Rumors of My Erasure are Premature

Tom Coughlin
Coughlin Associates
www.tomcoughlin.com

For the July Issue of Storage and Entertainment, A WestWorld Publication

Computer data storage devices are designed to protect user data at all costs. This includes protection against accidental erasure. This is done with “recycle” folders and unerase commands. Drives use elaborate error detection and correction techniques to insure they never return *incorrect* user data. All this implies that true file erasure is an abnormal situation.

Consequently, user data often remains stored on disk drives when they are discarded, transferred to another user, or returned off lease. Even if users delete their files, they can be recovered from “recycling” folders or by special programs such as Norton Unerase.

Because of the industry focus on saving data sometimes the data on disk drives falls into the hands of others. Besides theft of computers and disk drives often data can be easily recovered from discarded or sold disk drives. Earlier this year two students at MIT, Simson Garfinkel and Abhi Shelat reported in the journal IEEE Security & Privacy that they bought 158 used hard drives at secondhand computer stores and on eBay. 129 of these drives were functional. 69 of these still had recoverable files on them. 49 contained "significant personal information" including medical correspondence, love letters, pornography and 5,000 credit card numbers. One even had a year's worth of transactions with account numbers from a cash machine in Illinois.

There is a long history of personal information turning up on used hard drives. This history raises concerns about privacy and identity theft. In 2002 Pennsylvania sold used computers containing information about state employees. In 1997, a Nevada woman bought a used computer and found it contained prescription records for 2,000 customers of an Arizona pharmacy.

Gartner Dataquest estimates that 150,000 hard drives were "retired" in 2002. Many of these drives are thrown away, but a significant percentage find their way back onto the market.

Experience gained from supporting customers with computer security and information storage media issues have shown a tremendous need for the capability to reliably eradicate data from computer hard drives for security and privacy reasons. These needs arise for different reasons depending upon the application.

Mainframes and storage networks:

- When a user releases storage, a drive moves to a new storage server, is removed for maintenance, or returned from lease.
- Storage devices are re-configured for other uses or users, for instance in expiring leased data storage facilities at an SSP or data center
- A RAID drive backs up data to a hot spare

Individual user PCs and workstations:

- A computer (and hard drive) is replaced by a newer machine and the older machine is discarded, or sold
- A project is completed and the data must be purged to protect "need to know" or to prepare the drives for new users or applications.
- When a user departs an organization and either leaves sensitive/personal data on the computer or may take the computer (and the organization's data) with them.
- When a drive is to be returned to a drive manufacturer or a drive repair facility after a drive failure or near failure (for instance upon a SMART drive replacement after imminent failure is determined).
- Data on a drive must be erased to protect digital content from unauthorized access
- A virus has been detected and all possible traces of the offending code must be eliminated.
- An extreme virus or hacker attack where it is desirable to completely erase the data on some disks and reinstall back-up data

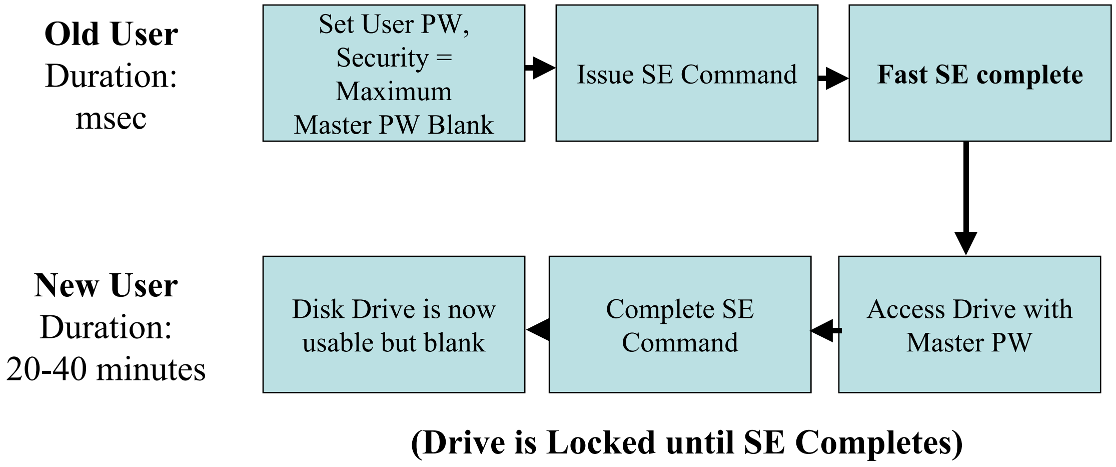
The elimination of unwanted data from a computer hard drive is not a simple task. Deleting a file merely removes its name from the directory structure; the data itself remains in the drive's data storage sectors where it can be retrieved until the sectors are overwritten. Reformatting a hard drive clears the file directory and severs the links between storage sectors, but the data can be recovered until the sectors are overwritten. Software utilities that overwrite individual data files (or an entire hard drive) are susceptible to error and require constant modifications to accommodate new hardware and evolving computer operating systems. As a consequence, computer users, system administrators, security personnel and service providers have spent considerable time in an endless game of technology catch-up while trying to develop solutions for the above problems. As an example of the vulnerability of traditional data security measures in the MIT study 51 of the 129 working drives in the had been reformatted, and 19 of them still contained recoverable data.

So what's a paranoid computer user or IT person supposed to do? Fortunately the common disk drive interface standards, ATA (also known as IDE) and SCSI contain a feature known as Secure Erase (SE). Secure erase is a positive, easy-to-use *data destroy* command, amounting to "Electronic data shredding." It completely erases all possible user data areas by overwriting , including the so-called g-lists that contain data in reallocated disk sectors (sectors that the drive quits using for data because they have hard errors in them). SE is a simple addition to the existing “format drive” command currently present in computer operating systems and storage systems, and consequently adds little or no cost to drives. In addition security erase does not require any additional software to implement.

Secure erasure capability will be required by the U.S. government for their disk drive purchases. Considering the security feature this capability offers to many users I expect there will be considerable commercial interest in this capability as well.

A full secure erase can be a lengthy process for today's high capacity disk drives, sometimes requiring more than 30 minutes to complete. For this reason an ATA disk drive user may want to do a "Fast Secure Erase" on a disk drive before disposing of it (see Figure 1). ATA disk drives can have a user "password" that is used to access certain features of the disk drive. If a secure erase is started using a user "password" even if it is stopped before completion and another user acquires the drive and uses the master "password" to reactivate the disk drive the disk drive must complete the secure erase before it can complete any other command. This fast secure erase prevents easy and casual access to data on discarded hard disk drives while allowing them to be available for resale, return to vendors, or donated. Data overwriting with more than one pass of the write head where the write head is mechanically set slightly off-track during each a first and second write on each track would provide an even higher level of security.

Figure 1. Fast Secure Erase


Old User
Duration:
msec
Set User PW, Security = Maximum Master PW Blank
Issue SE Command
Fast SE complete
New User
Duration:
20-40 minutes
Disk Drive is now usable but blank
Complete SE Command
Access Drive with Master PW
(Drive is Locked until SE Completes)

Secure erase is required by the ATA specification, although it is optional in SCSI. The new “serial ATA” drives will be able to advertise SE as a user feature, in their competition with SCSI and Fibre Channel drives for market share in low-end storage systems. SCSI and Fibre Channel disk drives have other system commands that allow erasure of data on the drives from the storage system level. These SCSI and Fibre Channel disk drives do not contain "Fast Secure Erase" capability.

If you want to make sure that your old data really is erased and unavailable to others you definitely want to look into Secure Erase!

For additional information on advanced disk drive monitoring (SMART), Secure Erase and other intelligent disk drive functions you can contact Dr. Gordon Hughes of CMRR at University of California, San Diego. Gordon's email address is gfhughes@ucsd.edu.

Biography

Thomas M. Coughlin

1665 Willowmont Ave.
San Jose, CA 95124
408-202-5098 (phone)
tom@tomcoughlin.com
www.tomcoughlin.com

Thomas M. Coughlin is President of Coughlin Associates, a data storage consulting firm specializing in data storage components, systems, and software used in consumer electronic, enterprise, and entertainment applications. He has over 20 years of industrial experience in data storage engineering, product development, program management, and market and technology assessment at such companies as 3M, Polaroid, Seagate Technology, Maxtor, Micropolis, Nashua Computer Products, Ampex and SyQuest. His consulting work has included clients such as Ampex, Raychem, Network Appliance, PriceWaterhouseCoopers, several start-ups, as well as Venture Capitalist and financial clients. Tom has over 50 articles, reports, and technical presentations to his credit and over 10 patents granted or pending. He is currently working on a Data Storage for Entertainment Report due out in Fall 2003. He has been active with IDEMA where he is chairman or a member of several committees, including overall Standards Chairman. He is a senior member of the IEEE, was publicity chairman of the 1992, 1996, and 2002 TMRC conferences and is the chairman of the Santa Clara Valley IEEE Magnetics Society. He organized magnetic recording Symposia for IIST from 1997-2002 at Santa Clara University where he is also an adjunct professor. Tom is the organizer of the annual Storage Visions Conference (www.storagevisions.com), which focuses on data storage and the digital content value chain. He is a senior member of the IEEE, a founding member of the Server Blade Trade Association and a member of IDEMA, ACM, APS, and AVS. The company website is www.tomcoughlin.com and the author's email is tom@tomcoughlin.com.