Oracle® Fusion Middleware

Installing and Configuring Oracle Enterprise Data Quality

12c (12.2.1.4.0)

E95106-07

April 2024

ORACLE®

Oracle Fusion Middleware Installing and Configuring Oracle Enterprise Data Quality, 12c (12.2.1.4.0)

E95106-07

Contents

Preface

1 Planning an Enterprise Data Quality Installation

2 Installing the Required External Software Components

3   Installing Enterprise Data Quality

4   Configuring Enterprise Data Quality with Oracle WebLogic Server

5   Configuring Enterprise Data Quality with Apache Tomcat

6   Setting Server Parameters to Support Enterprise Data Quality

7   Next Steps After Configuring Enterprise Data Quality

8   Upgrading Enterprise Data Quality

9   Removing Enterprise Data Quality from a System

A   Updating the JDK After Installing and Configuring an Oracle Fusion Middleware Product

Preface

This document describes how to install and configure Enterprise Data Quality on Linux, UNIX, and Windows platforms. It also provides instructions for configuring EDQ to work with an application server and database.

- Documentation Accessibility
- Audience
- Related Documents
- Conventions

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc`.

Accessible Access to Oracle Support

Oracle customers who have purchased support have access to electronic support through My Oracle Support. For information, visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info` or visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs` if you are hearing impaired.

Audience

This document is intended for system administrators or application developers who are installing Enterprise Data Quality. It is assumed that you have a basic understanding of application server and web technology and have a general understanding of Linux, UNIX, and Windows platforms. Throughout this guide, it is assumed that you are fully familiar with the components of the supported platform on which you want to install Enterprise Data Quality.

Related Documents

For more information about EDQ, see the following documents in the Enterprise Data Quality documentation set.

EDQ Documentation Library

The following publications are provided to help you install and use EDQ:

- *Release Notes for Enterprise Data Quality*
- *Installing and Configuring Enterprise Data Quality*

- *Administering Enterprise Data Quality*

- *Understanding Enterprise Data Quality*

- *Integrating Enterprise Data Quality With External Systems*

- *Securing Oracle Enterprise Data Quality*

- *Installation and Upgrade Guide*

- *Release Notes*

Find the latest version of these guides and all of the Oracle product documentation at

https://docs.oracle.com

Online Help

Online help is provided for all Enterprise Data Quality user applications. It is accessed in each application by pressing the **F1** key or by clicking the Help icons. The main nodes in the Director project browser have integrated links to help pages. To access them, either select a node and then press **F1**, or right-click on an object in the Project Browser and then select **Help**. The EDQ processors in the Director Tool Palette have integrated help topics, as well. To access them, right-click on a processor on the canvas and then select **Processor Help**, or left-click on a processor on the canvas or tool palette and then press **F1**.

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
`monospace`	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

1
Planning an Enterprise Data Quality Installation

This chapter helps to prepare you for your Enterprise Data Quality (EDQ) installation. Various topics are covered that should be reviewed thoroughly to help ensure that you do not encounter any problems either during or after the product installation and domain configuration.
This chapter includes the following sections:

- Overview of EDQ
- Overview of the Installation and Configuration Tasks
- Choosing EDQ Components and Versions
- Satisfying EDQ System Requirements
- Downloading EDQ

Overview of EDQ

EDQ provides a comprehensive data quality management environment that is used to understand, improve, protect and govern data quality. The software facilitates best practice primary data management, data integration, business intelligence, and data migration initiatives. It provides integrated data quality in customer relationship management (CRM) and other applications.

This documentation guides you through the selection, installation and configuration of the components that are needed to support EDQ.

Overview of the Installation and Configuration Tasks

This section (Table 1-1) provides an overview of the EDQ installation and configuration tasks that you will perform, in the order that they should be performed.

Table 1-1   EDQ Product Installation Procedure Tasks

Task	Action to Perform
Understand and select the external software components that support EDQ.	See Choosing EDQ Components and Versions
Satisfy EDQ system requirements.	See Satisfying EDQ System Requirements
Obtain an EDQ installation file from Oracle Software Delivery Cloud.	See Downloading EDQ
Install the JDK and your chosen application server and database components.	See Installing the Required External Software Components
Install the EDQ software.	See Installing Enterprise Data Quality

Table 1-1   (Cont.) EDQ Product Installation Procedure Tasks

Task	Action to Perform
Configure EDQ	For instructions see, Configuring Enterprise Data Quality with Oracle WebLogic Server or Configuring Enterprise Data Quality with Apache Tomcat
Set system parameters.	See Setting Server Parameters to Support Enterprise Data Quality
Next Steps	For log in and basic use information, see Next Steps After Configuring Enterprise Data Quality
(Optional) Upgrade from a previous release of EDQ.	If using Oracle WebLogic Server or Apache Tomcat, see Upgrading Enterprise Data Quality

Choosing EDQ Components and Versions

The following sections show you the components that are required to support EDQ and the supported versions of those components and EDQ.

- Choosing the Correct Combination of EDQ Required Components

- Supported Platforms and Component Versions

Choosing the Correct Combination of EDQ Required Components

EDQ is a Java Web Application that uses a Java Servlet Engine, a Java Web Start graphical user interface, and a data repository within a database. As such, it requires access to the following components:

- a Java Development Kit (JDK)

- a Java Application Server to supply web services. Oracle WebLogic Server and Apache Tomcat are supported.

- a structured query language (SQL) relational database management system (RDBMS) to store configuration data, working data, and the results of work performed by the processes. Oracle Database is supported.

The following application servers and databases are supported for use with EDQ:

- Oracle WebLogic 12.2.1 (from the Fusion Middleware Infrastructure 12.2.1 package), or Apache Tomcat 8.

- Oracle Database 11.2.0.4+ or 12.1.0.1+.

Instructions for installing these components are in Installing the Required External Software Components See Supported Platforms and Component Versions for supported versions of each of these components.

Supported Platforms and Component Versions

Review the list of certified platforms, JDKs, databases, application servers, and releases for EDQ prior to installation. This information is listed in *Enterprise Data Quality Certification Matrix* at

http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

Satisfying EDQ System Requirements

This section describes the hardware and software requirements of EDQ. These requirements represent the server configurations that are certified and supported by Oracle for the EDQ product.

- Disk and Memory Requirements
- EDQ Directory Requirements
- UNIX System Resource Requirements (ulimit)
- Virtual Hardware
- Operating System User

Disk and Memory Requirements

Depending on the tasks that EDQ is required to perform, it can place heavy demands on the hardware used to run it. A recommended *minimum* hardware specification for an EDQ server is:

- 8GB physical memory, with 4GB allocated to the EDQ Java Virtual Machine (JVM)
- At least 4 logical CPUs
- At least 500GB of hard disk space on the database server. The EDQ results schema, which contains the working data that EDQ generates, must have enough space to contain at least 20 times the volume of source data that you expect to process through EDQ. This size may increase if there are many EDQ users working on the same projects at the same time. The configuration schema remains small, normally less than 5GB, unless there is a large amount of user-modified reference data and Case Management data.

> **Note:**
>
> These recommendations do not represent sizing advice for any specific deployment, but rather a starting point for testing size requirements in your environment. It may be appropriate to deploy a larger machine or many machines, depending on the processing load placed on EDQ.

EDQ Directory Requirements

EDQ uses an installation directory and two configuration directories. You should record the location of these directories in case you need to apply manual updates to any of their contents.

EDQ Installation Directory

During the installation process, you must specify an installation directory to contain the EDQ installation files. This directory is known as the EDQ Home (*EDQ_HOME*) directory and is named as follows:

- If you are installing EDQ as part of the Oracle Fusion Middleware product stack, you must install the EDQ Home directory as a subdirectory of the Oracle Fusion Middleware home (installation) directory. A typical default Fusion Middleware home directory is as follows, depending on the platform:

  Linux and UNIX:

  ```
  /opt/Oracle/Middleware/FMW_HOME
  ```

  Windows:

  ```
  C:\Oracle\Middleware\FMW_HOME
  ```

  > ✎ **Note:**
  >
  > The `Middleware` home directory is referenced as *FMW_HOME* in this guide.

- If you use Apache Tomcat application server, you can install the EDQ Home directory in any directory.

The EDQ Home directory requires approximately 1GB of hard disk space.

EDQ Configuration Directories

EDQ requires two configuration directories, which are separate from the EDQ Home (installation) directory that contains the program files. The configuration directories are:

- The *base* configuration directory: This directory contains default configuration data. Once EDQ is installed, the files in the base configuration directory must not be altered, renamed, or moved.

- The *local* configuration directory: This directory contains overrides to the default configuration. EDQ looks for overrides in this directory first, before looking in the base configuration directory. Files in the local configuration directory can be modified to customize or extend EDQ.

The names and locations of the configuration directories are as follows:

- If you are using Oracle WebLogic Server, the Configuration Wizard automatically creates and populates the configuration directories in the EDQ domain with the names of `oedq.home` (base configuration directory) and `oedq.local.home` (local configuration directory). An example installation path is:

  ```
  WLS_HOME/user_projects/domains/edq/config/fmwconfig/edq/oedq.home
  WLS_HOME/user_projects/domains/edq/config/fmwconfig/edq/oedq.local.home
  ```

- If you are using Apache Tomcat, you create the configuration directories manually in any location, with any names, and the configuration utility will populate them. You are prompted to create the directories during the installation instructions.

On a default EDQ installation, the configuration directories occupy approximately 1MB in total. When an EDQ instance uses the `landingarea`, this total can increase as files are loaded up for processing and written out.

UNIX System Resource Requirements (ulimit)

Depending on how the UNIX system that hosts the application server is configured, you may find that the application server cannot create files larger than 1 GB. This

restricts your ability to work with large data sets if you intend to use files to transfer data to EDQ for processing. The most common limit that has to be changed is the number of processes that the application server user (for example. 'weblogic') can create. On large servers, this often needs to be increased, for example to 65536.

System resource limits are controlled by the `ulimit` command. Default `ulimit` values exist, which you can view by using the `ulimit -a` command. Look at the settings for file size, process limit, and file handle limits. The hard `ulimit` on file size may need to be adjusted upward or removed for your application server account. Consult your System Administrator for assistance, if needed.

Virtual Hardware

You can install EDQ on virtualized systems using a virtualization tool, such as Oracle VM Server. Both the virtual system and the physical system must fulfill the minimum hardware requirements listed in this documentation.

If load balancing software is used to deploy multiple virtual systems onto a single physical system, care must be taken to ensure that the load balancing software is carefully tuned. In general, EDQ requires considerable compute power and memory in the Middle Tier (where it performs most of its processing), and considerable tablespace in the Database (where it writes data and results). Between batches, very little load is imposed on the system. When processing a batch of data, EDQ rapidly drives hardware to be CPU or I/O bound. Unless the virtualized load balancing is correctly configured, suboptimal performance results.

Operating System User

An operating system user account is used to install and upgrade EDQ on your servers and install the application server. This user is required on all platforms. This user *must* have full permissions (read, write and execute) to the EDQ installation (*EDQ_HOME*) directory, configuration directories, and all database directories. This user account is referred to as the *EDQ installation user* in this documentation. For more information about the EDQ directories, see EDQ Directory Requirements

> **Note:**
>
> When installing on UNIX or Linux operating systems, do *not* use the root user as your EDQ installation user account.

Downloading EDQ

To download the EDQ installation and configuration files, obtain the generic package installer from the Oracle Technology Network website as follows:

1. Enter the following URL into a web browser:

   http://www.oracle.com/technetwork/middleware/oedq/overview/index.html

2. Click **Sign-in/Register**.

3. Locate and select Oracle Enterprise Data Quality Media Pack that you want to download.

4. Click the Download button.

5. Browse to the directory where you want to save the file. Click **Save** to start the file download. A ZIP file is downloaded.

6. Extract the ZIP file to a temporary directory.

2

Installing the Required External Software Components

This chapter describes how to install the software components that support EDQ.

> **Note:**
>
> These software components must be installed prior to installing EDQ.

This chapter includes the following sections:

- Prerequisites for these Procedures
- Installing a Java Development Kit to Support EDQ
- Installing an Application Server to Support EDQ
- Installing a Database to Support EDQ

Prerequisites for these Procedures

Before performing the procedures in this section, you must first read and satisfy the requirements in Planning an Enterprise Data Quality Installation

Installing a Java Development Kit to Support EDQ

EDQ and the application server both rely on the Java Development Kit (JDK). The JDK provides a Java run-time environment (JRE) and tools for compiling and debugging Java applications.

Use the following table to identify the appropriate JDK to install on the platform where you are installing EDQ.

Platform	JDK to Use
AIX	IBM JDK. This is the only supported JDK on AIX.
HP-UX	HP JDK
Solaris	Oracle JDK
Linux	Oracle JDK
Windows	Oracle JDK

For supported JDK versions per platform, see the *Enterprise Data Quality Certification Matrix* at

http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

Download and install the appropriate JDK using the instructions provided at

http://www.oracle.com/technetwork/java/javase/downloads/index.html

> **Note:**
>
> On Solaris systems, you must install both the 32-bit and 64-bit JDKs in order to run Java applications.

Make a note of the directory into which you installed the JDK. You will need to specify this directory during the installation of the application server. The path to this directory is referred to as the *JDK_HOME* directory in this documentation. For example, the JDK_HOME may be like one of the following:

Linux and UNIX:

/opt/jdk1.8

Windows:

C:\Program Files\Java\jdk1.8

> **Note:**
>
> If your JDK_HOME directory contains spaces, as in the default installation directory of C:\Program Files\Java\jdk1.8, you must set your JAVA_HOME environment variable to the Windows short name for the directory, as follows: C:\Progra~1\Java\jdk1.8.

Installing an Application Server to Support EDQ

For maximum security, you should configure the application server (whether Oracle WebLogic Server or Apache Tomcat) to use HTTPS only. You should disable unencrypted HTTP access. You can alternatively use a frontend such as Oracle HTTP Server or Apache HTTP Server, where HTTPS is provided.

Choose and download one of the following application servers:

- Oracle Fusion Middleware Infrastrucuture (includes WebLogic Server)
- Apache Tomcat

> **Note:**
>
> The application server that you choose determines the database that you can use. With Oracle WebLogic Server, you must use Oracle Database, and you must install EDQ within an Oracle Fusion Middleware environment. With Apache Tomcat, you can use either Oracle Database or PostgreSQL Database, and you can install EDQ as a standalone application. PostgreSQL is not supported for production use.

- Installing Oracle Fusion Middleware Infrastructure (includes Oracle WebLogic Server)
- Installing Apache Tomcat

Installing Oracle Fusion Middleware Infrastructure (includes Oracle WebLogic Server)

When installing on Oracle WebLogic Server, install both Oracle WebLogic Server and JRF. To view the WebLogic Server versions that are supported by EDQ in this release, see Enterprise Data Quality Certification Matrix at

http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

To download and install Oracle WebLogic Server and JRF, see Installing and Configuring Oracle WebLogic Server and Coherence.

> **Note:**
>
> Use the Fusion Middleware Infrastructure installer. Do not use the free WebLogic Server distribution for developers, because this installer does not contain all of the required files for EDQ.

The directory path to your WebLogic Server installation is referenced as the *WLS_HOME* directory in this documentation.

Installing Apache Tomcat

To view the Tomcat versions that are supported by EDQ in this release, see the Enterprise Data Quality Certification Matrix at

http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

To download and install Apache Tomcat, go to the Apache Software Foundation Server website at

http://tomcat.apache.org

After you install Tomcat, perform the following tasks:

1. To configure Tomcat to use the Java Development Kit (JDK) that you installed in Installing a Java Development Kit to Support EDQ, add the JDK path to the local setenv.sh file.

ORACLE®

2. (Recommended) Configure Tomcat to start as a service.

3. Create an operating system user who will run Tomcat. This user must own the EDQ installation and configuration directories. This account is used to administer your EDQ domain and to log into the EDQ application. For more information about the EDQ directories, see EDQ Directory Requirements

Installing a Database to Support EDQ

This section contains information that is specific to the installation and configuration of the database that will contain the EDQ repository.

- If using Oracle WebLogic Server as the EDQ application server, you must install Oracle Database.

- If using Apache Tomcat as the application server, the supported database for production is Oracle Database. PostgreSQL database is supported only for development or test purposes.

To view the supported database versions, see

http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

> **Note:**
>
> You will create three schemas and three users in this database when configuring EDQ in Configuring Enterprise Data Quality with Oracle WebLogic Server or Configuring Enterprise Data Quality with Apache Tomcat (depending on the application server you chose).

- Installing an Oracle Database to Support EDQ
- Installing a PostgreSQL Database to Support EDQ

Installing an Oracle Database to Support EDQ

You can download a supported Oracle Database product and installation instructions from the Oracle Database website at

http://www.oracle.com

- Configuring Oracle Database to Support EDQ

Configuring Oracle Database to Support EDQ

The following configuration elements are either required or recommended when installing or configuring an Oracle database for use with EDQ.

Required:

- Select the **Create and configure a database** installation option.

- Configure the database to use a Unicode character set to ensure that EDQ is able to capture and process data in the widest range of character sets. For more

information, see Supporting Multilingual Databases with Unicode in *Oracle Database Globalization Support Guide*.

- EDQ requires a database administrator (DBA) user account in the database. This database account is used to access the database during the installation and configuration processes to create database accounts and objects that are specific to EDQ.

Recommended:

- Oracle recommends making the following configuration selections when specifying the Oracle memory structure and tablespace configuration to support EDQ:
  - 4GB Program Global Area (PGA)
  - 4GB System Global Area (SGA)
  - 20GB undo tablespace
  - 20GB temp tablespace
  - 120GB tablespace (4 x 30GB files) for Results database (`EDQRESULTS` schema)
  - 30GB (1 x 30GB file) for Config database (`EDQCONFIG` schema)

- You may need to increase the values for the `SESSIONS` and `PROCESSES` parameters. The suggested values are as follows, but these parameters may need to be adjusted later for optimal performance:

  ```
  SESSIONS=500
  ```

  ```
  PROCESSES=500
  ```

  For more information about setting these values appropriately, contact your database administrator.

Installing a PostgreSQL Database to Support EDQ

You can download the PostgreSQL product and installation instructions from the official PostgreSQL website at

http://www.postgresql.org

> **Note:**
>
> If PostgreSQL was distributed with your operating system, verify that the release is one of EDQ supported releases by going to the certification matrix at
>
> http://www.oracle.com/technetwork/middleware/ias/downloads/fusion-certification-100350.html

- Configuring PostgreSQL Database to Support EDQ
- Checking Local Database Connections
- Enabling Password Authentication on Linux Systems
- Configuring Remote Connections

Configuring PostgreSQL Database to Support EDQ

The following configuration elements are either required or recommended when installing or configuring a PostgreSQL database for use with EDQ.

Required:

- Allow 403 maximum connections by editing the `postgresql.conf` file in the PostgreSQL `data` directory (for example, `var/lib/pgsql/data/postgresql.conf`).

- Configure the PostgreSQL database to use a Unicode character set to ensure that EDQ is able to capture and process data in the widest range of character sets. For more information, click the **Documentation** link on the PostgreSQL website at

  http://www.postgresql.org

- To use a PostgreSQL Database with EDQ, you must enable database listening. See Checking Local Database Connections

- On Linux systems, you must configure PostgreSQL to use password authentication. See Enabling Password Authentication on Linux Systems

Recommended:

If you are installing on Windows, Oracle recommends that you use the graphical installer that you can download from the PostgreSQL website at

http://www.postgresql.org

Each operating system on the Download page provides a link to the graphical installer.

Checking Local Database Connections

This section describes how to ensure that the PostgreSQL database local connections are properly configured to ensure database listening is on and that passwords are accepted for authentication.

1. Log in to the system as the EDQ database user.

2. (Windows only) Locate the MS-DOS Command Prompt (`cmd.exe`), and then run it.

3. Ensure that your database server is running. For more information, see the PostgreSQL documentation for your installed version on the PostgreSQL website at

   http://www.postgresql.org

4. Go to the PostgreSQL `data` directory. For example, `/var/lib/pgsql/data` on Linux or `C:\Program Files\PostgreSQL\9.3\data` on Windows.

5. Open the `pg_hba.conf` file by entering the one of the following commands, depending on the platform:

   On the Linux command line:

   `tail pg_hba.conf`

   On the Windows command line:

   `TYPE pg_hba.conf`

   The file contents are displayed and the following is an excerpt:

```
# TYPE  DATABASE        USER            ADDRESS                 METHOD
# "local" is for Unix domain socket connnections only
local  all             all                                     ident

# IPv4 local connections:
host    all             all             127.0.0.1/32            md5
# IPv6 local connections:
host    all             all             ::1/128                 md5
```

6. Ensure that the IPv4 and IPv6 local connections are configured as shown in the console output in the previous step. Use the address for your own database in the `ADDRESS` column.

Enabling Password Authentication on Linux Systems

On Linux systems, you must configure PostgreSQL to use password authentication by editing the `pg_hba.conf` file in the `data` subdirectory of the PostgreSQL installation directory (for example, `/var/lib/pgsql/data/pg_hba.conf`). Locate the `ident sameuser` entries, and then change them to a value of `md5`.

Configuring Remote Connections

If your PostgreSQL database is installed on a different system from the one where EDQ is installed, it must be configured to accept connections from other hosts. This is applicable to all PostgreSQL environments.

1. Locate the `data` directory for your PostgreSQL database.

2. Edit the `postgresql.conf` file.

3. Locate the following line:

   ```
   #listen_addresses = 'localhost'
   ```

4. Insert the following line to cause PostgreSQL to accept connections from remote hosts:

   ```
   listen_addresses = '*'
   ```

5. Edit the `pg_hba.conf` file then add the following line using *addr/mask* to identify the subnet of the host running EDQ:

   ```
   host all all addr/mask md5
   ```

   For example, `host all all 192.168.0.0/24 md5` allows connections from all hosts with the IP addresses 192.168.0.0 to 192.168.0.255 while `host all all 0.0.23.56/32 md5` accepts connections only from IP address 10.0.23.56.

3

Installing Enterprise Data Quality

This chapter describes how to start the Enterprise Data Quality installation program in graphical mode. It also describes the sequence of screens that appear in the installation process.
This chapter includes the following sections:

- Starting the Installation Program
- Navigating the EDQ Installation Screens
- Verifying the Installation

Starting the Installation Program

The EDQ installation program is delivered as a generic Java Archive (JAR) file that is used to install all supported operating systems.

To start the installation program, perform the following steps.

1. Log in to the target system as your EDQ installation user. This user must be able to run the command prompt as an administrator.

2. Go to the directory where you downloaded and unzipped the installation program.

3. Launch the installation program by invoking `java -jar` from the JDK installation directory (*JDK_HOME*) directory on your system as follows:

   On UNIX operating systems:

   ```
   JDK_HOME/bin/java -jar edq_generic.jar
   ```

   On Windows operating systems:

   ```
   JDK_HOME\bin\java -jar edq_generic.jar
   ```

In the above commands for UNIX and Windows operating systems, edq_generic.jar is a generic place-holder for the actual name of the installation jar file. When the installation program appears, you are ready to begin the installation. See Navigating the EDQ Installation Screens for a description of each installation program screen.

Navigating the EDQ Installation Screens

The installation program displays a series of screens. If you need additional help with an installation screen, click **Help** on the installation screens during installation. Complete the installation using the instructions in Table 3-1.

Table 3-1   Running the Installation Program

Screen	Action to Perform
Installation Inventory Setup	On UNIX operating systems, this screen opens if this is the first time you are installing an Oracle product on this host. Specify the location where you want to create your central inventory. Make sure that the operating system group name selected on this screen has write permissions to the central inventory location. For more about the central inventory, see "Understanding the Oracle Central Inventory" in Installing Software with the Oracle Universal Installer. This screen does not appear on Windows operating systems.
Welcome	Click **Next** to proceed with the installation. You may cancel the installation at any time by clicking **Cancel**.
Auto Updates	Use this screen to search for the latest software updates, including important security updates, through your My Oracle Support account.
Installation Location	Specify the Middleware installation directory (FMW_HOME), which will contain EDQ once installed. See EDQ Directory Requirements Click **Next** to continue. You may click **View** to verify the installation location.
Installation Type	Use this screen to select the type of installation and consequently, the products and feature sets you want to install.
Prerequisite Checks	This screen verifies that your system meets the minimum necessary requirements. Click **Next** to continue. You may click **View Successful Tasks**, to view the list of tasks that gets verified. You may click **View Log**, to view log details.
Installation Summary	Use this screen to verify the installation options you selected. If you want to save these options to a response file, click **Save Response File** and provide the location and name of the response file. Response files can be used later in a silent installation situation. Click **Install** to continue.
Installation Progress	This screen allows you to see the progress of the installation. When the installation program progress reaches 100%, click Finish to dismiss the installer or click **Next** to see a summary.
Installation Complete	This screen appears when the installation is complete. Click **Finish** to exit the installation program.

Verifying the Installation

After you complete the installation, verify it by successfully completing a series of tasks.

- Reviewing the Installation Files
- Checking the Directory Service
- Viewing the Contents of the Oracle Home

Reviewing the Installation Files

Review the contents of the installation log files to make sure that no problems were encountered. By default, the installer writes logs files to the

Oracle_Inventory_Location/logs (on UNIX operating systems) or Oracle_Inventory_Location\logs (on Windows operating systems) directory.

For a description of the log files and where to find them, see "Installation Log Files" in Installing Software with the Oracle Universal Installer.

Checking the Directory Service

The contents of your installation vary based on the options you selected during the installation.

For more information about the directory structure you should see after installation, see "What are the Key Oracle Fusion Middleware Directories?" in Understanding Oracle Fusion Middleware.

Viewing the Contents of the Oracle Home

You can also view the contents of the Oracle home using the viewInventory script. For more information, see "Viewing the Contents of an Oracle Home" in Installing Software with the Oracle Universal Installer.

4

Configuring Enterprise Data Quality with Oracle WebLogic Server

This chapter describes how to create the EDQ database repository, create an Oracle WebLogic Server domain for EDQ, and start Oracle WebLogic Server.

> **Note:**
>
> These instructions apply to Oracle WebLogic Server environments only. If you are using Apache Tomcat, you must follow the directions in Configuring Enterprise Data Quality with Apache Tomcat

This chapter includes the following sections:

- Prerequisites for these Procedures
- Creating an EDQ Database Repository
- Creating the WebLogic Server EDQ Domain
- Start Oracle WebLogic Server
- Running Multiple EDQ Servers in the Same Domain
- Configuring EDQ for High-Availability in a WebLogic Server Cluster

Prerequisites for these Procedures

Before performing the procedures in this section, you must first read and satisfy the steps in:

- Planning an Enterprise Data Quality Installation
- Installing the Required External Software Components
- Installing Enterprise Data Quality

Creating an EDQ Database Repository

EDQ makes use of some database schemas. These schemas are the configuration schema (`EDQCONFIG`), the results schema (`EDQRESULTS`), and the staging schema (EDQSTAGING). You create them with the Oracle Repository Creation Utility (RCU).

The person who runs RCU must be able to log into the database with DBA privileges. If you cannot run with DBA privileges, RCU can create a script for a DBA to run later.

> **✎ Note:**
>
> Do not use RCU to upgrade EDQ; use the instructions in Upgrading
> Enterprise Data Quality

To run RCU:

1. Make certain the repository database is running.

2. Run the command shell or console of the operating system.

3. Start RCU from the `FMW_HOME/oracle_common/bin` directory, where `FMW_HOME` is the Oracle Fusion Middleware installation directory.

   On Linux:

   `./rcu`

   On Windows:

   `rcu.bat`

Complete the RCU configuration screens by following the instructions in Table 4-1.

Table 4-1   Running the RCU Program

Screen	Action to Perform
Welcome	Click **Next** to proceed with the installation.
	You may cancel the installation at any time by clicking **Cancel**.
Create Repository	Click **Next** to continue.
	This uses the default **Create Repository** and **System Load and Product Load** options. This requires the person running RCU to have DBA privileges.
Database Connection Details	Select **Oracle Database** from the Database Type list.
	Specify the host name where your Oracle database is running.
	Enter the port number for your database. The default port number for Oracle Database is 1521.
	Specify the service name for the database. Typically, the service name is the same as the global database name. For example, `orcl.example.com`. Note that if you are using release 12c of the Oracle Database, you need to ensure that you enter the connection details of a pluggable database. For example, `pdborcl.example.com`.
	Enter the user name for your database. The user name could either be `SYS` or that of the user with DBA privileges.
	Enter the password for your database user.
	Select **SYSDBA** from the Role list. This is automatically selected when the user is `SYS`. Select **Normal** from the Role list if you are a user with DBA privileges.
	Click **Next** to continue.
Checking Global Prerequisites	When the prerequisites checking progress has reached completion, click **OK** to continue.

Table 4-1   (Cont.) Running the RCU Program

Screen	Action to Perform
Select Components	Select **Create new prefix** and enter a unique prefix name for all the database schemas you are creating in this session. For example, `edqprod` or the default of `DEV`.
	Select the **Oracle AS Repository Components** check box. The Oracle EDQ check boxes that create the EDQ configuration, results, and staging schemas in the database repository are then automatically selected.
	When you select the **Oracle EDQ** component, all the necessary schemas are installed. Not only will the EDQ schema be selected but the common schema needed to support audit and OPSS, which EDQ is integrated with will also be selected. This ensure that EDQ will function correctly with these integrated components.
	The three EDQ schemas that are to be installed are Config, Results and Staging. The Staging schema is also used by CDS, with or without the Fusion Connector. However installing the Staging schema when not necessary will not interfere with operation and will consume minimal database resources.
	Click **Next** to continue.
Checking Component Prerequisites	When the prerequisites checking progress has reached completion, click **OK** to continue.
Schema Passwords	Ensure that **Use same passwords for all schemas** is selected.
	Enter the password that you want to use for all of the EDQ database schemas in all password fields, then click **Next** to continue.
Map Tablespaces	The default EDQ tablespaces that will be created by RCU are displayed by component.
	You can change the tablespaces by clicking **Manage Tablespaces** and then modifying the information. Oracle recommends using one user tablespace for the `EDQCONFIG` schema, and a different user tablespace for the `EDQRESULTS` schema. The recommended minimum sizes (can be adjusted later) are:
	• `EDQRESULTS`: 120GB tablespace (4 x 30GB files)
	• `EDQCONFIG`: 30GB tablespace (1 x 30GB file)
	• `EDQSTAGING`: 30GB tablespace (1 x 30GB file)
	Click **Next** to continue.
Validating and Creating Tablespaces	Click **OK** to create any non-existent tablespaces in your schema, then click **OK** when the operation completes.
Summary	Review the database details, then click **Create** to continue.
	A status screen is displayed that shows the progress of creating the repository components.
Completion Summary	Click **Close** to exit the RCU program.

Creating the WebLogic Server EDQ Domain

These instructions use the Configuration Wizard to create a Basic WebLogic Server domain for EDQ, with the following:

- One Administration Server and one managed server (no additional managed servers or clusters).

> **Note:**
>
> If there is the possibility you may add more managed servers in the future, you should choose a cluster deployment, even if it starts with a single server. Otherwise it will be necessary to manually re-target the Data Sources, Applications and Deployments in WebLogic from a single server to the cluster.

- One (non-RAC) data source for the EDQ configuration schema, one data source for the results schema, and one data source for the staging schema. You can convert the datasources to RAC data sources with the Configuration Wizard, or you can do so later through the WebLogic Server Administration Console.

- A Node Manager configuration that is predefined within the EDQ domain as `edq/nodemanager`. You cannot edit the Node Manager home in this configuration. You can change this configuration during this procedure, if desired.

> **Note:**
>
> Oracle recommends the use of managed servers that are administered by Oracle WebLogic Node Manager. You can configure Managed Servers, Clusters, and other advanced features through the Configuration Wizard, but it may be more practical to do so by using the WebLogic Server Administration Console after the initial configuration process. For more information, see Running Multiple EDQ Servers in the Same Domain

- Starting the WebLogic Server Domain Configuration Wizard
- Navigating the Domain Configuration Wizard Screens
- Configuring Launchpad to Show the Managed Server

Starting the WebLogic Server Domain Configuration Wizard

To start the Domain Configuration wizard, follow these steps. You will run the configuration wizard in graphical mode.

1. Log in to the system as the EDQ installation user that you created in Operating System User

2. Go to `FMW_HOME/oracle_common/common/bin` directory, where `FMW_HOME` is the Fusion Middleware installation directory.

3. Start the wizard by entering the following command:

   On Linux or UNIX operating systems:

   ```
   ./config.sh
   ```

   On Microsoft Windows operating systems:

   ```
   config.cmd
   ```

The WebLogic Server Configuration Wizard is displayed.

Navigating the Domain Configuration Wizard Screens

Table 4-2 describes the screens in the configuration wizard. Certain screens are displayed only in certain situations depending on your selections. For help with any screen, click the **Help** button.

Table 4-2   Configuration Screens for Creating a New EDQ WebLogic Server Domain

Screen	Action to Perform
Create Domain	Select **Create a new domain**.
	In the **Domain Location** box, enter the path to the new domain (for example, *FMW_HOME*/user_projects/domains/edq_domain) or click **Browse** to create the domain directory.
	Click **Next** to continue.
Templates	Select **Oracle Enterprise Data Quality – 12.2.1.4.0 [edq]**. The Oracle JRF and WebLogic Coherence Cluster Extension are automatically selected. Keep these selections.
	Click **Next** to continue.
Administrator Account	Specify the user name and password for the EDQ domain's administrator account. This account is used to administer the domain and to log into the EDQ application.
	Click **Next** to continue.
Application Location	Specify the directory in which the applications of the EDQ domain are to be stored.
	Click **Next** to continue.
Domain Mode and JDK	**Domain Mode:** Select the startup operation mode for your domain from the following options:
	• **Development Mode**—In this mode, the security configuration is relatively relaxed and you utilize boot.properties for the username and password, allowing you to auto-deploy applications.
	• **Production Mode**—In this mode, the security configuration is relatively stringent, requiring a username and password to deploy applications. Before putting a domain into production, familiarize yourself with procedures for securing a production environment. For more information, see Securing a Production Environment for Oracle WebLogic Server.
	JDK: From the Available JDKs list, select the JDK that you installed in Installing a Java Development Kit to Support EDQ
	Click **Next** to continue.
Database Configuration Type	Ensure that **RCU Data** is selected. This populates the connection information you supplied when you ran the Repository Creation Utility (see Table 4-1).
	If you must change any of these fields, ensure that you use the schema prefix (DEV by default) and password that you specified when you ran RCU.
	When done, click **Get RCU Configuration** to connect to the Oracle Database and bind the EDQ schemas.
	Click **Next** to continue.
Component Datasources	Accept the defaults and then click **Next**.

Table 4-2   (Cont.) Configuration Screens for Creating a New EDQ WebLogic Server Domain

Screen	Action to Perform
JDBC Test	All schemas are selected and automatically tested.
	Return to the previous screen to alter the connection configuration if necessary.
	Click **Next** to continue.
Advanced Configuration	Select the **Administration Server**, **Node Manager**, **Topology** (which includes **Server Templates, Managed Servers, Clusters, Virtual Targets and Coherence**) option.
Administration Server	On the Administration Server screen, change the listen address from "All Local Addresses" to the IP address of the host where the Administration Server will reside.
	Do not use All Local Addresses.
	Do not specify any server groups for the Administration Server.
Node Manager	Let the Per Domain Default Location option remain selected, and enter a Username and Password for the Node Manager.
Managed Servers	**Clone** the EDQ server to create a copy of the server.
	If you do not want to clone the EDQ server, select **Add** to add additional EDQ servers.
	Select the IP address of the host on which the Managed Server will reside.
	For configuration procedures for a clustered mode installation, see Configuring EDQ for High-Availability in a WebLogic Server Cluster
Clusters You must use this screen if EDQ is deployed in a cluster.	Select **Add** to add a cluster. If both servers are being deployed on the same machine, enter the machine name as the cluster address.
	For additional configuration procedures for a clustered mode installation and a description of those configuration screens, see Configuring EDQ for High-Availability in a WebLogic Server Cluster
	To deploy EDQ in a non-clustered mode, select the appropriate components as needed.
Assign Servers to Clusters You must use this screen if EDQ is deployed in a cluster.	Assign the managed servers that you created to the cluster that you created.
	For configuration procedures for a clustered mode installation, see Configuring EDQ for High-Availability in a WebLogic Server Cluster
Coherence Clusters You must use this screen if EDQ is deployed in a cluster.	Configure the Coherence cluster that is automatically added to the domain.
	Leave the default port number 7574 as the Coherence cluster listen port.

Table 4-2   (Cont.) Configuration Screens for Creating a New EDQ WebLogic Server Domain

Screen	Action to Perform
Machines	Create a new machine in the domain. A machine is required so that the Node Manager can start and stop servers.
	Click **Add** to create a new machine.
	Specify `edq_machine_1` in the Name field.
	In the Node Manager Listen Address field, select the IP address of the machine in which the Managed Servers are being configured. You must select a specific interface and not "localhost." This allows Coherence cluster addresses to be dynamically calculated.
	Verify the port in the Node Manager Listen Port field. The port number `5556`, shown in this example, may be referenced by other examples in the documentation. Replace this port number with your own port number as needed.
Assign Servers to Machines	Assign the Administration Server and Managed Servers to the new machine you just created.
	In the Machines pane, select the machine you want to assign the servers to; in this case, `edq_machine_1`.
	In the Servers pane, assign `AdminServer` to `edq_machine_1` by doing one of the following:
	• Click once on `AdminServer` to select it, then click on the right arrow to move it beneath the selected machine (`edq_machine_1`) in the Machines pane.
	• Double-click on `AdminServer` to move it beneath the selected machine (`edq_machine_1`) in the Machines pane.
	Repeat to assign both `edq_server_1` and `edq_server_2` to `edq_machine_1`.
Virtual Targets This screen is optional and you can use the screen, if required.	Add or delete Virtual Targets for use in a WebLogic Server Multitenant (MT) environment.
	Virtual targets define the locations where a partition or resource group runs. They also provide instructions on how to route traffic, including addresses, protocol settings, and targeting. Request routing is determined by the host name and optional URI.
	WebLogic Server Multitenant virtual targets are deprecated in WebLogic Server *12.2.1.4.0* and will be removed in the next release.
Partitions	For each partition you want to add, click **Add** and enter the name you want to use for the partition. The partition name must be unique within the domain.
	WebLogic Server Multitenant domain partitions are deprecated in WebLogic Server *12.2.1.4.0* and will be removed in the next release.
Configuration Summary	Review the configuration for your domain by selecting a view and then selecting individual items in the list for that view.
	If the domain is configured as you want it, click **Create** to create the domain.
	If you need to make changes to the configuration, click **Back** to return to the appropriate screen for the settings you want to change, or click on the links on the left to go to that screen.
Configuration Progress	Shows the progress of the domain creation.
	When the process completes, click **Next**.

ORACLE

Table 4-2   (Cont.) Configuration Screens for Creating a New EDQ WebLogic Server Domain

Screen	Action to Perform
End of Configuration	Review the domain creation results. Click **Finish** to exit the Configuration Wizard.

Configuring Launchpad to Show the Managed Server

To configure a launchpad to show the managed server that it is connected to, add this line to the `director.properties` in the `local` home directory:

```
[expr]adf.headerextra = ': ' || weblogic.Name
```

Start Oracle WebLogic Server

You must start your Administration Server, Managed Servers, and clusters to complete the installation. For information about starting managed servers using Node Manager and Administration Servers, see "Starting and Stopping Oracle WebLogic Server Instances" in *Administering Server Startup and Shutdown for Oracle WebLogic Server*.

See also Setting Server Parameters to Support Enterprise Data Quality for important information about setting server parameters for startup.

Running Multiple EDQ Servers in the Same Domain

To support high availability scenarios, Oracle recommends that you configure a cluster of multiple EDQ servers to share the incoming load (for example, from a large number of simultaneous web service requests), and to provide continuous service in the event of failure of an individual server.

> **✎ Note:**
>
> You must have an appropriate license of WebLogic Server and Coherence to configure EDQ to operate in a cluster.

This section provides some basic guidance about how to configure EDQ to support such a model using Oracle WebLogic Server.

Multiple EDQ managed servers can be configured to run in the same WebLogic Server domain either in a cluster or not. If all the servers are on the same machine, each server must listen on a different port.

The Java Required Files (JRF) Template *must* be applied to any managed servers that were created using the WebLogic Server Administration Console. This is equivalent to the library targeting performed automatically by the WebLogic Server Configuration Wizard.

The final step is:

- Use the WebLogic Server Administration Console to modify the managed server settings for the additional EDQ servers.

Once multiple EDQ servers have been configured, you can leave them un-clustered and accessed directly using their respective Launchpad URLs to the relevant port, or you can configure them as part of a cluster using standard WebLogic Server practices. You can configure a separate front-end load balancer to handle incoming web service requests through a single cluster URL.

Configuring EDQ for High-Availability in a WebLogic Server Cluster

You can install and configure Oracle Enterprise Data Quality for high availability in an Oracle WebLogic Server cluster environment. The high availability features in Oracle Enterprise Data Quality have been enhanced to make the system function in a clustered environment, and tolerate individual RAC node failures and reconnect after complete database failures.

For more information on the high-availability features in Enterprise Data Quality, please see *Oracle Fusion Middleware Understanding Oracle Enterprise Data Quality*.

- Configuring EDQ for Clustered Mode Deployment
- Starting EDQ in a Cluster
- Enabling JMX API and Command Line for HA Clusters
- Landing Area

Configuring EDQ for Clustered Mode Deployment

To install and configure Oracle Enterprise Data Quality for high availability in an Oracle WebLogic Server cluster environment:

- Plan the installation, and install the prerequisites and software, as described in Prerequisites for these Procedures.
- Run the Repository Creation Utility (RCU), as described in Creating an EDQ Database Repository. To get a full high availability implementation, you must deploy the schema to an Oracle RAC instance. Deploying to a standard Oracle database does not provide HA capability for the database.
- Run the Domain Configuration wizard to provision a domain, as described in Creating the WebLogic Server EDQ Domain and with the configuration steps for a cluster noted in Running the Domain Configuration Wizard for a Clustered Deployment.
- Start EDQ in a cluster as described in Starting EDQ in a Cluster.
- Enable or configure options for high availability in EDQ. See the sections:
  - Enabling JMX API and Command Line for HA Clusters
  - Landing Area
- Running the Domain Configuration Wizard for a Clustered Deployment

Running the Domain Configuration Wizard for a Clustered Deployment

To deploy EDQ in a clustered mode using the schema you created in the RCU:

1. Follow the steps in Starting the WebLogic Server Domain Configuration Wizard and Navigating the Domain Configuration Wizard Screens to launch and configure the basic configuration screens in the Domain Configuration wizard.

2. On the Advanced Configuration screen, select the **Topology** option and then select **Next**. In the Topology option, you can Add, Delete or Modify settings for Server Templates, Managed Servers, Clusters, Virtual Targets and Coherence.

3. On the Manage Servers screen, select the EDQ server, for example "edq_server1", and select the **Clone** button to create a copy of the server. If you do not want to clone the server, select the **Add** button to add a new server and target it to the cluster.

   Name this server appropriately, for example "edq_server2".

   The listen port should not be the same port as in Server 1 and the Server Group should be EDQ Managed Server. The listen address should be the same as server 1, assuming they are both running on the same machine.

   Click **Next** and move to the Clusters screen.

4. Select the **Add** button to add a cluster.

   If servers are being deployed on the same machine, enter the machine name as the cluster address. If not, the cluster address should be a comma-separated list of the machine names.

   Leave the frontend host blank if you are not deploying a load balancer. If you are deploying with a load balancer, enter the URL of the load balancer.

   Click **Next**.

5. On the Assign Servers to Clusters screen, assign both the servers that have been created to the cluster that you have just created.

6. Step through the Coherence Clusters screen.

7. On the Machines screen you can add machines in the cluster, and allocate managed servers to them. The AdminServer does not need to be allocated to a machine since it is started manually via the `startWebLogic.sh` script and is not managed by the Node Manager.

8. Step through the remaining screens and select **Create** to create the cluster.

Starting EDQ in a Cluster

Once the domain has been created, start the Admin Server by running the following command:

```
FMW_HOME/user_projects/domains/edq_domain/bin/startWebLogic.sh
```

If you select Production Mode for the Domain Mode and JDK screen when you create the domain, you will be prompted for the Administrator user login credentials. Enter the same credentials that you received in the Administrator Account screen

Then start the Node Manager with the command:

```
FMW_HOME/user_projects/domains/edq_domain/bin/startNodeManager.sh
```

You can then log onto the Admin Console through a browser and start both of the Managed servers that were created during the configuration steps.

You can access the servers as follows:

- You can access any of the servers' launchpads. Connect to one of the servers and start the Director application. Or,

- You can connect via a load balancer in front of the cluster, and will be connected to whichever managed server the load balancer picks.

To display a dialog indicating which server the GUI is connected to, right-click the server in the project browser and select **Server Information**.

Enabling JMX API and Command Line for HA Clusters

By default, the EDQ high availability deployment templates come with the internal JMX server enabled, but each server uses a dynamically assigned port.

If you are running EDQ as a single managed server and not using a WebLogic cluster, the default management port is 8090.

If you are running EDQ in a WebLogic cluster, EDQ uses random ports which are allocated on server start. You can examine the Server Access MBean to find the current ports, but the ports change when the EDQ servers are restarted.

The standard EDQ JMX client has been updated to query these server access beans automatically, using the MBean server in the Administration server. Specify the host name and port for the Administration server using the command: `$ java -jar jmxtools.jar runjob ... adminhost:7001`, and the connector code redirects the request to one of the managed servers on the correct port.

The username and password with jmxtools must be valid in WebLogic and EDQ. Use a user from the authentication provider setup in OPSS.

If a single EDQ server runs on a host, do not run the servers in a cluster. In this case, the default port 8090 is used.

If multiple EDQ managed servers run on the same host, you can configure different fixed port for each server in a cluster. (If the servers are running on different machines, they can use the same port). Edit the `oedq.local.home` version of `director.properties` file and replace the current management port setting with: `[expr]management.port = 8090 + servernum - 1`

`servernum` is a precomputed value, which is "1" for edq_server1, "2" for edq_server2 and so on.

This setting causes the first EDQ managed server to run the JMX server on port 8090, second EDQ managed server to run the JMX server on port 8091, and so on.

The Siebel connector does not support the use of the admin server/port for JMX calls, but needs an explicit port like 8090.

You can determine the current JMX ports for each running server using the command:

```
$ java -jar jshell.har scripts/system/serveraccess.groovy -server host:port -user user -pw pw
```

`host:port` is the domain admin server.

Landing Area

The landing area is a feature that allows EDQ to read and write data to the server file system. If the landing area is to be used in a cluster then some consideration needs to be given as to

how the landing area is shared amongst the managed servers. By default the landing area is located in the oedq.local.home area. If this is shared amongst the various hosts supporting EDQ managed servers then the landing area will continue to work as in a non-clustered system.

If the landing area file system is not shared amongst the managed servers, but is required for use then a number of options are available:

- The location of the landing area can be changed using the `landingarea` property in the `director.properties` file to a location that is shared amongst the hosts running EDQ managed servers.

- Since an EDQ job runs all its tasks and processes on the same managed server, any files consumed or generated by the job are written to its local landing area. External tasks can be added to the job to transfer any incoming or outgoing files to an appropriate shared location.

- If an EDQ job is consuming external files then these could be copied to all managed servers before the job is started.

- If an EDQ job is generating files for consumption by further EDQ jobs then the landing area can be synchronized across the various managed server between the various job run by use of an external tool such as rsync.

If files are being generated and consumed within the same job then a shared landing area may not necessary. This is because the entire job will run on the same managed server and so access the local landing area.

5

Configuring Enterprise Data Quality with Apache Tomcat

This chapter describes how to configure EDQ and the database to operate with a Tomcat Application Server and perform the necessary additional configuration steps.

> **Note:**
>
> These instructions apply to Apache Tomcat environments only. If you are using Oracle WebLogic Server, you must follow the directions in Configuring Enterprise Data Quality with Oracle WebLogic Server

This chapter includes the following sections:

- Prerequisites for these Procedures
- Creating the EDQ Database Objects
- Configuring Tomcat Application Server
- Creating Directories for Use With Tomcat Application Server
- Configuring EDQ to work with Tomcat Application Server
- Verifying EDQ Functional Packs
- Deploying the EDQ Application on a Tomcat Application Server

Prerequisites for these Procedures

Before performing the procedures in this section, you must first read and satisfy the steps in:

- Planning an Enterprise Data Quality Installation
- Installing the Required External Software Components
- Installing Enterprise Data Quality

Creating the EDQ Database Objects

EDQ requires three database accounts and three schemas. The first schema contains configuration data, the second contains results data, and the third contains staging data. These objects must be created in the database that you installed as the EDQ repository in Installing the Required External Software Components

- Creating an EDQ Repository in an Oracle Database
- Creating an EDQ Repository in an Apache Derby Database

Creating an EDQ Repository in an Oracle Database

This procedure configures an Oracle database as the repository for EDQ.

1. Create three database accounts named `edqconfig`, `edqresults`, and `edqstaging` or other names of your choosing. When selecting names, consider that one schema will be the configuration schema and the other will be the results schema.

2. Grant the following privileges to each of these accounts:

   - `CREATE SESSION` (normally granted by `CONNECT`)

   - `CREATE INDEXTYPE` (normally granted by `RESOURCE`)

   - `CREATE SEQUENCE` (normally granted by `RESOURCE`)

   - `CREATE TABLE` (normally granted by `RESOURCE`)

   - `CREATE TYPE` (normally granted by `RESOURCE`)

   - `CREATE INDEX`

   - `DROP TABLE`

   - `DROP INDEX`

   - `UNLIMITED TABLESPACE` (recommended to allow the auto-extension of tablespace)

Creating an EDQ Repository in an Apache Derby Database

EDQ 12.2.1.4.3 and later on Apache Tomcat support the use of an embedded Apache Derby database for the configuration and results schemas.

This procedure configures an Apache Derby embedded database as the repository for EDQ.

1. Use the *migration.jar* tool that is shipped with EDQ to create and initialize an Apache Derby database. Run the tool using the following command:

   ```
   java -jar migration.jar init 'derby:(/path_to_db;create=true)'
   ```

   The value you enter in *path_to_db* should refer to a non-existent directory. This is the location where the database files will be created. After the files are created you can copy them for use with EDQ.

2. Edit `director.properties` file and use these properties to configure EDQ to use the Derby database:

   ```
   dataSource.driverClassName       =
   org.apache.derby.jdbc.EmbeddedDriver
   dataSource.url                   = jdbc:derby:/path_to_db
   dataSource.username              = APP
   dataSource.password              = app

   resultsDataSource.driverClassName =
   org.apache.derby.jdbc.EmbeddedDriver
   resultsDataSource.url             = jdbc:derby:/path_to_db
   ```

```
resultsDataSource.username      = APP
resultsDataSource.password      = app
```

Here, *path_to_db* is the absolute location of the directory containing the database files that you created in *Step 1*.

If you are using EDQ 12.2.1.4.3, you need to add the following additional settings:

```
indexing.maxthreads         = 1
workunitexecutor.commitSize = 1
```

Here,

- **indexing.maxthreads** prevents multiple indexes for a table being created in separate threads. In EDQ 12.2.1.4.4 and later this setting is not needed because the indexing logic has been updated.

- **workunitexecutor.commitSize** prevents multiple transactions on a single table happening in multiple threads. In EDQ 12.2.1.4.4 and later this value is set automatically.

Note that the setup process for the Oracle Enterprise Data Quality on Tomcat 12.2.1.4.3 and later Marketplace images has been updated to include Apache Derby as a schema configuration option. If selected, no further configuration is required.

Configuring Tomcat Application Server

> **Note:**
>
> If you are upgrading from Oracle Enterprise Data Quality (EDQ) 12.2.1.4.0 or earlier to 12.2.1.4.1 or later versions, edit the `catalina.properties` file to remove the reference to Oracle Application Development Framework (ADF) jars, and delete the ADF jars. Optionally, you can also remove `lib.adf` directory from their respective shared library path.
>
> If you are performing a fresh install of EDQ 12.2.1.4.1 or later versions, you need not set up ADF jars. However, if you are performing a fresh install of EDQ 12.2.1.4.0, follow the steps to set up ADF jars.

Create the necessary Oracle Application Development Framework (ADF) shared library:

1. Create a `lib.adf` ADF shared directory in your Tomcat directory.

2. Unzip the `adf-essentials.zip` file, delivered with the EDQ product, into the `lib.adf` directory. The JAR files in this zip file must be in the `lib.adf` directory, not the `adf-essentials` directory that it unzips to by default.

3. Configure the Tomcat `common.loader` property (typically located in the `catalina.properties` file) to load Oracle ADF from their respective shared library paths. For example:

   ```
   common.loader=${catalina.base}/lib,${catalina.base}/lib/*.jar,
   ${catalina.home}/lib,${catalina.home}/lib/*.jar,
   ${catalina.home}/lib.adf/*.jar
   ```

Creating Directories for Use With Tomcat Application Server

Create two empty directories to contain your EDQ configuration files. For example:

On Linux and UNIX operating systems:

`/opt/edqconfig/oedqhome`

`/opt/edqconfig/oedqlocalhome`

On the Windows operating system:

`C:\edqconfig\oedqhome`

`C:\edqconfig\oedqlocalhome`

The first directory (`oedqhome`) is the base configuration directory and will contain the configuration files that should not be changed post-installation. The second directory (`oedqlocalhome`) is the local configuration directory and will contain any custom settings that you create. Ensure that your application server user has read and write access to the two directories you create. For more information about these directories, see EDQ Directory Requirements

Configuring EDQ to work with Tomcat Application Server

Follow these steps to run the EDQ Configuration Application (`configapp`) to configure EDQ to work with Apache Tomcat and to populate the repository schemas with the required EDQ objects.

> ✎ **Note:**
>
> These instructions assume you have created the required database schemas as directed in Installing the Required External Software Components

To start the EDQ Configuration Application, follow these steps:

1. Log in to the system as your EDQ installation user.

2. Go to your *FMW_HOME*/edq/oracle.edq directory.

3. Start the Configuration Application by entering the following command:

   `java -jar configapp.jar`

   The EDQ Configuration Application is displayed.

Use Table 5-1 to configure EDQ:

Table 5-1   Running the EDQ Configuration Application

Screen	Action to Perform
EDQ Configuration Application Completion	Click **Begin** to proceed with the configuration. You may cancel the installation at any time by clicking **Cancel**.

Table 5-1   (Cont.) Running the EDQ Configuration Application

Screen	Action to Perform
Configuration Directory	Specify the two empty directories that you created in Creating Directories for Use With Tomcat Application Server to contain your EDQ configuration files. Click **Next** to continue.
Functional Packs	Select the functional packs you want and are licensed to use. Click **Next** to continue.
Configure the Configuration Schema	From the Selection Type list, select the correct option for the *configuration* schema that you created in Creating the EDQ Database Objects. • *Do not* use the **JNDI Connection** option. Enter the user name and password for the configuration schema. Select the option for the type of database, either **Oracle** or **Postgres**. Enter the host name of the system that contains the database, its port number, and the database unique id. (For Oracle the unique id is the SID not the service name. If you want to use a pluggable database, which requires a service name, you must use a raw JDBC URL or JNDI source.) Ensure that you configured the connection to the database correctly by clicking **Test**. Click **Next** to continue.
Configure the Results Schema	From the Selection Type list, select the correct option for the *results* schema that you created in Creating the EDQ Database Objects. • *Do not* use the **JNDI Connection** option. • For PostgreSQL, you may want to use the **Properties** option. Enter the user name and password for the results schema. Select the option for the type of database, either **Oracle** or **Postgres**. Enter the host name of the system that contains the database, its port number, and the database unique id. (For Oracle the unique id is the SID not the service name. If you want to use a pluggable database, which requires a service name, you must use a raw JDBC URL or JNDI source.) Ensure that you have configured the connection to the database correctly by clicking **Test**. Click **Next** to continue.
Selected Options	Review the summary of your selections. Use **Back** to modify any of your selections. Click **Finish** to complete the configuration and continue.
EDQ Configuration Application Completion	Click **Done** to exit the installation program.

Verifying EDQ Functional Packs

You can verify which EDQ functional packs have been installed by using the EDQ Configuration Application as described in Configuring EDQ to work with Tomcat Application Server Additionally, you should be aware of the following:

- All Functional Packs are needed if you want to use the full capabilities of the EDQ Customer Data Services Pack on your EDQ server, such as for integration with Siebel Customer Relationship Management or Universal Customer Master.
- All Functional Packs are needed if you want to install Oracle Watchlist Screening on your EDQ server.

Deploying the EDQ Application on a Tomcat Application Server

The EDQ application must be manually deployed on Tomcat Application Servers.

1. Stop the application server.

2. Deploy any one of the following files on your application server:

   - To deploy the EDQ application on Tomcat version 8, use `edq.war`.

     For example, see the *Tomcat Web Application Deployment* web page at https://tomcat.apache.org/tomcat-9.0-doc/deployer-howto.html.

   - To deploy the EDQ application on Tomcat version 10 or later, use `jakartaee/edq.war`.

     For example, see the *Tomcat Web Application Deployment* web page at https://tomcat.apache.org/tomcat-10.1-doc/deployer-howto.html.

   > **Note:**
   >
   > You must use the `jakartaee/edq.war` while using tomcat version 10.1. This is only applicable to 12.2.1.4.3 and later versions.

3. To link EDQ to the base and local configuration directories that you created in Creating Directories for Use With Tomcat Application Server, you can do one of the following:

   - On any of the EDQ-supported platforms, you can use a Java property named `-edq.config.path` that specifies the paths to those directories. The syntax is as follows, where *path_to_base_config* is the path to the base configuration directory and *path_to_local_config* is the path to the local configuration directory.

     Linux and UNIX (colon separates the paths):

     ```
     -Dedq.config.path=path_to_base_config:path_to_local_config
     -Doracle.mds.cache=simple
     ```

     Windows (semi-colon separates the paths):

     ```
     -Dedq.config.path=path_to_base_config;path_to_local_config
     -Doracle.mds.cache=simple
     ```

> **Note:**
>
> The `oracle.mds.cache` property is not needed in EDQ 12.2.1.4.1 and later.

- On Windows, you can append an environment variable named `EDQ_CONFIG_PATH` to your `JAVA_OPTS` environment variable, as shown in the following, where *path_to_base_config* is the path to the base configuration directory and *path_to_local_config* is the path to the local configuration directory.

  `EDQ_CONFIG_PATH="`*path_to_base_config*`;`*path_to_local_config*`"`

For more information about setting Java parameters, see the Apache Tomcat Documentation website at

http://tomcat.apache.org/

4. Restart your application server service and ensure that `edq.war` is successfully deployed.

6

Setting Server Parameters to Support Enterprise Data Quality

This chapter contains recommended Managed Server parameter settings for the framework that supports EDQ. These settings apply to both Oracle WebLogic Server and Apache Tomcat application server environments.
This chapter contains the following topics:

- How to Set the Server Parameters
- Recommended Parameter Values
- Example of Parameter Settings

How to Set the Server Parameters

This section describes how to set the recommended server parameters for WebLogic Server and Tomcat domains. Follow the guidelines that are appropriate for your EDQ installation.

- How to Set Server Parameters in a WebLogic Server Domain
- How to Set Server Parameters in a Tomcat Domain

How to Set Server Parameters in a WebLogic Server Domain

When you start a Managed Server in a WebLogic Server domain, the startup passes a set of configuration parameters to initialize a Java Virtual Machine (JVM). These parameters include, but are not limited to, the server's JVM settings and the path to the server configuration directories.

The method of setting the required server configuration parameters for EDQ varies, depending on the method that you use to start the Managed Servers. The following points explain the differences:

- When starting an EDQ Managed Server from the WebLogic Server Administration Console, the configuration settings for the Managed Server are pre-defined as server startup arguments, because they were set by the domain Configuration Wizard (see Creating the WebLogic Server EDQ Domain). When starting the server, these arguments are applied by default, but you can change them to the recommended settings described in Recommended Parameter Values.

- When starting an EDQ Managed Server by using a startup script, the server startup arguments that are defined in the WebLogic Server Administration Console are ignored. However, when the domain Configuration Wizard was run, it created an `edq_server1` Managed Server that contains default JVM settings. These settings are stored in the `setStartupEnv.cmd` (Windows) or `setStartupEnv.sh` (Linux) script. You can edit this script to add or change the name of the Managed Server and to apply the recommended JVM settings described in Recommended Parameter Values.

How to Set Server Parameters in a Tomcat Domain

Consult the Tomcat documentation for your version of the application server. Documentation for all versions of Tomcat are at

http://tomcat.apache.org

Recommended Parameter Values

This section describes the recommended server parameter settings.

• Setting JVM Parameters
• Setting Other Parameters

Setting JVM Parameters

You must set the following parameters for the Java Virtual Machine (JVM) that runs the EDQ environment. Most of these parameters apply to the HotSpot JVM, which is the central JVM of the Java Development Kit (JVM) from Oracle. This is the preferred JVM because it yields the best performance for EDQ.

• Maximum heap memory, `-Xmx,` should normally be set to approximately 50% of available physical memory on the host server. For example, on a server with 32GB of RAM, start by setting it to 16GB with `-Xmx16384m`. Depending on the relative usage of memory between the JVM and native processing, you may need to be adjust this setting. This parameter is common to most JVMs.

• Reserved Code Cache size should be set to 128m, `-XX:ReservedCodeCacheSize=128m`. However, for systems running CDS or OWS, where there are a lot of scripts, it is advisable to set the Reserved Code Cache size to 512m. This parameter is specific to HotSpot JVM.

Setting Other Parameters

The following additional parameter setting is recommended.

• Headless mode should always be set to true on UNIX systems, `-Djava.awt.headless=true`. This parameter is specific to the Java Abstract Window Toolkit (AWT) library. Setting this is recommended because some of the graphic drawing routines (such as around graphs) will attempt to connect to the display device on a client to size fonts and other attributes. If the machine does not have a display device, this may interfere with operation. The `headless=true` setting directs the graphics library to operate without connecting to a display.

Example of Parameter Settings

The following is an example of the full set of recommended parameter options for a typical UNIX or Linux server with 32GB of RAM:

```
-Xmx16384m
-XX:ReservedCodeCacheSize=128m
-Djava.awt.headless=true
```

7

Next Steps After Configuring Enterprise Data Quality

This chapter describes how to start using EDQ and is intended to help you become familiar with the main components of EDQ.
This chapter includes the following sections:

- Logging Into EDQ
- The EDQ Launchpad

Logging Into EDQ

You can access the EDQ Launchpad and client applications by starting a supported browser and enter the following URL:

```
http://server_name:port_number/edq
```

where *server_name* is the name of the server onto which you installed EDQ and *port_number* is the HTTP or HTTPS port that your application server is running against (8001 on WebLogic Server by default, 8080 on Tomcat). If you deployed the application server to run against a different port, you should use that port number.

Enter your login credentials for the appropriate application server:

WebLogic Server Login
Use the WebLogic Server user that you created in Navigating the Domain Configuration Wizard Screens

Tomcat Login
Use the default EDQ administrator account `dnadmin` with the password `dnadmin`. You are prompted to change the password the first time you log in. This password must meet the default security standards.

The EDQ Launchpad

The Launchpad provides access to the EDQ client applications, services, and system management. The following may be displayed on the Launchpad when EDQ is installed:

Options	Description
Director	Starts the Director client application, which is the main configuration application.
Server Console	Allows you to perform the following tasks: • Schedule jobs • View current tasks • View the event log • View job results

Options	Description
Dashboard	Starts the Dashboard, where published data quality metrics are displayed.
Match Review	Starts the Match Review application, which allows users to view an overview of the reviews assigned to them and to launch the review application.
Case Management	Case Management is an application designed to support the manual investigation of results from data quality processes. It is also used as the main investigation application in Oracle Watchlist Screening, for both batch and real time screening results. Using Case Management, privileged users can manage and review matching results using highly configurable workflows with a comprehensive audit history of all investigation work.
Case Management Administration	Case Management Administration provides the following areas of functionality: • Workflow Administration—allows you to create, edit, copy, import, export and delete workflows. • Case Source Administration—allows you to import, export and delete case source definitions. • Permission Administration—allows you to define and manage the data that users have permission to access.
Administration	This application allows a sufficiently privileged user to configure EDQ users, permission groups, password and security rules, extensions, functional packs, the applications that appear on the launchpad, and the ability to view and monitor sessions.
Web Services	Displays details of the web services configured in the EDQ server, and allows access to the web service tester.
Login	An explicit Login button in the upper right corner. Upon logging in the full list of pages available to the user is displayed. Once the user is logged in they have access to a user menu by clicking on their user name in the top right of the screen. This provides access to log out and change password functionality (change password is only shown if using an internal realm).
Change Password	Allows users to change their passwords. A user must log in with the existing password and then provide and confirm the new password. This option is not available when using a WebLogic Server.
Help	Allows you to retrieve information about how to use each EDQ application.

An Administrator can reconfigure the Launchpad to define the user applications and links that are displayed to users by using the Administration pages. For more information, see the *Enterprise Data Quality Online Help*.

8

Upgrading Enterprise Data Quality

If you are upgrading EDQ from any previous 12*c* release (12.1.3, or 12.2.1.x) to 12c (12.2.1.4.0), refer this section. If you are upgrading EDQ from a previous version on WebLogic, refer *Oracle Enterprise Data Quality Upgrade Guide*, and if you are upgrading EDQ on Tomcat, refer Support Article 2433058.1

> **Note:**
>
> You must review Oracle Fusion Middleware upgrade documentation before referring to the Oracle Fusion Middleware Oracle Enterprise Data Quality Upgrade Guide.
>
> Read *Planning an Upgrade of Oracle Fusion Middleware* for additional guidelines for preparing to upgrade to Oracle Fusion Middleware 12c (12.2.1.4.0). This documentation also includes descriptions of terminology changes that you must understand to move forward to a 12*c* environment.

- Upgrading from previous 12c releases to 12.2.1.4.0

Upgrading from previous 12c releases to 12.2.1.4.0

Upgrading from an installation of EDQ 12*c* that uses the Oracle database and the WebLogic application server as its platform to EDQ 12.2.1.4.0 using the Oracle Database and Oracle WebLogic as its platform is essentially a two-part process:

- Part 1: Perform a Partial Installation of EDQ 12.2.1.4.0
- Part 2: Upgrade your Old EDQ Environment

Part 1: Perform a Partial Installation of EDQ 12.2.1.4.0

To perform a partial installation:

1. Ensure that your environment satisfies all EDQ 12.2.1.4.0 System Requirements, as outlined in Satisfying EDQ System Requirements, and upgrade your environment, if necessary.
2. Download EDQ 12.2.1.4.0, as described in Downloading EDQ.
3. Download the Fusion Middleware Infrastructure release 12.2.1.4.0.
4. If your environment's current Java Development Kit does not support EDQ 12.2.1, install a Java Development Kit to support EDQ 12.2.1, as described in Installing a Java Development Kit to Support EDQ.
5. If your environment's current Oracle Fusion Middleware Infrastructure does not support EDQ 12.2.1, install Oracle Fusion Middleware Infrastructure to support EDQ 12.2.1, as described in Installing Oracle Fusion Middleware Infrastructure (includes Oracle WebLogic Server).

> **✎ Note:**
>
> The Oracle Fusion Middleware Infrastructure that you used to support your old release of EDQ will not support your new release of EDQ. Therefore, unless, you have already installed the new release of the Oracle Fusion Middleware Infrastructure for some other purpose, you will have to install it now.
>
> The new release of the Fusion Middleware Infrastructure will have to be installed in a different Fusion Middleware Home to your old release.

6. If your environment's current Oracle Database does not support EDQ 12.2.1, upgrade the Oracle Database to a release that will support EDQ 12.2.1.

7. Install Enterprise Data Quality 12.2.1.4.0 assets, as described in Installing Enterprise Data Quality , ensuring that you install EDQ in your new Fusion Middleware Infrastructure's Fusion Middleware Home. This step entails installing the Oracle Fusion Middleware EDQ distribution. This is usually downloaded as a `.zip` file that contains a single file with a name similar to `fmw_12.2.1.4.0_edq.jar`. Running this `.jar` file invokes the Oracle Universal Installer. You do not need to configure or extend a WebLogic domain now, since you will upgrade the WebLogic domain using the Reconfiguration Wizard at a later stage.

> **✎ Note:**
>
> This step places elements of Enterprise Data Quality in your environment's Fusion Middleware Home location. Some of these will be used later in the installation process. However, you will not have a working EDQ environment at this stage.

Part 2: Upgrade your Old EDQ Environment

Before you upgrade to the latest EDQ Environment, here are a few considerations:

• When upgrading EDQ instances that are configured for high availability:

Enterprise Data Quality 11*g* and higher can be enabled for high availability by configuring several WebLogic Managed Servers within the same WebLogic Cluster. In releases of EDQ 11, and also in EDQ 12.1.3, it was necessary for each of these Managed Servers to have its own dedicated database repository (in other words, its own pair of Config and Results datasources), and its own dedicated config directory.

In EDQ 12.2.1, configuration for High Availability was simplified. From release 12.2.1.0.0 and above, all the WebLogic managed servers in a cluster can share the same database repository. In other words, only one pair of Config and Results datasources is required and all the WebLogic managed servers in a cluster can share the same base and local config directories.

Oracle recommends that when you upgrade an instance of EDQ 11*g* or 12.1.3 that is configured for high availability, you change its configuration to take advantage of the simplified approach to High Availability introduced in EDQ 12.2.1. To do this, perform the following tasks:

- Upgrade only one database repository (in other words, only one pair of Config and Results schemas) using the Upgrade Assistant.

- Run the Oracle Fusion Middleware Reconfiguration Wizard a single time, supplying the connection details of the database repository that you upgraded. As you run the Wizard, on the Advanced Options screen, ensure that you select the Managed Servers, Clusters and Coherence option, and then, referring to the EDQ High Availability Guide for release 12.2.1, configure your upgraded EDQ instance for High Availability using the new, simplified architecture.

- Ensure that all your WebLogic Managed Server's server parameters, including, its 'startup arguments' or 'Java options', are configured to point towards the same base and local config directories.

- It is also possible to maintain the existing high availability configuration after an upgrade, by continuing to use multiple schemas and config areas. To do this:

  - Run the Upgrade Assistant separately against each pair of Config and Results schemas.

  - Disable coherence by setting coherence.clustering = false in the `director.properties` file.

To start upgrading your old environment:

1. Back up your old EDQ environment.

   Before you begin, Oracle recommends that you make a complete backup of your old EDQ environment, including your EDQ domain. For instructions, see *Backup and Recovery Strategies for Upgrade* in the *Planning an Upgrade of Oracle Fusion Middleware* documentation.

2. Stop the WebLogic Server Administration Server and all managed servers.

3. Start the Upgrade Assistant, to upgrade your EDQCONFIG (configuration), EDQRESULTS (results), EDQSTAGING (only if upgrading from 12.2.1.x. This will not be present if upgrading from 12.1.3) and OPSS, IAU and STB schemas in the EDQ repository database.

   a. Go to the following directory, where `12.2.1_FMW_HOME` is the version 12.2.1.4.0 Oracle home directory:

      UNIX:

      `12.2.1_FMW_HOME/oracle_common/upgrade/bin`

      Windows:

      `12.2.1_FMW_HOME\oracle_common\upgrade\bin`

   b. Run the following program:

      Linux or UNIX:

      `./ua`

      Windows:

      `ua.bat`

4. At the Welcome screen, Review Important Reminders before Proceeding. The Welcome screen contains reminders to consider before proceeding with your upgrade. Make sure you read these and verify that you are ready to proceed. For assistance or more information on any screen, click **Help**. Otherwise, click **Next**.

5. Select the type of upgrade and then select **All Schemas Used by a Domain** option. In the Domain Directory field, select your old EDQ environment's domain directory (this is likely to be similar to `/OLD_FMW_HOME/user_projects/domains/ edq_domain`). Click **Next**.

6. The Component List screen will display a list of the components (or in other words, schemas) to be upgraded in your old EDQ environment's domain. The following components will be listed:

   • Oracle Enterprise Data Quality Results

   • Oracle Audit Services

   • Oracle Platform Security Services

   • Common Infrastructure Services

   • Oracle Enterprise Data Quality Configuration

   Select the components and click **Next**.

7. In the Prerequisites screen, verify that you have met the prerequisites, and provided that you have, check the relevant boxes. Note that the Upgrade Assistant will not verify that the prerequisites have been met. Click **Next**.

8. Specify Database and Schema Credentials:

   a. The EDQ Results schema, IAU, OPSS, STB, and EDQ Configuration schema screens are displayed in succession.

   b. On the first screen, specify connection details for the database that contains the EDQ Results (EDQRESULTS) schema, and then click Connect. You are then prompted for the password for the schema user.

   c. The remaining screens are automatically populated with the database connection and schema credentials that you supplied on the EDQ Results Schema screen. If these entries are not correct for any schema, change the entries and ensure that a database connection is made.

9. Complete the upgrade validation. On the Examine screen, the Upgrade Assistant performs a series of validations before upgrading the selected components. Ensure that all validations have succeeded.

10. Initiate the Upgrade. Click **Upgrade** on the Upgrade Summary screen to initiate the upgrade. The Upgrade Progress screens shows information about the progress of the upgrade, and the Upgrade Success screen summarizes the upgrade.

11. Create the Required 12*c* Schemas.

> **✎ Note:**
>
> You only need to create the staging schema if upgrading from 12.1.3. If upgrading from a previous version of 12.2.1 this step can be skipped.

Oracle Fusion Middleware Release 12.2.1 introduced schemas and tables that must be created with the 12.2.1 version of the Oracle Repository Creation Utility (RCU). This step of the upgrade must be performed before Reconfiguring the EDQ Domain.

To create the required 12.2.1 schemas, see Creating an EDQ Database Repository for release 12.2.1. Follow all of those instructions, but only select the following schema on the Select Components screen:

- EDQ Staging Repository (note that you will find this under the Oracle EDQ option).

12. Reconfigure the EDQ Domain. In this step you will run the Oracle Fusion Middleware Reconfiguration wizard to complete the upgrade of your WebLogic Server domain environment.

13. Check that the Staging Mode is set to Nostage, by performing the following steps:

   a. Log in to the WebLogic Server Administration Server console from any web browser.

   b. In the left pane, expand **Environment** and then click **Servers**. The Summary of Servers screen will be displayed.

   c. In the Summary of Servers screen, click your EDQ Managed WebLogic Server (for example: `edq_server1`).

   d. Select the **Configuration** tab, then the **Deployment** tab.

   e. From the Staging Mode menu, ensure that select **Nostage** is selected, and, if it is not, select it. Do not change any other server settings.

   f. Click **Save**.

14. Stop Node Manager:

   a. If the Node Manager was not already running, and you started it using the command above, then stop it by closing the command shell from which it is running.

   b. If Node Manager was already running (for example, because it was already running as a service in your environment), then stop it using the appropriate means (for example, stop the Node Manager service).

15. Start the Reconfiguration Wizard to upgrade the EDQ domain:

   a. Go to the following directory, where `12.2.1_FMW_HOME` is the version 12.2.1 Oracle home directory:

   Windows:

   ```
   12.2.1_FMW_HOME\oracle_common\common\bin
   ```

   UNIX:

   ```
   12.2.1_FMW_HOME/oracle_common/common/bin
   ```

   b. Start the domain reconfiguration wizard.

   Windows:

   ```
   reconfig.cmd -log=log_file
   ```

   UNIX:

   ```
   ./reconfig.sh -log=log_file
   ```

16. Specify the full path and file name in place of `log_file`. Creating a log file can be helpful if you need to troubleshoot the reconfiguration process.

> **✎ Note:**
>
> If the following error message is displayed, it indicates that the default cache directory is not valid:
>
> ```
> *sys-package-mgr*: can't create package cache dir
> ```
>
> You can change the cache directory by setting the environment variable CONFIG_JVM_ARGS. For example:
>
> ```
> CONFIG_JVM_ARGS=-Dpython.cachedir=valid_directory
> ```

17. Specify the domain. On the Select Domain screen, specify the full path to the location of your old EDQ domain (it is likely to be similar to `FMW_HOME/user_projects/domains/edq_domain` or `FMW_HOME\user_projects\domains\edq_domain`). You can also click Browse and use the file manager window to help you select the domain location.

18. View the Reconfiguration Setup Progress.

    The Reconfiguration Setup Progress displays the reconfiguration progress and verifies whether the base domain that you selected can be reconfigured to the 12.2.1 domain. The message `Core WLS Infrastructure Reconfigured Successfully!` indicates that the domain can be reconfigured to the 12.2.1 domain, and you can click **Next** to go to the next step. If this message is not returned, the domain cannot be reconfigured to the 12.2.1 domain. If this happens, check to see if the EDQ version is earlier than version 11*g*. If so, then you must first upgrade to EDQ 11*g* and then upgrade to version 12.2.1.

19. Select the Domain Mode and JDK. It is not possible to change the Domain Mode from development to production or vice versa.

    On the Domain Mode and JDK screen, specify the location of the Java Development Kit (JDK) to use in the new EDQ domain. This must be a JDK that conforms to the system requirements of your new EDQ environment.

20. Select the Database Configuration Type:

    a. In the Database Configuration Type screen, the RCU Data option should be selected by default. The connection details for your EDQ repository database should be automatically populated, and the schema owner and password for the Common Infrastructure schema (STB) should be displayed. If this information is not automatically populated, or is not correct, make any necessary amendments.

    b. Click Get RCU Configuration. The Reconfiguration Wizard should report that it can connect to the Database Server and that it can retrieve schema data and bind local schema components with the retrieved data. If these steps fail, amend the connection data and click **Get RCU Configuration again**. Otherwise, click **Next**.

21. Check the JDBC Component Schema. On the JDBC Component Schema screen, the schemas associated with your old EDQ domain are listed in the lower half of the screen. If you need to make changes, select the check box next to the data source name and then make the changes.

> **Note:**
>
> The correct connection details for the EDQ Staging schema that you created using the Repository Creation Utility in step 5, above, may not be displayed. If they are not, amend them so that they are correct.

22. Test the JDBC Component Schema. On the JDBC Component Schema Test screen, test the data source connections that were detected. Select the schemas that you want to test and then click Test Selected Connections.

> **Note:**
>
> To test a connection, the database to which you are connecting must be running.

23. Select Optional Advanced Configuration Options to select additional domain options:

    a. To make changes to the Managed Server configuration, select **Managed Servers, Clusters and Coherence**. For help with any of the options, click the Help button.

    b. To skip making these changes, click `Next`.

24. Initiate the Domain Reconfiguration. On the Configuration Summary screen, review the configuration and then click **Reconfig** to start the reconfiguration process, or click **Back** to make changes.

25. Finish the Reconfiguration. Wait for the message that states, `Domain Reconfiguration Applied Successfully` and then click **Next**.

    A check mark and the message `Oracle WebLogic Server Reconfiguration Succeeded` indicate that the reconfiguration was successful. Click **Next**, and then click **Finish** to dismiss the Reconfiguration wizard.

26. Follow the instructions below to start the Upgrade Assistant and upgrade EDQ's domain configuration.

    a. Go to the following directory, where `12.2.1_FMW_HOME` is the version 12.2.1 Oracle home directory.

    Windows:

    `12.2.1_FMW_HOME\oracle_common\upgrade\bin`

    UNIX:

    `12.2.1_FMW_HOME/oracle_common/upgrade/bin`

    b. Run the following program:

    Linux or UNIX:

    `./ua`

    Windows:

    `ua.bat`

    c. At the Welcome screen, Review Important Reminders before Proceeding. The Welcome screen contains reminders to consider before proceeding with your upgrade. Make sure you read these and verify that you are ready to proceed. For assistance or more information on any screen, click Help. Otherwise, click **Next**.

> **✎ Note:**
>
> For more information about the following screens, see "Upgrading Oracle WebLogic Component Configurations" in Upgrading with the Upgrade Assistant.

d. At the Selected Schemas screen, select the All Configurations Used By a Domain option. At this point the name of the screen will change to All Configurations. In the Domain Directory field, select your old EDQ environment's domain directory (this is likely to be similar to `/OLD_FMW_HOME/user_projects/domains/edq_domain`), and click **Next**.

e. At the Component List screen, review the list of components to be upgraded (it should consist of Oracle JRF, System Components Infrastructure, and Oracle Enterprise Data Quality Configuration), and click **Next**.

f. At the Prerequisites screen, verify that you have met the prerequisites, and, provided that you have, check the relevant boxes (note that The Upgrade Assistant will not verify that the prerequisites have been met.). Then click **Next**.

g. At the Examine screen, the Upgrade Assistant will validate that the components can be upgraded, returning statuses for either upgrade not necessary or succeeded for each of them. When the examination has finished, click **Next**.

h. At the Upgrade Summary screen, click **Upgrade**.

i. At the Upgrade Progress screen, wait until the upgrade has finished, and then click **Next**.

j. At the Upgrade Success screen, click **Close**.

27. Apply the Upgrade Changes to the Base Domain, using the following steps:

a. Start the WebLogic Node Manager for the upgraded EDQ domain.

b. Start the WebLogic Server Administration Server from the upgraded EDQ domain. If it is running, stop and then restart it.

Windows:

`DOMAIN_HOME\bin\startWebLogic.cmd`

UNIX:

`DOMAIN_HOME/bin/startWebLogic.sh`

c. Log in to the WebLogic Server Administration Server console from any web browser.

d. In the Domain Structure list, click **Deployments**.

e. Under Deployments, select **edq** and then click **Update**.

f. Click **Next**.

g. Click **Finish**.

h. To complete the upgrade, start the EDQ managed server.

> **✎ Note:**
>
> The `startManagedWebLogic.sh` script is located in your EDQ WebLogic domain's `bin` directory. Running this script is one method of starting the EDQ WebLogic Managed Server. The `startManagedWebLogic.sh` script may be overwritten when the EDQ WebLogic Domain is reconfigured. Therefore, if you customized the script, any customization may have been lost during the upgrade process. This is a particularly important point to note if you embedded EDQ's Server Parameters (in other words, its 'startup arguments' or 'Java options') within this script. Without these, the EDQ Managed Server may not function correctly.
>
> See Setting Server Parameters to Support Enterprise Data Quality , for more information about EDQ's server parameters.

28. Open a Web Browser, and navigate to your upgraded EDQ environment's Launchpad. Unless you changed any of the default options when installing EDQ, its address would be `http://<your-server-name>:8001/edq`.

29. From the Launchpad, open the Director user interface and check to see the version of EDQ to confirm that the version is upgraded. From the menu in Director, click the **Help** > **About** option to check the version.

30. Verify that all your projects have been successfully migrated.

> **✎ Note:**
>
> On a correctly upgraded system, you do not need to specify the passwords for any data stores such as Oracle. No additional settings are required to ensure that you do not specify the passwords again.

31. If any of your processes use Match Review, you will need to run these (or jobs of which they are a part) with intelligent execution turned off before you can access Match Review. You can turn off Intelligent Execution in a job or process's Run Preferences.

9

Removing Enterprise Data Quality from a System

This chapter describes how to remove EDQ from Linux, UNIX, and Windows.

> **Note:**
>
> The procedures in the following sections remove the EDQ files that were installed by Oracle Universal Installer (OUI). They do not remove the EDQ application from Oracle WebLogic Server. Follow the instructions in "Removing Your Domain and Application Data" in *Installing and Configuring Oracle WebLogic Server and Coherence.* is chapter includes the following sections:

- Removing EDQ from a Linux or UNIX System
- Removing EDQ from a Windows System

Removing EDQ from a Linux or UNIX System

To remove EDQ from a Linux or UNIX system, follow these steps:

1. Log in to the target system as your EDQ installation user.

2. Go to the directory, `EDQ_HOME`/oui/bin.

   > **Note:**
   >
   > Ensure that all files in the directory have executable permissions for the user you logged in with before continuing.

3. Run the following command to start the `deinstall.sh` program:

   ```
   ./deinstall.sh
   ```

   The EDQ deinstallation program is displayed.

4. Select Enterprise Data Quality 12.2.1.1.0 as the distribution to remove, and then click **Uninstall**.

5. Navigate through the screens and then click **Finish** to complete the removal of EDQ from the system.

Removing EDQ from a Windows System

To remove EDQ from a Windows system, follow these steps:

1. Log in to the target Windows system as your EDQ installation user.

2. Locate the MS-DOS Command Prompt (`cmd.exe`), right-click on it, and then select **Run as administrator**.

3. Go to the installation directory *EDQ_HOME*`\oui\bin`.

4. Enter the following command:

   ```
   deinstall.cmd -deinstall
   ```

   The EDQ deinstallation program is displayed.

5. Select Enterprise Data Quality 12.2.1.1.0 as the distribution to remove, and then click **Uninstall**.

6. Navigate through the screens and then click **Finish** to complete the removal of EDQ from the system.

A

Updating the JDK After Installing and Configuring an Oracle Fusion Middleware Product

Consider that you have a JDK version jdk1.8.0_191 installed on your machine. When you install and configure an Oracle Fusion Middleware product, the utilities, such as Configuration Wizard (`config.sh|exe`), OPatch, or RCU point to a default JDK, for example, jdk1.8.0_191. After some time, Oracle releases a new version of the JDK, say jdk1.8.0_211 that carries security enhancements and bug fixes. From 12c (12.2.1.4.0) onwards, you can upgrade the existing JDK to a newer version, and can have the complete product stack point to the newer version of the JDK.

You can maintain multiple versions of JDK and switch to the required version on need basis.

- About Updating the JDK Location After Installing an Oracle Fusion Middleware Product
  The binaries and other metadata and utility scripts in the Oracle home and Domain home, such as RCU or Configuration Wizard, use a JDK version that was used while installing the software and continue to refer to the same version of the JDK. The JDK path is stored in a variable called JAVA_HOME which is centrally located in .globalEnv.properties file inside the *ORACLE_HOME*/oui directory.

About Updating the JDK Location After Installing an Oracle Fusion Middleware Product

The binaries and other metadata and utility scripts in the Oracle home and Domain home, such as RCU or Configuration Wizard, use a JDK version that was used while installing the software and continue to refer to the same version of the JDK. The JDK path is stored in a variable called JAVA_HOME which is centrally located in .globalEnv.properties file inside the *ORACLE_HOME*/oui directory.

The utility scripts such as config.sh|cmd, launch.sh, or opatch reside in the *ORACLE_HOME*, and when you invoke them, they refer to the JAVA_HOME variable located in .globalEnv.properties file. To point these scripts and utilities to the newer version of JDK, you must update the value of the JAVA_HOME variable in the .globalEnv.properties file by following the directions listed in Updating the JDK Location in an Existing Oracle Home .

To make the scripts and files in your Domain home directory point to the newer version of the JDK, you can follow one of the following approaches:

- Specify the path to the newer JDK on the Domain Mode and JDK screen while running the Configuration Wizard.

  For example, consider that you installed Oracle Fusion Middleware Infrastructure with the JDK version 8u191. So while configuring the WebLogic domain with the Configuration Assistant, you can select the path to the newer JDK on the Domain Mode and JDK screen of the Configuration Wizard. Example: `/scratch/jdk/jdk1.8.0_211`.

- Manually locate the files that have references to the JDK using `grep` (UNIX) or `findstr` (Windows) commands and update each reference. See Updating the JDK Location in an Existing Oracle Home .

> **✎ Note:**
>
> If you install the newer version of the JDK in the same location as the existing JDK by overwriting the files, then you don't need to take any action.

- Updating the JDK Location in an Existing Oracle Home
  The `getProperty.sh|cmd` script displays the value of a variable, such as JAVA_HOME, from the .globalEnv.properties file. The `setProperty.sh|cmd` script is used to set the value of variables, such as OLD_JAVA_HOME or JAVA_HOME that contain the locations of old and new JDKs in the .globalEnv.properties file.

- Updating the JDK Location in an Existing Domain Home
  You must search the references to the current JDK, for example 1.8.0_191 manually, and replace those instances with the location of the new JDK.

Updating the JDK Location in an Existing Oracle Home

The `getProperty.sh|cmd` script displays the value of a variable, such as JAVA_HOME, from the .globalEnv.properties file. The `setProperty.sh|cmd` script is used to set the value of variables, such as OLD_JAVA_HOME or JAVA_HOME that contain the locations of old and new JDKs in the .globalEnv.properties file.

The `getProperty.sh|cmd` and `setProperty.sh|cmd` scripts are located in the following location:

(UNIX) *ORACLE_HOME*`/oui/bin`
(Windows) *ORACLE_HOME*`\oui\bin`
Where, *ORACLE_HOME* is the directory that contains the products using the current version of the JDK, such as 1.8.0_191.

To update the JDK location in the .globalEnv.properties file:

1. Use the `getProperty.sh|cmd` script to display the path of the current JDK from the JAVA_HOME variable. For example:

   (UNIX) *ORACLE_HOME*`/oui/bin/getProperty.sh JAVA_HOME`
   (Windows) *ORACLE_HOME*`\oui\bin\getProperty.cmd JAVA_HOME`
   `echo JAVA_HOME`
   Where JAVA_HOME is the variable in the .globalEnv.properties file that contains the location of the JDK.

2. Back up the path of the current JDK to another variable such as OLD_JAVA_HOME in the .globalEnv.properties file by entering the following commands:

   (UNIX) *ORACLE_HOME*`/oui/bin/setProperty.sh -name OLD_JAVA_HOME -value` *specify_the_path_of_current_JDK*
   (Windows) *ORACLE_HOME*`\oui\bin\setProperty.cmd -name OLD_JAVA_HOME -value` *specify_the_path_of_current_JDK*

This command creates a new variable called OLD_JAVA_HOME in the .globalEnv.properties file, with a value that you have specified.

3. Set the new location of the JDK in the JAVA_HOME variable of the .globalEnv.properties file, by entering the following commands:

(UNIX) *ORACLE_HOME*/oui/bin/setProperty.sh -name JAVA_HOME -value *specify_the_location_of_new_JDK*
(Windows) *ORACLE_HOME*\oui\bin\setProperty.cmd -name JAVA_HOME -value *specify_the_location_of_new_JDK*

After you run this command, the JAVA_HOME variable in the .globalEnv.properties file now contains the path to the new JDK, such as jdk1.8.0_211.

Updating the JDK Location in an Existing Domain Home

You must search the references to the current JDK, for example 1.8.0_191 manually, and replace those instances with the location of the new JDK.

You can use the `grep` (UNIX) or `findstr` (Windows) commands to search for the jdk-related references.

You'll likely be required to update the location of JDK in the following three files:

(UNIX) *DOMAIN_HOME*/bin/setNMJavaHome.sh
(Windows) *DOMAIN_HOME*\bin\setNMJavaHome.cmd

(UNIX) *DOMAIN_HOME*/nodemanager/nodemanager.properties
(Windows) *DOMAIN_HOME*\nodemanager\nodemanager.properties

(UNIX) *DOMAIN_HOME*/bin/setDomainEnv.sh
(Windows) *DOMAIN_HOME*\bin\setDomainEnv.cmd