Oracle® Endeca Server

Migration Guide

Version 7.7.0 • January 2016

ORACLE

Copyright and disclaimer

Table of Contents

Copyright and disclaimer **2**

Preface **4**
About this guide 4
Who should use this guide 4
Conventions used in this guide 4
Contacting Oracle Customer Support 5

Chapter 1: Upgrading to Oracle Endeca Server Version 7.7.0 . . . . . **6**
Required reading 6
Upgrading from 7.6.x to 7.7.0 6
About upgrading client stubs 8

Chapter 2: Required Changes **9**
Platform support 9

Chapter 3: Behavioral Changes **10**
Changes to the EndecaServer.properties file 10
Hostname resolution improvements 11
Security of administrative and configuration operations is improved 12
Improved Endeca Server startup 12
Improved handling of invalid records by the Bulk Load Interface 12
Change to the data domain's idle timeout 12
A .profile file created for new data domains 12
A new EQL function added 13
Sorting records based on geocode 14
Automatic log cleanup by the Cluster Coordinator 14

Preface

Oracle® Endeca Server is a hybrid search-analytical engine that organizes complex and varied data from disparate sources. At the core of Endeca Information Discovery, the unique NoSQL-like data model and in-memory architecture of the Endeca Server create an extremely agile framework for handling complex data combinations, eliminating the need for complex up-front modeling and offering extreme performance at scale. Endeca Server also supports 35 distinct languages.

About this guide

This guide helps you upgrade your Oracle Endeca Server implementation by describing the upgrade procedure and listing the major changes between versions 7.6.x and 7.7.0.

Who should use this guide

This guide is intended for system administrators and developers who are upgrading Oracle Endeca Server on Windows or Linux.

Conventions used in this guide

The following conventions are used in this document.

Typographic conventions

This table describes the typographic conventions used when formatting text in this document.

Typeface	Meaning
User Interface Elements	This formatting is used for graphical user interface elements such as pages, dialog boxes, buttons, and fields.
`Code Sample`	This formatting is used for sample code phrases within a paragraph.
Variable	This formatting is used for variable values. For variables within a code sample, the formatting is *`Variable`*.
`File Path`	This formatting is used for file names and paths.

Symbol conventions

This table describes the symbol conventions used in this document.

Symbol	Description	Example	Meaning
>	The right angle bracket, or greater-than sign, indicates menu item selections in a graphic user interface.	File > New > Project	From the File menu, choose New, then from the New submenu, choose Project.

Path variable conventions

This table describes the path variable conventions used in this document.

Path variable	Meaning
`$MW_HOME`	Indicates the absolute path to your Oracle Middleware home directory, which is the root directory for your WebLogic installation.
`$DOMAIN_HOME`	Indicates the absolute path to your WebLogic domain home directory. For example, if `endeca_server_domain` is the name of your WebLogic domain, then the `$DOMAIN_HOME` value would be the `$MW_HOME/user_projects/domains/endeca_server_domain` directory.
`$ENDECA_HOME`	Indicates the absolute path to your Oracle Endeca Server home directory, which is the root directory for your Endeca Server installation.

Contacting Oracle Customer Support

Oracle Endeca Customer Support provides registered users with important information regarding Oracle Endeca software, implementation questions, product and solution help, as well as overall news and updates.

You can contact Oracle Endeca Customer Support through Oracle's Support portal, My Oracle Support at *https://support.oracle.com*.

Chapter 1

Upgrading to Oracle Endeca Server Version 7.7.0

This section provides instructions for upgrading to Oracle Endeca Server 7.7.0. Before you start the upgrade processes, check the remaining sections of this guide to learn about the changes that will affect you during or after an upgrade.

Required reading

Upgrading from 7.6.x to 7.7.0

About upgrading client stubs

Required reading

In addition to reading this document, it is recommended that you read the following documents for important information about the release.

Release Announcement

The Release Announcement outlines the new features that were added in Endeca Server version 7.7.0.

Release Notes

The Release Notes provide information about known issues and bug fixes for this release.

Installation Guide

The *Oracle Endeca Server Installation Guide* contains installation instructions, and information on how to verify your installation.

Upgrading from 7.6.x to 7.7.0

This topic describes how to upgrade from the 7.6.x version of Oracle Endeca Server to 7.7.0.

This procedure assumes that you have the following products installed:

- WebLogic Server 10.3.6
- Application Developer Framework Runtime
- Endeca Server 7.6.x

These three products will be removed. (Note that Endeca Server 7.7.0 doesn't require ADF Runtime, so it will not be reinstalled.) Before uninstalling them, it is recommended that you save any important data you have,

such as your Endeca Server data domains and any configuration you want to replicate in your new installation. Instructions on doing this are provided in the procedure below.

Note: If you are upgrading an Endeca Server cluster, follow the steps for uninstalling the Endeca Server instances on each machine, remove the Endeca Server WebLogic domain, and then remove the Endeca Server data domains.

To upgrade to Oracle Endeca Server 7.7.0 from Oracle Endeca Server 7.6.x:

1. Export your Endeca data domains from your WebLogic data domain according to the instructions in "Exporting and importing a data domain" in the *Oracle Endeca Server Administrator's Guide* (version 7.6.x).

   You'll be able to import these into your new installation.

   If your Endeca Server installation is earlier than 7.6.1.13 (that is, if you didn't apply the 7.6.1.13 hotfix), make a note of the data domain's profile name, as you will need this to import it. This isn't required for exporting a data domain from the version 7.6.1.13 and later. For more information, see *A .profile file created for new data domains*.

2. Make a note of any Endeca Server configuration you want to replicate in your new installation, including information in your properties files and your log level settings.

   Although you won't be able to reuse any of your 7.6.x configuration files directly, you can manually update your 7.7.0 installation with the information they contain.

3. Uninstall Endeca Server version 7.6.x according to the instructions in the *Oracle Endeca Server Installation Guide* (version 7.6.x).

   Use the appropriate uninstaller in the `$MW_HOME/EndecaServer7.6.x/oui/bin` directory: `runInstaller` for Linux or `setup.exe` for Windows.

4. Manually remove the WebLogic domain for Endeca Server according to the instructions in the *Oracle Endeca Server Installation Guide* (version 7.6.x).

5. Uninstall the Application Developer Framework Runtime.

   Note that this won't be reinstalled for Endeca Server 7.7.0.

6. Uninstall WebLogic Server 10.3.6.

   For instructions, see *Deinstalling Oracle WebLogic Server and Coherence*.

7. Install WebLogic Server 12.1.3 and Endeca Server 7.7.0 according to the instructions in the *Oracle Endeca Server Installation Guide* (version 7.7.0).

   You can install both products manually or with the orchestration script.

   As part of the installation, be sure to create a new WebLogic domain for Endeca Server 7.7.0. Additionally, if you had installed Endeca Server 7.6.x in SSL mode, do the same for the new version.

8. If you exported your Endeca Server 7.6.x data domain, import it into your new WebLogic domain according to the instructions in "Exporting and importing a data domain" in the *Oracle Endeca Server Administrator's Guide* (version 7.7.0).

   If you're importing a data domain from an Endeca Server version before version 7.6.1.13, you must provide the data domain profile name. This isn't required for version 7.6.1.13 and later.

To verify that the upgrade was successful, create an Endeca data domain (with no source data loaded).

About upgrading client stubs

In this release, each Web service packaged with the Oracle Endeca Server includes a version consisting of major and minor numbers; for example 3.0, where 3 is a major version number and 0 is a minor version number. Therefore, if you are planning to issue requests via stubs that utilize any of the Web services, you must generate new client stubs from the currently supported versions of each Web service.

Note: You only should generate new client stubs if you are planning to use direct requests to the Oracle Endeca Server utilizing any of its packaged Web services. Studio and Integrator use the supported versions of each Web service, therefore no special procedures apply in this case.

To obtain documentation on Web service topics:

- For information on changes that took place for each of the Web services, see the chapters in this guide.
- For information on versions of Web services supported in this release, see the topic in the *Oracle Endeca Server Installation Guide*.
- For information on how Web service versions are assigned and how to avoid version incompatibility, see the section on Web service versions in the *Oracle Endeca Server Developer's Guide*.

Chapter 2

Required Changes

This section describes required changes in the Oracle Endeca Server.

Platform support

Platform support

This release adds support for the following platforms:

- Windows Server 2012 R2
- Java 7 and 8
- Oracle WebLogic Server 12c (12.1.3)

This release also removes support for the following:

- Oracle Enterprise Linux 5 for x64
- Red Hat Enterprise Linux Server 5 for x64
- Red Hat Enterprise Linux Advanced Platform 5 for x64
- Windows Server 2008 R2 Enterprise
- Oracle WebLogic Server 11gR1 (10.3.6)

Chapter 3

Behavioral Changes

This section describes behavioral changes to Oracle Endeca Server and its interfaces.

Changes to the EndecaServer.properties file

Hostname resolution improvements

Security of administrative and configuration operations is improved

Improved Endeca Server startup

Improved handling of invalid records by the Bulk Load Interface

Change to the data domain's idle timeout

A .profile file created for new data domains

A new EQL function added

Sorting records based on geocode

Automatic log cleanup by the Cluster Coordinator

Changes to the EndecaServer.properties file

The `EndecaServer.properties` file sets global-wide parameters for Endeca Server, such as the default locations of mandatory files and directories. The file is located in the `$DOMAIN_HOME/config` directory.

The `EndecaServer.properties` file is automatically created when you create an Endeca Server domain in the WebLogic Server. If you are upgrading from the previous release of the Endeca Server, do not save or use this file from the previous release. Instead, use the new file that is available after the creation of the new WebLogic domain. For the most part, you will not be modifying this file. The file parameters are explained in the *Oracle Endeca Server Administrator's Guide*. If you change any parameters in this file, the Endeca Server should be previously stopped and then restarted.

Note: In the Endeca Server cluster, the configuration of this file for a specific data domain must be identical on all Endeca Server nodes that support this data domain.

In this release, the `EndecaServer.properties` file has the following new Endeca Server parameter:

Endeca Server parameter	Description
`endeca-disable-entity-validation-refresh`	(Optional). Specifies whether entity validation is enabled in the Endeca Server. If not present in the `EndecaServer.properties` file, validation happens (this is the default). Validation is skipped (disabled) only if you include this flag and set it to `true`. To change the value, edit the file and restart the Endeca Server.  **Note:** This flag is needed for performance improvement in cases when you are adding large numbers of entities to the data domain, during an update or a data loading phase. This flag should not be used otherwise.

In this release, the `EndecaServer.properties` file has the following new Cluster Coordinator parameters. Endeca Server uses these parameters to control how many snapshots of logs ZooKeeper should keep in the data directory, and how often ZooKeeper should run an automatic log cleanup:

Cluster Coordinator parameter	Description
`endeca-cluster-coordinator-auto-purge-snap-retain-count`	Specifies the number of log snapshots to retain in the Cluster Coordinator data directory. Is set to 3 (this is the default).
`endeca-cluster-coordinator-auto-purge-interval`	Specifies the interval for automatic deletion of log snapshots. Is set to 1 hour (this is the default). To disable the automatic deletion of ZooKeeper's log snapshots, set the task interval to 0 (zero) hours.

Hostname resolution improvements

In Endeca Server 7.7, either a Fully Qualified Domain Name (FQDN) or the first part of an FQDN is required for hostname resolution.

You can only use the first part of the FQDN if Endeca Server and its clients are running on the same domain network; otherwise, you must use the complete FQDN. This is used for TCP/IP communication between different Endeca Server nodes or between Endeca Server and its clients, such as instances of the Provisioning Service running on other nodes.

 **Note:** `localhost` is used in ES 7.7 internally for TCP/IP communication between Endeca Server and Dgraphs running on the same node.

Security of administrative and configuration operations is improved

In Endeca Server 7.7, most administrative and configuration operations require that you log in to the Endeca Server host and run the operations locally.

For additional information about the syntax of the operations, see the *Endeca Server Administrator's Guide*.

Improved Endeca Server startup

In the previous release, if you shut down the WebLogic Server before shutting down Endeca Server, and then restart the WebLogic, Endeca Server would start after a delay. In this release, if your Endeca Server deployment is on a single node, the Endeca Server starts without delay (in the cluster of more than one nodes, a short delay is still required).

Improved handling of invalid records by the Bulk Load Interface

Previously, if a batch of records sent for loading into the Dgraph contained a record with invalid XML characters, the Bulk Load Interface rejected the entire batch of records, including records with valid data. In this release, instead of rejecting the entire records batch, the Dgraph only rejects the individual records with invalid data and informs you why the record was rejected.

Change to the data domain's idle timeout

In this release, the default for the data domain's timeout value is 15 minutes (instead of 10 minutes in the previous release).

You can configure a data domain profile that allows you to automatically idle a data domain after a specified timeout, if the data domain does not receive queries during this timeout. The timeout allows for long-running queries to complete successfully before the data domain is turned idle. You can use the optional `idle-timeout` parameter to specify the timeout for automatically idling a data domain. If not specified, the default timeout of 15 minutes is used.

Also, if you create a data domain profile with the timeout that is less than 15 minutes, Endeca Server issues a warning and automatically resets it to 15 minutes, because 15 minutes is the lowest idle timeout value you can set. If you have any queries running longer than 15 minutes, increase the idle timeout, by creating a new data domain profile with the longer timeout.

For detailed information, see the *Oracle Endeca Server Administrator's Guide* and the *Endeca Server Cluster Guide*.

A .profile file created for new data domains

Starting with Endeca Server version 7.6.1.13, for each new data domain that you create, Endeca Server automatically creates an internal `.profile` file that is stored as part of the data domain's index. The `.profile` file stores the data domain profile information for each newly created data domain. This simplifies

the maintenance of data domains during their backup and restoration (exporting and importing), because you no longer need to keep track of a custom data domain profile so that you can use it when importing a previously exported data domain.

This means that when you back up and restore data domains, you only need to keep track of data domain's index files stored in an offline directory. You no longer are required to retain custom data domain's profiles. If you prefer to have a custom data domain profile file, you can still use it, but its creation and maintenance are now optional.

Here is the description of this change:

Each newly created data domain has a `.profile` file. If you update the data domain, its `.profile` is updated. If you export and import this data domain, the `.profile` file is used as the data domain's profile. If you delete the data domain, its `.profile` file is deleted.

The automatic creation of `.profile` file for each newly created data domain was added in Endeca Server version 7.6.1.13 (as part of the hotfix release for 7.6.1.13).

In Endeca Server releases before version 7.6.1.13, before you created a data domain, you needed to create and save a custom data domain profile. (Alternatively, the Endeca Server also included the default data domain profile which you could use, by specifying it, when creating a new data domain.) Next, if you exported this data domain's index with the purpose of restoring the data domain later (such as after an upgrade), you also had to keep track of the custom data domain profile file used to create this data domain. When importing a previously exported data domain, you had to specify the data domain's profile. If you didn't specify it, or didn't have this file in your system, the new data domain with the imported index would be created based on the default data domain profile (and not the custom profile that you might have had).

If you exported a data domain and its index file does not include `.profile`, then you still need to specify the custom data domain profile for it, when importing it. (This only affects data domains exported from Endeca Server before version 7.6.1.13). If you don't specify the custom data domain profile (or no longer have it stored on your system), the new data domain is created based on the default data domain profile.

Starting with the Endeca Server version 7.6.1.13, each data domain is created with the internal `.profile` file, as part of its index. The system relies on this file and automatically uses it when you export the data domain, and also when you later import it. If the data domain has the .profile file (you can check this in the index directory), then when you export it, this file is also exported and stored in the offline directory as part of the index. When you later import this data domain, you no longer need to specify a custom data domain profile; the system automatically relies on the `.profile` file to create a new data domain with the imported index, based on the data domain's profile stored in `.profile` file (from the exported data domain).

A new EQL function added

In this release, a new EQL function, `CountDistinctMembers`, is added.

The `CountDistinctMembers` function applies only to multi-assign attributes (these are managed attributes with multi-assign values). You can use this function instead of the `Cardinality` function. These two functions are equivalent, but `CountDistinctMembers` is faster.

For example, instead of using:

```
Cardinality(Set_Unions(multi-assign-attribute))
```

You can use:

```
 CountDistinctMembers(multi-assign-attribute)
```

For detailed information, see the *Oracle Endeca Server EQL Guide*.

Sorting records based on geocode

In this release, the Conversation Web Service lets you sort records that have geocode attribute values by the distance from a reference point you specify.

Geocode attributes represent latitude and longitude value pairs that can be assigned on your records. If the records in your result list have assignments on a geocode attribute, you can sort them by the distance to a geocode reference point that you specify.

You do this by adding the `Sort` type in a `RecordListConfig` configuration of your Conversation Web Service request:

```
<Sort Key="?" Direction="?">
<GeocodeReferencePoint latitude="?" longitude="?"/>
</Sort>
```

where `Key` is the name of the geocode attribute on which to sort. `Direction` (which is optional) is either `Ascending` for an ascending order (which is the default), or `Descending` for a descending order. The `GeocodeReferencePoint` element specifies the geocode coordinates as doubles.

Note: This feature is supported only if you use a Conversation Web Service request. It is not supported in Studio's user interface.

Geocode sort example

The following request uses the Location geocode attribute as the sort key:

```
<Request>
<Language>en</Language>
<State>
<Name>GeoQuery</Name>
<SelectionFilter Id="SelFlt">
<filterString>WineID > 10</filterString>
</SelectionFilter>
</State>
<RecordListConfig Id="Results" MaxPages="20">
<StateName>GeoQuery</StateName>
<RecordsPerPage>20</RecordsPerPage>
<Sort Key="Location" Direction="Ascending">
<GeocodeReferencePoint latitude="42.365615" longitude="-71.075647"/>
</Sort>
</RecordListConfig>
</Request>
```

For detailed information on sorting records, see the section "Sorting Records", and the topic "Geospatial sorting" in the *Endeca Server Developer's Guide*.

Automatic log cleanup by the Cluster Coordinator

In this release, the Cluster Coordinator has been updated to automatically clean up its logs.

The cleanup is enabled by default. For the logs cleanup, Endeca Server uses two new parameters in `EndecaServer.properties` to control how many snapshots of logs ZooKeeper should keep in the data directory, and how often ZooKeeper should run an automatic log cleanup:

- The `endeca-cluster-coordinator-auto-purge-snap-retain-count` sets the number of snapshot logs to retain in the ZooKeeper data directory, and is set to 3 by default.

- The `endeca-cluster-coordinator-auto-purge-interval` sets the logs cleaning task interval in hours, and is set to 1 by default. To disable the automatic logs cleanup for ZooKeeper, set it to 0 (zero).