# Recovering Deleted Files and Partitions

**MODULE 8**

# Contents

# Recovering Deleted Files and Partitions

## 8.1 LEARNING OBJECTIVES

After going through this unit, you will be able to:

- Explain the anatomy of a disk drive
- Explain data organization in Windows
- Retrieve deleted files
- Define slack space and swap file
- Implement technologies and tools for data and file recovery in windows system during forensic investigation.

## 8.2 RECOVERING DELETED FILES AND PARTITIONS

### 8.2.1 Anatomy of a disc drive

A hard disk drive (HDD), hard disk, hard drive or fixed disk is a data storage device used for storing and retrieving digital information using one or more rigid ("hard") rapidly rotating disks (platters) coated with magnetic material. The platters are paired with magnetic heads arranged on a moving actuator arm, which read and write data to the platter surfaces. Data is accessed in a random-access manner, meaning that individual blocks of data can be stored or retrieved in any order rather than sequentially. HDDs retain stored data even when powered off. The primary characteristics of an HDD are its capacity and performance. Capacity is specified in unit prefixes corresponding to powers of 1000: a 1-terabyte (TB) drive has a capacity of 1,000 gigabytes (GB; where 1 gigabyte = 1 billion bytes). Typically, some of an HDD's capacity is unavailable to the user because it is used by the file system and the computer operating system, and possibly inbuilt redundancy for error correction and recovery.

An HDD records data by magnetizing a thin film of ferromagnetic material on a disk. Sequential changes in the direction of magnetization represent binary data bits. The data is read from the disk by detecting the transitions in magnetization. User data is encoded using an encoding scheme, such as run-length limited encoding, which determines how the data is represented by the magnetic transitions.

In computer disk storage, a *sector* is a subdivision of a track on a magnetic disk or optical disc. Each sector stores a fixed amount of user-accessible data, traditionally 512 bytes for hard disk drives (HDDs) and 2048 bytes for CD-ROMs and DVD-ROMs. Newer HDDs use 4096-byte (4 KB) sectors, which are known as the Advanced Format (AF).

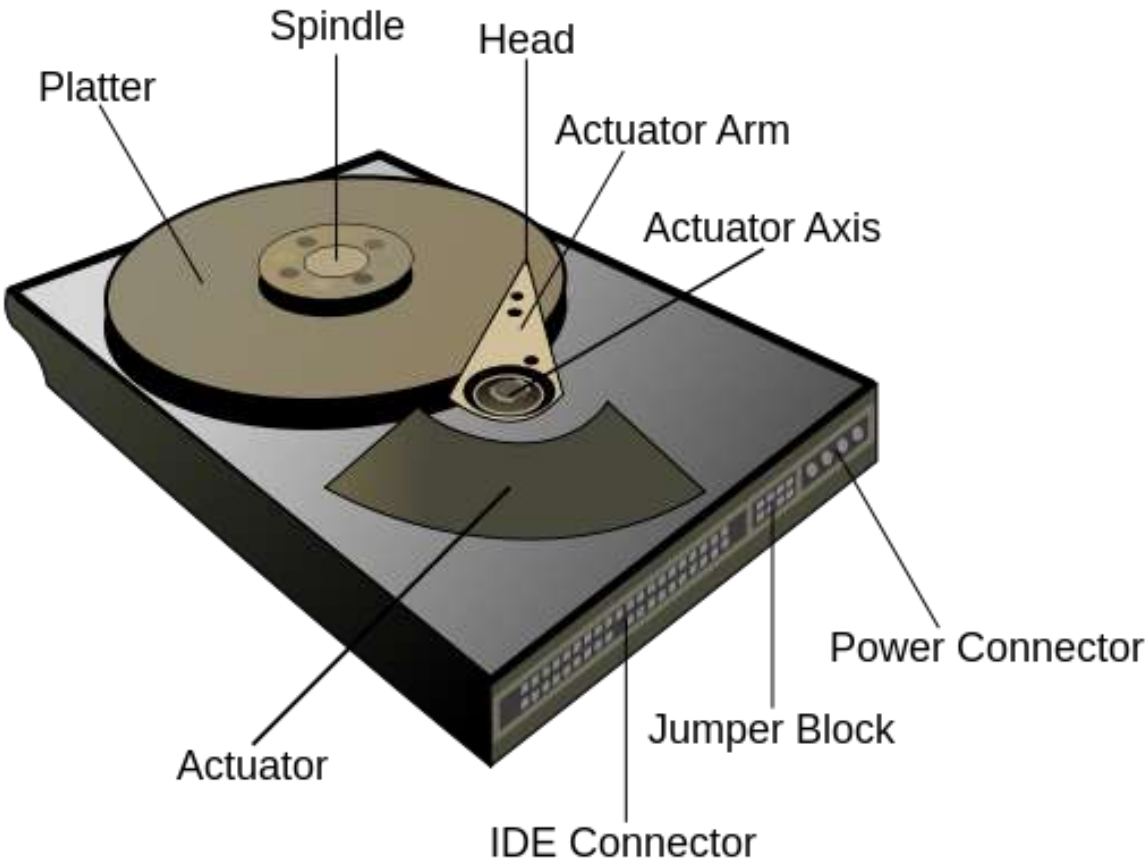



*Figure 1: Hard Disk parts*

Geometrically, the word sector means a portion of a disk between a center, two radii and a corresponding arc (see Figure 17, item B), which is shaped like a slice of a pie. Thus, the *disk sector* (Figure 17, item C) refers to the intersection of a *track* and geometrical *sector*.

In disk drives, each physical sector is made up of three basic parts, the sector header, the data area and the error-correcting code (ECC). The sector header contains information used by the drive and controller; this information includes sync bytes, *address identification*, flaw flag and header parity bytes. The header may also include an alternate address to be used if the data area is undependable. The *address identification* is used to ensure that the mechanics of the drive have positioned the read/write head over the correct location. The data area contains the recorded user data, while the ECC field contains codes based on the data field, which are used to check and possibly correct errors that may have been introduced into the data.

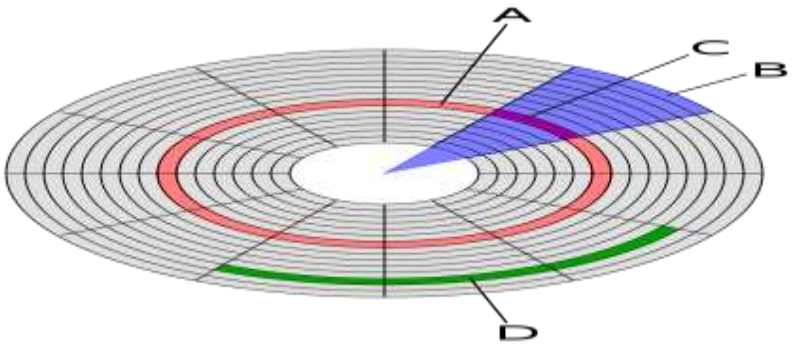



*Figure 2: disk layout*

### 8.2.2 Data organization in Windows

Windows organisation data using following structures or elements:

*Cluster:* Group of sectors form a cluster. Typically, clusters can be of 32 kB. Clusters use logical representation of sectors.

*Partition:* Logical division of the physical storage. A large physical storage needs to be partitioned in smaller size so that the OS can use each partition as separate entity. However, smart user hides data into hidden or temporarily deleted partitions.

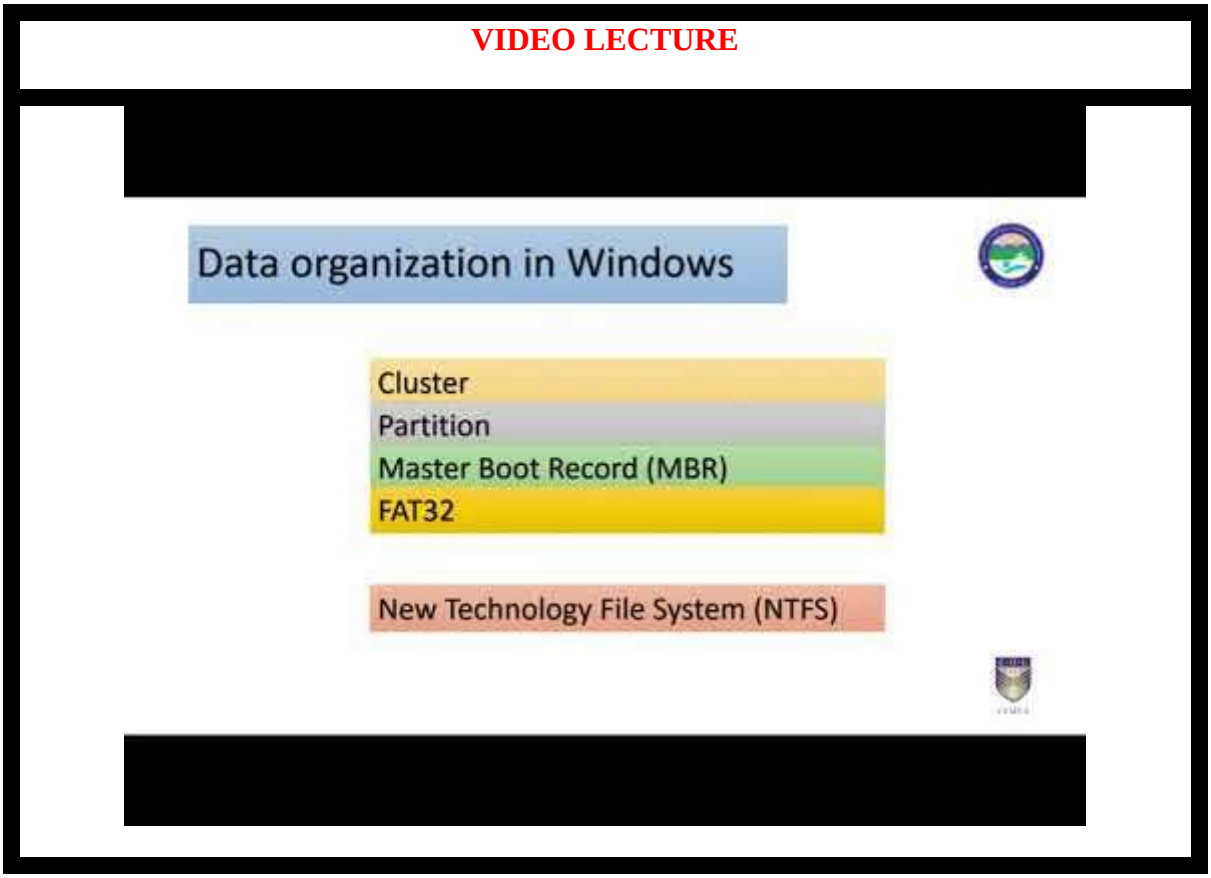
*Master Boot Record (MBR):* Every OS starts with reading a boot record or programme at the first location of a partition that is mapped to the OS hardware but up sequence.

*FAT32:* Initially FAT was the widely used allocation systems. FAT stands for file allocation table and it's a structure that keeps vital Meta data of a file that resides on the hard disc or any other storage. The FAT system uses a certain defines mechanisms to construct/store a file. These pre-defined mechanisms are used very nicely by forensics tools to reconstruct file by forensic tools.

*New Technology File System (NTFS):* The NTFS is a latest standard version introduced by Microsoft which is little advanced in terms of the allocation table structure as well as keeps the data compatible foe other OS to work upon. NTFS is currently used with Window OS. A file in NTFS is deleted in two steps.

a. The file moved to recycle bin and the meta-data is updated. These meta-data can be read by the forensic tools.
b. The clusters occupied (originally) by the file are then marked available for new file and the master file table is also updated. When the user empties the recycle bin the NTFS marks the clusters as available and considers the files to be permanently deleted.

**VIDEO LECTURE**




### 8.2.3 Retrieving deleted files

When a file is deleted, the file system removes the file logically. That is, it removes all the meta-data and stamps related to the file. However, the file still resides in the disk as a physical entity until it is overwritten. These physical areas can be very easily explored and read and converted to a readable file using forensic application. It is observed that data resides on a computer for a very long time and are retrieved to a good extent.

### 8.2.4 Retrieving cached files

One can find the webpage visited by the suspect or the victim by looking into the cache. The cache file of an application can be spread across in the system storage. We can confine only search by using typical keywords elated to the case or probable websites.

### 8.2.5 Retrieving files in unallocated space

In general, a deleted file can be searched sequentially or structurally by looking for file headers or extensions. However, certain tools help us to scan and look for broken headers and use supplementary headers to retrieve data or at least retrieve blocks of a lost file for unallocated space. These retrieved blocks can later be studied and reformed using other tools to retrieve lost files to a great extent. This is also called as file carving.

Meta data of the files can be found from the applications used to create the files however there can be certain tools available to view the metadata of a files like Meta Viewer, Metadata Analysis, iscrub etc.

## 8.3 MORE ABOUT RECOVERING LOST FILES/DATA

### 8.3.1 Slack space, swap file, deleted files

Even these days most of the users aren't careful and thus the forensic experts get more clues because of this. The user's ignorance of how computers manage memory, disks and related stuff leaves lots of spaces which are rather invisible to the user (who can be a subject of an investigation). Let us look at three potential locations where an investigator explores to find lost data as deleted files and slack space, swap space etc.

#### 8.3.1.1 Slack Space

Slack Space is the unused space in a disk cluster. The DOS and Windows file systems use fixed-size clusters. Even if the actual data being stored requires less storage than the cluster size, an entire cluster is reserved for the file. The unused space is called the slack space.

DOS and older Windows systems use a 16-bit file allocation table (FAT), which results in very large cluster sizes for large partitions. For example, if the partition size is 2 GB, each cluster will be 32 K. Even if a file requires only 4 K, the entire 32 K will be allocated, resulting in 28 K of slack space. In computer forensics, slack space is examined because it may contain meaningful data.

#### 8.3.1.2 Swap space

Swap space is the area on a hard disk which is part of the Virtual Memory of your machine, which is a combination of accessible physical memory (RAM) and the swap space. Swap space temporarily holds memory pages that are inactive. Swap space is used when your system decides that it needs physical memory for active processes and there is insufficient unused physical memory available. If the system happens to need more memory resources or space, inactive pages in physical memory are then moved to the swap space therefore freeing up that physical memory for other uses. On a Windows machine, the swap space is a file called Pagefile.sys.

Almost everything on a RAM can be swapped if necessary, because of this we can find very important and forensically interesting things in the swap space. Apart from plain-text data of an encrypted text in a disk file we can even find encryption keys! Thanks to flaw-full weaknesses in some applications that allow unencrypted keys to reside in memory. Also, part of e-mails or matter stored at remote locations might still reside in swap space. And to relief of all investigators, any standard disk maintenance utility can access this information.

On Windows, the swap file is a hidden file found in the root directory called pagefile.sys. The registry path for the swap file is (can be subject to change):

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\SessionManager\MemoryManagement.

Several tools are available to retrieve slack space and swap space on windows system. Slack space can retrieved using a popular tools like DriveSpy, Encase etc. The overall process of retrieving involves following steps:

a) Connect to the computer.
b) Have a bit level image of original media.
c) Keep a hashed value of all images.
d) Use key word searches and hash analysis etc. using tools like Encase.

Tools like DriveSpy can be used to do some of the above processes.

#### 8.3.1.3 File Carving

File carving can be used to recover data from a hard disk where the metadata is missing or damaged, especially by professional data recovery companies.

When a file is deleted, only the entry in the file system metadata is removed, while the actual data is still on the disk. After a format and even a repartitioning it might be that most of raw data is untouched and can be recovered using file carving.

All file systems contain some metadata that describes the actual file system. At a minimum the following is stored: the hierarchy of folders and files, with names for each. For each file is also stored the physical address on the hard disk where the file is stored. As explained below, a file might be scattered in fragments at different physical addresses.

File carving is the process of trying to recover files without this metadata. This is done by analysing the raw data and identifying what it is (text, executable, png, mp3, etc.). This can be done in different ways, but the simplest is to look for headers. For instance, every Java class file has as its first four bytes the hexadecimal value CA FE BA BE. Some files contain footers as well, making it just as simple to identify the ending of the file.

Most file systems, such as FAT and UNIX Fast File System, work with the concept of clusters of an equal and fixed size. For example, a FAT32 file system might be broken into clusters of 4 KB each. Any file smaller than 4 KB fits into a single cluster, and there is never more than one file in each cluster. Files that take up more than 4 KB are allocated across many clusters. Sometimes these clusters are all contiguous, while other times they are scattered across two or potentially many more so called fragments, with each fragment containing a number of contiguous clusters storing one part of the file's data. Obviously large files are more likely to be fragmented.

File carving is a highly complex task, with a potentially huge number of permutations to try. To make this task tractable, carving software typically makes extensive use of models and heuristics. This is necessary not only from a standpoint of execution time, but also for the

accuracy of the results. State of the art file carving algorithms use statistical techniques like sequential hypothesis testing for determining fragmentation points.

### 8.3.1.4 Event logs

Event logs are stored in Metadata files. The entries in these files can be retrieved on a good way depending upon how refining is carried out by investigators. The victim or suspect system log entries change rapidly as the new events are recorded. The event logs can also be configured minimal to maximum events and durations. We can use tools like Ps log list and EVT to retrieve event records. See figure 19,20, 21.

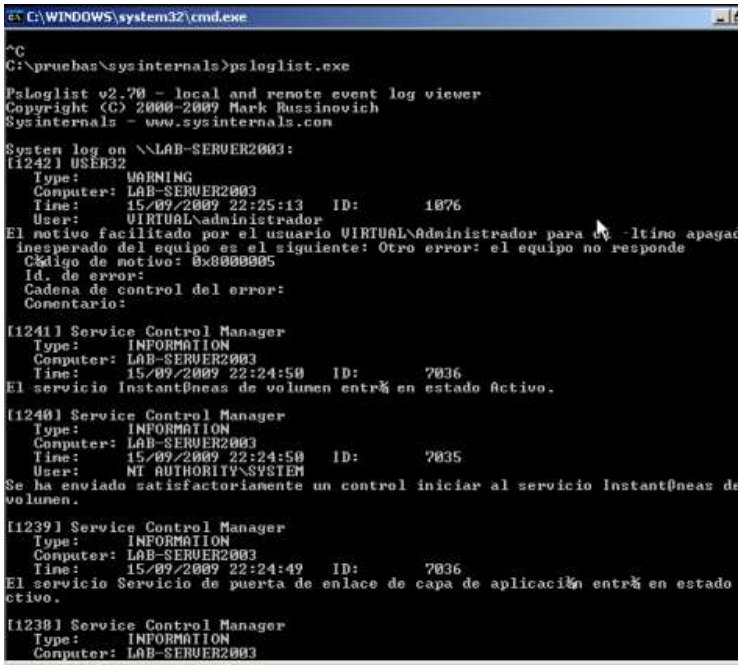



*Figure 3: PsLoglist output.*

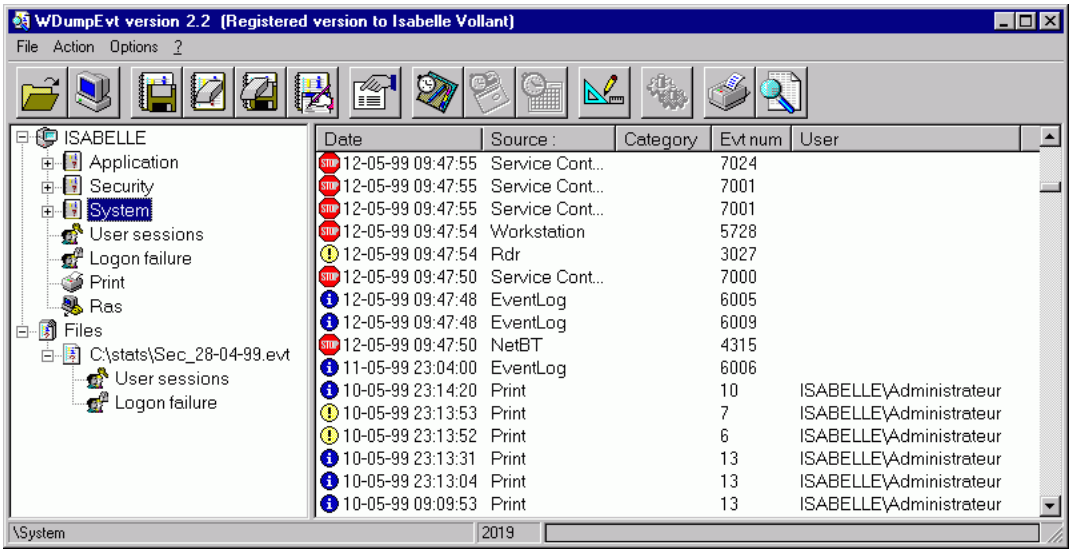



*Figure 4: WDumEvt window (showing system).*

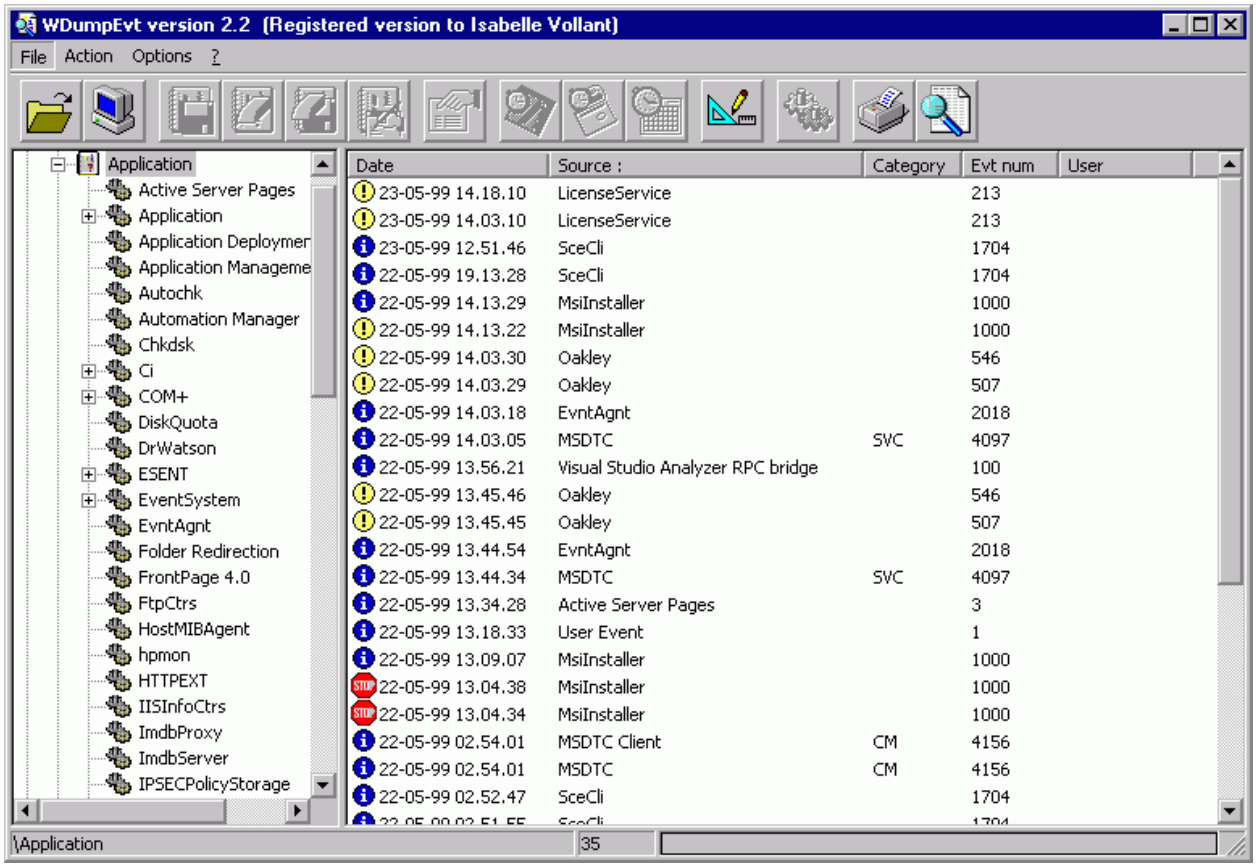



*Figure 5: WDumEvt window (showing apps).*

**VIDEO LECTURE**

## 8.4 SUMMARY

1. and collecting these information would be very vital.
2. Tools like *DevCon, Access Data tool kit, reg* and *regedit* helps in extracting non-volatile information in Windows.
3. Windows organises data using structures or elements like *Cluster, Partition, Master Boot Record, FAT32, New Technology File System.*
4. Files that are deleted, lost, cached or unallocated can be retrieved using various methods and tools.

## 8.5 CHECK YOUR PROGRESS

1. Fill in the blanks.

   a) _____ is the amount of on-disk file space from the end of the logical record information to the end of the physical disk record.
   b) _____ is the process of trying to recover files without a file system metadata.

2. State True or False.

   a) Registry information is an example of volatile information
   b) Group of sectors form a cluster.
   c) When a file is deleted, the file system removes the file logically i.e. it removes all the meta-data and stamps related to the file.

## 8.6 ANSWERS TO CHECK YOUR PROGRESS

1. Fill in the blanks.

   a) Slack space.
   b) File carving.

2. State True or False

   a) (F)
   b) (T)
   c) (T)

## 8.7 FURTHER READINGS


- Windows Forensic Analysis Toolkit, Third Edition: Advanced Analysis Techniques for Windows 73rd Edition, by Harlan Carvey.
- File system forensic analysis 1st edition, by Brian carrier
- http://www.gfi.com/blog/top-20-free-digital-forensic-investigation-tools-for-sysadmins/
- Linda Volonino, Reynaldo Anzaldua; Computer Forensics For Dummies, Wiley Publishing, Inc.
- Investigating Hard Disks, File and Operating Systems: EC-Council | Press

## 8.8 MODEL QUESTIONS

1. State the usage and forensic importance of PsLoggedon, Netsessions, logonsessions tools.
2. How the deleted and lost files are recovered in a windows system?
3. Describe the disk and file structure in a windows system.
4. What is a slack space, swap space and file carving?
5. How is registry information important in windows forensics?

### References, Article Source & Contributors


[1] Disk Sector, https://en.wikipedia.org/wiki/Disk_sector, retrieved Nov 2015

[2] DriveSpy, https://www.digitalintelligence.com/software/disoftware/drivespy/, retrieved Nov 2015

[3] File Carving, https://en.wikipedia.org/wiki/File_carving, retrieved Nov 2015

[4] Hard Disk Drive, https://en.wikipedia.org/wiki/Hard_disk_drive, retrieved Nov 2015

[5] Operating Systems, https://en.wikipedia.org/wiki/Operating_system, retrieved Nov 2015

[6] What is slack space, A Webopedia Definition, www.webopedia.com/TERM/S/slack_space


### Bibliography


[1] Windows System artefacts, http://resources.infosecinstitute.com/windows-systems-and-artifacts-in-digital-forensics-part-i-registry/, retrieved Nov 2015

[2] Tom Olzak, IT Security, http://www.techrepublic.com/blog/it-security/computer-forensics-finding-hidden-data/, May 21, 2007, retrieved Nov 2015.

# EXPERT PANEL

**Dr. Jeetendra Pande, Associate Professor- Computer Science, School of Computer Science & IT, Uttarakhand Open University, Haldwani**

**Dr. Ajay Prasad, Sr. Associate Professor, University of Petroleum and Energy Studies, Dehradun**

**Dr. Akashdeep Bharadwaj, Professor, University of Petroleum and Energy Studies, Dehradun**

**Mr. Sridhar Chandramohan Iyer, Assistant Professor- Universal College of Engineering, Kaman, Vasai, University of Mumbai**

**Mr. Rishikesh Ojha, Digital Forensics and eDiscovery Expert**

**Ms. Priyanka Tewari, IT Consultant**

**Mr. Ketan Joglekar, Assistant Professor, GJ College, Maharastra**

**Dr. Ashutosh Kumar Bhatt, Associate Professor, Uttarakhand Open University, Haldwani**

**Dr. Sangram Panigrahi, Assistant Professor, Siksha 'O' Anusandhan, Bhubaneswar**

This MOOC has been prepared with the support of