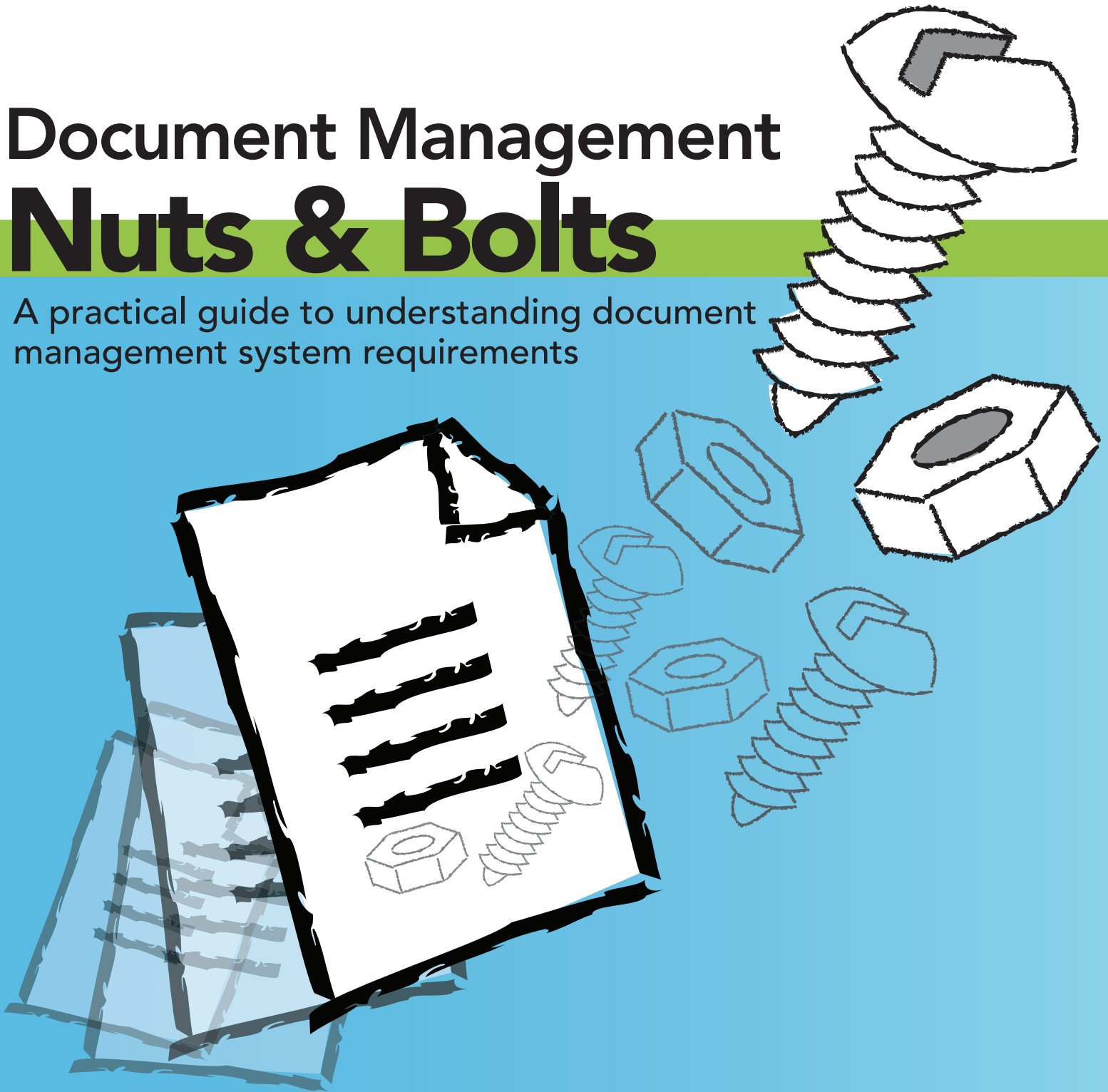


Document Management **Nuts & Bolts**

A practical guide to understanding document management system requirements



Contents

Part 1: Client Software Requirements

How your users will interact with and be accommodated by a document management system.

Part 2: Application Server Software

The brains of the document management system you install on the server and different ways to distribute the components.

Part 3: Vaulting

How documents get in and out of the document management system—and stay secure.

Part 4: System Hardware Requirements & Considerations

Make the most of your existing hardware and things to consider when investing in new hardware.

Part 5: Optimizing Performance

Be aware of factors that can positively affect or negatively impact the performance of your document management system.

About the Author

Todd Cummings is the Vice President of Research and Development for Synergis Software, the developer of Adept Engineering Document Management software. Todd has directed R&D at Synergis Software for over 14 years and is passionate about the benefits of document management for solving the most pressing business challenges for engineering companies. Todd is an expert in all areas of document management including: Needs analysis and ROI; hands-on troubleshooting and problem solving; managing implementation projects; designing and writing product requirements; and communicating complex ideas in a easy-to-understand manner to all levels of customers. Todd has been a guest speaker at Autodesk University, SolidWorks World, ASME, TAPPI, and the United State Coast Guard CADD and GIS Symposium and has also given hundreds of technical presentations to SMBs and Fortune 100 companies.



Introduction

This white paper intends to help business and high-level technical evaluators understand the breadth and scope of the components that make up an engineering document management system—whether it's on a LAN, a WAN, or the Internet.

Because engineering document management is a mission critical application that touches many people in an organization, it's my goal to give you a solid foundation in the high-level technical requirements and business considerations of a document management system as early as possible in your research/evaluation process. That way you are 1) better prepared to do a great job asking questions and comparing vendor products; and 2) you are on track for a truly successful implementation, knowing how all the pieces of the puzzle fit together.

Whether you are team of one or many, whose job it is to understand all the interacting parts of a document management system, this paper is a must read. For those of you who are new to document management, this material will help you connect your business knowledge with the technical requirements and give you the vocabulary you need for further exploration.

So, let's get started!

Part 1: Client Software Requirements

Client software requirements define how users interact with the system.

Data Creators vs. Data Consumers

To understand fully the ways users interact with a document management system, we need to appreciate the differences between data creators and data consumers. When it comes to client requirements, you'll need to consider both groups.

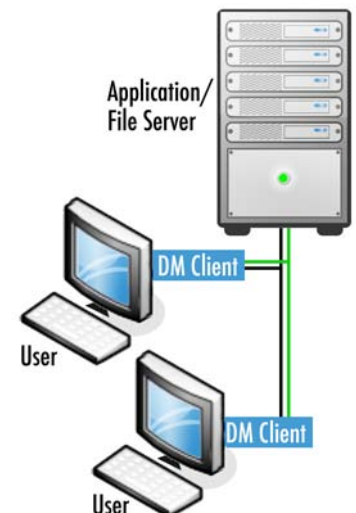
In a design engineering-centric world, data creators are those users who work inside the CAD applications. They create and revise existing CAD documents. In today's global organizations, data creators are certainly designers and can be engineers or project leads responsible for specifications, analysis, or project and product success. Creators contribute data and documents to the system. They are also your power users and require a more powerful workstation.

Data consumers are the people that need to work with the data submitted by the data creators. Data consumers have different needs and often access the system using a different interface than the data creators. Generally, data consumers access the system to view, print or redline via a streamlined user interface and then, often only part time.

Most document management deployments are accomplished using more than one type of client. It's rare that a company deploys just one type of user interface to access the system. So let's look at some of the different client deployment methodologies.

Client Workstation Deployment

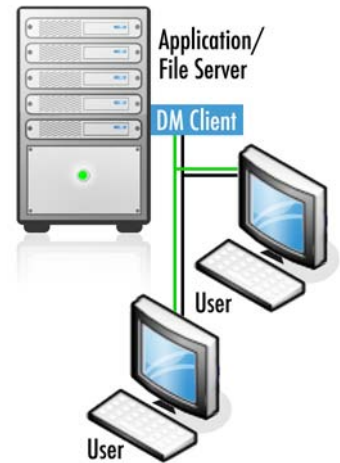
The first client we'll discuss is the client that is installed on a local workstation. It is a classic install where the actual program files – such as Word or Excel or Outlook – are installed on each individual's workstation, as seen in the figure to the right.



Network Client Deployment

A network client deployment is where the document management client is installed on the server in one location. Users in this method use a desktop shortcut to launch the application (see figure on right).

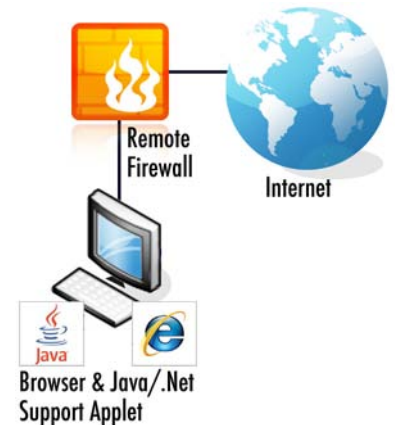
Generally, IT appreciates this lightweight deployment because they can install the application in one location and serve updates out to everyone easily. The trade-off is that user performance might take a hit. It's also important to keep in mind not all document management systems support a network deployment well.



Client Browser Deployment

For many companies today, it's becoming more important to deploy a client through a browser (see figure on right). Browser access can mean access via the Internet but many also deploy a browser client behind the firewall.

These deployments are relatively thin and from an IT perspective should be close to no-touch. For example, if a person can open a browser and point to a URL and the application of the data housed in it is immediately available.



Remote Client Options

A remote client deployment is yet another option. This entails Citrix, terminal services, or similar technologies, which are popular, centralized methods for IT to deploy access to a document management system. The important thing to understand when using this method is that not all document management systems are friendly to this environment. This method may also require that the CAD applications are deployed on the central server and not all CAD applications support this. So it's necessary to identify up front if this is a requirement for your organization and whether your chosen document management system supports it.

Part 2: Application Server Software

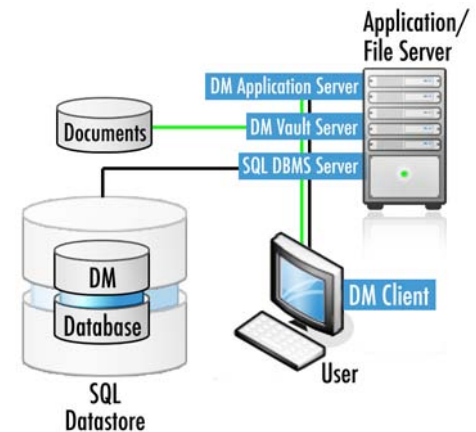
In the last section, I discussed the nuts and bolts of a document management solution as it relates to client workstation deployment. In this section, I'll talk about the server-side components, application server software, and the database engine components of the system.

Application Server Software

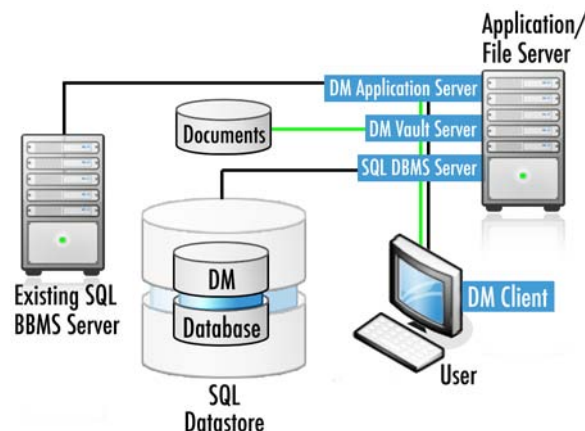
The application server software is the silent piece of the document management system that users don't see—but IT does. It's the brains of the document management solution installed on the server.

You can think of the application server software as an information broker. It's what every user connects to and it facilitates data exchange between each user, the database, and the document vaults.

There are a couple of ways you can deploy the application server software. The figure on the right illustrates a deployment where all the server side components are installed on one server, including the application server software, the vault server software, and the database engine.



The figure below illustrates another type of deployment where the application server software is installed on a different server than the database engine. Flexible and scalable document management systems support this modular deployment strategy.



In fact, often there are more server side components, which may be distributed across multiple servers. Architectural flexibility in this regard is solution dependent. In small-to medium-size deployments, I like to recommend keeping all the application server software components on one server if possible. In larger organizations where user and document counts are higher, multiple servers may be required.

When the database and the application server software are on different servers, make sure you have at least a gigabit ethernet connection between them. Otherwise, you are likely to encounter a performance bottleneck.

The Database Engine

There are three different categories of database engines that document management systems support: Well-known name brands like Microsoft SQL Server and Oracle; less familiar, yet commercially available databases; and proprietary databases developed and supported by the document management vendor.

I am often asked, “Which database is best?” In my mind, the most important question is not “Which database is best” rather, “Which database will best meet my company’s business needs?” If your company has a database standard that IT supports, then the decision is made. If your company has no mandated database engine, then you’ll need to select one.

If you’re a small company without consistent access to database experts, I don’t recommend choosing a “high-end” database solution for your implementation. If you do have DBA expertise in house, then a “high end” solution could be a good choice.

Beyond the basic criteria of price, performance and scalability, the main driver in today’s interconnected business environment is access to your data—whether it be through open database connectivity protocols like ODBC or vendor-provided APIs. You also want the document management systems’ schema to be open and accessible. Your data is your most critical knowledge asset—so make sure you can access it.

As a final thought, I encourage you to understand database engine licensing requirements, as they can be difficult to decipher. Most database vendors

offer varied licensing options that can make a significant difference in your company's overall investment. Take time to fully understand the options before making your final selection. In many cases, database licensing is un-metered and compliancy depends on your understanding of the licensing requirements and the honor system.

Part 3: Vaulting

Now we arrive at another critical document management component, the vault. Although it is also a server side component, it deserves discussion on its own. In everyday language, a “vault” conjures images of a steel-reinforced concrete room where banks store their cash. Ideally, your corporate electronic assets, the documents, are protected from unauthorized user access and secure. Vaulting, however, also explains how documents are moved in and out of that specified ‘vaulted’ location. Simply stated, vaulting manages the way documents get in and out of the document management system as well as where and how they are stored, protected, and secured.

Common Types of Vaults

Often, I am asked, “Which type of vault is best?” I believe the answer is self-evident once you truly understand the type of vaulting mechanism used by your document management system of choice. Once you understand this, you will be answering the question “Which type of vaulting mechanism is best for my company?”

A wraparound vault literally wraps around your existing file structure without changing the folder or file names it enfolds. It also secures all content from unauthorized “back door” access through Windows Explorer. Many companies prefer this type of vault because it doesn’t scramble file names or move file locations, which makes it easy to access and recover documents in human readable format in the event of a network failure. This type of fault tolerance is often popular with technology and business owners alike.

A proprietary or a scrambled vault copies the documents from your existing folder structure and moves them to a new location. This method of vaulting documents also scrambles or hashes the original filename, and more often than not, the folder names. Scrambling or hashing a file name or location is not the same as file encryption. When data is encrypted, an algorithm is used to make it unreadable to the user and to its editing application – until a file is decrypted.

Here’s an example. When you place the file “my-document.doc” into a proprietary vault, that document is given a different name. It is not encrypted, but rather the name and the folder in which it resides is named something that a user won’t easily recognize, for instance “PQ_123.ZFW”. The same goes

for file folders. In a proprietary or scrambled vault, the original folder name, let's say, "Acme Corporation" will be moved into a folder with a name that is undecipherable to the user.

An overlay vault sits on top of an existing file structure, but provides passive security—meaning it doesn't really protect documents from "back door" Windows Explorer access.

A blob vault places documents into the document management database in fields, within the database schema. In this way, a blob vault can grow your database exponentially and cause database backups to take significantly longer.

With any of these vault types, the size of the vaulted data, meaning the actual number of documents and their cumulative size in gigabytes, or these days even terabytes—is important to keep in mind. If you are managing 100K Word files, then a solution using a blob vault may be appropriate for your needs. However, if you have large numbers of CAD documents that are hundreds of megabytes or gigabytes in size, then the blob vault could hinder system performance.

Vault Attributes

All these vault types can have different attributes, which modify or enhance their usefulness in your organization. Common vault attributes are multiple or remote vaults; hybrid vaults; and replicated/synchronized vaults.

A multiple or remote vault system can accommodate multiple document stores across multiple servers in different geographic locations.

A hybrid vault gives some users access to documents through Windows Explorer (e.g., in a read-only mode) while others have limited or no access, except through the document management system.

A replicated/synchronized vault duplicates documents in remote locations and keeps document revisions synchronized across all locations. This attribute can eliminate the on demand need to transfer large amounts of data across the WAN during peak usage times.

And the Best Vault is...

As we discussed earlier, the most relevant question for you is, “What’s the best vault for my company?” Now that you understand different common approaches to document management vaulting methodologies, you may use these questions to help you further clarify the solution that is best for you:

- Which vaulting method best matches your company’s business goals and its technology requirements?
- What are your backup strategies?
- What are the replication strategies for each type of vault?
- What are your company’s uptime requirements and fault tolerance plans?

Keep these criteria in mind as well:

- Speed / document access
- Administrative / IT overhead
- Additional system resources / requirements
- Document transport methodologies
- Company security policies
- Ease of maintenance

Ultimately, the best vaulting method and attributes will depend on your specific environment.

Part 4: System Hardware Requirements & Considerations

Getting the Hardware Right

Now that we have discussed the different high-level aspects of a document management system—it's time to tie these concepts together and consider the hardware required to meet your business and deployment goals. The hardware you need to support your document management implementation will depend on several variables:

- Are your users in one location or distributed across multiple geographies?
- Does your company have a standard, approved database technology? Are there existing database instances already housed on servers that your document management system must utilize?
- Does the document management system have specific standard or nonstandard hardware requirements—whether it be on a LAN, across the WAN, or even the Internet?
- Do you have business justification to dedicate hardware to your document management system?
- Does the system support, and are there benefits to distributing server components across multiple pieces of hardware?
- Does your company currently have a server virtualization initiative underway; and does your chosen document management solution support virtual servers in production?

Once you answer these questions, you'll be able to determine the hardware requirements that are the best profile for your document management implementation. In truth there is really no "one size fits all" hardware solution. Hardware requirements also depend on the number of locations and people you want to support.

Remember that minimum requirements for a test bed or a sandbox implementation are probably not sufficient for a full production environment.

I remember several years ago when the minimum requirements for SQL Server were published. At that time, a Pentium 3 with a minimum of 256 MB of RAM was recommended as a minimum. Even then, no IT department would have ever deployed an enterprise application using those guidelines. On the other hand, there may be diminishing returns on over specifying your hardware. Be sure to discuss the requirements that will best serve the solution and your successful deployment. Your document management provider will be well versed in properly specifying hardware.

Most document management server side components and database engines are multi-threaded applications; they will take advantage of multiple processes and multi-core processors. This in turn, increases server side application performance and therefore the overall end user application response time. Because RAM is relatively inexpensive, you don't need to cut corners there. In fact, with the wider adoption of 64-bit operating systems, more RAM can be used efficiently than in 32-bit server operating systems. It's best to work with your document management provider to determine the optimal amount of RAM.

As you appreciate, when specifying the right hardware for your solution, details can be important. One overlooked area is in the communication between the application server software and the database engine. Typically, it's a rapid-fire communication, so it makes the best sense to invest in a quality, high-speed interface card between the motherboard and the hard drive to get the highest possible response time and the best overall performance. This looks beyond RAID strategy to the actual hard disk controller itself. A little extra here can make a difference in overall performance.

If you use virtual server technology, the recommended hardware-based server resources may be different from those recommended for standalone hardware servers. Here's a simple rule of thumb: if the document management system requirements say you need an application server with two gigabytes of RAM, you'll generally want to specify slightly more RAM in a virtual server environment. While folks who live deep inside virtual server technology might not agree, field experience says you'll need more RAM.

Perhaps the single, most important part of any discussion regarding the hardware requirements and planning is this: Be sure you have a backup solution in place and be sure to validate that it's working as intended. Although this may seem like common sense, the results of a failed backup strategy are both heartbreaking and costly to your business.

Vault Hardware Requirements

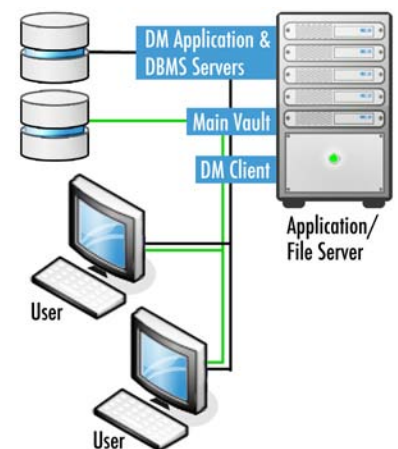
Vault server software isn't usually resource intensive enough that you will require a separate piece of hardware. That's not to say that you should ignore the vaulting hardware requirements entirely. Rather, note that additional hardware to support a vault is uncommon.

If you have multiple vaults across different geographic locations, it's best to have vault server software at each remote site. You will have one instance of the application server/business logic component of your document management solution at your company's main location, but you may have multiple instances of the vaulting software deployed across different pieces of hardware in your organization. If you require replication, you may need to invest in a dedicated piece of hardware--but that's not always the case. Your company's requirements coupled with the vault methodology you choose will ultimately dictate the hardware you require.

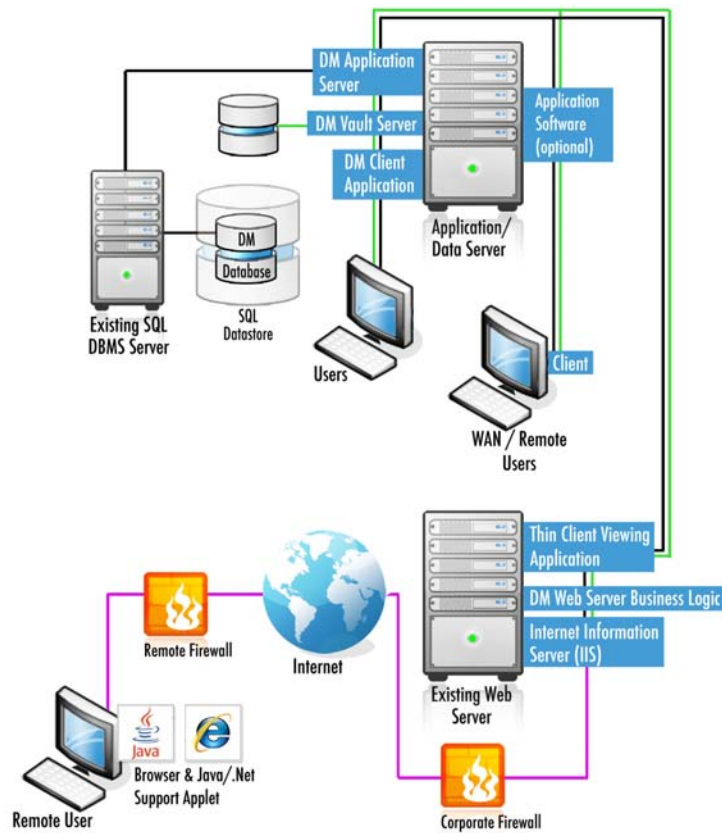
Deploying Hardware

Here are some variations on hardware deployments to consider. These diagrams do not illustrate all the variations that we see in the field; however, they should provide good examples of some common deployments.

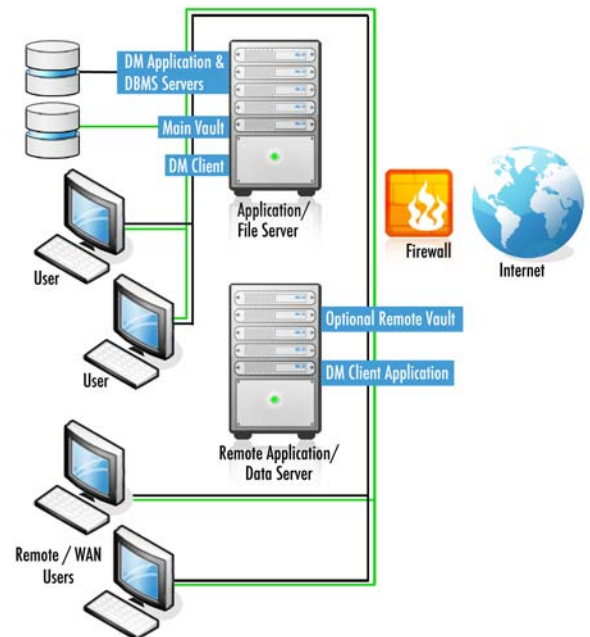
The figure to the right is a keep-it-simple deployment where everything for your document management deployment fits on a single piece of hardware.



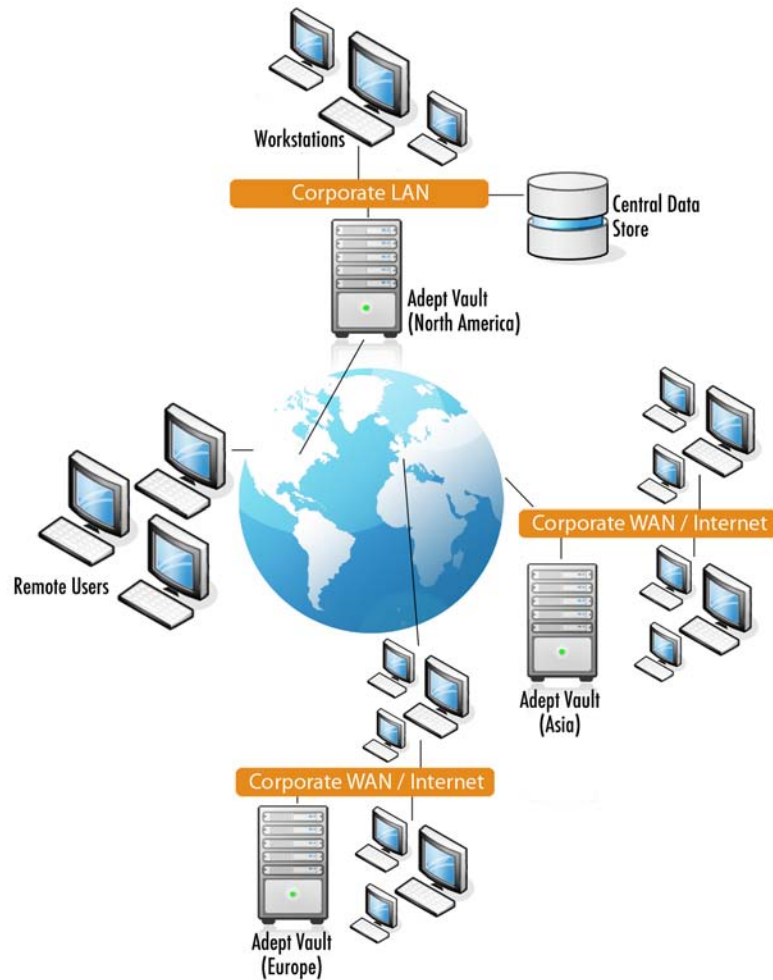
The figure below shows a deployment on a LAN. You have a central database, an application server, and a remote vault. When remote users need access to a document, it comes from a remote vault. In this particular example, the system is not accessible outside the firewall.



The figure shown on the right is another very common deployment where there are local area and wide area users as well as users accessing the system outside the firewall.



On a grander scale, the figure below shows a global implementation with multiple remote vaults and multiple users around the world accessing the system. As you might imagine with multiple locations across the world, users access the system 24/7. Not only do design teams across different time zones access the central database, they also access documents from remote vaults all over the globe. This is becoming a very popular methodology as companies collaborate and grow internationally.



Part 5: Optimizing Performance

Since the success of your document management system relies on your network infrastructure, next we'll focus on some key areas for you to be aware of as you finalize your plans.

It's important to make sure your network backbone isn't the weak link that causes poor performance and consequently a bad user experience. By today's standards, most deployments over a local area network environment perform well. Performance challenges tend to arise in extended Wide Area and Internet deployments. The factors that have the most impact in these environments are bandwidth, latency, and the way your document management system is optimized for both.

Bandwidth

Bandwidth is the maximum amount of data that can travel over a communications path in a given amount of time. It's measured in bits per second. A T1 line, for example, is 1.544 kilobits per second. Bandwidth is really all about measuring potential, not unlike the amount of electricity that can travel over a particular wire. When we talk about having a certain amount of bandwidth, like a T1, the common term refers to the maximum amount available. The tricky thing is that the maximum available bandwidth can diminish rapidly. If there's a lot of data being transferred over a fixed amount of bandwidth, the available or effective bandwidth is reduced. Additionally, a T1 in one company may have consistent adequate bandwidth while a T1 in another company may have less effective available bandwidth, e.g., a company with 25 employees vs. a company with 100 employees. With an additional 75 employees using the available bandwidth, the effective bandwidth will be less.

Most IT departments today have a very good handle on available and effective bandwidth. Greater challenges exist when connectivity is required between the US or Europe and emerging countries. In this situation, available bandwidth can vary dramatically.

Latency

Latency, or the measured time between data transmission and arrival across a network, is the hidden performance killer on a WAN. You might have adequate bandwidth but very slow data access times because latency is high.

Latency is measured in milliseconds. On today's LANs, the average access time between a workstation and the server is probably less than 1 millisecond. On a good performing WAN, it's in the 25 to 40 millisecond range. Across the Internet, latency can be 100 milliseconds or more depending on the path the data signal takes.

Reducing Latency

When it comes to dealing with latency there are good tools that monitor a network and help you define strategies to reduce the amount of lag time between transmission and arrival of data across the wire. "Chatty" or verbose applications can increase general latency and therefore, can negatively affect other applications.

If you have an inadequate bandwidth situation, then you certainly can increase the bandwidth. For instance, if you have a T1 you can introduce a T3. In some, but not all situations, increasing bandwidth will also reduce latency. You might also try to reduce the number of "hops" between locations or compress the data footprint that is being transmitted. The wide area acceleration market in the IT field is booming today and there are some very good technologies to help increase performance over the wide area network by directly reducing latency.

Poor application response in a wide area deployment could also be due to the way your document management system processes information. It could be that your document management system is not optimized for communication across lower bandwidth or higher latency environments.

Multiple Remote Vaults

Another way to improve performance in your document management system is to use remote vaults. With this methodology, you put the documents in a vault where they are most commonly accessed so that you reduce the amount of information that gets transferred across the wire.

You can also replicate vaults. If you have teams of people around the country who require access to most of the documents or even specific projects you are tracking and managing, then it is possible to replicate them so that everyone has immediate access to the data and nobody has to pull large amounts of documents across the wide area network.

Many variables contribute to optimum performance. Depending on the scope, business goals, and mandates of your document management initiative, it pays to consider them in advance. For example, a client called me and shared that their business team required IT to make connectivity over the WAN between the US, Europe and China perform as if all users, regardless of their location, were accessing information over the LAN.

Where challenges exist in your environment, solutions are available. Chances are good that should performance issues arise, your document management provider and/or your IT department can help.

This ends our whirlwind tour of the considerations for implementing a document management system. We've covered a lot of ground, including some high level concepts and some in the trenches details. For further information about engineering document management, please visit www.SynergisSoftware.com or call us at 800.836.5440 or +1 215.529.9900.

Copyright 2010 Synergis Software, a division of Synergis Technologies, Inc. All rights reserved.
All trademarks are the property of their respective owners.