CRAWFORD TECHNOLOGIES

Migrating Customer Communications Archives

Everything you should consider for a successful archive migration

Introduction

Customer communications drive the workflow and success of nearly every organization. Without transactional documents like invoices, statements and acknowledgements, most businesses would simply not operate. And for those organizations in particularly document-intensive industries like financial services, healthcare or insurance, the critical nature of customer communications is an underpinning of C-Level concerns surrounding regulatory compliance, financial disclosure and legal discovery. The production and distribution of many of these documents are mandated by regulations to communicate the status of the financial relationship between the company and their customers.

As a result, a wide variety of approaches and technologies have been developed over the years to save and archive critical customer communications. Everything from simple scans of filing cabinets full of paper, to advanced digital enterprise document management systems are used to secure, share and store the ever-growing convergence of customer communications that drive the fundamental workflow of business every day.

Next Generation Approaches

We find that organizations today are looking for more. They have a desire to not just store and archive critical customer documents, but also to leverage advancements in content management systems to be more agile, responsive and effective in the face of digital transformation. That's where we come in. Our expertise in both legacy and modern-day document systems is unparalleled and we understand the business implications of information management. CrawfordTech has the tools and expertise to help you effectively transform your legacy archives in ways that make a real difference.

Steps to Make the Move

Let's face it; a lot of customer communications archives are old, obsolete, and costly and complex to maintain. It is for those reasons, along with the desire to boost information governance, compliance and to enable better business process management that organizations decide to make the move to a next generation document archive. We work with them all, and we have a great deal of experience migrating critical customer communications from one archive to another regardless of which one you choose.

But navigating a successful archive migration requires more than technology; you must also have the right information, perspective and strategic vision to achieve your objective. It is for this reason that we have created this whitepaper to provide a guide to the most important things to consider for a successful archive migration. Use the concepts, approaches and structure to evaluate, plan and execute your migration effort.

> **"Organizations today are looking to not just store and archive critical documents, but also to leverage advancements in content management to be more agile, responsive and effective."**

Extraction

Any customer communications archive migration starts with a simple goal: Extracting documents and data out of the current system and placing it in the new system. It sounds easy, but the scope and complexity of your migration can vary greatly depending on the type of archive you are currently using. When it comes to extraction strategies, one size does not fit all; so it is important to begin your effort by carefully evaluating the factors involved with extraction.

Scoping the Migration

It can be tempting to jump right in and begin moving things before properly scoping the size, shape and complexity of the migration, but it is vitally important to take the time and make the effort to do a thorough discovery and analysis. Otherwise, you run the risk of bringing forward inefficiencies and redundancies that will make the entire migration process more cumbersome, time consuming and prone to error.

Formats

In many organizations today it is not uncommon for a customer communication archives to have been in existence for a long time and it's likely that there will be some surprising and challenging content in it. Even if you're distributing your customer communications electronically, on the web or in e-mail, it's still common to generate a "print file" to store in an archive. So begin your discovery by asking questions like:

- What types of data and documents are stored in the current archive?

- Are they all print files? And what format are those print files: AFP? Xerox Metacode? PDF?

- Is the information in the form of old reports, line data or text files?

- What surprises could the archive contain (business documents, e-mail, audio/video)?

With a complete inventory of the files and formats contained in your archive you will be in a much better position to plan the migration properly.

The scope and complexity of your migration can vary greatly depending on the type of archive you are currently using.

Document Types

Transactional documents and customer-facing correspondence come in many varieties and classifications. Some documents are frequently referenced while others live a life of obscurity because they were saved and forgotten years ago. On the other hand, many documents demand hyper-vigilance to ensure both security and privacy, while others need to be easily accessed and frequently revised. Different industries have different retention requirements that determine what information should be archived and for how long. Consider the following steps to map your document types:

- Identify each document type in the archive (e.g., statement, letter, confirmation, contract, etc).

- Assign a level of importance for each one (e.g., 1 = High, 2 = Medium, 3 = Low).

- Determine a strategy for handling the outliers (e.g., odd text files, business, desktop or alternative formats).

- Anticipate future needs for analytics and optimization (e.g., "Big Data" initiatives, BPM).

Sizing the Migration

The next step is to determine how much content you've got to manage and migrate. This is where some crucial and strategic decisions should be made regarding exactly what gets taken on board into the new system and what is left behind or eliminated.

Target the Content

It's unlikely that you'll need to migrate every single file and every single document; only the most recent correspondence and high-value documents may be needed. Consider these questions as you size your migration.

- Do you need to move all the content from the current archive?

- Have you been keeping absolutely everything in the past, but don't need to?

- How can you make thoughtful decisions regarding what to retain moving forward?

- What documents are you obliged to retain to be regulatory compliant?

- Are there ways to build efficiencies and eliminate wasted resources along the way?

You can use your assigned level of importance from above to help guide your sizing efforts, but there will certainly be any number of other factors that will influence your decisions and strategies. The important thing to remember is that without a clear and factual understanding of the size of the targeted archive migration, you'll be less prepared to plan for adequate time and computing resources to get it done correctly. And you will be more likely to leave something important behind by mistake.

Without a clear idea of the size of your archive migration, you'll be less prepared to plan for adequate time and computing resources

Time Requirements

The time that it will take to complete your archive migration will also vary greatly depending on the size and scope of your effort and the amount of computing power, temporary storage and throughput available. The extract and migrate process is very I/O intensive so it is important to perform a statistically sound sizing analysis because hitting the production systems with highly CPU-intensive task in the middle of a key batch process or high demand window is not something business users will accept.

Click here for a template spreadsheet you can use for more detailed analysis. Be careful that you keep the units consistent.

Whatever method you use to calculate the time requirements for your archive migration it is important to consider that you are unlikely to be given computing access for an entire 24 hours in a day. Therefore, if you have access to the mainframe for only one hour per day, 24 hours' worth of work is going to require at least 24 days, not 24 hours. Parallel processing may be an option to reduce processing time, but the trade-off is that you might encounter I/O throughput constraints.

Extracting from the Source

The methods used to extract files and documents from your current archive will vary depending on the source. In some cases it may be as simple as using an unload command, or if your archive was developed in-house you may need to write API's on your own. More commonly with larger enterprise systems the tools needed will be provided by the original archive supplier to extract your data but protect their intellectual property. In short, extraction varies source system by source system, so consider these questions carefully.

- What tools are needed for the extraction?

- Does the current supplier provide me with the tools?

- Is an API available, or will one need to be written.

- Do you have tools already on hand that are now reusable?

- Do you need to license more advanced tools for the period of the migration?

Staging Content

During the migration process you will need space upon which you can store the data as it moves from the previous archive and is written to its new destination. If you've only got 20GB disk available you will likely experience a very tight constraint at in the core infrastructure. In our experience you should count on the need to a couple of terabytes, possibly more, of disk space and consider this in your planning.

Also consider the need for appropriate governance controls in the extraction process. Your particular information governance policies should determine the balance point between two potentially divergent organizational goals: extracting the information from the archive and reducing the potential risk of that information being lost, stolen or corrupted. Use a consistent and logical framework to handle data through according to organizational information governance policies, best practices and procedures.

The methods used to extract files and documents from your current archive will vary depending on the source.

Transformation

You must determine which format you will use to store the documents in your new archive. You may elect to retain the original format or a more flexible format may be desired that will necessitate a transformation to incorporate support for accessibility or electronic distribution and presentment.

Another consideration are resources like fonts, images and machine-control metadata which can be overlooked and are especially important for legacy print streams which rely on these resources to be resident at the printer itself. Consider these questions:

- What format will be used for the new archive?

- How will these documents be accessed; in print or online, or both?

- Will all documents and files be in the same format, or will there be a mix?

- Where do all the fonts, graphics and other resources fit in the transformation?

- Will you convert everything to PDF or some other common format?

- Is there a requirement for 100% match in terms of document fidelity?

Optimization

It is not uncommon for legacy archives to contain things like duplicate pages, blank pages, and other arcane resources that are either obsolete or no longer needed. As a result, consider a number of modifications and transformation that can dramatically optimize the new archive system:

- Remove all duplicate pages and reduce the storage requirements of the new archive.

- Eliminate and optimize the resources and improve storage and retrieval time.

- Be able to present fully accessibly documents for blind and partially-sighted customers.

- Ensure additional levels of encryption to increased defense against data breaches, and cyber theft and to be complaint with applicable regulations for your industry.

- Use digital signature encryption, and records management technology and other techniques to protect against internal fraud and ensure proper governance.

- Automatic redaction to keep sensitive data hidden while otherwise providing more agile information access overall.

Indexing

All document management systems have some level of content indexing. Planning proper indexes is one of the key factors behind a successful document archive migration. Default system indexing include elements like Date, Document Type, Name, Account Number, or other common identifiers to ensure that the documents and files can be found in the new archive system. There are additional indexes that enable more advanced functions like full-text search and more advanced classification opens the door to Business Process Management applications and Big Data analytics. Consider the following for your migration efforts:

- Are you going to use the same indexes from the previous system?

- Will new indexing parameters and approaches be needed?

- Are you introducing a master data management system and therefore all the indexes need to be mapped to this new paradigm?

- How can a more advanced indexing schema bring more insight from the information contained in the new archive?

Load the New Archive

As you prepare to load the new archive remember that each system will have its own technique for ingesting the files and documents. Some archives use Web APIs for document ingestion others batch loading from specially formatted files. All repositories have their own set of tools and techniques, so be sure to consider all the requirements a head of time.

Transition State

Consider what state your system will be in when you make it available for business users. Will the entire archive be a part of the new application environment and all of your migrated documents made available? Or are you going to back-load the most recent segment (e.g., the past two years) and slowly bring across older content as needed over a period of time?

Quality Assurance and Reconciliation

Quality assurance should be the first priority of any customer communications archive migration. Two quality factors should be considered: The quality of the archive migration, and the quality of the documents themselves.

Archive Migration Quality

How do you guarantee that what you had to begin with is what you now have loaded into your new archive? Is it sufficient to count your documents going in to the extraction process, count them through the staging and transform activities, and count them a third time as they move through the load process or are more sophisticated techniques required? Consider also that the source system may have developed errors and inconsistencies over time; the indices could be wrong, out of date or corrupted.

Quality assurance should be the first priority of any customer communications archive migration.

Document Fidelity

Many documents require a high degree of image fidelity; retaining the exact appearance of the original. When transforming from one format to another -- AFP to PDF, for example – it is critically important to take the quality of the transformation into account. One the other hand, some documents are for internal use only, or don't carry forward specific fidelity requirements in particular. In those cases it might not make sense to retain full-color or high-resolution images and reduce the load time latency to customer service representatives.

Records Management

A proper approach to managing your archive files and documents should include governance throughout their entire life cycle -- from the time they are created and stored, to the time of their eventual disposal. This includes identifying, classifying, storing, securing, retrieving, tracking and destroying or permanently preserving the records. These activities should fall under your organization's broader policies regarding Governance, Risk, and Compliance and be primarily concerned with the managing the evidence of an organization's activities as well as the reduction or mitigation of risk associated with it.

Disposition and Dates

Archive database suppliers make it easy to support a formalized approach to document disposition. Determine and stipulate when a file or document will be eventually disposed, if at all. For example: "This record will be deleted in ten years" or "six months after we receive notification of an account cancellation." By solidifying disposition policies and dates during your migration effort you avoid having to revisit disposition policies and procedures later on; or worse, and have your archive eventually become cluttered with years of content taking up space that is no longer needed and represent increased cost and risk to the organization.

Moving Forward

We hope this white paper will be a helpful guide to your assessment of the most important things to consider for a successful customer communications archive migration. Use the concepts, approaches and structure to evaluate, plan and execute your strategies and activities moving forward. While technologies play a big part, remember that proper discovery, assessment and planning are essential. And look for proven partners and expert resources, like CrawfordTech, that can assist you in making the most of your archive migration.

We rely on our STEP™ migration services to ensure a fast, accurate and reliable migration between repositories. STEP is based on industry standard offerings and employs defined processes to extract, transform and load content and indices from original source archives to the target repository. Available as an onsite or remote service, STEP is platform agnostic, and provides the visibility and control that you need to guarantee an efficient transition.

For more information just visit CrawfordTech.com or call 866.679-0864

CrawfordTech Solutions

Crawford Technologies develops software and solutions to help enterprises optimize and improve the secure and accessible delivery, storage and presentment of their customer communications.

With over 1,800 customers on six continents, CrawfordTech solutions and know-how enable the largest banks, insurers, healthcare providers, utilities and print services companies to use their existing technologies, documents and data in new ways. We help them navigate the challenges in leveraging legacy applications in the platforms and applications of the future.

CrawfordTech's products, services and domain expertise reside at the nexus of content, data, and output management and are essential components of our customers' digital transformation, output management and document accessibility strategies.