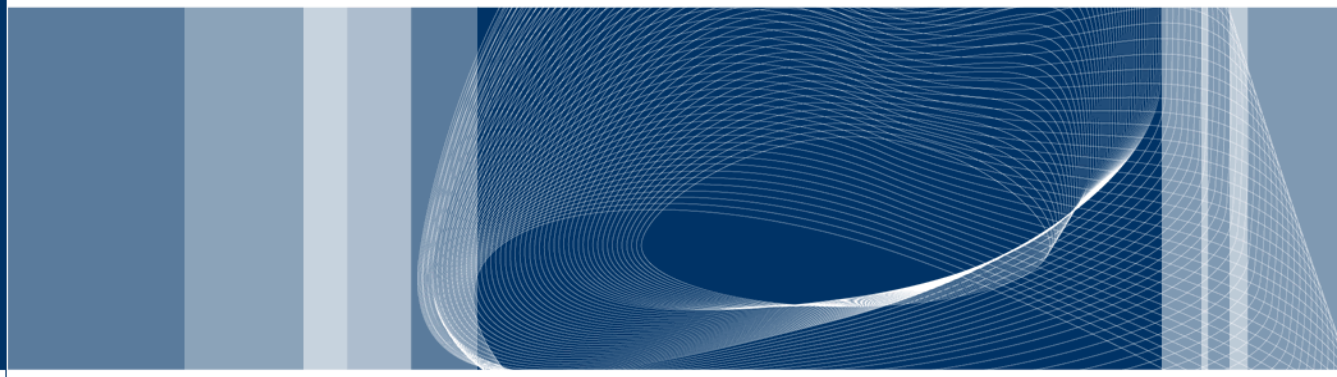
POLITECNICO DI MILANO



Forensics and Antiforensics 101

Ing. Stefano Zanero, PhD
Dip. Elettronica e Informazione – Politecnico di Milano

What does “Computer Forensics” mean?

- **Forensics** is the application of **scientific** analysis methods to reconstruct **evidence**
- **Computer (or Digital) Forensics** is the application of scientific analysis methods to digital data, computer systems and network data to reconstruct evidence
- **Scientific = Repeatable** (Galileo, circa 1650)
  - Beware: in Italian law, “repeatability” has a different meaning!
- **Scientific = Falsifiable** (Popper, 1934)
- Evidence: in the Italian legal framework, “evidence” is recognized as such in a court of law, so beware of the term

- IANAL, and neither are you. We are not here to discuss computer law.

Example of forensic engagements

Situation

- Internal investigations (inside an organization)
- Criminal investigations (defense or prosecution)
- Post-mortem of a system to assess damage / define recovery strategy
- Research (honeypot, etc)

Crimes and events

- Child pornography
- Fraud
- Cyber extortion / threats
- Espionage
- Copyright infringements
- Policy violations

4 phases of an investigation

- Source acquisition
- Evidence identification
- Evaluation
- Presentation

- Special Agent Mark M. Pollitt (FBI), “Computer Forensics: An Approach to Evidence in Cyberspace”
  http://www.rcmp-grc.gc.ca/tsb/pubs/bulletins/bull41_3.htm

Acquisition

- Key difference with the USA – beware, forensic procedures have been developed with the USA in mind
- Evidence in USA: “chain of custody”, and admissibility
- In Italy evidence is based on the evaluation performed by the judge
- Law 48/2008 (convention of Budapest on cybercrime) introduced at last in the Italian law sound computer forensic requirements
- In ancient times... (1994, Italian Crackdown)
- … but even in modern times …

- *I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. I watched C-beams glitter in the dark near the Tannhauser gate. All those moments will be lost in time... like tears in rain...*

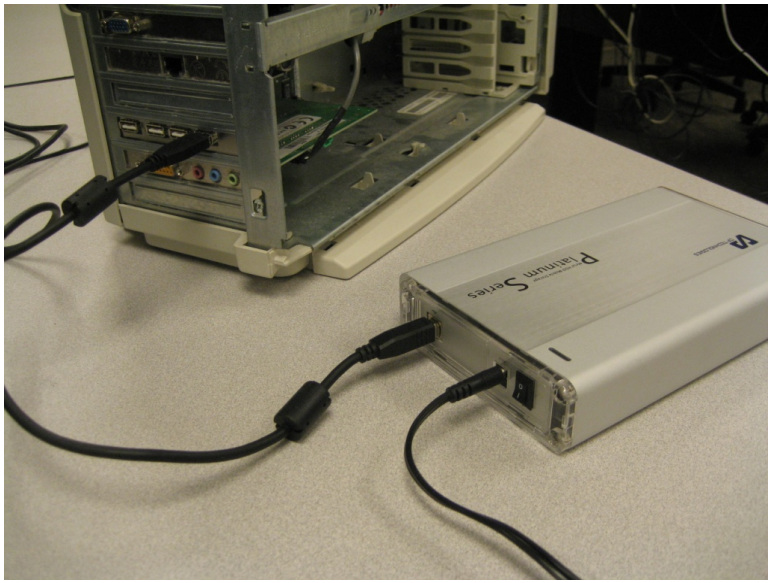
Repeatability problem

- In Italian law, a repeatable analysis (accertamento ripetibile) is one that does not cause an irreversible alteration of the object
  - E.g. of non repeatable analysis which is routinely performed: an autopsy; chemical analyses which require reagents; etc.
- Non repeatable analyses have a procedure which is slightly more complex
- Starting a computer or using it **alterates the evidence**
  - E.g. timestamps of files
- Digital evidence is *brittle*: if modified, there is no way to tell. I can theoretically create a perfect *fake*
- In order to seal the evidence, hashes and digital signatures are routinely performed. If the hash is recorded, and constantly checked, it can ensure on the identity, authenticity and non-tampered state of the evidence

Standard operating procedures

- We want to outline some best practices for handling computer evidence
- If the machine is shut down, or the media are disconnected, good practice is to perform a forensic copy as soon as technically possible
  - Connect the media, if possible with a *write blocker*
  - Compute the hash of the source
  - Copy
  - Compute and compare the hashes of the source and the clone(s)
- Further clones can and should be obtained, as working copies
- It would be good to compute both MD5 and SHA-1 hashes, both for redundancy and security
- All can be performed with open source software under the Linux or BSD operating systems (dd, md5sum, sha1sum)

Write blocker



+ external USB drive

Some useful commands to remember

- dd: bit per bit copy
  - dd if=/dev/hda of=immagine ...
- netcat (nc): network send
  - nc -l -p 5678 > file-dest
  - cat file-src | nc -p 5678
- md5sum/sha1sum
  - Checksum and hashing

Computer Forensics live

- Sometimes we need to work directly on the machine:
  - Laptop with weird hw and controllers (Toshiba, anyone?)
  - Peculiar hardware
  - Raid devices
  - Specific investigation constraints
- In this case we can use a live Linux distribution targeted to forensic analysis (NOT ANY LIVE)
- Examples that work:
  - Helix, http://www.e-fense.com/helix/
  - DEFT, http://www.deft-linux.it

What if the machine is turned on

- Can we turn it off? (hint: critical services?)
- Should we turn it off? (hint: live analysis of an intruder?)
- Network disconnect (to eject the intruder, if still connected)
- Work in volatility order
  - Dump of memory: if possible, and not costly; hardware tricks to perform the dump are available
  - Save runtime information: network, process information, etc.
  - Finally, disk acquisition
- It could be possible to perform the acquisition without a shutdown; if impossible, pull the plug (do not perform the shutdown procedure, unless it is really necessary to ensure the reboot of the machine)
- Document all activities executed before sealing the evidence

Some useful commands

- Network data
  - ifconfig -a ; netstat -anp ; route -n ; arp
- Process data
  - ps aux ; Lsof *file*
- Users data
  - who; last; lastlog

New challenges: memory cards and co.

- Memory Card
  - Small hard drives
  - Can be partitioned, reformatted... encrypted...
  - Can be hidden (just think of a microSD...)
- MP3 readers:
  - Hard drives interfaced with proprietary OS and interfaces
  - How to extract the drive w/o breaking the device
  - Proprietary file-systems?
- Even more so, problems with PDAs and smartphones

Analysis or identification

- Hardware:
  - Removable HD enclosures or connectors with different plugs
  - Write blockers
  - A DVD burner
  - External disks
  - USB2, firewire, SATA and e-SATA controllers, if possible
- Operating system:
  - Linux: extensive native file system support
  - Virtualization:
    - A set of Windows machines (2000, XP, Vista, 7)
    - At least a Freedos machine

    Networked with the host and sharing disks with samba. Wonder why?

Windows caged

- Windows MUST be confined because:
  - They tamper with the drives and modify evidence
  - They cannot handle images or hotswapping of drives
  - They do not handle properly any 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

Live network analysis

- In some cases we will want to observe an attacker “live”
  - Honeypots, e.g.
  - An intruder can react if he feels observed
  - Reminder: tools installed on a compromised machine may be unreliable (e.g. rootkit)
- Key observation points:
  - Logs
  - Network traffic
- Some tools
  - tcpdump
  - wireshark

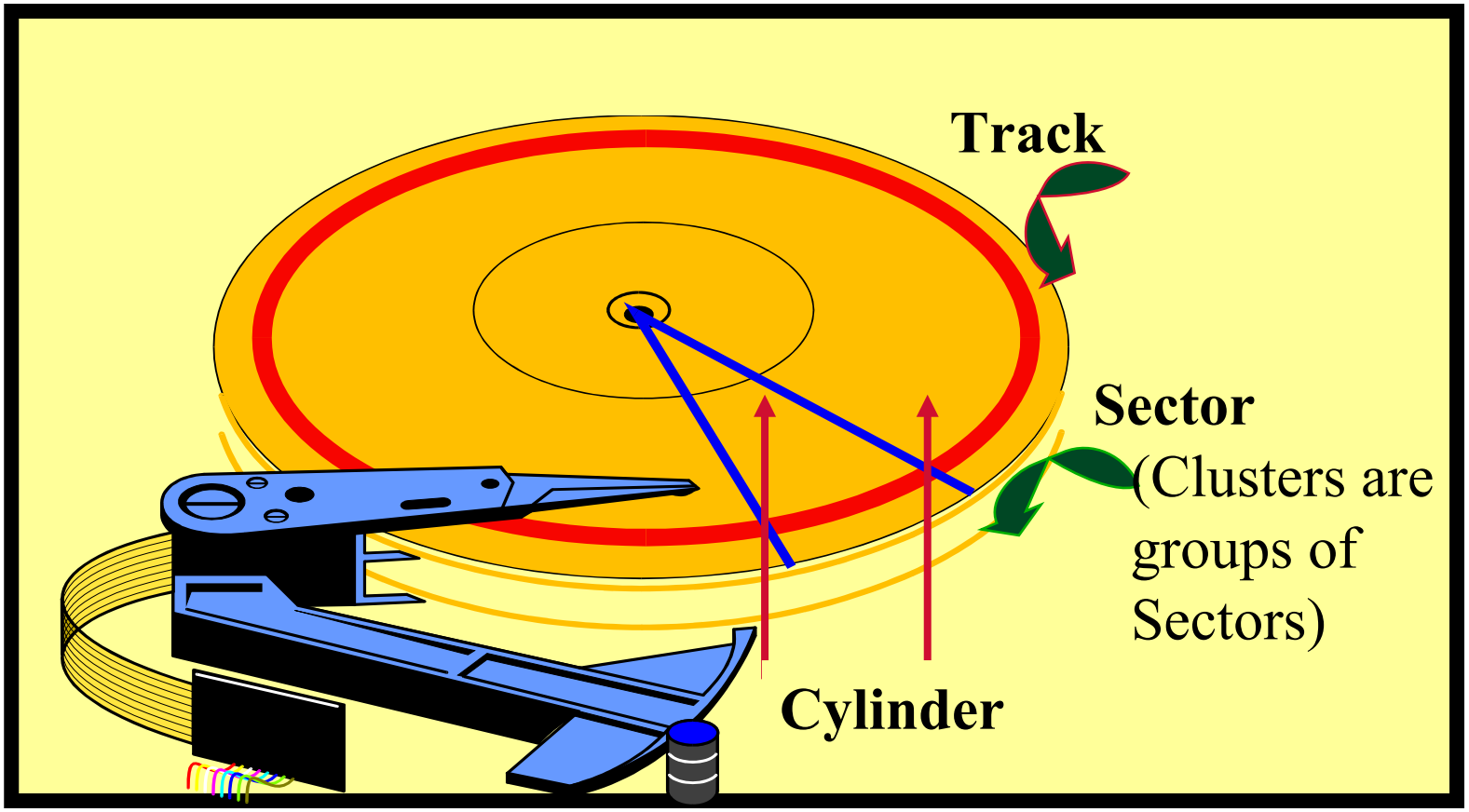
Typical analysis task: reading the ashes...

- Deleted file recovery
- Slack Space analysis
- Access to bad blocks
- Recovery of formatted/destroyed partitions
- Recovery of damaged drives

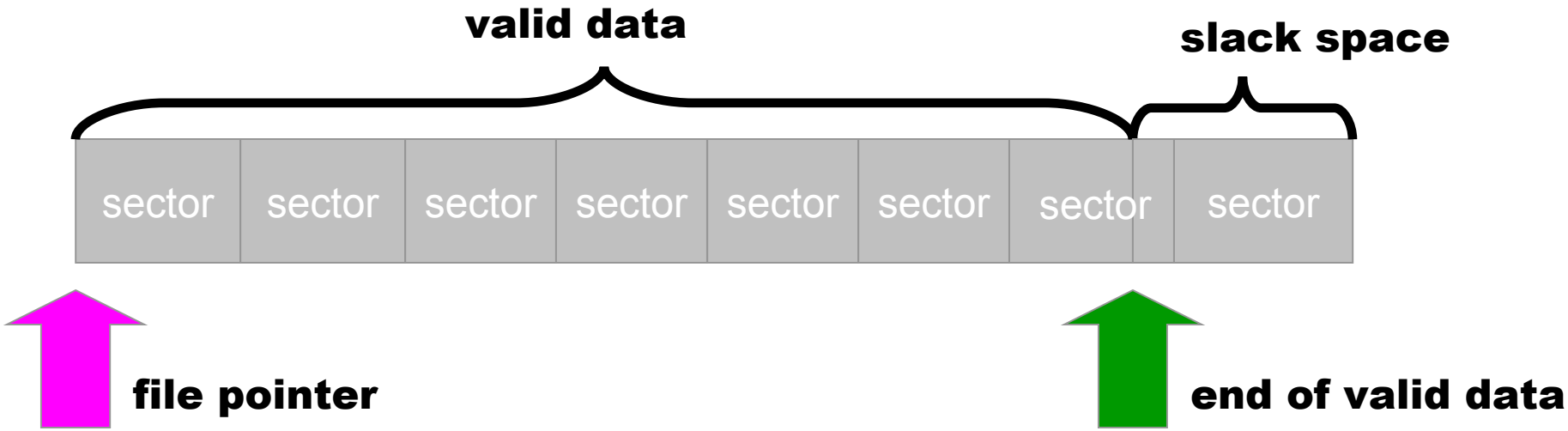
- **Black magic?**

- No, simple application of the data persistence and locality properties caused by OS optimizations
- When the OS deallocates a file (or a memory area) it doesn't actually remove contents
- Using specific tools, we can recover deleted files, and sometimes even their metadata!

Disk geometry



Sector, clusters and slack space



Fragments of deleted data accrete in slack space

Free Tools

- 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/
- Foremost – file recovery through file carving

  http://foremost.sourceforge.net/
- gpart, testdisk: partition recovery
- Active Uneraser: using DOS, analyze FAT, NTFS, searches the slack space, commercial but cheap :-)

Attacker tool analysis

- What binaries were installed ?
  - Chkrootkit and autopsy can help, but in general it may be difficult
- How were they compiled and executed ?
  - Xferlog and other sources, timeline of autopsy...
- Used languages ? Scripts ?
- Can you find this stuff on google, or was it custom built?
- What do they do? We cannot execute them
- Rough analysis: strings, file, nm
- Use a sandbox, e.g. anubis (anubis.iseclab.org)
- Reverse-engineering and decompilation (last resort)

Evaluation

- Understanding if the evidence supports the legal position, and how
- Lawyers, prosecutors and investigators are seldom computer experts (and let's not discuss this)
- Good questions make for better answers
- It may be difficult to find a common language

Presentation

- 90% of the stuff you did or found will be of little relevance
- You will need to present the remaining 10%
- Ethical and legal issues in interrogation
- What you will present will be challenged:
  - Your deductions on the meaning of the evidence
  - The evidence itself
  - The method you applied to gather evidence
  - The chain of custody
  - The acquisition
- Or... you might be the challenger as opposed to the challenged!

Recap

- Forensic analysts wish to reconstruct “what has happened”
- Reconstruction must hold up to scrutiny in court
- Phaese
  - Acquisition
  - Identification
  - Evaluation
  - Presentation

Critical 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 antiforensics: if we interfere with identification, in a way which can be defeated if detected
  - Definitive antiforensics: if we interfere with acquisition, by making evidence impossible to acquire, unreliable or tampered

Anti-forensics definition

- Techniques that aim to create confusion in the analyst, to lead him off track, or to defeat tools and techniques used by analysts
- Some are sci-fi, others are simple and effective
- Targets:
  - Timestamps
  - Log analysis
  - File recovery and carving
  - File and executable identification
  - Steganography and data hiding

Timeline...

- 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)
- You can bet your money that even costly tools such as EnCase cannot do much against this.

Log analysis

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

Deleted file recovery

- If forensics = reading the ashes, let's throw the ashes to the wind
  - Secure deletion (heide, sysinternals sdelete, etc)
  - Wiping unallocated space
  - Encryption
    - Note: some secure delete utilities are fake, be advised...
- Note: reading “residuals of magnetization”, a la Gutmann, are science fiction: overwritten means gone.

FISTing (cough...)

- Filesystem Insertion and Subversion Technologies
- We place data where there's no reason to look for them, in particular inside FS metadata
  - fsck is our enemy as it may “repair” metadata and trash our insertions
  - Inside partition table I can hide 32 KB of data
  - In EXT(2/3) I can do:
    - 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)

Partition table fun

- 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 n logical partitions in an extended
  - With n high enough tools die

Carving and filetype searches

- Most tools use two base methods for filetype detection
  - Extensions (oh, yeah !)
  - Signature on header&footer (not much better)
- ... couple of bash lines, and no more child porn images will be retrieved from a media
- Solution: using more flexible and advanced way to detect files (under research)

Ghost in the shell

- What if the traces are not on the disk?
- Example: Metasploit's meterpreted (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; Windows Memory Forensics Tool (M. Burdach) or memdump