32158 Camino Capistrano
Suite A   PMB 373
San Juan Capistrano, CA 92675
Sales@CaptureComponents.com
www.CaptureComponents.com

Transform Your Office with Document Capture

Overview

This White Paper explains how to Transform Your Office with Document Capture from one that is paper-centric to one that is electronic document-centric; from an office that consumes countless staff hours filing, retrieving, locating paper files and is controlled by significant reliance on paper to one that places lower reliance on paper and more on the ability to electronically access your documents quickly and effortlessly.

Start the transformation by analyzing your reliance on paper documents, estimating the current cost of paper filing and storage and locating misfiled paper documents. Use the introduction to page scanners and digital document storage to help understand the document capture techniques for scanning, identifying and storing your new digital documents. With this information you will learn how to eliminate lost documents, save many valuable square feet of file cabinet space and instantly share documents with associates.

The transformation officially begins with your commitment to reduce paper and use new document capture technology to dramatically improve your office productivity.

Get started transforming your office

The first step is to analyze your reliance on paper by taking a quick inventory of paper documents in your office. Excluding magazines, publications, books, and permanently bound documents, estimate the number of vertical and lateral file cabinets, vertical and lateral file drawers and depth of stacks of paper on desks, tables and shelves.

Estimate the amount of paper storage and handling

One typical vertical file cabinet occupies about 8 sq. ft. of floor space and a typical lateral file cabinet occupies about 13 sq. ft. including access space with drawers extended. Use table T1 to calculate the number of square feet of office space consumed by your file cabinets.

Estimate 1,500 pages per foot for loosely stored pages in stacks. For papers stored in file cabinets use 3,000 pages per drawer. These values vary with drawer depth, paper thickness, number of folders and how many folders are in use. For tightly packed file drawers use 4000 pages per drawer. Tables T-1 through T-3 may be helpful for these calculations.

Table T-1 -- Space Consumed by File Cabinets

File Cabinets	Count		sq.ft./cabinet		Area
Vertical	_____	x	8	=	_____
Lateral	_____	x	13	=	_____

Total Area (sq.ft.) Consumed by File Cabinets   =   ═══════════

    Your facility cost per mo/year per sq.ft.   =   _____

Potential Savings for Alternate Use of Space   =   ═══════════

Table T-2 -- Total Number of Paper Pages

File Drawers	Count		Pages/Drawer		Total Pages
Vertical	_____	x	3,000	=	_____
Lateral	_____	x	3,300	=	_____

Loose Stacks	Feet		pages/foot		
Stacks	_____	x	1,500	=	_____

Total Pages in File Cabinets and Stacks (P)   =   ═══════════

Table T-3 -- Digital Storage Space Requirement

% of Total		Total Pages (P)		KB/Page		Space Required
Black & White, letter size, single sided						
_____	x	_____	x	40	=	_____
Black & White, letter size, double sided						
_____	x	_____	x	80	=	_____
Color or Grayscale, letter size, single sided						
_____	x	_____	x	120	=	_____

Total Digital Storage Required in KB   =   _____

divide by   1,000,000

Total Digital Storage Required in GB   =   ═══════════

Estimate costs of handling paper and offsite data protection

The Total Area Consumed by File Cabinets and the Total Paper Pages in File Cabinets and Stacks are the basic data to justify document imaging. Additionally think about these questions and whether or not you can quantify a cost for each.

1. How much time have you and others spent searching for a document in stacks, in file cabinets or elsewhere in the office?

2. How important are your documents? Could you afford to lose any? What would it cost your organization if your file cabinets were damaged by a fire or water? Do you have protected offsite microfilm, digital or paper archival storage for your documents? Do you offset archival storage costs for your files?

3. Do you have a need to share documents with associates in your building or office? Do you need to share documents with others offsite? In your work process how much time do you or staff spend making copies of paper documents which are physically delivered to others? How much time do you spend sending faxes of paper documents?

4. What is a rough estimate of the cost of labor for filing, retrieving, checking-out and distributing hardcopy pages stored in file cabinets?

Based on the total number of pages, file cabinet square footage and potential savings or alternate use of space, filing labor, cost of lost documents and potential offsite archival cost information, do you want to eliminate most of the paper files and store documents electronically? If you answer Yes, please continue. If No, perhaps document imaging is not appropriate for you, but read on to understand more about document capture and its benefits.

Learning about Document Capture

Scanned documents are essentially digital images of paper documents. They look the same, can have one or two sides and display in color, black & white or grayscale. In the imaging business, a digital document is defined as one or more pages, cohesive in content, which are logically inseparable. A one or two page letter is a document. Contracts, articles, checks, invoices, loans, proposals, resumes, reports, quotes, tax returns, are examples of documents. Stored digital documents occupy space on a computer hard drive, USB hard drive, network storage system or removable media like CDs, DVDs or USB flash drives.

A page scanner is used to scan paper documents and convert them to digital documents. A computer program is used to control the scanner. A packet of pages placed in a scanner tray may contain one or many documents. After scanning a packet, documents are separated by the program, identified automatically or manually, and given a file name. Finally they are stored in some type of digital document storage location called a repository. This complete process of scanning, separating, identifying and storing images is called Document Capture.

Document Images

A typical black & white (bitonal) image is usually stored in the industry standard TIF (Tagged Image File) format. A letter size document image takes up about 40 KB of disk storage when scanned at 200 dpi. (dots per inch)and are of rather good quality. For higher quality documents scanned at 300 dpi, the image consumes 70 - 90 KB. Bitonal TIF images

are usually compressed according to the standard. A full color letter size document image may take 0.5 to 10+ MB, depending on the format type and compression/quality used. JPEG (Joint Photography Experts Group) is the most common color format, and offers a range of compression ratios so you can trade off size with quality. Grayscale documents, with hundreds or thousands of shades of black and white take up more space than bitonal images but usually less than color images. The common grayscale format is JPEG.

Another very popular document format becoming universally accepted is PDF (Portable Document Format) originally developed by Adobe Corporation. PDF documents can be pure text as from a word processing program, a TIF bitonal image, JPEG image or a combination of an image and text, where the text is obtained from the parts of the image containing text. This is called a PDF with image and searchable text. The other forms are called PDF text or PDF image only. PDF files are only slightly larger than the native format of the original document. The major PDF advantage is that it is a de facto document standard and is universally accepted around the world. PDFs may be viewed or printed by free readers, such as the popular Adobe Reader.

Document Scanners

Page scanners come in several different sizes, shapes and brands with many different capabilities. Document scanners typically range from small low or medium speed models with document feeders or flatbeds to large, high speed, high volume production scanners.

Flatbed scanners scan one page at a time by placing a single page on a glass platen (bed), like a copy machine, and pressing a button either on the scanner or in the scanning application. Lower speed (Desktop or Workgroup) scanners typically scan from 10 to 60 pages per minute (PPM), and usually have a 35 to 100-sheet automatic document feeder (ADF). Medium to high volume scanners, (Department or Production) scanners can handle a stack of 50 to 500 pages in the automatic document feeder and scan at rates up to 200 PPM.

Many scanners can scan both sides of the paper simultaneously and are known as *duplex* scanners. Duplex scanning is double the scanning rate of simplex scanning and that rate is known as images per minute (IPM) rather than PPM. Scanners that can scan only one side at a time are known as *simplex* scanners. Most scanner manufacturers offer a simplex scanner and an equivalent duplex scanner at a slightly higher cost.

Most of the home or workgroup scanners can scan business card sizes to letter size (8.5 x 11 in. or A4) paper. These scan letter size pages in *portrait* orientation where the top or smaller dimension of the page is fed into the scanner first. Some of the workgroup scanners also can handle legal (8.5 x 14 in) size documents. Many of the department and most production scanners can handle 11 x 17 in. or A3 paper sizes allowing letters size documents to be scanned in *landscape* mode with the long edge entering the scanner first. This decreases the transport distance and therefore the time for each page to feed through the scanner. In turn this increases the scanning speed for letter size documents by almost 30%. These scanners usually are rated in landscape orientation. *Large Format* scanners can handle architectural sizes B, C, D and E drawings and scan at a fairly low rate because of the paper size.

Modern scanners scan in bitonal (black and white), full color and grayscale. Some scan in color or grayscale and bitonal simultaneously which is referred to as dual-stream scanning.

A number of scanners support third party image enhancement products and greatly improve the look and consistency of older, damaged or poor quality documents. Others have their own image enhancement technology built into the scanner.

Some new copy machines contain printers and scanners, and are knows as Multi-Function Printers or Multi-Function Devices (MFPs or MFDs). Some of the newer models may be used in place of scanners, provided they meet some criteria: they must support TWAIN drivers and must be able to scan fast enough to meet your needs.

All scanners are designed for a specific *duty cycle*. Duty cycle is the number of pages per day which a scanner can scan reliably without abnormal wear of parts, dirt or ink buildup, or accelerating parts deterioration. A 60 ppm scanner, for example, may have a duty cycle of 5,000 to 7,500 pages per day. Duty cycle is related to the scanner warranty and is a number that should not be exceeded or you risk invalidating the warranty.

Finally, scanners are priced generally according to scanning speed, duty cycle (pages per day) and paper size capability. Higher speed, higher duty cycle scanners are the most expensive. Lower speed and lower duty cycle are the least expensive.

Choosing a Scanner

Scanning a large number of existing paper documents is known as *backfile conversion* or *backfiling*. Scanning document which arrive on a daily basis is known as *day forward* scanning. Both types of scanning must be considered in choosing a scanner.

What is your timetable to complete all of the backfiling as calculated in Table T-2 above? Days, weeks, months? Convert this time into workable hours. To determine the speed of requirement for your scanner, divide the total number of pages to scan by the number of hours allocated to obtain the scanning rate per hour. Divide this by 60 to obtain the ideal scanning rate. Then multiply this number by two to obtain a practical value of pages per minute for the scanner.

The reason for the factor of two is the practicality of achieving 100% of a scanner's advertised rate every minute of a work shift. The published scanner rates in pages per minute are valid, but in practice users spend time preparing documents, removing staples, paper clips, sticky notes and orienting pages. Documents must then be ordered, stacked properly in the feeder, scanned, removed from the outfeed tray and placed back in boxes or folders. Scanned documents may be identified or indexed, although most quality applications allow this to be accomplished post scanning. Occasionally documents jam in the scanner, which requires clearing the jam and rescanning.

Sample calculation:

Total pages to scan:	240,000
Time allocated:	20 days or 160 hours
Pages per hour:	240,000 / 160 = 1,500 pph
Required ppm:	1,500 / 60 = 25 ppm (pages per minute)

Practically you cannot achieve 100% efficiency, i.e. scan at 25 ppm for 160 hours to achieve the allocated time. Therefore, use a factor of two to increase the scanner speed requirement to achieve a practical scanner speed rating:

Practical scanner size = 25 ppm x 2 = 50 ppm.

Note 1: If this number exceeds 60 pages per minute consider extending the amount of time allocated for scanning all pages. Purchasing scanners with speeds greater than 60 ppm tends to be quite expensive.

Note 2: Consider purchasing two or more scanners to share the scanning load. This may offer some flexibility in locating scanners in different parts of the office.

There are three more considerations relating to scanner choice, the first being the duty cycle of the scanner. Although you could scan more pages per day, the duty cycle should be the limiting factor.

The second consideration is that it is possible to *outsource* scanning of your paper documents to a scanning service. Scanning services use production level scanners which have an unlimited duty cycle, and have organized their shop to achieve very close to the rated speed of the scanner hour after hour. Depending on volume, indexing and identifying need guidelines the cost for production scanning is nominally $0.10 per image or less.

The third consideration is the amount of day forward scanning. Estimate your day forward scanning load according to Table T-2. Then perform a scanner sizing calculation to determine a practical scanning rate in ppm. This number becomes another factor in your choice of scanners. Here are some rules to help with the calculations:

1. If your day forward scanning rate is above 30 pages per minute, consider multiple scanners instead of one or two faster scanner. This will make day forward scanning more convenient to many people in your office and will only slightly increase your overall costs.

2. If your day forward scanning is a small fraction of your backfiling, consider outsourcing a portion of your backfiling.

At this point you should have solid criteria on which you can base your scanning choice. An additional recommendation would be to pick a quality scanner manufacturer with good support and warranty provisions.

Computer Requirements for Document Capture

How do the possible scanner choices relate to modern computers? How powerful should your computer be to run a scanner at rated speed and do all the processing you require of it? How much memory does it need? How big should the hard drive be? What operating system is required? Here are a few guidelines you can use to match up your scanner and computer.

A general rule is that the widest selection of scanners and scanning applications are available for computers running the Windows XP, Vista, Windows 7 or later operating system, with Pentium processors.

Another rule: the faster your scanner, the faster your computer should be. A 15 – 20 page per minute (ppm) letter size scanner with an ADF or a Flatbed scanner requires only a computer with 1 – 1.5 GHz processor speed. A 30+ ppm workgroup scanner may require a 2 GHz (giga-Hertz) processor. A department scanner at 60 PPM may require a computer with a 2.4+ GHz single or dual core. Department or production level scanners at 100 to 200 pages per minute requires a very *multi-core* processor with speeds of 2.8 to 3+ GHz.

All computers controlling scanners can utilize lots of memory. Typically, Windows XP systems run very well with 768MB to 1 GB of RAM (random access memory), but Windows Vista and Windows 7 should have 2 GB. In all cases doubling the RAM to 4 GB is desirable. At present, 64-bit processors are of little benefit for most scanning projects, and may not improve performance. Most computers purchased within the last few years have more than one processor core. The processor is divided internally into one, two, three, four or more individual processing units called cores. These are loosely referred to as *multi-core* processors and can be a great advantage to capture applications designed to use multiple cores. In general, a dual core computer will run significantly faster than a single core computer with the same processor speed and a quad core is faster than a dual core system.

Document Capture Programs

There are many document scanning or capture applications on the market. Some are very basic applications with very simple features for casual scanning. Others are complex applications with many features and are quite expensive in addition to requiring significant user and administrator training. Some scanning applications are included in comprehensive, content or document management applications which are often used by large businesses and enterprises. These applications may use proprietary or third party databases for management and retrieval of your documents, increasing the functionality of the product at the expense of database management. Very few document capture applications fall between the low and high ends, yet offer high end and professional capabilities at reasonable prices.

The basic functions of a document capture application are to control the scanner by utilizing all or most of its capabilities, to separate single or multi-page documents from a stack of pages in the feeder, automate document identification as much as possible, provide Optical Character Recognition if needed on some documents, and store the scanned documents in a repository for future access.

The basic requirements of a document capture application for your business should fit with the number of documents to scan, the scanner meeting your needs, and the repository you favor priced within your budget. Beyond these, you expect the program to use the latest technology, be user friendly, cost-effective, highly secure and offer flexibility in your scanning jobs. You want capabilities beyond what you currently expect so as you learn more and want to accomplish more, the product you purchased will be able to handle your needs, and lastly a product that will evolve as technology evolves. Most of all, you want to protect your documents and use a capture program and repository which will ensure that your documents are secure, safe and always available.

Digital Document Storage

First you must estimate the amount of digital storage needed for all of your documents. Assuming you want good quality using 200 dpi which is used most often in the industry, figure 40,000 bytes per page. Double this if your pages are predominantly double sided. Also double the amounts if you need much higher quality by scanning at 300 dpi. When done, divide the number you computed by 1 billion (1,000,000,000 or to be exact 1,073,741,824) to provide your storage needs in GB (giga-bytes). Finally, estimate of the annual growth rate of your documents for the next 5 years. This is difficult, but you could base it on the current growth rate of your company.

Now a few questions about storage or the repository:

1. Does your organization have secure, reliable, frequently backed-up computer systems or networks containing sufficient excess capacity on hard drives? Do you believe that if your documents are stored in your computer system or network, they will be safe for as long as you need them? Local or network storage may be the most inexpensive and fastest storage available.

2. Does your organization have fast and reliable access to the Internet? Do you believe that with the right security precautions, information sent across the Internet is safe and secure? Do you believe that under the right conditions you can trust large internet companies to store your documents securely and reliably? If your answers are generally "yes" consider storing your document on a repository or content management system in the Cloud. Consider Google Docs, Box.net or Dropbox.

3. Does your organization use Microsoft SharePoint for document collaboration among members of the SharePoint sites? Would Windows SharePoint Services (WSS), SharePoint Server 2007 (MOSS), SharePoint 2010 or Office 365 SharePoint Online be a possible location for some or all of your document storage? If so there is an immediate advantage of using SharePoint for document storage because you already have it running, know how to manage it, and can instantly share documents among your associates and workers.

4. Are you using an existing custom or third party Content Management Systems CMS that is already setup to handle storage of many documents?

From these few questions you can decide which storage approach is best for your company. You can even use more than one scanner if you need higher availability and flexibility.

Accessing Stored Digital Documents

Now that you know a little more about digital document storage, scanners and document capture programs, how can you access all of these digital documents? Once they are stored how can you view, print, fax, email, copy, share, edit or even delete them? All of these questions are significant and crucial to making a decision to Transform Your Office. As previously stated, the primary function of document capture programs is to control the scanner, separate pages into documents, identify documents, provide Optical Character Recognition if needed on some documents, and finally store the documents in a repository.

To access these documents, you need to have an application capable of locating the desired document easily and quickly so they may be processed as needed. This could be as simple as using Windows Explorer to navigate to your main image folder and "drill down" though the hierarchical identifying folders to the name of the document you want. Once available in Explorer you can do almost anything with it that you would do with a paper document.

If you decided to store documents on Cloud systems like Box.net, Dropbox, Google Docs, or more you can use a browser or other recommended products to view, print, email or even share documents with others inside your office or around the world.

If you company uses Microsoft SharePoint and you decided to store digital documents there, you or any SharePoint site members can access your documents. Similarly with Box.net. Simply view, print, email, or share documents through the familiar user interface.

Finally, if your company is using a custom or third party Content Management System, check the requirements for importing or scanning directly into the CMS.

Our Document Capture Solution

Capture Components LLC was founded specifically to offer document capture products and components which will meet most if not all of your requirements. We have drawn from our experience in the business world, the document imaging business, and many hours designing and programming sophisticated software and hardware systems to create our **ccScan** product line. In every step of **ccScan's** development we tried to think of the end user's need for performance, security, reliability and ease of use.

ccScan was originally designed to take advantage of processor speed, memory and the number processor cores as well as disk access speed and in some cases Internet and intranet bandwidth. Depending on your capture requirements, **ccScan** can efficiently use up to 4 processors simultaneously which dramatically speeds up performance of all functions in the document capture process.

All existing document capture and document management applications, with very few exceptions, do not take advantage of the multi-core processor capability of new computers. They give up this advantage by hoping the operating system will make their application a little faster. But since you will pay for the whole computer your capture application should take advantage of all processor cores. **ccScan** does this elegantly, reliably, transparently and inexpensively.

ccScan currently provides secure communication to document repositories such as Amazon S3 and Windows Azure. **ccScan** also can store documents directly into Content Management Systems like Microsoft SharePoint, Box.net, Dropbox, Google Docs, Docstoc and Scribd. You can also store documents on local or network shared drives or folders or removable media. In all cases you may store documents in encrypted PDF format if needed.

From a reliability standpoint, we have one unbreakable rule: Images and electronic renderings of your paper documents are 100% yours. **ccScan** does nothing to prevent your access of these documents once scanned and stored. No proprietary formats, encryption or cryptic file naming. Finally, **ccScan** was designed to make life easy for users, both for your needs now and in the future.

If you would like to learn more about **ccScan**, please return to our website. If you choose other products, we hope this white paper has provided some value in your quest to improve paper flow in your office. We ask only that you take a good look at **ccScan**. If you have any questions, please visit our Sales or Service Forums, or send Sales an email, all of which are accessible from the web site.

Now, on to **Transforming Your Office**.

Let the transformation begin

Use this <u>Transform Your Office Action List</u> to determine your next step:

- ❑ Before any other decisions, you must make a firm commitment to use document imaging to reduce your paper load and improve your document workflow.

- ❑ Decide to concentrate on backfiling or day forward scanning first.

- ❑ Based on the above decision<u>s</u>, carefully investigate brands and models of scanner(s), storage repositories, and document capture applications.

- ❑ Make decisions on the selected scanners, repository and capture applications.

- ❑ Set up a pilot document imaging project to learn first-hand about the capture process and its methods such as document indexing and identification, choosing document formats, and deciding when to use OCR. Select a <u>person who is </u>computer-wise ~~person to run this project~~and ~~to~~ understand<u>s</u> <u>the </u>existing paper sources and paper flows.

- ❑ After learning from the pilot project, invite <u>other</u>~~different~~ members of your office to participate in the pilot project, to ask questions, ~~to~~offer comments and best of all ~~to~~ share your vision for higher productivity.

- ❑ Also after the pilot experience, plan an approach to the various categories of paper files, for example, departments, subjects, older documents first, day forward documents, first and most recent backfiled documents next. A small amount of serious planning will greatly smooth the transformation

- ❑ Be conservative with your plan. Start with just one work group or department and then expand to others. Be careful of attempting too much at first.

- ❑ Train at least two capable employees in the transformation process. Commit to their success. Provide the management and technical support they will need.

- ❑ Carefully review the transformation as it nears completion. Invite managers and workers to see demonstrations and participate in basic document capture training.

- ❑ Finally, Transform Your Office and enjoy the rewards that thousands of companies have already experienced with this transformation.

Afterword

We at Capture Components hope that this White Paper is a benefit to you. If it is and document capture helps you Transform Your Office, then we have achieved a measure of success. If you choose ccScan as your document capture application we will do our best to provide the highest quality product and support.

2012-02-07