Annual ADFSL Conference on Digital Forensics, Security and Law

2009
Proceedings

May 21st, 2:30 PM

Concerning File Slack

Stephen P. Larson
VCU School of Business, Richmond, VA, stephen.larson@sru.edu

Concerning File slack

Stephen P Larson
VCU School of Business, Richmond, VA
larsonsp@vcu.edu

ABSTRACT

In this paper we discuss the phenomena known as file slack. File slack is created each time a file is created on a hard disk, and can contain private or confidential data. Unfortunately, the methods used by Microsoft Windows operating systems to organize and save files require file slack, and users have no control over what data is saved in file slack. This document will help create awareness about the security issue of file slack and discuss research results concerning file slack.

Keywords : Computer Forensics, File Slack, Ram Slack, Disk Slack

1. INTRODUCTION

In this digital age, keeping personal or confidential data private is quite difficult. Regrettably, this problem is exacerbated by the very technology we use to create the digital data. It has already been established that a problem exists with users and companies selling off old hard disks that still contain commercial or personal data, even if the hard disk has been formatted (Garfinkel and Shelat, 2003; Jones, Valli, Sutherland, and Thomas, 2006; Jones, Valli, Dardick, and Sutherland, 2008). What has yet to be established is the extent of commercial, private, or personal data that can be transmitted via file slack. "File slack" can contain data dumped randomly from the computer's memory, data from previously deleted files, etc., and can potentially reveal prior uses of the computer such as fragments of email messages, network or internet site logon names and passwords, etc (Volonino, Anzaldua, and Godwin, 2006).

For this paper, we will limit our discussion to Microsoft Windows OS because "unlike Windows … file systems, UNIX does not have file slack space. When UNIX creates a new file, it writes the remainder of the block with zeros and sets them as unallocated. Therefore it is not possible to recover deleted data from slack space on UNIX systems" (Casey 2004). According to NetApplications, roughly 90% of operating systems on PCs are some version of Microsoft Windows (NetApplications 2008). Additionally, Steve Ballmer of Microsoft stated that "forty percent of servers run Windows" (Niccolai 2008).

"File slack is a data storage area most users are unaware of" (Vacca, 2002). "It is a source of significant *security leakage* and consists of raw memory dumps that occur during the work session as files are closed" (Vacca, 2005).

This paper will introduce how data is saved on hard disk drives, give definitions of file slack, ram slack, and disk slack, explain how file fragmentation affects file slack, and explain how file slack on file servers is shared among users, and determine whether file slack is "portable." We will also discuss future research needs for file slack.

2. HOW DATA IS SAVED ON HARD DISK DRIVES

During its manufacture, a low level format of the hard disk drive is done by the manufacturer to ready it to be partitioned into one or more logical partitions or volumes. During the low-level format tracks and sectors are created. A track is a concentric ring around the platter containing information. Hard drives typically contain several platters, with the tracks on each platter lining up. The tracks are then divided up into sectors. A sector is the smallest unit of the hard drive and is 512 bytes in size. A cluster contains one or more sectors. "All Microsoft operating systems read and write in blocks of data called *clusters*" (Volonino, et al, 2006).

More simply put (see Figure 1):

- A platter contains concentric tracks
- A track contains 512 byte sectors
- A cluster contains one or more sectors
- A cluster is the smallest unit on disk for storing a file

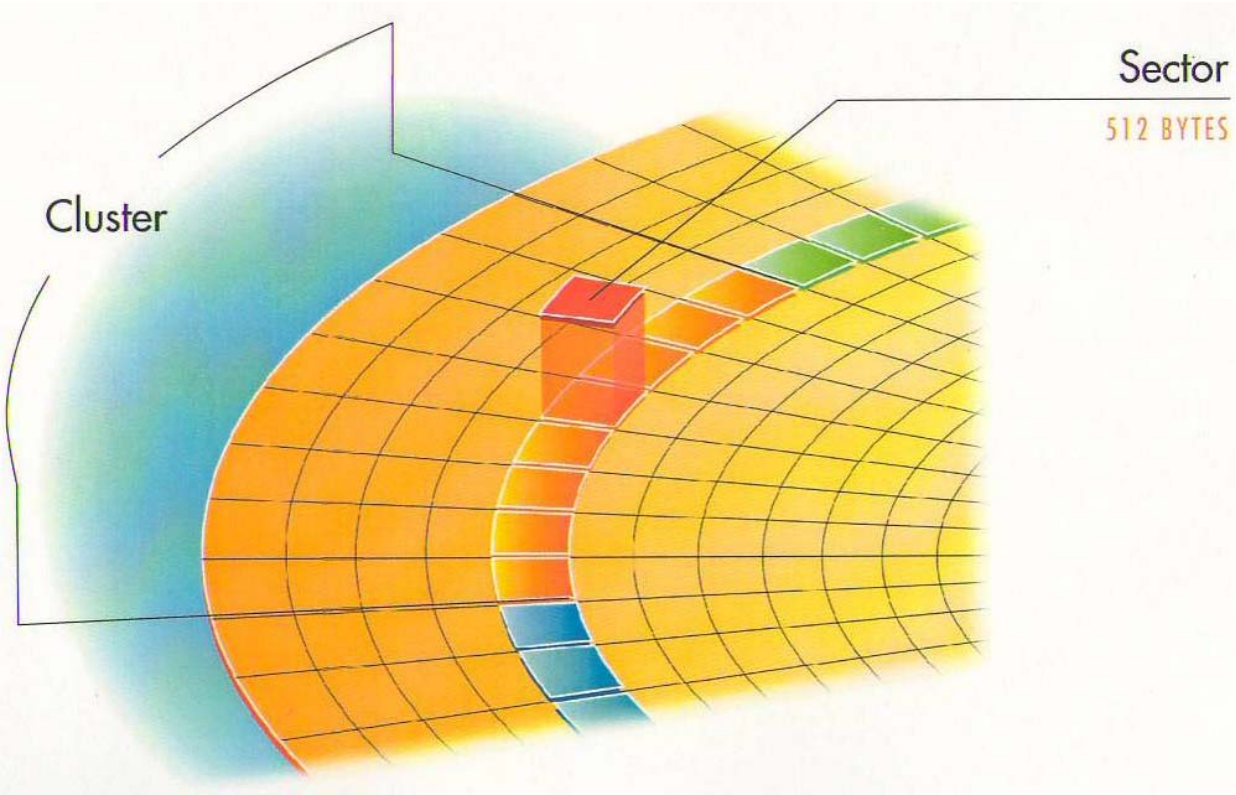


Figure 1. Sectors and Clusters (New Technologies, Inc. 2001b)

In Microsoft Windows operating systems, the default cluster sizes are shown in Table 1. As most PCs running a Windows OS have hard disk drives with partitions or volumes greater than 2 GB but less than 2 TB in size, the most common cluster size is 4 KB; each cluster contains 8 sectors that are 512 bytes in size.

Volume or partition size	NTFS cluster size
7 MB – 512 MB	512 bytes
513 MB – 1,024 MB (1 GB)	1 KB
1,025 MB – 2 GB	2 KB
2 GB – 2 TB	4 KB

Table 1. Default cluster sizes for the NT File System in Microsoft Windows (Microsoft 2007).

3. FILE SLACK, RAM SLACK, AND DRIVE SLACK

How does cluster size affect the size of files on your hard disk? A file on the hard disk must be the same size as a default cluster size; currently the most common cluster size is 4 KB, or multiple of 4 KB. But file sizes rarely exactly match the size of the clusters. The space that exists from the end of the file to the end of the last cluster assigned to the file is called "file slack". Larger cluster sizes mean more file slack and also the waste of storage space (Reyes & Wiles 2007). A cluster size of 4 KB means there is a potential for 3.9 KB of space wasted for a file; this also means a potential for 3.9 KB of unwanted file slack being attached to the end of a file.

3.1 File Slack

Let's create a text file with the word "hello" in it and save it on the desktop as a file named 5byte. Then we will view the properties of the file by right-clicking on the file and choosing "properties." The file's properties are shown in Figure 2.

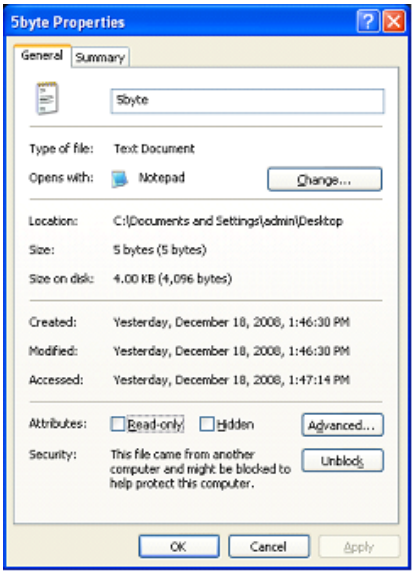


Figure 2. File properties

Notice the size of the file is 5 bytes, but the size on disk is 4 KB. This tells us that there is 4,091 bytes of file slack attached on the end of this file. What are the contents of file slack? "Ram slack" plus "disk slack."

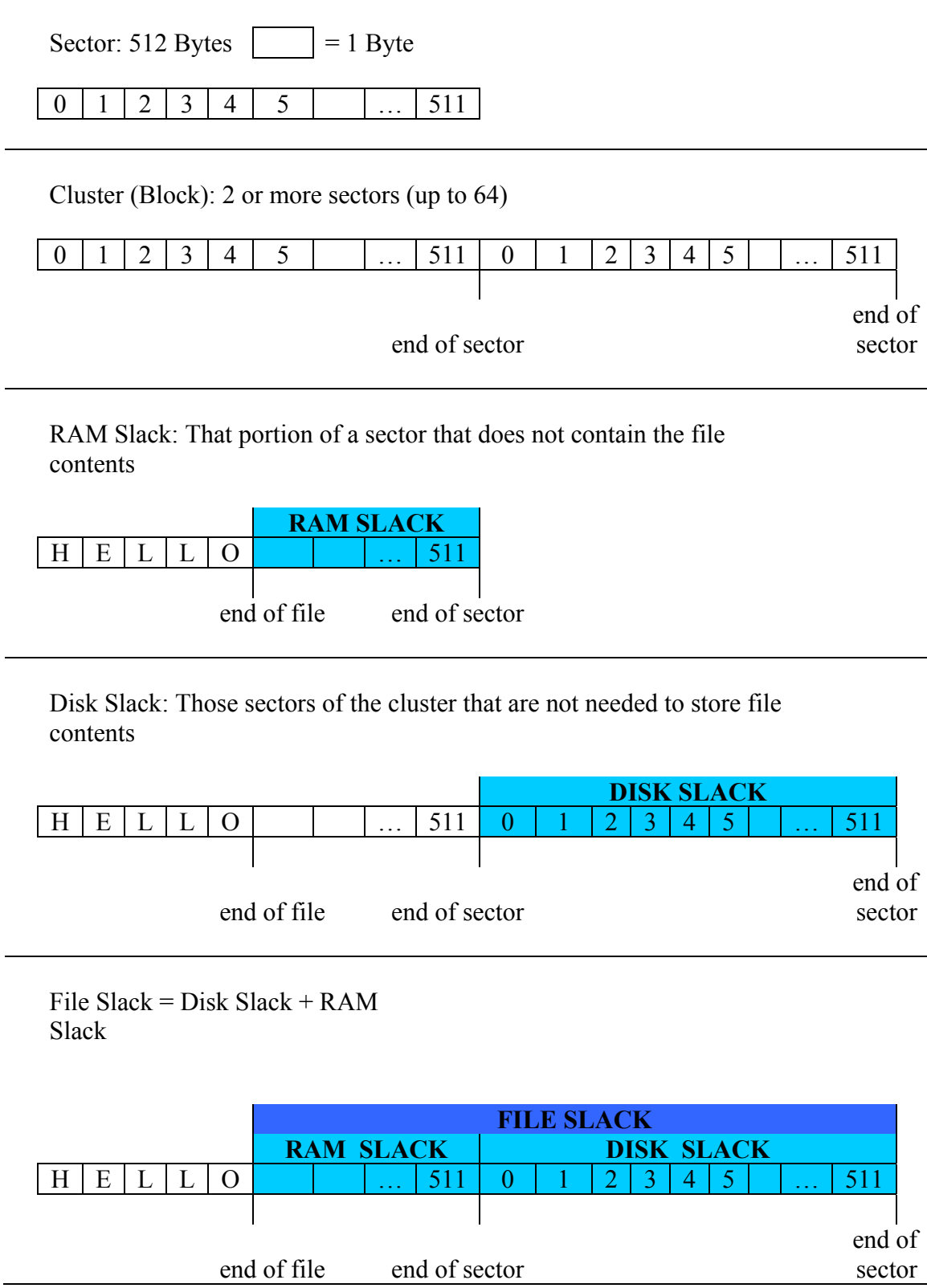
3.2 Ram Slack and Disk Slack

Using our example file, the 4,091 bytes of "file slack" contains "ram slack" and "disk slack."

The ram slack is the data required to fill in the space from the end of the file to the end of the sector. As previously discussed, a sector is 512 bytes. The first 5 bytes of the file are used by the text of the file. The next 507 bytes are filled with "ram slack". Why is this called ram slack? The Windows operating system used to fill in the space from the end of the file to the end of the sector with randomly selected data pulled in from RAM (New Technologies, Inc. 2001a), but now will fill in the space with zeros.

That leaves 3.5 KB, or 7 sectors of 512 bytes each, of space to fill in. Disk slack is the space from the end of the sector that contains the end of the file to the end of the cluster. This can contain one or more sectors. Because the file does not have a need for this space, but the file system needs to fill in the space with something, the data that was previously in the sectors is used and not overwritten. A graphical representation of file slack (containing ram slack and disk slack), using a cluster size of 2 sectors or 1KB, is shown in Figure 3.

Figure 3. Sectors, Clusters, RAM slack, Disk slack, and File Slack (adapted from Dampier 2008).



3.3 File Slack and Disk Fragmentation

Any time a file is deleted, the clusters on the hard disk which were used by the file is marked free or unallocated by the file system. These clusters either get overwritten by another file or become disk slack. As you use your hard disk, clusters belonging to a file can become scattered all over the hard disk. This phenomenon is called fragmentation. To illustrate, let's consider an example. You start off with a newly formatted hard disk using 4 KB clusters. If you save a file that is 8 KB in size, the file system will use clusters 1 and 2 for that file. The next file you create is 12 KB in size. Because clusters 1 and 2 are in use, the file system will start the new file in cluster 3 and end in cluster 5. All clusters in a file contain information about the physical location of each other on the hard disk. Next you delete the first file you created, and create another file that is 3 KB in size. The file system will see that the first available cluster is cluster 1, and will save the file in cluster 1. However, because the file is only 3 KB and the cluster is 4 KB in size, on KB of file slack, which contains data from the deleted file, is "attached" to the new file. You then create a fourth file that is 8 KB in size. The file system sees that the first available cluster is cluster 2, and starts saving the file in that cluster. When cluster 2 is filled, the file system finds the next available cluster, cluster 6, and saves the rest of the file there. File number 4 is fragmented. After normal use for even a few months, a hard disk can become heavily fragmented. The computer on which this paper was written has been in use for 5 months, and has 4279 fragmented files and 23892 excess fragments, with an average of 1.48 fragments per file. Anytime a file gets deleted, its clusters, which could be scattered all over the hard disk, become available for use by another file. Recent versions of MS Windows can be configured to run disk defragmentation automatically and continuously to minimize fragmentation. This process also moves parts of files around the disk. Thus, file slack could contain any type of data.

3.4 File slack on file servers

Until now, our discussion has focused on file slack on a personal computer. The file slack issue is compounded on a file server. File servers using a Windows operating system store information in clusters (Volonino, et al. 2006). Consequently, when a file is deleted, its clusters are marked as unallocated and can end up as file slack in another user's file. Thus, if all departments use the same file server to store their files, confidential Human Resources or Finance information, such as social security numbers or salaries, can become part of the file slack in another user's file.

4. DISCUSSION

Due to the nature of how Windows OS stores data in clusters, there is a potential for private, personal, or commercially sensitive data to be included with a file in the file slack area. File slack is a primary source of electronic evidence because the clusters that made up deleted files are released by the operating system and are unallocated until overwritten by new file content.

4.1 Questions about file slack

Exploring file slack brought to mind certain questions:

- Does file slack accompany a file when it is emailed?
- Does file slack accompany a file when it is copied to another disk or media (hard disk, USB, CDROM, etc.)?
- Does file slack accompany a file when FTP is used to copy the file to another location?
- Does file slack accompany a file when the file is saved under a different name?
- Does file slack from a file on a file server accompany the file when it is sent to a different location?
- Does the recipient's OS add file slack from its own hard disk or file server's hard disk?

To answer these questions, I started with a USB disk that had been prepared by formatting it, creating a file with easily viewable text on it, then erased that file, and created the file mentioned above (5byte.txt). Using a blank USB disk instead of a hard disk drive ensured the file wouldn't be created in the MFT of the Windows XP operating system's Master File Table. "If the data in the file is small (typically a few hundred bytes), then this data can be completely contained within the Master File Table (MFT) record of the file" (Microsoft, 2004). A file saved in the MFT will not exhibit the same file slack as a file saved elsewhere.

After creating the file, I then proceeded to run tests to answer the above questions. I used ProDiscover and FTK to verify the following results:

Question	Result
Does file slack accompany a file when it is emailed?	No (file contents copied, slack from recipient disk)
Does file slack accompany a file when it is copied to another disk or media (hard disk, USB, CDROM, etc.)?	No (file contents copied, slack from recipient disk)
Does file slack accompany a file when FTP is used to copy the file to another location?	No*
Does file slack accompany a file when the file is saved under a different name?	No (file contents copied, slack from new portion of disk)
Does file slack from a file on a file server accompany the file when it is sent to a different location?	Untested**
When copying a file from a file server to a PC, does the recipient's OS add file slack from the recipient's hard disk or file server's hard disk?	Recipient's hard disk
Does file slack occur on a CDRW?	Yes (contents from previously deleted files were found in file slack)

*An secure FTP client was not used. ** Examining files and file slack on a source file server and destination file server to compare the contents of file slack was beyond the scope of this investigation.

I invite readers to verify these findings and provide feedback.

5. FUTURE RESEARCH

Clearly this study is only the beginning of necessary research on file slack and its security implications. During a literature search, an article mentioned that versions of Linux using the ext2 file system can have file slack: "if the file is removed by /bin/rm, its content still remains on disk, unless overwritten by other files," and mentions the use of an obscure tool called bmap that can insert data into the slack space of files. (Chuvakin, 2002).

Other areas that need closer inspection is MAC's HFS (Hierarchical File System) and HFS+, the various "flavors" of Unix, transferring files with a secure ftp client, file slack in the MFT, etc.

REFERENCES

Casey, E (2004) Digital Evidence and Computer Crime, Second Edition, Academic Press, p 301.

Chuvakin A. (2002) LinuxSecurity.com, available from www.linuxsecurity.com/content/view/117638/49/, 2002, (accessed 15 April 2009).

Dampier, D. (2008) Introduction to Cyber Crime and Computer Forensics, at www.cse.msstate.edu/~dampier/CSE6273/Slides/CSE6273-Intro-2.ppt, (accessed 18 December 2008).

Garfinkel, S., and Shelat, A (2003) Remembrance of Data Passed: A Study of Disk Sanitization Practices, *IEEE Xplore*, at http://computer.org/security, (downloaded 4 December 2008).

Jones, A., Valli, C., Sutherland, I., and Thomas, P (2006) The 2006 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market, *Journal of Digital Forensics, Security and Law*, Vol. 1:3.

Jones, A., Valli, C., Dardick, G., and Sutherland, I. (2008) The 2007 Analysis of Information Remaining on Disks Offered for Sale on the Second Hand Market, *Journal of Digital Forensics, Security and Law*, Vol. 3:1.

Microsoft WinHec 2004. Local File Systems for Windows. Available from http://download.microsoft.com/download/5/b/5/5b5bec17-ea71-4653-9539-204a672f11cf/LocFileSys.doc, (accessed 15 April 2009).

Microsoft Support Article ID: 140365 Revision: 5.3 (2007) Default cluster size for FAT and NTFS, at http://support.microsoft.com/kb/140365, August 22, 2007, (accessed 15 December 2008).

NetApplications (2008) Top Operating System Share Trend, http://marketshare.hitslink.com/os-market-share.aspx?qprid=9, (accessed 17 December 2008).

New Technologies, Inc. (2001a) File Slack Defined, at www.forensics-intl.com/def6.html, (accessed 17 November 2008).

New Technologies, Inc. (2001b) Sectors, handout received during forensics training 2003.

Niccolai, J (2008) Ballmer Still Searching for an Answer to Google, *IDG News Service*, at http://www.pcworld.com/businesscenter/article/151568/ballmer_still_searching_for_an_answer_to_google.html, September 26, 2008, (accessed 15 December 2008).

Reyes, A. and Wiles, J. (2007) Best Damn Cybercrime and Digital Forensics Book Period, Syngress, p 495.

RAIDS.co.uk (2008) RAID 5, at http://www.raids.co.uk/raid_5.htm, (accessed 21 December 2008).

Vacca, J. (2002) The Essential Guide for Storage Area Networks, Prentice Hall.

Vacca, J. (2005) Computer Forensics: Computer Crime Scene Investigation, Charles River Media, p 244.

Volonino, L., Anzaldua, R., and Godwin, J (2006) Computer Forensics Principles and Practices, Pearson.