7. Identification

Digital Forensics and Cybercrime course

Prof. Zanero

Analysis or identification toolset

- Operating system
  - Linux
    - extensive native file system support
    - Native support of hot swapping drives and devices, mounting images, etc.
  - Virtualization:
    - A set of Windows machines with different versions, networked with the Linux host and using Samba to share drives

Why not Windows?

- Windows MUST be confined because:
  - It tampers with drives and modifies evidence
  - No image handling or hotswapping of drives
  - No support for non-Windows FS
- Using Linux as host, and Windows as guest, we can:
  - Work the images with Linux, mounting them read-only and then exporting them via Samba to Windows
  - Use specific Windows tools
- Not always doable to use Samba: if Windows must see the file system (e.g. file recovery tool or unallocated space analysis) we can mount the image as a read-only loop device under Linux, and/or use the “non-persistent” mode of VMWare

Scientific means...

- Repeatable
  - Any other expert will be able to perform the same experiment, on a clone of the image, obtaining the same results I obtained
- The experiment:
  - Not just a tool input and output, but also the logic!
  - Result validation, the “expert” must be able to perform the same analysis by hand (at least in theory)
- This means, to me
  - That analysis software needs to be open sourced, and possibly free
  - That proprietary or “law enforcement only” tools are not really fit for the job

Analysis means… everything?

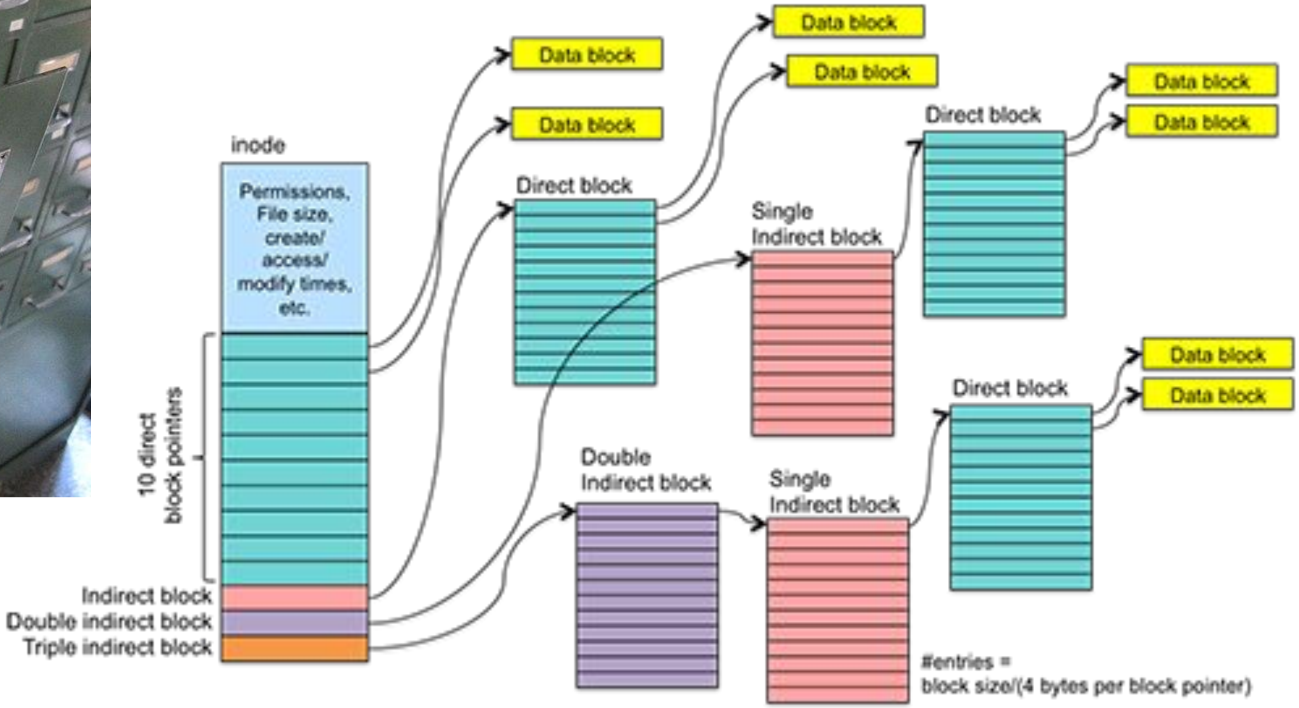
- During the analysis phase, we may need to apply a number of methodologies from computer science!
    - Opening files with appropriate viewers, or building some
    - Extracting, analyzing and mining data
    - Analyzing source code or object code
    - …
- #import <everything_else_you_learned.h>
- **In the following we will focus on tasks that happen only, or mostly, in forensics**

Recovery of deleted data

A typical challenge

- In many cases, information or data of interest has been (voluntarily or involuntarily) deleted
    - File deletion
    - Formatting or repartitioning of drives
    - Damaged drives/bad blocks
- One of the most typical tasks of computer forensics is the retrieval (complete or partial) of such deleted data
- To understand it, we need to recall basic elements on data storage by OSs

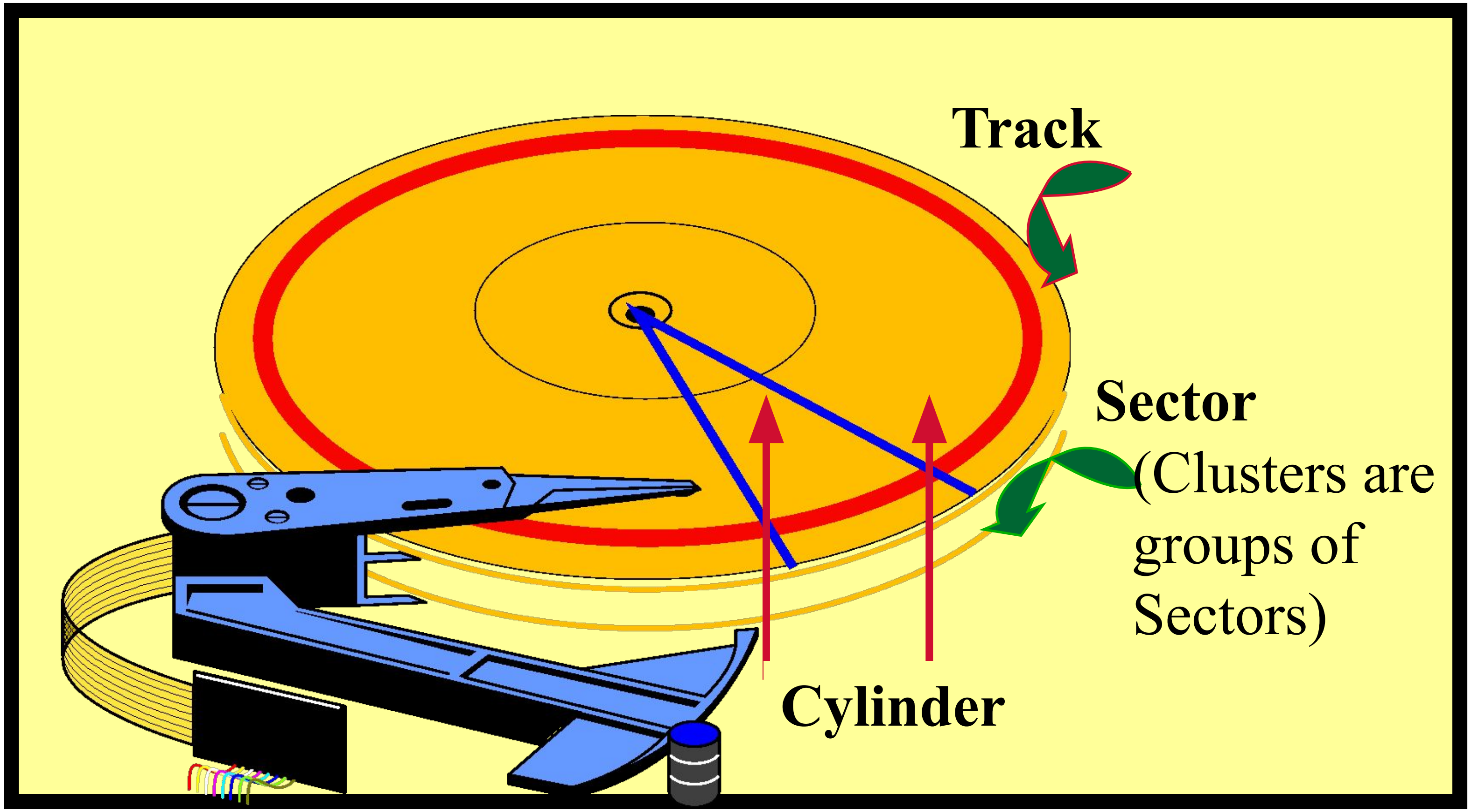
File system (UNIX)



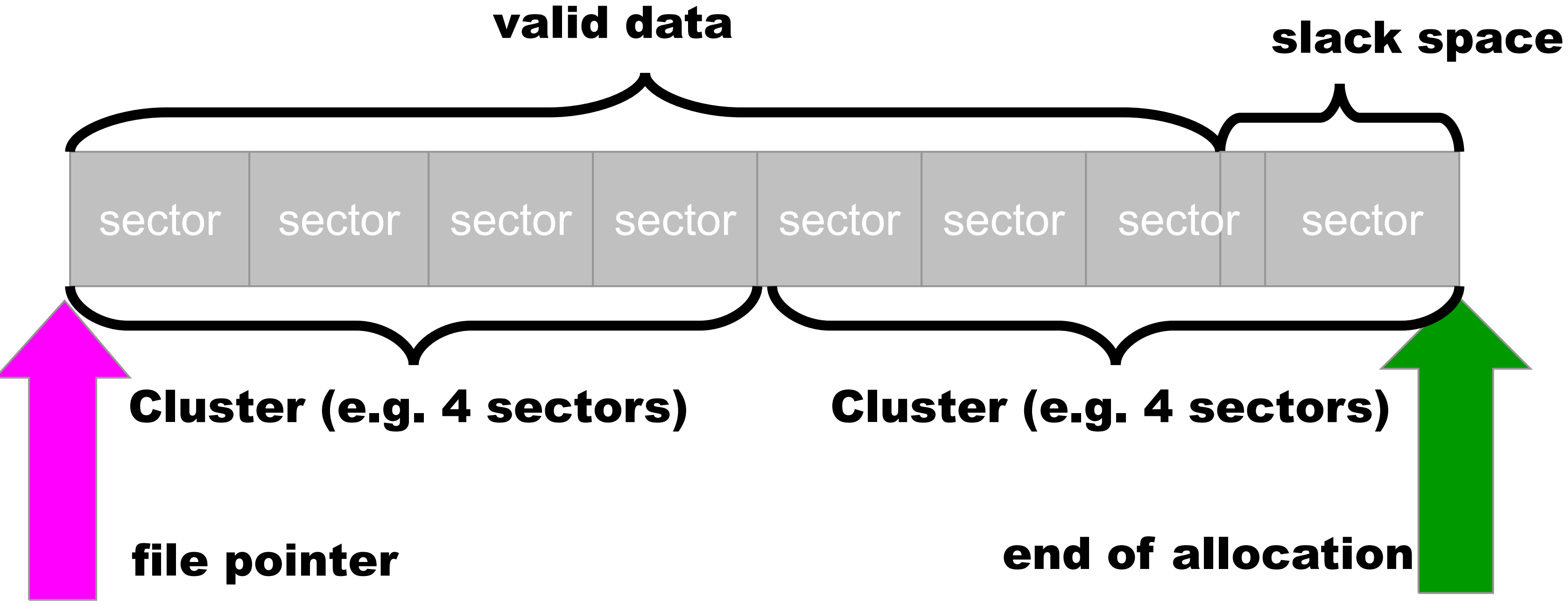
What happens on file deletion?

- OS is “lazy” and optimizations cause data persistence and locality
- When we delete a file
  - First, the file entry in the FS is flagged as deleted
    - Until here, it can be “undeleted” by simply removing the flag
  - Then, at some random time, the following two things will eventually happen, not in a particular order:
    - The FS entry will be removed (when FS structure is rewritten or rebalanced)
      - Until this happens we can find metadata on the file
    - The actual blocks (once) allocated to the file will be overwritten with other content
      - Until this happens, we can retrieve the actual blocks on disk

Disk geometry

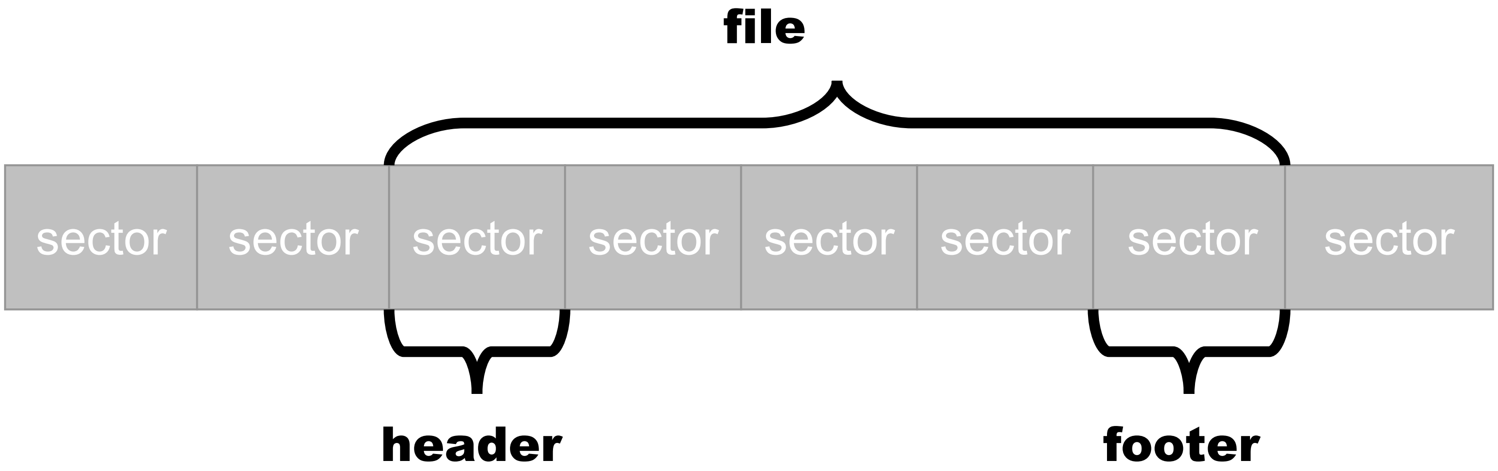


Sector, clusters and slack space



Fragments of deleted data accrete in slack space

File recovery through carving



- We scan the entire drive as a single bitstream
- We locate headers & footers of interesting filetypes
  - Anything in between, if not too large, is a candidate file
  - Techniques to determine filetype from content exist
- Issues:
  - Fragmentation (but on modern large drives this is not common, if fragmented mostly 2-fragmented)
  - (headerless) encryption and compression

Free software tools for data recovery

- TSK & Autopsy – Data recovery under linux: analyzes DD images, supports NTFS, FAT, FFS, EXT2, EXT3..., recovers deleted files, creates timelines, etc...

  http://www.sleuthkit.org/
  www.autopsy.com

- Foremost – file recovery through file carving

  http://foremost.sourceforge.net/

- gpart, testdisk: partition recovery

- photorec (self-explaining)

Antiforensic techniques

Anti-forensics definition

- Techniques that aim to create confusion in the analyst, to lead them off track, or to defeat tools and techniques used by analysts
  - Transient anti forensics: can be defeated if detected
  - Definitive anti forensics: destroying evidence, or making it impossible to acquire, unreliable or tampered
- Some techniques are sci-fi, but many are simple and effective

https://core.ac.uk/download/pdf/36736409.pdf

Critical failure points

- Which are the technology-dependent phases?
  - Acquisition (usage of tools for repeatable cloning and custody)
  - Identification (usage of tools for analysis of file systems, data reconstruction and carving)
- Interfering, we can compromise the process
  - **Transient** anti forensics if we interfere with **identification**
  - **Definitive** anti forensics if we interfere with **acquisition**

Timeline tampering (definitive)

- As we saw, analysis tools can display a timeline based on MAC(E) values: Modified, Accessed, Changed, (Entry Changed: check value on NTFS)
- We can therefore modify events by making them appear separated, or close, randomizing them or moving them completely out of scope
- Tool: “timestomp” (MACE) o “touch” (MAC)
- Once destroyed or modified, such data cannot be retrieved; modification not visible per se

Countering file recovery (definitive)

- File recovery uses data remnants
  - Secure deletion (heide, sdelete, ...)
    - Some secure deletion utilities are fake
  - Wiping unallocated space
  - Encryption
  - (Virtual machine usage)
- Note: reading “residuals of magnetization”, a la Gutmann, are science fiction: overwritten means gone
  https://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html#Epilogue

Fileless attacks (definitive)

People love machines in 2029A.D.
? Who slips into my robot body and whispers to my ghost?"
ST IN THE SHELL 2.0
0:25 / 0:50

- What if the traces are **not on the disk at all**?
- e.g.: Metasploit's meterpreter (or Mosdef, or IMPACT)
  - Injected in a process memory space
  - Gives attacker control
  - Doesn't write anything to disk
  - Can add thread, execute...
- So...
  - When the machine is shut down, evidence is lost!
  - ... and what is the first or second step of the regular S.O.P. when a machine is compromised?
  - Only hope: in-memory forensics; e.g. memdump, volatility

Filesystem Insertion and Subversion Technologies (transient)

- Don’t google for the acronym…
  https://www.blackhat.com/presentations/bh-asia-03/bh-asia-03-grugq/bh-asia-03-grugq.pdf
- We place data where there's no reason to look for them, in particular inside filesystem metadata
  - fsck is our enemy as it may “repair” metadata and trash our insertions
  - Inside a partition table there is space for ~32 KB of data
  - In EXT(2/3):
    - RuneFS: writing in bad block inodes (unlimited space)
    - WaffenFS: adds a fake EXT3 journal in an EXT2 partition (up to 32 MB storage)
    - KY FS: uses directory inodes (unlimited space)
    - Data Mule FS: puts data in padding and metadata structures of FS ignored by forensic tools (up to 1MB of space on a typical FS)

Log analysis (~transient)

- Typically you don't analyze logs by hand
- You typically use regular expressions or scripts
- If attackers can inject stuff in the logs (very likely), they can try to make your scripts fail, or even to exploit them

https://owasp.org/www-community/attacks/Log_Injection

Partition table tricks (transient)

- Partitions not correctly aligned
  - Using a partition restore tool we can read them, but they may escape a forensic analyst
- Adding multiple extended partitions
  - Windows and Linux manage them, many forensic tools don't
- Generate a high number n of logical partitions in an extended
  - With n high enough tools die