

Embedding EMC Documentum Content Server OEM Edition

Best Practices Planning

Abstract

There are several things to consider when planning to embed EMC[®] Documentum[®] Content Server OEM Edition within an application or product: What are the recommendations from EMC Documentum for minimum hardware requirements? How can I install Documentum as a part of my product installation? An EMC Documentum repository can contain millions of objects and finding the information you need is a major challenge. What is the best way to set up search and full-text indexing? What planning should I do to avoid unnecessary customizations, both now and in the future? This white paper is intended to assist you with these questions and presents best practices for achieving high performance and availability with EMC Documentum Content Server OEM Edition.

March 2009

Copyright © 2009 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

H4659

Table of Contents

Executive summary	4
Introduction	4
Audience	4
Planning your system deployment.....	4
Sizing your Documentum deployment.....	5
Default set configuration for Content Server OEM Edition	6
Installing Documentum Content Server OEM Edition from a command line.	7
Installing Documentum as a part of your installer	8
Setting job schedules based on your needs	8
Managing search and indexing.....	9
Tailored for OEMs	9
Conclusion	9

Executive summary

Information, including structured data (stored in relational databases) and unstructured content (stored in file systems, content management stems, and more), is growing 60 percent to 200 percent per year and is increasing storage and handling costs and compliance issues, and limiting employee productivity. Since the majority of this information is unstructured, content management represents a critical technology to help address this explosive growth.

EMC® Documentum® Content Server OEM Edition's unified, comprehensive, and scalable platform allows easy collaboration, sharing, and reuse of content. It is the most open, extensible information infrastructure platform for managing information within the applications one develops, distributes, or hosts.

Content Server OEM Edition is built as an embeddable platform for rapid application development. It includes the content-related services and application integration interfaces, together with the developer resources and support services, that are essential for building content-rich applications. Content Server OEM Edition includes a wide range of developer services for developing third-party applications with the content platform. Application developers and ISVs have several integration options, depending on the granularity of their integration efforts and the level of abstraction for representing application-level objects.

OEM Edition is an integrated platform that combines Content Server, a relational database management system, and a full-text index system in a single package. It includes the core technology, content-related services, and the developer resources for building content-rich applications. Installing Documentum Content Server OEM Edition automatically deploys all components and configures a content repository. Applications that require a content repository can “silently” install Documentum Content Server OEM Edition as part of application deployment.

Introduction

This white paper is intended to provide detailed examples of some of the best practices to keep in mind when embedding Documentum Content Server within your application or solution

Audience

This paper is written for developers and IT staff who are considering how to embed Documentum Content Server within their products or applications.

Planning your system deployment

The process of planning your Documentum deployment should include a thorough analysis of your existing system, listing the resources required to implement your planned application, and determining what additional resources or information may be required for the various tasks.

Documentum depends on the proper installation and configuration of your network, server hosts, and server and client operating systems. Test your system resources and software to confirm that they are available and functioning properly before installing Documentum Content Server.

For example, Documentum clients cannot connect to Content Server without a correctly installed and configured TCP/IP LAN or WAN. Application integrations, which allow you to access Content Server directly from popular software applications such as Microsoft Word, cannot function unless the corresponding applications are installed on clients.

Before you install any Documentum products, test the installation and configuration of the following key components:

- **Server operating system**

To install and configure Content Server, create an installation owner account. The installation owner controls the administrator account and password. For example, on Windows, the installation owner must be a member of the local Administrators group. However, the installation owner should not be the same account as the Windows Administrator.

The installation owner account is known as dmadmin on UNIX and Documentum on Windows NT. (The installation owner account names differ because UNIX restricts the account name to no more than eight characters.) Only the user who owns the Documentum installation owner account can directly access the files in the file system. Many organizations require security checks before creating new user accounts. There may also be a backlog for fulfilling system administration requests, so obtain any system privileges you need in advance of installation day.

- **Network and computer hardware**

Your network must have sufficient capacity to support the hosts running Documentum server products as well as Documentum users and their computers. Each computer in the installation must meet the minimum requirements for running the Documentum products to be installed.

Documentum's architecture allows considerable flexibility in designing your server configuration. Refer to the *EMC Documentum Content Server OEM Edition Version 6.5 SP1 Installation and Configuration Guide* ("Setting Up the Development Environment" chapter) for more information on configuring your installation to meet your specific business needs.

- **TCP/IP network**

TCP/IP must be installed on the same computer as Content Server. The client computers must be able to communicate with the server using TCP/IP. Here are some things to check before you install Documentum on the host machine:

- Use ping or another simple test to make sure the client computers can communicate with the computer that will be the Content Server host machine.
- Confirm that your network can transfer large files from the server host to a client machine. For example, verify that it is possible to FTP a file from the Content Server host machine to a client machine.
- For each host machine in your Documentum 5 system, verify that you have sufficient disk space and RAM.

Sizing your Documentum deployment

System sizing is the process of determining what hardware, software, and network configurations will provide the best performance for users at the lowest cost to the enterprise. Another term for system sizing is capacity planning.

The following tables list nominal machine resources required for Content Server installation and use. Your individual machine requirements will vary depending on factors such as the number of products installed, size of your deployment, number of users, and network latency.

Hard Disk Space	RAM*	CPU	Display
700 MB	512 MB	See the Supported Processor Type column for your operating system	A video card capable of displaying a minimum of 256 colors (SVGA), or 800 x 600
80 MB	512 MB	See the Supported Processor Type column for your operating system	A video card capable of displaying a minimum of 256 colors (SVGA), or 800 x 600

* The amount of RAM that is available after taking into consideration all other RAM utilization requirements.

Desktop Machine Requirements

Product	Hard Disk Space*	RAM	CPU	Display
Desktop	300 MB*	256 MB	800 Mhz	800x600 at 256 colors minimum. Recommend 1024x768 at 256 colors

Default set configuration for Content Server OEM Edition

The Content Server OEM Edition is preconfigured with the following operating values:

- Accepts a maximum of 500 concurrent users
- Can index a maximum of 5K objects per hour
- Can store a maximum of 10M objects
- Uses the ACL associated with a user as the default ACL for objects created by the user (dm_server_config.default_acl=3)
- Enforces folder security (dm_docbase_config.folder_security = true)

After installation, it is possible to change the allowed number of concurrent users by setting the concurrent_session key in Content Server's server.ini file. Resetting that key requires restarting the server after the change to make the change effective. You cannot change the tuned values for indexing ingestion or maximum object storage in the repository.

Installing Documentum Content Server OEM Edition from a command line

You can install Documentum Content Server OEM Edition by running a command on the command line. The installation program creates the repository and installs all of the Documentum components.

Use the following steps for installation:

1. Log in to the host system using the installation owner account.

Note: Windows user accounts are not case-sensitive, but Documentum Content Server OEM Edition is case-sensitive, and the installation fails if you connect to the host with the incorrect case in the username. For example, if the account is set up as JPSmith and you connect as jpsmith, you can log in to the host, but installation fails.

2. Run the installation program from the directory into which you expanded the files. The format of the command line is:

```
oemwinSuiteSetup.exe [silent]
target
programDir="program_directory"
userDir="user_directory"
passwords
databaseAdmin=database_admin_password
docbaseOwner=password
windowsInstallOwner=installation_owner_password
globalRegistryUser=password
oemConfig
file="full_path_of_config_file"
```

The brackets around the `-silent` parameter indicate that it is optional.

When you include it (without the surrounding brackets), the installation program does not display feedback to the screen. When you omit it, an installation wizard appears to provide progress information. Follow the instructions in the wizard to complete the installation.

Use quotation marks to enclose any parameter values containing one or more spaces.

If you are connecting to an existing database, the `database_admin_password` must be the password for the user identified in the configuration file as `databaseAdminName`. If the Documentum Content Server OEM Edition installation program is creating a new database, it will set the password for the user "sa" to the password you enter. The installation program sets the passwords for `docbaseOwner` and `globalRegistryUser` to the values you enter on the command line.

Refer to the *EMC Documentum Content Server OEM Edition Installation and Configuration Guide* for complete installation steps.

Installing Documentum as a part of your installer

You can configure Documentum CS OEM Edition to be installed as a part of your installer. Documentum Content Server OEM Edition is provided to you in the form of a "silent" installation program. The installation program does not display a user interface for entering the necessary configuration parameters. Instead, the installation program retrieves the necessary parameters from a configuration file and from values passed on its command line.

When the Documentum Content Server OEM Edition installation program runs, it refers to an XML configuration file to retrieve values for a variety of configuration settings. The values in the configuration file configure Content Server, the index agent, Java method server, and repository. You can optionally install Documentum DocApps and Documentum DAR files in the repository by identifying the DocApps in the configuration file. When you do so, the installer silently invokes Documentum Application Installer (DAI) to deploy DocApps and Documentum Composer to deploy DAR files. You can also edit the configuration file to "hard code" the values for parameters that will be the same for all installations of your application.

For example, any DocApps or DAR files required by your application can be specified in the config file.

Refer to the *EMC Documentum Content Server OEM Edition Installation and Configuration Guide* for the format and schema of the XML Config file.

Setting job schedules based on your needs

Configuring a repository automatically installs a suite of administration jobs. These jobs are installed with default schedules and many, including all those that delete objects from the repository, are installed in the inactive state. For a list of the installed administration jobs, refer to the "Methods & Jobs" chapter of the *EMC Documentum Content Server Administration Guide*.

You may choose to activate or deactivate one or more jobs while installing the embedding application.

You may also choose to reset a job's schedule. Both of these operations can be performed by setting the appropriate attributes for a job. Job schedules are controlled by several attributes. If you are not using Documentum Administrator to set job schedules, then you must set the attributes directly. The attributes that control a job schedule are:

- start_date
- a_next_invocation
- run_mode
- run_interval

The starting date is the earliest date at which the job can be executed. The a_next_invocation property defines the first (or next) scheduled execution of the job. When you set the job's schedule, set start_date and a_next_invocation to the same value. Thereafter, when the job runs, a_next_invocation is automatically reset to the next invocation time. However, for the first run of the job, the value in a_next_invocation must match the value in start_date.

The `run_mode` and `run_interval` values work in conjunction to define how often the job is run after its first execution. The `run_mode` attribute defines a unit of measure. The integer value you specify in the `run_interval` attribute is interpreted according to the unit of measure you specify in the `run_mode` attribute.

Managing search and indexing

Installing Documentum Content Server OEM Edition installs an index agent and the Lucene full-text indexing engine and creates the full-text index. The full-text indexing process is described in the *EMC Documentum Content Server Fulltext Indexing System Installation and Administration Guide*.

Documentum Content Server OEM Edition installs Lucene Search Engine by default. The Lucene software consists of an indexer plug-in and a searcher servlet. Lucene extracts the indexable content from the content file synchronously and updates the index. Lucene can index objects that have multiple associated content files. However, clients cannot specify a query against a particular object. All text extracted from the associated content files is indexed together, not discretely.

Lucene does not support zone searching and does not support chunked XML documents. Lucene wild card searches also do not return matches when there is a space in the result. To optimize the indexing process Documentum CS OEM Edition comes with a default job called `dm_FTOptimizeIndex`. The optimization process merges all existing file segments into a new segment and removes index entries that are marked for deletion. This results in a more compact index and faster search speed.

You can tune the size of the results set returned by the searcher for queries by setting the `result_batch_size` attribute in the `ft engine config` object. The default value for this attribute is 1000.

Tailored for OEMs

EMC's offering for content management OEMs is a comprehensive set of technologies, tools, and programs. Documentum OEM Edition comes with OEM-specific documentation, architectural advice, and other useful tips, including information on how to embed Documentum into a host application.

The licensing structure for Documentum OEM Edition is flexible and is configured by the EMC Documentum OEM Sales team to align with the licensing model of the host application.

Conclusion

If you are designing enterprise software solutions and want to know more about EMC Documentum Content Server OEM Edition, visit EMC.com or www.emcinside.com, or call **800-607-9546** (outside the U.S.: +1-925-600-5802).