



# **The Advantages of Digital Archiving**

## **Executive Overview**

Providing long-term access to an organization's paper and electronic records is becoming increasingly important. Government mandates such as HIPAA and new regulations from agencies such as the SEC require more careful tracking of many types of documents. The cost to securely and accurately maintain long-term records using existing paper or electronic systems is very high. Moving to a fully digital archive can reduce ongoing operational expenses more than 80%. In this paper we will discuss some of the operational and financial benefits of using a digital archive service.

---

## The cost of managing paper files

How much does it cost to store paper documents and manage them? Handling paper documents is such an ingrained part of a professional's job the true costs are hidden; we all just handle the paper as best we can. In fact, it is estimated that 8 hours out of every professional's workweek is wasted in paper document management, ½ hour of which is related to archival and retrieval.

In a typical office, a contract is signed by a principal, and then handed off to admin for filing. A bid is put together and then filed, a letter is sent to a customer and a copy is filed. The result is rows of filing cabinets filled with documents that are the lifeblood of the organization. Gartner Group estimates that the average company spends \$20 to file a document. A typical filing cabinet holds 1,000 documents or more, filling one cabinet can cost as much as \$20,000 per year.

What happens when you need to refer to a filed document? Either the principal or the admin goes to retrieve the appropriate file. Most of the time the file is right where you expect it to be, but Gartner found that 3% of all paper documents are misfiled, and 7.5% of all documents get lost or are never filed. In addition to the disruption caused when you can't get access quickly to the information there is real expense associated with finding or recreating files. On average, the cost of tracking down a misfiled document is \$120, and reproducing a lost file is \$220. Applying these costs and percentages to our "typical" filing cabinet of 1,000 documents adds another \$20,000 or so to the bill, for an annual total cost of \$40,000.

Why all the expense? Most of it comes from the fact that we need to physically handle the paper, create file folders, and sort out the pile of documents as they accumulate. A major source of errors is that a document can only be in a single file. The contract can be filed in the contracts drawer, or the product drawer, or the customer drawer, but not all three. Perhaps service contracts are filed separately from product contracts. When the need to review a transaction arises, it means that someone has to go to the customer file to retrieve correspondence and the contracts file(s) to pull the relevant contract. Occasionally, we forget how we filed a document so can't find it, or we forget to re-file the information after we have used it. In short, maintaining paper files is a manually intensive task, and understandably prone to error.

---

## **The digital alternative**

By converting paper documents to electronic format and storing information digitally an organization can drastically cut down on the cost of maintaining accurate files, and improve productivity for principals and administrative staff. To understand how digital archiving helps reduce costs, it is useful to consider a few points.

Archivists use the concept of meta-data to distinguish between a physical record (any piece of information that needs to be saved) and information about that record. Think of the distinction between a book on a library shelf and entries in the card catalog that help users locate the book. There are card catalog entries for the author, the title, and possibly multiple subject entries. These are the meta-data associated with the book. In our discussion of filing contracts, meta-data would be the customer name, the date of the contract, the product or service discussed in the contract, perhaps the name of the sales person who closed the deal, and (very importantly) the fact that the document is a contract.

Archivists are concerned about maintaining the legitimacy of a record: ensuring that the document that gets pulled out of the archive is in fact the original document and has not been changed. They also want to make sure that important documents don't disappear. While intentional theft or destruction of enterprise documents is rare, it does happen and represents a real liability to management, and a solid audit trail will cut down on potential abuses.

Within an archive, it is important to maintain a record of who has accessed the document. This is most useful when trying to find a document that someone else has pulled out of the file, but it is also helpful when looking for security violations.

Lastly, an archive must be protected against physical loss through fire, vandalism or other disaster. Most organizations recognize the importance of protecting long term paper records by placing them off-site in climate controlled and fireproof warehouses. But the data that is physically on the premises is the most critical information that a company has.

While all of the above points could be addressed with business process changes on how an organization handles paper files, the solution could easily double the cost of existing methods.

## **Operational benefits of a digital archive**

A digital archive can deliver all of the above capabilities while reducing the total costs of administering a company's vital information.

---

As part of the digital conversion process, an enterprise can create custom metadata to describe the documents being filed. Typical index fields are customer name, document type, subject or any other important information that can help locate the document quickly. When looking for a document, you can enter as much information needed, such as show me all the contracts that we have with XYZ Company. Or, show me all of the correspondence with XYZ Company related to Product A. The second search would not only return the contract, but also proposals, quotes and general communications. The effect is as if any particular document was stored in multiple folders. Once the records have been identified, the user can browse the documents on the system, download them to their local computer, or even print them out. This has a number of obvious benefits:

- Retrieving a document electronically is much faster than pulling a physical file;
- Misfiling is much less likely as files cannot be physically misplaced;
- Documents are less likely to be lost as only copies of the file are retrieved; the original is always in the archive.

The last point also helps ensure legitimacy. Because the original of the document always stays in the archive, the enterprise has absolute control over what happens to it. Security features such as password protection, digital signatures and user specific permissions provide a robust defense against theft or alteration of the original document.

Because all of the requests flow through the archiving system, it can automatically track who has viewed or downloaded a particular file. While such an audit trail is not needed to find missing documents, it does provide an additional measure of security. By tracking how often a particular document is accessed and by whom, the enterprise can also glean information about workflow patterns that can lead to more effective organization of operations.

A digital archive can also be easily preserved. The system can be backed up to tape and stored offsite, or the entire process of archiving can be outsourced to a 3<sup>rd</sup> party provider so that the data is not stored within the enterprise. Outsourcing provides an additional level of security that we will explore in more detail.

## **Replacing more than paper**

A digital archive does not have to be a repository just for scanned images. The same technologies that enable secure storage and easy access to images can also be used

---

to preserve records that began life digitally. Word documents, PDF files, JPEG images, and e-mails can be stored within the digital archive as well. By consolidating all of an organization's long-term records in one repository, a single search can retrieve related information such as contracts, proposals and e-mails, greatly simplifying tracking the business or responding to external audits.

Most enterprises have installed storage systems to support their day-to-day operations. These systems are optimized to provide lightning fast access to critical operational information, and to support 1000's of transactions per second. The amount of data stored this way has been (and is expected to continue) growing at rate exceeding 50% per year. Currently, Gartner estimates that the total cost of enterprise storage is 10 times the cost of storage hardware. That puts the total cost of storage at around \$1,000/GByte/year. But studies have shown that at least 75% of the information in these systems is no longer live data. Just as paper records are cleared out of filing cabinets and sent to the warehouse, data should be moved out of enterprise storage into long-term archives. Depending on the size of the archive, the annual cost of an outsourced service ranges from \$300 to \$600/GByte, or up to a 70% reduction in costs, enabling organizations to lower their overall spending on storage. This is in addition to the added security and ease of retrieval of information from the archive.

### **Operational cost of a digital archive**

The costs of converting to a digital archive are quickly repaid in lower ongoing costs. There will be investments in hardware and software and most importantly operational costs. But these costs are at least 80% lower than the paper costs shown above, with an ROI of six months or less! Upfront costs associated with conversion are low; a scanner capable of scanning 50 pages per minute with associated control software costs less than \$8,000. To understand the ongoing costs of conversion, we will examine the process in more detail.

Creating a digital record from an existing paper document has three steps: scanning, indexing, and verifying. Scanning covers the preparation of a document for scanning (e.g. removing staples) and physically running the document through an optical scanner. Indexing refers to the creation of meta-data descriptors such as customer name, document tracking number, creation date, etc. Verifying means reviewing the output of the scanning process to ensure that the file has been correctly scanned and is still legible. The cost of all of these steps can vary depending on the degree of automation and complexity of the document, but a reasonable estimate is that a 4-

---

page document would likely take 10 minutes or less to scan, index and verify. Assuming a clerical worker carrying out the scanning operation is paid \$20/hour, this translates to around \$3 per document. As a reference point, Kofax Imaging Products, the leading supplier of scanning software, estimates that even with low levels of automation scanning, indexing, and verifying takes about 19 seconds per page, or just over a minute for a 4-page document. For our cost comparisons, we will use the 10-minute estimate.

## **In House or Outsourced storage**

The other cost component of a digital archive is storage. An enterprise can decide to store the electronic archive in house, or outsource the archive to a third party. An outsourced solution offers benefits in three key areas: upfront expenses; ongoing operational costs; and security.

There are only minimal set-up costs for an outsourced solution. Depending on the level of performance and reliability an in-house enterprise level solution will cost upwards of \$70,000 in hardware and software. With a hosted service, there is no hardware to purchase, and if the service is delivered using standard web interfaces, no software to install.

As mentioned above, on-going management and operations is the major cost component of storage systems. While a storage system that is optimized for archival can lower the hardware cost by perhaps 50% compared with the cost of enterprise storage, the internal administrative costs remain the same. This translates into an annual expense of around \$850/GByte of storage for an archival system. Because a service provider can spread administrative and management costs across a larger installed base, annual service charges are in the range of \$300 to \$600/GByte, up to a 70% reduction.

While many would assume that a hosted service would be less secure than an in-house installation, the opposite is the case. According to Matt Barzowskas, vice president and research analyst at First Albany in Boston, "The No. 1 threat to corporate security is internal problems." Anything from accidental penetration of the network to deliberate acts by disgruntled employees can compromise your systems. (Boston Business Journal, March 8, 2002). According to recent research commissioned by Oracle, internal security breaches present the largest threat to compromising data within a company. (Internet World, June 15, 2003). By moving archiving out of the enterprise, the risk of intentional destruction or alteration of important records is greatly reduced.

---

## Costs – A simple example

To compare the difference in costs between paper and digital archiving, let's return to our "typical" filing cabinet. As we showed above, the cost to fill the filing cabinet costs about \$20,000, and the lost productivity due to misfiled or lost documents subtracts another \$20,000 from an organization's bottom line. What would a filing cabinet's worth of documents cost to digitize?

As mentioned above, an estimate for the cost of scanning a 4-page document is \$3.00, and about 1000 documents are in the filing cabinet, it would cost \$3,000 to scan, verify, and index the entire cabinet. An individual scanned page requires about 50,000 bytes of storage capacity, so the cabinet would need (1000 documents x 4 pages/document x 50,000 bytes) about 200 million bytes of storage or .2 GBytes. At \$600/GByte/year, this amount of storage would cost \$120.

## Summary

Converting paper documents into digital format reduces annual ongoing paper management costs from \$40,000 to less than \$4,000, over a 90% reduction in ongoing expense. With monthly savings of \$3,000, the \$15,000 spent initially to install a scanning workstation and set-up the archiving account would be recovered in 5 months.

Digital archives are also more secure, more easily searched, and provide for better preservation of records to ensure their long-term viability and legitimacy. A hosted archiving service costs less, is easier to implement, and provides additional levels of security. Without the need for costly hardware and software installation, a hosted service starts delivering profits faster, for an improved ROI.

For more information about our archiving services, please contact Image Fortress at: +1.978.399.0248, or [sales@imagefortress.com](mailto:sales@imagefortress.com).