Informatica®
10.5.1

Installation for Enterprise Data Preparation

Informatica Installation for Enterprise Data Preparation
10.5.1

Publication Date: 2022-07-14

Table of Contents

Preface

Follow the instructions in *Installation for Enterprise Data Preparation* to install Enterprise Data Preparation. The guide includes pre- and post-requisite tasks and steps to install the Informatica services and clients for the Informatica domain. Prerequisite tasks include planning the environment, setting up databases, and verifying system requirements. Post-requisite tasks include additional application services and configuring environment variables. Enterprise Data Preparation uses Enterprise Data Catalog for data discovery, lineage, and relationships. You must install, create, and configure Enterprise Data Catalog and its services before you install Enterprise Data Preparation.

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.

Part I: Installation Getting Started

This part contains the following chapter:

CHAPTER 1

Installation Getting Started

This chapter includes the following topics:

Checklist to Getting Started

This chapter contains high-level concepts and planning information related to installation. Use this checklist to track the completion of preliminary tasks.

❑ Understand high-level concepts:

- The installer description and process.
- Informatica domain terminology and components.

❑ Start high-level planning:

- Installation options. Review the installation options to know the product and options for installation.
- Installation components. Review the description of the installation components and the planning notes.

Installation Overview

Welcome to the Informatica installer Informatica domain services and clients. The Informatica domain services consist of core services to support the domain and application services. The Informatica clients consist of thick and web client applications.

When you install the Informatica domain services, you are prompted to create a domain or to join a domain. The domain is a collection of nodes that represent the machines on which the application services run. The first time you run the installer, you must create the domain. If you install on a single machine, you create the Informatica domain and a gateway node on the machine. If you install on multiple machines, you create an Informatica domain and a gateway node during the first installation. During the installation on the additional machines, you create gateway or worker nodes that you join to the domain.

When you run the installer, it installs files for services. You can optionally create application services during the installation process, or you can manually create application services when the installation completes.

If you have other Informatica products installed, verify that the installed version is compatible with the version of the product that you are installing.

Installation Process

The installation of the Informatica domain services and Informatica clients consists of multiple phases.

The installation process varies based on the products that you install. Consider the following high-level tasks of the installation process:

Perform pre-installation tasks.

1. Plan the Informatica installation. Determine the products that you want to run in your environment. If you are creating a domain, consider the number of nodes in the domain, the application services that will run on each node, the system requirements, and the type of user authentication that the domain will use.
2. Prepare the databases required for repositories, warehouses, and catalogs. Verify the database requirements and set up the databases.
3. Set up the machines to meet system requirements to ensure that you can successfully install and run the Informatica services.
4. Determine security requirements for the domain, services, and databases.

Run the installer.

When you run the installer, you can choose from different options based on your requirements.

Complete the configuration.

1. Verify code page compatibility.
2. Configure environment variables.
3. Complete tasks required by the type of user authentication used by the domain.
4. Optionally, configure secure communication for the domain.
5. Create and configure application services.
6. Configure connections required by the application services.
7. Create the users and connections required by the application services.

Install the Informatica client tools.

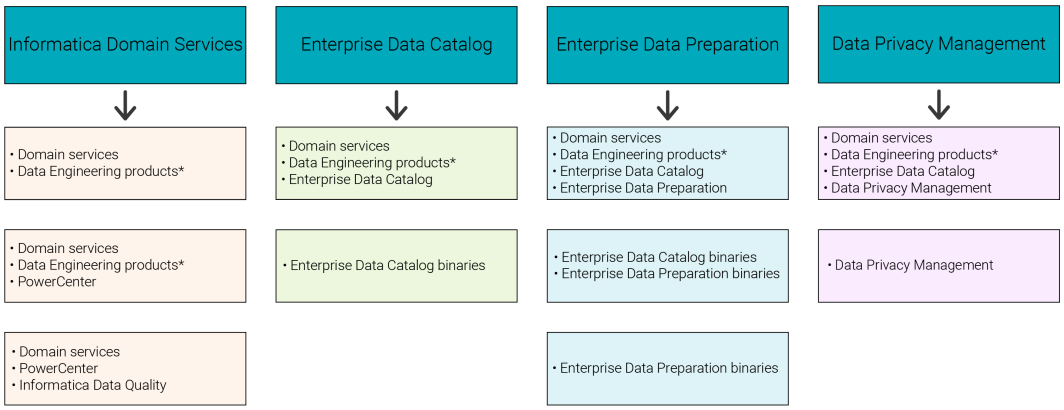
1. Verify the installation and third-party software requirements for the clients.
2. Use the client installer to install on Windows machines.
3. Configure required environment variables, and optionally install additional languages.

Plan the Installation Option

Before you begin the planning and preparation for install, determine the type of installation that you want to run.

When you run the installer, you can choose from options in the Welcome panel based on the product or products that you want to install. The Components panel appears based on your product selection so you can choose product components.

The following image shows the products that you can install based on the installation options:



*Data Engineering products include Data Engineering Integration, Data Engineering Quality, and Data Engineering Streaming.

Consider the different options available when you run the installer:

Informatica domain services

To install the Informatica domain services, you can select the installation option `1` in the Components panel to install and configure Informatica domain services.

With the Informatica domain services installation, install from one of the following product options:

- Only the Data Engineering products for Integration, Quality, and Streaming
- Traditional products and the aforementioned Data Engineering products
- Only traditional products such as PowerCenter and Informatica Data Quality

When you install Informatica domain services, you can choose to create a domain or join a domain. Test Data Management is installed with both traditional and Data Engineering products.

Enterprise Data Catalog

To install Enterprise Data Catalog, you can select the installation option `2` in the Components panel to install and configure Enterprise Data Catalog.

With the Enterprise Data Catalog installation, install from one of the following product options:

- Data Engineering products and Enterprise Data Catalog
- Enterprise Data Catalog binaries in an existing domain. After you install the binaries, you can run the installer again to configure the services.

Enterprise Data Preparation

To install Enterprise Data Preparation, you can select the following installation option `3` in the Components panel to install and configure Enterprise Data Preparation.

With the Enterprise Data Preparation installation, install from one of the following product options:

- Data Engineering products, Enterprise Data Catalog, and Enterprise Data Preparation.
- Enterprise Data Catalog and Enterprise Data Preparation binaries in an existing domain. After you install the binaries, you can run the installer again to configure the services.
- Only Enterprise Data Preparation binaries in an existing domain with Enterprise Data Catalog. After you install the binaries, you can run the installer again to configure the services.

Data Privacy Management

To install Data Privacy Management, you can select the following installation option `4` in the Components panel to install and configure Data Privacy Management.

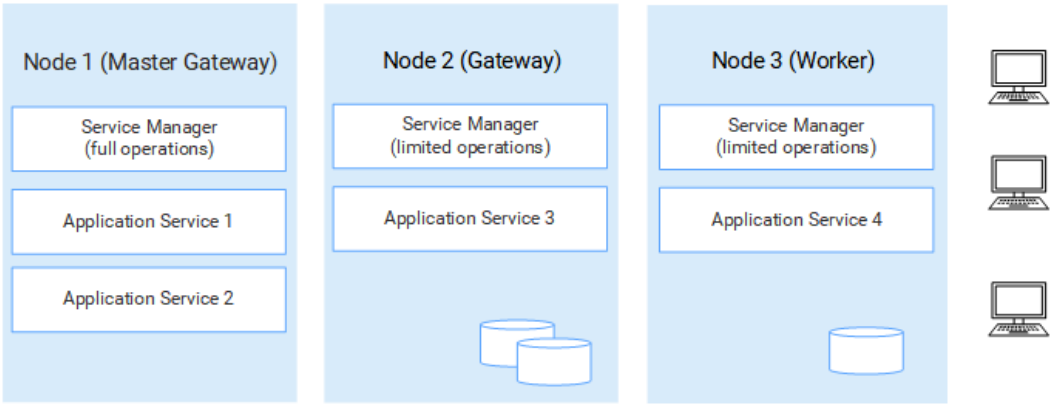
With the Data Privacy Management installation, install from one of the following product options:

- Data Engineering products, Enterprise Data Catalog, and Data Privacy Management.
- Data Privacy Management in an existing domain with Enterprise Data Catalog.

Plan the Installation Components

An Informatica domain is a collection of nodes and services. A node is the logical representation of a machine in a domain. Services include the Service Manager that manages all domain operations and a set of application services that represent server-based functionality. The domain and some services require databases to write metadata and run-time results.

The following image shows a high-level architecture of a domain on multiple nodes:



Nodes

The first time that you install the domain services, you create the Informatica domain and a gateway node. When you install the domain services on other machines, you create additional nodes that you join to the domain.

The domain has the following types of nodes:

- Gateway node. A gateway node is any node that you configure to serve as a gateway for the domain. A gateway node can run application services and it can serve as a master gateway node. The master gateway node is the entry point to the domain. You can configure more than one node as a gateway node, but only gateway node acts as the master gateway node at any given time.

- Worker node. A worker node is any node that you do not configure to serve as a gateway for the domain. A worker node can run application services, but it cannot serve as a gateway.

When you plan the installation: You need to plan the number and type of nodes that you need based on your service and processing requirements. If you have high availability, you will want to create more than one gateway node for fail-over functionality.

Service Manager

The Service Manager is a service that manages all domain operations. The Service Manager runs on each node in the domain and performs domain functions, such as authentication, logging and application service management. The Service Manager on a gateway node performs more tasks than the Service Manager on a worker node.

When you plan the installation: Note that the Service Manager functionality is associated with the type of node.

Application Services

Application services represent server-based functionality. An application service might be required or optional, and it might require access to a database.

When you run the installer, you can choose to create some services. After you complete the installation, you create other application services based on the license key generated for your organization.

When you plan the installation: When you plan the application services, you must account for the associated services that connect to the application service. You also must plan the relational databases that are required to create the application service.

Databases

Some application services require databases to store metadata and to write run-time results. You need to create databases for the application services in the domain.

You can create the following databases:

Domain configuration repository database

The domain configuration repository stores configuration and user information from a domain.

Data asset analytics repository database

The data asset analytics repository stores the analytical information collected from the catalog. You can view reports and statistical data on the Analytics tab in Enterprise Data Catalog.

Reference data warehouse database

The reference data warehouse stores the data values for reference table objects that you define in a Model repository. Configure a Content Management Service to identify the reference data warehouse and the Model repository.

Data object cache database

The data object cache stores cached logical data objects and virtual tables for the Data Integration Service. Data object caching enables the Data Integration Service to access pre-built logical data objects and virtual tables.

Profiling warehouse database

The profiling warehouse stores profiling and scorecard results. You need a profiling warehouse to perform profiling and data discovery.

Workflow database

The workflow database stores run-time metadata for workflows using the Data Integration Service.

Data Preparation repository database

The Data Preparation repository stores recipe and mapping metadata in the repository using the Interactive Data Preparation Service. The Data Preparation repository also stores worksheet metadata prepared by Enterprise Data Preparation users.

Model repository database

The Model repository stores data and metadata from the Informatica services and clients. Informatica client tools, such as Analyst tool and the Developer tool stores the data into the Model repository.

Monitoring Model repository database

The Monitoring Model repository stores statistics for ad hoc jobs, applications, logical data objects, SQL data services, web services, and workflows created by Informatica clients and application services.

When you plan the installation: You need to create databases and database users required by application services.

User Authentication

When you run the installer, you can choose the authentication to use for the domain.

The Informatica domain can use the following types of authentication to authenticate users in the domain:

- Native. Native user accounts are stored in the domain and can only be used within the domain. Native authentication is default.
- LDAP. LDAP user accounts are stored in an LDAP directory service and are shared by applications within the enterprise. You can configure LDAP authentication after you run the installer.
- SAML. You can configure Security Assertion Markup Language (SAML) authentication for the Administrator tool, the Analyst tool, and the Monitoring tool. You can configure SAML authentication after you run the installer.

When you plan the installation: You need to plan the type of authentication that you want to use in the domain.

Secure Data Storage

Informatica encrypts sensitive data before it stores the data in the Informatica repositories.

When you create a domain, you must specify the encryption key directory. The installer generates an encryption key file named siteKey and stores it in a default directory or the directory you specify. All nodes in a domain must use the same encryption key.

Important: The installer also generates a unique site key. If you lose the site key, you cannot generate the site key again. Make sure that you save a copy of this key and do not share the unique site key with others.

Domain Security

When you create a domain, you can enable options to configure security in the domain.

You can configure secure communication for the following domain components:

- Administrator tool. Configure a secure HTTPS connection for the Administrator tool. During installation, you can provide the keystore file to use for the HTTPS connection.
- Service Manager. Configure a secure connection between the Service Manager and other domain services. During installation, you can provide keystore and truststore files containing SSL certificates that you want to use.
- Domain configuration repository. You can secure the domain configuration repository with SSL protocol. During installation, you can provide the truststore file containing the SSL certificate that you want to use.

When you plan the installation: Determine the level of security that you want to configure for the domain components. If you decide to configure security for the domain, you must know the location and password for the keystore and truststore files.

Informatica Client Tools

You use Informatica clients to access underlying Informatica functionality in the domain. The clients make requests to the Service Manager and to application services.

The Informatica clients consist of thick client applications and thin or web client applications that you use to access services and repositories in the domain.

The following table describes the client tools for Enterprise Data Preparation:

Informatica Client	Description
Informatica Developer (the Developer tool)	A thick client application to create, import, and export data domains.
Informatica Administrator (the Administrator tool)	A web application to manage the domain and application services.
Informatica Catalog Administrator	A web application to administer resources, scanners, schedules, attributes, and connections.
Enterprise Data Catalog	A web application that displays a comprehensive view of metadata from configured data assets.
Enterprise Data Preparation application	A web application for Enterprise Data Preparation to search for, discover, and prepare data assets.

When you plan the installation: Determine how many instances of the Developer tool you want to install. You do not need to plan for the web client applications.

Part II: Before You Install the Services

This part contains the following chapters:

CHAPTER 2

Before You Install the Services on UNIX or Linux

This chapter includes the following topics:

Before You Begin Checklist

This chapter contains preliminary tasks that you must complete. Use this checklist to track preliminary tasks before you prepare for services.

- ❑ Read the Informatica Release Notes for updates to the installation and upgrade process.
- ❑ Verify system requirements:
  - Verify sizing requirements based upon your processing and concurrency requirements.
  - Review the patch requirements to verify that the machine has the required operating system patches and libraries.
  - Verify that the port numbers to use for application service processes are available on the machines where you install the Informatica services.
  - Verify that the operating system meets the file descriptor limit.
- ❑ Back up the Data Transformation files that were created in a previous installation.
- ❑ Review system environment variables.
- ❑ Create a system user account to run the installer.

- ❑ Set up keystore and truststore files if you want to configure secure communication for the domain and set up a secure connection to web client applications.
- ❑ Extract the installer files.
- ❑ Verify the license key.

Read the Release Notes

Read the Release Notes for updates to the installation and upgrade process. You can also find information about known and fixed issues for the release.

Find the Release Notes on the Informatica documentation portal.

Verify System Requirements

Verify that your environment meets the minimum system requirements for the installation process, temporary disk space, port availability, databases, and application service hardware.

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

Enterprise Data Catalog System Requirements

Verify that your machine meets the minimum system requirements to install the Enterprise Data Catalog services.

The system requirements differ based on the following conditions:

If the Informatica Domain, data nodes, and processing nodes are on the same machine with two million assets in the catalog

The minimum memory requirement for the Linux operating system is 56 GB RAM. The minimum disk space required is 185 GB. The number of CPU cores required is 24.

If the Informatica Domain, data nodes, and processing nodes are on different machines

The minimum memory requirement for the Linux operating system is 24 GB for a cluster node and 32 GB for the machine on which Informatica domain runs. The minimum disk space required is 125 GB. The number of CPU cores required is 8 cores for a cluster node and 16 cores for the machine on which the Informatica domain runs.

Data Asset Analytics Repository Database Server System Requirements

The minimum system requirements for the machine that hosts the Data Asset Analytics repository database is based on the number of assets in the catalog and the database type. For information about the hardware requirements such as the number of CPU cores and the memory, see the Performance Tuning Parameters for Data Asset Analytics section of the *Enterprise Data Catalog Performance Tuning Guide*.

Advanced Scanners Requirements

Before you generate PNG visualizations, perform the following steps to install the additional libraries and rebuild the fonts cache on the machine:

1. Run the following command:

```
# install required fonts
sudo yum install fontconfig dejavu-sans-fonts dejavu-serif-fonts
```

   to install the following libraries:

   - fontconfig
   - dejavu-sans-fonts
   - dejavu-serif-fonts

2. Run the following command:

```
# rebuild fonts cache
fc-cache -f -v
```

   to rebuild the fonts cache.

Enterprise Data Preparation System Requirements

Verify that your machine meets the minimum system requirements to install the Enterprise Data Preparation services.

The system requirements differ based on the following conditions:

If the Informatica Domain and Hadoop cluster are on the same machine

The minimum memory requirement for the Linux operating system is 32 GB RAM. The minimum disk space required is 150 GB. The number of CPU cores required is 24.

If the Informatica Domain and Hadoop cluster are on different machines

The minimum memory requirement for the Linux operating system is 24 GB for a cluster node and 32 GB for the machine on which Informatica domain runs. The minimum disk space required is 100 GB. The number of CPU cores required is 8 cores for a cluster node and 16 cores for the machine on which the Informatica domain runs.

Verify Temporary Disk Space and Permissions

Verify that your environment meets the minimum system requirements for the temporary disk space, permissions for the temporary files, and the Informatica client tools.

Disk space for the temporary files

The installer writes temporary files to the hard disk. Verify that you have 1 GB disk space on the machine to support the installation. When the installation completes, the installer deletes the temporary files and releases the disk space.

The following table describes the minimum disk space and memory requirements for PowerCenter or Data Engineering product installation:

Options	Minimum Requirements
Temporary disk space to run the installer	1 GB disk space
Install with application services	50 GB disk space, 8 GB RAM, and 8 cores. Out of the 50 GB, 25 GB is for the product installation binaries.

Permissions for the temporary files

Verify that you have read, write, and execute permissions on the `/tmp` directory.

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 Sizing Requirements

Allocate resources for installation and deployment of services based on the expected deployment type of your environment.

Before you allocate resources, you need to identify the deployment type based on your requirements for the volume of processing and the level of concurrency. Based on the deployment type, you can allocate resources for disk space, cores, and RAM. You can also choose to tune services when you run the installer.

Determine the Installation and Service Deployment Type

The following table describes the environment for the different deployment types:

Deployment Type	Environment Description
Sandbox	Used for proof of concepts or as a sandbox with minimal users.
Basic	Used for low volume processing with low levels of concurrency.
Standard	Used for high volume processing with low levels of concurrency.
Advanced	Used for high volume processing with high levels of concurrency.

Identify Sizing Requirements

The following table provides the minimum sizing requirements for the Informatica domain node:

Deployment Type	Disk Space per Node	Total Virtual Cores	RAM per Node
Sandbox	140 GB	16	32 GB
Basic	140 GB	24	64 GB

Deployment Type	Disk Space per Node	Total Virtual Cores	RAM per Node
Standard	140 GB	48	64 GB
Advanced	140 GB	96	128 GB

The following table provides the minimum sizing requirements for Enterprise Data Preparation:

	Sandbox	Basic	Standard	Advanced
Maximum Heap Size	2 GB	2 GB	2 GB	4 GB
Number of Catalog Objects	100,000	1 Million	20 Million	50 Million
Total Data Volume	100 GB	1 TB	10 TB	100 TB
Operational Data Volume	100 MB	1 GB	10 GB	100 GB
Number of Active Users	1	1-10	11-50	51-100
Total Number of CPU Cores for YARN (Hadoop Clusters running all jobs)	12	24	84	204
Total Number of CPU Cores for YARN (Hadoop Clusters dedicated to Data Integration Service)	8	16	60	156
Total Memory for YARN (Hadoop Clusters running all jobs)	32 GB	32 GB	128 GB	320 GB
Total Memory for YARN (Hadoop Clusters dedicated to Data Integration Service)	8 GB	32 GB	64 GB	192 GB

The sizing requirements account for the following factors:

- Disk space required to extract the installer
- Temporary disk space to run the installer
- Disk space required to install the services and components
- Disk space required for log directories

- Requirements to run the application services

The sizing numbers do not account for operational data processing and object caching requirements for native mode of execution.

Note: For cloud deployments, choose machines with a configuration that is closest to the sizing requirements.

Tune During Installation

When you run the installer, you can choose to tune the services based on the deployment size. If you create a Model Repository Service, a Data Integration Service, or a Content Management Service during installation, the installer can tune the services based on the deployment type that you enter. The installer configures properties such as maximum heap size and execution pool size.

You can tune services at any time after you install the services by using the infacmd autotune command. When you run the command, you can tune properties for other services as well as the Hadoop run-time engine properties.

Review Patch Requirements on UNIX or Linux

Before you install the Informatica services, verify that the machine has the required operating system patches and libraries.

Data Engineering on Linux

The following table lists the patches and libraries that the Informatica services require on Linux:

Platform	Operating System	Operating System Patch
AWS Linux	Linux 2 - 2.0.20210126	All of the following packages: - e2fsprogs-libs-1.42.9-12.amzn2.0.2.x86_64 - keyutils-libs-1.5.8-3.amzn2.0.2.x86_64 - libselinux-2.5-12.amzn2.0.2.x86_64 - libsepol-2.5-8.1.amzn2.0.2.x86_64
Ubuntu	20.04.1	All of the following packages: - e2fsprogs/focal,now 1.45.5-2ubuntu1 amd64 [installed] - libkeyutils1/focal,now 1.6-6ubuntu1 amd64 [installed,automatic] - libselinux1/focal,now 3.0-1build2 amd64 [installed,automatic] - libsepol1/focal,now 3.0-1 amd64 [installed,automatic]
Linux-x64	Red Hat Enterprise Linux 6.7	All of the following packages, where <version> is any version of the package: - e2fsprogs-libs-<version>.el6 - keyutils-libs-<version>.el6 - libselinux-<version>.el6 - libsepol-<version>.el6
Linux-x64	Red Hat Enterprise Linux 7.3	All of the following packages, where <version> is any version of the package: - e2fsprogs-libs-<version>.el7 - keyutils-libs-<version>.el7 - libselinux-<version>.el7 - libsepol-<version>.el7

Platform	Operating System	Operating System Patch
Linux-x64	Red Hat Enterprise Linux 8	All of the following packages, where <version> is any version of the package: - e2fsprogs-libs-<version>.el8 - keyutils-libs-<version>.el8 - libselinux-<version>.el8 - libsepol-<version>.el8
Linux-x64	SUSE Linux Enterprise Server 12	Service Pack 2
Linux-x64	SUSE Linux Enterprise Server 15	Service Pack 0 and Service Pack 1.

Verify Port Requirements

The installer sets up the ports for components in the Informatica domain, and it designates a range of dynamic ports to use for some application services.

You can specify the port numbers to use for the components and a range of dynamic port numbers to use for the application services. Or you can use the default port numbers provided by the installer. Verify that the port numbers are available on the machines where you run the installer.

Note: Services and nodes can fail to start if there is a port conflict.

The following table describes the port requirements for installation:

Port	Description
Node port	Port number for the node created during installation. Default is 6005.
Service Manager port	Port number used by the Service Manager on the node. The Service Manager listens for incoming connection requests on this port. Client applications use this port to communicate with the services in the domain. The Informatica command line programs use this port to communicate to the domain. This is also the port for the SQL data service JDBC/ODBC driver. Default is 6006.
Service Manager Shutdown port	Port number that controls server shutdown for the domain Service Manager. The Service Manager listens for shutdown commands on this port. Default is 6007.
Informatica Administrator port	Port number used by Informatica Administrator. Default is 6008.
Informatica Administrator shutdown port	Port number that controls server shutdown for Informatica Administrator. Informatica Administrator listens for shutdown commands on this port. Default is 6009.
Minimum port number	Lowest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6014.
Maximum port number	Highest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6114.

Port	Description
Range of dynamic ports for application services	Range of port numbers that can be dynamically assigned to application service processes as they start up. When you start an application service that uses a dynamic port, the Service Manager dynamically assigns the first available port in this range to the service process. The number of ports in the range must be at least twice the number of application service processes that run on the node. Default is 6014 to 6114. The Service Manager dynamically assigns port numbers from this range to the Model Repository Service.
Static ports for application services	Static ports have dedicated port numbers assigned that do not change. When you create the application service, you can accept the default port number, or you can manually assign the port number. The following services use static port numbers: - Catalog Service. Default is 9085 for HTTP. - Content Management Service. Default is 8105 for HTTP. - Data Integration Service. Default is 8095 for HTTP. - Interactive Data Preparation Service. Default is 8099 for HTTP. - Enterprise Data Preparation Service. Default is 9045 for HTTP.

Guidelines for Port Configuration

The installer validates the port numbers that you specify to ensure that there will be no port conflicts in the domain.

Use the following guidelines to determine the port numbers:

- The port number you specify for the domain and for each component in the domain must be unique.
- The port number for the domain and domain components cannot be within the range of the port numbers that you specify for the application service processes.
- The highest number in the range of port numbers that you specify for the application service processes must be at least three numbers higher than the lowest port number. For example, if the minimum port number in the range is 6400, the maximum port number must be at least 6403.
- The port numbers that you specify cannot be lower than 1025 or higher than 65535.

Verify the File Descriptor Limit

Verify that the operating system meets the file descriptor requirement.

Informatica service processes can use a large number of files. To prevent errors that result from the large number of files and processes, you can change system settings with the limit command if you use a C shell, or the ulimit command if you use a Bash shell.

List Operating System Settings

To get a list of the operating system settings, including the file descriptor limit, run the following command:

With C shell, run `limit`

With Bash shell, run `ulimit -a`

Set the File Descriptor Limit

Informatica service processes can use a large number of files. Set the file descriptor limit per process to 16,000 or higher. The recommended limit is 32,000 file descriptors per process.

To change system settings, run the limit or ulimit command with the pertinent flag and value. For example, to set the file descriptor limit, run the following command:

With C shell, run `limit -h filesize <value>`

With Bash shell, run `ulimit -n <value>`

Set Max User Processes

Informatica services use a large number of user processes. Use the ulimit -u command to adjust the max user processes setting to a level that is high enough to account for all the processes required by the Blaze engine.

To set the max user processes, run the following command: Run the following command to set the max user processes setting:

With C shell, run `limit -u processes <value>`

With Bash shell, run `ulimit -u <value>`

Back Up the Data Transformation Files

Before installation, you must back up the Data Transformation files that were created under previous versions. After you complete the installation, copy the files to the new installation directories to get the same repository and custom global components as in the previous version.

The following table lists the files or directories that you must back up:

File or Directory	Default Location
Repository	<Informatica installation directory>\DataTransformation\ServiceDB
Custom Global Components directory (TGP files)	<Informatica installation directory>\DataTransformation\autoInclude\user
Custom Global Components directory (DLL and JAR files)	<Informatica installation directory>\DataTransformation\externLibs\user
Configuration file	<Informatica installation directory>\DataTransformation\CMConfig.xml
License file	<Informatica installation directory>\DataTransformation\CDELicense.cfg

Do not copy the Data Transformation Library files. Instead, install the Data Transformation Libraries again.

Review the Environment Variables

Configure environment variables for the Informatica installation.

The following table describes the environment variables to review:

Variable	Description
IATEMPDIR	Location of the temporary files created during installation. Informatica requires 1 GB disk space for temporary files. Configure the environment variable if you do not want to create temporary files in the `/tmp` directory. If you want to change the default `/tmp` directory, you must set IATEMPDIR and _JAVA_OPTIONS environment variables to the new directory. For example, set the variable to export IATEMPDIR=/home/user. **Note:** Unset the IATEMPDIR variable after the installation.
_JAVA _OPTIONS	Configure the environment variable to change the temporary directory. If you want to change the default /tmp directory, you must set IATEMPDIR and _JAVA_OPTIONS the environment variables to the new directory. For example, set the variable to export _JAVA_OPTIONS=-Djava.io.tmpdir=/home/user. **Note:** Unset the _JAVA _OPTIONS variable after the installation.
LANG and LC_ALL	Change the locale to set the appropriate character encoding for the terminal session. For example, set the encoding to `Latin1` or `ISO-8859-1` for French, `EUC-JP` or `Shift JIS` for Japanese, or `UTF-8` for Chinese or Korean. The character encoding determines the types of characters that appear in the UNIX terminal.
DISPLAY	Unset the DISPLAY environment before you run the installer. Installation might fail if the DISPLAY environment variable has some value.

Note: Make sure that the NOEXEC flag is not set for the file system mounted on the /tmp directory.

Create a System User Account

Create a user account specifically to run the Informatica service.

Verify that the user account you use to install Informatica has write permission on the installation directory.

Verify that the user account that installs the Informatica service does not have any privileges and permissions to access sensitive files on the machine where you install the Informatica services.

Set Up a Keystore File

When you install the Informatica services, you can configure secure communication for the domain and set up a secure connection to Informatica Administrator (the Administrator tool). If you configure these security options, you must set up keystore and truststore files.

Before you install the Informatica services, set up the files for secure communication within the Informatica domain or for a secure connection to the Administrator tool. To create the required files, you can use the following programs:

keytool

You can use keytool to create an SSL certificate or a Certificate Signing Request (CSR) as well as keystores and truststores in JKS format.

OpenSSL

You can use OpenSSL to create an SSL certificate or CSR as well as convert a keystore in JKS format to PEM format.

For more information about OpenSSL, see the documentation on the following website: https://www.openssl.org/docs/

For a higher level of security, send your CSR to a Certificate Authority (CA) to get a signed certificate.

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.

Secure Communication Within the Informatica domain

Before you enable secure communication within the Informatica domain, verify that the following requirements are met:

You created a certificate signing request (CSR) and private key.

You can use keytool or OpenSSL to create the CSR and private key.

If you use RSA encryption, you must use more than 512 bits.

You have a signed SSL certificate.

The certificate can be self-signed or CA signed. Informatica recommends a CA signed certificate.

You imported the certificate into keystores.

You must have a keystore in PEM format named `infa_keystore.pem` and a keystore in JKS format named `infa_keystore.jks`.

The keystore files must contain the root and intermediate SSL certificates.

Note: The password for the keystore in JKS format must be the same as the private key pass phrase used to generate the SSL certificate.

You imported the certificate into truststores.

You must have a truststore in PEM format named `infa_truststore.pem` and a truststore in JKS format named `infa_truststore.jks`.

The truststore files must contain the root, intermediate, and end user SSL certificates.

The keystores and truststores are in the correct directory.

The keystore and truststore must be in a directory that is accessible to the installer.

For more information about how to create a custom keystore and truststore, see the Informatica How-To Library article "How to Create Keystore and Truststore Files for Secure Communication in the Informatica Domain".

Secure Connection to the Administrator tool

Before you secure the connection to the Administrator tool, verify that the following requirements are met:

You created a certificate signing request (CSR) and private key.

You can use keytool or OpenSSL to create the CSR and private key.

If you use RSA encryption, you must use more than 512 bits.

You have a signed SSL certificate.

The certificate can be self-signed or CA signed. Informatica recommends a CA signed certificate.

You imported the certificate into a keystore in JKS format.

A keystore must contain only one certificate. If you use a unique certificate for each web application service, create a separate keystore for each certificate. Alternatively, you can use a shared certificate and keystore.

If you use the installer-generated SSL certificate for the Administrator tool, you do not need to import the certificate into a keystore in JKS format.

The keystore is in the correct directory.

The keystore must be in a directory that is accessible to the installer.

Extract the Installer Files

The installer files are compressed and distributed as a compressed file.

You can get the installation file from the FTP link contained in your fulfillment email. Download the Informatica installation tar file from the Informatica Electronic Software Download site to a directory on your machine and then extract the installer files.

Extract the installer files to a directory on your machine. The user that runs the installer must have read and write permissions on the installer files directory and execute permissions on the executable file.

Note: Make sure that you download the file to a local directory or a shared network drive that is mapped on your machine. You can then extract the installer files. However, you cannot run the installer from a mapped file. Copy the extracted files to a local drive and then run the installer.

Installer Code Signing

You can verify the signature of the Informatica software code.

Informatica uses a certificate based digital signature to sign the Informatica software code. The code signing helps to validate the authenticity of the code and ensures that there has been no changes or corruptions to the code after Informatica signs the code. You can determine whether to trust the software based on whether the code sign is present or not.

You can request a code signing certificate that contains information that fully identifies Informatica LLC and a Certificate Authority (CA) that issues the certificate. The digital certificate binds the identity of Informatica to a public key and to a private key.

Digital signing of software begins with the creation of a cryptographic hash, or a digest. The digest has a one to one correspondence with the original data. Use the digest as there are no hints on how to recreate the original data, and even a small change in the original data results in a change in the hash value. Informatica uses its private key to sign the digest, or generates a signature in the form of a string of bits. Good digital signature algorithms allow a user with the public key to verify the creator of the signature.

To Verify the Signed Code is Authentic

After Informatica signs the software bundle, you can contact Informatica Global Customer Support to access the code signing certificate. Informatica ships the installer along with the signature file that contains the

hash of the installer binary encrypted with Informatica's private key. You can validate the integrity of digitally signed binaries using any available tools, such as OpenSSL.

For instance, if you have to verify the package authentication and confirm the code security, enter the following OpenSSL commands:

```
openssl base64 -d -in $signature -out /tmp/sign.sha256
openssl dgst -sha256 -verify <(openssl x509 -in <cert> -pubkey -noout) -signature /tmp/
sign.sha256 <file>
```

Where `<signature>` is the file containing the signature in Base64, `<cert>` is the code signing certificate, and `<file>` is the file to verify.

Based on verification process, OpenSSL displays a success or error message to validate if the installer code is genuine or not. Note that the verification for the installer might take around two minutes.

Verify the License Key

Before you install the software, verify that you have the license key available.

When you download the installation files from the Informatica Electronic Software Download (ESD) site, the license key is in an email message from Informatica. Copy the license key file to a directory accessible to the user account that installs the product.

Contact Informatica Global Customer Support if you do not have a license key or if you have an incremental license key and you want to create a domain.

Prepare for Cluster Import

When you run the installer, you can choose to configure the cluster. The cluster configuration enables the Data Integration Service to push mapping logic to the cluster. To integrate the Informatica domain with the non-native cluster, you must import a cluster configuration. You can import the cluster information directly from the cluster or from an archive file.

You can import cluster information from an archive file of any supported cluster into the domain. Your administrator might prefer to provide you with the archive file to protect sensitive connection information to the cluster. The archive file can be in a .zip or .tar format. Ensure that you store the archive file locally.

Prepare the Archive File for Hadoop Environment

To import the cluster configuration from Amazon EMR, MapR, or Google Dataproc cluster, you must import from an archive file. The Hadoop cluster configuration archive file can have the following contents based on the distribution:

- core-site.xml
- hbase-site.xml. hbase-site.xml is required only if you access HBase sources and targets.
- hdfs-site.xml
- hive-site.xml
- mapred-site.xml or tez-site.xml. Include the mapred-site.xml file or the tez-site.xml file based on the Hive execution type used on the Hadoop cluster.

- yarn-site.xml

Note: When you configure a CDP Public Cloud cluster, the hbase-site.xml file is on the Data Lake cluster. The other files are on the Data Hub cluster.

Prepare the Archive File for the Databricks Environment

To create the .xml file for import, you must get the required information from the Databricks administrator. You can provide any name for the file and store it locally.

The following table describes the cluster properties required to configure in the import file for the Databricks environment:

Property Name	Description
cluster_name	Name of the Databricks cluster.
cluster_ID	The cluster ID of the Databricks cluster.
base URL	URL to access the Databricks cluster.
accesstoken	Token ID created within Databricks required for authentication.

Optionally, you can include other properties specific to the Databricks environment. When you complete the .xml file, compress it into a .zip or .tar file for import.

CHAPTER 3

Prepare for Application Services and Databases

This chapter includes the following topics:

Checklist to Prepare for Application Services

This chapter contains information about application services and databases for the Informatica environment. Use this checklist to track service planning and database preparation.

- ❑ Identify the application services that you need in your environment.
- ❑ Identify the application services that you want the installer to create.

❑ Prepare databases for the services:

- Create the database.
- Create a user for the database.
- Create environment variables.
- Configure connectivity.

Prepare for Application Services and Databases Overview

When you plan the application services, you must account for the associated services that connect to the application service. You also must plan the relational databases that the application service requires.

The installer prompts you to optionally create some services during the installation. Some service properties require database information. If you want the installer to create a service that requires a database, you must prepare the database before you run the installer. To prepare the databases, verify the data base requirements, set up the database, and set up a user account. The database requirements depend on the application services that you create.

If you do not create services during installation, you can create them manually after you install.

Set Up Database User Accounts

Set up a database and user account for the repository databases.

Use the following rules and guidelines when you set up the user accounts:

- The database user account must have permissions to create and drop tables, indexes, and views, and to select, insert, update, and delete data from tables.
- Use 7-bit ASCII to create the password for the account.
- To prevent database errors in one repository from affecting any other repository, create each repository in a separate database schema with a different database user account. Do not create a repository in the same database schema as the domain configuration repository or any other repository in the domain.

Identify Application Services by Product

Each application service provides different functionality within the Informatica domain. You create application services based on the license key generated for your organization.

The following table lists the application services that each product uses:

Product	Application Services
Enterprise Data Catalog	- Analyst Service - Catalog Service * - Content Management Service * - Data Integration Service * - Metadata Access Service * - Model Repository Service * - monitoring Model Repository Service * - Informatica Cluster Service * - Search Service
Enterprise Data Preparation	- Catalog Service - Content Management Service * - Data Integration Service * - Interactive Data Preparation Service * - Enterprise Data Preparation Service * - Metadata Access Service * - Model Repository Service * - monitoring Model Repository Service *
Data Engineering Integration	- Analyst Service - Data Integration Service * - Mass Ingestion Service - Metadata Access Service * - Model Repository Service * - monitoring Model Repository Service *
Data Engineering Quality	- Analyst Service - Content Management Service * - Data Integration Service * - Mass Ingestion Service - Metadata Access Service * - Model Repository Service * - monitoring Model Repository Service * - Search Service
Data Engineering Streaming	- Analyst Service - Data Integration Service * - Mass Ingestion Service - Metadata Access Service * - Model Repository Service * - monitoring Model Repository Service *
** You can create these services when you install the product.*	

Domain Configuration Repository Database Requirements

Informatica components store metadata in relational database repositories. The domain stores configuration and user information in a domain configuration repository.

You must set up a database and user account for the domain configuration repository before you run the installation. The database must be accessible to all gateway nodes in the Informatica domain.

When you install Informatica, you provide the database and user account information for the domain configuration repository. The Informatica installer uses JDBC to communicate with the domain configuration repository.

The domain configuration repository supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- Oracle
- PostgreSQL
- Sybase ASE

Allow 200 MB of disk space for the database.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- If the repository is in an IBM DB2 database, verify that IBM DB2 Version 10.5 is installed.
- On the IBM DB2 instance where you create the database, set the following parameters to ON:
  - DB2_SKIPINSERTED
  - DB2_EVALUNCOMMITTED
  - DB2_SKIPDELETED
  - AUTO_RUNSTATS
- On the database, set the configuration parameters.

  The following table lists the configuration parameters that you must set:

Parameter	Value
logfilsiz	8000
maxlocks	98
locklist	50000
auto_stmt_stats	ON

- Set the tablespace pageSize parameter to 32768 bytes.

In a single-partition database, specify a tablespace that meets the pageSize requirements. If you do not specify a tablespace, the default tablespace must meet the pageSize requirements.

In a multi-partition database, specify a tablespace that meets the pageSize requirements. Define the tablespace in the catalog partition of the database.

- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.
- Verify that the database user has CREATETAB, CONNECT, and BINDADD privileges.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- In the DataDirect Connect for JDBC utility, update the DynamicSections parameter to 3000.

  The default value for DynamicSections is too low for the Informatica repositories. Informatica requires a larger DB2 package than the default. When you set up the DB2 database for the domain configuration repository or a Model repository, you must set the DynamicSections parameter to at least 3000. If the DynamicSections parameter is set to a lower number, you can encounter problems when you install or run Informatica services.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

  CREATE SEQUENCE

  CREATE SESSION

  CREATE SYNONYM

  CREATE TABLE

  CREATE VIEW

- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

PostgreSQL Database Requirements

Use the following guidelines when you set up the repository on PostgreSQL:

- Verify that the database user account has CONNECT, CREATE TABLE, and CREATE VIEW privileges.
- Specify the database schema name when you use PostgreSQL as the database.
- Ensure that PostgreSQL has sufficient disk space for the data files. By default, the data files are present in the following location:

  ```
  <PostgreSQL installation directory>/data
  ```

- On the database, set the configuration parameters.

  The following table lists the minimum and recommended values for the configuration parameters that you must set:

Parameter	Minimum Value	Recommended Value
max_connections	200	4000
shared_buffers	2 GB	16 GB
max_locks_per_transaction	1024	1024
max_wal_size	1 GB	8 GB
checkpoint_timeout	5 minutes	30 minutes

Sybase Database Requirements

Use the following guidelines when you set up the repository on Sybase ASE:

- Set the database server page size to 16K or higher. You must set the page size to 16K as this is a one-time configuration and cannot be changed afterwards.
- Set the database locking configuration to use row-level locking.

The following table describes the database locking configuration that you must set:

Database Configuration	Sybase System Procedure	Value
Lock scheme	sp_configure "lock scheme"	0, datarows

- Set the Sybase database option "ddl in tran" to TRUE.
- Set "allow nulls by default" to TRUE.
- Turn ON the Sybase database option select into/bulkcopy/pllsort.
- Enable the "select" privilege for the sysobjects system table.
- Create the following login script to disable the default VARCHAR truncation:

```
create procedure dbo.sp_string_rtrunc_proc as set string_rtruncation on
sp_modifylogin "user_name", "login script", sp_string_rtrunc_proc
```

  The login script is executed every time the user logs into the Sybase instance. The stored procedure sets the parameter at the session level. The sp_modifylogin system procedure updates "user_name" with the stored procedure as its "login script". The user must have permission to invoke the stored procedure.

- Verify that the database user has CREATE DEFAULT, CREATE PROCEDURE, CREATE RULE, CREATE TABLE, and CREATE VIEW privileges.
- Set the database configurations to the recommended baseline values.
  The following table lists the database memory configuration parameters that you must set:

Database Configuration	Sybase System Procedure	Value
Maximum amount of total physical memory	sp_configure "max memory"	2097151
Procedure cache size	sp_configure "procedure cache size"	500000
Number of open objects	sp_configure "number of open objects"	5000
Number of open indexes	sp_configure "number of open indexes"	5000
Number of open partitions	sp_configure "number of open partitions"	5000
Heap memory per user	sp_configure "heap memory per user"	49152
Number of locks	sp_configure "number of locks"	100000

Analyst Service

The Analyst service runs the Analyst tool. It manages the connections between service components and the user that have access to the Analyst tool. When you create the service, you need to associate other application services with it.

The following table summarizes some dependencies that are associated with the Analyst Service:

Dependency	Summary
Products	The following products use the Analyst Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Enterprise Data Catalog - Informatica Data Quality - PowerCenter - Test Data Management
Services	The Analyst Service requires a direct association with the following services: - Data Integration Service - Model Repository Service
Databases	The Analyst Service does not have any associated database.
Installer	You cannot create the Analyst Service during installation.

Catalog Service

The Catalog Service is an application service that runs Enterprise Data Catalog in the Informatica domain. It manages the connections between service components and the users that have access to Enterprise Data Catalog search interface and Catalog Administrator.

The catalog represents an indexed inventory of all the configured data assets in an enterprise. You can find metadata and statistical information, such as profile statistics, data asset ratings, data domains, and data relationships, in the catalog.

The following table summarizes the dependencies for products, services, and databases that are associated with the Catalog Service:

Dependency	Summary
Products	The following products use the Catalog Service: - Enterprise Data Catalog - Enterprise Data Preparation - Data Privacy Management
Services	The Catalog Service depends on the following services: - Content Management Service - Data Integration Service - Informatica Cluster Service - Model Repository Service - Analyst Service

Dependency	Summary
Databases	The Catalog Service does not have any associated database. If you configure Data Asset Analytics for the Catalog Service, you can select from one of the following databases that you want to configure for Data Asset Analytics: - Oracle - SQL Server - PostgreSQL
Installer	You can create the Catalog Service when you install Enterprise Data Catalog.

Data Asset Analytics Repository Database Requirements

The Data Asset Analytics repository database stores the analytical information collected from the catalog. The analytical information is displayed as reports and statistical data in the **Analytics** tab in Enterprise Data Catalog. You can access the **Analytics** tab after you configure Data Asset Analytics in the Catalog Service and assign the required privileges for users to access the **Analytics** tab.

You can configure any of the following databases as the repository database for Data Asset Analytics:

- Oracle
- SQL Server
- PostgreSQL

Oracle Database Requirements

Perform the following steps before you configure Oracle as the repository database:

- Create a user name and password to access the Oracle database using the CREATE USER <Database Schema> command.
- Provide the following privileges to the user name:
  - GRANT CONNECT TO <Database Schema>;
  - GRANT RESOURCE TO <Database Schema>;
  - GRANT CREATE VIEW TO <Database Schema>;
  - GRANT CREATE MATERIALIZED VIEW TO <Database Schema>;
  - GRANT UNLIMITED TABLESPACE TO <Database Schema>. Alternatively, use the command ALTER USER <Database Schema> QUOTA <SIZE> ON <tablespace name>;

SQL Server Database Requirements

Perform the following steps before you configure SQL Server as the repository database:

- Make sure that you use the dbo schema for SQL Server.
- Run the following commands to create the database, the user name, password and assign the required privileges for the user:
  - USE master
  - GO
  - CREATE DATABASE <new database name>;
  - GO
  - use <new database name>

- CREATE LOGIN <new login name> WITH PASSWORD = '<password>';
- CREATE USER <new user name> FOR LOGIN <new login name> WITH DEFAULT_SCHEMA = [dbo];
- GRANT CREATE TABLE TO <new user name>;
- GRANT CREATE VIEW TO <new user name>;
- GRANT SELECT, ALTER, INSERT, DELETE, UPDATE on schema::dbo to <new user name> ;
- GRANT REFERENCES to <new user name>;

PostgreSQL Database Requirements

Perform the following steps before you configure PostgreSQL as the repository database:

- Make sure that you use the public schema for PostgreSQL.
- Create the credentials to access the database using the following command: CREATE USER <new user name> WITH PASSWORD '<password>';
- Create the database and assign the ownership to the user name that you created. Use the following command to complete this step: CREATE DATABASE <NEW db NAME> owner=<new user name>;

Note:

- Make sure that you do not create a schema in the new database with the same name as the user name that you created.
- After you connect to the new database, make sure that you run the command `SHOW search_path`. The command must return the value `"$user", public`.

Content Management Service

The Content Management Service manages reference data for data domains that use reference tables. It uses the Data Integration Service to run mappings to transfer data between reference tables and external data sources. When you create the service, you need to associate other application services with it.

The following table summarizes the dependencies for products, services, and databases that are associated with the Content Management Service:

Dependency	Summary
Products	The following products use the Content Management Service: - Data Engineering Quality - Data Privacy Management - Enterprise Data Catalog - Enterprise Data Preparation - Informatica Data Quality - Test Data Management
Services	The Content Management Service requires a direct association with the following services: - Model Repository Service - Data Integration Service

Dependency	Summary
Databases	The Content Management Service uses the following database: - Reference data warehouse. Stores data values for the reference table objects that you define in the Model repository. When you add data to a reference table, the Content Management Service writes the data values to a table in the reference data warehouse.
Installer	You can create the Content Management Service when you run the installer. **Note:** You must create the Content Management Service on the same node as the Data Integration Service.

Reference Data Warehouse Requirements

The reference data warehouse stores the data values for reference table objects that you define in a Model repository. You configure a Content Management Service to identify the reference data warehouse and the Model repository.

You associate a reference data warehouse with a single Model repository. You can select a common reference data warehouse on multiple Content Management Services if the Content Management Services identify a common Model repository. The reference data warehouse must support mixed-case column names.

The reference data warehouse supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- Oracle
- PostgreSQL, using a JDBC driver

Allow 200 MB of disk space for the database.

Note: Ensure that you install the database client on the machine on which you want to run the Content Management Service.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- Verify that the database user account has CREATETAB and CONNECT privileges.
- Verify that the database user has SELECT privileges on the SYSCAT.DBAUTH and SYSCAT.DBTABAUTH tables.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- Set the tablespace pageSize parameter to 32768 bytes.
- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Verify that the database user account has CONNECT and CREATE TABLE privileges.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

  ALTER SEQUENCE

  ALTER TABLE

  CREATE SEQUENCE

  CREATE SESSION

  CREATE TABLE

  CREATE VIEW

  DROP SEQUENCE

  DROP TABLE

- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

PostgreSQL Database Requirements

Use the following guidelines when you set up the repository on PostgreSQL:

- Use a JDBC connection to connect to the PostgreSQL database.

  Informatica installs a DataDirect JDBC driver for PostgreSQL that you can use to connect to the database. Find the driver in the clients/DeveloperClient/infacmd installation directory, and copy the driver to the clients/externaljdbcjars directory .

- Specify the database schema name. Do not leave the schema name blank.

  If the database uses the default PostgreSQL schema name of `public`, you can specify `public` as the schema name.

- Verify that the database user has the CONNECT and CREATE TABLE privileges.

Data Integration Service

The Data Integration Service receives requests from Informatica client tools to run integration, profile, and data preparation jobs. It writes results to different databases, and it writes run-time metadata to the Model repository. When you create the service, you need to associate another application service with it.

The following table lists the dependencies for products, services, and databases that are associated with the Data Integration Service.

Dependency	Summary
Products	The following products use the Data Integration Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Data Privacy Management - Enterprise Data Catalog - Enterprise Data Preparation - Informatica Data Quality - PowerCenter - Test Data Management
Services	The Data Integration Service requires a direct association with the following service: - Model Repository Service
Databases	The Data Integration Service uses the following databases: - Data object cache. Stores cached logical data objects and virtual tables. - Profiling warehouse. Stores profiling information, such as profile and scorecard results. - Workflow database. Stores run-time metadata for workflows.
Installer	You can create the Data Integration Service when you run the installer.

Rules and Guidelines

Consider the following rules and guidelines for Data Integration Service creation:

Informatica recommends creating a dedicated Data Integration Service and a dedicated Model Repository Service for Enterprise Data Preparation.

Data Object Cache Database Requirements

The data object cache database stores cached logical data objects and virtual tables for the Data Integration Service. You specify the data object cache database connection when you create the Data Integration Service.

The data object cache database supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- Oracle

Allow 200 MB of disk space for the database.

Note: Ensure that you install the database client on the machine on which you want to run the Data Integration Service.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- Verify that the database user account has CREATETAB and CONNECT privileges.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- Set the tablespace pageSize parameter to 32768 bytes.
- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Verify that the database user account has CONNECT and CREATE TABLE privileges.

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Verify that the database user account has CONNECT and CREATE TABLE privileges.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

  CREATE INDEX

  CREATE SESSION

  CREATE SYNONYM

  CREATE TABLE

  CREATE VIEW

  DROP TABLE

  INSERT INTO TABLE

  UPDATE TABLE
- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

Profiling Warehouse Requirements

The profiling warehouse database stores profiling and scorecard results. You specify the profiling warehouse connection when you create the Data Integration Service.

The profiling warehouse supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Oracle

Allow 10 GB of disk space for the database.

Note: Ensure that you install the database client on the machine on which you want to run the Data Integration Service. You can specify a JDBC connection as the profiling warehouse connection for IBM DB2 UDB, Microsoft SQL Server, and Oracle database types.

For more information about configuring the database, see the documentation for your database system.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- The database user account must have the CREATETAB, CONNECT, CREATE VIEW, and CREATE FUNCTION privileges.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- Set the tablespace pageSize parameter to 32768 bytes.
- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.

Note: Informatica does not support the partitioned database environment for IBM DB2 databases when you use a JDBC connection as the profiling warehouse connection.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- The database user account must have the CONNECT, CREATE TABLE, CREATE VIEW, and CREATE FUNCTION privileges.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

  ALTER TABLE

  CREATE ANY INDEX

  CREATE PROCEDURE

  CREATE SESSION

  CREATE TABLE

  CREATE VIEW

  DROP TABLE

  UPDATE TABLE

- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

- Set the following parameters to the Informatica recommended values:

Parameter	Recommended Value
open_cursors	4000
Sessions	1000
Processes	1000

Workflow Database Requirements

The Data Integration Service stores run-time metadata for workflows in the workflow database. Before you create the workflow database, set up a database and database user account for the workflow database.

You specify the workflow database connection when you create the Data Integration Service.

The workflow database supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- Oracle
- PostgreSQL

Allow 200 MB of disk space for the database.

Note: Ensure that you install the database client on the machine on which you want to run the Data Integration Service.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- Verify that the database user account has CREATETAB and CONNECT privileges.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- Set the tablespace pageSize parameter to 32768 bytes.
- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.
- Set the connection pooling parameters.

  The following table lists the connection pooling parameters that you must set:

Parameter	Value
Maximum Connection Pool Size	128
Minimum Connection Pool Size	0
Maximum Idle Time	120 seconds

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Verify that the database user account has CONNECT and CREATE TABLE privileges.
- Set the connection pooling parameters.

  The following table lists the connection pooling parameters that you must set:

Parameter	Value
Maximum Connection Pool Size	128
Minimum Connection Pool Size	0
Maximum Idle Time	120 seconds

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

  ALTER TABLE

  ALTER VIEW

  CREATE SEQUENCE

  CREATE SESSION

  CREATE SYNONYM

  CREATE TABLE

  CREATE VIEW

  DROP TABLE

  DROP VIEW

- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.
- Set the connection pooling parameters.

The following table lists the connection pooling parameters that you must set:

Parameter	Value
Maximum Connection Pool Size	128
Minimum Connection Pool Size	0
Maximum Idle Time	120 seconds

PostgreSQL Database Requirements

Use the following guidelines when you set up the repository on PostgreSQL:

- Verify that the database user account has CONNECT, CREATE TABLE, and CREATE VIEW privileges.
- Specify the database schema name when you use PostgreSQL as the database.
- Ensure that PostgreSQL has sufficient disk space for the data files. By default, the data files are present in the following location:

  ```
  <PostgreSQL installation directory>/data
  ```

- On the database, set the configuration parameters.

  The following table lists the minimum and recommended values for the configuration parameters that you must set:

Parameter	Minimum Value	Recommended Value
max_connections	200	4000
shared_buffers	2 GB	16 GB
max_locks_per_transaction	1024	1024
max_wal_size	1 GB	8 GB
checkpoint_timeout	5 minutes	30 minutes

Interactive Data Preparation Service

The Interactive Data Preparation Service manages data preparation within Enterprise Data Preparation. When an analyst prepares data in a project, the Interactive Data Preparation Service stores worksheet metadata in

the Data Preparation repository. When you create the service, you can associate other application services with it.

The following table summarizes the dependencies for products, services, and databases that are associated with the Interactive Data Preparation Service:

Dependency	Summary
Products	The following products use the Interactive Data Preparation Service: - Enterprise Data Preparation
Services	If you plan to use rules during data preparation, you can provide a direct association with the following services: - Model Repository Service - Data Integration Service
Databases	The Interactive Data Preparation Service uses the following database: - Data Preparation repository. Stores worksheet metadata created when users prepare data assets for publication.
Installer	You can create the Interactive Data Preparation Service when you run the installer.

Rules and Guidelines

Consider the following rules and guidelines for Interactive Data Preparation Service creation:

- If you use the installer to create the Enterprise Data Preparation and Interactive Data Preparation Service, you must create both of the application services on the same node.
- If you configure Enterprise Data Preparation to use rules, you must associate a Data Integration Service with the Interactive Data Preparation Service. The Data Integration Service is required to run rules during data preparation.
- If you configure Enterprise Data Preparation to use rules, you must associate the Model Repository Service that manages the Model repository in which rule objects and metadata are stored with the Interactive Data Preparation Service.
- If you want to create and enable the Interactive Data Preparation Service when you run the installer, the domain must contain connections associated with the Hadoop environment. For more information, see "Prepare to Create the Enterprise Data Preparation Services" on page 62.

Data Preparation Repository Database Requirements

The Interactive Data Preparation Service stores recipe and mapping metadata in the repository.

Set up one of the following databases to use as the Data Preparation repository:

- Oracle
- MySQL
- MariaDB

Allow 5 GB of disk space for the repository database. Allocate more space based on the amount of metadata you want to store.

MySQL and MariaDB Database Requirements

You can use a MySQL database or a MariaDB database as the Data Preparation repository.

Set the required system variables on the database server.

- For MySQL version 5.6.26 and higher, set lower_case_table_names=1.
- For MySQL version 5.7 and higher, set explicit_defaults_for_timestamp=1.

Set the same system variable values for a MariaDB database.

Set the required permissions on the MySQL or MariaDB database.

- Create tables and views.
- Drop tables and views.
- Insert, update, and delete data.

Download and copy the MySQL connector .jar file to the node where you create the Interactive Data Preparation Service.

Download the file and copy it to the following directory before you start the installer:

```
$USER_INSTALL_DIR$/services/shared/jars/thirdparty/
```

Make sure the name of the file is in the following format:

```
mysql-connector-java-<versiondetails>.jar
```

You must also set the mysql_connector_jar_path environment variable to the location of the MySQL connector .jar file.

Oracle Database Requirements

You can use an Oracle database as the Data Preparation repository.

Ensure that the database has the following permissions:

- Create tables and views.
- Create sequence, session, and synonyms.
- Drop tables and views.
- Insert, update, and delete data.

Enterprise Data Preparation Service

The Enterprise Data Preparation Service runs the Enterprise Data Preparation application in the Informatica domain. When you create the service, you need to associate other application services with it.

The following table summarizes the dependencies for products, services, and databases that are associated with the Enterprise Data Preparation Service:

Dependency	Summary
Products	The following products use the Enterprise Data Preparation Service: - Enterprise Data Preparation
Services	The Enterprise Data Preparation Service requires a direct association with the following services: - Model Repository Service - Data Integration Service - Interactive Data Preparation Service
Databases	The Enterprise Data Preparation Service uses the following database: - Model repository. Stores mappings generated for published data assets.
Installer	You can create the Enterprise Data Preparation Service when you run the installer.

Rules and Guidelines

Consider the following rules and guidelines for Enterprise Data Preparation Service creation:

- If you create the Enterprise Data Preparation Service and Interactive Data Preparation Service during installation, you must create both of the application services on the same node.
- If you want to create and enable the Enterprise Data Preparation Service when you run the installer, the domain must contain connections associated with the Hadoop environment. For more information, see "Prepare to Create the Enterprise Data Preparation Services" on page 62.

Informatica Cluster Service

The Informatica Cluster Service runs and manages Enterprise Data Catalog and the associated services.

The following table summarizes the dependencies for products, services, and databases that are associated with the Informatica Cluster Service:

Dependency	Summary
Products	The following products use the Informatica Cluster Service: - Data Privacy Management - Enterprise Data Catalog - Enterprise Data Preparation
Services	The Informatica Cluster Service must be associated with the Catalog Service.
Databases	Mongo DB as the metadata store.
Installer	You can create the Informatica Cluster Service when you install Enterprise Data Catalog.

Metadata Access Service

The Metadata Access Service allows the Developer tool to access Hadoop connection information to import and preview metadata from the Hadoop environment. The Metadata Access Service is required for design-time access to the Hadoop environment.

The following table summarizes the dependencies for products, services, and databases that are associated with the Metadata Access Service:

Dependency	Summary
Products	The following products use the Metadata Access Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Enterprise Data Catalog - Enterprise Data Preparation
Services	The Metadata Access Service does not require an association with another application service.
Databases	The Metadata Access Service does not have an associated database.
Installer	You can create the Metadata Access Service when you run the installer.

Model Repository Service

The Model Repository Service manages the Model repository. It receives requests from Informatica clients and application services to store or access metadata in the Model repository.

The following table summarizes the dependencies for products, services, and databases that are associated with the Model Repository Service.

Dependency	Summary
Products	The following products use the Model Repository Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Data Privacy Management - Enterprise Data Catalog - Enterprise Data Preparation - Informatica Data Quality - PowerCenter - Test Data Management
Services	The Model Repository Service does not require an association with another application service.
Databases	The Model Repository Service uses the following database: - Model repository. Stores metadata created by Informatica clients and application services.
Installer	You can create the Model Repository Service when you run the installer.

Model Repository Database Requirements

Informatica services and clients store data and metadata in the Model repository. Configure a monitoring Model repository to store statistics for ad hoc jobs, applications, logical data objects, SQL data services, web services, and workflows. Before you create the Model Repository Service, set up a database and database user account for the Model repository. It is recommended that you use different database configuration for Model repository and monitoring Model repository.

The Model repository supports the following database types:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- Oracle
- PostgreSQL

When you configure Microsoft SQL Server, you can choose to configure the Microsoft Azure SQL Database as the Model repository.

Allow 3 GB of disk space for DB2. Allow 200 MB of disk space for all other database types.

For more information about configuring the database, see the documentation for your database system.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- Specify the tablespace name when you use IBM DB2 as the Model Repository database.
- If the repository is in an IBM DB2 database, verify that IBM DB2 Version 10.5 is installed.
- On the IBM DB2 instance where you create the database, set the following parameters to ON:
  - DB2_SKIPINSERTED
  - DB2_EVALUNCOMMITTED
  - DB2_SKIPDELETED
  - AUTO_RUNSTATS
- On the database, set the configuration parameters.

  The following table lists the configuration parameters that you must set:

Parameter	Value
logfilsiz	8000
maxlocks	98
locklist	50000
auto_stmt_stats	ON

- Set the tablespace pageSize parameter to 32768 bytes.

  In a single-partition database, specify a tablespace that meets the pageSize requirements. If you do not specify a tablespace, the default tablespace must meet the pageSize requirements.

In a multi-partition database, specify a tablespace that meets the pageSize requirements. Define the tablespace in the catalog partition of the database.

- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.
- Verify that the database user has CREATETAB, CONNECT, and BINDADD privileges.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- In the DataDirect Connect for JDBC utility, update the DynamicSections parameter to 3000.

  The default value for DynamicSections is too low for the Informatica repositories. Informatica requires a larger DB2 package than the default. When you set up the DB2 database for the domain configuration repository or a Model repository, you must set the DynamicSections parameter to at least 3000. If the DynamicSections parameter is set to a lower number, you can encounter problems when you install or run Informatica services.

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Note: The guidelines to set up the repository for Azure SQL Database with Active Directory authentication is the same.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Specify the database schema name when you use Microsoft SQL Server as the Model Repository database.
- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.
  To set the isolation level for the database, run the following commands:

  ```
  ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
  ```

  ```
  ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
  ```

  To verify that the isolation level for the database is correct, run the following commands:

  ```
  SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
  ```

  ```
  SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
  ```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Note: The guidelines to set up the repositories for Microsoft Azure SQL Database and Azure SQL Database with Active Directory authentication is the same.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Set the OPEN_CURSORS parameter to 4000 or higher.
  Verify that the database user has the following privileges:

  CREATE SEQUENCE

  CREATE SESSION

  CREATE SYNONYM

  CREATE TABLE

  CREATE VIEW
- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

PostgreSQL Database Requirements

Use the following guidelines when you set up the repository on PostgreSQL:

- Verify that the database user account has CONNECT, CREATE TABLE, and CREATE VIEW privileges.
- Specify the database schema name when you use PostgreSQL as the database.
- Ensure that PostgreSQL has sufficient disk space for the data files. By default, the data files are present in the following location:

  ```
  <PostgreSQL installation directory>/data
  ```
- On the database, set the configuration parameters.

  The following table lists the minimum and recommended values for the configuration parameters that you must set:

Parameter	Minimum Value	Recommended Value
max_connections	200	4000
shared_buffers	2 GB	16 GB
max_locks_per_transaction	1024	1024
max_wal_size	1 GB	8 GB
checkpoint_timeout	5 minutes	30 minutes

Monitoring Model Repository Service

The monitoring Model Repository Service is a Model Repository Service that monitors statistics for Data Integration Service jobs. You configure the monitoring Model Repository Service in the domain properties.

Note: If you want to generate monitoring statistics, you must create a dedicated Model Repository Service for monitoring. You cannot store run-time monitoring statistics in the same repository where you store object metadata.

The following table summarizes the dependencies for products, services, and databases that are associated with the monitoring Model Repository Service:

Dependency	Summary
Products	The following products use the monitoring Model Repository Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Data Privacy Management - Enterprise Data Catalog - Enterprise Data Preparation - Informatica Data Quality - PowerCenter - Test Data Management
Services	The monitoring Model Repository Service does not require an association with another application service.
Databases	The monitoring Model Repository Service uses the following database: - Model repository. Stores run-time monitoring statistics that you can view in the Administrator tool.
Installer	You can create the monitoring Model Repository Service when you run the installer.

Search Service

The Search Service manages searches in the Analyst tool and returns search results from the Model repository. When you create the service, you need to associate another application service with it.

The following table summarizes the dependencies for products, services, and databases that are associated with the Search Service:

Dependency	Summary
Products	The following products use the Search Service: - Data Engineering Integration - Data Engineering Quality - Data Engineering Streaming - Enterprise Data Catalog - Enterprise Data Preparation - Informatica Data Quality - PowerCenter
Services	The Search Service requires a direct association with the following service: - Model Repository Service
Databases	The Search Service is not associated with any database.
Installer	You cannot create the Search Service when you run the installer.

Prepare to Create the Enterprise Data Preparation Services

To create the Enterprise Data Preparation services, the domain must be integrated with the Hadoop environment through a domain cluster configuration object.

The Enterprise Data Preparation services require connections to the environment. The connections are associated with the environment through a cluster configuration. The process to integrate the environments and create the services can vary based on the type of installation you choose.

Install Enterprise Data Preparation with Informatica Domain Services

If you install the Informatica domain services when you install Enterprise Data Preparation and you want to create the Enterprise Data Preparation services, you must provide the cluster information during the installation. The installer can import the cluster configuration from the Hadoop environment, and create the connections required by the Enterprise Data Preparation services.

Before you run the installer, you need the information you need to import the cluster configuration from the Hadoop administrator. The cluster administrator can provide the information to you in one of the following formats:

- Cluster authentication information. The Hadoop administrator can provide you with cluster authentication information to connect to the cluster for the import process.
- Archive file. The Hadoop administrator can provide you an archive file that contains properties from *-site.xml files on the cluster. If you are importing from Amazon EMR, MapR, or Google Dataproc you can import only from an archive file.

Note: When the installation completes, you must fully integrate the domain with the Hadoop environment, including a task to refresh the cluster configuration. If you want to complete all integration tasks at one time, you can skip creating the services during installation and create them manually after you integrate the domain with the Hadoop environment.

Prepare for Archive File Import with a Full Installation

The Hadoop administrator might choose to provide you with a .zip or .tar archive file instead of with direct connection information.

If you are integrating with an Amazon EMR, MapR, or Google Dataproc cluster, you must import the cluster configuration through an archive file.

Get an archive file that contains the following *-site.xml files from the cluster:

- core-site.xml
- hbase-site.xml. Required only if you access HBase sources and targets.
- hdfs-site.xml
- hive-site.xml
- mapred-site.xml or tez-site.xml. Include the mapred-site.xml file or the tez-site.xml file based on the Hive execution type used on the Hadoop cluster.
- yarn-site.xml

Note: Verify that the Hadoop administrator creates an archive file from all the listed *-site.xml files.

After creating the archive file, the Hadoop administrator needs to edit it for the following distributions:

Azure HDInsight

Edit the Hortonworks Data Platform (HDP) version string wherever it appears in the archive file. Search for the string `${hdp.version}` and replace all instances with the HDP version that HDInsight includes in the Hadoop distribution.

Hortonworks HDP

Edit the Hortonworks Data Platform (HDP) version string wherever it appears in the archive file. Search for the string `${hdp.version}` and replace all instances with the HDP version that Hortonworks includes in the Hadoop distribution.

Prepare for Direct Import with a Full Installation

If you want to create the Enterprise Data Preparation services when you perform a full installation, you must import properties from the *-site.xml files into the domain. You can get connection information for the cluster or an archive file from the Hadoop administrator to import cluster configuration from the non-native environment.

The following table describes information that you need from the Hadoop administrator to create the cluster configuration directly from the cluster:

Property	Description
Host	IP address of the cluster manager.
Port	Port of the cluster manager.
User ID	Cluster user ID.
Password	Password for the user.
Cluster name	Name of the cluster. Use the display name if the cluster manager manages multiple clusters. If you do not provide a cluster name, the installer imports information based on the default cluster. **Note:** To find the correct Cloudera cluster name when you have multiple clusters, the Hadoop administrator can add the string `/api/v8/clusters` to the URL and provide you with the name that appears in the browser tab.

Configure Native Connectivity on Service Machines

To establish native connectivity between an application service and a database, install the database client software for the database that you want to access.

Native drivers are packaged with the database server and client software. Configure connectivity on the machines that need to access the databases. To ensure compatibility between the application service and the database, install a client software that is compatible with the database version and use the appropriate database client libraries.

The Data Integration Service uses native database drivers to connect to the following databases:

- Source and target databases. Reads data from source databases and writes data to target databases.
- Data object cache database. Stores the data object cache.
- Profiling source databases. Reads from relational source databases to run profiles against the sources.

- Profiling warehouse. Writes the profiling results to the profiling warehouse.
- Reference tables. Runs mappings to transfer data between the reference tables and the external data sources.

When the Data Integration Service runs on a single node or on primary and back-up nodes, install database client software and configure connectivity on the machines where the Data Integration Service runs.

When the Data Integration Service runs on a grid, install database client software and configure connectivity on each machine that represents a node with the compute role or a node with both the service and compute roles.

Install Database Client Software

You must install the database clients on the required machines based on the types of databases that the application services access.

To ensure compatibility between the application service and the database, use the appropriate database client libraries and install a client software that is compatible with the database version.

Install the following database client software based on the type of database that the application service accesses:

IBM DB2 Client Application Enabler (CAE)

Configure connectivity on the required machines by logging in to the machine as the user who starts Informatica services.

Microsoft SQL Server 2014 Native Client

Download the client from the following Microsoft website: http://www.microsoft.com/en-in/download/details.aspx?id=42295.

Oracle client

Install compatible versions of the Oracle client and Oracle database server. You must also install the same version of the Oracle client on all machines that require it. To verify compatibility, contact Oracle.

Sybase Open Client (OCS)

Install an Open Client version that is compatible with the Sybase ASE database server. You must also install the same version of Open Client on the machines hosting the Sybase ASE database and Informatica. To verify compatibility, contact Sybase.

Configure Database Client Environment Variables

Configure database client environment variables on the machines that run the Data Integration Service processes.

After you configure the database environment variables, you can test the connection to the database from the database client.

The following table lists the database environment variables you need to set:

Database	Environment Variable Name	Database Utility	Value
Oracle	ORACLE_HOME PATH LD_LIBRARY_PATH TNS_ADMIN INFA_TRUSTSTORE	sqlplus	Set to: <*Client InstallDatabasePath*> Add: <*DatabasePath*>/bin and `USER_INSTALL_DIR/server/bin:$PATH` Set to: `$Oracle_HOME/lib` and `USER_INSTALL_DIR/server/bin:$LD_LIBRARY_PATH` Set to location of the tnsnames.ora file: `$ORACLE_HOME/network/admin` For default SSL domain, add to: USER_INSTALL_DIR/services/shared/security For custom SSL domain, set INFA_TRUSTSTORE and INFA_TRUSTSTORE_PASSWORD
IBM DB2	DB2DIR DB2INSTANCE PATH	db2connect	Set to: <*database path*> Set to: <*DB2InstanceName*> Add: <*database path*>/bin
Sybase ASE	SYBASE15 SYBASE_ASE SYBASE_OCS PATH	isql	Set to: <*database path*>/sybase<*version*> Set to: ${SYBASE15}/ASE-<*version*> Set to: ${SYBASE15}/OCS-<*version*> Add: ${SYBASE_ASE}/bin:${SYBASE_OCS}/bin:$PATH
PostgreSQL	PGSERVICEFILE PGHOME PATH LD_LIBRARY_PATH INFA_TRUSTSTORE		Set to the location of the pg_service.conf file: *<pg_service.conf file directory>/pg_service.conf* Set to: */usr/pgsql -10* Add to: *$PGHOME:${PATH}* Add to: *$PGHOME/lib:${LD_LIBRARY_PATH}* For default SSL domain, add to: <InstallationDirectory>/services/shared/security For custom SSL domain, set INFA_TRUSTSTORE and INFA_TRUSTSTORE_PASSWORD
SQL Server	ODBCHOME ODBCINI ODBCINST PATH LD_LIBRARY_PATH INFA_TRUSTSTORE		Set to: *USER_INSTALL_DIR/ODBC7.1* Set to: *$ODBCHOME/odbc.ini* Set to: *$ODBCHOME/odbcinst.ini* Add to: */opt/mssql-tools/bin:$PATH $PATHUSER_INSTALL_DIR/ODBC7.1:$PATHUSER_INSTALL_DIR/server/bin:$PATH* Add to: *$ODBCHOME/lib* *USER_INSTALL_DIR/server/bin:$LD_LIBRARY_PATH* For default SSL domain, add to: USER_INSTALL_DIR/services/shared/security For custom SSL domain, set INFA_TRUSTSTORE and INFA_TRUSTSTORE_PASSWORD

CHAPTER 4

Prepare for Enterprise Data Catalog Deployment

This chapter includes the following topics:

Checklist to Prepare for Enterprise Data Catalog Deployment

You can deploy Enterprise Data Catalog on a single data node or alternatively on three or six data nodes in a cluster.

This chapter contains tasks that you must complete before you deploy Enterprise Data Catalog.

Complete the following tasks:

- ❑ Understand about Enterprise Data Catalog deployment and the Informatica Cluster Service that runs and manages Enterprise Data Catalog.
- ❑ Understand about services associated with Informatica Cluster Service.
- ❑ Complete prerequisites and prepare the environment.

Deployment Planning

For an external cluster, you can plan to deploy Enterprise Data Catalog on one, three, or six nodes.

Deployment of Enterprise Data Catalog involves creating and configuring the Informatica Cluster Service. The Informatica Cluster Service runs and manages your Enterprise Data Catalog deployment. If you plan to deploy Enterprise Data Catalog on more than one node, high-availability is enabled by default to provide resiliency, fail over, and recovery for the Informatica Cluster Service.

The Informatica Cluster Service uses a set of applications and associated services bundled with the installer to manage and run Enterprise Data Catalog.

When you configure the Informatica Cluster Service, you can specify one, three, or six nodes on which the associated applications and services can run. These nodes are referred to as the data nodes in the cluster. The nodes on which you want to run profiling jobs or metadata scan jobs are referred to as processing nodes in the cluster. There are no restrictions on the number of processing nodes that you can configure.

Important: For the associated application services, you cannot configure more than three service instances for a deployment.

Informatica Cluster Service and Associated Services

The Informatica Cluster Service uses the following services to run and manage Enterprise Data Catalog:

Service	Description
Mongo DB	Service to manage the Mongo DB database used as the metadata store.
Nomad	Service to manage the Nomad orchestration service.
Apache Solr	Service to manage Apache Solr used to index and search catalog assets.
ZooKeeper	Service to manage ZooKeeper used as the service co-ordination application
PostgreSQL	Service to manage PostgreSQL database used as the store for similarity profiling data.

Common Operating System Prerequisites

You can install Enterprise Data Catalog on a machine that runs on Red Hat Enterprise Linux Server or SUSE Linux Enterprise Server. The prerequisites in this section are common for Red Hat Enterprise Linux and SUSE Linux Enterprise Server.

Operating System Configuration Prerequisites

Make sure that you verify the following operating system prerequisites:

- `Bash` is the default shell.
- Verify that ntpd is synchronized between the Informatica domain node and the cluster nodes.

- The Linux base repository is set up and working.
- Set the soft limit and hard limit for max user processes and file descriptors to 65,000 or more for the machines where you plan to host the Informatica domain.
- Set the soft limit and hard limit for open file descriptor to 65,000 or more for the machines where you deploy the Informatica Cluster Service.
- Set UMASK to 022 (0022).

Applications and Utility Prerequisites

Verify that the following applications and utilities are installed:

- JDK 1.8
- Bash shell
- xz-libs
- systemctl
- rsync
- netstat
- YUM
- Zypper
- scp
- curl
- rpm
- zip
- unzip
- tar
- wget
- libcurl
- nslookup
- md5sum
- ping
- ifconfig
- cksum
- dnsdomainname
- OpenSSL version 1.0.1e-30.el6_6.5.x86_64 or later. OpenSSL 3.0 is not supported.

  Note: Verify that the $PATH variable points to the `/usr/bin` directory to use the correct version of Linux OpenSSL.

Directory Prerequisites

Configure the following directory prerequisites:

root directory (/)

The root directory (/) must have a minimum of 10 GB of free disk space. If you plan to create the data directory for Informatica Cluster Service in the root directory, verify that the root directory has a minimum of 50 GB additional free disk space available. If you want to configure a separate directory for the Informatica Cluster Service log files, verify that the directory has a minimum of 50 GB of free disk space.

If you configure the `workingDir` to `/`, validate if the file system mounted on `/tmp` and `/var` directories have the `EXEC` flag set.

If the `workingDir` is not configured to `/`, validate if the `workingDir` directory has `read`, `write`, and `execute` permissions configured. Validate if the `EXEC` flag is set for the directory.

The directory must not have the read, write, and execute permissions configured.

/var directory

The directory must not have the write permission for everyone.

The directory must have a minimum of 2 GB of free disk space.

The directory must not have the read, write, and execute permissions configured.

/tmp directory

The directory must have the read, write, and execute permissions configured.

/usr directory

The directory must have a minimum of 2 GB of free disk space.

Custom directory

Verify that a minimum of 60 GB of free disk space is available in any custom directory that you plan to configure.

DNS Prerequisites

Verify the following DNS prerequisites in the `/etc/hosts` file for all the nodes in the cluster:

- An entry for the loopback address, `127.0.0.1 localhost localhost.domain.com` in the file.
- Each machine in the cluster includes the `127.0.0.1 localhost localhost.localdomain` entry in the file.
- The file includes the fully-qualified host names for all the cluster nodes. Alternatively, make sure that reverse DNS lookup returns the fully-qualified host names for all the cluster nodes.

User Account Prerequisites

Verify the following prerequisites for all the user accounts:

- Create a user account without root privileges and ensure that the user has sudo privileges.

  Note: The services associated with the Informatica Cluster Service such as Apache Solr, PostgreSQL, and Nomad require a non-root user with sudo privileges to run the services. Sudo allows a user to run programs or commands with elevated privileges for a specific time frame. Enterprise Data Catalog requires a non-root user with sudo permissions to run certain commands when performing administrative tasks such as installation, upgrade, and service status monitoring.

- Update the sudoers file. Certain commands require sudo privileges for the gateway user when you enable the Informatica Cluster Service for the first time. Ensure that the user has sudo privileges for the `mkdir, chown, chmod, echo, systemctl, cp, mv, sysctl, rm` commands.

  Note: You must configure sudo permissions for the commands if you plan to change the properties for the Informatica Cluster Service or replace the SSL certificates configured for the Informatica Cluster Service.

  To configure sudo privileges for the commands, you must add the commands to the `/etc/sudoers` file as shown in the following example:

  ```
  %<Gateway user name> ALL =(ALL) NOPASSWD: /bin/mkdir, /bin/chown, /bin/chmod, /bin/
  echo, /bin/systemctl, /bin/cp, /bin/mv, /usr/sbin/sysctl, /bin/rm
  ```

  Note: You can determine the directory where each command is located using the which <command name> command.

After you enable the Informatica Cluster Service for the first time, you can choose to disable the validation for sudo permissions for the Informatica Cluster Service. To disable the validation, configure the following custom property for the Informatica Cluster Service: `IcsCustomOptions.IcsGatewayUserSudoEnabled` and set the value to false. After you disable the validation, sudo privileges are not required when you restart the Informatica Cluster Service. However, to shut down the Informatica Cluster Service, sudo permission is still required for the `systemctl` command.

To run the `infacmd ics cleanCluster` command to clean the Informatica Cluster Service, sudo permission must be configured for the `systemctl` and `rm` commands.

- Add the following entries in the `/etc/security/limits.d/20-nproc.conf` file for the root user:
  - soft nproc 65000
  - hard nproc 65000
  - soft nofile 65000
  - hard nofile 65000
- Add the following entries in the `/etc/security/limits.d/20-nproc.conf` file for the non-root user:
  - <non root user name> soft nproc 65000
  - <non root user name> hard nproc 65000
  - <non root user name> soft nofile 65000
  - <non root user name> hard nofile 65000
- For the non-root user account with sudo privileges that you use to install Enterprise Data Catalog, configure the following ulimit values:
  - -f (file size): unlimited
  - -t (cpu time): unlimited
  - -v (virtual memory): unlimited
  - -l (locked-in-memory size): unlimited
  - -n (open files): 64000
  - -m (memory size): unlimited
  - -u (processes/threads): 64000
- Disable the password prompt from the domain host to the cluster gateway host and from the cluster gateway host to all the agent nodes.
- Verify that the gateway user has the required privileges to run the ping command.
- Disable the password prompt for the gateway user.
- If you use a user account without root privileges and if you want to remove sudo access, comment `defaults requiretty` in `/etc/sudoers` file.

Port Prerequisites

Verify that the following ports are available:

Service	Default Port
HTTP/HTTPS	9075
Nomad Serf	4648
Nomad HTTP	4646

Service	Default Port
Nomad RPC	4647
ZooKeeper	2181
ZooKeeper peer	2888
ZooKeeper leader	3888
Solr	8983
Mongo DB is not configured as a shard member or configuration server.	27017
Mongo DB is configured as a shard member	27018
Mongo DB is configured as a configuration server.	27019
PostgreSQL	5432

Operating System Prerequisites for Red Hat Enterprise Linux

Verify the following prerequisites for a Red Hat Linux Enterprise Server if you plan to install Enterprise Data Catalog on a Red Hat Enterprise Linux Server:

Operating System	Prerequisite
Red Hat Enterprise Linux	- Sudo version 1.8.16 or later. - Install openssl version v1.01 build 16 or later or v1.0.2k. - Verify that the `/etc/sysconfig/network` directory exists and configure read permission for the directory. - Verify that `/etc/sysconfig/network` includes the same entry as the entry configured for hostname -f. - Disable SSL certificate validation. - For RHEL 6.x, install lsb_release. - For RHEL 8.x, install ncurses-c++-libs and ncurses-compat-libs. - For RHEL 8.3, install libidn.so.11.

See the Informatica Product Availability Matrix for 10.5.1 for more details.

Operating System Prerequisites for SUSE Linux Enterprise Server

Verify the following prerequisites for a SUSE Linux Enterprise Server if you plan to install Enterprise Data Catalog on a SUSE Linux Enterprise Server:

Operating System	Prerequisite

SUSE Linux Enterprise Server	- Install netcat-openbsd.
	- Verify that the `/etc/HOSTNAME` directory exists and configure read permission for the directory.
	- Verify that the `/etc/HOSTNAME` directory includes the same entry as the entry configured for hostname -f
	- Install the following RPM Package Manager (RPMs) on all the cluster nodes:
	  - openssl-1.0.1c-2.1.3.x86_64.rpm
	  - libopenssl1_0_0-1.0.1c-2.1.3.x86_64.rpm
	  - libopenssl1_0_0-32bit-1.0.1c-2.1.3.x86_64.rpm
	- Install libncurses5.
	Note: The pre-validation utility does not validate this prerequisite.
	- Do not install libsnappy if you install Enterprise Data Catalog on SUSE Linux Enterprise Server.

See the Informatica Product Availability Matrix for more details.

Node Prerequisites

Host Node Prerequisites

Verify the following prerequisites for the host nodes:

- Disable the firewall on each host in the cluster.
- Enable passwordless SSH between the following nodes:
  - Node that hosts the Informatica domain and node that hosts the gateway.
  - Gateway node and all data nodes and processing nodes.
  - Backup nodes and gateway node.
  - All nodes in the cluster and data nodes.

Cluster Node Prerequisites

Verify that the cluster nodes meet the following minimum requirements per node:

Requirement	Value
CPU	4
Unused memory	12 GB

Requirement	Value
Total memory	16 GB
Disk space	60 GB

CHAPTER 5

Record Information for Installer Prompts

This chapter includes the following topics:

Checklist to Record Installer Prompts

This chapter contains information that you need to enter when you run the installer. Use this checklist to track the recording tasks before you run the installer.

- ❑ Record the names of nodes that you want to create and the services that you want to create on each node.
- ❑ Record basic database information for each database associated with a service that you are creating.
- ❑ If the domain configuration and Model repository databases are secure, record the JDBC connection string with required security parameters.
- ❑ Record the site key for the installer.
- ❑ If you want to enable Kerberos authentication when you run the installer, record Kerberos information for each node in the domain.

Record Information for Installer Prompts Overview

When you install the Informatica services, you need to know information about the domain, nodes, application services, and databases that you plan to create.

This section lists information that you need to provide when you run the installer. Informatica recommends recording installer prompts before you start the installation process. For example, you might want to create a text file of information so you can copy into the installer.

Domain Object Naming Conventions

You cannot change domain, node, and application service names. Use names that continue to work if you migrate a node to another machine or if you add additional nodes and services to the domain. In addition, use names that convey how the domain object is used. Naming conventions are provided in applicable topics.

Domain

When you create a domain, you must provide a domain name and gateway node name.

The following table describes the domain information that you need to enter during the installation process:

Domain Information	Description
Domain name	Name of the domain that you plan to create. The name must not exceed 128 characters and must be 7-bit ASCII only. It cannot contain a space or any of the following characters: ` % * + ; " ? , < > \ / Consider one of the following naming conventions: DMN, DOM, DOMAIN, _<ORG>_<ENV>
Master gateway node host name	Fully qualified host name of the machine on which to create the master gateway node. If the machine has a single network name, use the default host name. The node host name cannot contain the underscore (_) character. If the machine has multiple network names, you can modify the default host name to use an alternate network name. If the machine has a single network name, use the default host name. **Note:** Do not use localhost. The host name must explicitly identify the machine.
Master gateway node name	Name of the master gateway node that you plan to create on this machine. The node name is not the host name for the machine. Consider the following naming convention: Node<node##>_<ORG>_<optional distinguisher>_<ENV>

Nodes

When you install the Informatica services, you add the installation machine to the domain as a node. You can add multiple nodes to a domain.

The following table describes the node information that you need to enter when you join a domain.

Node Information	Description
Node host name	Fully qualified host name of the machine on which to create nodes. If the machine has a single network name, use the default host name. The node host name cannot contain the underscore (_) character. If the machine has multiple network names, you can modify the default host name to use an alternate network name. If the machine has a single network name, use the default host name. **Note:** Do not use localhost. The host name must explicitly identify the machine.
Node name	Name of the nodes that you plan to create on this machine. The node name is not the host name for the machine. Consider the following naming convention: Node<node##>_<ORG>_<optional distinguisher>_<ENV>

Application Services

Record the application service names and the nodes where you want to create them.

The following table lists the application services that you can create when you run the installer:

Application Service	Naming Convention
Catalog Service	CS_<ORG>_<ENV>
Content Management	CMS_<ORG>_<ENV>
Data Integration Service	DIS_<ORG>_<ENV>
Data Privacy Management Service	DPM_<ORG>_<ENV>
Interactive Data Preparation Service	DPS_<ORG>_<ENV>
Enterprise Data Preparation Service	EDLS_<ORG>_<ENV>
Metadata Access Service	MAS_<ORG>_<ENV>
Informatica Cluster Service	ICS_<ORG>_<ENV>
Model Repository Service	MRS_<ORG>_<ENV>
monitoring Model Repository Service	mMRS_<ORG>_<ENV>

Application Service	Naming Convention
PowerCenter Repository Service	PCRS, RS _<ORG>_<ENV>
PowerCenter Integration Service	PCIS, IS _<ORG >_<ENV>

For more information about all service naming conventions, see the following Informatica Velocity Best Practice article available on the Informatica Network: Velocity Naming Conventions

Databases

When you plan the installation, you also need to plan the required relational databases. The domain requires a database to store configuration information and user account privileges and permissions. Some application services require databases to store information processed by the application service.

Domain

The following table describes the information that you need to enter during the installation process:

Database Information	Description
Domain configuration database type	Database type for the domain configuration repository. The domain configuration repository supports IBM DB2 UDB, Microsoft SQL Server, Oracle, PostgreSQL, or Sybase ASE.
Domain configuration database host name	The name of the machine hosting the database.

Content Management Service

The following table describes the information that you need to enter during the installation process:

Database Information	Description
Reference data warehouse database type	Database type for the reference data warehouse. The reference data warehouse supports IBM DB2 UDB, Microsoft Azure SQL Database, Microsoft SQL Server, Oracle, or PostgreSQL.
Reference data warehouse database host name	The name of the machine hosting the database.

Data Integration Service

The following table describes the information that you need to enter during the installation process:

Database Information	Description
Data object cache database type	Database type for the data object cache database. The data object cache database supports IBM DB2 UDB, Microsoft SQL Server, or Oracle.
Data object cache database host name	The name of the machine hosting the database.
Profiling warehouse database type	Database type for the profiling warehouse. The profiling warehouse supports IBM DB2 UDB, Microsoft SQL Server, or Oracle.
Profiling warehouse database host name	The name of the machine hosting the database.
Workflow database type	Database type for the workflow database. The workflow database supports IBM DB2 UDB, Microsoft Azure SQL Database, Microsoft SQL Server, Oracle, or PostgreSQL.
Workflow database host name	The name of the machine hosting the database.

Data Preparation Service

The following table describes the information that you need to enter during the installation process:

Database Information	Description
Data Preparation repository database type	Database type for the Data Preparation repository. The Data Preparation repository supports MySQL, MariaDB, or Oracle.
Data Preparation repository database host name	The name of the machine hosting the database.

Model Repository Service

The following table describes the information that you need to enter during the installation process:

Database Information	Description
Model repository database type	Database type for the Model repository. The Model repository supports IBM DB2 UDB, Microsoft SQL Server, PostgreSQL, or Oracle.
Model repository database host name	The name of the machine hosting the database.

Connection String to a Secure Database

If you create a repository on a secure database, you must provide the truststore information for the database and a JDBC connection string that includes the security parameters for the database.

During installation, you can create the domain configuration repository in a secure database. You can also create the Model repository and PowerCenter repository in a secure database.

You can configure a secure connection to the following databases:

- IBM DB2 UDB
- Microsoft SQL Server
- Microsoft Azure SQL Database
- PostgreSQL
- Oracle

Note: You cannot configure a secure connection to a Sybase database.

When you configure the connection to the secure database, you must specify the connection information in a JDBC connection string. In addition to the host name and port number for the database server, the connection string must include security parameters.

The following table describes the security parameters that you must include in the JDBC connection string:

Parameter	Description
EncryptionMethod	Required. Indicates whether data is encrypted when transmitted over the network. This parameter must be set to `SSL`.
ValidateServerCertificate	Optional. Indicates whether Informatica validates the certificate that is sent by the database server. If this parameter is set to True, Informatica validates the certificate that is sent by the database server. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to false, Informatica does not validate the certificate that is sent by the database server. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Optional. Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate. If SSL encryption and validation is enabled and this property is not specified, the driver uses the server name specified in the connection URL or data source of the connection to validate the certificate.
cryptoProtocolVersion	Required. Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to `cryptoProtocolVersion=TLSv1.1` or `cryptoProtocolVersion=TLSv1.2` based on the cryptographic protocol used by the database server.

You can use the following syntax in the JDBC connection string to connect to a secure database:

IBM DB2

```
jdbc:Informatica:db2://<host name>:<port number>;DatabaseName=<database
name>;EncryptionMethod=SSL;HostNameInCertificate=<database host
name>;ValidateServerCertificate=<true or false>
```

Oracle

```
jdbc:Informatica:oracle://<host name>:<port number>;ServiceName=<service
name>;EncryptionMethod=SSL;HostNameInCertificate=<database host
name>;ValidateServerCertificate=<true or false>
```

Use the following connection string to connect to the Oracle database through the Oracle Connection Manager:

```
jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora
file>;TNSServerName=<TNS server name>;
```

Microsoft SQL Server

```
jdbc:Informatica:sqlserver://<host name>:<port
number>;SelectMethod=cursor;DatabaseName=<database
name>;EncryptionMethod=SSL;HostNameInCertificate=<database host
name>;ValidateServerCertificate=<true or false>
```

Microsoft Azure SQL

```
jdbc:Informatica:sqlserver://<host name>:<port
number>;SelectMethod=cursor;DatabaseName=<database
name>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.win
dows.net;ValidateServerCertificate=false
```

PostgreSQL

```
jdbc:Informatica:postgresql://<host name>:<port number>;DatabaseName=<database
name>;EncryptionMethod=SSL;HostNameInCertificate=<database host
name>;ValidateServerCertificate=<true or false>
```

Note: The installer does not validate the connection string. Verify that the connection string contains all the connection parameters and security parameters required by your database.

Cluster Configuration

You import configuration properties from the non-native cluster to create a cluster configuration. The cluster configuration enables the Data Integration Service to push jobs to the non-native environment.

You can import the properties from an archive file that the Hadoop administrator creates, or you can import the properties directly from the cluster. When you create the cluster configuration, you can also choose to create Hadoop, Hive, HBase, HDFS, or Databricks connections associated with the cluster. The installer appends the connection type to the cluster configuration name to create each connection name.

The following table describes the initial information that you need to enter during the installation process:

Cluster Information	Description
Cluster configuration name	Name of the cluster configuration to create.
Distribution type	Type of non-native cluster distribution.
Cluster configuration import method	Method to import the cluster configuration. You can choose to import the cluster configuration from an archive file or from the cluster.

Import Cluster Configuration from an Archive File

To import the cluster configuration properties from an archive file, specify the path of the configuration archive file.

Import Cluster Configuration from the Cluster

The following table describes the cluster properties for Cloudera, Hortonworks, or Azure HDInsight that you need to enter when you import from cluster during the installation process:

Property	Description
Host	The host name or IP address of the cluster manager.
Port	Port of the cluster manager.
User ID	Cluster user name.
Password	Password for the cluster user.
Cluster Name	Name of the cluster. Use the display name if the cluster manager manages multiple clusters. If you do not provide a cluster name, the wizard imports information based on the default cluster.
Engine type	For a Cloudera cluster, the installer prompts for the engine type. If you are on a CDP cluster, accept the default engine type of Tez. If you are on a CDH cluster, set the engine type to MRv2.

The following table describes the cluster properties for Databricks that you need to enter during the installation process:

Property	Description
Databricks domain	URL of the Databricks cluster.
Databricks token ID	Token ID of the Databricks cluster.
Databricks cluster ID	Cluster ID of the Databricks cluster.

Secure Data Storage

When you install the Informatica services, you must back up the site key that the installer generates and ensure that you save the site key. If you lose the site key, you cannot generate the site key again.

Use the following table to record the information that you need to configure secure data storage:

Property	Description
Encryption key directory	Directory in which to store the encryption key for the domain. By default, the encryption key is created in the following directory: `<Informatica installation directory>/isp/config/keys`.
Specify if you want to back up the site key that the installer generates or not:	Specify if you want to back up the site key that the installer generates or not: - Select **1** for No. If you choose No, the installer exits. - Select **2** for Yes. If you choose Yes, you agree to back up the file manually. A unique site key is generated. If you lose the site key, you cannot generate the site key again. Make sure that you save a copy of this key and do not share the unique site key with others.

CHAPTER 6

Configure Custom SSL Certificates

This chapter includes the following topics:

Configure Custom SSL Certificates (Optional)

During an install or an upgrade, you can use SSL certificates of your choice, referred to as custom SSL certificates, to secure the Informatica Cluster Service. You can also choose to use default SSL certificates as the client and cluster certificates to secure the Informatica Cluster Service.

Generate CA-signed SSL Certificates

You have a custom SSL certificate for the Informatica domain, and plan to configure custom SSL for Enterprise Data Catalog. In this scenario, you must generate CA-signed certificates.

Use the `generate_csr.sh` and `generate_certs.sh` scripts to generate the Certificate Signing Request (CSR) to send to a CA and generate the required custom SSL certificates. You can download the scripts from the Akamai Download Manager.

Perform the following steps to use the scripts to generate the custom SSL certificates:

1. Extract the generate_csr.sh and generate_certs.sh scripts from the following location: `<Location of installer files>/properties/utils/CustomSSLScriptsUtil_ExternalCA`.
2. Set the JAVA_HOME environment variable to point to JDK 8.

3. In the gen_csr.properties file, provide the values for the following parameters:

Parameter	Description
InfaDomainKeystorePassword	The Informatica domain keystore password in plain text.
ServerHosts	The Informatica Cluster Service hosts that include the data nodes, processing nodes, and gateway node. Enter a comma-separated list of FQDNs of cluster nodes.
ClientHosts	Comma-separated list of unique host names of domain nodes and cluster nodes.
InfaDomainName	The Informatica domain name.
ICSServiceName	The name of the Informatica Cluster Service.
KeysOutputDir	The directory to store the generated keys. Specify the $CUSTOM KEYSTORE LOC directory to avoid the additional steps to copy the generated keys. The $ICS SERVICENAME/client_certs and the $ICS SERVICENAME/cluster_certs directories are created under the $CUSTOM KEYSTORE LOC directory. $CUSTOM KEYSTORE LOC is the directory where the custom keystore for the Informatica domain (infa_keystore.jks) is located. $ICS SERVICENAME is the name of the Informatica Cluster Service.
CertsOutputDir	The $CUSTOM TRUSTSTORE LOC directory to store the generated truststore files. The $ICS SERVICENAME/client_certs and the $ICS SERVICENAME/cluster_certs directories are created under the $CUSTOM TRUSTSTORE LOC directory. The $CUSTOM TRUSTSTORE LOC is the directory where the custom truststore for the Informatica domain (infa_truststore.jks) is located.
DNSDomainName	The DNS domain name for the cluster nodes.
ClusterCert_OrganizationUnit	Optional. The value for the OrganizationUnit for the cluster nodes certificate.
ClusterCert_Organization	The value for the Organization for the cluster nodes certificate. **Note:** Verify that the combination of the Organization (O) and the Organizational Unit (OU) parameters in the certificate subject is distinct for the cluster and client certificates
ClusterCert_Location	The value for the Location for the cluster nodes certificate.
ClusterCert_State	The value for the State for the cluster nodes certificate.
ClusterCert_CountryCode	The value for the Country Code for the cluster nodes certificate.
DomainCert_OrganizationUnit	Optional. The value for the Organization Unit for the domain nodes certificate.
DomainCert_Organization	The value for the Organization for the domain nodes certificate. **Note:** Verify that the combination of the Organization (O) and the Organizational Unit (OU) parameters in the certificate subject is distinct for the cluster and client certificates

Parameter	Description
DomainCert_Location	Optional. The value for the Location for the domain nodes certificate. Default is the ClusterCert_Location parameter.
DomainCert_State	Optional. The value for the State for the domain nodes certificate. Default is the ClusterCert_State parameter.
DomainCert_CountryCode	Optional. The value for the Country Code for the domain nodes certificate. Default is the ClusterCert_CountryCode parameter.
Custom_Server_Certificate_CN	Optional. The value for the Common Name in the cluster nodes certificate that can be used instead of the default $InfaDomainName-$ICSServiceName.$DNSDomainName value. **Note:** You must enter RFC2253 compliant values. The following special characters are supported: `, + " \ < > ;`. Double quotes (") must be used in pairs. The characters \ and " must not be used together. The value cannot contain a space.
Custom_Client_Certificate_CN	Optional. The value for the Common Name in the domain nodes certificate that can be used instead of the default $InfaDomainName-$ICSServiceName.$DNSDomainName value. Default is the Custom_Server_Certificate_CN parameter. **Note:** You must enter RFC2253 compliant values. The following special characters are supported: `, + " \ < > ;`. Double quotes (") must be used in pairs. The characters \ and " must not be used together. The value cannot contain a space.

Note: If the values contain spaces or special characters, you must enclose the values within double-quotes.

4. Run the generate_csr.sh script using the following command to generate the .csr files to send to an external CA: `./generate_csr.sh gen_csr.properties`
   The following files are generated for the cluster:

   infa_nodecert.csr

   infa_privkey.key

   infa_privkey.pem

   keystore.jks

   The following files are generated for the client:

   infa_nodecert.csr

   infa_privkey.key

   infa_privkey.pem

   keystore.jks

   browser_cert.csr

   browser_keystore.jks

   browser_privkey.key

   browser_privkey.pem
5. Validate the contents of the .csr files. Run the following command to view the contents: `keytool -printcertreq -file $PATH TO CSR`

6. Send the following .csr files to an external CA for signing:
   - <CertsOutputDir>/<ICSServiceName>/client_certs/infa_nodecert.csr
   - <CertsOutputDir>/<ICSServiceName>/cluster_certs/infa_nodecert.csr
   - <CertsOutputDir>/<ICSServiceName>/client_certs/browser_cert.csr

   Note: The browser_cert.csr file is required if you want to create the browser certificates to view the scan job logs on Nomad.
7. After you receive the certificates or certificate chains from the CA in .pem format, "Validate the CA-signed Certificates" on page 87 and store the certificates to a location under the $INFA HOME directory on your machine.

   Note: If you receive the certificates or certificate chains from the CA in .cer format, run the following command to convert the files to pem format: `openssl x509 -inform der -in <certificate file name>.cer -outform pem -out <certificate file name>.pem`.

   If you receive a certificate chain from the CA, you must extract the root certificate, intermediate certificates, and the end user certificate.
8. In the gen_certs.properties file, provide the values for the following parameters:

Parameter	Description
InfaDomainKeystorePassword	The Informatica domain keystore password in plain text.
InfaDomainTruststorePassword	The Informatica domain truststore password in plain text.
ClusterCertificate	The path to the cluster certificate signed by the CA in .pem format. This is an end user certificate.
ClientCertificate	The path to the client certificate signed by the CA in .pem format. This is an end user certificate.
BrowserCertificate	The path to the browser certificate signed by the CA in .pem format.
ICSServiceName	The name of the Informatica Cluster Service.
IsCACertificateChainAvailable	Specify if the CA certificate chain is available as a single .pem file. Enter true or false. **Note:** The certificate chain must contain only the root and intermediate certificates.
SingleCACertificateChain	The path to the CA certificate chain in .pem format.
IndividualCertificatesFromCAChain	Optional. Only required if the IsCACertificateChainAvailable parameter is set to false. Comma-separated paths to the public certificates in the CA certificate chain in .pem format if the complete CA certificate chain is available as individual .pem files.

Parameter	Description
KeysOutputDir	The $CUSTOM KEYSTORE LOC directory store the generated keys. The $ICS SERVICENAME/client_certs and $ICS SERVICENAME/cluster_certs directories are created under the $CUSTOM KEYSTORE LOC directory. $CUSTOM KEYSTORE LOC is the directory where the custom keystore for the Informatica domain (infa_keystore.jks) is located. $ICS SERVICENAME is the name of the Informatica Cluster Service
CertsOutputDir	The $CUSTOM TRUSTSTORE LOC directory to store the generated truststore files. The $ICS SERVICENAME/client_certs and the $ICS SERVICENAME/cluster_certs directories are created under the $CUSTOM TRUSTSTORE LOC directory. The $CUSTOM TRUSTSTORE LOC is the directory where the custom truststore for the Informatica domain (infa_truststore.jks) is located.

9. Run the generate_certs.sh script using the following command to generate the certificates: `./generate_certs.sh gen_certs.properties`
   The keystore.jks keystore and the infa_privkey.pem private keys are stored at $CUSTOM KEYSTORE LOC/$ICSServiceName/client_certs and $CUSTOM KEYSTORE LOC/$ICSServiceName/cluster_certs directories.

   The truststore.jks truststore and the infa_nodecert.pem, infa_nodecertkey.pem, and infa_pubcert.pem public keys are stored at $CUSTOM TRUSTSTORE LOC/$ICSServiceName/client_certs and $CUSTOM TRUSTSTORE LOC/$ICSServiceName/cluster_certs directories.
10. Optional. The directories $CUSTOM KEYSTORE LOC and $CUSTOM TRUSTSTORE LOC are generally the same. If the <KeysOutputDir> location is not the same as $CUSTOM KEYSTORE LOC and <CertsOutputDir> location is not the same as $CUSTOM TRUSTSTORE LOC, move the keys and certificates to the respective directories.

    Note: Verify that the $CUSTOM KEYSTORE LOC and the $CUSTOM TRUSTSTORE LOC directories have the required user privileges. Also, validate that the user has minimum chmod 700 permissions configured for the directories and chmod 600 permissions configured for the files that are copied to the directories.

To access the Nomad Web UI and Solr Admin UI when the Informatica Cluster Service is SSL enabled, you must import the browser certificates. To know more about how to import the browser certificates, see the Access Nomad Web UI and Solr Admin UI when Informatica Cluster Service is SSL enabled in EDC KB article.

Validate the CA-signed Certificates

You must use a single CA for the cluster and client certificates. After you receive the signed certificates from the CA, you must verify that each certificate is an X.509 certificate in .pem format.

Run the following command to view the contents of the signed cluster and client certificates: `keytool -printcert -file $PATH TO CERTIFICATE PEM FILE.`

Validate the following requirements for the cluster and client certificates:

Prerequisite	Certificate Requirement
Mandatory fields	For the cluster certificate, consider the following key usage requirements: keyUsage = digitalSignature,keyEncipherment extendedKeyUsage = serverAuth,clientAuth For the client certificate, consider the following key usage requirements: keyUsage = digitalSignature extendedKeyUsage = clientAuth
Subject Alternate Name (SAN)	For the cluster certificate, the SAN must include the list of cluster nodes in the following format: SAN=DNS:$CLUSTER HOST1 FQDN,DNS:$CLUSTER HOST2 FQDN,DNS:$CLUSTER HOST3 FQDN The client certificate must contain the FQDNs for the cluster nodes. For the client certificate, the SAN must include the list of all Informatica nodes in the following format: SAN=DNS:$INFA DOMAIN HOST1 FQDN,DNS:$INFA DOMAIN HOST2 FQDN The client certificate must contain the FQDNs for both the domain and cluster nodes.

CLUSTER HOST FQDN represents the fully qualified domain name for the cluster gateway host, processing nodes, and data nodes in the cluster.

INFA DOMAIN HOST FQDN represents the fully qualified domain name of the Informatica domain gateway host, domain nodes, cluster gateway host, processing nodes, and data nodes in the cluster.

Note: Verify that the custom certificate location for the domain nodes contains the infa_truststore.pem file. Also, verify that all the certificates in the CA certificate chain are present in the truststore.jks and the infa_pubcert.pem files.

Custom SSL Utility for Informatica Cluster Service

You have a custom SSL certificate for the Informatica domain, but you want to use default SSL certificates as the client and cluster certificates for the Informatica Cluster Service. In this scenario, you can use the custom SSL utility bundled with the installer to generate the required SSL certificates.

Perform the following steps to generate the required certificates using the custom SSL utility:

1. Extract `GenerateCustomSslUtility.zip` from the following location: `<Location of installer files>/properties/utils/CustomSslCertsUtility/`.
2. Set the JAVA_HOME environment variable to point to JDK 8.

3. Configure the following parameters in the `input.properties` file that you extracted from the `GenerateCustomSslUtility.zip` file:

Parameter	Description
KeystoreFile	Path to a custom keystore file along with file name. The keystore type must be in JKS. X509 format. The file must contain a single PrivateKeyEntry with the complete certificate chain. Verify that the file has CA-signing capability.
KeystorePassword	Password of the custom KeystoreFile in plain text format.
TruststoreFile	Path to a custom truststore file along with file name. the truststore type must be in JKS. X509 format. The file must contain the public certificates corresponding to the PrivateKeyEntry in the KeystoreFile.
TruststorePassword	Password of the custom TruststoreFile in plain text format.
ISPDomainKeystorePassword	Password of Informatica domain keystore in plain text format. The utility uses the password for the generated Informatica Cluster Service cluster and client keystore.jks file.
ISPDomainTruststorePassword	Password of the Informatica domain truststore in plain text format. The utility uses the password for the generated Informatica Cluster Service cluster and client truststore.jks file
KeystoreOutputDir	Represents the location of the Informatica domain custom keystore that you provided when you installed Enterprise Data Catalog.
TruststoreOutputDir	Represents the location of the Informatica domain custom truststore that you provided when you installed Enterprise Data Catalog. **Note:** Verify that the KeystoreOutputDir and TruststoreOutputDir parameters point to the same directory.
ServerNodes	Comma-separated list of fully qualified domain names of nodes that you plan to configure as data nodes, processing nodes, service hosts, and gateway node when you configure the Informatica Cluster Service.
IcsServiceName	Name of the Informatica Cluster Service.
ClientNodes	Comma-separated list of fully qualified domain names of nodes that you plan to configure as Informatica domain hosts, data nodes, processing nodes, service hosts, and gateway node.
IspDomainName	The Informatica domain name.
ClusterNodeDNSDomain	Domain name of the gateway host that you plan to configure for the Informatica Cluster Service.

4. Run the utility using the following command: `java -jar GenerateCustomSslUtility.jar -in input.properties`. The utility generates the following keys and client and cluster certificates:
   - Keys:
     - keystore.jks

- infa_privkey.pem

  Note: The cluster keys are generated in the following directory: `<KeystoreOutputDir>/<IcsServiceName>/cluster_certs`. The client keys are generated in the following directory: `<KeystoreOutputDir>/<IcsServiceName>/client_certs`

- Certificates:

  - truststore.jks
  - infa_pubcert.pem
  - infa_nodecert.pem
  - infa_nodecertkey.pem

  Note: The cluster certificates are generated in the following directory: `<TruststoreOutputDir>/<IcsServiceName>/cluster_certs`. The client certificates are generated in the following directory: `<TruststoreOutputDir>/<IcsServiceName>/client_certs`

5. Copy the generated certificates to the required folders.
6. Assign the ownership of the following directories to the Informatica domain user:
   - <KeystoreOutputDir>/<IcsServiceName>/cluster_certs
   - <TruststoreOutputDir>/<IcsServiceName>/cluster_certs
   - <KeystoreOutputDir>/<IcsServiceName>/client_certs
   - <TruststoreOutputDir>/<IcsServiceName/client_certs

Note: Verify that all the client certificates are included in a single directory. Similarly, verify that all the cluster certificates are included in a single directory. You must also verify that the directories have the `chmod 700` permission configured and the files under the directories have the `chmod 600` permission configured.

CHAPTER 7

Introduction to the Services Installer

This chapter includes the following topics:

Services Installer Tasks

The installer performs install tasks based on the product or products that you install.

The installer can perform the following tasks:

1. Perform pre-install validation and system check.
2. Create a domain or join a node to an existing domain.
3. Install binaries for service support.
4. Create application services.
5. Configure security between the domain and services.
6. Start the domain and application services that you created.
7. Write message to the log file.

Secure Files and Directories

When you install or upgrade Informatica, the installer creates directories to store Informatica files that require restricted access, such as the domain encryption key file and the nodemeta.xml. The installer assigns different permissions for the directories and the files in the directories.

By default, the installer creates the following directories within the Informatica installation directory:

<Informatica installation directory>/services/shared/security

If you enable secure communication for the domain, the /security directory contains the keystore and truststore files for the default SSL certificates.

To maintain the security of the directories and files, the installer restricts access to the directories and the files in the directories. The installer assigns specific permissions to the group and user account that own the directories and files.

For more information about permissions assigned to the directories and files, see the Informatica Security Guide.

Pre-install Utilities

Informatica provides utilities to facilitate the Informatica services installation process. You can use the Informatica installer to run the utilities.

Run the Pre-Installation (i10Pi) System Check Tool in Console Mode

Run the Pre-installation (i10Pi) System Check Tool to verify whether the machine meets the system requirements for installation or upgrade.

Ensure that you verified the system requirements and prepared the domain configuration repository database.

1. Log in to the machine with a system user account.
2. Close all other applications.
3. On a shell command line, run the install file.

   The installer displays the message to verify that the locale environment variables are set.
4. If the environment variables are not set, press **n** to exit the installer and set them as required.

   If the environment variables are set, press **y** to continue.
5. Press **1** to install or upgrade Informatica.
6. Press **1** to run the Pre-Installation (i10Pi) System Check Tool that verifies whether the machine meets the system requirements for the installation or upgrade.
7. From the Informatica Pre-Installation (i10Pi) System Check Tool **Welcome** section, press **Enter**.

   The **System Information** section appears.
8. Type the absolute path for the installation directory.

   The directory names in the path must not contain spaces or the following special characters: @|* $ # ! % ( ) { } [ ] , ; '

   Note: Informatica recommends using alphanumeric characters in the installation directory path. If you use a special character such as á or €, unexpected results might occur at run time.
9. Press **Enter**.

10. Enter the starting port number for the node that you will create or upgrade on the machine. The default port number for the node is 6005.
11. Press **Enter**.

    The **Database and Connection Information** section appears.
12. To enter the JDBC connection information using a custom JDBC connection string, press **1**. To enter the JDBC connection information using the JDBC URL information, press **2**.

    To connect to a secure database, you must enter the JDBC connection using a custom JDBC connection string.
13. Enter the JDBC connection information.
    - To enter the connection information using a custom JDBC connection string, type the connection string and specify the connection parameters.
      Use the following syntax in the JDBC connection string:
      IBM DB2

      ```
      jdbc:Informatica:db2://<host name>:<port number>;DatabaseName=
      ```

      Oracle

      ```
      jdbc:Informatica:oracle://<host name>:<port number>;ServiceName=
      ```

      Use the following connection string to connect to the Oracle database through the Oracle Connection Manager:

      ```
      jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora file>;TNSServerName=<TNS name>;
      ```

      Microsoft SQL Server

      ```
      jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=
      ```

      Microsoft Azure SQL

      ```
      jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=<database name>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net;ValidateServerCertificate=false
      ```

      Azure SQL Database with Active Directory authentication

      ```
      jdbc:informatica: sqlserver://<host_name>:<port_number>;database=<database_name>;encrypt=true;AuthenticationMethod=ActiveDirectoryPassword;trustServerCertificate=false;hostNameInCertificate=*.database.windows.net;loginTimeout=<seconds>
      ```

      PostgreSQL

      ```
      jdbc:Informatica:postgresql://<host name>:<port number>;DatabaseName=
      ```

      Sybase

      ```
      jdbc:Informatica:sybase://<host name>:<port number>;DatabaseName=
      ```

      Verify that the connection string contains all the connection parameters required by your database system.
    - To enter the connection information using the JDBC URL information, specify the JDBC URL properties.

The following table describes the connection information:

Prompt	Description
Database type	Type of database for the domain configuration repository. Select from the following database types: - 1 - Oracle - 2 - Microsoft SQL Server - 3 - IBM DB2 - 4 - Sybase ASE - 5 - PostgreSQL
Database user ID	User ID for the database user account for the domain configuration repository.
Database user password	Password for the database user account.
Database host name	Host name for the database server.
Database port number	Port number for the database.
Database service name	Service name for Oracle and IBM DB2 databases, or database name for PostgreSQL, Microsoft SQL Server, and Sybase ASE.

- To connect to a secure database, select **1** to use a custom string and type the connection string. You must include the security parameters in addition to the connection parameters. For information about the security parameters you must include in the JDBC connection for a secure database, see "Connection String to a Secure Database" on page 79.

The tool checks the settings of the hard drive, the availability of the ports, and the configuration of the database. After the system check is complete, the **System Check Summary** section displays the results of the system check.

14. Analyze the results of the system check.

    Each requirement is listed, along with one of the following check statuses:

    - [Pass] - The requirement meets the criteria for the Informatica installation or upgrade.
    - [Fail] - The requirement does not meet the criteria for the Informatica installation or upgrade. Resolve the issue before you proceed with the installation or upgrade.
    - [Information] - Verify the information and perform any additional tasks as outlined in the details.

    The results of the system check are saved to the following file: `...<Informatica installation directory>/Server/I10PI/I10PI/en/I10PI_summary.txt`

15. Press **Enter** to close the Pre-Installation (i10Pi) System Check Tool.

    You can continue to the Informatica service installation or upgrade immediately or end the system check and continue with the installation or upgrade later. If you continue to the installation or upgrade immediately, you do not have to restart the installer.

16. To continue to the Informatica service installation or upgrade immediately, press **y**.

    To end the system check and continue with the installation or upgrade later, press **n**.

If the Pre-Installation (i10Pi) System Check Tool finishes with failed requirements, resolve the failed requirements and run the Pre-Installation (i10Pi) System Check Tool again.

Note: If the Informatica Pre-Installation (i10Pi) System Check Tool check finishes with failed requirements, you can still perform the Informatica installation or upgrade. However, Informatica highly recommends that you resolve the failed requirements before you proceed.

Run the Pre-Installation (i10Pi) System Check Tool in Silent Mode

Run the Pre-installation (i10Pi) System Check Tool in silent mode to verify system requirements for installation without user intervention.

1. Extract the Informatica services installer file.
2. Navigate to the following location:

   `<Informatica installation directory>/Server/I10PI`
3. To specify the properties for the I10PI system check tool in silent mode, update the `SilentInput.properties` file in the `I10PI` folder.
4. To run the i10Pi in silent mode, run the `silentInstall` file in the `I10PI` folder.

You can view the results of the i10Pi system check tool in silent mode from the `I10PI_summary.txt` file in the following location:
`<Informatica installation directory>/Server/I10PI/I10PI/en`

If the Pre-Installation (i10Pi) System Check Tool finishes with failed requirements, resolve the failed requirements and run the Pre-Installation (i10Pi) System Check Tool again.

Note: If the Informatica Pre-Installation (i10Pi) System Check Tool check finishes with failed requirements, you can still perform the Informatica installation or upgrade. However, Informatica highly recommends that you resolve the failed requirements before you proceed.

Part III: Run the Services Installer

This part contains the following chapters:

CHAPTER 8

Create a Domain with Catalog and Data Engineering Products

This chapter includes the following topics:

Begin the Install

This task includes installer prompts to begin the installation. You will provide basic information such as acceptance of terms, installation option, and the installation directory.

You can install with the following options:

- Enterprise Data Preparation binaries in an existing domain with Enterprise Data Catalog
- Enterprise Data Catalog and Enterprise Data Preparation binaries in an existing domain
- Data Engineering products, Enterprise Data Catalog, and Enterprise Data Preparation

When you complete the preliminary tasks, you will continue with the installer prompts and will provide information to configure the domain.

Run the Installer

Perform the following steps to run the installer:

1. Log in to the machine with a system user account.
2. Close all other applications.
3. On a shell command line, run the `install.sh` file.

   The installer displays the message to verify that the locale environment variables are set.
4. If the environment variables are not set, press **n** to exit the installer and set them as required.

   If the environment variables are set, press **y** to continue.

Welcome - Accept Terms and Conditions

- Read the terms and conditions for Informatica installation and the product usage toolkit and select **I agree to the terms and conditions**.
  a. Press **1** if you do not want to accept the terms and conditions.
  b. Press **2** to accept the terms and conditions.

The **Component Selection** sections appears.

Choose the Installation Option

After you accept terms and conditions, you can install Informatica domain services supporting Data Engineering and Catalog products.

1. Press **3** to install Informatica Enterprise Data Preparation.

   When you select this option, you can choose to install Only Enterprise Data Preparation, Enterprise Data Catalog and Enterprise Data Preparation, and Informatica domain services, supporting Data Engineering services, Enterprise Data Catalog, and Enterprise Data Preparation.
2. Select whether current version of the Informatica domain services is installed on the node.
   a. Press **1** if current version of the Informatica domain services is not installed on the node.
   b. Press **2** if current version of the Informatica domain services is installed on the node.
3. Select whether you have read and accepted terms and conditions to use Java SE Development Kit software.
   a. Press **1** to not accept the terms and conditions to use Java SE Development Kit software.
   b. Press **2** to accept the terms and conditions to use Java SE Development Kit software.

The **Installation Prerequisites** section displays the installation requirements. Verify that all requirements are met before you continue the installation.

Tune the Application Services

After you review the installation prerequisites, you can choose to tune the application services for better performance based on the deployment type in your environment. If you do not tune now, you can tune the services later through infacmd.

1. Select if you want the installer to tune the application services:
   - Press 1 if you do not want to tune the services.
   - Press 2 if you want to tune the services.

   If you are joining the node to existing domain, ensure the deployment type you select here is same deployment type as the gateway nodes.
2. Select the deployment type associated with the Informatica environment.

Deployment Type	Disk Space per Node	Total Virtual Cores	RAM per Node
Sandbox	140 GB	16	32 GB
Basic	140 GB	24	64 GB

Deployment Type	Disk Space per Node	Total Virtual Cores	RAM per Node
Standard	140 GB	48	64 GB
Advanced	140 GB	96	128 GB

3. Select whether you want to change the deployment type or continue with the current deployment selection.
   a. Press **1** to change the deployment type.
   b. Press **2** to continue with the current deployment selection.

The **License and Installation Directory** section appears.

Specify the Installation Directory

After you verify the installation prerequisites, you can specify the installation directory.

1. Enter the installation directory.

   The directory names in the path must not contain spaces or the following special characters: @|* $ # ! % ( ) { } [ ] , ; ' Default is /<home>/Informatica/10.5.1.

   Note: Informatica recommends using alphanumeric characters in the installation directory path. If you use a special character such as á or €, unexpected results might occur at run time.
2. Enter the path to the license key file and press **Enter**.
3. Choose an installation environment and press **Enter**.
   - Press **1** to set Sandbox environment for a basic environment used for proof of concept with minimal users.
   - Press **2** to set Development environment for the design environment.
   - Press **3** to set Test environment for high volume processing that is closest to a production environment.
   - Press **4** to set Production environment for high volume processing with high levels of concurrency meant for end users. Advanced production environments are typically multi-node setups.

   Default is 1 for Sandbox.
4. Select whether you want to run the pre-validation utility.
   a. Press **1** to skip the pre-validation utility.
   b. Press **2** to run the pre-validation utility.

If you choose to skip the pre-validation utility, the **Pre-Installation Summary** section appears. Review the installation summary.

If choose to run the pre-validation utility, the Pre-validation section appears.

Prepare the Pre-validation Utility

You can use the pre-validation utility to verify the prerequisites to create the Informatica Cluster Service.

1. Select if you want to run the pre-validation utility:
   a. Press **1** to skip running the pre-validation utility.
   b. Press **2** to run the pre-validation utility.

2. If you chose to run the pre-validation utility, you must provide the details listed in the following table:

Property	Description
Informatica cluster gateway host	Fully qualified domain name of the node that you want to configure as the gateway host.
Gateway user	User name for the gateway host.
Enable advanced configuration	- Select 1 to skip validation of advanced configuration properties for associated services. - Select 2 to validate advanced configuration properties of associated services. If you select this option, you must specify values for all the properties.
Data nodes	Comma-separated list of fully qualified domain names of nodes that you want to configure as data nodes.
Processing nodes	Comma-separated list of fully qualified domain names of nodes that you want to configure as processing nodes.
Working directory path	Directory for the Informatica Cluster Service. Default is `/opt/informatica/ics`. **Note:** The permission on the directory must be `u=rwx (0700) or u=rwx,g=rx (0750)`. The Postgres service does not start if the directory does not have the required permission.

3. Applies if you selected the option to validate advanced configuration properties of associated services.

   Provide the following values for validation of the Nomad server configuration parameters:

Property	Description
Nomad Server Hosts	Comma-separated list of fully qualified domain names of nodes that host the Nomad servers.
Nomad HTTP Port	HTTP port number configured for the Nomad server. Default is 4646.
Nomad Serf Port	Serf port configured as the gossip protocol for the Nomad servers. Default is 4648.
Nomad RPC Port	The Remote Procedure Call (RPC) port configured for communication. Default is 4647.
Nomad Server Working Directory	The directory that includes sub directories with tasks running on the Nomad server. Default is $clusterCustomDir/nomad/nomadserver
Nomad Client Working Directory	The directory configured for tasks in the Nomad client. Default is $clusterCustomDir/nomad/nomadclient
Nomad Custom Options	Specify any custom options for the service in the following format: [OptionGroupName.OptionName=OptionValue]. You can separate multiple options using a white space character. If the OptionValue includes a white space character, you must enclose the OptionValue within double quotes as shown in the following sample: “sample value”.

Provide the following values for validation of the Apache ZooKeeper server configuration parameters:

Property	Description
ZooKeeper Hosts	Comma-separated list of fully qualified domain names of nodes that host the Apache ZooKeeper server.
ZooKeeper Port	Port number configured for the Apache ZooKeeper Server. Default is 2181.
ZooKeeper Peer Port	Port number configured for Apache ZooKeeper peer communication. Default is 2888.
ZooKeeper Leader Port	Port number configured for the ZooKeeper Sever identified as the Leader. Default is 3888.
ZooKeeper Installation Directory	Specify the path to the directory where you want to install Apache ZooKeeper. Default is $clusterCustomDir/zk/install
ZooKeeper Data Directory	Specify the path to the directory where you want to store data from Apache ZooKeeper. Default is $clusterCustomDir/zk/data
ZooKeeper Custom Options	Specify any custom options for the service in the following format: [OptionGroupName.OptionName=OptionValue]. You can separate multiple options using a white space character. If the OptionValue includes a white space character, you must enclose the OptionValue within double quotes as shown in the following sample: “sample value”.

Provide the following values for validation of the Apache Solr server configuration parameters:

Property	Description
Solr Hosts	Comma-separated list of fully qualified domain names of nodes that host the Apache Solr server.
Solr Port	Port number configured for Apache Solr Server. Default is 8983.
Solr Installation Directory	Specify the path to the directory where you want to install Apache Solr Server. Default is $clusterCustomDir/solr/install.
Solr Data Directory	Specify the path to the directory where you want to store data from Apache Solr. Default is $clusterCustomDir/solr/data
Solr Custom Options	Specify any custom options for the service in the following format: [OptionGroupName.OptionName=OptionValue]. You can separate multiple options using a white space character. If the OptionValue includes a white space character, you must enclose the OptionValue within double quotes as shown in the following sample: “sample value”.

Provide the following values for validation of the MongoDB database configuration parameters:

Property	Description
MongoDB Hosts	Comma-separated list of fully qualified domain names of nodes that host the MongoDB database.
MongoDB Port	Port number configured for MongoDB. Default is 27017.
MongoDB Log Directory	Specify the path to the directory where you want to store the log files. Default is $clusterCustomDir/mongo/log
MongoDB Data Directory	Specify the path to the directory where you want to store data from the MongoDB database. Default is $clusterCustomDir/mongo/data
MongoDB Custom Options	Specify any custom options for the service in the following format: [OptionGroupName.OptionName=OptionValue]. You can separate multiple options using a white space character. If the OptionValue includes a white space character, you must enclose the OptionValue within double quotes as shown in the following sample: “sample value”.

Provide the following values for validation of the PostgreSQL database configuration parameters:

Property	Description
PostgreSQL DB Host	Fully qualified domain name of the machine that hosts the PostgreSQL database. Default is the gateway host. **Note:** If you did not select the Enable Advanced Configuration option, the service uses the gateway host value specified as the host value
PostgreSQL DB Port	Port number configured for PostgreSQL. Default is 5432.
PostgreSQL DB Installation Directory	Specify the path to the directory where you want to install the PostgreSQL database. Default is $clusterCustomDir/postgres/install
PostgreSQL DB Log Directory	Specify the path to the directory where you want to store the log files from the PostgreSQL database. Default is $clusterCustomDir/postgres/log
PostgreSQL DB Data Directory	Specify the path to the directory where you want to store PostgreSQL data. Default is $clusterCustomDir/postgres/data
PostgreSQL DB Custom Options	Specify any custom options for the service in the following format: [OptionGroupName.OptionName=OptionValue]. You can separate multiple options using a white space character. If the OptionValue includes a white space character, you must enclose the OptionValue within double quotes as shown in the following sample: “sample value”.

Note: The details for the Data Privacy Management, Elasticsearch, and Spark services are not validated by the pre-validation utility.

Configure the Domain

This task includes installer prompts to configure the domain. You will provide information to create a domain, configure the domain security, domain repository, and application services.

Configure the Domain Options

After you review the Pre-Installation summary, you can enter the domain information.

1. Press **1** to create a domain.

   When you create a domain, the node that you create becomes a gateway node in the domain. The gateway node contains a Service Manager that manages all domain operations.

2. Select whether you want to enable secure communication for services in the domain.

   a. Press **1** to disable secure communication for the domain.

   b. Press **2** to enable secure communication for the domain.

   By default, if you enable secure communication for the domain, the installer sets up an HTTPS connection for the Informatica Administrator. You can also create a domain configuration repository on a secure database.

3. Specify the connection details for Informatica Administrator.

   a. If you do not enable secure communication for the domain, you can specify whether to set up a secure HTTPS connection for the Informatica Administrator.

   The following table describes the options available to enable or disable a secure connection to Informatica Administrator:

Option	Description
Enable HTTPS for Informatica Administrator	Set up a secure connection to Informatica Administrator.
Disable HTTPS	Do not set up a secure connection to Informatica Administrator.

   b. If you enable secure communication for the domain or if you enable HTTPS connection for the Informatica Administrator, enter the keystore file and port number for the HTTPS connection to Informatica Administrator.

The following table describes the connection information you must enter if you enable HTTPS:

Option	Description
Port	Port number for the HTTPS connection.
Keystore file	Select whether to use a keystore file generated by the installer or a keystore file you create. You can use a keystore file with a self-signed certificate or a certificate signed by a certification authority. 1 - Use a keystore generated by the installer 2 - Specify a keystore file and password If you select to use a keystore file generated by the installer, the installer creates a self-signed keystore file named Default.keystore in the following location: `<Informatica installation directory>/tomcat/conf/`

c. If you specify the keystore, enter the password and location of the keystore file.

d. If you enabled secure communication for the domain, the **Domain Security - Secure Communication** section appears.

e. If you did not enable secure communication for the domain, the **Domain Configuration Repository** section appears. Skip to "Configure Domain Repository Details" on page 107.

4. Select whether to enable SAML authentication to configure Security Assertion Markup Language (SAML)-based single sign-on (SSO) support for web-based Informatica applications in an Informatica domain.

   Press **1** to disable SAML authentication and skip to #unique_165. Press **2** to enable and configure SAML authentication.

5. Enter the Identity Provider URL for the domain.
6. Specify the relying party trust name or the service provider identifier for the domain as defined in the identity provider. If you choose No, the service provider identifier is set to "Informatica".
7. Specify whether IdP will sign SAML assertion or not.
8. Enter the identity provider assertion signing certificate alias name.
9. Select whether to use the default Informatica SSL certificates or to use your SSL certificates to enable secure communication in the domain.

   The following table describes the SSL certificate options for SAML authentication:

Option	Description
Use the default Informatica SSL certificate file.	Select to use the default Informatica truststore file for SAML authentication.
Enter the location of the SSL certificate file.	Select to use a custom truststore file for SAML authentication. Specify the directory containing the custom truststore file on gateway nodes within the domain. Specify the directory only, not the full path to the file.

10. Select whether to use the default Informatica SSL certificates or to use your SSL certificates to enable SAML authentication in the domain.

The following table describes the SSL certificate options for SAML authentication:

Option	Description
Use the default Informatica SSL certificate file.	Select to use the default Informatica truststore file for SAML authentication.
Enter the location of the SSL certificate file.	Select to use a custom truststore file for SAML authentication. Specify the directory containing the custom truststore file on gateway nodes within the domain. Specify the directory only, not the full path to the file.

11. If you provide the security certificates, specify the location and passwords of the keystore and truststore files.

    The following table describes the location and password of the truststore and keystore files:

Property	Description
Truststore Directory	Specify the directory containing the custom truststore file on gateway nodes within the domain. Specify the directory only, not the full path to the file.
Truststore Password	The password for the custom truststore file.
Keystore Directory	Specify the directory containing the custom keystore file.
Keystore Password	The password for the custom keystore file.

12. To specify the Authentication Context Comparison, specify the strength comparison of the authentication mechanism used by the user with the IdP server.

    Supported values are MINIMUM, MAXIMUM, BETTER, or EXACT option. Default is MINIMUM.

13. To set the Authentication Context Class, specify the expected mechanism of first time authentication of the user with the IdP server.

    Supported values are PASSWORD or PASSWORDPROTECTEDTRANSPORT. Default is PASSWORD.

14. Specify if you want to enable the webapp to sign the SAML authentication request or not?

    Default is disabled.

15. Specify the alias name of the private key that was imported to the node SAML keystore using which the SAML request should be signed.

16. Specify the password to access the private key used for signing the SAML request.

17. Specify the algorithm that the web application uses to sign the SAML request.

    Supported values are `RSA_SHA256, DSA_SHA1, DSA_SHA256, RSA_SHA1, RSA_SHA224, RSA_SHA384, RSA_SHA512, ECDSA_SHA1, ECDSA_SHA224, ECDSA_SHA256, ECDSA_SHA384, ECDSA_SHA512, RIPEMD160, or RSA_MD5`.

18. Specify whether you want IdP to sign the SAML response or not?

    Choose to select to enable the webapp to receive the signed SAML response or not. Default is disabled.

19. Specify whether IdP will encrypt SAML assertion or not.

    Select to enable the webapp to receive an encrypted SAML assertion. Default is enabled.

20. Specify the alias name of the private key present in the gateway nodes gateway node SAML truststore that used for Informatica uses to decrypt decrypting the SAML assertion.

21. Provide the password to access the private key to use when decrypting the assertion encryption key.
22. Click **Next**.

The **Configure Domain Security** appears.

Configure Domain Security

After you configure the domain, you can configure domain security.

- In the Domain Security - Secure Communication section, specify whether to use the default Informatica SSL certificates or to use your SSL certificates to secure domain communication.
  a. Select the type of SSL certificates to use.

     The following table describes the options for the SSL certificates that you can use to secure the Informatica domain:

Option	Description
Use the default Informatica SSL certificates	Use the default SSL certificates contained in the default keystore and truststore. **Note:** If you do not provide an SSL certificate, Informatica uses the same default private key for all Informatica installations. If you use the default Informatica keystore and truststore files, the security of your domain could be compromised. To ensure a high level of security for the domain, select the option to specify the location of the SSL certificate files.
Use custom SSL certificates	Specify the path for the keystore and truststore files that contain the SSL certificates. You must also specify the keystore and truststore passwords. You can provide a self-signed certificate or a certificate issued by a certificate authority (CA). You must provide SSL certificates in PEM format and in Java Keystore (JKS) files. Informatica requires specific names for the SSL certificate files for the Informatica domain. You must use the same SSL certificates for all nodes in the domain. Store the truststore and keystore files in a directory accessible to all the nodes in the domain and specify the same keystore file directory and truststore file directory for all nodes in the same domain.

  b. If you provide the SSL certificate, specify the location and passwords of the keystore and truststore files.

     The following table describes the parameters that you must enter for the SSL certificate files:

Property	Description
Keystore file directory	Directory that contains the keystore files. The directory must contain files named infa_keystore.jks.
Keystore password	Password for the keystore infa_keystore.jks.

Property	Description
Truststore file directory	Directory that contains the truststore files. The directory must contain files named infa_truststore.jks and infa_truststore.pem.
Truststore password	Password for the infa_truststore.jks file.

The **Domain Configuration Repository** section appears.

Configure Domain Repository Details

After you configure domain security, you can configure domain repository details.

1. Select the database to use for the domain configuration repository details.

   The following table lists the databases you can use for the domain configuration repository:

Prompt	Description
Database type	Type of database for the domain configuration repository. Select from the following options: 1 - Oracle 2 - Microsoft SQL Server 3 - IBM DB2 4 - Sybase ASE 5 - PostgreSQL

   The Informatica domain configuration repository stores metadata for domain operations and user authentication. The domain configuration repository must be accessible to all gateway nodes in the domain.

2. Enter the properties for the database user account.

   The following table lists the properties for the database user account:

Property	Description
Database user ID	Name for the domain configuration database user account.
User password	Password for the domain configuration database user account.

3. Select whether to create a secure domain configuration repository.

   You can create a domain configuration repository in a database secured with the SSL protocol. To create a domain configuration repository in a secure database, press 1 and skip to step 5.

   To create a domain configuration repository in an unsecure database, press 2.

4. If you do not create a secure domain configuration repository, enter the parameters for the database.

   a. If you select IBM DB2, select whether to configure a tablespace and enter the tablespace name.

The following table describes the properties that you must configure for the IBM DB2 database:

Property	Description
Configure tablespace	Select whether to specify a tablespace: 1 - No 2 - Yes In a single-partition database, if you select No, the installer creates the tables in the default tablespace. In a multi-partition database, you must select Yes.
Tablespace	Name of the tablespace in which to create the tables. Specify a tablespace that meets the pageSize requirement of 32768 bytes. In a single-partition database, if you select Yes to configure the tablespace, enter the name of the tablespace in which to create the tables. In a multi-partition database, specify the name of the tablespace that resides in the catalog partition of the database.

b. If you select Microsoft SQL Server, enter the schema name for the database.

The following table describes the properties that you must configure for the Microsoft SQL Server database:

Property	Description
Schema name	Name of the schema that will contain domain configuration tables. If this parameter is blank, the installer creates the tables in the default schema.

c. To enter the JDBC connection information using the JDBC URL information, press **1**. To enter the JDBC connection information using a custom JDBC connection string, press **2**.

d. Enter the JDBC connection information.

- To enter the connection information using the JDBC URL information, specify the JDBC URL properties.
  The following table describes the database connection information:

Prompt	Description
Database host name	Host name for the database.
Database port number	Port number for the database.

Prompt	Description
Database service name	Service or database name: - Oracle: Enter the service name. - Microsoft SQL Server: Enter the database name. - IBM DB2: Enter the service name. - Sybase ASE: Enter the database name. - PostgreSQL: Enter the database name.
Configure JDBC Parameters	Select whether to add additional JDBC parameters to the connection string: 1 - Yes 2 - No If you select Yes, enter the parameters or press Enter to accept the default. If you select No, the installer creates the JDBC connection string without parameters.

- To enter the connection information using a custom JDBC connection string, type the connection string.
  Use the following syntax in the JDBC connection string:
  IBM DB2

  ```
  jdbc:Informatica:db2://<host name>:<port number>;DatabaseName=
  ```

  Oracle

  ```
  jdbc:Informatica:oracle://<host name>:<port number>;ServiceName=
  ```

  Use the following connection string to connect to the Oracle database through the Oracle Connection Manager:

  ```
  jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora file>;TNSServerName=<TNS name>;
  ```

  Microsoft SQL Server

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=
  ```

  Microsoft Azure SQL

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=<database name>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net;ValidateServerCertificate=false
  ```

  Azure SQL Database with Active Directory authentication

  ```
  jdbc:informatica: sqlserver://<host_name>:<port_number>;database=<database_name>;encrypt=true;AuthenticationMethod=ActiveDirectoryPassword;trustServerCertificate=false;hostNameInCertificate=*.database.windows.net;loginTimeout=<seconds>
  ```

  PostgreSQL

  ```
  jdbc:Informatica:postgresql://<host name>:<port number>;DatabaseName=
  ```

Sybase

```
jdbc:Informatica:sybase://<host name>:<port number>;DatabaseName=
```

Verify that the connection string contains all the connection parameters required by your database system.

5. If you create a secure domain configuration repository, enter the parameters for the secure database.

   If you create the domain configuration repository on a secure database, you must provide the truststore information for the database. You must also provide a JDBC connection string that includes the security parameters for the database.

   The following table describes the options available to create a secure domain configuration repository database:

Property	Description
Database truststore file	Path and file name of the truststore file for the secure database.
Database truststore password	Password for the truststore file.
Custom JDBC Connection String	JDBC connection string to connect to the secure database, including the host name and port number and the security parameters for the database.

In addition to the host name and port number for the database server, you must include the following secure database parameters:

EncryptionMethod

Required. Indicates whether data is encrypted when transmitted over the network. This parameter must be set to`SSL`.

ValidateServerCertificate

Optional. Indicates whether Informatica validates the certificate that the database server sends.

If this parameter is set to True, Informatica validates the certificate that the database server sends. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate.

If this parameter is set to False, Informatica does not validate the certificate that the database server sends. Informatica ignores any truststore information that you specify

Default is True.

HostNameInCertificate

Optional. Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.

cryptoProtocolVersion

Required. Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to `cryptoProtocolVersion=TLSv1.1` or `cryptoProtocolVersion=TLSv1.2` based on the cryptographic protocol used by the database server:

You can use the following syntax for the secure JDBC connection strings:

- IBM DB2: `jdbc:Informatica:db2://<host name:port number>;DatabaseName=<database name>;EncryptionMethod=SSL;HostNameInCertificate=<database host name>;ValidateServerCertificate=<true or false>`
- Oracle: `jdbc:Informatica:oracle://<host name:port number>;ServiceName=<service name>;EncryptionMethod=SSL;HostNameInCertificate=<database host name>;ValidateServerCertificate=<true or false>`
- Microsoft SQL Server: `jdbc:Informatica:sqlserver://<host name:port number>;SelectMethod=cursor;DatabaseName=<database name>;EncryptionMethod=SSL;HostNameInCertificate=<database host name>;ValidateServerCertificate=<true or false>`
- PostgreSQL: `jdbc:Informatica:postgresql://<host name:port number>;DatabaseName=<database name>;EncryptionMethod=SSL;HostNameInCertificate=<database host name>;ValidateServerCertificate=<true or false>`

Note: The installer does not validate the connection string. Verify that the connection string contains all the connection parameters and security parameters required by your database.

6. If the database contains a domain configuration repository for a previous domain, select to overwrite the data or set up another database.

   The following table describes the options of overwriting the data or setting up another database when you create a domain configuration repository for a previous domain:

Option	Description
1 - OK	Enter the connection information for a new database.
2 - Continue	The installer overwrites the data in the database with new domain configuration.

The **Domain Security - Encryption Key** section appears.

Domain Security - Encryption Key

After you configure domain repository, you can configure encryption key.

▶ In the **Domain Security - Encryption Key** section, enter the directory for the encryption key for the Informatica domain.

The following table describes the encryption key parameters that you must specify when you create a domain:

Property	Description
Encryption key directory	Directory in which to store the encryption key for the domain. By default, the encryption key is created in the following directory: `<Informatica installation directory>/isp/config/keys`.
Specify if you want to back up the site key that the installer generates or not	A unique site key is generated. If you lose the site key, you cannot generate the site key again. Make sure that you save a copy of this key and do not share the unique site key with others. Specify if you want to back up the site key that the installer generates or not: - Select **1** for No. If you choose No, the installer generates an error. Press **Enter** to continue. - Select **2** for Yes. If you choose Yes, you agree to back up the file manually.

The installer sets different permissions to the directory and the files in the directory. For more information about the permissions for the encryption key file and directory, see "Secure Files and Directories" on page 91.

The **Domain and Node Configuration** section appears.

Configure the Domain and Node

After you configure the encryption key, you can configure the domain and node.

1. Enter the information for the domain and the node that you want to create.

   The following table describes the properties that you set for the domain and gateway node.

Property	Description
Domain name	Name of the Informatica domain to create. The default domain name is Domain_<MachineName>. The name must not exceed 128 characters and must be 7-bit ASCII only. The name cannot contain a space or any of the following characters: `` ` % * + ; " ? , < > \ / ``
Node name	Name of the node to create.
Node host name	Host name or IP address of the machine on which to create the node. If the machine has a single network name, use the default host name. If the a machine has multiple network names, you can modify the default host name to use an alternate network name. **Note:** The node host name cannot contain the underscore (_) character. Do not use localhost. The host name must explicitly identify the machine.

Property	Description
Node port number	Port number for the node. The default port number for the node is 6005. If the port number is not available on the machine, the installer displays the next available port number.
Domain user name	User name for the domain administrator. You can use this user name to initially log in to Informatica Administrator. Use the following guidelines: - The name is not case sensitive and cannot exceed 128 characters. - The name cannot include a tab, newline character, or the following special characters: % * + / ? ; < >\n- The name can include an ASCII space character except for the first and last character. Other space characters are not allowed.

2. Select whether you want to enable password complexity to secure sensitive data in the domain.

   The following table describes the password complexity:

Prompt	Description
Password complexity	Select whether you want to enable password complexity. 1 - Yes 2 - No If you select Yes, the password must meet the following requirements: It must be at least eight characters long and contain at least one alpha character, one numeric character, and one special character.
Domain password	Password for the domain administrator. The password must be more than 2 characters and must not exceed 16 characters. Not available if you configure the Informatica domain to run on a network with Kerberos authentication.
Confirm password	Enter the password again to confirm. Not available if you configure the Informatica domain to run on a network with Kerberos authentication.
Security domain name	Name of the security domain to which the domain administrator account belongs.

3. Select whether to display the default ports for the domain and node components assigned by the installer.

The following table describes the advanced port configuration page:

Prompt	Description
Display advanced port configuration page	Select whether to display the port numbers for the domain and node components assigned by the installer: 1 - No 2 - Yes If you select Yes, the installer displays the default port numbers assigned to the domain components. You can specify the port numbers to use for the domain and node components. You can also specify a range of port numbers to use for the service process that will run on the node. You can use the default port numbers or specify new port numbers. Verify that the port numbers you enter are not used by other applications.

4. If you display the port configuration page, enter new port numbers at the prompt or press Enter to use the default port numbers.

   The following table describes the ports that you can set:

Port	Description
Service Manager port	Port number used by the Service Manager on the node. The Service Manager listens for incoming connection requests on this port. Client applications use this port to communicate with the services in the domain. The Informatica command line programs use this port to communicate to the domain. This is also the port for the SQL data service JDBC/ODBC driver. Default is 6006.
Service Manager Shutdown port	Port number that controls server shutdown for the domain Service Manager. The Service Manager listens for shutdown commands on this port. Default is 6007.
Informatica Administrator port	Port number used by Informatica Administrator. Default is 6008.
Informatica Administrator HTTPS port	No default port. Enter the required port number when you create the service. Setting this port to 0 disables an HTTPS connection to the Administrator tool.
Informatica Administrator shutdown port	Port number that controls server shutdown for Informatica Administrator. Informatica Administrator listens for shutdown commands on this port. Default is 6009.
Minimum port number	Lowest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6014.
Maximum port number	Highest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6114.

5. Specify if you want to create the Enterprise Data Catalog application services.

   Press 1 to create the application services, or press 2 if you want to create the application services from the Administrator tool.

6. Specify if you want to create a monitoring Model Repository Service to monitor domain statistics.

   Press 1 to create a monitoring Model Repository Service, or press 2 if you want to create a monitoring Model Repository Service from the Administrator tool.

7. Specify if you want to create a Content Managment Service for data domain discovery.

Press 1 to create a Content Management Service, or press 2 if you want to create a Content Management Service from the Administrator tool.

8. Specify if you want to configure a profiling warehouse connection.

   Press 1 to configure a profiling warehouse connection, or press 2 if you want to configure a profiling warehouse connection from the Administrator tool.

9. Specify if you want to create a cluster configuration to enable the Data Integration Service to push mapping logic to the cluster.

   Press 1 to create a cluster configuration, or press 2 if you want to create a cluster configuration from the Administrator tool.

10. Specify if you want to create an Informatica Metadata Access Service.

    Press 1 to create an Informatica Metadata Access Service, or press 2 if you want to create an Informatica Metadata Access Service from the Administrator tool.

11. Specify if you want to configure the repository for Advanced Scanners.

    Press 1 to configure the repository for Advanced Scanners, or press 2 if you want to configure the repository for Advanced Scanners after installation using the steps specified in the following section: Configure Advanced Scanners Repository After You Install Enterprise Data Catalog .

If you choose to create Enterprise Data Catalog Application Services, the **Model Repository Database** section appears. If you choose not to create Enterprise Data Catalog Application Services, the **Post-Installation Summary** section indicates whether the installation completed successfully. The summary also shows the status of the installed components and their configuration.

Configure the Model Repository Database

After you configure the domain and the node, you can configure the Model repository database properties.

1. Enter the Model Repository Service name.

   Enter the name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters:

   ```
   ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( ) ] [
   ```

   You cannot change the name of the service after you create it.

   If you selected process level SPN, specify the Model Repository Service keytab file. The keytab file for the Model Repository Service process. The keytab file must have the following name: .keytab

2. Select the database to configure Model repository database.

   The following table lists the database type for the Model repository:

Prompt	Description
Database type	Type of database for the Model repository. Select from the following options: 1 - Oracle 2 - Microsoft SQL Server 3 - IBM DB2 4 - PostgreSQL

3. Enter the properties for the database user account.

The following table lists the properties for the database user account:

Property	Description
Database user ID	Name for the Model repository database user account.
User password	Password for the Model repository user account.

4. Select whether to create a secure Model repository database.

   You can create a model repository service in a database secured with the SSL protocol. To create a model repository service in a secure database, press 1 and skip to step to enter the JDBC connection information.

   To create a Model repository in an unsecure database, press 2.

5. If you do not create a secure Model repository, enter the parameters for the database.

   a. If you select IBM DB2, select whether to configure a tablespace and enter the tablespace name.

      The following table describes the properties that you must configure for the IBM DB2 database:

Property	Description
Configure tablespace	Select whether to specify a tablespace: 1 - No 2 - Yes In a single-partition database, if you select No, the installer creates the tables in the default tablespace. In a multi-partition database, you must select Yes.
Tablespace	Name of the tablespace in which to create the tables. Specify a tablespace that meets the pageSize requirement of 32768 bytes. In a single-partition database, if you select Yes to configure the tablespace, enter the name of the tablespace in which to create the tables. In a multi-partition database, specify the name of the tablespace that resides in the catalog partition of the database.

   b. If you select Microsoft SQL Server or PostgreSQL, enter the schema name for the database.

      The following table describes the properties that you must configure for the Microsoft SQL Server or PostgreSQL database:

Property	Description
Schema name	Name of the schema that will contain domain configuration tables. If this parameter is blank, the installer creates the tables in the default schema.

   c. To enter the JDBC connection information using the JDBC URL information, press **1**. To enter the JDBC connection information using a custom JDBC connection string, press **2**.

d. Enter the JDBC connection information.

- To enter the connection information using the JDBC URL information, specify the JDBC URL properties.
  The following table describes the database connection information:

Prompt	Description
Database host name	Host name for the database.
Database port number	Port number for the database.
Database service name	Service or database name : - Oracle: Enter the service name. - Microsoft SQL Server: Enter the database name. - IBM DB2: Enter the service name. - PostgreSQL: Enter the database name.
Configure JDBC Parameters	Select whether to add additional JDBC parameters to the connection string: 1 - Yes 2 - No If you select Yes, enter the parameters or press Enter to accept the default. If you select No, the installer creates the JDBC connection string without parameters.

- To enter the connection information using a custom JDBC connection string, type the connection string.
  Use the following syntax in the JDBC connection string:
  IBM DB2

  ```
  jdbc:Informatica:db2://<host name>:<port number>;DatabaseName=
  ```

  Oracle

  ```
  jdbc:Informatica:oracle://<host name>:<port number>;ServiceName=
  ```

  Use the following connection string to connect to the Oracle database through the Oracle Connection Manager:

  ```
  jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora file>;TNSServerName=<TNS name>;
  ```

  Microsoft SQL Server

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=
  ```

  Microsoft Azure SQL

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=<database name>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net;ValidateServerCertificate=false
  ```

Azure SQL Database with Active Directory authentication

```
jdbc:informatica: sqlserver://
<host_name>:<port_number>;database=<database_name>;encrypt=true;Authentication
Method=ActiveDirectoryPassword;trustServerCertificate=false;hostNameInCertific
ate=*.database.windows.net;loginTimeout=<seconds>
```

PostgreSQL

```
jdbc:Informatica:postgresql://<host name>:<port number>;DatabaseName=
```

Verify that the connection string contains all the connection parameters required by your database system.

The **Service Parameters** section appears.

Configure the Monitoring Model Repository Database

If you chose to create a monitoring Model Repository Service, you can provide connection information about the repository database.

1. Enter the monitoring Model Repository Service name.

   Enter the name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters:

   ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [

   You cannot change the name of the service after you create it.

   If you selected process level SPN, specify the monitoring Model Repository Service keytab file. The keytab file for the monitoring Model Repository Service process. The keytab file must have the following name: .keytab

2. Select the database for the monitoring Model repository.

   The following table lists the database type for monitoring Model repository:

Prompt	Description
Database type	Type of database type for monitoring Model repository. Select from the following options: 1 - Oracle 2 - Microsoft SQL Server 3 - IBM DB2 4 - PostgreSQL

3. Enter the properties for the database user account.

   The following table lists the properties for the database user account:

Property	Description
Database user ID	Name for the monitoring Model repository database user account.
User password	Password for the Monitoring model repository database user account.

4. Select whether to create a secure monitoring Model repository.

   You can create a monitoring Model repository in a database secured with the SSL protocol. To create a monitoring Model repository in a secure database, press 1 and skip to step to enter the JDBC connection information.

   To create a monitoring Model repository in an unsecured database, press 2.

5. If you do not create a secure monitoring Model repository, enter the parameters for the database.

   a. If you select IBM DB2, select whether to configure a tablespace and enter the tablespace name.

      The following table describes the properties that you must configure for the IBM DB2 database:

Property	Description
Configure tablespace	Select whether to specify a tablespace: 1 - No 2 - Yes In a single-partition database, if you select No, the installer creates the tables in the default tablespace. In a multi-partition database, you must select Yes.
Tablespace	Name of the tablespace in which to create the tables. Specify a tablespace that meets the pageSize requirement of 32768 bytes. In a single-partition database, if you select Yes to configure the tablespace, enter the name of the tablespace in which to create the tables. In a multi-partition database, specify the name of the tablespace that resides in the catalog partition of the database.

   b. If you select Microsoft SQL Server or PostgreSQL, enter the schema name for the database.

      The following table describes the properties that you must configure for the Microsoft SQL Server or PostgreSQL database:

Property	Description
Schema name	Name of the schema that will contain domain configuration tables. If this parameter is blank, the installer creates the tables in the default schema.

   c. To enter the JDBC connection information using the JDBC URL information, press **1**. To enter the JDBC connection information using a custom JDBC connection string, press **2**.

   d. Enter the JDBC connection information.

      - To enter the connection information using the JDBC URL information, specify the JDBC URL properties.

The following table describes the database connection information:

Prompt	Description
Database host name	Host name for the database.
Database port number	Port number for the database.
Database service name	Service or database name : - Oracle: Enter the service name. - Microsoft SQL Server: Enter the database name. - IBM DB2: Enter the service name. - PostgreSQL: Enter the database name.
Configure JDBC Parameters	Select whether to add additional JDBC parameters to the connection string: 1 - Yes 2 - No If you select Yes, enter the parameters or press Enter to accept the default. If you select No, the installer creates the JDBC connection string without parameters.

- To enter the connection information using a custom JDBC connection string, type the connection string.
  Use the following syntax in the JDBC connection string:
  IBM DB2

  ```
  jdbc:Informatica:db2://<host name>:<port number>;DatabaseName=
  ```

  Oracle

  ```
  jdbc:Informatica:oracle://<host name>:<port number>;ServiceName=
  ```

  Use the following connection string to connect to the Oracle database through the Oracle Connection Manager:

  ```
  jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora file>;TNSServerName=<TNS name>;
  ```

  Microsoft SQL Server

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=
  ```

  Microsoft Azure SQL

  ```
  jdbc:Informatica:sqlserver://<host name>:<port number>;SelectMethod=cursor;DatabaseName=<database name>;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net;ValidateServerCertificate=false
  ```

  Azure SQL Database with Active Directory authentication

  ```
  jdbc:informatica: sqlserver://<host_name>:<port_number>;database=<database_name>;encrypt=true;Authentication
  ```

```
Method=ActiveDirectoryPassword;trustServerCertificate=false;hostNameInCertific
ate=*.database.windows.net;loginTimeout=<seconds>
```

PostgreSQL

```
jdbc:Informatica:postgresql://<host name>:<port number>;DatabaseName=
```

Sybase

```
jdbc:Informatica:sybase://<host name>:<port number>;DatabaseName=
```

Verify that the connection string contains all the connection parameters required by your database system.

The **Service Parameters** section appears.

Configure the Data Integration Service Parameters

After you configure the Model Repository database, you can configure the service parameters for the Data Integration Service.

1. Enter the following service parameter information:

Port	Description
Data Integration Service name	Name of the Data Integration Service to create in the Informatica domain.
HTTP protocol type	Type of connection to the Data Integration Service. Select one of the following options: - HTTP. Requests to the service uses an HTTP connection. - HTTPS. Requests to the service uses a secure HTTP connection. - HTTP&HTTPS. Requests to the service can use either an HTTP or HTTPS connection.
HTTP port	Port number to used for the Data Integration Service. Default is 9085.
HTTPS port	Port number to used for the Data Integration Service. Default is 9085.

2. Select the SSL certificates to use to secure the Data Integration Service.

Option	Description
Use the default Informatica SSL certificate files	Use the default Informatica SSL certificates contained in the default keystore and truststore. **Note:** If you do not provide an SSL certificate, Informatica uses the same default private key for all Informatica installations. If you use the default Informatica keystore and truststore files, the security of your domain could be compromised. To ensure a high level of security for the domain, select the option to specify the location of the SSL certificate files.
Use custom SSL certificates	Use custom SSL certificates. You must specify the location of the keystore and truststore files. You can provide a self-signed certificate or a certificate issued by a certificate authority (CA). You must provide SSL certificates in PEM format and in Java Keystore (JKS) files. Informatica requires specific names for the SSL certificate files for the Informatica domain. You must use the same SSL certificates for all nodes in the domain. Store the truststore and keystore files in a directory accessible to all the nodes in the domain and specify the same keystore file directory and truststore file directory for all nodes in the same domain.

If you choose to use custom SSL certificates, enter the following information.

Property	Description
Keystore file directory	Directory that contains the keystore files. The directory must contain files named infa_keystore.jks and infa_keystore.pem.
Keystore password	Password for the keystore infa_keystore.jks.
Truststore file directory	Directory that contains the truststore files. The directory must contain files named infa_truststore.jks and infa_truststore.pem.
Truststore password	Password for the infa_truststore.jks file.

3. Do you want to enable data engineering recovery for the Data Integration Service?
   - Yes
   - No

   If you choose Yes, you can recover mapping jobs that the Data Integration Service pushes to the Spark engine for processing. Default is No.
4. Do you want to create a cluster configuration?

   The cluster configuration enables the Data Integration Service to push mapping logic to the cluster. If you are integrating with a Hadoop or Databricks environment, you can create a cluster configuration.

   Press 1 if you want to create a cluster configuration.

   Press 2 if you do not want to create a cluster configuration. Default is 1.

   If you want to create Enterprise Data Preparation services during the installation, you must create a cluster configuration. After installation, refer to the *Data Engineering Integration Guide* to fully integrate the domain with the Hadoop environment.
5. Select whether you want to configure profiling warehouse connection.
   a. Press **1** to configure the profiling warehouse connection.
   b. Press **2** to skip configuring the profiling warehouse connection.
   c. If you choose to configure the profiling warehouse connection, the **Profiling Warehouse Connection Database** section appears.
   d. If you choose to skip the profiling warehouse connection, the **Post-Installation Summary** section indicates whether the installation completed successfully. The summary also shows the status of the installed components and their configuration.
6. Select whether you want to configure the Content Management Service for data domain discovery.
   a. Press **1** to configure the Content Management Service for data domain discovery.
   b. Press **2** to skip configuring the Content Management Service for data domain discovery.
   c. If you choose to configure the Content Management Service for data domain discovery, the **Content Management Service Parameters and Database** section appears.
   d. If you choose to skip configure the Content Management Service for data domain discovery, the **Post-Installation Summary** section indicates whether the installation completed successfully. The summary also shows the status of the installed components and their configuration.

Create the Cluster Configuration

Create the cluster configuration, which contains configuration information about the non-native environment. The cluster configuration enables the Data Integration Service to push jobs to the non-native environment. You must create a cluster configuration if you plan to use Enterprise Data Preparation.

You import configuration properties from the non-native environment to create a cluster configuration. You can import the properties from an archive file that the Hadoop administrator creates, or you can import the properties directly from the cluster.

When you create the cluster configuration, you can also choose to create Hadoop, HDFS, Hive, HBase, or Databricks connections to the non-native environment. The installer appends the connection type to the cluster configuration name to create each connection name.

1. Enter the name of the cluster configuration to create.
2. Specify the non-native distribution for the cluster.

   The following table describes the options you can specify:

Prompt	Description
1	Cloudera. You can create a cluster configuration for a Cloudera cluster on Cloudera Distribution Hadoop (CDH).
2	Hortonworks
3	Azure HDInsight
4	MapR. You must import MapR cluster configuration properties from an archive file.
5	Amazon EMR. You must import Amazon EMR cluster configuration properties from an archive file.
6	Databricks
7	Google Dataproc

3. Import configuration properties from the non-native environment to create the cluster configuration.
   - To import the properties from an archive file, press **1**. If you create a cluster configuration for an Amazon EMR, MapR, or Google Dataproc cluster, you must import the properties from an archive file.
   - To import the properties directly from the cluster, press **2**.
4. If you choose to import the properties from an archive file, you must choose the configuration archive file name and path to the file.
5. If you choose to import the properties directly from the cluster, specify the connection properties.

The following table describes the Cloudera, Hortonworks, or Azure HDInsight cluster properties you specify:

Property	Description
Host	The host name or IP address of the cluster manager.
Port	Port of the cluster manager.
User ID	Cluster user name.
Password	Password for the cluster user.
Cluster Name	Name of the cluster. Use the display name if the cluster manager manages multiple clusters. If you do not provide a cluster name, the wizard imports information based on the default cluster.
Engine type	If you specified a Cloudera cluster, the installer prompts for the engine type. If you are on a CDP cluster, accept the default engine type of Tez. If you are on a CDH cluster, press **2** to set the engine type to MRv2. Default is **1**.

6. To create the Hadoop, HDFS, Hive, HBase, or Databricks connections to the cluster, press **1**.

   The installer appends the connection type to the cluster configuration name to create a connection name.

Configure Enterprise Data Catalog

This task includes installer prompts to configure Enterprise Data Catalog. You will provide basic information for configuring the application services and Hadoop cluster.

When you complete the preliminary tasks, you will continue with the installer prompts to configure Enterprise Data Catalog.

Configure Profiling Warehouse Database Details

If you chose to configure the service parameters, you can provide warehouse information.

1. Select the database type for the profiling warehouse.

   The following table lists the database type for the profiling warehouse.

Prompt	Description
Database type	Type of database for the profiling warehouse connection. Select from the following options: 1 - Oracle 2 - Microsoft SQL Server 3 - IBM DB2

2. Enter the properties for the database user account.

   The following table lists the properties for the database user account:

Property	Description
Database user ID	Name for the profiling warehouse database user account.
User password	Password for the profiling warehouse database user account.

3. Based on the database type selected, enter the parameters for the database.
   a. If you select IBM DB2, select whether to configure a tablespace and enter the tablespace name.

      The following table describes the properties that you must configure for the IBM DB2 database:

Property	Description
Configure tablespace	Select whether to specify a tablespace: 1 - No 2 - Yes In a single-partition database, if you select No, the installer creates the tables in the default tablespace. In a multi-partition database, you must select Yes.
Tablespace	Name of the tablespace in which to create the tables. Specify a tablespace that meets the pageSize requirement of 32768 bytes. In a single-partition database, if you select Yes to configure the tablespace, enter the name of the tablespace in which to create the tables. In a multi-partition database, specify the name of the tablespace that resides in the catalog partition of the database.

   b. If you select Microsoft SQL Server, enter the schema name for the database.

      The following table describes the properties that you must configure for the Microsoft SQL Server database:

Property	Description
Schema name	Name of the schema that will contain domain configuration tables. If this parameter is blank, the installer creates the tables in the default schema.

   c. To enter the JDBC connection information using the JDBC URL information, press **1**. To enter the JDBC connection information using a custom JDBC connection string, press **2**.

   d. Enter the JDBC connection information.

      - To enter the connection information using the JDBC URL information, specify the JDBC URL properties.

The following table describes the database connection information:

Prompt	Description
Database address	Host name and port number for the database.
Database service name	Service or database name : - Oracle: Enter the service name. - Microsoft SQL Server: Enter the database name. - IBM DB2: Enter the service name.
Configure JDBC Parameters	Select whether to add additional JDBC parameters to the connection string: 1 - Yes 2 - No If you select Yes, enter the parameters or press Enter to accept the default. If you select No, the installer creates the JDBC connection string without parameters.

- To enter the connection information using a custom JDBC connection string, type the connection string.
  Use the following syntax in the JDBC connection string:
  IBM DB2

  ```
  jdbc:Informatica:db2://host_name:port_no;DatabaseName=
  ```

  Oracle

  ```
  jdbc:Informatica:oracle://host_name:port_no;ServiceName=
  ```

  Microsoft SQL Server

  ```
  jdbc:Informatica:sqlserver://
  host_name:port_no;SelectMethod=cursor;DatabaseName=
  ```

  Microsoft Azure SQL Database

  ```
  jdbc:Informatica:sqlserver://
  host_name:port_no;SelectMethod=cursor;DatabaseName=database_name;SnapshotSeria
  lizable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net
  ;ValidateServerCertificate=false
  ```

  Verify that the connection string contains all the connection parameters required by your database system.

e. Enter the data access connection string.

The **Content Management Service Parameters and Database** section appears.

Configure the Content Management Service Parameters and Database

After you configure the profiling warehouse, you can configure the content management service parameters and database properties.

1. Enter the name of the Content Management Service.

2. Enter the following service parameter information:

Port	Description
HTTP protocol type	Type of connection to the Data Integration Service. Select one of the following options: - HTTP. Requests to the service uses an HTTP connection. - HTTPS. Requests to the service uses a secure HTTP connection.
HTTP port	Port number to used for the Content Management Service. Default is 8105.
HTTPS port	Port number to used for the Content Management Service. Default is 8105.

3. If you select a keystore for the Content Management Service, enter the keystore file and port number for the HTTPS connection to the Content Management Service.

   Select whether to use a keystore file generated by the installer or a keystore file you create. You can use a keystore file with a self-signed certificate or a certificate signed by a certification authority.

   - Use the default keystore generated by the installer.
   - Specify the location and password of a custom keystore file.

   If you select to use a keystore file generated by the installer, the installer creates a self-signed keystore file named Default.keystore in the following location: `<Informatica installation directory>/tomcat/conf/`

4. Enter database information for the reference data warehouse.

   The following table lists the database information for the reference data warehouse.

Prompt	Description
Database type	Type of database for reference data warehouse. Select from the following options: 1 - Oracle 2 - Microsoft SQL Server 3 - IBM DB2

5. Enter the properties for the database user account.

   The following table lists the properties for the database user account:

Property	Description
Database user ID	Name for the reference data warehouse database user account.
User password	Password for the profiling warehouse database user account.

6. Based on the database type selected, enter the parameters for the database .
   a. If you select IBM DB2, select whether to configure a tablespace and enter the tablespace name.

The following table describes the properties that you must configure for the IBM DB2 database:

Property	Description
Configure tablespace	Select whether to specify a tablespace: 1 - No 2 - Yes In a single-partition database, if you select No, the installer creates the tables in the default tablespace. In a multi-partition database, you must select Yes.
Tablespace	Name of the tablespace in which to create the tables. Specify a tablespace that meets the pageSize requirement of 32768 bytes. In a single-partition database, if you select Yes to configure the tablespace, enter the name of the tablespace in which to create the tables. In a multi-partition database, specify the name of the tablespace that resides in the catalog partition of the database.

b. If you select Microsoft SQL Server, enter the schema name for the database.

The following table describes the properties that you must configure for the Microsoft SQL Server database:

Property	Description
Schema name	Name of the schema that will contain domain configuration tables. If this parameter is blank, the installer creates the tables in the default schema.

c. To enter the JDBC connection information using the JDBC URL information, press **1**. To enter the JDBC connection information using a custom JDBC connection string, press **2**.

d. Enter the JDBC connection information.

- To enter the connection information using the JDBC URL information, specify the JDBC URL properties.

The following table describes the database connection information:

Prompt	Description
Database address	Host name and port number for the database.
Database service name	Service or database name : - Oracle: Enter the service name. - Microsoft SQL Server: Enter the database name. - IBM DB2: Enter the service name.
Configure JDBC Parameters	Select whether to add additional JDBC parameters to the connection string: 1 - Yes 2 - No If you select Yes, enter the parameters or press Enter to accept the default. If you select No, the installer creates the JDBC connection string without parameters.

- To enter the connection information using a custom JDBC connection string, type the connection string.
  Use the following syntax in the JDBC connection string:
  IBM DB2

  ```
  jdbc:Informatica:db2://host_name:port_no;DatabaseName=
  ```

  Oracle

  ```
  jdbc:Informatica:oracle://host_name:port_no;ServiceName=
  ```

  Microsoft SQL Server

  ```
  jdbc:Informatica:sqlserver://
  host_name:port_no;SelectMethod=cursor;DatabaseName=
  ```

  Microsoft Azure SQL Database

  ```
  jdbc:Informatica:sqlserver://
  host_name:port_no;SelectMethod=cursor;DatabaseName=database_name;SnapshotSeria
  lizable=true;EncryptionMethod=SSL;HostNameInCertificate=*.database.windows.net
  ;ValidateServerCertificate=false
  ```

  Verify that the connection string contains all the connection parameters required by your database system.

e. Enter the data access connection string.

The **Cluster and Application Service Options** section appears.

Create and Configure the Informatica Cluster Service

1. Configure the following properties to create the Informatica Cluster Service:

Property	Description
User name for the gateway host	User name for the gateway host. The gateway user must be a non-root user with sudo access. You must enable passwordless SSH for the following nodes: - Between the Informatica domain and the gateway host for the gateway user. - Between gateway host and data nodes and processing nodes. - If you plan to enable Advanced Configuration for the service, enable passwordless SSH between the gateway node and service nodes.
Enter the Informatica Cluster Service Name	Name of the service. The name is not case-sensitive and must be unique within the domain. The name cannot exceed 128 characters or begin with @. The name cannot contain character spaces. The characters in the name must be compatible with the code page of the Model repository. The name cannot contain the following special characters:` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Informatica Cluster Gateway Host	Fully qualified domain name of the node that you want to configure as the gateway host. The node that you configure as the gateway host must be a data node or a processing node.
Enable Advanced Configuration	Select `2` if you want to configure the properties of the applications and associated services. If you select `2`, the associated services use the values that you specify for them instead of using the values specified for the data nodes. Provide the parameters listed in the #unique_172 section.
Data Nodes	Comma-separated list of fully qualified domain names of nodes that you want to configure as data nodes.
Processing Nodes	Comma-separated list of fully qualified domain names of nodes that you want to configure as processing nodes.
Enable Transport Layer Security (TLS)	Select the option to enable TLS for the Informatica Cluster Service.
Enter the Informatica Cluster Service HTTPS Port	Port number for the HTTPS connection. Required if you selected **Enable Transport Layer Security**.
Enable Secure Communication for the Service	Select `1` to use to default Informatica SSL certificates or select 2 to use the custom SSL certificates. If the Informatica domain is enabled for SSL, you must provide the following details: - HTTPS port. The HTTPS port to access the Informatica domain node. - Informatica keystore file. The fully qualified path to the Informatica domain keystore file. - Keystore password. The password for the keystore file.
SSL protocol to use	Optional. Provide the SSL protocol that you want to use for the service.
Cluster Custom Directory	Directory for the service. Default is `/opt/informatica/ics`

2. Press **Enter**.

Create and Configure the Catalog Service

1. Configure the following properties to create the Catalog Service:

Property	Description Catalog
Catalog Service name	Name of the service. The name is not case-sensitive and must be unique within the domain. The name cannot exceed 128 characters or begin with @. The name cannot contain character spaces. The characters in the name must be compatible with the code page of the Model repository that you associate with the Catalog Service. The name cannot contain the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Catalog Service HTTPS port	Applies if the Informatica domain is SSL enabled. Port number for the HTTPS connection.
Do you want to enable secure communication for the service?	Select 1 to use the default SSL certificates or select 2 to use custom SSL certificates. If the Informatica domain is enabled for SSL, you must provide the following details: - HTTPS port. The HTTPS port to access the Informatica domain node. - Informatica keystore file. The fully qualified path to the Informatica domain keystore file. - Keystore password. The password for the keystore file.
Do you want to enable Asset Change Email Notifications?	Select 2 to receive email notifications when there are updates for assets.
Do you want to enable Data Asset Analytics?	Select 2 to enable Data Asset Analytics.

2. Press **Enter**.
3. Configure the following properties if you enabled Data Asset Analytics:

Property	Description
Select Database	Select the repository database that you want to use for Data Asset Analytics from the following options: - Oracle - SQLServer - PostgreSQL
User Name	The database user name for the repository.
Password	The password for the database user name.

Property	Description

Database Connection String	Enter the JDBC connection string to connect to the repository database.
	Use the following syntax for the connection string based on the database selected:
	- Oracle.
	`jdbc:informatica:oracle://<host name>:<port number>;ServiceName=<database name>`
	- SQLServer:
	  - SQL Server.
	`jdbc:informatica:sqlserver://<host name>: <port number>;DatabaseName=<database name>;SnapshotSerializable=true`
	- PostgreSQL.
	`jdbc:informatica:postgresql://<host name>: <port number>;DatabaseName=<database name>`

Secure JDBC Parameters	If the repository database is secured with the SSL protocol, you must enter the secure database parameters as name=value pairs separated by semicolon characters (;). For example: `param1=value1;param2=value2`

Configure the Advanced Scanners Repository

Configure the following properties to create the repository for Advanced Scanners.

▶ If you selected the option to configure the repository for Advanced Scanners, configure the following properties:

Property	Description
Database type	Specify the database that you want to configure as the repository for Advanced Scanners. You can select from the following options: - Oracle - SQLServer
Database user name	User name configured to access the database before upgrade.
Database password	Password configured for the user name.
Schema name	Not applicable if you selected Oracle as the database. Specify the schema that you want to use.
Is the database secure?	Specify if the database is enabled for SSL.
Path to the database truststore file	Applies if you specified that the database is enabled for SSL. Provide the fully qualified path to the database truststore file.
Truststore password	Applies if you specified that the database is enabled for SSL Password to access the truststore file.

Property	Description
Secure JDBC parameters	Applies if you specified that the database is enabled for SSL. Specify the secure database parameters as shown in the following sample: `EncryptionMethod=SSL;HostNameInCertificate=ORATLS.informatica.com; ValidateServerCertificate=true;CryptoProtocolVersion=TLSv1.2;` **Note:** If you want to use secure parameters that are not listed by default, add them using the separator (;). Enterprise Data Catalog appends the secure JDBC parameters to the JDBC connect string. If you include the secure JDBC parameters directly in the connection string, do not enter any parameter in the Secure JDBC Parameters field.
JDBC connect string	The JDBC connection string to connect to the database. Use the following formats to specify the connection strings: - Oracle. jdbc:Informatica:oracle ://<host name>:<port number>;ServiceName= - SQLServer. jdbc:Informatica: sqlserver://<host name>:<portnumber> ;SelectMethod=cursor;DatabaseName=
Advanced Scanners repository server port	A unique port for the Advanced Scanners repository server. Default is 48090.
Enable secure communication for the Advanced Scanners repository ?	Specify if TLS is enabled for communication with the repository.
SSL certificates to secure the Advanced Scanner repository?	Specify if you want to use the default SSL certificates or custom SSL certificates to secure the repository. Specify the following properties if you want to use custom SSL certificates: - Keystore type. Specify the type of keystore. - Keystore alias. Unique alias to access the keystore entry. - Keystore file. Specify the keystore file that contains the required keys and certificates. - Keystore password. Password to access the keystore file.

Configure Enterprise Data Preparation

This task includes installer prompts to configure Enterprise Data Preparation. You will provide basic information for configuring the application services, Hadoop cluster, and creating the Enterprise Data Preparation services.

When you complete the tasks, you will complete the installation.

Configure the Model Repository Service and Model Repository Database Details

Choose whether to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation. If you create a Model Repository Service, specify the connection details for the Model repository database.

If you choose to create a Model Repository Service, specify the connection details for the Model repository database.

1. Choose to either create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation.
   - To create to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service on the node, press **1**.
   - To associate an existing Model Repository Service and an existing Data Integration Service with the Enterprise Data Preparation Service, press **2**, and then enter the name of each service.
2. Enter the name of the Model Repository Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
3. Specify the connection details for the Model repository database.

   The following table describes the parameters you set:

Property	Description
Database Type	Database for the Model repository managed by the Model Repository Service.
Database User ID	User name of the database user account to use to log in to the Model repository database.
User Password	Password for the Model repository database user account.
Tablespace	Configure for a IBM DB2 database. Name of the tablespace in which to create the tables. The tablespace must be defined on a single node, and the page size must be 32K. This option is required for a multi-partition database. If this option is not selected for a single-partition database, the installer creates the tables in the default tablespace.
Schema Name	Configure for a Microsoft SQL Server database. Name of the schema that will contain domain configuration tables. If not selected, the installer creates the tables in the default schema.

4. Specify the truststore details required to access a secure Model repository database.

   The following table describes the properties you set:

Property	Description
Database truststore file	Path and file name of the truststore file for the secure database.
Database truststore password	Password for the truststore.

5. Choose whether to configure the database connection using a JDBC URL or a custom JDBC connection string.

   - Press **1** to configure the database connection using a JDBC URL.
     The following table describes the properties you set:

Property	Description
Database address	Host name and port number for the database in the format <host name>:<port>.
Database service name	Service name for Oracle and IBM DB2 databases, or database name for Microsoft SQL Server.
JDBC parameters	Optional parameters to include in the database connection string. Use the parameters to optimize database operations for the Model repository. You can use the default parameters or add or modify the parameters based on your database requirements. Verify that the parameter string is valid. The installer does not validate the parameter string before it adds the string to the JDBC URL. If not selected, the installer creates the JDBC URL without additional parameters. The following examples show the default connection strings for each database: - Oracle. jdbc:Informatica:oracle://host_name:port_no ;ServiceName= - IBM DB2. jdbc:Informatica:db2://host_name:port_no ;DatabaseName= - Microsoft SQL Server. jdbc:Informatica:sqlserver://host_name:port_no ;SelectMethod=cursor;DatabaseName= - Microsoft Azure. jdbc:informatica:sqlserver://host_name:port_number ;DatabaseName=<database_name>;SnapshotSerializable=true;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.<hostnameincertificate>;ValidateServerCertificate=true - PostgreSQL. jdbc:informatica:postgresql://<host name>:<port number>;DatabaseName=

   - Press **2** to configure the database connection using a custom JDBC connection string.
     The following table describes the properties you set:

Property	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that is sent by the database server. If this parameter is set to True, Informatica validates the certificate that is sent by the database server. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to false, Informatica does not validate the certificate that is sent by the database server. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.

The Data Integration Service section appears.

Configure the Data Integration Service Properties

If you create a Data Integration Service to associate with Enterprise Data Preparation during installation, specify the properties required to create the Data Integration Service.

1. Specify the name of the Data Integration Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. Specify the HTTP protocol type for the Data Integration Service, and then enter the port for each protocol you select.
   - To select HTTP only, press **1**.
   - To select HTTPS only, press **2**.
   - To select both HTTP and HTTPS, press **3**.
3. If you select HTTPS or both HTTP and HTTPS, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in custom keystore and truststore files, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.

The Data Preparation Repository Database section appears.

Configure the Data Preparation Repository Database Details

Specify the Data Preparation repository database connection details. You can choose to use an Oracle database or a MySQL database for the Data Preparation repository database.

If you do not have the database details, you can enter placeholder values, and then create the Interactive Data Preparation Service. If you continue without specifying the database connection details, you cannot enable the Interactive Data Preparation Service.

Oracle

1. To use an Oracle database for the Data Preparation repository, press **1**.
2. Enter the connection properties for the database.

   The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the database.
Database Port Number	Port number for the database.
JDBC Parameters	Parameters required to connect to the database.
Custom JDBC Connection String	JDBC connection string to connect to the database. Format the string as follows: `jdbc:informatica:oracle://<database host name>:<port>;ServiceName=<service name>`

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

The following table describes the secure connection properties:

Property	Description
Truststore File	Path and file name for the database truststore file.
Truststore Password	Password for the database truststore file.
Secure JDBC Parameters	List of secure database parameters to connect to the database. Format the parameters as follows: `EncryptionMethod=SSL;HostNameInCertificate=<secure database host name>;ValidateServerCertificate=true`

4. Press **2** to continue.

The Interactive Data Preparation Service Details section appears.

MySQL

1. To use a MySQL database or a MariaDB database for the Data Preparation repository, press **2**.
2. Enter the connection properties for the database.
   The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the Data Preparation repository database.
Database User Name	Database user account to use to connect to the Data Preparation repository.
Database User Password	Password for the Data Preparation repository database user account.
Database Port Number	Port number for the database.
Database Name	Schema or database name of the Data Preparation repository database.

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

   The following table describes the secure connection properties:

Property	Description

Custom JDBC Connection String	Connection string to connect to the database.
	To connect to a non-secure database, format the string as follows:
	`jdbc:mysql://<database host name>:<port>`
	The connection string is optional if you connect to a non-secure database.
	To connect to an SSL-enabled database, format the string as follows:
	`verifyServerCertificate=true&useSSL=true&requireSSL=true`

Secure JDBC Parameters	String containing the path and file name for the database truststore file, and the truststore password. Format the string as follows: `trustCertificateKeyStoreUrl=file://<truststore path/truststore file name>&trustCertificatekeyStorePassword=<truststore password>`

4. Press **Enter** to continue.

The Interactive Data Preparation Service Details section appears.

Create the Interactive Data Preparation Service

When you install Enterprise Data Preparation on a gateway node, you can create the Interactive Data Preparation Service and the Enterprise Data Preparation Service during installation.

If you do not create the Enterprise Data Preparation Service and the Interactive Data Preparation Service during installation, or if you install Enterprise Data Preparation on another node in the domain, you can use the Administrator tool to create the services after you install the Enterprise Data Preparation binaries.

1. Specify the name of the Interactive Data Preparation Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. If you plan to use rules, you must associate a Model Repository Service and a Data Integration Service with the Interactive Data Preparation Service.
   - To skip associating a Model Repository Service and a Data Integration Service with the service, press **1**.
   - To associate a Model Repository Service and a Data Integration Service with the service, press **2**, and then enter the service names.
3. Choose whether to enable secure communication for the service.
   - To enable secure communication, press **1**.
   - To disable secure communication, press **2**.
4. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.
5. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
6. Select the Hadoop authentication mode.
   - To select the simple authentication mode, press **1**.
   - To select Kerberos authentication, press **2**.
7. If you select Kerberos, enter the authentication parameters.

The following table describes the authentication parameters that you must set if you select Kerberos:

Property	Description
HDFS Principal Name	Service Principal Name (SPN) for the data preparation Hadoop cluster.
Hadoop Impersonation User Name	User name to use in Hadoop impersonation as shown in the Impersonation User Name property for the Hadoop connection in the Administrator tool. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Kerberos Keytab File	Path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the Interactive Data Preparation Service runs.
Fully Qualified Path to the Kerberos Configuration File	Path to the krb5.conf Kerberos configuration file.

8. Specify the HDFS storage location, HDFS connection, and local storage location details.

   The following table describes the properties you must set:

Property	Description
HDFS Storage Location	HDFS location for data preparation file storage. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Local Storage Location	Directory for data preparation file storage on the node where the Interactive Data Preparation Service runs. If the connection to the local storage fails, the service recovers data preparation files from the HDFS storage location.

9. Specify the logging options.

   The following table describes the properties you must set:

Property	Description
Log Severity	Severity of messages to include in the logs. Select from one of the following values: - FATAL. Writes FATAL messages to the log. FATAL messages include nonrecoverable system failures that cause the service to shut down or become unavailable. - ERROR. Writes FATAL and ERROR code messages to the log. ERROR messages include connection failures, failures to save or retrieve metadata, service errors. - WARNING. Writes FATAL, WARNING, and ERROR messages to the log. WARNING errors include recoverable system failures or warnings. - INFO. Writes FATAL, INFO, WARNING, and ERROR messages to the log. INFO messages include system and service change messages. - TRACE. Write FATAL, TRACE, INFO, WARNING, and ERROR code messages to the log. TRACE messages log user request failures. - DEBUG. Write FATAL, DEBUG, TRACE, INFO, WARNING, and ERROR messages to the log. DEBUG messages are user request logs.
Log Directory	Location of the directory of log files.

10. Specify the advanced options.

    The following table describes the properties you must set:

Property	Description
Model Repository Service Name	Name of the Model Repository Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [You cannot change the name of the service after you create it.
Data Integration Service Name	Name of the Data Integration Service.
HDFS Connection	HDFS connection for data preparation file storage.

11. Choose whether to enable the service.
    - If you want to enable the service at a later time using the Administrator tool, press **1**.
    - If you want the installer to enable the service after you complete the installation, press **2**.

The Enterprise Data Preparation Service section appears.

Create the Enterprise Data Preparation Service

When you install Enterprise Data Preparation on the master gateway node for the domain, you can create the Enterprise Data Preparation Service and the Interactive Data Preparation Service during installation.

If you do not create the Enterprise Data Preparation Service and the Interactive Data Preparation Service during installation, or if you install Enterprise Data Preparation on another gateway node in the domain, you can use the Administrator tool to create the services after you install the Enterprise Data Preparation binaries.

1. Specify the details for the Enterprise Data Preparation Service.

   The following table describes the properties that you set:

Property	Description
Enterprise Data Preparation Service Name	Name of the Enterprise Data Preparation Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Interactive Data Preparation Service Name	Name of the Interactive Data Preparation Service to associate with the Enterprise Data Preparation Service.
Model Repository Service Name	Name of the Model Repository Service to associate with the Enterprise Data Preparation Service.
Data Integration Service Name	Name of the Data Integration Service associated with the Enterprise Data Preparation Service.

2. Choose whether to enable secure communication for the service.

- To enable secure communication, press **1**.
- To disable secure communication, press **2**.

3. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.
4. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
5. Specify the data lake connection properties.

   The following table describes the properties that you set for the data lake connections:

Property	Description
HDFS Connection	HDFS connection for the data lake. If you selected the option to create the connection when creating the cluster configuration, the installer sets this value to the name created for the connection.
HDFS Working Directory	HDFS directory where the Enterprise Data Preparation Service copies temporary data and files necessary for the service to run.
Hadoop Connection	Hadoop connection for the data lake. If you selected the option to create the connection when creating the cluster configuration, the installer sets this value to the name created for the connection.

6. Select the Hadoop authentication mode.
   - To select the simple authentication mode, press **1**.
   - To select Kerberos authentication, press **2**.
7. If you select Kerberos, enter the authentication parameters.

   The following table describes the authentication properties that you set if you select Kerberos:

Property	Description
Kerberos Principal	If the Hadoop cluster uses Kerberos authentication, specify the Service Principal Name (SPN) of the user account to impersonate when connecting to the data lake Hadoop cluster.
Kerberos KeyTab File	If the Hadoop cluster uses Kerberos authentication, specify the path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the Enterprise Data Preparation runs.
Local System Directory	Local directory that contains the files downloaded from the Enterprise Data Preparation application, such as .csv and .tde files.

8. The following table describes the authentication parameters that you set if you select Kerberos:
   - If you want to enable the service at a later time using the Administrator tool, press **1**.
   - If you want the installer to enable the service after you complete the installation, press **2**.
9. Choose whether to enable logging of user activity events.

- To disable logging of user activity events, press **1**.
- To enable logging of user activity events, press **2**.

CHAPTER 9

Join a Domain with Catalog and Data Engineering Products

This chapter includes the following topics:

Begin the Installation

This task includes installer prompts to begin the installation. You will provide basic information such as acceptance of terms, installation option, and the installation directory.

When you complete the preliminary tasks, you will continue with the installer prompts and will provide information to configure the domain.

Run the Installer

Perform the following steps to run the installer:

1. Log in to the machine with a system user account.
2. Close all other applications.
3. On a shell command line, run the `install.sh` file.

   The installer displays the message to verify that the locale environment variables are set.
4. If the environment variables are not set, press **n** to exit the installer and set them as required.

   If the environment variables are set, press **y** to continue.

Welcome - Accept Terms and Conditions

- Read the terms and conditions for Informatica installation and the product usage toolkit and select **I agree to the terms and conditions**.
  a. Press **1** if you do not want to accept the terms and conditions.
  b. Press **2** to accept the terms and conditions.

The **Component Selection** sections appears.

Choose the Installation Option

After you accept terms and conditions, you can install Informatica domain services supporting Data Engineering and Catalog products.

1. Press **3** to install Informatica Enterprise Data Preparation.

   When you select this option, you can choose to install Only Enterprise Data Preparation, Enterprise Data Catalog and Enterprise Data Preparation, and Informatica domain services, supporting Data Engineering services, Enterprise Data Catalog, and Enterprise Data Preparation.
2. Select whether current version of the Informatica domain services is installed on the node.
   a. Press **1** if current version of the Informatica domain services is not installed on the node.
   b. Press **2** if current version of the Informatica domain services is installed on the node.
3. Select whether you have read and accepted terms and conditions to use Java SE Development Kit software.
   a. Press **1** to not accept the terms and conditions to use Java SE Development Kit software.
   b. Press **2** to accept the terms and conditions to use Java SE Development Kit software.

The **Installation Prerequisites** section displays the installation requirements. Verify that all requirements are met before you continue the installation.

Tune the Application Services

After you review the installation prerequisites, you can choose to tune the application services for better performance based on the deployment type in your environment. If you do not tune now, you can tune the services later through infacmd.

1. Select if you want the installer to tune the application services:
   - Press 1 if you do not want to tune the services.
   - Press 2 if you want to tune the services.

   If you are joining the node to existing domain, ensure the deployment type you select here is same deployment type as the gateway nodes.
2. Select the deployment type associated with the Informatica environment.

Deployment Type	Disk Space per Node	Total Virtual Cores	RAM per Node
Sandbox	140 GB	16	32 GB
Basic	140 GB	24	64 GB
Standard	140 GB	48	64 GB
Advanced	140 GB	96	128 GB

3. Select whether you want to change the deployment type or continue with the current deployment selection.
   a. Press **1** to change the deployment type.
   b. Press **2** to continue with the current deployment selection.

The **License and Installation Directory** section appears.

Specify the Installation Directory

After you verify the installation prerequisites, you can specify the installation directory.

1. Enter the installation directory.

   The directory names in the path must not contain spaces or the following special characters: @|* $ # ! % ( ) { } [ ] , ; ' Default is /<home>/Informatica/10.5.1.

   Note: Informatica recommends using alphanumeric characters in the installation directory path. If you use a special character such as á or €, unexpected results might occur at run time.
2. Enter the path to the license key file and press **Enter**.
3. Choose an installation environment and press **Enter**.
   - Press **1** to set Sandbox environment for a basic environment used for proof of concept with minimal users.
   - Press **2** to set Development environment for the design environment.
   - Press **3** to set Test environment for high volume processing that is closest to a production environment.
   - Press **4** to set Production environment for high volume processing with high levels of concurrency meant for end users. Advanced production environments are typically multi-node setups.

   Default is 1 for Sandbox.
4. Select whether you want to run the pre-validation utility.
   a. Press **1** to skip the pre-validation utility.
   b. Press **2** to run the pre-validation utility.

If you choose to skip the pre-validation utility, the **Pre-Installation Summary** section appears. Review the installation summary.

If choose to run the pre-validation utility, the Pre-validation section appears.

Configure the Domain

This task includes installer prompts to configure the domain. You will provide information to join a domain, configure the domain security, domain repository, and the encryption key for the domain.

When you complete the tasks, you will complete the installation.

Configure the Domain

After you review the Pre-Installation summary, you can enter the domain information.

1. Press **2** to join a domain.

   The installer joins a node on the machine where you install.
2. Specify whether the domain you want to join has the secure communication option enabled.

   Press 1 to join an insecure domain or press 2 to join a secure domain.
3. Select the type of node you want to create.

The following table describes that types of nodes that you can create:

Property	Description
Configure this node as a gateway	Select whether to configure the node as a gateway or worker node. 1 - Yes 2 - No Select 1 to configure a gateway node or 2 to configure a worker node.

If you configure the node as a gateway, you can enable a secure HTTPS connection to the Informatica Administrator.

4. Specify the connection details to Informatica Administrator.
   a. Specify whether to set up a secure HTTPS connection to the Informatica Administrator.

Option	Description
1 - Enable HTTPS for Informatica Administrator	Set up a secure connection to Informatica Administrator.
2 - Disable HTTPS	Do not set up a secure connection to Informatica Administrator.

   b. If you enable HTTPS connection for the Informatica Administrator, enter the keystore file and port number to use to secure the connection.

Option	Description
Port	Port number for the HTTPS connection.
Keystore file	Select whether to use a keystore file generated by the installer or a keystore file you create. You can use a keystore file with a self-signed certificate or a certificate signed by a certification authority. 1 - Use a keystore generated by the installer 2 - Specify a keystore file and password If you select to use a keystore file generated by the installer, the installer creates a self-signed keystore file named Default.keystore in the following location: `<Informatica installation directory>/tomcat/conf/`

   c. If you specify the keystore, enter the password and location of the keystore file.
   d. If you enabled secure communication for the domain, the **Domain Security - Secure Communication** section appears.
   e. If you did not enable secure communication for the domain, the **Domain Configuration Repository** section appears. Skip to "Configure the Domain Repository" on page 148.

5. Select if SAML authentication is enabled to configure Security Assertion Markup Language (SAML)-based single sign-on (SSO) support for web-based Informatica applications in an Informatica domain.

The following table describes the information you must enter to enable SAML authentication:

Prompt	Description
Does the domain use SAML authentication?	Select if the domain uses SAML authentication: 1 - No If you select No, skip to "Domain Security" on page 147 2 - Yes If you select Yes, configure the SAML authentication.

6. Enter the Identity Provider URL for the domain.
7. Enter the identity provider assertion signing certificate alias name.
8. Select whether to use the default Informatica SSL certificates or to use your SSL certificates to enable SAML authentication in the domain.

   The following table describes the SSL certificate options for SAML authentication:

Option	Description
Use the default Informatica SSL certificate file	Select to use the default Informatica truststore file for SAML authentication.
Enter the location of the SSL certificate file	Select to use a custom truststore file for SAML authentication. Specify the directory containing the custom truststore file on gateway nodes within the domain. Specify the directory only, not the full path to the file.

9. If you provide the security certificates, specify the location and passwords of the keystore and truststore files.

   The following table describes the location and password of the truststore file:

Property	Description
Truststore Directory	Specify the directory containing the custom truststore file on gateway nodes within the domain. Specify the directory only, not the full path to the file.
Truststore Password	The password for the custom truststore file.

The **Domain Security - Secure Communication** appears.

Domain Security

After you configure the domain, you can configure domain security.

- In the Domain Security - Secure Communication section, specify whether to use the default Informatica SSL certificates or to use your SSL certificates to secure domain communication.
  a. Select the type of SSL certificates to use.

The following table describes the options for the SSL certificates that you can use to secure the Informatica domain:

Option	Description
Use the default Informatica SSL certificates	Use the default SSL certificates contained in the default keystore and truststore. **Note:** If you do not provide an SSL certificate, Informatica uses the same default private key for all Informatica installations. If you use the default Informatica keystore and truststore files, the security of your domain could be compromised. To ensure a high level of security for the domain, select the option to specify the location of the SSL certificate files.
Use custom SSL certificates	Specify the path for the keystore and truststore files that contain the SSL certificates. You must also specify the keystore and truststore passwords. You can provide a self-signed certificate or a certificate issued by a certificate authority (CA). You must provide SSL certificates in PEM format and in Java Keystore (JKS) files. Informatica requires specific names for the SSL certificate files for the Informatica domain. You must use the same SSL certificates for all nodes in the domain. Store the truststore and keystore files in a directory accessible to all the nodes in the domain and specify the same keystore file directory and truststore file directory for all nodes in the same domain.

b. If you provide the SSL certificate, specify the location and passwords of the keystore and truststore files.

The following table describes the parameters that you must enter for the SSL certificate files:

Property	Description
Keystore file directory	Directory that contains the keystore files. The directory must contain files named infa_keystore.jks.
Keystore password	Password for the keystore infa_keystore.jks.
Truststore file directory	Directory that contains the truststore files. The directory must contain files named infa_truststore.jks and infa_truststore.pem.
Truststore password	Password for the infa_truststore.jks file.

The **Domain Configuration Repository** section appears.

Configure the Domain Repository

After you configure the domain, you can configure domain repository.

- Enter the information for the domain that you want to join.

The following table describes the properties that you specify for the domain:

Property	Description
Domain name	Name of the domain to join.
Gateway node host	Host name of the machine that hosts the gateway node for the domain.
Gateway node port	Port number of the gateway node.
Domain user name	User name of the administrator for the domain you want to join.
Domain password	Password for the domain administrator.

The **Domain Security - Encryption Key** section appears.

Domain Security - Encryption Key

After you configure the domain repository, you can configure the encryption key.

- Enter the directory for the encryption key for the Informatica domain.

  The following table describes the encryption key parameters that you must specify when you join a domain:

Prompt	Description
Select the encryption key	Path and file name of the encryption key for the Informatica domain that you want to join. All nodes in the Informatica domain use the same encryption key. You must specify the encryption key file created on the gateway node for the domain that you want to join. If you copied the encryption key file to a temporary directory to make it accessible to the nodes in the domain, specify the path and file name of the encryption key file in the temporary directory.
Encryption key directory	Directory in which to store the encryption key on the node created during this installation. The installer copies the encryption key file for the domain to the encryption key directory on the new node.

  The installer sets different permissions to the directory and the files in the directory. For more information about the permissions for the encryption key file and directory, see "Secure Files and Directories" on page 91.

The **Join Domain Node Configuration** section appears.

Configure the Domain and Node

After you configure the encryption key, you can configure the domain and node.

1. Enter the information for the domain and the node that you want to join.

The following table describes the properties that you set for the domain and gateway node.

Property	Description
Node Host name	Host name for the node. The node host name cannot contain the underscore (_) character.Note: Do not use localhost. The host name must explicitly identify the machine.
Node name	Name of the node to join.
Node port number	Port number for the node. The default port number for the node is 6005. If the port number is not available on the machine, the installer displays the next available port number.
Database truststore file	Path and file name of the truststore file for the secure database. Select the same database truststore file used by the master gateway node in the domain. Available when you join a gateway node to a domain that uses a domain configuration repository database that is secured with the SSL protocol.
Truststore password	Password for the database truststore file for the secure database. Available when you join a gateway node to a domain that uses a domain configuration repository database that is secured with the SSL protocol.

2. Select whether to display the default ports for the domain and node components assigned by the installer.

   The following table describes the advanced port configuration page:

Prompt	Description
Display advanced port configuration page	Select whether to display the port numbers for the domain and node components assigned by the installer: 1 - No 2 - Yes If you select Yes, the installer displays the default port numbers assigned to the domain components. You can specify the port numbers to use for the domain and node components. You can also specify a range of port numbers to use for the service process that will run on the node. You can use the default port numbers or specify new port numbers. Verify that the port numbers you enter are not used by other applications.

3. If you display the port configuration page, enter new port numbers at the prompt or press Enter to use the default port numbers.

The following table describes the ports that you can set:

Port	Description
Service Manager port	Port number used by the Service Manager on the node. The Service Manager listens for incoming connection requests on this port. Client applications use this port to communicate with the services in the domain. The Informatica command line programs use this port to communicate to the domain. This is also the port for the SQL data service JDBC/ODBC driver. Default is 6006.
Service Manager Shutdown port	Port number that controls server shutdown for the domain Service Manager. The Service Manager listens for shutdown commands on this port. Default is 6007.
Informatica Administrator port	Port number used by Informatica Administrator. Default is 6008.
Informatica Administrator HTTPS port	No default port. Enter the required port number when you create the service. Setting this port to 0 disables an HTTPS connection to the Administrator tool.
Informatica Administrator shutdown port	Port number that controls server shutdown for Informatica Administrator. Informatica Administrator listens for shutdown commands on this port. Default is 6009.
Minimum port number	Lowest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6014.
Maximum port number	Highest port number in the range of dynamic port numbers that can be assigned to the application service processes that run on this node. Default is 6114.

The **Post-Installation Summary** section indicates whether the installation completed successfully. The summary also shows the status of the installed components and their configuration.

CHAPTER 10

Install Enterprise Data Catalog and Enterprise Data Preparation Binaries

This chapter includes the following topics:

Overview

If Data Engineering products are installed and configured in a domain, you can install the Enterprise Data Catalog and Enterprise Data Preparation binaries on a node in the domain and configure the associated services.

Complete the following tasks to install and configure the Catalog products on an existing node in the domain:

1. Perform prerequisite tasks for application services and databases.
2. Run the installer to install the Enterprise Data Catalog and Enterprise Data Preparation binaries on a node.
3. Run the installer again to configure the services associated with Enterprise Data Catalog and Enterprise Data Preparation.

Complete the Prerequisites

The installation of Enterprise Data Catalog and Enterprise Data Preparation consists of multiple phases.

When you plan to install Enterprise Data Catalog and Enterprise Data Preparation, you must account for the dependencies for each product. You also must plan the relational databases that each product requires.

See the following list of services and review the requirements associated with the services:

Informatica Cluster Service

The Informatica Cluster Service is an application service that runs and manages the nodes and services associated with Enterprise Data Catalog. For more information, see Informatica Cluster Serviceon page 131.

Catalog Service

The Catalog Service manages connections between service components and the users that have access to Enterprise Data Catalog search interface and Catalog Administrator. For more information, see "Catalog Service" on page 43.

Data Integration Service

The Data Integration Service receives requests from Informatica client tools to run integration, profile, and data preparation jobs. It writes results to different databases, and it writes run-time metadata to the Model repository. You also need to prepare databases associated with the service. For more information, see "Data Integration Service" on page 48.

Model Repository Service

The Model Repository Service manages the Model repository. It receives requests from Informatica clients and application services to store or access metadata in the Model repository. You also need to prepare the Model Repository database. For more information, see Model Repository Service "Model Repository Service" on page 57.

Interactive Data Preparation Service

Review the Interactive Data Preparation Service requirements.

You must prepare the Data Preparation repository database before you create the Interactive Data Preparation Service. For more information, see *Data Preparation Repository Database Requirements*.

If you plan to use rules during data preparation, you can create a Model Repository Service to associate with the Interactive Data Preparation Service when you configure the Enterprise Data Preparation services. Prepare the Model repository database that contains rule metadata before you create the Interactive Data Preparation Service.

Enterprise Data Preparation Service

Review the Enterprise Data Preparation Service requirements. For more information, see *Enterprise Data Preparation Service*.

You can create a Model Repository Service to associate with the Enterprise Data Preparation Service when you configure the Enterprise Data Preparation services. Prepare the Model repository database before you create the Enterprise Data Preparation Service.

Install Binaries for the Catalog Products

You can install the Enterprise Data Catalog and Enterprise Data Preparation binaries on a node on which the Informatica services are installed.

1. On a shell command line, run the `install.sh` file from the root directory.

   The installer displays the message for documentation and copyright information.

2. Press **Y** to continue the installation.

3. Press **1** to install Informatica 10.5.1 products.
4. Press **3** to run the installer.
5. Press **1** to indicate that you do not want to resume an installation.
6. Press **2** to agree to the terms and conditions.
7. Press **3** to install Enterprise Data Preparation.
8. Press **2** to indicate that the Informatica services are installed on the node.
9. Press **1** to indicate that Enterprise Data Catalog is not installed on the node.
10. Press **2** to tune the application services for better performance based on your deployment type.
11. Enter the directory where you want to install the Enterprise Data Catalog and Enterprise Data Preparation binaries.

    To enable Informatica to register Enterprise Data Preparation with the domain, you must install the Enterprise Data Catalog and Enterprise Data Preparation binaries in the same directory on any gateway node.
12. Choose how to proceed if Enterprise Data Preparation is already installed in the specified directory.
    - Press **1** to change the installation directory.
    - Press **2** to overwrite the existing installation.
13. Review the pre-installation summary, and then press **Enter**.
14. Ensure the current node is shut down, and then press **Enter**.

After you install the Enterprise Data Catalog and Enterprise Data Preparation binaries, run the installer to configure the Enterprise Data Catalog and Enterprise Data Preparation services.

For more information about configuring the Enterprise Data Catalog services, see "Create Services for Enterprise Data Catalog" on page 154

For more information about configuring the Enterprise Data Preparation services, see "Configure Enterprise Data Preparation" on page 163.

Create Services for Enterprise Data Catalog

Perform the following steps to create the application services using the installer after you install the Enterprise Data Catalog binaries:

1. Log in to the machine with a system user account.
2. Close all applications running on the machine.
3. On a shell command line, run the `./install.sh` command to start the installer.
4. Press **y** to proceed with the installation.
5. Press **3** to select the option to install the application services for Enterprise Data Catalog or Enterprise Data Lake.
6. Press **2** to agree to the terms and conditions.
7. Press **2** to accept that you want to proceed with the installation of Informatica 10.5.1.
8. Press **1** to configure services for Enterprise Data Catalog.
9. Enter the directory where you installed Enterprise Data Catalog and press **Enter**.

10. Enter the following domain details that you had configured when you installed Enterprise Data Catalog:
    a. Domain name. Provide the name of the Informatica domain that you created and press **Enter**.
    b. Node name. Provide the name of the node that you created on the machine where you installed Enterprise Data Catalog and press **Enter**.
    c. Domain user password. Provide the password you configured for the Informatica domain administrator and press **Enter**.
11. Press **2** to specify that you do not want to create a cluster configuration. To create Enterprise Data Preparation and Interactive Data Preparation services, you must create a cluster configuration in the Informatica domain.
12. Press **1** to create the profiling warehouse connection.
13. Press **1** to confirm that you want to create the Model Repository Service and Data Integration Service.
14. Press **1** if you want to create the monitoring Model Repository Service to monitor the Informatica domain statistics.
15. Press **1** if you want to create the Content Management Service to discover data domains.
16. Press **1** to configure the Content Management Service.
17. You need to create the Informatica Cluster Service to manage the services required to run Enterprise Data Catalog. Press **1** if you want to configure the Informatica Cluster Service.
    See the following points to decide how you want to create the Informatica Cluster Service:
    - Select 2 to specify that you do not want to create the Informatica Cluster Service. The installer prompts you to specify if you want to associate an Informatica Cluster Service with the Catalog Service. If you select this option, the installer does not create a new Informatica Cluster Service. The installer prompts you for an Informatica Cluster Service that you want to associate with the Catalog Service.
    - Select the options to specify that you do not want to create the Informatica Cluster Service and associate an existing Informatica Cluster Service with the Catalog Service. The installer does not create the Informatica Cluster Service and proceeds to create the Catalog Service.
18. Press **1** if you want to configure the Catalog Service.
19. Perform the steps in the following sections to create the application services.

Specifying the Informatica Domain Details

Perform the following steps to specify the Informatica domain details that you had configured:

1. Enter the following domain details that you had configured when you installed Enterprise Data Catalog:
    a. Domain name. Provide the name of the Informatica domain that you created and press **Enter**.
    b. Node name. Provide the name of the node that you created on the machine where you installed Enterprise Data Catalog and press **Enter**.
    c. Domain User Name. Username to access the Informatica domain administrator.
    d. Domain user password. Provide the password you configured for the Informatica domain administrator and press **Enter**.

Creating the Model Repository Service

Perform the following steps to create the Model Repository Service:

1. Name of the Model Repository Service.
2. Name of the node on which the Model Repository Service must run.

3. The license that you want to associate with the Model Repository Service.
4. Select the database that you want to configure for the Model repository from the following options:
   - Oracle
   - SQL Server
   - DB2
   - PostgreSQL

   Default is Oracle.
5. Type the username to access the database in the **Database user ID** parameter and press **Enter**. Default is **admin**.
6. Type the password for the username in the **User password** parameter and press Enter.
7. Press **1** if the database is secured with SSL.
   If you selected the option to specify that the database is SSL-enabled, provide the following parameters:

Secure Database Parameter	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that the database server sends. If this parameter is set to True, Informatica validates the certificate that the database server sends. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to False, Informatica does not validate the certificate that the database server sends. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.
cryptoProtocolVersion	Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to cryptoProtocolVersion=TLSv1.1 or cryptoProtocolVersion=TLSv1.2 based on the cryptographic protocol used by the database server.
TrustStore	Path and file name of the truststore file that contains the SSL certificate for the database. If you do not include the path for the truststore file, Informatica looks for the file in the following default directory: `<Informatica installation directory>/tomcat/bin`
TrustStorePassword	Password for the truststore file for the secure database.

8. Press **1** to specify the JDBC URL to connect to the database.
9. Specify the database address in the following format for the **Database address** parameter: `<Fully qualified domain name of the host>:<port>`
10. Specify the database service name in the following format for the **Database service name** parameter: `<Fully qualified domain name of the service>`
11. Press **1** to specify that you want to configure the JDBC parameters.

12. Specify the required values for the parameters or press **Enter** to apply the default values. Press **Enter** to retain the default values.
Use the following syntax for the connection string of the selected database type:

Database Type	Connection String Syntax
IBM DB2	`jdbc:informatica:db2://<host_name>:<port_number>;DatabaseName=<database_name>;BatchPerformanceWorkaround=true;DynamicSections=3000`

Microsoft SQL Server	- Microsoft SQL Server that uses the default instance
	`jdbc:informatica:sqlserver://<host_name>:<port_number>;DatabaseName=<database_name>;SnapshotSerializable=true`
	- Microsoft SQL Server that uses a named instance
	`jdbc:informatica:sqlserver://<host_name>\<named_instance_name>;DatabaseName=<database_name>;SnapshotSerializable=true`

Oracle	`jdbc:informatica:oracle://<host_name>:<port_number>;SID=<database_name>;MaxPooledStatements=20;CatalogOptions=0;BatchPerformanceWorkaround=true`

The installer validates the node name and license, and then creates the Model Repository Service. The installer proceeds to create the Data Integration Service.

Creating the Data Integration Service

Perform the following steps to create the Data Integration Service:

1. Name of the Data Integration Service.
2. Name of the node on which the Data Integration Service must run.
3. The license that you want to associate with the Data Integration Service.
4. The name of the Model Repository Service that you want to associate with the Data Integration Service.
5. Specify the protocol that you want to use for the service from the following options:
   - http
   - https
   - http&https
     If you select **https** or **http&https** as the protocol for the service, provide the following details:
     1. HTTPS port. Default is 18095.
     2. Specify the SSL certificate that you want to use to secure the Data Integration Service. You can use the default SSL certificates in the default keystore and truststore or use custom SSL certificates. If you choose custom SSL certificates, specify the path that includes the filename of the keystore and truststore files and the passwords to access the keystore and truststore files.
6. Press **1** if you want to enable data engineering recovery for the Data Integration Service.

The installer validates the node name and the license and creates and enables the Data Integration Service. The installer proceeds to create the profiling warehouse.

Configuring the Profiling Warehouse

Provide the following details to configure the database for the profiling warehouse:

1. Name of the Data Integration Service that you want to associate with the profiling warehouse.
2. Select the database that you want to configure for the profiling warehouse from the following options:
   - Oracle
   - SQL Server
   - DB2

   Default is Oracle.
3. Type the username to access the database in the **Database user ID** parameter and press **Enter**. Default is **admin**.
4. Type the password for the username in the **User password** parameter and press Enter.
5. Press **1** if the database is secured with SSL.
   If you selected the option to specify that the database is SSL-enabled, provide the following parameters:

Secure Database Parameter	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that the database server sends. If this parameter is set to True, Informatica validates the certificate that the database server sends. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to False, Informatica does not validate the certificate that the database server sends. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.
cryptoProtocolVersion	Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to cryptoProtocolVersion=TLSv1.1 or cryptoProtocolVersion=TLSv1.2 based on the cryptographic protocol used by the database server.
TrustStore	Path and file name of the truststore file that contains the SSL certificate for the database. If you do not include the path for the truststore file, Informatica looks for the file in the following default directory: `<Informatica installation directory>/tomcat/bin`
TrustStorePassword	Password for the truststore file for the secure database.

6. Press **1** to specify the JDBC URL to connect to the database.
7. Specify the database address in the following format for the **Database address** parameter: `<Fully qualified domain name of the host>:<port>`
8. Specify the database service name in the following format for the **Database service name** parameter: `<Fully qualified domain name of the service>`

9. Press **1** to specify that you want to configure the JDBC parameters.
10. Specify the required values for the parameters or press **Enter** to apply the default values. Press **Enter** to retain the default values.
    Use the following syntax for the connection string of the selected database type:

Database Type	Connection String Syntax
IBM DB2	`jdbc:informatica:db2://<host_name>:<port_number>;DatabaseName=<database_name>;BatchPerformanceWorkaround=true;DynamicSections=3000`

Microsoft SQL Server	- Microsoft SQL Server that uses the default instance
	`jdbc:informatica:sqlserver://<host_name>:<port_number>;DatabaseName=<database_name>;SnapshotSerializable=true`
	- Microsoft SQL Server that uses a named instance
	`jdbc:informatica:sqlserver://<host_name>\<named_instance_name>;DatabaseName=<database_name>;SnapshotSerializable=true`

Oracle	`jdbc:informatica:oracle://<host_name>:<port_number>;SID=<database_name>;MaxPooledStatements=20;CatalogOptions=0;BatchPerformanceWorkaround=true`

The installer creates the data profiling warehouse and proceeds to create the Content Management Service.

Creating the Content Management Service

Provide the following details to create the Content Management Service:

1. Name of the Model Repository Service that you want to associate with the service.
2. Name of the Data Integration Service that you want to associate with the service.
3. Name of the node on which the Content Management Service must run.
4. The license that you want to associate with the Content Management Service.
5. Name of the Content Management Service.
6. Choose to use http or https for the service.
   If you select **https** as the protocol for the service, provide the following details:
   - HTTPS port. Default is 17466.
   - Specify the SSL certificate that you want to use to secure the Content Management Service. You can use the default SSL certificates in the default keystore or use custom SSL certificates. If you choose custom SSL certificates, specify the path that includes the filename of the keystore file and the password to access the keystore file.
7. Select the database that you want to configure for the Content Management Service from the following options:
   - Oracle
   - SQL Server
   - DB2

   Default is Oracle.

8. Type the username to access the database in the **Database user ID** parameter and press **Enter**. Default is **admin**.
9. Type the password for the username in the **User password** parameter and press Enter.
10. Press **1** if the database is secured with SSL.
    If you selected the option to specify that the database is SSL-enabled, provide the following parameters:

Secure Database Parameter	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that the database server sends. If this parameter is set to True, Informatica validates the certificate that the database server sends. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to False, Informatica does not validate the certificate that the database server sends. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.
cryptoProtocolVersion	Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to cryptoProtocolVersion=TLSv1.1 or cryptoProtocolVersion=TLSv1.2 based on the cryptographic protocol used by the database server.
TrustStore	Path and file name of the truststore file that contains the SSL certificate for the database. If you do not include the path for the truststore file, Informatica looks for the file in the following default directory: `<Informatica installation directory>/tomcat/bin`
TrustStorePassword	Password for the truststore file for the secure database.

11. Press **1** to specify the JDBC URL to connect to the database.
12. Specify the database address in the following format for the **Database address** parameter: `<Fully qualified domain name of the host>:<port>`
13. Specify the database service name in the following format for the **Database service name** parameter: `<Fully qualified domain name of the service>`
14. Press **1** to specify that you want to configure the JDBC parameters.
15. Specify the required values for the parameters or press **Enter** to apply the default values. Press **Enter** to retain the default values.

Use the following syntax for the connection string of the selected database type:

Database Type	Connection String Syntax
IBM DB2	`jdbc:informatica:db2://<host_name>:<port_number>;DatabaseName=<database_name>;BatchPerformanceWorkaround=true;DynamicSections=3000`

Microsoft SQL Server	- Microsoft SQL Server that uses the default instance
	`jdbc:informatica:sqlserver://<host_name>:<port_number>;DatabaseName=<database_name>;SnapshotSerializable=true`
	- Microsoft SQL Server that uses a named instance
	`jdbc:informatica:sqlserver://<host_name>\<named_instance_name>;DatabaseName=<database_name>;SnapshotSerializable=true`

Oracle	`jdbc:informatica:oracle://<host_name>:<port_number>;SID=<database_name>;MaxPooledStatements=20;CatalogOptions=0;BatchPerformanceWorkaround=true`

The installer creates and enables the Content Management Service and proceeds to configure the cluster and application service options.

Creating the Catalog Service

Provide the following details to configure the Catalog Service:

1. Name of the Catalog Service.
2. Enter the name of the Model Repository Service that you want to associate with the Catalog Service.
3. Enter the name of the node on which the service must run.
4. HTTP port number for the Catalog Service. Default is 9085.
5. Press **2** to enable asset change notifications to receive email notifications in Enterprise Data Catalog when there are updates for assets.
6. Press **2** if you want to configure the properties to enable Data Asset Analytics for Enterprise Data Catalog. You can use Data Asset Analytics with Enterprise Data Catalog to gain analytical insights into asset details, such as values, enrichment, and collaboration using reports and charts.
   Configure the following properties if you enabled Data Asset Analytics:

Property	Description
Select Database	Select the repository database that you want to use for Data Asset Analytics from the following options: - Oracle - SQLServer - PostgreSQL
User Name	The database user name for the repository.
Password	The password for the database user name.

Property	Description

Database Connection String	Enter the JDBC connection string to connect to the repository database.
	Use the following syntax for the connection string based on the database selected:
	- Oracle.
	`jdbc:informatica:oracle://<host name>:<port number>;ServiceName=<database name>`
	- SQLServer:
	  - SQL Server.
	`jdbc:informatica:sqlserver://<host name>: <port number>;DatabaseName=<database name>;SnapshotSerializable=true`
	- PostgreSQL.
	`jdbc:informatica:postgresql://<host name>: <port number>;DatabaseName=<database name>`

Secure JDBC Parameters	If the repository database is secured with the SSL protocol, you must enter the secure database parameters as name=value pairs separated by semicolon characters (;). For example: `param1=value1;param2=value2`

Configuring the Advanced Scanners Repository

If you selected the option to configure the repository for Advanced Scanners, configure the following properties:

Property	Description
Database type	Specify the database that you want to configure as the repository for Advanced Scanners. You can select from the following options: - Oracle - SQLServer
Database user name	User name configured to access the database before upgrade.
Database password	Password configured for the user name.
Schema name	Not applicable if you selected Oracle as the database. Specify the schema that you want to use.
Is the database secure?	Specify if the database is enabled for SSL.
Path to the database truststore file	Applies if you specified that the database is enabled for SSL. Provide the fully qualified path to the database truststore file.
Truststore password	Applies if you specified that the database is enabled for SSL Password to access the truststore file.
Secure JDBC parameters	Applies if you specified that the database is enabled for SSL. Specify the secure database parameters as shown in the following sample: `EncryptionMethod=SSL;HostNameInCertificate=ORATLS.informatica.com ; ValidateServerCertificate=true;CryptoProtocolVersion=TLSv1.2;` **Note:** Enterprise Data Catalog appends the secure JDBC parameters to the JDBC connect string. If you include the secure JDBC parameters directly in the connection string, do not enter any parameter in the Secure JDBC Parameters field.

Property	Description
JDBC connect string	The JDBC connection string to connect to the database. Use the following formats to specify the connection strings: - Oracle. jdbc:Informatica:oracle ://<host name>:<port number>;ServiceName= - SQLServer. jdbc:Informatica: sqlserver://<host name>:<portnumber> ;SelectMethod=cursor;DatabaseName=
Advanced Scanners repository server port	A unique port for the Advanced Scanners repository server. Default is 48090.
Enable secure communication for the Advanced Scanners repository ?	Specify if TLS is enabled for communication with the repository.
SSL certificates to secure the Advanced Scanner repository?	Specify if you want to use the default SSL certificates or custom SSL certificates to secure the repository. Specify the following properties if you want to use custom SSL certificates: - Keystore type. Specify the type of keystore. - Keystore alias. Unique alias to access the keystore entry. - Keystore file. Specify the keystore file that contains the required keys and certificates. - Keystore password. Password to access the keystore file.

Configure Enterprise Data Preparation

After you install the Enterprise Data Preparation binaries on a node, run the installer to create and enable the Enterprise Data Preparation services on the node.

You can create the Interactive Data Preparation Service and the Enterprise Data Preparation Service during the configuration process. To create the services, the domain must be integrated with the Hadoop environment before you run the installer. For more information about integrating the domain with the Hadoop environment, see the *Data Engineering Integration Guide*.

If you plan to use rules, you must associate the Interactive Data Preparation Service with the Model Repository Service that manages the Model repository that contains the rule objects and metadata. You must also associate a Data Integration Service with the Interactive Data Preparation Service that runs rules during data preparation. You can create a Model Repository Service and Data Integration Service to associate with the Interactive Data Preparation Service, or you can associate existing services with the Interactive Data Preparation Service.

You must associate a Model Repository Service and a Data Integration Service with the Enterprise Data Preparation Service. You can create a Model Repository Service and Data Integration Service to associate with the Enterprise Data Preparation Service, or you can associate existing services with the Enterprise Data Preparation Service.

If you create a Model Repository Service, you must specify the details for the Model repository database used by the Model Repository Service.

Configure the Enterprise Data Preparation Services

When you configure the Enterprise Data Preparation services on a domain node on which the application binaries are already installed, you indicate that domain services and Enterprise Data Catalog are already installed on the node.

1. On a shell command line, run the `install.sh` file from the root directory.
2. Press **3** to configure the Enterprise Data Catalog or Enterprise Data Preparation services.
3. Press **2** to agree to the terms and conditions.
4. Press **2** to continue with the installation.
5. Press **2** to configure the Enterprise Data Preparation services.
6. Press **2** to indicate that the Enterprise Data Catalog services exist on the node.
7. Enter the directory containing the Enterprise Data Preparation binaries.

The Domain Details section appears.

Configure the Domain Details

Provide the domain authentication details.

- Enter the domain administrator user name and password.

The Application Services for Enterprise Data Preparation section appears.

Configure the Model Repository Service and Model Repository Database Details

Choose whether to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation. If you create a Model Repository Service, specify the connection details for the Model repository database.

If you choose to create a Model Repository Service, specify the connection details for the Model repository database.

1. Choose to either create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation.
   - To create to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service on the node, press **1**.
   - To associate an existing Model Repository Service and an existing Data Integration Service with the Enterprise Data Preparation Service, press **2**, and then enter the name of each service.
2. Enter the name of the Model Repository Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
3. Specify the connection details for the Model repository database.

The following table describes the parameters you set:

Property	Description
Database Type	Database for the Model repository managed by the Model Repository Service.
Database User ID	User name of the database user account to use to log in to the Model repository database.
User Password	Password for the Model repository database user account.
Tablespace	Configure for a IBM DB2 database. Name of the tablespace in which to create the tables. The tablespace must be defined on a single node, and the page size must be 32K. This option is required for a multi-partition database. If this option is not selected for a single-partition database, the installer creates the tables in the default tablespace.
Schema Name	Configure for a Microsoft SQL Server database. Name of the schema that will contain domain configuration tables. If not selected, the installer creates the tables in the default schema.

4. Specify the truststore details required to access a secure Model repository database.

   The following table describes the properties you set:

Property	Description
Database truststore file	Path and file name of the truststore file for the secure database.
Database truststore password	Password for the truststore.

5. Choose whether to configure the database connection using a JDBC URL or a custom JDBC connection string.
   - Press **1** to configure the database connection using a JDBC URL.

The following table describes the properties you set:

Property	Description
Database address	Host name and port number for the database in the format <host name>:<port>.
Database service name	Service name for Oracle and IBM DB2 databases, or database name for Microsoft SQL Server.
JDBC parameters	Optional parameters to include in the database connection string. Use the parameters to optimize database operations for the Model repository. You can use the default parameters or add or modify the parameters based on your database requirements. Verify that the parameter string is valid. The installer does not validate the parameter string before it adds the string to the JDBC URL. If not selected, the installer creates the JDBC URL without additional parameters. The following examples show the default connection strings for each database: - Oracle. jdbc:Informatica:oracle://host_name:port_no ;ServiceName= - IBM DB2. jdbc:Informatica:db2://host_name:port_no ;DatabaseName= - Microsoft SQL Server. jdbc:Informatica:sqlserver://host_name:port_no ;SelectMethod=cursor;DatabaseName= - Microsoft Azure. jdbc:informatica:sqlserver://host_name:port_number ;DatabaseName=<database_name>;SnapshotSerializable=true;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.<hostnameincertificate>;ValidateServerCertificate=true - PostgreSQL. jdbc:informatica:postgresql://<host name>:<port number>;DatabaseName=

- Press **2** to configure the database connection using a custom JDBC connection string.
  The following table describes the properties you set:

Property	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that is sent by the database server. If this parameter is set to True, Informatica validates the certificate that is sent by the database server. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to false, Informatica does not validate the certificate that is sent by the database server. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.

The Data Integration Service section appears.

Data Integration Service Details

If you create a Model Repository Service and a Data Integration Service to associate with Enterprise Data Preparation during installation, specify the properties required to create the Data Integration Service.

1. Specify the name of the Data Integration Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. Enter the name of the node where the service runs.
3. Enter the name of the Informatica license to associate with the service.
4. Enter the name of the Model Repository Service to associate with the Data Integration Service.
5. Specify the HTTP protocol type for the Data Integration Service, and then enter the port for each protocol you select.
   - To select HTTP only, press **1**.
   - To select HTTPS only, press **2**.
   - To select both HTTP and HTTPS, press **3**.
6. If you select HTTPS or both HTTP and HTTPS, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in custom keystore and truststore files, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.

The Data Preparation Repository Database section appears.

Configure the Data Preparation Repository Database Details

Specify the Data Preparation repository database connection details. You can choose to use an Oracle database or a MySQL database for the Data Preparation repository database.

If you do not have the database details, you can enter placeholder values, and then create the Interactive Data Preparation Service. If you continue without specifying the database connection details, you cannot enable the Interactive Data Preparation Service.

Oracle

1. To use an Oracle database for the Data Preparation repository, press **1**.
2. Enter the connection properties for the database.

   The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the database.
Database Port Number	Port number for the database.

Property	Description
JDBC Parameters	Parameters required to connect to the database.
Custom JDBC Connection String	JDBC connection string to connect to the database. Format the string as follows: `jdbc:informatica:oracle://<database host name>:<port>;ServiceName=<service name>`

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

   The following table describes the secure connection properties:

Property	Description
Truststore File	Path and file name for the database truststore file.
Truststore Password	Password for the database truststore file.
Secure JDBC Parameters	List of secure database parameters to connect to the database. Format the parameters as follows: `EncryptionMethod=SSL;HostNameInCertificate=<secure database host name>;ValidateServerCertificate=true`

4. Press **2** to continue.

The Interactive Data Preparation Service Details section appears.

MySQL

1. To use a MySQL database or a MariaDB database for the Data Preparation repository, press **2**.
2. Enter the connection properties for the database.
   The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the Data Preparation repository database.
Database User Name	Database user account to use to connect to the Data Preparation repository.
Database User Password	Password for the Data Preparation repository database user account.
Database Port Number	Port number for the database.
Database Name	Schema or database name of the Data Preparation repository database.

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

The following table describes the secure connection properties:

Property	Description

Custom JDBC Connection String	Connection string to connect to the database.
	To connect to a non-secure database, format the string as follows:
	`jdbc:mysql://<database host name>:<port>`
	The connection string is optional if you connect to a non-secure database.
	To connect to an SSL-enabled database, format the string as follows:
	`verifyServerCertificate=true&useSSL=true&requireSSL=true`

Secure JDBC Parameters	String containing the path and file name for the database truststore file, and the truststore password. Format the string as follows: `trustCertificateKeyStoreUrl=file://<truststore path/truststore file name>&trustCertificatekeyStorePassword=<truststore password>`

4. Press **Enter** to continue.

The Interactive Data Preparation Service Details section appears.

Interactive Data Preparation Service Details

Create the Interactive Data Preparation Service. If you run the installer to create the Enterprise Data Preparation Service and the Interactive Data Preparation Service, you must create both services on the same node.

1. Specify the name of the Interactive Data Preparation Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. If you plan to use rules, you must associate a Data Integration Service and a Model Repository Service with the Interactive Data Preparation Service.
   - To skip associating a Model Repository Service and a Data Integration Service with the Interactive Data Preparation Service, press **1**.
   - To associate a Model Repository Service and a Data Integration Service with the Interactive Data Preparation Service, press **2**, and then enter the service names.
3. Enter the name of the Informatica license to associate with the service.
4. Choose whether to enable secure communication for the service.
   - To enable secure communication for the service, press **1**.
   - To disable secure communication, press **2**.
5. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.
6. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
7. Select the Hadoop authentication mode.

- To select the simple authentication mode, press **1**.
- To select Kerberos authentication, press **2**.

8. If you select Kerberos, enter the authentication parameters.

   The following table describes the authentication parameters that you set if you select Kerberos:

Property	Description
HDFS Principal Name	Service Principal Name (SPN) for the data preparation Hadoop cluster.
Hadoop Impersonation User Name	User name to use in Hadoop impersonation as shown in the Impersonation User Name property for the Hadoop connection in the Administrator tool. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Kerberos Keytab File	Path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the service runs.
Fully Qualified Path to the Kerberos Configuration File	Path to the krb5.conf Kerberos configuration file.

9. Specify the HDFS storage location, HDFS connection, and local storage location details.

   The following table describes the properties you set:

Property	Description
HDFS Storage Location	HDFS location for data preparation file storage. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Local Storage Location	Directory for data preparation file storage on the node on which the Interactive Data Preparation Service runs. If the connection to the local storage fails, the service recovers data preparation files from the HDFS storage location.

10. Specify the logging options.

The following table describes the properties you must set:

Property	Description
Log Severity	Severity of messages to include in the logs. Select from one of the following values: - FATAL. Writes FATAL messages to the log. FATAL messages include nonrecoverable system failures that cause the service to shut down or become unavailable. - ERROR. Writes FATAL and ERROR code messages to the log. ERROR messages include connection failures, failures to save or retrieve metadata, service errors. - WARNING. Writes FATAL, WARNING, and ERROR messages to the log. WARNING errors include recoverable system failures or warnings. - INFO. Writes FATAL, INFO, WARNING, and ERROR messages to the log. INFO messages include system and service change messages. - TRACE. Write FATAL, TRACE, INFO, WARNING, and ERROR code messages to the log. TRACE messages log user request failures. - DEBUG. Write FATAL, DEBUG, TRACE, INFO, WARNING, and ERROR messages to the log. DEBUG messages are user request logs.
Log Directory	Location of the directory of log files.

11. Specify the advanced options.

    The following table describes the properties you must set:

Property	Description
Model Repository Service Name	Name of the Model Repository Service. You cannot change the name of the service after you create it.
Data Integration Service Name	Name of the Data Integration Service.
HDFS Connection	HDFS connection for data preparation file storage.

12. Choose whether to enable the Interactive Data Preparation Service.
    - If you want to enable the service at a later time using the Administrator tool, press **1**.
    - If you want the installer to enable the service after you complete the installation, press **2**.

The Enterprise Data Preparation Service Details section appears.

Enterprise Data Preparation Service Details

Create the Enterprise Data Preparation Service. If you run the installer to create the Enterprise Data Preparation Service and the Interactive Data Preparation Service, you must create both services on the same node.

1. Specify the details for the Enterprise Data Preparation Service.

The following table describes the properties that you set:

Property	Description
Enterprise Data Preparation Service Name	Name of the Enterprise Data Preparation Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Interactive Data Preparation Service Name	Name of the Interactive Data Preparation Service to associate with the Enterprise Data Preparation Service.
Model Repository Service Name	Name of the Model Repository Service to associate with the Enterprise Data Preparation Service.
Data Integration Service Name	Name of the Data Integration Service associated with the Enterprise Data Preparation Service

2. Choose whether to enable secure communication for the service.
   - To enable secure communication for the service, press **1**.
   - To disable secure communication, press **2**.
3. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.
4. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
5. Specify the data lake connection options.

   The following table describes the properties that you set for the data lake connections:

Property	Description
HDFS Connection	HDFS connection for the data lake.
HDFS Working Directory	HDFS directory where the Enterprise Data Preparation Service copies temporary data and files necessary for the service to run.
Hadoop Connection	Hadoop connection for the data lake.

6. Select the Hadoop authentication mode.
   - To select the simple authentication mode, press **1**.
   - To select Kerberos authentication, press **2**.
7. If you select Kerberos, enter the authentication parameters.

The following table describes the authentication properties that you must set if you select Kerberos:

Property	Description
Kerberos Principal	If the Hadoop cluster uses Kerberos authentication, specify the Service Principal Name (SPN) of the user account to impersonate when connecting to the data lake Hadoop cluster.
Kerberos KeyTab File	If the Hadoop cluster uses Kerberos authentication, specify the path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the Enterprise Data Preparation Service runs.
Local System Directory	Local directory that contains the files downloaded from the Enterprise Data Preparation application, such as .csv and .tde files.

8. Choose whether to enable the Enterprise Data Preparation Service immediately after you create the service.
   - If you want to enable the service at a later time using the Administrator tool, press **1**.
   - If you want the installer to enable the service after the installation process completes, press **2**.
9. Choose whether to enable logging of user activity events.
   - To disable logging of user activity events, press **1**.
   - To enable logging of user activity events, press **2**.

CHAPTER 11

Install Enterprise Data Preparation Binaries

This chapter includes the following topics:

Installation Overview

You can run the installer to install the Enterprise Data Preparation binaries on a node on which Enterprise Data Catalog is installed.

You can create the Interactive Data Preparation Service and the Interactive Data Preparation Service during the installation process. If you want the installer to create the services, it creates both services on a node. The installer prompts for connection objects associated with the Hadoop environment.

Before you run the installer, verify that the domain is integrated with the Hadoop environment and that the Hadoop, HDFS, and Hive connections are associated with the cluster configuration. If the cluster configuration does not exist, you can use the Administrator tool to create the connections manually after you integrate the domain with the Hadoop environment. For more information about integrating the domain with the Hadoop environment, see the *Integration Guide*.

Informatica recommends that you associate dedicated Model Repository Service and Data Integration Service instances with the Interactive Data Preparation Service. You can create a Model Repository Service and Data Integration Service to associate with the Enterprise Data Preparation Service during installation, or you can associate existing services with the Interactive Data Preparation Service.

If you create a Model Repository Service, you must specify the details for the Model repository database used by the Model Repository Service.

If you create the Enterprise Data Preparation Service and the Interactive Data Preparation Service during the installation process, you must create both services on the same node.

You can also create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service during installation.

Complete the Prerequisites

The installation of Enterprise Data Preparation consists of multiple phases.

When you plan to install Enterprise Data Preparation, you must account for the dependencies for the product. You also must plan the relational databases that the product requires.

For information about application service and database dependencies:

Interactive Data Preparation Service

Review the Interactive Data Preparation Service requirements. For more information, see "Interactive Data Preparation Service" on page 53.

You must prepare the Data Preparation repository database before you create the Interactive Data Preparation Service. For more information, see "Data Preparation Repository Database Requirements" on page 54.

If you plan to use rules during data preparation, you can create a Model Repository Service to associate with the Interactive Data Preparation Service when you configure the Enterprise Data Preparation services. Prepare the Model repository database that contains rule metadata before you create the Interactive Data Preparation Service.

Enterprise Data Preparation Service

Review the Enterprise Data Preparation Service requirements. For more information, see " Enterprise Data Preparation Service" on page 56.

You can create a Model Repository Service to associate with the Enterprise Data Preparation Service when you configure the Enterprise Data Preparation services. Prepare the Model repository database before you create the Enterprise Data Preparation Service.

Install the Enterprise Data Preparation Binaries

When you install the Enterprise Data Preparation binaries on a domain node on which Enterprise Data Catalog is already installed, you indicate that domain services and Enterprise Data Catalog are already installed on the node.

After you install the binaries, run the installer again to configure the services.

If the installer stops or is interrupted, the installer prompts you to continue installing Enterprise Data Preparation at the point at which it stopped.

1. On a shell command line, run the `install.sh` file from the root directory.
2. Press **1** to install the Informatica Big Data products.
3. Press **3** to run the installer.
4. Press **2** to agree to the terms and conditions.

5. Press **2** to continue with the installation.
6. Press **3** to install Enterprise Data Preparation.
7. Press **2** to indicate that the Informatica services are installed on the node.
8. Press **2** to indicate that Enterprise Data Catalog is installed on the node.
9. Press **2** to tune the application services for better performance based on your deployment type.
10. Enter the directory where you want to install Enterprise Data Preparation.

    The first time you install Enterprise Data Preparation, enter the Enterprise Data Catalog installation directory.
11. Choose how to proceed if Enterprise Data Preparation is already installed in the specified directory.
    - Press **1** to change the installation directory.
    - Press **2** to overwrite the existing installation.
12. Review the pre-installation summary, then click **Enter**.
13. Ensure the current node is shut down, then click **Enter**.

The Domain Details section appears.

Configure the Domain Details

Configure the domain details.

1. Press **2** if the current node is the gateway node for the domain.
2. Enter the domain name.
3. Enter the name of the current node.
4. Enter the domain administrator user name and password.
5. Press **1** to continue with the installation.

The Application Services for Enterprise Data Preparation section appears.

Configure the Model Repository Service and Model Repository Database Details

Choose whether to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation. If you create a Model Repository Service, specify the connection details for the Model repository database.

If you choose to create a Model Repository Service, specify the connection details for the Model repository database.

1. Choose to either create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service, or to associate existing application services with Enterprise Data Preparation.

- To create to create a Model Repository Service and a Data Integration Service to associate with the Enterprise Data Preparation Service on the node, press **1**.
- To associate an existing Model Repository Service and an existing Data Integration Service with the Enterprise Data Preparation Service, press **2**, and then enter the name of each service.

2. Enter the name of the Model Repository Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [

3. Specify the connection details for the Model repository database.

   The following table describes the parameters you set:

Property	Description
Database Type	Database for the Model repository managed by the Model Repository Service.
Database User ID	User name of the database user account to use to log in to the Model repository database.
User Password	Password for the Model repository database user account.
Tablespace	Configure for a IBM DB2 database. Name of the tablespace in which to create the tables. The tablespace must be defined on a single node, and the page size must be 32K. This option is required for a multi-partition database. If this option is not selected for a single-partition database, the installer creates the tables in the default tablespace.
Schema Name	Configure for a Microsoft SQL Server database. Name of the schema that will contain domain configuration tables. If not selected, the installer creates the tables in the default schema.

4. Specify the truststore details required to access a secure Model repository database.

   The following table describes the properties you set:

Property	Description
Database truststore file	Path and file name of the truststore file for the secure database.
Database truststore password	Password for the truststore.

5. Choose whether to configure the database connection using a JDBC URL or a custom JDBC connection string.
   - Press **1** to configure the database connection using a JDBC URL.

The following table describes the properties you set:

Property	Description
Database address	Host name and port number for the database in the format <host name>:<port>.
Database service name	Service name for Oracle and IBM DB2 databases, or database name for Microsoft SQL Server.
JDBC parameters	Optional parameters to include in the database connection string. Use the parameters to optimize database operations for the Model repository. You can use the default parameters or add or modify the parameters based on your database requirements. Verify that the parameter string is valid. The installer does not validate the parameter string before it adds the string to the JDBC URL. If not selected, the installer creates the JDBC URL without additional parameters. The following examples show the default connection strings for each database: - Oracle. jdbc:Informatica:oracle://host_name:port_no ;ServiceName= - IBM DB2. jdbc:Informatica:db2://host_name:port_no ;DatabaseName= - Microsoft SQL Server. jdbc:Informatica:sqlserver://host_name:port_no ;SelectMethod=cursor;DatabaseName= - Microsoft Azure. jdbc:informatica:sqlserver://host_name:port_number ;DatabaseName=<database_name>;SnapshotSerializable=true;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.<hostnameincertificate>;ValidateServerCertificate=true - PostgreSQL. jdbc:informatica:postgresql://<host name>:<port number>;DatabaseName=

- Press **2** to configure the database connection using a custom JDBC connection string.
  The following table describes the properties you set:

Property	Description
EncryptionMethod	Indicates whether data is encrypted when transmitted over the network. This parameter must be set to SSL.
ValidateServerCertificate	Indicates whether Informatica validates the certificate that is sent by the database server. If this parameter is set to True, Informatica validates the certificate that is sent by the database server. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to false, Informatica does not validate the certificate that is sent by the database server. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.

The Data Integration Service section appears.

Data Integration Service Details

If you create a Model Repository Service and a Data Integration Service to associate with Enterprise Data Preparation during installation, specify the properties required to create the Data Integration Service.

1. Specify the name of the Data Integration Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. Enter the name of the node where the service runs.
3. Enter the name of the Informatica license to associate with the service.
4. Enter the name of the Model Repository Service to associate with the Data Integration Service.
5. Specify the HTTP protocol type for the Data Integration Service, and then enter the port for each protocol you select.
   - To select HTTP only, press **1**.
   - To select HTTPS only, press **2**.
   - To select both HTTP and HTTPS, press **3**.
6. If you select HTTPS or both HTTP and HTTPS, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in custom keystore and truststore files, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.

The Data Preparation Repository Database section appears.

Configure the Data Preparation Repository Database Details

Specify the Data Preparation repository database connection details. You can choose to use an Oracle database or a MySQL database for the Data Preparation repository database.

If you do not have the database details, you can enter placeholder values, and then create the Interactive Data Preparation Service. If you continue without specifying the database connection details, you cannot enable the Interactive Data Preparation Service.

Oracle

1. To use an Oracle database for the Data Preparation repository, press **1**.
2. Enter the connection properties for the database.

The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the database.
Database Port Number	Port number for the database.
JDBC Parameters	Parameters required to connect to the database.
Custom JDBC Connection String	JDBC connection string to connect to the database. Format the string as follows: `jdbc:informatica:oracle://<database host name>:<port>;ServiceName=<service name>`

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

   The following table describes the secure connection properties:

Property	Description
Truststore File	Path and file name for the database truststore file.
Truststore Password	Password for the database truststore file.
Secure JDBC Parameters	List of secure database parameters to connect to the database. Format the parameters as follows: `EncryptionMethod=SSL;HostNameInCertificate=<secure database host name>;ValidateServerCertificate=true`

4. Press **2** to continue.

The Interactive Data Preparation Service Details section appears.

MySQL

1. To use a MySQL database or a MariaDB database for the Data Preparation repository, press **2**.
2. Enter the connection properties for the database.
   The following table describes the connection properties:

Property	Description
Database Host Name	Host name of the machine that hosts the Data Preparation repository database.
Database User Name	Database user account to use to connect to the Data Preparation repository.
Database User Password	Password for the Data Preparation repository database user account.
Database Port Number	Port number for the database.
Database Name	Schema or database name of the Data Preparation repository database.

3. To connect to a secure database, press **2**, and then enter the secure connection properties.

The following table describes the secure connection properties:

Property	Description

Custom JDBC Connection String	Connection string to connect to the database.
	To connect to a non-secure database, format the string as follows:
	`jdbc:mysql://<database host name>:<port>`
	The connection string is optional if you connect to a non-secure database.
	To connect to an SSL-enabled database, format the string as follows:
	`verifyServerCertificate=true&useSSL=true&requireSSL=true`

Secure JDBC Parameters	String containing the path and file name for the database truststore file, and the truststore password. Format the string as follows: `trustCertificateKeyStoreUrl=file://<truststore path/truststore file name>&trustCertificatekeyStorePassword=<truststore password>`

4. Press **Enter** to continue.

The Interactive Data Preparation Service Details section appears.

Interactive Data Preparation Service Details

Create the Interactive Data Preparation Service. If you run the installer to create the Enterprise Data Preparation Service and the Interactive Data Preparation Service, you must create both services on the same node.

1. Specify the name of the Interactive Data Preparation Service.

   The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
2. If you plan to use rules, you must associate a Data Integration Service and a Model Repository Service with the Interactive Data Preparation Service.
   - To skip associating a Model Repository Service and a Data Integration Service with the Interactive Data Preparation Service, press **1**.
   - To associate a Model Repository Service and a Data Integration Service with the Interactive Data Preparation Service, press **2**, and then enter the service names.
3. Enter the name of the Informatica license to associate with the service.
4. Choose whether to enable secure communication for the service.
   - To enable secure communication for the service, press **1**.
   - To disable secure communication, press **2**.
5. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.

6. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
7. Select the Hadoop authentication mode.
   - To select the simple authentication mode, press **1**.
   - To select Kerberos authentication, press **2**.
8. If you select Kerberos, enter the authentication parameters.

   The following table describes the authentication parameters that you set if you select Kerberos:

Property	Description
HDFS Principal Name	Service Principal Name (SPN) for the data preparation Hadoop cluster.
Hadoop Impersonation User Name	User name to use in Hadoop impersonation as shown in the Impersonation User Name property for the Hadoop connection in the Administrator tool. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Kerberos Keytab File	Path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the service runs.
Fully Qualified Path to the Kerberos Configuration File	Path to the krb5.conf Kerberos configuration file.

9. Specify the HDFS storage location, HDFS connection, and local storage location details.

   The following table describes the properties you set:

Property	Description
HDFS Storage Location	HDFS location for data preparation file storage. If the Hadoop cluster uses Kerberos authentication, the Hadoop impersonation user must have read, write, and execute permissions on the HDFS storage location folder.
Local Storage Location	Directory for data preparation file storage on the node on which the Interactive Data Preparation Service runs. If the connection to the local storage fails, the service recovers data preparation files from the HDFS storage location.

10. Specify the logging options.

The following table describes the properties you must set:

Property	Description
Log Severity	Severity of messages to include in the logs. Select from one of the following values: - FATAL. Writes FATAL messages to the log. FATAL messages include nonrecoverable system failures that cause the service to shut down or become unavailable. - ERROR. Writes FATAL and ERROR code messages to the log. ERROR messages include connection failures, failures to save or retrieve metadata, service errors. - WARNING. Writes FATAL, WARNING, and ERROR messages to the log. WARNING errors include recoverable system failures or warnings. - INFO. Writes FATAL, INFO, WARNING, and ERROR messages to the log. INFO messages include system and service change messages. - TRACE. Write FATAL, TRACE, INFO, WARNING, and ERROR code messages to the log. TRACE messages log user request failures. - DEBUG. Write FATAL, DEBUG, TRACE, INFO, WARNING, and ERROR messages to the log. DEBUG messages are user request logs.
Log Directory	Location of the directory of log files.

11. Specify the advanced options.

    The following table describes the properties you must set:

Property	Description
Model Repository Service Name	Name of the Model Repository Service. You cannot change the name of the service after you create it.
Data Integration Service Name	Name of the Data Integration Service.
HDFS Connection	HDFS connection for data preparation file storage.

12. Choose whether to enable the Interactive Data Preparation Service.
    - If you want to enable the service at a later time using the Administrator tool, press **1**.
    - If you want the installer to enable the service after you complete the installation, press **2**.

The Enterprise Data Preparation Service Details section appears.

Enterprise Data Preparation Service Details

Create the Enterprise Data Preparation Service. If you run the installer to create the Enterprise Data Preparation Service and the Interactive Data Preparation Service, you must create both services on the same node.

1. Specify the details for the Enterprise Data Preparation Service.

The following table describes the properties that you set:

Property	Description
Enterprise Data Preparation Service Name	Name of the Enterprise Data Preparation Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Interactive Data Preparation Service Name	Name of the Interactive Data Preparation Service to associate with the Enterprise Data Preparation Service.
Model Repository Service Name	Name of the Model Repository Service to associate with the Enterprise Data Preparation Service.
Data Integration Service Name	Name of the Data Integration Service associated with the Enterprise Data Preparation Service

2. Choose whether to enable secure communication for the service.
   - To enable secure communication for the service, press **1**.
   - To disable secure communication, press **2**.
3. If you enable secure communication for the service, select the SSL certificate to use.
   - To use the default Informatica SSL certificate contained in the default keystore and the default truststore, press **1**.
   - To use a custom SSL certificate contained in a custom keystore and truststore, press **2**, and then enter the path and file name for the keystore and truststore files. You must also enter the keystore and truststore passwords.
4. If you enable secure communication for the service, enter the port number for the HTTPS connection. If you enable non-secure communication for the service, enter the port number for the HTTP connection.
5. Specify the data lake connection options.

   The following table describes the properties that you set for the data lake connections:

Property	Description
HDFS Connection	HDFS connection for the data lake.
HDFS Working Directory	HDFS directory where the Enterprise Data Preparation Service copies temporary data and files necessary for the service to run.
Hadoop Connection	Hadoop connection for the data lake.

6. Select the Hadoop authentication mode.
   - To select the simple authentication mode, press **1**.
   - To select Kerberos authentication, press **2**.
7. If you select Kerberos, enter the authentication parameters.

The following table describes the authentication properties that you must set if you select Kerberos:

Property	Description
Kerberos Principal	If the Hadoop cluster uses Kerberos authentication, specify the Service Principal Name (SPN) of the user account to impersonate when connecting to the data lake Hadoop cluster.
Kerberos KeyTab File	If the Hadoop cluster uses Kerberos authentication, specify the path and file name of the SPN keytab file for the user account to impersonate when connecting to the Hadoop cluster. The keytab file must be in a directory on the machine where the Enterprise Data Preparation Service runs.
Local System Directory	Local directory that contains the files downloaded from the Enterprise Data Preparation application, such as .csv and .tde files.

8. Choose whether to enable the Enterprise Data Preparation Service immediately after you create the service.
   - If you want to enable the service at a later time using the Administrator tool, press **1**.
   - If you want the installer to enable the service after the installation process completes, press **2**.
9. Choose whether to enable logging of user activity events.
   - To disable logging of user activity events, press **1**.
   - To enable logging of user activity events, press **2**.

CHAPTER 12

Run the Silent Installer

This chapter includes the following topics:

Installing in Silent Mode

To install without user interaction, install in 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 the services on multiple machines on the network or to standardize the installation across machines.

Copy the installation files to the hard disk on the machine where you plan to install the services. If you install on a remote machine, verify that you can access and create files on the remote machine.

To install in silent mode, complete the following tasks:

1. Run the password encryption utility to encrypt the passwords in the installation properties file.
2. Configure the installation properties file and specify the installation options in the properties file.
3. Run the installer with the installation properties file.

Configure the Properties File

Configure the properties file that contains the configuration properties required to install the Informatica services in silent mode.

Informatica provides two versions of the properties file. Use either file to specify the options for your installation.

Silent input properties file

The silent input properties file contains the configuration properties required to install the Informatica services in silent mode. Use the file if you want to consider the appropriate value to set for each property in the file.

Default silent input properties file

The default silent input properties file contains default values for many configuration properties. The properties are listed in the bottom portion of the file. Use the file if you plan to install the Informatica services using the default property values.

The file contains properties set to the default value for the following options:

- Application service names.
- Secure Sockets Layer authentication.
- Kerberos authentication.
- Port number assignment for domain and node components.

To configure the properties file that contains the configuration properties required to install the Informatica services in silent mode, complete the following steps:

1. Go to the root of the directory that contains the installation files.
2. Optionally, run the password encryption utility to encrypt passwords in the .properties file.
3. Create a backup copy of the `SilentInput.properties` file.
4. Open either the `SilentInput.properties` file or the `SilentInput_Default.properties` file.
5. Configure the properties in the file.
6. Save the file with the name `SilentInput.properties`.

Run the Installer

After you configure the properties file, open a command prompt to start the silent installation.

1. Open a command prompt.
2. Go to the root of the directory that contains the installation files.
3. Verify that the directory contains the file SilentInput.properties that you edited and resaved.
4. Run the silent installation. On Linux, run silentInstall.sh.

The silent installer runs in the background. The process can take a while. The silent installation is complete when the Informatica_<Version>_Services_InstallLog<timestamp>.log file is created in the installation directory.

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.

Encrypting Passwords in the Properties File

The installer includes a utility that you can use to encrypt passwords you set in the properties file you use to specify options when you run the installer in silent mode. Informatica uses AES encryption with multiple 128-bit keys to encrypt passwords.

You run the utility for each password you want to encrypt. When you run the utility, you specify the value of the password in plain text at the command prompt. The utility generates the password in encrypted format as output. The output includes the following prefix: `=INSTALLER:CIPHER:AES:128=`

Copy the complete output string, including the prefix, and then paste it into the properties file as the value for the password property. When you run the installer in silent mode, the installation framework decrypts the password.

1. Go to the utility directory:

```
<Installer directory>/properties/utils/passwd_encryption
```

2. Run the utility. Specify the plain text password you want to encrypt as the value for <password>.
   - On Linux and UNIX, run the following command:

     ```
     sh install.sh <password>
     ```
   - On Windows, run the following command:

     ```
     install.bat <password>
     ```
3. Copy the encrypted password string from the output, and then paste the string into the .properties file as the value for the corresponding password.

   The following example shows the encrypted password set as the value for the DOMAIN_PSSWD property:

   ```
   DOMAIN_PSSWD==INSTALLER:CIPHER:AES:128=mjkjmDR2kzFJiizfRWIOPg==
   ```

CHAPTER 13

Troubleshooting

This chapter includes the following topics:

Installation Troubleshooting Overview

The topics in this section provides you information on troubleshooting probable issues that you might encounter during Informatica installation process. The examples included in the topics describe general troubleshooting strategies and are not a comprehensive list of possible causes of installation issues.

Resuming a Failed Installer Process

When the installation process stops midway, you can resume the installation from the point of failure or exit.

When the service installation process fails on UNIX or Linux, you can resume from the previous service configuration and recover the last entered details for that service installation. The install process might fail for reasons such as network outage, when you exit the installation before completing the entire installation process, or because of incorrect information entered.

Consider the following guidelines for resuming the installation:

You can resume the installer

If a service fails or if the installation process fails during a service creation, you can resume the installation process with the server installer. To resume the installation process, ensure that at least one of the services is created and that the domain is up and running from the installation log. For example, if you want to check whether the Model Repository Service is created, check if you have a service creation success text in the server log in the following format:

```
SUCCESS: MRS Service [mrs_name] is created. Command ran successfully.
```

To resume the installation, run the installer again.

When you resume the installer while creating a service, the installer retains all the service and database specific information, such as the create service status, service name, service enabled or disabled status. You can confirm and use the previously entered values or specify new values for the service and resume the installation process.

You cannot resume the installer

You cannot resume the installer in the following situations:

- You run installer to configure services after the services are created.
- You run the service configuration wizard.
- You join a domain.

Before You Resume the Installer

When the installation process stops midway, you can resume the installation from the point of failure or exit.

Before you can resume the installer, complete the following prerequisites:

1. In the installation log file present in the installation directory, verify that at least the domain and one service is created. The installer log file name appears in the following syntax: Informatica_<Version>_Services_<timestamp>.log
2. Ensure that you do not delete the installInst.obj object file present in the tools folder of the user installation directory.
3. If you are going to resume through the silent installer, ensure that RESUME_INSTALLATION is set to true in the SilentInput.properties file.

Resume the Installer

After you complete prerequisite tasks, you can resume the installer.

1. Open a command prompt and navigate to the location of the installation files.
2. Run the console installer or the silent installer.
3. When the regular installer runs, you might get a prompt confirming whether you want to resume previous installer or not.
   - If you do not want to resume installation, enter 1 for No. Default is 1.
   - If you want to resume installation, enter 2 for Yes.

   Before you can resume the installation, the services get validated.

Troubleshooting with Installation Log Files

You can use the following log files to troubleshoot an Informatica installation:

Installation log files

The installer produces log files during and after the installation. You can use these logs to get more information about the tasks completed by the installer and errors that occurred during installation. The installation log files include the following logs:

- Debug logs

- File installation logs

Service Manager log files

Log files generated when the Service Manager starts on a node.

Debug Log Files

The installer writes actions and errors to the debug log file. The name of the log file depends on the Informatica component you install.

The debug log contains output from the infacmd and infasetup commands used to create the domain, node, and application services. It also contains information about starting the application services.

The following table describes the properties of the debug log files:

Property	Description
Log File Name	- Informatica_<Version>_Services_<timestamp>.log - Informatica_<Version>_Client_<timestamp>.log - Informatica_<Version>_Services_Upgrade_<timestamp>.log - Informatica_<Version>_Client_Upgrade_<timestamp>.log
Location	Installation directory.
Usage	Get more information about the actions performed by the installer and get more information about installation errors. The installer writes information to this file during the installation. If the installer generates an error, you can use this log to troubleshoot the error.
Contents	Detailed summary of each action performed by the installer, the information you entered in the installer, each command line command used by the installer, and the error code returned by the command.

File Installation Log File

The file installation log file contains information about the installed files.

The following table describes the properties of the installation log file:

Property	Description
Log File Name	- Informatica_<Version>_Services_InstallLog.log - Informatica_<Version>_Client_InstallLog.log
Location	Installation directory.
Usage	Get information about the files installed and registry entries created.
Contents	Directories created, names of the files installed and commands run, and status for each installed file.

Service Manager Log Files

The installer starts the Informatica service. The Informatica service starts the Service Manager for the node. The Service Manager generates log files that indicate the startup status of a node. Use these files to

troubleshoot issues when the Informatica service fails to start and you cannot log in to Informatica Administrator. The Service Manager log files are created on each node.

The following table describes the files generated by the Service Manager:

Property	Description
catalina.out	Log events from the Java Virtual Machine (JVM) that runs the Service Manager. For example, a port is available during installation, but is in use when the Service Manager starts. Use this log to get more information about which port was unavailable during startup of the Service Manager. The catalina.out file is in the following directory: `<Informatica installation directory>/logs/<node name>/catalina.out`
node.log	Log events generated during the startup of the Service Manager on a node. You can use this log to get more information about why the Service Manager for a node failed to start. For example, if the Service Manager cannot connect to the domain configuration database after 30 seconds, the Service Manager fails to start. The node.log file is in the /tomcat/logs directory.

Note: The Service Manager also uses node.log to record events when the Log Manager is unavailable. For example, if the machine where the Service Manager runs does not have enough available disk space to write log event files, the Log Manager is unavailable.

Troubleshooting Domains and Nodes

The installer can generate errors when creating and configuring domains and nodes during the Informatica installation.

Creating the Domain Configuration Repository

If you create a domain, the installer creates a domain configuration repository to store domain metadata. The installer uses the options you enter during installation to add configuration metadata to the domain configuration repository. The installer uses JDBC to communicate with the database. You do not need to configure ODBC or native connectivity on the machine where you install the Informatica services.

The installer creates and drops a table in the domain configuration repository database to verify the connection information. The user account for the database must have create privileges on the database. Each domain must have a separate domain configuration repository.

Creating or Joining a Domain

The installer completes different tasks depending on whether you create a domain or join a domain:

- **Creating a domain.** The installer runs the infasetup DefineDomain command to create the domain and the gateway node for the domain on the current machine based on the information you enter in the Configure Domain window.
- **Joining a domain.** The installer runs the infasetup DefineWorkerNode command to create a node on the current machine, and runs the infacmd AddDomainNode command to add the node to the domain. The installer uses the information you enter in the Configure Domain window to run the commands.

The infasetup and infacmd commands fail if the gateway node is unavailable. If the gateway node is unavailable, you cannot log in to Informatica Administrator.

For example, the DefineDomain command fails if you click Test Connection and the connection test passes but the database becomes unavailable before you click Next. The DefineDomain command can also fail if the host name or IP address does not belong to the current machine. Verify that the database for the domain configuration is available and that the host name is correct and try again.

If the AddDomainNode command fails, verify that the Informatica service is running on the gateway node and try again.

Starting Informatica

The installer runs infaservice to start the Informatica service. To troubleshoot issues when Informatica fails to start, use the information in the installation debug log and the node.log and catalina.out Service Manager log files to identify the cause of the error.

If you create a domain, log in to Informatica Administrator after the Informatica service starts to verify that the domain is available. If you join a domain, log in to Informatica Administrator after the Informatica service starts to verify that the node was successfully created and started.

Informatica can fail to start for the following reasons:

- **The Service Manager is out of system memory.** The Java Runtime Environment (JRE) that starts Informatica and runs the Service Manager may not have enough system memory to start. Set the INFA_JAVA_OPTS environment variable to configure the amount of system memory used by Informatica. On UNIX, you can set the memory configuration when you start Informatica.
- **The domain configuration database is not available.** Informatica fails to start on a node if the Service Manager on a gateway node cannot connect to the domain configuration database within 30 seconds. Verify that the domain configuration repository is available.
- **Some of the folders in the Informatica installation directory do not have the appropriate execute permissions.** Grant execute permission on the Informatica installation directory.

Pinging the Domain

The installer runs the *infacmd* Ping command to verify that the domain is available before it continues the installation. The domain must be available so that license objects can be added to the domain. If the Ping command fails, start Informatica on the gateway node.

Adding a License

The installer runs the *infacmd* AddLicense command to read the Informatica license key file and create a license object in the domain. To run the application services in Informatica Administrator, a valid license object must exist in the domain.

If you use an incremental license and join a domain, the serial number of the incremental license must match the serial number for an existing license object in the domain. If the serial numbers do not match, the AddLicense command fails.

You can get more information about the contents of the license key file used for installation, including serial number, version, expiration date, operating systems, and connectivity options in the installation debug log. You can get more information about existing licenses for the domain in Informatica Administrator.

Troubleshooting Informatica Developer

Consider the following tips when you work with the Informatica Developer:

Informatica Developer fails to launch

This issue might occur if the jvm.dll of java requires the MSVCR100.dll.

To resolve this issue, download Microsoft Visual C++ Studio 2010 Redistributable Package from the Microsoft website.

Part IV: After You Install the Services

This part contains the following chapters:

CHAPTER 14

Complete the Domain Configuration

This chapter includes the following topics:

Checklist to Complete the Domain Configuration

This chapter contains information about domain configuration tasks that you need to complete after installation. Use this checklist to track domain configuration tasks.

❑ Integrate the domain with the Hadoop environment.

❑ Verify locale settings and code page compatibility:

- Verify that the domain configuration database is compatible with the code pages of the application services that you create in the domain.
- Verify that the locale settings on machines that access the Administrator tool and the Informatica client tools is compatible with the code pages of repositories in the domain.
- Configure the locale environment variables.

❑ Configure the following environment variables:

- Informatica environment variables to store memory, domain, and location settings.
- Library path environment variables on the machines that run the Data Integration Service.
- Kerberos environment variables if you configure the Informatica domain to run on a network with Kerberos authentication.

Complete the Domain Configuration Overview

After you install Informatica services and before you create the application services, complete the configuration for the domain services.

Domain configuration includes tasks such as verifying code pages, configuring the environment variables for the domain, and configuring the firewall.

Integrate the Domain with the Hadoop Environment

If you imported the cluster configuration from the Hadoop environment during installation, you must complete the integration between the domain and the Hadoop environment. Integration tasks are required in both the Hadoop environment and the Informatica domain environment.

For information on how to import a Hadoop cluster configuration, refer to the "Cluster Configuration" on page 80 topic and the Hadoop Integration section of the *Data Engineering Integration* Guide.

To integrate the domain with the Hadoop environment, you complete the following high-level tasks:

1. Prepare directories, users, and permissions.
2. Configure *-site.xml files on the Hadoop environment. The properties *-site.xml files must be updated with values required for Informatica processing in the Hadoop environment.
3. Refresh the cluster configuration in the Administrator tool. Refresh the cluster configuration to get the updated properties from the *-site.xml files on the cluster.
4. Update connections in the Administrator tool. Update connections if you want to use property values other than the default values. You will also need to configure environment variables in the Hadoop connection.

Verify Locale Settings and Code Page Compatibility

The code pages for application services must be compatible with code pages in the domain.

Verify and configure the locale settings and code pages:

Verify that the domain configuration database is compatible with the code pages of the application services that you create in the domain.

The Service Manager synchronizes the list of users in the domain with the list of users and group in each application service. If a user name in the domain has characters that the code page of the application service does not recognize, characters do not convert correctly and inconsistencies occur.

Verify that the locale settings on machines that access the Administrator tool and the Informatica client tools are compatible with code pages of repositories in the domain.

If the locale setting is not compatible with the repository code page, you cannot create an application service.

Configure Locale Environment Variables

Verify that the locale setting is compatible with the code page for the repository. If the locale setting is not compatible with the repository code page, you cannot create an application service.

Use LANG, LC_CTYPE, or LC_ALL to set the UNIX or Linux code page.

Different operating systems require different values for the same locale. The value for the locale variable is case sensitive.

Use the following command to verify that the value for the locale environment variable is compatible with the language settings for the machine and the type of code page you want to use for the repository:

```
locale -a
```

The command returns the languages installed on the operating system and the existing locale settings.

Set the following locale environment variables:

Locale on Linux

All UNIX operating systems except Linux have a unique value for each locale. Linux allows different values to represent the same locale. For example, “utf8,” “UTF-8,” “UTF8,” and “utf-8” represent the same locale on a Linux machine. Informatica requires that you use a specific value for each locale on a Linux machine. Make sure that you set the LANG environment variable appropriately for all Linux machines.

Locale for Oracle database clients

For Oracle database clients, set NLS_LANG to the locale that you want the database client and server to use with the login. A locale setting consists of the language, territory, and character set. The value of NLS_LANG depends on the configuration.

For example, if the value is american_america.UTF8, set the variable in a C shell with the following command:

```
setenv NLS_LANG american_america.UTF8
```

To read multibyte characters from the database, set the variable with the following command:

```
setenv NLS_LANG=american_america.AL32UTF8
```

You must set the correct variable on the Data Integration Service machine so that the Data Integration Service can read the Oracle data correctly.

Configure Environment Variables on UNIX or Linux

Informatica uses environment variables to store configuration information when it runs the application services and connects to the clients. Configure the environment variables to meet the Informatica requirements.

Incorrectly configured environment variables can cause the Informatica domain or nodes to fail to start or can cause connection problems between the Informatica clients and the domain.

To configure environment variables, log in with the system user account you used to install Informatica.

Configure Informatica Environment Variables

You can configure Informatica environment variables to store memory, domain, and location settings.

Set the following environment variables:

INFA_JAVA_OPTS

By default, Informatica uses a maximum of 512 MB of system memory.

The following table lists the minimum requirement for the maximum heap size settings, based on the number of users and services in the domain:

Number of Domain Users	Maximum Heap Size (1-5 Services)	Maximum Heap Size (6-10 Services)
1,000 or less	512 MB (default)	1024 MB
5,000	2048 MB	3072 MB
10,000	3072 MB	5120 MB
20,000	5120 MB	6144 MB
30,000	5120 MB	6144 MB

Note: The maximum heap size settings in the table are based on the number of application services in the domain.

If the domain has more than 1,000 users, update the maximum heap size based on the number of users in the domain.

You can use the INFA_JAVA_OPTS environment variable to configure the amount of system memory used by Informatica. For example, to configure 1 GB of system memory for the Informatica daemon in a C shell, use the following command:

```
setenv INFA_JAVA_OPTS "-Xmx1024m"
```

Restart the node for the changes to take effect.

INFA_DOMAINS_FILE

The installer creates a domains.infa file in the Informatica installation directory. The domains.infa file contains the connectivity information for the gateway nodes in a domain, including the domain names, domain host names, and domain host port numbers.

Set the value of the INFA_DOMAINS_FILE variable to the path and file name of the domains.infa file.

Configure the INFA_DOMAINS_FILE variable on the machine where you install the Informatica services.

INFA_HOME

Use INFA_HOME to designate the Informatica installation directory. If you modify the Informatica directory structure, you need to set the environment variable to the location of the Informatica installation directory or the directory where the installed Informatica files are located.

For example, you use a softlink for any of the Informatica directories. To configure INFA_HOME so that any Informatica application or service can locate the other Informatica components it needs to run, set INFA_HOME to the location of the Informatica installation directory.

INFA_TRUSTSTORE

If you enable secure communication for the domain, set the INFA_TRUSTSTORE variable with the directory that contains the truststore files for the SSL certificates. The directory must contain truststore files named infa_truststore.jks and infa_truststore.pem.

You must set the INFA_TRUSTSTORE variable if you use the default SSL certificate provided by Informatica or a certificate that you provide.

INFA_TRUSTSTORE_PASSWORD

If you enable secure communication for the domain and you specify the SSL certificate to use, set the INFA_TRUSTSTORE_PASSWORD variable with the password for the infa_truststore.jks that contains the SSL certificate. The password must be encrypted. Use the command line program pmpasswd to encrypt the password.

Configure Library Path Environment Variables

Configure library path environment variables on the machines that run the Data Integration Service processes. The variable name and requirements depend on the platform and database.

Configure the LD_LIBRARY_PATH environment variable.

The following table describes the values that you set for the LD_LIBRARY_PATH for the different databases:

Database	Value
Oracle	<Database path>/lib
IBM DB2	<Database path>/lib
Sybase ASE	"${SYBASE_OCS}/lib:${SYBASE_ASE}/lib:${LD_LIBRARY_PATH}"
Teradata	<Database path>/lib
ODBC	<CLOSEDODBCHOME>/lib
PostgreSQL	$PGHOME/lib:$ {LD_LIBRARY_PATH}

CHAPTER 15

Prepare to Create the Application Services

This chapter includes the following topics:

Checklist for Preparing to Create Application Services

This chapter contains tasks that you need to complete before you create or configure the Data Integration Service and the Content Management Service. When you configure the services you configure properties based on the connections and directories that you create. Use this checklist to track the configuration tasks.

❑ Create the following connections for the Data Integration Service:

- Data object cache database
- Workflow database
- Profiling warehouse

❑ Create the following connection for the Content Management Service:

- Reference data warehouse

Create a Keystore for a Secure Connection to a Web Application Service

You can secure the connection between the Informatica domain and a web application service, such as the Analyst service. Informatica uses the SSL/TLS protocol to encrypt network traffic. To secure the connection, you must create the required files.

Before you can secure the connection to a web application service, verify that the following requirements are met:

You created a certificate signing request (CSR) and private key.

You can use keytool or OpenSSL to create the CSR and private key.

If you use RSA encryption, you must use more than 512 bits.

You have a signed SSL certificate.

The certificate can be self-signed or CA signed. Informatica recommends a CA signed certificate.

You imported the certificate into a keystore in JKS format.

A keystore must contain only one certificate. If you use a unique certificate for each web application service, create a separate keystore for each certificate. Alternatively, you can use a shared certificate and keystore.

If you use the installer-generated SSL certificate for the Administrator tool, you do not need to import the certificate into a keystore in JKS format.

The keystore is in an accessible directory.

The keystore must be in a directory that is accessible to the Administrator tool.

Log In to Informatica Administrator

You must have a user account to log in to the Informatica Administrator web application.

In Microsoft Internet Explorer and Google Chrome, add the URL of the Informatica web application to the list of trusted sites. If you are using Chrome version 80.0.39x or later, you must also set the AuthServerWhitelist and AuthNegotiateDelegateWhitelist policies.

1. Start a Microsoft Internet Explorer or Google Chrome browser.
2. In the **Address** field, enter the URL for the Administrator tool:
   - If the Administrator tool is not configured to use a secure connection, enter the following URL:

     ```
     http://<fully qualified host name>:<http port>/administrator/
     ```
   - If the Administrator tool is configured to use a secure connection, enter the following URL:

     ```
     https://<fully qualified host name>:<http port>/administrator/
     ```

   Host name and port in the URL represent the host name and port number of the master gateway node. If you configured secure communication for the domain, you must use HTTPS in the URL to ensure that you can access the Administrator tool.

3. Enter the user name, password, and security domain for your user account, and then click **Login**.

   The **Security Domain** field appears when the Informatica domain contains an LDAP security domain. If you do not know the security domain that your user account belongs to, contact the Informatica domain administrator.

   Note: If this is the first time you log in with the user name and password provided by the domain administrator, change your password to maintain security.

Create Connections

In the Administrator tool, create connections to the databases that the application services use. You need to specify the connection details while you configure the application service.

When you create the database connection, specify the database connection properties and test the connection.

The following table describes the database connections that you need to create before the application services can access the associated databases.

Database Connection	Description
Data object cache database	To access the data object cache, create the data object cache connection for the Data Integration Service.
Workflow database	To store run-time metadata for workflows, create the workflow database connection for the Data Integration Service.
Profiling warehouse database	To create and run profiles and scorecards, create the profiling warehouse database connection for the Data Integration Service. **Note:** To use the Microsoft SQL Server database as the profiling warehouse, choose ODBC as the provider type, and clear the **use DSN** option in the **Microsoft SQL Server connection properties** dialog box when you configure the Microsoft SQL Server connection.
Reference data warehouse	To store reference table data, create the reference data warehouse connection for the Content Management Service.

IBM DB2 Connection Properties

Use a DB2 for LUW connection to access tables in a DB2 for LUW database.

The following table describes the DB2 for LUW connection properties:

Property	Description
User name	Database user name.
Password	Password for the user name.
Connection String for metadata access	Connection string to import physical data objects. Use the following connection string: `jdbc:informatica:db2://<host>:50000;databaseName=<dbname>`

Property	Description
Connection String for data access	Connection string to preview data and run mappings. Enter `dbname` from the alias configured in the DB2 client.
Code Page	Database code page.
Environment SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the connection environment SQL each time it connects to the database.
Transaction SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the transaction environment SQL at the beginning of each transaction.
Retry Period	This property is reserved for future use.
Tablespace	Tablespace name of the DB2 for LUW database.
SQL Identifier Character	The type of character used to identify special characters and reserved SQL keywords, such as WHERE. The Data Integration Service places the selected character around special characters and reserved SQL keywords. The Data Integration Service also uses this character for the Support Mixed-case Identifiers property.
Support Mixed-case Identifiers	When enabled, the Data Integration Service places identifier characters around table, view, schema, synonym, and column names when generating and executing SQL against these objects in the connection. Use if the objects have mixed-case or lowercase names. By default, this option is not selected.

Microsoft SQL Server Connection Properties

Use a Microsoft SQL Server connection to access tables in a Microsoft SQL Server database.

The following table describes the Microsoft SQL Server connection properties:

Property	Description
User name	Database user name.
Password	Password for the user name.
Use Trusted Connection	Optional. When enabled, the Data Integration Service uses Windows authentication to access the Microsoft SQL Server database. The user name that starts the Data Integration Service must be a valid Windows user with access to the Microsoft SQL Server database.
Connection String for metadata access	Connection string to import physical data objects. Use the following connection string: `jdbc:informatica:sqlserver://<host>:<port>;databaseName=<dbname>`
Connection String for data access	Connection string to preview data and run mappings. Enter `<ServerName>@<DBName>`
Domain Name	Optional. Name of the domain where Microsoft SQL Server is running.

Property	Description
Packet Size	Required. Optimize the ODBC connection to Microsoft SQL Server. Increase the packet size to increase performance. Default is 0.
Code Page	Database code page.
Owner Name	Name of the schema owner. Specify for connections to the profiling warehouse database or data object cache database.
Schema Name	Name of the schema in the database. Specify for connections to the profiling warehouse or data object cache database. You must specify the schema name for the profiling warehouse if the schema name is different from the database user name. You must specify the schema name for the data object cache database if the schema name is different from the database user name and you manage the cache with an external tool.
Environment SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the connection environment SQL each time it connects to the database.
Transaction SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the transaction environment SQL at the beginning of each transaction.
Retry Period	This property is reserved for future use.
SQL Identifier Character	The type of character used to identify special characters and reserved SQL keywords, such as WHERE. The Data Integration Service places the selected character around special characters and reserved SQL keywords. The Data Integration Service also uses this character for the Support Mixed-case Identifiers property.
Support Mixed-case Identifiers	When enabled, the Data Integration Service places identifier characters around table, view, schema, synonym, and column names when generating and executing SQL against these objects in the connection. Use if the objects have mixed-case or lowercase names. By default, this option is not selected.

Note: When you use a Microsoft SQL Server connection to access tables in a Microsoft SQL Server database, the Developer tool does not display the synonyms for the tables.

Oracle Connection Properties

Use an Oracle connection to access tables in an Oracle database.

The following table describes the Oracle connection properties:

Property	Description
User name	Database user name.
Password	Password for the user name.

Property	Description
Connection String for metadata access	Connection string to import physical data objects. Use the following connection string: `jdbc:informatica:oracle://<host>:1521;SID=<sid>` Use the following connection string to connect to Oracle through Oracle Connection Manager: `jdbc:Informatica:oracle:TNSNamesFile=<fully qualified path to the tnsnames.ora file>;TNSServerName=<TNS server name>;`
Connection String for data access	Connection string to preview data and run mappings. Enter `dbname.world` from the TNSNAMES entry.
Code Page	Database code page.
Environment SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the connection environment SQL each time it connects to the database.
Transaction SQL	Optional. Enter SQL commands to set the database environment when you connect to the database. The Data Integration Service executes the transaction environment SQL at the beginning of each transaction.
Retry Period	This property is reserved for future use.
Parallel Mode	Optional. Enables parallel processing when loading data into a table in bulk mode. Default is disabled.
SQL Identifier Character	The type of character used to identify special characters and reserved SQL keywords, such as WHERE. The Data Integration Service places the selected character around special characters and reserved SQL keywords. The Data Integration Service also uses this character for the Support Mixed-case Identifiers property.
Support Mixed-case Identifiers	When enabled, the Data Integration Service places identifier characters around table, view, schema, synonym, and column names when generating and executing SQL against these objects in the connection. Use if the objects have mixed-case or lowercase names. By default, this option is not selected.

Creating a Connection

In the Administrator tool, you can create relational database, social media, and file systems connections.

1. In the Administrator tool, click the **Manage** tab.
2. Click the **Connections** view.
3. In the Navigator, select the domain.
4. In the Navigator, click **Actions** > **New** > **Connection**.

   The **New Connection** dialog box appears.
5. In the **New Connection** dialog box, select the connection type, and then click **OK**.

   The **New Connection** wizard appears.
6. Enter the connection properties.

The connection properties that you enter depend on the connection type. Click **Next** to go to the next page of the **New Connection** wizard.

7. When you finish entering connection properties, you can click **Test Connection** to test the connection.
8. Click **Finish**.

CHAPTER 16

Create and Configure Application Services

This chapter includes the following topics:

Checklist to Create and Configure Application Services

This chapter contains instructions to create and configure application services. Even if you created services during installation, you might still need to configure some services. Use this checklist to track completion of application service configuration.

- ❑ Review your notes for planning the application services.
- ❑ Identify the services that you created during installation, and complete additional configuration for the service.
- ❑ Create and configure other services that you want in the domain.

Create and Configure the Application Services Overview

If you did not create services with you ran the installer, use the Administrator tool to create the application services.

Some application services depend on other application services. When you create these dependent application services, you must provide the name of other running application services. Review the application service dependencies to determine the order that you must create the services. For example, you must create a Model Repository Service before you create a Data Integration Service.

Before you create the application services, verify that you have completed the prerequisite tasks required by the installation and configuration process.

Create and Configure the Model Repository Service

The Model Repository Service is an application service that manages the Model repository. The Model repository stores metadata created by Informatica clients and application services in a relational database to enable collaboration among the clients and services.

When you access a Model repository object from an Informatica client tool or application service, the client or service sends a request to the Model Repository Service. The Model Repository Service process fetches, inserts, and updates the metadata in the Model repository database tables.

Create the Model Repository Service

Use the service creation wizard in the Administrator tool to create the service.

1. In the Administrator tool, click the **Manage** tab.
2. Click **Actions** > **New** > **Model Repository Service**.

   The **New Model Repository Service** dialog box appears.
3. On the **New Model Repository Service - Step 1 of 2** page, enter the following properties:

Property	Description
Name	Name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the service. The description cannot exceed 765 characters.
Location	Domain and folder where the service is created. Click **Browse** to choose a different folder. You can move the service after you create it.
License	License object that allows use of the service.

Property	Description
Node	Node on which the service runs.
Backup Nodes	If your license includes high availability, nodes on which the service can run if the primary node is unavailable.

4. Click **Next**.

   The **New Model Repository Service - Step 2 of 2** page appears.

5. Enter the following properties for the Model repository database:

Property	Description
Database Type	The type of the repository database.
Username	The database user name for the repository.
Password	Repository database password for the database user.
Database Schema	Available for Microsoft SQL Server and PostgreSQL. Name of the schema that will contain Model repository tables.
Database Tablespace	Available for IBM DB2. Name of the tablespace in which to create the tables. For a multi-partition IBM DB2 database, the tablespace must span a single node and a single partition.

6. Enter the JDBC connection string that the service uses to connect to the Model repository database.

Use the following syntax for the connection string for the selected database type:

Database Type	Connection String Syntax
IBM DB2	`jdbc:informatica:db2://<host name>:<port number>;DatabaseName=<database name>;BatchPerformanceWorkaround=true;DynamicSections=3000`
Microsoft SQL Server	- Microsoft SQL Server that uses the default instance `jdbc:informatica:sqlserver://<host name>:<port number>;DatabaseName=<database name>;SnapshotSerializable=true` - Microsoft SQL Server that uses a named instance `jdbc:informatica:sqlserver://<host name>\<named instance name>;DatabaseName=<database name>;SnapshotSerializable=true` - Microsoft Azure. `jdbc:informatica:sqlserver://<host_name>:<port_number>;DatabaseName=<database_name>;SnapshotSerializable=true;SnapshotSerializable=true;EncryptionMethod=SSL;HostNameInCertificate=*.<hostnameincertificate>;ValidateServerCertificate=true` - Azure SQL Database with Active Directory authentication. `jdbc:informatica: sqlserver://<host_name>:<port_number>;database=<database_name>;encrypt=true;AuthenticationMethod=ActiveDirectoryPassword;trustServerCertificate=false;hostNameInCertificate=*.database.windows.net;loginTimeout=<seconds>`
Oracle	`jdbc:informatica:oracle://<host name>:<port number>;SID=<database name>;MaxPooledStatements=20;CatalogOptions=0;BatchPerformanceWorkaround=true`
PostgreSQL	`jdbc:informatica:postgresql://<host name>:<port number>;DatabaseName=`

7. If the Model repository database is secured with the SSL protocol, you must enter the secure database parameters in the **Secure JDBC Parameters** field.

   Enter the parameters as `name=value` pairs separated by semicolon characters (;). For example:

   ```
   param1=value1;param2=value2
   ```

Enter the following secure database parameters:

Secure Database Parameter	Description
EncryptionMethod	Required. Indicates whether data is encrypted when transmitted over the network. This parameter must be set to `SSL`.
ValidateServerCertificate	Optional. Indicates whether Informatica validates the certificate that the database server sends. If this parameter is set to True, Informatica validates the certificate that the database server sends. If you specify the HostNameInCertificate parameter, Informatica also validates the host name in the certificate. If this parameter is set to False, Informatica does not validate the certificate that the database server sends. Informatica ignores any truststore information that you specify.
HostNameInCertificate	Optional. Host name of the machine that hosts the secure database. If you specify a host name, Informatica validates the host name included in the connection string against the host name in the SSL certificate.
cryptoProtocolVersion	Required. Specifies the cryptographic protocol to use to connect to a secure database. You can set the parameter to `cryptoProtocolVersion=TLSv1.1` or `cryptoProtocolVersion=TLSv1.2` based on the cryptographic protocol used by the database server.
TrustStore	Required. Path and file name of the truststore file that contains the SSL certificate for the database. If you do not include the path for the truststore file, Informatica looks for the file in the following default directory: `<Informatica installation directory>/tomcat/bin`
TrustStorePassword	Required. Password for the truststore file for the secure database.

Note: Informatica appends the secure JDBC parameters to the JDBC connection string. If you include the secure JDBC parameters directly in the connection string, do not enter any parameter in the **Secure JDBC Parameters** field.

8. Click **Test Connection** to verify that you can connect to the database.
9. Select **No content exists under specified connection string. Create new content.**
10. Click **Finish**.

    The domain creates the Model Repository Service, creates content for the Model repository in the specified database, and enables the service.

After you create the service through the wizard, you can edit the properties or configure other properties.

After You Create the Model Repository Service

After you create the Model Repository Service, perform the following tasks:

- Create the Model repository user if the domain does not use Kerberos authentication.
- Create other application services.

Create the Model Repository User

When you create an application service that depends on the Model Repository Service, you provide the name of the Model Repository Service and of this Model repository user.

1. In the Administrator tool, click the **Security** tab.
2. On the Security Actions menu, click **Create User** to create a native user account.

   Note: If you set up LDAP authentication in the domain, you can use an LDAP user account for the Model repository user.
3. Enter the following properties for the user:

Property	Description
Login Name	Login name for the user account. The login name for a user account must be unique within the security domain to which it belongs. The name is not case sensitive and cannot exceed 128 characters. It cannot include a tab, newline character, or the following special characters: , + " \ < > ; / * % ? & The name can include an ASCII space character except for the first and last character. All other space characters are not allowed.
Password	Password for the user account. The password can be from 1 through 80 characters long.
Confirm Password	Enter the password again to confirm. You must retype the password. Do not copy and paste the password.
Full Name	Full name for the user account. The full name cannot include the following special characters: < > “
Description	Description of the user account. The description cannot exceed 765 characters or include the following special characters: < > “

4. Click **OK**.

   The user properties appear.
5. Click the **Privileges** tab.
6. Click **Edit**.

   The **Edit Roles and Privileges** dialog box appears.
7. On the **Roles** tab, expand the Model Repository Service.
8. Under **System Defined Roles**, select Administrator and click **OK**.

Create Other Services

After you create the Model Repository Service, create the application services that depend on the Model Repository Service.

Create the dependent services in the following order:

1. Data Integration Service
2. Content Management Service

Create and Configure the Data Integration Service

When an analyst uploads data or publishes prepared data, the Enterprise Data Lake Service connects to the Data Integration Service to write the data to a Hive table in the Hadoop cluster.

Create the Data Integration Service

Use the service creation wizard in the Administrator tool to create the service.

Before you create the Data Integration Service, verify that you have created the following service:

Model Repository Service

1. In the Administrator tool, click the **Manage** tab.
2. Click the **Services and Nodes** view.
3. In the Domain Navigator, select the domain.
4. Click **Actions** > **New** > **Data Integration Service**.

   The **New Data Integration Service** wizard appears.
5. On the **New Data Integration Service - Step 1 of 14** page, enter the following properties:

Property	Description
Name	Name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the service. The description cannot exceed 765 characters.
Location	Domain and folder where the service is created. Click **Browse** to choose a different folder. You can move the service after you create it.
License	License object that allows use of the service.
Assign	Select **Node** to configure the service to run on a node. If your license includes grid, you can create a grid and assign the service to run on the grid after you create the service.
Node	Node on which the service runs.
Backup Nodes	If your license includes high availability, nodes on which the service can run if the primary node is unavailable.
Model Repository Service	Model Repository Service to associate with the service.
Username	User name that the service uses to access the Model Repository Service. Enter the Model repository user that you created.

Property	Description
Password	Password for the Model repository user.
Security Domain	LDAP security domain for the Model repository user. The field appears when the Informatica domain contains an LDAP security domain. Not available for a domain with Kerberos authentication.

6. Click **Next**.

   The **New Data Integration Service - Step 2 of 14** page appears.

7. Enter the HTTP port number to use for the Data Integration Service.
8. Accept the default values for the remaining security properties. You can configure the security properties after you create the Data Integration Service.
9. Select **Enable Service**.

   The Model Repository Service must be running to enable the Data Integration Service.

10. Verify that the **Move to plugin configuration page** is not selected.
11. Click **Next**.

    The **New Data Integration Service - Step 3 of 14** page appears.

12. Set the **Launch Job Options** property to one of the following values:
    - In the service process. Configure when you run SQL data service and web service jobs. SQL data service and web service jobs typically achieve better performance when the Data Integration Service runs jobs in the service process.
    - In separate local processes. Configure when you run mapping, profile, and workflow jobs. When the Data Integration Service runs jobs in separate local processes, stability increases because an unexpected interruption to one job does not affect all other jobs.

    If you configure the Data Integration Service to run on a grid after you create the service, you can configure the service to run jobs in separate remote processes.

13. Accept the default values for the remaining execution options and click **Next**.

    The **New Data Integration Service - Step 4 of 14** page appears.

14. If you created the data object cache database for the Data Integration Service, click **Select** to select the cache connection. Select the data object cache connection that you created for the service to access the database.
15. Accept the default values for the remaining properties on this page and click **Next**.

    The **New Data Integration Service - Step 5 of 14** page appears.

16. For optimal performance, enable the Data Integration Service modules that you plan to use.

    The following table lists the Data Integration Service modules that you can enable:

Module	Description
Web Service Module	Runs web service operation mappings.
Mapping Service Module	Runs mappings and previews.
Profiling Service Module	Runs profiles and scorecards.

Module	Description
SQL Service Module	Runs SQL queries from a third-party client tool to an SQL data service.
Workflow Orchestration Service Module	Runs workflows.

17. Click **Next**.

    The **New Data Integration Service - Step 6 of 14** page appears.

    You can configure the HTTP proxy server properties to redirect HTTP requests to the Data Integration Service. You can configure the HTTP configuration properties to filter the web services client machines that can send requests to the Data Integration Service. You can configure these properties after you create the service.

18. Accept the default values for the HTTP proxy server and HTTP configuration properties and click **Next**.

    The **New Data Integration Service - Step 7 of 14** page appears.

    The Data Integration Service uses the result set cache properties to use cached results for SQL data service queries and web service requests. You can configure the properties after you create the service.

19. Accept the default values for the result set cache properties and click **Next**.

    The **New Data Integration Service - Step 8 of 14** page appears.

20. If you created the profiling warehouse database for the Data Integration Service, select the Profiling Service module.

21. If you created the workflow database for the Data Integration Service, select the Workflow Orchestration Service module.

22. Verify that the remaining modules are not selected.

    You can configure properties for the remaining modules after you create the service.

23. Click **Next**.

    The **New Data Integration Service - Step 11 of 14** page appears.

24. If you created the profiling warehouse database for the Data Integration Service, click **Select** to select the database connection. Select the profiling warehouse connection that you created for the service to access the database.

25. Select whether or not content exists in the profiling warehouse database.

    If you created a new profiling warehouse database, select **No content exists under specified connection string**.

26. Click **Next**.

    The **New Data Integration Service - Step 12 of 14** page appears.

27. Accept the default values for the advanced profiling properties and click **Next**.

    The **New Data Integration Service - Step 14 of 14** page appears.

28. If you created the workflow database for the Data Integration Service, click **Select** to select the database connection. Select the workflow database connection that you created for the service to access the database.

29. Click **Finish**.

    The domain creates and enables the Data Integration Service.

After you create the service through the wizard, you can edit the properties or configure other properties.

After You Create the Data Integration Service

After you create the Data Integration Service, perform the following tasks:

- Verify the host file configuration.
- Create other application services.

Verify the Host File Configuration

If you configured the Data Integration Service on UNIX or Linux to launch jobs as separate processes, verify that the host file on the node that runs the service contains a localhost entry. Otherwise, jobs fail when the **Launch Jobs as Separate Processes** property for the Data Integration Service is enabled.

Create Other Services

After you create the Data Integration Service, create the application services that depend on the Data Integration Service.

Create the dependent services in the following order:

1. Content Management Service

Create and Configure the Content Management Service

The Content Management Service is an application service that manages reference data. A reference data object contains a set of data values that you can search while performing data quality operations on source data. The Content Management Service also compiles rule specifications into mapplets. A rule specification object describes the data requirements of a business rule in logical terms.

The Content Management Service uses the Data Integration Service to run mappings to transfer data between reference tables and external data sources. The Content Management Service also provides transformations, mapping specifications, and rule specifications with the following types of reference data:

- Address reference data
- Identity populations
- Probabilistic models and classifier models
- Reference tables

Create the Content Management Service

Use the service creation wizard in the Administrator tool to create the service.

Before you create the Content Management Service, verify that you have created and enabled the following services:

Model Repository Service

Data Integration Service

1. In the Administrator tool, click the **Manage** tab.

2. Click **Actions** > **New** > **Content Management Service**.

   The **New Content Management Service** dialog box appears.

3. On the **New Content Management Service - Step 1 of 2** page, enter the following properties:

Property	Description
Name	Name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the service. The description cannot exceed 765 characters.
Location	Domain and folder where the service is created. Click **Browse** to choose a different folder. You can move the service after you create it.
License	License object that allows use of the service.
Node	Node on which the service runs.
HTTP Port	HTTP port number to use for the Content Management Service.
Data Integration Service	Data Integration Service to associate with the service. The Data Integration Service and the Content Management Service must run on the same node.
Model Repository Service	Model Repository Service to associate with the service.
Username	User name that the service uses to access the Model Repository Service. Enter the Model repository user that you created.
Password	Password for the Model repository user.
Security Domain	LDAP security domain for the Model repository user. The field appears when the Informatica domain contains an LDAP security domain. Not available for a domain with Kerberos authentication.
Reference Data Location	Reference data warehouse connection that you created for the Content Management Service to access the reference data warehouse. Click **Select** to select the connection.

4. Click **Next**.

   The **New Content Management Service - Step 2 of 2** page appears.

5. Accept the default values for the security properties.

6. Select **Enable Service**.

   The Model Repository Service and Data Integration Service must be running to enable the Content Management Service.

7. Click **Finish**.

   The domain creates and enables the Content Management Service.

After you create the service through the wizard, you can edit the properties or configure other properties.

After You Create the Content Management Service

After you create the Content Management Service, create the Catalog Service that depends on the Content Management Service.

Create and Configure the Interactive Data Preparation Service

The Interactive Data Preparation Service manages data preparation within Enterprise Data Preparation. When an analyst prepares data in a project, the Interactive Data Preparation Service stores worksheet metadata in the Data Preparation repository.

The service connects to the Hadoop cluster to read sample data from Hive tables. The service connects to the HDFS system in the Hadoop cluster to store the sample data being prepared in the worksheet.

Create the Interactive Data Preparation Service before you create the Enterprise Data Preparation Service. You must associate the Enterprise Data Preparation Service with a Interactive Data Preparation Service.

Create the Interactive Data Preparation Service

If you did not create the Interactive Data Preparation Service service during the console, or if you ran the silent installer, create the service through the Administrator tool.

Before you create the Interactive Data Preparation Service, verify that you have created and enabled the following services:

1. In the Administrator tool, click the **Manage** tab.
2. Click the **Services and Nodes** view.
3. In the Domain Navigator, select the domain.
4. Click **Actions** > **New** > **Interactive Data Preparation Service.**
5. Enter the following properties:

Property	Description
Name	Name of the Interactive Data Preparation Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the Interactive Data Preparation Service. The description cannot exceed 765 characters.
Location	Location of the Interactive Data Preparation Service in the Informatica domain. You can create the service within a folder in the domain.
License	License object with the data lake option that allows the use of the Interactive Data Preparation Service.

Property	Description
Node Assignment	Type of node in the Informatica domain on which the Interactive Data Preparation Service runs. Select **Single Node** if a single service process runs on the node or **Primary and Backup Nodes** if a service process is enabled on each node for high availability. However, only a single process runs at any given time, and the other processes maintain standby status. The **Primary and Backup Nodes** option will be available for selection based on the license configuration. Select the **Grid** option to ensure horizontal scalability by using a grid with multiple Interactive Data Preparation Service nodes. Improved scalability supports high performance, interactive data preparation during increased data volumes and number of users. Each user is assigned a node in the grid using round-robin method to distribute the load across the nodes. Default is Single Node.
Node	Name of the node on which the Interactive Data Preparation Service runs.

6. Click **Next**.
7. Enter the following Data Preparation repository properties:

Property	Description
Database Type	Type of database to use for the Data Preparation repository.
Database User Name	Database user account to use to connect to the database.
Database User Password	Password for the database user account.
Host Name	Host name of the database machine.
Port Number	Port number for the database.
Schema Name	Schema or database name of the Data Preparation repository database.

Property	Description

Connection String	Connection string used to connect to the database.
	To connect to an Oracle database, format the string as follows:
	`jdbc:informatica:oracle://<database host name>:<port>;ServiceName=<service name>`
	To connect to a non-secure MySQL or MariaDB database, format the string as follows:
	`jdbc:mysql://<database host name>:<port>`
	The connection string is optional if you connect to a non-secure database.
	To connect to an SSL-enabled MySQL or MariaDB database, format the string as follows:
	`verifyServerCertificate=true&useSSL=true&requireSSL=true`

Secure JDBC Parameters	Secure JDBC parameters required to access a secure database.
	To connect to a secure Oracle database, format the string as follows:
	`EncryptionMethod=SSL;HostNameInCertificate=<secure database host name>;ValidateServerCertificate=true`
	To connect to as secure MySQL or MariaDB database, format the string as follows:
	`trustCertificateKeyStoreUrl=file://<truststore path/truststore file name>&trustCertificatekeyStorePassword=<truststore password>`

8. Click **Next**.
9. Enter the following storage properties:

Property	Description
Local Storage Location	Directory for data preparation file storage on the node where the service runs.
Durable Storage Type	Storage type for the data preparation file.
Durable Storage Connection	Connection for the data preparation file storage.
Durable Storage Location	Location for the data preparation file storage. If the connection to the local storage fails, the service recovers data preparation files from the location. If the Hadoop cluster uses Kerberos authentication, the impersonation user name must have read, write and execute permission on the HDFS storage location directory. The default location is: /datalake/dps_durable_storage.
Hadoop Authentication Mode	Security mode enabled for the Hadoop cluster for data preparation storage. If the Hadoop cluster uses Kerberos authentication, you must set the required Hadoop security properties for the cluster.

10. Click **Next**.
11. Enter the logging properties.
12. If you plan to use rules, you must associate the Interactive Data Preparation Service with the Model Repository Service that manages the Model repository that contains the rule objects and metadata. You must also associate a Data Integration Service that runs rules during data preparation with the Interactive Data Preparation Service.

Enter the following properties required to enable rules:

Property	Description
Enable Rule Execution	Enables the execution of the validation rule objects.
Model Repository Service Name	Name of the Model Repository Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [You cannot change the name of the service after you create it.
Model Repository Service User Name	User name to access the Model Repository Service.
Model Repository Service Password	Password to access the Model Repository Service.
Security Domain	Select the security domain to access the Model Repository Service.
Data Integration Service Name	Name of the Data Integration Service.

13. Click **Next**.
14. Enter the following HTTP configuration properties:

Property	Description
HTTP Port	Port number for the HTTP connection to the Interactive Data Preparation Service.
Enable Secure Communication	Use a secure connection to connect to the Interactive Data Preparation Service. If you enable secure communication, you must set all required HTTPS properties, including the keystore and truststore properties.
HTTPS Port	Port number for the HTTPS connection to the Interactive Data Preparation Service.
Keystore File	Path and the file name of keystore file that contains key and certificates required for HTTPS communication.
Keystore Password	Password for the keystore file.
Truststore File	Path and the file name of truststore file that contains authentication certificates for the HTTPS connection.
Truststore Password	Password for the truststore file.

15. Click **Next**.
16. Enter the following rules execution property:

Property	Description
Rules Server Port	Port used by the rules server managed by the Interactive Data Preparation Service. Set the value to an available port on the node where the service runs.

17. Click **Finish**.
18. Select the Interactive Data Preparation Service in the Domain Navigator, and then select **Actions** > **Create Repository** to create the repository contents.
19. Select **Actions** > **Enable Service** to enable the Interactive Data Preparation Service.

Install Python for Enterprise Data Preparation

Enterprise Data Preparation uses the Apache Solr indexing capabilities to provide recommendations of related data assets. Apache Solr requires Python modules.

You must install Python with the following modules on every node that hosts the Interactive Data Preparation Service associated with Enterprise Data Preparation:

argparse

sys

getopt

os

urllib

httplib2

ConfigParser

Enable Data Preparation of JSON Files on Cloudera CDH

If you integrate Enterprise Data Preparation with a Cloudera CDH Hadoop cluster, you must specify the location of a Hive .jar file in the Hive Auxiliary JARs Directory in CDH to enable data preparation of JSON files.

1. Search for hive-hcatalog-core.jar in the Cloudera CDH installation directory.

   You can generally find the .jar file in the following directory:

   `/opt/cloudera/parcels/CDH/lib/hive-hcatalog/share/hcatalog/`
2. Log in to Cloudera Manager.
3. Select **Hive** in the cluster.
4. Click the **Configuration** tab, then select the **Advanced** category.
5. Enter the path to the directory containing the .jar file in the Hive Auxiliary JARs Directory property.

   If the file is in the location noted, the path to the directory is:

   `/opt/cloudera/parcels/CDH/lib/hive-hcatalog/share/hcatalog/`
6. Restart the Hive server.

Create and Configure the Enterprise Data Preparation Service

The Enterprise Data Preparation Service runs the Enterprise Data Preparation application in the Informatica domain. Enterprise Data Preparation requires the Enterprise Data Preparation Service to complete operations.

When an analyst uploads data, the Enterprise Data Preparation Service connects to the HDFS system in the Hadoop cluster to temporarily stage the data. When an analyst previews data, the Enterprise Data Preparation Service connects to the Hadoop cluster to read the data.

Create the Enterprise Data Preparation Service

If you did not create the Enterprise Data Preparation Service service during the console, or if you ran the silent installer, create the service through the Administrator tool.

Before you create the Enterprise Data Preparation Service, verify that you have created and enabled the following services:

Catalog Service

Interactive Data Preparation Service

1. In the Administrator tool, click the **Manage** tab.
2. Click the **Services and Nodes** view.
3. In the Domain Navigator, select the domain.
4. Click **Actions** > **New** > **Enterprise Data Preparation Service.**
5. Enter the following properties:

Property	Description
Name	Name of the Enterprise Data Preparation Service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of theEnterprise Data Preparation Service. The description cannot exceed 765 characters.
Location	Location of the Enterprise Data Preparation Service in the Informatica domain. You can create the service within a folder in the domain.
License	License object that allows the use of the Enterprise Data Preparation Service.
Node Assignment	Type of node in the Informatica domain on which the Enterprise Data Preparation Service runs. Select **Single Node** if a single service process runs on the node or **Primary and Backup Nodes** if a service process is enabled on each node for high availability. However, only a single process runs at any given time, and the other processes maintain standby status. The **Primary and Backup Nodes** option is available based on the license configuration. Default is Single Node.
Node	Name of the node on which theEnterprise Data Preparation Service runs.

6. Click **Next**.

7. Enter the following properties for the Model Repository Service:

Property	Description
Model Repository Service	Name of the Model Repository Service associated with the Enterprise Data Preparation Service.
Model Repository Service User Name	User account to use to log in to the Model Repository Service.
Model Repository Service User Password	Password for the Model Repository Service user account.

8. Click **Next**.
9. Enter the following properties for the Interactive Data Preparation Service, Data Integration Service, and Catalog Service:

Property	Description
Interactive Data Preparation Service	Name of the Interactive Data Preparation Service associated with the Enterprise Data Preparation Service.
Data Integration Service	Name of the Data Integration Service associated with the Enterprise Data Preparation Service.
Catalog Service	Name of the Catalog Service associated with the Enterprise Data Preparation Service.
Catalog Service User Name	User account to use to log in to the Catalog Service.
Catalog Service User Password	Password for the Catalog Service user account.

10. Click **Next**.
11. Enter the following execution properties:

Property	Description
Execution Engine	Engine to run the mappings.
Hadoop Connection	Hadoop connection for the data lakehouse.
HDFS Connection	HDFS connection for the Hadoop working directory.
Hadoop Working Directory	HDFS directory where the Enterprise Data Preparation Service copies temporary data and files necessary for the service to run. This directory must have permissions to enable users to upload data.

Property	Description
Hadoop Authentication Mode	Security mode of the Hadoop cluster for the data lake. If the Hadoop cluster uses Kerberos authentication, you must set the required Hadoop security properties for the cluster.
Local Working Directory	Local directory that contains the files downloaded from the Enterprise Data Preparation Service application, such as .csv or .tde files

12. Click **Next**.
13. Enter the following user event logging properties:

Property	Description
Log User Activity Events	Indicates whether the Enterprise Data Preparation Service logs user activity events.
Solr JVM Options	Solr JVM options required to connect to the specified JDBC port used to retrieve data from Zookeeper. Set to connect to Zookeeper from an external client.
Index Directory	Location of a shared NFS directory used by primary and secondary nodes in a multiple node installation.

14. Click **Next**.
15. Enter the logging properties.
16. Click **Next**.
17. Enter the following advanced properties:

Property	Description
Maximum Concurrent Upload/Download Activities	Maximum concurrent upload or download activities. You can specify a maximum of 2,000,000,000 activities to run concurrently. Enter a value of -1 (default) to run unbounded number of activities concurrently.

18. Click **Next**.
19. Enter the following properties:

Property	Description
HTTP Port	Port number for the HTTP connection to the Enterprise Data Preparation Service.
Enable Secure Communication	Use a secure connection to connect to the Enterprise Data Preparation Service. If you enable secure communication, you must enter all required HTTPS options.
HTTPS Port	Port number for the HTTPS connection to the Enterprise Data Preparation Service.
Keystore File	Path and the file name of keystore file that contains key and certificates required for the HTTPS connection.

Property	Description
Keystore Password	Password for the keystore file.
Truststore File	Path and the file name of the truststore file that contains authentication certificates for the HTTPS connection.
Truststore Password	Password for the truststore file.

20. Select **Enable Service** if you want to enable the service immediately after you create the service.

    If you want to enable the service at a later time, in the Domain Navigator, select the service and then select **Actions** > **Enable Service**.

21. Click **Finish**.

Create and Configure the Catalog Service

Create a Catalog Service to run Enterprise Data Catalog and manage the connections between the Enterprise Data Catalog components. You can configure the general, application service, and security properties of the Catalog Service

If you plan to deploy Enterprise Data Catalog on multiple nodes, ensure that you configure the Informatica Cluster Service and Catalog Service on separate nodes.

Before you create the Catalog Service, verify that you have created and enabled the following services:

Model Repository Service

Content Management Service

Data Integration Service

Informatica Cluster Service

Note: The Catalog Service has the same privileges as the user account that creates it. Ensure that the user account does not have privileges to read or modify sensitive files on the system.

1. In the Administrator tool, select a domain, and click the **Services and Nodes** tab.
2. On the Actions menu, click **New** > **Catalog Service**.

   The **New Catalog Service Step 1 of 5** dialog box appears.
3. Configure the general properties in the dialog box.

The following table describes the properties:

Property	Description
Name	Name of the service. The name is not case-sensitive and must be unique within the domain. The name cannot exceed 128 characters or begin with @. The name cannot contain character spaces. The characters in the name must be compatible with the code page of the Model repository that you associate with the Catalog Service. The name cannot contain the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the service. The description cannot exceed 765 characters.
Location	Domain in which the service runs.
License	License to assign to the Catalog Service. Select the license that you installed with Informatica.
Node	Node in the Informatica domain on which the Catalog Service runs. If you change the node, you must recycle the Catalog Service.
Backup Nodes	If your license includes high availability, nodes on which the service can run if the primary node is unavailable.

4. Click **Next**.

   The **New Catalog Service - Step 2 of 5** dialog box appears.

5. Configure the application service properties in the dialog box.

   The following table describes the properties:

Property	Description
Model Repository Service	Model Repository Service to associate with the Catalog Service. The Model Repository Service manages the Model repository that Enterprise Data Catalog uses. If you update the property to specify a different Model Repository Service, recycle the Catalog Service.
User name	The database user name for the Model repository.
Password	An encrypted version of the database password for the Model repository.
Security Domain	Name of the security domain that includes the **User name**.

6. Click **Next**.

   The **New Catalog Service - Step 3 of 5** dialog box appears.

7. Configure the security properties in the dialog box.

The following table describes the properties:

Property	Description
HTTP Port	A unique HTTP port number used for each Data Integration Service process. Default is 8085.
Enable Transport Layer Security	Indicates that the Catalog Service must use HTTPS. If you did not configure the Data Integration Service to use HTTPS, the Catalog Service does not start. If the cluster is enabled for SSL, make sure that you enable SSL for the Informatica domain and the application services.
HTTPS Port	Port number for the HTTPS connection.
Keystore File	Path and file name of the keystore file. The keystore file contains the keys and certificates required if you use the SSL security protocol with Catalog Administrator. Required if you select Enable Transport layer Security. When Enterprise Data Catalog creates the Catalog Service, Enterprise Data Catalog exports the keystore to a certificate and stores the certificate in the keystore directory. Ensure that you configure the read and write permissions on the directory for Enterprise Data Catalog to successfully store the certificate. **Note:** Verify that you specified the correct keystore file for the Catalog Service. The certificates in the keystore must be trusted by the domain truststore. The keystore file must contain CA-signed certificates for custom SSL configuration.
Keystore Password	Password for the keystore file. Required if you select Enable Transport layer Security.
SSL Protocol	Secure Sockets Layer protocol to use.

8. Click **Next**.

   The **New Catalog Service - Step 4 of 5** dialog box appears.

9. Configure the following properties in the dialog box.

   The following table describes the properties:

Property	Description
Informatica Cluster Service	Name of the Informatica Cluster Service that you must associate with the Catalog Service.
Receive Alerts through Email	Choose to receive email notifications on the Catalog Service status. **Note:** If you select this option, you must enable the Email Service. For more information about enabling Email Service, see the Administrator Reference for Enterprise Data Catalog guide.
Enable Catalog Service	Select the option to enable the Catalog Service.
Enable Email Notifications for Asset Changes	Select the option to receive email notifications in Enterprise Data Catalog when there are updates for assets.

10. Click **Next**.

    The **New Catalog Service - Step 5 of 5** dialog box appears.

11. Optional. Click **Enable Data Asset Analytics** to configure the properties to enable Data Asset Analytics for Enterprise Data Catalog. You can use Data Asset Analytics with Enterprise Data Catalog to gain analytical insights into asset details, such as values, enrichment, and collaboration using reports and charts.

    Configure the following properties to enable Data Asset Analytics:

Property	Description
Select Database	Select the repository database that you want to use for Data Asset Analytics from the following options: - Oracle - SQLServer - PostgreSQL
User Name	The database user name for the repository.
Password	The password for the database user name.

Database Connection String	Enter the JDBC connection string to connect to the repository database.
	Use the following syntax for the connection string based on the database selected:
	- Oracle.
	`jdbc:informatica:oracle://<host name>:<port number>;ServiceName=<database name>`
	- SQLServer:
	  - SQL Server.
	`jdbc:informatica:sqlserver://<host name>: <port number>;DatabaseName=<database name>;SnapshotSerializable=true`
	- PostgreSQL.
	`jdbc:informatica:postgresql://<host name>: <port number>;DatabaseName=<database name>`

Secure JDBC Parameters	If the repository database is secured with the SSL protocol, you must enter the secure database parameters as name=value pairs separated by semicolon characters (;). For example: `param1=value1;param2=value2`

    Data Asset Analytics supports the following schemas for the databases listed:

    - dbo schema for SQL Server.
    - public schema for PostgreSQL.

    Optional. Click **Test Connection** if you want to validate the configuration details.

12. Click **Finish**.

Configure the Advanced Scanners Server

If you configured the repository for Advanced Scanners, you must configure Enterprise Data Catalog details in the Advanced Scanners server.

Perform the following steps to configure the details:

1. Start the Catalog Service.
2. Log in to the Advanced Scanners web interface as an administrator. You can access the web interface by providing the URL in the <host>:<port> format. <host> represents the host name configured for the Informatica domain and <port> represents the port number configured for the Advanced Scanners repository server.

Note: For enhanced security, you must change the password after you log in for the first time.

3. Click **Administration** > **Global Variables**.
4. Configure the following variables with the values shown:

Variable	Value
EDC_USER	The user name configured as an administrator in Enterprise Data Catalog.
EDC_PASSWORD	The password configured for the administrator.
EDC_URL	The host name and port number configured for Enterprise Data Catalog in the following format: <host>:<port>

5. Shut down the Advanced Scanners server using the `server.sh stop` command available in the following directory: `<INFA_HOME>/services/CatalogService/AdvancedScannersApplication/app/`
6. Start the Advanced Scanners repository server using the following command: `server.sh &`.

Create and Configure the Metadata Access Service

The Metadata Access Service is an application service that allows the Developer tool to access Hadoop connection information to import and preview metadata.

The Metadata Access Service contains information about the Service Principal Name (SPN) and keytab information if the Hadoop cluster uses Kerberos authentication.

Create the Metadata Access Service

The Metadata Access Service allows the Developer tool to access Hadoop connection information to import and preview metadata from the Hadoop environment. The Metadata Access Service is required for design-time access to the Hadoop environment.

1. In the Administrator tool, click the **Manage** tab.
2. Click the **Services and Nodes** view.
3. In the Domain Navigator, select the domain.
4. Click **Actions** > **New** > **Metadata Access Service**.

   The **New Metadata Access Service** wizard appears.

5. On the **New Metadata Access Service - Step 1 of 3** page, enter the following properties:

Property	Description
Name	Name of the service. The name is not case sensitive and must be unique within the domain. It cannot exceed 128 characters or begin with @. It also cannot contain spaces or the following special characters: ` ~ % ^ * + = { } \ ; : ' " / ? . , < > | ! ( )] [
Description	Description of the service. The description cannot exceed 765 characters.
Location	Domain and folder where the service is created. Click **Browse** to choose a different folder. You can move the service after you create it.
License	License object that allows use of the service.
Node	Node on which the service runs.
Backup Nodes	If your license includes high availability, nodes on which the service can run if the primary node is unavailable.

6. Click **Next**.

   The **New Metadata Access Service - Step 2 of 3** page appears.

7. Select the HTTP Protocol Type and enter the respective port number to use for the Metadata Access Service.
8. Accept the default values for the remaining security properties. You can configure the security properties after you create the Metadata Access Service.
9. Select **Enable Service**.

   The Metadata Access Service does not have any other service dependency.

10. Click **Next**.

    The **New Metadata Access Service - Step 3 of 3** page appears.

11. If applicable, specify the execution options for impersonation user, Kerberos cluster, logging options, and click **Next**.
12. Click **Finish**.

    The domain creates and enables the Metadata Access Service.

Part V: Informatica Client Installation

This part contains the following chapters:

CHAPTER 17

Install Informatica Developer

This chapter includes the following topics:

Before You Install Informatica Developer

Before you install the Informatica Developer, verify that the minimum system and third-party software requirements are met. If the machine where you install the Informatica Developer is not configured correctly, the installation can fail.

Verify System Requirements

Before you install the client, verify the following installation requirements to install and run the client are met:

Disk space for the temporary files

The installer writes temporary files to the hard disk. Verify that you have 1 GB disk space on the machine to support the installation. When the installation completes, the installer deletes the temporary files and releases the disk space.

Permissions to install

Verify that the user account that you use to install the client has write permission on the installation directory and Windows registry.

Minimum system requirements

The following table lists the minimum system requirements to run the client:

Processor	RAM	Disk Space
1 CPU	1GB	6 GB

Verify Third-party Requirements for Informatica Developer

Before you install the Developer tool, verify the following third-party installation requirements:

- Install the .NET Framework 4.0 or later. If you plan to use Data Processor or Hierarchical-To-Relational transformations, you must install the .NET Framework before you install the Developer tool.
- Install the latest version of Microsoft Visual C++ Redistributable Package (x64) before you use or install the Developer tool. You can download it from the Microsoft website.

Install the Developer tool

Perform the following steps to install the Developer tool:

1. Close all other applications.
2. Go to the root of the directory for the installation files and run install.bat as administrator.

   To run the file as administrator, right-click the install.bat file and select **Run as administrator**.

   Note: If you do not run the installer as administrator, the Windows system administrator might encounter issues when accessing files in the Informatica installation directory.

   If you encounter problems when you run the install.bat file from the root directory, run the following file:
   `<installer files directory>\client\install.exe`
3. Select **Install Informatica <Version> Clients** and click **Next**.
4. Read the terms and conditions for Informatica installation and the product usage toolkit and select **I agree to the terms and conditions**.
   a. Press **1** if you do not want to accept the terms and conditions.
   b. Press **2** to accept the terms and conditions.
5. Version 10.5.1 is for installing Informatica 10.5.1 products.
   a. Press **1** and type **quit** to quit the installation.
   b. Press **2** to continue the installation.

   If you choose to not accept the terms and condition, the installer prompts you to accept the terms and conditions.
6. The **Installation Pre-requisites** page displays the system requirements. Verify that all installation requirements are met before you continue the installation.
7. On the **Installation Directory** page, enter the absolute path for the installation directory.

   The installation directory must be on the current computer. The maximum length of the path must be less than 260 characters. The directory names in the path must not contain spaces or the following special characters: @|* $ # ! % ( ) { } [ ] , ; '

   Note: Informatica recommends using alphanumeric characters in the installation directory path. If you use a special character such as á or €, unexpected results might occur at run time.
8. Click **Next**.
9. On the **Pre-Installation Summary** page, review the installation information, and click **Install**.

   The installer copies the Developer tool files to the installation directory.

   The **Post-installation Summary** page indicates whether the installation completed successfully.
10. Click **Done** to close the installer.

You can view the installation log files to get more information about the tasks performed by the installer.

After You Install Informatica Developer

After you install Informatica Developer, you can install other languages, enable secure communication within the domain, and start the Developer tool.

Install Languages

To view languages other than the system locale and to work with repositories that use a UTF-8 code page, install additional languages on Windows for use with the Informatica clients.

You also must install languages to use the Windows Input Method Editor (IME).

1. Click **Start** > **Settings** > **Control Panel**.
2. Click **Regional Options**.
3. Under Language settings for the system, select the languages you want to install.
4. Click **Apply**.

If you change the system locale when you install the language, restart the Windows machine.

Configure the Client for a Secure Domain

When you enable secure communication within the domain, you also secure connections between the domain and Informatica client applications. Based on the truststore files used, you might need to specify the location and password for the truststore files in environment variables on each client host.

You might need to set the following environment variables on each client host:

INFA_TRUSTSTORE

Set this variable to the directory that contains the truststore files for the SSL certificates. The directory must contain truststore files named `infa_truststore.jks` and `infa_truststore.pem`.

INFA_TRUSTSTORE_PASSWORD

Set this variable to the password for the `infa_truststore.jks` file. The password must be encrypted. Use the command line program pmpasswd to encrypt the password.

Informatica provides an SSL certificate that you can use to secure the domain. When you install the Informatica clients, the installer sets the environment variables and installs the truststore files in the following directory by default: `<Informatica installation directory>\clients\shared\security`

If you use the default Informatica SSL certificate, and the `infa_truststore.jks` and `infa_truststore.pem` are in the default directory, you do not need to set the INFA_TRUSTSTORE or INFA_TRUSTSTORE_PASSWORD environment variables.

You must set the INFA_TRUSTSTORE and INFA_TRUSTSTORE_PASSWORD environment variables on each client host in the following scenarios:

You use a custom SSL certificate to secure the domain.

If you provide an SSL certificate to use to secure the domain, copy the `infa_truststore.jks` and `infa_truststore.pem` truststore files to each client host. You must specify the location of the files and the truststore password.

You use the default Informatica SSL certificate, but the truststore files are not in the default Informatica directory.

If you use the default Informatica SSL certificate, but the `infa_truststore.jks` and `infa_truststore.pem` truststore files are not in the default Informatica directory, you must specify the location of the files and the truststore password.

Configure the Developer Tool Workspace Directory

Configure Informatica Developer to write the workspace metadata to the machine where the user is logged in.

1. Go to the following directory: `<Informatica installation directory>\clients\DeveloperClient\configuration\`
2. Locate the config.ini file.
3. Create a backup copy of the config.ini file.
4. Use a text editor to open the config.ini file.
5. Add the `osgi.instance.area.default` variable to the end of the config.ini file and set the variable to the directory location where you want to save the workspace metadata. The file path cannot contain non-ANSI characters. Folder names in the workspace directory cannot contain the number sign (#) character. If folder names in the workspace directory contain spaces, enclose the full directory in double quotes.
   - If you run Informatica Developer from the local machine, set the variable to the absolute path of the workspace directory:

     ```
     osgi.instance.area.default=<Drive>/<WorkspaceDirectory>
     ```

     or

     ```
     osgi.instance.area.default=<Drive>\\<WorkspaceDirectory>
     ```
   - If you run Informatica Developer from a remote machine, set the variable to the directory location on the local machine:

     ```
     osgi.instance.area.default=\\\\<LocalMachine>/<WorkspaceDirectory>
     ```

     or

     ```
     osgi.instance.area.default=\\\\<LocalMachine>\\<WorkspaceDirectory>
     ```

     The user must have write permission to the local workspace directory.

Informatica Developer writes the workspace metadata to the workspace directory. If you log into Informatica Developer from a local machine, Informatica Developer writes the workspace metadata to the local machine. If the workspace directory does not exist on the machine from which you logged in, Informatica Developer creates the directory when it writes the files.

You can override the workspace directory when you start Informatica Developer.

Starting the Developer Tool

When you start the Developer tool, you connect to a Model repository. The Model repository stores metadata created in the Developer tool. The Model Repository Service manages the Model repository. Connect to the repository before you create a project.

1. From the Windows Start menu, click **Programs** > **Informatica[Version]** > **Client** > **Developer Client** > **Launch Informatica Developer**.

   The first time you run the Developer tool, the Welcome page displays several icons. The Welcome page does not appear when you run the Developer tool subsequently.
2. Click **Workbench.**

   The first time you start the Developer tool, you must select the repository in which to save the objects you create.
3. Click **File** > **Connect to Repository**.

The **Connect to Repository** dialog box appears.

4. If you have not configured a domain in the Developer tool, click **Configure Domains** to configure a domain.

   You must configure a domain to access a Model Repository Service.

5. Click **Add** to add a domain.

   The **New Domain** dialog box appears.

6. Enter the domain name, host name, and port number.
7. Click **Finish**.
8. Click **OK**.
9. In the **Connect to Repository** dialog box, click **Browse** and select the Model Repository Service.
10. Click **OK**.
11. Click **Next**.
12. Enter a user name and password.
13. Click **Finish**.

    The Developer tool adds the Model repository to the Object Explorer view. When you run the Developer tool the next time, you can connect to the same repository.

CHAPTER 18

Install in Silent Mode

This chapter includes the following topics:

Overview of Install in Silent Mode

To install the Informatica clients without user interaction, install in 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 the Informatica clients on multiple machines on the network or to standardize the installation across machines.

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.

Configure the Properties File

Informatica provides a sample properties file that includes the properties required by the installer. Customize the sample properties file to create a properties file and specify the options for your installation. Then run the silent installation.

The sample SilentInput.properties file is stored in the installer download location.

1. Go to the root of the directory that contains the installation files.
2. Locate the sample `SilentInput.properties` file.
3. Create a backup copy of the SilentInput.properties file.
4. Use a text editor to open and modify the values of the properties in the file.

The following table describes the installation properties that you can modify:

Property Name	Description
INSTALL_TYPE	Indicates whether to install or upgrade the Informatica clients. If the value is 0, the Informatica clients are installed in the directory you specify. If the value is 1, the Informatica clients are upgraded. Default is 0.
USER_INSTALL_DIR	Informatica client installation directory.
DXT_COMP	Indicates whether to install Informatica Developer. If the value is 1, the Developer tool will be installed. If the value is 0, the Developer tool will not be installed. Default is 1.

5. Save the properties file.

Run the Silent Installer

After you configure the properties file, open a command prompt to start the silent installation.

1. Open a command prompt.
2. Go to the root of the directory that contains the installation files.
3. Verify that the directory contains the file SilentInput.properties that you edited and resaved.
4. To run the silent installation, run silentInstall.bat.

   The silent installer runs in the background. The process can take a while. The silent installation is complete when the Informatica_<Version>_Client_InstallLog<timestamp>.log file is created in the installation directory.

   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.

Part VI: Uninstallation

This part contains the following chapter:

CHAPTER 19

Uninstallation

This chapter includes the following topics:

Informatica Uninstallation Overview

Uninstall Informatica to remove the Informatica server or clients from a machine.

The Informatica uninstallation process deletes all Informatica files and clears all Informatica configurations from a machine. The uninstallation process does not delete files that are not installed with Informatica. For example, the installation process creates temporary directories. The uninstaller does not keep a record of these directories and therefore cannot delete them. You must manually delete these directories for a clean uninstallation.

Important: If you install the Informatica services and the PowerCenter Client in the same install directory, the service binaries will be uninstalled when you uninstall the PowerCenter Client.

Rules and Guidelines for Uninstallation

Use the following rules and guidelines when you uninstall Informatica components:

- The Informatica server uninstallation mode depends on the mode you use to install Informatica server. For example, you install Informatica server in console mode. When you run the uninstaller, it runs in console mode. The Informatica clients uninstallation mode does not depend on the mode you use to install Informatica clients. For example, you install Informatica clients in silent mode. When you run the uninstaller, it can run in graphical or silent mode.
- Uninstalling Informatica does not affect the Informatica repositories. The uninstaller removes the Informatica files. It does not remove repositories from the database. If you need to move the repositories, you can back them up and restore them to another database.
- Uninstalling Informatica does not remove the metadata tables from the domain configuration database. If you install Informatica again using the same domain configuration database and user account, you must manually remove the tables or choose to overwrite the tables. You can use the infasetup BackupDomain

command to back up the domain configuration database before you overwrite the metadata tables. To remove the metadata tables manually, use the infasetup DeleteDomain command before you run the uninstaller.

- Uninstalling Informatica removes all installation files and subdirectories from the Informatica installation directory. Before you uninstall Informatica, stop all Informatica services and processes and verify that all of the files in the installation directory are closed. At the end of the uninstallation process, the uninstaller displays the names of the files and directories that could not be removed.
- The Informatica server installation creates the following folder for the files and libraries required by third party adapters built using the Informatica Development Platform APIs:
  `<Informatica installation directory>/services/shared/extensions`

  Uninstalling the Informatica server deletes this folder and any subfolders created under it. If you have adapter files stored in the /extensions folder, back up the folder before you start uninstallation.
- If you perform the uninstallation on a machine, you must back up the ODBC folder before you uninstall. Restore the folder after the uninstallation completes.

Uninstalling the Informatica Server in Console Mode

If you installed the Informatica server in console mode, uninstall the Informatica server in console mode.

Before you run the uninstaller, stop all Informatica services and processes and verify that all files in the installation directory are closed. The uninstallation process cannot remove files that are open or are being used by a service or process that is running.

1. Go to the following directory:

   `<Informatica installation directory>/Uninstaller_Server`
2. Type the following command to run the uninstaller:

   ```
   ./uninstaller.sh
   ```

   If you installed the Informatica server in console mode, the uninstaller launches in console mode.

Uninstalling Informatica Server in Silent Mode

If you installed the Informatica server in silent mode, uninstall the Informatica server in silent mode.

Before you run the uninstaller, stop all Informatica services and processes and verify that all files in the installation directory are closed. The uninstallation process cannot remove files that are open or are being used by a service or process that is running.

1. Go to the following directory:

   `<Informatica installation directory>/Uninstaller_Server`
2. Type the following command to run the silent uninstaller:

   ```
   ./uninstaller.sh
   ```

   If you installed the Informatica server in silent mode, the uninstaller launches in silent mode. The silent uninstaller runs in the background. The process can take a while. The silent uninstallation fails if the installation directory is not accessible.

After you uninstall the the Informatica server, delete any remaining folders and files from the Informatica installation directory. For example:

- Informatica _<Version>_Services_InstallLog.log file
- Informatica_<Version>_Services_<timestamp>.log file

APPENDIX A

Starting and Stopping Informatica Services

This appendix includes the following topics:

Starting and Stopping Informatica Services Overview

The Informatica service runs the Service Manager on the node. The Service Manager manages all domain functions and starts application services configured to run on the node. The method you use to start or stop Informatica depends on the operating system. You can use Informatica Administrator to shut down a node. When you shut down a node, you stop Informatica on the node.

The Informatica service also runs Informatica Administrator. You use Informatica Administrator to administer the Informatica domain objects and user accounts. Log in to Informatica Administrator to create the user accounts for users of Informatica and to create and configure the application services in the domain.

Stopping Informatica in Informatica Administrator

When you shut down a node using Informatica Administrator, you stop the Informatica service on that node.

You can abort the processes that are running or allow them to complete before the service shuts down. If you shut down a node and abort the repository service processes running on the node, you can lose changes that have not yet been written to the repository. If you abort a node running integration service processes, the workflows will abort.

1. Log in to Informatica Administrator.
2. In the Navigator, select the node to shut down.
3. On the Domain tab **Actions** menu, select **Shutdown Node**.

Rules and Guidelines for Starting or Stopping Informatica

Consider the following rules and guidelines when starting and stopping Informatica on a node:

- When you shut down a node, the node is unavailable to the domain. If you shut down a gateway node and do not have another gateway node in the domain, the domain is unavailable.
- When you start Informatica, verify that the port used by the service on the node is available. For example, if you stop Informatica on a node, verify that the port is not used by any other process on the machine before you restart Informatica. If the port is not available, Informatica will fail to start.
- If you do not use Informatica Administrator to shut down a node, any process running on the node will be aborted. If you want to wait for all processes to complete before shutting down a node, use Informatica Administrator.
- If you have two nodes in a domain with one node configured as a primary node for an application service and the other node configured as a backup node, start Informatica on the primary node before you start the backup node. Otherwise, the application service will run on the backup node and not the primary node.

APPENDIX B

Connecting to Databases from UNIX or Linux

This appendix includes the following topics:

Connecting to an IBM DB2 Universal Database

For native connectivity, install the version of IBM DB2 Client Application Enabler (CAE) appropriate for the IBM DB2 database server version. To ensure compatibility between Informatica and databases, use the appropriate database client libraries.

Configuring Native Connectivity

You can configure native connectivity to an IBM DB2 database to increase performance.

The following steps provide a guideline for configuring native connectivity. For specific instructions, see the database documentation.

1. To configure connectivity on the machine where the Data Integration Service, PowerCenter Integration Service, or PowerCenter Repository Service process runs, log in to the machine as a user who can start a service process.
2. Set the DB2INSTANCE, INSTHOME, DB2DIR, and PATH environment variables.

   The UNIX IBM DB2 software always has an associated user login, often db2admin, which serves as a holder for database configurations. This user holds the instance for DB2.

   DB2INSTANCE. The name of the instance holder.

   Using a Bourne shell:

   ```
   $ DB2INSTANCE=db2admin; export DB2INSTANCE
   ```

Using a C shell:

```
$ setenv DB2INSTANCE db2admin
```

INSTHOME. This is db2admin home directory path.

Using a Bourne shell:

```
$ INSTHOME=~db2admin
```

Using a C shell:

```
$ setenv INSTHOME ~db2admin>
```

DB2DIR. Set the variable to point to the IBM DB2 CAE installation directory. For example, if the client is installed in the /opt/IBM/db2/V9.7 directory:

Using a Bourne shell:

```
$ DB2DIR=/opt/IBM/db2/V9.7; export DB2DIR
```

Using a C shell:

```
$ setenv DB2DIR /opt/IBM/db2/V9.7
```

PATH. To run the IBM DB2 command line programs, set the variable to include the DB2 bin directory.

Using a Bourne shell:

```
$ PATH=${PATH}:$DB2DIR/bin; export PATH
```

Using a C shell:

```
$ setenv PATH ${PATH}:$DB2DIR/bin
```

3. Set the shared library variable to include the DB2 lib directory.

   The IBM DB2 client software contains a number of shared library components that the Data Integration Service, PowerCenter Integration Service, and PowerCenter Repository Service processes load dynamically. Set the shared library environment variable so that the services can find the shared libraries at run time.

   The shared library path must also include the Informatica installation directory *(server_dir)*.

   Set the shared library environment variable based on the operating system.

   The following table describes the shared library variables for each operating system:

Operating System	Variable
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

   For example, use the following syntax for Linux:

   - Using a Bourne shell:

     ```
     $ LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:$HOME/server_dir:$DB2DIR/lib; export
     LD_LIBRARY_PATH
     ```

   - Using a C shell:

     ```
     $ setenv LD_LIBRARY_PATH ${LD_LIBRARY_PATH}:$HOME/server_dir:$DB2DIR/lib
     ```

   For AIX:

   - Using a Bourne shell:

     ```
     $ LIBPATH=${LIBPATH}:$HOME/server_dir:$DB2DIR/lib; export LIBPATH
     ```

   - Using a C shell:

     ```
     $ setenv LIBPATH ${LIBPATH}:$HOME/server_dir:$DB2DIR/lib
     ```

4. Edit the .cshrc or .profile to include the complete set of shell commands. Save the file and either log out and log in again or run the source command.

   Using a Bourne shell:

   ```
   $ source .profile
   ```

   Using a C shell:

   ```
   $ source .cshrc
   ```

5. If the DB2 database resides on the same machine on which the Data Integration Service, PowerCenter Integration Service, or PowerCenter Repository Service process runs, configure the DB2 instance as a remote instance.

   Run the following command to verify if there is a remote entry for the database:

   ```
   DB2 LIST DATABASE DIRECTORY
   ```

   The command lists all the databases that the DB2 client can access and their configuration properties. If this command lists an entry for “Directory entry type” of “Remote,” skip to 7.

6. If the database is not configured as remote, run the following command to verify whether a TCP/IP node is cataloged for the host:

   ```
   DB2 LIST NODE DIRECTORY
   ```

   If the node name is empty, you can create one when you set up a remote database. Use the following command to set up a remote database and, if needed, create a node:

   ```
   db2 CATALOG TCPIP NODE <nodename> REMOTE <hostname_or_address> SERVER <port number>
   ```

   Run the following command to catalog the database:

   ```
   db2 CATALOG DATABASE <dbname> as <dbalias> at NODE <nodename>
   ```

   For more information about these commands, see the database documentation.

7. Verify that you can connect to the DB2 database. Run the DB2 Command Line Processor and run the command:

   ```
   CONNECT TO <dbalias> USER <username> USING <password>
   ```

   If the connection is successful, clean up with the `CONNECT RESET` or `TERMINATE` command.

Connecting to a Microsoft SQL Server Database

Use the Microsoft SQL Server connection to connect to a Microsoft SQL Server database from a UNIX or Linux machine.

Connecting to an Oracle Database

For native connectivity, install the version of Oracle client appropriate for the Oracle database server version. To ensure compatibility between Informatica and databases, use the appropriate database client libraries.

You must install compatible versions of the Oracle client and Oracle database server. You must also install the same version of the Oracle client on all machines that require it. To verify compatibility, contact Oracle.

Configuring Native Connectivity

You can configure native connectivity to an Oracle database to increase performance.

The following steps provide a guideline for configuring native connectivity through Oracle Net Services or Net8. For specific instructions, see the database documentation.

1. To configure connectivity for the Data Integration Service, PowerCenter Integration Service, or PowerCenter Repository Service process, log in to the machine as a user who can start the server process.
2. Set the ORACLE_HOME, NLS_LANG, TNS_ADMIN, and PATH environment variables.

   ORACLE_HOME. Set the variable to the Oracle client installation directory. For example, if the client is installed in the /HOME2/oracle directory. set the variable as follows:

   Using a Bourne shell:

   ```
   $ ORACLE_HOME=/HOME2/oracle; export ORACLE_HOME
   ```

   Using a C shell:

   ```
   $ setenv ORACLE_HOME /HOME2/oracle
   ```

   NLS_LANG. Set the variable to the locale (language, territory, and character set) you want the database client and server to use with the login. The value of this variable depends on the configuration. For example, if the value is american_america.UTF8, set the variable as follows:

   Using a Bourne shell:

   ```
   $ NLS_LANG=american_america.UTF8; export NLS_LANG
   ```

   Using a C shell:

   ```
   $ NLS_LANG american_america.UTF8
   ```

   To determine the value of this variable, contact the administrator.

   TNS_ADMIN. If the tnsnames.ora file is not in the same location as the Oracle client installation location, set the TNS_ADMIN environment variable to the directory where the tnsnames.ora file resides. For example, if the file is in the /HOME2/oracle/files directory, set the variable as follows:

   Using a Bourne shell:

   ```
   $ TNS_ADMIN=$HOME2/oracle/files; export TNS_ADMIN
   ```

   Using a C shell:

   ```
   $ setenv TNS_ADMIN=$HOME2/oracle/files
   ```

   Note: By default, the `tnsnames.ora` file is stored in the following directory: `$ORACLE_HOME/network/admin`.

   PATH. To run the Oracle command line programs, set the variable to include the Oracle bin directory.

   Using a Bourne shell:

   ```
   $ PATH=${PATH}:$ORACLE_HOME/bin; export PATH
   ```

   Using a C shell:

   ```
   $ setenv PATH ${PATH}:ORACLE_HOME/bin
   ```

3. Set the shared library environment variable.

   The Oracle client software contains a number of shared library components that the Data Integration Service, PowerCenter Integration Service, and PowerCenter Repository Service processes load dynamically. To locate the shared libraries during run time, set the shared library environment variable.

   The shared library path must also include the Informatica installation directory `(server_dir)`.

   Set the shared library environment variable to LD_LIBRARY_PATH.

For example, use the following syntax:

- Using a Bourne shell:

```
$ LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:$HOME/server_dir:$ORACLE_HOME/lib; export
LD_LIBRARY_PATH
```

- Using a C shell:

```
$ setenv LD_LIBRARY_PATH ${LD_LIBRARY_PATH}:$HOME/server_dir:$ORACLE_HOME/lib
```

4. Edit the .cshrc or .profile to include the complete set of shell commands. Save the file and either log out and log in again, or run the source command.

   Using a Bourne shell:

   ```
   $ source .profile
   ```

   Using a C shell:

   ```
   $ source .cshrc
   ```

5. Verify that the Oracle client is configured to access the database.

   Use the SQL*Net Easy Configuration Utility or copy an existing `tnsnames.ora` file to the home directory and modify it.

   The `tnsnames.ora` file is stored in the following directory: `$ORACLE_HOME/network/admin`.

   Enter the correct syntax for the Oracle connect string, typically `databasename.world`.

   Here is a sample `tnsnames.ora` file. Enter the information for the database.

   ```
   mydatabase.world =
      (DESCRIPTION
      (ADDRESS_LIST =
           (ADDRESS =
              (COMMUNITY = mycompany.world
              (PROTOCOL = TCP)
              (Host = mymachine)
              (Port = 1521)
           )
      )
      (CONNECT_DATA =
         (SID = MYORA7)
         (GLOBAL_NAMES = mydatabase.world)
   ```

   Here is a sample `tnsnames.ora` file to connect to Oracle using Oracle Connection Manager:

   ```
   ORCL19C_CMAN =
   (description=
   (address_list=
   (source_route=yes)
   (address=(protocol=tcp)(host=inrh74ocm.mycompany.com)(port=1521))
   (address=(protocol=tcp)(host=inrh74oradb.mycompany.com)(port=1521))
   )
   (connect_data=
   (service_name=ORCL19C.mycompany.com)
   )
   )
   ```

6. Verify that you can connect to the Oracle database.

   To connect to the Oracle database, launch SQL*Plus and enter the connectivity information. If you fail to connect to the database, verify that you correctly entered all of the connectivity information.

   Enter the user name and connect string as defined in the `tnsnames.ora` file.

Connecting to a Teradata Database

Install and configure native client software on the machines where the Data Integration Service or PowerCenter Integration Service process runs. To ensure compatibility between Informatica and databases, use the appropriate database client libraries.

Install the Teradata client, the Teradata ODBC driver, and any other Teradata client software that you might need on the machine where the Data Integration Service or PowerCenter Integration Service runs. You must also configure ODBC connectivity.

Note: Based on a recommendation from Teradata, Informatica uses ODBC to connect to Teradata. ODBC is a native interface for Teradata.

Configuring ODBC Connectivity

You can configure ODBC connectivity to a Teradata database.

The following steps provide a guideline for configuring ODBC connectivity. For specific instructions, see the database documentation.

1. To configure connectivity for the integration service process, log in to the machine as a user who can start a service process.
2. Set the TERADATA_HOME, ODBCHOME, and PATH environment variables.

   TERADATA_HOME. Set the variable to the Teradata driver installation directory. The defaults are as follows:

   Using a Bourne shell:

   ```
   $ TERADATA_HOME=/opt/teradata/client/<version>; export TERADATA_HOME
   ```

   Using a C shell:

   ```
   $ setenv TERADATA_HOME /opt/teradata/client/<version>
   ```

   ODBCHOME. Set the variable to the ODBC installation directory. For example:

   Using a Bourne shell:

   ```
   $ ODBCHOME=$INFA_HOME/ODBC<version>; export ODBCHOME
   ```

   Using a C shell:

   ```
   $ setenv ODBCHOME $INFA_HOME/ODBC<version>
   ```

   PATH. To run the *ddtestlib* utility, to verify that the DataDirect ODBC driver manager can load the driver files, set the variable as follows:

   Using a Bourne shell:

   ```
   PATH="${PATH}:$ODBCHOME/bin:$TERADATA_HOME/bin"
   ```

   Using a C shell:

   ```
   $ setenv PATH ${PATH}:$ODBCHOME/bin:$TERADATA_HOME/bin
   ```
3. Set the shared library environment variable.

   The Teradata software contains multiple shared library components that the integration service process loads dynamically. Set the shared library environment variable so that the services can find the shared libraries at run time.

   The shared library path must also include installation directory of the Informatica service *(server_dir)*.

   Set the shared library environment variable based on the operating system.

The following table describes the shared library variables for each operating system:

Operating System	Variable
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

For example, use the following syntax for Linux:

- Using a Bourne shell:

```
$ LD_LIBRARY_PATH="${LD_LIBRARY_PATH}:$HOME/server_dir:$ODBCHOME/lib:

$TERADATA_HOME/lib64:$TERADATA_HOME/odbc_64/lib";

export LD_LIBRARY_PATH
```

- Using a C shell:

```
$ setenv LD_LIBRARY_PATH "${LD_LIBRARY_PATH}:$HOME/server_dir:$ODBCHOME/
lib:$TERADATA_HOME/lib64:

$TERADATA_HOME/odbc_64/lib"
```

For AIX

- Using a Bourne shell:

```
$ LIBPATH=${LIBPATH}:$HOME/server_dir:$ODBCHOME/lib:$TERADATA_HOME/
lib64:$TERADATA_HOME/odbc_64/lib; export LIBPATH
```

- Using a C shell:

```
$ setenv LIBPATH ${LIBPATH}:$HOME/server_dir:$ODBCHOME/lib:$TERADATA_HOME/lib64:

$TERADATA_HOME/odbc_64/lib
```

4. Edit the existing odbc.ini file or copy the odbc.ini file to the home directory and edit it.

   This file exists in $ODBCHOME directory.

```
$ cp $ODBCHOME/odbc.ini $HOME/.odbc.ini
```

   Add an entry for the Teradata data source under the section [ODBC Data Sources] and configure the data source.

   For example, for Teradata Parallel Transporter utilities, version 15.10:

```
MY_TERADATA_SOURCE=Teradata Driver
[MY_TERADATA_SOURCE]
Driver=/opt/teradata/client/15.10/lib64/tdata.so
Description=NCR 3600 running Teradata V1R5.2
DBCName=208.199.59.208
DateTimeFormat=AAA
SessionMode=ANSI
DefaultDatabase=
Username=
Password=
```

   For example, for Teradata Parallel Transporter utilities, version 16.20:

```
MY_TERADATA_SOURCE=Teradata Driver
[dwtera]
Driver=/opt/teradata/client/16.20/lib64/tdataodbc_sb64.so
Description=NCR 3600 running Teradata V1R5.2
DBCName=tdvbe1510
LastUser=
Username=
Password=
Database=
```

```
DefaultDatabase=
UseNativeLOBSupport=Yes
CharacterSet=UTF8
SessionMode=ANSI
```

5. Set the DateTimeFormat to AAA in the Teradata data ODBC configuration.
6. Optionally, set the SessionMode to ANSI. When you use ANSI session mode, Teradata does not roll back the transaction when it encounters a row error.

   If you choose Teradata session mode, Teradata rolls back the transaction when it encounters a row error. In Teradata mode, the integration service process cannot detect the rollback, and does not report this in the session log.
7. To configure connection to a single Teradata database, enter the DefaultDatabase name. To create a single connection to the default database, enter the user name and password. To connect to multiple databases, using the same ODBC DSN, leave the DefaultDatabase field empty.

   For more information about Teradata connectivity, see the Teradata ODBC driver documentation.
8. Verify that the last entry in the odbc.ini is InstallDir and set it to the odbc installation directory.

   For example:

   ```
   InstallDir=<Informatica installation directory>/ODBC<version>
   ```
9. Edit the .cshrc or .profile to include the complete set of shell commands.
10. Save the file and either log out and log in again, or run the source command.

    Using a Bourne shell:

    ```
    $ source .profile
    ```

    Using a C shell:

    ```
    $ source .cshrc
    ```
11. For each data source you use, make a note of the file name under the Driver=<parameter> in the data source entry in odbc.ini. Use the *ddtestlib* utility to verify that the DataDirect ODBC driver manager can load the driver file.

    For example, if you have the driver entry:

    ```
    Driver=/u01/app/teradata/td-tuf611/odbc/drivers/tdata.so
    ```

    run the following command:

    ```
    ddtestlib /u01/app/teradata/td-tuf611/odbc/drivers/tdata.so
    ```
12. Test the connection using BTEQ or another Teradata client tool.

Connecting to a JDBC Data Source

To enable the the Data Integration Service to write to relational targets, download JDBC driver .jar files to the Data Integration Service host and to all client machines that run mappings that have relational targets.

Obtain the driver .jar file from the database vendor. For example, to access an Oracle database, download the file `ojdbc.jar` from the Oracle website.

1. Place the JDBC driver .jar file in the following directory on the Data Integration Service machine `<Informatica installation directory>/externaljdbcjars`. Then recycle the Data Integration Service.
2. Place the JDBC driver .jar file in the following directory on machines that host the Developer tool: `<Informatica installation directory>/clients/externaljdbcjars`. Then recycle the Developer tool.

Connecting to an ODBC Data Source

Install and configure native client software on the machine where the Data Integration Service, PowerCenter Integration Service, and PowerCenter Repository Service run. Also install and configure any underlying client access software required by the ODBC driver. To ensure compatibility between Informatica and the databases, use the appropriate database client libraries.

The Informatica installation includes DataDirect ODBC drivers. If the odbc.ini file contains connections that use earlier versions of the ODBC driver, update the connection information to use the new drivers. Use the System DSN to specify an ODBC data source on Windows.

1. On the machine where the application service runs, log in as a user who can start a service process.
2. Set the ODBCHOME and PATH environment variables.

   ODBCHOME. Set to the DataDirect ODBC installation directory. For example, if the install directory is `/export/home/Informatica/10.0.0/ODBC7.1`.

   Using a Bourne shell:

   ```
   $ ODBCHOME=/export/home/Informatica/10.0.0/ODBC7.1; export ODBCHOME
   ```

   Using a C shell:

   ```
   $ setenv ODBCHOME /export/home/Informatica/10.0.0/ODBC7.1
   ```

   PATH. To run the ODBC command line programs, like *ddtestlib*, set the variable to include the odbc bin directory.

   Using a Bourne shell:

   ```
   $ PATH=${PATH}:$ODBCHOME/bin; export PATH
   ```

   Using a C shell:

   ```
   $ setenv PATH ${PATH}:$ODBCHOME/bin
   ```

   Run the *ddtestlib* utility to verify that the DataDirect ODBC driver manager can load the driver files.
3. Set the shared library environment variable.

   The ODBC software contains a number of shared library components that the service processes load dynamically. Set the shared library environment variable so that the services can find the shared libraries at run time.

   The shared library path must also include the Informatica installation directory *(server_dir)*.

   Set the shared library environment variable based on the operating system.

   The following table describes the shared library variables for each operating system:

Operating System	Variable
Linux	LD_LIBRARY_PATH
AIX	LIBPATH

   For example, use the following syntax for Linux:

   - Using a Bourne shell:

     ```
     $ LD_LIBRARY_PATH=${LD_LIBRARY_PATH}:$HOME/server_dir:$ODBCHOME/lib; export
     LD_LIBRARY_PATH
     ```

   - Using a C shell:

     ```
     $ setenv LD_LIBRARY_PATH $HOME/server_dir:$ODBCHOME:${LD_LIBRARY_PATH}
     ```

For AIX

- Using a Bourne shell:

```
$ LIBPATH=${LIBPATH}:$HOME/server_dir:$ODBCHOME/lib; export LIBPATH
```

- Using a C shell:

```
$ setenv LIBPATH ${LIBPATH}:$HOME/server_dir:$ODBCHOME/lib
```

4. Edit the existing odbc.ini file or copy the odbc.ini file to the home directory and edit it.

   This file exists in $ODBCHOME directory.

   ```
   $ cp $ODBCHOME/odbc.ini $HOME/.odbc.ini
   ```

   Add an entry for the ODBC data source under the section [ODBC Data Sources] and configure the data source.

   For example:

   ```
   MY_MSSQLSERVER_ODBC_SOURCE=<Driver name or data source description>
   [MY_SQLSERVER_ODBC_SOURCE]
   Driver=<path to ODBC drivers>
   Description=DataDirect 8.0 SQL Server Wire Protocol
   Database=<SQLServer_database_name>
   LogonID=<username>
   Password=<password>
   Address=<TCP/IP address>,<port number>
   QuoteId=No
   AnsiNPW=No
   ApplicationsUsingThreads=1
   ```

   This file might already exist if you have configured one or more ODBC data sources.

5. Verify that the last entry in the odbc.ini is InstallDir and set it to the odbc installation directory.

   For example:

   ```
   InstallDir=/export/home/Informatica/10.0.0/ODBC7.1
   ```

6. If you use the odbc.ini file in the home directory, set the ODBCINI environment variable.

   Using a Bourne shell:

   ```
   $ ODBCINI=/$HOME/.odbc.ini; export ODBCINI
   ```

   Using a C shell:

   ```
   $ setenv ODBCINI $HOME/.odbc.ini
   ```

7. Edit the .cshrc or .profile to include the complete set of shell commands. Save the file and either log out and log in again, or run the source command.

   Using a Bourne shell:

   ```
   $ source .profile
   ```

   Using a C shell:

   ```
   $ source .cshrc
   ```

8. Use the *ddtestlib* utility to verify that the DataDirect ODBC driver manager can load the driver file you specified for the data source in the odbc.ini file.

   For example, if you have the driver entry:

   ```
   Driver = /export/home/Informatica/10.0.0/ODBC7.1/lib/DWxxxxnn.so
   ```

   run the following command:

   ```
   ddtestlib /export/home/Informatica/10.0.0/ODBC7.1/lib/DWxxxxnn.so
   ```

9. Install and configure any underlying client access software needed by the ODBC driver.

   Note: While some ODBC drivers are self-contained and have all information inside the .odbc.ini file, most are not. For example, if you want to use an ODBC driver to access Sybase IQ, you must install the Sybase IQ network client software and set the appropriate environment variables.

   To use the Informatica ODBC drivers (DWxxxxnn.so), manually set the PATH and shared library path environment variables. Alternatively, run the odbc.sh or odbc.csh script in the $ODBCHOME folder. This script will set the required PATH and shared library path environment variables for the ODBC drivers provided by Informatica.

Sample odbc.ini File

The following sample shows the entries for the ODBC drivers in the ODBC.ini file:

```
[ODBC Data Sources]
SQL Server Legacy Wire Protocol=DataDirect 7.1 SQL Server Legacy Wire Protocol
DB2 Wire Protocol=DataDirect 7.1 DB2 Wire Protocol
Informix Wire Protocol=DataDirect 7.1 Informix Wire Protocol
Oracle Wire Protocol=DataDirect 8.0 Oracle Wire Protocol
Sybase Wire Protocol=DataDirect 7.1 Sybase Wire Protocol
SQL Server Wire Protocol=DataDirect 8.0 SQL Server Wire Protocol
MySQL Wire Protocol=DataDirect 7.1 MySQL Wire Protocol
PostgreSQL Wire Protocol=DataDirect 7.1 PostgreSQL Wire Protocol
Greenplum Wire Protocol=DataDirect 7.1 Greenplum Wire Protocol

[ODBC]
IANAAppCodePage=4
InstallDir=/<Informatica installation directory>/ODBC7.1
Trace=0
TraceFile=odbctrace.out
TraceDll=/<Informatica installation directory>/ODBC7.1/lib/DWtrc27.so

[DB2 Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWdb227.so
Description=DataDirect 7.1 DB2 Wire Protocol
AccountingInfo=
AddStringToCreateTable=
AlternateID=
AlternateServers=
ApplicationName=
ApplicationUsingThreads=1
AuthenticationMethod=0
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
CatalogSchema=
CharsetFor65535=0
ClientHostName=
ClientUser=
#Collection applies to z/OS and iSeries only
Collection=
ConcurrentAccessResolution=0
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
CurrentFuncPath=
#Database applies to DB2 UDB only
Database=<database_name>
DefaultIsolationLevel=1
DynamicSections=1000
EnableBulkLoad=0
```

```
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
GrantAuthid=PUBLIC
GrantExecute=1
GSSClient=native
HostNameInCertificate=
IpAddress=<DB2_server_host>
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
#Location applies to z/OS and iSeries only
Location=<location_name>
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=
PackageCollection=NULLID
PackageNamePrefix=DD
PackageOwner=
Pooling=0
ProgramID=
QueryTimeout=0
ReportCodePageConversionErrors=0
TcpPort=50000
TrustStore=
TrustStorePassword=
UseCurrentSchema=0
ValidateServerCertificate=1
WithHold=1
XMLDescribeType=-10

[Informix Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWifcl27.so
Description=DataDirect 7.1 Informix Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
CancelDetectInterval=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
HostName=<Informix_host>
LoadBalancing=0
LogonID=
Password=
PortNumber=<Informix_server_port>
ServerName=<Informix_server>
TrimBlankFromIndexName=1
UseDelimitedIdentifiers=0

[Oracle Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWora28.so
Description=DataDirect 8.0 Oracle Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
AccountingInfo=
Action=
ApplicationName=
ArraySize=60000
AuthenticationMethod=1
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
CachedCursorLimit=32
CachedDescLimit=0
CatalogIncludesSynonyms=1
```

```
CatalogOptions=0
ClientHostName=
ClientID=
ClientUser=
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
DataIntegrityLevel=0
DataIntegrityTypes=MD5,SHA1
DefaultLongDataBuffLen=1024
DescribeAtPrepare=0
EditionName=
EnableBulkLoad=0
EnableDescribeParam=0
EnableNcharSupport=0
EnableScrollableCursors=1
EnableStaticCursorsForLongData=0
EnableTimestampWithTimeZone=0
EncryptionLevel=0
EncryptionMethod=0
EncryptionTypes=AES128,AES192,AES256,DES,3DES112,3DES168,RC4_40,RC4_56,RC4_128,
RC4_256
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
GSSClient=native
HostName=<Oracle_server>
HostNameInCertificate=
InitializationString=
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LocalTimeZoneOffset=
LockTimeOut=-1
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Module=
Password=
Pooling=0
PortNumber=<Oracle_server_port>
ProcedureRetResults=0
ProgramID=
QueryTimeout=0
ReportCodePageConversionErrors=0
ReportRecycleBin=0
ServerName=<server_name in tnsnames.ora>
ServerType=0
ServiceName=
SID=<Oracle_System_Identifier>
TimestampeEscapeMapping=0
TNSNamesFile=<tnsnames.ora_filename>
TrustStore=
TrustStorePassword=
UseCurrentSchema=1
ValidateServerCertificate=1
WireProtocolMode=2

[Sybase Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWase27.so
Description=DataDirect 7.1 Sybase Wire Protocol
AlternateServers=
ApplicationName=
ApplicationUsingThreads=1
ArraySize=50
AuthenticationMethod=0
BulkBinaryThreshold=32
```

```
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadFieldDelimiter=
BulkLoadRecordDelimiter=
Charset=
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
CursorCacheSize=1
Database=<database_name>
DefaultLongDataBuffLen=1024
EnableBulkLoad=0
EnableDescribeParam=0
EnableQuotedIdentifiers=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
GSSClient=native
HostNameInCertificate=
InitializationString=
Language=
LoadBalancing=0
LoadBalanceTimeout=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
NetworkAddress=<Sybase_host,Sybase_server_port>
OptimizePrepare=1
PacketSize=0
Password=
Pooling=0
QueryTimeout=0
RaiseErrorPositionBehavior=0
ReportCodePageConversionErrors=0
SelectMethod=0
ServicePrincipalName=
TruncateTimeTypeFractions=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
WorkStationID=

[SQL Server Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWsqls28.so
Description=DataDirect 8.0 SQL Server Wire Protocol
AlternateServers=
AlwaysReportTriggerResults=0
AnsiNPW=1
ApplicationName=
ApplicationUsingThreads=1
AuthenticationMethod=1
BulkBinaryThreshold=32
BulkCharacterThreshold=-1
BulkLoadBatchSize=1024
BulkLoadOptions=2
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
EnableBulkLoad=0
EnableQuotedIdentifiers=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=1
GSSClient=native
HostName=<SQL_Server_host>
```

```
HostNameInCertificate=
InitializationString=
Language=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
PacketSize=-1
Password=
Pooling=0
PortNumber=<SQL_Server_server_port>
QueryTimeout=0
ReportCodePageConversionErrors=0
SnapshotSerializable=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
WorkStationID=
XML Describe Type=-10

[MySQL Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWmysql27.so
Description=DataDirect 7.1 MySQL Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=1024
EnableDescribeParam=0
EncryptionMethod=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
HostName=<MySQL_host>
HostNameInCertificate=
InteractiveClient=0
LicenseNotice=You must purchase commercially licensed MySQL database software or
a MySQL Enterprise subscription in order to use the DataDirect Connect for ODBC
for MySQL Enterprise driver with MySQL software.
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LogonID=
LoginTimeout=15
MaxPoolSize=100
MinPoolSize=0
Password=
Pooling=0
PortNumber=<MySQL_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TreatBinaryAsChar=0
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1

[PostgreSQL Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWpsql27.so
Description=DataDirect 7.1 PostgreSQL Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=2048
```

```
EnableDescribeParam=1
EncryptionMethod=0
ExtendedColumnMetadata=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=0
HostName=<PostgreSQL_host>
HostNameInCertificate=
InitializationString=
KeyPassword=
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=
Pooling=0
PortNumber=<PostgreSQL_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TransactionErrorBehavior=1
TrustStore=
TrustStorePassword=
ValidateServerCertificate=1
XMLDescribeType=-10

[Greenplum Wire Protocol]
Driver=/<Informatica installation directory>/ODBC7.1/lib/DWgplm27.so
Description=DataDirect 7.1 Greenplum Wire Protocol
AlternateServers=
ApplicationUsingThreads=1
ConnectionReset=0
ConnectionRetryCount=0
ConnectionRetryDelay=3
Database=<database_name>
DefaultLongDataBuffLen=2048
EnableDescribeParam=0
EnableKeysetCursors=0
EncryptionMethod=0
ExtendedColumnMetadata=0
FailoverGranularity=0
FailoverMode=0
FailoverPreconnect=0
FetchTSWTZasTimestamp=0
FetchTWFSasTime=0
HostName=<Greenplum_host>
InitializationString=
KeyPassword=
KeysetCursorOptions=0
KeyStore=
KeyStorePassword=
LoadBalanceTimeout=0
LoadBalancing=0
LoginTimeout=15
LogonID=
MaxPoolSize=100
MinPoolSize=0
Password=
Pooling=0
PortNumber=<Greenplum_server_port>
QueryTimeout=0
ReportCodepageConversionErrors=0
TransactionErrorBehavior=1
XMLDescribeType=-10
```

Note: You might have to customize the DSN entries in the `ODBC.ini` file based on the third-party driver that you use. For more information about the DSN entries, see the corresponding third-party driver documentation.

INDEX

A

B

C

D

E

G

H

I

J

K

L

M

N

O

P

R

S

T

U

W