

# Designing a Multi-Tenant Content Management System with EMC Documentum

*A Detailed Review*



---

## **Abstract**

Companies can drastically reduce the cost of their content management services by sharing resources across user communities through multi-tenancy. This white paper will highlight potential benefits and outline pitfalls to avoid when implementing a multi-tenant solution. EMC<sup>®</sup> Documentum<sup>®</sup> provides a strong, flexible platform for developing a multi-tenant solution. It offers a variety of options to minimize the risks and maximize the potential gains associated with developing and deploying a multi-tenant content management solution.

October 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](http://EMC.com)

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

[H4735]

---

## Table of Contents

<b>About Blue Fish Development Group .....</b>	<b>4</b>
Company profile .....	4
Author biography .....	4
<b>About EMC partner-contributed papers.....</b>	<b>4</b>
<b>Executive summary .....</b>	<b>5</b>
<b>Introduction .....</b>	<b>5</b>
Audience .....	5
<b>What is multi-tenancy?.....</b>	<b>6</b>
Benefits of multi-tenant solutions .....	6
Allows for sharing of resources .....	6
Enables service providers to roll out the same solution to multiple clients quickly .....	6
Allows for isolation of data/environment between users .....	6
<b>Challenges of multi-tenant solutions .....</b>	<b>6</b>
Data segregation.....	7
Administration .....	7
Provisioning.....	7
Security .....	7
Scaling .....	7
Tenant personalization.....	8
<b>How to design a Documentum CMS solution for multi-tenancy.....</b>	<b>8</b>
Shared repository.....	8
Dedicated repository .....	8
Virtualization.....	9
<b>Conclusion .....</b>	<b>10</b>
Documentum: The industry-leading choice for MT content management.....	10

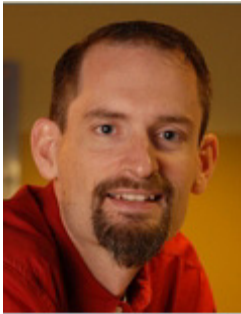


## About Blue Fish Development Group

### Company profile

Blue Fish Development Group helps Fortune 1000 companies plan, design, and implement enterprise content management (ECM) solutions. Based in Austin, Texas, Blue Fish has been delivering services and products that address the full spectrum of ECM challenges since 1999. Blue Fish can be found at <http://www.bluefishgroup.com/>.

### Author biography



Pete Nevin, a senior consultant in Blue Fish's Content Migrations practice, has the challenging goal of knowing more about our clients' data than they do. He recently joined the Blue Fish team after more than 15 years of consulting and engineering at such companies as Avaya and Groundwater Services.

## About EMC partner-contributed papers

Business professionals respect the opinions of their peers. In fact, for many of you, your first resource when researching an organizational purchase is someone who has used the product or service and has knowledge to share.

EMC® Documentum® partners have a wealth of expertise and innovative technologies to share so you can successfully address the needs of your information infrastructure. These partners are very familiar with EMC Documentum products and services and can help you achieve maximum flexibility from your EMC solution by offering helpful insights and tips.

The new TechFocus series of EMC partner-contributed white papers provides a forum for these knowledge leaders to offer information in a brief, unbiased format that will be educational for our partner, developer, and user audiences.

---

## Executive summary

In today's business climate, IT organizations are compelled to reduce the cost of solution delivery while at the same time offering users increased control and personalization. Some organizations deliver solutions to disparate internal groups through a shared services approach, tailoring features and functionality to a variety of internal needs. Other businesses are looking outside their organizations and finding IT solutions delivered by service providers that provide common solutions for multiple customers. A third characteristic of the industry is the need for greater information security, whether for customer and employee information or intellectual property. These three elements together have helped to create a demand for enterprise software applications that can be deployed quickly for a variety of clients without extensive setup. This type of solution is often referred to as a *multi-tenant* solution.

Many companies are seeking to reduce the cost associated with deploying and maintaining their enterprise content management systems by sharing resources among multiple groups of users. Since these groups will often have disparate requirements in a content management system, there is a trade-off to make between having a uniform platform to reduce costs, and offering custom features to meet each group's requirements. Multi-tenancy is a system architecture that seeks to meet both needs: cost reductions through extensive resource sharing as well as customizable functionality for each group.

Deploying an effective multi-tenant solution poses significant challenges. These include the following:

- Segregating and securing data between tenants
- Administering the differing groups within the application
- Providing an easy method for adding tenants
- Scaling to large numbers of users
- Offering custom behavior per tenant with minimal administrative effort

The EMC Documentum platform offers a number of options for designing a multi-tenant content management system that addresses the challenges listed above. An EMC Documentum solution can be designed to have multiple tenants share data within a single content repository (shared repository architecture), or the individual tenants may each have their own content repository (dedicated repository architecture). A third option is to have each tenant function within its own instance of EMC Documentum hosted on a virtualization platform like EMC VMware. The architectural flexibility offered by the EMC Documentum platform makes it the leading choice for designing and implementing multi-tenant content management systems.

## Introduction

This white paper describes multi-tenant applications including the benefits and challenges of such a solution. It includes a section on the three models that can be used in Documentum to design a content management system enabled by multi-tenancy.

## Audience

This paper is targeted toward IT executives and managers who are looking to reduce the costs associated with deploying and maintaining their content management systems.

---

## What is multi-tenancy?

In a multi-tenant (MT) application users and data are segregated into partitions (tenants) that are independent of each other. Tenants can be added or modified without the need for extensive software installation, and ideally are able to be provisioned with a minimum amount of configuration.

Multi-tenant applications offer the promise of substantial cost savings by making repeated application deployment less expensive than traditional point solutions and by allowing for sharing of computing resources across tenants. In many cases the MT application also allows for a good deal of customization between tenants in terms of functionality and branding, enabling a single service provider to support many differing clients.

We see two primary types of application providers leveraging multi-tenancy. The first are corporate IT groups providing applications to internal users in a shared services environment. The second are external organizations that provide software as a service (SaaS) services to their customers. These customers may range in size from small workgroups up to large corporations.

### ***Benefits of multi-tenant solutions***

Multi-tenant solutions provide answers to a lot of the demands of today's IT buyers. Multi-tenancy allows differing groups of users to share IT resources to reduce cost, while at the same time providing data security by isolating each tenant's data from the others. Let's take a look at how multi-tenancy provides these benefits.

#### **Allows for sharing of resources**

The chief cost benefit supplied by multi-tenancy occurs as a result of resource sharing across the tenants. Instead of deploying a set of servers for each workgroup, a MT application may support multiple different groups with one installation process, one application development phase, one set of servers, and one common support staff.

#### **Enables service providers to roll out the same solution to multiple clients quickly**

A MT application does not require a full hardware rollout each time a tenant is added. This means that the provisioning time for a new tenant can be much shorter than for traditional applications.

#### **Allows for isolation of data/environment between users**

A well-designed MT solution provides the ability to share resources while at the same time offering data isolation between user groups such as departments within a single company, or competitors sharing a common outsourcer. Data isolation is a key feature particularly in a SaaS environment where different tenants are from different companies and may even be competitors. Multi-tenancy provides a way for a SaaS service provider to isolate their customers so that the customers cannot view or modify each others' data. In a well-implemented MT system, tenants should not be aware that other tenants exist and also should not be aware that their computing resources are being shared.

## Challenges of multi-tenant solutions

Multi-tenant solutions offer a number of benefits to service providers trying to deploy a single solution to disparate customers with minimal cost. However, designing and building an effective multi-tenant application entails a number of significant challenges that need to be overcome.

---

## **Data segregation**

One of the primary challenges in deploying an application in a MT environment is designing and enforcing rules that properly segregate the data between tenants. The term “data” here is referring to all administrative and transactional data that may be associated with a tenant: documents, application properties, user accounts and preferences, for example. A multi-tenant solution needs to provide a failsafe mechanism for preventing users of one tenant from seeing other tenants’ data regardless of how the users are accessing data, whether by browsing or searching, or even when performing administrative tasks. The two primary ways of achieving data segregation are physically or logically. When data is segregated logically the data resides in common tables and isolation is managed by the application. With physical segregation tenant data is physically separate and stored in completely different data structures.

## **Administration**

A MT application needs to give consideration to the problem of how tenants administer their user accounts, business process settings, preferences, and other data. Some administrative tasks are necessarily managed by the owner of the application, but most data-related administrative tasks are delegated to users in each of the tenants themselves. Because MT applications can support large numbers of tenants, it is critical to have an administration design that allows users to easily understand the tasks needed to perform administrations functions without significant training.

## **Provisioning**

Having an efficient provisioning process for new tenants is a key consideration in a MT environment, particularly when the number of tenants is high. Creating a new tenant, importing users and data, applying configurations, and deploying customizations are all aspects of provisioning. For shared services deployments where the number of tenants is in the dozens or hundreds, it may be possible to justify a process containing manual steps. Once the number of tenants numbers into the thousands it is imperative to have a very streamlined provisioning process. For SaaS it is often necessary to have a process that can be performed by the tenant themselves via a self-service web portal.

## **Security**

Designing a security model to accommodate a MT application can be very challenging. Disparate requirements from multiple tenants place flexibility demands on a security architecture. Furthermore, a successful security architecture needs to be able to provide isolation between tenants’ data and still be manageable by tenant administrators who may not have much training. It is a difficult task to find a balance between these competing objectives.

## **Scaling**

Even more than other types of IT solutions, MT applications need to consider scale. Growth can occur as a result of adding tenants, or because of growth in the tenants themselves. Solutions that work for hundreds of large tenants may not work well for a similarly-sized deployment that has thousands of small tenants.

Each tenant’s infrastructure footprint typically depends on some fixed amount of resources needed for the tenant, plus whatever is needed to handle the number of users and the amount of data for that tenant. The smaller the footprint that is required for each tenant alone, the better that solution scales to handle the large numbers of tenants needed for a typical SaaS deployment.

---

## ***Tenant personalization***

In addition to being a way to segregate data multi-tenancy can be used by a service provider to provide different business rules, different client experiences, and even differing security models to different users of the same application. This ability to configure differing tenants with different characteristics is called personalization. The degree to which an application provides rich options for personalization can make a single application useful for a large number of potential clients. However, personalization options can require significant development effort and lead to much more complicated provisioning and administrative processes.

## **How to design a Documentum CMS solution for multi-tenancy**

EMC's Documentum platform provides a number of useful tools to build and deploy a CMS with MT capabilities. First, Documentum is built around a flexible hierarchical object model. This provides a simple mechanism for providing data segregation as well a flexible tool for matching differing client's design needs. Documentum also has a powerful security model that uses aliases to apply a consistent security model across multiple tenants. Finally, Documentum also supports an array of architectures that allows it to scale up to handle very large sites, as well as multiple sites situated across the globe.

There are three primary models that can be used in Documentum to build a MT-enabled CMS: the shared repository model, the dedicated repository model, and the virtualization model. We will compare these three options as they relate to the challenges noted above and determine which architectures work best for different types of applications.

### ***Shared repository***

The first option for developing a MT solution in Documentum is the use of a shared repository. As the name implies this model involves having some or all tenants sharing the same data repository. This shared repository model employs the smallest resource overhead per tenant. As a result, it is generally the best choice for large numbers of tenants that have relatively few users per tenant. Some current EMC customers have SaaS deployments based upon the shared repository model that services thousands of tenants.

In the shared repository model data segregation occurs by partitioning within the same database. Since there is no physical segregation, logical segregation needs to be achieved using the Documentum object model and application customizations. Because this partitioning of data, users, and business logic must occur within the application layer, this model likely requires the largest amount of custom application development. As a result, there are some important design considerations to follow to ensure proper data security. First, using tenant-specific extensions to Documentum's base object model is a very good way to provide data segregation. Each object type has its data stored in a separate table and queries can be developed to run against those specific types and their subtypes only. A second consideration is to use the DFS web services to ensure that all client applications are enforcing the same data segregation rules and business logic across tenants. EMC has built the DFS platform to incorporate MT concepts, so making use of these web services for all content server interactions creates a "choke-point" to prevent leakage of data across tenants.

### ***Dedicated repository***

The second MT architecture model we will consider is the dedicated repository model. In this model a separate content repository is created for each tenant housing all tenant-specific data, while other components of the solution — storage infrastructure, client application tier, and so on — are shared. This

model requires relatively few application customizations, because the separate repositories mean that data segregation is achieved without any additional effort.

The dedicated repository model does require additional resources per tenant above and beyond those found in the shared repository model. A dedicated content server process is required for each repository and therefore for each tenant. This provides additional data security and isolation at the expense of having less resource sharing than with the shared repository model. Deploying these components also adds to the effort required when provisioning the tenant. As a result, this model is probably most appropriate for deployments with fewer, relatively larger tenants, or those requiring a greater degree of flexibility and security.

## Virtualization

Virtualization is the third option for designing MT applications in Documentum, and this option has been gaining in acceptance and popularity as of late. This model creates an entirely separate instance of the Documentum services for each tenant on one or more virtual machines (VM) and offers complete isolation of all data and services. Advances in virtualization technology, combined with increasingly inexpensive computing resources, have made this option increasingly attractive for many deployments. Virtualization affords tremendous flexibility, requires minimal application customization, and still offers excellent utilization of shared computing, storage, and support resources.

The virtualization architecture lends itself well to handling the challenges of data segregation and security. This option also allows for a great deal of personalization per tenant. Different tenants are hosted on completely separate instances of the Documentum platform.

Deploying an instance of the Documentum services for each tenant obviously makes this option have the highest resource overhead per tenant of the three. However, having everything hosted on a virtual machine allows for very easy and flexible options for high availability and disaster recovery that are not possible with shared machines. Perhaps the greatest advantage of the virtualization approach is that it provides flexible options for handling tenants as they grow. If a particular VM is handling an increasing load, it is relatively simple to move it to a larger capacity server or move it to a dedicated hosting service.

Documentum provides several choices architecting a solution that best fits your MT needs. Table 1 provides a snapshot of some of the distinctions between these options to help you decide which is best for your application.

**Table 1. Comparison of multi-tenant deployment options**

	<b>Shared repository</b>	<b>Dedicated repository</b>	<b>Virtualization</b>
<b>Application customizations required for deployment</b>	Significant application modifications needed	Can be designed with minimal application changes	Maximum flexibility. Minimal changes needed
<b>Tenant overhead</b>	Minimal	Moderate	Significant
<b>Typical deployment scenario</b>	Large deployment with many, small tenants	Moderate to large deployments where the number of tenants is not too large	Used when expected growth is high or unknown

	<b>Shared repository</b>	<b>Dedicated repository</b>	<b>Virtualization</b>
<b>Representative tenant size</b>	Appropriate for tenants as small as 10 users. Can scale up to tens of thousands of tenants	Appropriate for tenants as small as 100 users. Scales up to hundreds of tenants	Appropriate for systems having tenants with hundreds of users or more, and a few hundred tenants

## Conclusion

### ***Documentum: The industry-leading choice for MT content management***

We have seen that Documentum can be used effectively as a platform for developing and delivering multi-tenant capable solutions, whether in a shared services or SaaS environment. Documentum offers a number of architectural options that allow it to be tailored and optimized to your unique needs. Perhaps the most important advantage offered by the Documentum platform is its ability to scale. Documentum is designed to support a varied array of content server topologies that support very large numbers of geographically scattered users. Additionally, it has a powerful combination of a robust object model, flexible security, numerous prebuilt client applications, and an MT-aware web services development framework that makes it an ideal choice.

In addition to the capabilities available in Documentum today, EMC has put together a roadmap that incorporates key MT concepts in upcoming releases. Future releases of the Documentum platform promise to significantly decrease the overhead associated with the dedicated repository model, making it applicable for deployments with a much larger number of tenants. Additionally, all the standard Documentum client platforms—CenterStage<sup>®</sup>, TaskSpace, Media WorkSpace, Webtop, and WDK—are undergoing significant changes to improve their multi-tenant capabilities. So whether you are considering current capabilities, or looking forward at future needs, Documentum delivers a uniquely flexible and powerful platform to create solutions that support a wide variety of multi-tenant applications.