Informatica

Informatica® Data Integration Hub
10.5

Installation and Configuration Guide

Informatica Data Integration Hub Installation and Configuration Guide
10.5
April 2021

© Copyright Informatica LLC 1993, 2021

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, PowerCenter, and PowerExchange are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at https://www.informatica.com/trademarks.html. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

See patents at https://www.informatica.com/legal/patents.html.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2021-06-17

Table of Contents

Preface 7

Informatica Resources 7

- Informatica Network 7
- Informatica Knowledge Base 7
- Informatica Documentation 7
- Informatica Product Availability Matrices 8
- Informatica Velocity 8
- Informatica Marketplace 8
- Informatica Global Customer Support 8

Chapter 1: Installation Overview 9

Data Integration Hub Installation 9

Additional Installation Components 10

Chapter 2: Before You Begin 11

User Accounts 12

Port Numbers 14

Chapter 3: Pre-Installation Tasks 15

Pre-Installation Tasks Overview 15

Verify the Minimum System Requirements 16

Verify the Database Requirements 17

Install the Prerequisite Software 18

Install and Configure the PowerCenter pmrep Command Line Utility 19

Configure Environment Variables 19

Create a Data Source Name for Microsoft Azure SQL Database 20

- Creating a Data Source Name on a Windows Operating System 20
- Creating a Data Source Name on a Unix Operating System 21

Set Up a Directory for the Document Store 22

Configure the Microsoft SQL Server Database 22

Chapter 4: Data Integration Hub Installation 23

Installing Data Integration Hub on a Windows Operating System 23

- Step 1. Run the Installer 23
- Step 2. Define Installation Settings 25
- Step 3. Configure Data Integration Hub Repositories 28
- Step 4. Set Up the Operational Data Store 31
- Step 5. Configure User Authentication 33
- Step 6. Configure Document Store and Web Server 38
- Step 7. Configure PowerCenter Settings 41

Step 8. Configure Informatica Data Quality Settings. . . . 45
Step 9. Complete the Installation. . . . 46
Installing Data Integration Hub on a UNIX Operating System in Console Mode . . . 47
Step 1. Run the Installer. . . . 47
Step 2. Define Installation Settings. . . . 47
Step 3. Configure Data Integration Hub Repositories. . . . 49
Step 4. Set Up the Operational Data Store. . . . 50
Step 5. Configure User Authentication. . . . 51
Step 6. Configure Document Store and Web Server . . . 53
Step 7. Configure PowerCenter Settings. . . . 55
Step 8. Configure Processing Engine Settings. . . . 56
Step 9. Complete the Installation. . . . 57
Installing Data Integration Hub in a Silent Mode. . . . 57
Configuring the Installation Properties. . . . 58
Sample of the Installer Properties Configuration. . . . 64
Running the Silent Installer. . . . 66

Chapter 5: Post-Installation Tasks. . . . 68

Post-Installation Tasks Overview. . . . 68
Register the Data Integration Hub Server Plug-in for PowerCenter. . . . 69
Set Up the Run Publication Subscription Web Service . . . 70
Importing the Web Service to PowerCenter. . . . 70
Verifying the Server Settings. . . . 71
Configure Credentials for Windows Authentication. . . . 71
Set Up Database Partitions on the Publication Repository. . . . 71
Configuring the Environment for a Hadoop Publication Repository. . . . 72
Log in to the Operation Console. . . . 75
Configure Connections to the Data Integration Hub Repositories. . . . 75
Configure Connections to Microsoft Azure SQL Database. . . . 77
Synchronize Data Integration Hub Users. . . . 78
Activate the Dashboard and Reports Component. . . . 79
Register the Dashboard and Reports License. . . . 79
Import the Operational Data Store Event Loader Workflow to PowerCenter. . . . 79
Synchronize Data Integration Hub Users. . . . 80

Chapter 6: Upgrading Data Integration Hub . . . 81

Upgrading Data Integration Hub Overview. . . . 81
Before You Upgrade. . . . 82
Upgrading Data Integration Hub on a Windows Operating System. . . . 83
Step 1. Run the Installer. . . . 83
Step 2. Define Installation Settings. . . . 84
Step 3. Configure Data Integration Hub Repositories . . . 86
Step 4. Set Up the Operational Data Store. . . . 90

Step 5. Configure the Web Server . . . 92
Step 6. Configure PowerCenter Settings. . . . 93
Step 7. Configure Data Quality Settings. . . . 96
Step 8. Complete the Installation. . . . 97
Upgrading Data Integration Hub on a UNIX Operating System. . . . 98
Step 1. Run the Installer. . . . 98
Step 2. Define Installation Settings. . . . 98
Step 3. Configure Data Integration Hub Repositories. . . . 99
Step 4. Set Up the Operational Data Store. . . . 101
Step 5. Configure the Web Server . . . 102
Step 6. Configure PowerCenter Settings. . . . 103
Step 7. Configure Processing Engine Settings. . . . 104
Step 8. Complete the Installation. . . . 105
After You Upgrade. . . . 105
Reapplying Configuration Modifications. . . . 107
Registering the Dashboard and Reports License. . . . 107
Replacing the Operational Data Store Loader Workflow. . . . 107
Configure Credentials for Windows Authentication. . . . 108
Updating the Security Configuration Files . . . 108
Updating the Run Publication Subscription Web Service API. . . . 109
Restart the Data Integration Hub Services . . . 109
Validating Invalidated Subscriptions. . . . 109

Chapter 7: Starting and Stopping Data Integration Hub. . . . 110

Overview of Starting and Stopping Data Integration Hub. . . . 110
Starting and Stopping Data Integration Hub on Windows. . . . 110
Starting and Stopping Data Integration Hub from the Start Menu. . . . 110
Starting and Stopping Data Integration Hub with Batch Scripts. . . . 111
Starting and Stopping Data Integration Hub on Linux. . . . 111

Chapter 8: Optional Data Integration Hub Configuration. . . . 112

Optional Data Integration Hub Configuration Overview. . . . 112
Modifying the Data Integration Hub Server RMI Port Number. . . . 113
Logs. . . . 113
Default Log Files. . . . 114
Customizing the Destination for Log Messages. . . . 114
Changing the Maximum Java Heap Size. . . . 117
Changing the Credentials for a Database User Account. . . . 118
Changing the Credentials for the Data Integration Hub Hadoop Service. . . . 119
Updating the Dashboard Configuration File. . . . 120
Configuring a PowerCenter Integration Service to Access Data Integration Hub. . . . 121
Configuring Repository Connections on PowerCenter Version 10. . . . 122

Chapter 9: Installing and Configuring the Data Integration Hub Accelerator for Data Archive 123

Installing and Configuring Data Integration Hub Accelerator for Data Archive Overview 123

Pre-Installation Steps 123

Installing the Data Integration Hub Accelerator for Data Archive 124

Configuring the Data Archive Source Connection 125

Chapter 10: Troubleshooting the Data Integration Hub Installation 126

Troubleshooting the Data Integration Hub Installation Overview 126

Troubleshooting the Data Integration Hub Hadoop Service 126

Installation Log Files 129

Chapter 11: Uninstallation 130

Uninstallation Overview 130

Uninstalling Data Integration Hub from Windows Operating Systems 130

Uninstalling Data Integration Hub from UNIX Operating Systems 131

Index 132

Preface

Follow the instructions in the *Data Integration Hub Installation and Configuration Guide* to install and configure Data Integration Hub. The guide also includes information about the pre-install tasks and post-install tasks that you need to perform to complete the installation.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit https://network.informatica.com.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at https://network.informatica.com/community/informatica-network/product-availability-matrices.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at http://velocity.informatica.com. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
https://www.informatica.com/services-and-training/customer-success-services/contact-us.html.

To find online support resources on the Informatica Network, visit https://network.informatica.com and select the eSupport option.

CHAPTER 1

Installation Overview

This chapter includes the following topics:

- Data Integration Hub Installation, 9
- Additional Installation Components, 10

Data Integration Hub Installation

Data Integration Hub consists of the core application component and additional required and optional components. If you install Data Integration Hub, PowerCenter® services, and the PowerCenter Client on the same machine, you can select all of the components in the installer. Otherwise, install each component on the required machine.

The Data Integration Hub installation includes the following components:

Data Integration Hub

Core application component. Includes the Operation Console, Data Integration Hub server, the Data Integration Hub repository, and the publication repository. The PowerCenter services must be running when you install Data Integration Hub. You must set up the database user accounts before you install this component.

Data Integration Hub Dashboard and Reports

You can make the Dashboard available in Data Integration Hub in two ways. When you install Data Integration Hub, you can enable the dashboard that uses operational data store. Otherwise, you can use the dashboard that uses metadata repository which is available by default.

This is a business activity monitoring component. Includes the dashboard application and the operational data store repository. You must set up a different user account from the user account that you use for the Data Integration Hub repository.

You must install the Data Integration Hub component to install this component.

Data Integration Hub PowerCenter server plug-in

PowerCenter repository plug-in that Data Integration Hub uses to run Data Integration Hub transformations in PowerCenter. The installation includes files to add to the classpath of the PowerCenter Integration Service. You must install this plug-in on the same machine as the PowerCenter services.

After you install this component, you must register the plug-in to the PowerCenter repository before you create and run Data Integration Hub workflows.

Data Integration Hub PowerCenter Client plug-in

PowerCenter Client plug-in that displays Data Integration Hub transformation properties in PowerCenter mappings. You install this plug-in on all PowerCenter Client machines that you plan to use to build mappings and workflows for Data Integration Hub transformations.

Data Integration Hub Hadoop Service

Available when you install Data Integration Hub on a Linux operating system.

Connecting module between Data Integration Hub and the Hadoop cluster. Enables Data Integration Hub to perform operations on the Hadoop publication repository.

Select this component if you want to define a Hadoop-based publication repository and manage some of your topics on Hadoop.

Note: The Data Integration Hub Hadoop Service must be installed on an edge node in the Hadoop cluster. Unless you are installing Data Integration Hub on a Hadoop edge node, do not select this component during Data Integration Hub installation. In this case, after you complete the Data Integration Hub installation, run the installation again on the edge node and select to install only the Data Integration Hub Hadoop Service component on the node.

Data Integration Hub Data Engineering Integration

Enables Data Integration Hub to use Data Engineering Integration mappings for custom publications and subscriptions that publish and subscribe to a Hadoop-based publication repository.

Available when you install Data Integration Hub on a Linux operating system.

Data Integration Hub Informatica Data Quality

Enables Data Integration Hub to use Data Quality mappings for custom publications and subscriptions from and to on-premises applications.

Additional Installation Components

Data Integration Hub requires additional components to run. The components are installed as part of the Data Integration Hub installation process.

The Data Integration Hub includes the following additional applications and components:

Data Integration Hub Server

Server environment that manages publication and subscription processing in Data Integration Hub.

Operation Console

Web interface to customize and monitor processing, manage users, and set preferences.

Apache Tomcat

Web server environment that runs the Operation Console client.

Java JDK

Java run-time environment in which the Data Integration Hub server, Data Integration Hub Operation Console, and Data Integration Hub command line client tools run.

CHAPTER 2

Before You Begin

This chapter includes the following topics:

- User Accounts, 12
- Port Numbers, 14

User Accounts

Before you install, verify that you have the user names and passwords for the required database and domain accounts.

The following table describes the required user accounts:

User Account	Description

Database	Database user account that you use to log in to the database server and create tables and views for the Data Integration Hub repository and the publication repository. If you install the Dashboard and Reports component, you also use a user account for the operational data store.
	You must install all the repositories on the same type of database server. You must create a separate user account for each repository.
	The user accounts must have privileges to perform the following actions:
	- Select data from tables and views.
	- Insert data into tables, delete data from tables, and update data in tables.
	- Create, change, and delete the following elements:
	  - Tables
	  - Views
	  - Synonyms
	  - Indexes
	  - Custom data types
	  - Triggers
	- Create, change, delete, and run stored procedures and functions.
	If you use a Microsoft SQL Server database, you must set up separate databases for each repository. It is recommended that you grant database owner privileges to the user accounts.

If you use Data Integration Hub with Informatica domain authentication: Informatica domain administrator	Administrator account for the Informatica domain.

User Account	Description
If you use Data Integration Hub with Informatica domain authentication: Informatica security domain	User account for Informatica domain authentication. The user account must be created in the Informatica Administrator tool with the **manage roles/groups/users** privileges. The Data Integration Hub administrator synchronizes the user account after the installation.
PowerCenter Repository Service	User account that Data Integration Hub uses to perform operations in the PowerCenter Repository Service. The user accounts must have privileges and permissions to perform the following actions: **General** - Access Repository Manager privilege **Folders** - Read on folder permission - Create privilege - Copy privilege **Design Objects** - Read on folder permission - Read on shared folder permission - Read and Write on destination folder permission - Create, Edit, and Delete privilege with the Read on original folder, Read and Write on destination folder, and Read and Write on folder permissions **Sources and Targets** - Read on folder permission - Read on shared folder permission - Read and Write on destination folder permission - Create, Edit, and Delete privilege with the Read on original folder, Read and Write on destination folder, and Read and Write on folder permissions **Run-time Objects** - Read on folder permission - Create, Edit, and Delete privilege with the Read on original folder, Read and Write on destination folder, Read on connection object, and Read and Write on folder permissions - Monitor privilege with the Read on folder permission - Execute privilege with the Read and Execute on folder permission **Global Objects** - Read on connection object permission - Read and Write on connection object - Create Connections privilege - Execute privilege with the Read and Execute on folder permission

Port Numbers

The installer sets the default port numbers for the installation components. If another application uses the same port number as one of the installation components, a port conflict might prevent the component from running correctly or cause errors.

You can change the port numbers after installation. Before you start Data Integration Hub, verify that the port numbers do not conflict with other applications and change the port numbers in Data Integration Hub to prevent port conflicts.

The following table describes the default port numbers:

Port Number	Description
18000	UDP multicast port that Data Integration Hub uses for internal communications.
18005	Operation Console shutdown port. Only required to be available on the machine where Data Integration Hub is installed.
18050	Port that the Operation Console uses for internal communications.
18080	Operation Console HTTP port. Required only if you use an HTTP port for the Operation Console.
18095 and 18096	RMI ports that the Operation Console and PowerCenter workflows use to communicate with the Data Integration Hub server.
18100	Port that the Data Integration Hub server uses for internal communications.
18443	Operation Console HTTPS port. Required only if you use an HTTPS port for the Operation Console.
18616	Port for the Data Integration Hub JNDI provider.

CHAPTER 3

Pre-Installation Tasks

This chapter includes the following topics:

- Pre-Installation Tasks Overview, 15
- Verify the Minimum System Requirements, 16
- Verify the Database Requirements, 17
- Install the Prerequisite Software, 18
- Install and Configure the PowerCenter pmrep Command Line Utility, 19
- Configure Environment Variables, 19
- Create a Data Source Name for Microsoft Azure SQL Database, 20
- Set Up a Directory for the Document Store, 22
- Configure the Microsoft SQL Server Database, 22

Pre-Installation Tasks Overview

Before you install Data Integration Hub, set up the machines to meet the installation requirements, verify that you have all the user account credentials, and prepare your environment for installing and running Data Integration Hub.

Note: Data Integration Hub and the PowerCenter Integration Service that Data Integration Hub uses must be installed on the same type of operating system. Both must be installed either on a machine or machines that are running Windows operating systems, or on a machine or machines that are running non-Windows operating systems.

The following components must reside on machines with the same locale and the same time zone:

- Data Integration Hub server
- Data Integration Hub repositories
- Data Integration Hub Operation Console clients
- PowerCenter Repository Service that Data Integration Hub uses
- PowerCenter Integration Service that Data Integration Hub uses

Verify the Minimum System Requirements

Verify that your system meets the minimum requirements.

The following table describes the minimum system requirements:

System	Requirement
Operating system	- Microsoft Windows - IBM AIX - Red Hat Linux - SUSE Linux
Processor	- Minimum: 2 CPU cores - Recommended: 4 CPU cores
RAM	8 GB
Disk space	- Minimum: 3 GB - Recommended: 8 GB
Browser	- Microsoft Internet Explorer - Google Chrome - Microsoft Edge

The following table describes the minimum system requirements to run the installer:

System	Requirement
RAM	512 MB
Disk space	1 GB

For more information about product requirements and supported platforms, see the Product Availability Matrix on Informatica Network:
https://network.informatica.com/community/informatica-network/product-availability-matrices

Verify the Database Requirements

Verify that your database meets the requirements for running Data Integration Hub.

The following table describes the database requirements for Data Integration Hub:

Database Component	Description
Database System	Type of database on which to install the repositories. You can use one of the following database systems: - Amazon RDS for Oracle - Oracle - Microsoft SQL Server - Microsoft Azure SQL Server **Note:** You must have the Oracle Enterprise Edition to use Filter Accelerator
Database Instances	Two database instances. Data Integration Hub uses one database for the Data Integration Hub repository and one database for the publication repository. Both database instances must be on the same type of database system. If you use Oracle databases, it is recommended that each database has a unique user account and a unique schema. If you use Microsoft Azure SQL Server, it is recommended that you create a separate Data Source Name (DSN) for each database. If you install the Dashboard and Reports component, an additional database instance is required for the operational data store. The operational data store must be on the same type of database system as the Data Integration Hub repositories. **Note:** If you install the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, you must configure the repository connections in PowerCenter Workflow Manager. For details, see "Configuring Repository Connections on PowerCenter Version 10" on page 122.
Disk space	The Data Integration Hub repository database requires at least 512 MB of disk space for the core application. You also need additional space on the publication repository database based on the number of publications and publication instances that you need to retain. **Note:** Unicode data requires twice as much storage than single-byte character sets.
Database connections	Multiple database connections for each repository must always be available. The number of required connections for each repository depends on the number of publications and subscriptions that run concurrently. Use the following formula to calculate the number of required database connections for each repository: NumberOfConcurrentPublicationsOrSubscriptions X 3 + 2 If you do not have enough database connections available, Data Integration Hub might fail or encounter database deadlocks.

Database Unicode Support

If you require Unicode support, create the Data Integration Hub repository database and the Data Integration Hub publication database with the following settings:

- Oracle databases: use the AL32UTF8 Unicode character set.
- Microsoft SQL Server: it is recommended that you use data types that support Unicode data: nchar, nvarachar, and ntext.

Microsoft SQL Server Collation

If you use Microsoft SQL Server or Microsoft Azure SQL Server, the collation for the Data Integration Hub repository and for the Data Integration Hub publication repository must not be case sensitive.

Install the Prerequisite Software

Install the prerequisite software on your machine.

- PowerCenter. Install PowerCenter before you install Data Integration Hub. Make sure to install PowerCenter services on a machine that is accessible to Data Integration Hub. After you install PowerCenter, verify that the PowerCenter Web Services Hub is running.
  If you do not install the PowerCenter services on the same machine that you install Data Integration Hub, install the PowerCenter pmrep command line utility on the machine where you install Data Integration Hub. Verify that Data Integration Hub and PowerCenter can be accessed with the same drive and file path.

  Note: Verify that repository agent caching on the PowerCenter Repository Service is disabled. In Informatica Administrator, access the advanced properties of the Repository Service and verify that the option **Enable Repagent caching** is cleared.
- Java Development Kit (JDK). On IBM AIX operating systems, install the IBM JDK version 8.0.5.16 (8.0 Service Refresh 5 Fix Pack 16) and configure the INFA_JDK_HOME environment variable before you install Data Integration Hub. Verify that the login shell can access the INFA_JDK_HOME environment variable.
  For more information about Java installation, see the Java website at the following address: https://www.ibm.com/developerworks/java/jdk/fixes/8/index.html
  The software available for download at the referenced links belongs to a third party or third parties, not Informatica LLC. The download links are subject to the possibility of errors, omissions or change. Informatica assumes no responsibility for such links and/or such software, disclaims all warranties, either express or implied, including but not limited to, implied warranties of merchantability, fitness for a particular purpose, title and non-infringement, and disclaims all liability relating thereto.

  For more information about product requirements and supported platforms, see the Product Availability Matrix on Informatica Network:
  https://network.informatica.com/community/informatica-network/product-availability-matrices
- Microsoft Visual C++ 2008 Redistributable Package (x86). Install this package if you use the Data Integration Hub PowerCenter Client plug-in on a Windows Server 2008 64-bit operating system.
  The software available for download at the referenced links belongs to a third party or third parties, not Informatica LLC. The download links are subject to the possibility of errors, omissions or change. Informatica assumes no responsibility for such links and/or such software, disclaims all warranties, either express or implied, including but not limited to, implied warranties of merchantability, fitness for a particular purpose, title and non-infringement, and disclaims all liability relating thereto.
- Java Cryptography Extension (JCE). Install this package if you are installing Data Integration Hub on an IBM AIX operating system.

Install and Configure the PowerCenter pmrep Command Line Utility

If you do not install the PowerCenter services on the same machine that you install Data Integration Hub, you install and configure the PowerCenter pmrep command line utility on the machine where you install Data Integration Hub.

To download the utility, contact Informatica Shipping. The utility version must match the PowerCenter version.

1. Extract the ZIP file on your local machine to a directory that is accessible by the Data Integration Hub installer.

   By default, the installer searches for the utility in the following directory: `<LocalDrive>\Informatica\version`

2. Configure the utility settings based on your operating system.

   For information about the utility settings, see the *Informatica Repository Guide*.

To test the utility settings, run the utility from the command line and verify that no errors appear in the run results.

Note: If you upgrade the pmrep command line utility at a later time, clean up all CNX files from the `Tmp` folder on your home directory.

Configure Environment Variables

After you install PowerCenter or the PowerCenter pmrep command line utility on the machine where you install Data Integration Hub, configure environment variables.

1. Set the INFA_HOME environment variable to point to the Informatica installation directory.
2. Set the INFA_DOMAINS_FILE environment variable to the path and the file name of the `domains.infa` file.
3. On Solaris and Linux, add `<INFA_HOME>/server/bin` to the LD_LIBRARY_PATH environment variable.
4. On AIX, add `<INFA_HOME>/server/bin` to the LIBPATH environment variable.
5. Verify that the pmrep utility code page matches the PowerCenter Repository Service code page. You specify the code page in the INFA_CODEPAGENAME environment variable of the utility.
6. To reduce the length of time to wait before the pmrep utility reports an error when connecting to PowerCenter, change the value of the INFA_CLIENT_RESILIENCE_TIMEOUT environment variable in the utility.

   The default timeout waiting time is 180 seconds.

7. On Linux, to use SQL repositories installed on Windows as the Data Integration Hub and publication repositories, set the ODBCINST environment variable to `<INFA_HOME>/ODBC7.1/odbcinst.ini`.

   Note: This step is relevant to systems running PowerCenter version 10 or higher.

Create a Data Source Name for Microsoft Azure SQL Database

If you install Data Integration Hub repositories on a Microsoft Azure SQL database, create data source names on the operating system of the machine on which PowerCenter is installed.

Create data source names on the operating system as provided in the following sections:

- "Creating a Data Source Name on a Windows Operating System" on page 20
- "Creating a Data Source Name on a Unix Operating System" on page 21

Creating a Data Source Name on a Windows Operating System

This section describes how to create a data source name on the Windows operating system on which Data Integration Hub is installed.

Note: Ensure that you configure a separate Data Source Name (DSN) entry for every database on which Data Integration Hub is installed.

1. Open **ODBC Data Sources** from the Windows environment and select **64-bit ODBC Data Source for DataDirect 7.1 New SQL Server Wire Protocol**.

   The **ODBC DataSource Administrator** window is displayed.
2. Select **System DSN** and click **Add New**.

   The **Create New DataSource** window is displayed.
3. Select the following driver: **DataDirect 7.1 New SQL Server Wire Protocol**.

   The **ODBC SQL Server Wire Protocol Driver Setup** window is displayed.
4. Enter the following details in the **General** tab:
   - **DataSource Name**. Name of the data source.
   - **Host Name**. Host name of the machine where the database server is installed.
   - **Port Number**. Port number of the database. The default port number for Microsoft Azure SQL Database is 1433.
   - **Database**. Name of the database instance.
5. Select **Advanced** > **Extended Options** and add **WorkArounds2=2**.

   Note: Enabling the WorkArounds2=2 option, causes the driver to ignore the column size, decimal digits, or datetime values specified by the application and use the database defaults instead. Some applications incorrectly specify the column size or decimal digits when binding timestamp parameters.
6. Select **Security** and update the following information:
   - **User Name**. Name of the Microsoft Azure SQL Database user.
   - Select any **Encryption Method**.
7. Click **Save**.

   Data Source Name (DSN) details are saved.
8. Test the driver with credentials of the user that you have provided in the procedure and ensure that the connection passes.

Creating a Data Source Name on a Unix Operating System

This section describes how to create data source names on a Unix operating system, on which Data Integration Hub is installed.

Note: Ensure that you configure a separate Data Source Name (DSN) entry for every database on which Data Integration Hub is installed. The driver should point to the `DWsqls27.so` driver by using an absolute path.

Perform the following steps in the PowerCenter installation directory:

1. Navigate to the `<pwc_install_path>/ODBC7.1` folder in the PowerCenter installation directory and edit the `odbc.ini` driver to update the following information:
   - **Driver**. Enter a path to the driver.
     An example of the path is as follows: `/data/Informatica/10.1.1/ODBC7.1/lib/DWsqls27.so`.
   - **Description**. Enter the description of the DSN entry.
     An example of the description is as follows: `Azure SQL DATABASE Connection for ODL`
   - **Address**. Enter the host name of the machine where the database server is installed.
   - **LogonID**. Enter name for the Microsoft Azure SQL Database user.
   - **Password**. Enter a password for the Microsoft Azure SQL Database user.
   - **QuotedId**. Select `No`.
   - **AnsiNPW**. Select `Yes`.
   - **EncryptionMethod**. Enter a numerical value that corresponds to the encryption method that you want to select.
   - Enter `WorkArounds2=2`.

     Note: Enabling the WorkArounds2=2 option causes the driver to ignore the column size or decimal digits specified by the application and use the database defaults instead. Some applications incorrectly specify the column size or decimal digits when binding timestamp parameters.
   - If the `$ODBINI` environment variable pointing to the `odbc.ini` file was not configured, then configure the `$ODBINI` environment variable as follows:`$ODBCINI=<pwc_install_path>/ODBC7.1/odbc.ini`.

     Note: ODBC environment variables are configured before installing PowerCenter.

   An example of the configuration is as follows:

   ```
   Driver=<PwC_Install_Loc>/ODBC7.1/lib/DWsqls27.so
   Description=Azure SQL DATABASE Connection for ODL
   Address=<server_name>
   Database= <db_name>
   LogonID=<usn>
   Password=<pwd>
   QuotedId=No
   AnsiNPW=Yes
   EncryptionMethod=1
   ValidateServerCertificate=0
   WorkArounds2=2
   ```
2. If you configure the environment variable while creating data source name on the operating system as described in this procedure, then restart PowerCenter services.

Set Up a Directory for the Document Store

Set up a directory for the Data Integration Hub document store.

The document store directory must be accessible to Data Integration Hub, Apache Tomcat, and PowerCenter with the same drive and file path.

Configure the Microsoft SQL Server Database

If you install Data Integration Hub repositories on a Microsoft SQL Server database, enable the `READ_COMMITTED_SNAPSHOT` database option. If you install the Dashboard and Reports component, enable the `READ_COMMITTED_SNAPSHOT` option on the operational data store as well.

Note: If you use Microsoft SQL Server 2012, you can set the option **Is read committed snapshot on** in Microsoft SQL Server Management Studio to **true** instead.

1. Open an SQL query for the database server with rights to set database options.
2. Run the following SQL statements:

```
ALTER DATABASE [<database_name>] SET SINGLE_USER WITH ROLLBACK IMMEDIATE
```

3. Run the following SQL query:

```
ALTER DATABASE <database_name> SET READ_COMMITTED_SNAPSHOT ON
```

4. To verify that this option is set, run the following SQL query:

```
SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name =
'<database_name>'
```

   If the option is set, the query returns the value 1. If the option is not set, the query returns the value 0.

5. Run the following SQL statement to forcefully disconnect all users from the system:

```
ALTER DATABASE [<database_name>] SET MULTI_USER
```

CHAPTER 4

Data Integration Hub Installation

This chapter includes the following topics:

- Installing Data Integration Hub on a Windows Operating System, 23
- Installing Data Integration Hub on a UNIX Operating System in Console Mode , 47
- Installing Data Integration Hub in a Silent Mode, 57

Installing Data Integration Hub on a Windows Operating System

Install Data Integration Hub on Windows operating systems in graphical mode. On UNIX operating systems, install Data Integration Hub in console mode.

Before you install, verify that your environment meets the minimum system requirements, perform the pre-installation tasks, and verify that the PowerCenter services are running.

Note: During the installation, Data Integration Hub saves log files in the home directory of the user in the subdirectory named `DXLogs`. If the installation does not complete successfully, you can view the log files in this location.

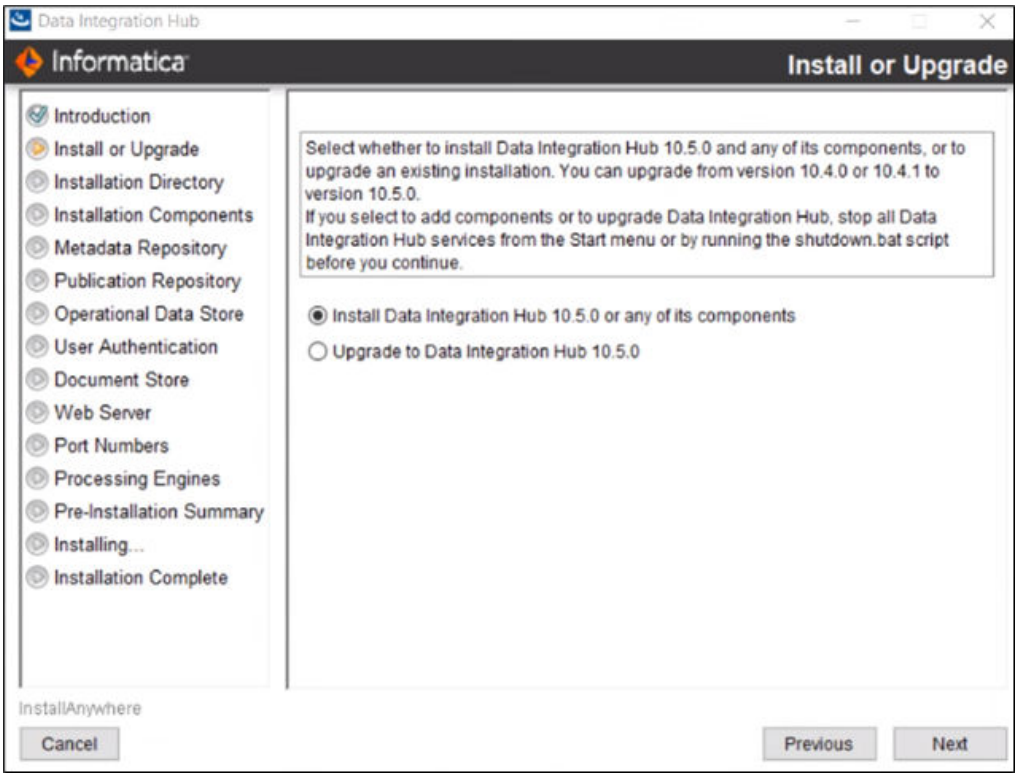
Step 1. Run the Installer

1. Log in to the machine with the user account that you want to use to install Data Integration Hub.

   To prevent permission errors, use the same account to install Data Integration Hub and PowerCenter.
2. Close all other applications.
3. Run `Install.exe` from the directory where you downloaded the installer.

   The **Introduction** page appears.
4. Read the instructions, and then click **Next**.

The **Install or Upgrade** page appears.

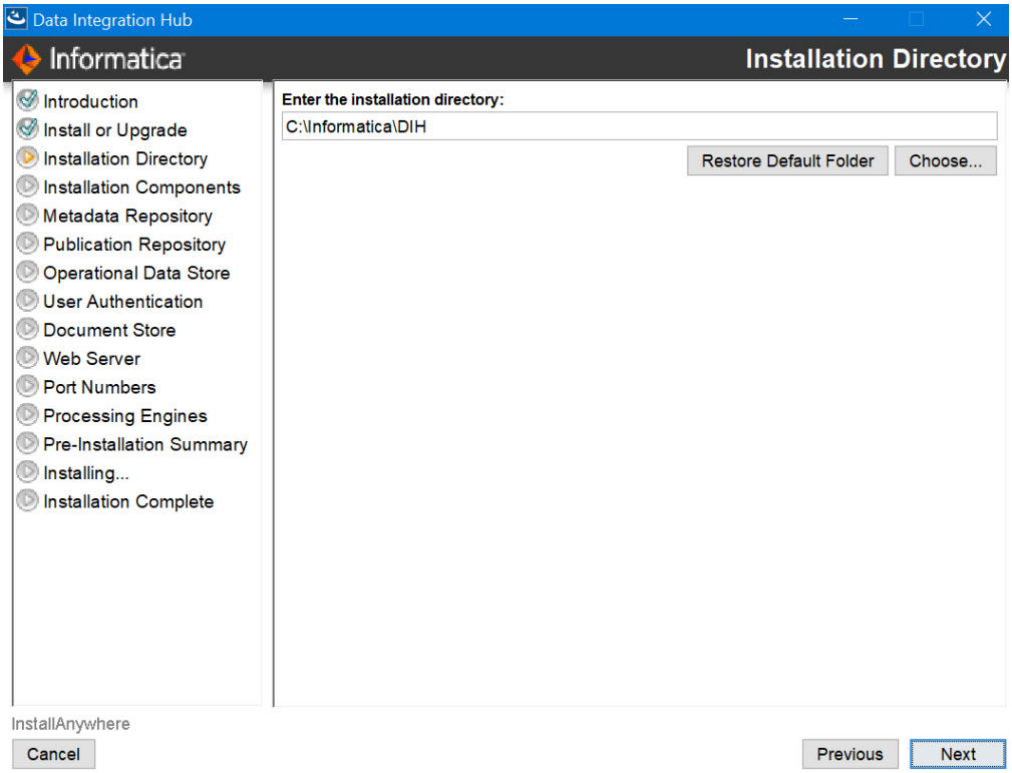


5. Select the option to install Data Integration Hub, and then click **Next**.

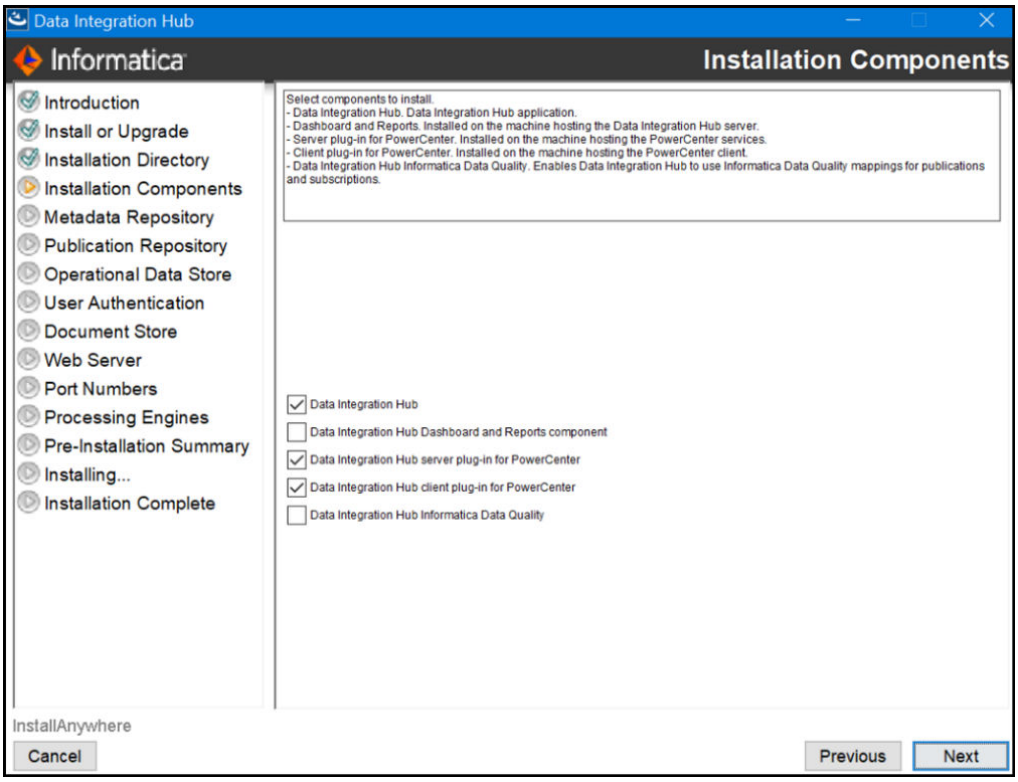
   The **Installation Directory** page appears.

Step 2. Define Installation Settings

1. On the **Installation Directory** page, enter the absolute path to the installation directory or accept the default directory, and then click **Next**.



The **Installation Components** page appears:



2. Select the components to install:

Data Integration Hub

Installs the core Data Integration Hub application.
Selected by default.

Data Integration Hub Dashboard and Reports

Installs the Data Integration Hub Dashboard and Reports component.
You can make the Dashboard available in Data Integration Hub in two ways. When you install Data Integration Hub, you can enable the dashboard that uses operational data store. Otherwise, you can use the dashboard that uses metadata repository which is available by default.

Select this component to view the reports in the Dashboard using operational data store.You must install Data Integration Hub to install the Dashboard and Reports component.
Cleared by default.

Note:

- The Dashboard using metadata directory is installed by default. Select this component to view the reports in the Dashboard using operational data store. After installation, you must ensure to set the `dx.dashboard.ods.page.show` system property to `true` to view the reports in the Dashboard using operational data store. For more information, see the *Data Integration Hub Administrator Guide* and *Data Integration Hub Operator Guide*.
- If you install the Dashboard and Reports component, you must import the operational data store event loader after you install Data Integration Hub.

- If you install the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, you must configure the repository connections in PowerCenter Workflow Manager. For details, see "Configuring Repository Connections on PowerCenter Version 10" on page 122.

Data Integration Hub PowerCenter server plug-in

Installs the Data Integration Hub plug-in for the PowerCenter services. After the installation, you register the plug-in to the PowerCenter repository.
Selected by default.

Data Integration Hub PowerCenter client plug-in

Installs the Data Integration Hub plug-in for the PowerCenter Client. Install this component on every machine that runs the PowerCenter Client.
Selected by default.

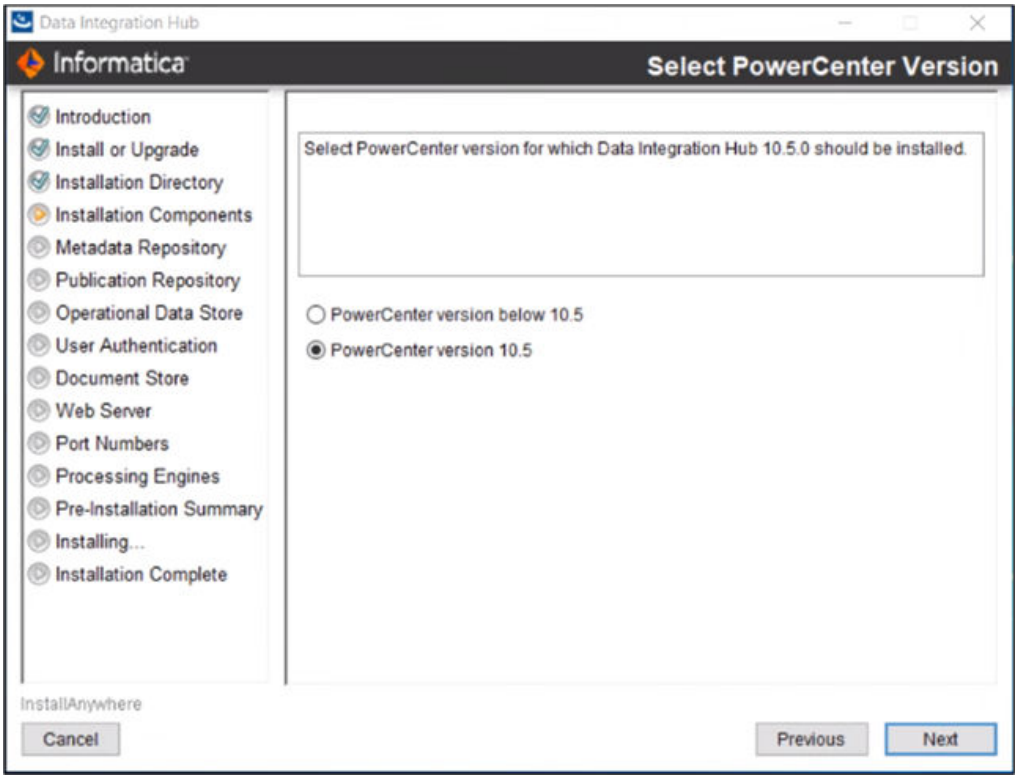
Data Integration Hub Informatica Data Quality

Enables Data Integration Hub to use Data Quality mappings for custom publications and subscriptions.
Select this component if you want to create custom publications and subscriptions that use Data Quality mappings in Data Integration Hub.

Cleared by default.

3. Click **Next**.

   The **Select PowerCenter Version** page appears.

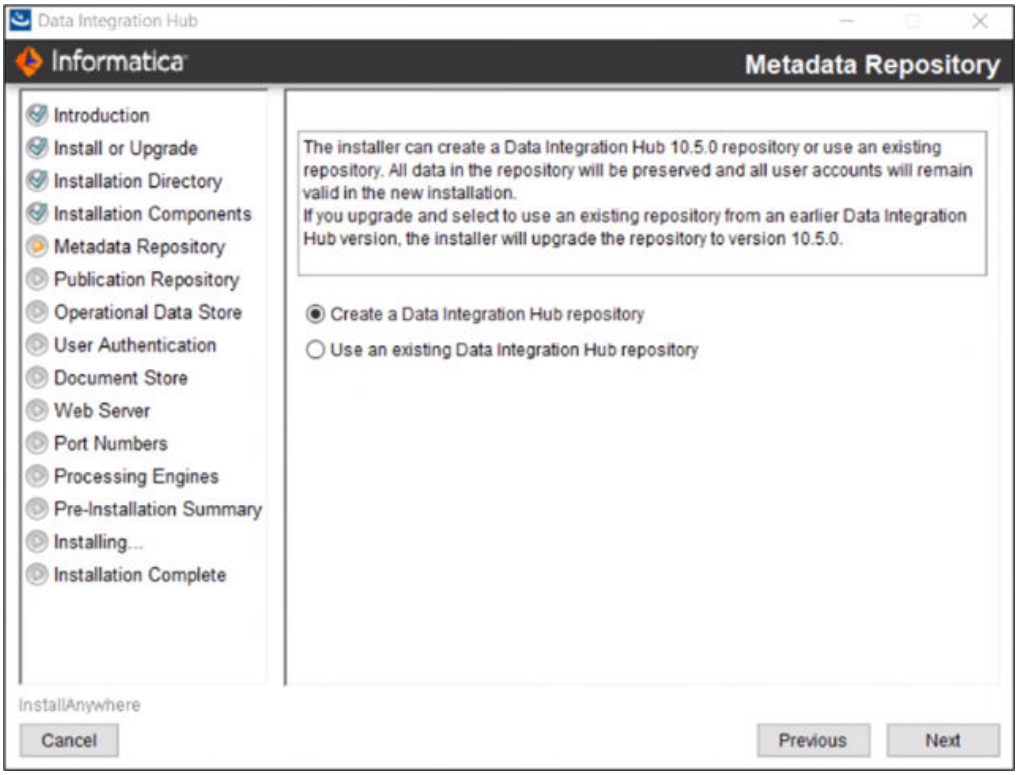


4. Select the PowerCenter version for which to install Data Integration Hub and then click **Next**.

   The **Metadata Repository** page appears.

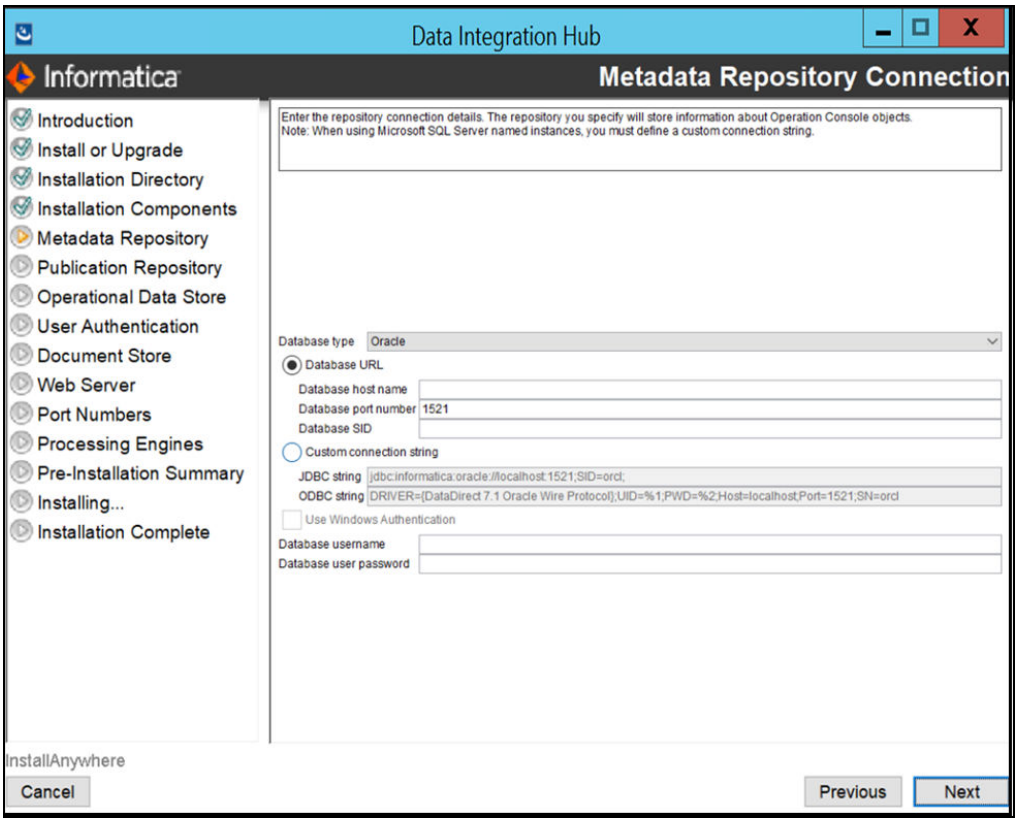
Step 3. Configure Data Integration Hub Repositories

1. On the **Metadata Repository** page, select one of the following options:
   - **Create a Data Integration Hub repository**. Creates a repository in the database.
   - **Use an existing Data Integration Hub repository**. Uses the tables and data in an existing Data Integration Hub repository and upgrades the repository.



2. Click **Next**.

The **Metadata Repository Connection** page appears.



3. Enter values in the following fields:

Database type

Type of database to use for the Data Integration Hub metadata repository. You can choose one of the following options:

- Oracle
- Microsoft SQL Server

Database URL

Location of the database.
If you select this option, enter the values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port**. Port number for the database. The default port number for Oracle is 1521. The default port number for Microsoft SQL Server or Microsoft Azure SQL Database is 1433.
- **Database SID**. System identifier for the database if the database is Oracle. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name.

Custom Connection String

Connection string to the database.

If you select this option, enter values in one of the following fields:

- **JDBC string**. JDBC connection string to the metadata repository.
- **ODBC string**. ODBC connection string to the metadata repository. Available if you install the PowerCenter Client plug-in. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

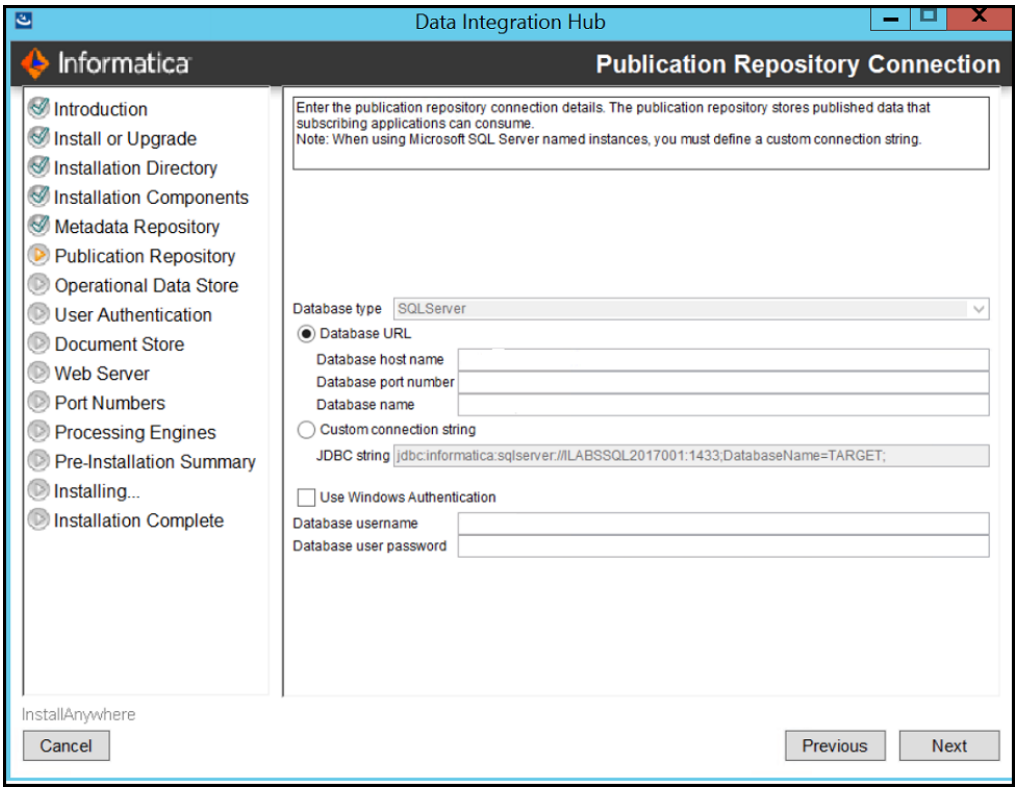
Name of the database user account for the database where you do not use Windows authentication.

Database user password

Password for the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

4. Click **Next**.

The **Publication Repository Connection** page appears.



5. Enter values in the following fields:

Database type

Type of database to use for the publication repository. The database type must match the Data Integration Hub metadata repository database type and appears in read-only mode.

Database URL

Location of the database.
If you select this option, enter values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database 1433.
- **Database SID**. System identifier for the database if the database is Oracle. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Name of the database instance.

Custom Connection String

JDBC connection string to the database.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

Name of the database user account for the database where you do not use Windows authentication.

Database user password

Password of the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

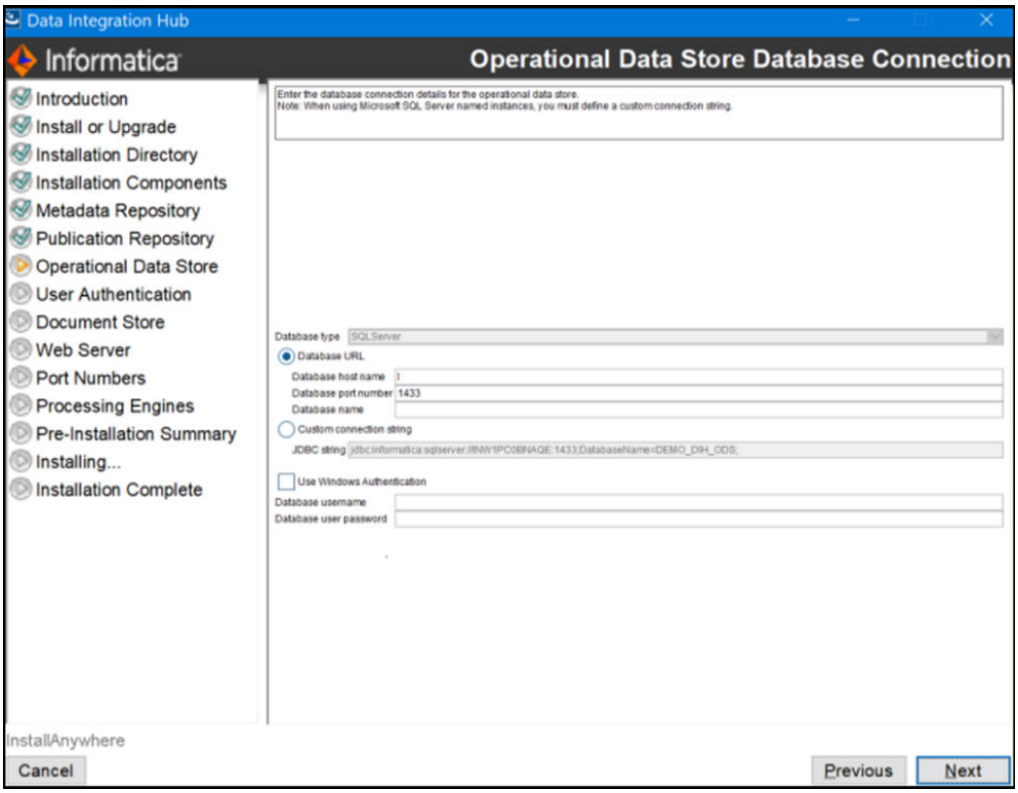
6. Click **Next**.

   If you selected the **Data Integration Hub Dashboard and Reports** component, the **Operational Data Store** page appears. If you did not select the Dashboard and Reports component, go to "Step 5. Configure User Authentication" on page 33.

Step 4. Set Up the Operational Data Store

1. This applies only if you want to use the legacy dashboard. On the **Operational Data Store** page, select one of the following options:
   - **Create an operational data store repository**. Creates an operational data store repository in the database.
   - **Use an existing operational data store repository**. Uses the tables and data in an existing operational data store repository.
2. Click **Next**.

The **Operational Data Store Database Connection** page appears.



3. Enter values in the following fields:

Database URL

Location of the database. If you select this option, enter the values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port number**. Port number for the database. The default port number for an Oracle database is 1521. The default port number for a Microsoft SQL server or a Microsoft Azure SQL database is 1433.
- **Database SID**. System identifier for the database if you select Oracle as the database. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name.

Custom Connection String

Connection string to the database. If you select this option, enter values in one of the following fields:

- **JDBC string**. JDBC connection string to the Operational Data Store.
- **ODBC string**. ODBC connection string to the Operational Data Store. Available if you install the PowerCenter Client plug-in. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

Name of the operational data store user account for the database where you do not use Windows authentication.

Database user password

Password for the operational data store account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

4. Click **Next**.

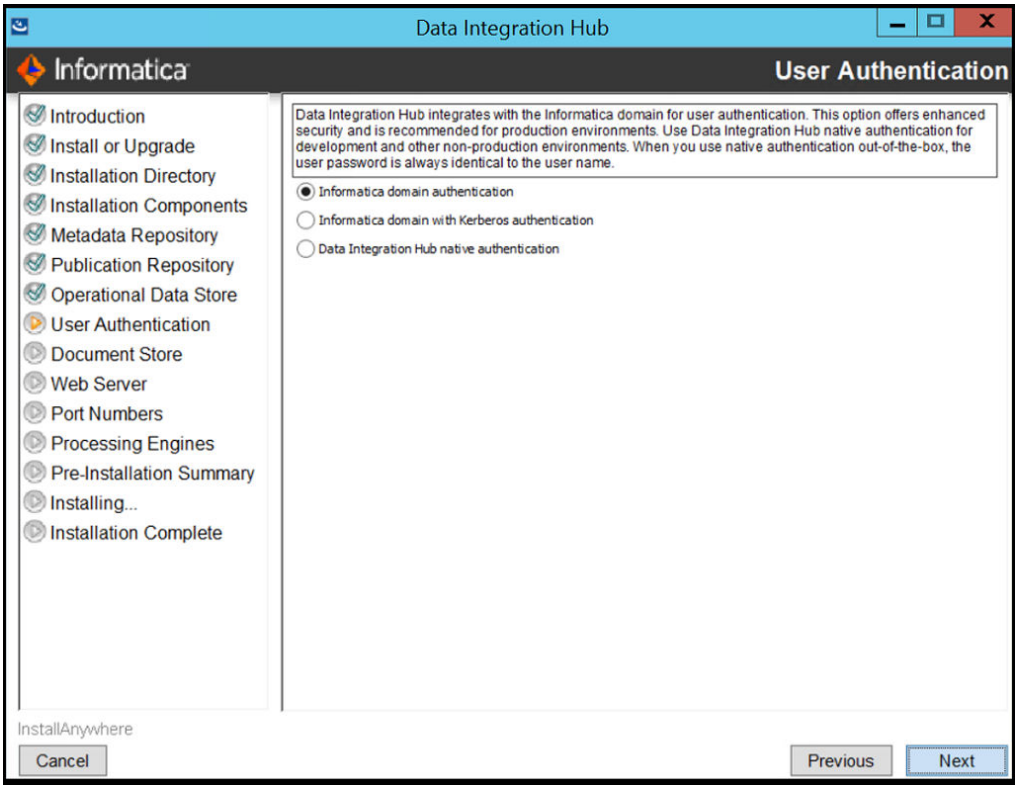
   The **User Authentication** page appears.

Step 5. Configure User Authentication

1. On the **User Authentication page**, choose the type of user authentication that you want to use.
   - Choose **Informatica domain authentication** to manage user credentials in the Informatica domain and synchronize user information with Data Integration Hub. Use Informatica domain authentication for production environments. For more information, see "Configure Settings for Informatica Domain Authentication " on page 34.

     Note: If your Informatica domain uses Kerberos authentication, choose the option **Informatica domain with Kerberos authentication**.
   - Choose **Informatica domain with Kerberos authentication** if the Informatica domain uses Kerberos authentication. Use Informatica domain with Kerberos authentication for production environments. For more information, see "Configure Settings for Informatica Domain with Kerberos Authentication" on page 36.

- Choose **Data Integration Hub native authentication** to manage user credentials locally in Data Integration Hub. Use native authentication in development and staging environments. For more information, see "Configure Settings for Data Integration Hub Native Authentication" on page 37.



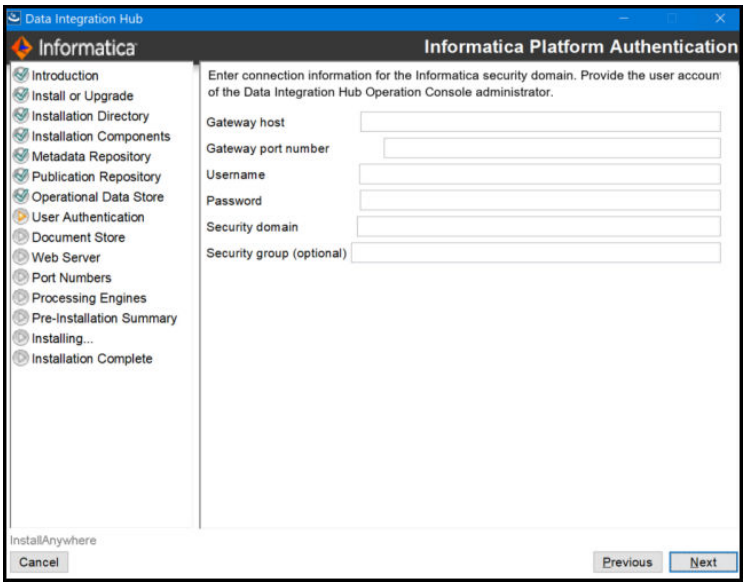
2. Enter the authentication information, and then click **Next**.

   The **Data Integration Hub Document Store** page appears.

Configure Settings for Informatica Domain Authentication

If you select the **Informatica Domain Authentication** option on the **User Authentication** page, you can configure the Informatica domain authentication settings on the **Informatica Platform Authentication** page.

The following image shows the **Informatica Platform Authentication** page.



The following table describes the settings that you need to configure for the **Informatica Platform Authentication** page:

Gateway host

Host name of the Informatica security domain server. Data Integration Hub stores the host name in the `pwc.domain.gateway` system property.

Gateway port number

Port number for the Informatica security domain gateway. Data Integration Hub stores the port number in the `pwc.domain.gateway` system property. Use the gateway HTTP port number to connect to the domain from the PowerCenter Client. You cannot use the HTTPS port number to connect to the domain.

Username

User name to access the Administrator tool. You must create the user in the Administrator tool and assign the **manage roles/groups/user** privilege to the user.

Password

Password of the Informatica security domain user.

Security domain

Name of the Informatica security domain where the user is defined.

Security group

Optional. Security group within the Informatica security domain where Data Integration Hub users are defined in the following format:

```
<security group>@<domain>
```

If you leave the field empty, the Informatica security domain synchronizes only the Data Integration Hub administrator user account.

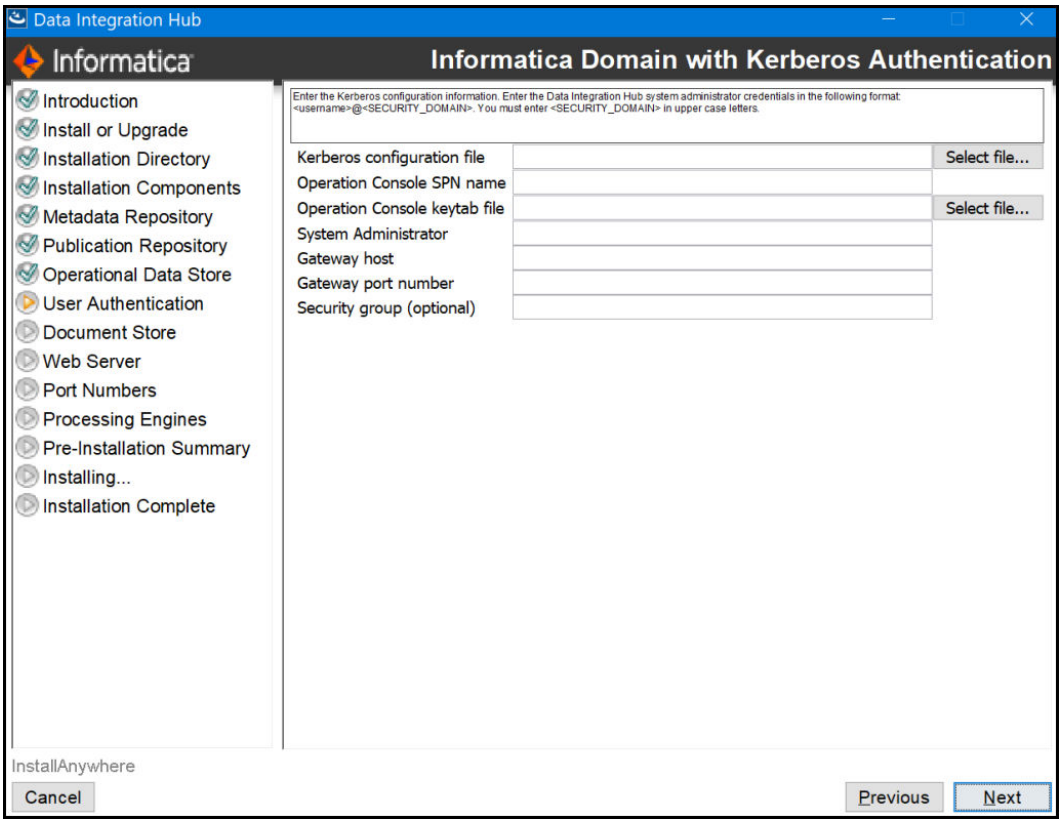
Data Integration Hub stores the security group in the dx.authentication.groups system property in the following format:

```
<group name>@<security group>[;<groupname>@<security group>]
```

Configure Settings for Informatica Domain with Kerberos Authentication

If you select the **Informatica domain with Kerberos authentication** option on the **User Authentication** page, you can configure the authentication settings on the **Informatica Domain with Kerberos Authentication** page.

The following image shows the **Informatica Domain with Kerberos Authentication** page.



Enter the configuration information.

Kerberos configuration file

File that stores Keberos configuration information, usually named `krb5.conf`

The installation copies the file to the following location:

```
<DIHInstallationDir>/shared/conf/security/krb5.conf
```

Operation Console SPN name

Service Principal Name (SPN) for the Data Integration Hub Operation Console.

Data Integration Hub stores the SPN in the `dx-security-config.properties` property file, in the dx.kerberos.console.service.principal.name property.

Operation Console keytab file

Location of the keytab file for the Data Integration Hub Operation Console SPN.

The installer copies the file to the following location:

```
<DIHInstallationDir>/shared/conf/security/HTTP_console.keytab
```

Data Integration Hub stores the location of the keytab file in the `dx-security-config.properties` property file, in the dx.kerberos.console.keytab.file property.

If you change the property to point to a different file, you must enter the absolute path to the file using the following format:

```
file://<full_path>
```

System Administrator

Data Integration Hub system administrator credentials.

Enter the credentials in the following format:

```
<username>@<SECURITY_DOMAIN>
```

Note: You must enter `<SECURITY_DOMAIN>` in uppercase letters.

Gateway host

PowerCenter domain gateway host.

Gateway port number

PowerCenter domain gateway port number.

Security group

Optional. Security group within the Informatica security domain where Data Integration Hub users are defined in the following format:

```
<security group>@<domain>
```

If you leave the field empty, the Informatica security domain synchronizes only the Data Integration Hub administrator user account.

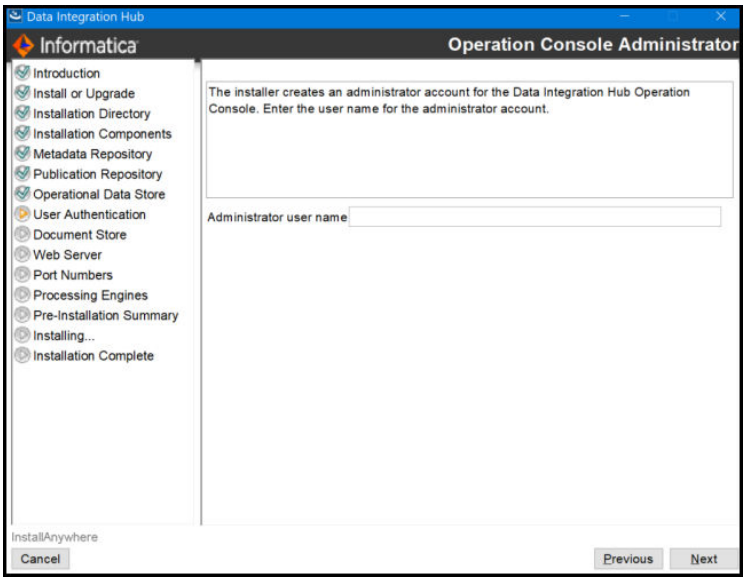
Data Integration Hub stores the security group in the dx.authentication.groups system property in the following format:

```
<group name>@<security group>[;<groupname>@<security group>]
```

Configure Settings for Data Integration Hub Native Authentication

If you select the **Data Integration Hub native authentication** option on the **User Authentication** page, you need to enter the Data Integration Hub administrator user name on the **Operation Console Administrator** page. Data Integration Hub uses this value for the user name and password when you log in to the Operation Console.

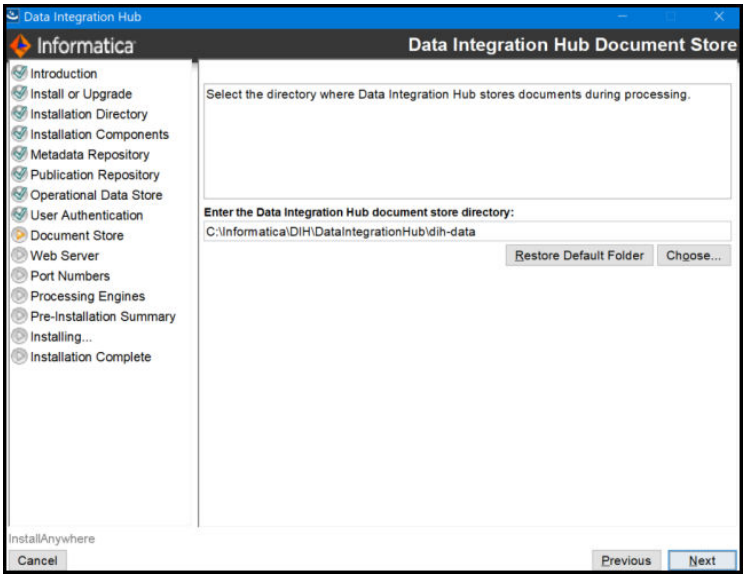
The following image shows the **Operation Console Administrator** page.



Step 6. Configure Document Store and Web Server

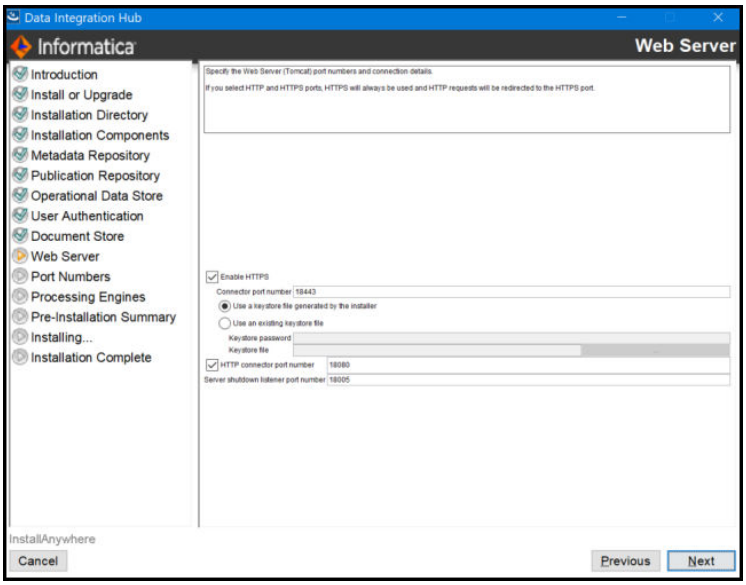
1. On the **Data Integration Hub Document Store** page, accept the default directory or enter the directory where you want to create the document store directory.

   Data Integration Hub stores documents and files in the document store during processing. The document store directory must be accessible to Data Integration Hub, PowerCenter services, and Data Transformation.



2. Click **Next**.

The **Web Server** page appears.



3. Enter values in the following fields:

Enable HTTPS

Instructs Data Integration Hub to use secure network communication when you open the Operation Console in the browser. If you select HTTPS and HTTP, the Operation Console switches existing HTTP connections with HTTPS connections.

Connector port number

Port number for the Tomcat connector to use when you open the Operation Console with HTTPS. The default value is 18443.

Use a keystore file generated by the installer

Instructs the installer to generate a keystore file with an unregistered certificate. If you select this option, ignore the security warning that you receive from the browser the first time you open the Operation Console.

Use an existing keystore file

Instructs the installer to load an existing keystore file. Enter values in the following fields:

- Keystore password. Password for the keystore file.
- Keystore file. Path to the keystore file.

The keystore file must be in the Public Key Cryptography Standard (PKCS) #12 format.

HTTP connector port number

Port number for the HTTP connector. If you clear this field, your browser must connect to the Data Integration Hub server with HTTPS when you log in to the Operation Console.
The default value is 18080.

Server shutdown listener port number

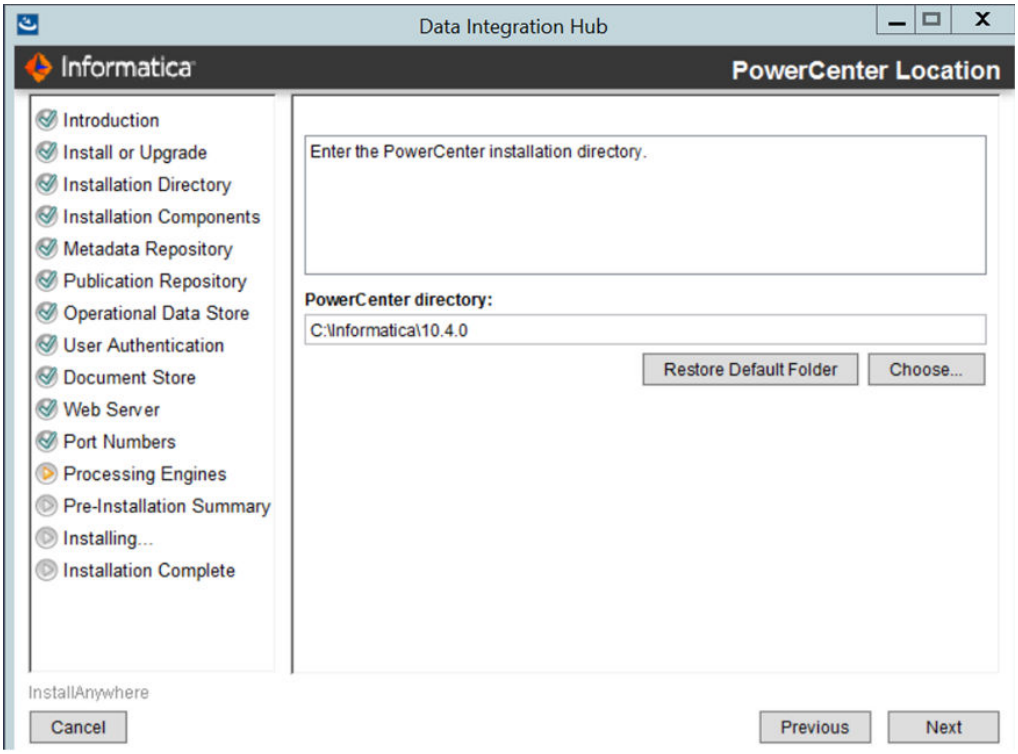
Port number for the listener that controls the Tomcat server shutdown.
The default value is 18005.

4. Click **Next**.

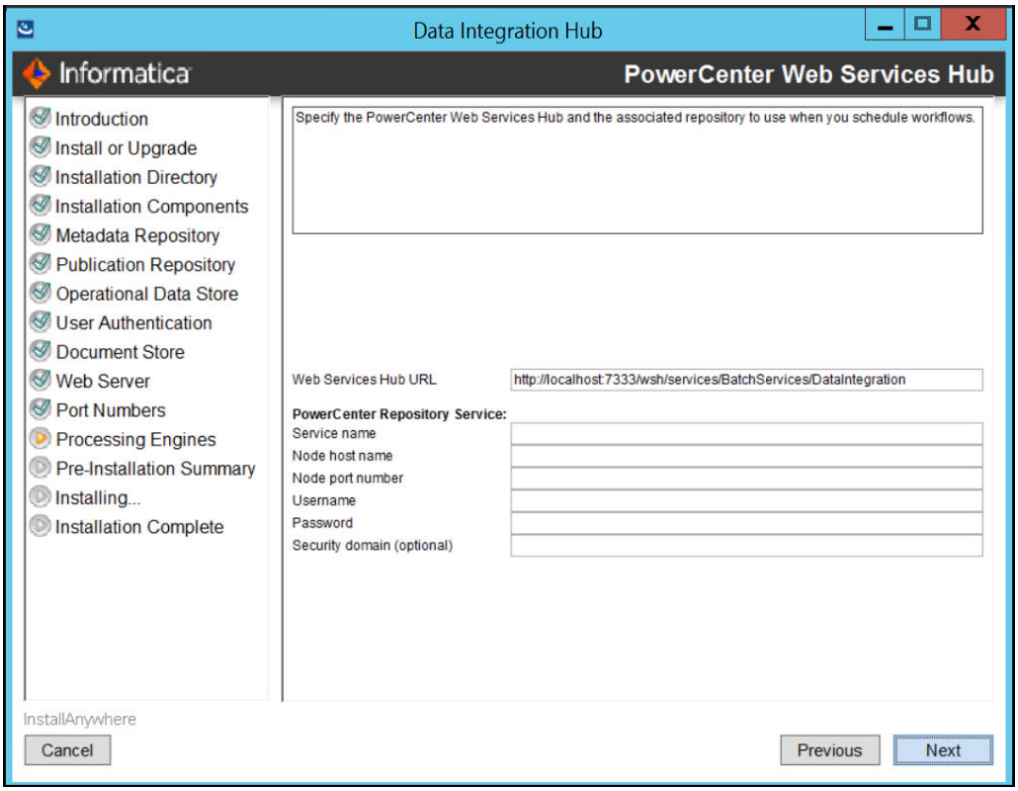
   If you selected to install the Data Integration Hub server plug-in for PowerCenter or the Data Integration Hub client plug-in for PowerCenter components, the **PowerCenter Location** page appears. If you did not select the PowerCenter server or client components, the **PowerCenter Web Services Hub** page appears.

Step 7. Configure PowerCenter Settings

1. If you selected to install the Data Integration Hub server plug-in for PowerCenter or the Data Integration Hub client plug-in for PowerCenter components, on the **PowerCenter Location** page, enter the directory where you installed PowerCenter or accept the default directory, and then click **Next**.



The **PowerCenter Web Services Hub** page appears.



2. On the **PowerCenter Web Services Hub** page, enter the PowerCenter web services details.

Web Services Hub URL

URL that the PowerCenter Web Services Hub uses to process publication and subscription workflows.

Service name

Name of the PowerCenter Repository Service.

Node host name

Host name of the node that runs the PowerCenter Repository Service.

Node port number

Port number of the node that runs the PowerCenter Repository Service.

Username

Name of the PowerCenter Repository Service user.

Password

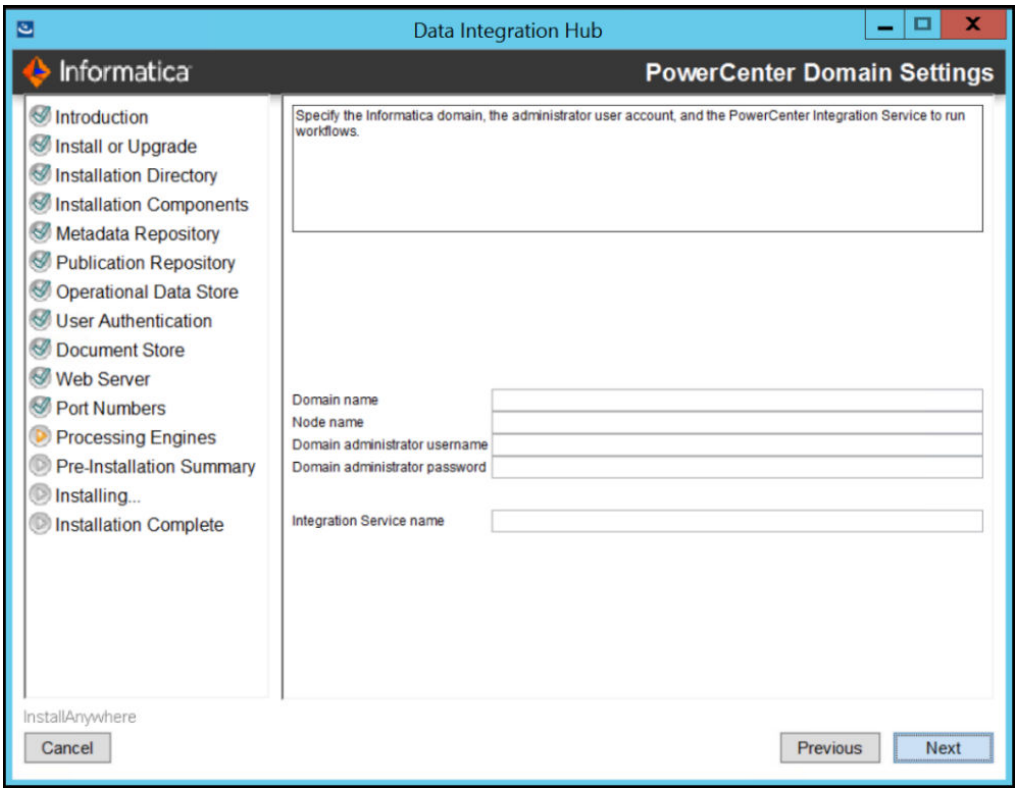
Password for the PowerCenter Repository Service user. Data Integration Hub stores the password as an encrypted string.

Security domain

Optional. Name of the Informatica security domain in which the PowerCenter Repository Service user is stored.
Default is Native.

3. Click **Next**.

If you selected to install the Data Integration Hub server plug-in for PowerCenter component, the **PowerCenter Domain Settings** page appears.
If you did not select the PowerCenter server component, the **PowerCenter pmrep Command Line Utility Location** page appears. Go to step 6.



4. Enter values in the following fields:

Domain name

Name of the Informatica domain that contains the PowerCenter Integration Service that runs Data Integration Hub workflows.

Node name

Node in the Informatica domain on which the PowerCenter Integration Service runs.

Domain administrator username

Name of the Informatica domain administrator.

Domain administrator password

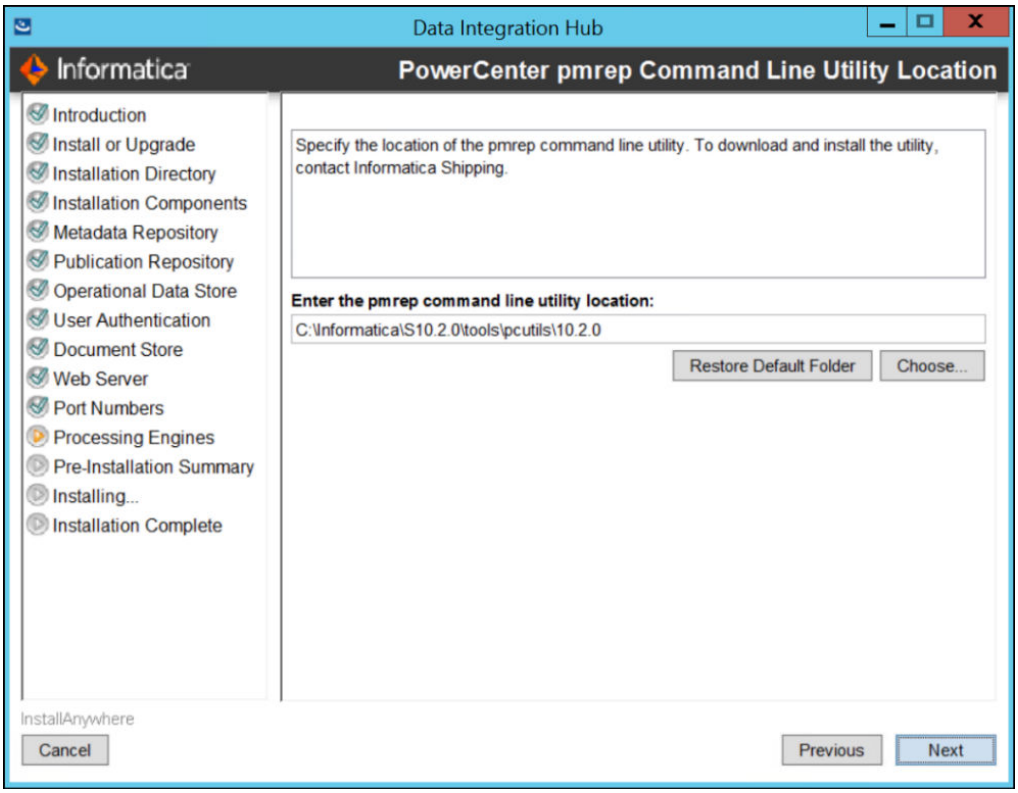
Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.

Integration Service name

The name of the PowerCenter Integration Service that Data Integration Hub uses to run workflows.

5. Click **Next**.

The **PowerCenter pmrep Command Line Utility Location** page appears.



6. Specify the location of the pmrep command line utility.

   The location of the utility depends on whether or not you install Data Integration Hub on the machine where the PowerCenter services are installed.

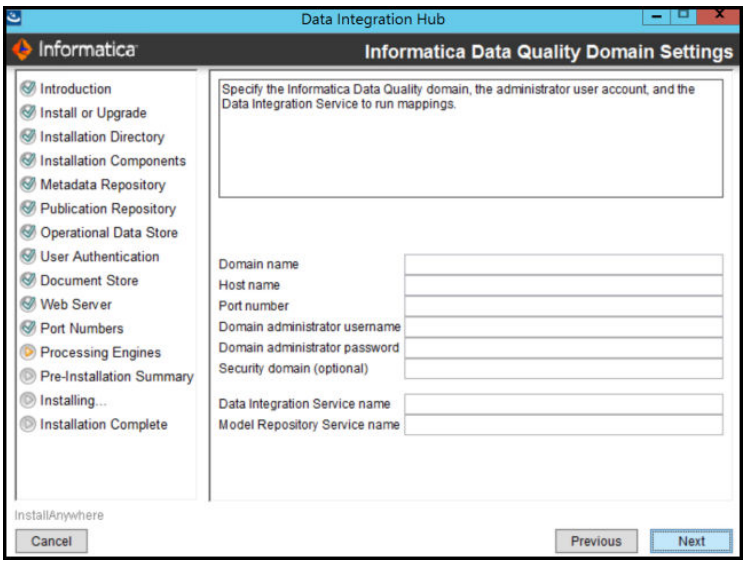
Environment	Location of the pmrep command line utility
Data Integration Hub installed on the machine where the PowerCenter services are installed	`<PowerCenter_services_installation_folder>\<PowerCenter_version>\tools\pcutils\<PowerCenter_version>`
Data Integration Hub and PowerCenter services installed on different machines	`<PowerCenter_client_installation_folder>\<PowerCenter_version>\clients\PowerCenterClient\client\bin`

7. Click **Next**.

   If you selected to install the Data Integration Hub Informatica Data Quality component, the **Informatica Data Quality Domain Settings** page appears. If you did not select to install the Data Integration Hub Informatica Data Quality component, the **Pre-Installation Summary** page appears.

Step 8. Configure Informatica Data Quality Settings

1. If you selected to install the Data Integration Hub Informatica Data Quality component, the **Informatica Data Quality Domain Settings** page appears.



Enter values in the following fields:

Domain name

Domain name of the node on which the Data Integration Service runs.

Host name

Host name of the node on which the Data Integration Service runs.

Port number

Port number of the node on which the Data Integration Service runs.

Domain administrator username

User name for the node on which the Data Integration Service runs.

Domain administrator password

Password for the node on which the Data Integration Service runs.

Security Domain

Optional. Name of the Informatica security domain in which the Model Repository Service user is stored.

Data Integration Service name

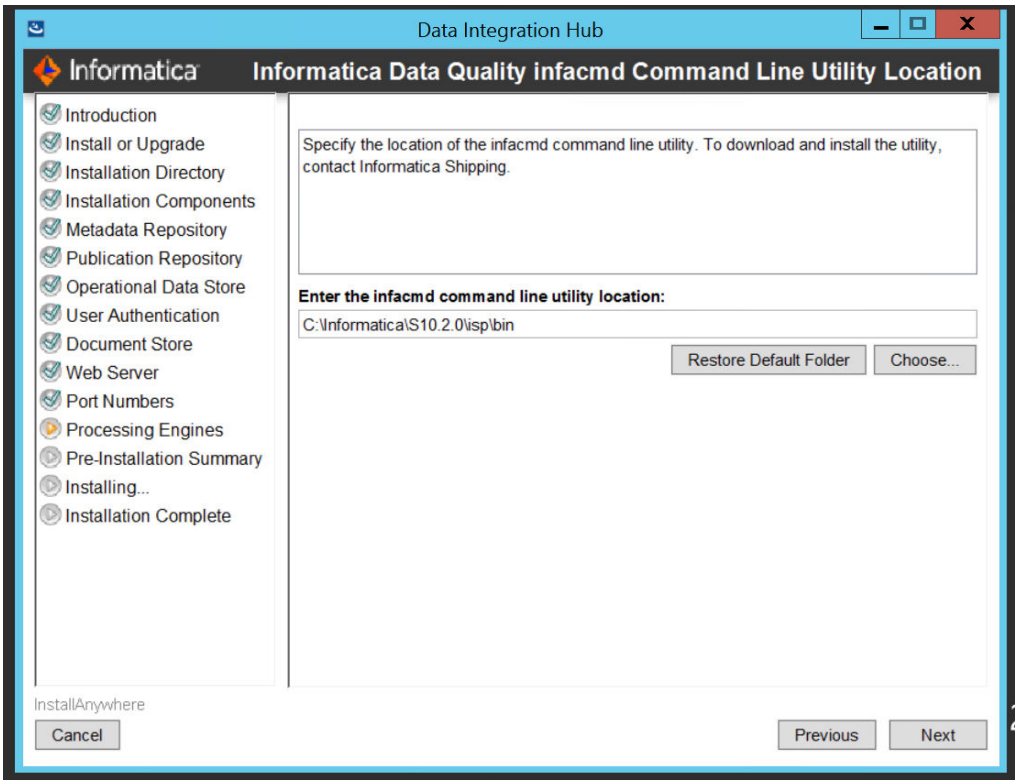
Name of the Data Integration Service.

Model Repository Service name

Name of the Model Repository Service.

2. Click **Next**.

The **Informatica Data Quality infacmd Command Line Utility** page appears.



3. Enter the location of the infacmd command line utility or accept the default directory, and then click **Next**.

   The **Pre-Installation Summary** page appears.

Step 9. Complete the Installation

1. On the **Pre-Installation Summary** page, verify that the installation information is correct, and then click **Install**.

   During the installation process, the installer displays progress information. When the installation process ends, the **Post-Installation Actions** page appears.

2. If you installed the Data Integration Hub PowerCenter server plug-in, follow the wizard instructions to register the plug-in to the PowerCenter repository, and then click **Next**.

   The **Installation Complete** page appears.

3. Click **Done** to close the installer.

4. To view the log files that the installer generates, navigate to the following directory: `<DIHInstallationDir>\logs`.

5. Perform the required post-installation tasks.

   For more information, see Chapter 5, “Post-Installation Tasks” on page 68.

   Note: Perform only the tasks that are relevant for your environment.

6. Optionally, perform additional configuration tasks. For more information, see Chapter 8, “Optional Data Integration Hub Configuration” on page 112.

Installing Data Integration Hub on a UNIX Operating System in Console Mode

Install Data Integration Hub on UNIX operating systems in console mode. On Windows operating systems, install Data Integration Hub in graphical mode.

Before you install, verify that your environment meets the minimum system requirements, perform the pre-installation tasks, and verify that the PowerCenter services are running.

During the installation, Data Integration Hub saves log files in the home directory of the user, in the subdirectory named `DXLogs`. If the installation does not complete successfully, you can view the log files in this location.

Step 1. Run the Installer

1. Log in to the machine with the user account that you want to use to install Data Integration Hub.

   To prevent permission errors, use the same account to install Data Integration Hub and PowerCenter.
2. Close all other applications.
3. Run `Install.bin -i console` from the directory where you downloaded the installer.

   The **Introduction** section appears.
4. Read the instructions, and then press **Enter**.

   The **Install or Upgrade** section appears.
5. Enter **1** to install Data Integration Hub, and then press **Enter**.

   The **Installation Directory** section appears.

Step 2. Define Installation Settings

1. In the **Installation Directory** section, enter the absolute path to the installation directory or accept the default directory, and then press **Enter**.

   The **Installation Components** section appears and displays a numbered list of the components to install.
2. Enter a comma-separated list of numbers for the components to install or accept the default components:

   1- Data Integration Hub

   Installs the core Data Integration Hub application.
   Selected by default.

   2- Data Integration Hub Dashboard and Reports

   You can make the Dashboard available in Data Integration Hub in two ways. When you install Data Integration Hub, you can enable the dashboard that uses operational data store. Otherwise, you can use the dashboard that uses metadata repository which is available by default.

   Installs the Data Integration Hub Dashboard and Reports component. You must install Data Integration Hub to install the Dashboard and Reports component.
   Cleared by default.

Note:

- The Dashboard using metadata directory is installed by default. Select this component to view the reports in the Dashboard using operational data store. After installation, you must ensure to set the `dx.dashboard.ods.page.show` system property to `true` to view the reports in the Dashboard using operational data store. For more information, see the *Data Integration Hub Administrator Guide* and *Data Integration Hub Operator Guide*.
- If you install the Dashboard and Reports component, you must import the operational data store event loader after you install Data Integration Hub.
- If you install the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, you must configure the repository connections in PowerCenter Workflow Manager. For details, see "Configuring Repository Connections on PowerCenter Version 10" on page 122.

3- Data Integration Hub Server Plug-in for PowerCenter

Installs the Data Integration Hub PowerCenter server plug-in component. After the installation, register the plug-in to the PowerCenter repository.
Selected by default.

4- Data Integration Hub Hadoop Service

Connecting module between Data Integration Hub and the Hadoop cluster. Enables Data Integration Hub to perform operations on the Hadoop publication repository.
Select this component if you want to define a Hadoop-based publication repository and manage some of your topics on Hadoop.

Note: The Data Integration Hub Hadoop Service must be installed on an edge node in the Hadoop cluster. Unless you are installing Data Integration Hub on a Hadoop edge node, do not select this component during Data Integration Hub installation. In this case, after you complete the Data Integration Hub installation, run the installation again on the edge node and select to install only the Data Integration Hub Hadoop Service component on the node.

5- Data Integration Hub Data Engineering Integration/Informatica Data Quality

Enables Data Integration Hub to use Data Engineering Integration and Data Quality mappings for custom publications and subscriptions.
Select this component if you want to create custom publications and subscriptions that use Data Engineering Integration or Data Quality mappings in Data Integration Hub.

3. Press **Enter**.

   The **Select PowerCenter Version** section appears.

4. Select the PowerCenter version for which to install Data Integration Hub or accept the default selection:

   1- PowerCenter version below 10.5.

   Select this option for PowerCenter versions below 10.5.

   2 - PowerCenter version 10.5.

   Select this option for PowerCenter version 10.5.

5. Press **Enter**.

   The **Metadata Repository** section appears.

Step 3. Configure Data Integration Hub Repositories

1. In the **Metadata Repository** section, enter the number for the metadata repository database configuration option or accept the default option:

   1- Create a Data Integration Hub repository

   Creates a repository in the database.

   2- Use an existing Data Integration Hub repository

   Uses the tables and data in an existing repository.

2. Press **Enter**.

   The **Metadata Repository Connection** section appears.

3. Enter one of the following numericals depending on the database you plan to use as the Data Integration Hub metadata repository database:

   - Enter `1` to use an Oracle database as the Data Integration Hub metadata repository database.
   - Enter `2` to use Microsoft SQL Server database as the Data Integration Hub metadata repository database.

4. Enter the number for the metadata repository database connection type or accept the default connection type:

   1- Database URL

   Location of the database. If you select this option, enter values in the following fields:

   - **Database Host Name**. Host name of the machine where the database server is installed.
   - **Database Port Number**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server and Microsoft Azure SQL Database is 1433.
   - **Database SID**. System identifier for the database.
   - **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

   2- Custom Connection String

   Connection string to the database. If you select this option, enter values in one of the following fields:

   - JDBC string. JDBC connection string to the metadata repository.
   - ODBC string. ODBC connection string to the metadata repository. Applicable if you install the PowerCenter client plug-in. The installer cannot verify the validity of the ODBC string.

   Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

   For example:

   ```
   jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
   \MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
   ```

5. Enter values in the following fields:

   Database username

   Name of the database user account.

   Database user password

   The password for the database account for the database. Data Integration Hub stores the password as an encrypted string.

6. Press **Enter**.
7. Enter the number for the publication repository connection type or accept the default connection type.

   Note: When using Microsoft SQL Server named instances, you must define a custom connection string.

   1- Database URL

   Location of the database. If you select this option, enter values in the following fields:

   - **Database host name**. Host name of the machine where the database server is installed.
   - **Database port number.** Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database.
   - **Oracle database**. Database SID. System identifier for the database.
   - **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

   2- Custom Connection String

   Connection string to the database. If you select this option, enter values in one of the following fields:

   - **JDBC string**. JDBC connection string to the publication repository.
   - **ODBC string**. Applicable if you install the PowerCenter client plug-in. ODBC connection string to the publication repository. The installer cannot verify the validity of the ODBC string.

   Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

   For example:

   ```
   jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
   \MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
   ```
8. Press **Enter**.
9. Enter values in the following fields:

   Database username

   Name of the database user account for the database.

   Database user password

   The password for the database account. Data Integration Hub stores the password as an encrypted string.
10. Press **Enter**.

    If you selected to install the Data Integration Hub Dashboard and Reports component, the **Operational Data Store** section appears. If you did not select to install the Dashboard and Reports component, go to "Step 5. Configure the Web Server " on page 102.

Step 4. Set Up the Operational Data Store

1. In the **Operational Data Store** section, enter the number for the database configuration option for the operational data store or accept the default option:

   1- Create an operational data store repository

   Creates an operational data store repository in the database.

   2- Use an existing operational data store repository

   Uses the tables and data in an existing operational data store repository.

2. Enter the number for the database connection type for the operational data store or accept the default connection type:

   1- Database URL

   Location of the database. If you select this option, enter values in the following fields:

   - **Database host name**. Host name of the machine where the database server is installed.
   - **Database port number**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database is 1433.
   - **Oracle database**. Database SID. System identifier for the database.
   - **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

   2- Custom Connection String

   Connection string to the database. If you select this option, enter values in one of the following fields:

   - **JDBC string.** JDBC connection string to the Operational Data Store.
   - **ODBC string**. ODBC connection string to the Operational Data Store. If you install the PowerCenter client plug-in, the installer cannot verify the validity of the ODBC string.

   Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

   For example:

   ```
   jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
   \MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
   ```

3. Enter values for the operational data store in the following fields:

   Database username

   Name of the database user account for the database.

   Database user password

   The password for the database account for the database. Data Integration Hub stores the password as an encrypted string.

4. Press **Enter**.

   The **User Authentication** section appears.

Step 5. Configure User Authentication

1. In the **User Authentication** section, choose the type of user authentication that you want to use.
   - Choose Informatica domain authentication to manage user credentials in the Informatica domain and synchronize user information with Data Integration Hub. Use Informatica domain authentication for production environments. For more information, see "Configure Settings for Informatica Domain Authentication" on page 52.

     Note: If your Informatica domain uses Kerberos authentication, choose the option **Informatica domain with Kerberos authentication**.
   - Choose Informatica domain with Kerberos authentication if your Informatica domain uses Kerberos authentication. Use Informatica domain with Kerberos authentication for production environments. For more information, see "Configure Settings for Informatica Domain with Kerberos Authentication" on page 52.

- Choose **Data Integration Hub native authentication** to manage user credentials locally in Data Integration Hub. Use native authentication in development and staging environments. For more information, see "Configure Settings for Data Integration Hub Native Authentication " on page 53.

2. Press **Enter**.

   The **Document Store** section appears.

Configure Settings for Informatica Domain Authentication

If you choose Informatica domain authentication, enter values in the following fields:

Gateway host

Host name of the Informatica security domain server. Data Integration Hub stores the host name in the `pwc.domain.gateway` system property.

Gateway port

Port number for the Informatica security domain gateway. Data Integration Hub stores the port number in the `pwc.domain.gateway` system property. Use the gateway HTTP port number to connect to the domain from the PowerCenter Client. You cannot use the HTTPS port number to connect to the domain.

Username

User name to access the Administrator tool. You must create the user in the Administrator tool and assign the **manage roles/groups/user** privilege to the user.

Password

Password of the Informatica security domain user.

Security domain

Name of the Informatica security domain where the user is defined.

Security group

Optional. Security group within the Informatica security domain where Data Integration Hub users are defined in the following format:

```
<security group>@<domain>
```

If you leave the field empty, the Informatica security domain synchronizes only the Data Integration Hub administrator user account.

Data Integration Hub stores the security group in the dx.authentication.groups system property in the following format:

```
<group name>@<security group>[;<groupname>@<security group>]
```

Configure Settings for Informatica Domain with Kerberos Authentication

If you choose Informatica domain with Kerberos authentication, enter values in the following fields:

Kerberos configuration file

File that stores Keberos configuration information, usually named `krb5.conf`

The installation copies the file to the following location:

```
<DIHInstallationDir>/shared/conf/security/krb5.conf
```

Operation Console SPN name

Service Principal Name (SPN) for the Data Integration Hub Operation Console.

Data Integration Hub stores the SPN in the `dx-security-config.properties` property file, in the dx.kerberos.console.service.principal.name property.

Operation Console keytab file

Location of the keytab file for the Data Integration Hub Operation Console SPN.

The installation copies the file to the following location:

```
<DIHInstallationDir>/shared/conf/security/HTTP_console.keytab
```

Data Integration Hub stores the location of the keytab file in the `dx-security-config.properties` property file, in the dx.kerberos.console.keytab.file property.

If you change the property to point to a different file, you must enter the absolute path to the file using the following format:

```
file://<full_path>
```

System Administrator

Data Integration Hub system administrator credentials.

Enter the credentials in the following format:

```
<username>@<SECURITY_DOMAIN>
```

Note: You must enter `<SECURITY_DOMAIN>` in uppercase letters.

Gateway host

PowerCenter domain gateway host.

Gateway port number

PowerCenter domain gateway port number.

Security group

Optional. Security group within the Informatica security domain where Data Integration Hub users are defined in the following format:

```
<security group>@<domain>
```

If you leave the field empty, the Informatica security domain synchronizes only the Data Integration Hub administrator user account.

Data Integration Hub stores the security group in the dx.authentication.groups system property in the following format:

```
<group name>@<security group>[;<groupname>@<security group>]
```

Configure Settings for Data Integration Hub Native Authentication

If you choose **Data Integration Hub native authentication**, enter the Data Integration Hub administrator user name. Data Integration Hub uses this value for the user name and password when you log in to the Operation Console.

Step 6. Configure Document Store and Web Server

1. In the **Document Store** section, enter the directory where Data Integration Hub stores documents and files during processing or accept the default directory, and then press `Enter`.

   The document store directory must be accessible to Data Integration Hub, PowerCenter services, and Data Transformation.

2. Press **Enter**.

   The **Web Server** section appears.

3. Configure the Web Server connection.

   a. Enter the number for the network communication protocol or accept the default protocol:

      1- Enable HTTPS

      Instructs Data Integration Hub to use secure network communication when you open the Operation Console in the browser.
      If you select HTTPS and HTTP, the Operation Console switches existing HTTP connections with HTTPS connections.

      2- Enable HTTP

      Instructs Data Integration Hub to use regular HTTP network communication when you open the Operation Console in the browser.

   b. If you selected **Enable HTTPS**, enter values in the following fields:

      Connector port number

      Port number for the Tomcat connector to use when you open the Operation Console with HTTPS.
      The default value is 18443.

      Use a keystore file generated by the installer

      Instructs the installer to generate a keystore file with an unregistered certificate. If you select this option, ignore the security warning that you receive from the browser the first time you open the Operation Console.

      Use an existing keystore file

      Instructs the installer to load an existing keystore file. Enter values in the following fields:

      - Keystore password. Password for the keystore file.
      - Keystore file. Path to the keystore file.

      The keystore file must be in the Public Key Cryptography Standard (PKCS) #12 format.

   c. If you selected **Enable HTTP**, enter values in the following fields:

      HTTP connector port number

      Port number for the HTTP connector. If you clear this field, your browser must connect to the Data Integration Hub server with HTTPS when you log in to the Operation Console.
      The default value is 18080.

      Server shutdown listener port number

      Port number for the listener that controls the Tomcat server shutdown.
      The default value is 18005.

4. Press **Enter**.

   If you selected to install the Data Integration Hub PowerCenter server plug-in or the Data Integration Hub PowerCenter Client plug-in components, the **PowerCenter Location** section appears. If you did not select the PowerCenter server or client components, the **PowerCenter Web Services Hub** section appears.

Step 7. Configure PowerCenter Settings

1. If you selected to install the Data Integration Hub PowerCenter server plug-in or the Data Integration Hub PowerCenter Client plug-in components, in the **PowerCenter Location** section, enter the directory where you installed PowerCenter or accept the default directory, and then press **Enter**.

   The **PowerCenter Web Services** section appears.
2. In the **PowerCenter Web Services** section, press **Enter** to accept the default URL or enter the URL that the PowerCenter Web Services Hub uses to process publication and subscription workflows and then press **Enter**.
3. Enter the name of the PowerCenter Repository Service, and then press **Enter**.
4. Enter values in the following fields:

   Node host name

   Host name of the node that runs the PowerCenter Repository Service.

   Node port number

   Port number of the node that runs the PowerCenter Repository Service.

   Username

   Name of the PowerCenter Repository Service user.

   Password

   Password for the PowerCenter Repository Service user. Data Integration Hub stores the password as an encrypted string.

   Security domain

   Optional. Name of the Informatica security domain in which the PowerCenter Repository Service user is stored.
   Default is Native.
5. Press **Enter**.

   If you selected to install the Data Integration Hub server plug-in for PowerCenter component, the **PowerCenter Domain Settings** section appears. If you did not select the PowerCenter server component, the **PowerCenter pmrep Command Line Utility Location** section appears. Go to step 9.
6. Enter values in the following fields:

   Domain name

   Name of the Informatica domain that contains the PowerCenter Integration Service that runs Data Integration Hub workflows.

   Node name

   Node in the Informatica domain on which the PowerCenter Integration Service runs.

   Domain administrator user name

   Name of the Informatica domain administrator.

   Domain administrator password

   Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.
7. Press **Enter**.
8. Enter the name of the PowerCenter Integration Service that Data Integration Hub uses to run workflows, and then press **Enter**.

9. Enter the location of the pmrep command line utility and then press **Enter**. The location of the utility depends on whether or not you install Data Integration Hub on the machine where the PowerCenter services are installed.

   Note: On Linux operating systems, pmrep must be executable.

Environment	Location of the pmrep command line utility
Data Integration Hub installed on the machine where the PowerCenter services are installed	`<PowerCenter_services_installation_folder>/<PowerCenter_version>/tools/pcutils/<PowerCenter_version>`
Data Integration Hub and PowerCenter services installed on different machines	`<PowerCenter_client_utility_directory>/PowerCenter/server/bin`

10. Press **Enter**.

    If you selected to install the Data Integration Hub Data Engineering Integration/Informatica Data Quality component, the **Data Engineering Integration/Informatica Data Quality Domain name** section appears. If you did not select to install the Data Integration Hub Data Engineering Integration/Informatica Data Quality component, the **Pre-Installation Summary** section appears.

Step 8. Configure Processing Engine Settings

1. In the **Data Engineering Integration/Informatica Data Quality Domain name** section, press **Enter** to accept the default domain name or enter the name of the Informatica domain, and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Host name** section appears.

2. Press **Enter** to accept the default host name or enter the host name of the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Port number** section appears.

3. Press **Enter** to accept the default port number or enter the port number of the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality administrator user name** section appears.

4. Press **Enter** to accept the default user name or enter the user name for the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality administrator password** section appears.

5. Press **Enter** to accept the default password or enter the password for the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Security Domain** section appears.

6. Optionally, enter the name of the Informatica security domain in which the Model Repository Service user is stored. Press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Data Integration Service name** section appears.

7. Press **Enter** to accept the default name or enter the name of the Data Integration Service and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Model Repository Service name** section appears.

8. Press **Enter** to accept the default name or enter the name of the Model Repository Service and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality infacmd Command Line Utility Location** section appears.

9. Press **Enter** to accept the default location or enter the location of the infacmd command line utility then press **Enter**.

   The **Pre-Installation Summary** section appears.

Step 9. Complete the Installation

1. In the **Pre-Installation Summary** section, verify that the installation information is correct, and then press **Enter**.

   During the installation process, the installer displays progress information.

2. If you installed the Data Integration Hub PowerCenter server plug-in, follow the on-screen instructions to register the plug-in to the PowerCenter repository, and then press **Enter**.
3. To view the log files that the installer generates, navigate to the following directory:
   `<DIHInstallationDir>\logs`
4. Perform the required post-installation tasks.

   For more information, see Chapter 5, “Post-Installation Tasks” on page 68.

   Note: Perform only the tasks that are relevant for your environment.

5. Optionally, perform additional configuration tasks. For more information, see Chapter 8, “Optional Data Integration Hub Configuration” on page 112.

Installing Data Integration Hub in a Silent Mode

To install Data Integration Hub without user interaction, install in a silent mode. Use a properties file to specify the installation options. The installer reads the file to determine the installation options. You can use silent mode installation to install Data Integration Hub on multiple machines on the network or to standardize the installation across machines.

The silent install supports fresh installation of all Data Integration Hub components on a single node. Upgrade of Data Integration Hub and High Availability multi-node installations of Data Integration Hub using the silent installer is supported. Before you install, verify that your environment meets the minimum system requirements, perform the pre-installation tasks, and verify that the PowerCenter services are running.

To install in silent mode, complete the following tasks

1. Configure the installation properties file and specify the installation options in the properties file.
2. Run the installer with the installation properties file.

Configuring the Installation Properties

The installation properties file includes the parameters that are required by the installer.

The following table describes parameters that you add in the installation properties file:

Parameter	Description
Specify whether to install or upgrade Data Integration Hub.	
IS_INSTALL	To install Data Integration Hub, set the parameter to 1. Else, set the parameter to 0.
IS_UPGRADE	To upgrade Data Integration Hub, set the parameter to 1. Else, set the parameter to 0. **Note:** If you set this parameter to 1, you must set the BLANK_USER parameter to 0 and the CONFIGURED_USER parameter to 1.
Configure PowerCenter version using the following parameters:	
PMREP_HOME_1	Absolute location of the pmrep command line utility folder.
POWERCENTER_HOME_1	Absolute directory where you installed PowerCenter.
Configure the installation directory using the following parameter:	
USER_INSTALL_DIR	Absolute path to the installation directory.
Configure the components to install using the following parameters:	
DX_SERVER	To install the Data Integration Hub application, set the parameter to 1. Else, set the parameter to 0.
DX_DASHBOARD	To install the Data Integration Hub Dashboard and Reports component, set the parameter to 1. Else, set the parameter to 0.
PC_SERVER_PLUGIN	To install the Data Integration Hub server plug-in for the PowerCenter services, set the parameter to 1. Else, set to 0.
PC_CLIENT_PLUGIN	To install the Data Integration Hub plug-in for the PowerCenter client, set the parameter to 1. Else, set the parameter to 0. This applies only for Windows.
DIH_BIG_DATA_MANAGEMENT	To install the Data Integration Hub Big Data Management component, set the parameter to 1. Else, set the parameter to 0.
DIH_HADOOP_SERVICE	To install the Data Integration Hub Hadoop service, set the parameter to 1. Else, set the parameter to 0. This applies only for Linux.
Configure the Data Integration Hub repository using the following parameters:	
BLANK_USER	To create a new Data Integration Hub repository, set the parameter to 1. Else, set the parameter to 0.

Parameter	Description
CONFIGURED_USER	To configure an existing Data Integration Hub repository, set the parameter to 1. Else, set the parameter to 0.
DB_TYPE_1	Type of database to use for the Data Integration Hub metadata repository. Enter either of the following options: - Oracle - SQLServer
DB_CONNECTION_STRING_1	To use a custom connection string, set the parameter to 1. Else, set the parameter to 0. **Note:** If you set this parameter to 1, ignore DB_HOST, DB_PORT and DB_SID.
DB_HOST_1	Host name of the machine where the database server is installed.
DB_PORT_1	Port number for the database server. The default port number for Oracle is 1521. The default port number for Microsoft SQL Server is 1433.
DB_SID_1	System identifier for the database. Enter a fully qualified Service Name or a fully qualified SID.
DB_CONNECTION_STRING_VALUE_1	The JDBC value of the connection string. Configure this parameter only if DB_CONNECTION_STRING_1 is set to 1.
DB_ODBC_STRING_VALUE_1	Mandatory. The driver value of the connection string. Use one of the following driver connection strings: - Oracle: {DataDirect 7.1 Oracle Wire Protocol};UID=%1;PWD=%2;Host=<Server_Host>;Port=1521;SN=<SID> - SQL server: {DataDirect 7.1 SQL Server Wire Protocol};UID=%1;PWD=%2;Server=<server>,<port_no>;Database=<REPO_DB>
DB_WINDOWS_AUTHENTICATION_1	To authenticate user names against the Microsoft Windows authentication mechanism, set the parameter to 1. Else, set the parameter to 0.
DB_USER_1	Name of the database user account for the database where you do not use Windows authentication.
DB_PASSWORD_1	Password for the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.
Configure the publication repository using the following parameters:	
STAGING_TYPE_1	Type of database to use for the Data Integration Hub publication repository. Enter either of the following options: - Oracle - SQLServer

Parameter	Description
STAGING_CONNECTION_STRING_1	To use a custom connection string, set the parameter to 1. Else, set the parameter to 0. **Note:** If you set this parameter to 1, ignore STAGING_HOST, STAGING_PORT and STAGING_SID.
STAGING_HOST_1	Host name of the machine where the staging server is installed.
STAGING_PORT_1	Port number for the staging server. The default port number for Oracle is 1521. The default port number for Microsoft SQL Server is 1433.
STAGING_SID_1	System identifier for the staging if the database. Enter a fully qualified Service Name or a fully qualified SID.
STAGING_WINDOWS_AUTHENTICATION_1	To authenticate user names against the Microsoft Windows authentication mechanism, set the parameter to 1. Else, set the parameter to 0.
STAGING_USER_1	Name of the staging user account for the database where you do not use Windows authentication.
STAGING_PASSWORD_1	Password for the database account for the staging where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.
STAGING_CONNECTION_STRING_VALUE_1	The JDBC value of the connection string. Configure this parameter only if STAGING_CONNECTION_STRING_1 is set to 1.
Configure the operational data store using the following parameters:	
ODS_DB_USE_NEW	To create an operational data store repository, set the parameter to 1.
ODS_DB_USE_EXISTING	To configure an existing operational data store repository, set the parameter to 1. Else, set the parameter to 0.
ODS_DB_TYPE_1	Type of database to use for the operational data store. Enter one of the following options: - Oracle - SQLServer
ODS_DB_CONNECTION_STRING_1	To use a custom connection string, set the parameter to 1. Else, set the parameter to 0. **Note:** If you set this parameter to 1, ignore ODS_HOST, ODS_PORT and ODS_SID.
ODS_DB_HOST_1	Host name of the machine where the database server is installed.
ODS_DB_PORT_1	Port number for the database. The default port number for an Oracle database is 1521. The default port number for a Microsoft SQL server is 1433.

Parameter	Description
ODS_DB_SID_1	System identifier for the database. Enter either a fully qualified Service Name or a fully qualified SID.
ODS_DB_JDBC_STRING_VALUE_1	The JDBC value of the connection string. Configure this parameter only if ODS_CONNECTION_STRING_1 is set to 1.
ODS_DB_WINDOWS_AUTHENTICATION_1	To authenticate user names against the Microsoft Windows authentication mechanism., set the parameter to 1. Else, set the parameter to 0.
ODS_DB_USER_1	Name of the database user account for the database where you do not use Windows authentication.
ODS_DB_PASSWORD_1	Password for the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.
Configure settings for Informatica Domain with Kerberos authentication using the following parameters:	
INTERNAL_AUTH	To use Data Integration Hub native authentication, set the value to 1. Else, set the value to 0.
INTERNAL_AUTH_DEFAULT	To use the Data Integration Hub native authentication as the default authentication, set the value to True. Else, set the value to False.
ISF_AUTH	To use Informatica domain authentication, set the value to 1. Else, set the value to 0.
ISF_AUTH_DEFAULT	To use the Informatica domain authentication as the default authentication, set the value to True. Else, set to False.
KERBEROS_AUTH	To use Informatica domain with Kerberos authentication, set the value to 1. Else, set the value to 0.
KERBEROS_AUTH_DEFAULT	To use the Informatica domain with Kerberos authentication as the default authentication, set the value to True. Else, set the value to False.
Configure settings for Informatica domain authentication using the following parameters:	
INFA_HOST_1	Host name of the Informatica security domain server.
INFA_PORT_1	Port number for the Informatica security domain gateway.
INFA_USERNAME_1	User name to access the Administrator tool.
INFA_PASSWORD_1	Password of the Informatica security domain user.
INFA_SECURITY_DOMAIN_1	Name of the Informatica security domain where the user is defined.

Parameter	Description
INFA_SECURITY_GROUP_1	Optional. Security group within the Informatica security domain where Data Integration Hub users are defined in the following format: <security group>@<domain>
Configure Data Integration Hub native authentication using the following parameter:	
CONSOLE_ADMIN_1	User name of the administrator account to be created during installation.
Configure the Data Integration Hub document store using the following parameter:	
DATA_STORE_FOLDER_1	The absolute path of the directory where the Data Integration Hub stores documents during processing.
Configure the Web Server using the following parameters:	
TOMCAT_ENABLE_HTTPS_1	To use the HTTPS secure network communication when you open the Operation Console in the browser, set the parameter to 1. Else, set the parameter to 0.
TOMCAT_HTTPS_PORT_1	The Port number for the Tomcat connector to use when you open the Operation Console with HTTPS. The default value is 18443.
TOMCAT_EXISTING_KEYSTORE_FILE_1	To use an existing keystore file, set the parameter to 1. Else, set the parameter to 0.
TOMCAT_KEYSTORE_PASSWORD_1	If you chose to use an existing keystore, enter the password for the keystore file.
TOMCAT_KEYSTORE_FILE_PATH_1	If you chose to use an existing keystore, enter the path to the keystore file.
TOMCAT_ENABLE_HTTP_1	To use the HTTP network communication when you open the Operation Console in the browser, set the parameter to 1. Else, set the parameter to 0.
TOMCAT_PORT_1	Port number for the HTTP connector.
TOMCAT_SERVER_LISTENER_PORT_1	Port number for the listener that controls the Tomcat server shutdown. The default value is 18005.
Configure the PowerCenter settings using the following parameters:	
PWC_VERSION_BEFORE_10_5	To install Data Integration Hub with PowerCenter version below 10.5, set the parameter to 1. Else, set the parameter to 0.
PWC_VERSION_10_5	To install Data Integration Hub with PowerCenter version 10.5, set the parameter to 1. Else, set the parameter to 0.

Parameter	Description
POWERCENTER_HOME_1	Absolute path where PowerCenter is installed.
PC_SCHEDULING_WORKFLOWS_1	Mandatory. For Data Integration Hub to use the PowerCenter Services for processing batch workflows, you must set the parameter to 1.
PC_WEB_SERVICES_URL_1	URL that the PowerCenter Web Services Hub uses when Data Integration Hub transfers documents to PowerCenter for processing with batch workflows.
PC_REPOSITORY_NAME_1	Name of the PowerCenter Repository Service.
PC_REPOSITORY_HOST_1	Host name of the node that runs the PowerCenter Repository Service.
PC_REPOSITORY_PORT_1	Port number of the node that runs the PowerCenter Repository Service.
PC_REPOSITORY_USER_1	Name of the PowerCenter Repository Service user.
PC_REPOSITORY_PASSWORD_1	Password for the PowerCenter Repository Service user. Data Integration Hub stores the password as an encrypted string.
PC_REPOSITORY_SECURITY_DOMAIN_1	Optional. Name of the Informatica security domain in which the PowerCenter Repository Service user is stored. Default is Native.
PC_DOMAIN_NAME_1	Name of the Informatica domain that contains the PowerCenter Integration Service that runs Data Integration Hub workflows.
PC_NODE_NAME_1	Node in the Informatica domain on which the PowerCenter Integration Service runs.
PC_ADMIN_USER_1	Name of the Informatica domain administrator.
PC_ADMIN_PASSWORD_1	Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.
PC_INTEGRATION_SERVICE_1	The name of the PowerCenter Integration Service that Data Integration Hub uses to run workflows.
PMREP_HOME_1	The absolute location of the pmrep command line utility.
Configure Informatica Data Quality domain settings using the following parameters:	
DIS_DOMAIN_NAME_1	Name of the Informatica domain that contains the Data Integration Service.
DIS_HOST_NAME_1	Host name of the node that runs the Data Integration Service.

Parameter	Description
DIS_PORT_NUMBER_1	Port number of the node that runs the Data Integration Service.
DIS_ADMIN_SECURITY_DOMAIN_1	Optional. Name of the Informatica security domain in which the Data Integration Service user is stored. Default is Native.
DIS_ADMIN_USER_1	Name of the Informatica domain administrator.
DIS_ADMIN_PASSWORD_1	Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.
DIS_INTEGRATION_SERVICE_1	The name of the Data Integration Service that Data Integration Hub uses to run batch workflows.
DIS_MODEL_REPOSITORY_SERVICE_1	The name of the Data Integration Service that Data Integration Hub uses to store batch workflows.
Configure Informatica Data Quality Command Line Utility using the following parameters:	
INFACMD_HOME_1	The absolute location of the infacmd command line utility.

Sample of the Installer Properties Configuration

Use the following sample to configure the installation properties file to install Data Integration Hub in a silent mode:

```
#Install or Upgrade
#------------------
IS_INSTALL=1
IS_UPGRADE=0

#PowerCenter pmrep Command Line Utility Location
#-------------------
PMREP_HOME_1=C:\\Informatica\\105\\server\\bin

#PowerCenter Version
#--------------------
PWC_VERSION_BEFORE_10_5=0
PWC_VERSION_10_5=1

#PowerCenter Location
#--------------------
POWERCENTER_HOME_1=C:\\Informatica\\105

#Installation Directory#Mandatory
#--------------------------------
USER_INSTALL_DIR=C:\\Informatica\\DIH

#Installation Components
#-----------------------
DX_SERVER=0
DX_DASHBOARD=0
PC_SERVER_PLUGIN=0
PC_CLIENT_PLUGIN=1
DIH_HADOOP_SERVICE=0
```

```
DIH_BIG_DATA_MANAGEMENT=0

#Metadata Repository
#-------------------
BLANK_USER=0
CONFIGURED_USER=1

#Metadata Repository Connection
#------------------------------
DB_TYPE_1=SQLServer
DB_CONNECTION_STRING_1=0
DB_HOST_1=<serverhost>
DB_PORT_1=<port_no>
DB_SID_1=<REPO_SID>
DB_CONNECTION_STRING_VALUE_1=jdbc:informatica:sqlserver://
<server>:<port_no>;DatabaseName=<DB_Name>;
DB_ODBC_STRING_VALUE_1=DRIVER={DataDirect 7.1 SQL Server Wire Protocol};UID=%1;PWD=
%2;Server=<server>, <port_no>;Database=<REPO_DB>
DB_WINDOWS_AUTHENTICATION_1=0
DB_USER_1=<db_username>
DB_PASSWORD_1=<db_password>

#Publication Repository Connection
#---------------------------------
STAGING_TYPE_1=SQLServer
STAGING_CONNECTION_STRING_1=0
STAGING_HOST_1=<serverhost>
STAGING_PORT_1=<port_no>
STAGING_SID_1=<STAGING_SID>
STAGING_WINDOWS_AUTHENTICATION_1=0
STAGING_USER_1=<db_username>
STAGING_PASSWORD_1=<db_password>
STAGING_CONNECTION_STRING_VALUE_1=jdbc:informatica:sqlserver://
<server>:<port_no>;DatabaseName=<DB_Name>;


#Operational Data Store
#----------------------
ODS_DB_USE_NEW=1
ODS_DB_USE_EXISTING=0

#Operational Data Store Database Connection
#------------------------------------------
ODS_DB_TYPE_1=SQLServer
ODS_DB_CONNECTION_STRING_1=0
ODS_DB_HOST_1=<server_host>
ODS_DB_PORT_1=<port_no>
ODS_DB_SID_1=<ODS_SID>
ODS_DB_JDBC_STRING_VALUE_1=
ODS_DB_WINDOWS_AUTHENTICATION_1=0
ODS_DB_USER_1=<db_username>
ODS_DB_PASSWORD_1=<db_password>

#User Authentication
#-------------------
INTERNAL_AUTH=1
INTERNAL_AUTH_DEFAULT=true
ISF_AUTH=0
ISF_AUTH_DEFAULT=false
KERBEROS_AUTH=0
KERBEROS_AUTH_DEFAULT=false

#Informatica Platform Authentication
#-----------------------------------
INFA_HOST_1=$INFA_HOST_1$
INFA_PORT_1= $INFA_PORT_1$
INFA_USERNAME_1=$INFA_USERNAME_1$
INFA_PASSWORD_1=
INFA_SECURITY_DOMAIN_1=$INFA_SECURITY_DOMAIN_1$
INFA_SECURITY_GROUP_1=$INFA_SECURITY_GROUP_1$
```

```
#Operation Console Administrator
#-------------------------------
CONSOLE_ADMIN_1=Administrator

#Data Integration Hub Document Store
#---------------------------
DATA_STORE_FOLDER_1=C:\\Informatica\\DIH\\DataIntegrationHub\\dih-data

#Web Server
#----------
TOMCAT_ENABLE_HTTPS_1=1
TOMCAT_HTTPS_PORT_1=18443
TOMCAT_EXISTING_KEYSTORE_FILE_1=0
TOMCAT_KEYSTORE_PASSWORD_1=
TOMCAT_KEYSTORE_FILE_PATH_1=
TOMCAT_ENABLE_HTTP_1=1
TOMCAT_PORT_1=18080
TOMCAT_SERVER_LISTENER_PORT_1=18005

#PowerCenter Location
#--------------------
POWERCENTER_HOME_1=C:\\Informatica\\105

#PowerCenter Web Services Hub
#----------------------------
PC_SCHEDULING_WORKFLOWS_1=1
PC_WEB_SERVICES_URL_1=http://<hostname>:7333/wsh/services/BatchServices/DataIntegration
PC_REPOSITORY_NAME_1=<PowerCenter_Repository_Name>
PC_REPOSITORY_HOST_1=<hostname>
PC_REPOSITORY_PORT_1=<PwC_Port_no>
PC_REPOSITORY_USER_1=<PwC_username>
PC_REPOSITORY_PASSWORD_1=<PwC_password>
PC_REPOSITORY_SECURITY_DOMAIN_1=<Optional>

#PowerCenter Domain Settings
#---------------------------
PC_DOMAIN_NAME_1=<PwC_Domain_Name>
PC_NODE_NAME_1=<PwC_node_Name_inDomain>
PC_ADMIN_USER_1=<PwC_username>
PC_ADMIN_PASSWORD_1=<PwC_password>
PC_INTEGRATION_SERVICE_1=<PwC_Integration_Service_Name>

#PowerCenter pmrep Command Line Utility Location
#-----------------------------------------------
PMREP_HOME_1=C:\\Informatica\\105\\server\\bin

#Informatica Data Quality Domain Settings
#----------------------------------------
DIS_DOMAIN_NAME_1=<IDQ_Domain_name>
DIS_HOST_NAME_1=<host_name>
DIS_PORT_NUMBER_1=<IDQ_Port>
DIS_ADMIN_SECURITY_DOMAIN_1=<Optional>
DIS_ADMIN_USER_1=<IDQ_Domain_username>
DIS_ADMIN_PASSWORD_1=<IDQ_Domain_password>
DIS_INTEGRATION_SERVICE_1=<DIS_Name>
DIS_MODEL_REPOSITORY_SERVICE_1=<MRS_Name>

#Informatica Data Quality infacmd Command Line Utility Location
#--------------------------------------------------------------
INFACMD_HOME_1=C:\\Informatica\\105\\isp\\bin
```

Running the Silent Installer

Before you run the installer in the silent mode, ensure that you configure the installer configuration file:

1. Configure the installer properties in a text file.

Note:

- For more information about parameters to configure in the installer properties file, refer to "Configuring the Installation Properties" on page 58.
- For using a sample of the installer properties file, refer to "Sample of the Installer Properties Configuration" on page 64.

2. Run the following command in the command prompt to silent install B2B Data Exchange using the installer properties file:
   - If you use the Windows operating system, run the following command:
     Install.exe -i silent -f <installer.properties>, where filename is the name of the file that contains the installer properties.
   - If you are using the UNIX operating system, run the following command: ./Install.bin -i silent -f <location>/installer.properties , where location is the location of the file that contains the installer properties. For example, /data/username/installers/105/installer.properties.

The silent installer runs in the background.

The silent installation fails if you incorrectly configure the properties file or if the installation directory is not accessible. View the installation log files and correct the errors. Then run the silent installation again.

Note: If an error occurs during pre-installation, the log is created in `USER_HOME\DXLogs\B2B_DX_InstallLog.txt` path. You must fix the error and trigger the silent mode installation.

CHAPTER 5

Post-Installation Tasks

This chapter includes the following topics:

- Post-Installation Tasks Overview, 68
- Register the Data Integration Hub Server Plug-in for PowerCenter, 69
- Set Up the Run Publication Subscription Web Service , 70
- Configure Credentials for Windows Authentication, 71
- Set Up Database Partitions on the Publication Repository, 71
- Configuring the Environment for a Hadoop Publication Repository, 72
- Log in to the Operation Console, 75
- Configure Connections to the Data Integration Hub Repositories, 75
- Configure Connections to Microsoft Azure SQL Database, 77
- Synchronize Data Integration Hub Users, 78
- Activate the Dashboard and Reports Component, 79
- Synchronize Data Integration Hub Users, 80

Post-Installation Tasks Overview

After you install Data Integration Hub, perform the steps that are relevant for your environment.

1. If you installed the Data Integration Hub server plug-in for PowerCenter, register the plug-in to the PowerCenter repository.
2. Configure PowerCenter to access Data Integration Hub.
3. If you want to use the Run Publication Subscription web service API, set up the web service.
4. If you installed the Data Integration Hub repositories on a Microsoft SQL Server and you selected to use Windows authentication, configure credentials for Windows authentication.
5. Set up partitions on the Data Integration Hub publication repository database that stores published data sets. Setting up partitions is highly recommended.
6. If you installed the Data Integration Hub Hadoop Service component, configure your environment for a Hadoop-based publication repository.
7. Set the environment variable `JRE_HOME` to `<DIH_HOME>/DataIntegrationHub/jdk1.8/jre`, if you installed Data Integration Hub on SUSE Linux operating system.
8. Start the Data Integration Hub services. For more information, see Chapter 7, “Starting and Stopping Data Integration Hub” on page 110.

Note: If you installed the Data Integration Hub Hadoop Service component on a different machine than the machine where you installed Data Integration Hub, start the services on both machines. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111.

9. Log in to the Data Integration Hub Operation Console.
10. Configure connections to the Data Integration Hub repositories in the Data Integration Hub Operation Console.
11. If you installed Data Integration Hub repositories on Microsoft Azure SQL databases, configure connections to the Microsoft Azure SQL databases.
12. If you installed Data Integration Hub with Informatica domain authentication or with Informatica domain with Kerberos authentication, synchronize Data Integration Hub users in the Data Integration Hub Operation Console.
13. If you installed the Dashboard and Reports component, activate the component.
14. If you want to configure subscriptions with the delivery option **Insert new rows and update changed rows**, and if the column values are expected to contain NULLs, you must configure the PowerCenter Integration Service property **Treat Null In Comparison Operators As** to `Low`.

Related Topics:

- "Overview of Starting and Stopping Data Integration Hub" on page 110

Register the Data Integration Hub Server Plug-in for PowerCenter

If you installed the Data Integration Hub server plug-in for PowerCenter, register the plug-in to the PowerCenter repository where you create Data Integration Hub workflows. If you do not register the plug-in, you cannot create Data Integration Hub transformations in the PowerCenter Designer, and Data Integration Hub cannot run PowerCenter workflows. Under these conditions, when Data Integration Hub attempts to run a PowerCenter workflow and fails, Data Integration Hub logs the following error in the PowerCenter workflow session log: `[REP_12400 Repository Error ([REP_57140] Object does not exist).`

The PowerCenter Repository Service must be running in exclusive mode when you register the plug-in. After the registration, restart the PowerCenter Repository Service in normal mode.

1. Log in to the Administration tool.
2. In the Navigator, select the Repository Service for which you want to register the plug-in.
3. On the Properties tab, edit the General Properties section and set the operating mode to Exclusive.
4. Restart the Repository Service.
5. After the Repository Service restarts, click the Plug-ins tab.
6. Click the link to register a Repository Service plug-in.
7. On the Register Plug-in for <Repository Service> page, click the **Browse** button to find the plug-in file.

   Select the following file in the directory where you installed the Data Integration Hub server plug-in for PowerCenter:

   `<DIHInstallationDir>/powercenter/plugin`*`VERSION`*`/dxplugin.xml`

   Note: The Data Integration Hub installer creates separate plug-in folders for each PowerCenter version. Make sure to select the plug-in folder for the PowerCenter version that you are using.

8. Enter the Repository Service administrator user name and password to log in to the repository.

   If the security group field appears, select the security group for the Repository Service administrator.

9. Click **OK**.

   The Repository Service registers the plug-in. Verify that the list of registered plug-ins for the Repository Service includes the Data Integration Hub transformations.

10. On the Properties tab, edit the General Properties section and set the operating mode to Normal.
11. Restart the PowerCenter Integration Service.

Set Up the Run Publication Subscription Web Service

If you want to use the Run Publication Subscription web service API, import the web service workflow into PowerCenter.

If the PowerCenter services and the Data Integration Hub server run on separate machines, verify that the settings for the Data Integration Hub server are set correctly.

Importing the Web Service to PowerCenter

When you install Data Integration Hub, the Data Integration Hub web service workflow is installed in the `<DIH_InstallationDir>\powercenter\webservices` folder. You must import the Data Integration Hub web service workflow into PowerCenter before you can access the web service.

To use the web service, the Informatica domain must contain the following services:

- PowerCenter Repository Service
- Web Services Hub
- PowerCenter Integration Service

1. Use the PowerCenter Repository Manager to import the following workflow file into the PowerCenter repository: `wf_m_DIH_WS_StartPublicationSubscription.xml`.
2. In the Web Services Hub console, verify that the Data Integration Hub web service is correctly imported into PowerCenter. If the import process is successful, the list of valid services includes the Data Integration Hub web service.
3. You can use the Try-It application in the Web Services Hub console to test the Data Integration Hub web service. On the **XML Input** tab, enter the data into the SOAP message and click **Send**. To avoid authentication errors, do not use the **Form Input** page to test the Data Integration Hub web service.

   After you verify that the web service is working, you can create a client application to send requests to the web service.

Verifying the Server Settings

If the PowerCenter services and the Data Integration Hub server run on separate machines, verify that the settings for the Data Integration Hub server are set correctly.

In the Informatica Administrator, select the PowerCenter Integration Service that runs Data Integration Hub workflows. Verify the following environment variable settings:

Environment Variable	Value
DX_SERVER_URL	The RMI URL for Data Integration Hub server. For example: `rmi://<DIHServerHostname>:<RMIPort>/TSSARuntime` **Note:** <RMIPort> must match the *dx.rmi.port* parameter in the `dx-configuration.properties` file. The default port number is 18095.

Configure Credentials for Windows Authentication

If you installed any of the Data Integration Hub repositories on a Microsoft SQL Server and you selected to use Windows authentication, configure the credentials that Data Integration Hub uses to access the Microsoft SQL Server instance.

Before you start the configuration process, verify that all Data Integration Hub Windows services are stopped and that the Data Integration Hub Operation Console and the Data Integration Hub server are not running.

1. Access the Windows **Services** window.
2. Double-click the service **Informatica Data Integration Hub Server *version***.

   The Data Integration Hub Server Properties window appears.
3. Select the **Log On** tab.
4. Select **This account**, click **Browse**, and then specify a user account in the **Select User** dialog box. When you are finished, click **OK**.
5. Type the password for the user account in **Password** and in **Confirm password**, and then click **OK.**
6. Repeat steps 2 through 5 for the service **Informatica Data Integration Hub Console *version***.

Set Up Database Partitions on the Publication Repository

It is highly recommended that you set up partitions on the publication repository database that stores published data sets. Partitions divide the stored published data set according to creation date, where each partition stores data sets that all applications publish on a specific date. By default, the publication repository does not use partitions.

When you set up partitioning, Data Integration Hub reduces fragmentation by deleting expired data sets and reducing the time of data set creation and data consumption. You can set up partitions on Oracle and Microsoft SQL Server databases.

Use the Data Integration Hub repository utility command enablePartitioning to create and manage the partitions in the publication repository database. The partitions state appears in the dih.staging.use.partitions.default system property. For more information about the Data Integration Hub repository utility see the *Data Integration Hub Administrator Guide*.

If you switch between a partitioned database and a non-partitioned database, the change affects topics, publications, and subscriptions that you create after the switch. Therefore, it is recommended that you choose a partition state before you start to use Data Integration Hub. To switch the partition state for existing topics, publications, or subscriptions, contact Informatica Global Customer Service.

Configuring the Environment for a Hadoop Publication Repository

If you installed the Data Integration Hub Hadoop Service component, configure the environment for a Hadoop publication repository.

To complete the task you perform the following actions:

1. Configure settings in the `dx-configuration.properties` file.
2. Verify references in the `dih-hadoop-service.xml` file.
3. Enable an option in Cloudera Manager.

1. On the machine where the Data Integration Hub Hadoop Service is installed, use a text editor to open the `dx-configuration.properties` file from the following location:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/
   ```

2. Set the properties in the following sections of the file and then save the file:

   HIVE settings

Property	Description
dih.hadoop.hive.username	User name for connecting to the Apache Hive server.
dih.hadoop.hive.password	Password of the user that connects to the Apache Hive server.

Property	Description
dih.hadoop.hive.url	URL to the Apache Hive server in the following format: `jdbc:hive2://<hostname>:<port>/<schema>` Where: - `hostname` is the host name or IP number of the server. - `port` is the port number of the server. Default is 10000. - `schema` is the schema used with the Hive warehouse. Default is default. If the Hive warehouse uses a non default schema, set the property `dih.hadoop.service.warehouse.dir`.

dih.hadoop.service.warehouse.dir	Path to the Hive warehouse directory. Required if the Apache Hive server uses a non default schema. If the Apache Hive server uses a default schema, do not enter a value for this property.
	For example:
	`dih.hadoop.hive.username=hive`
	`dih.hadoop.hive.password=password`
	`dih.hadoop.hive.url=jdbc:hive2://hive_host:10000/myschema`
	`dih.hadoop.service.warehouse.dir=/user/hive/mydatawarehousedir`

SPARK settings

Property	Description
dih.hadoop.service.spark.version	Apache Spark version. Takes the following values: - 1.2. Apache Spark version 1.2. - 1.3. Apache Spark version 1.3 and higher. Default is 1.3.

dih.hadoop.service.spark.url	Apache Spark URL.
	- If Apache Spark is running in YARN mode, use the default value:
	`dih.hadoop.service.spark.url=yarn`
	- If Apache Spark is running in standalone mode, enter the URL in the following format:
	`spark://master_host:<port_number>`
	Where:
	- `master_host` is the Master daemon which coordinates the operations of the Workers, which run the executors.
	- `<port_number>` is the port number of the Master daemon. Default is 7077.
	For example:
	`spark://Mymasterhost:7077`
	Note: The value you enter here must be identical to the value that is shown in the Spark console. By default, the Spark console is located at `http://<host_name>:18080`.

dih.hadoop.service.spark.additional.args	Additional arguments for running jobs. For example: `--executor-memory 20G --total-executor-cores 100` For a complete list of arguments see the Spark documentation.

KERBEROS settings

If the Hadoop cluster uses Kerberos authentication, configure the following settings:

Property	Description
dih.hadoop.principal	Kerberos principal name in the following format: `<principal>/<domain>@<realm>`

dih.hadoop.keytab.path	Location and name of the keytab file.
	For example:
	`dih.hadoop.principal=infa/admin@informatica.com`
	`dih.hadoop.keytab.path=/etc/security/keytabs/infa.keytab`
	Note: The file name does not have to be `infa.keytab`.

3. On the machine where the Data Integration Hub Hadoop Service is installed, use a text editor to open the `dih-hadoop-service.xml` file from the following location:

   `<DIHInstallationDir>/DataIntegrationHub/tomcat/conf/Catalina/localhost`

4. Verify that the file contains the correct references to your Hadoop classpath configuration. By default, the file references a Cloudera VM configuration in the section `Cloudera VM - sample default configurations`.

   The file also contains references to Cloudera Manager and to Hortonworks VM, in the following commented out sections: `Cloudera CDH - sample cloudera manager configurations` and `Hortonworks VM - sample default configurations`. If required, comment out the `Cloudera VM - sample default configurations` section and uncomment the section that is appropriate to your configuration.
5. In Cloudera Manager, enable the option **Bind NameNode to Wildcard Address** and then restart the HDFS service.

Log in to the Operation Console

You log in to the Operation Console with the administrator user account that you defined during installation.

If you use Informatica platform authentication, verify that all user accounts and user passwords exist on the authentication server.

1. Make sure that the Operation Console service is running.
2. Access the Operation Console login page.
   - On Microsoft Windows or UNIX operating systems, open a browser window and enter the URL for the Operation Console in one of the following formats:

     ```
     HTTP: http://<HostName>:<HTTPPortNumber>/
     HTTPS: https://<HostName>:<HTTPSPortNumber>/
     ```
   - On Microsoft Windows operating systems, click the Desktop shortcut to open a new browser window to the Operation Console URL.
3. Enter your user name and password and click **Log In**.

Configure Connections to the Data Integration Hub Repositories

Configure connections to the Data Integration Hub repositories. After you configure the connections, Data Integration Hub creates the connections in the PowerCenter repository. Data Integration Hub uses the connections details to create workflows for publications and for subscriptions, and to create data sources and data targets in PowerCenter .

To configure the connections, you must be logged in to the Data Integration Hub Operation Console with the administrator user account.

1. In the Navigator, click **Hub Management** > **Connections**.

   The **Connections** page appears, showing the DIH__REPO and DIH__STAGING connections.
2. Click the **Edit** icon next to the DIH__REPO connection.

   The **Edit Connection** page appears, showing the DIH__REPO connection details. DIH__REPO is the connection to the Data Integration Hub metadata repository that you defined during installation.

3. Configure the **Database Name** if applicable and the **Data Access** properties of the connection. For more information, see the *Data Integration Hub Administrator Guide*.
4. In the **Edit Connection** page, click **Save**.

   The **Edit Connection** page closes.
5. Click the **Test Connection** icon next to the DIH__REPO connection.

   Data Integration Hub tests the connection. The process might take a few moments.
6. When the message "Connection successful" shows, click **OK** in the **Test Connection** dialog box.

   Note: The first time that you test connections in Data Integration Hub, the test might fail with an error. Wait one minute and then test the connection.

   If the test fails after you wait a few seconds, delete the folder `DIH__SYSTEM_WORKFLOWS` in PowerCenter, and then test the connection.
7. In the Navigator, click **Connections**.

   The **Connections** page appears.
8. Click the **Edit** icon next to the DIH__STAGING connection.

   The **Edit Connection** page appears, showing the DIH__STAGING connection details. DIH__STAGING is the connection to the Data Integration Hub publication repository that you defined during installation.
9. Repeat steps 3 through 6 to configure and test the DIH__STAGING connection.
10. If you installed the Data Integration Hub Hadoop Service component, click the **Edit** icon next to the DIH__STAGING__HADOOP connection.

    The **Edit Connection** page appears, showing the DIH__STAGING__HADOOP connection details. DIH__STAGING__HADOOP is the connection to the Data Integration Hub Hadoop-based publication repository.
11. Configure the following Hadoop Settings properties of the connection:

    Hadoop Cluster User

    Name of the Hadoop cluster user account.

    NameNode URI

    Use the following connection URI:

    ```
    hdfs://<namenode>:<port>/<path>
    ```

    Where

    - `<namenode>` is the host name or IP address of the NameNode.
    - `<port>` is the port on which NameNode listens for remote procedure calls (RPC).
    - `<path>` is a working directory in HDFS. The Hadoop cluster user account must have full permissions to the working directory.

    For example:

    ```
    hdfs://mycluster:8020/user/claudera
    ```

    Note: `hdfs://<namenode>:<port>` must be identical to property `fs.defaultFS` as it appears in the file `core-site.xml`.

    Hadoop Distribution

    Type of Hadoop distribution that the Hadoop cluster uses.
12. In the **Edit Connection** page, click **Save**.

    The **Edit Connection** page closes.

13. Click the **Test Connection** icon next to the DIH__STAGING__HADOOP connection.

    Data Integration Hub tests the connection. The process might take a few moments.

14. When the message "Connection successful" shows, click **OK** in the **Test Connection** dialog box.

Configure Connections to Microsoft Azure SQL Database

If you install Data Integration Hub repositories on a Microsoft Azure SQL database, configure connections to the database in the Data Integration Hub and PowerCenter applications. After you configure connections, Data Integration Hub creates corresponding connections in the PowerCenter repository. Data Integration Hub uses connection details to create workflows for publications and for subscriptions, and to create data sources and data targets in PowerCenter.

To configure the connections, you must log in to the following applications:

- The Data Integration Hub operation console with the administrator user account.
- The PowerCenter client operation console with the administrator user account.

In order to configure connections to the Microsoft Azure SQL database, edit the DIH_REPO connection in the Data Integration Hub console, enable DSN for the DIH_REPO connection in PowerCenter, and then test the connection in Data Integration Hub. Perform the same procedure for the DIH_STAGING connection.

1. Perform the following steps in the Data Integration Hub operation console to edit the DIH___REPO connection:

    a. In the Navigator, click **Hub Management** > **Connections**.

    b. Click the **Edit** icon next to the DIH___REPO connection.

    The **Edit Connection** page appears, showing the DIH___REPO connection details. DIH___REPO is the connection to the Data Integration Hub metadata repository that you defined during installation.

    c. Configure **Data Access** properties as follows:

    - Enter **Database Name** in the following format: `dbname`
    - Enter **Server Name** in the following format: `hostname`

    For more information, see the *Data Integration Hub Administrator Guide*.

    d. In the **Edit Connection** page, click **Save**.

    The **Edit Connection** page closes.

2. Perform the following steps in PowerCenter to create a DSN name for the DIH__REPO connection:

    a. Select **PwC Workflow Manager** > **Connections** > **Relational**.

    b. Select **Microsoft SQL Server Connections** and click **Edit**.

    c. Enable DSN and enter the name of the data source for the DIH_Repo conection, that you created during the environment configuration.
    For more information about the data source name, see "Create a Data Source Name for Microsoft Azure SQL Database" on page 20.

    d. Click **Save**.

3. Perform the following steps to test the DIH_REPO connection:
   a. Click the **Test Connection** icon next to the DIH___REPO connection in the Data Integration Hub console.

      Data Integration Hub tests the connection. The process might take a few moments.
   b. When the message "Connection successful" shows, click **OK** in the **Test Connection** dialog box.

      Note: The first time that you test connections in Data Integration Hub, the test might fail with an error. Wait one minute and then test the connection.

      If the test fails after you wait a few seconds, delete the folder `DIH__SYSTEM_WORKFLOWS` in PowerCenter, and then test the connection.
4. Repeat the above procedure for DIH_STAGING connection to configure connections to the DIH_STAGING repository.

Synchronize Data Integration Hub Users

If you installed Data Integration Hub with Informatica domain authentication or with Informatica domain with Kerberos authentication, synchronize Data Integration Hub users in the Data Integration Hub Operation Console.

To synchronize users in the Informatica security domain with Data Integration Hub, the following conditions must be true:

- The Informatica security domain is configured on the **Security** page of Informatica Administrator.
- At least one security group in the Informatica security domain contains the Data Integration Hub users to synchronize.
- The Data Integration Hub system property `dx.authentication.groups` contains the list of groups from the Informatica security domain to synchronize, in the following format:

  ```
  <group name>@<security domain> [;<groupname>@<security domain>]
  ```
- One of the groups that are defined in `dx.authentication.groups` contains the user that performs the synchronization.
- The user that is defined in the Data Integration Hub system property `pwc.repository.user.name` has privileges to manage users, groups, and roles.
- The Data Integration Hub user has privileges to synchronize users.

1. In the Navigator, click **Administration** > **Users**.

   The **Users** page appears.
2. Click **Synchronize users** and follow the instructions on the screen.
3. For each user that is added to the **Users** page, assign the required privileges. For more information see the *Data Integration Hub Administrator Guide*.

Activate the Dashboard and Reports Component

To view the reports in the dashboard using operational data store, you must install the Dashboard and Reports component and perform the following tasks to activate the component:

- Register the license of the Dashboard and Reports component.
- Import the operational data store event loader to PowerCenter.

Register the Dashboard and Reports License

This section applies only if you use the Dashboard using operational data store.Register the license of the Data Integration Hub Dashboard and Reports component.

1. Contact Informatica Global Customer Support to receive the Logi Info Dashboard license files.
2. Start the Data Integration Hub services.
3. Move the file `_Settings.lgx` from the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\webapps\dih-dashboard\_Definitions`

   To the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\shared\classes`

   Rename the file to the following name:

   `dx_dashboard_configuration.xml`
4. Copy the Logi Info Dashboard license file `_Settings_encrypted.lgx` to the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\webapps\dih-dashboard\_Definitions`
5. Rename the file `_Settings_encrypted.lgx` to `_Settings.lgx`.
6. Restart the Data Integration Hub services.

If the IP addresses of the machine that hosts Data Integration Hub change any time after the installation, you must update the IP addresses in the Logi Info Dashboard license file. For more information, see "Updating the Dashboard Configuration File" on page 120.

Import the Operational Data Store Event Loader Workflow to PowerCenter

This section applies only if you use the Dashboard using operational data store. Import the operational data store (ODS) event loader workflow to load event information from the Data Integration Hub repository to the Data Integration Hub ODS.

If you use an existing workflow with the name DX_ETL, rename the existing workflow in PowerCenter Repository Manager before you import the ODS event loader workflow, or import the workflow to a different folder.

Note: After you import the ODS event loader workflow, do not run the workflow manually. The workflow must start at the scheduled time. If you start the workflow manually it might fail to store aggregated events in the Data Integration Hub ODS.

1. In the PowerCenter Workflow Manager, select **Connections** > **Relational**.
2. Add the **DX_REPO** connection for the Data Integration Hub repository.
3. Add the **DX_ODS** connection for the Data Integration Hub ODS.

4. In the PowerCenter Repository Manager, import the Data Integration Hub ODS workflow file. The name of the workflow file depends on type of database on which the ODS is installed.

Database Type	Workflow Location and Name
Oracle	<DIHInstallationDir>\powercenter\ETL\DX_ETL.xml
Microsoft SQL Server	<DIHInstallationDir>\powercenter\ETL\DX_ETL_SQLSERVER.xml

5. In the PowerCenter Workflow Manager, connect to the PowerCenter repository.
6. If Data Integration Hub was previously installed at your site, and you want to exclude old events from the dashboard reports, instruct the workflow to load only events that finished processing after a specific date and time. Click **Workflow** > **Edit** > **Variables** and change the value of the **$$WF_Last_Load_End_Time** variable.

   Note: Do not change the variable after the first time the workflow runs.
7. By default, the workflow runs every 15 minutes. You can schedule the workflow start time.
8. Right-click the PowerCenter Integration Service that you want to assign to the ODS event loader workflow and select **Assign to Workflows**.

   The **Assign Integration Service** dialog box appears.
9. Select the **DX_ETL** check box and then click **Assign**.

   PowerCenter assigns the Data Integration Service to the ODS event loader workflow.

Synchronize Data Integration Hub Users

If you installed Data Integration Hub with Informatica domain authentication or with Informatica domain with Kerberos authentication, synchronize Data Integration Hub users in the Data Integration Hub Operation Console.

1. In the Navigator, click **Administration** > **Users**.

   The **Users** page appears.
2. Click **Synchronize users** and follow the instructions on the screen.
3. For each user that is added to the **Users** page, assign the required privileges. For more information see the *Data Integration Hub Administrator Guide*.

CHAPTER 6

Upgrading Data Integration Hub

This chapter includes the following topics:

- Upgrading Data Integration Hub Overview, 81
- Before You Upgrade, 82
- Upgrading Data Integration Hub on a Windows Operating System, 83
- Upgrading Data Integration Hub on a UNIX Operating System, 98
- After You Upgrade, 105

Upgrading Data Integration Hub Overview

You can directly upgrade the following versions of to the latest version:

- 10.4.1
- 10.4.0

When you upgrade Data Integration Hub, the installer backs up the files of the previous version of Data Integration Hub and installs the new version. If the document store is found under the Data Integration Hub Installation folder, you must move the document store to its original location after the upgrade process completes and before you start the DIH service. You can create a new repository for the new version or you can use the existing repository. If you use the repository from the previous version, the installer upgrades the repository to the latest version. When you upgrade the Data Integration Hub repository, you cannot change the database server type and server location.

Before you start the upgrade process, perform the procedures that are described in "Before You Upgrade" on page 82. Then run the Data Integration Hub installer. After the upgrade, perform the procedures that are described in "After You Upgrade" on page 105.

Note: During the upgrade you cannot change the user authentication method that Data Integration Hub uses. To change the user authentication method you must first upgrade the system and then switch to the required authentication method. For information about switching between authentication methods see the *Data Integration Hub Administrator Guide*.

Before You Upgrade

To prepare for the upgrade, perform the following tasks:

1. Verify that you have the user names and passwords for the required database accounts.
2. Verify that no publications or subscriptions are running and that all events are in a Final state.

   Tip: You can view the state of all events in the Data Integration Hub Operation Console, on the **Event List** page.
3. Ensure that the all topics are in a Valid status.

   Tip: You can view the status of all topics in the Data Integration Hub Operation Console, on the **Topics** page.
4. Ensure that the retention management job is not running. If the retention management job is running, you will see an incomplete System Event for the retention management job.

   Note: If the publication repository is used by applications other than Data Integration Hub, you might lose data in the tables that are not managed by Data Integration Hub when the upgrade is complete.
5. Ensure that the `dx.endpoint.file.prefix` property is empty.

   Note: This property is the landing zone path. Earlier releases of Data Integration Hub do not support this property. If this property value is not empty, then set it to an empty value.
6. Stop all Data Integration Hub services. The Data Integration Hub upgrade modifies the Data Integration Hub files. The installer cannot proceed if the files are in use.

   Note: If the Data Integration Hub Hadoop Service component is installed on a different machine than the machine where you installed Data Integration Hub, stop the services on both machines. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111.
7. Unschedule all custom workflows in PowerCenter. In PowerCenter, unschedule all custom workflows and verify that they do not run until the upgrade is complete.
8. Back up the Data Integration Hub repository to be upgraded. Use the database server backup utility to back up the repository. This ensures that you can recover from any errors that you encounter during the upgrade.
9. Back up the Data Integration Hub publication repository to be upgraded. Use the database server backup utility to back up the repository. This ensures that you can recover from any errors that you encounter during the upgrade.
10. Back up the existing Data Integration Hub installation folder. Perform this action to help ensure that you can recover from any errors encountered during the upgrade, and that, after the upgrade, you can reapply modifications that were made to the configuration in previous versions.
11. If the PowerCenter services are not installed on the same machine where Data Integration Hub is installed and you have upgraded the pmrep command line utility after you installed the previous version of Data Integration Hub, clean up all CNX files from the `Tmp` folder on your root directory.

Upgrading Data Integration Hub on a Windows Operating System

Upgrade Data Integration Hub on Windows operating systems in graphical mode. On UNIX operating systems, upgrade Data Integration Hub in console mode.

Before you install, verify that your environment meets the minimum system requirements, perform the pre-installation tasks, and verify that the PowerCenter services are running.

Note: During the upgrade, Data Integration Hub saves log files in the home directory of the user in the subdirectory named `DXLogs`. If the upgrade does not complete successfully, you can view the log files in this location.

Step 1. Run the Installer

1. Log in to the machine with the user account that you want to use to install Data Integration Hub.

   To prevent permission errors, use the same account to install Data Integration Hub and PowerCenter.
2. Close all other applications.
3. Run `Install.exe` from the directory where you downloaded the installer.

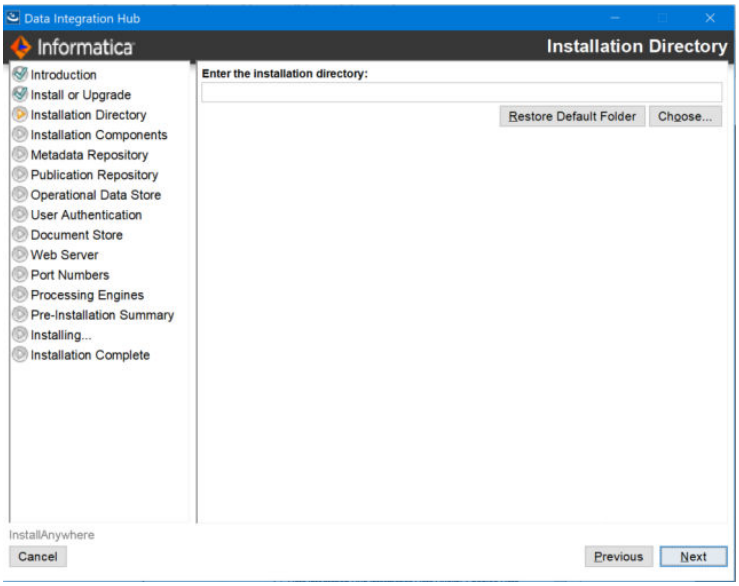
   The **Introduction** page appears.
4. Read the instructions, and then click **Next**.

   The **Install or Upgrade** page appears.
5. Select the option to upgrade Data Integration Hub, and then click **Next**.

   The **Installation Directory** page appears.

Step 2. Define Installation Settings

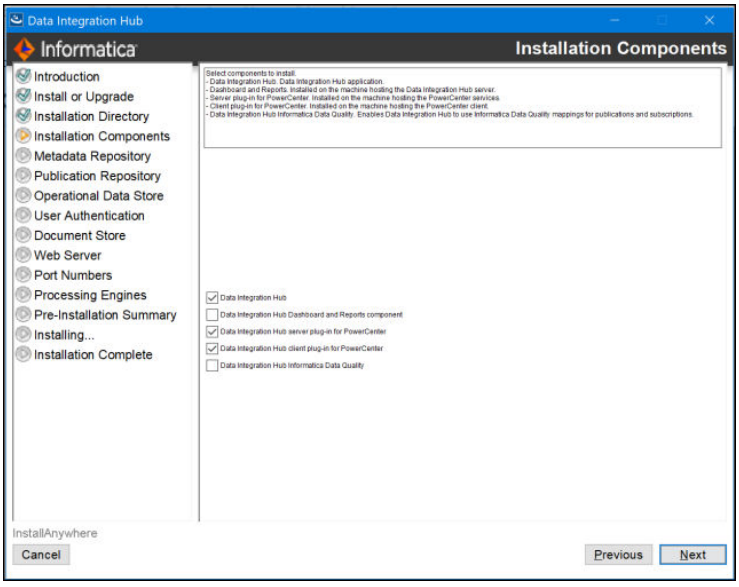
1. On the **Installation Directory** page, enter the absolute path to the installation directory or accept the default directory.



Note: You must select the same installation directory where you installed the previous Data Integration Hub version.

2. Click **Next**.

The **Installation Components** page appears.



3. Select the components to install:

Data Integration Hub

Installs the core Data Integration Hub application.
Selected by default.

Data Integration Hub Dashboard and Reports

Installs the Data Integration Hub Dashboard and Reports component. Select this component to view the reports in the Dashboard using operational data store.You must install Data Integration Hub to install the Dashboard and Reports component.
Cleared by default.

Note:

- The Dashboard using metadata directory is installed by default. Select this component to view the reports in the Dashboard using operational data store. After installation, you must ensure to set the `dx.dashboard.ods.page.show` system property to `true` to view the reports in the Dashboard using operational data store. For more information, see the *Data Integration Hub Administrator Guide* and *Data Integration Hub Operator Guide*.
- If you install the Dashboard and Reports component, you must import the operational data store event loader after you install Data Integration Hub.
- If you install the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, you must configure the repository connections in PowerCenter Workflow Manager. For details, see "Configuring Repository Connections on PowerCenter Version 10" on page 122.

Data Integration Hub PowerCenter server plug-in

Installs the Data Integration Hub plug-in for the PowerCenter services. After the installation, you register the plug-in to the PowerCenter repository.
Selected by default.

Data Integration Hub PowerCenter client plug-in

Installs the Data Integration Hub plug-in for the PowerCenter Client. Install this component on every machine that runs the PowerCenter Client.
Selected by default.

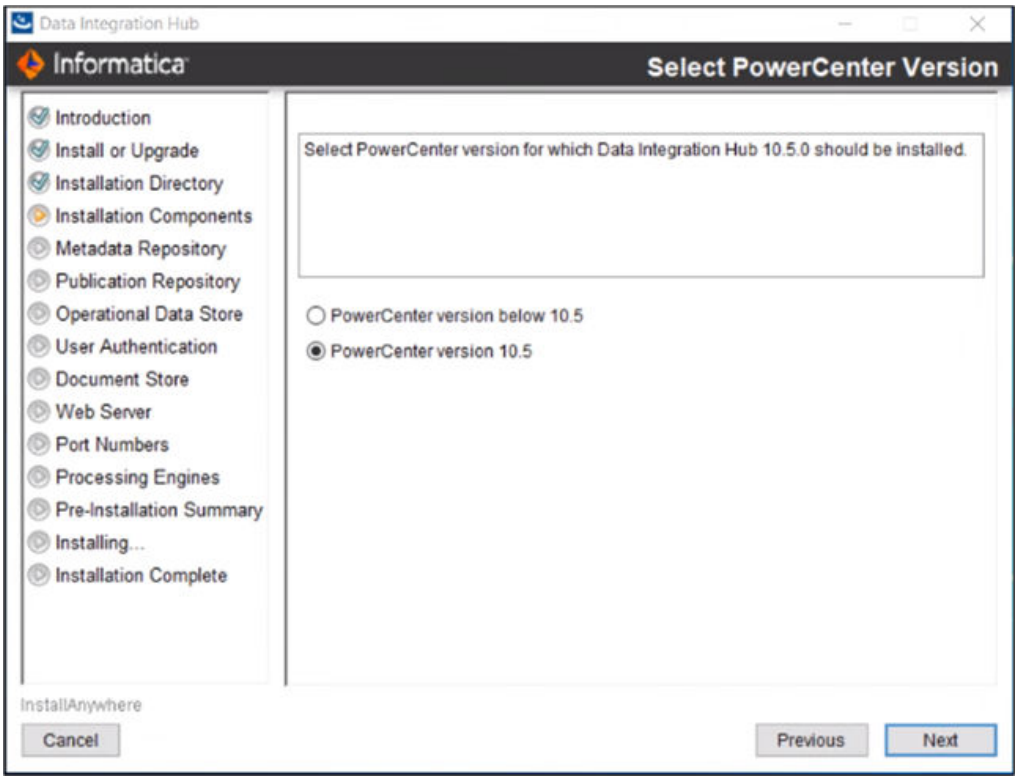
Data Integration Hub Informatica Data Quality

Enables Data Integration Hub to use Data Quality mappings for custom publications and subscriptions.
Select this component if you want to create custom publications and subscriptions that use Data Quality mappings in Data Integration Hub.

Cleared by default.

4. Click **Next**.

The **Select PowerCenter Version** page appears.



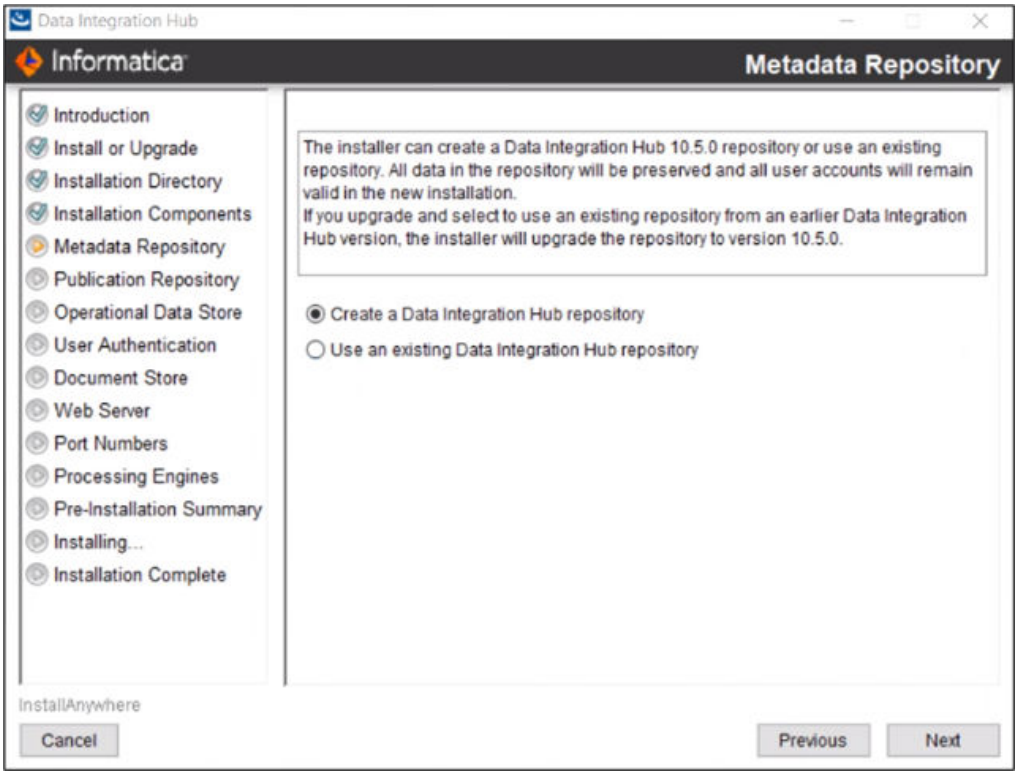
5. Select the PowerCenter version for which to install Data Integration Hub and click **Next**.

   The **Metadata Repository** page appears.

Step 3. Configure Data Integration Hub Repositories

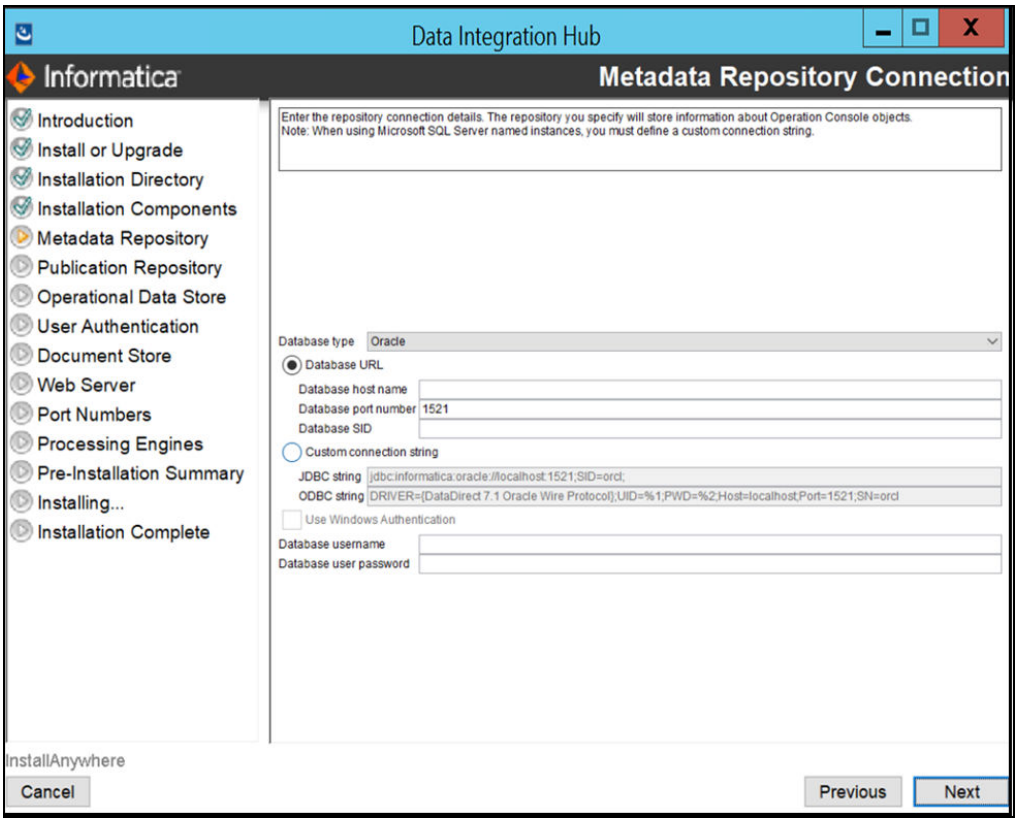
1. On the **Metadata Repository** page, select one of the following options:
   - **Create a Data Integration Hub repository**. Creates a repository in the database.

- **Use an existing Data Integration Hub repository**. Uses the tables and data in an existing Data Integration Hub repository and upgrades the repository.



2. Click **Next**.

The **Metadata Repository Connection** page appears.



3. Enter values in the following fields:

Database type

Type of database to use for the Data Integration Hub metadata repository. You can choose one of the following options:

- Oracle
- Microsoft SQL Server

Database URL

Location of the database.
If you select this option, enter the values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port**. Port number for the database. The default port number for Oracle is 1521. The default port number for Microsoft SQL Server or Microsoft Azure SQL Database is 1433.
- **Database SID**. System identifier for the database if the database is Oracle. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name.

Custom Connection String

Connection string to the database.

If you select this option, enter values in one of the following fields:

- **JDBC string**. JDBC connection string to the metadata repository.
- **ODBC string**. ODBC connection string to the metadata repository. Available if you install the PowerCenter Client plug-in. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

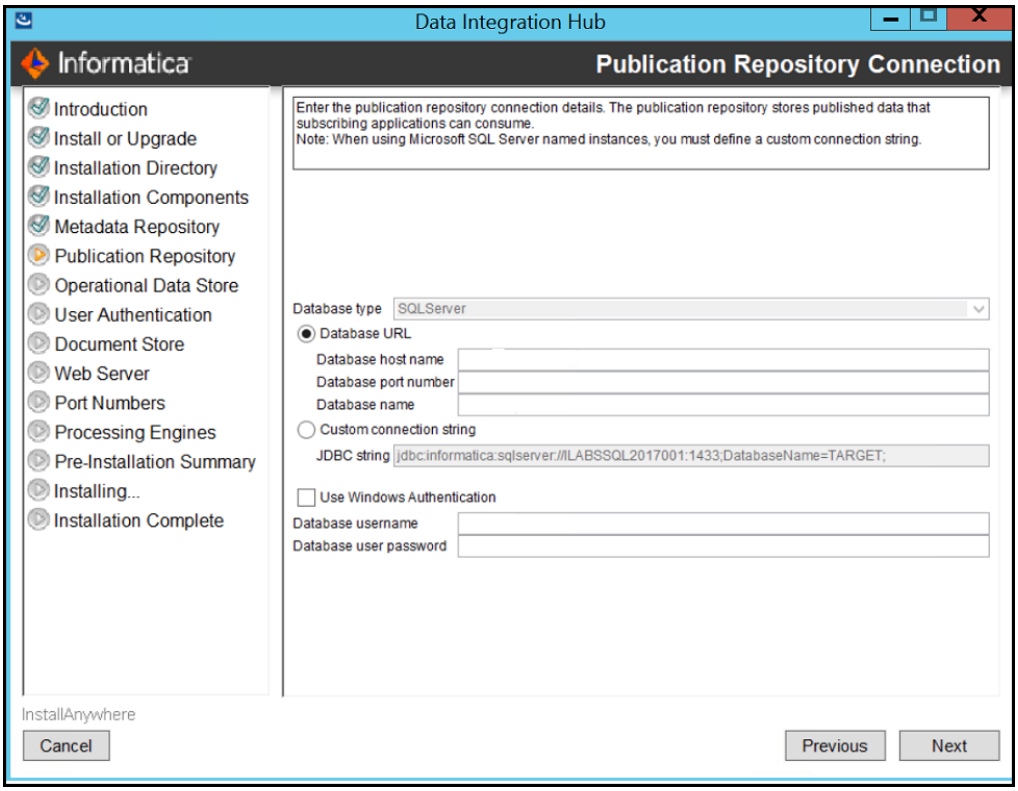
Name of the database user account for the database where you do not use Windows authentication.

Database user password

Password for the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

4. Click **Next**.

The **Publication Repository Connection** page appears.



5. Enter values in the following fields:

Database type

Type of database to use for the publication repository. The database type must match the Data Integration Hub metadata repository database type and appears in read-only mode.

Database URL

Location of the database.
If you select this option, enter values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database 1433.
- **Database SID**. System identifier for the database if the database is Oracle. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Name of the database instance.

Custom Connection String

JDBC connection string to the database.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

Name of the database user account for the database where you do not use Windows authentication.

Database user password

Password of the database account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

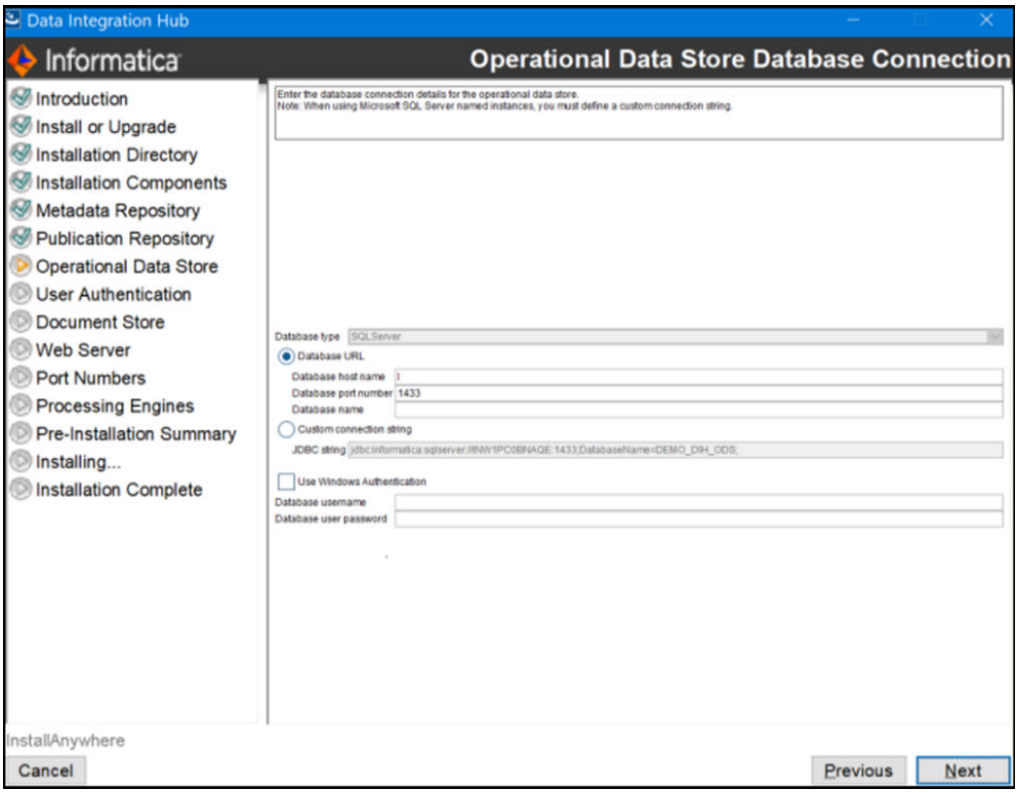
6. Click **Next**.

   If you selected the **Data Integration Hub Dashboard and Reports** component, the **Operational Data Store** page appears. If you did not select the Dashboard and Reports component, go to "Step 5. Configure the Web Server " on page 92.

Step 4. Set Up the Operational Data Store

1. This applies only if you want to use the legacy dashboard. On the **Operational Data Store** page, select one of the following options:
   - **Create an operational data store repository**. Creates an operational data store repository in the database.
   - **Use an existing operational data store repository**. Uses the tables and data in an existing operational data store repository.
2. Click **Next**.

The **Operational Data Store Database Connection** page appears.



3. Enter values in the following fields:

Database URL

Location of the database. If you select this option, enter the values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port number**. Port number for the database. The default port number for an Oracle database is 1521. The default port number for a Microsoft SQL server or a Microsoft Azure SQL database is 1433.
- **Database SID**. System identifier for the database if you select Oracle as the database. Enter either a fully qualified ServiceName or a fully qualified SID.

  Note: It is recommended that you enter a ServiceName in this field.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name.

Custom Connection String

Connection string to the database. If you select this option, enter values in one of the following fields:

- **JDBC string**. JDBC connection string to the Operational Data Store.
- **ODBC string**. ODBC connection string to the Operational Data Store. Available if you install the PowerCenter Client plug-in. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

Use Windows Authentication

Instructs Data Integration Hub to authenticate user names against the Microsoft Windows authentication mechanism. Available when you select a Microsoft SQL Server database.

Database username

Name of the operational data store user account for the database where you do not use Windows authentication.

Database user password

Password for the operational data store account for the database where you do not use Windows authentication. Data Integration Hub stores the password as an encrypted string.

4. Click **Next**.

   The **Web Server** page appears.

Step 5. Configure the Web Server

1. On the **Web Server** page enter values in the following fields:

   Enable HTTPS

   Instructs Data Integration Hub to use secure network communication when you open the Operation Console in the browser. If you select HTTPS and HTTP, the Operation Console switches existing HTTP connections with HTTPS connections.

   Connector port number

   Port number for the Tomcat connector to use when you open the Operation Console with HTTPS. The default value is 18443.

   Use a keystore file generated by the installer

   Instructs the installer to generate a keystore file with an unregistered certificate. If you select this option, ignore the security warning that you receive from the browser the first time you open the Operation Console.

   Use an existing keystore file

   Instructs the installer to load an existing keystore file. Enter values in the following fields:

   - Keystore password. Password for the keystore file.
   - Keystore file. Path to the keystore file.

   The keystore file must be in the Public Key Cryptography Standard (PKCS) #12 format.

2. Click **Next**.

   If you selected to install the Data Integration Hub server plug-in for PowerCenter or the Data Integration Hub client plug-in for PowerCenter components, the **PowerCenter Location** page appears. If you did not select the PowerCenter server or client components, the **PowerCenter Web Services Hub** page appears.

Step 6. Configure PowerCenter Settings

1. On the **PowerCenter Web Services Hub** page, enter the PowerCenter web services details.

   Web Services Hub URL

   URL that the PowerCenter Web Services Hub uses to process publication and subscription workflows.

   Service name

   Name of the PowerCenter Repository Service.

   Node host name

   Host name of the node that runs the PowerCenter Repository Service.

   Node port number

   Port number of the node that runs the PowerCenter Repository Service.

   Username

   Name of the PowerCenter Repository Service user.

   Password

   Password for the PowerCenter Repository Service user. Data Integration Hub stores the password as an encrypted string.

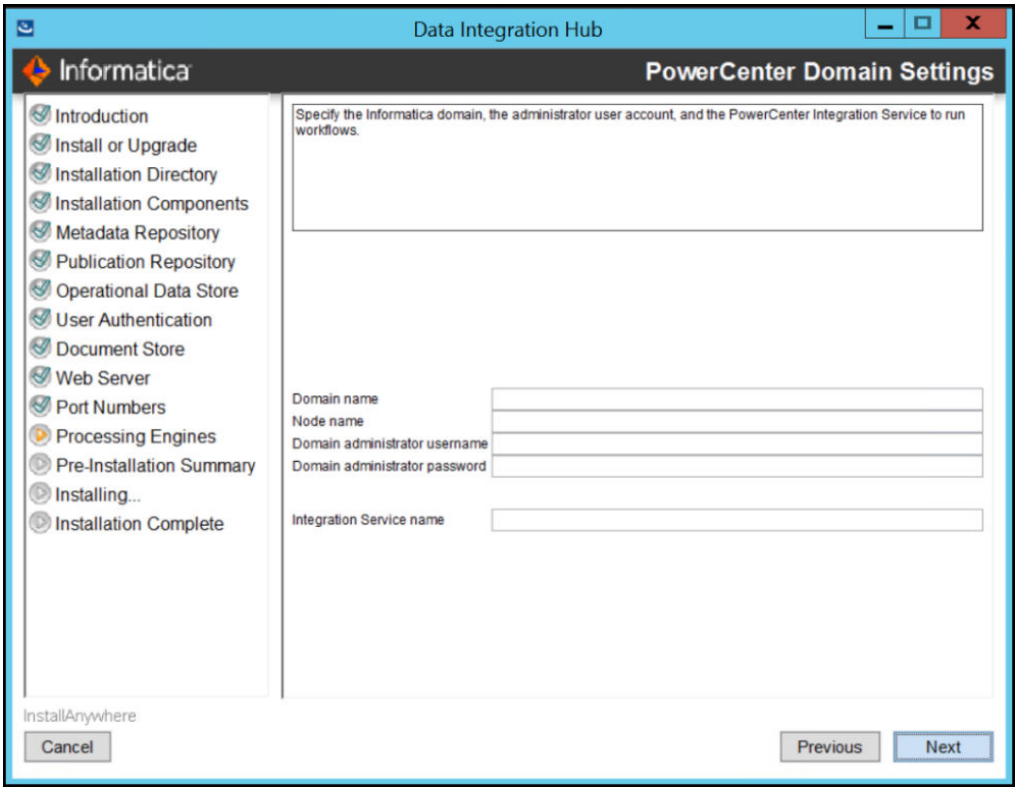
   Security domain

   Optional. Name of the Informatica security domain in which the PowerCenter Repository Service user is stored.
   Default is Native.

2. Click **Next**.

   If you selected to install the Data Integration Hub server plug-in for PowerCenter component, the **PowerCenter Domain Settings** page appears.

If you did not select the PowerCenter server component, the **PowerCenter pmrep Command Line Utility Location** page appears. Go to step 5.



3. Enter values in the following fields:

Domain name

Name of the Informatica domain that contains the PowerCenter Integration Service that runs Data Integration Hub workflows.

Node name

Node in the Informatica domain on which the PowerCenter Integration Service runs.

Domain administrator username

Name of the Informatica domain administrator.

Domain administrator password

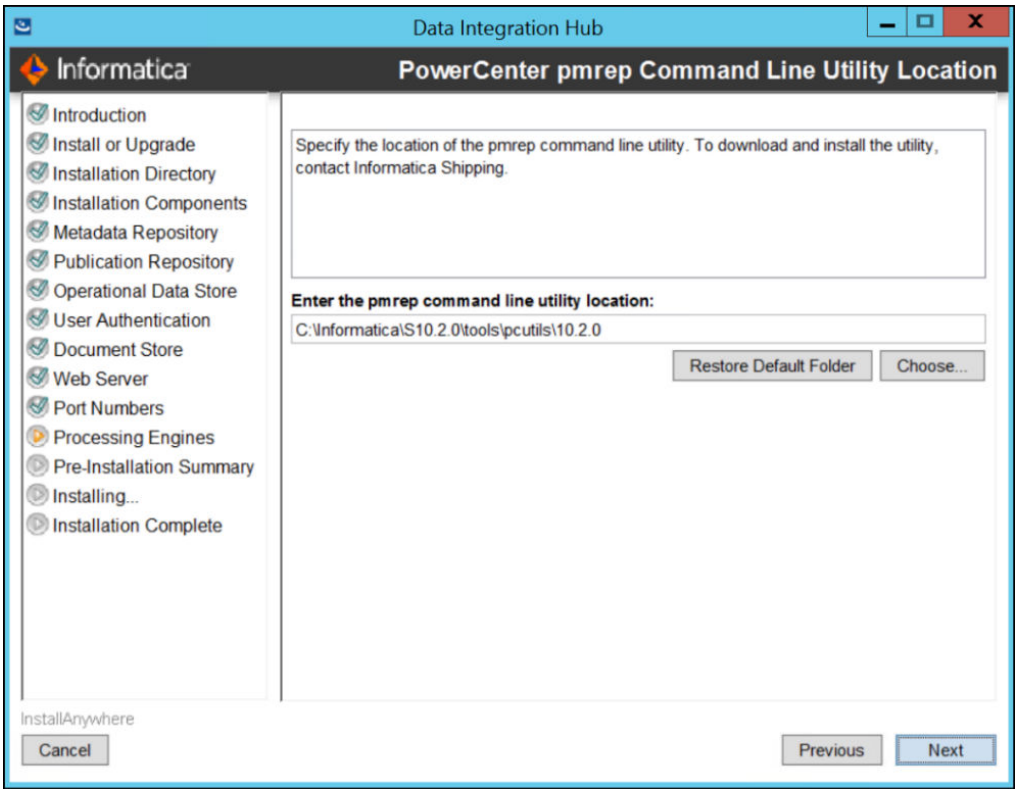
Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.

Integration Service name

The name of the PowerCenter Integration Service that Data Integration Hub uses to run workflows.

4. Click **Next**.

The **PowerCenter pmrep Command Line Utility Location** page appears.



5. Specify the location of the pmrep command line utility.

   The location of the utility depends on whether or not you install Data Integration Hub on the machine where the PowerCenter services are installed.

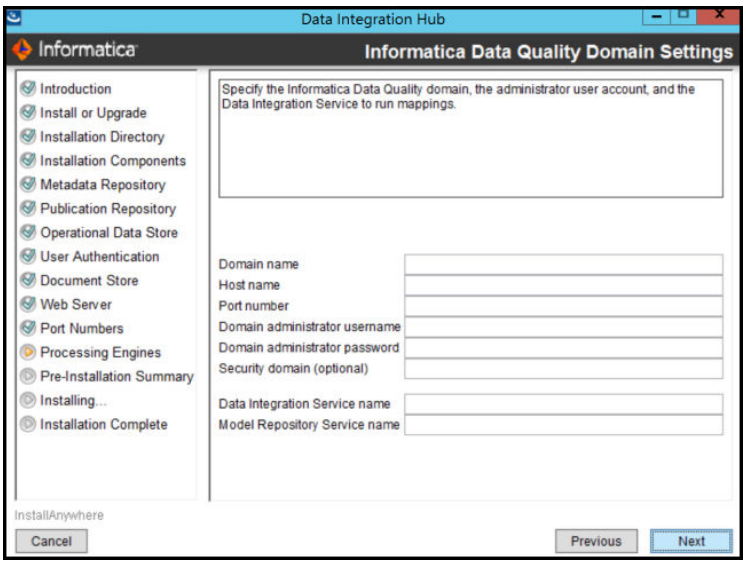
Environment	Location of the pmrep command line utility
Data Integration Hub installed on the machine where the PowerCenter services are installed	`<PowerCenter_services_installation_folder>\<PowerCenter_version>\tools\pcutils\<PowerCenter_version>`
Data Integration Hub and PowerCenter services installed on different machines	`<PowerCenter_client_installation_folder>\<PowerCenter_version>\clients\PowerCenterClient\client\bin`

6. Click **Next**.

   If you selected to install the Data Integration Hub Informatica Data Quality component, the **Informatica Data Quality Domain Settings** page appears. If you did not select to install the Data Integration Hub Informatica Data Quality component, the **Pre-Installation Summary** page appears.

Step 7. Configure Data Quality Settings

1. If you selected to install the Data Integration Hub Informatica Data Quality component, the **Informatica Data Quality Domain Settings** page appears.



Enter values in the following fields:

Domain name

Domain name of the node on which the Data Integration Service runs.

Host name

Host name of the node on which the Data Integration Service runs.

Port number

Port number of the node on which the Data Integration Service runs.

Domain administrator username

User name for the node on which the Data Integration Service runs.

Domain administrator password

Password for the node on which the Data Integration Service runs.

Security Domain

Optional. Name of the Informatica security domain in which the Model Repository Service user is stored.

Data Integration Service name

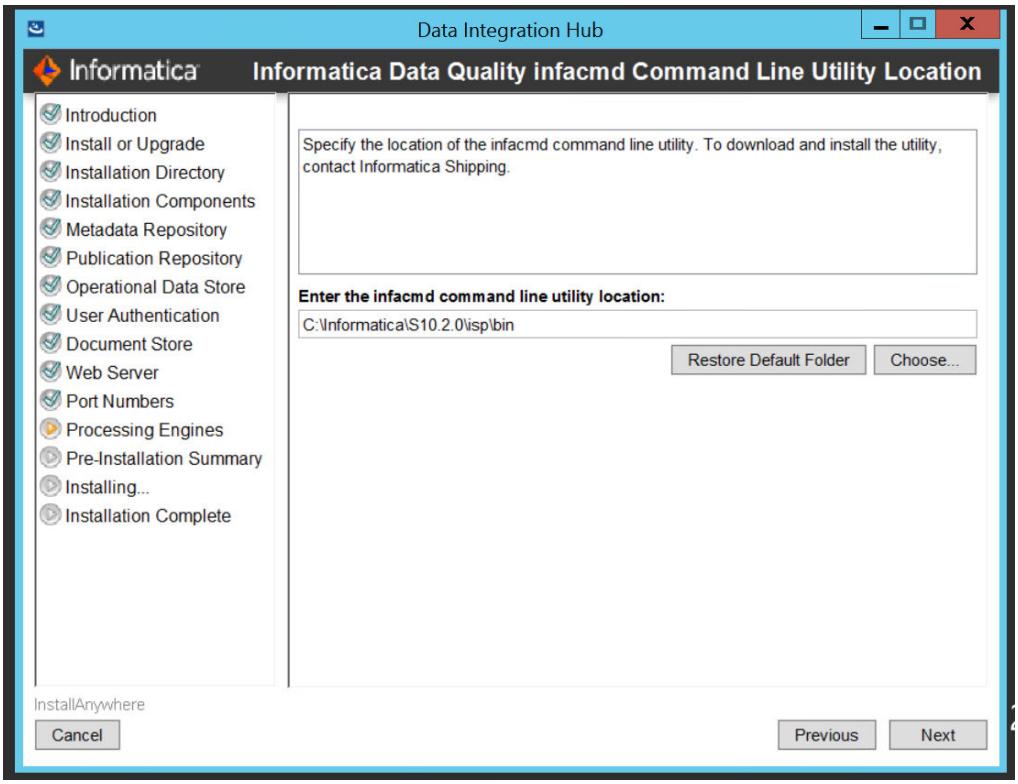
Name of the Data Integration Service.

Model Repository Service name

Name of the Model Repository Service.

2. Click **Next**.

The **Informatica Data Quality infacmd Command Line Utility** page appears.



3. Enter the location of the infacmd command line utility or accept the default directory, and then click **Next**.

   The **Pre-Installation Summary** page appears.

Step 8. Complete the Installation

1. On the **Pre-Installation Summary** page, verify that the installation information is correct, and then click **Install**.

   During the installation process, the installer displays progress information. When the installation process ends, the **Post-Installation Actions** page appears.

2. If you installed the Data Integration Hub PowerCenter server plug-in, follow the wizard instructions to register the plug-in to the PowerCenter repository, and then click **Next**.

   The **Installation Complete** page appears.

3. Click **Done** to close the installer.

4. To view the log files that the installer generates, navigate to the following directory: `<DIHInstallationDir>\logs`.

5. Perform the required post-installation tasks.

   For more information, see Chapter 5, “Post-Installation Tasks” on page 68.

   Note: Perform only the tasks that are relevant for your environment.

6. Optionally, perform additional configuration tasks. For more information, see Chapter 8, “Optional Data Integration Hub Configuration” on page 112.

Upgrading Data Integration Hub on a UNIX Operating System

Upgrade Data Integration Hub on UNIX operating systems in console mode. On Windows operating systems, upgrade Data Integration Hub in graphical mode.

Before you install, verify that your environment meets the minimum system requirements, perform the pre-installation tasks, and verify that the PowerCenter services are running.

Note: During the upgrade, Data Integration Hub saves log files in the home directory of the user, in the subdirectory named `DXLogs`. If the upgrade does not complete successfully, you can view the log files in this location.

Step 1. Run the Installer

1. Log in to the machine with the user account that you want to use to install Data Integration Hub.

   To prevent permission errors, use the same account to install Data Integration Hub and PowerCenter.
2. Close all other applications.
3. Run `Install.bin -i console` from the directory where you downloaded the installer.

   The **Introduction** section appears.
4. Read the instructions, and then press **Enter**.

   The **Install or Upgrade** section appears.
5. Enter **2** to upgrade Data Integration Hub, and then press **Enter**.

   The **Installation Directory** section appears.

Step 2. Define Installation Settings

1. In the **Installation Directory** section, enter the absolute path to the installation directory or accept the default directory, and then press **Enter**.

   The **Installation Components** section appears and displays a numbered list of the components to install.
2. Enter a comma-separated list of numbers for the components to install or accept the default components:

   1- Data Integration Hub

   Installs the core Data Integration Hub application.
   Selected by default.

   2- Data Integration Hub Dashboard and Reports

   Installs the Data Integration Hub Dashboard and Reports component. You must install Data Integration Hub to install the Dashboard and Reports component. Select this component to view the reports in the Dashboard using operational data store.
   Cleared by default.

   3- Data Integration Hub Server Plug-in for PowerCenter

   Installs the Data Integration Hub PowerCenter server plug-in component. After the installation, register the plug-in to the PowerCenter repository.
   Selected by default.

4- Data Integration Hub Hadoop Service

Connecting module between Data Integration Hub and the Hadoop cluster. Enables Data Integration Hub to perform operations on the Hadoop publication repository.
Select this component if you want to define a Hadoop-based publication repository and manage some of your topics on Hadoop.

Note: The Data Integration Hub Hadoop Service must be installed on an edge node in the Hadoop cluster. Unless you are installing Data Integration Hub on a Hadoop edge node, do not select this component during Data Integration Hub installation. In this case, after you complete the Data Integration Hub installation, run the installation again on the edge node and select to install only the Data Integration Hub Hadoop Service component on the node.

5- Data Integration Hub Data Engineering Integration/Informatica Data Quality

Enables Data Integration Hub to use Data Engineering Integration and Data Quality mappings for custom publications and subscriptions.
Select this component if you want to create custom publications and subscriptions that use Data Engineering Integration or Data Quality mappings in Data Integration Hub.

3. Press **Enter**.

   The **Select PowerCenter Version** section appears.

4. Select the PowerCenter version for which to install Data Integration Hub or accept the default selection:

   1- PowerCenter version below 10.5.

   Select this option for PowerCenter versions below 10.5.

   2 - PowerCenter version 10.5.

   Select this option for PowerCenter version 10.5.

5. Press **Enter**.

   The **Metadata Repository** section appears.

Step 3. Configure Data Integration Hub Repositories

1. In the **Metadata Repository** section, enter the number for the metadata repository database configuration option or accept the default option:

   1- Create a Data Integration Hub repository

   Creates a repository in the database.

   2- Use an existing Data Integration Hub repository

   Uses the tables and data in an existing repository.

2. Press **Enter**.

   The **Metadata Repository Connection** section appears.

3. Enter one of the following numericals depending on the database you plan to use as the Data Integration Hub metadata repository database:

   - Enter `1` to use an Oracle database as the Data Integration Hub metadata repository database.
   - Enter `2` to use Microsoft SQL Server database as the Data Integration Hub metadata repository database.

4. Enter the number for the metadata repository database connection type or accept the default connection type:

1- Database URL

Location of the database. If you select this option, enter values in the following fields:

- **Database Host Name**. Host name of the machine where the database server is installed.
- **Database Port Number**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server and Microsoft Azure SQL Database is 1433.
- **Database SID**. System identifier for the database.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

2- Custom Connection String

Connection string to the database. If you select this option, enter values in one of the following fields:

- JDBC string. JDBC connection string to the metadata repository.
- ODBC string. ODBC connection string to the metadata repository. Applicable if you install the PowerCenter client plug-in. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

5. Enter values in the following fields:

Database username

Name of the database user account.

Database user password

The password for the database account for the database. Data Integration Hub stores the password as an encrypted string.

6. Press **Enter**.
7. Enter the number for the publication repository connection type or accept the default connection type.

Note: When using Microsoft SQL Server named instances, you must define a custom connection string.

1- Database URL

Location of the database. If you select this option, enter values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port number.** Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database.
- **Oracle database**. Database SID. System identifier for the database.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

2- Custom Connection String

Connection string to the database. If you select this option, enter values in one of the following fields:

- **JDBC string**. JDBC connection string to the publication repository.

- **ODBC string**. Applicable if you install the PowerCenter client plug-in. ODBC connection string to the publication repository. The installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

8. Press **Enter**.
9. Enter values in the following fields:

Database username

Name of the database user account for the database.

Database user password

The password for the database account. Data Integration Hub stores the password as an encrypted string.

10. Press **Enter**.

If you selected to install the Data Integration Hub Dashboard and Reports component, the **Operational Data Store** section appears. If you did not select to install the Dashboard and Reports component, go to "Step 5. Configure the Web Server " on page 102.

Step 4. Set Up the Operational Data Store

1. In the **Operational Data Store** section, enter the number for the database configuration option for the operational data store or accept the default option:

1- Create an operational data store repository

Creates an operational data store repository in the database.

2- Use an existing operational data store repository

Uses the tables and data in an existing operational data store repository.

2. Enter the number for the database connection type for the operational data store or accept the default connection type:

1- Database URL

Location of the database. If you select this option, enter values in the following fields:

- **Database host name**. Host name of the machine where the database server is installed.
- **Database port number**. Port number for the database. The default port number for Oracle is 1521. The default port for Microsoft SQL Server or Microsoft Azure SQL Database is 1433.
- **Oracle database**. Database SID. System identifier for the database.
- **Microsoft SQL Server database** or **Microsoft Azure SQL Database**. Database name. Name of the database instance.

2- Custom Connection String

Connection string to the database. If you select this option, enter values in one of the following fields:

- **JDBC string.** JDBC connection string to the Operational Data Store.

- **ODBC string**. ODBC connection string to the Operational Data Store. If you install the PowerCenter client plug-in, the installer cannot verify the validity of the ODBC string.

Note: If you use a named Microsoft SQL Server database instance, you cannot connect to the database instance using the **Database URL** option. Use the **Custom Connection String** option.

For example:

```
jdbc:informatica:sqlserver://MYSQLSERVERCOMPUTERHOSTNAME
\MYDBINSTANCENAME;DatabaseName=MYDATABASENAME;
```

3. Enter values for the operational data store in the following fields:

Database username

Name of the database user account for the database.

Database user password

The password for the database account for the database. Data Integration Hub stores the password as an encrypted string.

4. Press **Enter**.

The **Web Server** section appears.

Step 5. Configure the Web Server

1. Configure the Web Server connection.
   a. Enter the number for the network communication protocol or accept the default protocol:

1- Enable HTTPS

Instructs Data Integration Hub to use secure network communication when you open the Operation Console in the browser.
If you select HTTPS and HTTP, the Operation Console switches existing HTTP connections with HTTPS connections.

2- Enable HTTP

Instructs Data Integration Hub to use regular HTTP network communication when you open the Operation Console in the browser.

   b. If you selected **Enable HTTPS**, enter values in the following fields:

Connector port number

Port number for the Tomcat connector to use when you open the Operation Console with HTTPS.
The default value is 18443.

Use a keystore file generated by the installer

Instructs the installer to generate a keystore file with an unregistered certificate. If you select this option, ignore the security warning that you receive from the browser the first time you open the Operation Console.

Use an existing keystore file

Instructs the installer to load an existing keystore file. Enter values in the following fields:

- Keystore password. Password for the keystore file.

- Keystore file. Path to the keystore file.

  The keystore file must be in the Public Key Cryptography Standard (PKCS) #12 format.

c. If you selected **Enable HTTP**, enter values in the following fields:

HTTP connector port number

Port number for the HTTP connector. If you clear this field, your browser must connect to the Data Integration Hub server with HTTPS when you log in to the Operation Console.
The default value is 18080.

Server shutdown listener port number

Port number for the listener that controls the Tomcat server shutdown.
The default value is 18005.

2. Press **Enter**.

   If you selected to install the Data Integration Hub PowerCenter server plug-in or the Data Integration Hub PowerCenter Client plug-in components, the **PowerCenter Location** section appears. If you did not select the PowerCenter server or client components, the **PowerCenter Web Services Hub** section appears.

Step 6. Configure PowerCenter Settings

1. If you selected to install the Data Integration Hub PowerCenter server plug-in or the Data Integration Hub PowerCenter Client plug-in components, in the **PowerCenter Location** section, enter the directory where you installed PowerCenter or accept the default directory, and then press **Enter**.

   The **PowerCenter Web Services** section appears.

2. In the **PowerCenter Web Services** section, press **Enter** to accept the default URL or enter the URL that the PowerCenter Web Services Hub uses to process publication and subscription workflows and then press **Enter**.
3. Enter the name of the PowerCenter Repository Service, and then press **Enter**.
4. Enter values in the following fields:

Node host name

Host name of the node that runs the PowerCenter Repository Service.

Node port number

Port number of the node that runs the PowerCenter Repository Service.

Username

Name of the PowerCenter Repository Service user.

Password

Password for the PowerCenter Repository Service user. Data Integration Hub stores the password as an encrypted string.

Security domain

Optional. Name of the Informatica security domain in which the PowerCenter Repository Service user is stored.
Default is Native.

5. Press **Enter**.

   If you selected to install the Data Integration Hub server plug-in for PowerCenter component, the **PowerCenter Domain Settings** section appears. If you did not select the PowerCenter server component, the **PowerCenter pmrep Command Line Utility Location** section appears. Go to step 9.

6. Enter values in the following fields:

   Domain name

   Name of the Informatica domain that contains the PowerCenter Integration Service that runs Data Integration Hub workflows.

   Node name

   Node in the Informatica domain on which the PowerCenter Integration Service runs.

   Domain administrator user name

   Name of the Informatica domain administrator.

   Domain administrator password

   Password for the Informatica domain administrator. Data Integration Hub stores the password as an encrypted string.

7. Press **Enter**.
8. Enter the name of the PowerCenter Integration Service that Data Integration Hub uses to run workflows, and then press **Enter**.
9. Enter the location of the pmrep command line utility and then press **Enter**. The location of the utility depends on whether or not you install Data Integration Hub on the machine where the PowerCenter services are installed.

   Note: On Linux operating systems, pmrep must be executable.

Environment	Location of the pmrep command line utility
Data Integration Hub installed on the machine where the PowerCenter services are installed	`<PowerCenter_services_installation_folder>/<PowerCenter_version>/tools/pcutils/<PowerCenter_version>`
Data Integration Hub and PowerCenter services installed on different machines	`<PowerCenter_client_utility_directory>/PowerCenter/server/bin`

10. Press **Enter**.

    If you selected to install the Data Integration Hub Data Engineering Integration/Informatica Data Quality component, the **Data Engineering Integration/Informatica Data Quality Domain name** section appears. If you did not select to install the Data Integration Hub Data Engineering Integration/Informatica Data Quality component, the **Pre-Installation Summary** section appears.

Step 7. Configure Processing Engine Settings

1. In the **Data Engineering Integration/Informatica Data Quality Domain name** section, press **Enter** to accept the default domain name or enter the name of the Informatica domain, and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Host name** section appears.

2. Press **Enter** to accept the default host name or enter the host name of the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Port number** section appears.

3. Press **Enter** to accept the default port number or enter the port number of the node on which the Data Integration Service runs and then press **Enter**.

The **Data Engineering Integration/Informatica Data Quality administrator user name** section appears.

4. Press **Enter** to accept the default user name or enter the user name for the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality administrator password** section appears.

5. Press **Enter** to accept the default password or enter the password for the node on which the Data Integration Service runs and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Security Domain** section appears.

6. Optionally, enter the name of the Informatica security domain in which the Model Repository Service user is stored. Press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Data Integration Service name** section appears.

7. Press **Enter** to accept the default name or enter the name of the Data Integration Service and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality Model Repository Service name** section appears.

8. Press **Enter** to accept the default name or enter the name of the Model Repository Service and then press **Enter**.

   The **Data Engineering Integration/Informatica Data Quality infacmd Command Line Utility Location** section appears.

9. Press **Enter** to accept the default location or enter the location of the infacmd command line utility then press **Enter**.

   The **Pre-Installation Summary** section appears.

Step 8. Complete the Installation

1. In the **Pre-Installation Summary** section, verify that the installation information is correct, and then press **Enter**.

   During the installation process, the installer displays progress information.

2. If you installed the Data Integration Hub PowerCenter server plug-in, follow the on-screen instructions to register the plug-in to the PowerCenter repository, and then press **Enter**.

3. To view the log files that the installer generates, navigate to the following directory:
   `<DIHInstallationDir>\logs`

4. Perform the required post-installation tasks.

   For more information, see Chapter 5, “Post-Installation Tasks” on page 68.

   Note: Perform only the tasks that are relevant for your environment.

5. Optionally, perform additional configuration tasks. For more information, see Chapter 8, “Optional Data Integration Hub Configuration” on page 112.

After You Upgrade

To complete the upgrade process, perform the following steps:

1. Reapply modifications that were made to Data Integration Hub configuration files in previous versions.

2. If you installed the Dashboard and Reports component, register the license of the Dashboard and Reports component.
3. If you installed the Dashboard and Reports component, import the operational data store (ODS) loader workflow.
4. If any of the Data Integration Hub repositories are installed on a Microsoft SQL Server and use Windows authentication, configure credentials for Windows authentication.
5. Update the security configuration files.
6. If any of the Data Integration Hub repositories are installed on a Microsoft SQL Server and do not have partitions, delete all Data Integration Hub partition functions. For example: `DIH__DATE_PARTITION_FUNCTION`.
7. If you are using the Data Integration Hub Run Publication Subscription Web Service API, update the API.
8. If you installed the Data Integration Hub Hadoop Service component for the first time, configure your environment for a Hadoop-based publication repository. For more information, see "Configuring the Environment for a Hadoop Publication Repository" on page 72.
9. Start the Data Integration Hub services. For more information, see Chapter 7, "Starting and Stopping Data Integration Hub" on page 110.

   Note: If you installed the Data Integration Hub Hadoop Service component on a different machine than the machine where you installed Data Integration Hub, start the services on both machines. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111.
10. If you installed the Data Integration Hub Hadoop Service component for the first time, log in to the Operation Console and configure the DIH__STAGING__HADOOP connection. For more information, see "Configure Connections to the Data Integration Hub Repositories" on page 75.
11. Clear the browser cache on each of the client machines.
12. If you use real-time publication workflows, enable the property **Enable high precision** in the session that writes data to the Data Integration Hub publication repository in all the real-time publication workflows.
13. If you use custom workflows, update the Data Integration Hub publication repository in all the real-time publication workflows.
14. If the `B2B_DX_InstallLog.txt` log file lists invalidated subscriptions, validate them. To validate the invalidated subscriptions navigate through the subscription wizard of the corresponding subscription and click **Finish** on the summary page.
    For more information about finding the invalidated subscriptions and validating them, see "Validating Invalidated Subscriptions" on page 109.

Note: The Data Integration Hub installer does not delete the previous version of Data Integration Hub. The installer renames the folder with the suffix `_Backupn.n.n` where n.n.n is the version number that you upgraded. To ensure that you update the configuration files correctly, see the configuration files in the directory of the previous version of Data Integration Hub.

Related Topics:

- "Importing the Web Service to PowerCenter" on page 70

Reapplying Configuration Modifications

Reapply modifications that were made to Data Integration Hub configuration files in previous versions.

To perform this procedure, you must have backed up the Data Integration Hub installation folder.

1. Open the following file from the location where you backed up the Data Integration Hub installation folder:

   `<BackupDir>/conf/dx-configuration.properties`
2. On the machine where Data Integration Hub is installed, open the server and console copies of the `dx-configuration.properties` files in a text editor from the following locations:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\shared\classes\`

   `<DIHInstallationDir>\conf\`
3. Copy any relevant configuration changes from the file that you backed up to both the `dx-configuration.properties` files.
4. Save the `dx-configuration.properties` files.

Registering the Dashboard and Reports License

Register the license of the Data Integration Hub Dashboard and Reports component.

1. Start the Data Integration Hub services.
2. Move the file `_Settings.lgx` from the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\webapps\dih-dashboard\_Definitions`

   To the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\shared\classes`

   Rename the file to the following name:

   `dx_dashboard_configuration.xml`
3. Reapply modifications that were made to the file `dx_dashboard_configuration.xml` in previous versions.
4. Copy the Logi Info Dashboard license file `_Settings_encrypted.lgx` to the following location:

   `<DIHInstallationDir>\DataIntegrationHub\tomcat\webapps\dih-dashboard\_Definitions`
5. Rename the file `_Settings_encrypted.lgx` to `_Settings.lgx`.
6. Restart the Data Integration Hub services.

Replacing the Operational Data Store Loader Workflow

If you installed the Dashboard and Reports component for the first time, import the operational data store (ODS) loader workflow. If the Dashboard and Reports component was installed in the previous version of Data Integration Hub and you are upgrading to the current version from a version earlier than 9.6.2, replace the existing workflow.

Note: If the Dashboard and Reports component was installed in the previous version of Data Integration Hub, you must have unscheduled the Data Integration Hub ODS loader workflow before you upgraded Data

Integration Hub to the current version. For more information, see GUID-8BD86874-0566-47F6-99C8-16427FA8E887.

1. In PowerCenter Repository Manager, run the **Import Wizard**.
2. Select the Data Integration Hub ODS loader workflow file. The name of the workflow file depends on the type of database on which the ODS is installed:

Database Type	Workflow Location and Name
Oracle	<DIHInstallationDir>\powercenter\ETL\DX_ETL.xml
Microsoft SQL Server	<DIHInstallationDir>\powercenter\ETL\DX_ETL_SQLSERVER.xml

3. If the Dashboard and Reports component was installed in the previous version of Data Integration Hub and you are upgrading to the current version from version 9.6.1, select the PowerCenter repository folder that contains the Data Integration Hub ODS loader workflow from the previous version as the import folder target.
4. If the Dashboard and Reports component was installed in the previous version of Data Integration Hub, in the **Conflict Resolution Wizard**, select **Replace**. In the **Apply this resolution to** list, select **All Conflicts**. Click **Next**.
5. In the Global Copy Options area select the options **Retain Sequence Generator, Normalizer, or XML key current values** and **Retain Persistent Mapping Variable Values**.
6. Follow the instructions in the **Import Wizard** to complete the import of the workflow.

Configure Credentials for Windows Authentication

If you installed any of the Data Integration Hub repositories on a Microsoft SQL Server and you selected to use Windows authentication, configure the credentials that Data Integration Hub uses to access the Microsoft SQL Server instance.

Before you start the configuration process, verify that all Data Integration Hub Windows services are stopped and that the Data Integration Hub Operation Console and the Data Integration Hub server are not running.

1. Access the Windows **Services** window.
2. Double-click the service **Informatica Data Integration Hub Server *version***.

   The Data Integration Hub Server Properties window appears.
3. Select the **Log On** tab.
4. Select **This account**, click **Browse**, and then specify a user account in the **Select User** dialog box. When you are finished, click **OK**.
5. Type the password for the user account in **Password** and in **Confirm password**, and then click **OK.**
6. Repeat steps 2 through 5 for the service **Informatica Data Integration Hub Console *version***.

Updating the Security Configuration Files

To perform this procedure, you must have backed up the Data Integration Hub installation folder.

1. Open the following folder from the location where you backed up the Data Integration Hub installation folder:

   ```
   <BackupDir>\shared\conf\security\
   ```

2. Copy all the files from the backup `security` folder to the following folder:

   `<DIHInstallationDir>\shared\conf\security\`

   Overwrite existing files.

Updating the Run Publication Subscription Web Service API

If you are using the Data Integration Hub Run Publication Subscription Web Service API, update the API before you start the Data Integration Hub services.

1. Import the `runpublicationsubscription` Data Integration Hub web service workflow to PowerCenter.
2. Update your client application to use the new web service.

Restart the Data Integration Hub Services

Restart the Data Integration Hub Server and the Data Integration Hub Console services.

Validating Invalidated Subscriptions

During the upgrade, Data Integration Hub invalidates automatic relational database subscriptions that have the following configurations:

- Data Integration Hub system fields are mapped in the field mapping page of the subscription wizard.
- Subscription delivery strategy is to insert new rows and update rows that exist in the target.

You can view the list of invalidated subscriptions in the following log file: `<Data Integration Hub installation directory>/logs/B2B_DX_InstallLog.txt`.

The following sample shows the list of invalidated subscriptions in the log file:

```
The invalidated subscriptions with delivery option 'Insert new rows and update changed rows'
and system fields mapped are: [employeedata, coursedescriptions, departmentdetails]
```

To validate an invalidated subscription, perform the following steps:

1. Access the subscription wizard of the invalidated subscription.
2. Click **Next** on each page to navigate to the next page.
3. Click **Finish** on the Summary page.

CHAPTER 7

Starting and Stopping Data Integration Hub

This chapter includes the following topics:

- Overview of Starting and Stopping Data Integration Hub, 110
- Starting and Stopping Data Integration Hub on Windows, 110
- Starting and Stopping Data Integration Hub on Linux, 111

Overview of Starting and Stopping Data Integration Hub

Stop or start the Data Integration Hub services.

For example, start the services after you install Data Integration Hub, or stop the services before you upgrade Data Integration Hub.

Starting and Stopping Data Integration Hub on Windows

Start and stop the Data Integration Hub services from the Start menu or run the startup and shutdown scripts.

The installer creates shortcuts in the Start menu to start and stop all Data Integration Hub services.

Starting and Stopping Data Integration Hub from the Start Menu

On Windows operating systems, you can use the Start menu to start and stop all Data Integration Hub services. You cannot start or stop a single service from the Start menu.

1. In the Start menu, click **Informatica > Data Integration Hub**.
2. Choose one of the following options:

- Start Services. Starts all Data Integration Hub services.
- Stop Services. Stops all Data Integration Hub services.
- Operation Console. Opens the Operation Console in a Web browser.

Starting and Stopping Data Integration Hub with Batch Scripts

On Windows operating systems, you can run scripts to start and stop one or more Data Integration Hub services.

1. Navigate to the following directory:

   ```
   <DIHInstallationDir>\bin
   ```

2. Choose the script to run.
   - startup.bat. Starts all Data Integration Hub services.
   - shutdown.bat. Stops all Data Integration Hub services.
   - Start each of these services separately in the listed order:
     - dihconsole.bat. Starts the Operation Console.
     - dihserver.bat. Starts the Data Integration Hub server.

Starting and Stopping Data Integration Hub on Linux

Run the scripts to start or stop the Data Integration Hub services. The installer creates shell scripts that you can use to start or stop all the Data Integration Hub services or to start each service separately. You cannot stop each service separately.

1. Navigate to the following directory:

   ```
   <DIHInstallationDir>/bin
   ```

2. Choose the script to run.
   - startup.sh. Starts all Data Integration Hub services.
   - shutdown.sh. Stops all Data Integration Hub services.
   - Start each of these services separately in the listed order:
     - dihconsole.sh. Starts the Operation Console.
     - dihserver.sh. Starts the Data Integration Hub server.

Note: If you installed the Data Integration Hub Hadoop Service component on a different machine than the machine where you installed Data Integration Hub, to stop or to start the Hadoop service on the machine where the service is installed, run the `startup.sh` or `shutdown.sh` script from the following location:

```
<DIHInstallationDir>/DataIntegrationHub/tomcat/bin/
```

CHAPTER 8

Optional Data Integration Hub Configuration

This chapter includes the following topics:

- Optional Data Integration Hub Configuration Overview, 112
- Modifying the Data Integration Hub Server RMI Port Number, 113
- Logs, 113
- Changing the Maximum Java Heap Size, 117
- Changing the Credentials for a Database User Account, 118
- Changing the Credentials for the Data Integration Hub Hadoop Service, 119
- Updating the Dashboard Configuration File, 120
- Configuring a PowerCenter Integration Service to Access Data Integration Hub, 121
- Configuring Repository Connections on PowerCenter Version 10, 122

Optional Data Integration Hub Configuration Overview

Optional configuration includes tasks that you might want to perform after you install or upgrade Data Integration Hub, or at a later date.

- The Data Integration Hub components send information through ports. You can change the default port numbers based on the requirements of your network environment.
- When different components process information or encounter errors, log files contain information that you can use to analyze and troubleshoot the installed components. You can change the location of the log files or define custom logs.
- To increase performance and reliability, you can change the maximum memory allocation for the embedded Data Integration Hub server broker, or the embedded Data Integration Hub console broker.
- If you change the database user credentials for the Data Integration Hub repositories or for the operational data store, you must update the Data Integration Hub configuration files. If you are running the Dashboard and Reports component, you must also update the relevant PowerCenter connections.
- If you use a Hadoop-based publication repository and you change the credentials for the user account of the Data Integration Hub Hadoop Service, you must update the Data Integration Hub configuration files.

- If you use the Dashboard and Reports component, and the IP addresses of the machine that hosts Data Integration Hub change any time after the installation, you must update the IP addresses in the dashboard configuration file.
- During the Data Integration Hub installation or upgrade, you define a PowerCenter Integration Service that Data Integration Hub uses to run workflows. If required, you can configure a different PowerCenter Integration Service to access Data Integration Hub.
- If you use the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, configure the repository connections in PowerCenter Workflow Manager.

Modifying the Data Integration Hub Server RMI Port Number

Replace the RMI port number in the dx-configuration.properties files and in the PowerCenter Integration Service.

1. On the machine where Data Integration Hub is installed, open the server and console copies of the dx-configuration.properties files in a text editor from the following locations:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/
   <DIHInstallationDir>/conf/
   ```

2. Enter the port number in the following property:

   ```
   dx.rmi.port=
   ```

3. Save the dx-configuration.properties files.
4. In the Administrator tool, select the PowerCenter Integration Service that runs Data Integration Hub transformations.
5. On the Processes tab of the PowerCenter Integration Service, add or edit the DX_SERVER_URL environment variable and set the URL of the Data Integration Hub server in the following format:

   ```
   rmi://<HostName>:<PortNumber>
   ```

6. Save the changes and restart the Data Integration Hub services.

Logs

The Data Integration Hub log files include information that you can use to analyze activity and troubleshoot.

You can configure the following logs:

- Debug logs
- RMI server logs
- Database debug logs
- Import logs

To send log messages to a different log file destination, you can create an SNMP appender to redirect the logs to a custom destination.

Default Log Files

Data Integration Hub creates log files that record diagnostic information regarding system and user operations. The installer also creates log files that record installation selections and configuration.

You can configure log settings in the log4j.xml file located in the Data Integration Hub configuration directory.

The following log files are available:

Server

The dxserver.log file is located in the following directory:

```
<DIHInstallationDir>/logs
```

You can change the log mode to debug to generate more messages while you troubleshoot server issues.

Operation Console

The log files are located in the following directory:

```
<DIHInstallationDir>/DataIntegrationHub/tomcat/logs
```

Installer

The log files are located in the following directory:

```
<DIHInstallationDir>/logs
```

Data Integration Hub Hadoop Service

When you use the Data Integration Hub Hadoop Service, the log files are located in the following directory:

```
<DIHInstallationDir>/DataIntegrationHub/tomcat/logs
```

Customizing the Destination for Log Messages

By default, the log4j logging utility sends log messages to files. You can configure the log4j utility to send log messages to a destination that is different from the default log files with the Simple Network Management Protocol (SNMP). The installer installs the file that the log4j utility requires to work with SNMP.

Complete the following tasks to change the destination:

1. Add an SNMP appender to the log4j properties file and set the logging level. Change the sample SNMP appender in the log4j.xml file to the appender that you want to use. You can add multiple appenders to the log4j.xml file that send different types of log messages to different SNMP outputs.
2. Configure an SNMP manager to listen for messages. For information about configuring the SNMP manager to handle log4j messages, see the documentation for your SNMP network management software.

For general information about the log4j utility, see the Apache Web site:
http://logging.apache.org/log4j/1.2/manual.html

SNMP Appender Parameters

The parameters of the SNMP appender in the log4j.xml file define the output destination and settings for log messages.

The following table describes the SNMP parameters that you can define for Data Integration Hub:

Parameter	Description
ManagementHost	IP address of the monitoring system host. Default is **127.0.0.1**
ManagementHostTrapListenPort	Port number of the monitoring system host. Default is **162**
LocalIPAddress	IP address of local SNMP embedded agent. You do not normally need to modify this value. Default is **127.0.0.1**
LocalTrapSendPort	Port number of the local SNMP embedded agent. Default is **161**
CommunityString	Name of the SNMP community. Default is **public**
GenericTrapType	Type of the trap. Set one of the following values: - 0=cold start - 1=warm start - 2=link down - 3=link up - 4=authentication failure - 5=egp neighbor loss - 6=enterprise specific Default is **6=enterprise specific**
ApplicationTrapOID	Identifier of the application object that sends the trap messages. You can set the value of this parameter to the name of the application object in Data Integration Hub. Default is **1.3.6.1.2.1.2.0.0.0.0**
EnterpriseOID	Identifier of the organization object sending the trap message. You can set this parameter to any value that identifies the message in Data Integration Hub. Default is **1.3.6.1.2.1.2.0**
ForwardStackTraceWithTrap	Determines whether to include the stack trace in the log message. Default is **False**

Parameter	Description
Threshold	Level of details to report. Set one of the following values: - FATAL - ERROR - WARN - INFO - DEBUG Threshold values that are lower than INFO or WARN might cause heavy network traffic. For fewer notifications, set the threshold value to FATAL. For a larger number of notifications, set the threshold value to WARN.
SysUpTime	Amount of time that the application is running. Set the value to 0 to calculate the system up time when a message is sent. Default is **0**

Adding an SNMP Appender to the log4j.xml File

Add an SNMP appender to the Data Integration Hub server or Operation Console copies of the log4j.xml file to customize the output destination for log messages.

1. Back up the log4j.xml file that you want to edit from one of the following locations:
   - Data Integration Hub server: `<DIHInstallationDir>/conf`
   - Operation Console: `<DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes`
2. Open the file in a text editor and search for the following text:

   ```
   SNMP_TRAP is a sample appender
   ```
3. To edit the sample appender with the actual values of your appender, remove the comment indicators from the SNMP_TRAP appender and edit the appender parameters and values based on your requirements.

   Note: You can also add an appender below the sample appender instead of editing the sample appender.
4. To set the formatting of the log messages, edit the layout element.

   The following example shows the layout element of the sample appender:

   ```
   <layout class="org.apache.log4j.PatternLayout">
      <param name="ConversionPattern" value="%d{ISO8601} %-5p [%c] {%t} %m%n"/>
   </layout>
   ```

   For information about the layout pattern options, see the description on the Apache Website: http://logging.apache.org/log4j/1.2/apidocs/org/apache/log4j/PatternLayout.html
5. To activate the appender, search for the following text:

   ```
   <root>
   ```
6. Add the appender name to the appender list.

   The following example shows the appender list after you add the appender name:

   ```
   <root>
      <priority value="INFO"/>
      <appender-ref ref="BROKER-LOG"/>
      <appender-ref ref="CONSOLE"/>
      <appender-ref ref="SNMP_TRAP"/>
   </root>
   ```
7. Save the log4j.xml file.

After you add the SNMP appender, configure your SNMP manager to listen for the log messages.

Changing the Maximum Java Heap Size

You can change the maximum memory allocation for the embedded Data Integration Hub server broker, or the embedded Data Integration Hub console broker.

Embedded Data Integration Hub server broker

To change the maximum Java heap size of the embedded Data Integration Hub server broker, open one of the following files:

Operating System	File Location
Microsoft Windows	<Data Integration Hub installation directory>\bin\setenv.bat
UNIX	<Data Integration Hub installation directory>/bin/setenv.sh

Change the maximum heap size in MB in the DX_SERVER_OPTS property. The default maximum heap size is 1024 MB.

Embedded Data Integration Hub client broker

To change the maximum Java heap size of the Data Integration Hub client, open one of the following files:

Operating System	File Location
Microsoft Windows	<Data Integration Hub installation directory>\bin\setenv.bat
UNIX	<Data Integration Hub installation directory>/bin/setenv.sh

Change the maximum heap size in MB in the CATALINA_OPTS property. The default minimum heap size is 128 MB and the default maximum heap size is 2048 MB.

Data Integration Hub server service

To change the maximum Java heap size of the Data Integration Hub server service on Microsoft Windows operating systems, open the following file:

```
<Data Integration Hub installation directory>\conf\wrapper.conf
```

Change the maximum heap size in MB in the wrapper.java.maxmemory property. The default maximum heap size is 1024 MB.

Data Integration Hub client service

To change the maximum Java heap size of the Data Integration Hub client service on Microsoft Windows operating systems, run the following command:

```
<Data Integration Hub installation directory>\DataIntegrationHub\tomcat\bin
\tomcat7.exe //US//InfaDIHOpConsole<i>version</i> --JvmMx <max_heap_size>
```

Enter the maximum heap size in MB with integers and without letters. The default maximum heap size is 2048 MB.

Changing the Credentials for a Database User Account

When you install Data Integration Hub, you specify a user name and a user password for the user accounts of the Data Integration Hub repository database and of the publication repository database. If you are running the Dashboard and Reports component, you also specify a user name and a user password for the operational data store user account.

Perform the following steps if you change the credentials for a database user account after you install Data Integration Hub. Perform only the steps that are relevant to the changes that you are making. If you are not running the Dashboard and Reports component, skip the steps that are only relevant to this component.

1. Stop the Data Integration Hub services and close the Operation Console.
2. Verify that the PowerCenter Integration Service is not running any Data Integration Hub workflows.
3. If you are running the Dashboard and Reports component, and you are changing credentials for the Data Integration Hub repository or for the operational data store user account, use the PowerCenter Workflow Manager to update the credentials in the following connections:
   - For the Data Integration Hub repository, update the DX_REPO connection.
   - For the operational data store, update the DX_ODS connection.
4. If you are changing a password, perform the following steps:
   a. Run the password encryption utility and enter the new password in the following syntax:
      - On Windows operating systems: `<DIHInstallationDir>\dx-tools\dxpasswd.bat -p <NewPassword>`
      - On UNIX operating systems: `<DIHInstallationDir>/dx-tools/dxpasswd.sh -p <NewPassword>`

      The password encryption utility encrypts the password and displays an encrypted string. For example, `-->ywo+o3cw8+O3iLdlhPprW2YA==<--`.
   b. Copy the encrypted string without the `--><--` indicators to the clipboard.
5. Open both copies of the `dx-configuration.properties` file from the following locations in a text editor:

   `<DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/`

   `<DIHInstallationDir>/conf/`
6. In both copies of the `dx-configuration.properties` file, perform the following steps:
   a. Search for the text that is relevant to the changes that you are making:
      - Data Integration Hub repository:

        `dx.jdbc.username=<CurrentUsername>`

        `dx.jdbc.password=<CurrentPassword>`
      - Publication repository:

        `dih.staging.jdbc.username=<CurrentUsername>`

        `dih.staging.jdbc.password=<CurrentPassword>`
      - Operational data store:

        `dx.dashboard.jdbc.username=<CurrentUsername>`

        `dx.dashboard.jdbc.password=<CurrentPassword>`
   b. Replace the relevant value with the new value. If you are replacing a password, enter the encrypted string.
   c. Save and close the files.

   Note: The content in both copies of the `dx-configuration.properties` file must be identical.

7. Start the Data Integration Hub Operation Console.
8. If you are changing the credentials for the operational data store user account, perform the following steps:
    a. In the Navigator click **Administration** > **System Properties**.
    b. Change the values of the `dx.dashboard.jdbc.username` and the `dx.dashboard.jdbc.password` properties to the new values. For the password, enter the encrypted string.
9. Verify that the PowerCenter Integration Service is running.
10. If you are changing the credentials for the Data Integration Hub repository or for the publication repository, perform the following steps:
    a. In the Navigator, click **Hub Management** > **Connections**.

       The **Connections** page appears.
    b. Click the **Test Connection** icon next to the connection to the repository for which you want to change the credentials.
        - DIH__REPO. Connection to the Data Integration Hub repository.
        - DIH__STAGING. Connection to the publication repository.

       Data Integration Hub tests the connection. The process might take a few moments.
    c. When the message "Connection successful" shows, click **OK** in the **Test Connection** dialog box.
11. Start the Data Integration Hub Server service.
12. If you are running the Dashboard and Reports component, perform the following steps to schedule the Data Integration Hub operational data store loader workflow:
    a. In PowerCenter Workflow Manager, expand the folder where the operational data store loader workflow is located and then right-click **DX_ETL**.
    b. Select **Schedule** from the menu and wait until PowerCenter Workflow Manager schedules and runs the workflow.

Changing the Credentials for the Data Integration Hub Hadoop Service

When you install Data Integration Hub, the installation assigns a user name and a user password for the user account of the Data Integration Hub Hadoop Service.

Perform the following steps if you change the credentials of the Data Integration Hub Hadoop Service at a later date.

1. Stop the Data Integration Hub services and close the Operation Console. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111
2. Verify that the PowerCenter Integration Service is not running any Data Integration Hub workflows.
3. If you are changing a password, perform the following steps:
    a. Run the password encryption utility and enter the new password in the following syntax:

```
<DIHInstallationDir>/dx-tools/dxpasswd.sh -p <NewPassword>
```

The password encryption utility encrypts the password and displays an encrypted string. For example, `-->ywo+o3cw8+O3iLdlhPprW2YA==<--`.

b. Copy the encrypted string without the `--><--` indicators to the clipboard.

4. Open both copies of the `dx-configuration.properties` file in a text editor from the following locations on the machine where Data Integration Hub is installed:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/
   <DIHInstallationDir>/conf/
   ```

5. If the Data Integration Hub Hadoop Service component is installed on a different machine than the machine where Data Integration Hub is installed, on the machine where the Data Integration Hub Hadoop Service is installed, open the `dx-configuration.properties` file in a text editor from the following location:

   ```
   <DIHInstallationDir>/conf/
   ```

6. In all the copies of the `dx-configuration.properties` file, perform the following steps:

   a. Search for the following text :

      ```
      internal.service.username=
      internal.service.password=
      ```

   b. Replace the relevant value with the new value. If you are replacing a password, enter the encrypted string.

   c. Save and close the files.

   Note: The content in all copies of the `dx-configuration.properties` file must be identical.

7. Start the Data Integration Hub Operation Console.
8. Verify that the PowerCenter Integration Service is running.
9. In the Navigator, click **Hub Management** > **Connections**.

   The **Connections** page appears.

10. Click the **Test Connection** icon next to the DIH__STAGING__HADOOP connection.

    Data Integration Hub tests the connection. The process might take a few moments.

11. When the message "Connection successful" shows, click **OK** in the **Test Connection** dialog box.
12. Start the Data Integration Hub Server services. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111

Updating the Dashboard Configuration File

If you use the Dashboard and Reports component, and the IP addresses of the machine that hosts Data Integration Hub change any time after the installation, you must update the IP addresses in the dashboard configuration file.

1. Open the dashboard configuration file from the following location:

   ```
   <Data Integration Hub installation directory>\DataIntegrationHub\tomcat\shared
   \classes\dx_dashboard_configuration.xml
   ```

2. In the Security section, in the `AuthenticationClientAddresses` attribute, update the IP addresses of all the Network Interface Cards that provide access to the machine that hosts Data Integration Hub, including both IPv4 and IPv6 addresses. For example:

```
<Security AuthenticationSource="SecureKey" SecurityEnabled="True"
AuthenticationClientAddresses="127.0.0.1,0:0:0:0:0:0:0:1,10.36.8.34,fe80:0:0:0:c1f:
167a:bc3c:8307%11,10.36.112.186,fe80:0:0:0:5c9a:af6e:
87b9:3c08%12,fe80:0:0:0:7c26:627:71d2:4784%14,fe80:0:0:0:0:5efe:a24:822%16,fe80:0:0:0
:0:5efe:a24:70ba%16,fe80:0:0:0:0:5efe:c0a8:7701%17,192.168.119.1,fe80:0:0:0:45cf:
7bcf:f673:4392%18,192.168.162.1,fe80:0:0:0:3516:cd0c:
6f8:df39%19,fe80:0:0:0:0:5efe:c0a8:a201%20" RestartSession="False"
CacheRights="Session" LogonFailPage="https://HBW180084:18443/dih-console/logout.jsp">
```

3. Restart the Data Integration Hub services.

Configuring a PowerCenter Integration Service to Access Data Integration Hub

During the Data Integration Hub installation or upgrade, you define a PowerCenter Integration Service that Data Integration Hub uses to run workflows. If required, you can configure a different PowerCenter Integration Service to access Data Integration Hub.

In the Java classpath for the PowerCenter Integration Service, add the path to the Data Integration Hub class files.

1. Log in to the Administrator tool and select the PowerCenter Integration Service that runs the workflows for Data Integration Hub.
2. On the **Processes** tab, edit the Java SDK ClassPath property and add the location of the Data Integration Hub Java classes at the beginning of the ClassPath property:

```
<Data Integration Hub installation directory>/powercenter/lib/dx-client-
powercenter-10.5.0.jar;
<Data Integration Hub installation directory>/powercenter/lib/commons-
logging-1.1.3.jar;
<Data Integration Hub installation directory>/powercenter/lib/log4j-1.2.17.jar;
<Data Integration Hub installation directory>/powercenter/lib/activemq-all-5.15.9.jar
```

   If the installation log file lists invalidated subscriptions, manually run them. You can find the installation log file in the following directory: <Data Integration Hub installation directory>/logs

   During the upgrade, Data Integration Hub invalidates the automatic relational database subscriptions that have the following configurations:•

   Subscription field mapping.•

   Subscription delivery strategy is to insert new rows and update rows that exist in the target.

   For more information about running the subscriptions manually, see the Data Integration Hub Operator

3. Add environment variables to the Data Integration Hub console and server integration services.

Integration Service	Environment Variable
DX_CONSOLE_URL	`rmi://<HostName>:<dx.tpm.rmi.port>`
DX_SERVER_URL	`rmi://<HostName>:<dx.rmi.port>`

You can find the RMI port numbers for the console and server in the following location:

```
<Data Integration Hub installation directory>\conf\dx-configuration.properties
```

By default:

- `dx.tpm.rmi.port`: 18096
- `dx.rmi.port`: 18095

4. Save the changes.

Configuring Repository Connections on PowerCenter Version 10

If you use the Dashboard and Reports component, your Data Integration Hub and operational data store repositories are installed on Microsoft SQL Servers, and you use PowerCenter version 10, configure the repository connections in PowerCenter Workflow Manager.

1. In the Workflow Manager, access the DX_REPO database connection and open the **Connection Object Definition** dialog box.
2. Perform the following actions and then click **OK**:
   a. Select **Use DNS**.
   b. In the **Connect String** text box enter the connection name. The name is defined in the ODBC Data Source Administrator interface, in ODBC SQL Server Wire Protocol Setup, in the **Data Source Name** field.
3. Repeat steps 1 and 2 for the DX_ODS connection.

CHAPTER 9

Installing and Configuring the Data Integration Hub Accelerator for Data Archive

This chapter includes the following topics:

- Installing and Configuring Data Integration Hub Accelerator for Data Archive Overview, 123
- Pre-Installation Steps, 123
- Installing the Data Integration Hub Accelerator for Data Archive, 124
- Configuring the Data Archive Source Connection, 125

Installing and Configuring Data Integration Hub Accelerator for Data Archive Overview

The Data Integration Hub accelerator is a plug-in for Data Archive that enables you to purge events that Data Integration Hub generates with Data Archive.

After you install Data Integration Hub Accelerator for Data Archive, configure the Data Archive source connection.

Pre-Installation Steps

Before you install the Data Integration Hub accelerator in Data Archive, make sure that your system meets the minimum requirements and follow the pre-installation steps.

1. Make sure you have an active installation of Data Archive. A limited Data Archive version is provided with Data Integration Hub.
2. Create and assign privileges to the Data Integration Hub repository users.

3. In the Data Archive installation directory, open the `conf.properties` file and set the value of the following properties:
   - Set the value of the `informia.useDbaViewsInSeamlessAccess` property to `false`.
   - Set the value of the `informia.proceduresToExecute.inArchiveFromHistory` property to `java://com.informatica.b2b.dx.ilm.MoveDXDocStoreDatabaseDAOImpl`.

   Note: Back up the `conf.properties` file before you modify the property.
4. Restart Data Archive.

Installing the Data Integration Hub Accelerator for Data Archive

Install the Data Integration Hub accelerator for Data Archive after you install Data Integration Hub and Data Archive.

Before you install the accelerator, install Data Archive and follow the pre-installation steps.
The minimum supported version of Data Archive is 6.1. If you have Data Archive version 6.1 installed, also install EBF 11801 and EBF 11672.

1. Log in to Data Archive with administrator privileges and click **Accelerators** > **Enterprise Data Manager**.

   The Enterprise Data Manager appears.
2. In the Enterprise Data Manager, click **File** > **Import** > **Accelerator**.

   The **Import Metadata Options** window appears.
3. Select **Continue Import through EDM** and click **OK**.
4. Navigate to the location `<DIHInstallationDir>/ILM-accelerator` and select to import all the `XML` files in the directory.

   Note: Do not import the sub-folders in the directory.

   The Enterprise Data Manager displays a progress window during the import process.
5. To verify the import process, restart the Enterprise Data Manager and make sure that the accelerator appears in the `Data Integration Hub` node of the Explorer pane.
6. To add the accelerator to drop-down lists in Data Archive, log in to the database with the Data Integration Hub repository credentials and run the SQL script on the Data Archive repository. The script is located in one of the following locations:

Database	Path
Oracle	<DIHInstallationDir>/ILM-accelerator/sql/oracle_ilm_repository_update.sql
Microsoft SQL Server	<DIHInstallationDir>ILM-accelerator/sql/sqlserver_ilm_repository_update.sql

Configuring the Data Archive Source Connection

Configure the Data Archive source connection before you purge events from the database.

For information about creating source connections, see the *Data Archive Administrator Guide*.

1. Log in to Accelerator.
2. Click **Administration** > **New Source Connection**.
3. Enter the database-specific connection properties.
4. Set the application version of the connection to **Data Integration Hub 10.5.0**.
5. Set the property **Source / Staging Attachment Location** to the root path for the production document store.

   The path is the same as the path defined for the Data Integration Hub system property `dx.system.document.store.folder`.
6. If the connection is to an SQL Server database, select the option **Compile ILM Functions**.

Note: The location specified in the **Source / Staging Attachment Location** property must be accessible to the Data Archive purge process.

CHAPTER 10

Troubleshooting the Data Integration Hub Installation

This chapter includes the following topics:

- Troubleshooting the Data Integration Hub Installation Overview, 126
- Troubleshooting the Data Integration Hub Hadoop Service, 126
- Installation Log Files, 129

Troubleshooting the Data Integration Hub Installation Overview

The installer can generate errors when the installer creates and configures Data Integration Hub. Use the installation log files to troubleshoot the installation and upgrade.

If you install the Data Integration Hub Hadoop Service, troubleshoot the installation of the service.

Troubleshooting the Data Integration Hub Hadoop Service

Consider the following troubleshooting tips when you install the Data Integration Hub Hadoop Service:

When I test the DIH__STAGING__HADOOP connection the test succeeds but when I try to run a publication the running fails.

The following example shows a sample error message:

```
Wrong FS: hdfs://10.40.40.96:8020/user/infa_user/TEST_FILTER_87003, expected: hdfs://
tavdxcdh53n1:8020
```

In the Data Integration Hub Operation Console, in the Connections page, edit the DIH__STAGING__HADOOP connection so that **NameNode URI** is identical to the setting of the property fs.default.name or the property fs.defaultFS in the `core-site.xml` file.

When I test the DIH__STAGING__HADOOP connection the test fails with the following error: `Data Access Connection test failed with the following error:DXServerException: Cannot establish connection to Apache Spark.`

Perform the following actions:

1. Check the Data Integration Hub Hadoop Service log file for additional details. For more information, see "Default Log Files" on page 114.
2. Verify that Apache Spark is running.
3. On the machine where the Data Integration Hub Hadoop Service is installed, open the `dx-configuration.properties` file in a text editor from the following location:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/
   ```

   - Verify that the value of dih.hadoop.service.spark.url is correct. For more information, see "Configuring the Environment for a Hadoop Publication Repository" on page 72.
   - Verify that the value of dih.hadoop.service.spark.version is correct.

When I test the DIH__STAGING__HADOOP connection the test fails with the following error: `Data Access Connection test failed with the following error:DXServerException: Exception : org/apache/hadoop/fs/FileSystem.`

The definition of the classpath of the Data Integration Hub Hadoop Service is incorrect.

Perform the following actions:

1. On the machine where the Data Integration Hub Hadoop Service is installed, open the `dih-hadoop-service.xml` file in a text editor from the following location:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/conf/Catalina/localhost
   ```

2. Configure the correct locations of the `JAR` files.

When I test the DIH__STAGING__HADOOP connection the test fails with the following error: `Data Access Connection test failed with the following error:DXServerException: Cannot establish connection to Apache Hive.`

Perform the following actions:

1. Check the Data Integration Hub Hadoop Service log file for additional details. For more information, see "Default Log Files" on page 114.
2. Verify that Apache Hive is running.
3. On the machine where the Data Integration Hub Hadoop Service is installed, open the `dx-configuration.properties` file in a text editor from the following location:

   ```
   <DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/
   ```

   Verify that the value of dih.hadoop.service.hive.url is correct. For more information, see "Configuring the Environment for a Hadoop Publication Repository" on page 72.
4. If the message java.lang.ClassNotFoundException: org.apache.hive.jdbc.HiveDriver appears in the log file, or if a similar message appears in the file, this is an indication that the definition of the classpath of the Data Integration Hub Hadoop Service is incorrect. Perform the following actions:
   a. On the machine where the Data Integration Hub Hadoop Service is installed, open the `dih-hadoop-service.xml` file in a text editor from the following location:

      ```
      <DIHInstallationDir>/DataIntegrationHub/tomcat/conf/Catalina/localhost
      ```

b. Configure the classpath correctly. For more information, see "Configuring the Environment for a Hadoop Publication Repository" on page 72.

When I test the DIH__STAGING__HADOOP connection the test fails with the following error: `Data Access Connection test failed with the following error:DXServerException: The Hadoop file system is not available.`

Perform the following actions:

1. Verify that Apache Hadoop is running.
2. On the machine where the Data Integration Hub Hadoop Service is installed, open the `dih-hadoop-service.xml` file in a text editor from the following location:

   `<DIHInstallationDir>/DataIntegrationHub/tomcat/conf/Catalina/localhost`
3. Verify that all settings in the file are correct, including the correct locations of the `JAR` files. For more information, see "Configuring the Environment for a Hadoop Publication Repository" on page 72.

When I test the DIH__STAGING__HADOOP connection the test fails with the following error: `Data Access Connection test failed with the following error:DXServerException: ResourceAccessException: The connection is not valid. The Data Integration Hub Hadoop service is not running.`

The Data Integration Hub Hadoop Service is not running.

Log in to the machine where the Data Integration Hub Hadoop Service is installed and run the service. For more information, see "Starting and Stopping Data Integration Hub on Linux" on page 111.

Publication or subscription events stay in Processing status indefinitely

Verify that the configuration of the Data Integration Hub Hadoop Service is correct.

Perform the following actions:

1. On the machine where the Data Integration Hub server is installed, open the `dx-configuration.properties` file in a text editor from the following location:

   `<DIHInstallationDir>/DataIntegrationHub/tomcat/shared/classes/`

   Verify that the value of dx.server.url is correct.
2. If you installed the Data Integration Hub Hadoop Service on a different machine than the machine where you installed Data Integration Hub, open the `dx-configuration.properties` file in a text editor from the same location on the machine where the Data Integration Hub Hadoop Service is installed and verify that the value of dx.server.url is correct.
3. On the machine where the Data Integration Hub Hadoop Service is installed, ping the URL that is defined in dx.server.url and verify that it is accessible.

The running of publications and subscriptions fails. The Data Integration Hub Hadoop Service log shows that the service repeatedly tries to access `localhost:8020` and fails each time.

In Cloudera Manager, enable the option **Bind NameNode to Wildcard Address** and then restart the HDFS service.

Installation Log Files

Use the installation log files to troubleshoot a Data Integration Hub installation or upgrade.

The installer creates an installation log file during and after the installation. You can find the installation log file in the following directory:

```
<DIHInstallationDir>/logs
```

CHAPTER 11

Uninstallation

This chapter includes the following topics:

- Uninstallation Overview, 130
- Uninstalling Data Integration Hub from Windows Operating Systems, 130
- Uninstalling Data Integration Hub from UNIX Operating Systems, 131

Uninstallation Overview

Uninstall Data Integration Hub to remove the core application and additional components that you installed on the machine.

The uninstallation process does not delete the repositories or the Data Integration Hub document store.

The uninstallation process depends on the operating system on which Data Integration Hub is installed, Windows or UNIX.

Uninstalling Data Integration Hub from Windows Operating Systems

1. Stop all Data Integration Hub services.
2. In the Add/Remove Programs control panel, right-click **Data Integration Hub** and select **Uninstall**.

   The **Uninstall Data Integration Hub** wizard appears.
3. Click **Next**.

   The **Pre-Uninstall Summary** screen appears.
4. Click **Uninstall**.

   The **Uninstall Data Integration Hub** screen displays the progress of the uninstallation process. When the uninstallation process ends, the **Uninstall Complete** screen appears.
5. Click **Done** to close the wizard.

Uninstalling Data Integration Hub from UNIX Operating Systems

1. Stop all Data Integration Hub services.
2. Run the Uninstall.exe file from the Data Integration Hub installation directory.

   The **Uninstall Data Integration Hub** section appears.
3. Click **Next**.

   The **Pre-Uninstall Summary** section appears.
4. Click **Uninstall**.

   The uninstaller displays the progress of the uninstallation process. When the uninstallation process ends, the **Uninstall Complete** section appears.
5. Click **Done** to exit the uninstaller.

INDEX

A

after you upgrade
  description 105
  reapply configuration modifications 107
  tasks 105

B

before you upgrade
  description 82
  tasks 82
browser
  minimum system requirements 16

C

configuration
  Java heap size 117
  SNMP logs 114
connections
  repositories 75, 77
credentials
  changing for Data Integration Hub Hadoop Service 119
  changing for repository user account 118

D

Dashboard and reports
  importing operational data store event loader workflow 79
data archive
  configure source connection 125
Data Archive
  installing the Data Integration Hub accelerator 123, 124
Data Integration Hub Hadoop Service
  changing the credentials 119
document store
  setting up 22

E

event purging
  Data Archive 123, 124

H

Hadoop
  configuration files 72

I

Installation
  additional components 10
  components 9
  uninstalling from UNIX 131
  uninstalling from Windows 130
installer requirements
  minimum system requirements 16

L

log files
  location 114

M

minimum system requirements
  installer 16

O

operating system
  minimum system requirements 16
Operation Console
  logging in 75

P

partitions
  description 71
port numbers
  default 14
post-installation
  description 68
  Hadoop publication repository 72
  PowerCenter Integration Service 121
  registering PowerCenter server plug-in 69
  set up database partitions 71
  tasks 68
PowerCenter
  connections 75, 77
prerequisite
  Microsoft SQL Database 22
  pmrep 19
  software 18

R

repositories
  connections 75, 77

repository user account
changing the credentials 118
requirements
database 17
RMI port number
modifying 113

S

services
starting and stopping on Windows 110, 111
starting on Linux 111
SNMP appender
add to file 116
SNMP appender *(continued)*
parameters 115
SNMP logs
configuration 114
source connection
configure 125
system requirements
user accounts 12

U

user accounts
installation 12