Large-scale image systems: USAA case study

by C. A. Plesums
R. W. Bartels

A large-scale, optical disk-based, operational image system for office-size documents has been implemented to support an insurance customer service application. Images stored on optical disk can be displayed on any of the more than 1400 workstations or printed on any of 22 printers. This system was the prototype for the IBM MVS/ESA™ ImagePlus™ product. Each day over 25 000 pages of incoming mail are scanned, stored, and delivered to users for processing. In addition, computer-generated data (soon expected to reach one million pages per day) are stored for display or print on image-stored overlay forms. The system is described in the context of any large-scale office document application. The discussion includes some of the business factors that created an environment for success and the business issues that led to the development of the system. The paper discusses the discoveries and lessons learned from use of two pilot systems. The authors conclude that the present level of technology makes this a good time to move forward with the installation of large-scale operational document image systems.

Historically the mobile American military officer had difficulty obtaining quality insurance. USAA (United Services Automobile Association) was formed as a member-owned organization to provide insurance and financial services to military officers, former officers, and their families. Serving the two million members and associate members has made USAA one of the largest insurers in the United States, the largest mail order company in America, and a diverse financial association with over $16 billion in owned or managed assets.

USAA members (customers) deal directly with the association, primarily at its home office—there are no agents in local communities. As a result, USAA has a huge quantity of mail and telephone calls. To support this workload, the management of USAA has an explicit and continuing policy to be a user of leading-edge technology. This includes the use of optical disk-based image systems to maintain customer files.

Currently USAA receives 80 000 to 125 000 letters each day. Over half are payments; others are directed to the different services, which include a bank, investment companies, claims, staff, and so forth; and over 10 000 a day are for the Property and Casualty Policy Service and Underwriting functions. The last is the area selected for the first image project. Each of the letters directed to this function may contain a variety of different types of information. It may consist of several pages of plain text, forms, appraisals, applications, and even the envelope that records the postmark date. The letter, considered a "document," ranges from one page to hundreds of pages, with an average size of 2.45 pages.

In 1982 a project was begun to study the image concepts and develop the requirements and plans for a large-scale document image system. Several offers to build a custom system were considered and rejected. A standard product was required for maintainability, reconfigurability, and expandability.

By the early 1980s, image technology was starting to emerge.

Concepts were tested and requirements were identified from a percentage of actual customer mail during 1984 to 1987, using a prototype 3M system and a FileNet™ system. In 1987 IBM was selected to implement the MVS/ESA™ ImagePlus™ system to fulfill these requirements.

The image system

Background. In the late 1960s, USAA recognized that it had a severe paper problem. The association was drowning in a sea of paper that was interfering with service to its members. In January 1969, the chief executive officer at USAA challenged the association to become paperless. Paper-handling procedures improved so dramatically during the 1970s that two attempts to convert the primary files to microfilm were unsuccessful; the savings did not justify the inconvenience of film.

By the early 1980s, image technology was starting to emerge. It was considered by USAA for the property and casualty insurance areas—such as automobile and homeowner insurance—which had the largest files, both in claims and in underwriting. The property and casualty underwriting files, serving sales, service, accounting, underwriting, and production, were selected as the flagship application for analysis and implementation.

When the research project began in 1982, 196 people were dedicated to delivering the mail and maintaining the 1.3 million files. By 1987, as full-scale implementation took place, paper-handling procedures had improved to the extent that only 161 people were required to sort and deliver the mail and maintain 1.7 million files. Today the image system supports about 2 million files, with a staff of roughly 40 users and technicians to operate and maintain the image system.

In the original system, active documents were stored in a paper folder, one for each member-customer. These files took 37 000 to 39 000 square feet (3500 square meters) of prime office space, an area larger than a football field. A crew of clerks continuously reviewed the files, looking at the contents of each file every year or two, to remove the older documents. The inactive documents filled 80 000 boxes in a warehouse-like records center, where they are stored for the remainder of their seven-year life.

The paper files were both labor- and space-intensive. The goals of the image system, therefore, were to

- Provide compact storage for the files. If all the files were placed on optical disks, and the disks were put on shelves, the shelves would take less than 100 square feet (10 square meters) of floor space.
- Identify and index each document on the mainframes, so it can be accessed exactly like any other corporate data, and can be used with *existing* business applications. (USAA could not rewrite all its programs simultaneously to use the new form of customer file.)
- Deliver the documents electronically. Immediate access through terminals located anywhere in the country improves service and productivity. The electronic movement insures correct delivery and reduces the chance for loss, as well as saving clerical staff.

Benefits of implementation. Image systems provide extraordinary advantages to users with large document filing problems. The first application at USAA supports 2000 professionals accessing 2 million files that will eventually encompass almost 2 billion pages. The scope and cost of such a system is very large, but the benefits are similarly great and fall into several major categories, as described in the following paragraphs.

Reduced support costs. Support costs are reduced by the savings in clerical staff, office (file) space, warehouse space, equipment, and supplies. These direct savings can be estimated with reasonable confidence prior to implementation of the project.

The direct savings were calculated to be $5 million per year for the first system at USAA. The cost of operating the image system over the first five years, including equipment, staff, and supplies, averaged $5 million per year at USAA. These direct savings were projected to pay for the system, but did not in themselves provide a sufficient return on investment to justify the system.

Improved productivity. The productivity of the professionals using the system is improved through immediate access to the files necessary to serve a customer, as well as better management and control of the work in process. Paper-intensive shops often claim 30 percent or more productivity improvement.

Images of documents require very large amounts of data compared to conventional data processing records.

These benefits are very large and important to the justification of the system, but cannot be projected accurately prior to the installation of the system.

At USAA, 85 percent of the work is received by telephone rather than by mail, and sophisticated on-line systems provide information to handle most calls without access to the files. Improving access to the files through an image system has a much smaller advantage at USAA than at many sites. USAA managers estimated that savings among the 2000 professionals who initially used the system would be 2 to 10 percent. Since this is a large group of professional users, the savings at USAA total about $1 million per year for each percent of productivity improvement, and will probably amount to an additional $2 to $10 million per year.

Improved service. Customer service is improved by the immediate, simultaneous access to files from anywhere in the association. All the information needed to serve the customer at once is always there. These savings are intangible but important.

Improved security. Document and file security is improved since papers are not removed; there are no missing files or misfiled documents. An on-line audit trail records the processing of a document, the modifications, and who handled it. Documents that are accidentally or maliciously deleted can be recovered. Off-site back-up copies of the records are practical. Secure and encrypted data channels can be used, if desired, providing even more privacy than is possible in moving paper. These benefits are also intangible, but are valuable even though it is difficult to assign a cash value to each one.

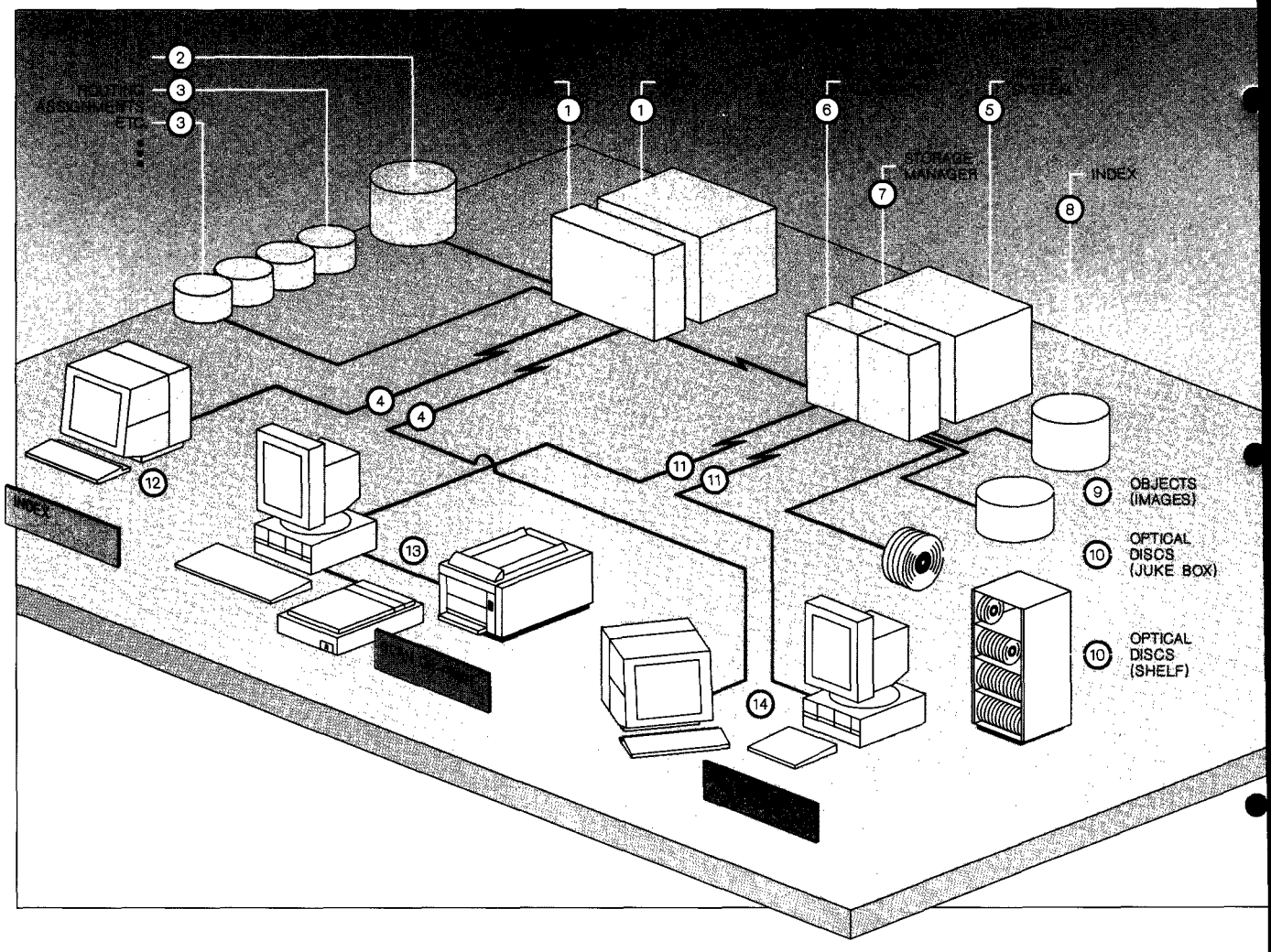
System architecture. Images of documents require very large amounts of data compared to conventional data processing records. Each page of a business-size document starts as a 500 000-byte record (about 4 million bits). After compression, each page averages 50 000 bytes and might range in size (depending on the amount of detail on the page) from 10 000 to 200 000 bytes or more. The several pages of a document (such as a two-page form or letter, and an envelope with postmark) are stored together as a single "object" that requires tens of thousands of bytes. For large documents, millions of bytes are required.

Not only is each image object large, but in a typical operational image system there are very large numbers of these records, so that the total amount of data to be stored is extremely large. The large-scale ImagePlus architecture allows this information to be stored and managed, if desired, on a separate computer system that can be configured and tuned to the different type of data.

In addition to the very large image objects, there must be an index describing each document. There will be tens or hundreds of millions of documents in an operational image system; the index database will have a comparable number of records. Despite its large overall size, this index is a conventional database in record size, performance, recoverability, and other requirements, and so is run on the regular mainframe computer.

The user interface to access image files needs to be integrated with the existing data processing systems, so users do not have to interrupt their work and move to a different system to retrieve data (as with most paper and microfilm systems). This interface can be through consecutive Information Management System (IMS) or Customer Information Control System (CICS) transactions, or through integration

Figure 1 MVS/ESA ImagePlus architecture



within the business application. In either case, the index database needs to be close to the business applications (on the same machine), rather than on a separate image computer system.

Figure 1 shows the general layout of the MVS/ESA ImagePlus architecture. The following numbered paragraphs are associated with the numbers in Figure 1.

1. The data system pictured in the center is the existing on-line data processing system, running CICS or IMS for other business applications. The image system is driven by the new "folder" or "case" transactions, or by calls from the application programs, that run on this system. If there are several processors, such as in regional computer centers, the image system can be run from all of the processors—generally from whatever CPU runs the business transactions for that customer.
2. The image index on the data system includes business information about all of the documents in the system, such as date received, line of business, customer number, form number of the document, description, and the internal codes (names) necessary to find the document. This is a very large logical database, and may be split into smaller physical databases by image application or customer.
3. There are several smaller databases on the same system, with information on the work waiting

to be processed, work waiting to be scanned, the qualifications of employees, supervisor assignments, and so forth.

4. Communications to the data terminals can be through existing data lines to existing terminals, or new terminals and communications approaches can be used. The image system will operate with any IBM 3270-type data terminal connection compatible with other on-line software.
5. The image computer system is shown in Figure 1 as a separate central processor. In a pilot or small operation, this can reside on the same system as other business applications, such as the data system (Figure 1, 1). In a large-scale production operation, images are sufficiently different that it may be more efficient to use one or several separate computer systems customized to image system requirements.
6. The distribution manager is a CICS application that receives requests from one or several folder applications on the data system (Figure 1, 1), finds the object on the specified storage manager (on the same or other machines), and sends that image to the workstation.
7. The storage manager is part of Multiple Virtual Storage/Data Facility Product (MVS/DFP™), providing operating system level management of and access to the optical storage system. It uses DATABASE 2™ (DB2™) to manage its index and other storage.
8. A separate index of each document is required on the image system processor, containing storage information such as the location of each copy of the image.
9. Objects (images) are temporarily stored on magnetic disk when they first enter the system, and sometimes during the life of the document when exceptional performance is required. They are generally on magnetic disk while they are queued in the routing system for processing.
10. The optical storage is connected to the image system. It consists of disks that are readily accessible in the library (juke box), as well as the disks that have become inactive and have been moved to a shelf for manual mounting.
11. Communications from the image system to the image workstations consist of several local and/or remote Token-Ring Networks. This provides the required high-capacity connections while using standard components. There are no peer-to-peer communications on a particular ring, so there are no constraints on the network configuration. There is no capacity limitation if a particular ring becomes full; it can simply be split into multiple rings.
12. Indexing (for mixed type documents) is performed from paper rather than after an image is scanned, so is done on a standard data terminal, as shown on the left in Figure 1.
13. Images are scanned or printed at any workstation that is configured with the necessary scanner or printer. The request to print and the indexing information for the scanned document come from any data terminal, so the scanner/printer workstation does not have to have a dedicated data terminal or data connection.
14. The image workstation can be viewed logically as two separate terminals, sometimes (but not always) implemented as separate devices. The *data terminal* is a standard IBM 3270-type device, either real or emulated. The *image terminal* receives several documents, each consisting of several pages. It is responsible for displaying them in the best way possible, whether in a Presentation Manager™ window or (preferably) as a full-size full-page image. The image terminal is also responsible for page turn, zoom, display of multiple pages (if desired), and any other manipulation of the several pages in use. It can be configured with a custom keypad, as shown, providing specific image functions.

Usage. Three aspects of hands-on experience are discussed, the entry of documents into the system, the initial processing, and the means used for file references.

Document entry. Entering the document into the system proved to be much easier than expected. Existing procedures are followed as far as possible, such as opening the mail, extracting and arranging the contents, and identifying the customer. The actual indexing is done wherever the mail is first read to determine the routing, whether in the mail room or in the user area.

To perform the indexing, the customer is identified through a standard data terminal on the data system. Then an "Index" transaction is executed to collect data about the document. If a form number or code for the type of document can be entered, the indexing screen can be largely prefilled to simplify entry, review, and editing. When the indexing information is complete, it is saved in a temporary database, and the assigned key is written on the document.

After the document is indexed, it is moved to the scanner, which is a separate workstation connected

to the image system. The temporary ID written on the document is entered, which simultaneously turns on the scanner and begins the verification of the number entered.

Operating the scanner is similar to making an office copy. Each page is entered, and the quality of each image is checked as it is scanned. The scanner is adjusted and the pages rescanned as required to provide a good image.

When all the pages are entered, checked, and in order, the document is registered. Registration: first, moves the image from the scanner workstation (Figure 1, 13) to the image host (Figure 1, 5); second, causes the image to be stored by the storage manager (Figure 1, 7); third, advises the data system (Figure 1, 1) that the document has been registered; which, fourth, updates the index (Figure 1, 2) to show the document is in the system and available. If the document requires processing, the routing databases (Figure 1, 3) are simultaneously updated. The temporary ID and database entries are then deleted from the system. Within seconds of completion of the scanning, the document is available to any authorized user anywhere on the system across the country.

A priority for each document is computed by the system.

Initial processing. In a very small office, work is typically delivered to a single person who then allocates and distributes the work to the rest of the staff. All work is assigned to someone, but if that person falls behind (or is absent), the work is easily reassigned. With everyone working together, equitable distribution of the work is not much of a problem, and high priority and difficult work is rarely lost.

The problem is different in certain larger offices. Work is typically kept in a central pool and allocated in smaller batches, so that equitable distribution can be measured and so that high priority and difficult work can be tracked. This image system works on the latter basis, i.e., work is normally stored in unique departmental queues for each type of work. A supervisor can manually assign work to a specific person's queue, and work that has been started by a user and suspended remains in the user's personal queue to be completed.

A priority for each document is computed by the system, based on the type of work and age. A two-dimensional table provides the needed flexibility and simplicity. The typical entries are line of business (automobile, homeowner, boat insurance) and type of transaction (quote, new policy, change, complaint, estate, etc.). The values in the table include the initial priority, amount the priority will increase each day, and the queue in which the work should be placed for management.

When a system user is ready to begin work, a command is issued, in the form of a transaction on the data terminal, to "Send Work." The system selects the highest priority active task in the personal work queue or, if none, selects the highest priority work among the several queues that the supervisor has assigned.

Once the one "best" document is selected by the system, the user receives a picture of the document to be processed on the image display. The first page will be visible, with the other pages available by pressing the "Next Page" key. The data screen shows the processing history—an electronic "buck slip" with notes equivalent to "The letter arrived yesterday afternoon," or "Finished the auto insurance request, but need help with the boat insurance request." The date, time, and employee identification are also recorded each time the document is handled, together with any notes that the employee may add (such as why the work was not completed).

The existing data processing application programs were designed to process the information on the document (i.e., letter or form) efficiently. Nothing has changed but the medium. The document is now presented, just as before, but on a display rather than a piece of paper. With the instructions and document at hand, processing can take place using the existing applications programs. The existing programs need not be changed in any way, although they can be modified, if desired, to interface directly to the image system.

When the user is finished working on the document, another command is issued to "Wind Up" process-

ing. Three choices are available: First, if all steps have been completed, the document can be filed (it actually has been "in the file" available to all users since it was scanned, but there is no more routing). Second, it can be routed elsewhere for further processing. Third, work can be suspended until a later date and time (for example, if the customer needs to be called).

File reference (archive processing). While handling current mail or telephone calls, it is sometimes necessary to view a document from the files. Entering the command "File" will show the index of documents in any file, for the current customer (or other customer, if specified), for the most recent 18 months (or any specified year), including all types of documents (or just the specified file "tab"—line of business). This is based on the index database (Figure 1, 2).

While looking at the index on the data display, the user may select one or several documents to be viewed, printed, or routed for processing. The folder application program on the data system (Figure 1, 1) sends the required commands to the image computer system (Figure 1, 5). The storage manager (Figure 1, 7) finds the image for the distribution manager (Figure 1, 6), which sends it over the token ring (Figure 1, 11) to the image workstation (Figure 1, 14).

Computer output. A normal customer file is not limited to only the received documents, but also includes replies and copies of documents sent to the customer, exactly as sent. This permanent record might also include documents received electronically, such as credit reports and motor vehicle reports.

One method to store the documents that were in the computer system would be to print them and scan them. Not only would this consume much paper and labor, but there is a more efficient way to store the image.

The information can be formatted, as if it were to be printed, and then captured. By doing so, a typical page takes a few hundred or thousand bytes, rather than the 50 000 bytes required for a typical image. The original data processing programs are no longer required, so the documents will be usable "forever." However, it is not practical to revise or process the information, just to read it on a display or print it on paper.

Forms on which the presentation text would have been printed can be stored in a separate library, so the image of the form does not have to be stored hundreds or thousands of times per day. Each page of the presentation text can point to the form on which it will be displayed. These forms include:

- Letterheads with just a logo and return address, on which a large amount of variable information is printed
- Traditional forms with many "boxes" for data, on which a modest amount of information is printed
- Contract pages (the "fine print"), where the variable information might include only a page number and/or customer number

Supervisor functions. In a paper environment, the backlog of work can be counted periodically, and the productivity of an employee can be estimated by how often the in-box must be refilled. Between formal counts, supervisors can estimate the progress of their workers by looking at the stacks of paper waiting to be processed. With an image system, all the visual clues used by experienced supervisors are gone. The stacks of paper must be replaced by comparable information provided by the image system.

A holdover display provides a summary of the work waiting for each department or other work group. A queue display lists the characteristics of each document waiting to be processed for any customer, user, work group, or type of work. From the queue display, supervisors can examine the document (What did the letter say?) or the processing history (Why isn't it processed yet?), without leaving the desk.

Supervisors also use the system to assign work to their employees. They can assign all the different types of work for which the employee is qualified, or limit the employee to specific types. To assist another department, an employee can even be assigned certain work normally handled by that department.

Performance. The users required that the basic operations take about the same amount of time as they would on paper. A "page turn" to the next, previous, first, last, or a specific page in the current document or set of documents occurs in about one second. Page turns of even a few seconds interfere with productivity when a large set of documents is being reviewed.

Active documents, i.e., those that would be in the "in basket" for processing, are displayed in less than

five seconds. This requires that they be on magnetic storage, including a prefetch if they have been migrated to optical disk.

Random retrievals of documents in the file may be a little slower. Most retrievals occur in the first year, while the paper would be kept in the office area or while the optical disks are in the robotic library or juke box. The users agreed to a goal of 15 to 20 seconds, which is fast enough to serve customers while they are still on the phone.

Requests for older documents are rare enough that storage space is more important than speed. Users were willing to keep older optical disks on a shelf rather than in a juke box, and wait minutes for retrieval (even though it will mean calling the customer back) while an operator manually mounts the disk. When the optical disk is mounted, the requested document is moved to magnetic storage for a few days so that the disk does not need to remain mounted, yet the document will be available quickly during the call-back. Users recall that retrieval of older paper or microfilm documents took many hours, or even days.

Discoveries and lessons learned

Use of a separate machine. At USAA, the images require several terabytes (million megabytes) for just the first image application, which is more storage than all of the other data processing systems combined. The management of this tremendous amount of information will present some interesting challenges.

The ImagePlus architecture allows use of a separate computer to store and manage the images, or allows storage on the primary data processing computer system. If a low-volume pilot system is being built, or if the primary machine supports a widely varying workload, that existing computer system can also support image. In a large installation, where the primary computer system is finely tuned to a particular on-line workload, it may be more effective to use a separate image machine that can be configured and tuned to this very unusual workload. USAA chose to keep the images on a separate system.

Need to index before scanning. One of the key discoveries in the several years of image research at USAA concerned the sequence of indexing and scanning for mixed types of documents. Like most beginners, USAA intuitively concluded that the documents should be scanned before indexing. This would minimize paper handling and allow a visual quality check as the image was used to index the document.

As the procedures were developed, there were many cases where the paper was required after identifying

The indexing and scanning operations were separated, since different skills are required.

the document and the customer. (When scanning is done first, the purpose is to handle the paper quickly without pausing to identify what it is.) Representative cases include the following:

- *Business decision required before acceptance of the document.* For example, when an account is in arrears, partial payment may not be acceptable. When a check arrives with a letter, it is separated and deposited with the thousands of other checks received that day. Hours later, when the customer is identified and the document is indexed, the warning not to accept the check is finally noted, but it is too late.
- *Need to return document.* The original document, such as appraisals and certificates, may need to be returned to the sender. The document is not examined during scanning; by the time the image is indexed, the original paper document is buried among thousands of other documents and is difficult to recover. Furthermore, it is not apparent from the image whether a copy of a certificate was submitted (which does not need to be returned) or the original.
- *Use of turn-around form.* Some blank forms or return envelopes are received. If not examined first, they might be scanned into the system and destroyed. The appropriate action may be to process the document first, then enter the completed form into the system, then return it as originally requested.The solution to all these and many similar problems was to index the document before scanning. Although counter-intuitive, the paper only traveled slightly farther in the process. By doing the indexing from paper, the analysis of the

document and identification of the customer is completed while the paper is still available for any special handling.

Standardizing index descriptions. It was feared that it would be burdensome to capture a description of each document. Mail room personnel found that they reused a few descriptions frequently, so standard descriptions were entered in tables. (Thirty-five descriptions covered 90 percent of the cases; 50 descriptions covered 95 percent.) System tables now allow hundreds of standard descriptions for practically all forms.

Key to the success seems to be the abstract description of a document without any business judgment. For example, a good description might be "Correspondence from the insured about auto insurance," not "Request to increase collision coverage on '87 Oldsmobile to ...".

Separating indexing and scanning operations. The indexing and scanning operations were separated, since different skills are required. Indexing is performed by a person highly trained in office skills; scanning is largely a mechanical process performed by a clerical employee. Because there are many more indexing clerks than scanner operators, separating the functions also greatly reduces the investment in scanning equipment.

When the indexing and scanning operations are separated, there must be some way of linking the image with the descriptive and routing information. As the document is indexed, the system generates and returns the key (temporary ID) to the database that temporarily holds the indexing and routing information. That key is written on the document and is entered when the document is scanned. By using the customer number, normally found on the documents, as part of the temporary ID, the writing is reduced and error recovery simplified.

Adjustment of scanner speed. Scanning is similar to using an office copier, a comparison that provided many helpful ideas during the requirements study. Correspondence and other office documents received by mail come in large quantities with mixed sizes, weights, and folds. No one would consider using a high-speed stack-fed copier to reproduce such a volume of mail, yet the scanner vendors felt that would be appropriate. Likewise, copiers that require each page to be aligned under a lid for a 10-second or longer scan would not be practical for 25 000 pages per day.

After experimenting with various copiers, the scanner requirements were set at 1 to 3 seconds per page, with an automatic feed, and the original returned near the input station to allow rescanning (or reverse side scanning) if required. An occasional document, such as an express mail envelope, needs a flat-bed scanner.

Ability to view a list rather than a folder. Some image systems move the entire contents of a folder to the user, simulating the delivery of a paper folder, so that the file can be browsed. This approach requires that a file be kept on a contiguous area of the optical disk to make retrieval time tolerable. As new documents are added, the optical disks have to be reorganized to consolidate documents within a file and make space for additional documents. With the relatively slow performance of write-once optical storage, this is expensive.

Use of image systems provides an opportunity to manage the work more closely.

The USAA analysis showed that normally only a few documents actually have to be viewed. With a good list of what is in the file, it is very rare that the entire file has to be read. This makes it practical to store the documents in the sequence received, optimizing use of the optical disks.

In 24 percent of the cases in the first application at USAA, the users satisfied their inquiry without reference to any documents. Simply knowing whether a document was (or was not) in the file, based on the index, was sufficient.

Routing single items. Use of image systems provides an opportunity to manage the work more closely. Rather than distributing a package of work (such as 25 documents) to a user, most work is kept in a common queue. There is less chance that a high priority piece of work will be buried behind other

work in someone's work packet. As described above, when a user requests a piece of mail via the "Send Work" command, the system examines the employee qualifications, the supervisor assignments, and the priorities of all the waiting work, and selects the one best document for this user to process at this instant. That one item, with both the image and the processing history, is sent to the user to be processed.

Integration within application programs. Image systems can be an integral part of a business application. For example, when renewing the record of an insurance policy, the insurance processing programs could use a function key to automatically display the image of the original application for that insurance policy. In the extreme case, users have even asked that the relevant postmark be displayed when inputting the effective date of a policy.

Although the integration is certainly possible, this ties the implementation of a new technology to the redesign of existing computer systems. The alternative is to deliver a document for use with existing computer systems. The document appears vertically on a display screen, rather than lying horizontally as a piece of paper on the desk, but the same programs work as well (or as poorly) as they ever did.

The decision for the first system at USAA was to use the existing application programs because they had recently been rewritten at a cost of over $100 million over a five-year period. At the same time, the image utility feels integrated to the user, since the same data terminal is used, and common information, such as the customer number, is automatically picked up by the image system.

Storage of computer-generated data. The image system was justified by the high cost and slow processing of paper mail. Once the system was available, there was a desire to include the outgoing correspondence with the rest of the file, and perhaps the insurance policies and other documents. It was a great surprise to find how much computer-generated data were produced "to be filed." The initial system supported about 10 000 letters received each day. The computer output, however, included about 140 000 documents.

Since computer-generated data can be stored very efficiently in the image system as presentation text, a new opportunity emerged. For example, only the declaration page of a policy, summarizing the insurance coverages, was stored on microfilm. The extra cost of storing the additional standard contract pages on the image system was extremely small, since each page consisted primarily of a pointer to that page in the form library. The business benefits of being able to refer to exactly the same document as the customer is substantial. Therefore, in addition to a large number of documents, there is also a larger number of pages.

Solving display resolution requirements. The amount of information in a picture is very large. Information is required for the placement of each ink dot for character representation, including the specific style and position. This occurs because the image system must represent the original document exactly as it appeared, rather than optimizing the character style for the image system hardware.

Practically all business documents can be read if a minimum of 100 pixels are displayed for each inch of the original document. If the pixel is less than the ideal square that just touches its neighbor, the 100 pixels per inch (4 pixels per millimeter) is not adequate. Likewise, anything less than 100 pixels per inch proved inadequate.

Since reading "practically any" document isn't good enough, the decision was made to scan and store at 200 pixels per inch of the original document (8 pixels per millimeter). This insured the capture of the information in the smallest print ordinarily used in business. These conclusions are confirmed by the international facsimile standards for business documents at 98 or 196 pixels per inch vertically by 203 pixels per inch horizontally (7.7 by 8 pixels per millimeter).[1,2]

The economics of displays led to the selection of image displays that could support 100 pixels per inch of the original document. At this level, at least 1150 lines are required on the screen to display a full 8 1/2- by 11-inch (or A-4 size) page with titles and status lines as usual. This is two to four times as much detail as on normal displays, and 10 percent more than most "high resolution" displays. Therefore either special displays are required, or only a partial page can be displayed.

It is occasionally necessary to print a page (at a higher resolution) or "zoom in" to display a larger image. With the 200 pixels per inch available, the zoom shows more detail and not just a larger version of the same image.

Simplifying system functions. When interviewed, or when seeing technology demonstrations, users were

initially anxious to have many image-handling features. To our surprise, after they tried the features, the users did not like the windows, mice, or other "high tech" functions. For example, after working with a mouse long enough to become proficient, they asked that it be removed. There was a strong preference for a simple keypad with dedicated keys for page turn, zoom, rearranging pages, and other simple functions.

Need for full-page display. In most paper factories much of the processing consists of entering information from the paper (letter or application) into the computer system, or comparing that information with data previously entered. In order to use existing programs to process the data, a full screen of data and a standard keyboard must be available while simultaneously viewing the image. To minimize the "paper handling" effort, users asked for a full page of paper to be visible at once. They did not want to take time to move around a page of a letter looking for data needed for each field.

USAA operates a very centralized data processing operation.

Two test systems were installed: one with a separate display for image, the size of a sheet of paper, located adjacent to a conventional data terminal, and the other with a single large display showing data on one side and image on the other. Management and analysts vigorously preferred the single display. It seemed more logical. The users did not care whether they had a single large display or specialized but separate image and data displays. They were far more concerned with simplicity, speed, and ease of use.

Timely access to filed documents. Access to a paper file typically took long enough that, if file information were necessary during a phone call, users preferred to call the customer back. A significant productivity gain could be achieved if the customer could be served during the initial phone call. Recognizing that immediate access was not practical, users asked for 10-second access, and agreed to tolerate up to 15 or 20 seconds. They are concerned about "what to talk about" while waiting for the document.

More study may show that this requirement can be somewhat relaxed. In the occasional periods of poor performance as the system was being developed, users almost never recorded problems until response time exceeded 30 to 45 seconds, indicating, perhaps, a viable service level.

Need for service to remote users. USAA operates a very centralized data processing operation. About two-thirds of the employees are located in the home office, and all data processing is provided from the centralized location. The field offices are responsible for terminal installation and use, but do not provide a 24-hour computer operation.

USAA wants to maintain a similar operation for image. It would be very costly to build and operate a high-security computer room in each field office. To provide remote service comparable to the home office, very high-speed communication lines can be used. A 256 Kbps data channel adds only a second or two to the access time for the first page of a document. Since the first page is displayed as soon as it arrives, that time is the relevant consideration; the other pages need to arrive only after the first page has been read.

In the type of use at USAA (light use throughout the day with most mail processing concentrated in a few hours while the volume of phone calls is low), a single 256 Kbps data channel supports almost 200 image terminals. At a smaller office, the same communications speed is required to maintain response time, even though the utilization of the lines may be low.

In addition to large offices, the image system must support the "one-person office," such as a remote claims staff or "walk-in" service area. When images are needed, they will be requested through the existing data terminal (which might be a laptop computer with a cellular phone). The image would be delivered through a facsimile machine. The same facsimile would also be used to enter images. The image technology in ImagePlus is virtually identical to the facsimile technology.

The midsize office (such as an insurance agency or claims operation) is one of the remaining challenges.

The offices are too small to justify high-speed communications lines to support immediate access from central storage. There are too many users in the office for individual facsimile machines and phone lines. Distributed image storage is less secure than a regular computer facility unless special precautions are taken.

Several solutions are being explored to provide temporary image storage locally with primary storage on the host, including networks of personal computers with shared communications, a new configuration of AS/400 system, placing a special "object distribution manager" on a smaller remote machine, and even networks of the RISC System/6000™ machines. The question is only which of these several possible approaches, or others, will be supported.

Ability to decentralize image storage. At some time USAA may want to decentralize its computer operations. A requirement for the ImagePlus system was that the image storage could be decentralized, which has been tested.

Images can be stored in the cities where the document is most likely to be used (each field office), or in the same cities where the data are stored, whether centrally or distributed. The plan is that old images will normally stay where they are initially stored, even if the rest of the data for the customer (and any new images) move to a different system. The system records where the documents are stored and retrieves them from that original location if they are later needed.

Conversion of existing files. It was assumed that the existing files of documents would have to be converted. How else would the benefits of immediate access accrue? However, two factors argue against conversion:

1. In most applications, documents undergo the highest activity at the beginning of their life; converting old documents has less benefit than making the new documents available on the system.
2. The cost of conversion is very high, primarily since each document must be read and handled again. This is not a major cost for new documents being entered into the image system, since the indexing can be done during the regular processing of the documents.

At USAA the requirement for a wholesale conversion was dropped when it was recognized that the conversion would take over 10 years, longer than the seven-year life of the documents. In practice, archived paper documents being retrieved for the field offices are being entered in the system as requested (called "on-demand conversion"), rather than being copied and mailed. There will probably be other applications in the future where conversion is required.

Concluding remarks

Getting started. The temptation is always to start small and simply: find a group of five people with a filing problem and install the new technology with minimal risk. The difficulty with this approach is that such a small system may not be cost effective compared to the use of microfilm or paper. For example, it may be cheaper to expand the file capacity for that test group by installing additional file cabinets near all the users. When the image system is not cost effective (compared to a few file cabinets), companies have wrongly stopped considering image systems.

In contrast to the small systems, there are unique problems in maintaining a large set of files. Two thousand USAA users cannot all "sit next to" two million customer files. If a separate professional staff is required to maintain the files, the problem is probably at a level that image may be easy to justify. A large-scale image system can solve file access problems that are unique to such a large-scale operation, and can be extraordinarily cost effective. Furthermore, a large installation can have sufficient benefits to cover the overhead of introducing a new technology.

The decision to start with a large application does not require a sudden installation. Image systems can be implemented for a small group, and procedures can be tested and users trained before proceeding. The key is that no one should expect to see major benefits with just a few users, before the installation is complete.

Outlook for image systems. Small image systems have been available "off the shelf" for years, but are difficult to justify based on cost, compared to paper or microfilm.

Large custom systems have been available for years. There are legitimate concerns about flexibility, maintainability, and expandability, but they demonstrate that there are no hidden technical limitations to prevent building a large system.

With the USAA prototype of the IBM MVS/ESA ImagePlus system, large-scale "off the shelf" image systems are generally available.

The next step at USAA. At USAA today over 1400 image terminals are in daily use for over 2000 users in the Property and Casualty Policy Service and Underwriting area. All of the incoming mail (an average of over 10 000 letters or 25 000 pages per day) has been stored in the system for about a year. Most of the outgoing mail is or will soon be stored, which is an additional 140 000 documents totaling nearly a million pages per day. Correspondence and policies are presently stored on the system.

The second project is well under way. It will include an additional 3000 to 4500 terminals on the enterprise image system to support claims processing. In addition to office documents, it will include facilities to store approximately one million color photos per year and recordings of a half-million telephone conversations.

Work has begun on a third project to store applications and correspondence for our bank credit card center. An estimated 120 image terminals will be required.

Countless other application areas are waiting to be provided with image services, such as life and health insurance, real estate investments, transfer agency, tax records, personnel records, accounts payable, and even the general office files.

Image systems have been in operation at USAA for several years. Even though the technology will continue to advance rapidly, we believe that the time to move to image systems is now! The operators like using the system more than working with paper. Always having the needed documents when a customer calls will improve customer service. The systems are cost justified. Therefore, the outlook is great—great for the users of the system, great for the customers, and great for the "bottom line" of the business.

Acknowledgments

The authors wish to acknowledge several people whose support made this image project a resounding success. In 1969 General Robert McDermott, Chairman and CEO of USAA, challenged his organization to become paperless and supported image systems through its many years of gestation. John Akers, Chairman of the Board of IBM, personally recognized the business opportunity and challenged his organization to find a solution. Mark Morin, Vice President of the IBM Application Solutions Division, had the courage and persistence needed to forge the USAA/IBM partnership. Without the support and guidance of all these people, and countless more, large-scale image systems would still be a dream rather than a reality.

Cited references

1. R. Hunter and A. H. Robinson, "International Digital Facsimile Coding Standards," *Proceedings of the IEEE* **68,** No. 7, 854–867 (July 1980).
2. "Standardization of Group 3 Facsimile Apparatus for Document Transmission," *Red Book,* International Telegraph and Telephone Consultative Committee (CCITT), Fascicle VII. 3, Rec. T. 4, 1984.

Charles A. Plesums *USAA Information Services, USAA Building, San Antonio, Texas 78288.* Mr. Plesums is senior director of image technology at USAA. He managed the installation of the IBM MVS ImagePlus prototype system described in this paper. From 1982 through 1987 he led the USAA research effort in optical storage for image and data (including the installation of two test image systems and an optical data storage system), and the studies in intelligent workstations and expert systems. Prior to joining USAA, Mr. Plesums was a principal consultant with a national firm and director of two computer centers. He has taught computer science in the United States and as a World Trade Corporation visiting professor in Brazil. He has a B.E.E. degree and an M.S. degree in computer science from Union College in Schenectady, New York.

Ralph W. Bartels *IBM Corporation, 300 Convent, San Antonio, Texas 78205.* Mr. Bartels is a consulting industry specialist for image systems with IBM. He has worked with USAA for the last five years on image systems requirements and was instrumental in developing the IBM solution for large-scale image systems. Mr. Bartels has 29 years of systems engineering experience with IBM in several different industries. He has a B.S. degree in engineering from Iowa State University.

Reprint Order No. G321-5404.