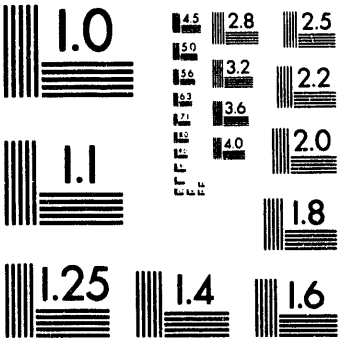
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
7.1
8.0
9.0

1 of 1

Conf-930749--16

NETWORKED MS WINDOWS 3.1 BASED CLASSIFIED DOCUMENT CONTROL SYSTEM (CDOCS)

Lawrence M. Desonier
Sandia National Laboratories
Insider Technology Department
Albuquerque, New Mexico 87185

Abstract

Current classified document management systems require a tremendous amount of space and extensive manpower to account for, inventory, and protect the documents. Comprehensive analysis of current control and accountability procedures reveal the main problem is the actual handling of the paper itself. The purpose of the Networked Microsoft Windows 3.1 based Classified Document Control System (CDOCS) is to eliminate the paper by scanning and storing images of pages on a personal computer using "write once read mostly" (WORM) high density optical media. By saving images on the computer, not only can manpower and space requirements be reduced, but the chance of compromise is diminished. As an added benefit, the information is now more readily available to the authorized user and is provided to the user at the user's PC. The network target for CDOCS is Microsoft Windows for Workgroups. Thus, the system is also readily applicable to unclassified document imaging uses.

Introduction - The Problem

The management of paper in both industry and government requires an enormous amount of personnel, time, and money. Besides the normally large amounts of unclassified documents, our government agencies must account for, inventory, archive, and destroy millions of classified documents. Private and proprietary documents also increase the magnitude of this challenge.

The Department of Energy has sponsored a program for the investigation and development of better ways to protect classified information. One of the main problems with managing and protecting the paper is the existence of the paper itself. To lessen this problem, Sandia developed a Classified Document Control System (CDOCS) to better manage and account for classified and critical information. Images of the pages of documents are stored within the computer itself. (Figure 1) In addition to handling classified, the system can also be used as a regular document control system.

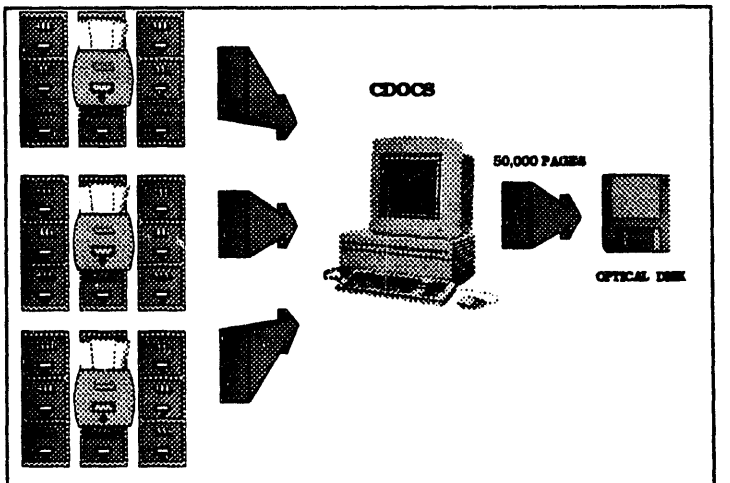

Figure 1: The CDOCS Concept

The CDOCS is based on a PC/AT type system. The basic design requirements were to provide for better protection of classified documents and to make data readily available through fast data retrieval. In addition, the system had to store many thousands of document images in compact form and be easy to use. The CDOCS system consists of a high speed personal computer, a high resolution display, a dual optical drive tower, a high speed scanner, and a laser printer. (Figure 2)

This work was supported by the U.S. Department of Energy under Contract DE-AC04-76DP00789,

CDOCS Hardware

Scanning

The first requirement, for such a system, is to get the information into the system. The CDOCS system uses a high-speed scanning process to accomplish this task. This scanning process employs a scanner capable of

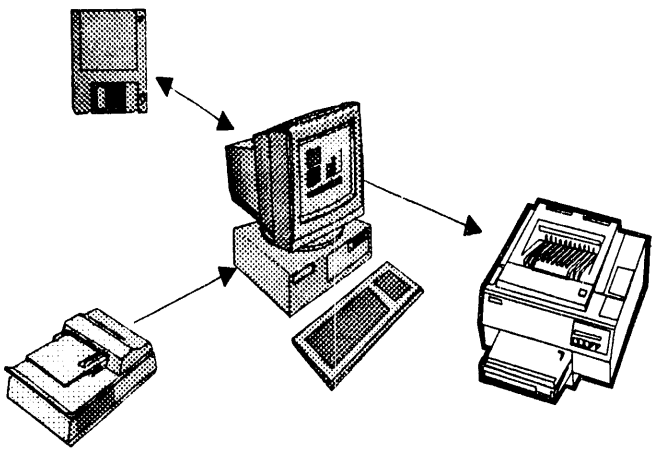

Figure 2: CDOCS Station

handling roughly 40 to 60 pages per minute. The scanner range of resolutions is from 200 dots per inch (dpi) to 400 dpi. This allows the user to determine the density and sharpness of the image. At 200 dpi we can optimize the storage of the image, while at 400 dpi we optimize the image resolution but the required storage is greater. We decided to scan at a resolution of 200 dpi since we found no visibly discernible difference between 200 dpi and 400 dpi on either the display or printed output.

Storage

The image storage medium needed as large a capacity as possible for a sensible price. In addition, the storage medium chosen had to assure that no one could alter the document. To meet both of these requirements we selected an optical Write Once Read Many (WORM) drive made by Literal Corporation. The 5¼" WORM platter has a capacity of 1.2 gigabytes, which is 1200 million bytes or characters of information. This equates to approximately 50,000 pages of storage at 200dpi scanned resolution. Not only is the optical drive capable of immense storage but the access speed is at least equivalent to or even greater than that of a typical PC magnetic hard disk.

Display

There were several options used during the development and implementation of CDOCS for display of the actual scanned images, the menus, and the document record data. In the first option we used a two display system: a VGA display, for the menus and data, and a high resolution dual page monochrome 19" display for the actual scanned image. The dual page display was chosen over a portrait display to allow the option of rotating images into landscape mode (i.e., for viewgraphs). The second option, which we are now employing, was to use a single high resolution, high speed color display. This allows the user to see both the data and the image in separate windows on the screen at the same time. Feedback has shown that this is the preferred method, to have the data on the display at the same time as the image.

Printing

The system must guarantee replacement of the paper since usually the user is going to give up ready access to the original piece of paper. A laser printer provides all hard copy output and special circuitry scales the document image and prints the image at 300 dpi resolution. Experience has shown that the printed output is visually indiscernible from the original.

CDOCS Software

Security

The CDOCS system employs many levels of checks to offer enhanced security. The different security levels include a password system for login, a comparison of classification levels, access level checking for data access, a complete audit trail of all actions on classified documents and images, and a Trusted Interface for operation on the network.

The first level of security is a login/password requirement. The password is stored inside the data base encrypted in Data Encryption Standard (DES) format. The key used to encrypt the password is the password itself. This implies that to decrypt the password would require the original password.

The next level employed is the use of multiple levels of checking before granting access to the information. The first level checked is the security level. The user must have a security level equal to or greater than the security clearance level of the information. For example, the system manager must have entered a user classification of Confidential or greater to allow the access to Confidential information. The second checking level is to verify there are no access restrictions. A document may have a single access type or multiple access types (e.g., SUCI or CNWDI). The user must have the same access type or types to retrieve information. The third level to be added is need-to-know checking. A document may have certain need-to-know or program codes. A user must

match in at least one of the need-to-knows to gain access to the document.

Another area of security is the audit trail. The system maintains a complete audit trail of all classified paper documents checked into or out of the system, all classified images viewed, view requests denied and the reason for denial, and all changes to record information on classified documents. In addition, once a user has scanned a document into the system the classification level of that document may not change. This is because of the requirement that the system compartmentalize the different classifications on different optical disks.

The most recent security software added to CDOCS is the network Trusted Interface. The idea of the Trusted Interface is to prevent the normal network user from accessing any CDOCS files, except through the CDOCS program itself. This consists of preventing the user from running multiple programs on the workstation while the CDOCS program is running. Any attempt to bypass this terminates the CDOCS program and locks the user from the server information.

The Windows Interface

The user interface consists of a data window with pull-down menus and appropriate data entry forms (See Figure 3). The Graphical User Interface (GUI) selected for this was Microsoft Windows 3.1. A special key on the mouse allows users to access the menu. Using the TAB key or REVTAB key allows the users to traverse the forms used in the system for data entry and entry of search criteria. The user types the data into appropriate fields and accepts the data by hitting the ENTER key.

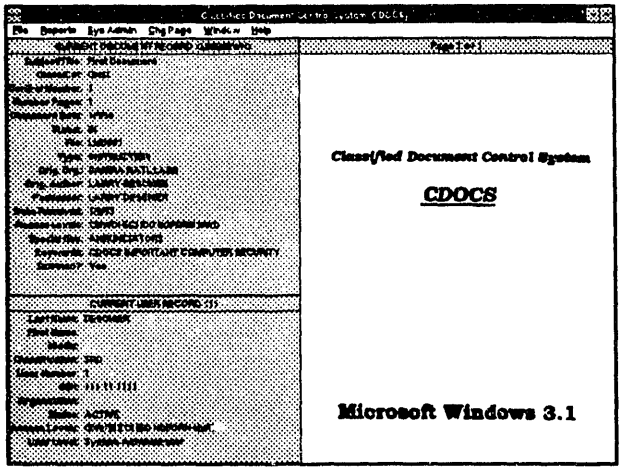

Figure 3: The CDOCS Screen

Retrieval

The scanned image may be retrieved in several ways. The user has an option of storing a programmable number of key words on the form. The user may search for the document on any single keyword or multiple keywords. In addition, the operator may assign a document control number, a file location, a type, document date, and(or) a status to any document. The operator may then use any of the above criteria to search for a document image. For example, one search combination could be all ACTIVE images that have the type assignment of MEMO. The user would see a list, on screen, of all documents that meet the search criteria. The user then could choose from that list and the system would display the document record and image.

Reporting Capabilities

The CDOCS system provides extensive auditing or logging of events to management personnel. The manager may display the desired reports of transactions on the screen, to a text file, or provide a printout.

Paper Management

Even with most of the paper scanned and stored in the CDOCS system there will still be some paper left to manage. Human nature is to want that piece of paper to carry around or use as a reference when writing another document. The CDOCS system also provides for the management of these paper documents, through data records similar to those kept for scanned images.

In addition to the management of the paper left over, CDOCS allows for changes in the images when there are changes in the pages of a document. The system allows for deletion of a page, replacement of a page, and the insertion of a page in the document. There is also a feature that allows for the appending of pages to the end of a document. In addition, the system allows for scanning of double sided documents by turning over the paper in the scanner when one side is finished.

The Future

We have a considerable amount of work to do to get CDOCS to a more viable product. Two of the main features needed at this time are Optical Character Recognition (OCR) and Full Text Retrieval. The OCR and full text indexing would take place as a background process and there would be no intervention required by the user. When the user would desire to search for something, he/she would merely ask for any text, phrase, etc. and the system would show relative documents. Another important feature needed is Markups of the image. This would entail the concept of "yellow stickups" on an image, the insertion of margin text, boxes around items, lines, and electronic white out or black out. These markups could also be controlled by access level, thus giving a

much finer grade of information control.

Users have also expressed a need for a more secure operating system. We have chosen Microsoft Windows NT for this system since it is approved for C2 level of security by NSA. Another feature needed for the system is work flow control. This would allow documents to be forwarded to individual users on the network for comments or review. One final item of development would be a hierarchical network scheme for storing all documents on a jukebox system. Only working documents would be stored on the local net, but all information would be available for download from the archival network.

System Advantages

The most obvious advantage of CDOCS is the reduction of paper, thus decreasing the storage requirements and the paper management costs. Once the system has scanned this paper, and the paper is either destroyed or archived, there is no longer a reason to perform any type of inventory or accountability on anything other than the optical disk itself. The reductions in the cost of managing the old paper could be tremendous.

The second advantage is the added security. Since the system separates the data by classification, the most critical data could be more secure. In addition, with the login requirements, the operating system limitations, and the various levels of checking, access to the actual information is more difficult. Finally, the audit trails should deter any insider from trying to access unauthorized or unrequired information.

One disadvantage may be that if a disk is lost, stolen, or destroyed a large amount of data would be lost. Therefore, we have developed the capability to perform periodic backups for the maintenance of duplicate data.

DATE
FILMED
9/29/93

END