GIAC
CERTIFICATIONS

Global Information Assurance Certification Paper

Copyright SANS Institute
Author Retains Full Rights

This paper is taken from the GIAC directory of certified professionals. Reposting is not permited without express written permission.

Interested in learning more?

Check out the list of upcoming events offering
"Security Essentials: Network, Endpoint, and Cloud (Security 401)"
at http://www.giac.org/registration/gsec

# A Case for Forensics Tools in Cross-Domain Data Transfers

© SANS Institute 2003, Author retains full rights.

Dwane Knott

GSEC Practical Assignment
Version 1.4B (August 29, 2002)

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

Abstract.

Corporate and government organizations dependence on computers and networks for storage and movement of data raises significant security issues. Two of these are movement of data across security domains (cross-domain) and computer reuse.

The cross-domain transfer problem must address the contents of the file space as well as the contents of slack and free space. Three options are presented and discussed. One is selected as most practical and more fully discussed. Since this option involves the use of forensics software, a software tool is selected and its application discussed. The final discussion is protecting against inadvertent data compromise when reusing computers or salvaging them. Forensics software has a role here also.

Background.

Government and business reliance upon computers presents challenging problems for management, computer support personnel, and security professionals. These challenges include designing an effective network; selecting the right hardware; software selection and configuration; defending against hackers, viruses, worms, and Trojan horse attacks; training users; developing and enforcing security policies; and etc. Many of these issues were simple in the early days of the computer revolution. Computers were slower, applications less robust, internal networks were a novelty and threats fewer and less sophisticated. Today, however, the situation is radically different. Today's computers are many times faster and more powerful, most organizations are connected to the Internet exposing them to hackers and other bad things, and most have internal networks of multiple computers.

Early network design for most organizations was to connect all the computers together in one network. This works well for a small network but as the number of computers and usage increases, performance drops. To regain the performance, organizations have subdivided large networks into smaller units - subnets. These subnets usually serve a functional area or location. For example, a corporation might create subnets aligned by department; i.e. executives, finance, human resources, sales, purchasing, research and development (R&D), etc or by location such as the Boston, New York City, and Baltimore offices. Some serve both function and location. Each of these subnets support personnel with like responsibilities and permits or limits sharing data, e.g. documents, email, and databases, with members of that subnet. The result is the creation of network security domains.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

"A network security domain is a contiguous region of network that operates under a single, uniform security policy."[1] The subnet security policy answers the who, what, where, when, and how questions regarding access to that domain and its information. The organizational network security policy needs to address the issue of the transfer of information between security domains - cross-domain transfers. External entities should be considered just another security domain and the release of data to them included in this policy. Examples of when the cross-domain policy would be applied: 1) the Director of R&D needs to see the resume for the prospective new hire that human resources has; 2) the Director of R&D needs to send Finance his recommendation for a bonus for the development team responsible for the new product; or 3) Human Resources needs to provide the life insurance company with personal information on corporate personnel. The first example does not contain sensitive data while the other two do, one has information regarding the proposed bonus and the other has privacy act information (inferred).

The value and application of security domains is clearly demonstrated in their implementation within the Federal government. Government organizations mark data with security classification markings of which unclassified, secret, and top secret are the most common. The purpose of security classifications is to indicate the sensitivity of the information and the level of protection it should be afforded. Protection includes restricting access to personnel with that level of clearance or higher and only in facilities approved for that level of information or higher. Computers and networks that hold or move the information also must be approved to handle the minimum or higher levels of information. Each constitutes a security domain since they operate under different security policies. The Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP)[2] defines a process for certifying and accrediting networks. This certification becomes part of the security domain policy.

As in the commercial world, government organizations sometimes need to move data between security domains. A person working on a computer in the top-secret security domain creates files of which some will not contain top-secret information. Some will contain data for which secret or unclassified handling is sufficient. However, because of the domain the file resides on, the top-secret policy applies. The data would need to be moved to a lower security domain for less stringent handling to be permitted. A cross-domain transfer would be done in accordance with the organization's policy.

[1] "Design the Firewall System". Carnegie Mellon Software Engineering Institute CERT Coordination Center. URL:http://www.cert.org/security-imporvement/practices/p053.html (5/14/2003).
[2] Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP), URL: http://iase.disa.mil/ditscap/ditsdoc.html, 23 May 2003.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

Problem.

The requirement to move data from a network with a particular security policy to a network with a different security policy presents a challenge that extends throughout government and the commercial arenas. This problem becomes even more significant when the data is to be released to an external organization. How does one ensure that the desired information and only the desired information is moved? The result of this process is called ‘sanitization’ and the resultant ‘sanitized data’.

Issues.

The data selected should to be approved for movement. A knowledgeable person should attest or certify that the information is authorized to be moved. The person should consider not only the actual information in the file but also whether it can be combined with other information to form a different classified picture. For example, the anecdote that one could tell when there was a international crisis by the number of pizzas delivered after hours to the Pentagon. Sometimes it is possible to build a picture using facts found in various files. The certifier must consider this when determining classification and authorizing files to be moved. The procedure for this process should be referenced in the security domain policy.

A member of the security staff should be the person authorizing movement following completion of the established review process. This person should not be the same person as the one who certifies that the data can be moved. Separation of responsibilities is practical and conducive to improved security.

The security person, one who will authorize and perform the data movement between security domains, should be able to view all of the material to be moved. The files should be uncompressed and unencrypted. If they are not, the person cannot evaluate the data manually and may not be able to using computer programs. If there is a requirement for compression or encryption, this should be done after security review.

The process for evaluating the data and determining suitability for movement to the new domain must be clearly defined. The process may be unique for different domains. It would be appropriate that the process for moving data within different internal security domains be different from that of data moving from an organizational domain to an outside domain.

Electronic movement of data between domains requires that one consider the contents of the file and its “slack space”, slack space being the space between the end of the file data and the end of the space allocated for file storage. This space exists because hard disks are divided into clusters, i.e. fixed

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

segments of storage space, when they are formatted and partitioned and only one file at a time may be assigned use of a cluster. If a file does not fully fill a cluster, most operating systems will fill the unused portion of the cluster with data from random areas of the computer's memory or cache. If the cluster size is large, as occurs on newer, larger hard drives, and the file data actually written to the cluster is small then it is possible that large amounts of personal or corporate data could be attached to the file when transferred.

The transfer of data using magnetic media presents an additional challenge in that one must also evaluate the contents of the unassigned or free space on the media. These are clusters on the media that are not currently assigned to file storage. This is significant when the media is not new. If the media has ever been used, it is likely that there is residual data present. This is because Microsoft simplified the process for deleting files for the sake of speed. Files are deleted by changing their status in the file location table but the space on the media is not over-written. Therefore, with the proper programs it is possible to retrieve the old contents from any cluster that had been used. Only new magnetic media should be used to send files to external organizations.

Microsoft's files present particularly challenging issues that include the ability to embed differing types of Microsoft files into other types of Microsoft files. For example, a Microsoft Excel spreadsheet can be embedded into a Microsoft Word document. The typical use would be to create a spreadsheet, select the spreadsheet or part of it, select copy from the menu, open a Word document, and paste the spreadsheet area. One thinks that only the selected area is being copied but in reality all of the file's spreadsheets and objects[3] are copied. This problem exists with other Microsoft products.

Process.

Data should not be moved until a certification from a properly designated individual that the data is authorized for release to the designated security domain or organization is received. The approval authority need not be the data owner but should have an awareness of the organization's policy regarding the specific information being considered for release. The individual should have access to corollary information that permits him/her to evaluate if privileged or classified information can be intimated using the data being released or transferred.  This is especially important when releasing multiple files at the same time. Sometimes one can collate unclassified bits and pieces of information into a classified picture.

The authorization is provided to the security staff and the security review is performed. The security review consists of evaluating the data for specific key

[3] Naval Surface Warfare Center, URL: http://www.nswc.navy.mil/wwwDL/XD/ISSEC/Docs/Lessons_Learned/paste.html, 22 May 2003.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

words or text strings. The options for review include a fully manual review, a fully automated review, and a hybrid manual/automated review.

Manual review is performed when a person opens each file and reviews the file's contents. This is the most time consuming. Imagine reviewing 650mb of documents by reading them (an encyclopedia fits on one CDrom). A danger associated with manual review occurs when the releaser or the security person uses a printed copy to determine suitability for transfer and the file will be moved electronically. This process does not permit evaluating the data that may be stored in the slack space or free space. Paper copy review only meets security concerns when the paper copy is the product moved.

On the opposite end of the spectrum is an automated review. This process requires a computer program be used to search the target files and compare the contents against a rule set. The rule set is prepared and controlled by the organization. The program will take actions against the target file according to the rule set. Some options are to reject and delete the file or modify the file's content. The United States Navy contracted the development of just such a program for use with formatted messages. The product is called Radiant Mercury[4]. Radiant Mercury takes a formatted message input and processes it against a rule set. This type of program is not effective with files that do not follow a rigid format.

The hybrid process combines human review with computer program review. This is the most practical since it uses automation to quickly find target keywords and the human component to determine what action to take. The automated process can do what the human reviewer cannot, that is search the slack space and free space for hidden data on the media. Hidden data can be evaluated against the same rule set used for the target files and allow for an informed decision on whether to release the media outside the organization.

Following review the data is moved to the target security domain. This may involve an electronic transfer within the organization's domains if the domains are at the same security level or placing the data onto magnetic media for "sneakernetting" when the domains are at different levels. Sneakernet has a "person in the loop" in that the data is downloaded onto tape or disk on a workstation on one security domain and manually transported to a workstation on the second domain where it is uploaded onto that domain. New magnetic media should be used for sneakernetting for reasons previously discussed – delete leaves data in the storage clusters.

Hybrid Evaluation

The process of using computer programs to search files (data) against a keyword list has been used by law enforcement agencies for years. For

[4] FAS, "Radiant Mercury [RM]", URL: http://www.fas.org/irp/program/disseminate/radiant_mercury.htm, 22 May 2003.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

example, New Technologies, Inc has been providing such software to law enforcement since 1996.[5] These programs are referred to as computer forensics, a term coined in 1991[6], software since they permit a throrough investigation of the computer's random-access memory, cache memory, and hard drive. Most of the programs can be used against media such as floppy and zip disks, jazz drives, and external hard drives. These programs have provided the computer forensics scientist with the capabilities to extract data that has led to the successful prosecution for crimes such as child pornography, fraud, etc.

These programs are suitable for use by organizations needing to do cross-domain transfers or release information outside the organization. New Technologies, Inc (NTI) TextSearch PLUS is a program that possesses the capabilities to search files, slack space and free space on floppy, zip and hard drives. TextSearch PLUS searches the target media "for key words or specific patterns of text"[7] in the file and its slack space. The key words or patterns of text used in the search are either created as a file that TextSearch PLUS references or entered directly when starting the program. Depending on the configuration settings set by the user; TextSearch PLUS will either stop after finding a keyword or complete the search and then provide a list of all occurrences of the target words or text patterns. Additional information is provided which allow the user to find and evaluate the context of the keyword. The human reviewer must locate the word or pattern and determine if its presence is cause to reject the file. If configured to, TextSearch PLUS will include the file's slack space and/or the magnetic media's free space in the search. The returns are included into the total returns for the selected file(s).

This process should be all that is required to move a document between security domains but is not always the case. The selected file, including its slack space and free space when appropriate, may been evaluated and all instances of the key words and text strings noted. The human reviewer notes that there is no "hits" (found instances of the keyword returns) or determined that the file is okay to move despite the hits. Though the hybrid process concludes the file appropriate to move, it may not be due to improper evaluation criteria.

The keyword list is essential to successfully evaluating a file for transfer between security domains. TextSearch PLUS returns hits or does not return hits based on a comparison of the file contents against the keyword list. If the list contains an error or the file contains a misspelled target word, no hit will be generated. No hits will be generated, obviously, if the keyword list does not contain the word or text string. Therefore, the reviewer may note the hits or lack of them, verify those that are there and determine they are innocuous, and authorize the file to be moved using a faulty word list. Despite the best intentions, the process will have failed and the file should have not been moved.

[5] New Technologies, Inc., URL: http://www.forensics-intl.com/aboutus.html, 22 May 2003.
[6] New Technologies, Inc., URL: http://www.forensics-intl.com/def4.html, 22 May 2003.
[7] New Technologies, Inc., URL: http://www.secure-data.com/txtsrchp.html, 22 May 2003.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

The key word list is not always easy to create since considerable thought must go into the selection of the text strings/keywords. TextSearch PLUS only permits fifty (50) search terms that can be text strings or words. An example of a search term is the text string ‘sec’. This term yields hits on any occurrence of secret, secretary, top secret, security, seclusion, secular, and many additional words. This is probably not a good search string since it could return many irrelevant hits. “Secret” would be a better term since it significantly reduces the number of words by eliminating security, seclusion, secular, and many others. The keyword list should not be a static list, but it should reflect current issues and be updated as areas of emphasis change. Using multiple lists and specifying which to use for a specific search can overcome the limitation of 50 entries.

Hidden Data and Computer Reuse.

Organizations face the problem of what to do with computers that have been used. Sometimes, they are reallocated within the same security domain while other times they are designated to be salvaged. (“Salvaged” could be donating the use equipment to charity or selling it). Reusing the computer within the same security domain does not present significant security issues. However, salvaging can create security issues when the computer being salvaged is believed to have been used only for unclassified (non-sensitive) processing.

Most organizations have policies that require magnetic media to be removed before salvaging computers that have been used on sensitive security domains. The hard drives and sometimes the computer memory are removed before sending the computer to salvage. The policy for computers used in unclassified or non-sensitive applications normally only requires the hard drive to be formatted before being sent to salvage. Such a policy is ripe for disaster.

Noted previously, Microsoft’s delete function does not overwrite or delete information from disk clusters. Microsoft’s format function does not either. The data is still there waiting for a program such as TextSearch PLUS to be used to recover the data.  Place a floppy disk with TextSearch PLUS into the floppy drive, reboot the computer, configure and begin search. Voila, the secrets are revealed. TextSearch PLUS does not search Microsoft NTFS formatted drives, but NTI’s DiskSearch NT[8] will.

Why the concern, when the computer was used for non-sensitive or unclassified work anyway? The problem is that computers collect information and retain them without the user being aware. For example, you receive a proposal for new work to review. The originator mails it to you with the caveat that it is unclassified but contains some proprietary information. The mail program shows the file as an attachment and since you use a POP3 mail account, the attachment gets saved to your hard drives in the mail programs attachment

[8] NEW Technologies, Inc., URL: http://www.forensics-intl.com/suite9.html, 23 May 2003.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

folder. You select the attachment and ask it be opened but not saved to disk. Since it is a Microsoft Word file, Word creates a temp file and saves the contents of the attachment in it. You finish the review, noting on paper your comments, close Word and delete the message. Word deletes the temp file and the mail program deletes the message and the attachment. No harm, you didn't save any sensitive data to disk but the computer did - twice. This is an example of how hidden data can get onto a hard drive as a result of normal usage. This data will remain on the drive until over-written which might not occur for an extended period of time.

Most applications use some type of temporary storage as part of normal operations. Sometimes this temp space will be returned for system use and sometimes it isn't. When the files are not deleted, one can search for them but finding and deleting them does not remove the data. Any data written to the hard drive will be on it until that cluster is overwritten by another set of data.

The only way to get rid of the data is to use a program specifically designed to overwrite free space with a preset value. Such a program is M-Sweep Pro[9]. A careful user could then use TextSearch PLUS as a quick means to verify that the data was successfully removed by M-Sweep Pro or use NTI's GetSlack NT and GetFree NT[10] for a more thorough check.

## TextSearch PLUS.

TextSearch PLUS is a DOS based forensics tool used to locate text strings on floppy, zip, and hard drives. This makes it ideal for use in evaluating files for cross-domain transfers. The program will start in a DOS window and uses the arrow keys to move through the menus. Highlighting a menu option and hitting the <enter> key will enable selecting that option for change. Some menu options are toggles and some are text entry.

TextSearch PLUS allows the user to search file space, slack space and free space either separately or in the same pass. The user can direct the program to continue search when finding a target text string or to stop after each occurrence. The program output can be logged to a file or sent to a printer. File types can be selected to be excluded from the search. The simplicity of the interface and clarity of descriptions in the 19-page manual provide adequate guidance for even a novice user. The only complicated part is creating the keyword list and that is covered adequately in the manual[11].

## Summary.

[9] New Technologies, Inc., URL: http://www.secure-data.com/ms.html, 22 May 2003.
[10] NEW Technologies, Inc., URL: http://www.forensics-intl.com/suite9.html, 22 May 2003.
[11] New Technologies Inc., "TextSearch Plus (™)", ©1996-1999.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

Computers have become an essential appliance in most government and corporate offices. Today almost every office worker has one. Computers are networked and information flows freely between workers. At one time, networks existed only in the largest corporations or government offices. Now subnets are common in both large and small organizations and used to separate personnel into workgroups with each workgroup having its own security policy and forming its own security domain.

The requirement to move data between security domains or to release to external organizations is readily appreciated. The problems with this requirement are also valued. The volume and source media prevent manual review since the media cannot be evaluated and is too time consuming. Fully automated review is only practical when the source is strictly structured. The hybrid review is the only practical answer to the problem. Software must be used to evaluate data on magnetic media, finding the target information, with the human reviewer deciding the importance of the found information.

Ultimately the policies and procedures must account for all these requirements while enabling the organization to meet its mission.

© SANS Institute 2003, Author retains full rights.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.

# References

1. Design the Firewall System”. Carnegie Mellon Software Engineering Institute CERT Coordination Center. URL:http://www.cert.org/security-imporvement/practices/p053.html (5/14/2003).

2. Department of Defense Information Technology Security Certification and Accreditation Process (DITSCAP), URL: http://iase.disa.mil/ditscap/ditsdoc.html, 23 May 2003.

3. Naval Surface Warfare Center, URL: http://www.nswc.navy.mil/wwwDL/XD/ISSEC/Docs/Lessons_Learned/paste.html, 22 May 2003.

4. FAS, “Radiant Mercury [RM]”, URL: http://www.fas.org/irp/program/disseminate/radiant_mercury.htm, 22 May 2003.

5. New Technologies, Inc., URL: http://www.forensics-intl.com/aboutus.html, 22 May 2003.

6. New Technologies, Inc., URL: http://www.forensics-intl.com/def4.html, 22 May 2003.

7. New Technologies, Inc., URL: http://www.secure-data.com/txtsrchp.html, 22 May 2003.

8. NEW Technologies, Inc., URL: http://www.forensics-intl.com/suite9.html, 23 May 2003.

9. New Technologies, Inc., URL: http://www.secure-data.com/ms.html, 22 May 2003.

10. New Technologies Inc., “TextSearch Plus (™)”, ©1996-1999.

© SANS Institute 2003, Author retains full rights.

© SANS Institute 2003, As part of GIAC practical repository. Author retains full rights.