Upgrade DI Systems

Help and Support Resources

If you do not find answers to your quesions here, 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, visit http://www.pentaho.com/training.

Liability Limits and Warranty Disclaimer

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.

Third-Party Open Source Software

For a listing of open source software used by each Pentaho component, navigate to the folder that contains the Pentaho component. Within that folder, locate a folder named licenses. The licenses folder contains HTML.files that list the names of open source software, their licenses, and required attributions.

Contact Us

Global Headquarters 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

Sales Inquiries: sales@pentaho.com

Contents

DI Upgrade Introduction

This section explains how to upgrade version 4.4.x of the Pentaho Data Integration (DI) Server, plugins, and design tools to version 5.0. Version 5.0 contains major usability, stability, and architectural improvements. This version also includes many bug fixes and new features, as well as enhanced support for Big Data.

Note: To upgrade the Pentaho Business Analytics (BA) Server, plugins, and design tools see *Upgrade BA Systems*.

We recommend that you read through all upgrade documentation before you perform tasks. Familiarity with the entire process can help you plan your strategy and increase the likelihood of a successful upgrade experience.

Note: If you have a pre-4.4.x version of DI Server and you want to upgrade to version 5.0, you must do so incrementally, version by version, until you reach version 4.4.x. Upgrade instructions appear in the documentation for each version of the software. After you have upgraded to version 4.4.x, you can use this guide to migrate to version 5.0.

Prerequisites

You need a working version of DI Server version 4.4.x.

Expertise

The topics in this section are written for IT administrators who know where data is stored, how to connect to it, details about the computing environment, and how to use the command line to issue commands for Microsoft Windows or Linux. You also need the appropriate accesses and permissions to install and upgrade software.

Tools

You need these tools to perform upgrade tasks.

- Pentaho-supplied Migration Tool. Instructions for how to obtain this tool appear in the *Prepare Environment* section.
- File Difference tool of your choice.
- Zip or Archive tool of your choice.
- Text editor of your choice.

> **Note: This guide is only for Enterprise Edition upgrades.** You cannot upgrade a PDI 4.4 Community Edition deployment to PDI 5.0 Enterprise Edition through this process.

Login Credentials

You need access to an account that has administrative privileges. Additionally, Linux users might need to use the **root** account for some tasks.

DI Upgrade Overview

The upgrade process consists of the steps indicated in the *guidepost*. Each step consists of one or more tasks.

- *Prepare Upgrade Plan*: Prepare your upgrade strategy.
- *Prepare Environment*: Install the new version of the server, download and unpack the Migration Tool.
- *Export Data From Old Server*: Use the Migration Tool to export data from the old server to migration bundle file.
- *Import Data To New Server*: Use the Migration Tool to import migration bundle file to the new server. Manually migrate additional data, as needed.
- *Verify Upgrade*: Use checklists to test that upgrades were completed properly.
- *Next Steps*: Indicates what to do after the upgrade process is complete.

Prepare Upgrade Plan

Planning essential for a successful upgrade experience. As you formulate your strategy, consider these questions.

- *Should You Upgrade To This Version?*
- *What Resources Do You Need?*
- *When Should You Upgrade?*
- *How Much Time Do You Need?*
- *What Data Will Be Migrated?*
- *Where Are The Components In The Old System Located?*
- *Where Will The Components In The New System Be Located?*
- *How Will You Prepare Users For The Upgrade Process?*
- *How Will You Document Your Upgrade Plan?*

Should You Upgrade To Version 5.0?

As you determine whether to upgrade your system to this version of the software, consider these questions.

- Does this version of the software contain new features or bug fixes that your organization needs? Review *Explore What's New for Pentaho Data Integration* for more details.
- Is your current system near or at End of Life? See the Support section of the *Pentaho Website* to determine if your version of the system is still being supported.

What Resources Do You Need?

Resources are a key component in an upgrade plan. Consider both technical and human resources.

- Do your systems meet the hardware and software requirements indicated in the *Supported Components* matrix?
- Who should be involved in the upgrade process? Smaller systems that are housed on a single computer can sometimes be upgraded by one person. But, larger or geographically distributed systems might involve more people. For example, the web application server and security might be handled by one group while reporting databases might be handled by another. List those who can provide needed information or help troubleshoot problems. Remember to include stakeholders.
- Who will help verify that data has been migrated correctly? Keep in mind that some information, such as user favorites, are sometimes best verified by end users.
- You will need to download the Migration Tool from our site. You also need a text editor, and a tool to compare the differences between text files.

How Much Time Will The Upgrade Process Take?

The amount of time it takes to perform the upgrade process varies based on how much data you have and how much you have configured the old system. Allot time to install and configure the licenses for the new system, run the Migration Tool, perform manual configuration tasks, and to test and troubleshoot the system. After testing, allow time to configure the new system, deprovision the old server, and provision the new one. Use these time estimates as a rough guide.

Table 1: Upgrade Process Time Estimates

Task	Time Estimate	Description
Installing new system, licenses, and the Migration Tool	1 to 1 1/2 Hours	If you decide to use the installation wizard to new server, allot about one hour to run the installation wizard and the license, and to also download and unpack the Migration Tool. If you need to customize your DI Repository database, web application server, or both, set aside anywhere from an hour to an hour and a half to work through the custom installation

Task	Time Estimate	Description
		procedures, install licenses, and to download and unpack the Migration Tool.
Run Migration Tool	TBD.	*TBD.*
Verify Migration	1 Hour to Several Days	Plan to spend anywhere from a few hours to several days for validation testing. The amount of time that you spend depends on your content, the numbers of transformations, shared objects, and jobs that you have, and how exhaustively you want to test the system.

When Should You Upgrade?

Once you have identified your resources and have estimated how much time it might take to upgrade your system, consider appropriate timeframes for upgrade.

- Some parts of the process require that the old and new servers be offline. How much downtime does your organization allow?
- Consider performing the migration process during off-peak hours.
- Consider organizational events and seasonal reporting needs as you select a timeframe.

What Data Will Be Migrated?

It is important to understand what types of data is migrated from the old system to the new one. Much of the data from the old server can be migrated automatically using the Migration Tool, but some must be migrated manually. The Migration Tool is a standalone utility that you can download from our website. If you use the Migration Tool, it will attempt to migrate all of the data in the table. You cannot choose to exclude.

Table 2: Data Migrated by the Migration Tool

Type of Data	Data Migrated with the Migration Tool	Exceptions
Solution Content	*All solution content in the `pentaho-solutions` directory. This includes reports, properties, graphics , xactions, as well as XML files. Localization files for directories and content files. Permission settings for the solution content files.*	*The following subdirectories that are in the `pentaho-solution` directory: `Admin`, `Systems`, and `Temp`. Any Pentaho system localization data.*
Transformation and Jobs	All versions of transformations and jobs, all comments related to each version. This includes all KTR and KJB files.	
Shareable Objects	All enterprise repository shareable objects and associated connection information.	File based shareable objects that are stored in the shared.xml file, database shareable objects. Only the latest version of the shareable objects are migrated.
Schedules		
System Configuration	*Information in these files: applicationContext-security-ldap.properties, applicationContext-spring-security-hibernate.properties, pentaho.xml, quartz.properties, mondrian.properties, settings.xml for various plugins, analyzer.properties, email_config.xml. Also the following tomcat files: web.xml, context.xml, classic-engine.properties, logging.properties, log4j.xml*	*JNDI data sources named "Hibernate" or "Quartz" will not be migrated. Only the following web.xml settings are migrated: Solution-path, Base-url, Fully-qualified-server-url, Locale-language, contextClass, contextConfigLocation, pentahoObjectFactory, encoding, hsqldb-databases, Proxy Trusting Filter, TrustedIpAddrs. No custom servlets or filters will be migrated.*
User Preference Settings	*Recent files list, global settings, and whether to show the tooltips in the User Console. If you have*	

Type of Data	Data Migrated with the Migration Tool	Exceptions
	set up favorites in Mobile, favorites are migrated and appear in the Pentaho User Console.	
Users and Roles	*Usernames, passwords, username description field information, role names, role descriptions, user role assignment*	This data is only migrated if the system is using Pentaho Security to authenticate users. If not, see the Configuration section for additional detail. Note that role permissions are not migrated, regardless of authentication method, and must be manually reset.

Data that must be migrated manually includes:

- *JDBC Drivers for your data source.*
- *Memory settings. These are set during the configuration step.*
- *Changes made to security files, pentaho.xml, web.xml, and context.xml.*
- *Any additional JSPs, servlets, custom libraries, plugins, java API extensions that require a change to the pentaho-objects-spring.xml file, themes, or other resources integrated as part of the custom pentaho system.*
- *Pentaho system localization information.*
- *Generated content for deleted or otherwise inactive users.*

Where Are The Components In The Old System Located?

Identify the paths to the these components so that you can supply them when you export data from the old server.

- *Web Application Server*
- *DI Repository Database*
- *Any additional servlets, and files needed to customize your system.*

Obtain the location of design tools that are not on the server.

Where Will The Components In The New System Be Located?

Identify the paths to the these components so that you can supply them when you import data into the new server.

- *Web Application Server*
- *DI Repository Database*

Additionally, obtain the location of design tools that are not on the server.

Note: After you migrate the data from the old server into the new one, you must verify that the data in both systems match. This is easier to do if you can see the user interfaces for the old and new versions of the server at the same time.

How Will You Prepare Users For The Upgrade?

Consider how you will communicate with users, obtain feedback, and provide training.

- How will you handle customer feedback about the upgrade process?
- How you will communicate with internal and external customers?
- How much advanced notice will you provide?
- How will you train users? Note that tutorials and documentation are available in the Infocenter. Additionally, Pentaho offers professionally-led training classes on many aspects of the new system. See the Pentaho website for details.

How Will You Document Your Upgrade Plan?

Documenting your responses to the questions raised in this section can be an invaluable troubleshooting aid should you need to contact technical support.

- How will you document upgrade plan?
- Who will have access to the plan and where will it be located? How will this plan be updated throughout the process?

Prepare Environment
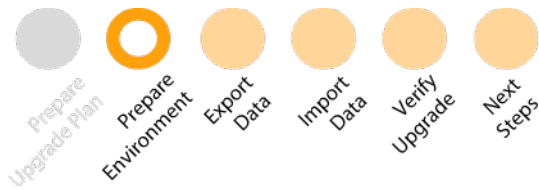
After you have *prepared your upgrade plan*, prepare your environment. To prepare your environment, do these things.

- *Install New DI Server and Licenses*
- *Download and Install Migration Tool*

Install New DI Server and Licenses

Complete these steps to install the new DI Server and licenses.

1. If you are going to instsall the new DI Server on the same machine as the old one, make sure the old version of DI Server is running as you complete these instructions. If you decide to run the Installation Wizard, the software will sense that another system is running and will make port assignments that will enable both systems to run simultaneously. This enables you to use both systems to verify that data has migrated properly near the end of the process.
2. Install the new version of the DI Server. See *Select DI Installation Options* to determine the best installation option. Make sure that you install the samples as well.
3. Ensure that *license keys* have been installed. Do not configure the system further, just install the license keys.
4. Start the DI Server and log in as the admin user. Make sure to clear your cache. Do not configure the DI Server or make any additional changes.

   Note: If you have problems with starting the DI Server, be sure to fix it before proceeding further.

Download and Install the Migration Tool

The Migration Tool automates the data migration process. Download the Migration Tool from the Pentaho Customer Support Portal. (TBD. Will there be a CE only version of this tool made available?)

1. Download the migrator.zip file from the Customer Support Portal on the new version of the server, then unzip the file.
2. If you use Windows, open the **Command Prompt** window. If you use Linux, open the **Terminal** window.
3. Install the Migration Tool by doing one of the following.

   a) If you use Windows, enter this command.

   ```
   install.bat
   ```

   b) If you use Linux, enter this command.

   ```
   ./install.sh
   ```

4. Read the EULA that appears. **Select I accept the terms of this license agreement**, then click **Next**.
5. Indicate where you want the file to be unpacked. It doesn't matter where you download the file. Click the **Next** button.
6. The **Installation Progress** window appears. Progress bars indicate the status of the process. When the process is complete, click **Quit** to exit.
7. If the old version of the server is installed on another computer, repeat these steps to install the Migration Tool on that computer.

Next Steps
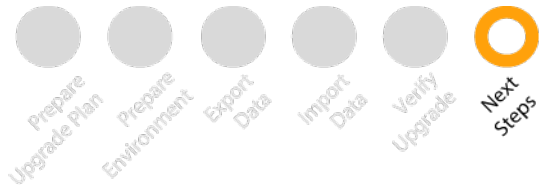
After you have *verified the upgrade* and upgraded Spoon, perform post-migration tasks. These include completing the following tasks.

- Configure the DI Server.
- To learn more about how to administer the DI Server, read Administer DI Server.
- Deprovision the old server and provision the new one in accordance with your organization's procedures.
- Review the What's New Guide if you have not done so already.
- Review Getting Started with Pentaho to learn how to use Spoon.