COLD, COOL, COM, Greenbar, and Your Bank Statement

Introduction

Many organizations meet their record retention requirements for computer records by printing the computer records out and storing the printouts for seven or more years. This may be inefficient, and require forests of trees, but it is easily understood and cannot be affected by errors in DP. DP (Data Processing) aka (also known as) traditionally EDP and ADP (Electronic and Automatic DP), more recently as IT (Information Technology), IS (Information Systems), and MIS (Management Information Systems) and most recently as KM (Knowledge Management).

Mechanics

The stored printouts are often on fanfolded (continuous) 14 inch wide paper that is folded at 11 inch intervals creating 11 by 14 inch pages. The text on these pages is in 66 rows of 132 characters each, for a total of 8,712 character per page. Because the number of characters per line is fixed, the fonts used are called monospaced fonts.

In PCs the monospaced format can be recreated by using the Courier font, a monospaced font. Many reports are emailed in a monospaced format, to maintain compatibility, but become unreadable when reproduced in a proportional spaced font such as the Arial font used in Microsoft Internet Explorer (IE). Converting the text of an emailed report to the Courier font can often make the email readable.

Most text is printed in IBM's EBCDIC (Extended Binary Coded Decimal Interchange Code) which can easily be converted into the ASCII (ANSI (American National Standards Institute) Standard Code for Information Interchange)) characters. The 14 inch long lines of text on the computer printout sheets are hard to read, so lightly tinted green bars, three rows of text high, are preprinted on the pages. The green bars extended the width of the pages, and are separated by white spaces thee rows of text high. These green bars make it much easier to follow a row of text across the page. The bars have become widely used, and have given all 11 by 14 inch paper the name 'greenbar' paper. When laid flat, greenbar paper also fits nicely into records storage cartons which are designed to store 11 inch long letter size paper in one dimension and 14 inch long legal paper in the other dimension.

Efficiency

The first step toward efficiency was to microfilm the printouts. Then, tractor feeders were added to the microfilm cameras for automatic operation. To cut out the paper, the computer output was written directly to film using CRT (Cathode Ray Tube) or EBR (Electron Beam Recorders) recording. This process was called Computer Output to Microfilm (COM). Because 35 mm (millimeter) microfilm is the same format as 35 mm movie film, COM eventually lead to the digital dinosaurs in *Jurassic Park* and colorized movies like *Gone With the Wind*. An early COM equipment vendor, Information International, Inc. (Triple I), in Culver City, CA, (where the first *King Kong* movie was filmed) made parts of the first computer generated movie, *Tron*.

Soon firms were scanning paper images and placing the images on optical discs. Microfilm was also scanned and placed on optical disc. Again, a step was cut out by writing the documents directly to optical disc, creating COLD (Computer Output to Laser Disc). Some of these files were written as images, but most were written as characters in ASCII or EBCDIC characters, greatly reducing the space required for storage.

As magnetic storage became less expensive, the information on the optical discs was transferred to magnetic storage and put online creating Computer Output On-Line (COOL).

Frequently, a full text search of the reports in COOL is easier than electronically flipping through the pages as though they were on printed paper. This obvious success lead to the creation of databases to make COOL easier to search. By highlighting portions of the COOL images on a CRT (Cathode Ray Tube display), a relational database can be built up on the fly using WYSIWYG (What You See Is What You Get) point and click mouse operations. Because this reverses the process used to write the reports from the MIS database, the process can be called reverse engineering of the database.

The reverse engineered database can then be used to generate advertising mailings and many other functions without having to interact with MIS or wait for their often one or two year response cycle. This access to data has made COOL very popular in many organizations.

Print File Spooling

To avoid the need to schedule MIS time for producing output for the implementation of a COOL system, most COOL systems use the same files used to print existing greenbar reports or statements. Computers do not print directly to printers because there is a speed mismatch between the computers and the printers. The computers create a spool file, which is sent to a magnetic disk. It is this file that is queued up and printed when the printer is available. COOL systems use this 'spool' or 'print files' as input. Users can then generate the input for their COOL systems and subsequent reverse engineering of databases by merely specifying report formats and contents. Most mainframe systems are built with extensive report format and content specification tools.

A Twist

With PCs and laser printers, the 11 by 14 paper has been reduced by 8 ½ by 11 letter size paper. The higher resolution of the laser printers and the higher quality fonts has actually made the reports more readable in some cases. However, the proportional fonts and multiple font sizes can make the simple database reverse engineering much more difficult than with the monospaced, fixed size mainframe report fonts.

Beyond special fonts, some COOL documents use a bit-mapped format. In bit-mapping, every pixel or dot on a page is addressable creating APA (All Points Addressable) formats. These formats are something like PC laser printers or PC screens with their elaborate graphics. For APA printing, IBM (International Business Machines) uses AFP (Advanced Function Printing) and Xerox has the Metacode format. Being more complex, page images created in these formats require more COOL storage.

How Big Is a COOL Image?

(How Much Computer Storage Does it Require?)

Most COOL systems use compression. Some use a simple blank space removal compression algorithm. (Any program is an algorithm, but algorithm is usually used to describe a specific well-defined operation such as compression.) Other systems add more sophisticated compression that can double the compression achieved with simple space removal.

The estimates below can be used to do the initial sizing for the storage for a COOL system. After processing even one week of output, the exact size of an average COOL image can easily be determined. It is important to remember that there are often special cases. For example, end of month and end of year reports are frequently much different that daily or weekly reports.

A statement, such as a utility or a bank statement may only contain a few line items. Most of these COOL page images are letter size or smaller and are less than 2 KiloBytes (2 thousand bytes).

Statements make use of a forms overlay, overlaying an image of a standard form that includes such things as the logo of the billing organization. The overlay is placed over the displayed or printed statement information so that entries such as the dollar amount due can be easily distinguished from the dollar amount previously paid. This forms overlay must only be stored once for all statements printed or displayed, so the forms overlay size is very small relative to the size of the COOL files.

Greenbar 11 by 14 inch COOL report page images are usually less than 4 KiloBytes when compressed. The bit mapped COOL report page images are usually less than 10 KiloBytes.

Less than optimal conversion algorithms (from the print file to the COOL file format) can make the page image sizes much larger, but a change of conversion algorithms can almost achieve the estimated sizes listed above.

Indices and Database

Because the indices and database entries in COOL databases are much smaller than the page images, these estimates of page storage requirements include the storage required for the indices and database entries.

Interfiling Statements and File Copies of Checks

In many organizations, file copies of checks written to suppliers are manually filed by supplier name or number. This is extremely labor

intensive. Also, the NCR (no carbon required) duplicate paper file copies are more expensive than the checks alone and cannot be recycled because of the ink capsules embedded in the paper. The manually filed copies must be de-collated (separated) from the checks and burst (torn apart) and then manually interfiled.

A COOL application can easily interfile the documents using several techniques ranging from a computer search of all COOL data, merging of indices, or even merging of the COOL images. Images of the canceled checks can be received from the bank in electronic image format and the checks themselves could be eliminated using EFT (Electronic Funds Transfer).

From Boxes to COOL

A record storage box of fanfolded greenbar paper has about 2,500 sheets. With each sheet producing 4 KiloBytes of COOL page image data, each box represents 10 MegaBytes of computer storage. A CD can store 50 boxes of fanfolded greenbar; a DVD can store 500 boxes.

To use COOL, the print files must be used. If the greenbar is scanned, 50 KiloBytes are required per page if the pages are plain white paper, more than 12 times the COOL storage requirement. If the paper is preprinted with green bars, or some other pattern, and the pattern cannot be removed by thresholding the image so that the preprinted pattern is considered to be part of the white background, then the scanned pages can require up to 500 KiloBytes of storage each, or 120 time the COOL storage requirement. Scanning representative samples can show whether the scanned pages require 50 or 500 KiloBytes of storage.

When converting from boxes to COOL it is frequently better to do day-forward, stopping the storage of new boxes, but continuing the storage of previously stored boxes. As the stored boxes reach the end of their retention period, they can be discarded. For a 7 year retention period and for the same volume of reports stored each year, one-seventh of the boxes can be discarded each year until no storage is required for the converted report series at the end of 7 years.

An additional advantage of COOL is that the cost of storage is extremely small, and the records can be kept forever if there is even the slightest advantage in an unlimited retention period.

Other Electronic Documents

COOL documents can be electronically interfiled with all the other forms of electronic documents in an organization: the word processor and spreadsheets, scanned documents, and operating data such as fetal heart monitor data in hospitals.

Summary

COOL gives the records manager control of transaction and report data. The records manager can certify that a given page image was produced by an organization on a given day, with no reference to, or influence from, a database maintained by the MIS department. This control was previously achieved by filling a warehouse with paper. Now records can be managed on a small amount of computer disk storage with all the controls necessary for the successful operation of an organization.

Sidebar:

Content-Free and Greeking:

Greeking is used by artists when creating a display ad. The elements of the ad: pictures, borders and other graphic elements, and text much all contribute to the visual effect of the ad. Because the artist may be creating the image of the ad before the copy (text) for the ad has been completed, Greeking is used. Greeking is text that looks like text from a distances, but does not say anything. An example of Greeking is 'lum birnuma arumi om palium' which might later be replaced with 'makes teeth whiter', and still later with 'can make teeth whiter under some circumstances'.

Content-free can be used to describe some instruction manuals or brochures created under special circumstances. Content-free documents are created when the product might not do what could be expected from the description. Content-free documents are also used when an explanation is in order, but the explanation would not contribute to the marketability of a product. Occasionally the marketing department must describe a product before the engineering department has explained what the product does, or before the engineering department has determined what the product can do. This is accomplished with a content-free description.

Updates and More Detailed Descriptions

When using the information in this article, please check the website www.ArchiveBuilders.com for updates. The version number of this article is just before the page number below. The website also has articles that provide more details on some of the terms and concepts in this article.

Comments

Please let us know how you like this paper, or if you had any questions. What would you like to see in the future? For more, and the most recent version of this article, please visit our web site at www.ArchiveBuilders.com.

Please send your comments via email to SteveGilheany@ArchiveBuilders.com +1 310-937-7000. Also, please let us know where you saw this article.

Acknowledgements

Reprinted from Archive Planning, Volume 2, number 6, 1998, Archive Builders' analysis newsletter for document management.

See www.ArchiveBuilders.com

Note to Editors

Paper 22007v006

We will continue to update these articles as we get comments. Please contact us for the most current version before you publish. Also, please request permission to publish the article. Permission will be given freely for most purposes.

Steve Gilheany
Archive Builders
1209 Manhattan Ave., PMB C-14
Manhattan Beach, CA 90266
Tel: +1 310-937-7000 Fax: +1 310-937-7001
SteveGilheany@ArchiveBuilders.com

Bio

Steve Gilheany, BA in Computer Science, MBA, MLS Specialization in Information Science, CDIA (Certified Document Imaging System Architect), AIIM Maser, and AIIM Laureate, of Information Technologies, CRM (Certified Records Manager, ARMA) has seventeen years experience in document imaging and is a Sr. Systems Engineer at Archive Builders.

Author

Steve Gilheany is a Sr. Systems Engineer at Archive Builders. He has worked in digital document management and document imaging for seventeen years.

His experience in the application of document management and document imaging in industry includes: aerospace, banking, manufacturing, natural resources, petroleum refining, transportation, energy, federal, state, and local government, civil engineering, utilities, entertainment, commercial records centers, archives, non-profit development, education, and administrative, engineering, production, legal, and medical records management. At the same time, he has worked in product management for hypertext, for windows based user interface systems, for computer displays, for engineering drawing, letter size, microform, and color scanning, and for xerographic, photographic, newspaper, engineering drawing, and color printing.

In addition, he has nine years of experience in data center operations and database and computer communications systems design, programming, testing, and software configuration management. He has an MLS Specialization in Information Science and an MBA with a concentration in Computer and Information Systems from UCLA, a California Adult Education teaching credential, and a BA in Computer Science from the University of Wisconsin at Madison. His industry certifications include: the CDIA (Certified Document Imaging System Architect) and the AIIM Master, and AIIM Laureate, of Information Technologies (from AIIM International, the Association of Information and Image Management, www.AIIM.org), and the CRM (Certified Records Manager) (from the ICRM, the Institute of Certified Records Managers, an affiliate of ARMA International, the Association of Records Managers and Administrators, www.ARMA.org).

Contact:

SteveGilheany@ArchiveBuilders.com
Tel: +1 310-937-7000 Fax: +1 310-937-7001

For more information, courses, and papers:

http://www.ArchiveBuilders.com