	http://www.firebirdsql.org	
	Driver URL *http://www.firebirdsql.org/en/jdbc-driver/*	
	JDBC URL Syntax by Type Server—jdbc:firebirdsql:<server>[/<port>]:/<database-file> (JDBC Type 4, official format) Server—jdbc:firebirdsql://<server>[:<port>]/<database-file> (JDBC Type 4, compatibility format) Server—jdbc:firebirdsql:native//<server>[/<port>]:/<database-file> (JDBC Type 2, compatibility format) Server—jdbc:firebirdsql:native://<server>[:<port>]/<database-file> JDBC Type 2, compatibility format. Requires libraries) Embedded—jdbc:firebirdsql:embedded:/<local-database-file> (JDBC Type 2, compatibility format. Requires libraries)	**Default Port** 3050 3050 3050 3050 N/A
	JDBC Class org.firebirdsql.jdbc.FBDriver	**JDBC JAR File Name** jaybird-full-xxx.jar
	Shipped with Pentaho Products Pentaho Data Integration	

FrontBase

Vendor Name	Details	
Recommended Native Driver		
FrontBase	**Company URL** *http://www.frontbase.com*	
	Driver URL *http://www.frontbase.com/cgi-bin/WebObjects/FBWebSite*	
	JDBC URL Syntax by Type Server— jdbc:FrontBase://<host>[:<port>]/<databaseName>	**Default Port** N/A
	JDBC Class	**JDBC JAR File Name**

Vendor Name	Details	
	com.frontbase.jdbc.FBJDriver	frontbasejdbc.jar

Greenplum

Vendor Name	Details	
Recommended Native Driver		
Greenplum	**Company URL** *http://www.greenplum.com*	
	Driver URL *http://jdbc.postgresql.org/download.html*	
	JDBC URL Syntax by Type Server—jdbc:postgresql://<server>[:<port>]/<databaseName>	**Default Port** 5342
	JDBC Class org.postgresql.Driver	**JDBC JAR File Name** postgresql-8.x-xxx.jdbc4.jar
	Comments Greenplum uses the Postgresql JDBC driver	

H2 Database

Vendor Name	Details	
Recommended Native Driver		
H2	**Company URL** *http://www.h2database.com*	
	Driver URL *http://www.h2database.com*	
	JDBC URL Syntax by Type Server—jdbc:h2:tcp://server[:port]/file-path Embedded—jdbc:h2:file-name	**Default Port** 9092 N/A
	JDBC Class jdbc:h2:tcp://server[:port]/file-path org.h2.Driver	**JDBC JAR File Name** h2-x.x.xxx.jar
	Shipped with Pentaho Products • Pentaho Business Analysis Server • Pentaho Data Integration • Pentaho Metadata-Editor Pentaho Report-Designer	
	Comments	

Vendor Name	Details
	Open source Java SQL database

Hive

Vendor Name	Details	
Recommended Native Driver		
Apache	**Company URL** *http://hive.apache.org/*	
	Driver URL N/A	
	JDBC URL Syntax by Type Server—jdbc:hive://<server>[:<port>]/default	**Default Port** 10000
	JDBC Class org.apache.hadoop.hive.jdbc.HiveDriver	**JDBC JAR File Name** hive-jdbc-x.x.x-pentaho-y.y.y.jar Example: hive-jdbc-0.7.0-pentaho-1.0.1.jar*
	Shipped with Pentaho Products • Pentaho Business Analysis Server • Pentaho Data Integration • Pentaho Metadata-Editor Pentaho Report-Designer	
	Comments Data warehouse infrastructure that provides data summarization and ad hoc querying *x.x.x is the Hive version, y.y.y is the Pentaho version. Pentaho has enhanced the standard Hive JDBC driver	

HSQLDB

Vendor Name	Details	
Recommended Native Driver		
HyperSQL	**Company URL** *http://www.hsqldb.org*	
	Driver URL *http://sourceforge.net/projects/hsqldb/*	
	JDBC URL Syntax by Type Server—jdbc:hsqldb:hsql://<server>[:<port>]/<databaseName> Embedded Memory—jdbc:hsqldb:mem:<databaseName>	**Default Port** 9001 N/A N/A

Vendor Name	Details	
	Embedded File — jdbc:hsqldb:file:<database-file>	
	JDBC Class org.hsqldb.jdbcDriver	**JDBC JAR File Name** hsqldb.jar
	Shipped with Pentaho Products • Pentaho Enterprise-Console • Pentaho Business Analysis Server • Pentaho Aggregation-Designer • Pentaho Data Integration • Pentaho Metadata-Editor • Pentaho Report-Designer	

Informix

Vendor Name	Details	
Recommended Native Driver		
IBM	**Company URL** *http://www.ibm.com*	
	Driver URL =	
	JDBC URL Syntax by Type Server—jdbc:informix-sqli://<server>[:<port>]/<databaseName>:informixserver=<dbservername>	**Default Port** 1533
	JDBC Class com.informix.jdbc.IfxDriver	**JDBC JAR File Name** ifxjdbc.jar
	Shipped with Pentaho Products Pentaho Data Integration	

Ingres

Vendor Name	Details	
Recommended Native Driver		
Actian	**Company URL** *http://www.actian.com/*	
	Driver URL *http://esd.actian.com/product/drivers/JDBC/java*	
	JDBC URL Syntax by Type Server—jdbc:ingres://<server>[:<port>]/<databaseName>	**Default Port** 21071
	JDBC Class	**JDBC JAR File Name**

Vendor Name	Details	
	com.ingres.jdbc.IngresDriver	iijdbc.jar
	Comments Open source relational database management system	

InterBase

Vendor Name	Details	
Recommended Native Driver		
Embarcadero	**Company URL** *http://edn.embarcadero.com*	
	Driver URL N/A	
	JDBC URL Syntax by Type Server—jdbc:interbase://<server>/<full_db_path>	**Default Port** N/A
	JDBC Class interbase.interclient.Driver	JDBC JAR File Name interclient.jar
	Shipped with Pentaho Products Pentaho Data Integration	

jTDS Free MS SQL Sybase

Vendor Name	Details	
Recommended Native Driver		
jTDS	**Company URL** *http://jtds.sourceforge.net/*	
	Driver URL N/A	
	JDBC URL Syntax by Type SQL Server—jdbc:jtds:<server_type>//<server>[:<port>][/<database>][;<property>=<value>[;...]]] Sybase—jdbc:jtds:<server_type>://<server>[:<port>][/<database>]	**Default Port** 1433 7100
	JDBC Class interbase.interclient.Driver	**JDBC JAR File Name** jtds-x.x.x.jar
	Shipped with Pentaho Products • Pentaho Enterprise-Console • Pentaho Business Analysis Server	

Vendor Name	Details
	• Pentaho Aggregation-Designer • Pentaho Data Integration • Pentaho Metadata-Editor • Pentaho Report-Designer

LucidDB

Vendor Name	Details	
Recommended Native Driver		
DynamoDB	**Company URL** *http://www.dynamobi.com*	
	Driver URL *http://www.dynamobi.com/c/downloads/stable/*	
	JDBC URL Syntax by Type Server—jdbc:luciddb:http://<server>[:<port>]	**Default Port** 8034
	JDBC Class org.luciddb.jdbc.LucidDbClientDriver	**JDBC JAR File Name** LucidDbClient-x.x.x.jar
	Shipped with Pentaho Products Pentaho Data Integration	
	Comments **Open source BI solution for Big Data**	

MaxDB

Vendor Name	Details	
Recommended Native Driver		
SAP	**Company URL** *http://www.sap.com*	
	Driver URL *http://maxdb.sap.com*	
	JDBC URL Syntax by Type Server—jdbc:sapdb://<server>[:<port>]/<databaseName>	**Default Port** 7210
	JDBC Class com.sap.dbtech.jdbc.DriverSapDB	**JDBC JAR File Name** sapdbc.jar
	Comments Database management system developed and supported by SAP AG	

Mckoi SQL Database

Vendor Name	Details	
Recommended Native Driver		
Mckoi	**Company URL** *http://www.mckoi.com*	
	Driver URL *http://www.mckoi.com/originalmckoisql/index.html*	
	JDBC URL Syntax by Type Server—jdbc:mckoi://<server>[:<port>][/<schema>]/	**Default Port** 9157
	JDBC Class com.mckoi.JDBCDriver	**JDBC JAR File Name** mckoidb.jar
	Comments Open source SQL database written in Java	

Mimer

Vendor Name	Details	
Recommended Native Driver		
Mimer Information Technology	**Company URL** *http://www.mimer.com*	
	Driver URL N/A	
	JDBC URL Syntax by Type Server—jdbc:mimer:<protocol>://<server>[:<port>]/<database>	**Default Port** 1360
	JDBC Class com.mimer.jdbc.Driver	**JDBC JAR File Name** mimer.jar

MonetDB

Vendor Name	Details	
Recommended Native Driver		
MonetDB	**Company URL** *http://www.monetdb.org*	
	Driver URL N/A	
	JDBC URL Syntax by Type	**Default Port** 50000

Vendor Name	Details	
	Server—jdbc:monetdb://<server>[:<port>]/<database>	
	JDBC Class nl.cwi.monetdb.jdbc.MonetDriver	**JDBC JAR File Name** monetdb-jdbc-x.x.jar
	Shipped with Pentaho Products Pentaho Data Integration	
	Comments An open source database system	

MY SQL

Vendor Name	Details	
Recommended Native Driver		
Oracle	**Company URL** *http://www.mysql.com*	
	Driver URL *http://dev.mysql.com/downloads/connector/j/*	
	JDBC URL Syntax by Type Server—jdbc:mysql://<hostname>[,<failoverhost>][:<port>] /<dbname>[?<URL attribute>=<value>[&<URL attribute>=<value>] ...]	**Default Port** 3306
	JDBC Class com.mysql.jdbc.Driver (official class name) org.gjt.mm.mysql.Driver (older class name)	**JDBC JAR File Name** mysql-connector-java-5.x.xx-bin.jar
	Shipped with Pentaho Product • Pentaho Enterprise-Console • Pentaho Business Analysis Server • Pentaho Aggregation-Designer • Pentaho Metadata-Editor • Pentaho Report-Designer	

Neoview

Vendor Name	Details
Recommended Native Driver	
HP	**Company URL** *http://www.hp.com*

Vendor Name	Details	
	Driver URL *https://h20392.www2.hp.com/portal/swdepot/displayProductInfo.do?productNumber=NEO10*	
	JDBC URL Syntax by Type Server—jdbc:hpt4jdbc://<system>[:<port>]/[:][<URL attribute>= <value>[;<URL attribute>=<value> …]	**Default Port** 18650
	JDBC Class com.hp.t4jdbc.HPT4Driver	**JDBC JAR File Name** N/A

Netezza

Vendor Name	Details	
Recommended Native Driver		
IBM	**Company URL** *http://www.netezza.com*	
	Driver URL N/A	
	JDBC URL Syntax by Type jdbc:netezza://<server>[:<port>]/<database>	**Default Port** 5480
	JDBC Class org.netezza.Driver	**JDBC JAR File Name** N/A

OpenBase SQL

Vendor Name	Details	
Recommended Native Driver		
OpenBase International	**Company URL** *http://www.openbase.com*	
	Driver URL *http://www.openbase.com/index.php/products/downloads*	
	JDBC URL Syntax by Type Server—jdbc:openbase://<server>/<databaseName>	**Default Port** N/A
	JDBC Class com.openbase.jdbc.ObDriver	**JDBC JAR File Name** OpenBaseJDBC.jar

Oracle

Vendor Name	Details	
Recommended Native Driver		
Oracle	**Company URL** *http://www.oracle.com*	
	Driver URL *http://www.oracle.com/technetwork/database/features/jdbc/index.html*	
	JDBC URL Syntax by Type Thin Server—jdbc:oracle:thin:@<server>[:<port>]:<sid> OCI Server—jdbc:oracle:oci:@<server>[:<port>]:<sid>	**Default Port** 1521
	JDBC Class oracle.jdbc.driver.OracleDriver oracle.jdbc.OracleDriver	**JDBC JAR File Name** ojdbcx.jar, ora18n.jar
	Comments The OCI server requires OCI libraries	

Pervasive

Vendor Name	Details	
Recommended Native Driver		
Pervasive	**Company URL** *http://www.pervasivedb.com/Pages/default.aspx*	
	Driver URL *http://www.pervasivedb.com/download/Pages/PDBDownloads.aspx*	
	JDBC URL Syntax by Type Server—jdbc:pervasive://<server>[:<port>]/<datasource>	**Default Port** 1583
	JDBC Class com.pervasive.jdbc.v2.Driver	**JDBC JAR File Name** N/A
	Shipped with Pentaho Products • Pentaho Data Integration • Pentaho Report-Designer	
	Comments The data source is the ODBC DSN	

PostgreSQL

Vendor Name	Details	
Recommended Native Driver		
PostgreSQL Global Development Group	**Company URL** *http://www.postgresql.org/*	
	Driver URL *http://jdbc.postgresql.org/*	
	JDBC URL Syntax by Type Server—jdbc:postgresql://<server>[:<port>]/<databaseName>	**Default Port** 5342
	JDBC Class org.postgresql.Driver	**JDBC JAR File Name** postgresql-8.x-xxx.jdbc4.jar
	Shipped with Pentaho Products • Pentaho Data Integration • Pentaho Report-Designer	

SAP DB

Vendor Name	Details	
Recommended Native Driver		
SAP DB	**Company URL** N/A	
	Driver URL *http://www.sapdb.org/sap_db_jdbc.htm*	
	JDBC URL Syntax by Type Server—jdbc:sapdb://<server>/<database_name>	**Default Port** N/A
	JDBC Class com.sap.dbtech.jdbc.DriverSapDB	**JDBC JAR File Name** sapdbc-x.x.x.jar
	Shipped with Pentaho Products Pentaho Data Integration	
	Comments FREE Enterprise Open Source Database	

SQLite

Vendor Name	Details
Recommended Native Driver	
Xerial	**Company URL**

Vendor Name	Details	
	N/A	
	Driver URL *http://www.xerial.org/trac/Xerial/wiki/SQLiteJDBC*	
	JDBC URL Syntax by Type Server—jdbc:sqlite:<filename.db>	**Default Port** N/A
	JDBC Class org.sqlite.JDBC	**JDBC JAR File Name** sqlite-jdbc-x.x.x.jar
	Shipped with Pentaho Products Pentaho Data Integration	

SQL Server

Vendor Name	Details	
Recommended Native Driver		
Microsoft	**Company URL** *http://www.microsoft.com*	
	Driver URL *http://msdn.microsoft.com/en-us/sqlserver/aa937724.aspx*	
	JDBC URL Syntax by Type Server—jdbc:sqlserver:// <server>[:<port>];DatabaseName=<databaseName>	**Default Port** 1433
	JDBC Class com.microsoft.sqlserver.jdbc.SQLServerDriver	**JDBC JAR File Name** sqljdbc4.jar
	Comments The open source jtds driver also works with MSSQL	

Sybase ASE

Vendor Name	Details	
Recommended Native Driver		
SAP	**Company URL** *http://www.sybase.com*	
	Driver URL *http://www.sybase.com/products/allproductsa-z/softwaredeveloperkit/jconnect*	
	JDBC URL Syntax by Type Server— jdbc:sybase:Tds:<server>[:<port>]/ <databaseName>	**Default Port** 5000

Vendor Name	Details	
	JDBC Class com.sybase.jdbc4.jdbc.SybDriver	**JDBC JAR File Name** N/A
	Comments The open source jTDS driver works with Sybase as well	

Sybase SQL Anywhere

Vendor Name	Details	
Recommended Native Driver		
SAP	**Company URL** *http://www.sybase.com*	
	Driver URL *http://www.sybase.com/products/allproductsa-z/softwaredeveloperkit/jconnect*	
	JDBC URL Syntax by Type Server— jdbc:sybase:Tds:<server>[:<port>]/ <databaseName>	**Default Port** 2638
	JDBC Class com.sybase.jdbc4.jdbc.SybDriver	**JDBC JAR File Name** N/A
	Comments This open source jTDS driver works with Sybase as well	

SmallSQL

Vendor Name	Details	
Recommended Native Driver		
SmallSQL	**Company URL** *http://www.smallsql.de/*	
	Driver URL *http://www.smallsql.de/download.html*	
	JDBC URL Syntax by Type Embedded— jdbc:smallsql:databaseName[?URL attribute=value[URLattribute=value] …]	**Default Port** N/A
	JDBC Class smallsql.database.SSDriver	**JDBC JAR File Name** smallsql.jar
	Comments	

Vendor Name	Details
	Java desktop SQL database engine

Teradata

Vendor Name	Details	
Recommended Native Driver		
Teradata	**Company URL** *http://www.teradata.com*	
	Driver URL *http://downloads.teradata.com/download/connectivity/jdbc-driver*	
	JDBC URL Syntax by Type Server—jdbc:teradata://<dbshost>[/<URL attribute>[;<URL attribute>]…]	**Default Port** N/A
	JDBC Class com.teradata.jdbc.TeraDriver	**JDBC JAR File Name** terajdbc4.jar

Vertica

Vendor Name	Details	
Recommended Native Driver		
HP	**Company URL** *http://www.vertica.com*	
	Driver URL TBD	
	JDBC URL Syntax by Type Server—jdbc:vertica://<server>[:<port>]/<databaseName>	**Default Port** 5433
	JDBC Class com.vertica.Driver	**JDBC JAR File Name** N/A