Oracle Universal Content Management

Fenton Travers
fenton.travers@oracle.com

September 17, 2010

Contents

List of Figures

0.1 Content Management Overview

0.1.1 Introduction

When people hear the word 'Content Management' a common reaction is: "Oh no not another buzz word!" Then they mentally switch off, freak out or have some other reaction that doesn't end in understanding. The purpose of this document is to change that. Understanding what content management is, is very easy. It is the combination of two concepts you probably already know about. Content Management at it's essence is the combination of a file system with a database table. A file system is just the place where you have been saving your documents for years, all those files and folders on your C: drive!

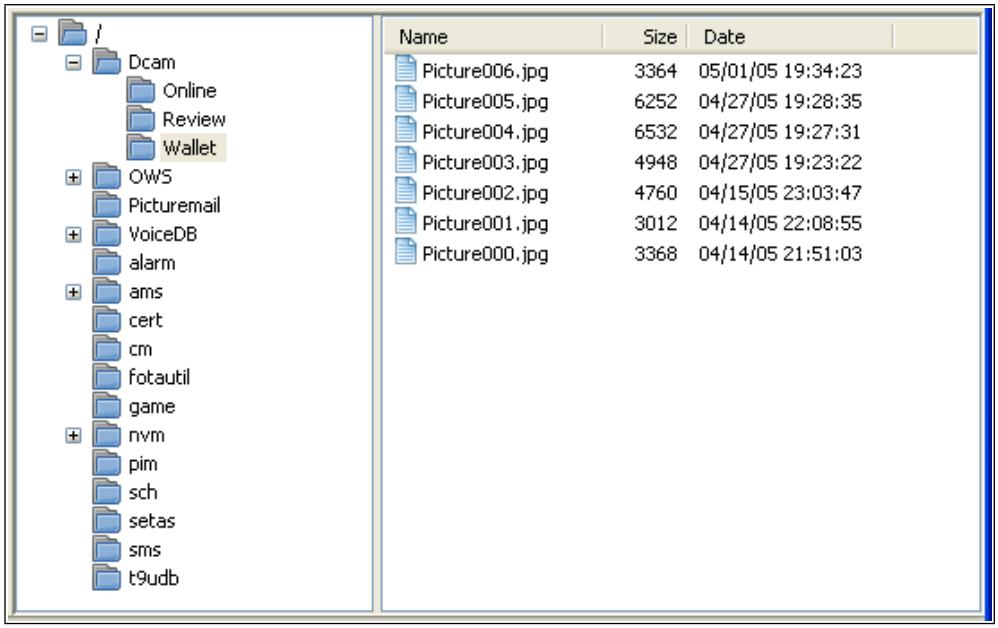

Figure 1: The concept of a file system: files and folders

The concept of a database table is really just like a spreadsheet, with rows and columns.

Don't look at these images too closely, just keep the concepts in mind. So how do these two concepts form the basis of the Content Management concept? Basically when you are talking about content management there two parts. The first part is the actual file, or piece of content. The second part is

	A	B	C	D	E	F
1						
2						
3	**Date**	**Start time**	**End time**	**Pause**	**Sum**	**Comment**
4	2007-05-07	9,25	10,25	0	1	Task 1
5	2007-05-07	10,75	12,50	0	1,75	Task 1
6	2007-05-07	18,00	19,00	0	1	Task 2
7	2007-05-08	9,25	10,25	0	1	Task 2
8	2007-05-08	14,50	15,50	0	1	Task 3
9	2007-05-08	8,75	9,25	0	0,5	Task 3
10	2007-05-14	21,75	22,25	0	0,5	Task 3
11	2007-05-14	22,50	23,00	0	0,5	Task 3
12	2007-05-15	11,75	12,75	0	1	Task 3
13						
14						
15						
16						
17						
18						

Figure 2: Information stored in rows and columns is a natural to keep information sorted.

the information about that piece of content. So what is common information people keep about a piece of content. Well some obvious information to store about a document might be:

- The author of the document
- The size of the document
- The type of document: word, pdf, spreadsheet, power-point

However, we might want to store even more information about a piece of content. We might want to say WHO can access the document. This is often a very important piece of information about the document. Think about your employement contract... you wouldn't want all your colleagues to know how much you make would you?

So up to now we are really talking about two kinds of information. The information inside the document and the information *ABOUT* the document. People often use the word *Content* to refer to the information that is *inside*

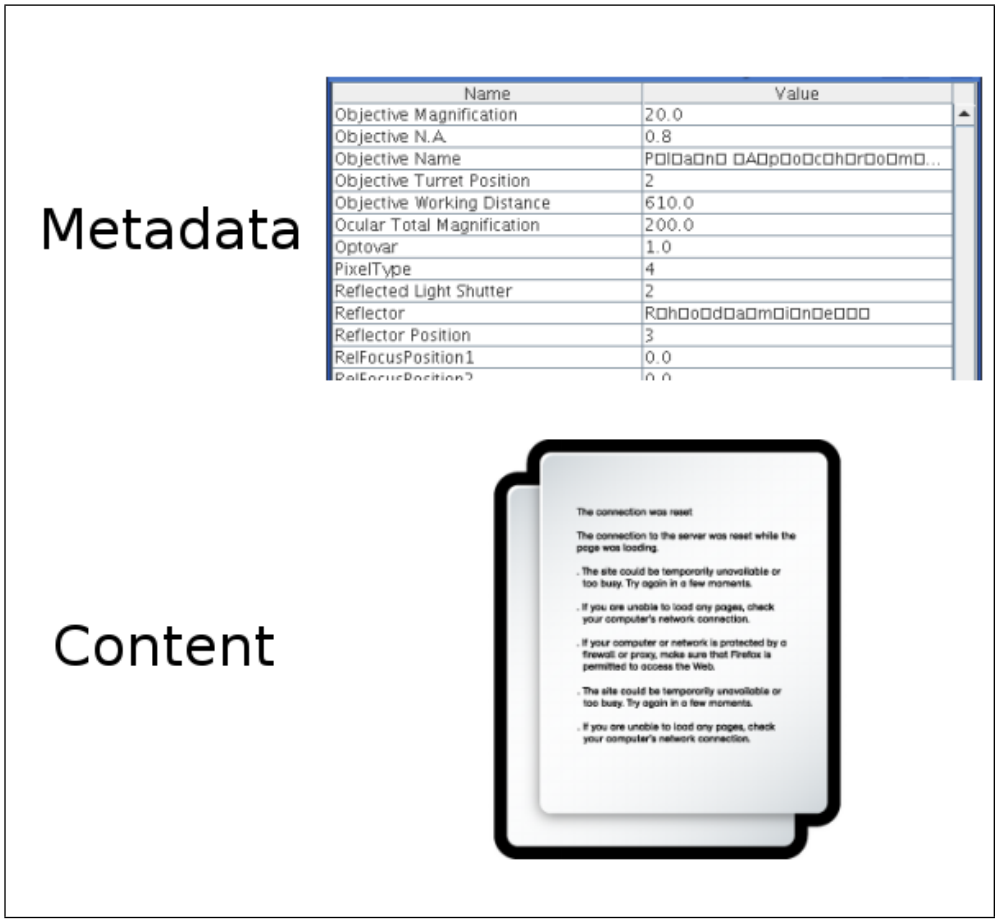

Figure 3: The concept of a file system: files and folders

the document and *Metadata* to refer to the information *about* the document. Lets look at an picture to help us conceptualize this.

So that is the essence of Content Management. You have content, a document, file, etc... and you have information about that document, metadata.

0.2 Web Content Management

0.2.1 Overview

What is Web Content Management (WCM)? The essence of WCM is that it allows non-technical people to rapidly edit content on a web site. Traditionally, a web site needs highly skilled HTML web *masters* to update the web site. However, these web master, after they have built the web site, tend to move on to building other web sites. Without these people available to keep the web site relevant, the web site information tends to become stale and out dated.

WCM seeks to address this problem. It allows web masters to build a WCM site that can be *maintained* by non-technical users on an ongoing basis. This is achieved by providing the non-technical users familiar WYSIWIG (What You See Is What You Get) editors that look much like familiar word-processing tools.

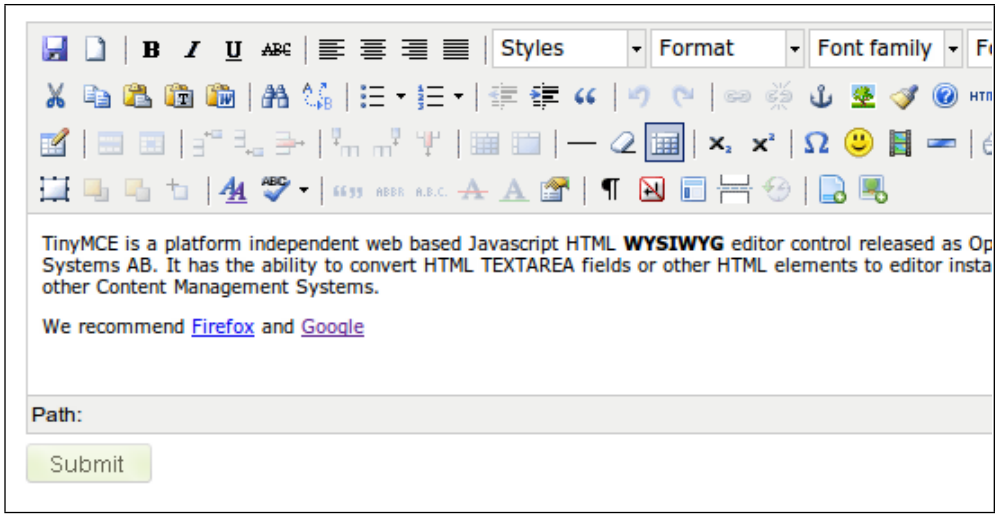

Figure 4: WYSIWYG Editor

0.2.2 How does this work in practice?

In a day to day usage scenario, a business user would decide he wanted to change some content on a web page. They would navigate to that page and

press a special set of keys and the web page then displays little pencil icons everywhere the user is allowed to modify content. They simply click the pencil icon and then they are presented with the *WYSIWIG* like the one shown above, and then they can edit the content. Once they save the section of the web page they edited, the content is either automatically updated on the production site, or alternatively it may be subjected to a workflow approval process.

0.2.3 The changed role of the Web Master

3 Obviously for all this 'magic' to occur, there needs to be something in-place other than plain old static HTML content. The basic architecture is that 90Universal Content Management (UCM) system. Additionally you place a Web Server in front of UCM to send the HTML the UCM system generates to the client browser. The following diagram lays out the basic architecture of an Oracle WCM site.

Now the web master creates special 'Site Studio' web sites, using the tools provided by Oracle WCM, and finally we have a WebContent Management website, versus some other type of web site.

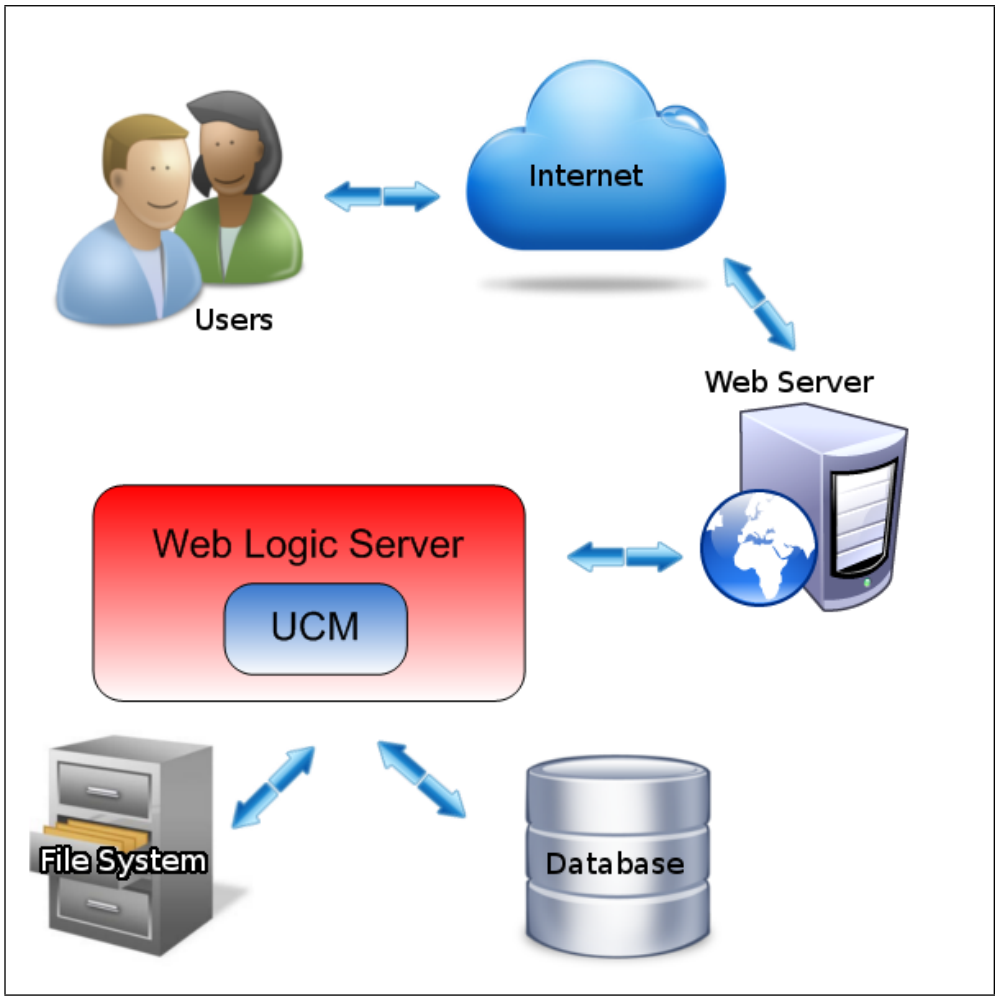

Figure 5: UCM Basic Architecture

0.3 Content Management to Virtual Machines

I'm going to segue into a discussion about virtual machines now. Although this conversation could be considered very tangential to previous content management system (CMS) topics, the overall motivation for the information provided here, is to give you the basis for actually implementing and using the entire architecture that is presented.

To do this on the small budget I'm assuming you have, it is necessary to leverage Virtual Machine technology. So lets get that topic under our belt.

0.4 Virtualization Overview

0.4.1 Motivation

I can hardly say enough about virtualization. It's pervasiveness and impact on the IT world can hardly be overstated. Also, IT environments are getting more and more complicated and having a wholistic understanding of the whole topology begins to get too complicated. We have experts in fields of a given domain who have trouble communicating with each other. It's my assertion that every member of the IT team should be able to install all the key pieces of their organizations infrastruture so they can test the effects of making changes etc. When you use a Virtual infrastructure to do this, you can set it up and tear it down without worring about having to buy a whole lot of hardware.

Organization that can understand the value of this will make huge strides in taming the wild animal of their infrastructure and begin to create much better infrastructures. So lets roll up our sleeves and come to understand about the types of virtualization so we can get busy using it.

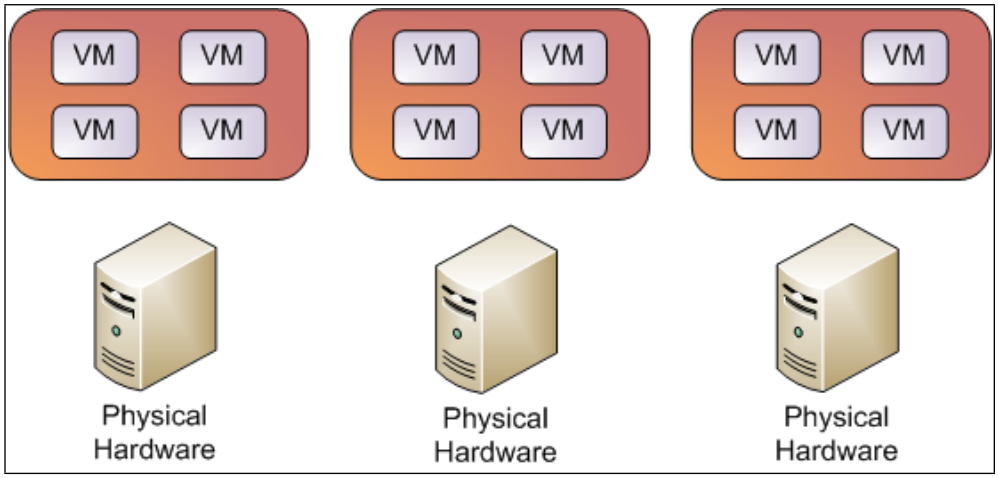

Figure 6: Virtual Machines

0.4.2 Overview

Virtualization comes in two primary forms. The type that runs on your personal laptop, that Sales people take out to demonstrate software to cus-

tomers, and the type that runs in the back office data centers...the 'real' virtualization.

These two type have names. The one you run on your lap top is called 'Fully Virtualized' virtualization. The one that runs in the data centers is called 'Para-Virtualized'

If you have some familiarity with virtualization you'll probably be familiar with fully hosted virtualization. Products like virtual-box, VM Ware Workstation (Player), etc... fall into this camp. Datacenter, para-virtualized (PV), products are VM Ware ESXi, and XEN. There are many many more but we'll focus mainly on Oracle Virtualization which is based on XEN.

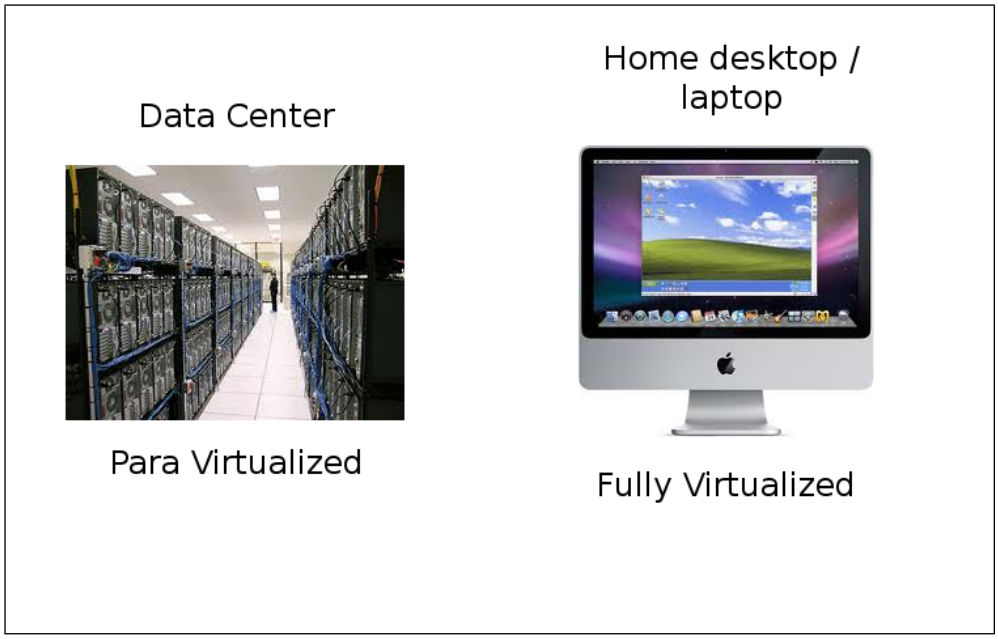

Figure 7: Para versus Fully virtualized

0.4.3 Whats appropriate for you?

When deciding which route you want to go, you will for sure want to become familiar with fully virtualized solutions, such as Virtual Box. The reason for this is that it is simple to setup and use on the operating system you currently have. They will all support Windows, Linux or Mac OS's.

I would highly recommend that you also try to do PV virtualization as well, since if you want to run this in production that is the way you would go. You need to have dedicated hardware for this. A spare laptop will do, and having the maximum amount of RAM is also highly recommended.

0.5 Architecture Version 1

We are going to work on several versions of a basic architecture. In the future we may cluster a given virtual machine, we may add a security layer, we may add some user interface layers, etc. But for now try to understand the components of this basic architecture.

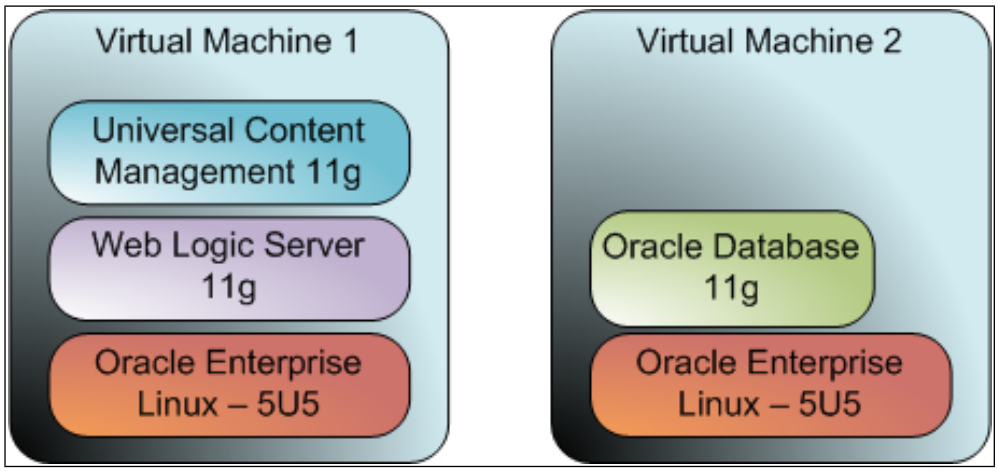

Figure 8: Basic Architecture

As you can see this is a two Virtual Machine system. This can live on either one or two physical machines. The RAM foot print is:

- Virtual Machine 1: (OEL5U5, WLS11g & UCM11g) - 1.5 GigaBytes
- Virtual Machine 2: (OEL5U5 & ODb11g) - 1 GigaByte

We'll supply 10 GB disk space for each of the images.