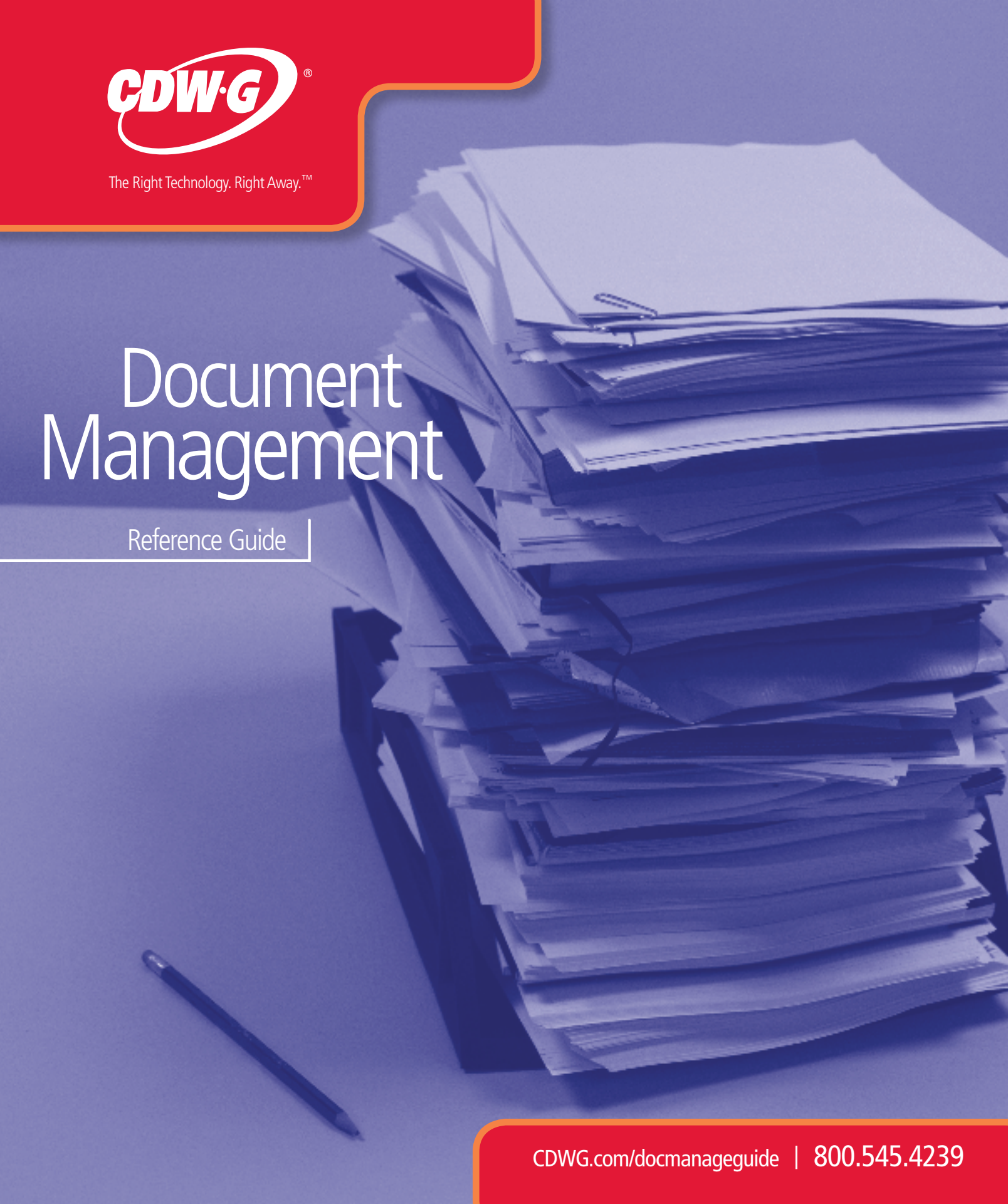
CDW·G®

The Right Technology. Right Away.™

# Document Management

## Reference Guide



CDWG.com/docmanageguide | 800.545.4239

Table of Contents

# Document Management

## Reference Guide



## What is a CDW•G Reference Guide?

At CDW•G, we're committed to getting you everything you need to make the right purchasing decisions — from products and services to information about the latest technology. Our Reference Guides are designed to provide you with an in-depth look at topics that relate directly to the IT challenges you face. Consider them an extension of your account manager's knowledge and expertise, to help your organization succeed. We hope you find this guide to be a useful resource.

Keep an eye on your mailbox for the **CDW•G IT Investment Guide**. It has even more products and information to help you realize your document management objectives.

CDW·G

# Chapter 1

# Introduction to Document Management in Government and Education

IN THIS CHAPTER:

- Motivators in Government
- Motivators in Education
- Document Management Roles in Government
- Document Management Roles in Education
- Benefits of Content Management

Being efficient is a never-ending process. Organizations are always striving to increase efficiency in operations while minimizing process defects and improving overall performance. In addition to the environmental benefits and cost reductions associated with moving to a paperless environment, going paperless decreases the time involved in nearly every aspect of traditional document processing.

Just as it is generally understood that streamlining and consolidating servers and infrastructure will maximize data center utilization and enable organizations to more quickly respond to changing IT needs, Document Management Systems (DMSs) offer a solution to paper-intensive operations that provides similar economies of scale.

The last U.S. census reported that there are at least 285 million Americans. When you consider that each of these Americans has a birth certificate, 74 percent are of voting age, and 11 percent have served in the U.S. military, you can begin to grasp the enormous volume of documents that federal, state, and local governments and educational institutions have to manage. Consider that many of these Americans also have:

- **Checking accounts** — Checks are paper that require processing.
- **Children** — Registering them for classes, extra curricular activities and much more.
- **Social networks** — Gym memberships, club memberships, ticket purchases, flights and so on, all generate paper.
- **Homes** — Mortgages require a large amount of paper including deed transfers, loan documentation, home owners association memberships and so on.
- **Vehicles** — Complete with title, insurance, registration and a great deal of other paperwork depending on what kind of vehicle and how it is used.

## Paper Use Grows

The depth and breadth of paper use in our society is overwhelming. The more society grows, the more complicated it becomes to gain access to appropriate information at the appropriate time, and so the process slows. Conversely, the public's expectations for quick, immediate access to information continue to increase.

Regardless of sector and whether you call the person whose documents you're handling student or constituent, people desire their information to be maintained confidentially, with secured integrity and available when needed. These three tenants of information protection are difficult, at best, to maintain in a paper-only environment. This further drives the need for DMSs.

DMSs come in many shapes and sizes, from systems designed to archive local office documents to full-scale enterprise solutions. They all serve the same goal: to provide a systematic method for creating, categorizing, storing, locating and retrieving documents. An effective DMS facilitates the sharing and secure distribution of documents and helps reduce the costs of distribution and filing.

A DMS should also improve and protect access to information, comply with government regulations and improve operational efficiencies. Whether your organization is a small local college or a massive federal agency, a DMS can meet demands to efficiently organize and manage the flow of information both internally and externally, and may substantially reduce the cost of operations.

This section will touch on some of the additional key motivators beyond the need for greater efficiency that are driving the adoption of DMSs across the government and educational sectors, and examine some of the roles of such systems. The remainder of this guide will explore different processes and technology components that comprise a DMS to help you make an informed purchasing decision.

# Motivators in Government

Much of corporate America has embraced DMSs and their supporting technologies as a means of reducing costs by collapsing existing document handling infrastructure and increasing the speed and availability of document delivery. The U.S. government took notice of the advantages of following a corporate approach. In order to comply with the Clinger-Cohen Act, the Office of Management and Budget (OMB) began an initiative in 2002 entitled the Federal Enterprise Architecture (FEA).

## Federal Enterprise Architecture

The Clinger-Cohen Act was written to improve the way the federal government acquires and manages information technology (IT), and the FEA provides a common methodology for IT acquisition within the U.S. federal government that is designed to ease sharing of information and resources across federal agencies. This also has the benefits of reducing costs and improving citizen services.

Two architectural components of FEA that directly relate to DMSs are the Data Reference Model (DRM) and the Technology Reference Model (TRM). The DRM describes data and information that supports government programs and business line operations. The TRM provides components that document standards for data interchange, data management, data presentation and security.

As the FEA continues to develop, it's clear that these standards will continue to be a driving force behind the adoption of DMSs within the government sector; any organization that needs to communicate via documents with the federal government will undoubtedly need to consider FEA to ensure compliance to this framework.

## Compliance (Section 508)

The Workforce Investment Act of 1998, commonly referred to as Section 508, is a series of amendments to the Rehabilitation Act of 1973 that require all electronic and Internet technologies developed, procured, used or maintained by the federal government to be accessible to both federal employees and members of the public with disabilities.

Document management systems allow organizations to remain 508 compliant by enabling documents to be retrieved from the system in many different forms. For example, once a document is in digital form, the document can be formatted as a larger text document or read aloud by screen reader equipped computers for the visually impaired.

### Specific State Guidelines

Many states have passed legislation requiring electronic and IT accessibility. In Maryland, for example, there are at least three statutes that address application development, another four related to IT procurement and two more for Web site development. These requirements not only reinforce compliance with Section 508, but are often expanding upon it.

In another example from Maryland, the state requires that technology-based instructional products procured or developed for the Maryland public school system meet the technical standards for electronic and Internet technologies issued under Section 508, which now also applies to teacher-developed instructional materials.

In California and many other states, organizations that violate these requirements face financial consequences, such as the curtailing of funding to the contractor, grantee or local agency responsible for a lack of compliance. With these kinds of negative consequences, it's apparent that documents need to be managed in a way that supports the goals of Section 508.

DMS solutions not only enable compliance, but also offer enhanced features to end users, such as the ability to zoom in on portions of documents or even read documents aloud to the visually impaired.

Additional information about accessibility legislation passed by individual states is available at **http://www.ittatc.org/laws/**.

## Consistent, Reliable and Fast

In 2006, the tech research firm Gartner reported that, "A desire for consistent, reliable results from the process of selecting documents for use in law cases will prompt U.S. courts to favor electronic content filtering by 2011." This desire is being realized in many states and counties across the United States as courthouses work to implement DMSs, with the clear motivators being first, *consistent* results, and second, *reliable* results.

Unlike traditional paper-based solutions, DMSs enable federal, state and local government agencies to share documents at the click of a mouse — as opposed to previous manual processes that required documents to be physically transported. Although the invention and adoption of fax machines did much to decrease delays associated with document transfer, they still require a request for the original document, its location and gathering, and its transmission and receipt — which is, aside from the transmission/receipt of the document, a scarcely shorter time in terms of personnel hours than a complete manual transfer of the document.

# Motivators in Education

Within the education sector, there are multiple motivators for adopting DMSs beyond simple efficiency. Of course, one motivator is the compliance needs put forth in Section 508 that are ultimately pushed down to state and local educational bodies. Second, there is the need to comply with the No Child Left Behind (NCLB) Act of 2001. There is also a clear need, as schools struggle financially, to achieve greater efficiencies and consolidate operations.

## NCLB

NCLB is a U.S. federal law signed on January 8, 2002, that reauthorizes a number of federal programs aiming to improve the performance of U.S. primary and secondary schools by increasing the standards of accountability for states, school districts and schools, as well as providing parents more flexibility in choosing which schools their children will attend.

NCLB's strong emphasis on accountability increases the demand for accessible documentation. With the right document management system in place, any size school can manage their NCLB documentation easily and efficiently.

## Greater Efficiencies and Consolidation

Many educational institutions today struggle to meet the needs of their students and communities. Budget and teacher shortfalls plague many school districts throughout the United States. Bridging the gap between what is needed and what can be afforded, school administrators work very hard to maintain as much efficiency within the organization as possible.

Administrators know that schools generate a great deal of paper. Aside from homework and other assignments, the sheer volume of paper required to manage and operate a school can be staggering.

DMSs enable schools to function smarter, not harder, by offering a uniform platform for document storage. Administrators and teachers can share virtually any kind of document from lesson plans, attendance reports, and past requests for financial support to admissions and health forms.

### Food for Thought

- The average document gets copied 9 times.
- It costs an average of $20 in labor to file a document, $120 in labor to find a misfiled document and $220 in labor to reproduce a lost document.
- 7.5 percent of all documents get lost, 3 percent of the remainder get misfiled.
- Staff spend 5-15 percent of their time reading information and up to 50 percent of their time looking for it.
- There are over 4 trillion paper documents in the U.S. alone — growing at a rate of 22 percent per year.
- The average worker spends 13 hours a year searching for lost documents.

*(above courtesy of Pricewaterhouse Coopers)*

The teacher of the very near future may be able to access a poorly performing student's grades and progress reports from other classes to determine if the poor class performance is an isolated or widespread problem.

Administrators and guidance counselors will have instant access to any document that comprises a student's history with a few simple clicks of a mouse, and will be able to make better, more informed decisions about a student's disposition than ever before. All of this consolidation means that less money is being spent to physically house documents — money that can now be spent on building more, smaller classrooms or hiring more teachers.

## Document Management Roles in Government

Within the government sector, many examples of organizations leveraging DMSs to increase efficiency and decrease operation expenditures can be cited:

- **Tax forms** — The IRS has in recent years expanded electronic filing capabilities that essentially allow citizens to directly input their documents into the IRS's DMS.
- **Drivers licenses** — Many states are now capable of reprinting a drivers license on demand and some even have the capability of providing police officers in the field with a copy of a drivers license in near-real time.
- **Local records** — It was not long ago that courthouse fires were events that wreaked havoc in the local government sector. Birth, marriage and death certificates would all be lost along with deeds, land plots and other important local information. Fire suppression systems and building codes have improved over the years, but neither offers a full assurance that all records will be kept safe.

E-mail transcripts, instant messaging records, Web pages, blogs and podcasts are a few examples of the types of media that need to be stored...

The best way to keep the records safe is to keep a backup copy offsite; a task that done manually, can be very time consuming and require additional expenses (transportation, storage rental, administrative staff). Today, documents can be sent electronically and entire warehouses full of documents can fit on devices that literally fit in the palm of a hand.

Beyond paper documents, electronic media needs to be stored and remain accessible. E-mail transcripts, instant messaging records, Web pages, blogs and podcasts are a few examples of the types of media that may need to be stored and managed in the government sector. Retention policies for documents need to be considered, as many of the government records are meant to be stored indefinitely because of their historical or other intrinsic value. A birth certificate, for example, needs to be kept on file throughout the entire life of the individual.

## Document Management Roles in Education

One environment that reaps great benefits from document management is education. Schools of all sizes are starting to see the benefits in consolidating paper operations and leveraging technology to capture, process and present routine data. In education, document management is about more than just the management of paper.

Document management has opened doors of opportunity that capitalize on collaboration. Parents in many school districts now have the capability to go online to a school Web site and view weekly progress reports on their children. School teachers can issue assignments via e-mail, and students can post and receive feedback on assignments through content management systems that enable the child, parent, teacher and administrator to interact in ways they never have before.

## Benefits of Content Management

Content management is a set of processes and technologies that support the lifecycle of digital information. This digital information is often referred to as content or, to be more precise, digital content. Digital content may take the form of text, such as documents; multimedia files, such as audio or video files; or any other file type that follows a content lifecycle that requires management.

DMSs come in many different forms, from small, open-source systems to large-scale enterprise systems. Benefits of using a DMS include:

- Relief of paperwork burdens through improved efficiency
- Minimization of the possibility of lost information and liability
- Increased productivity (which reduces costs)
- Sharing of information among many users simultaneously
- Improvement of service and response to customers

All DMS systems are not created equal, so you'll want to be discerning when determining the right DMS solution for your organization. When chosen correctly, a sound DMS will ease continuity and disaster recovery planning efforts, and help your organization to maintain compliance and avoid penalties. ■

### Quick ROI calculation:

- An organization generates or receives 100 important paper documents per day that must be filed.
- 100 documents x 6 minutes to file each = 600 minutes filing per day.
- 600 minutes per day = 10 work-hours per day.
- 10 work-hours per day x $5 per hour (inclusive of benefits and related costs) = $150 per day.
- $150 per day x 250 working days per year = $37,500 per year.

*(above courtesy of Pricewaterhouse Coopers)*

# Meet **Steve**

**Enterprise Content Management Specialist**

Steve has been a CDW•G specialist since January 2005, and brings eight years total experience in designing paperless environments and content management solutions to the CDW•G account management team. He has also earned CDIA+ certification from CompTIA and ECM-P certification from AIIM. Steve will help you develop document, content and digital asset management solutions to improve business process automation, cut customer inquiry response times, reduce paper storage and postage costs, and empower your company to meet compliance standards for retaining records.

## Turn piles of paperwork into usable electronic data with a document management system from CDW•G.

Today, in spite of all the technology, companies still have massive amounts of information that exist as individual pieces of paper and photographs. Sharing this information — and reducing the cost and complexity of managing it — is difficult because the various reports, proposals, brochures and photographs are all in different formats. But by converting these paper documents and photos into a single format — a high-quality PDF — you can simplify both the sharing and the management of the information.

## CDW•G can help you manage your paper documents.

Your CDW•G account manager will work with a document management specialist and your IT department to ensure that your solution integrates seamlessly with your current technology. Technology is only as good as the experts behind it. That's because it takes real people to solve real business problems. And at CDW•G, we take your IT issues seriously — bringing you the best technology, the best people and the best expertise the industry has to offer.

**Call your CDW•G account manager** or visit us online at CDWG.com to find out more about what an effective document management system can do for your organization.

Chapter 2

# Document Capture

IN THIS CHAPTER:

- Hardware Capture
- Software Capture
- Software Capture Formats
- Transmission

All document management begins with the gathering of information in the form of a document. A document may be an actual physical piece of paper that is electronically captured or a form that is filled out through electronic means.

The most common method of document capture is via a piece of equipment that captures physical documents, such as a document scanner. This is called a hardware capture. The second most common method is from a customer filling out an electronic form on a Web site, which is a form of software capture.

## Hardware Capture

The use of hardware to capture documents is quite common and the equipment used typically varies only in the scale at which documents need to be captured. The generic term *document scanner* is used to refer to the wide array of document imaging devices available, from small office flatbed scanners to larger paper-fed scanners that are capable of handling dozens or even hundreds of documents per minute.

Multifunction devices that combine the roles of printer, copier, scanner and fax into one device are useful in small office settings, such as local libraries or county coroner offices, that don't handle a large volume of documents and can most benefit from a small, relatively inexpensive document capture solution.

### Document Scanners

There are a wide variety of document scanners on the market today that are ideally suited for document capture and offer many features. When deciding upon a document scanner consider the following:

- Optical resolution — The optical resolution is a measure of the detail level at which a scanner can capture a document in pixels. The higher the optical resolution, the more detailed the document image will be when it's captured. Optical resolution is measured in dots per inch (dpi). Three hundred dpi has been the standard in optical resolution for documents for some time, but most scanners today are capable of greatly exceeding this resolution.
- Interpolated resolution — The interpolated resolution is a measure of the scanner software's ability to enhance an image and increase the resolution by guessing at the color of a dot in between dots of a known color. This is generally only a factor where images are concerned; the relationship being the higher the interpolated resolution, the smoother the image will appear when enlarged.
- Color depth — The color depth is the number of pieces of information (bits) that can be captured for each pixel. If your primary focus is the capture of black-and-white documents, color depth is not a decision-swaying factor. It would be uncommon today to find a document scanner that wasn't at least functioning at 24-bit. When capturing photographs or documents with highly detailed color graphics, higher bit color depth is desired.
- Simplex vs. duplex — As their names suggest, a simplex scanner can scan only one side of a document, and a duplex scanner is capable of scanning both sides of a document in a single pass. In today's financial world, where digital images of checks are just as legal tender as the original checks, financial institutions use high-speed duplex scanners to capture check items as they are processed.

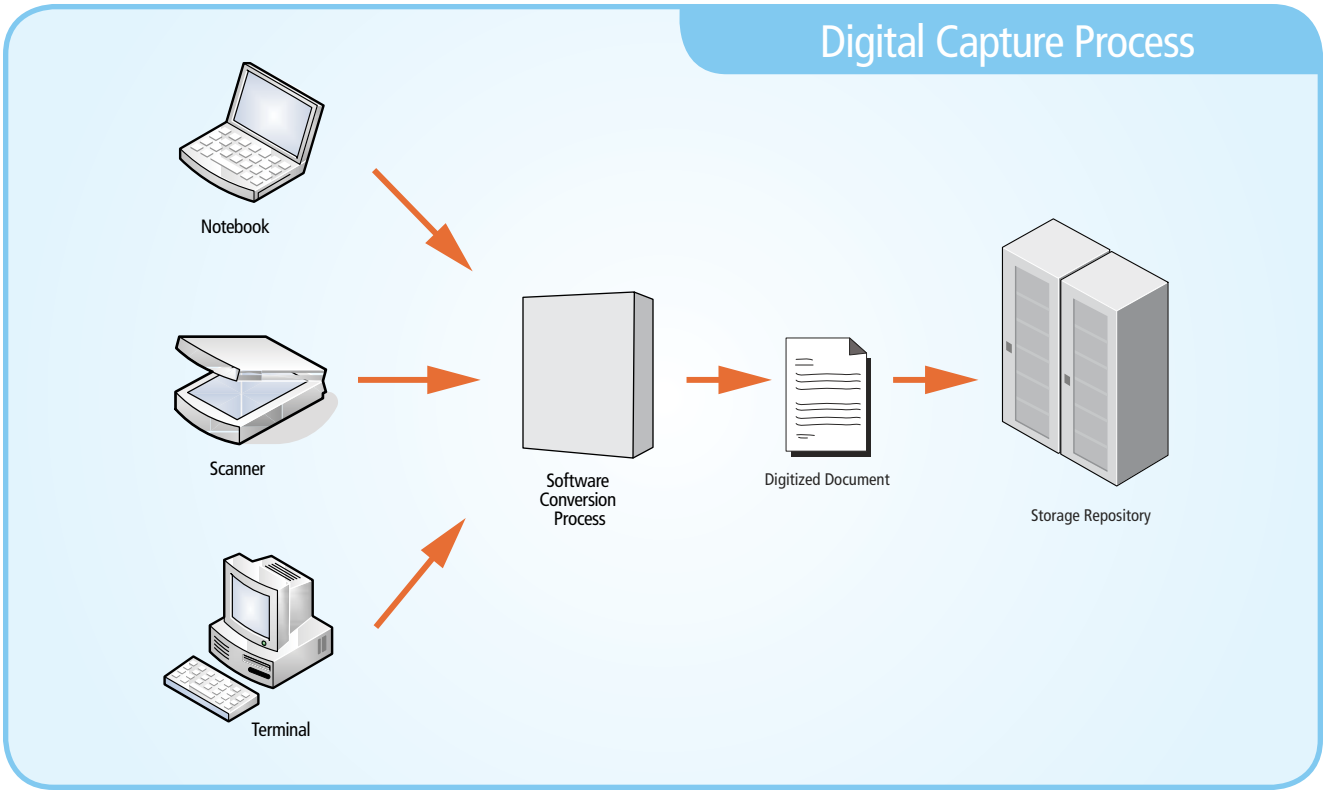
- Speed — Measured in pages per minute (ppm) or images per minute (ipm), speed refers to the rate at which scanners can capture data. Respectively, ppm is used as a measure for simplex scanning and ipm is used as a measure for duplex scanning. If speed is your primary requirement, large-scale document capture devices such as the Kodak i840 Color Duplex Scanner, which offers 160 ppm in simplex mode and 320 ipm in duplex, can make short work of tall stacks of documents.
- Flatbed vs. sheet-fed vs. multifunction — Flatbed scanners, such as the Cannon CanoScan LiDE 70, operate like a very basic copier, with only a flat glass surface and a lid to cover the image for scanning. Flatbeds are typically used as a means of document capture in environments where very few documents are captured or the documents being captured are items that cannot be sheet fed, such as capturing a few pages out of a book.

  Sheet-fed scanners such as the Xerox DocuMate 250 can offer higher volume and operate like copiers, with an automatic document feeder (ADF) that can be filled with the original documents and then scanned in order as fed from the ADF to the scanning surface.

  The key selling point on sheet-fed scanners is the ppm it's capable of scanning. This is important, but don't neglect to look at other important decision items such as resolution and color depth. A variation on the use of a sheet-fed scanner is the portable sheet-fed scanner. Devices like the Visioneer Roadwarrior scanner offer portable scanning solutions for notebook users on the go.

Multifunction devices may come with scanners that are flatbed, sheet-fed or both, and combine the secondary and tertiary functions of printing and faxing. Generally, the maximum size documents that can be handled are standard paper sizes, such as letter (8.5 inches by 11 inches) and legal (8.5 inches by 14 inches).

- Duty Cycle — Sometimes referred to as a Daily Duty Cycle (DDC), this is a measure of how many scans a scanner can process in a day. This factor is particularly important in large-scale document scanning solutions because manufacturers often design scanners to capture this data internally, and exceeding the Duty Cycle can mean voiding a warranty. In procurement, the DDC is a good way to determine how many of a particular model you should purchase to meet the needs of your document capture environment.
- Interface — In small office environments, the Universal Serial Bus (USB) interface dominates the market. This is because no additional hardware is required (most PCs have a USB port). But USB can handle only relatively slow speeds, document scans of approximately 20 ppm. In larger, faster environments other interfaces, such as the Small Computer Systems Interface (SCSI), may be warranted. SCSI is not only faster than USB, but can also be buffered using SCSI hardware, meaning that the scanner can



be used at maximum speed to the interface at all times.

A few variables specific to multifunction devices that should be taken into account are the features of the printer, copier and fax functions of the device.

## Printer

- Type of printer — Choosing between laser printing and inkjet printing is a primary decision. Laser printers are typically more expensive, but have a lower cost per print. So in large volumes of printing, laser is the more cost-effective option. Inkjet printers are less expensive, but the cost per print is typically higher than laser.
- Cost per page — Many manufacturers list an approximate cost per page, which takes into account the paper and ink or toner that is used.
- Print speed — Measured in ppm, this refers to the speed at which the device prints. Note that printers usually print noncolor documents faster, so check the ppm speed for both color and gray scale documents.
- Print quality — Measured in dpi. As with digital capture, this is the resolution at which the printed document will be rendered on paper.

## Copier

- Copy speed — Measured in copies per minute (cpm), copy speed is the rate at which documents can be copied.
- Resolution — Refers to the resolution of copied documents in dpi.
- Multicopying/stack/sort — If the device has an ADF, the multicopying, stack and sort specifications will show how many documents the device can make per page loaded in the ADF, and whether it will stack or sort those copies.
- Enlargement/reduction ratio — Enlargement permits documents to be made larger for printing. Reduction allows the user to compress two or four pages into a single page.

## Fax

- Transmission speed — 33.6Kbps is the standard transmission speed for most fax devices.
- Error Correction Mode (ECM) — ECM detects errors during fax transmission and will resend any pages that had an error. This feature must be on both the sending and receiving fax to function.
- Color or black-and-white — If color faxes are important, you will need to ensure that your fax machine is capable of sending and receiving color faxes.
- Speed dial — Take note of the number of speed dial numbers that can be stored in the device.
- Dual access — Allows the user to perform two actions on the machine at the same time, such as sending a fax and printing a document.
- Broadcasting — This refers to the ability to send a single fax to multiple numbers.
- Fax forwarding — This refers to the ability to send a fax that is received in memory to another device. If the device runs out of ink, is unable to print or sent to the wrong office, fax forwarding allows the user to send the fax to another machine.
- PC faxing — This refers to the ability to send faxes directly from the desktop.



# Software Capture

One of the biggest document management concerns is that of usability. Beyond capture and storage, the documents that enter into the DMS are going to be indexed and eventually used by someone. The ease with which the end user can find and access documents within the DMS should be a paramount concern for any organization.

As stated previously, one of the primary ways documents enter into the DMS is through software capture, such as a form filled out on a Web site. Beyond the capture of documents through electronic means, software capture includes the process of converting the image of a document into a format that can be indexed, edited and used in electronic form.

Many software applications on the market today enable software capturing of documents to make digitally imaged documents editable and usable. The right software choice can make all the difference in creating a usable DMS.

## Adobe Acrobat Capture 3.0

The predominant name in document management is Adobe. Its Portable Document Format (PDF) has become the de facto standard document format. Adobe Acrobat Capture 3.0 software is a professional production tool that teams with your scanner to convert volumes of paper documents into searchable PDF files. Features include Optical Character Recognition (OCR), which identifies printed characters and translates them into digital characters, advanced page and content recognition, and powerful cleanup tools.

There are two editions of Adobe Acrobat Capture:

- Acrobat Capture Personal Edition for Microsoft Windows or Mac environment. Acrobat Capture is a professional production tool for turning paper-based information into high-quality knowledge documents optimized for electronic publication. A hardware dongle, which is a security device that authenticates a particular software program, allows conversion of as many as 20,000 searchable PDF pages. Additional dongle packs are available for

purchase if more pages are needed.

- Acrobat Capture Cluster Edition: This software suite provides the same professional production features as the personal edition, but is scalable for high-volume requirements, offers multiprocessor licensing and converts an unlimited number of pages.

## Captiva Products

EMC software's Captiva QuickScan Pro (QSP) is a standalone, out-of-the-box imaging solution that provides capabilities for high-speed scanning, image enhancement, viewing, annotation, printing and storage. Its strengths include flexible job separation, bar code recognition (up to 30 bar codes per page) and usability.

In addition to being a standalone product, QSP can be easily integrated into DMSs, serving as a "scan module" that feeds scanned, cleaned-up images into the system. This flexibility makes it possible to streamline the incorporation of accurate, useable content into the overall document workflow.

Being resolution independent is an important quality, particularly for Section 508 compliance, because it allows the [PDF] document to be accessible on many different kinds of devices at many different resolutions.

## Kofax Image Products

Kofax offers a variety of image products serving the needs of end users, government organizations and software developers. Three of Kofax's main software capture solutions are:

- Kofax Ascent Capture — This product automates the capture of information from documents and forms in one software suite. Kofax Ascent Capture is available in a variety of forms, from the standard Ascent, which is a multipurpose suite, to software designed specifically to capture invoices (Ascent for Payables) or health care forms (Ascent for Health Insurance).
- Kofax Capio — Provides document imaging and management capabilities for individual users. Ideal for small offices where document management volume is low. Capio includes Kofax VirtualReScan.
- Kofax VirtualReScan — VirtualReScan is a software capture technology from Kofax that automatically detects document problems such as skew, upside-down capture and background pigment problems and corrects the document image.

## Nuance Products

Nuance offers products that enable the users of DMSs to convert, organize and share documents in various formats. OmniPage Professional enables end users to precisely and efficiently convert documents and forms into many popular office applications. This allows documents that are scanned to be made instantly available as Microsoft Office files or quickly formatted into PDF files and distributed through the enterprise. ScanSoft PaperPort Professional enables small- to midsize-organizations to organize, find and share their documents, including paper, PDF, application files and photographs.

# Software Capture Formats

Documents can be captured and stored in a variety of formats, each with their own benefits. At the most basic level, document capture involves the digital imaging of documents into a digital format. A software capture may simply be a graphic image file, such as a photograph of a document, which is useful at a basic level if you ever need to recall and reprint a document, but does little for an organization that needs to work with the information contained within the document.

Software capture converts what was essentially a digital image of a captured document into a more usable format from which information can be extracted. Adobe PDF is one of the most popular formats, but a wide variety of document formats exist that can be leveraged at various points within an enterprise or organizational DMS.

## Adobe PDF

Adobe PDFs are device independent, so they can be viewed on virtually any computer platform. Each PDF file includes a complete description including the text, fonts, images and graphics that compose the document. They are also resolution independent, so PDF documents can be easily resized and displayed on virtually any type of device.

Being resolution independent is an important quality, particularly for Section 508 compliance, because it allows the document to be accessible on many different kinds of devices at many different resolutions. This is important for 508 compliance because it allows people with visual impairments to scale documents to larger sizes without losing document quality.

Subsets of PDF have been or are being standardized under International Organization for Standardization (ISO) for several constituencies:

- PDF/A for archiving in corporate/government/library/ environments as ISO 19005 (work done in ISO 171)
- PDF/E for exchange of engineering drawings (work done in ISO 171)
- PDF/UA for universally accessible PDF files
- PDF/X for the printing and graphic arts as ISO 15930 (working in ISO TC130)

## TIFF

The Tagged Image File Format (TIFF) is the standard image file format for document capture and, like PDF, is device independent. TIFF files can store multiple document pages within a single file, making them a better storage standard for digitally imaged documents than other graphics file formats designed specifically to handle images (such as Windows Bitmap files — BMPs). TIFF is an image format, which means that it's generally not editable by word processing applications.

## Microsoft Suite

Microsoft Office applications (Word, Excel, PowerPoint) have an extremely large footprint within government and education offices and must be considered in the document management equation. Though usually just referred to as a Word, Excel or PowerPoint files, the Microsoft family of document formats consists of more than just the common file types most people equate to the applications.

Microsoft Office 2007 has recently introduced a new file format known as XML Paper Specification (XPS). XPS is designed to provide users with a consistent document appearance regardless of where and how the document is viewed, much like an open-source PDF system. XPS has the support of printing companies such as Canon, Epson, HP, Ricoh and Xerox as well as many software and hardware companies.

# Transmission

Once the documents have been digitally imaged and entered into the DMS, you might need to transmit the documents to other agencies or individuals that don't have direct access to the DMS. Many DMSs include features that easily allow for the export of documents from the DMS for transmission via fax, e-mail or other means.

When transmitting documents there are a few key points that need consideration:

- Portability — It's important that the person receiving the document be able to access and use it without any special hardware or software configuration. Adobe PDF, TIFF and the newly introduced Microsoft XPS format are all standard document formats that can be accessed regardless of operating system (OS). The Adobe PDF format does require a PDF reader that is freely available at http://www.adobe.com.
- Security — When sending a sensitive document it's important to protect the document from tampering. Both Adobe and Microsoft offer security features for their native documents that limit users' ability to edit or modify the document content.
- Compliance — Many compliance issues are raised in the transmission of documents. The Health Insurance Portability and Accountability Act (HIPPA), for example, sets forth policy on who may receive personal health-related information and, depending upon which compliance your organization needs to adhere to, there may be requirements for auditing who received what information and when.

In addition to traditional e-mail and fax transmission, another kind of device has emerged to manage the transmission of documents. The HP Digital Sender looks a lot like a multifunction device, but is actually a bit more. This device will convert paper documents into digital information to be sent by e-mail. The documents are compressed and converted to PDF or TIFF files.

Once the transmission is completed, the information can be reformatted, edited, printed or even stored by the receiving party, and documents can be sent to different people in different locations at the same time. ■

# Enhance your operations.

## Do it with a document management solution from CDW•G.

**An effective document management system** must accommodate every type of document that is part of your organization's processes and procedures.

There are three primary methods of bringing files into a document management system:
- Scanning or imaging — for paper files
- Importing — for archiving such electronic documents as word-processing files, spreadsheets and faxes
- Conversion — for creating unalterable images of electronic documents

Batch processing of documents and forms is also recommended if high-volume processing is part of your operations — otherwise, imagine the utter inefficiency that would ensue were each document to be processed individually.

A full-featured document management system will allow files and records to be brought into the system in one batch to speed up your process. Once all the pages have been captured, the system should enable users to easily annotate, arrange or otherwise sort them to ›

FUJITSU

### Fujitsu Scansnap™ S510
Color Document Scanner

Features "one-touch" scanning to searchable PDF and ships with Adobe® Acrobat® 8.0 Standard and WIndows Vista™-supported version of ABBYY FineReader 3.0.

- Scan speed: up to 15 and 18 ppm simplex
- Scan resolution: 600 dpi optical
- 50-page Automatic Document Feeder (ADF) with double-feed detection and blank-page detection
- Automatic color detection and automatic de-skew

$473.80 CDWG 1238024

Fujitsu 1 year Advance Exchange In-warranty — extended service agreement $63.83 CDWG 1071099

CDW·G®

800.545.4239
CDWG.com/docmanageguide

meet the particular needs of your organization. The system should also enable users to update or add index fields at a later date. If your organization has multiple offices, your document management system should permit authorized users at all locations to capture and access documents as necessary. In this case, your system should allow for documents to be scanned into the system and transferred into the database at different times to alleviate network bottlenecks during peak traffic hours.

Major advancements in scanning technology have made paper document conversion fast, inexpensive and easy. CDW•G offers a wide variety of scanners to meet your conversion needs. And we have the software to help you manage your digital assets, too.

**Call your CDW•G account manager today or visit us online at CDWG.com to find the document management system that's right for your organization.**

visioneer

GSA Compliant

### Visioneer Patriot 780 Duplex Production Flatbad Scanner

The Visioneer Patriot 780 Duplex Production Flatbad Scanner can be programmed to automatically scan to different destinations and be customized for each individual or workgroup in your organization.

- 60 ppm/120 ipm at 200 dpi black and white
- 120-page Automatic Document Feeder (ADF)
- Includes ScanSoft OmniPage, Kofax® VRS®, TWAIN Driver, and Certified ISIS Driver

$3905.20 CDWG 1043094

Canon

NEW model

### Canon ScanFront 220P

Versatile Compact Full-Color Network Scanning Console for Day-to-Day Needs

- Scan speeds up to 25 ppm
- Easy operation with 8.4" built-in color touch screen monitor
- 50-sheet Automatic Document Feeder (ADF)
- Capable of scanning simplex and duplex, as well as a variety of media types

$2195.00 CDWG 1254968

However you choose to distribute your documents, CDW•G has the solutions to help you make it happen.

# Get the tools to get the job done.

**Volume Software Licensing Advantages: Easier Purchasing** — As your organization grows, multiple users are easily added to your network by purchasing necessary license certificates instead of boxed product. For example, if you want to run a software program on 250 different workstations in multiple locations, you would purchase one Media Kit and a 250-user license. This alleviates the need for increased bulk shipping costs and saves you time distributing multiple boxed products to multiple locations.

**Centralized Asset Management** — Many leading vendors will provide asset management and help you track the number of licenses you have purchased. This helps you maintain compliance with manufacturer specifications and helps you determine how many software copies are needed when updating your network.

NUANCE


LICENSING AVAILABLE[1]

## PDF Converter Professional 4

PDF Converter Professional 4 is a complete PDF solution. The efficiency and power of PDF Converter Professional 4 is priced for every agency, every desktop. It's a Better PDF for Government™.

- Create PDFs from any application and form into a single multi-page PDF document
- Convert PDF files into editable Microsoft Word and Excel® documents while keeping the look exactly like the original PDF
- Turn static PDF forms into fillable PDF forms or Microsoft Word forms

Full version $79.99 CDWG 998383

Five-user license $325.18 CDWG 1031707

[1]Contact your CDW•G account manager for details

Adobe


LICENSING AVAILABLE[2]

## Adobe® Acrobat® 8 Professional

Adobe® Acrobat® 8 Professional software provides effective end-to-end electronic work processes for government agencies and their mobile workers by replacing paper with Adobe PDF documents and forms.

- Auto-recognizable form fields
- Enable advanced features in Adobe Reader®
- Save in Microsoft® Word

For Windows®

Upgrade $99.74 CDWG 1063257

Full version $209.47 CDWG 1063256

For Mac®

Upgrade $99.74 CDWG 1063254

Full version $209.47 CDWG 1063253

[2]Adobe Open Options (AOO) volume licensing; call your CDW•G account manager for details

**Upgrade Protection** — The purchase of software Upgrade Protection — sometimes known as Upgrade Insurance, Subscription or Maintenance — guarantees that you will obtain the latest version upgrades throughout the duration of your license agreement. This ensures that your organization has the most up-to-date versions of the licensed software as soon as it becomes available.

**Saves Money** — Volume software licensing offers discounts for multiple software product purchases. Generally, the greater the quantity required, the deeper the discount. Discounts vary between vendor programs and products. You can be sure that if more than five users in your organization require the same software program, there is a licensing alternative available which will lower your cost of ownership.

### Why Purchase Volume Software Licensing From CDW•G?

Licensing your organization's software programs can be expensive and time-consuming. It requires keeping track of the software being used by each department, in addition to staying on top of the licensing programs associated with various manufacturers.

At CDW•G, we understand the value of your budget and time. That's why we provide you with a Software Licensing Specialist who works directly with your dedicated CDW•G account manager to ensure you receive the most cost-effective software licensing solution. Our Software Licensing Specialists are available to help you understand complex volume licensing programs and answer any questions you may have.

EMC² | captiva


[ Increasingly fast response times make it even easier to implement a storage system to accommodate your document storage needs. ]

CALL FOR PRICING

**EMC® Captiva® QuickScan Pro v.4.5**

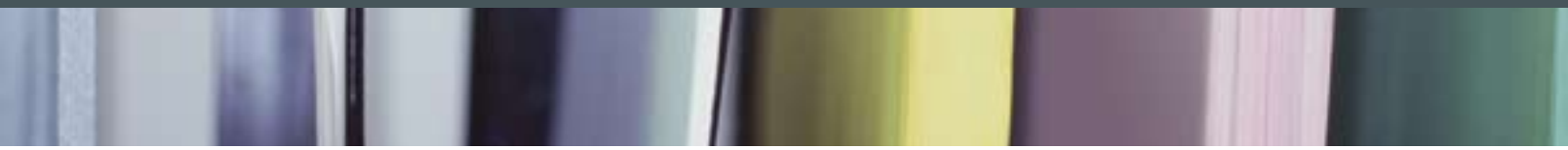
Easily integrates into content-management systems

QuickScan Pro is EMC® Captiva's out-of-the-box document scanner software and imaging solution that provides all the necessary capabilities for high-speed scanning, image enhancement, OCR, viewing, annotation, printing, and storing images for both black-and-white and color pages. This software supports bar-code and patchcode recognition for automatic document separation.

CDWG 1234446

# Document This

## The California Animal Health and Food Safety Laboratory System gets a grip on its data, thanks to Canon document scanners.



The next time you throw a few steaks on the barbeque or bite into a burger from your favorite fast-food restaurant, you may want to give thanks to an unlikely source: Canon DR-2580C color document scanners.

That's because the California Animal Health and Food Safety Laboratory System (CAHFS) — which is tasked with helping to ensure that animal diseases such as Mad Cow, Avian Influenza, Exotic Newcastle and West Nile virus do not pervade and endanger the public — has implemented the scanners as part of a comprehensive document management solution.

Operating within the School of Veterinary Medicine at the University of California, Davis, the CAHFS serves as the veterinary diagnostic lab for the state of California and is considered the backbone of the state's warning system. The organization helps protect the health of livestock and poultry, while safeguarding the public health with rapid and reliable diagnosis for animal diseases affecting humans.

Since deploying the nine scanners as part of a Laboratory Information Management System, the devices have significantly aided the organization in better automating many of its processes.

### Less Paperwork

"The Canon paper scanners help us reduce the amount of paperwork that we copy and circulate within the laboratory," reveals Jay Ross, IT manager for the CAHFS.

Often taking in 250 to 500 cases per day, staffers must document a full clinical history on each animal it diagnoses, such as the vaccination history, symptoms exhibited, suspected illness, how the specimen was collected and stored, and administrative details such as contact and billing information.

Prior to purchasing the scanners, CAHFS staffers were forced to fill out countless forms by hand, which had to be maintained indefinitely and produced in multiple copies.

Furthermore, if an employee had a question about a previous case, he or she had to sift through an enormous paper trail in search of the document. "Now that everything is being scanned in, they can look it up in the system and see it without having to pull any file copies," notes Ross.

### Small Size, Big Features

Compact, powerful and fast, the Canon DR-2580C is a versatile scanner designed to meet diverse imaging needs, including the ability to accommodate a wide range of source materials such as plastic ID or insurance cards.

In addition, when paired with the optional user-installable flatbed unit, the DR-2580C is capable of scanning books or fragile documents, while large documents up to 11" x 17" can easily be scanned with the device's Folio mode function.

"We selected this Canon based on its speed, price and size," Ross reports. "It's only slightly larger than a tissue box, and we were facing some space constraints."

Indeed, weighing just 4.2 lbs., the compact, portable DR-2580C is not only one of the smallest scanners in its class, but also one of the easiest to use. One-touch scanning is achieved by a user simply assigning the most common functions to the device's customizable Scan-To Job buttons.

Built to boost productivity, the DR-2580C scans at a rapid rate of 25 pages or 50 images per minute (letter/200 dpi/portrait). And validating the old adage that good things come in small packages, the DR-2580C doesn't skimp on performance, boasting some of the most advanced features in document scanning.

For example, users can enhance black-and-white scans with the contrast adjustment feature, which sharpens images and makes text more distinctive on low-contrast documents.

Furthermore, the scanner's punch-hole removal function erases the black dots that appear when pages are scanned from booklets or binders, while the DR-2580C's Rapid Recovery System ensures

seamless scanning operation if a feeding error occurs.

In addition to the color document scanners, the CAHFS also purchased 60 Symbol bar code scanners and 27 Zebra bar code label printers, which enable staff members to quickly and efficiently process and track the numerous vials and tubes within the lab.

## Greater Productivity

The technology has not only reduced paperwork and streamlined processes for the CAHFS, but it has also dramatically increased the organization's response time.

During a peak emergency, for example, it used to take eight people four hours to process 1500 specimens. Since the implementation of the document management system, however, a single individual can process 1000 specimens in just 15 minutes, according to Ross.

You don't have to be a mathematician to figure out the magnitude of this technology's return on investment. But Ross emphasizes that the value of the devices extends far beyond the savings on resources, paper, time and money.

"Our product is data, because we provide the test results and the knowledge gained from diagnosis," Ross explains. "Increasing our ability to manage this data and streamline the process improves the overall quality of that data."

## What's Next

"We are making full use of the system to automate our processes," Ross explains, adding that the organization is planning to further enhance its document management system with the addition of tablet PCs.

"We are often in wet and dirty environments, and having a tablet will eliminate the need to capture all of that information on paper," he points out. "We're right there, and we can capture the data directly at the source."

## Right on Target

Whatever technology needs arise, the CAHFS knows it can trust CDW•G to deliver the optimum solution.

"When we go to buy anything, we always contact CDW•G," says Ross. "They always have a group of technology experts available, and robust solutions to offer. It makes it really easy for us to choose and obtain the products we need."

Ross also praises CDW•G Account Manager Dejan Perovic for his ongoing support. "I contacted all three manufacturers that were selected for this proposal and negotiated for more aggressive pricing," notes Perovic. "By doing so, we helped them get under their target price for this proposal." ■

# Choosing a Document Management system.

### Key Points to Consider.

A document management system is a collection of integrated tools that convert paper documents into electronic form to be stored, shared and easily accessed. It can help streamline everyday business processes and improve operational efficiencies.

Since choosing the right solution is paramount to the success of your document management system, it's important to understand both your present and future needs up-front. Consider the following guidelines:

### Understand where you are now.

Before implementing a new system, it's beneficial to understand where your organization is right now. In addition to evaluating current processes and your overall IT environment, survey staff members to help identify your requirements.

Be sure to define the amount of information you need to manage and map out your current workflows, as well as how your files will be indexed, routed or archived.

### Determine where you want to be.

A well-defined requirements document that includes workflows will help you choose the right technology for your document management system. The more time you spend developing your requirements, the better your chances of creating a system that meets your needs. Make sure that everyone who will be impacted by the project is on board prior to implementation.

### Research and select your technology.

With requirements and architecture documents in hand, you can begin to compare the available technology solutions and choose the products that best fit your system.

### Assemble and pilot your system.

Once you've determined your needs and chosen your products, you can begin assembling your document management system. Consider starting with a pilot program. You can then tweak your system as needed, based on the results of the pilot program.

### Continuously evaluate your system.

As your organization and your document management system grow and change, it's critical to evaluate the effectiveness of your system, and be ready to make changes in scope and scale as your organization's needs dictate.

# It is vitally important for information to be continually secure and available.

## SURROUND YOURSELF with a document management solution from CDW•G and surround yourself with success.

As a government organization you have a lot of documents to manage, if you're not using the right software to help you manage all those documents, you quickly find yourself behind the game instead of ahead. At CDW•G we understand the challenges that face your organization on a daily basis and we carry innovative products designed around the way you work.

To help increase efficiencies, reduce costs and optimize interactions among co-workers and with customers in today's highly competitive environments, organizations need to extract every possible advantage from technology — especially in their communication activities.

At CDW•G we offer more than just brand name technology. We offer dedicated account managers who will work with you every step of the way to help you determine the best document management solution for your organization. Call your account manager today to get started on a document management solution that fits your needs, your style and your budget.

> When it comes to technology we know that you have a lot of options, and that's why we're passionate about responding to your needs immediately.

ESKER
Document Process Automation

CALL FOR PRICING

### Esker Fax™ Server Solutions

Version 4.0

Esker Fax™ provides true large-scale organization solutions for sending, receiving and routing high volumes of end-user and application-generated documents, all from a single server and without any modification to your existing applications.

Esker Fax integrates with mainframe and midrange host computing systems, multiple network types, SAP applications, GroupWare and more.

CDWG 986053


CDW·G®

800.545.4239
CDWG.com/docmanageguide

QUESTYS SOLUTIONS

CALL FOR PRICING

## Questys Pro

### Document management software

You can have any document at your fingertips in seconds — making it easy and affordable for organizations to go "paperless" and be more efficient. This intuitive solution delivers an advanced out-of-the-box feature set that includes all the tools an organization needs to digitally manage documents, including indexing, search and retrieval, OCR, access management, version control, and many more.

CDWG 776658


FileBound

CALL FOR PRICING

## FileBound Workflow

### Manage your data efficiently and effectively

FileBound allows you to enjoy all the advantages of a complete records management system without the upfront capital investment. With the FileBound on-demand, software-as-service model, you get all the benefits of a unified, scalable solution without the high costs, risks and complexity normally associated with records management software.

CDWG 827147


KOFAX
A DICOM GROUP COMPANY

## Kofax® Virtual ReScan® Plus

### For desktop scanners

VRS® technology from Kofax® ensures that your scanning is as efficient and easy as possible, while also improving both the quality of the scanned images and the automated capture of information from your paper documents and forms. VirtualReScan®, scanning productivity tool, offers reduced document preparation, intelligent color handling, and improved image quality and information capture.

$409.99 CDWG 798419

# Building a storage solution
## that's right for your organization.

**THINK OF YOUR STORAGE LANDSCAPE** as a spectrum. At one end, multiple forms of stored backup systems help protect data availability over the long run. At the other end, local backups provide fast, convenient access to critical data.

Identifying the right combination of storage solutions for your organization depends upon many factors including your infrastructure, industry-specific requirements and the different ways users need to work with and access critical data now and in the future. To determine where and how tape might best fit your data protection and storage management needs, first consider your current and anticipated data growth — how you acquire, use and manage your data as well as how much of it you'll need to manage should all factor into your buying decisions.

CDW•G's team of experts has the knowledge and experience to assist you with any type of storage issue from server consolidation and virtualization to the most complex compliancy-related needs. Our storage specialists work closely with your account manager to help you define how to store, manage and backup your data more efficiently and effectively.

CDW•G's team of experts has the knowledge and experience to assist you with any type of storage issue.

CDW•G storage solutions make it easier, faster and more cost-effective to deliver information to users when they need it, comply with government regulations when required and prepare your organization for growth.

**Do you know your options?** To learn more call your account manager or visit us online at www.cdwg.com.

EMC²
where information lives

CALL FOR PRICING

### EMC® CLARiiON® CX3 Model 40c

**Consolidate more, without extra costs or complexity, with FC/iSCSI options**

The powerful CLARiiON® CX3 UltraScale™ series of networked storage systems delivers maximum benefits with new levels of performance, scalability, flexibility and ease of use. CX3 UltraScale systems give you the flexibility to mix and match high-performance and low-cost drives to meet different service levels for tiered storage.

The CLARiiON CX3 Model 40c extends the CX3 UltraScale series into the heart of the mid-tier storage market. It's the OLTP workhorse of the CX3 UltraScale series. With support for up to 128 highly available, dual-connected hosts and the capability to scale up to 240 disk drives for a maximum capacity of 117TB, it outperforms other midrange arrays at an affordable price. The CX3-40c is available in a Fibre Channel-only configuration or with four 4 Gbps FC and eight iSCSI ports . The CX3 Model 40c is ideal for demanding OLTP workloads, and for larger requirements for array-based replication.

CALL FOR PRICING

## HP StorageWorks® MSA1000 Small Business SAN G2 Kit

### Reduces complexity and cost

- Supports Windows® and Linux®, both 32- and 64-bit
- Provides performance with dual/active controllers increases up to 40K I/Ops (cache), and throughput of up to 357MB/s
- Compatible with HP's ProLiant x86 and x64 servers, both Intel and AMD-based, with PCI-E slots
- Supports fourteen 1-inch drives in a 4U rack space. It is easily expanded to 12TB using forty-two 300GB hard drives in 10U of rack space
- Supports hot plug drives and enclosures, controllers, fans, and power supplies
- Supports the industry's highest level of fault tolerance (RAID 6 with HP's Advanced Data Guarding technology)
- Kit includes: (1) MSA1000 (1) MSA SAN Switch 2/8 (2) AE311A HBAs; (4) SFP transceivers + cables

CDWG 1092105

HP Smart Buy[1]

## HP 1/8 G2 Ultrium 448 Tape Autoloader

Provides high capacity data storage in a compact 1U form factor, with advanced features as a bar code reader and web based remote management, all included in an entry-level automated backup solution.

- Storage capacity: up to 1.6TB native, 3.2TB compressed[2]
- Includes HP 1/8 G2 Tape Autoloader Rack Kit
- One drive and eight slots
- Includes HP StorageWorks® Data Protector Express, Single Server

$3399.00 CDWG 1152161


HP Smart Buy[1]

CALL FOR PRICING

## HP DL100 G2 2TB Storage Server

### Add storage to your network quickly and easily

- Consolidate to a single centralized network storage appliance to manage and protect
- Microsoft® Windows Storage Server 2003 Release 2 Operating System pre-installed
- Multi-protocol file serving: SMB/CIFS, NFS, HTTP, FTP, WebDAV, AppleTalk, and NetWare
- Packs storage (up to 2TB raw), memory, and I/O into only 1U of rack space
- Print server support for up to 5 printers

CDWG 1227736

[1]After HP Smart Buy instant savings; savings may vary based on channel and/or direct standard pricing; available as open market purchases only; call your CDW•G account manager for details [2]Assumes 2:1 compression

# Why CDW•G?

**This reference guide offers an overview of Document Management Systems (DMSs) and how they can be utilized within government and education.**

Whether your office serves the public, or educates, CDW•G has specialists on hand that understand you. We know how important it is that you are able to do what you do and we offer assistance during all stages of design and deployment of your new DMS.

We offer resources far beyond the hardware and software necessary to build a DMS. We focus on providing the best solutions in the industry — not simply the latest solutions from a single provider. By building strong partnerships with industry leaders and leveraging our talent with their resources, we are able to drive out the best possible solutions for our customers.

At CDW•G, we have the technology and the talent to meet your needs. We offer a complete selection of IT hardware and software at competitive prices. The offerings range from scanners and software for document input and processing to secure storage devices and access, control and information lifecycle management tools. We also have resources available to deploy to your location for onsite consultations and demonstrations.

Wherever we go we are never alone. Behind each and every CDW•G account manager is a talented team of document management specialists who are on hand to offer in-depth industry knowledge, real-world experience and one-stop access to world-class partners, ensuring that every interaction with CDW•G is a positive interaction.

At CDW•G, we're committed to providing you with the very best in consultation, design, delivery and support. Our overriding goal is to ensure your satisfaction over the long run. At CDW•G, we have designed our company with your goals in mind.

Chapter 3

# Document Management

IN THIS CHAPTER:

- Store and Find
- Access
- Auditing
- Indexing
- Archiving

The actual management of documents, whether it's in a government organization or on a school campus, can be challenging. Many different document users, each with their own expectations, will depend upon the DMS to ensure that the documents they need are easily accessible and presented in a manner that is useful.

## Store and Find

DMSs are systems that store and find documents based upon a wide range of criteria. Management of the system involves determining the types of files being stored and searched for, as well as how often this operation occurs.

For example, if you work in a small government office such as a county commissioner or city council office, you might have a fair amount of documents (deeds, records, constituent information, local laws and ordinances) that need to be searched frequently by staff. For the most part, these are documents that are changed or updated infrequently. In this situation a document repository may be what is most warranted. Although not an official term, *repository* is often used to denote a DMS that changes infrequently.

A Content Management System (CMS) is a type of DMS that offers content to users as well as collects information from them, and can generate documents from that information. In contrast to a document repository, a CMS involves a higher volume of document turnover or more rapid document gathering via electronic means.

For example FileBound, a Web site service that allows users to interact, upload documents, check out documents, manage lists (as in Microsoft Excel) and display pictures, is a CMS widely used in the commercial and educational sectors. Because of the ease of access (via Web interface), it's difficult to discern where the Web site ends and the document management functions begin.

FileBound offers an intuitive, organic means of document management that handles high-volume document activity well. Users upload documents to share with the team in a collaborative environment, and all team members can search through all documents for pertinent information.

Whether your organization may benefit most from a document repository, a CMS or both, keep in mind that future retrieval of documents is no longer hindered by file type or version. To illustrate this point, consider any word-processing document created prior to 1990. An office suite known as PFS: First Choice was popular in the late 1980s, but its documents are all but inaccessible with the latest version of many popular word-processing systems.

Only in recent years has a conscious decision been made by digital document leaders, such as Adobe, to ensure standards are built that will remain open and compatible with future systems. The previous chapter discussed PDF/A and Microsoft's new XPS format, both promising to remain compatible for many years to come.

## Access

Regarding document access, there are two universal maxims in document management:

- Documents that cannot be accessed when needed cost the organization money.
- Documents accessed by nonauthorized personnel represent

a liability.

Thus, at this stage of document management, you must be concerned about security, control and compliance.

Security can be broken down into three basic concepts that form what is known in information security circles as the Confidentiality, Integrity and Availability (CIA) Triad.

- Confidentiality — Confidentiality refers to the prevention of unauthorized disclosure of information. Disclosure of data through the loss of confidentiality can occur in a number of ways including deliberate release, inaccurate or ineffective document security controls or the failure to apply properly defined security controls.
- Integrity — Integrity means ensuring that no unauthorized alterations are made to data, i.e. no unauthorized changes to documents. Integrity loss can occur through deliberate, unauthorized tampering, software or hardware malfunction and human error.
- Availability — Availability refers to the timely delivery of data to those who require access. This means ensuring all who need access receive access when they require it. Loss of availability, like integrity, can result from deliberate action, unauthorized use or misuse, software or hardware malfunction and human error.

Maintaining confidentiality, integrity and availability is essential to document management security, and failure to do so can open your organization to legal and compliance risks.

To protect documents from unauthorized disclosure your organization may wish to adopt a few high-level policies:

- Acceptable Use Policy (AUP) — The purpose of an AUP is to outline the acceptable use of computer equipment and resources. These rules are put into place to protect staff as well as the organization. Inappropriate use of a computer system can expose an organization to risks that include virus attacks, compromised network systems and services, and legal issues.

  You may want to develop a separate AUP specifically for acceptable use of documents to meet regulatory compliance, and to help prevent inappropriate use of sensitive information. An AUP is an important document to have on file should any rules violations require corrective action from a human resource or legal perspective.
- Access policies — Access policies clearly define the credentialing process for users: who may authorize use and under what conditions access may be granted. Many organizations use differing levels and types of access policies, from those that grant physical access to a location to those that grant access to a particular database or DMS.
- E-mail policies — E-mail policies define how e-mail is sent, received and retained. An auto-forward e-mail policy may be used to define the use of auto-forwarding and what conditions need to be met for this to be allowed. An e-mail retention policy may dictate which e-mail needs to be retained, where and for what duration.
- Information sensitivity policies — An information sensitivity policy is intended to help staff determine what information can be disclosed to unauthorized personnel, as well as the relative sensitivity of information that should not be disclosed outside of the organization without proper authorization.

  If your organization collects Social Security or criminal record information or handles documents from closed-session meetings of federal, state or local officials, an information sensitivity policy should be used to guide staff in the handling of these types of documents.
- Password policies — Passwords are the front line of protection for user accounts. A poorly chosen password may result in compromised sensitive information. A password policy sets forth standards for the creation of strong passwords, the protection of those passwords and the frequency of change.

# Auditing

To reduce the risk of falling outside any legal or compliance concern faced by your organization, regular audits of DMS policies should be performed to ensure that confidentiality, integrity and availability are being maintained. One of the most important aspects of DMS auditing is document access. Granting access to a document is not typically handled as an all-or-nothing equation. Different users within your organization may require varying degrees of access, from simply viewing a document to managing document creation, modification, deletion and archiving.

In addition to using Access Control Lists (ACLs) to limit the capabilities of users within the operating system (OS) or within a network environment like Active Directory, certain document formats can control user rights. Adobe Acrobat, for example, enables you to protect documents with passwords and apply restrictions on printing, copying and altering. These restrictions are applied to the documents themselves, so you can be assured that your documents are protected even after they leave your DMS.

# Indexing

As pointed out in the previous chapter, the simplest form of document management is the capture of document images that can later be printed. From an information perspective, these images are relatively useless because virtually none of the information contained within the document has been indexed. Once the information in the document is read and recognized either through optical character recognition, DMS input or user intervention, that data can then be indexed.

Indexing is the process of converting a collection of data into a database suitable for easy search and retrieval. Automating the capture and indexing of document data is important to improving the efficiency of the process. It's during indexing that the document tags are created that will enable the search for documents within the DMS.

A good DMS will allow you to define indexing priorities to help minimize disruption to existing document filing procedures. If your traditional document filing system includes a combination of date, time and a user identification number for the file name, you might want to index documents based upon these familiar qualities at first, to ease the transition into the new system.

Although it is possible to index the entire content of a document, care should be exercised to index only the information required to locate and retrieve the document. Aside from the security and storage concerns presented by indexing the contents of every document, too much information can slow search response and make it more difficult to locate individual documents.

Outside of a full DMS, individual documents can be indexed in other ways. Adobe Acrobat Reader allows users to search a collection of PDF files, and Microsoft Windows Indexing Service extracts information from a set of documents and organizes it for easy access through the OS.

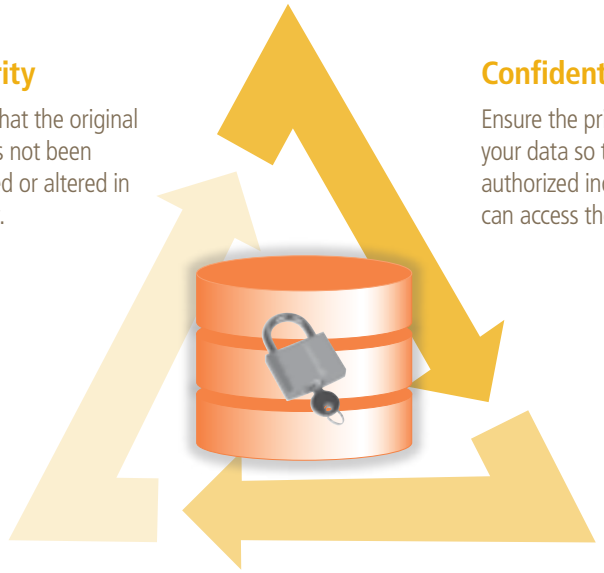
# Archiving

Archiving is the process of backing up selected documents that no longer need to be accessed on a regular basis within the DMS. Effective archiving means safely storing critical data in a highly accessible manner.

During the archiving process, documents often go through different stages and types of storage, which is referred to as Hierarchical Storage Management (HSM). HSM is the process of automatically



## The CIA Triad



storing data on the lowest-cost devices that can support the performance required by the applications.

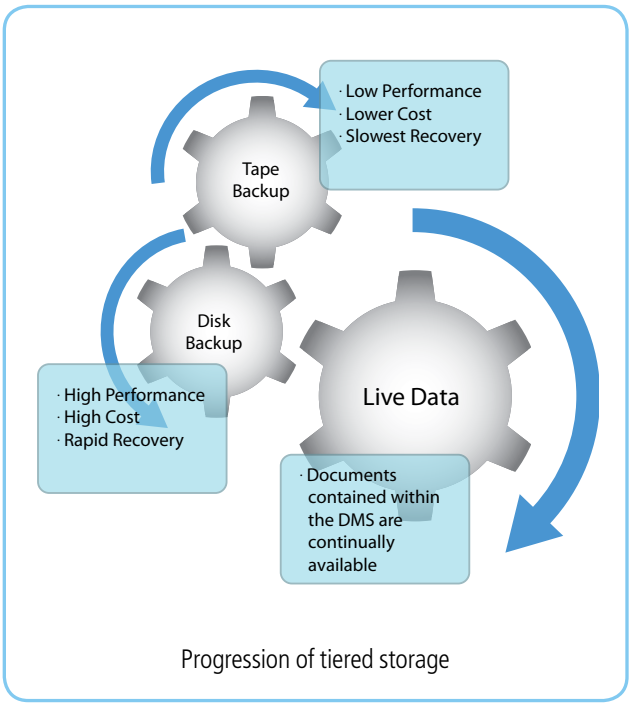
In HSM, documents are removed from online disk storage (live storage) to a lower-cost media such as tape or optical disk. If the hardware and software used for archiving supports on-demand recovery, archiving offers a near-online solution, allowing documents to be stored on slower, less costly media (such as tape) and accessed within a few minutes rather than a few seconds.

The architecture of an archive is highly dependent upon your organization's need to retain data. Going back to the repository versus CMS example, archiving in a DMS repository may mean storing lesser-used records — such as old laws and ordinances, retired teacher files or records held only for regulatory compliance — in an offline or near-offline state.

CMS archiving, however, may retain access to information as recent as last month that no longer needs to be displayed at the forefront of the CMS system. No matter what archiving solution you require, it must be secure, scalable and integrated into every facet of document management so that all content, from regulated documents to e-mail, can be easily recovered when needed.

## Archiving Storage

Archiving storage can be as easy as backing up to tape in smaller offices. But for larger environments, long-term storage will typically happen through a series of tiers. Tiered storage is a data storage environment that is comprised of two or more storage types that are used in concert to meet the short- and long-term needs of the DMS. The differences in the types of storage used are usually along one of four lines: price, performance, capacity and function.




Progression of tiered storage

For example, if your organization requires rapid recovery of the DMS due to a failure during business hours, but isn't likely to require a restoration past 24 hours (which would mean that your organization is a short-term document user), a high-performance disk storage solution may be used to perform frequent local or network backups during the day. Then each night a backup may be written to a less expensive (slower) tape media. This is an example of tiered storage commonly referred to as Disk-to-Disk-to-Tape (D2D2T). This hybrid storage solution is primarily appropriate for medium-to-large environments.

### D2D2T

In D2D2T, data is first copied to storage on another disk (D2D), and then periodically copied again to tape (D2T). Large scale, high-performance DMSs, which could potentially consist of many individual DMS servers, may have a production assurance or continuity need to have their documents immediately available for recovery from secondary disk.

Any significant change in performance, capacity or function can be implemented as a secondary tier. In the D2D2T example, there is a change in all four categories: performance, capacity, function and price. But tiers could easily be made up of low-capacity media for short-term storage and large-capacity media for long-term storage, or high-performance media for short-term storage and low-capacity media for long-term storage. Tiered storage can be used to cut costs by implementing storage solutions that directly meet the size of the need at the appropriate time by utilizing

Chapter 4

# Storage Management

IN THIS CHAPTER:

- Storage Media
- Storage Considerations

Beneath the DMS lies a vast array of choices in storage devices. Storage management encompasses both physical and virtual storage components. When relating storage management to DMS, you need to assess the overall storage requirements, where the documents will be stored, and how they will need to be accessed, by whom, how often and from what location.

Storage management of a DMS begins with the database. A database is a structural way of housing information. In DMSs, databases drive much of the presentation and storage. In a CMS, the data and which particular pieces of the data are displayed (for example, the fields you display or the order in which you display the data) do not have to mirror the way the data is actually stored within the storage infrastructure. Take particular care when choosing a DMS to note the database infrastructure, as each is going to have its own particular storage needs.

The following list highlights common questions to ask when looking at the underlying database in a DMS:

- Will it run on my hardware or commonly available hardware? — If your organization already uses Microsoft SQL Server and you have a budget and staff to support Microsoft SQL, it might be wise to stick within the realm of current expertise.
- How fast can documents be recovered?
- How fast is the indexing?
- Can you set your own index parameters or customize the index within the database itself?
- Will the performance degrade significantly over time and, if so, can you scale up to larger hardware solutions to meet the demand? How complicated or costly will it be to do so?
- How much does it cost? — Cost is always a factor, but failing to take aspects of the cost beyond the sticker price will itself be costly. Be sure to evaluate the Total Cost of Ownership (TCO) in hardware, software and ongoing technical support as a whole.

A few common databases being used in DMSs today include Microsoft SQL Server, MySQL and Oracle — all of which have their own particular advantages and disadvantages. Microsoft SQL Server and Oracle, for example, are both commercial products, whereas MySQL is an open-source solution.

## Storage Media

Choosing the proper storage media for your DMS is critical to the ongoing success and adoption of the system. Failure to plan the storage architecture can lead to future problems with access, scalability, availability and resiliency. From the single workstation installation to complete enterprise storage solutions, DMS storage solutions can take a variety of forms to meet the needs and size of virtually any document management environment.

### Direct Access Storage

Direct Access Storage (DAS) is a general term for disk storage devices such as hard disk drives. Direct access refers to how all the data is accessed directly in about the same amount of time, as opposed to sequential access methods (those used by magnetic tape).

In its most basic form, a DAS device attaches directly to a computer as either an internal hard disk drive or external device, such as a USB-attached hard drive. A standalone DAS device often makes a

poor choice for storing documents except for the smallest of single office/single workstation applications, because DASD devices lack the architecture needed to provide a sufficient amount of redundancy, scalability and resiliency.

More generally, DAS devices are the foundation for many types of storage solutions including Redundant Array of Independent Disks (RAID), as well as Network Attached Strorage (NAS) and Storage Area Network (SAN) solutions.

## RAID Arrays

A RAID is a collection of disk drives that offers increased performance and/or fault tolerance. There are a number of RAID levels. Those most commonly used in DMS are RAID 0, RAID 1, RAID 5 and RAID 10.

- RAID 0 (also known as a striped set) splits data evenly across two or more disks with no parity information for redundancy. RAID 0 is used primarily to increase performance.
- RAID 1 (also known as a mirror) creates an exact copy of a set of data on two or more disks. This is useful when write performance is more important than minimizing the storage capacity used for redundancy.
- RAID 5 (a stripe set with parity) uses block-level striping with parity data distributed across all member disks. RAID 5 has achieved popularity due to its low cost of redundancy.
- RAID 10 (mirroring combined with striping) uses a set of drives combined into a striped array. This striped array is then mirrored to another identical set of striped drives. RAID 10 provides the performance benefits of striping with the redundancy of mirroring. It is ideal for multiuser and server systems, but is the most expensive in terms of hard drive cost.

Small- and medium-size office DMS server infrastructures will typically demand RAID 1 or RAID 5 for fault tolerance. That is not to say that RAID 0 doesn't have its place. If document access speed and retrieval is your primary concern, a RAID 0 configuration, when combined with a sound (and frequent) backup methodology, may meet your needs. Rather than dedicate a server to storage, with all of its accompanying support and licensing costs, network-based storage can provide a good alternative for small-office document storage.

## NAS

NAS devices can be accessed over a computer network rather than being connected directly to a computer. This setup enables multiple computers to share the same storage space at once, minimizing overhead by centrally managing hard disks. NAS systems usually contain one or more hard disks, often arranged into RAID arrays.

NAS was developed to deliver ease of use, provide for data sharing and enable organizations to simplify their data management. To do so NAS solutions address several problems inherent to DAS. These include the administrative effort required to maintain servers and the challenges of scalability, reliability, availability and performance presented by server-based DAS solutions.

In recent years, the cost of NAS devices has plummeted and high-capacity NAS solutions — such as the Buffalo TeraStation Pro 2TB network attached storage, which offers 2 terabytes (TB) of storage — are available for very reasonable prices.

NAS solutions thrive in small- to midsize organizations as a cost-effective way to rapidly add storage. Larger organizations, however, often run into difficulty managing the number of NAS devices needed to meet their needs. If left unchecked, NAS devices can "sprawl" through an environment, creating challenges in storage management.

## SAN

A SAN is a network designed specifically for storage and is usually built on infrastructure designed especially for that purpose. Because of their dedicated architecture, SANs tend to provide faster and more reliable access to storage than NAS offers. SANs are scalable, highly available and come with a clear benefit over NAS solutions in being easier to centrally manage.

Fibre Channel is the most common SAN technology and utilizes a SCSI command set. A Fibre Channel SAN is usually made up of a number of Fibre Channel switches that are connected to form a SAN "fabric."

Internet SCSI (iSCSI), which uses the same SCSI command set over Transmission Control Protocol/Internet Protocol (TCP/IP), is another

### NAS vs. SAN

NAS solutions are useful for single small offices that never intend to outgrow a small 1 to 2TB solution. But in larger enterprises, NAS can represent a problem in the form of sprawl. Sprawl is a term used to describe the unintended proliferation of differing hardware or software components.

There are many forms of NAS from many different vendors, and over time it's easy for NAS systems to start piling up. Managing different storage technologies from different vendors can be a real headache for administrators. SAN offers a centralized, reliable and manageable solution that can scale to meet the needs of any size infrastructure.

popular type of SAN architecture. In an iSCSI architecture, the switches are Ethernet switches and the network is not referred to as fabric. iSCSI is a generally less-expensive alternative to Fibre Channel because it can use existing copper-based networks.

Small organizations that plan to grow storage requirements can start with one of the many SAN solutions geared to the entry-level SAN community. The HP StorageWorks Modular Smart Array 1500 SAN SCSI Starter G2 Kit is a cost-effective and easy way to get started with SAN and offers expandable storage to meet the growing demands of your DMS.

By now you should have a general idea of which storage solution might work best for your environment. If your DMS is simply going to be for a few documents that do not need to be archived, and your workstation is backed up regularly, a workstation-based DAS solution may meet your needs.

If you intend to share the documents among many users, DAS on a server in a RAID configuration, or as a more affordable alternative, a NAS device, may meet your needs. However, if you're going to be dealing with thousands of documents shared among many users and demand the utmost in scalability, performance, resilience and data protection, implementing a SAN is the best way to go.

> Continuity and disaster recovery planning... should be taken into account as a critical storage consideration.

## Storage Considerations

For a storage solution to be considered viable for document management it must be measured against your document management needs. Taking time to consider how your documents will be accessed, by who and how often are good starting points. You will also want to consider the overall volume of documents to be managed, and remember to review how legal and regulatory compliance will come into play for document archiving and retention in your workspace.

### Continuity/Disaster Recovery Planning

Continuity and disaster recovery planning, planning specifically focused on creating a comprehensive plan of actions to be taken before, during and after a significant loss of information systems resources, should be taken into account as a critical storage consideration.

During the continuity planning process a detailed analysis should be conducted of both threats and impacts. For example, if a threat of natural disaster, such as a hurricane or earthquake, has the potential to impact a data center where the DMS is stored, a solution should be aligned to mitigate the impact.

### Consolidation

There are essentially three stages of consolidation that any enterprise IT resource can go through:

- Logical consolidation — Focuses on the centralization of management of IT resources and is commonly referred to as unified management.
- Centralized consolidation (co-location) — Is the physical consolidation of servers, storage devices or other infrastructure components to a central location — usually a data center.
- Physical consolidation (compiled workload) — Involves the

Physical consolidation, which is often realized hand-in-hand with centralized consolidation, is an easy entry point for standardization efforts...

compiling of workload from multiple servers or storage devices onto a single server platform; in terms of storage, this may be the equivalent of replacing multiple independent RAID 5 storage arrays on many separate servers within a SAN.

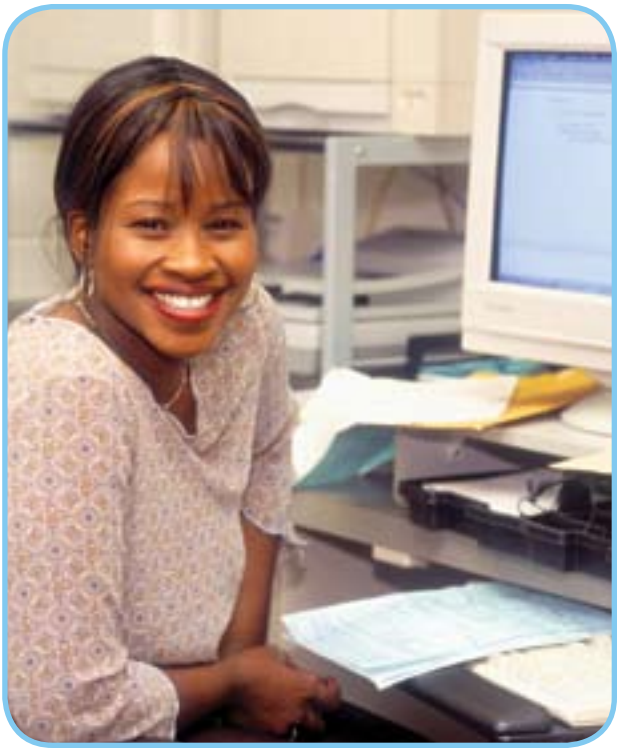
Logical consolidation is essentially what is already being done to documents by implementing a DMS. In doing so the DMS consolidates the management of documents to one unified system. That same concept applies directly to the storage infrastructure as well. As noted earlier, if your organization has already consolidated to a single solution, such as Microsoft SQL Server, the most cost-effective approach may be to continue to leverage that existing state of consolidation.

The most cost-effective step to implement is usually logical consolidation. For some organizations this might simply mean the purchase and implementation of software capable of simplifying the management of all storage assets under one system.

Centralized consolidation involves the co-location of physical storage devices and is usually more costly to implement than unified management, depending on the size and complexity of the storage infrastructure.

If your organization, for example, currently has data stored sporadically across the infrastructure, consolidating the data to one data center will simplify the management of the data to that physical infrastructure. Physical consolidation, which is often realized hand-in-hand with centralized consolidation, is an easy entry point for standardization efforts, as this involves the simplification and standardization of an infrastructure, often to a single platform.

## Legal and Regulatory Compliance



Legal and regulatory compliance may also encourage appropriate backup and recovery systems be put into place to protect information from accidental loss or destruction. This may warrant a larger scale media storage architecture than was originally forecast, such as a SAN, which provides a consolidated, fault-tolerant and scalable storage solution. ■

# Glossary

This glossary should serve as a quick guide to the most important terms touched on in this guide. Note acronyms are commonly used in the IT field.

## Acceptable Use Policy (AUP)

Organizational guidelines that specify the appropriate use of equipment and resources. These guidelines may cover access, e-mail, information sensitivity and password policies.

## Access Control List (ACL)

A list that specifies what access privileges users have to a directory, specifically, the right to look up, insert, delete, read, write, lock or administer files.

## Archiving

The intelligent backup of selected documents that no longer need to be accessed on a regular basis.

## Auditing

The process of regularly reviewing DMS policies to ensure continued confidentiality, integrity and availability.

## Clinger-Cohen Act

A 1996 U.S. federal law written to improve the way government acquires and manages information technology.

## Color depth

The number of pieces of information (bits) that can be captured for each pixel.

## Confidentiality, Integrity and Availability (CIA) Triad

An IT security framework that stresses the importance of securing confidentiality, integrity and availability in document management.

## Content Management System (CMS)

A system designed for the creation, modification, archiving and removal of digital information from an organized repository. A CMS may include tools for publishing, format management, revision control, indexing, search, retrieval and collaboration.

## Data Reference Model (DRM)

A component of the Federal Enterprise Architecture (FEA) that describes the data and information that supports government programs and business line operations.

## Direct Access Storage (DAS)

A very basic approach to disk storage, often consisting of a device that attaches directly to a computer and allows access to all of the data in the same amount of time.

## Disaster recovery planning

A comprehensive plan of actions taken before, during and after a significant loss of information systems resources that addresses both threats and impacts.

## Disk-to-Disk-to-Tape (D2D2T)

A backup methodology in which data is copied from one disk to a second disk, and from the second disk to magnetic tape.

## Document repository

A DMS approach for data that is searched frequently but changed or updated infrequently.

## Document scanner

A term used to describe a wide array of document imaging devices, including flatbed, multifunction and paper-fed scanners.

## Dots per inch (dpi)

A measurement unit that defines resolution.

## Duty cycle

A measure of how many scans a scanner can process in a day.

## Federal Enterprise Architecture (FEA)

A business-based framework for government-wide improvement.

## Fibre Channel

A network technology that utilizes a SCSI command set and is made up of a series of switches that all connect to form an interconnected SAN "fabric."

## Hardware capture

Physical document capture that is facilitated via a capture device, such as a scanner.

## Health Insurance Portability and Accountability Act (HIPAA)

Legislation passed in 1996 that includes a privacy rule creating national standards to protect personal health information.

## Hierarchical Storage Management (HSM)

An approach to document archiving that emphasizes the storage of data on the lowest-cost devices that can support all of the connected applications.

## Indexing

The act of describing or identifying a document in terms of its subject content.

## Interface

A hardware or software link that connects two computer systems, or a computer and its peripherals.

## Internet Small Computer Systems Interface (iSCSI)

A SAN architecture that uses the SCSI command set over TCP/IP in conjunction with Ethernet switches.

## Interpolated Resolution

A measure of the scanner software's ability to enhance an image and increase the resolution by guessing at the color of a dot in between dots of a known color.

## No Child Left Behind (NCLB) Act

A United States federal law that reauthorizes a number of federal programs aiming to improve the performance of U.S. primary and secondary schools.

## Optical Character Recognition (OCR)

A software process that takes a scanned document and attempts to recognize the letters and words on the page.

## Optical resolution

A measure of the detail level at which a scanner can capture a document in pixels.

## Portable Document Format (PDF)

The de facto standard format for digital documents. PDF's popularity stems from its device and resolution independence, which allow it to be read on most any computer platform and device.

## Redundant Array of Independent Disks (RAID)

An approach to data storage that emphasizes increased performance and fault tolerance and involves a collection of disk drives arranged in different levels.

## Pixel

The smallest picture element of a digital image. A digital image is composed of millions of these tiny dots.

## Resolution independent

A feature of the PDF document format that allows it to be resized and displayed on virtually any device. An important quality for Section 508 compliance.

## Section 508

A series of amendments to the Rehabilitation Act of 1973. Section 508 sets forth regulations requiring documents to be accessible to people with disabilities.

## Software capture

Document capture process that begins with a user entering data into a digital form, followed by the conversion of an image of the created document into a format that can be indexed, edited and used within the DMS.

## Storage Area Network (SAN)

An enterprise network designed specifically for data storage that emphasizes fast, reliable access, scalability and centralized management.

## Storage management

The process of overseeing both the physical and virtual aspects of data storage, including the database itself, overall storage requirements, document storage and document access.

## Tagged Image File Format (TIFF)

An industry standard file format developed for the purpose of storing high-resolution, bitmapped, gray-scale and color images.

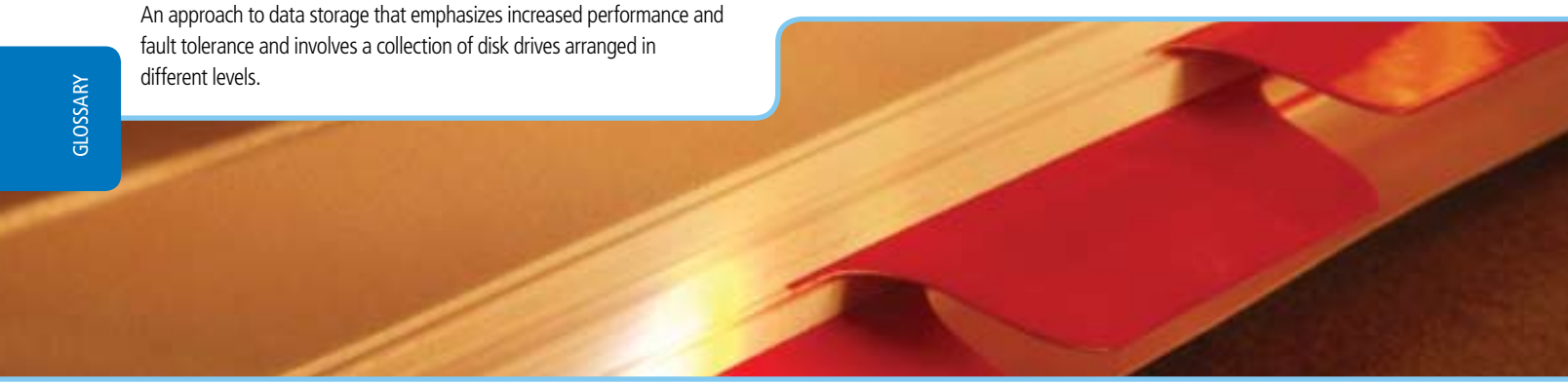
## Technical Reference Model (TRM)

A component-driven, technical framework used to categorize the standards, specifications and technologies that support and enable the delivery of service components and capabilities within the Federal Enterprise Architecture.

## XML Paper Specification (XPS)

A specification designed to provide users with a consistent document appearance regardless of where and how the document is viewed, much like an open-source PDF system.





# Index

# Document Management

Reference Guide

## Look inside for more information on:

- The benefits of content management for Government and Education
- Choosing the right document-capture solution
- Document management software that best meets your organization's needs
- Guidelines for document-access security
- Creating a tiered-storage solution

CDWG.com/docmanageguide | 800.545.4239

**CDW Government, Inc.**
One CDW Way
230 N. Milwaukee Avenue
Vernon Hills, IL 60061

ISO 9001:2000 certified

SmartPay™ GSA

GSA Contract Holder

PRSRT STD
U.S. POSTAGE
PAID
MT. PROSPECT, IL
PERMIT NO. 87

**Your account number is**

**Key Code**

071008
Flyer 31777