pentaho

Version 5.3

ExploreWhat'sNew in DI

Copyright Page

Help and Support Resources

To view the most up-to-date help content, visit https://help.pentaho.com.

If you do not find answers to your questions 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

The trademarks, logos, and service marks ("Marks") displayed on this website are the property of Pentaho Corporation or third party owners of such Marks. You are not permitted to use, copy, or imitate the Mark, in whole or in part, without the prior written consent of Pentaho Corporation or such third party. Trademarks of Pentaho Corporation include, but are not limited, to "Pentaho", its products, services and the Pentaho logo.

Trademarked names may appear throughout this website. Rather than list the names and entities that own the trademarks or inserting a trademark symbol with each mention of the trademarked name, Pentaho Corporation 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 460

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

New Features in Pentaho Data Integration 5.3

Pentaho Data Integration 5.3 delivers many exciting and powerful features that help you quickly and securely access, blend, transform, and explore data.

Simplify Deployment, Maintenance, and Configuration of Big Data Hadoop Clusters with the Named Clusters Feature

The Named Hadoop Clusters feature lets you store cluster configuration information, such as hostnames and port numbers. You can then reuse them in your HBase Input, HBase Output, Pentaho Map Reduce, Oozie Job Executor, Hadoop Job Executor, and Pig Script Executor transformation steps and job entries. With Named Hadoop Clusters, maintenance is a breeze! Simply change cluster configuration in one place. It updates wherever it is referenced when the job or transformation is opened or is run. You can even share your named Hadoop cluster configurations with any other user connected to the same DI Repository that you are using.

Related Content:

- Using Named Hadoop Clusters

Harness the Power of PDI with Our New Developer Documentation

The PDI Server API documentation is now available for developers. File management, user management, and carte web services are documented for use. The Pentaho SDK is also available for download from the Pentaho Support site.

Related Content:

- PDI Server API
- Pentaho Support Portal

New Cloudera and MapR Big Data Hadoop Distribution Support

CDH 5.2 and MapR 4.0.1 are now supported! To learn about other Hadoop Distributions we support, check out the 5.3 support matrix.

Related Content

- 5.3 Support Matrix

Better Support for High Load Environments and Large Deployments with Carte

Carte slave servers can use the same kettle variable values as the master server. This simplifies the deployment of additional server nodes because you no longer need to set up kettle.variable information for each node. You can also adjust Carte's Jetty server settings to support high load environments.

Related Content:

- Use Carte

Minor Functionality Changes

To learn more about minor functionality changes that might impact your upgrade or migration experience, see the Minor Functionality Changes for PDI 5.3 article. If you are migrating from an earlier version than 5.2, check the What's New articles and Minor Functionality changes for each intermediate version of the software.

PDI Version 5.3 Minor Functionality Changes

Manually Migrating Big Data Cluster Configurations Stored in Hadoop Steps and Entries

If you are migrating or upgrading to PDI 5.3 or greater, and you have transformations or jobs that use the following Big Data steps or entries, you might need to convert the existing cluster configuration information to use the named clusters feature.

- HBase Input
- HBase Output
- Pentaho Map Reduce
- Oozie Job Exec
- Hadoop Job Exec
- Pig Script Exec

You only need to perform the conversion process if you edit a step or entry in Spoon. Otherwise, you do not need to complete the conversion process. Note that you can continue to run scheduled transformations and jobs without the conversion, as long as you do not manually edit one of the above steps or entries.

To convert, create a new named configuration using the instructions in the Using Named Hadoop Clusters article.