SWINBURNE UNIVERSITY OF TECHNOLOGY

ICT30010 eForensic Fundamentals

Lecture 6

File Systems and Pattern Matching

Troy Pretty

Digital Forensic Analyst

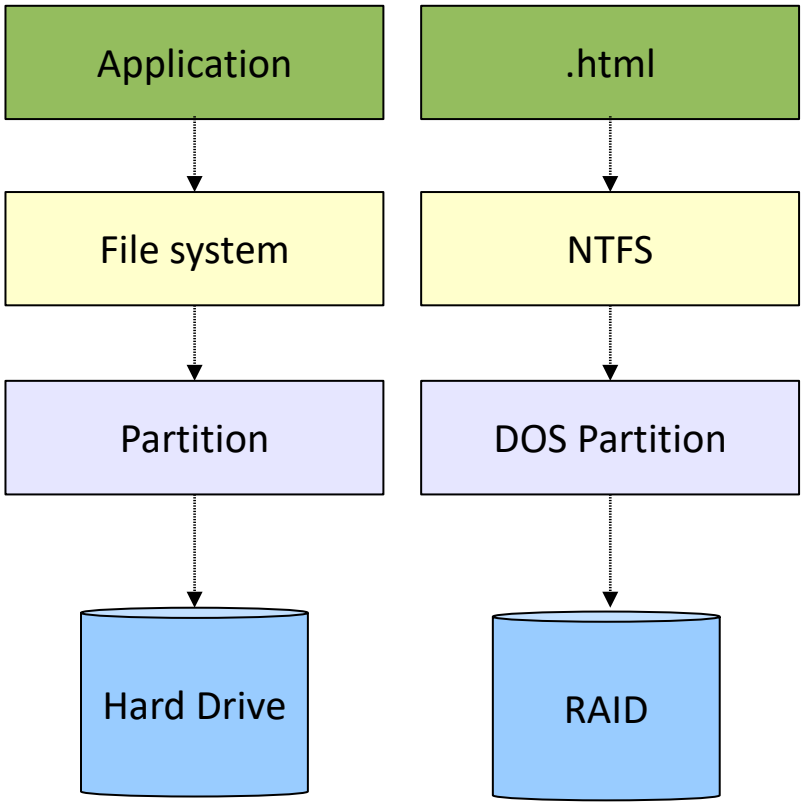
SWINBURNE UNIVERSITY OF TECHNOLOGY

Outline and Learning Goals

- General features of file systems
- FAT and NTFS
- Searching file systems
- File deletion vs. file erasing

File Systems

- A file system is a way of organising and store data on a disk according to well defined standards
- Imposes a structure on the partition scheme



Widely used file systems

- FAT
  - File Allocation Table
  - Simple and one of the earliest widely used file systems
  - Still widely used in flash memory devices such as thumb drives and digital cameras
  - Design goal is simplicity
  - Limits:
    - FAT12 – 16MB Volume
    - FAT16 – 4GB Volume
    - FAT32 – 2TB Volume, 4GB File Size

Widely used file systems

- exFAT
  - Extended File Allocation Table
  - Introduced around the time of Windows Vista
  - Designed to be lightweight like FAT but without the extra overhead of NTFS
  - Overcomes volume and file size limitations of FAT

Widely used file systems

- NTFS
  - New Technologies File System
  - Default system for Microsoft Windows since NT
  - One of the more complex file systems
  - Design goal is flexibility
    - Security
    - Compression
    - Encryption

Widely used file systems

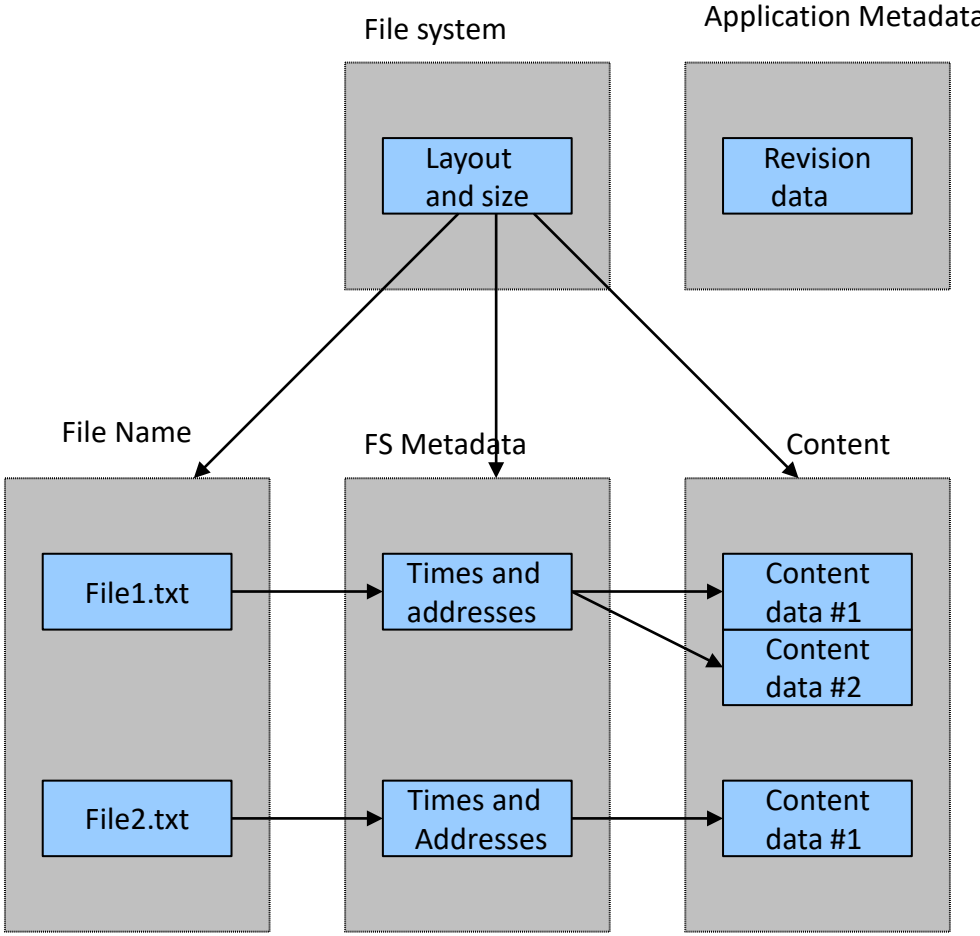
- EXT
  - Extended File System (EXT2, EXT3, EXT4)
  - Default file system for most Linux distributions
  - Based on UFS (but considerably simplified)
  - Design goals are speed and reliability

Widely used file systems

- HFS/HFS+
  - Hierarchical File System/Hierarchical File System Plus
  - Apple proprietary filesystem
  - Used in OSX and on Apple iDevices
  - Linus Torvalds the creator of the Linux kernel has been quoted to say:
    - “Quite frankly, HFS+ is probably the worst file-system ever”
- APFS
  - Apple File System
  - Replaces HFS+
  - Universal across all Apple devices
  - First released on iPhone running iOS 10.3 (March 2017)
  - Increase in speed and better disk utilisation

File systems

- The major file systems have some common features
  - File system information – A map of the file system
  - File name – Human readable label
  - FS Metadata – Data that describes files
  - Content – Data stored in file system
  - Application Metadata – Specific information depending on file type



File systems

- Analysis of a file system is usefully carried out according to the categories described in the previous slide
  - File system information, content, FS metadata, file name, application metadata
- Depending on what we are searching for, one particular category might be more appropriate as the basis of the search than the others

File system category

- A map of the file system
- Enables identification of the location of other important data
- Usually located in a standard data structure in the first few sectors of the system
- An analysis of a file system will usually start here

File system category

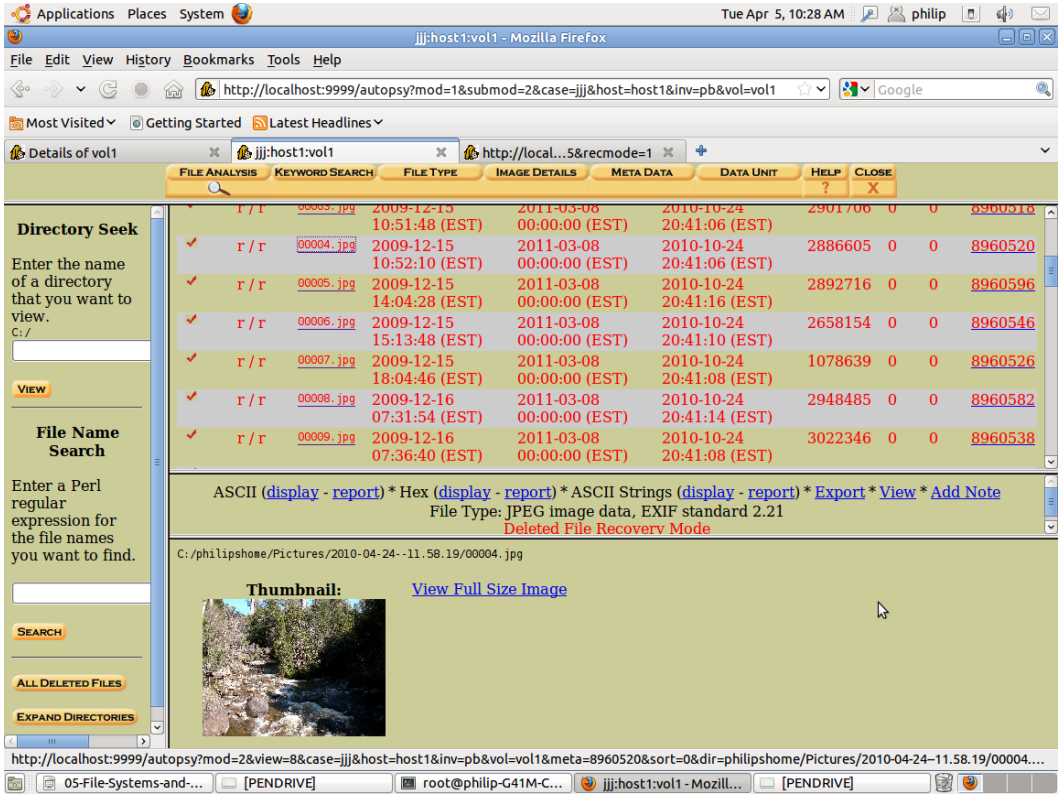
- File system is FAT32
- Volume name is PENDRIVE
- Boot sector is 0
- Data area is sector 7840 to 4004319
- Obtained using the fsstat command

```
philip@philip-G41M-Combo: ~/Desktop
File  Edit  View  Search  Terminal  Help
philip@philip-G41M-Combo:~/Desktop$ fsstat -f fat fat.dd | more
FILE SYSTEM INFORMATION
--------------------------------------------
File System Type: FAT32

OEM Name: MSDOS5.0
Volume ID: 0x6864dea4
Volume Label (Boot Sector): NO NAME
Volume Label (Root Directory): PENDRIVE
File System Type Label: FAT32
Next Free Sector (FS Info): 26000
Free Sector Count (FS Info): 2201232

Sectors before file system: 32

File System Layout (in sectors)
Total Range: 0 - 4004319
* Reserved: 0 - 33
** Boot Sector: 0
** FS Info Sector: 1
** Backup Boot Sector: 6
* FAT 0: 34 - 3936
* FAT 1: 3937 - 7839
* Data Area: 7840 - 4004319
** Cluster Area: 7840 - 4004319
*** Root Directory: 7840 - 7847

METADATA INFORMATION
--------------------------------------------
Range: 2 - 63943686
Root Directory: 2

CONTENT INFORMATION
--------------------------------------------
Sector Size: 512
Cluster Size: 4096
Total Cluster Range: 2 - 499561

FAT CONTENTS (in sectors)
--------------------------------------------
7840-7847 (8) -> EOF
8384-8831 (448) -> EOF
8832-9415 (584) -> EOF
9416-15431 (6016) -> EOF
```

SWINBURNE UNIVERSITY OF TECHNOLOGY

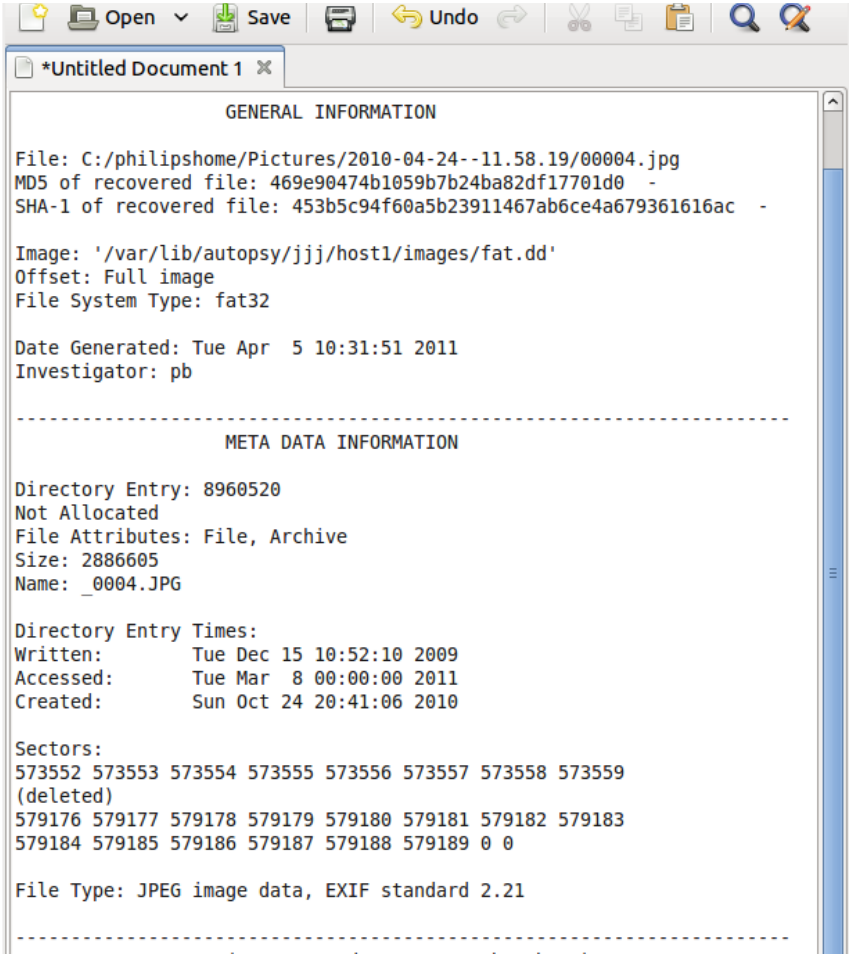
Content category

- File and directory contents
- Figure shows a deleted file from fat.dd obtained using Sleuth Kit / Autopsy



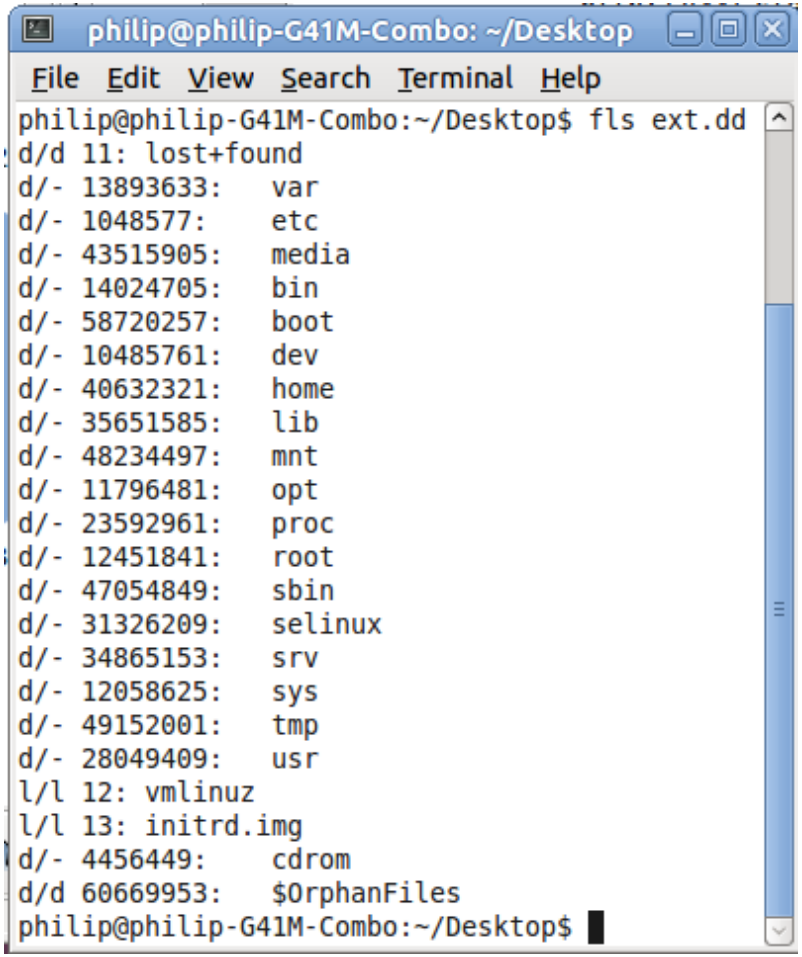
Metadata category

- Metadata from deleted file obtained using TSK/Autopsy
- Shows attributes of file _0004.jpg



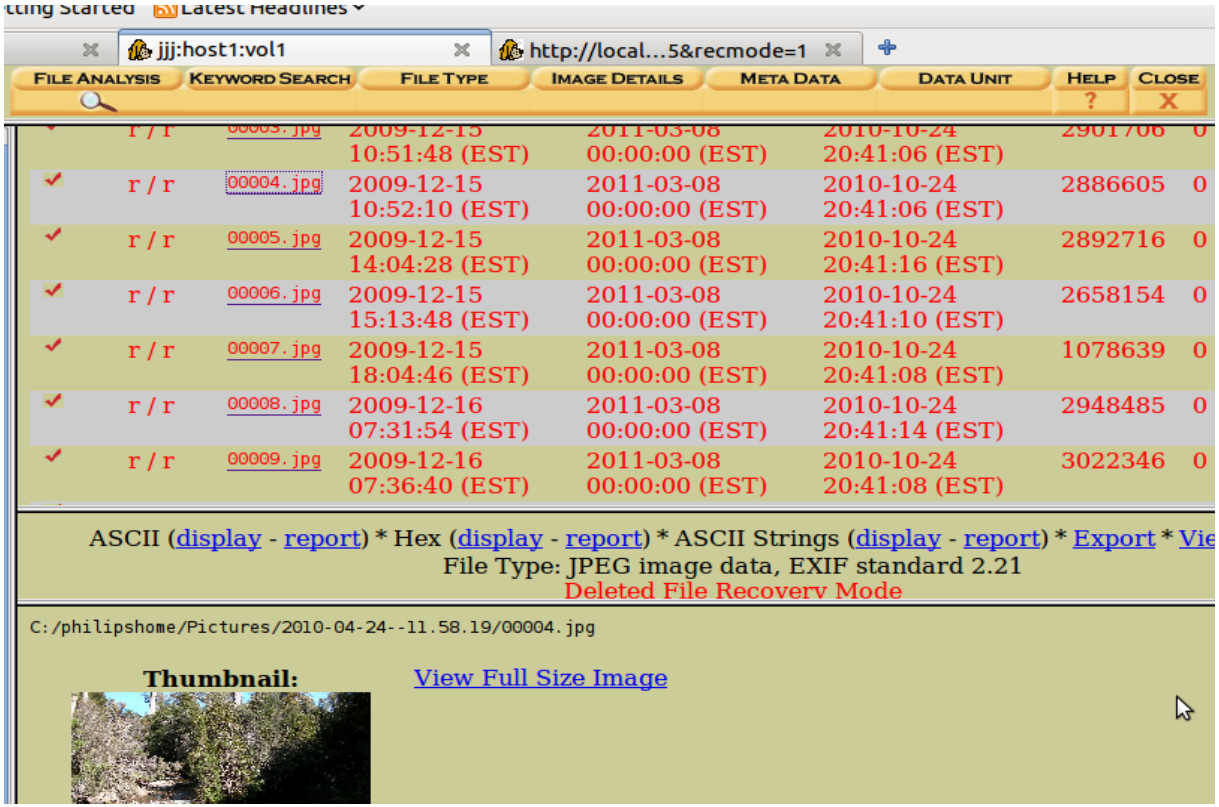
Filename Category

- Associates a name with a filesystem entry
- In example we see directory /var associated with address 13803633
  - An ext file system running Linux



Application category

- Interprets file at the application level



FAT concepts and analysis

- File Allocation Table (FAT) one of the simplest file systems in common use
  - Used in older versions of Windows but main use now is with flash memory systems such as thumb drives and digital camera cards
- Each file and directory is allocated a directory entry that contains the file name, size, starting address of the file (and additional metadata)
- File and directory content is stored in disk clusters

FAT system layout

- Reserved area
  - Boot sector
- FAT area
  - Location of File Allocation Tables
  - Links to root directory entries
  - May be two FAT tables for reliability

root@philip-G41M-Combo: /home/philip/Deskto

File Edit View Search Terminal Help

```
root@philip-G41M-Combo:/home/philip/Desktop# fsstat
FILE SYSTEM INFORMATION
--------------------------------------------
File System Type: FAT32

OEM Name: MSDOS5.0
Volume ID: 0x6864dea4
Volume Label (Boot Sector): NO NAME
Volume Label (Root Directory): PENDRIVE
File System Type Label: FAT32
Next Free Sector (FS Info): 26000
Free Sector Count (FS Info): 2201232

Sectors before file system: 32

File System Layout (in sectors)
Total Range: 0 - 4004319
* Reserved: 0 - 33
** Boot Sector: 0
** FS Info Sector: 1
** Backup Boot Sector: 6
* FAT 0: 34 - 3936
* FAT 1: 3937 - 7839
* Data Area: 7840 - 4004319
** Cluster Area: 7840 - 4004319
*** Root Directory: 7840 - 7847

METADATA INFORMATION
--------------------------------------------
Range: 2 - 63943686
Root Directory: 2

CONTENT INFORMATION
--------------------------------------------
Sector Size: 512
Cluster Size: 4096
Total Cluster Range: 2 - 499561

FAT CONTENTS (in sectors)
--------------------------------------------
7840-7847 (8) -> EOF
8384-8831 (448) -> EOF
```

FAT system layout

- FAT table consists of entries for each file
- Size of each entry specified by FAT type
  - FAT12 has 12 bits, FAT16 has 16, FAT32 32 bits
  - In FAT32 28 bits are used to address the starting cluster
- Root directory
  - Link to clusters in data area
  - In FAT12/16 preallocated
  - In FAT32 part of data structure (so no bounds on size)
- Data area

FAT system root directory

- Root directory contains
  - File name (up to 11 characters) in FFFFFFFF.XXX format
    - Long file names stored in separate entries
  - File attributes such as RW, hidden, label, whether it has a long file name, whether it is a directory
  - Time / date created/written/accessed
  - First cluster address of the file

Recovering deleted files in FAT systems

- It is often possible to recover deleted files in a FAT system
- When a file is deleted the system only makes a deletion mark on the file's directory entry
- Clusters in the DATA area are marked as empty in the directory BUT are not immediately overwritten
  - When writing new data the original file clusters might be overwritten
  - But often the old data is still available

FAT deleted file recovery in TSK/Autopsy

Most Visited | Getting Started | Latest Headlines

jjj:host1:vol1

FILE ANALYSIS | KEYWORD SEARCH | FILE TYPE | IMAGE DETAILS | META DATA | DATA UNIT | HELP | CLOSE

Directory Seek

Enter the name of a directory that you want to view.
C:/

VIEW

File Name Search

Enter a Perl regular expression for the file names you want to find.

SEARCH

Current Directory: C:/ /philipshome/ /Pictures/
/2010-04-24--11.58.19/

ADD NOTE | GENERATE MD5 LIST OF FILES

DEL	Type dir / in	NAME	WRITTEN	ACCESSED	CREATED	SIZE	UID	GID	META
✓	d / d	../	2010-04-26 12:31:32 (EST)	2010-10-24 00:00:00 (EST)	2010-10-24 20:41:16 (EST)	4096	0	0	942495
✓	d / d	./	2011-03-20 04:32:40 (EST)	2011-03-20 00:00:00 (EST)	2010-10-24 20:41:16 (EST)	4096	0	0	5855250
✓	r / r	00001.jpg	2009-12-04 13:29:56 (EST)	2011-03-08 00:00:00 (EST)	2010-10-24 20:41:08 (EST)	3162665	0	0	8960534
✓	r / r	00002.jpg	2009-12-04 13:36:00 (EST)	2011-03-08 00:00:00 (EST)	2010-10-24 20:41:14 (EST)	3063448	0	0	8960586
✓	r / r	00003.jpg	2009-12-15	2011-03-08	2010-10-24	2901706	0	0	8960518

ASCII (display - report) * Hex (display - report) * ASCII Strings (display - report) * Export * View * Add Note
File Type: JPEG image data, EXIF standard 2.21
Deleted File Recovery Mode

C:/philipshome/Pictures/2010-04-24--11.58.19/00004.jpg

Thumbnail:

View Full Size Image

NTFS concepts and analysis

- New Technologies File System
  - Microsoft designed and default for MS Windows since Win NT
- Designed for flexibility, scalability, reliability, security and large storage devices
- Key concept of NTFS is that everything in an NTFS system is a file
  - Basic file system administrative data is a file unlike in other systems where it is hidden
  - Entire file system is a data area
  - Only consistent layout is first sectors of the volume contain the boot sector and boot code

NTFS concepts and analysis

- NTFS is a **journaling** file system
  - Changes to the files in the system are entered in a transaction list (a journal) before carrying them out
  - Provides resilience to file system corruption since, in the event of a system failure, the journal entries can be replayed
  - Journal entries stored in attribute $LogFile
  - $LogFile can contain forensic artifacts!

Master File Table (MFT)

- Volume Boot Record in sector 0 (of partition)
- The Volume Boot Record (boot sector) has a pointer to the first record of the MFT
- MFT is the heart of NTFS
  - Contains information about all files and directories
- The MFT also has an entry in the MFT

Master File Table (MFT)

- Each file has (at least) one entry in the Master File Table (MFT)
- Each entry in the MFT contains a number of attributes
- Attributes are either resident or non-resident
  - If small enough stored in the MFT (resident) otherwise a pointer to the location (non-resident)

MFT Entries

- The size of each MFT entry is defined in the boot sector but all versions of Windows use 1024 bytes
- First 42 bytes contain 12 fields
- Remaining 982 bytes used for attributes
- If attribute cannot fit in 1024 bytes then the attribute is marked 'non-resident' and allocated to additional clusters
  - One attribute is the file's contents
    - Usually won't fit in the 1024 bytes

Standard attribute types

- Seventeen standard attributes
- Important ones
  - $STANDARD_INFORMATION
    - Flags, dates and times last accessed, owner
  - $FILE_NAME
    - File name in unicode
  - $DATA
    - File content

Deleting and erasing files

- Regardless of the file system it is surprisingly difficult to completely remove files from a disk
- Most 'delete' operations merely flag the entry in the file table (or equivalent) signifying that the area is available for reuse
- SSD Drives add complexity (Trim)
- Study at MIT in 2003 looked at 158 used disk drives purchased from eBay
  - http://web.mit.edu/newsoffice/2003/diskdrives.html
  - Found more than 5,000 credit card numbers, detailed personal and corporate financial records, numerous medical records, gigabytes of personal email and pornography.
  - Only 12 were properly sanitized

Deleting and erasing files

- Deleting and erasing quite different
  - Deleted does not guarantee erasure of the data
- To erase data need to use a utility such as scrub (in Unix systems)
  - Writes a random pattern of 0s and 1s to the disk
  - Does it several times, with newer hard drives 1 wipe is sufficient
  - An interesting paper on erasing data from solid state devices at http://www.usenix.org/events/fast11/tech/full_papers/Wei.pdf
- Zero Only
  - Single pass across drive writing zero values across each sector
- US Airforce standard 5020
  - First fills the drive with binary zeros, then binary ones, and finally an arbitrary character.

Deleting and erasing files

- Australia Signals Directorate (ASD) requires physical destruction for Secret+ data
- ASD standards require 3mm or 6mm pieces depending on sensitivity (e.g. Top Secret)

Hard Drive Destruction

SWINBURNE UNIVERSITY OF TECHNOLOGY

Deleting and Erasing Files for the Paranoid

Image Source: http://eecue.com/c/driveslag

Pattern Matching

- Simplest level is matching particular character strings
  - Already seen this in lab 2
  - Made use of piping of text into `grep`
- Unix has a rich set of pattern matching operations
  - Regular expressions
  - Lots of wild cards and options

- `xxd image.dd | grep secret`
- `strings --encoding=l --radix=d image.dd | grep "\.jpg$"`
- `cat inputfile | grep phil*`

Pattern matching – regular expressions

Meta character	Description
.	Matches any single character. Eg pic.hols Will match pic1hols, pic2hols and pic3hols
[ ]	Matches a single character contained within the brackets pic[12]hols will match pic1hols, pic2hols but not pic3hols
[^]	Matches a single character that is not contained within the brackets pic[^12]hols will match pic3hols but not pic1hols or pic2hols
$	Matches the last character of the string Eg pic1hol$ will match pic1hols
*	Matches preceding element 0 or more times eg ab*c will match ac, abc, abbbc

Pattern Matching

- Pattern matching not just restricted to alpha-numeric strings
  - Image matching an important area
  - Facial pattern matching
- Not necessarily matching the same image but similar images
  - Usually do not match pixel for pixel but match some attributes of the image
  - Example, facial recognition matches relative position of ears, nose, eyes, chin and mouth
    - Makes use of mathematical transforms such as Fourier, Discrete Cosine , Wavelets and similar

Summary

- Major file systems
- Detailed look at FAT and NTFS
- A little on pattern matching