

universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Course Presentation

Artur Varanda

School Year 2021-2022

- I. Context
- II. Objectives
- III. Syllabus
- IV. Evaluation
- V. Resources
- VI. Bibliography

Computer Systems Forensic Analysis:

Optional – 1st year, 1st Semester – 39 contact hours

Lecturer:

Artur Varanda (artur.varanda@sapo.pt)

Office hours:

send an email first to schedule a meeting (VTC)

This class aims to provide students with sound knowledge of digital forensics such as

- ✓ the collection, identification, preservation, documentation, analysis and presentation of digital evidence;
- ✓ digital evidence acquired from computers, cell phones and other electronic devices;
- ✓ this knowledge will be taught in the various areas of forensic discipline, forensic computing and forensic data analysis.

This course aims to address the transversal concepts to all areas of digital forensics such as:

- ✓ the scientific method of digital forensic investigation;
- ✓ the different types of digital forensic evidences: data, computers, mobile devices, ...
- ✓ the students will apply the knowledge acquired in the classroom to several laboratory assignments and will be able to produce a digital forensic report

Upon completion of this course, students should be able to:

- ✓ identify the different types of digital forensic evidence
- ✓ know the terminology, techniques and processes of a digital forensic investigation
- ✓ collect digital evidence from storage media
- ✓ know the limitations of digital forensics current techniques
- ✓ understand the scientific method and the need for its use
- ✓ apply the scientific method in a digital forensics investigation
- ✓ use some forensic tools and techniques
- ✓ comprehend forensic analysis reports

1 - Overview of cybercrime investigation

- ✓ Information security principles
- ✓ AAA Services concept
- ✓ Cybercrime vs Computer Crime
- ✓ Penal framework of cybercrime
- ✓ Applicable legislation

2 - Introduction to digital forensics

- ✓ Digital investigation
- ✓ Digital evidence
- ✓ Investigation process
- ✓ Digital evidence handling
- ✓ Ethical code

3 - Obtaining evidences

- ✓ Boot process
- ✓ Forensic boot tools
- ✓ Forensic sorting tools
- ✓ Forensic acquisition tools
- ✓ *FTK Imager* overview

4 - Data organization

- ✓ Data storage devices
- ✓ File system analysis
- ✓ Binary and hexadecimal numbers
- ✓ Endianess
- ✓ Character encoding
- ✓ Data structures

5 -Autopsy

- ✓ *Autopsy* workflow
- ✓ Create cases and add data sources
- ✓ Automated processing with ingest modules
- ✓ Manual content analysis
- ✓ Report generation

6 – Storage devices

- ✓ Hard disk geometry
- ✓ ATA and SCSI interfaces
- ✓ Flash memory drives
- ✓ Solid State Drives (SSD)

7 – Volumes and partitions

- ✓ Partition tables
- ✓ Logical addresses
- ✓ Volume analysis
- ✓ Common partitions
- ✓ Volume partition tools

8 - RAM Analysis

- ✓ General computer architecture
- ✓ Memory acquisition tools
- ✓ Memory analysis tools
- ✓ *Volatility* overview

9 - Mobile Forensics

- ✓ Mobile devices
- ✓ SIM cards
- ✓ Forensic value and potential evidence
- ✓ Mobile data acquisition
- ✓ Hardware and Software tools
- ✓ *XRY* and *XAMN* overview

10 – OSINT (Open-source Intelligence)

- ✓ History of OSINT
- ✓ Information sources
- ✓ Information to intelligence cycle
- ✓ Open-source possibilities
- ✓ Automated processing
- ✓ Social media OSINT
- ✓ Dark Net OSINT

11 – Documentation and Reporting

- ✓ Physical examination
- ✓ Computer examination
- ✓ Media examination
- ✓ What to report
- ✓ Windows forensic report
- ✓ Forensic report structure

Learned knowledge will be evaluated through one individual written test and 1 team project.

Final grade = 50% Individual written test + 50% Team

Project

Dates:

2021-01-15 09:00 - Individual written test

2021-01-08 23:59 - Team Project submission (Moodle)

2021-01-15 13:00 - Team Project presentation

Classes

Dates:

16/10/2020 - Class 1 (via zoom)

23/10/2021 - Classes 2 and 3

06/11/2021 - Classes 4 and 5

20/11/2021 - Classes 6 and 7

04/12/2021 - Classes 8 and 9

18/12/2021 - Classes 10 and 11

15/01/2022 - Test and Team Project Presentation

October			
16	23	30	
November			
6	13	20	27
December			
4	11	18	
January			
8	15	22	

Teams:

Three (3) students per Team

Exceptions must be approved by the teacher

1 week to create the teams
random pool if needed

Each team will choose just a **different** topic about digital forensic analysis:

- 1 - Computer Networks
- 2 - IoT devices
- 3 - Android devices
- 4 - RAM
- 5 - OSINT techniques
- 6 - Malicious software
- 7 - Dark Net
- 8 - Virtual Machines

Organization:

- ✓ create and discuss a plan with the team members and the teacher
- ✓ check the available resources on the Internet
- ✓ class resources will be available on Moodle

1 - Submit one PDF file, named TeamX-report1.pdf, with a maximum of 10 pages

write and introduction and the state of the art about the chosen topic, as well as the experimental part, results, conclusion and bibliography with [IEEE citation style](#).

the document should be written like a research paper:

must follow the IEEE template (for A4 two columns)

2 – The PDF file will be published on Moodle for all students

3 - Prepare a presentation of up to 20 minutes

all team members must participate

present an overview of the state of the art

the presentation should focus on the experimental part, results and conclusions

Project Team Evaluation

50% – Presentation

explanation of the concepts and technical details

clarity and communications skills

argumentation in the discussion phase

50% – Report

description of concepts and procedures

expected results and tested results of forensic interest

description and usage of tools and techniques

document formatting and references

Do not commit any crime for the purpose of this project

Do not include images or videos that may violate someone's privacy

- instead, use fake images

Do not use illegal content or software to achieve your goals

Do not hack any computer without written permission

- use only virtual machines that you control and setup for this purpose

If you have any doubt about the legality of an action, ask **first**

Think thoroughly

In a real-world case, your conclusions will influence the outcome of a trial.

Write clearly

Digital forensic reports are meant to be read by nontechnical individuals:

lawyers, judges, etc.

Always follow the digital forensics investigator code of ethics

Your team should

split tasks among the team members in a fair way, but

all team members have the responsibility to review the report before delivery

Software:

Virtual machines (VMware or Virtual Box)

Windows and Linux VMs

Windows Software

Free: FTK Imager, Autopsy 4, Volatility, XAMN Viewer

Hardware:

Computers

RAM: 8GB or more recommended

Lots of disc space

Large capacity USB HDD or SSD drive (\geq 250 GB)

Low capacity USB Pen drive (\geq 8GB)

USB, SATA and IDE write blocker (can be simulated by software)

Camera and graduated set square (for scale purposes when taking pictures of equipment)

Main Bibliography

- **Mário Antunes, Baltazar Rodrigues**, Introdução à Cibersegurança - A Internet, os aspectos legais e a análise digital forense, FCA, 2018, ISBN: 978-972-722-861-4
- John Sammons, The Basics of Digital Forensics: The Primer for Getting Started in Digital Forensics, 2nd edition. Amsterdam ; Boston: Syngress, 2014.
- B. Carrier, File System Forensic Analysis, 1st edition. Boston, Mass.; London: AddisonWesley Professional, 2005.
- Cory Altheide and Harlan Carvey, Digital Forensics with Open Source Tools, 1st edition. Burlington, MA: Syngress, 2011.
- Brett Shavers, Placing the Suspect Behind the Keyboard: Using Digital Forensics and Investigative Techniques to Identify Cybercrime Suspects, 1st edition. Waltham, MA: Syngress, 2013.
- Barrett, D., & Kipper, G. (2010). Virtualization and forensics: A digital forensic investigator's guide to virtual environments. Syngress.
- Davidoff, S., & Ham, J. (2012). Network forensics: tracking hackers through cyberspace (Vol. 2014). Upper Saddle River: Prentice hall.
- Polstra, P. Linux Forensics CreateSpace Independent Publishing Platform, 2015
- Ligh, M. H., Case, A., Levy, J., & Walters, A. (2014). The art of memory forensics: detecting malware and threats in windows, linux, and Mac memory. John Wiley & Sons.
- Mahalik, H., Tamma, R., & Bommisetty, S. (2016). Practical Mobile Forensics. Packt Publishing Ltd.
- Ligh, M., Adair, S., Hartstein, B., & Richard, M. (2010). Malware analyst's cookbook and DVD: tools and

PLEASE DOWNLOAD

Please Download:

[“Bandido” Virtual Machine Disk](#) bit.ly/3aEmj9n

Ubuntu Bionic releases.ubuntu.com/bionic/

Please Install:

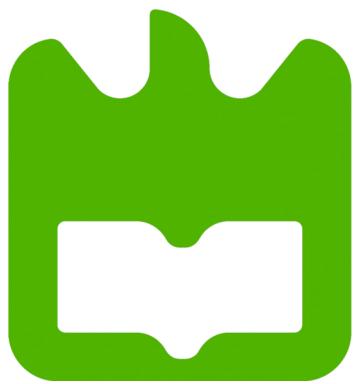
VirtualBox 6.1 virtualbox.org

7-Zip 19.0 7-zip.org

FTK Imager 4.5.0 accessdata.com

ANY QUESTIONS?





universidade
de aveiro

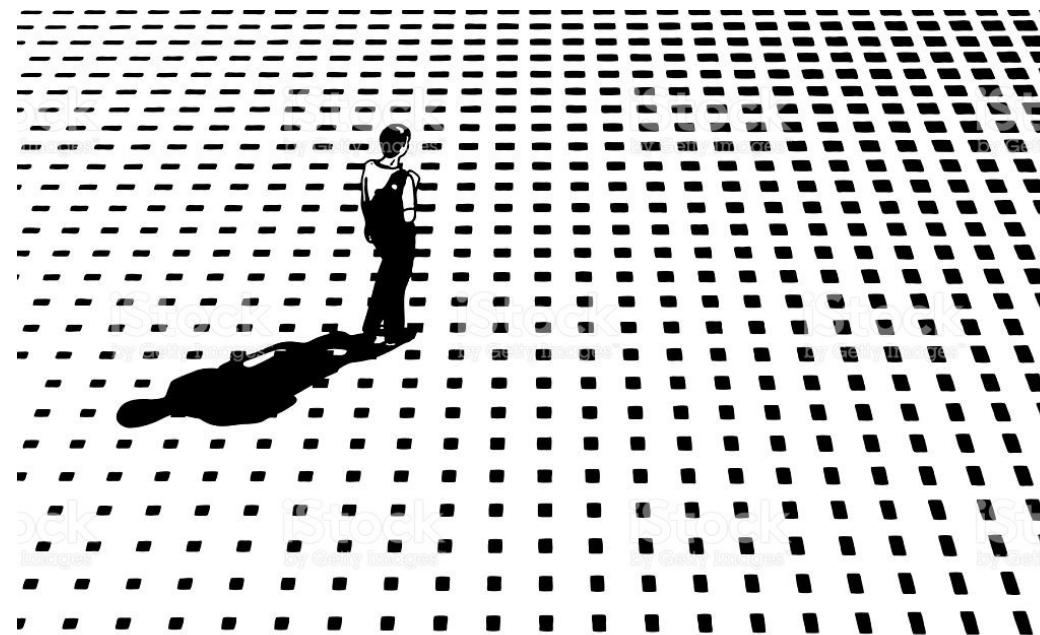
Computer Systems Forensic Analysis

AFSC

1 - Overview of Cybercrime

Artur Varanda
School Year 2021-2022

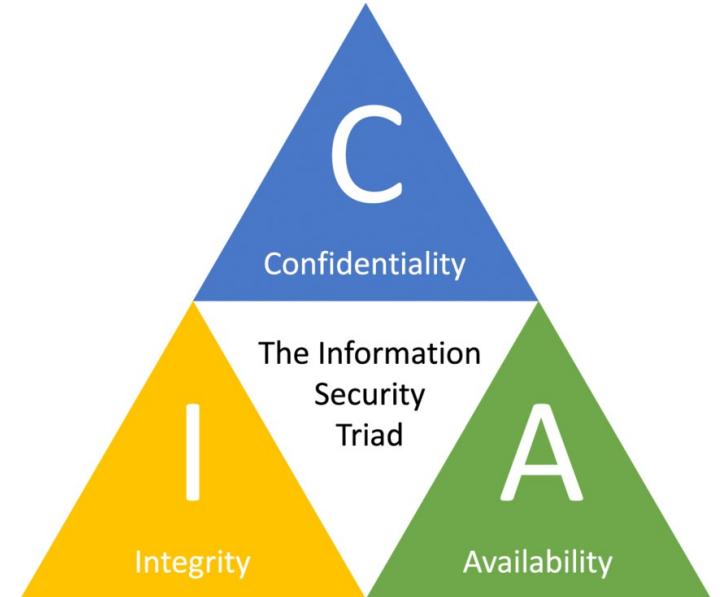
Cyberspace is the human sensation of space, offered by current communication technologies, supported by new business models, social networks, cloud computing, blogs, online stores,...



Information Security Principles

The principles of information security are based on the CIA concept:

- **Confidentiality:** ensures restriction access to information;
- **Integrity:** ensures consistency and inalterability of data;
- **Availability:** ensures data availability;



Also:

- **Non Repudiation:** ensures that the sender of information is provided with proof of delivery and the recipient is provided with proof of the sender's identity, so neither can later deny having processed the information.

AAA Services Concept:

- **Authentication:** identity verification Ex: login and password;
- **Authorization:** user privileges;
- **Accounting:** generation of logs on user actions in the system.

Cybercrime VS Computer Crime

- Cybercrime is any illicit act practiced in cyberspace, whether it is a computer crime or any other committed by computer means.



Cybercrime VS Computer Crime

- Computer Crime – an action that violates one of the CIA or AAA principles



Cybercrime

Slang

BBS's (Bulletin Board System)

Total or partial availability of information related to:

- Explosives
- Credit cards
- Description of ways to carry out crimes
- Copyright protected software
- In Portugal, BBSs are neither prohibited nor regulated, except with regard to the technical means used, which must comply with what is recommended by ANACOM – Autoridade Nacional de Comunicações.
- However, the content of BBS's and Portuguese Newsgroups cannot incite, help, facilitate or make available data or information that contravenes the law or in any way constitute a risk to personal, national or international safety.

BBS's (Cont.)

- Depending on the case, it assumes the figure of irregular practice or crime, who posts or makes available, in whole or in part, data relating to explosives, credit card numbers, description of ways of committing crimes, software protected by copyright, even if this is compressed by other programs or even if it is made available in parts or incomplete.

BlackBoxing and BlueBoxing

Blueboxing

Making unpaid phone calls using electronic devices.

Blackboxing

Interconnection of electronic components that when attached to home phones, allow all incoming calls to be received without charge to the caller.

BlackBoxing e BlueBoxing (Cont.)

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 221

Computer and communications fraud (Burla informática e nas comunicações)

1 - Whoever, with the intention of obtaining for himself or a third party illegitimate enrichment, causes another person to lose property, interfering with the result of data processing or by incorrectly structuring a computer program, incorrect or incomplete use of data, unauthorized use of data or intervention by any other unauthorized means of processing, is punishable by imprisonment for up to 3 years or with a fine.

2 - The same penalty applies to anyone who, with the intention of obtaining an illegitimate benefit for themselves or for a third party, causes damage to another person, using programs, electronic devices or other means that, specifically or together, are intended to reduce or change or prevent, in whole or in part, the normal operation or exploitation of telecommunications services.

Carding

- Handling and obtaining personal data from the face or from magnetic strips of credit, debit or telecommunications cards.
- All forms of data manipulation or identification elements, whether on the face or contained in magnetic strips of credit, debit or telecommunications cards, as well as the implantation of data or identification elements in other technical supports, constitute a crime of forgery, punishable by up to 3 years imprisonment.
- The use of identification elements or third-party bank details is a crime of fraud, punishable by a prison sentence of up to 3 years and is aggravated if the amount in question is high or if they continue this conduct more than once.
- Abuse of the possibility conferred by the possession of a credit card, even if only for the attempted form, is punishable by up to 3 years imprisonment, which may be aggravated up to 5 years or from 2 to 8 years, if the value is high or pretty high.

Portuguese Cybercrime Law (Lei n.º 109/2009)

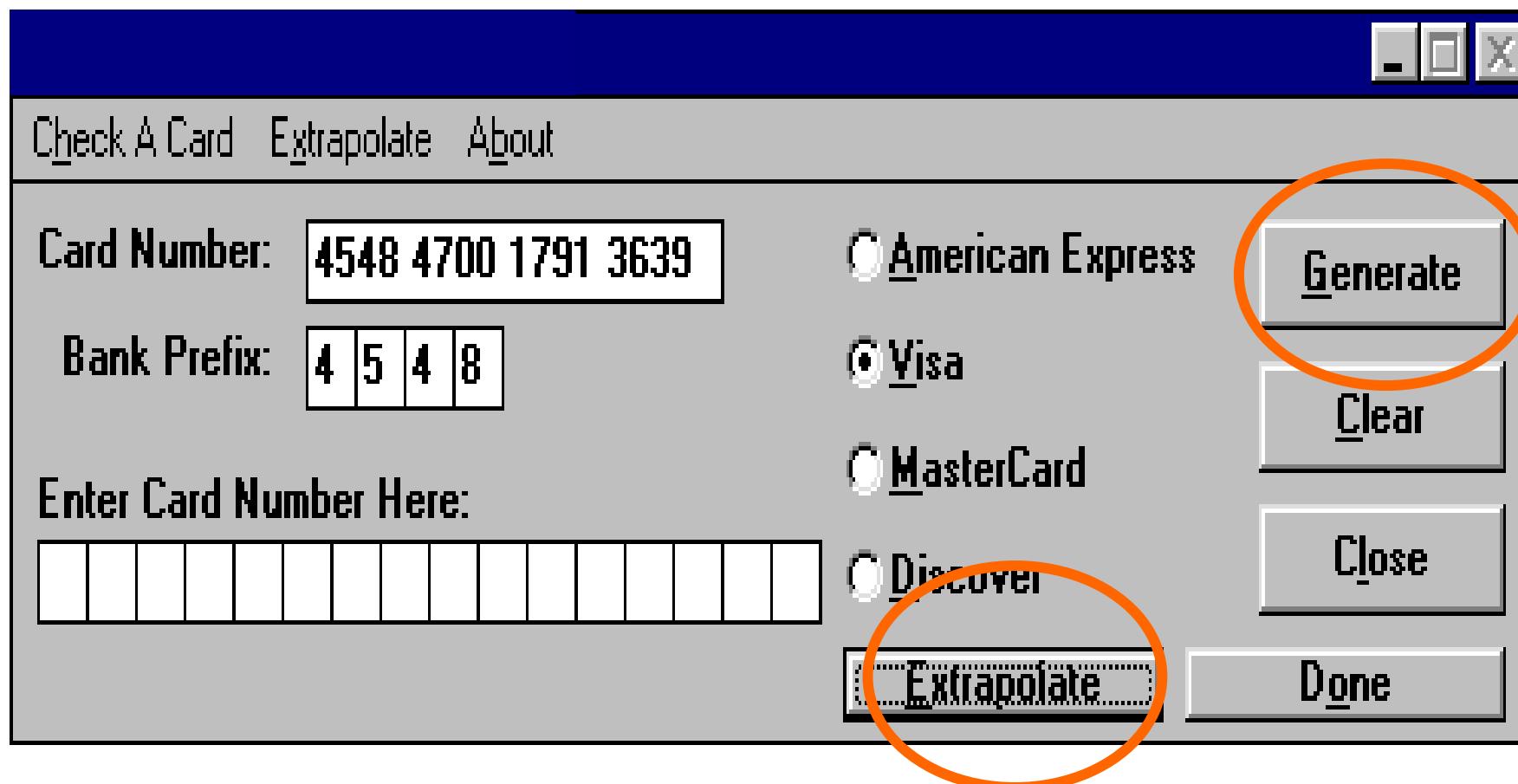
Article 3

Electronic falsification

(Falsidade informática)

1 - Who, with the intention of causing deception in legal relations, introduce, modify, delete or delete computer data or in any other way interfere in the computer processing of data, producing non-genuine data or documents, with the intention that they are considered or used for legally relevant purposes as if they were genuine, is punishable by imprisonment of up to 5 years or a fine of 120 to 600 days.

2 - When the actions described in the previous number concern the data registered or incorporated into a payment bank card or any other device that allows access to a payment system or means, to a communication system or to a conditioned access service, the penalty is of 1 to 5 years in prison.



Cracking

- The process of recovering passwords from data that has been stored in or transmitted by a computer system in scrambled form.
- Modification of software to remove or disable features which are considered undesirable to the cracker, especially copy protection systems or software annoyances, like adware.
- A cracker uses the capabilities to his own advantage while belittling damages to third parties
- The decompilation of programs is punished by the Legal Protection of Computer Programs Law and by the Portuguese Cybercrime Law, by the article 8 on illegitimate reproduction of protected program.
- This legislation covers memory resident programs (TSRs), which allow the use of utility software and games in violation of copyright.

Legal Protection of Computer Programs Law (Decreto-Lei n.º 252/94)

Article 7

Decompilaton (Descompilação)

1 - The decompilation of the parts of a program necessary for the interoperability of this computer program with other programs is always lawful, even if it involves operations provided for in the previous articles, when it is the indispensable way to obtain the information necessary for such interoperability.

2 - The holder of the user license or another person who can lawfully use the program, or persons authorized by them, have the legitimacy to carry out the decompilation, if this information is not readily and quickly available.

3 - Any stipulation contrary to the provisions of the previous numbers is null and void.

4 - The information obtained cannot:

a) Be used for an act that infringes copyright on the originating program;

b) Damaging the normal exploitation of the originating program or causing unjustified harm to the legitimate interests of the holder of the right;

c) Be communicated to others when not necessary for the interoperability of the independently created program.

5 - The program created under the terms of subparagraph c) of the previous number cannot be substantially similar, in its expression, to the original program.

Portuguese Cybercrime Law (Lei n.º 109/2009)

Artigo 8.º

**Illegitimate reproduction of protected program
(Reprodução ilegítima de programa protegido)**

- 1 - Anyone who illegitimately reproduces, discloses or communicates to the public a computer program protected by law is punishable with up to 3 years imprisonment or a fine.
- 2 - Those who illegitimately reproduce a topography of a semiconductor product or who commercially exploit or import, for these purposes, a topography or a semiconductor product made from that topography are incurred in the same penalty.
- 3 - The attempt is punishable.

Hacking

- Intrusion on computer systems in order to understand how they work and gain more knowledge about it.
- A **hacker** is a computerized intellectual who loves to break into other people's systems to simply fill his ego.
- **Hackers** do not destroy, steal or spy on information for money, unlike **crackers**.
- **Cracking** activities (illegitimate access for the purpose of data destruction) are, under Portuguese law, a crime of illegitimate access.

Portuguese Cybercrime Law (Lei n.º 109/2009)

Article 6

Illegitimate access

(Acesso ilegítimo)

1 - Anyone who, without legal permission or without being authorized by the owner, by another holder of the right to the system or part of it, in any way access a computer system, is punished with up to 1 year imprisonment or with a fine up to 120 days.

2 - The same penalty is incurred by anyone who illegitimately produces, sells, distributes or otherwise disseminates or introduces in one or more computer systems devices, programs, an executable set of instructions, a code or other computer data intended to produce the unauthorized actions described in the previous paragraph.

3 - The penalty is imprisonment for up to 3 years or a fine if the access is gained through violation of security rules.

4 - The penalty is imprisonment from 1 to 5 years when :

a) Through access, the agent has become aware of commercial or industrial secret or confidential data, protected by law; or

b) The benefit or equity advantage obtained is of a considerably high value.

5 - The attempt is punishable, except in the cases provided for in paragraph 2.

6 - In the cases provided for in paragraphs 1, 3 and 5, the criminal procedure depends on a complaint.

[General Data Protection Regulation \(GDPR\) \(Lei n.º 58/2019\)](#)

Article 47

**Improper access
(Acesso indevido)**

- 1 - Anyone who, without proper authorization or justification, accesses personal data in any way is punishable with a prison sentence of up to 1 year or a fine of up to 120 days.
- 2 - the penalty is doubled in its limits when dealing with personal data referred to in articles 9 and 10 of the GDPR.
- 3 - The penalty is also increased to double its limits when accessing:
 - a) Is achieved through violation of technical safety rules; or
 - b) Has provided the agent or third parties with a benefit or equity advantage.

General Data Protection Regulation (GDPR) (Lei n.º 58/2019)

Article 48

Data deviation

(Desvio de dados)

1 - Anyone who copies, subtracts, assigns or transfers, for a consideration or free of charge, personal data without legal provision or consent, regardless of the purpose pursued, is punishable with a prison sentence of up to 1 year or a fine of up to 120 days.

2 - the penalty is doubled in its limits when dealing with personal data referred to in articles 9 and 10 of the GDPR.

3 - The penalty is also increased to double its limits when accessing:

- a) Is achieved through violation of technical safety rules; or
- b) Has provided the agent or third parties with a benefit or equity advantage.

General Data Protection Regulation (GDPR) (Lei n.º 58/2019)

Article 49

Data tampering or destruction

(Viciação ou destruição de dados)

- 1 - Anyone who, without proper authorization or justification, deletes, destroys, damages, hides, suppresses or modifies personal data, making them unusable or affecting their potential for use, is punishable with imprisonment for up to 2 years or with a fine up to 240 days.
- 2 - The penalty is doubled in its limits if the damage produced is particularly serious.
- 3 - In the situations provided for in the preceding paragraphs, if the agent acts negligently, he is punished with imprisonment:
 - a) Up to 1 year or a fine of up to 120 days, in the case provided for in paragraph 1;
 - b) Up to 2 years or a fine of up to 240 days, in the case provided for in paragraph 2.

[General Data Protection Regulation \(GDPR\) \(Lei n.º 58/2019\)](#)

Article 50

Entering false data

(Inserção de dados falsos)

- 1 - Anyone who enters or facilitates the entry of false personal data, with the intention of obtaining undue advantage for himself or a third party, or to cause harm, is punishable with a prison sentence of up to 2 years or a fine of up to 240 days.
- 2 - The penalty is doubled in its limits if the insertion referred to in the previous number results in an effective loss.

VUI's e NUI's (Virtual User Interface e Network User Identification)

- Improper use of the so-called NUIs and VUIs to access x.25 networks is a crime of illegitimate access, punishable by the Portuguese Cybercrime Law.

Nuke

- Name given to programs that prematurely terminate a TCP/IP connection by sending ICMP packets with error messages. Such packages can be directed to the server (server-side nuke) or to the client (client-side nuke).

Phreaking

- Act of circumventing public telephones, copying telephones, tapping and even breaking into telephone exchanges by individuals with high knowledge of telephone systems (**Phreakers**).

Phreaking (Cont.)

- In addition to applying the same principles relating to blueboxing activities, the use of communication networks based on the manipulation of telephone exchanges accessed without authorization for that purpose, constitutes the crime of illegitimate access under the Portuguese Cybercrime Law.

Sniffing

- Act of listening to or intercepting other people's communications.
- It is generally used to discover passwords.
- It falls under the crime of illegitimate interception.
- The trafficking of wiretapping instruments is also a crime.

Portuguese Cybercrime Law (Lei n.º 109/2009)

Article 7

Illegitimate interception (Interceção ilegítima)

- 1 - Whoever, without legal permission or without being authorized by the owner, by another holder of the right to the system or part of it, and through technical means, intercept transmissions of computer data that take place within a computer system, the he intended or derived from it, is punishable with up to 3 years imprisonment or with a fine.
- 2 - The attempt is punishable.
- 3 - Incurs the same penalty provided for in paragraph 1 anyone who unlawfully produces, sells, distributes or in any other way disseminates or introduces in one or more computer systems devices, programs or other computer data intended to produce the unauthorized actions described in the same paragraph.

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 276

Telephone tapping instruments

(Instrumentos de escuta telefónica)

Anyone who imports, manufactures, stores, buys, sells, assigns or acquires for any reason, transports, distributes or holds an instrument or apparatus specifically intended for the assembly of telephone tapping, or for the violation of correspondence or telecommunications, outside the legal conditions or otherwise according to the provisions of the competent authority, he is punishable with a prison sentence of up to 2 years or a fine of up to 240 days.

Spam

- It consists of sending a large number of unsolicited e-mail messages.
- E-mails to publicize products and services, requests for donations of assistance works, lucky chains, proposals to earn easy money, lying rumors, among others.
- Basically, it is the simultaneous sending of an e-mail message to several users at the same time. It generally has the following characteristics:
 - a) is not requested by the recipient;
 - b) the sender's identification is false;
 - c) a victim's email server machine is used, be it an ISP or a public or private entity.

Spam (Cont.)

- In Portugal, it is this third paragraph c) that gives the criminal classification to those who send “spam”, since those who use a third-party email server in those terms can be accused of committing the crime of illegitimate access.
- The crime of Electronic falsification may also coexist if the falsified address identification referred to in paragraph b) (sender's identification is false) belongs to a specific person.
- If the purpose of "spam" is to interfere with the normal functioning of a computer system, it may be considered a crime of computer sabotage, punishable by a five-year prison sentence or a fine.

Portuguese Cybercrime Law (Lei n.º 109/2009)

Article 5

Computer sabotage

(Sabotagem informática)

1 - Whoever, without legal permission or without being authorized by the owner, by another holder of the right to the system or part of it, hinders, prevents, interrupts or seriously disturbs the operation of a computer system, through the introduction, transmission, deterioration , damage, alteration, deletion, impediment of access or deletion of programs or other computer data or any other form of interference in a computer system, is punishable with a prison sentence of up to 5 years or a fine of up to 600 days.

2 - The same penalty is incurred by anyone who illegitimately produces, sells, distributes or in any other way disseminates or introduces into one or more computer systems devices, programs or other computer data intended to produce the unauthorized actions described in the preceding paragraph.

Phishing

- It is an attempt to trick Internet service users into providing their confidential information, such as the username and password to access Home Banking.
- Often, these attempts use apparently legitimate emails or instant messages, combined with fake websites, to make fraudulent requests for information (ie, they will "fish" data).
- Phishing is a type of online fraud and phishers are nothing less than tech savvy crooks. In a typical phishing scam, phishers send emails that appear to come from a legitimate company in an attempt to trick users into providing their personal information, which will be used for "identity theft".
- Phishers use a variety of sophisticated devices to get the information they want, including pop-up windows, URL masks that simulate real web addresses, and keyboard action readers (keyLoggers, AxisLoggers, ScreenLoggers) that capture account names and passwords.

Social engineering: Electronic falsification

Phishing email dissemination; Hosting of Phishing sites; Aggregation of information collected in Phishing scams.

Intrusion: Illegitimate access

Exploits; SQL Injections; XSS; File Inclusion; Illegal login (Brute-force; Password cracking; Dictionary attacks); Bypass control system. Theft of access credentials.

Pharming

- It is an attempt to deceive Internet users by misappropriating or misusing a website's domain name or URL and redirecting its visitors to a fake website where fraudulent requests for information are made.
- Phishing and pharming activities are punishable in Portugal, depending on the applicable legal framework, such as Illegitimate access or computer sabotage crimes under the Portuguese Cybercrime Law and also as computer and communications fraud under the Portuguese Penal Code.

Internet Grooming

- Internet grooming is the English expression used to generically define the process used by sexual predators on the Internet, ranging from initial contact to sexual exploitation of children and young people.
- It is a complex, carefully individualized process, patiently developed through assiduous and regular contacts developed over time and which may involve flattery, sympathy, offering gifts, money or supposed modelling work, but also blackmail and intimidation.
- It is, in most situations, the preparatory act for another illegal activity: **Child Sexual Abuse (Pedophilia)**

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 171

Exhibitionist acts

(Atos exibicionistas)

Anyone who harasses another person, performing exhibitionist acts in front of him or her, is punishable with up to 1 year imprisonment or a fine of up to 120 days.

Article 172

Child sexual abuse

(Abuso sexual de crianças)

3 - Who:

- a) Carrying out an exhibitionist act in front of children under 14 years of age; or
- b) Acting on a minor under the age of 14, through obscene conversation or writing, pornographic performance or object, or using it in pornographic photography, film or recording;

is punishable with up to 3 years imprisonment.

4 - Anyone who performs the acts described in the preceding paragraph with a profit motive is punishable by imprisonment from 6 months to 5 years.

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 173

**Sexual abuse of adolescents and dependents
(Abuso sexual de adolescentes e dependentes)**

2 - Anyone who performs an act described in the paragraphs of paragraph 3 of article 172, in relation to a minor included in the paragraphs of the previous paragraph of this article and under the conditions described therein (*), shall be punished with up to 1 year imprisonment.

3 - Anyone who practices or takes to practice the acts described in the previous number with profit intention is punishable with up to 3 years imprisonment.

(*) A minor between 14 and 16 years of age who has been entrusted with education or assistance; or a minor between 16 and 18 years of age who has been entrusted with education or assistance, with abuse of the function he or she holds.

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 176

Children's pimping

(Lenocínio de menor)

1 - Anyone who encourages, favors or facilitates the exercise of prostitution of minors between 14 and 16 years of age, or the practice of relevant sexual acts, is punishable with imprisonment from 6 months to 5 years.

2 - If the agent uses violence, serious threat, ruse or fraudulent maneuver, acts professionally or with profit intention, or takes advantage of the victim's psychological incapacity, or if the victim is under 14 years of age, he is punished with a prison sentence of 2 to 10 years.

Sextortion

- Sextortion is a form of blackmail where someone threatens to share intimate images online unless the victim give in to their demands.
- These demands are typically for money, more intimate images or sexual favors.
- Blackmailers often target people through dating apps, social media, webcams or adult pornography sites.

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 222	Article 153	Article 154	Article 155
Extortion Threat	Duress	Severe duress	
(Extorsão)	(Ameaça)	(Coação)	(Coação grave)

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 222

Extortion

1 - Whoever, with the intention of obtaining for himself or for a third-party illegitimate enrichment, constrains another person, through violence or threat with important harm, to a patrimonial disposition that entails, for him or for others, damage is punishable with the penalty of imprisonment for up to 5 years.

2 - If the threat consists in the disclosure, through the media, of facts that could seriously damage the reputation of the victim or another person, the agent is punished with imprisonment from 6 months to 5 years.

Datajacking / Ramsomware

- Extortion of companies through the action of a hacker who after illegally access the system of that company proceeds to the encryption of the data stored there.
- The company is then contacted, and a ransom is required so that they can regain access to the system and information.
- It is punished in Portugal, depending on the applicable legal framework, as a crime of illegitimate access or computer sabotage under the Portuguese Cybercrime Law and as a **document extortion** under de Portuguese Penal Code.

Portuguese Penal Code (Decreto-Lei n.º 48/95)

Article 223

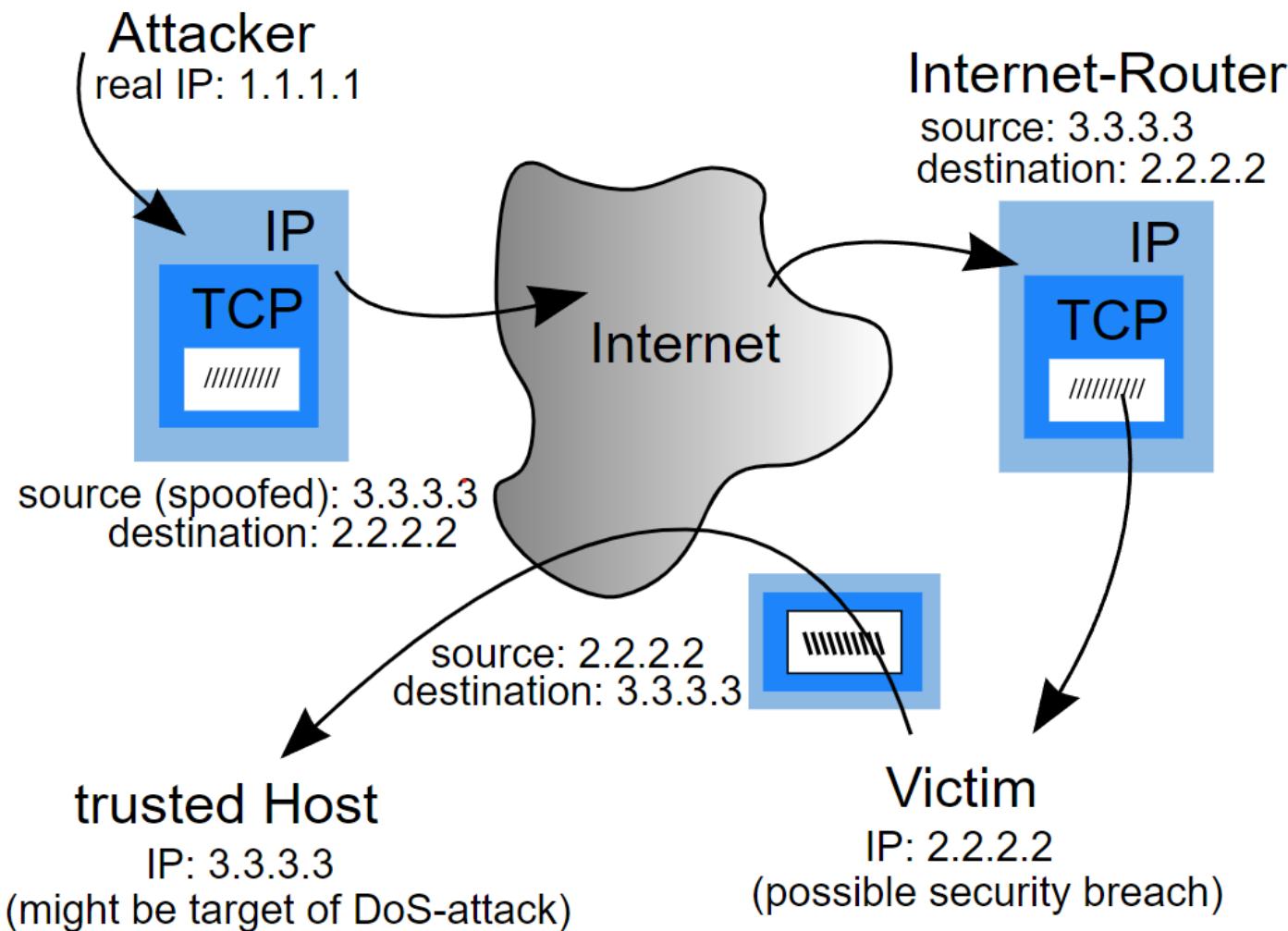
Document extortion

(Extorsão de documento)

Anyone who obtains, as a guarantee of debt and abusing another person's situation of need, a document that could give rise to criminal proceedings, is punishable with a prison sentence of up to 2 years or a fine of up to 240 days.

IP spoofing

- It is a computer systems subversion technique that consists of masking (spoof) IP packets using spoofed sender addresses.
- In the IP protocol, the forwarding of packets is based on a very simple premise: the packet must go to the recipient (destination-address) and there is no verification of the sender — there is no validation of the IP address nor its relationship with the previous router (who forwarded the package). Thus, it becomes trivial to spoof the source address through simple manipulation of the IP header.
- In Portugal, it can be penalized as a crime of Electronic falsification under the Portuguese Cybercrime Law or as a computer and communications fraud under the Portuguese Penal Code.

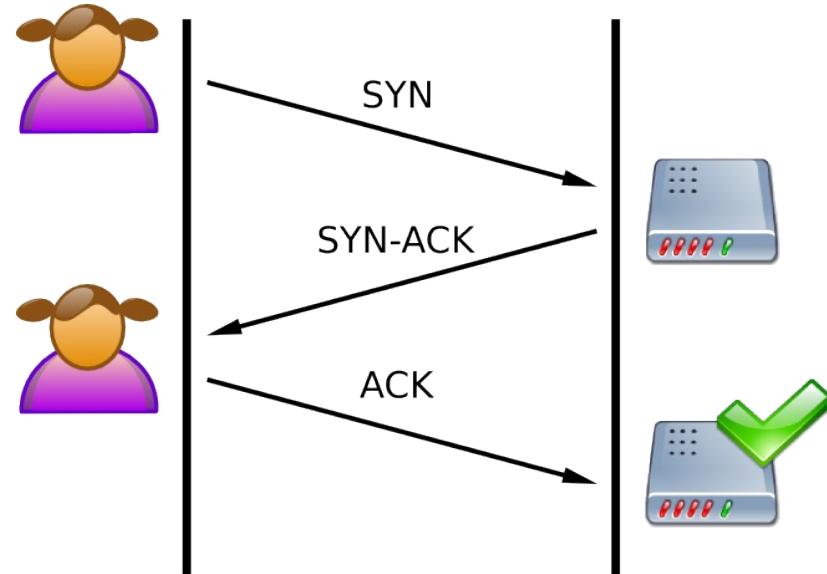


SYN flood (Denial of Service – DoS or DDoS)

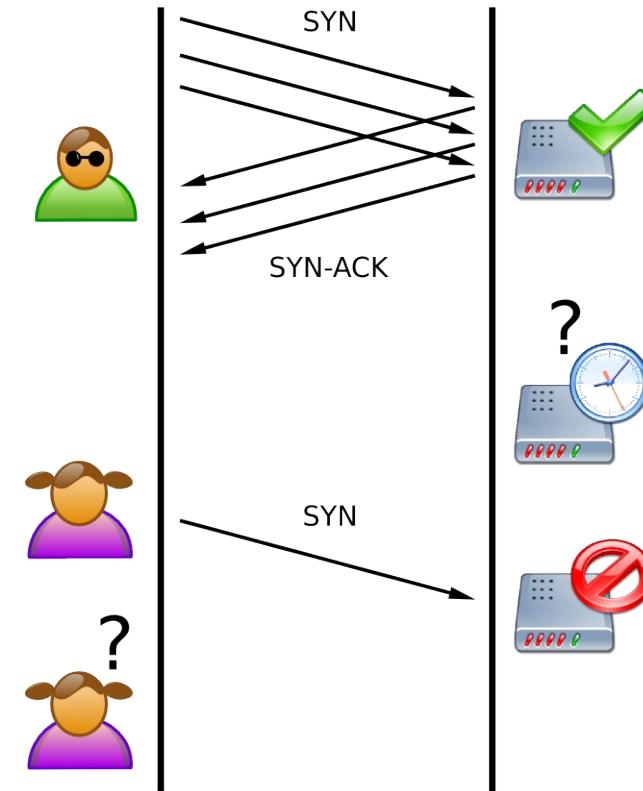
- SYN attack is a form of denial-of-service attack on computer systems, in which the attacker sends a sequence of SYN (synchronization) requests to a target system.
- When a client tries to initiate a TCP connection with a server, the client and server exchange a series of messages, which are typically:
 1. The client requests a connection by sending a SYN (synchronize) to the server.
 2. The server confirms this request by sending a SYN-ACK back to the client.
 3. The client in turn responds with an ACK, and the connection is established.
- This is called the Three-Way Handshake.

SYN flood (Cont.)

Normal Connection (TCP 3-Way Handshake)



SYN Flood Attack



SYN flood (Cont.)

- A malicious client may not send this last ACK message.
- The server will wait for this for a while, as simple network congestion can be the cause of the missing ACK.
- This so-called semi-open connection can take up resources on the server or cause losses for companies using licensed software per connection.
- It might be possible to occupy all the resources of the machine, with several SYN packages.
- Once all resources are occupied, no new connections (legitimate or otherwise) can be made, resulting in a denial of service.
- It is penalized as a crime of computer sabotage under the Portuguese Cybercrime Law.

Cyberterrorism

- Cyberterrorism is the use of the Internet to conduct violent acts that result in, or threaten, the loss of life or significant bodily harm, in order to achieve political or ideological gains through threat or intimidation.
- The 2007 cyberattacks on Estonia were a series of cyberattacks which began on 27 April 2007 and targeted websites of Estonian organizations, including Estonian parliament, banks, ministries, newspapers and broadcasters. Most of the attacks that had any influence on the general public were distributed denial of service type attacks ranging from single individuals using various methods like ping floods to expensive rentals of botnets usually used for spam distribution.

- Stuxnet is a malicious computer worm first uncovered in 2010 and thought to have been in development since at least 2005. Stuxnet targets supervisory control and data acquisition (SCADA) systems and is believed to be responsible for causing substantial damage to the nuclear program of Iran. Although neither country has openly admitted responsibility, the worm is widely understood to be a cyberweapon built jointly by the United States and Israel in a collaborative effort known as Operation Olympic Games
- In late November 2014, Sony Pictures Entertainment was hacked by a group calling itself the Guardians of Peace. The hackers, who are widely believed to be working in at least some capacity with North Korea, stole huge amounts of information off of Sony's network.

Malware : Crime of computer sabotage under the Portuguese Cybercrime Law.

Infection; Dissemination; Web hosting or Server; Replication.

Non-availability and sabotage: Crime of computer sabotage under the Portuguese Cybercrime Law.

DoS; Disruption of processing and response capacity; Package Flood; Exploit;

Illicit information collection: Crime of illegitimate interception under the Portuguese Cybercrime Law.

Scan; Probe to system; Network scan; DNS zone transfer; Sniffing; Wiretapping

Smurf Attack

- Like the SYN flood attack, although it involves a forged ICMP (Ping Service Protocol) packet sent to a broadcast address, targeted to most operating systems and routers.
 - The Smurf attack is a category of network-level attacks perpetrated against hosts with the aim of denying services.
 - The attacker sends large amounts of ICMP traffic echo requests (ping) to a network broadcast IP using a source address (spoofed IP) of the victim.
 - In a multi-access broadcast network, it can cause a few hundred computers on the network to respond to the request for each packet, which causes the computers on the network to bombard the victim with a response to the forged request.
- The Fragle attack is a variation of the Smurf attack that sends large amounts of UDP packets to ports 7 (Echo) and 19 (Chargen).
 - Currently, the machines most affected by this type of attacks are IRC servers and their suppliers. This type of attack is penalized in Portugal just like SYN flood attacks.

Cybersquatting

- Malicious practice which consists of registering domains relating to large companies or famous people (domain name) with the intention of taking advantage of the popularity of the person or the company's trademark, also known as domain trafficking.
- Cybersquatters often register these domains before the target company, thus forcing the target company to buy the domain from them at a higher price.
- Cybersquatting comes from the term “squatting”, which describes the act of occupying a space or building, abandoned or uninhabited, without permission from its legal owners.
- In some cases, the domain name is used to post derogatory comments about the target company. The legitimate company or person has no other option than to buy the domain name at ridiculously high prices.
- This practice can be penalized as extortion in the Portuguese Penal Code.

Website defacement

- Website defacement is an attack on a website that changes the visual appearance of a website or a web page.
- These are typically the work of defacers, who break into a web server and replace the hosted website with one of their own.
- It is penalized as a crime of damage relating to programs or other computer data under the Portuguese Cybercrime Law.

Portuguese Cybercrime Law (Lei n.º 109/2009)

Article 4

Damage related to programs or other computer data (Dano relativo a programas ou outros dados informáticos)

- 1 - Whoever, without legal permission or without being authorized by the owner, by another holder of the right to the system or part of it, deletes, alters, destroys, in whole or in part, damages, suppresses or renders unusable or inaccessible programs or other computer data from others or in any way affecting their ability to use, is punishable with up to 3 years imprisonment or a fine.
- 2 - The attempt is punishable.

Warez

- Term derived from the English language, second half of the word software in the plural, under an I33t (elitist) pronunciation: wares /'wɛərz/.
- It is Software that is illegally distributed over the Internet. The "Z" is purposeful, serving to indicate something that is illegal. It can also be used in other terms such as Gamez (pirated games), Romz (video games for PC through emulators, but also illegal), i.e., the term refers to the illegal trade (**piracy**) of copyrighted products used in general in the within organized groups, making use of **peer-to-peer** networks, sharing files among friends or among large groups of people with similar interests.
- Penalized in Portugal as the crime of usurpation under the Portuguese Code of Copyright and Related Rights and as a crime of illegitimate reproduction of a protected program, under the the Portuguese Cybercrime Law.
- Copying and distributing to third party computer programs protected by law - copyright - are prohibited and punishable by law up to three years in prison. Attempted copying or distribution is also punishable.
- This law covers the total or partial distribution of computer programs, even if compressed by other programs, in newsgroups, IRC's, www, ftp, etc.

Portuguese Code of Copyright and Related Rights (Decreto-Lei n.º 63/85)

Article 195

Usurpation

1 - Any person who, without authorization from the author or the artist, the producer of phonogram and videogram or the broadcasting organization, uses a work or service in any of the ways provided for in this Code, commits the crime of usurpation.

2 - Also commits the crime of usurpation:

a) Anyone who improperly discloses or publishes a work not yet disclosed or published by its author or not intended for dissemination or publication, even if it is presented as belonging to the respective author, whether or not intending to obtain any economic advantage;

b) Whoever collects or compiles published or unpublished works without the author's authorization;

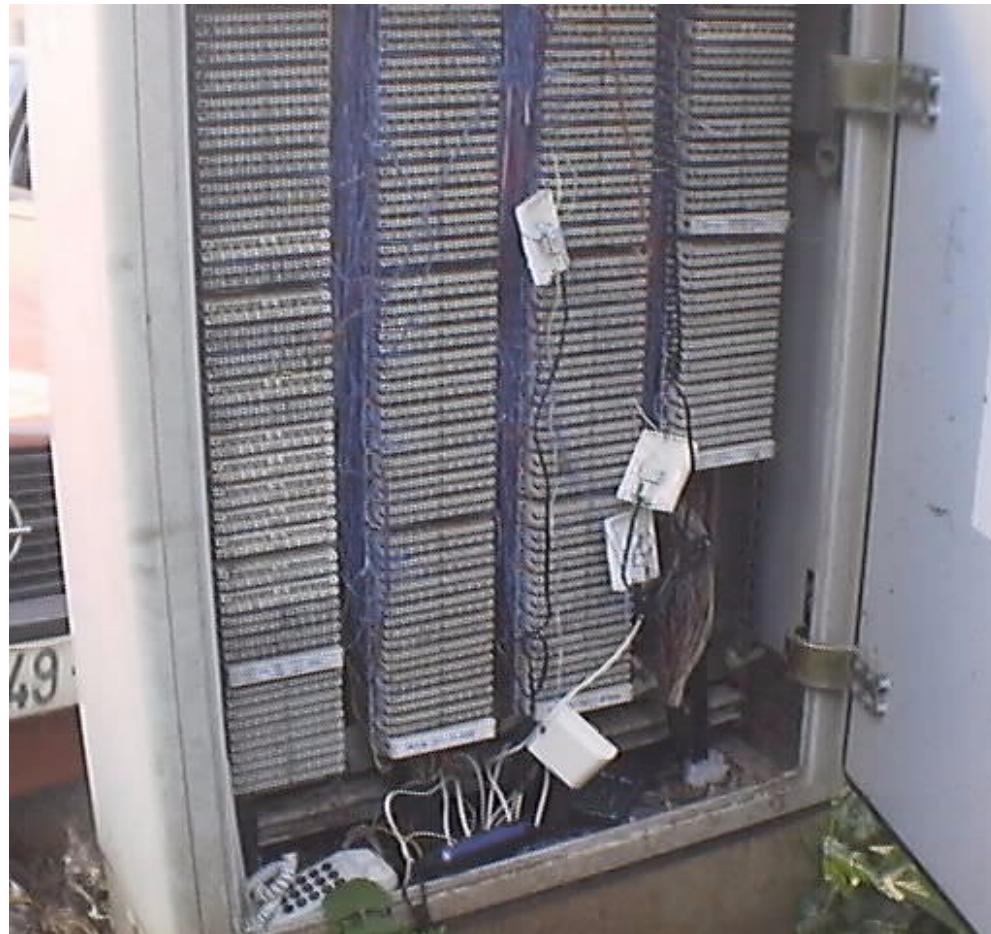
c) Who, being authorized to use a work, artist, phonogram, videogram or broadcast broadcast, exceeds the limits of the authorization granted, except in the cases expressly provided for in this Code.

3 - Will be punished with the penalties provided for in article 197, the author who, having transmitted, in whole or in part, the respective rights or having authorized the use of his work in any of the ways provided for in this Code, to use it directly or indirectly with offense of the rights attributed to others.

Examples (Images)

Examples - Dialers



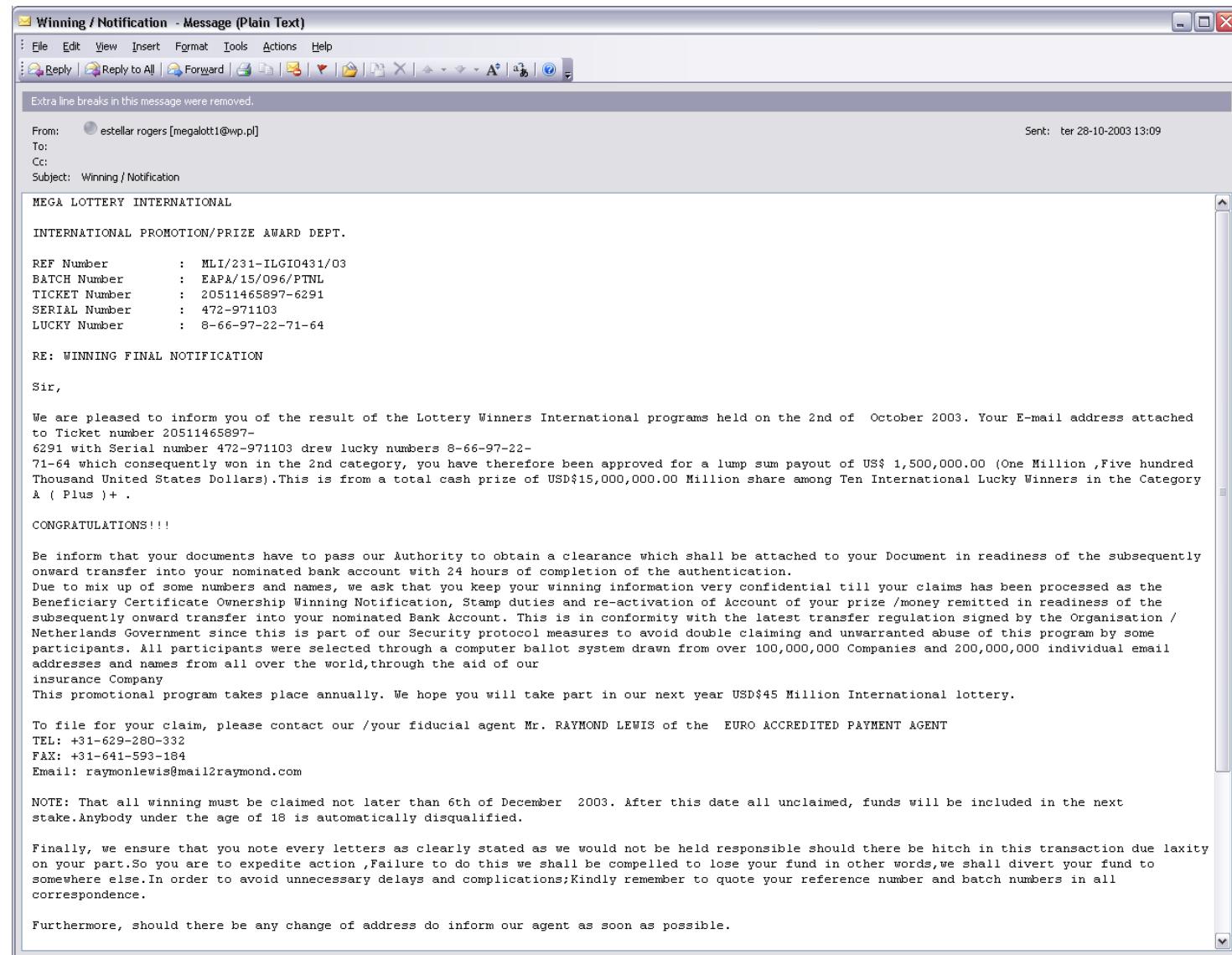


Distribution Box



Call-box

Examples - Nigerian Letter or “419” Fraud (Advance-fee scam)



Examples - Warez





1 card type

2.6 BIN

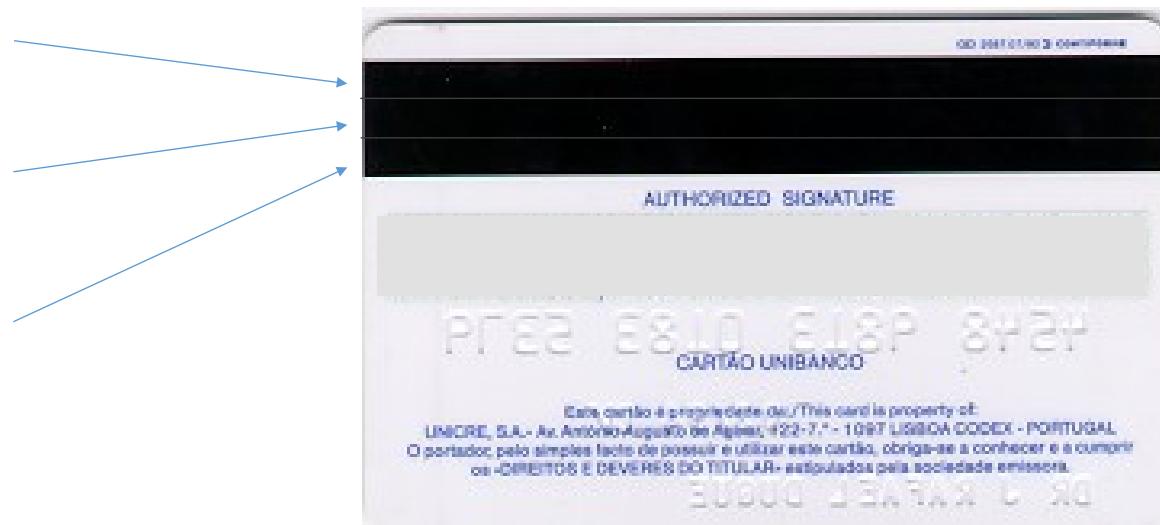
7.15 acc

check digit

Track 1: Holder Name, Primary Account Number, Expiration Date, Verification Values

Track 2: Primary Account Number, Expiration Date, Verification Values

Track 3: Primary Account Number, User and Security Data, Additional Data



Examples – Credit Cards



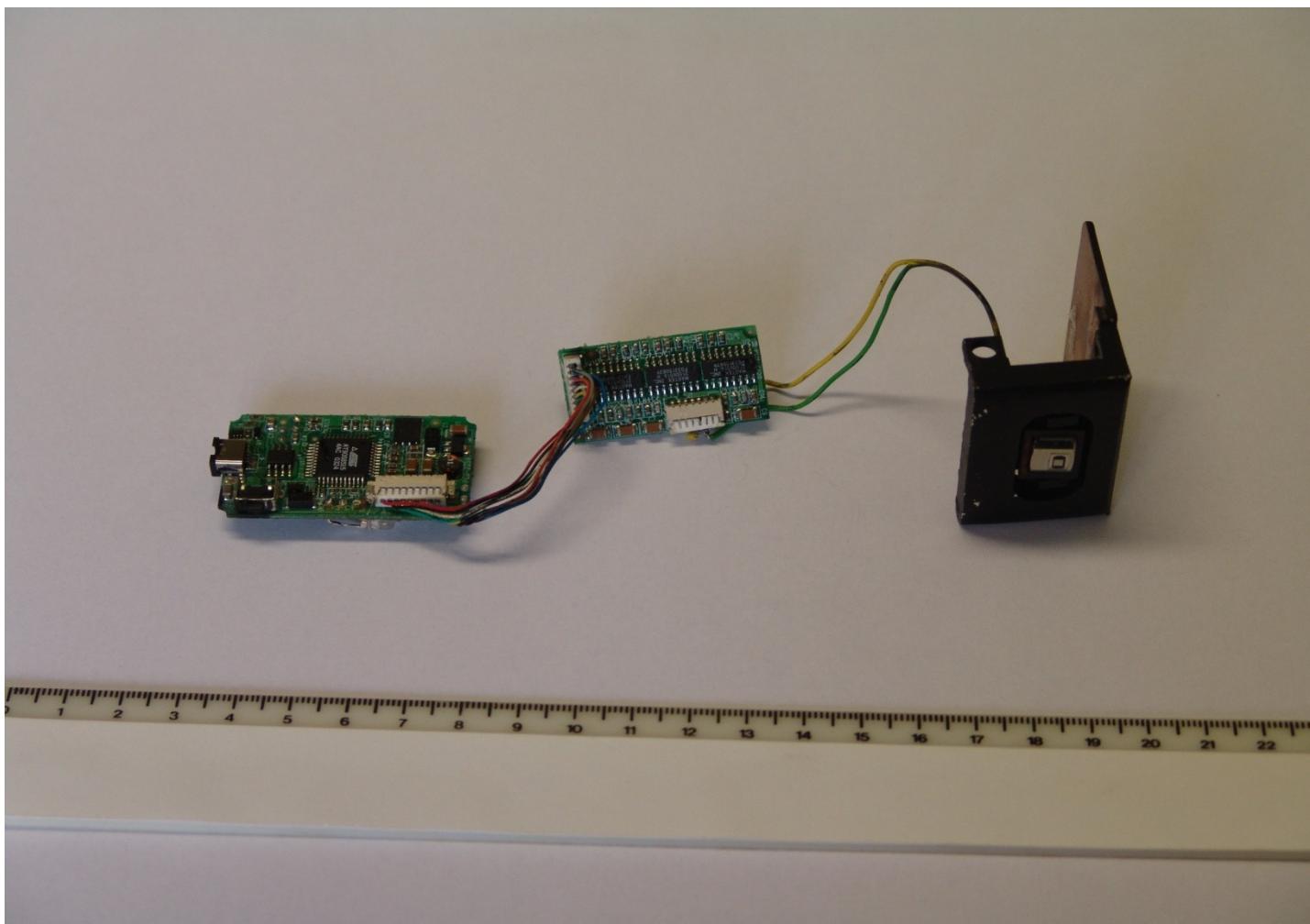
The screenshot shows a Windows application window titled "Portable Reader Data". The window has a blue header bar with the title and a red close button. Below the header is a toolbar with several buttons: "Record Transfer", "Poll Reader Data", "Stop Polling", "Purge Reader Data", "Save (text file)", "Save (Database)", "Delete Record", and "Close".
The main area is divided into two sections:

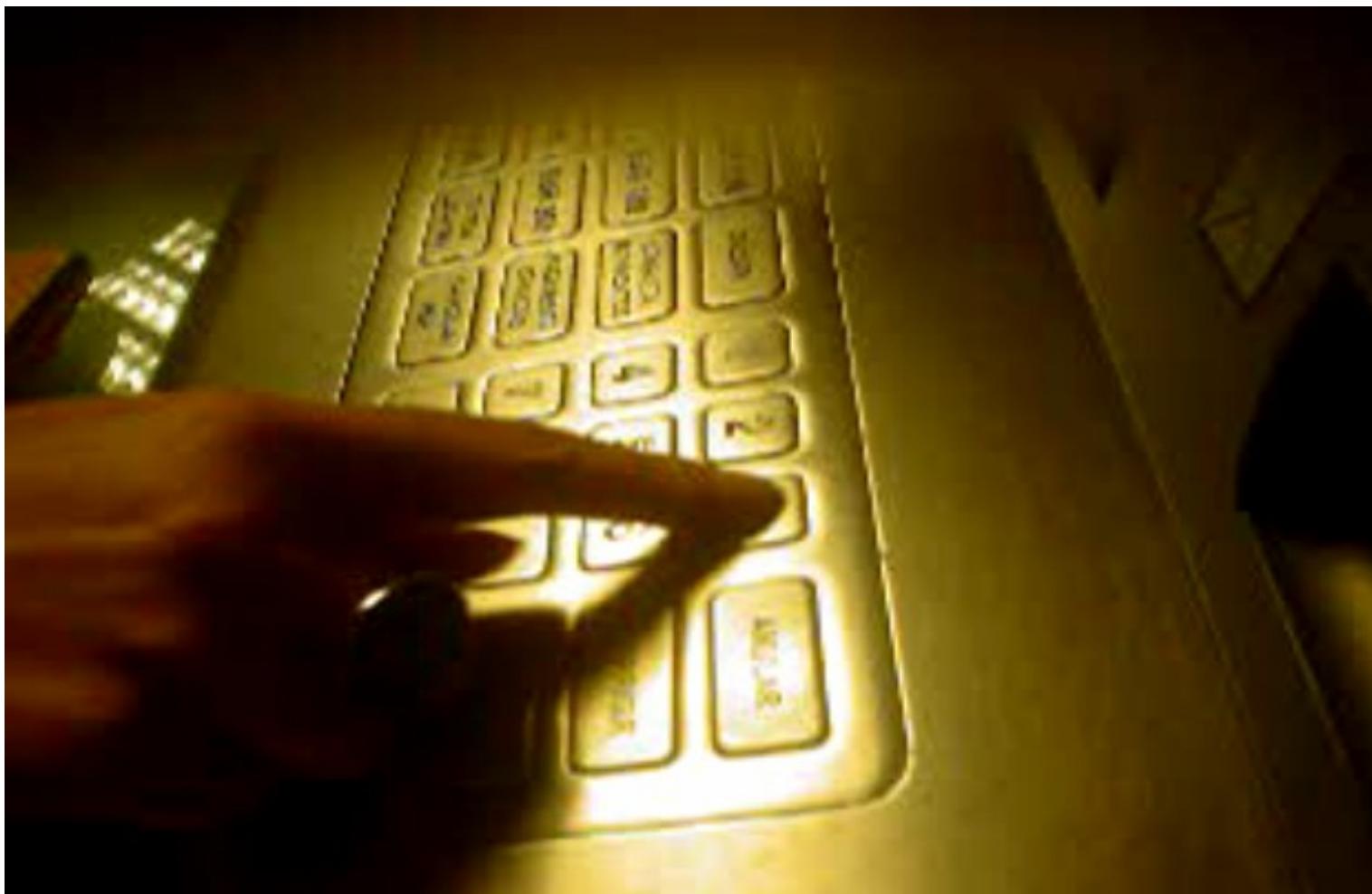
- Record Transfer:** Displays "Total Records in Reader: 15" and "Total Records Transferred: 15".
- Records to View:** Contains a radio button group for "All Records" (selected) or "Date Range", and two date pickers showing "07-06-2006" and "08-06-2006". To the right are "Load" and "Preview" buttons.

A large table below lists records from 1 to 15. The columns are labeled "Rec", "Track 1", "Track 2", "Track 3", and "Date/Time". The "Track 2" column contains binary data strings. Record 11 is highlighted with a blue background, indicating it is currently selected or being viewed.

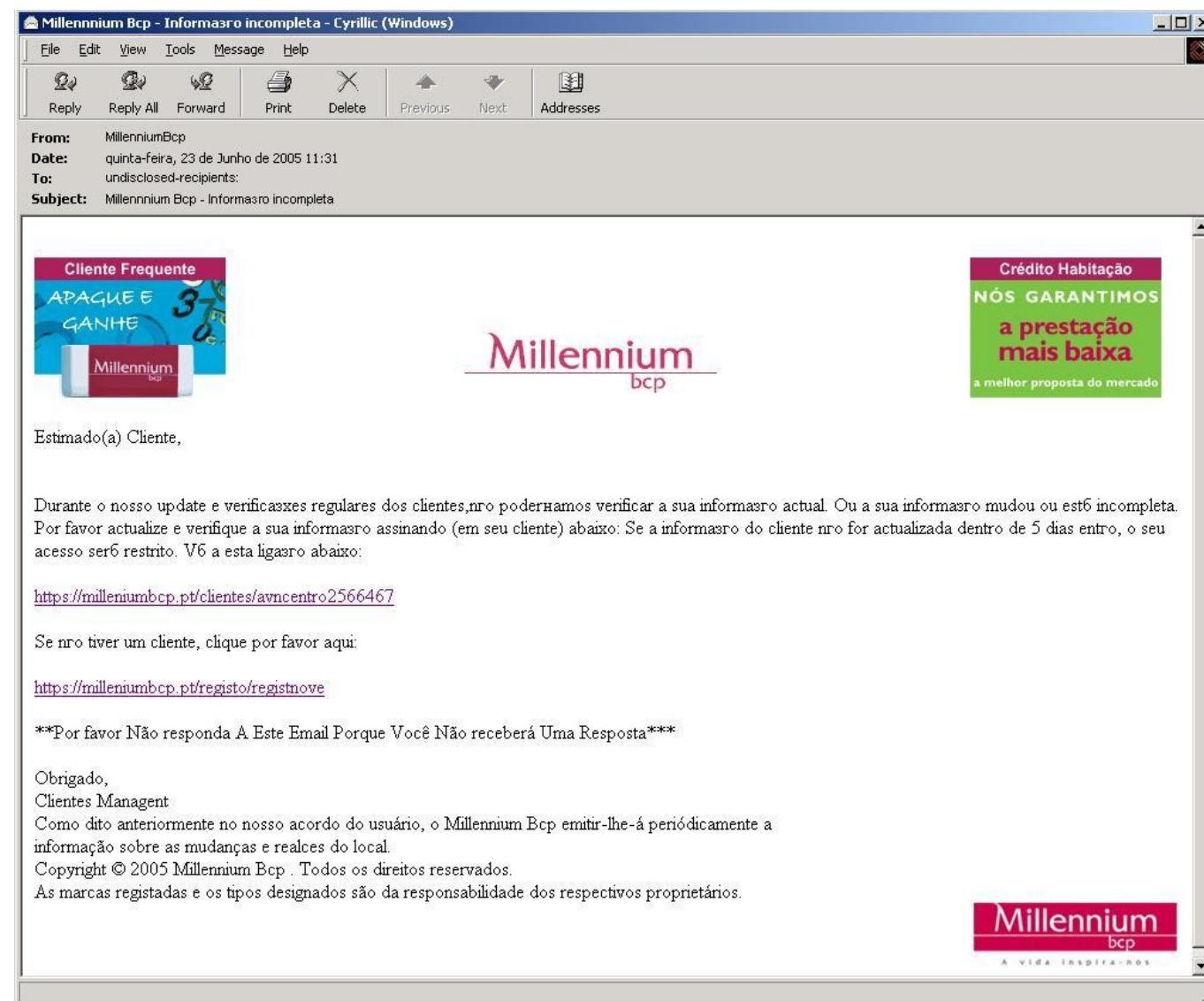
Rec	Track 1	Track 2	Track 3	Date/Time
1		0290100231320=0000032228070000000000		05-25-2006, 07:25:01
2		6337020210020569522=4912561014000000		05-25-2006, 07:25:58
3		6337020210020569522=4912561014000000		05-25-2006, 07:26:01
4		4406440526902200=08021261000787400000		05-25-2006, 07:50:05
5		0440000016001		05-25-2006, 11:19:09
6				,2006/05/2,5 11:19:3
7				,2006/05/2,5 11:19:3
8				,2006/05/2,5 11:19:3
9				,2006/05/2,5 11:19:3
10				,2006/05/2,5 11:19:4
► 11				,2006/06/0,9 04:58:5
12				,2006/06/0,9 04:59:0
13				,2006/06/0,9 04:59:3
14				,2006/06/0,9 04:59:4
15				,2006/06/0,9 05:02:1
*				

Examples - Carding

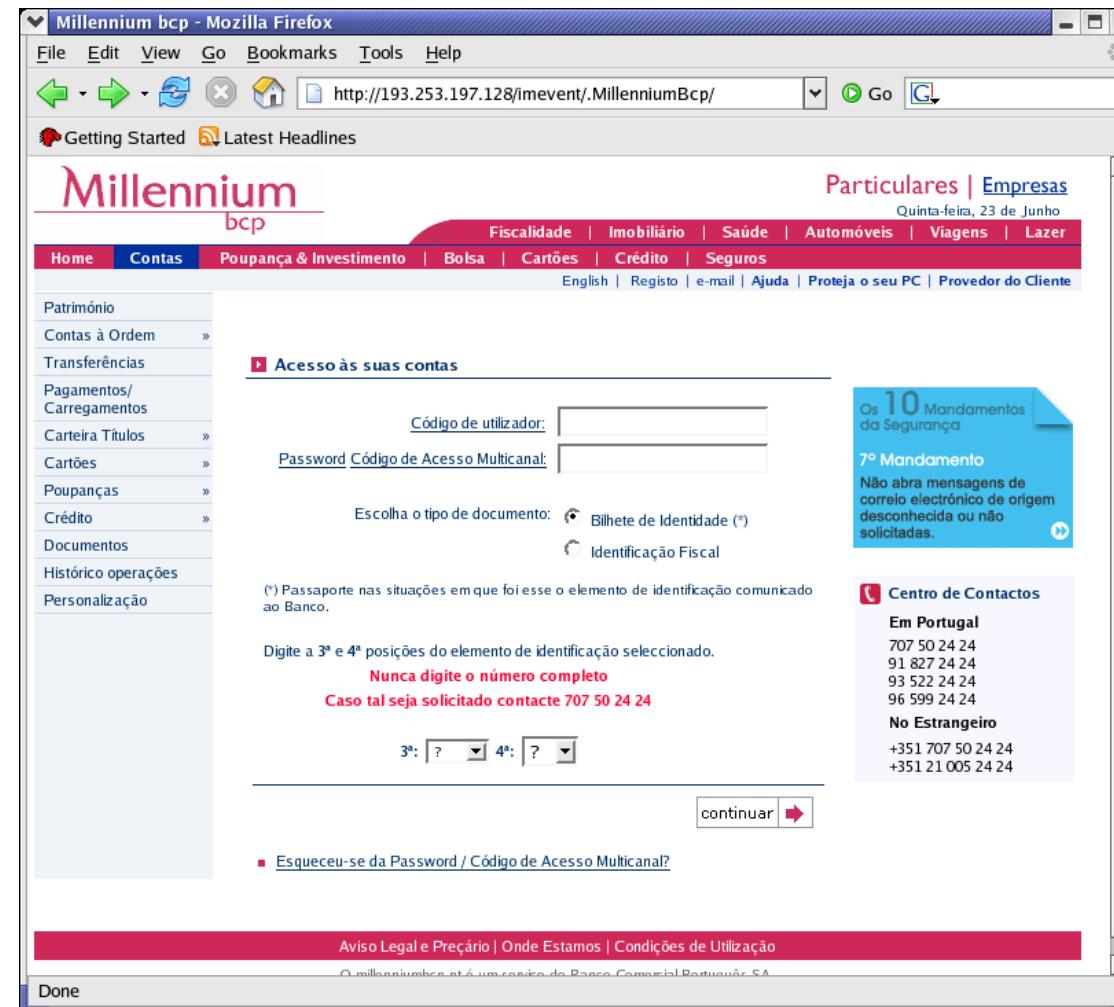




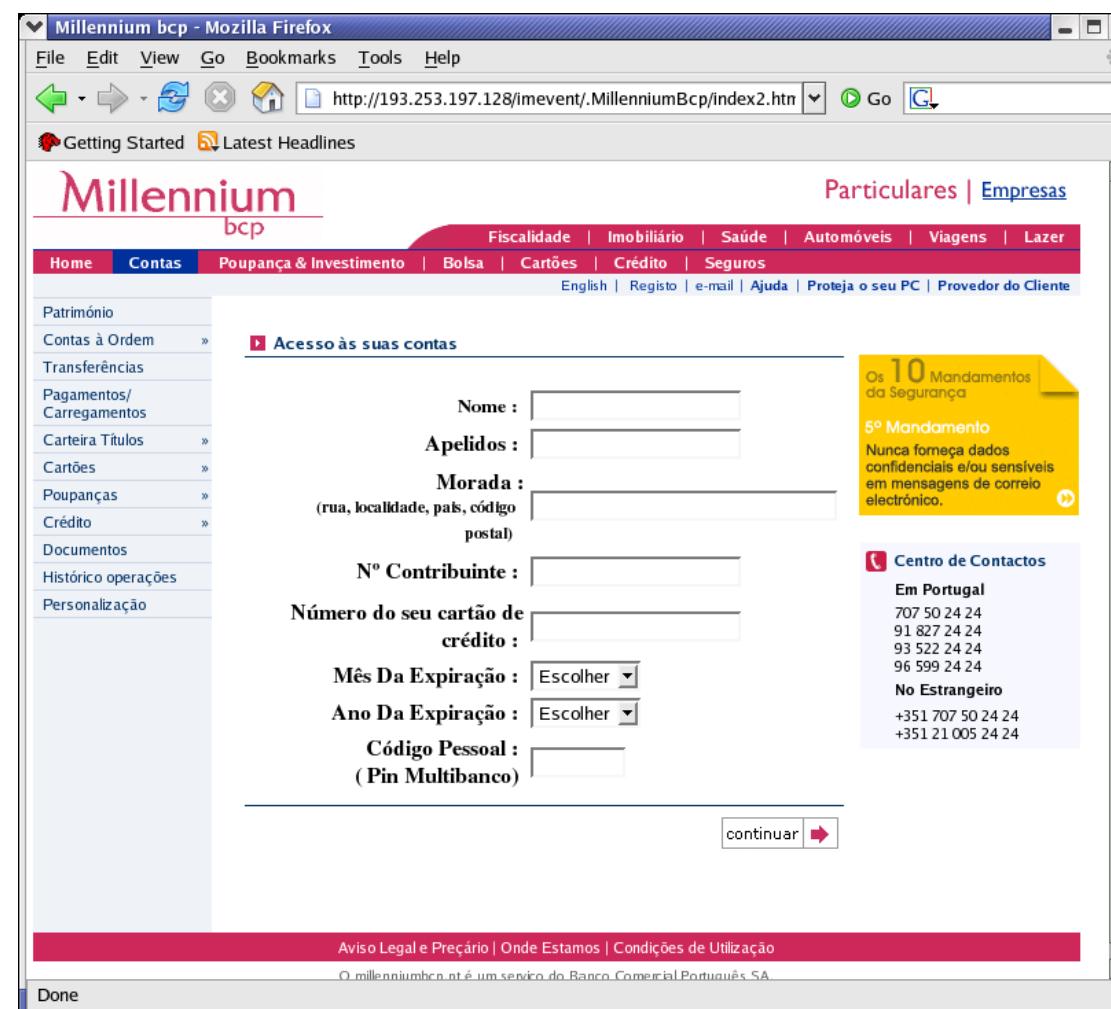
Examples - Phishing

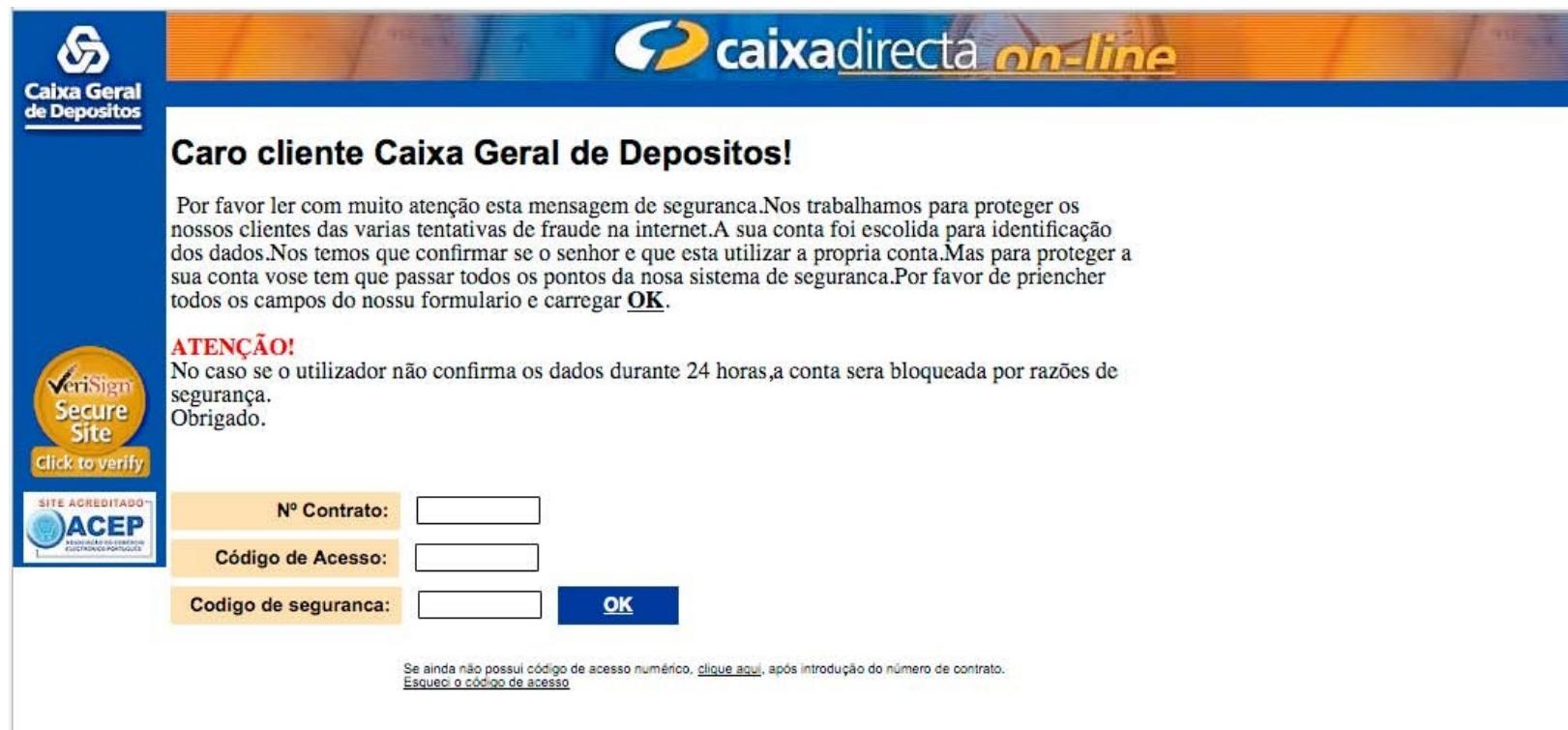


Examples - Phishing



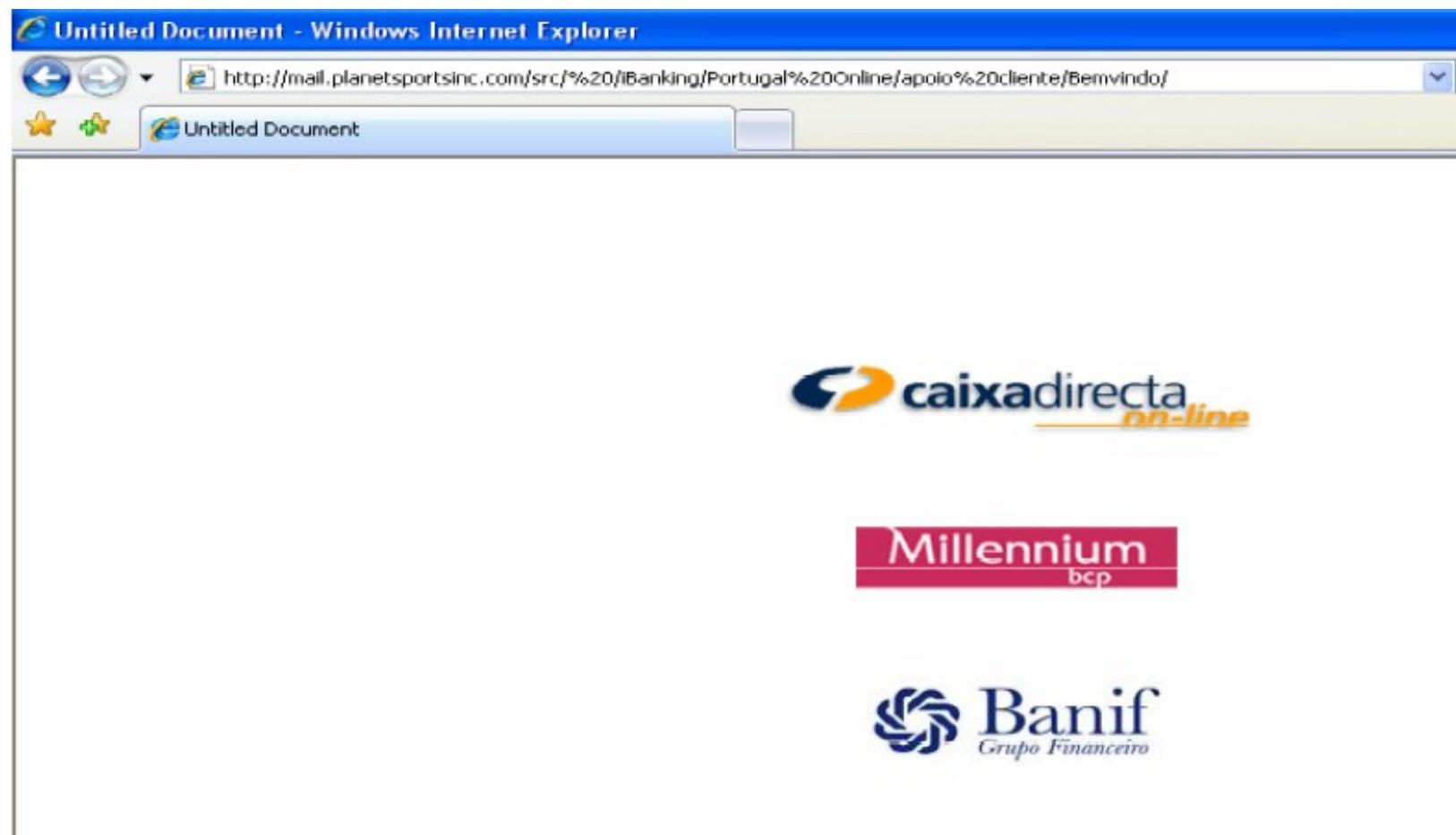
Examples - Phishing



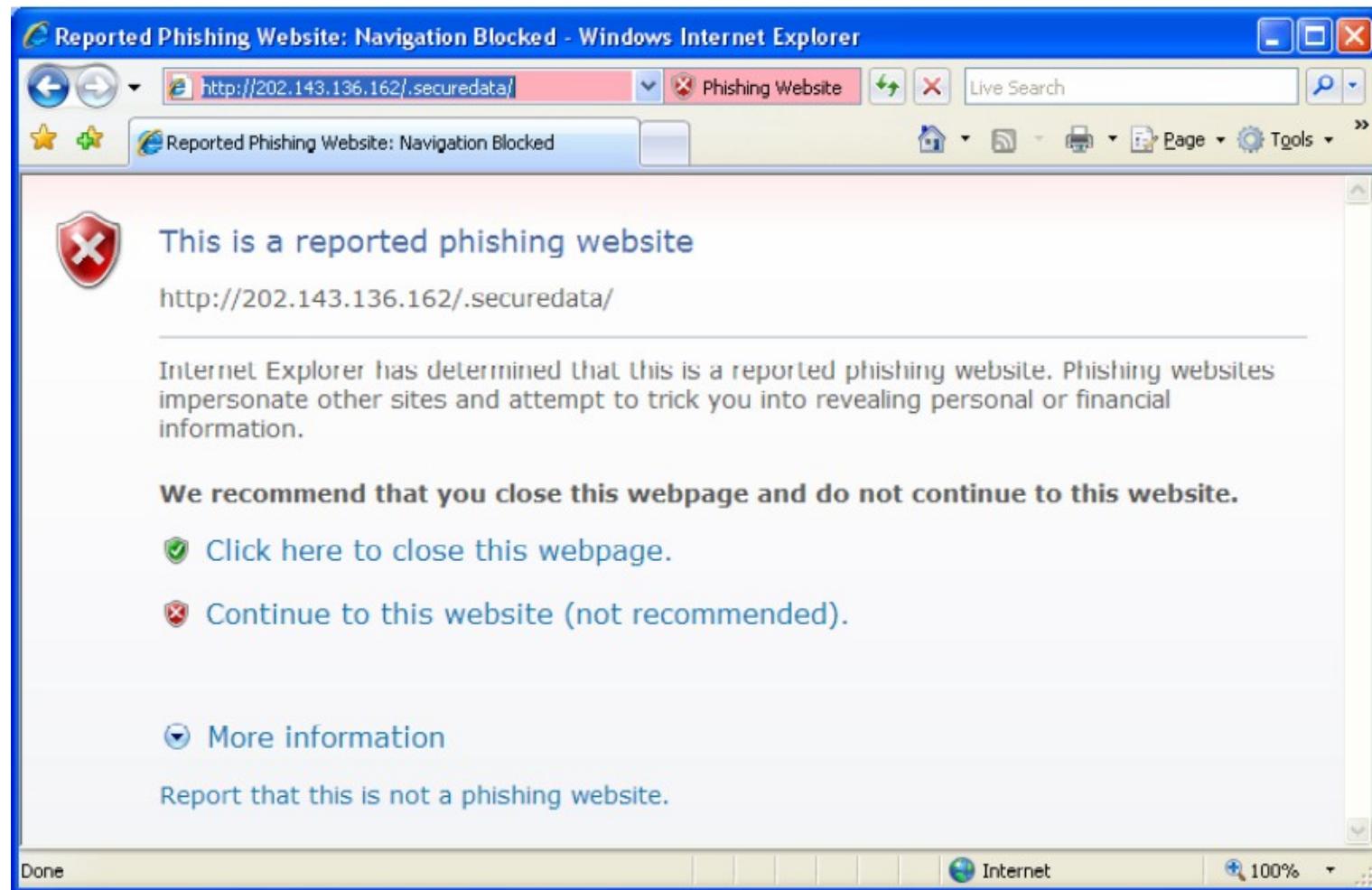




Examples - Phishing



Examples - Phishing



Digital Evidence



Digital evidence, like any other evidence, must be:

- Admissible
- Authentic
- Precise
- Complete

Digital Evidence - Requirements

- legally admissible
 - how it is obtained
- technically irrefutable
 - source
 - integrity
 - certification (digital signature)
 - dual control

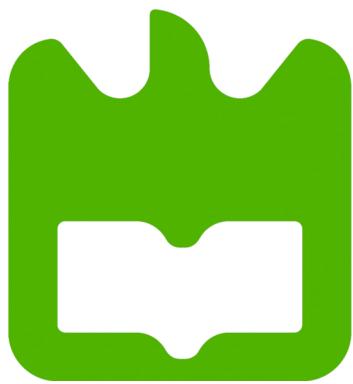
Adapted from BLAKESLEEL, Hyechin - USE OF
COMPUTER FORENSICS TECHNOLOGIES
IN CRIME INVESTIGATION, 2009; KSANDER, 2006;

The process of a forensic analysis is divided into four stages:

1. Identify the source of the digital evidence;
2. Preserving the evidence (involves the duplication of evidence according to technical-legal processes);
3. Analysis and investigation of evidence;
4. Presentation of reports and documentation of results.

ANY QUESTIONS?





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Introduction to digital forensics

Artur Varanda

School Year 2021-2022

Digital investigation focus:

- digital devices that has been involved in an incident or crime
- device used to:
 - ✓ commit a physical crime – e. g. a suspect used the Internet to conduct research about a physical crime
 - ✓ execute a digital event that violates a policy or law – e. g. an attacker gains unauthorized access to a computer, a user downloads contraband material, or a user sends a threatening e-mail, etc;
- When the violation is detected, an investigation is started to answer:
 - ✓ what, who, when, how
 - ✓ in some cases “where” and “why”

A digital investigation is

- a process where we develop and test hypotheses that answer questions about digital events
 - ✓ a scientific method
 - ✓ where we develop a hypothesis using evidence that we find
 - ✓ and then test the hypothesis by looking for additional evidence that shows the hypothesis is impossible

Digital evidence

Is a digital object that contains **reliable** information that supports or refutes a hypothesis. The digital evidence must be:

- admissible, authentic, accurate and complete

Digital evidence is:

- information stored or transmitted in digital formats or media, the content of which is evidence, whether material or merely indicative, of a particular incident or event;
- It is fragile and volatile, so the attention of a certified expert is required in order to ensure that the data of probative value are effectively isolated and extracted correctly and lawfully.

Digital evidence challenges:

- **hard to control** – it is very easy to create, modify, transmit or delete data in short amount of time
- **diversity and complexity** – sometimes is hard to identify the digital evidences because information systems evolve too fast

Forensic means

it has legal requirements to be accepted into a court of law

Note:

A **digital forensic investigation** is a more restricted form of digital investigation

Definition by Brian Carrier

Process that uses science and technology to analyze digital objects and that develops and tests theories, which **can be entered into a court of law**, to answer questions about events that occurred.

Another definition

The systematic and technological inspection of a computer system and its contents in order to obtain evidence of a crime or any other use that is under investigation.

Types of analysis to find evidences:

- *live analysis* – when the operating system or other resources of the system being investigated is used to find evidence
 - ✓ advantages: get data from RAM of a running process
 - ✓ disadvantages: risk of getting false information because the software could maliciously hide or falsify data
- *post-mortem analysis* – when trusted applications in a trusted operating system are used to find evidence (lab environment)
 - ✓ advantages: fully controlled environment
 - ✓ disadvantages: information from RAM is lost, e. g. key to decrypt a file, ...

A post-mortem analysis is more ideal, but not always possible.

A server has been compromised, how it occurred and who did it?

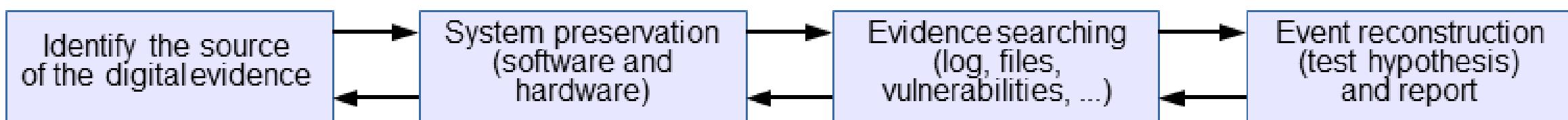
- find data that were created by events related to the incident
 - recover deleted log entries from the server
- find attack tools
- find the vulnerabilities that existed on the server
- using this data, and more, we develop an hypotheses
 - ✓ which vulnerability the attacker used to gain access
 - ✓ what he/she did afterwards
- later, examine the firewall configuration and logs
 - ✓ determine that some of the scenarios in our hypotheses are impossible because that type of network traffic could not have existed
 - ✓ evidence was found that refutes one or more hypotheses

Digital Crime Scene Investigation Methodology

Investigation process

- There is no single way to conduct an investigation
- It does not matter which process is used,
 - As long as we find the right person and do not break any laws in the process
- However, some are more efficient than others

The four major phases – based on the physical crime scene investigation process



1 – Preparation

- physically identifying the origin of the digital evidence
- choose the best approach to analyze it
- equipment seizure

2 – System Preservation

- goals
 - ✓ preserve the state of the digital crime scene
 - ✓ reduce the amount of evidence that may be lost
- actions vary depending on the legal, business, or operational requirements of the investigation
 - ✓ legal requirements may cause you to unplug the system and make a full copy of all data or,
 - ✓ could be a case involving a spyware infection or a honeypot and no preservation is performed
 - ✓ if it's not going to court, techniques in between can be used

Preservation Techniques

post-mortem analysis

- ✓ pull the plug to reduce the amount of evidence that is overwritten
- ✓ make duplicate copies of all data
- ✓ use write blockers to prevent evidence from being overwritten

live analysis

- ✓ kill or suspend suspect processes unplug or limit network connection
- ✓ use an empty hub or switch to prevent log messages about a dead link
- ✓ use network filters to avoid a remote connection from perpetrator to delete data
- ✓ backup important data (logs, files, etc)

Data integrity

- when important data are saved during a *post-mortem* or live analysis, a cryptographic hash should be calculated to later show that the data have not changed

Cryptographic hash algorithms



Data integrity – MD5 cryptographic hash

- this algorithm is broken since 2004
- use only for retro compatibility purposes
- it is possible to create collisions – different files with the same hash value examples:

<http://www.mscs.dal.ca/~selinger/md5collision/>

Demonstration

Data integrity – Hash values by itself are not enough

- given a message M , its hash value is $H(M) = h$
- someone can change both M and h , because h doesn't depend on a secret

Possible solution:

Digital Signatures

- depends on a private key
- better if done with a secure device,
 - e. g. the Portuguese Citizen Card (Cartão de Cidadão)

3 – Evidence searching

- goal: find data that support or refute hypotheses about an incident
- typically starts with a survey of common locations based on the type of incident:
 - ✓ Web-browsing habits: look at the Web browser cache, history file, and bookmarks
 - ✓ Linux intrusion: look for signs of a rootkit or new user accounts

It is important to look also for evidence that **refutes** your hypothesis instead of only looking for evidence that only supports your hypothesis.

The searching process:

1. define the general characteristics of the object for which we are searching
2. look for that object in a collection of data
3. two key steps:
 - determining for what we are looking
 - where we expect to find it

Example:
search all files with pictures

Search techniques:

- most searching for evidence is done in a file system and inside files
- search for files based on:
 - ✓ their names, or patterns in their names
 - ✓ a keyword in their content
 - ✓ temporal data, such as the last accessed or written time
 - ✓ hash values and compare them against a database
 - allows to find all files of a given type even if someone has changed their name
 - National Software Reference Library (NSRL) database <http://www.nsrl.nist.gov>
- analyzing network data based on:
 - ✓ packet headers, such as IP addresses, port number, protocol, . . .
 - ✓ keywords inside packets content

4 – Event reconstruction and report

- goal: try to answer questions about digital events in the system
- during the Evidence Searching Phase we might find several files that violate a law
 - but it doesn't answer questions about events
 - the file may have been the effect of an event, but what application downloaded it?
 - a web browser?
 - a malicious software? — **several cases have used malware as a defense**
- it may be possible to correlate the digital events with physical events

Event reconstruction requires knowledge about the applications and the OS that are installed on the system so that you can create hypotheses based on their capabilities

Examples:

- different versions of Windows OS (XP, 7, 8, 10) can cause different events
- different versions of Firefox, or Chrome Web browsers can cause different events

Exercises

Automate comparison of hash values

1. calculate the SHA256 values of all files inside a directory, e. g. C:\Windows or /etc and store the result in a file:

```
sha256sum * > SHA256.txt # works on Linux
```

2. verify the values:

```
sha256sum -c SHA256.txt # works on Linux
```

Tip for Windows OS:

hash calculation tool <https://www.slavasoft.com/hashcalc/>

Crack hash values

1. calculate the SHA256 of a common word, e. g. Baltazar:

```
echo-n "Baltazar" | sha256sum# this is a Linux command
```

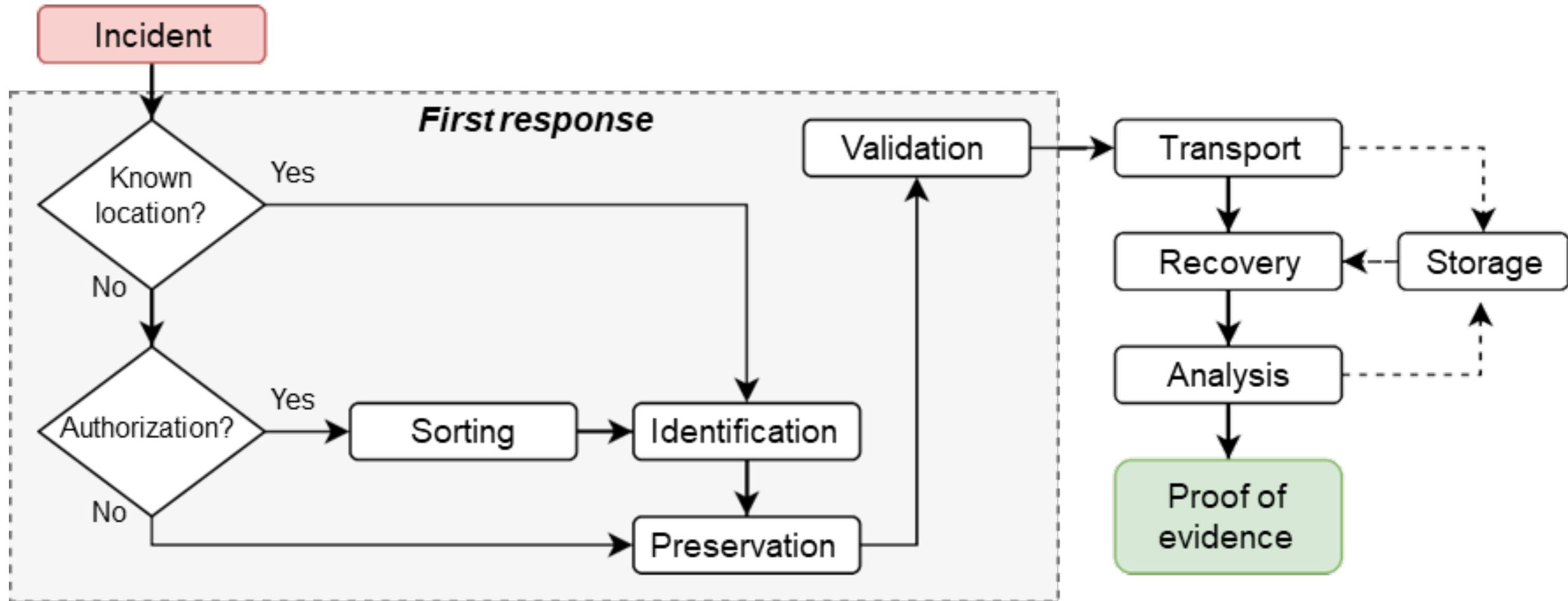
2. copy the hash value and paste it to crackstation.net or hashes.com

was the site able to find your word?

Digital Evidence Handling

Principles By EU-OLAF recommendation

1. The actions triggered by the first responders should not alter the data maintained on a computer or in a storage device that may be submitted to a court as evidence;
2. In exceptional circumstances, if it is considered necessary to access the original data stored on a computer or in a storage device, those who do so must have skills to be able to provide evidence, explaining the relevance and implications of his actions;
3. A chain of custody, or other record of all processes applied to digital evidence must be created and preserved. An independent third party should be able to examine these processes and obtain the same result;
4. The person responsible for the investigation must assume overall responsibility for compliance with the law and the present principles.



1 – Identification

Data states

- stored – data permanently stored in a storage device, e. g. an hard drive;
- in transit – data being sent through a local network, or Internet, to a reception device;
- in reception – data being received in a device, but not yet available to the user;
- in creation – data being locally produced and only partially available to the user;

Data sources

- storage devices – hard drives, SSD, USB drive, tapes, . . .
- temporary location – RAM, page files, swap partitions, cache files, . . .
- peripheral devices – printers, plotters, memory card readers, . . .
- active network devices – switches, routers, modems, print servers, . . .

1 – Identification

logical and physical location of the data

- local or remote devices
- dedicated storage systems (usually on data centers)
- computational systems (e. g. PC, laptop) has at least one storage device

main data types

- simple and human readable, e. g. photos, text documents, spread sheets
- complex and/or structured data, e. g. data bases, file system
- raw data, streams of data

2 – Preservation

do not modify any data that could have been evidence

- Copy important data, put the original in a safe place, and analyze the copy so that you can restore the original if the data is modified;
- Calculate hash values of important data so that you can later prove that the data has not changed – better yet if you do a digital signature;
- Use a write-blocking device during procedures that could write to the suspect data;
- Minimize the number of files created during a live analysis because they could overwrite evidence in unallocated space;
- Be careful when opening files on a live analysis because you could be modifying data, such as the last access time;

2 – Preservation: Prioritize the evidence to be collected

Order of volatility:

1. CPU, cache, and register content;
2. routing table, ARP cache, process table, kernel statistics
3. RAM
4. temporary file system, swap space
5. data on local storage media
6. remotely logged data
7. data contained on archival media (backups)

2 – Preservation: Sorting

Some times it is not possible to bring all devices due to several constrains: legal, time, technically unfeasible, . . .

Data sorting

In those situations a pre-analysis is required, but can only be performed if authorized accordingly to the country's laws.

3 – Isolate

isolate yourself from the suspect data because you do not know what it might do

- an executable from the suspect system could delete all files on your computer, or it could communicate with a remote system;
- opening an HTML file from the suspect system could cause your Web browser to execute scripts and download files from a remote server;

Create an isolated environment

- use virtual machines (VMware, VirtualBox, Xen, . . .)
- use an analysis network that is not connected to the outside world, or that is connected using a firewall that allows only limited connectivity
- isolation is very difficult, or impossible, with live analysis

4 – Correlate

correlate data with other independent sources

- helps reduce the risk of forged data
- timestamps can be easily changed in most systems
- find log entries, network traffic, or other events that can confirm the file activity times

This task is time consuming, specially if done without the help of software

5 – Log

log and document **all** of your actions

- helps identify what you have already done and what your results
- helps identify what searches you have not done yet
- in a live analysis, or performing techniques that will modify data, it is important to document what you do so that you can later document what changes in the system were made because of your actions

5 – Log: Identify devices:

- create tags to uniquely identify devices e. g. PC01, PC01.1, PC01.2, . . .
- take photographs
 - ✓ after placing tags
 - ✓ should be easy to read any relevant information, if needed take an overview photo and then a close up shot
 - computer brand, model, serial number, . . .
 - network cable connections, etc

Examples



Are there any problems with these photos?

5 – Log: Templates

Create templates with the required info to identify devices and services

Tag ID	PC01.HD03
Device	Hard disc drive, 2.5"
Brand	Seagate
Model	Momentus 5400.6 ST9250315AS
Serial number	5VC9CWTT
Interface	SATA
Capacity	250,0 GB
Type of intervention	Forensic copy
Working condition	Normal
Pictures	Yes, see Fig. 1 and 2
Observations	None

Tag ID	DNS01
Domain name	lotreur.com
Type of information	DNS history
Registrar	Center of Ukrainian Internet Names (UKR-NAMES)
Creation date	01-Mar-2013
Current state	expired
Registrant email	c152136@rmqkr.net
Annexes	Yes, see Annex A

5 – Log: Reception and tagging

- verify the list of devices delivered for analysis
- tag the devices, taking into account that one process may have:
 - ✓ one suspect with several devices,
 - ✓ or several suspects and many devices
 - ✓ there are processes with 30+ suspects and 200+ devices!



5 – Log: Tagging rules (as an example)

- letters to identify the owners of the devices in a given process
 - ✓ A, B, C, ...
- set of acronyms for each device type, followed by a 2 digits number
 - ✓ FPHxx, SPHxx, SIMxx, MCxx, GPSxx, CAMxx, PCxx, LAPxx, HDDxx, SSDxx, PENxx, ...
- a tag is composed by

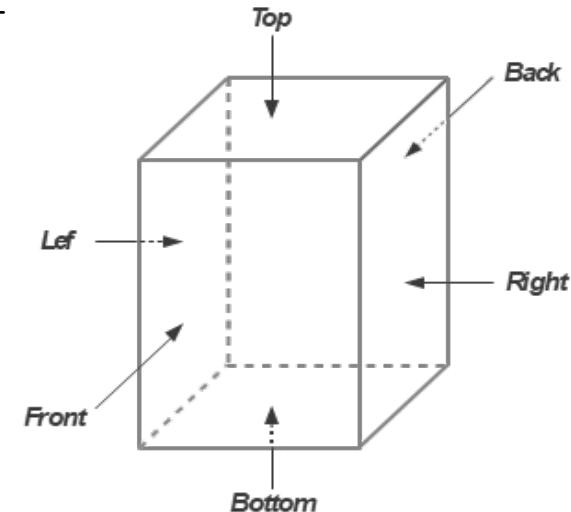
ownerID.deviceID[.inside deviceID]

example of a smartphone with 2 SIM cards and a memory card:

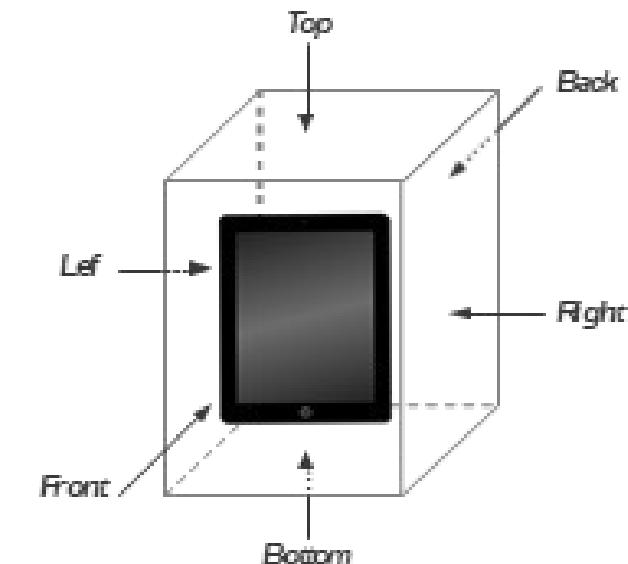
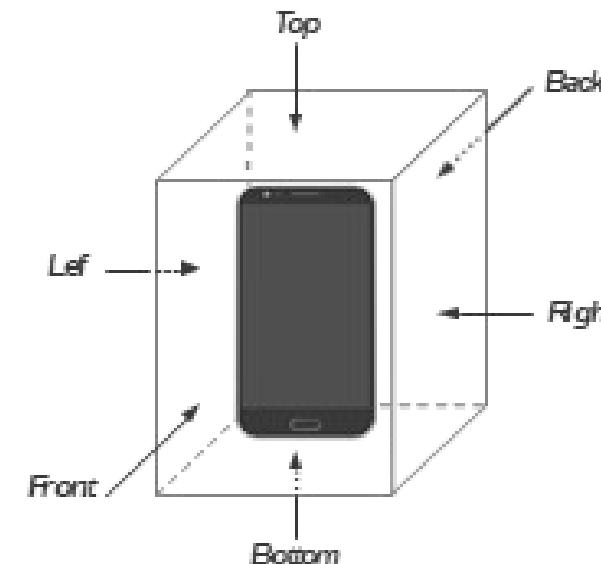
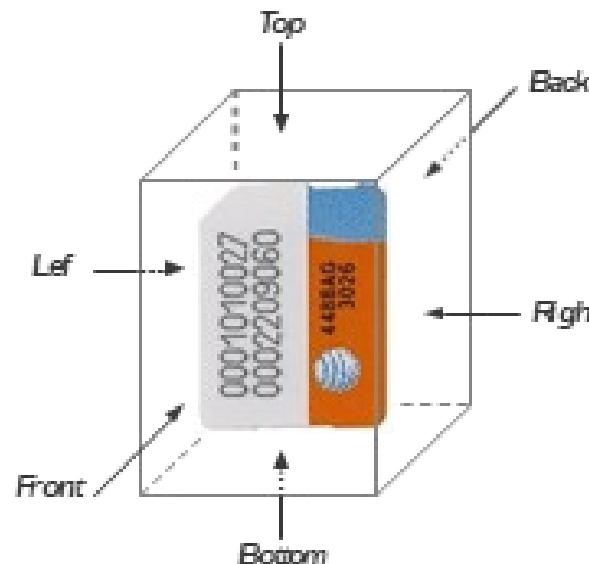
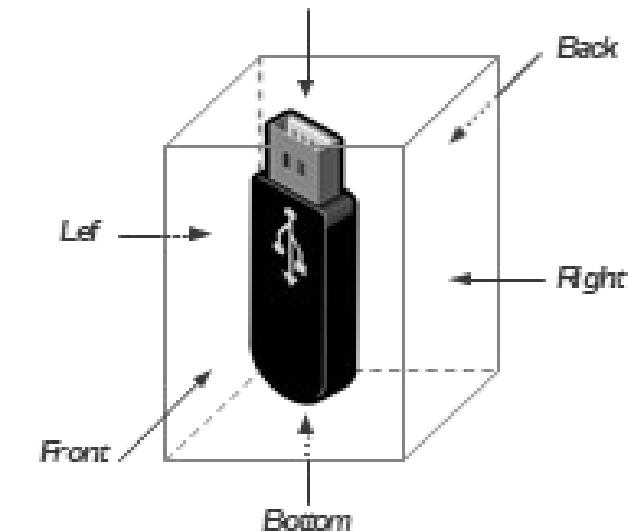
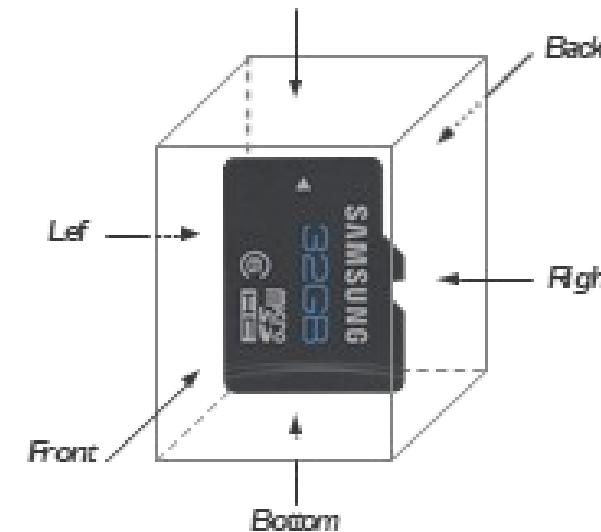
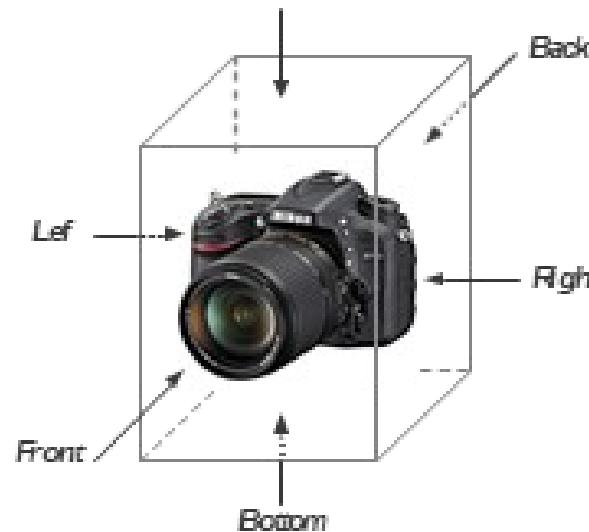
- ✓ A.SPH01 – the handset from owner A
- ✓ A.SPH01.SIM01 – first SIM card
- ✓ A.SPH01.SIM02 – second SIM card
- ✓ A.SPH01.MC01 – memory card

5 – Log: photograph rules (as an example)

- **include the scale of a ruler** in the photo
- photograph all important views
 - there are mandatory views by device type and some optional attention to details, like serial numbers, IMEI, etc
- apply the views names accordingly to the 3D box
 - position of the device is important
 - for some devices might be ambiguous which is the front, e. g. SIM card
- photos filenames: tagID-view name [-detail].jpg
 - examples:
 - A.SPH01-front.jpg,
 - A.SPH01-back-serial.jpg,
 - A.SPH01.MC01-front.jpg,



DIGITAL EVIDENCE HANDLING GUIDELINES



5 – Log: Catalog the device to **uniquely** identify it,

e. g.:

- tag ID device type
- brand and model number
- serial number, IMEI, UICCID, . . .
- type of intervention (logical or physical acquisition)
- device's condition (working / non-working)
- contents, e. g. has SIM or memory card
- worthy observations
- etc

Tag ID	A.SPH01
Device type	Smartphone
Brand	Samsung GSM
Model	GT-S5310
Serial n.	RV1D737C8AT
IMEI	356 431 051 982 186
SIM card	2: A.SPH01.SIM01 and A.SPH01.SIM02
Memory card	1: A.SPH01.MC01
Photos	Fig. 1, 2 and 3
Condition	Working
Observations	Battery not working
Intervention	Logic acquisition

ETHICAL CODE

Intent of the Ethical Code

- necessary to protect the integrity of the digital investigation process
- there are several codes

Example: International Society of Forensic Computer Examiners (ISFCE)

<https://www.isfce.com/ethics2.htm>

A computer examiner will **always**:

- Demonstrate commitment and diligence in performance of assigned duties
- Demonstrate integrity in completing professional assignments
- Maintain the utmost objectivity in all forensic examinations and accurately present findings
- Conduct examinations based on established, validated procedures
- Abide by the highest moral and ethical standards and abide by this Code
- Testify truthfully in all matters before any board, court or proceeding
- Avoid any action that would knowingly present a conflict of interest
- Comply with all legal orders of the courts
- Thoroughly examine all evidence within the scope of the engagement

A computer Examiner will **never**:

- Withhold any relevant evidence
- Reveal any confidential matters or knowledge learned in an examination without an order from a court of competent jurisdiction or with the express permission of the client
- Express an opinion on the guilt or innocence of any party
- Engage in any unethical or illegal conduct
- Knowingly undertake an assignment beyond his or her ability
- Misrepresent education, training or credentials
- Show bias or prejudice in findings or examinations
- Exceed authorization in conducting examinations

Exercises

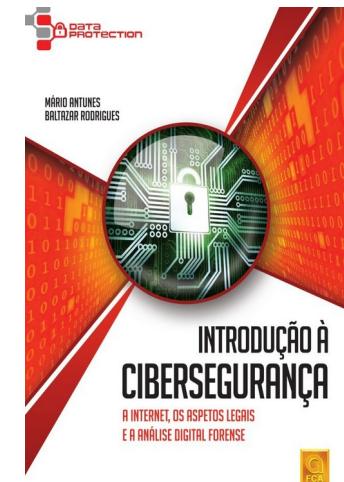
- Make a list of the information to be registered to uniquely identify these devices or services:
 - memory cards, computers, hard drives, solid state drives, phones, smartphones, GPS devices, routers, switches, modems
 - services: web, DNS, POP3, IMAP, SMTP, SSH
- Create a tagging system
 - easy to memorize, that reflects hierarchy if needed (1 PC with several hard drives)
 - e. g. computer: PC01, HD inside computer: PC01.1, ...
- Write your findings on a document
 - create a table template for each device or service
 - add real photos for the devices: use your phones, your own PCs, ... (don't forget the ruler)

This information will be used for the report of the Team Project

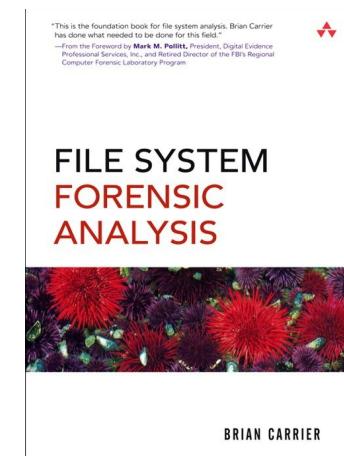
ANY QUESTIONS?

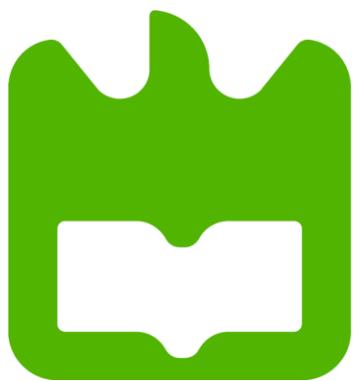


Antunes, M. & Rodrigues, B. (2018) Introdução à Cibersegurança: A Internet, os Aspetos Legais e a Análise Digital Forense. FCA (ISBN-13: 978-9727228614)



Carrier, B. (2005). File system forensic analysis.
Addison-Wesley Professional (ISBN-13: 978-0321268174)





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Obtaining Evidences

Artur Varanda

School Year 2021-2022

- It is normal forensic practice to remove a hard drive from a computer, write-block it and then image that hard drive
- But sometimes that is not possible:
 - ✓ some thin laptops have SSD chips soldered to the motherboard
 - ✓ the storage device has a non standard data interface and the examiner doesn't have the appropriate connector:
 - in these cases the imaging of the storage device needs to be done with the drive connected to the computer;

Use a forensic boot device on the computer:

- boot diskette, bootable CD-ROM/DVD, or bootable USB device to ensure the storage drive is not altered either during the boot or the acquisition phase.

The normal startup of a computer alters data on the primary storage drive during the boot process

- it is required to protect the integrity of the original evidence
- we must modify the start-up process in order to prevent any changes to the data on the storage drive

Boot process

- the normal boot process begins within the computer's hardware and moves to the boot device
- there are no changes made until the computer transfers control to the boot device

Boot process steps

- most systems have 2 phases:
 1. configure and start the hardware
 2. find the operating system and run it
- boot code – machine instructions used by the computer
- when it is starting when power is applied to a CPU:
 - ✓ it reads instructions from a specific location in memory – typically ROM
 - ✓ the instructions in ROM force the system to probe for and configure hardware
 - ✓ then searches for a device that may contain additional boot code
 - disks reserve space for boot code, but it isn't always used
 - its boot code is executed, and attempts to locate and load an operating system
 - the process after the bootable disk is found is platform-specific

Boot code characteristics

- has a specific location
- the instructions are in machine code

`0xB400 // machine code`

`MOV AH,00 // machine code representation in Assembly`

on a storage device it is difficult to distinguish random data from machine code

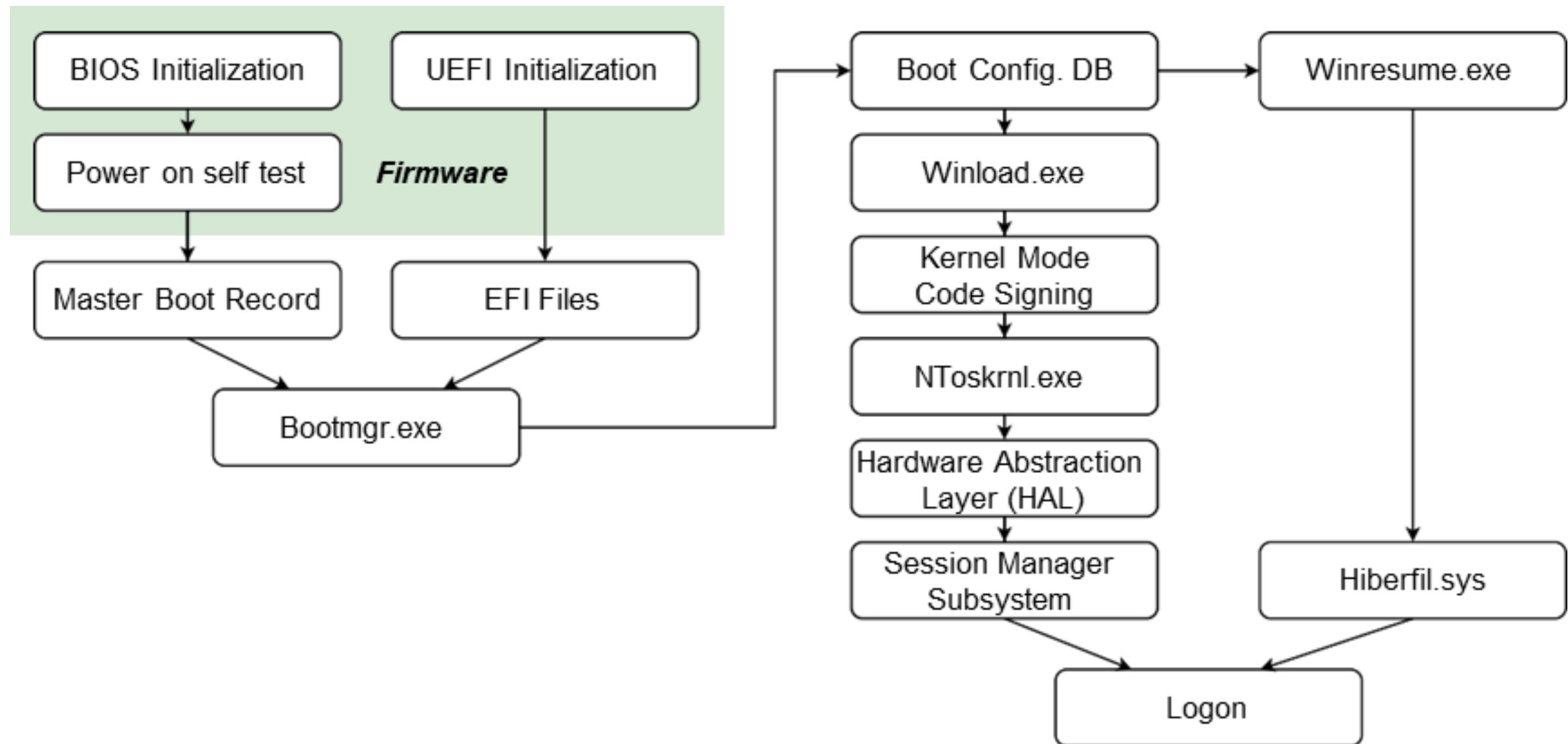
Systems with BIOS (Basic Input/Output System):

- BIOS boots by reading the first sector on a hard disk and executing it which has space limitations
- this boot sector code in turn locates and runs additional code in the first sector of the partition
- which locates and loads the actual operating system

UEFI (Unified Extensible Firmware Interface)

- boots by loading EFI program files (with .efi filename extensions)
 - ✓ stored in a special disk partition, known as the EFI System Partition (ESP)
- it can read files from a partition on the hard disk
- it is not limited to the size of the first sector
- it allows booting OS on disks with more than 2TB
 - ✓ regardless of the CPU architecture
- it supports graphics user interfaces on startup
- **secure boot** – is a feature of UEFI
 - ✓ the boot code must be digitally signed to prevent the installation of malware in the boot code
 - ✓ the examiner must use a forensic boot device that supports secure boot

WINDOWS BOOT PROCESS



Forensic Boot Tools

DOS boot disk (obsolete, but sometimes required)

- there are three files required to boot a computer into MS-DOS:

O.SYS, MSDOS.SYS and COMMAND.COM

- if present are also used in the boot process:

DRVSPACE.BIN or DBLSPACE.BIN, CONFIG.SYS and AUTOEXEC.BAT

How to create a **forensic** bootable diskette:

- on the command line of Windows 98: format a: /U /S

/U unconditional format

/S copy the necessary system files over to the diskette, in order to make it a boot disk

- then remove every file from the diskette except the mandatory three (see above)

- remove special attributes from the files to be deleted: attrib -H -R -S filename

later, it is possible to customize the forensic boot disk by adding CONFIG.SYS and AUTOEXEC.BAT files write-blocking utilities and other forensic tools

If you don't have a Windows 98 running

- HP makes an easy to use utility called HP USB Disk Format Tool, which includes a “Create a DOS Startup Disk” option
 - ✓ It's available for free download at <http://www.19systems.net/HP-USB-Tool-v2.1.8.exe> along with the Windows 98/DOS boot files <http://www.19systems.net/Win98-Boot-Files.zip>
- once the bootable diskette is created follow the same procedures to make it “forensic”:
 - ✓ remove every file from the diskette except the mandatory three O.SYS, MSDOS.SYS and COMMAND.COM later, it is possible to customize the forensic boot disk by adding CONFIG.SYS and AUTOEXEC.BAT files write-blocking utilities and other forensic tools

There are many Linux based bootable CD-ROMs (or Live CDs) with forensic tools, such as:

- Paladin (www.sumuri.com) – linux distro with many forensic GUI tools, including Autopsy
- DEFT (Digital Evidence & Forensic Toolkit <https://www.deftlinux.it/index.html>) is a customized distribution of the Ubuntu live Linux CD
- Caine (Computer Aided INvestigative Environment <https://www.caine-live.net>) is an **Italian** GNU/Linux live distribution created as a Digital Forensics project
- Kali Linux (<https://www.kali.org>) is an open-source, Debian-based Linux distribution geared towards various information security tasks, such as Penetration Testing, Security Research, Computer Forensics and Reverse Engineering

Create a bootable CD-ROM:

- find and download the ISO file, e.g. Paladin-6.1.iso
- use a CD-burning program to write the ISO file to the CD

Linux based bootable CD-ROM disadvantage

Not all first responders are comfortable using Linux

There are several Windows based bootable CD-ROMs (or Live CDs):

- WinFE (Windows Forensic Environment) created by Brett Shavers – Free
- System Acquisition Forensic Environment (SAFE) Boot Disk by Forensic Soft – Commercial
- Gargoyle Investigator by Wetstone – Commercial

WinFE (<https://www.winfe.net>)

Advantages

- it's free, but requires a Windows license
- runs windows software, *e. g.* FTK Imager, RegRipper, ... can be customized

Disadvantages

- requires configuration on the part of the user prior to use
- must be built to customize with your own set of tools

Nowadays, most computers don't have CD/DVD drives.

Tools to create a bootable USB device:

on Windows

- UNetBootIn (<https://unetbootin.github.io/>)
- Rufus (<http://rufus.akeo.ie/>)

on Linux

- Gnome Multi-Writer (<https://gitlab.gnome.org/GNOME/gnome-multi-writer>)
- Etcher – USB and SD Card Writer (<https://etcher.download/>)
- UnetBootIn for Linux (https://unetbootin.github.io/linux_download.html)
- DD command line tool: `sudo dd bs=4M if=input.iso of=/dev/sdx conv=fdatasync` (replace `sdx` by the drive letter of the USB device)

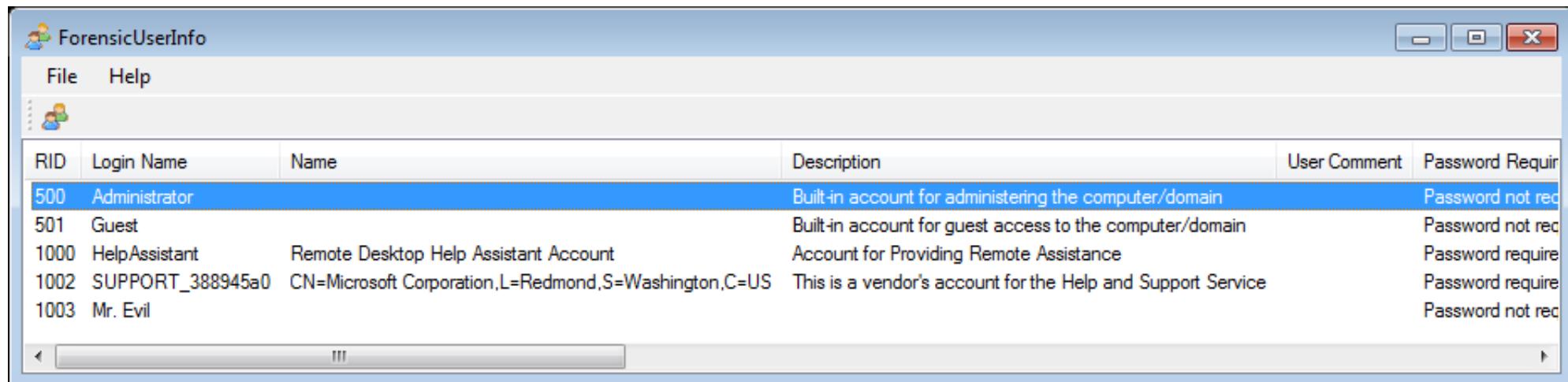
Forensic Sorting Tools

RegistryReport

- homepage: <https://www.gaijin.at/en/files?dir=old-software>
- requires the SAM, SOFTWARE, SYSTEM and NTUSER.DAT registry files
- doesn't process the registry files of the running operating system
- shows information about (Windows 2000 or higher)
 - ✓ the operating system
 - ✓ installed software
 - ✓ the last user activity
 - ✓ the user settings
 - ✓ and many other details
- the amount of information for each category can be configured in the settings dialog
- it allows to save, print and search the generated report

ForensicUserInfo

- homepage: <https://github.com/woanware/ForensicUserInfo>
- requires the SAM, SOFTWARE and SYSTEM files
- extracts the following information:
 - ✓ RID, Login Name, Name, Description, User Comment
 - ✓ LM Hash, NT Hash
 - ✓ Last Login Date, Password Reset Date, Account Expiry Date, Login Fail Date
 - ✓ Login Count, Failed Logins, Profile Path, Groups

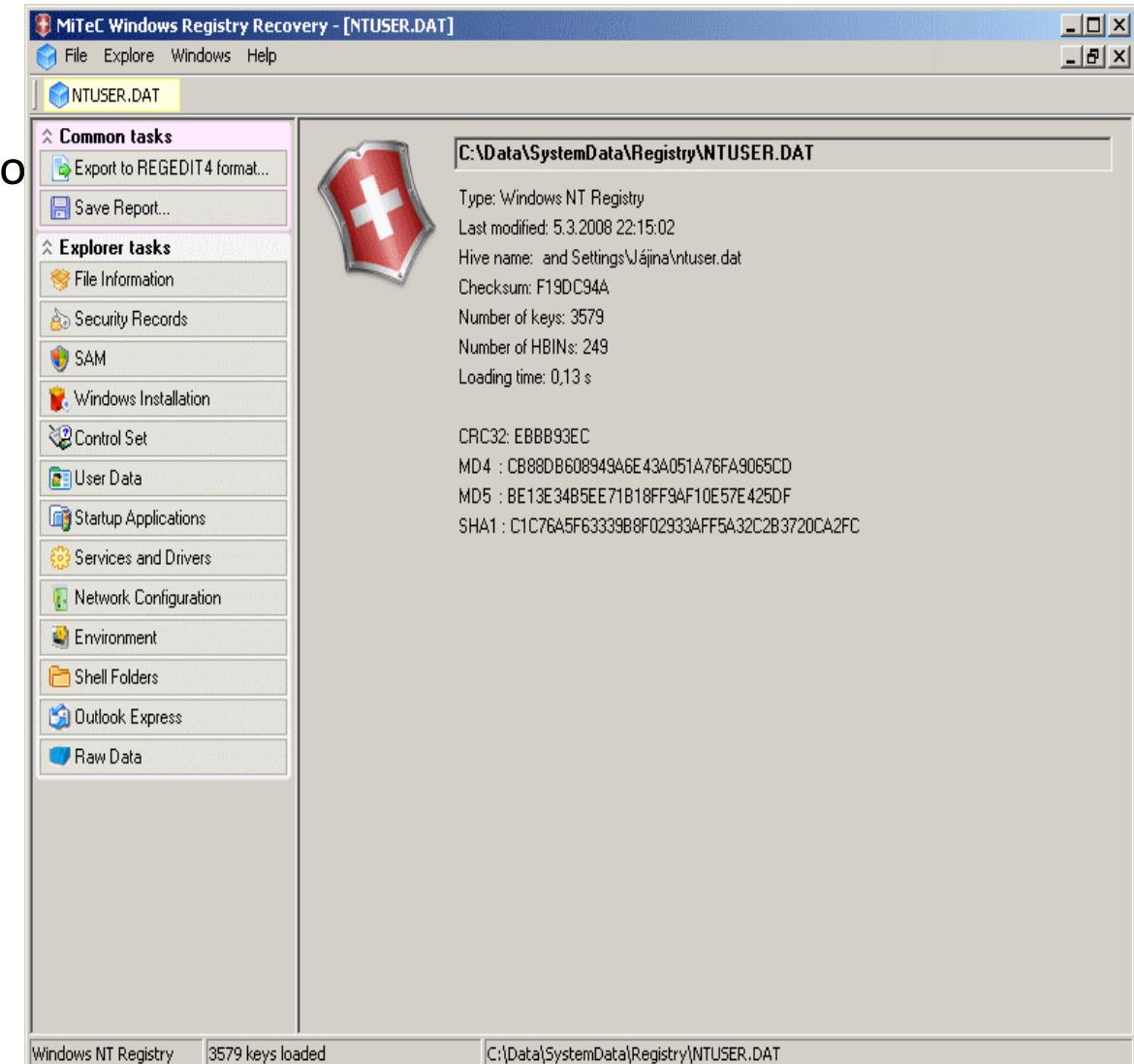


The screenshot shows a Windows application window titled "ForensicUserInfo". The window has a standard title bar with minimize, maximize, and close buttons. Below the title bar is a menu bar with "File" and "Help" options. The main area of the window contains a table with the following data:

RID	Login Name	Name	Description	User Comment	Password Requir
500	Administrator		Built-in account for administering the computer/domain		Password not req
501	Guest		Built-in account for guest access to the computer/domain		Password not req
1000	HelpAssistant	Remote Desktop Help Assistant Account	Account for Providing Remote Assistance		Password require
1002	SUPPORT_388945a0	CN=Microsoft Corporation,L=Redmond,S=Washington,C=US	This is a vendor's account for the Help and Support Service		Password require
1003	Mr. Evil				Password not req

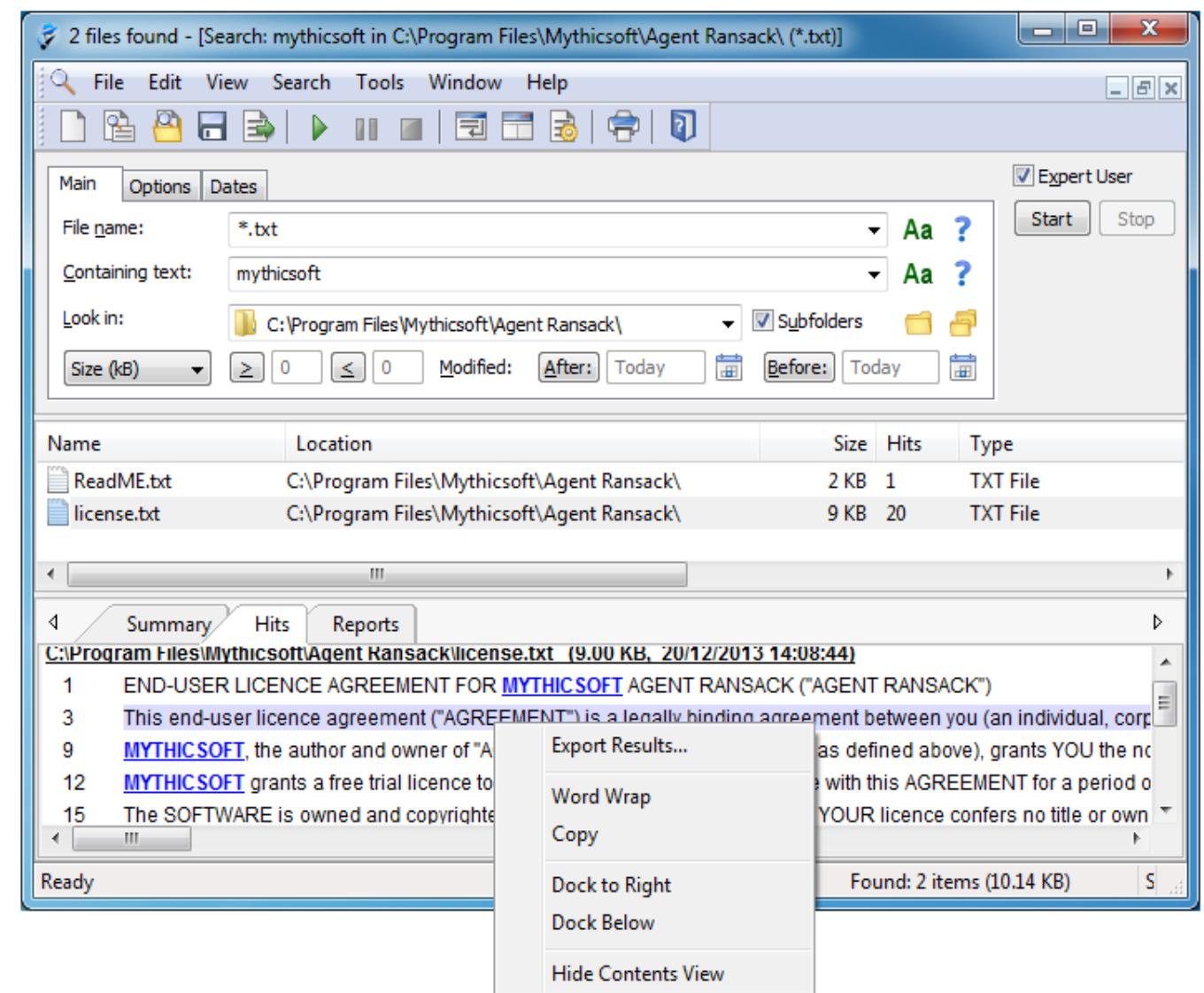
Mitec WRR (Windows Registry Recovery)

- homepage: <http://www.mitec.cz/wrr.html>
- for crashed machine, registry configuration, data recovery
- it allows to explore:
 - ✓ File Information
 - ✓ SAM
 - ✓ Security Record Explorer
 - ✓ Windows Installation
 - ✓ Hardware
 - ✓ User Data
 - ✓ Startup Applications
 - ✓ Services and Drivers
 - ✓ Network Configuration
 - ✓ Windows Firewall Settings
 - ✓ Environment,
 - ✓ Shell Folders
 - ✓ Outlook Express,
 - ✓ Raw Data



RanSack

- homepage:
<https://www.mythicsoft.com/agentransack/>
- free software program for finding files on your PC or network drives
 - ✓ fast search (less time waiting)
 - ✓ powerful search capabilities (Boolean expressions, Perl regex)
 - ✓ supports Microsoft Office and Libre Office files formats



Portable Forensic Tools

- collection of freeware tools, such as:
 - ✓ DataProtectionDecryptor – decrypts passwords of Microsoft Outlook accounts, credentials files of Windows, wireless network keys, passwords in some versions of Internet Explorer, passwords and cookies of Chrome Web browser
 - ✓ JumpListsView – displays the information stored by the 'Jump Lists'
 - ✓ Windows File Analyzer – decodes and analyzes to provide cached information
 - ✓ BinText – extracts strings from binary files
 - ✓ Data Converter – converts numbers, hexadecimal values or dates
 - ✓ EXIF Viewer – displays EXIF informations from JPEG images
 - ✓ eMule MET Viewer – shows various information from the eMule
 - ✓ ...
- to find more portable tools: <https://www.portablefreeware.com>

FTK Imager

- homepage: <https://www.exterro.com/ftk-imager>
- very powerful and user-friendly tool
 - ✓ runs as portable application, ideal to include in WinFE
 - ✓ search files
 - ✓ look for deleted files
 - ✓ copy files (*e. g.* cache and registry files)
 - ✓ identify ADS (Alternate Data Stream)
 - ✓ acquire storage devices and RAM
 - ✓ mount E01 files
 - ✓ ...

Forensic Acquisition

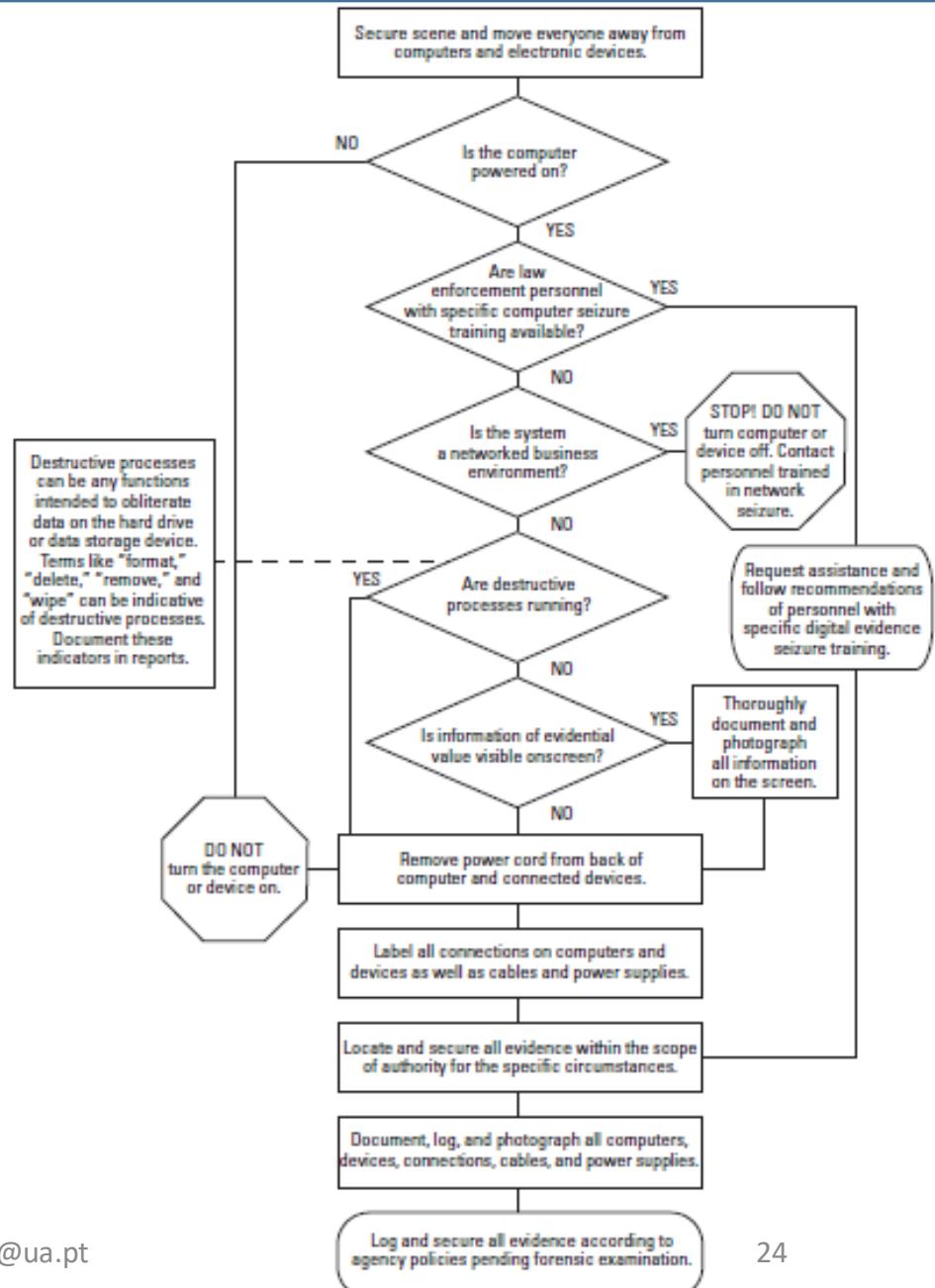
Data Acquisition

- typically occurs in the “system preservation” phase
- although it might also occur on a running system
- this is an import phase
 - ✓ if not done properly data can be lost forever
 - ✓ it must be done in a way that does not undermine its legal validity

INTRODUCTION

What to do if:

- the computer is off → remove power cord
- the computer is on:
 - ✓ take a picture of the screen
 - ✓ are destructive processes running? → remove power cord
 - ✓ do a memory dump and get network connections status → this may destroy or contaminate evidences
 - when you cannot turn off a server
 - to get passwords or encryption keys stored in RAM
 - to monitor malicious software network activities



Source: Electronic Crime Scene Investigation: A Guide for First Responders, Second Edition (U.S. Department of Justice)

Information analysis layers on storage media:

- physical – from the first to the last bit of the storage media
- volume – it is not possible to get unallocated sectors, partition table or hidden areas
- file – file copies (*e. g.* backup tools) less likely to retrieve deleted files
- application – each application has its own encoding or file format

The higher the acquisition layer, the less information can be retrieved

Whenever possible, data acquisition should be made at the physical level.

Other media:

- network and volatile memory
- each medium as its own recommend procedures

Copying storage media

- the bigger the block size, the faster the acquisition,
- but if there are sectors with errors, the all block will be invalid
- the acquisition block size should match the sector size
 - ✓ for HDD the sector size is 512 bytes
 - ✓ for SSD sector size depends on the brand, model and capacity
- data acquisition should include the complete storage medium (physical level)
 - ✓ including unallocated sectors, and
 - ✓ hidden areas: HPA or DCO – in this case 2 acquisitions are recommended
 - one with the hidden area in place, and another with the hidden areas disabled

Data acquisitions from storage media

make a storage medium forensic copy

- requires another storage medium of equal or bigger size, although many tools can create compressed images files

reading the data

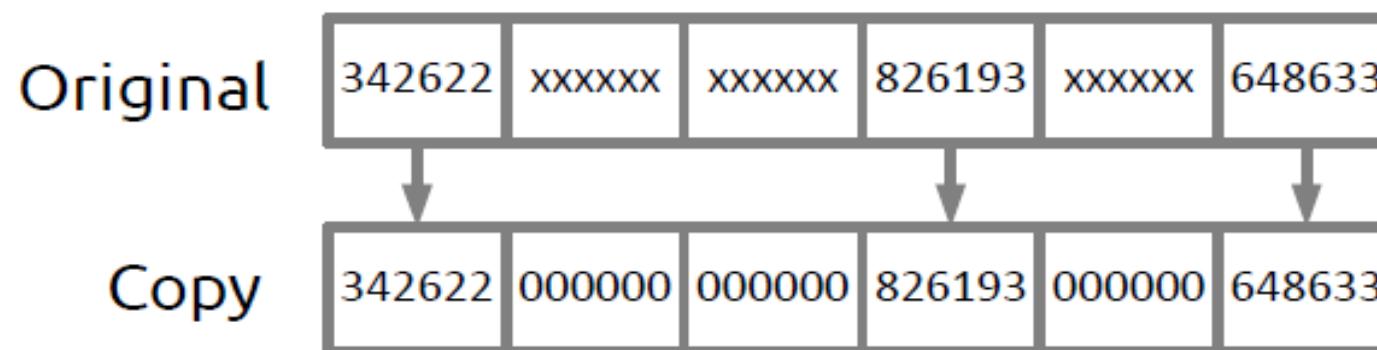
- through the BIOS – old BIOS don't support large storage drives → may report wrong drive size
- direct access – the best choice, but not supported by all tools

Post mortem versus alive data

- acquisition post mortem
 - ✓ the OS is shutdown
 - ✓ suspect hardware can be used using a trusted OS to boot it
 - **Caution:** new PCs boot too fast and we might not be able to change boot order
 - ✓ the NSA scandal showed that we cannot always trust the hardware
 - spyware inside HDDs' firmware
 - ✓ although it is less likely to happen than software tampering
- alive
 - ✓ the OS is running and used to perform the acquisition
 - there is the risk of the OS have been tampered and return wrong data
 - *e. g.* rootkits that hide processes and files to avoid detection
 - ✓ online acquisition should be performed only in special situations

What happens if the drive has bad sectors?

- acquisitions is still possible, if the percentage of bad sectors is small
- the tool must be able to deal with the errors:
 - ✓ place zeros on the bad sectors, so the data keeps its alignment
 - ✓ otherwise the forensic copy would be smaller and the analysis tool could trigger errors
 - ✓ the tool must register in a log file all the identified bad sectors
 - ✓ tools should automatically decrease the acquisition block size to the sector size



Host Protected Area (HPA)

- this area should be copied also, it may contain hidden data
- few tools support reading HPA
 - ✓ we can use the `hdparm` tool to temporarily access it (a reboot will restore the HPA)
 - ✓ as a precaution measure, we should make first one forensic copy with HPA in place

Device Configuration Overlay (DCO)

- the removal of the DCO is permanent, so as a precaution measure, we should make first one forensic copy with DCO in place
- few tools support reading DCO areas
 - list of tools is available in <https://forensicswiki.xyz>

Tableau Imager is able to identify and read both HPA and DCO

Write blockers stop any write operation to the storage media under investigation

- **Hardware**

- ✓ specific for each medium interface: ATA, SATA, SCSI, Firewire (IEEE 1394) or USB
- ✓ stops write operations regardless of the used OS
- ✓ specialized hardware provides better acquisition performance
- ✓ some hardware works like a proxy and monitors all operations
- ✓ this is the best option, but it is also the most expensive
- ✓ list of tests on hardware write blockers:
 - <https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt/cftt-technical/hardware>

- **Software**

- ✓ this is the less expensive solution
- ✓ but may be less effective, some apps can bypass the software write block
- ✓ list of tests on software write blockers:
 - <https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt/cftt-technical/software>

There are 2 main approaches to acquire

- data: cloning
 - ✓ it is recommended to use drives of the same size
 - ✓ if the clone driver is bigger, where the clone data ends?
 - ✓ it is highly recommended to zero out first the drive before cloning
 - ✓ drive geometry of the clone might be different
 - ✓ some OS, namely Windows, by default *auto mount* drives, so you need write blockers to analyze the clones
- imaging the drive – the most common approach
 - ✓ it is not vulnerable to *auto mount* by the OS
 - ✓ an image will always be mounted as read only, no need for a write blocker
 - ✓ it is possible to simulate read/write operation
 - the changes will be stored in a cache file, leaving the original intact
 - ✓ this way one drive can store image files from several different media
 - ✓ the image file can be split into smaller files to fit in a DVD

There are several image formats for acquired data:

- *raw image* – most flexible format, supported by all analysis software
- *raw image + external metadata* – like the raw image, but adds another file with description data, hash values and time
- *embedded image* – proprietary format, the metadata is embedded inside the image
- some image file formats support compression
 - ✓ save storage space, but the acquisition process takes longer to complete
 - ✓ not all tools support compressed image files – it might be required to uncompress first
 - ✓ good solution for long term storage of the image files

Acquisition

- local acquisition – implies physical access to the storage drive
- remote acquisition
 - ✓ when it is not possible to have physical access to the storage drive
 - ✓ when there are no adequate adapters for the storage medium
 - ✓ this process is slower, usage of compression is recommended
 - ✓ if there are no full control of the network encryption should be used

To guarantee integrity, hash values should be stored

- hash block of small size to prevent a single error to invalidate the all drive image
 - ✓ e. g. the same size of the acquisition block
- in a RAW file hash values are stored in a separate file
- **I recommend to do a digital signature on the hash values files**
 - ✓ Why do you think this is good practice?

Digital signatures

- best tool is GnuPG - <https://www.gnupg.org/>
- if possible, use also a time stamping server

Tools for data acquisition on storage media there are many tools

- Windows – with graphical user interface
 - ✓ FTK imager, Tableau imager, ... (<https://www.exterro.com/ftk-imager>)
- Linux – many are command line
 - ✓ GUI – Guymager (<https://guymager.sourceforge.io/>)
 - ✓ CMD – ewftools, dd and derived tools: dcfldd, dc3dd, ddrescue, dd rescue, rdd, ...

Computer Forensic Tool Testing (CFTT) project

- develops test-cases for digital forensic tools
- tests the tools and publishes the results

<https://www.nist.gov/itl/ssd/software-quality-group/computer-forensics-tool-testing-program-cftt>

Expert Witness Format (EWF)

- proprietary format from EnCase tool, supports compression
 - ✓ file extension: E01, E02, ...
- there are also open-source tools: `apt-get install ewf-tools`
 - ✓ `ewfacquire` – acquire drive data locally
 - ✓ `ewfacquirestream` – acquire drive data remotely
 - ✓ `ewfmount` – mount EWF images (will be mounted as a RAW file)
 - ✓ `ewfexport` – convert image file formats
 - ✓ `ewfinfo` – get info from a an EWF image file
 - ✓ `ewfrecover` – tries to recover corrupted EWF files
 - ✓ `ewfverify` – validate EWF file integrity → **very important** before start an analysis

```
sudo apt install ewf-tools          # Install EWF tools  
ewfexport -f raw filename.E??    # EnCase → RAW
```

Acquisition example:

```
ewfacquire /dev/sdd                                # issued command line
ewfacquire 20130416                                # info generated by the tool
```

Device information:

```
Bus type:                                         ATA
Vendor:                                           VMware Virtual I
Model:                                            00000000000000000000000000000001
Serial:
```

Storage media information:

```
Type:                                              Device
Media type:                                       Fixed
Media size:                                        53 MB (53477376 bytes)
Bytes per sector:                                  512
```

Acquiry parameters required, please provide the necessary input

```
Image path and filename without extension: example      # interactive part
```

```
Case number: 001
```

```
Description: This is just a test
```

```
Evidence number: 000001
```

```
Examiner name: Miguel Fraude                         # who did the acquisition
```

```
Notes: Visrtual disc drive
```

```
(...)
```

```
Compression level (none, empty-block, fast, best) [none]: best # compress option
```

```
(...)
```

Example

- download EWF files:
<http://downloads.digitalcorpora.org/corpora/drives/nps-2008-m57-jean/nps-2008-jean.E01>
<http://downloads.digitalcorpora.org/corpora/drives/nps-2008-m57-jean/nps-2008-jean.E02>
- get info: `ewfinfo nps-2008-jean.E0*`
- mount the EWF image is a 2-step process:
 - ✓ mount the EWF to be seen as RAW disk
 - ✓ mount the raw disk
- Tools:
 - ✓ Windows: FTK Imager (very simple)
 - ✓ Linux has several options:
 - `ewfmount` or `xmount` + `mount` → requires `sudo`, but allows to see MFT as a file
 - `xmount` + `udisksctl` → the safest option, can be done by a regular user

Mount EWF image with ewfmount on Linux

1. sudo ewfmount nps-2008-jean.E?? /mnt/raw

- maps as a RAW file
- read only, it is not possible to emulate write operations

2. sudo mount -t ntfs -o

ro,loop,show sys files,streams interface=windows,offset=\$[63*512]
/mnt/raw/ewf1/mnt/loop/

- mounts file system in the RAW image
- show sys files – allows to see NFTS structures as files
- streams interface=windows – allows access to *Alternate Data Streams (ADS)* data
- offset=\$[63*512] – beginning of the partition to mount (in bytes)

Mount EWF with xmount on Linux (without sudo)

1. `xmount --in ewf nps-2008-jean.E?? --out raw --cache cachefile /mnt`

- maps EWF as a RAW file
- xmount emulates write operations using a cache file

2. `udisksctl loop-setup -f /mnt/nps-2008-jean.dd`

- creates a loop device for each partition
- `ls /dev/loop0*` → check how many partitions were identified
- this command requires the user to belong to the fuse and disk groups:

- ✓ `sudo usermod -a -G fuse username`
- ✓ `sudo usermod -a -G disk username`

3. `udisksctl mount -b /dev/loop0p1`

- mount the first partition

dcfldd

- developed by *Department of Defense Computer Forensics Lab (DCFL)*
- forensic tool derivate from the `dd` command line tool
- differences with `dd`
 - ✓ calculates the hash values and supports md5, sha1, sha256 and sha512
 - ✓ can write the identified bad sectors to a separate file
 - ✓ can aggregate bad sectors errors Had 1,023 'Input/ouput errors' between blocks 17–233'
 - ✓ checks the hash values
 - ✓ reports the progress
 - ✓ allows to split the image file in smaller files
 - ✓ to ensure reproducibility bad sectors are written with zeros in the image file
- there are many forensic tools derivate from `dd`
 - ✓ `dc3dd` (very similar), `ddrescue`, `dd rescue`, `rdd`, ...

To get more info about the dcfldd tool

- dcfldd --version → installed version
- dcfldd --help → list all the supported options
- man dcfldd → man page

I couldn't find the official homepage of this tool, but you can get more info:

<https://forensicswiki.xyz/wiki/index.php?title=Dcfldd>

complement with info from [https://en.wikipedia.org/wiki/Dd_\(Unix\)](https://en.wikipedia.org/wiki/Dd_(Unix))

```
dcfldd if=/dev/sourcedrive hash=md5,sha256 hashwindow=1G md5log=md5.txt  
sha256log=sha256.txt hashconv=after bs=512 conv=noerror,sync split=10G  
splitformat=aa of=image.dd
```

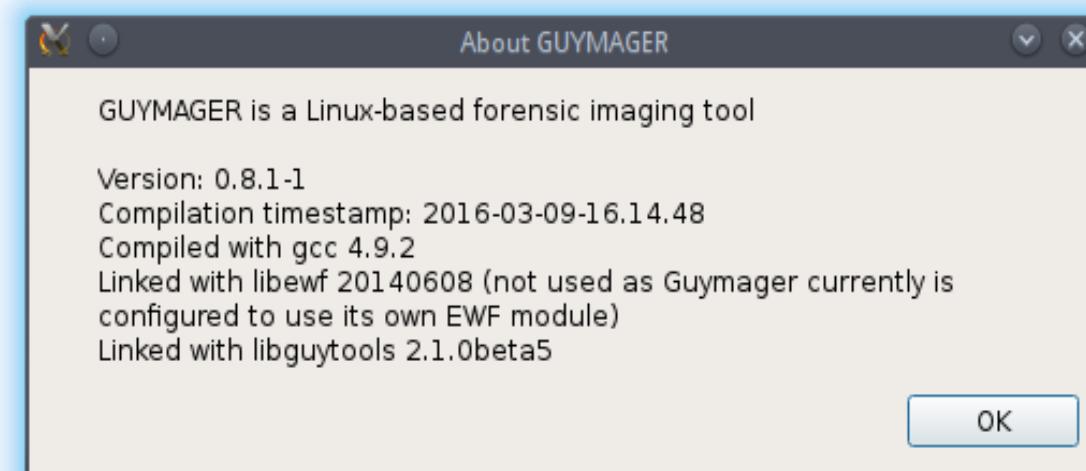
- **if=/dev/sourcedrive** → file that represents the drive to acquire
- **hash=md5,sha256** → request md5 and sha256 hash values
- **hashwindow=1G** → calculate hash values at each 1 GB
- **md5log=md5.txt sha256log=sha256.txt** → files name to store hash values
- **hashconv=after** → calculate hash values after error checking
- **bs=512** → use block size of 512 bytes
- **conv=noerror, sync** → noerror – doesn't stop in case of reading errors, sync – keeps image synced when errors show up
- **split=10G** → split RAW image into 10 GB files
- **of=image.dd** → base filename
- **splitformat=aa** → format of individual filenames image.dd.aa,image.dd.ab,...

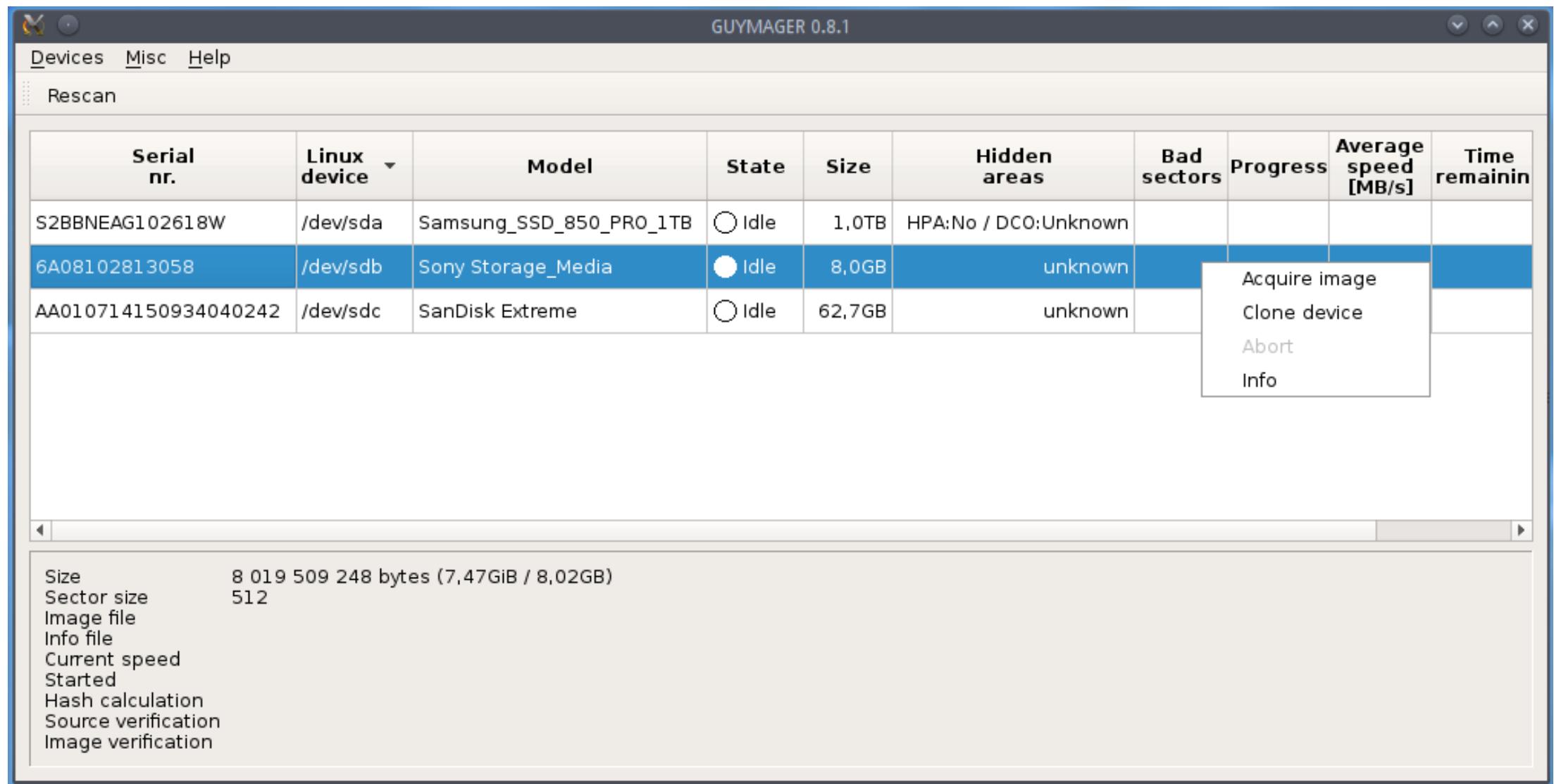
```
cat SHA256.txt
```

```
0 - 1048576:1756ad07245b68744644fa147bbed4dbf5b148ba61839d01e7421ff20098c681  
1048576 - 2097152:908348ee7e44372531d1311143d9ef3a2829fe30f93831b5a81e450a9d366168  
2097152 - 3145728:30e14955ebf1352266dc2ff8067e68104607e750abb9d3b36582b8af909fcb58  
3145728 - 4194304:30e14955ebf1352266dc2ff8067e68104607e750abb9d3b36582b8af909fcb58  
4194304 - 5242880:30e14955ebf1352266dc2ff8067e68104607e750abb9d3b36582b8af909fcb58  
5242880 - 6291456:30e14955ebf1352266dc2ff8067e68104607e750abb9d3b36582b8af909fcb58  
...  
52428800 - 53477376:30e14955ebf1352266dc2ff8067e68104607e750abb9d3b36582b8af909fcb58  
Total (sha256):5767d9dcd2e48b3a0dce1b9a143ecf7a664364660637c5f348d1054afb4a1784
```

Guymager <https://guymager.sourceforge.io>

- very user friendly
- supports RAW, EWF and AFF file formats
- faster than known commercial imagers running under Windows.
- does not support logical acquisitions:
 - ✓ /etc/guymager/guymager.cfg – default configuration file, do not change!
 - ✓ /etc/guymager/local.cfg – do all your configuration in this file, e. g. EwfCompression=BEST





Recommended option for Linux

Paladin Toolbox <https://sumuri.com/software/paladin>

- included in the live Linux Paladin from Sumuri (forensics distro)
- free, but not open source and requires registration to download
- supports logical acquisitions, which is the recommend option for SSDs

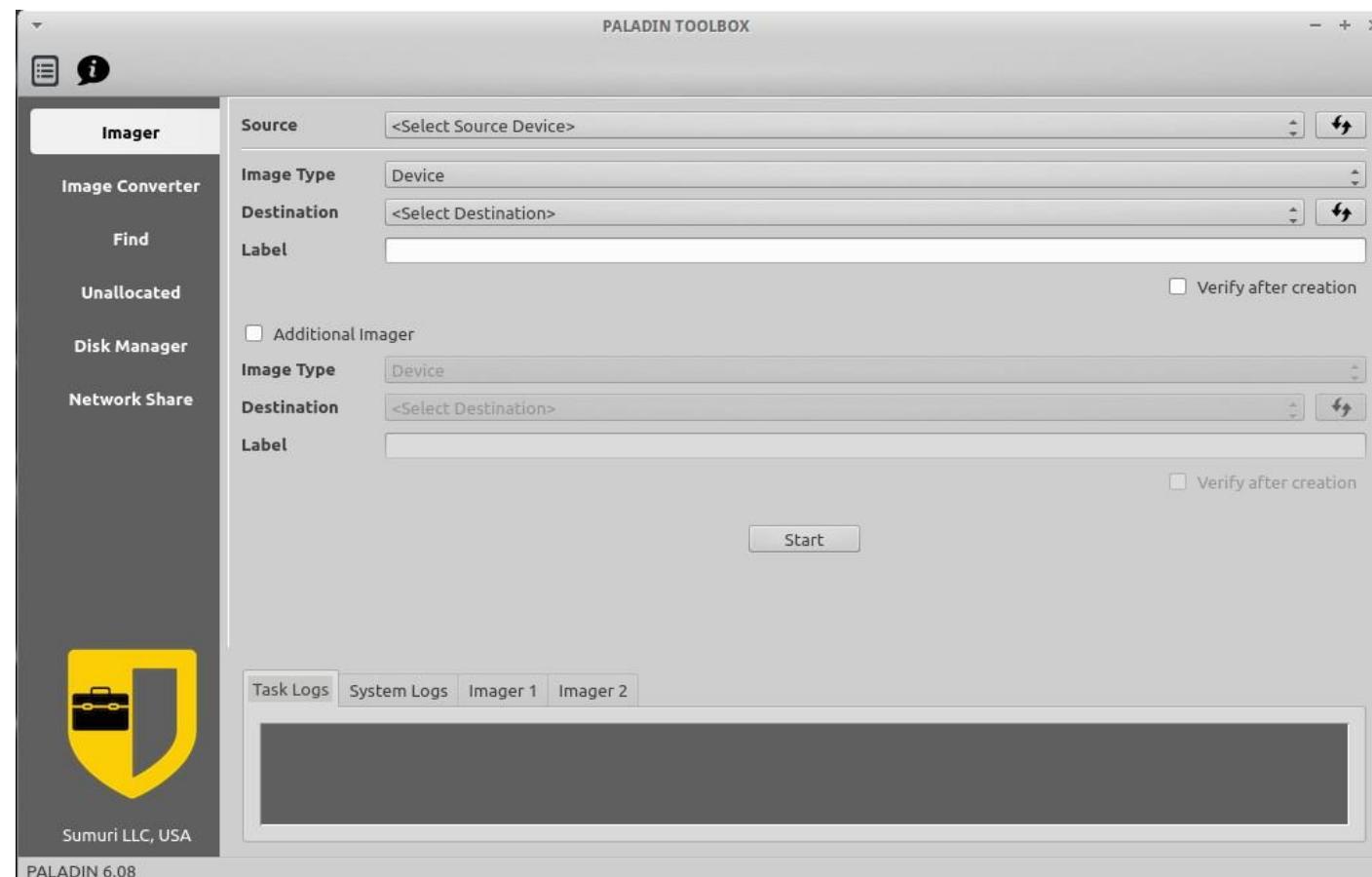
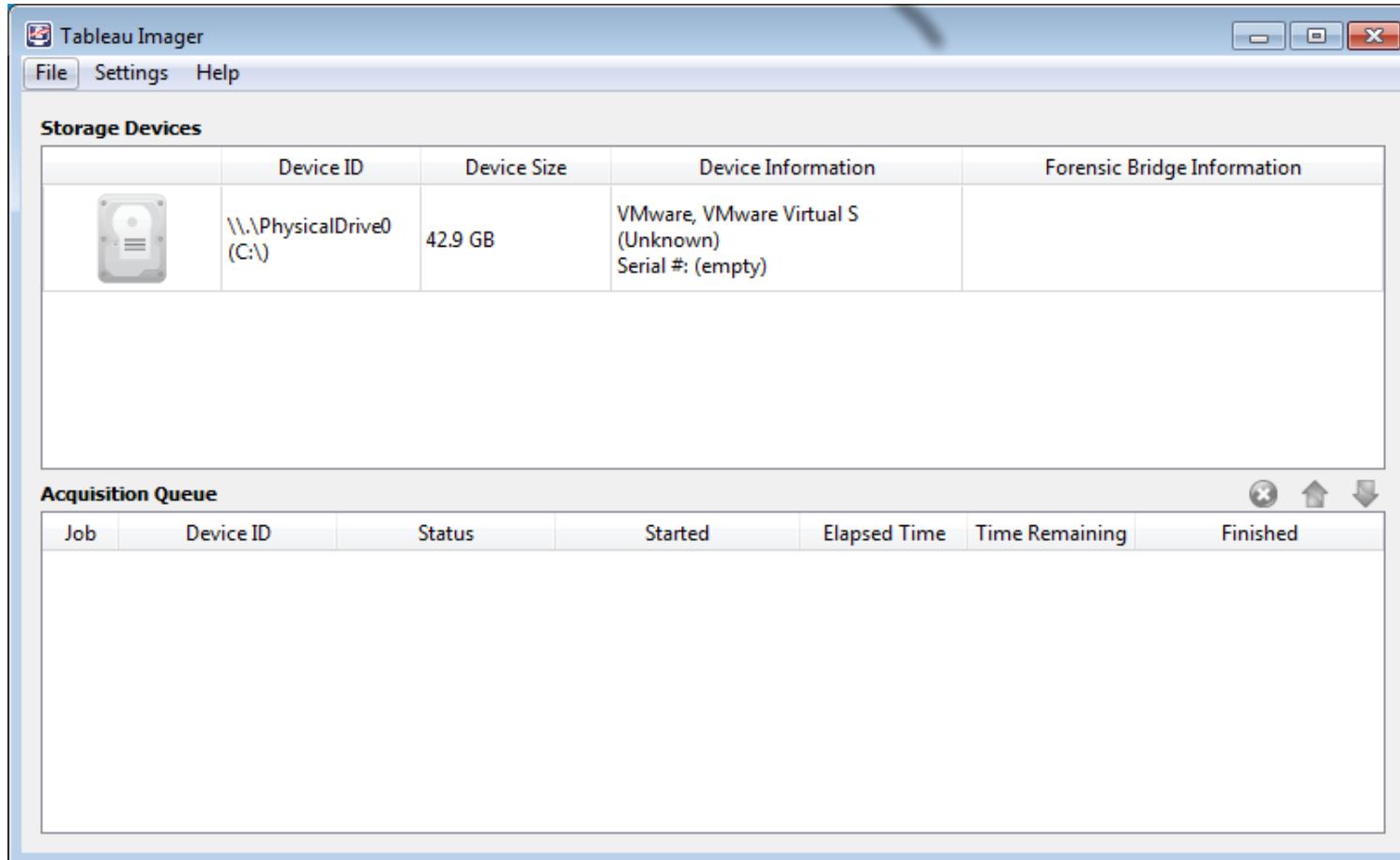


Tableau Imager <https://security.opentext.com/tableau/hardware/details/tx1>

- works only with Tableau write blockers
- free, but not open source and requires registration to download
- does not support logical acquisitions, which is the recommend option for SSDs

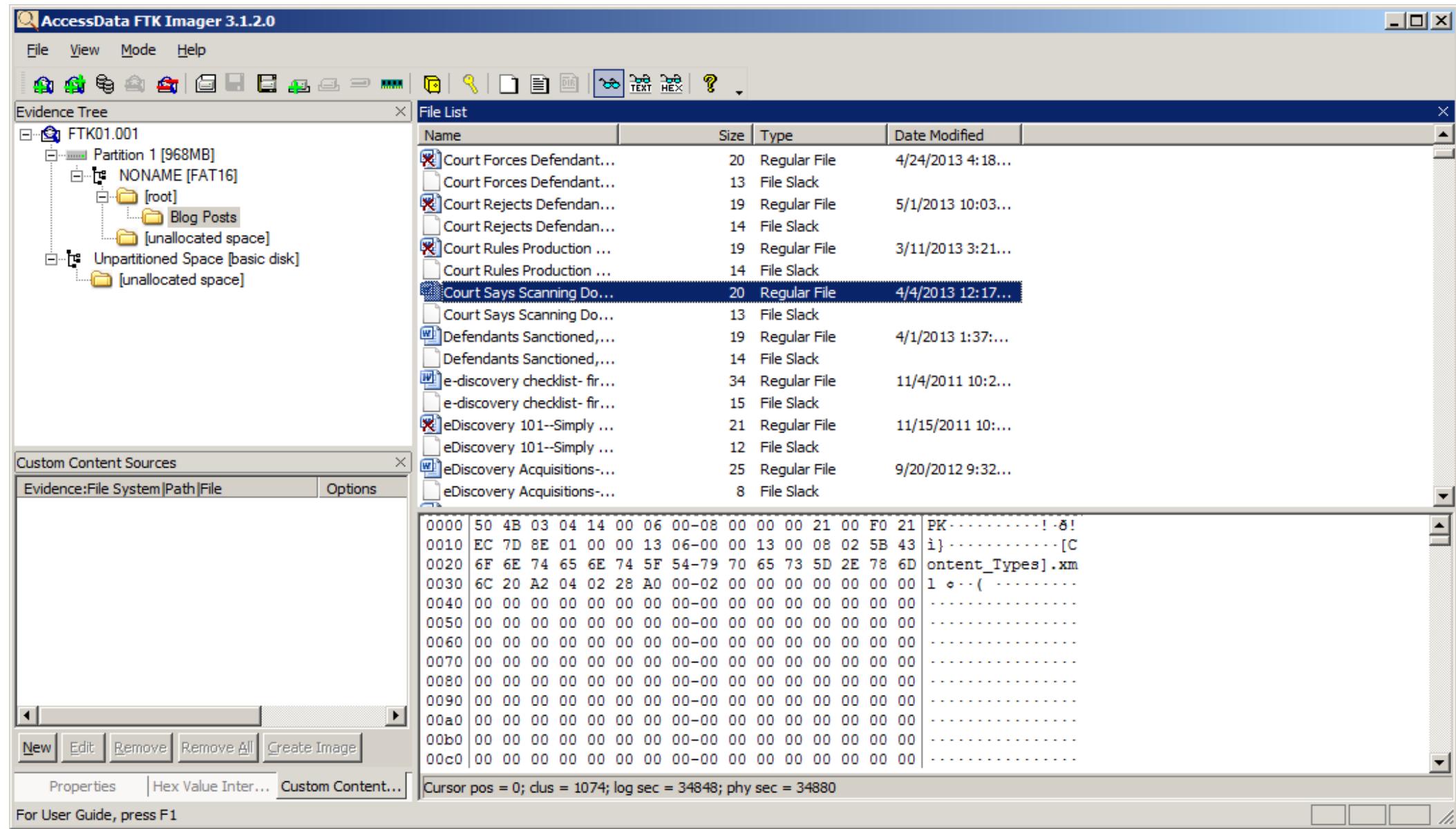


FTK Imager <https://accessdata.com/product-download/ftk-imager-version-4-5>

- Create forensic images of local hard drives, CDs and DVDs, thumb drives or other USB devices, entire folders, or individual files from various places within the media
- Preview files and folders
- Preview the contents of forensic images
- Mount an image for a read-only view
- Export files and folders from forensic images
- See and recover files that have been deleted from the Recycle Bin
- Create hashes of files to check the integrity of the
- Generate hash reports for regular files and disk images
- Lite version runs as portable application

Recommended option for windows

WINDOWS – FTK IMAGER



- Lab 01 - Build WinFE and boot a virtual machine with it

- Lab 02 – Create a Forensic Image

1. Without extracting the virtual machine assigned to your team, add it as an evidence source to FTK Imager running on your computer
2. Create a .e01 forensic image from the virtual machine with FTK Imager
 - ✓ make sure you have enough disk space for the acquisition
 - ✓ split into 4 096 MB files to fit into FAT32 file systems if needed
 - ✓ enable compression, it's slower, but takes less space
 - ✓ validate the forensic image





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

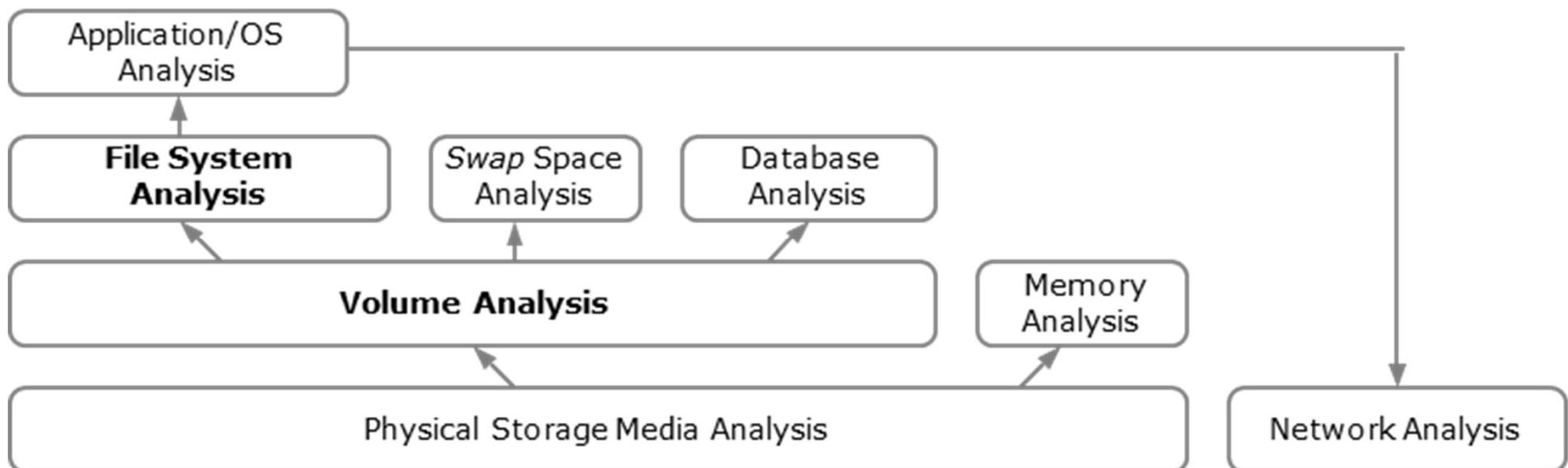
Data Organization

Artur Varanda

School Year 2021-2022

Data Organization

Data storage systems have several layers



Data Organization

Layer 1 – requires specialized laboratories

communication devices: Ethernet, 3G, UMTS, . . .

physical store mediums: hard disks, memory chips, CD-ROMs, . . .

Layer 2 – reading logical data (streams of 0s and 1s)

volatile memory (RAM) – data typically organized by processes

non-volatile storage – data typically organized into volumes

typically organized into volumes (partitions, RAID arrays, . . .)

analyze data at the volume level to find possible hidden data

Data Organization

Layer 3

file system (most common content)

temporary space: swap space in Linux or pagefile in Windows

direct database (without traditional file system), such as *Google file system*, ...

Layer 4

Operating systems (Windows, Mac OSX, Linux, Android, iOS, ...)

Applications (operating system dependent)

Focus of this course

File System Analysis

File system analysis:

collection of data structures that allow an application to create, read, and write files

Analise file system to:

find files

recover deleted files

find hidden data

the result can be:

file content

data fragments

metadata associated with files

File system

organizes data inside a volume

associate file names to file content

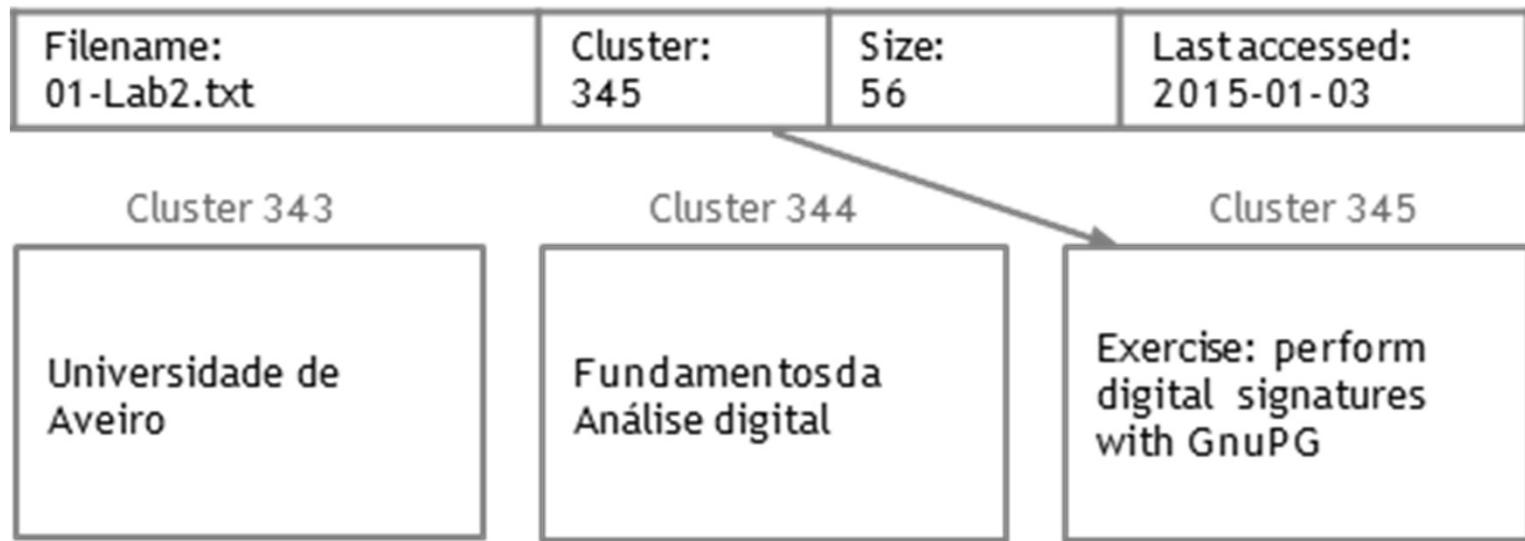
essential data: file names and content location

trustful data – however content may be invalid, *e. g.* deleted files

non essential data: last access time – even if it is wrong the file content still is valid

we may not be able to trust non essential data, *e. g.* system time may be inaccurate, the user may have changed the time, *etc*

we should try to find additional data sources to support an incident hypothesis



File content analysis:

data structure depends on the application or OS that created the file analysis tools may vary accordingly to the application that created the file

HTML is data structure differs from .jpeg

analysis of configuration files is important to determine what programs were running

Data analysis process from the physical level to the application level:



NUMERICAL REPRESENTATION

Numbers can be represented in several ways:

decimal – human system (10 fingers)

10 symbols: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]

binary – computational representation (2 levels of voltage)

2 symbols: [0, 1]

hexadecimal – compact representation of binary numbers

16 symbols: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F]

there are more numerical systems *e. g. octal*

Each symbol has a value depending on its position, e. g. $35\ 812_d$:

$$3 \times 10^4 + 5 \times 10^3 + 8 \times 10^2 + 1 \times 10^1 + 2 \times 10^0 =$$

$$3 \times 10\ 000 + 5 \times 1\ 000 + 8 \times 100 + 1 \times 10 + 2 \times 1 =$$

$$30\ 000 + 5\ 000 + 800 + 10 + 2 = 35\ 812$$

most significant symbol → leftmost value: 3

less significant symbol → rightmost value: 2

BINARY NUMBERS

Example: Convert $1001\ 0011_b$ to decimal:

$$1 \times 2^7 + 0 \times 2^6 + 0 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 =$$

$$1 \times 128 + 0 \times 64 + 0 \times 32 + 1 \times 16 + 0 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1 =$$

$$128 + 0 + 0 + 16 + 0 + 0 + 2 + 1 = 147_d$$

which is the **most** significant digit? which is the **less** significant digit?

Generic formula to convert from any base system to decimal:

$$s_p \times b^p + \dots + s_1 \times b^1 + s_0 \times b^0$$

s – symbol value

b – original base (binary, octal, hexadecimal, ...)

p – symbol position, begins at zero and increases from right to left

Convert to decimal $8BE4_h$, which can also be represented as $0x8BE4$

$$A = 10, B = 11, C = 12, D = 13, E = 14, F = 15$$

$$8 \times 16^3 + 11 \times 16^2 + 14 \times 16^1 + 4 \times 16^0 =$$

$$8 \times 4096 + 11 \times 256 + 14 \times 16 + 4 \times 1 =$$

$$32768 + 2816 + 224 + 4 = 35812_d$$

How do we convert from binary to hexadecimal?

CONVERTING BETWEEN BINARY AND HEXADECIMAL

Conversion Table

Decimal	Binary	Hexadecimal	Decimal	Binary	Hexadecimal
0	0000	0x0	8	1000	0x8
1	0001	0x1	9	1001	0x9
2	0010	0x2	10	1010	0xA
3	0011	0x3	11	1011	0xB
4	0100	0x4	12	1100	0xC
5	0101	0x5	13	1101	0xD
6	0110	0x6	14	1110	0xE
7	0111	0x7	15	1111	0xF

CONVERTING BETWEEN BINARY AND HEXADECIMAL

Example:

$$1001\ 0011_b = 0x93$$

direct conversion:

$$1001_b = 0x9$$

$$0011_b = 0x3$$

1 Byte = 8 bits = 2 hexadecimal digits

hexadecimal → binary compact representation

FLOATING POINT NUMBERS

Floating point number

- format IEEE 754 standard
- exponent in excess allows direct comparisons of floating-point numbers
- mantissa (or significand):
 - normalized → the binary digit 1 to the left of the comma is omitted
 - in additions and subtractions is denormalized, but there is a gradual loss of accuracy
- single precision: 32 bits — exponent in excess of 127 → $[2^{-126}, 2^{+127}]$
- double precision : 64 bits — exponent in excess of 1023 → $[2^{-1022}, 2^{+1023}]$
- conversion tools: <http://www.h-schmidt.net/FloatConverter/IEEE754.html> ; <https://www.binaryconvert.com>

Signal 1 bit	Exponent 8 bits	Mantissa (or significand) 23 bits
------------------------	---------------------------	---

Signal 1 bit	Exponent 11 bits	Mantissa (or significand) 52 bits
------------------------	----------------------------	---

ENDIANCESS – STORAGE ORDER

Binary data unit:

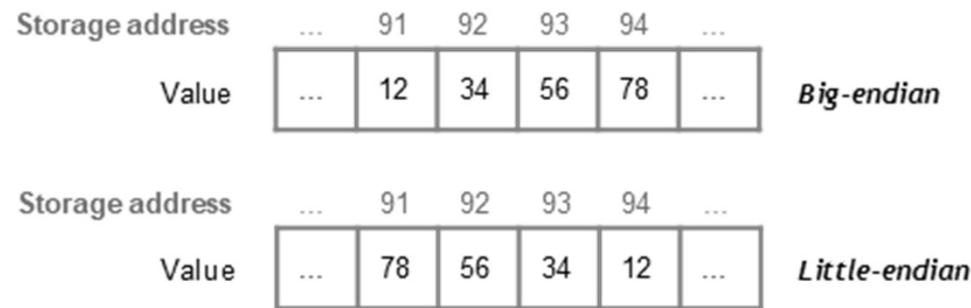
1 Byte (B) = 8 bits (b): $2^8 = 256$ possible values to store large numbers it is necessary to group bytes typically 2, 4 or 8 bytes (16, 32 or 64 bits)

Problem: systems differ in how they store multi-byte values

big-endian: place the **most** significant byte in the 1st (or lower) storage address

little-endian: place the **less** significant byte in the 1st (or lower) storage address

Example: 0x12345678



Big-endian processors

- SPARC, PowerPC, MIPS, Motorola 68k, Alpha, ...

Little-endian processors

- it seems that there are some optimization advantages in pipelined architectures
- z80, VAX, x86, x86-64, amd64, ...

Programmable Big/Little-endian

- ARM, ...

Data networks – *network order*

- IP: **Big-endian**, but there are some exceptions
- very important to guarantee systems interoperability

More info: <http://en.wikipedia.org/wiki/Endianness>

CHARACTER ENCODING – ASCII

Advantages:

- is the simplest way to encode the characters
- doesn't have *endianness* problems – uses 1 byte at a time original version uses only 7 bits – 128 different characters
 - ✓ ASCII table: <https://www.asciitable.com/asciifull.gif>
- it takes up less storage space than unicode

Disadvantages:

- very limited capacity to represent non-English characters
- there are several extended versions – 8 bits (ISO 8859)
 - ✓ the best known is *Latin-1 (ISO 8859-1)* (<http://cs.stanford.edu/people/miles/iso8859.html>)

Address	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
Example	Hex.	...	43	61	70	69	74	E3	6F	20	41	6D	E9	72	69	63	61	...
	Text		C	a	p	i	t	ã	o	A	m	é	r	i	c	a		

Advantages:

- the latest version has more than 137 000 characters
- covers 146 modern and historic scripts, as well as multiple symbol sets

Disadvantages:

- there are several implementations
 - ✓ UTF-8: compatible with ASCII, has variable length from 1 to 4 bytes and is used in *nix systems, www, HTML, etc
 - not subject to endianness problems – it reads 1 byte at a time
 - ✓ UTF-16: has variable length from 2 to 4 bytes, is used in Windows, Mac OS, Java, .Net, KDE, etc
 - ✓ UTF-32: fixed length of 4 bytes
- more complex processing

Example: UTF-8 vs Latin-1

Address	10	11	12	13	14	15
Hex.	...	4F	6C	C3	A1	...
UTF-8		O	l	á		

Address	10	11	12	13	14	15
Hex.	...	4F	6C	E1
Latin-1		O	l	á		

Data structures:

- specifies how data is placed
- it is like a map
- data structure itself is not recorded

Data structure:

```
typedef struct{
    char name[32];                      // 32 bytes
    char postalcode[16];                  // 16 bytes
    uint32_t partner_num;                // 4 bytes
    float quota;                         // 4 bytes
    char empty[8];                       // 8 bytes to lineup the structure to multiples of 16 bytes
};                                     // Total = 64 bytes
```

Data in hexadecimal:

0000000: 4d69 6775 656c 2046 7261 6465 0000 0000	Miguel Frade....
0000010: 0000 0000 0000 0000 0000 0000 0000 0000
0000020: 3234 3131 2d39 3031 0000 0000 0000 0000	2411-901.....
0000030: b300 0d00 c976 9e3f 0000 0000 0000 0000v.?.....

What is the partner number (in hex.)? Is it 0xb300 0d00 or 0x000d 00b3?

What is the quota value (in hex.)? Is it 0xc976 9e3f or 0x3f9e 76c9?

Original source code

```
strcpy(partners[0].name,"Miguel Frade");
strcpy(partners[0].postalcode,"2411-901");
partners[0].partner_num=852147;
partners[0].quota=1.238;
```

Question

Is this system *Little-endian* or *Big-endian*?

if it is *Big-endian*, then:

partner number: 0xb300 0d00 = 3 003 124 992; quota: 0xc976 9e3f = -1 010 147, 94

if it is *Little-endian*, then:

partner number: 0x000d 00b3 = 852 147; quota: 0x3f9e 76c9 = 1, 238

There are many different ways to storage date and time:

- as a string:

- easy to read by humans,
- but hard for computers' operations, for example to compare dates in a database
- many different representations and some times language dependent, *e. g.*:
 - Wednesday, January 9, 11:13:48 UTC 2019
 - 2018-12-23 08:23:55
 - 23-12-2018 08:23:55(PT)
 - 12-23-2018 08:23:55(US)
 - ...

- as binary represented with numbers or hexadecimal

- difficult to read by humans, but easier for computers' operations
- unfortunately, not all software uses the same representation

How time is counted:

Unix time = POSIX time = UNIX Epoch time (<https://www.epochconverter.com>)

- number of elapsed seconds since 1970-01-01 00:00:00
- Unix, Linux, **Firefox**, Java, JavaScript, Perl, PHP, Python, Ruby, Tcl, etc

400-year Gregorian calendar cycle

- number of microseconds since 1601-01-01 00:00:00
- **Google Chrome**, Windows 32 and 64 bits, NTFS, Cobol, etc

Examples of date and time storage formats:

Software	Representation	Example	Software	Representation	Example
Win 64 bits BE	16 Hex chars	01 CC A6 3C 91 B9 72 00	Firefox SQLite	16 digits	1 321 653 236 000 000
Win 64 bits LE	16 Hex chars	00 72 B9 91 3C A6 CC 01	Chrome SQLite	17 digits	12 966 126 836 000 000
Unix numeric	10 digits	1 321 653 236	Coockie hi:low	18:19 digits	3 923 586 186:30 188 999
Unix numeric m. seconds	13 digits	1 321 653 236 000 000	DOS wFAT	8 Hex chars	F0 48 64 40
Unix 32 bits BE	8 Hex chars	4E C6 D3 F4	GSM	14 digits	99 309 251 619 580
Unix 32 bits LE	8 Hex chars	F4 D3 C6 4E	Samsung Swift	8 Hex chars	E7 8C 94 7D
HFS+ 32 bits BE	8 Hex chars	CA EC 84 74	Nokia Series 40	14 Hex chars	07 D9 04 11 13 27 09
HFS+ 32 bits LE	8 Hex chars	74 84 EC CA			
Mac Absolute time	9 digits	343 346 036			

Chrome SQLite/1000000 = Unix time + 11644473600 sec

(1970-01-01 00:00:00) - (1601-01-01 00:00:00) = 11644473600 sec

Date and time convert tools

MFT Stampede <https://mft-stampede.software.informer.com/download/>

- ✓ Windows GUI tool, very easy to use

Nirsoft tool *BrowsingHistoryView* https://www.nirsoft.net/utils/browsing_history_view.html

- ✓ supports browsing history of Internet Explorer, Mozilla Firefox, Google Chrome, and Safari

SQL – cool for automation of tasks <https://sqlitebrowser.org/dl/>

- Firefox SQLite database overview of the visited sites:
 - Ubuntu -> [user_home_directory]/.mozilla/firefox/xxxxxxxx.default/places.sqlite
 - Windows -> [user_home_directory]\AppData\Roaming\Mozilla\Firefox\Profiles\xxxxxxxx.default\places.sqlite

```
SELECT datetime(moz_historyvisits.visit_date/1000000, 'unixepoch', 'localtime') AS Date_Time, moz_places.url  
FROM moz_places, moz_historyvisits  
WHERE moz_places.id = moz_historyvisits.place_id  
ORDER BY Date_Time ASC
```

DATE AND TIME

- Google Chrome SQLite database overview of the visited sites:

Ubuntu -> [user_home_directory]/.config/google-chrome/Default/databases

Windows -> [user_home_directory]\AppData\Local\Google\Chrome\User Data\Default\History

```
SELECT datetime(((visits.visit_time/1000000)-11644473600), 'unixepoch', 'localtime') AS Date_Time, urls.url, urls.title  
FROM urls, visits  
WHERE urls.id = visits.url  
ORDER BY Date_Time ASC
```

Binary-to-text encoding

- is encoding of data in plain text, or in other words it is an encoding of binary data in a sequence of printable characters
- the encoding is necessary for transmission of data when the channel does not allow binary data (e. g. email)
- encoding inflates the original data size, the inflate rate depends on the used technique encoding is a reversible operation
- it can also be applied to plain text

Not to be confused with encrypting, **it's not** encryption:

- encryption requires a key, usually secret
- encoding doesn't depend on a key

COMMON ENCODING TECHNIQUES

Common encoding techniques:

hexadecimal (also known as base16)

- ✓ used chars: [0..9] and [A..F] (or [a..f])
- ✓ hash values (MD5, SHA245, etc) are usually displayed in hexadecimal
- ✓ example: "Hello World" → 48656c6c6f20776f726c64

base64

- ✓ used chars: [A..Z, [a..z], [0..9], and [+ , /]
- ✓ base64 string size must be a multiple of 4, so char = can be used at the end as padding
- ✓ used on: email servers (MIME), OpenPGP, etc
- ✓ example: "Hello World" → SGVsbG8gV29ybGQ=
<https://cryptii.com/pipes/binary-to-base64> <https://www.browserling.com/tools/base64-encode>

base58:

- ✓ similar to base64, but modified to avoid both non-alphanumeric characters and letters which might look ambiguous when printed
- ✓ used on bitcoins,
- ✓ example: bitcoin public key 1ZNz2KDm8epACBA5bjgKQbRyaGcDt3XV2
- ✓ bitcoin private key: Ky1ZcCSMziFtdxfDEjANw3PZUZQLjh6hKpX1CinVtJscnAFnvcn
<https://learnmeabitcoin.com/technical/base58> https://en.bitcoin.it/wiki/Base58Check_encoding

Exercises

Please do the following exercises:

Lab 1 – Endianness, read binary file

Lab 2 – Identify different character encodings

Lab 3 – Character encoding conversions

ANY QUESTIONS?





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Autopsy

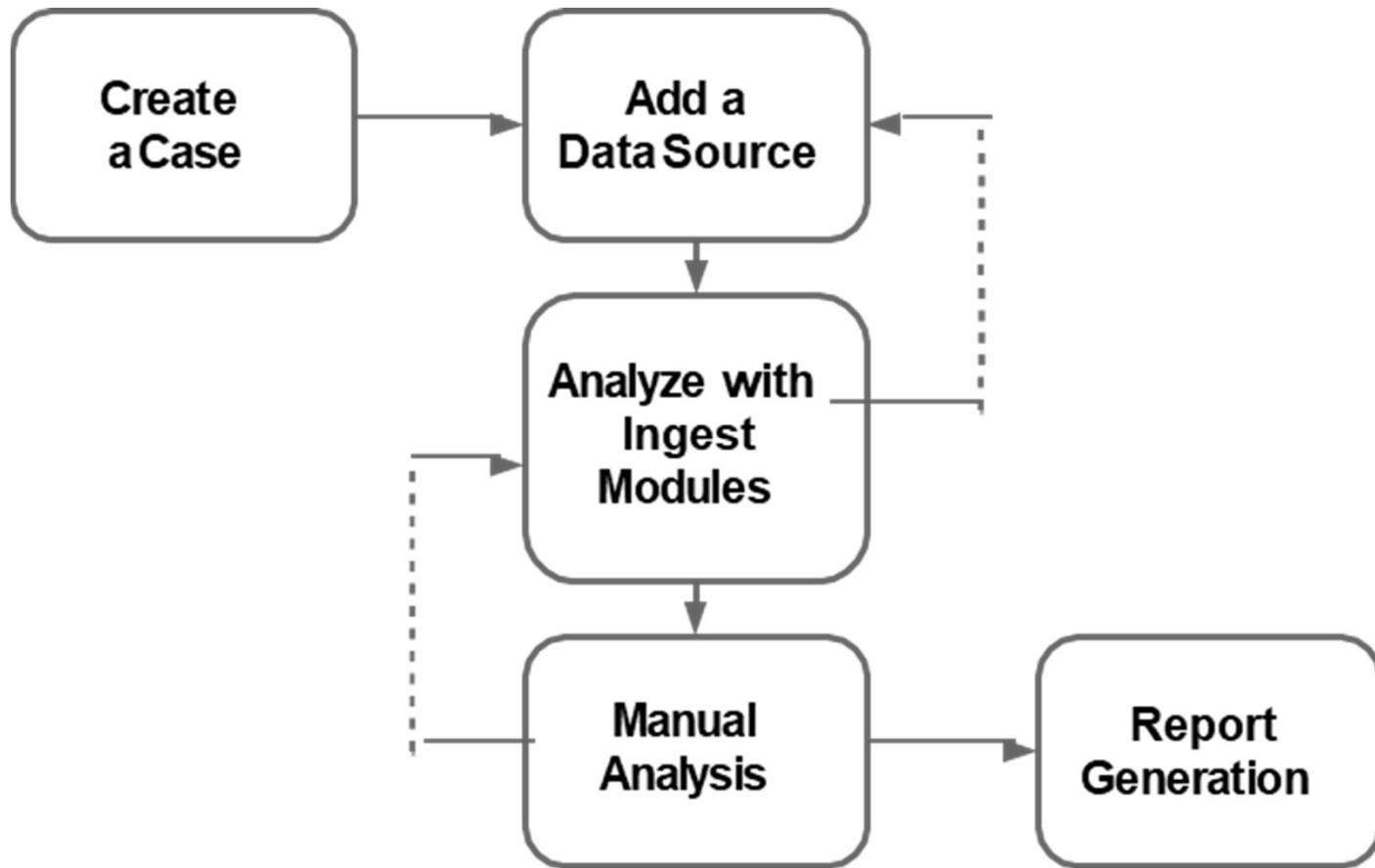
Artur Varanda

School Year 2021-2022

url: <https://www.autopsy.com>

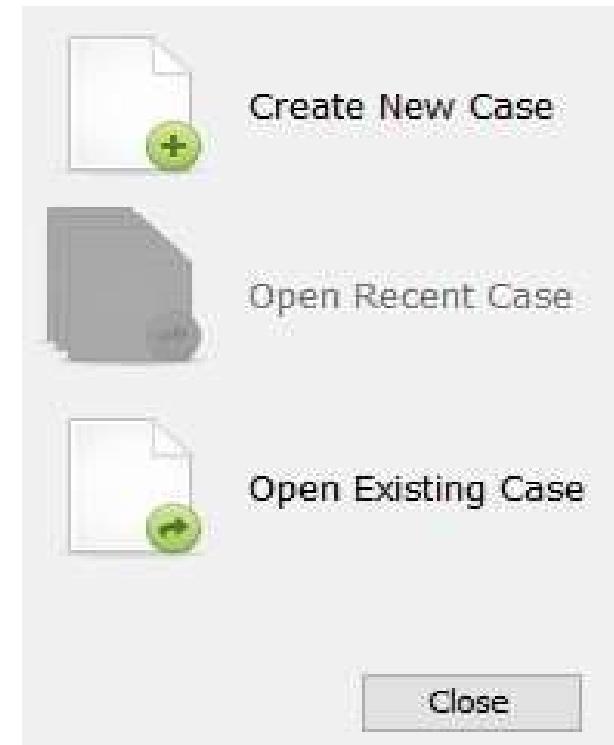


- *Autopsy* is a graphical tool aimed at the digital investigation of images of storage media
- It is developed in Java, mainly for *Windows*
- It is expandable (supports modules developed in *Python* for Java)
- It has limited support for *Android*



CREATE A CASE

1 - Create a Case



1 - Create a Case

- Case Information

Enter New Case Information:

Case Name: 2017.02.C001.M57BIZ

Base Directory: C:\Autopsy\Formação\

Case Type: Single-user Multi-user

Case data will be stored in the following directory:

C:\Autopsy\Formação\2017.02.C001.M57BIZ

< Back

Next >

1 - Create a Case

- Case Information
- Case number, examiner

Optional Information	
Case	
Number:	NUIPC xxxxxx
Examiner	
Name:	MF
Phone:	
Email:	
Notes:	Test
Organization	
Organization analysis is being done f... <input type="checkbox"/>	
< Back Next > Finish	

1 - Create a Case

- Case Information
- Case number, examiner

2 - Add a data source

- Raw (dd) or EnCase (E01) image
- Drives, files or local folders
- Virtual machine drives (vmdk, vhd)



1 - Create a Case

- Case Information
- Case number, examiner

2 - Add a data source

- Raw (dd) or EnCase (E01) image
- Drives, files or local folders
- Virtual machine drives (vmdk, vhd)

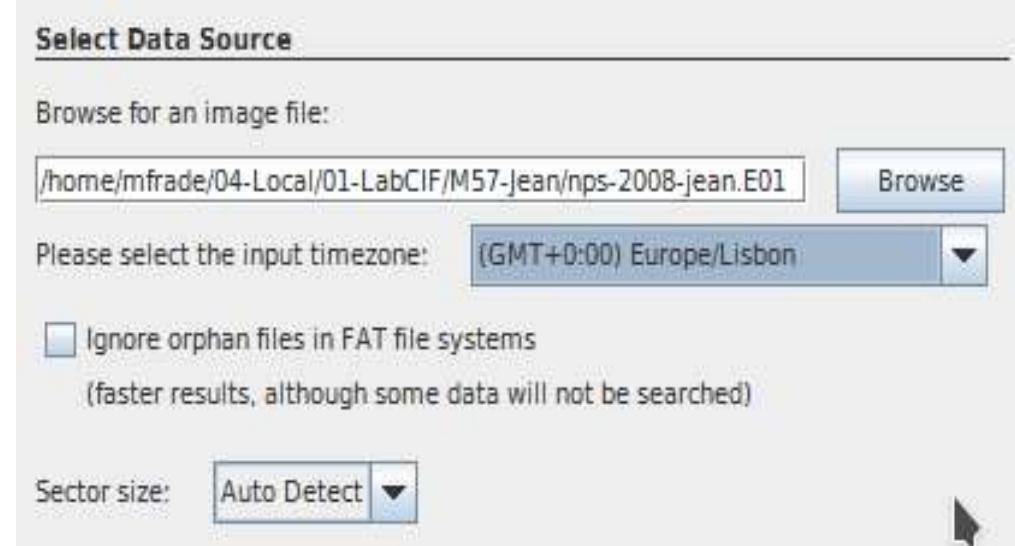
Select Data Source

Browse for an image file:

Please select the input timezone:

Ignore orphan files in FAT file systems
(faster results, although some data will not be searched)

Sector size:



AUTOMATED PROCESSING – WITH INGEST MODULES

Configure Ingest Modules

Run ingest modules on:

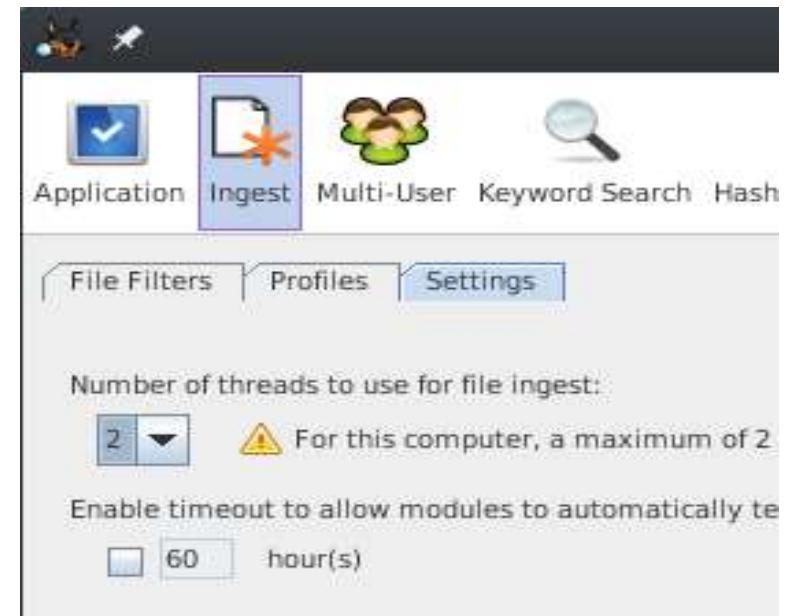
All Files, Directories, and Unallocated Space

<input checked="" type="checkbox"/> Recent Activity	The selected module has no per-run settings.
<input checked="" type="checkbox"/> Hash Lookup	
<input checked="" type="checkbox"/> File Type Identification	
<input checked="" type="checkbox"/> Embedded File Extractor	
<input checked="" type="checkbox"/> Exif Parser	
<input checked="" type="checkbox"/> Keyword Search	
<input checked="" type="checkbox"/> Email Parser	
<input checked="" type="checkbox"/> Extension Mismatch Detector	
<input checked="" type="checkbox"/> E01 Verifier	
<input checked="" type="checkbox"/> Encryption Detection	
<input checked="" type="checkbox"/> Interesting Files Identifier	
<input checked="" type="checkbox"/> PhotoRec Carver	
<input checked="" type="checkbox"/> Correlation Engine	
<input checked="" type="checkbox"/> Virtual Machine Extractor	
<input checked="" type="checkbox"/> Android Analyzer	

Select All Deselect All History Global Settings

Extracts recent user activity, such as Web browsing, recently ...

- Autopsy supports *multi-thread* execution of *file ingest*
- Aims to reduce the processing time
- Requires setting of the number of *threads* to use
- **Tools → Options → Ingest → Settings**

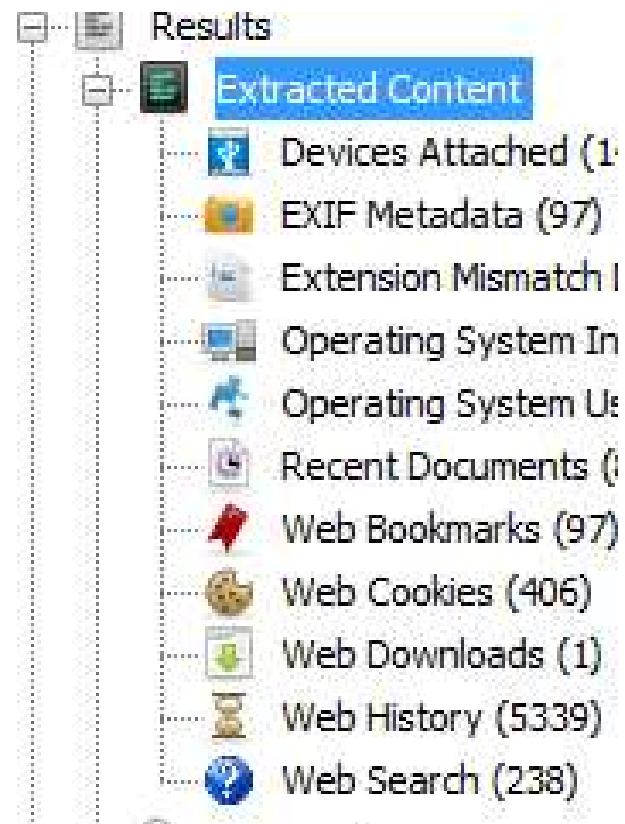


Extracts information from the last 7 days

- Internet usage (including searches)
- Installed programs
- Connected devices (USB)
- Processes the *Registry hive*

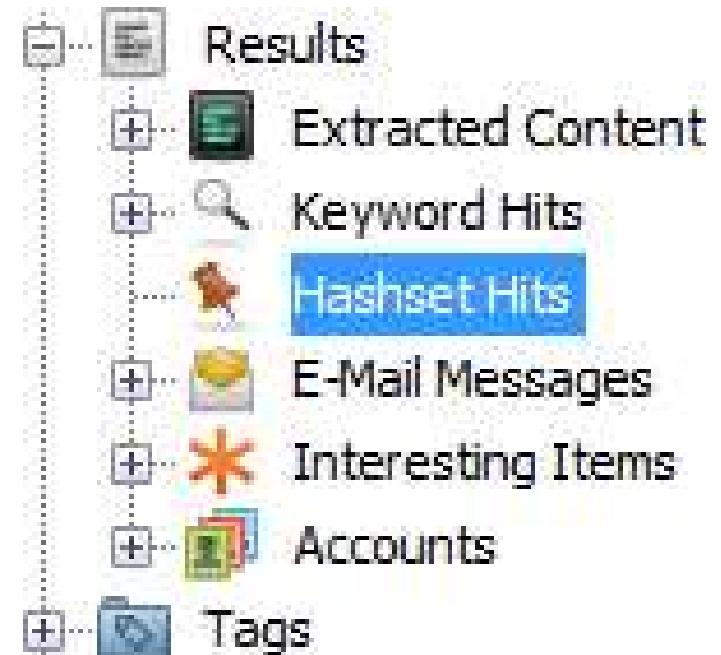
The information is displayed in

Results → Extracted Content



Computes hash values of all found files and compares them with an existing database of *MD5* hashes

- Known bad hashsets
 - ✓ Files that must be validated
- Known good hashsets
 - ✓ Files that can be ignored
- Known hashsets
 - ✓ Files that can be good or bad (depending on the context)



Mainly available only for police forces (*i.e. hash sets of child pornography pictures*)

List of hash can be *good, bad or just known*

National Software Reference Library (NSRL) from NIST

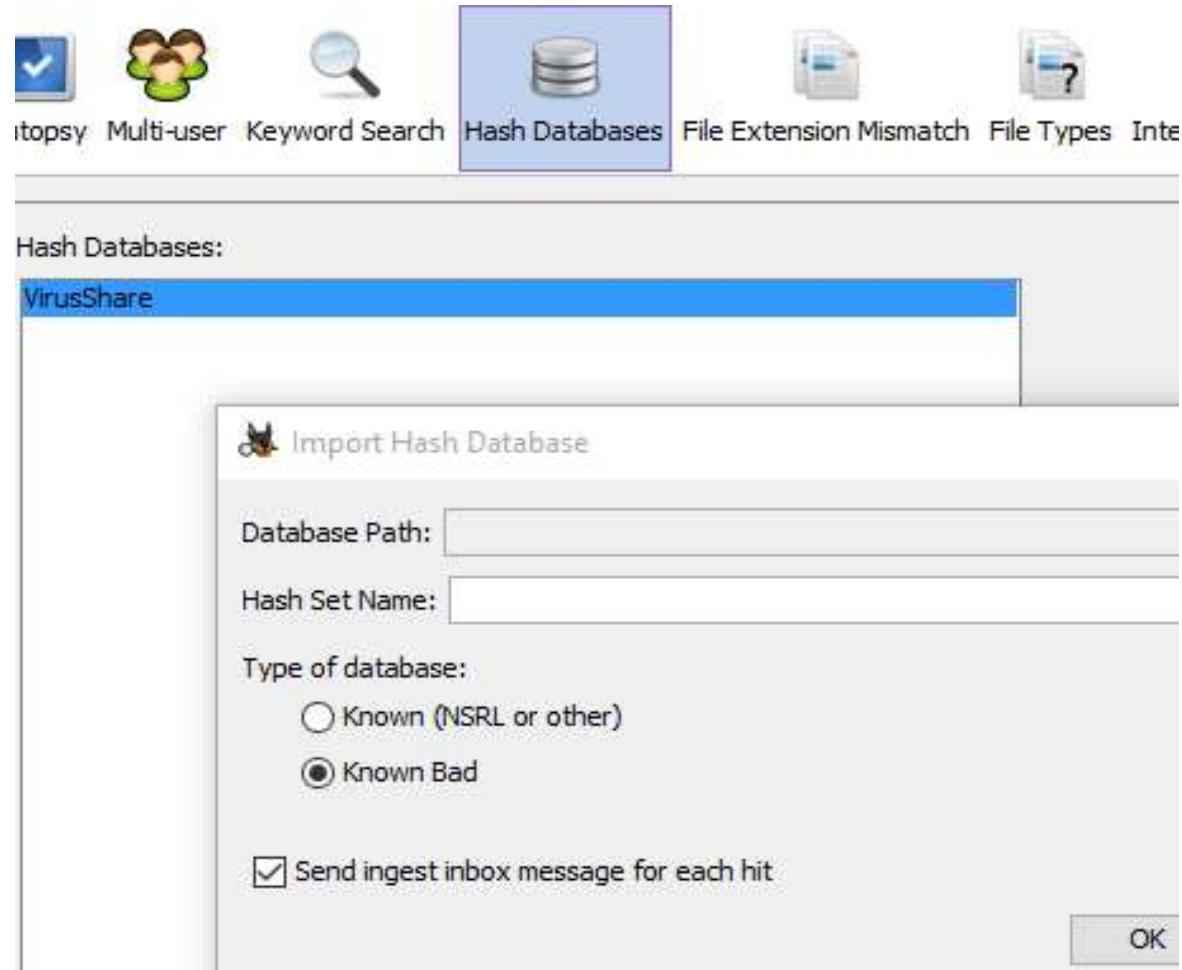
URL: <http://www.nsrl.nist.gov/>

URL: <http://sourceforge.net/projects/autopsy/files/NSRL/>

VirusShare

URL: <https://virusshare.com/hashes.4n6>

MODULE: HASH LOOKUP – HASH SETS

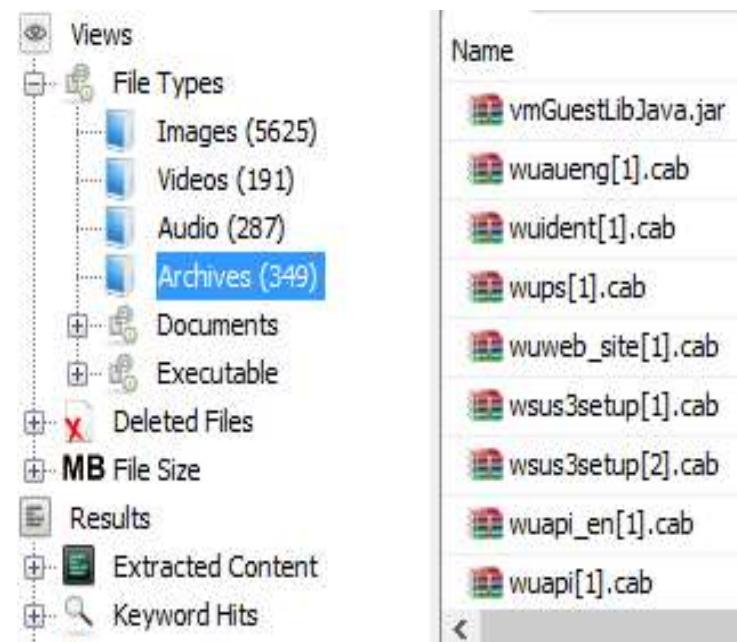


Checks the file type according to its characteristics and collects meta data

- Uses *Tika* (<http://tika.apache.org/>)
- Indexing module without its own *output*
- Generates information for other modules
 - ✓ Extension Mismatch Detector
 - ✓ Keyword Search

Uncompress files (ZIP, RAR) or embedded files (DOC, DOCX, PPT, PPTX, XLS and XLSX), processing them again.

- Enables analysis of files included in these files
- Results are displayed in **File types → Archives**



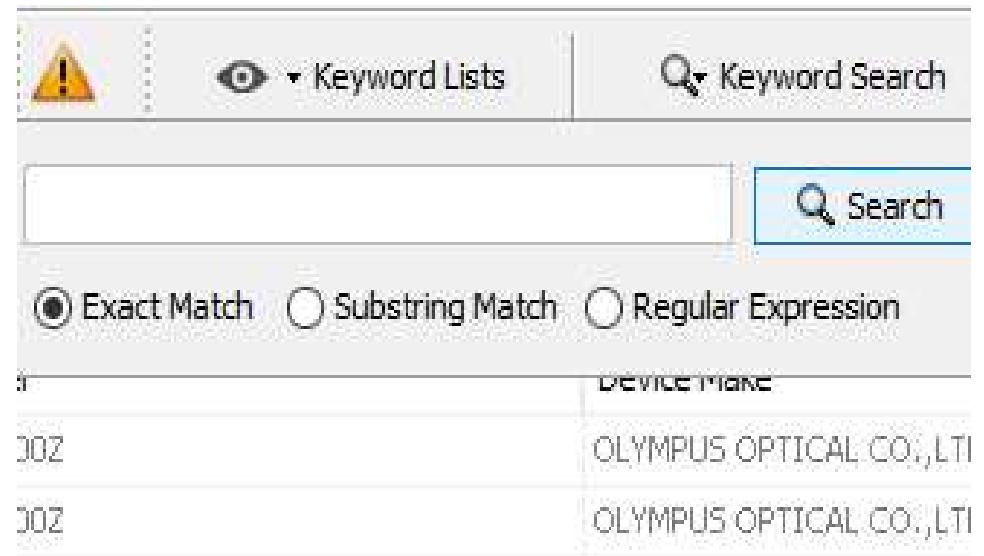
Extracts EXIF (*Exchangeable Image File Format*) information stored on images

- Geolocation, date and time
- Camera model, setup (exposure, resolution, . . .)
- Results are displayed in **Extracted content → EXIF Metadata**

Results			
	File	Date	Device
Extracted Content	yhst-39930517073039_2007_147143019[1].jpg	0002-11-30 00:00:00 GMT	C4100Z,C4000Z
Devices Attached (14)	yhst-39930517073039_2007_267809[1].jpg	0002-11-30 00:00:00 GMT	C4100Z,C4000Z
EXIF Metadata (97)	HPIM1361[1].jpg	2008-06-15 22:17:02 BST	Photosmart M525
Extension Mismatch Dete	HPIM1360[1].jpg	2008-06-15 22:16:52 BST	Photosmart M525
Operating System Inform			

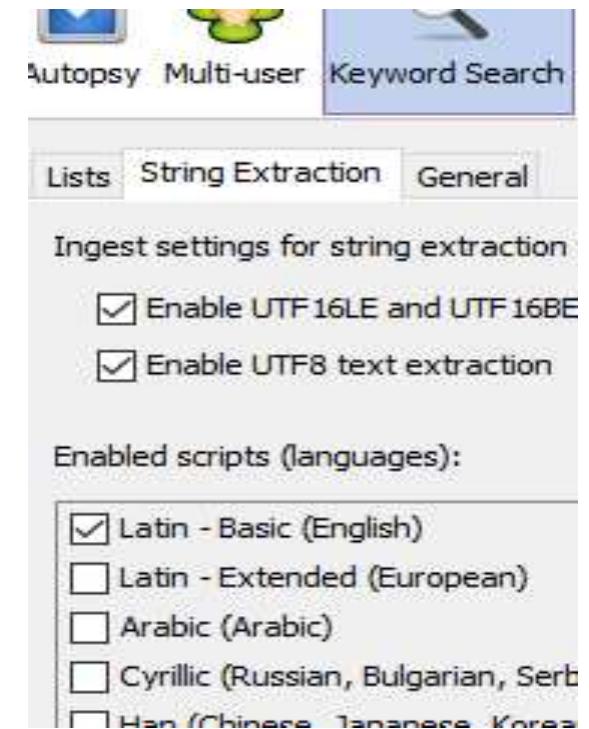
Search by keywords during initial or on-demand processing

- Extracts text from the files being processed and adds them to an index (Solr)
- Supports several formats (Text, MS Office, PDF, Emails)



Search by keywords during initial or on-demand processing

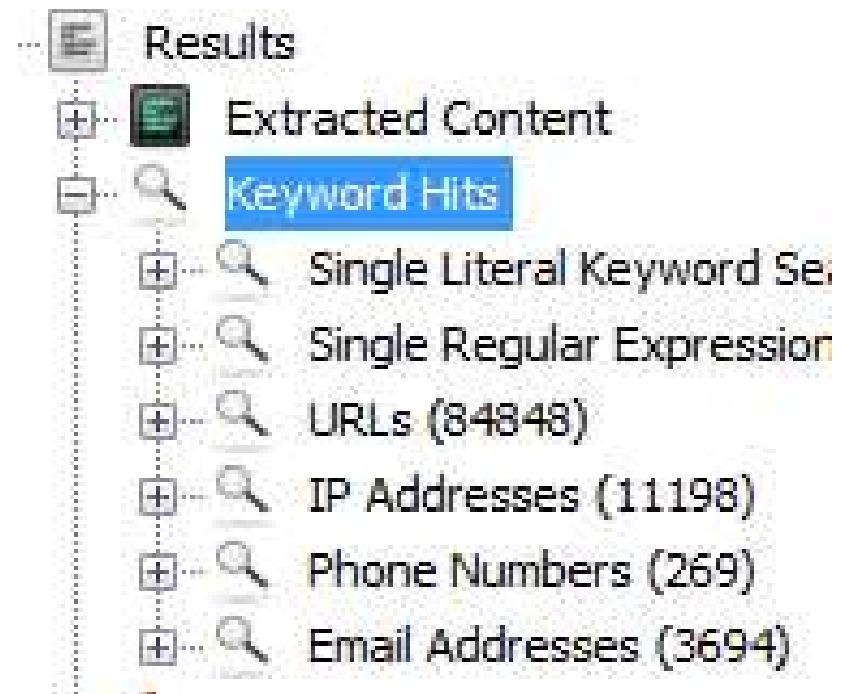
- Extracts text from the files being processed and adds them to an index (Solr)
- Supports several formats (Text, MS Office, PDF, Emails)
- For non-supported formats
 - ✓ String Extraction algorithm
 - ✓ Is able to identify encodings and languages



Autopsy includes a set of predefined lists of common expressions

- Web addresses (URLs)
- IP addresses
- Phone numbers E-mail addresses

Unfortunately, they generate a huge amount of false positives



Identifies and processes e-mail program files (MBOX, PST)

- Extract contained e-mails
- Processes its attachments

The screenshot shows a software interface for email parsing. On the left, there is a tree view of analysis results:

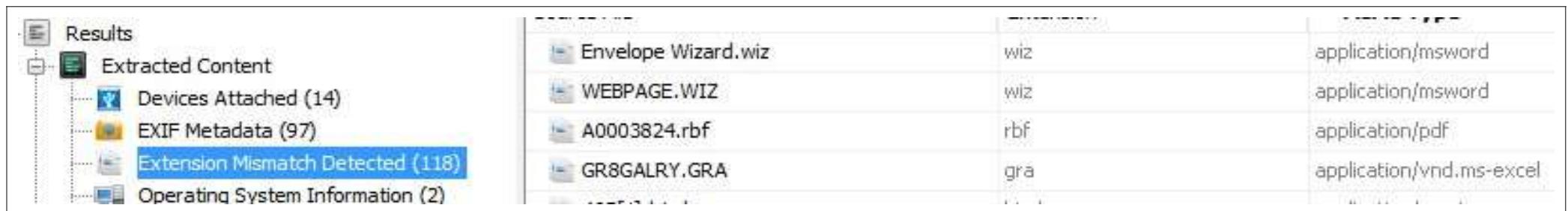
- Results
 - Extracted Content
 - Keyword Hits
 - Hashset Hits
 - E-Mail Messages
 - Default ([Default])
 - Default (261)

On the right, there is a table listing extracted email messages:

	outlook.pst	jean@m57.biz	Google Alerts: googlealerts-noreply@google.com
	outlook.pst	jean@m57.biz	Google Alerts: googlealerts-noreply@google.com
	outlook.pst	jean@m57.biz	alex; alex@m57.biz
	outlook.pst	jean@m57.biz	alex; alex@m57.biz
	outlook.pst	jean@m57.biz	alex; alex@m57.biz

Identifies files that have a file pattern that doesn't matches the filename extension

- Attempts to identify camouflaged files
 - ✓ may generate some false positives



The screenshot shows a software interface for forensic analysis. On the left, there's a tree view under the 'Results' section. The 'Extension Mismatch Detected' item is highlighted with a blue selection bar. To the right is a table displaying four rows of detected files:

FILENAME	EXTENSION	MIME TYPE
Envelope Wizard.wiz	wiz	application/msword
WEBPAGE.WIZ	wiz	application/msword
A0003824.rbf	rbf	application/pdf
GR8GALRY.GRA	gra	application/vnd.ms-excel

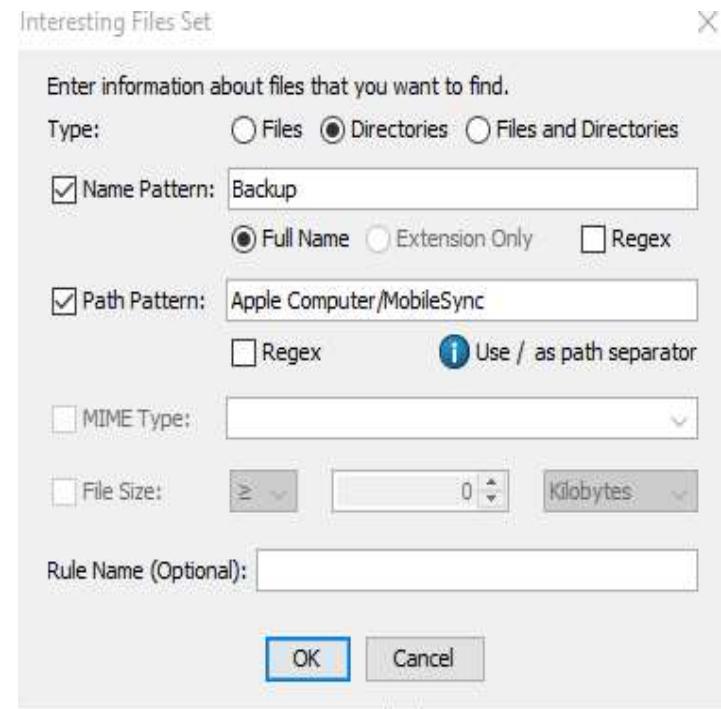
Verifies the hash value of the data stored in EWF files

- Calculates the hash and compares it with the values stored in the E01 metadata
- Aims to identify corrupted EWF files and prevents its automated process

MODULE: INTERESTING FILES IDENTIFIER

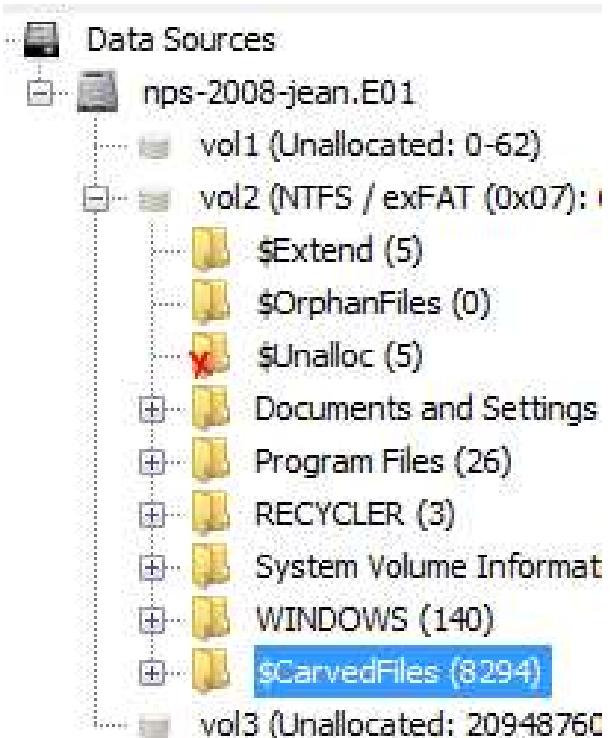
Generate alerts when it detects files and folders with certain characteristics

- Type (file / folder)
- Size, extension
- Name, path
- MIME type



Extract files from unallocated spaces

- Supports multiple file types
- Allows the discovery of recently deleted files
- Allows custom addition of file patterns
- “Process Unallocated Space” option must be selected

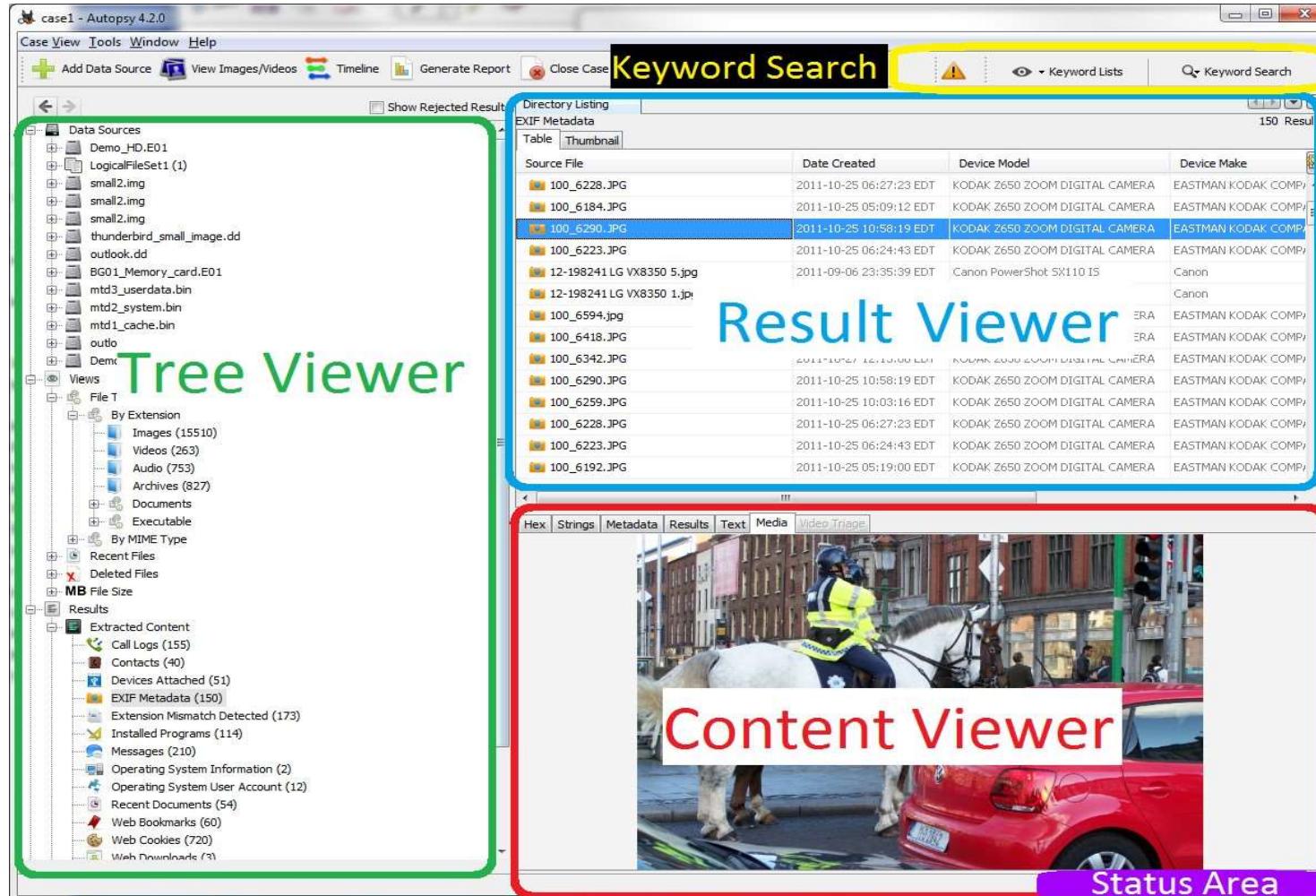


Identifies virtual machine disks and adds them directly as new data sources

- ✓ Supports VMWare (vmdk) and Microsoft Virtual Hard Drives (vhd) files

FTK Imager can read also virtual disks files and convert them to E01

MANUAL CONTENT ANALYSIS AUTOPSY GRAPHIC INTERFACE



Tree viewer indexes information resulting from automated processing and gives access to four large areas:

- **Data sources:** Indicates the data source, allowing navigation within the respective file systems
- **Views:** Shows the found files under multiple views (type, size, state). The same file can appear here several times (in different views).
- **Results:** Shows the results found by the several modules.
- **Reports:** Indicates the several produced reports, either manually or automatically by the modules.

The **Views** area has:

- **File type:** Sorts files by extension or MIME type.
- **Recent files:** Files accessed in the last 7 days.
- **Deleted files:** Deleted files deleted, it tries to recover their original name.
- **File size:** Sorts files by size.

Useful when image analysis is relevant to the case under consideration. It is available in the *Tools*

- Group images by folder, compressed file
- Allows viewing of images when detected
- Functionality can be activated /deactivated in the options
- Allows cataloging of images (for child pornography and similar tasks)

Useful when searching for a file with specific characteristics.

It is available in the *Tools* menu

- Name
- Size
- MIME type
- Date
- Good/Bad

File Search by Attributes X

Search for files that match the following criteria:

Name:

*Note: Name match is case insensitive and matches any part of the file name. Regular expressions are not currently supported.

Size: equal to 0 Byte(s)

MIME Type:

application/activemessage
application/andrew-inset
application/applefile
application/applicware

*Note: Multiple MIME types can be selected

Date: to

*Empty fields mean "No Limit" *The date format is mm/dd/yyyy

Timezone: (GMT+0:00) Europe/London

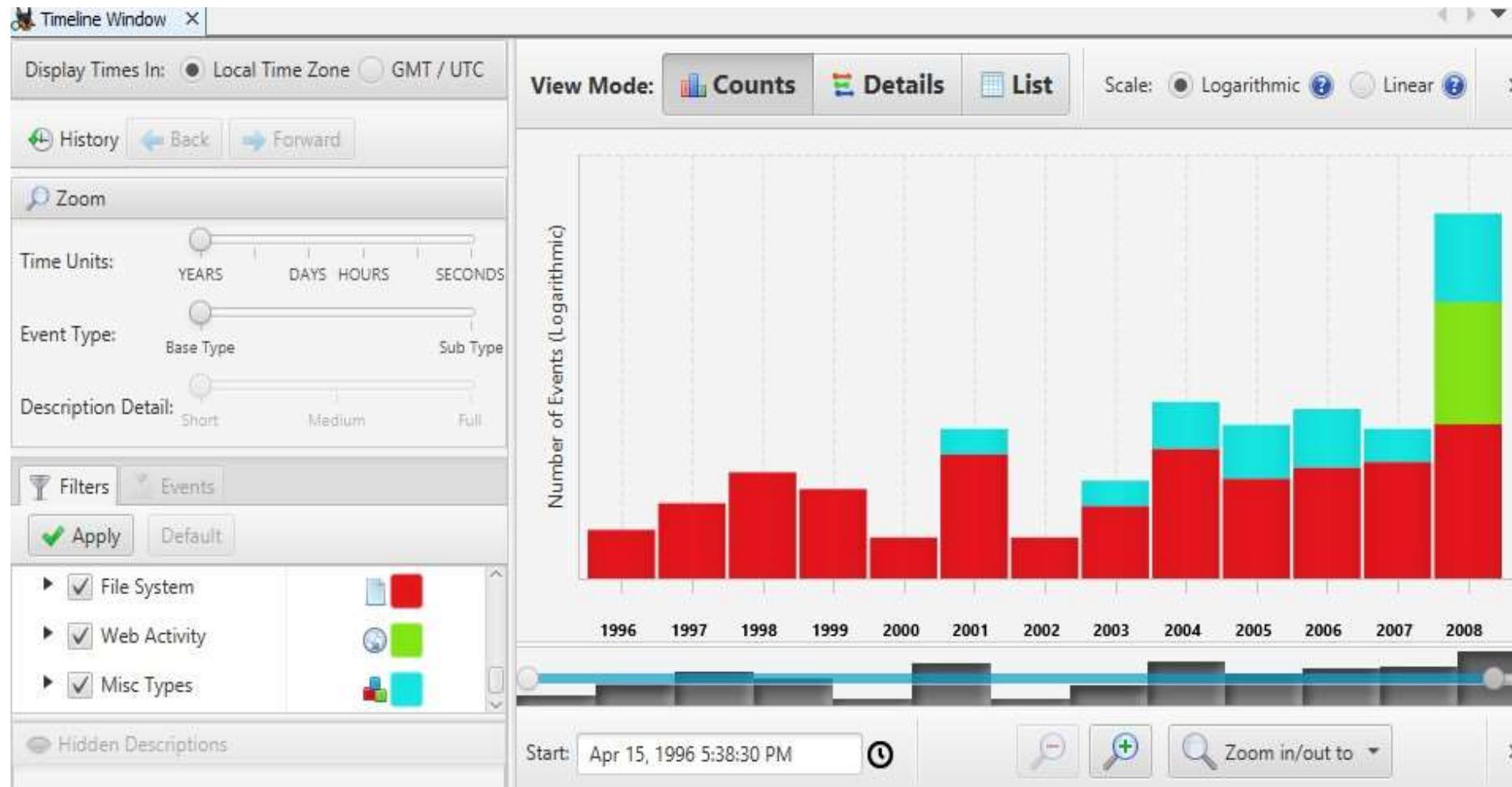
Modified Accessed Created Changed

Known Status:

Unknown Known (NSRL or other) Known bad

TIMELINES

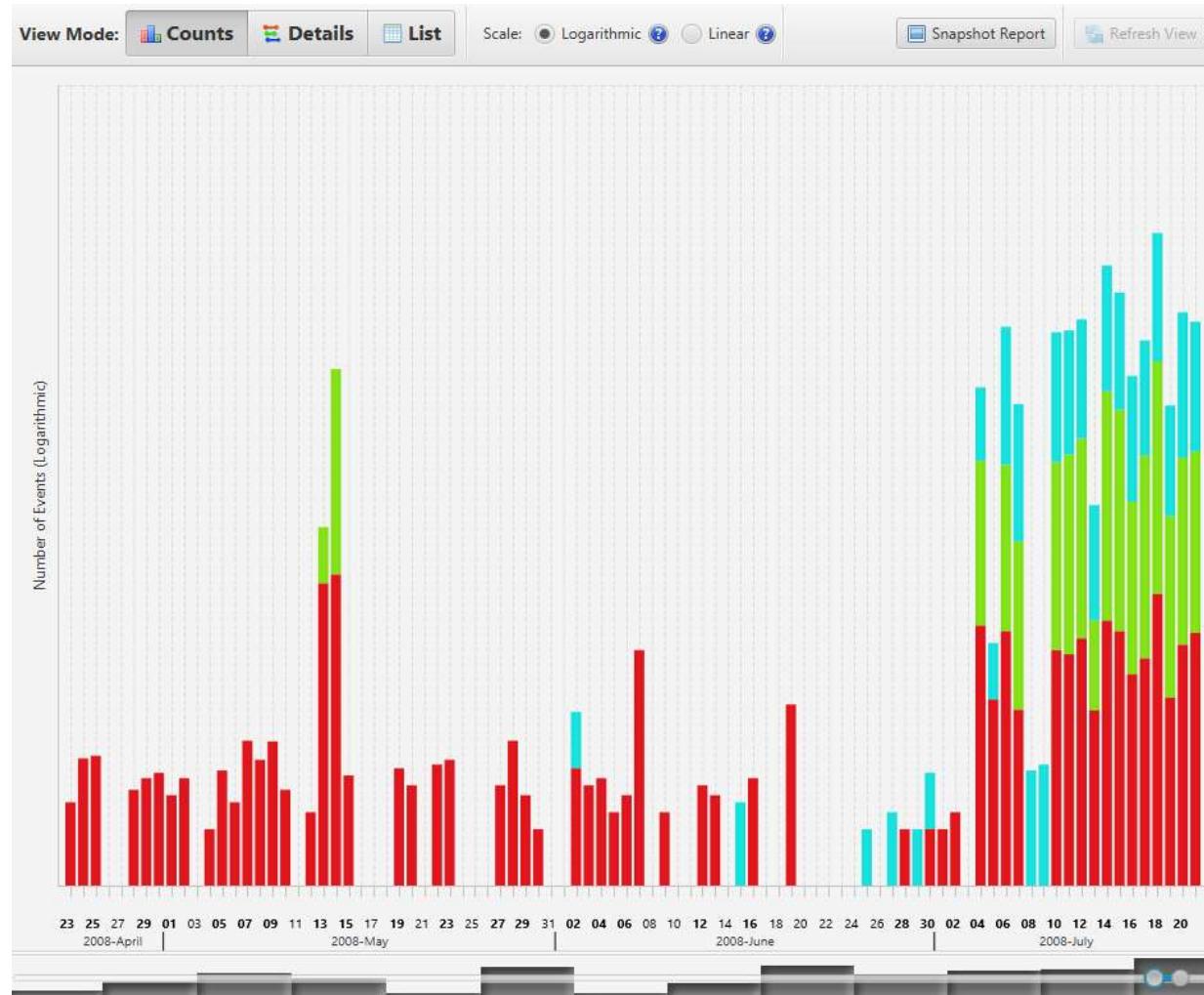
After indexing events, Autopsy allows you to create timelines based on the dates on which such events occurred



Autopsy recognizes events, such as

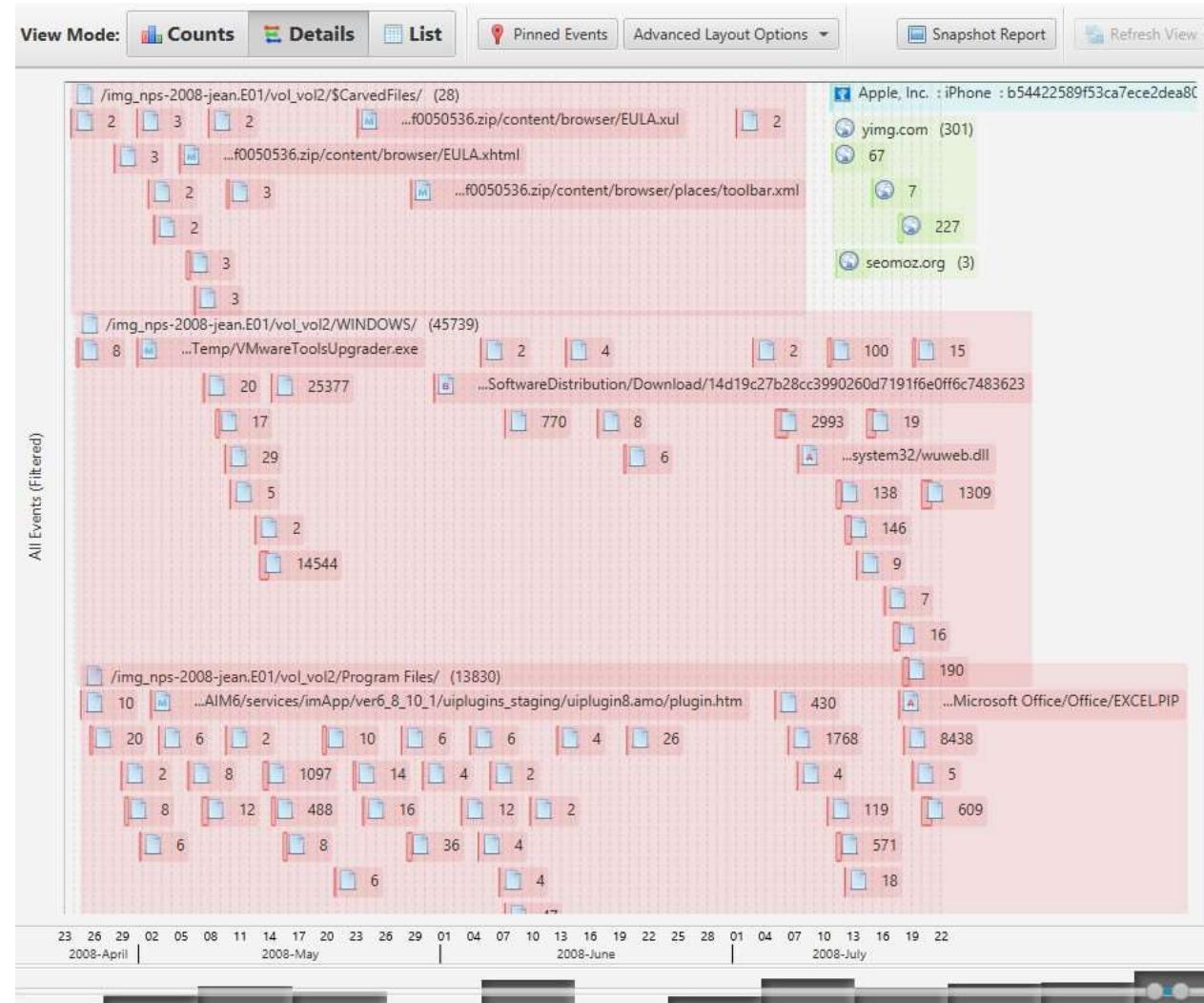
- Files (modification, access, creation, change)
- Internet access (downloads, cookies, bookmarks, searches, browser history)
- Others (messages, phone calls, e-mails, GPS tracks, . . .)

TIMELINE VISUALIZATION HISTOGRAM



TIMELINE VISUALIZATION

DETAILED VIEW

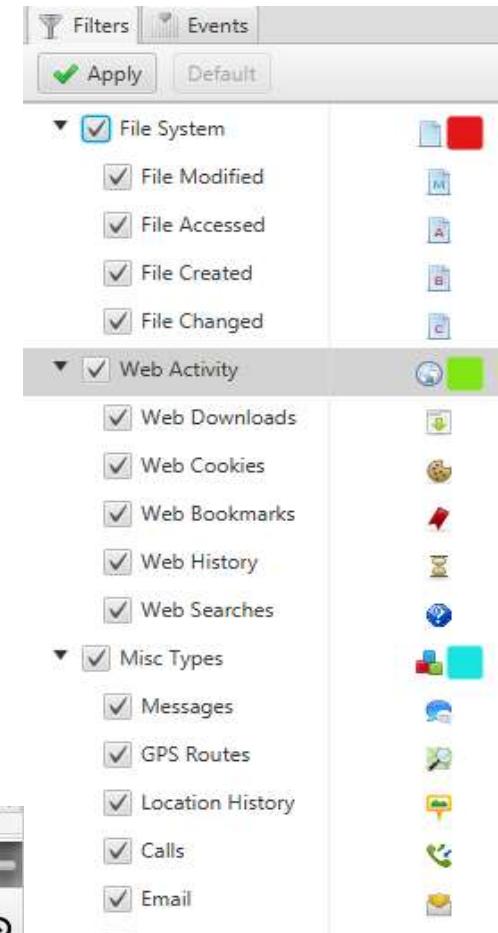


TIMELINE VISUALIZATION

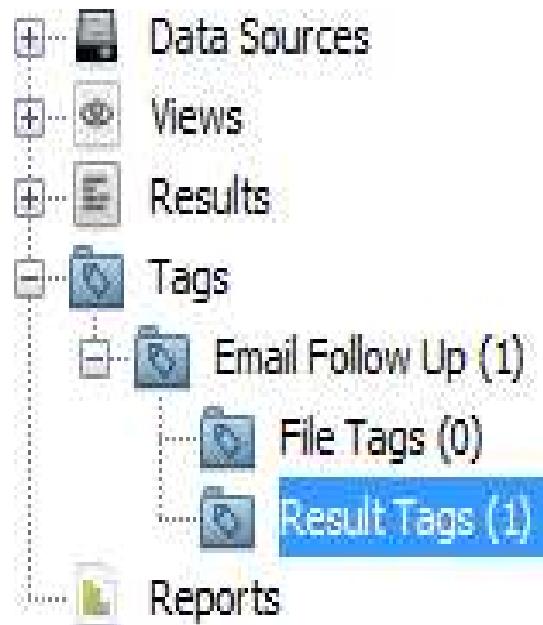
FILTERS

Autopsy allows to reduce the number of elements in a timeline using filters

- Filter known files
- Filter by text
- Event type
- Time windows



- Tag results with labels
- Items for future reference
- Enables the marking of files or results
- Tag name set by investigator
- Tags appear as a sub-area of **Results**



Several types of reports are available

Results: Applies to the items of the results view,



Generate Report

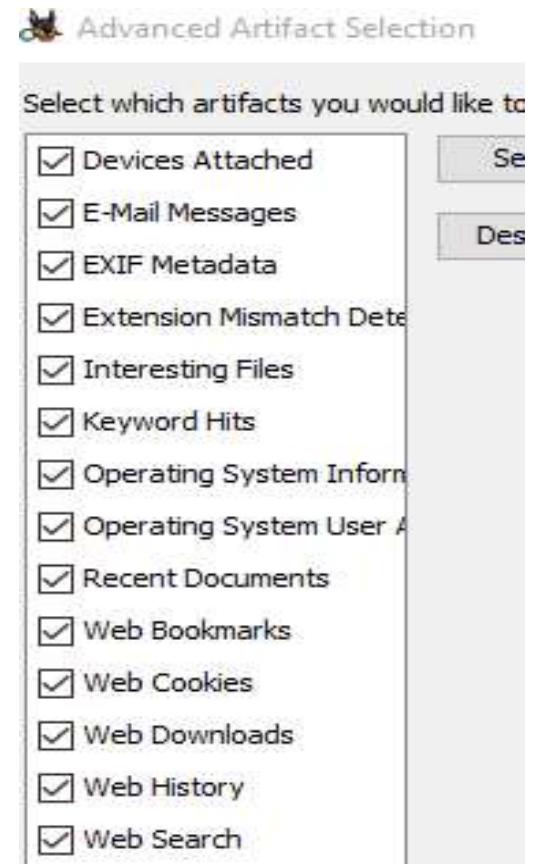
Select and Configure Report

Report Modules:

- Results - HTML
- Results - Excel
- Add Tagged Hashes
- Files - Text
- Google Earth/KML
- STIX
- TSK Body File

Several types of reports are available

Results: Applies to the items of the results view, can be filtered



Several types of reports are available

Results: Applies to the items of the results view, can be filtered

Tagged: Applies to the tagged items

Configure Artifact Reports

Select which data to report on:

All Results

Tagged Results

Email Follow Up

REPORT GENERATION

Several types of reports are available

Results: Applies to the items of the results view, can be filtered

Tagged: Applies to the tagged items

Files: List of files under analysis

Configure File Report

Select items to include in File Report:

- Name
- File Extension
- File Type
- Is Deleted
- Last Accessed
- File Created
- Last Modified
- Size
- Address
- Hash Value
- Known Status
- Permissions
- Full Path

Select All

Deselect All

Several types of reports are available

Results: Applies to the items of the results view, can be filtered

Tagged: Applies to the tagged items

Files: List of files under analysis

KML: List of GPS coordinates in *Google Earth* format



Select and Configure Report Modules

Report Modules:

- Results - HTML
- Results - Excel
- Add Tagged Hashes
- Files - Text
- Google Earth/KML
- STIX
- TSK Body File

Several types of reports are available

Results: Applies to the items of the results view, can be filtered

Tagged: Applies to the tagged items

Files: List of files under analysis

KML: List of GPS coordinates in *Google Earth* format

TSK: MAC timeline list of all files

Select and Configure Report

Report Modules:

- Results - HTML
- Results - Excel
- Add Tagged Hashes
- Files - Text
- Google Earth/KML
- STIX
- TSK Body File

Several types of reports are available

Results: Applies to the items of the results view, can be filtered

Tagged: Applies to the tagged items

Files: List of files under analysis

KML: List of GPS coordinates in *Google Earth* format

TSK: MAC timeline list of all files

STIX: Compares the results obtained with a threat file

Select and Configure Report Modules

Report Modules:

- Results - HTML
- Results - Excel
- Add Tagged Hashes
- Files - Text
- Google Earth/KML
- STIX
- TSK Body File

- Structured language for describing cyber threat information so it can be shared (XML)
- Accepts indicators like:
 - ✓ IP address, URL, Names
 - ✓ TCP, UDP connections
 - ✓ Filenames, *hashes*
 - ✓ ...
- More information: <https://stix.mitre.org/>
<https://stix.mitre.org/language/version1.0.1/samples.html>
<https://oasis-open.github.io/cti-documentation/stix/examples.html>

STRUCTURED THREAT INFORMATION EXCHANGE (STIX)

Example: IP address

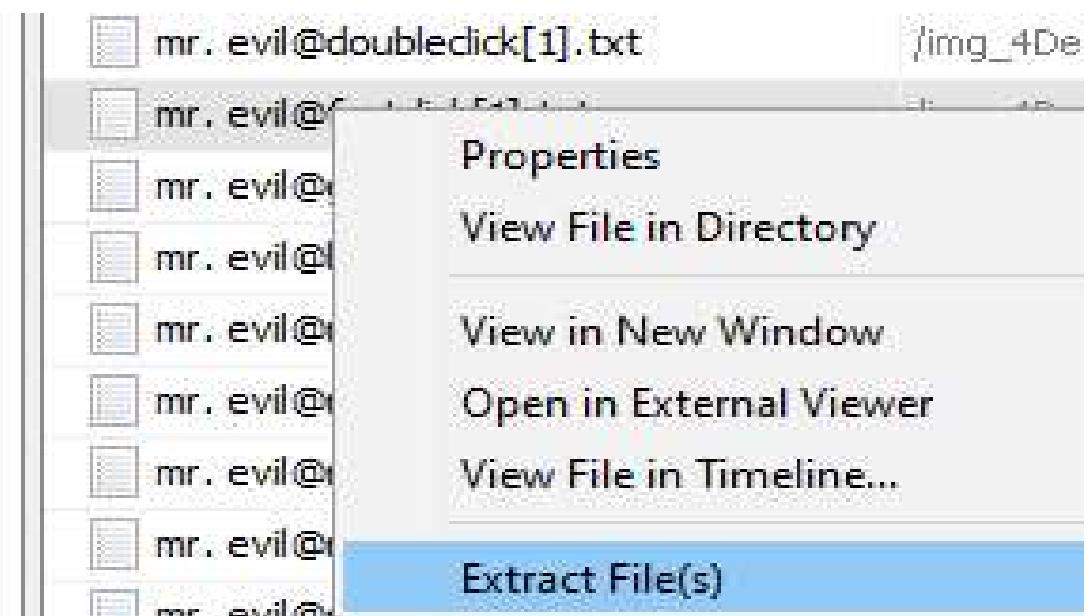
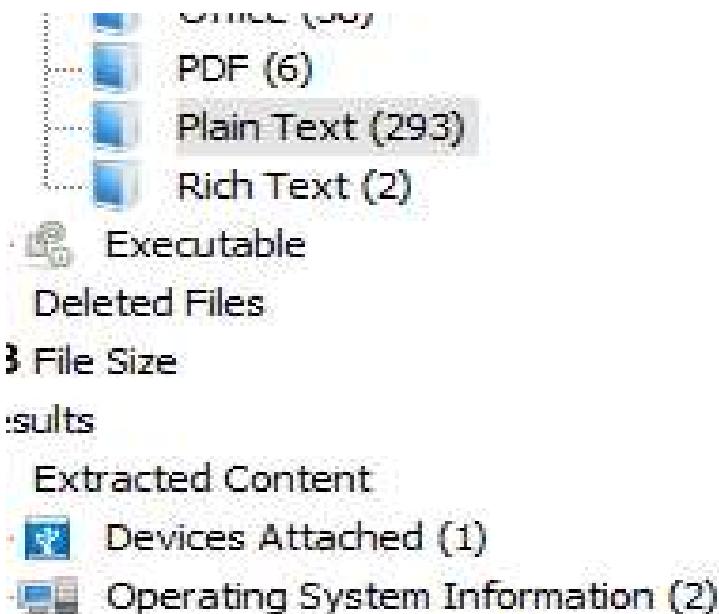
```
...
<stix:Indicators>
  <stix:Indicator xsi:type="indicator:IndicatorType" id="apinto:Indicador-83f51b6a-8512-4194-84bb-65744ad6604f"
    timestamp="2017-01-13T00:00:00.000000Z">
    <indicator:Title>Known IP address</indicator:Title>
    <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">IP Watchlist</indicator:Type>
    <indicator:Observable id="apinto:Observable-7b9e4a6f-513a-407d-9456-62f078cfdf0b">
      <cybox:Object id="apinto:Object-de674b6f-a5f4-4ee4-9360-1b65877354d7">
        <cybox:Properties xsi:type="AddressObject:AddressObjectType" category="ipv4-addr">
          <AddressObject:Address_Value condition="Equals">192.168.1.111</AddressObject:Address_Value>
        </cybox:Properties>
      </cybox:Object>
    </indicator:Observable>
  </stix:Indicator>
</stix:Indicators>
<stix:TTPs>
  <stix:TTP xsi:type="ttp:TTPType" id="apinto:TTP-83fe262c-0f34-4178-be3f-e96328fa1ee6" timestamp="2017-01-13T00:00:00.000000Z">
    <ttp:Title>Potentially dangerous equipment!</ttp:Title>
  </stix:TTP>
</stix:TTPs>
...

```

EXPORTING EVIDENCES

Autopsy allows to export files to:

- Analyse with other tools
- Compare
- Archive



Bibliography

Autopsy User's Guide, Autopsy User Documentation (version 4.19.2)

<https://github.com/sleuthkit/autopsy/tree/develop/docs/doxygen-user>

Autopsy User Documentation

<https://sleuthkit.org/autopsy/docs/user-docs/4.19.2>

Credits

The original author of these slides is António Pinto, adapted and updated by Miguel Frade, Baltazar Rodrigues and Artur Varanda

On 20-09-2004 a computer was found abandoned and it is suspected that this computer was used for hacking purposes. The suspect, Greg Schardt, uses the nickname "Mr. Evil" and some of his associates have said that he would park his vehicle within range of Wireless Access Points where he would then intercept Internet traffic, attempting to get credit card numbers, usernames & passwords.

Class 05 - LAB01 – Image Analysis with Autopsy

1. Download the PC drive images from link available on *Moodle*
2. Create a new case in Autopsy and start automated processing
3. Answer the questions
4. Generate a report by running the STIX sample file against the data sources

ANY QUESTIONS?





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Storage Devices

Artur Varanda

School Year 2021-2022

HDD

The importance of hard drives:

- are the primary form of non-volatile data storage
- they are the main source of digital evidence
 - ✓ but progressively replaced by SSD (discussed later on)
 - ✓ SSDs present new challenges to digital investigation

Main topics:

- physical interfaces and their main characteristics
- hidden areas

Direct access (without BIOS):

- reading and writing data directly through the hard disk controller
 - ✓ the software needs to know how to address the controller and how to issue commands to it
 - ✓ it needs to know the commands code for: read, write, . . .
 - ✓ it needs also how to query the hard disk for details such as type and size
 - ✓ this method is more complex, but also faster
 - ✓ modern OS perform direct accesses to disks

Access with BIOS

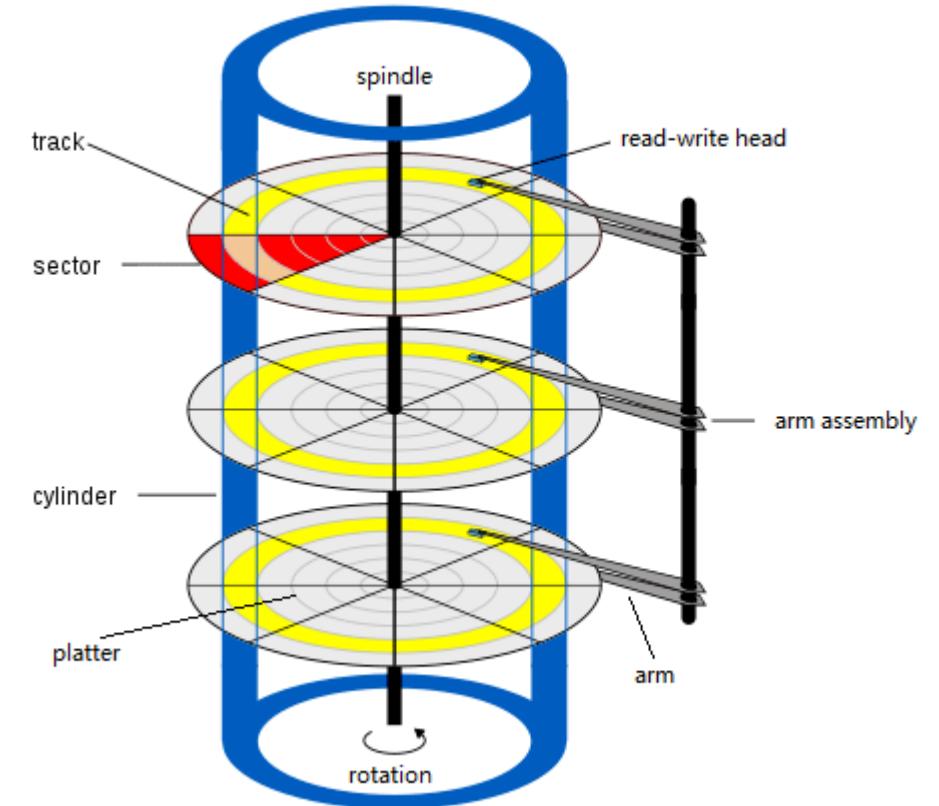
- slower than direct access
- but simpler, the BIOS does all the work
- the BIOS provides services to the software to communicate with the hardware
 - ✓ INT 13h and extended INT 13h
- nowadays it is only used in the boot process

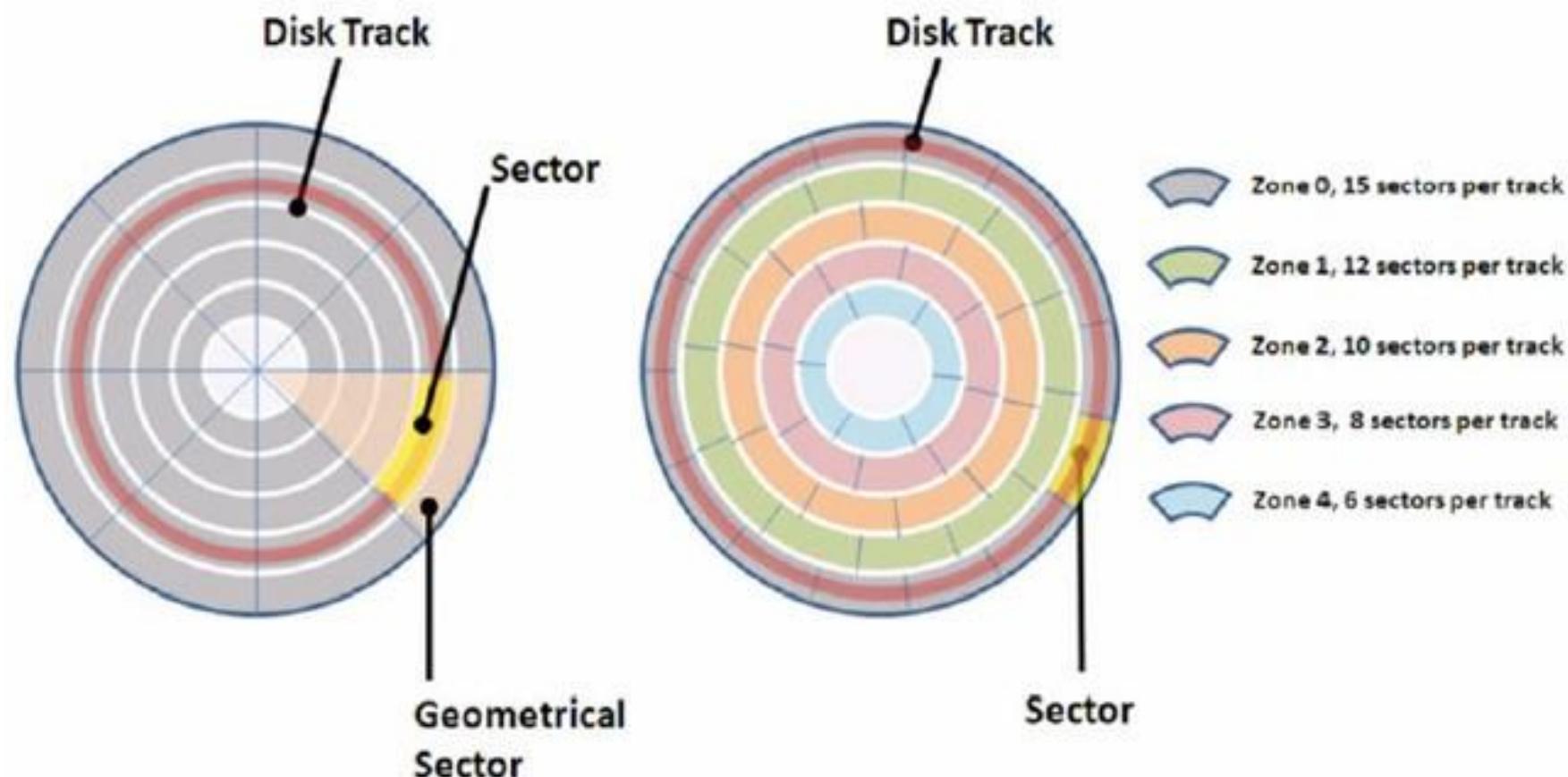
Low level format

- to create data structures
 - ✓ tracks - addresses start from outside
 - ✓ cylinders - all tracks at a given address on all platters $C \in [0, \max_C]$
 - tracks can be addressed by the head number $H \in [0, \max_H]$
 - ✓ sectors - subdivision of tracks, typically 512 bytes $S \in [1, \max_S]$

Get one sector CHS

- Cylinder address (C)
- Head number (H)
- Sector address (S)





Cylinder, Head, Sector (CHS) – used only on older systems

- maximum addressable capacity 504 MB
- way around the problem with fake geometry
- but this translation was limited to address a maximum of 8,1 GB

Logical Block Address (LBA)

- each sector has a unique address
- the software doesn't need to know the disk geometry
 - ✓ used in some file systems: Linux, BSD, MAC OS, . . .
- still use CHS: FAT, NTFS
- *LBA/CHS* conversion:
 - ✓ $LBA = (C \times max_H + H) \times max_S + (S - 1)$

(http://en.wikipedia.org/wiki/Logical_block_addressing)

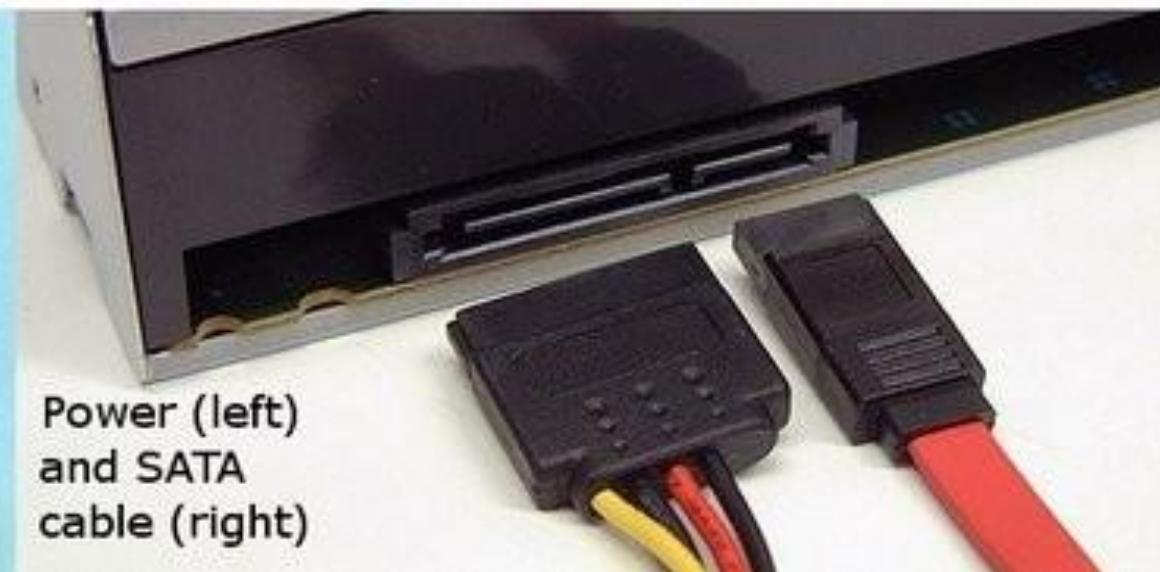
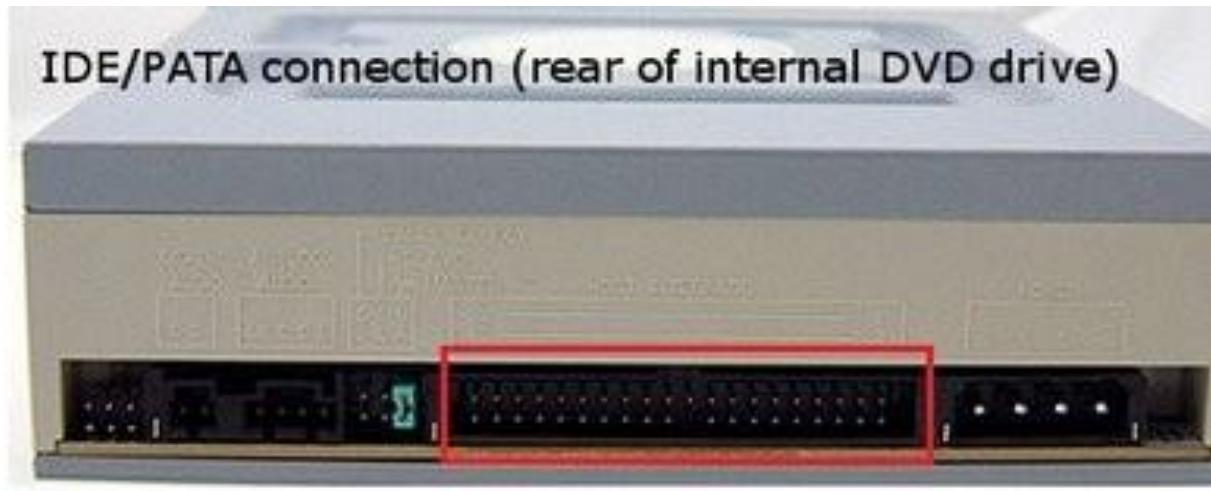
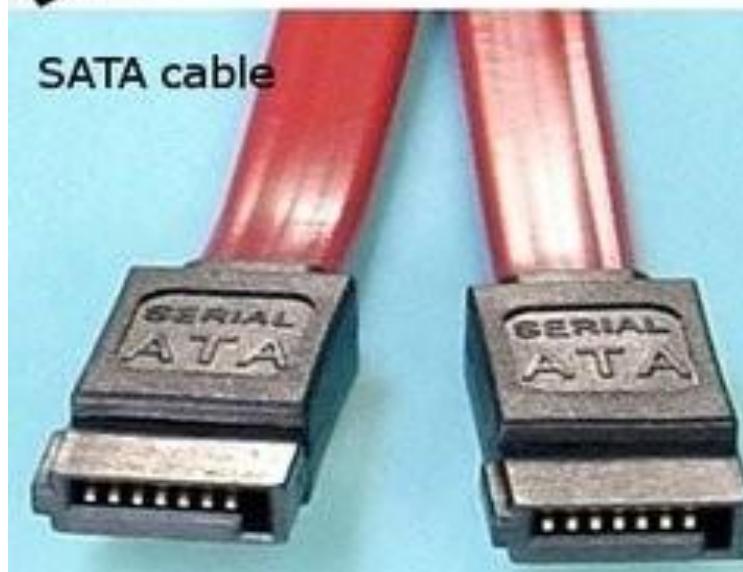
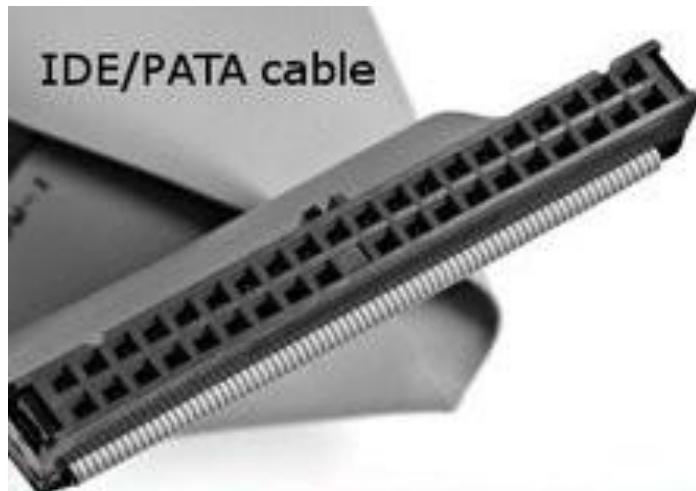
ATA Interface

(Advanced Technology Attachment)

Evolution of Advanced Technology Attachment (ATA) interface

Name (year)	Synonyms	Max. Mbps	Observations
ATA-1 (1994)	IDE	8,3	cable with 40 wires, 16 bits in parallel
ATA-2 (1996)	EIDE	16,7	2 devices per cable
ATA-3 (1997)		16,7	add SMART and passwords
ATA/ATAPI-4 (1998)		33,3	support for removable devices (CD-ROM, DVD, . . .)
ATA/ATAPI-5 (2000)		66,7	80 wires cable, to lower interferences
ATA/ATAPI-6 (2001)		100	LBA addresses with 48 bits
ATA/ATAPI-7 (2002)		133	
SATA 1.0 (2003)		1 500	serial cable, 1 bit after the other
SATA 2.0 (2004)		3 000	
SATA 3.0 (2009)		6 000	
SATA 3.1 (2011)		6 000	added mini-SATA (for SSD)
SATA 3.2 (2013)	SATA Express	16 000	added PCI Express

IDE and SATA connections

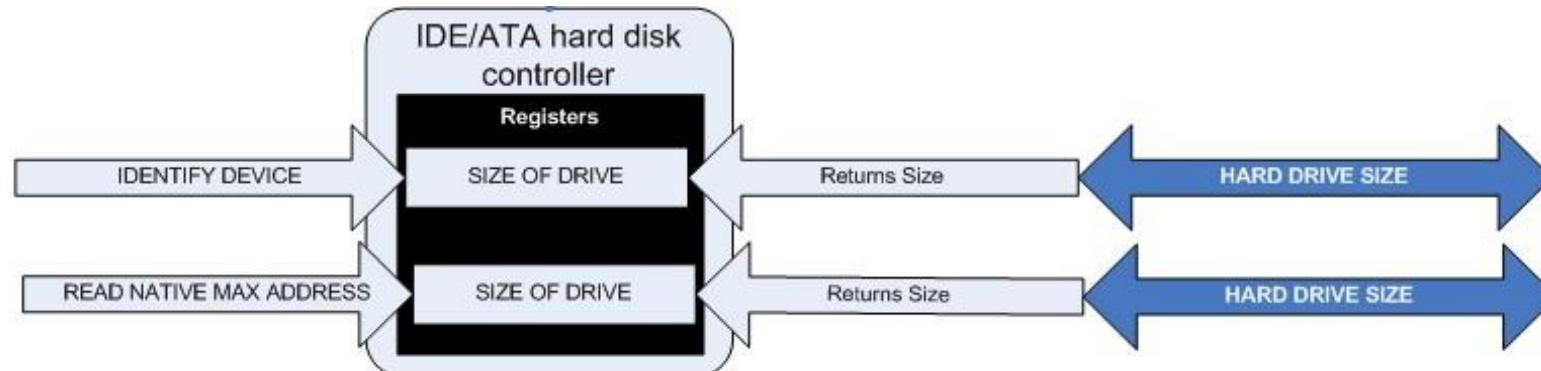


Host Protected Area (HPA)

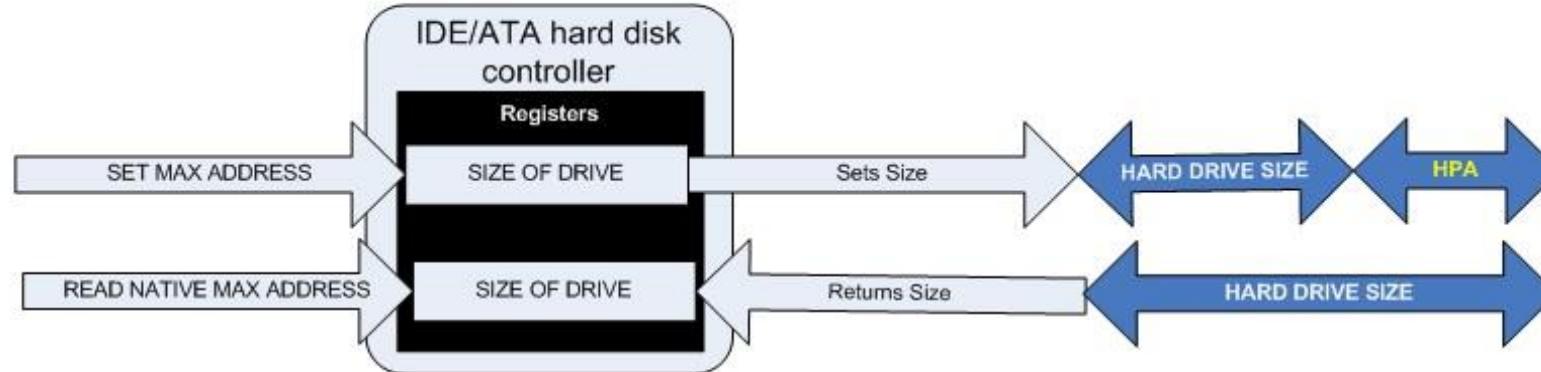
- added with ATA-4
- special area to store vendors data
 - ✓ size can be zero bytes
 - ✓ guaranteed persistence – it won't be erased with a format
- it is located at the end of the disk
- requires reconfiguration of the disk to be accessible
- it can be used to:
 - ✓ reduce the disk size for the old BIOS to recognize the drive
 - ✓ to store diagnostic applications
 - ✓ pre-loaded OS (e. g. dedicated buttons to web OS)
 - ✓ system recovery (e. g. IBM, LG, . . .)
 - ✓ anti-theft tools
 - ✓ but, it can also be used to hide illegal files
 - ✓ some rootkits are able to hide themselves to avoid detection by anti-virus
 - ✓ some NSA exploits are known to use HPA to guarantee persistence

How to create and check for HPA

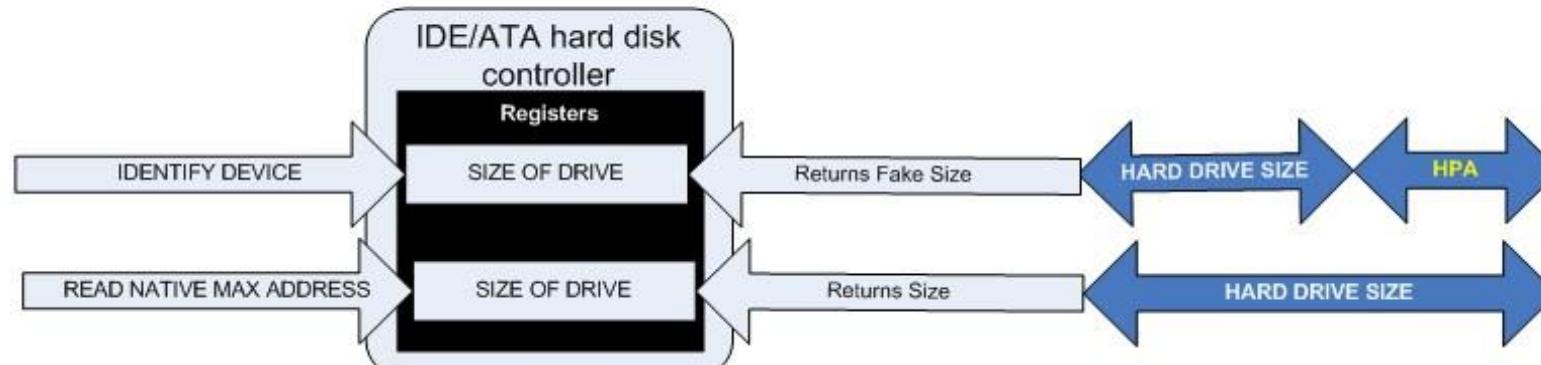
1



2



3



How to identify HPA

On Linux command line:

- at boot time

dmesg | less

[...]

hdb: Host Protected Area detected.

current capacity is 12000 sectors (6 MB)

native capacity is 120103200 sectors (61492 MB)

- by comparing size values

sudo hdparm -N /dev/sdX # replace X with the device letter, $X \in \{a, b, c, \dots\}$

/dev/sdX:

max sectors = 976773168/976773168, HPA is disabled

- to create an HPA

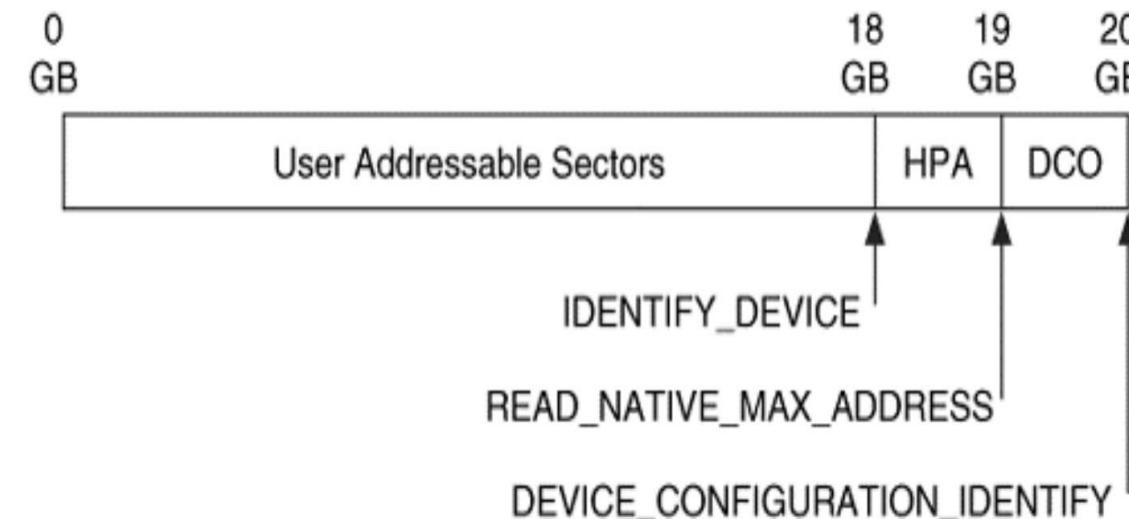
sudo hdparm -N pZZZZZ /dev/sdX # ZZZZZ is the number of visible sectors

Linux tools are free, but there are many more:

http://en.wikipedia.org/wiki/Host_protected_area

Device Configuration Overlay (DCO)

- added in ATA-6
- with DCO, both BIOS and OS see the same size
- DCO removable is permanent (HPA remotion can be temporary)
- allows to hide the disks real capacity
 - ✓ PC makers can buy different brands of discs with different sizes and set them to have exactly the same size
- HPA and DCO can coexist on the same disk



- on the Linux command line

```
hdparm --dco-identify /dev/sdX          # replace X with the device letter, X ∈ {a, b, c, . . .}
```

/dev/sdX:

DCO Revision: 0x0002

The following features can be selectively disabled via DCO:

(...)

Real max sectors: 976773168

DCO can be created

(...)

- compare values with

```
hdparm -Iv /dev/sdX
```

/dev/sdX:

multcount = 0 (off)

IO_support = 1 (32-bit)

readonly = 0 (off)

readahead = 256 (on)

geometry = 60801/255/63, *sectors* = 976773168, *start* = 0

(...)

LBA48 user addressable sectors: 976773168 # if smaller, there is a DCO area

(...)

Linux command line:

- with `hdparm` tool
- it is possible to remove, but not to create a new one
- **WARNING – it can destroy data permanently**
- to remove DCO and set the disk with the real size

```
hdparm --yes-i-know-what-i-am-doing --dco-restore /dev/sdX
```

Windows tools

[TAFT \(The ATA Forensics Tool\)](#) says it can detect and modify HPA and DCO (old, it mentions floppy disks!!)

<https://vidstromlabs.com/freetools/taft/>

[SAFE-Block](#) says it can detect HPA and DCO and put them back

<https://www.softpedia.com/get/Security/Security-Related/SAFE-Block.shtml>

more information and tools:

http://www.forensicswiki.org/wiki/DCO_and_HPA

SCSI Interface (Small Computer Systems Interface)

Small Computer Systems Interface <https://en.wikipedia.org/wiki/SCSI>

- can connect up to 8 (or 16) devices on the cable
- commands error checking, with parity (SCSI-1, SCSI-2), or CRC32 (SCSI-3)
- are common in servers and high-performance systems
 - ✓ *SCSI-over-Fibre Channel Protocol* (FCP) – NAS systems
 - ✓ *Serial Attached SCSI* (SAS) – serial cables (allows connection of SATA-2+ devices)
 - ✓ *USB Attached SCSI* (UAS) – external disks
- more expensive than ATA disks
- many kinds of connectors – generates some confusion

SCSI interface evolution

Version	Max. length	Max. throughput	# devices
SCSI-1	6 m	5 MBps	8
Fast SCSI	3 m	10 MBps	8
Fast Wide SCSI	3 m	20 MBps	16
Ultra SCSI	3 m	20 MBps	4
Wide Ultra SCSI	1,5 m	40 MBps	8
Wide Ultra SCSI	3 m	40 MBps	4
Ultra2 SCSI	4 m	40 MBps	8
Wide Ultra2 SCSI	4 m	80 MBps	16
Ultra160 SCSI	4 m	160 MBps	16
Ultra320 SCSI	4 m	320 MBps	16
SAS (2006)	—	3 Gbps	65 535
SAS (2009)	—	6 Gbps	65 535
SAS (2013)	—	12 Gbps	65 535



DB25m (Mac-SCSI)

Aprox: 39mm



C50m (SCSI-1)

Aprox: 65mm



IDC50m (SCSI-1)

Aprox: 70mm



IDC50f (SCSI-1)

Aprox: 67mm



HD50m (SCSI-2)

Aprox: 35mm



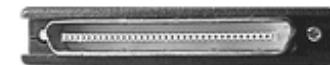
HD68m (SCSI-3)

Aprox: 47mm



HD68f (SCSI-3)

Aprox: 45mm



VHDC68m (SCSI-4)

Aprox: 32mm

Main differences between SCSI and ATA

Feature	ATA	SCSI
Devices per cable	up to 2	up to 8 (or 16)
Communication	by controller	direct by bus
Parallel communication	yes, 16 bits	yes, 8 or 16 bits
Wires per cable	40, or 80	50, or 68
Serial communication	> SATA-1	Serial Attached SCSI (SAS)
Availability	common	high availability system
Fault tolerance	—	power supply
Disk size	limited (older versions)	LBA of 32 or 64 bits
Rotations/minute	4,5k 7,2k 10k	10k 15k
Hidden areas	HPA, DCO	—

NAND Flash memory

Hard Disks Drives (HDD)

- few manufactures:
 - ✓ concentration of manufacturers through purchases and mergers over the years
- mature technology, with many aspects in common:
 - ✓ between disks models and sizes
 - ✓ between manufacturers
- digital research in hard drives is almost the same in all models and brands

Solid-State Drives (SSD)

- basic components are the same or very similar
 - ✓ between manufacturers
 - ✓ between flash memory and SSDs
- but there are important differences:
 - ✓ a flash memory requires driver software – uses CPU
 - ✓ SSD has its own processing unit – doesn't use CPU
 - ✓ firmware between models or manufacturers can be very different

Solid state drives (SSD):

- are mechanically more reliable
 - ✓ have no moving parts and are more resistant to falls
- read speed is independent of the data location (which doesn't happen with HDD)
- power consumption is lower (1h to 2h of increased battery autonomy on a laptop)
- emits no noise or vibrations
- heat less than HDD – *HDD can reach very high temperatures*
- are lighter – don't require a metallic structure as HDDs

DRAM

- older solid state disk (they exist for more than 30 years)
- based on volatile DRAM memory
- require battery or other power source to ensure redundancy
- need of a traditional drive to store data permanently
- used in high-performance systems such as banks, stock exchange, military assets, . . .
- the cost of flash memory is falling more than DRAM → the crossing point was reached in 2004

Flash memory

- non-volatile
- there are 2 categories:
 - ✓ NOR gates NAND gates

With NOR gates

- used for small amounts of memory (< 16MB), e. g. BIOS
- allows very fast readings, but is slow to write and erase (up to 5 seconds)
- supports fewer write cycles (10× less than NAND gates)
- allows to read or write a single byte at a time
- allows local execution, without having to use RAM
 - ✓ uses a SRAM interface that enables to address all bytes

With NAND gates

- provides large bit density → ideal for replacing HDD
- erase and write faster than NOR (up to 4 ms), but slightly slower readings
- reads and writes are made in large blocks of bytes
- disadvantages:
 - ✓ internal management complexity
 - ✓ serial access to data, wear leveling, garbage collection, . . .

NAND Flash memory – is the most common type of flash

- USB pen drives
- Solid State Drives (SSD)

Management of bad blocks

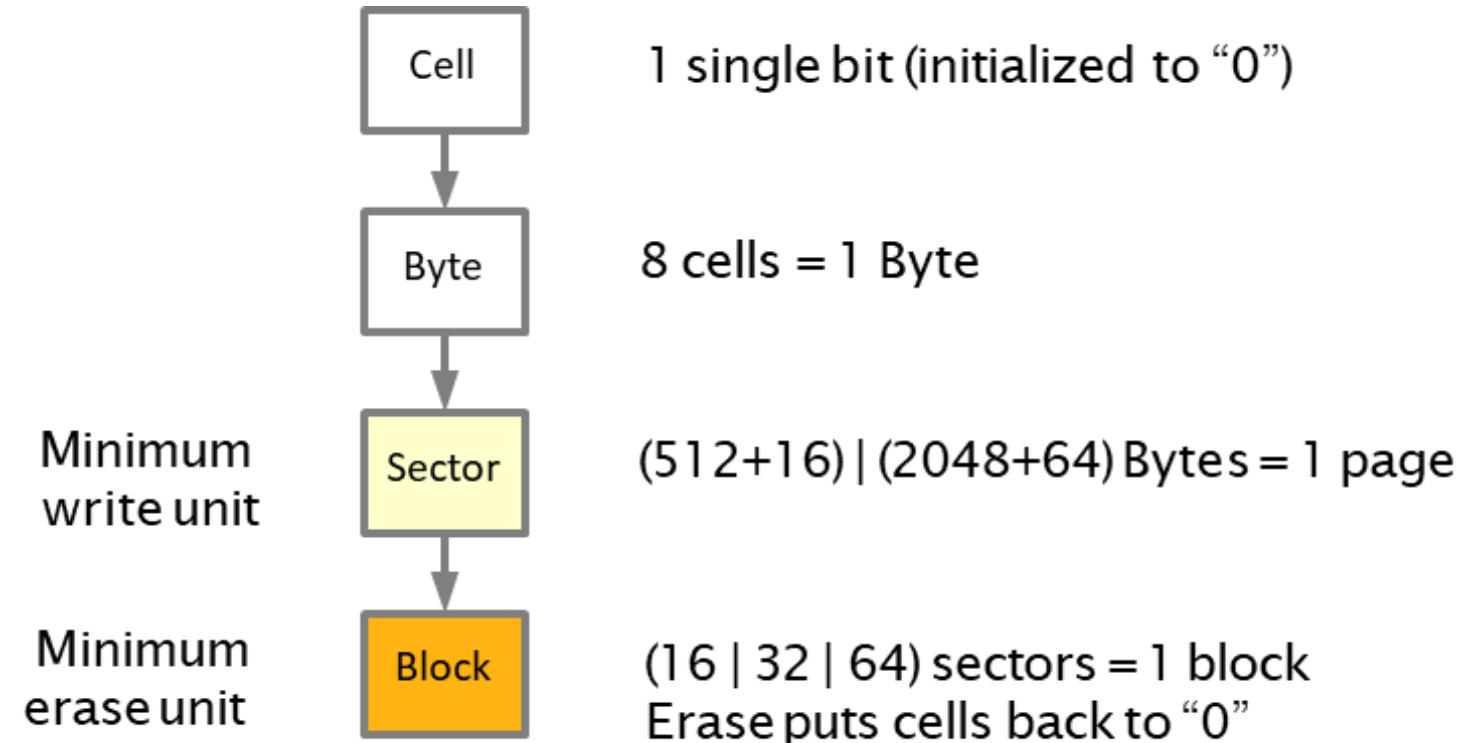
- all devices have bad blocks
- an initial test to identify bad blocks is required
 - ✓ the cost of creating chips without defects does not pay off
 - ✓ it is preferable to put capacity in excess and then remove the addresses with bad blocks

Inner working of a NAND chip

- at rest = 1 (stores the value 1) at load = 0 (stores the value zero)
- to increase density, they can be produced in layers: MLC (multi layer chip)
 - ✓ several bits have to be read/written simultaneously
 - ✓ allows more capacity, but has lower performance than the single layer chip (SLC)
 - ✓ cheaper

Data access

- data access in grid with word lines (16 bits)
- minimum writing unit is a sector with a size multiple of word lines
 - ✓ HDD: 1 sector = 512 Bytes → minimum read and write unit
 - ✓ SSD: 1 sector = [512, 2048] Bytes
 - ✓ depends on several factors, such as a manufacturer and disk capacity
 - ✓ minimum writing unit differs from minimum erasing unit
 - ✓ it is not possible to erase a single sector
 - ✓ data as to be erased by blocks – the electric charge to erase is similar to a photographic flash



What is level wearing?

- SSDs change data location to level the number of writing operations across all cells
- if the location was always the same for changing date, those cells would burn out quickly
 - ✓ each cells stands $\approx 100\,000$ erasing cycles
- firmware is responsible for doing the level wearing in an automated fashion
- it is also required a garbage collection system to identify freed sectors, that weren't erased yet

How garbage collection works?

- it keeps a list of freed sectors
- one block is erase only when all sectors are marked as free
- even without a data connection, the garbage collection keeps running on its own and restarts in case of a power outage

On HDD

- it is possible to read and change the information in a specific sector
- reading specific sectors is common in digital investigation

On SSD

- when a file is modified it is not possible to save it in the same sector
- because it is not possible to write into cells before erasing them
- the file is saved into a new empty sector and the original one goes to the garbage collection list
- the old sector is erased only when all sectors in the same block are freed

SSD erasing and wear leveling mechanisms consequences:

- when we ask the OS to delete, the data is not actually deleted, it goes to the garbage collection list
- only a few flags change
- it is effectively deleted only when the erase routine is executed by the garbage collection algorithm
 - ✓ for example: 1 block of 64 sectors × 2048 B = 128 kB
- the physical location of a sector changes over time → **this is a problem for data acquisition at the physical level**
- even without an OS commanding, garbage collection operations can happen (power on is enough)
 - ✓ a write blocker does not prevent garbage collection or wear leveling operations
 - ✓ this limits the availability of “deleted” or interesting slack data

Pen USB vs SSD Comparison

Main characteristics:

- doesn't have its own processor
- so, it requires a mass storage software driver to manage operations:
 - ✓ file system → block device services → mount/read/write/delete virtual sectors
 - ✓ identity/read/write/erase → flash memory
- uses the CPU to:
 - ✓ calculate ECC
 - ✓ bad blocks management
 - ✓ wear leveling, . . .

Main characteristics:

- has its own processor to manage operations:
 - ✓ wear leveling, bad blocks
 - ✓ erasing cycles counts, sectors initial location
 - ✓ error checking code
- so, it doesn't require a software driver
- SATA commands are emulated to guarantee compatibility
- garbage collection only needs power to start operations
 - ✓ a write blocker doesn't stop this operations
- repairing
 - ✓ easy on HDD: you can replace controller cards, heads, . . . and use the same platters
 - ✓ difficult on SSD: too complex, only possible on highly specialized labs

SSD Connectors, Interfaces and Transfer Protocols

Connectors:

- layer 1 – physical interface to connect devices. Examples: M.2, RJ45, . . .

Link interfaces:

- layer 2 – handles data encoding. Examples: PCIe and SATA

Transport protocols:

- layer 3 – handles data communication. Examples: NVMe, AHCI and IDE.

M.2 – One connector, several transport protocols:

- M.2 connector = SATA link interface + SATA transport protocol
- M.2 connector = PCIe link interface + AHCI transport protocol
- M.2 connector = PCIe link interface + NVMe transport protocol

M.2, formerly known as the Next Generation Form Factor (NGFF)

- specification for internally mounted computer expansion cards
- replaces the mSATA standard
- more flexible physical specification make it more suitable to:
 - ✓ solid-state storage
 - ✓ particularly for the use in small devices such as ultrabooks or tablets

protocols:

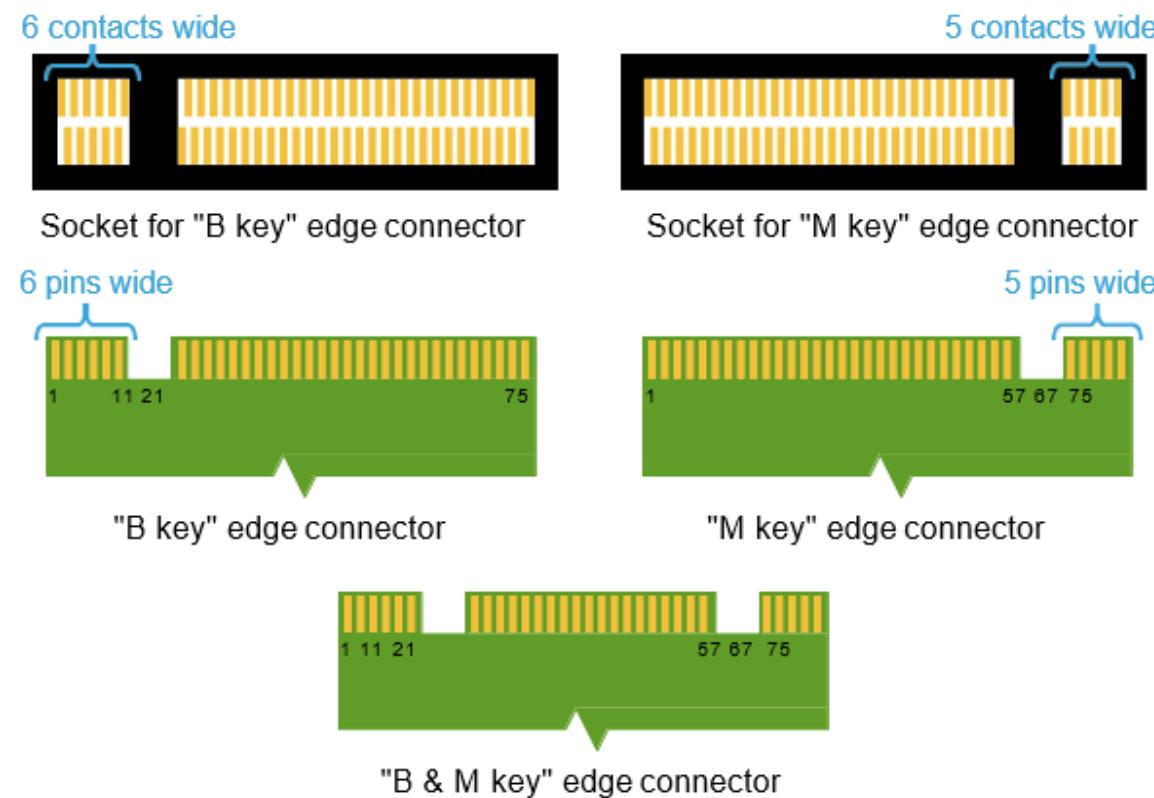
- link interface: PCI Express 3.0 (up to four lanes)
- transport protocol: Non-Volatile Memory Express (NVMe) as the logical device interface for

M.2 PCI Express SSDs

- ✓ NVMe is designed to fully utilize the capability of high-speed PCIe storage devices to perform many I/O operations in parallel
- Serial ATA 3.0 and USB 3.0 (a single logical port for both)
- the manufacturer selects which interfaces are supported

The M.2 connector has different keying notches:

- to denote various purposes and capabilities of M.2 hosts (SSD, WiFi, 4G modem, . . .)
- to prevent plugging into feature-incompatible host connectors



Source: Wikipedia, <https://en.wikipedia.org/wiki/M.2>

- PCIe is a high-speed serial transfer interface standard
- M.2 connector supports up to four PCIe channels

Evolution of PCIe

Version	Year	Transfer rate	Throughput (Channel width × transfers/second)				
			× 1	× 2	× 4	× 8	× 16
1.0	2003	2,5 GT/s	250 MB/s	500 MB/s	1,00 GB/s	2,00 GB/s	4,00 GB/s
2.0	2007	5,0 GT/s	500 MB/s	1,00 GB/s	2,00 GB/s	4,00 GB/s	8,00 GB/s
3.0	2010	8,0 GT/s	984,6 MB/s	1,97 GB/s	3,94 GB/s	7,88 GB/s	15,75 GB/s
4.0	2017	16,0 GT/s	1969 MB/s	3,94 GB/s	7,88 GB/s	15,75 GB/s	31,51 GB/s
5.0	2019(?)	32,0 GT/s	3938 MB/s	7,88 GB/s	15,75 GB/s	31,51 GB/s	63,02 GB/s

GT/s = Gigatransfers per second

Apple SSD Proprietary Connectors

Generation	Year	Connector	Interface
G1	2010	6+12	mSATA 3
G2	2011	7+17	mSATA 3
G3	2012	12+16	PCIe 2.0 ×2
G4	2013	12+16	PCIe 3.0 ×4
G5A	2015	22+34	PCIe 3.0 ×4 NVMe
G5B		12+16	

More info.: <https://beetstech.com/blog/apple-proprietary-ssd-ultimate-guide-to-specs-and-upgrades>

Some examples of Apple SSD connectors

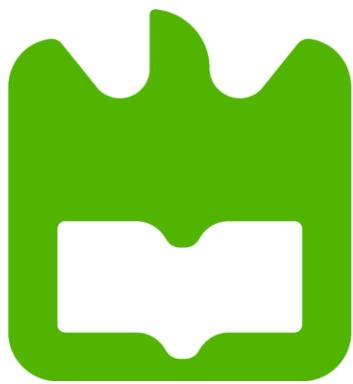


Do the following exercise:

06-Lab 1 – Add a RAW disk to a virtual machine

Do the following exercise:
06-Lab 2 - Smartmontools





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Volumes and Partitions

Artur Varanda

School Year 2021-2022

Volumes

- allows joining several small volumes into a larger one
- or split the physical storage space into smaller spaces
- is a collection of sectors
 - ✓ for the OS those sectors are consecutive – volume level
 - ✓ but at the physical level they **may not be consecutive**

Partitions

- is a particular case of volumes
- a partition is a set of **consecutive** sectors
- the confusion between partitions and volumes is common

Volumes

- are structures that define the space occupied by the file system
- you need to know the volume structure to analyze its contents
 - ✓ if a drive is corrupted you may not be able to read the volume structure
 - ✓ a volume might have been deleted in an attempt to hide data

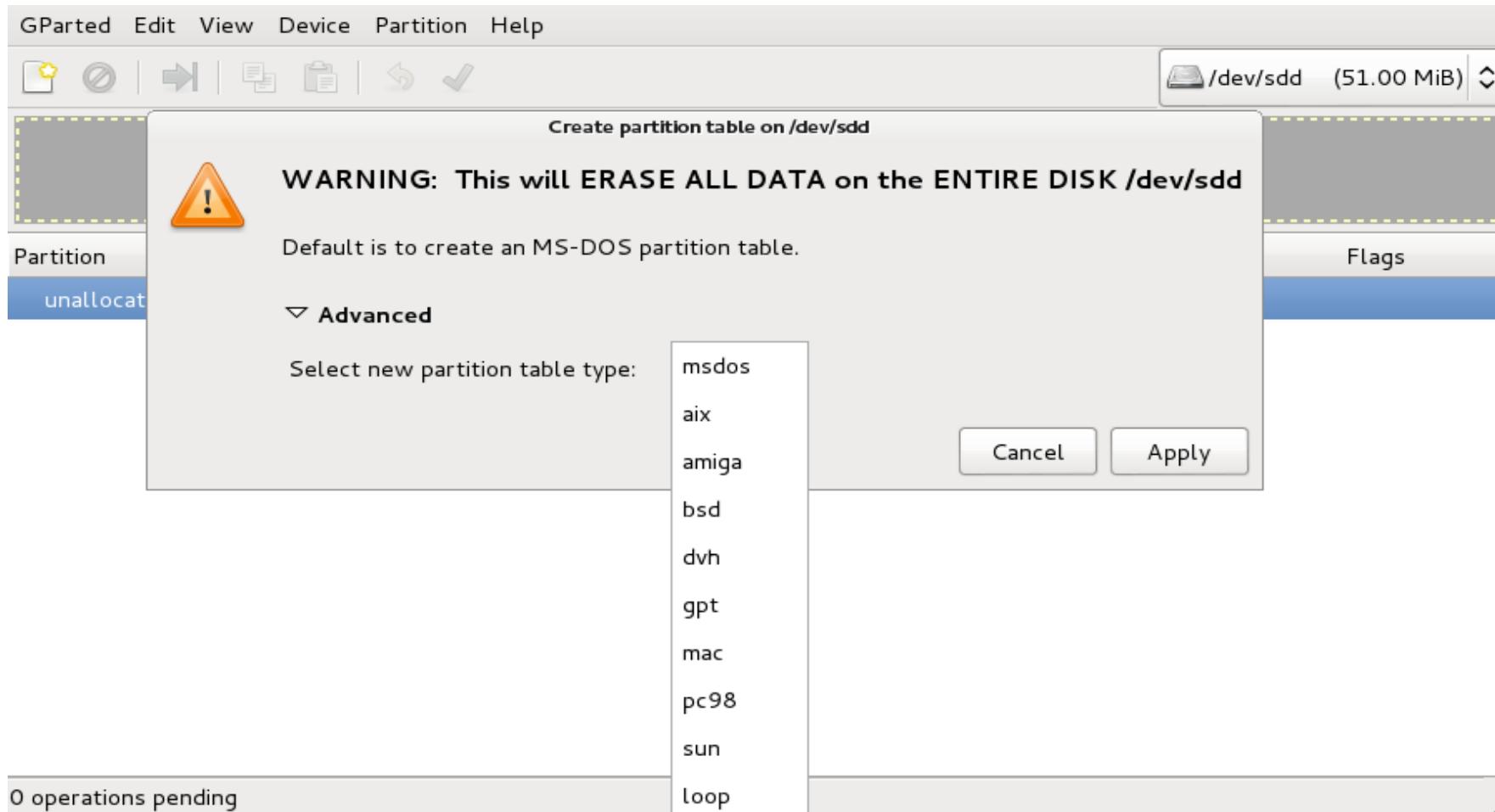
Why use volumes?

- some OS use a volume to store RAM data when they hibernate, *e. g.* linux
- to separate the OS files from the users' files
- to allow dual boot, *e. g.* windows and linux
- to aggregate volumes
 - ✓ to get more space for the file system
 - ✓ to get redundancy and prevent data loss due to drive failures

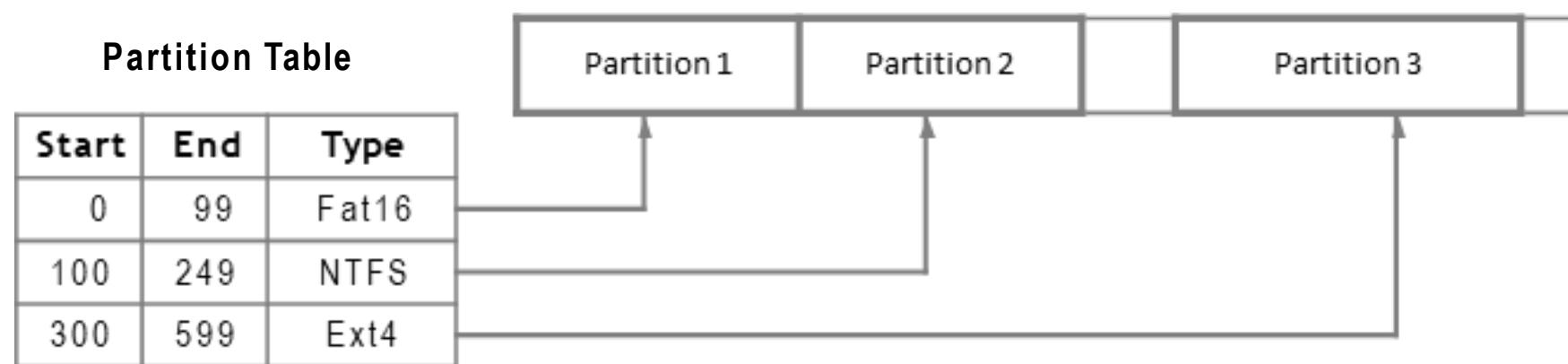
Partition Tables

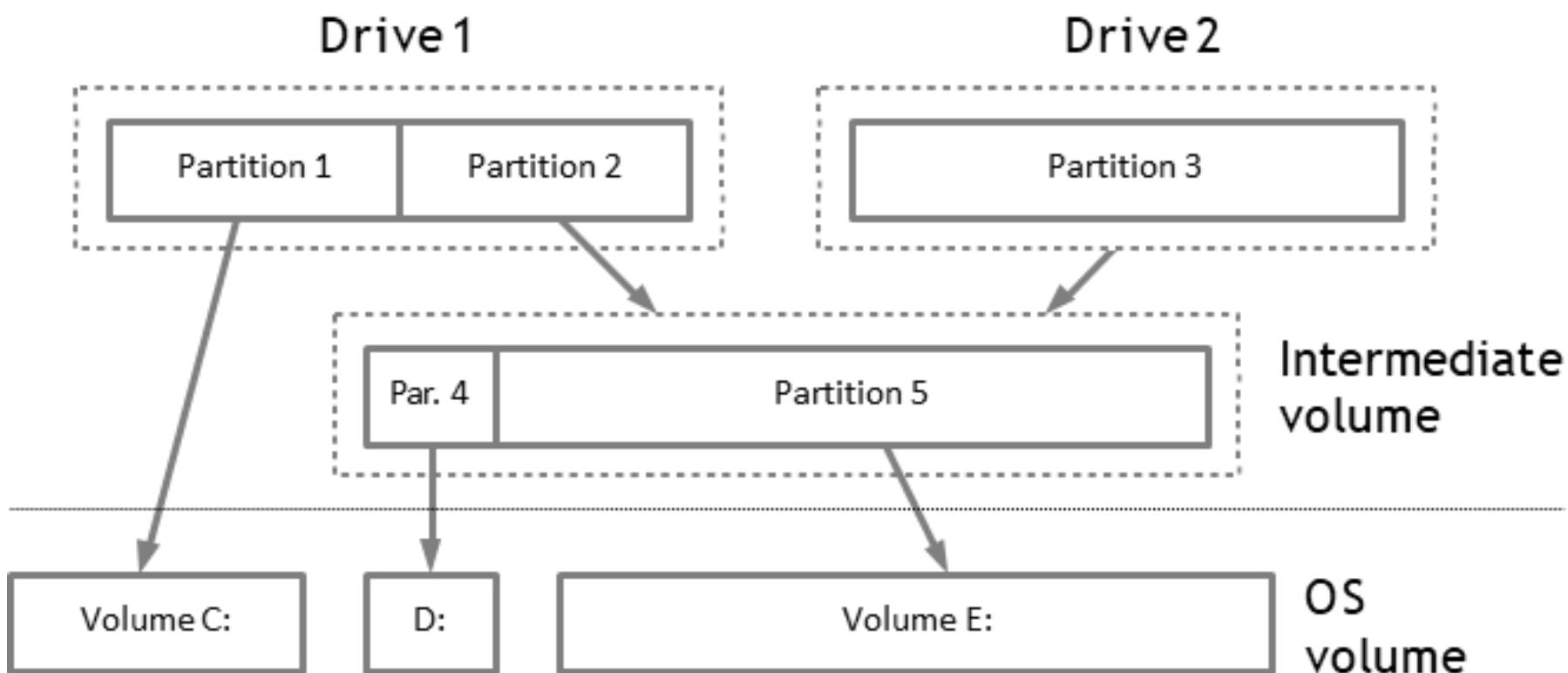
What is the partition table?

- is a data structure to store information about the partitions on a given drive
- there are several types of partition tables
- partition table depend on the used OS and not on the drives' physical interface

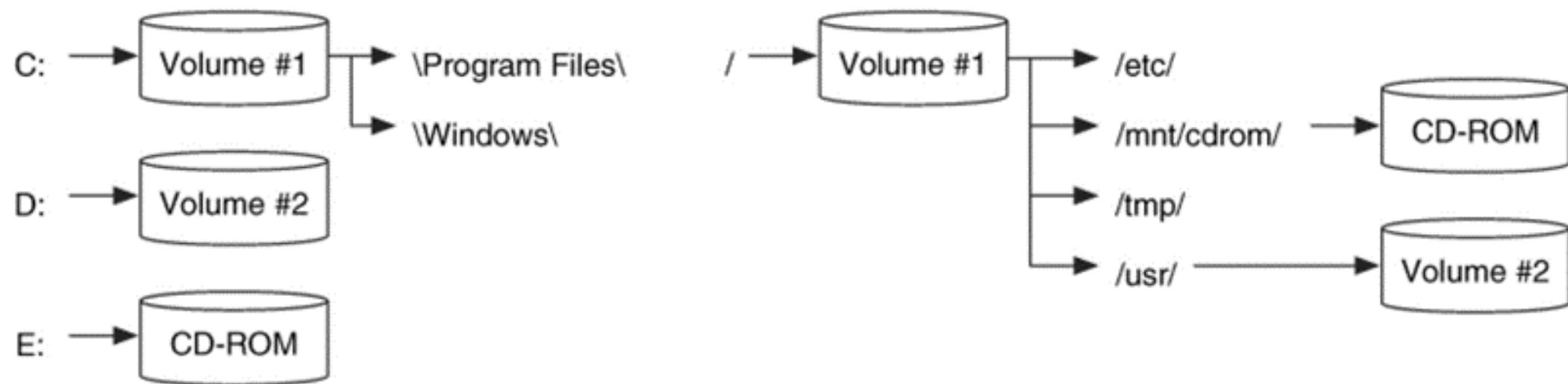


- **essential data:** begin and end sectors
- **non essential data:** type of partitions and description (can be fake)
- the start and end sectors are usually indistinguishable
 - ✓ if the partition table is corrupted it can be estimated based on prior knowledge
- there may be unallocated sectors between partitions





Partitions and volumes – Windows vs Linux



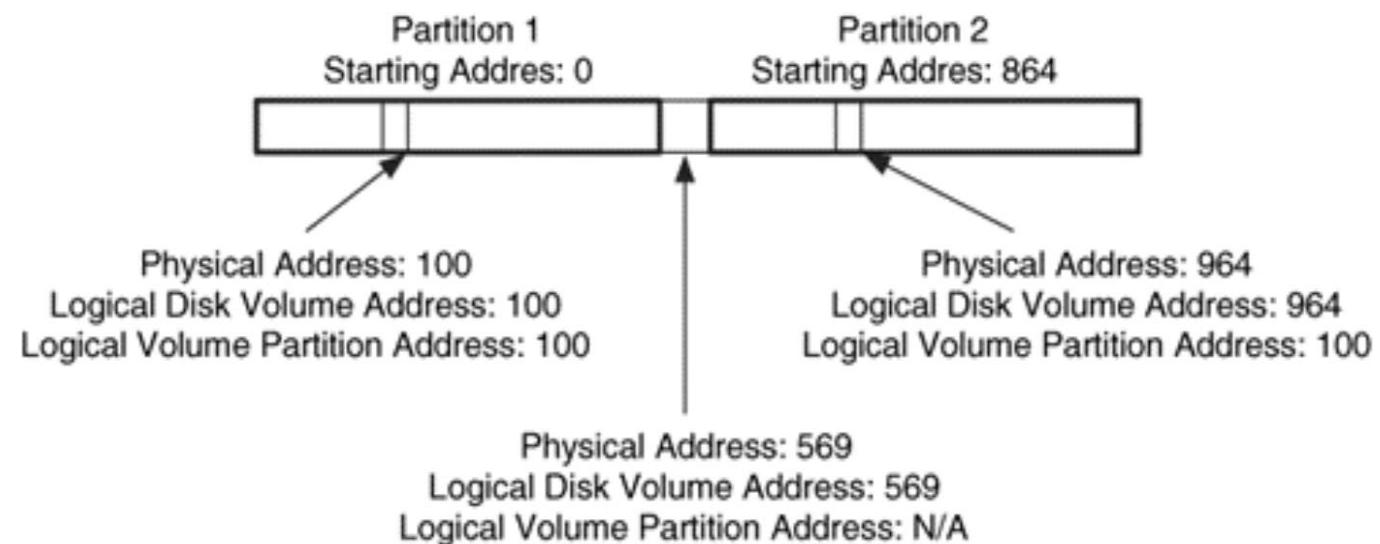
Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

LBA addresses

- maps the physical sectors of the drive
- the first sector is always 0 (zero)
- it cannot be used to address volume sectors
 - ✓ a volume is a collections of sectors, but
 - ✓ may not be consecutive and can even be in different drives

Layers of addresses

- physical address
- logical disk volume addresses – equal to the physical address
- logical volume partition addresses – each partition has its own address space



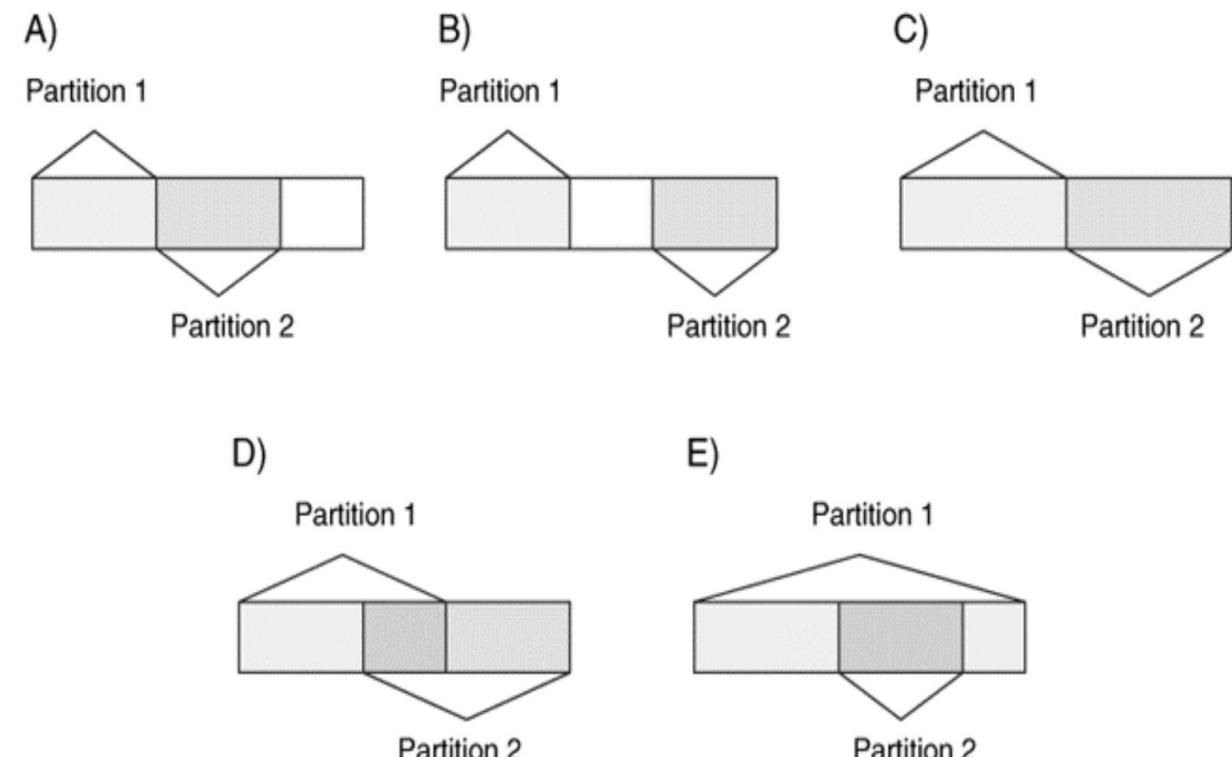
Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

Procedure

- performed automatically by tools most of the times
 - ✓ except if corruption has occurred
- steps that must be performed (by software or manually):
 - ✓ read partition table
 - ✓ identify the partition layout (start and end sectors)
 - ✓ analyze the unallocated space – it may contain data from a previous OS
 - ✓ a partition may be part of a volume with multiple partitions

Consistency checks

- does the last partition ends at the end of the parent volume?
- are the partitions consecutive?
- are there any overlap between partitions?
 - ✓ may happen if the partition table is corrupted



Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

How to recover partitions

- they may have been deleted to hinder the investigation
- or the partition table may have become corrupted
- usually partitions have file system, so we can search for their patterns
 - ✓ FAT has the values 0x55 and 0xAA on bytes 510 and 511 of the first sector
- gpart tool tries to identify partitions based on patterns: `gpart -v disco.dd`
- testdisk is another tool to recover partition tables

Types of Partition Tables

- on personal computers (PCs)
 - ✓ MBR, Apple, removable storage media, GPT, ...
 - GPT is required for the UEFI secure boot on servers
 - ✓ GPT, FreeBSD, Sun Solaris, ...
 - ✓ PC partitions can also be used on servers
 - ✓ the main difference is the frequent use of logic volumes

Common Partitions of PCs

DOS partitions

https://en.wikipedia.org/wiki/Master_boot_record

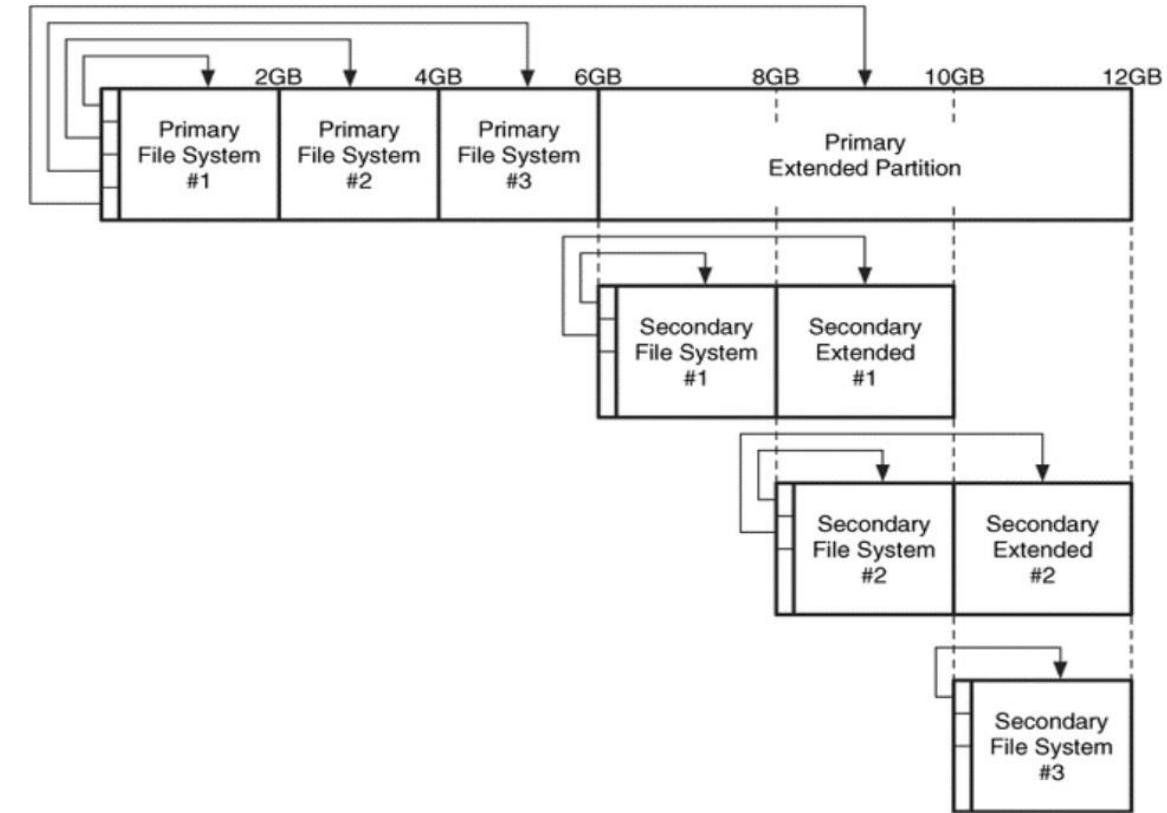
- also known as *Master Boot Record* (MBR)
- created by Microsoft (after Windows 2000 they call them *basic discs*)
- it's the most common partition table type
- it's used in: Microsoft DOS, Microsoft Windows, Linux, FreeBSD and OpenBSD
- MBR:
 - ✓ is located in the first sector (512 bytes)
 - ✓ *boot code* – instructions to process the partition table and to find the OS
 - ✓ partition table
 - ✓ pattern `0xAA55` – to identify the partition table

Structure of a DOS partition

- 4 entries – 4 primary partitions is the limit
- each one has:
 - ✓ begin and end address in CHS (< 8 GB) LBA address for large drives (several TB) amount of sectors in the partition
 - ✓ file system type stored in the partition (FAT, NTFS, EXT4, . . .)
 - Windows depends on this to mount the partition
 - it can be used to hide partitions from Windows OS
 - Linux ignores this value and supports a different FS from the one stored in the partition table
 - ✓ *flags* – allows to mark the boot partition (*bootable*)

Extended Partition

- to overcome the 4 primary partition limit
- always the last entry in the MBR
- allows to create several logical partitions
- types of extended partitions:
 - ✓ *DOS Extended, Windows 95 Extended and Linux Extended*
- usually there is only one extended partition
 - ✓ but it is possible to create more than one
 - ✓ few forensic tools support this



Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

Characteristics:

- located in the first 446 bytes of the first sector of 512 bytes (MBR)
- the boot code from Microsoft processes the partition table:
 - ✓ searches the bootable partition (with *boot flag* on)
 - ✓ the partition code is specific to each OS
- some virus are known to install themselves in MBR
- multiple OS:
 - ✓ boot selector from Windows
 - ✓ or MBR replaced by a specific application, like GRUB, LILO, . . .

Command to extract MBR in Linux

```
dd if=disco.dd bs=512 skip=0 count=1 | xxd
```

```
0000000: eb63 90d0 bc00 7c8e c08e d8be 007c bf00 .c....|.....|...
...
0000180: 7de8 2e00 cd18 ebf 4752 5542 2000 4765 }.....GRUB .Ge
0000190: 6f6d 0048 6172 6420 4469 736b 0052 6561 om.Hard Disk.Rea
00001a0: 6400 2045 7272 6f72 0d0a 00bb 0100 b40e d. Error.....
00001b0: cd10 ac3c 0075 f4c3 1996 6b49 0000 0020 ...<.u....kI...
00001c0: 2100 1cf e ffff 0008 0000 0000 2003 80fe !..... .
00001d0: ffff 07fe ffff 0008 2003 00c0 4917 00fe ..... .I...
00001e0: ffff 05fe ffff fecf 691a 0288 ce1f 0000 .....i.....
00001f0: 0000 0000 0000 0000 0000 0000 55aa .....U.
```

- MBR with GRUB boot loader from Linux

Partitions on removable storage:

- floppy disks – don't have partition table
 - ✓ has a FAT12 file system as a single partition
- memory cards – typically use DOS partitions
 - ✓ FAT32 – max volume size 32 GB, max file size 4 GB
 - ✓ exFAT (or FAT64) – max volume size 128 PB, max file size 128 PB
- external hard drives
 - ✓ DOS partitions, commonly sold already with a NTFS file system in place

Optical disks:

- CDs – ISO 9660
 - ✓ the ISO 9660 is very strict about file names
 - ✓ there are extensions (Joliet, Rock Ridge) to overcome these limitations
 - ✓ there are also hybrid CDs:
 - ISO 9660 + Joliet
 - ISO 9660 + Apple HPS+
 - ✓ bootable CDs for *intel* PCs can contain a DOS partition
- CD-R
 - use sessions, each one can be treated as a partition
 - each time data is added to the CD-R a new session is created
 - most OS only show the last created session
 - ✓ on OS X it's possible to see all sessions
 - ✓ on Linux it's possible to manually *mount* previous sessions
 - ✓ on Windows specific software is required e. g. *Iso Buster* - <https://www.isobuster.com/>

Common Partitions of Servers

GPT (GUID Partition Table)

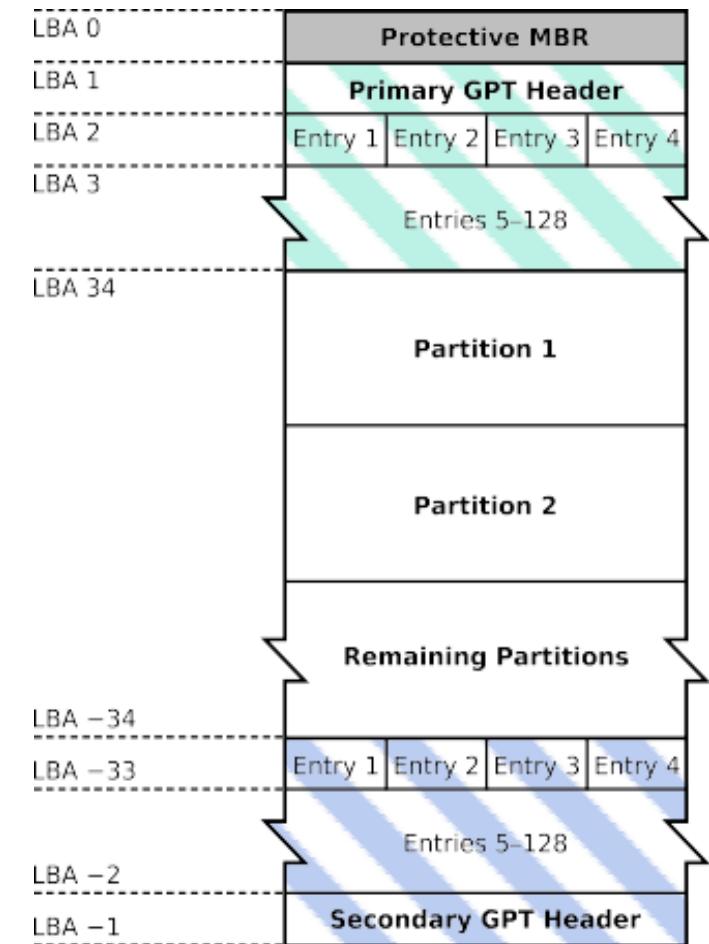
https://en.wikipedia.org/wiki/GUID_Partition_Table

- GUID – globally unique identifiers
- introduced on system with 64-bit Intel Itanium (IA64) processors
- is part of the *Unified Extensible Firmware Interface* (UEFI) standard
 - ✓ replaces the BIOS and can also be used in PCs
- drives are identified with *globally unique identifiers* (GUID)
- uses 64 bits LBA
- Microsoft added support since Windows 2008

Structure:

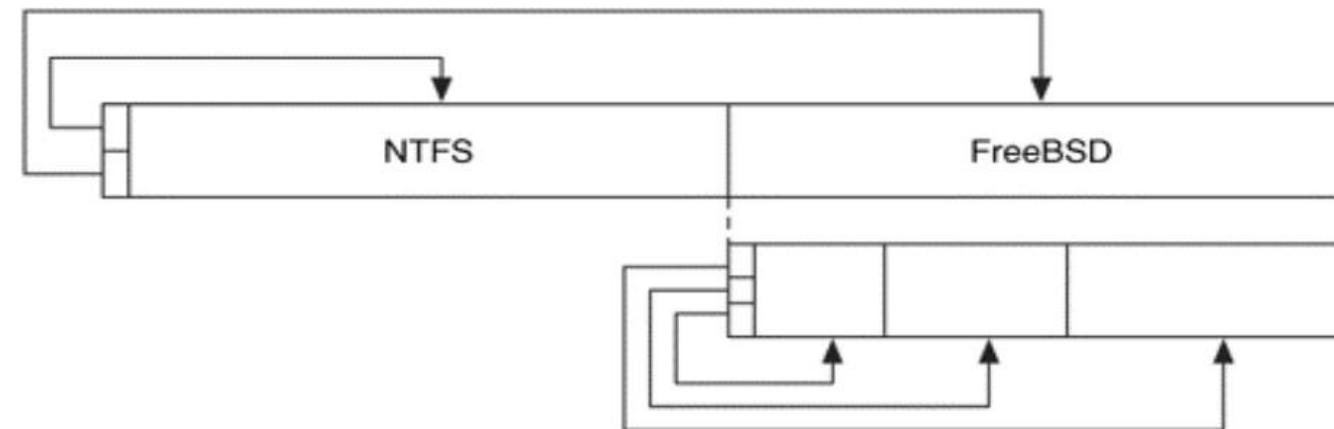
- protective MBR – to prevent non-compatible OS to format the drive
- GPT header – starts on sector 1, defines the size and location of the partition table and has also a checksum
- partition table – supports up to 128 partitions, contains begin and end sectors, type, name, attributes and GUID values (128 bits)
- partition area – the main drive area
- backup area – located in the last sectors of the drive

GUID Partition Table Scheme



Source: https://en.wikipedia.org/wiki/GUID_Partition_Table

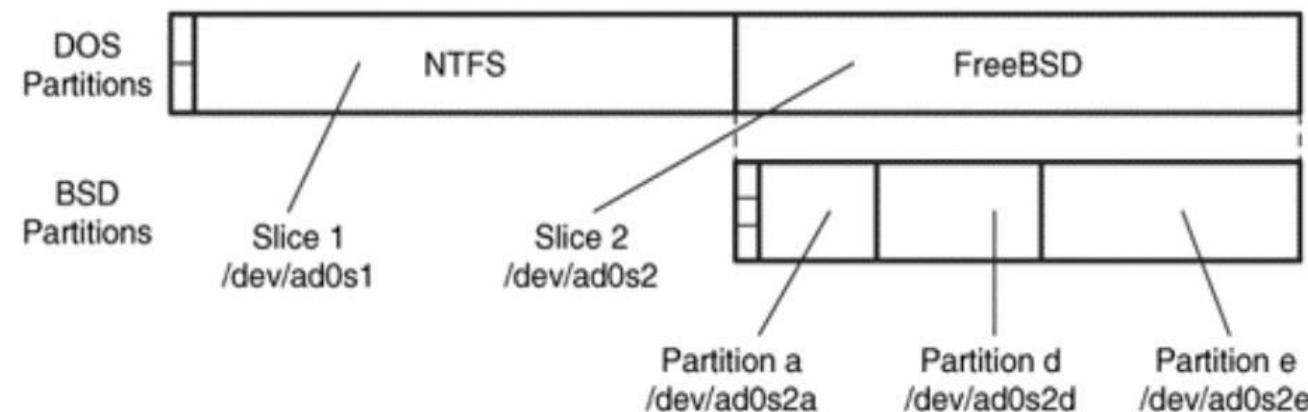
- used on BSD Unix: FreeBSD, OpenBSD and NetBSD
- BSD partitions can coexist with DOS partitions
 - ✓ BSD partition table will be located inside one of the DOS primary partition
 - ✓ FreeBSD allows access to DOS partitions



Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

Partitions naming scheme:

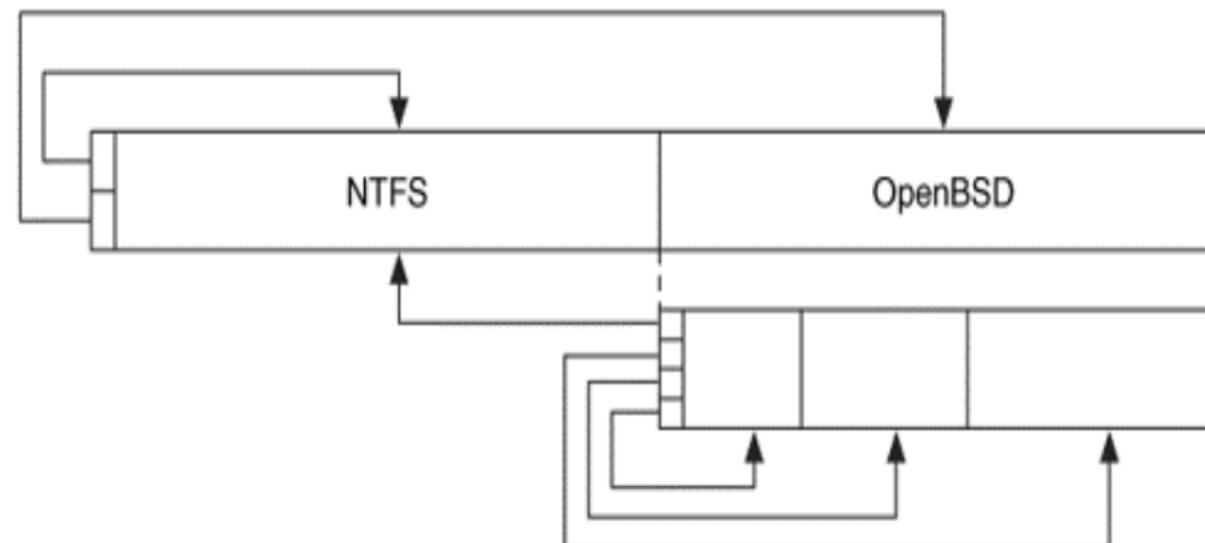
- ATA devices: /dev/ad0
- each DOS partition is a slice (s) → /dev/ad0s1, /dev/ad0s2, ...
- each FreeBSD slice has a letter:
 - ✓ a – root partition
 - ✓ b – swap space
 - ✓ c – full FreeBSD partition
 - ✓ d, e, . . . – the remaining partitions



Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

Main differences with FreeBSD partitions

- after boot, the OS ignores the DOS partition table
 - allows to refer partitions outside its main area
 - base name for the ATA devices: /dev/wd0
 - ✓ there are no *slices*
 - ✓ attributes letters to partitions like FreeBSD: a (root partition), b (swap), . . .



Source: Brian Carrier, "File System Forensic Analysis", Addison-Wesley Professional, March 27, 2005

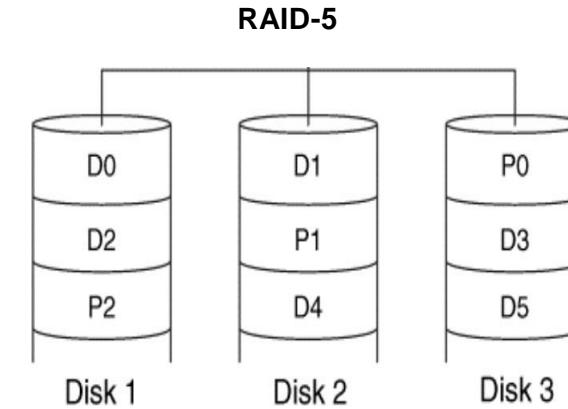
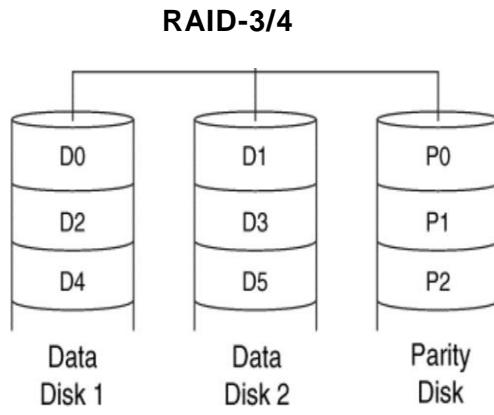
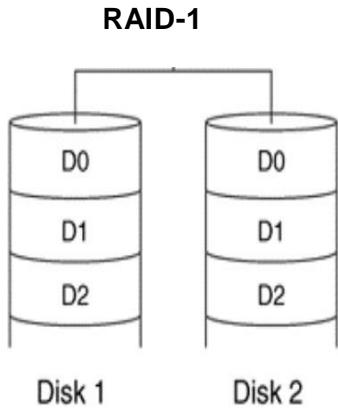
Why?

- improve performance
- prevent faults by adding redundancy
- gather free space from several drives

Types:

- RAID (Redundant Arrays of Inexpensive Disks)
 - ✓ common in high performance systems has many variants: RAID-1, RAID-5, . . .
 - ✓ can be implemented both in hardware, or software (usually at OS level)
- *spanning*
 - ✓ creates a logical volume by adding free space from several smaller volumes
 - ✓ just sums up space, it doesn't have any redundancy or performance gains

- RAID-0 – has no redundancy, but may increase performance
- RAID-1 – consists of an exact copy (or mirror) of a set of data on two or more disks
- RAID-2 – stripes data at the bit (rather than block) level, and uses a Hamming code for error correction
rarely used
- RAID-3 – consists of byte-level striping with a dedicated parity disk **rarely used**
- RAID-4 – consists of block-level striping with a dedicated parity disk, provides good read performance
- RAID-5 – block-level striping with distributed parity among the drives
- RAID-6 – extends RAID 5 by adding another parity block and support up to 2 drives failures
- RAID hybrid – combination of more than one RAID level, e. g. RAID 10



Calculator of usable free space available on RAID volumes - <http://www.raid-calculator.com/default.aspx>

Characteristics:

- requires specific controller
- guarantees best performance
- but is more expensive
- may require installation of drivers

Data acquisition:

- it's easier to acquire at logical level, as if it was a single disc
- acquisition OS must support RAID controllers
- individual acquisition of RAID discs:
 - ✓ only when the OS doesn't support RAID controllers
 - ✓ analysis is more complex – the RAID volume must be rebuilt

Characteristics:

- implemented in the OS (supported by most modern OS)
- less efficient, depends on the CPU to calculate the parity bits data splits
- Windows – *Logical Disk Management* (LDM)
 - ✓ requires dynamic volumes
 - ✓ RAID volume configuration is stored on each drive
 - ✓ supports RAID 0, 1 and 5

Linux

- ✓ uses *Logical Volume Manager*
- ✓ saves meta data of the volume inside the drives
- ✓ uses volumes on DOS partitions
- ✓ supports RAID 0, 1, 5 and 6
- ✓ supports Windows LDM (may require kernel recompilation)
- ✓ allows the creation of snapshots – records only the changes and can be reverted to previous state

Acquisition:

- it's easier to acquire at logical level, as if it was a single disc
- individual acquisition of RAID discs:
 - ✓ it's easier than hardware RAID systems
 - ✓ there are some tools to automatically rebuild the RAID volume
- Windows – with this OS a write blocker must be used
- Linux – it's possible to do a read only mount and it also supports LDM from Windows OS

Volume and Partition Tools

disktype

- identifies many partitions types – Homepage: <http://disktype.sourceforge.net/>
- run this command to install it under Linux: apt-get install disktype
- download disk images for testing: <http://dftt.sourceforge.net/>

Example of a Joliet partition:

```
disktype iso9660-joliet.image
```

```
-- iso9660-joliet.image
```

```
Regular file, size 256 KiB (262144 bytes)
```

```
ISO9660 file system
```

```
Volume name "ISO9660 Joliet"
```

```
Application "MKISOFS ISO 9660/HFS FILESYSTEM BUILDER & CDRECORD CD-R/DVD CREATOR (C) 1993 E.YOUNGDALE
```

```
Data size 256 KiB (262144 bytes, 128 blocks of 2 KiB)
```

```
Joliet extension, volume name "ISO9660 Joliet"
```

Example of a DOS partition

```
disktype /dev/sda

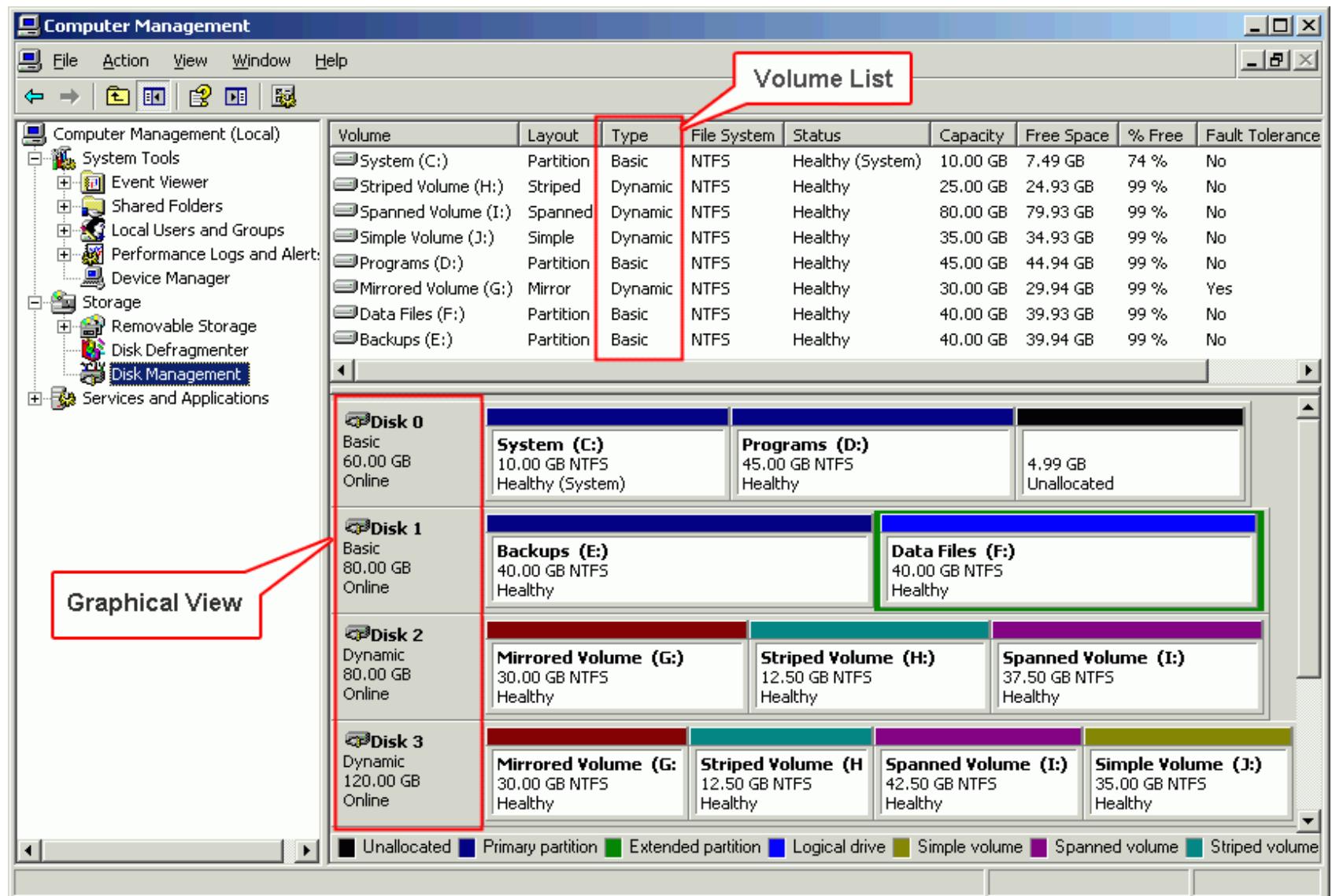
--- /dev/sda
Block device, size 20 GiB (21474836480 bytes)
DOS/MBR partition map
Partition 1: 19.14 GiB (20548943872 bytes, 40134656 sectors from 2048, bootable)
    Type 0x83 (Linux)
    Ext3 file system
        UUID FD935D1A-E410-4F97-BA7A-AE8A6B5C6E84 (DCE, v4)
        Last mounted at "/"
        Volume size 19.14 GiB (20548943872 bytes, 5016832 blocks of 4 KiB)
Partition 2: 880.0 MiB (922747904 bytes, 1802242 sectors from 40138750)
    Type 0x05 (Extended)
    Partition 5: 880 MiB (922746880 bytes, 1802240 sectors from 40138750+2)
        Type 0x82 (Linux swap / Solaris)
        Linux swap, version 2, subversion 1, 4 KiB pages, little-endian
        Swap size 880.0 MiB (922738688 bytes, 225278 pages of 4 KiB)
```

Example of a GTP partition

```
disktype gpt.image

--- gpt.image
Regular file, size 8 MiB (8388608 bytes)
DOS/MBR partition map
Partition 1: 8.000 MiB (8388096 bytes, 16383 sectors from 1)
    Type 0xEE (EFI GPT protective)
GPT partition map, 128 entries
Disk size 8 MiB (8388608 bytes, 16384 sectors)
Disk GUID 96117FCE-B25F-7D42-876F-8D5CB784DCF7
Partition 1: 7.967 MiB (8354304 bytes, 16317 sectors from 34)
    Type Basic Data (GUID A2A0D0EB-E5B9-3344-87C0-68B6B72699C7)
    Partition Name "test-ext2"
    Partition GUID 500E199D-6002-1248-8A72-88FB112FA191
    Ext2 file system
        UUID 726EA38B-087D-4FD9-9DD8-E42DD8D2E930 (DCE, v4)
        Volume size 7.967 MiB (8353792 bytes, 8158 blocks of 1 KiB)
Partition 2: unused
```

Windows Disc Management



Command line tool - See tutorial: <https://www.thegeekstuff.com/2010/08/how-to-create-lvm/>

- **install:** apt-get intall lvm2

- **phase 1**

`pvcreate` – initialize drive or partition

`pvscan` – search physical volumes

`pvdisplay` – shows physical volumes attributes

- **phase 2**

`vgcreate` – creates an aggregated volume – *volume group*

`vgdisplay` – shows the attributes of the aggregated volume

- **phase 3**

`lvcreate` – creates a logical volume

`lvdisplay` – shows the attributes of the logic volume

- **optional**

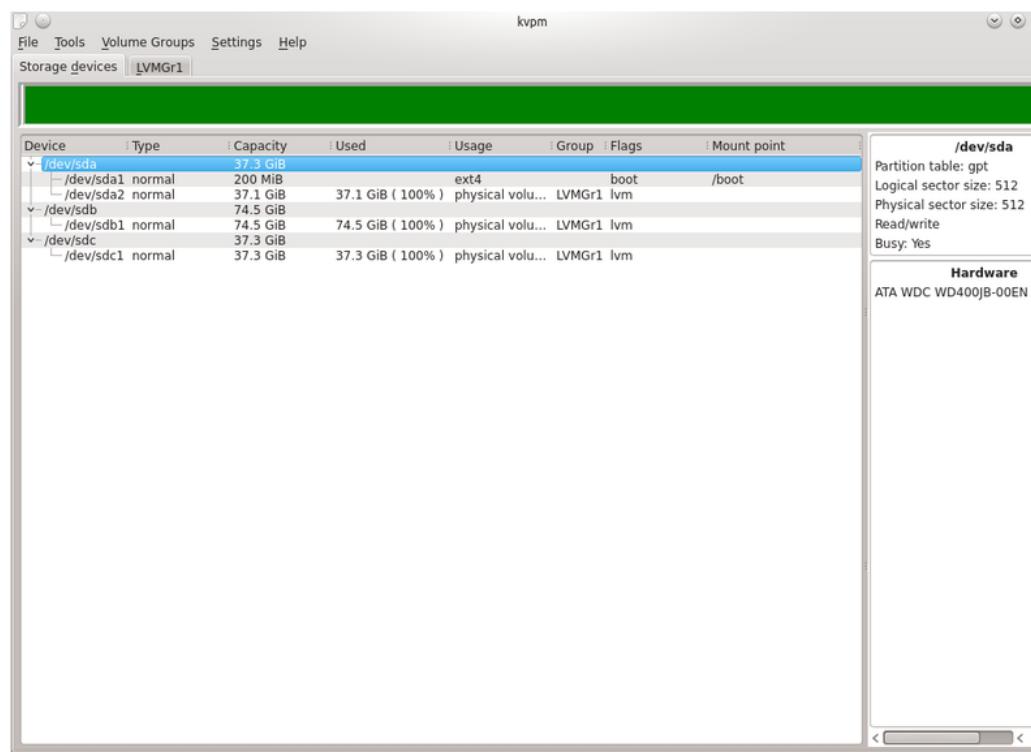
`lvextend` – change the size of a logic volume

GUI tools

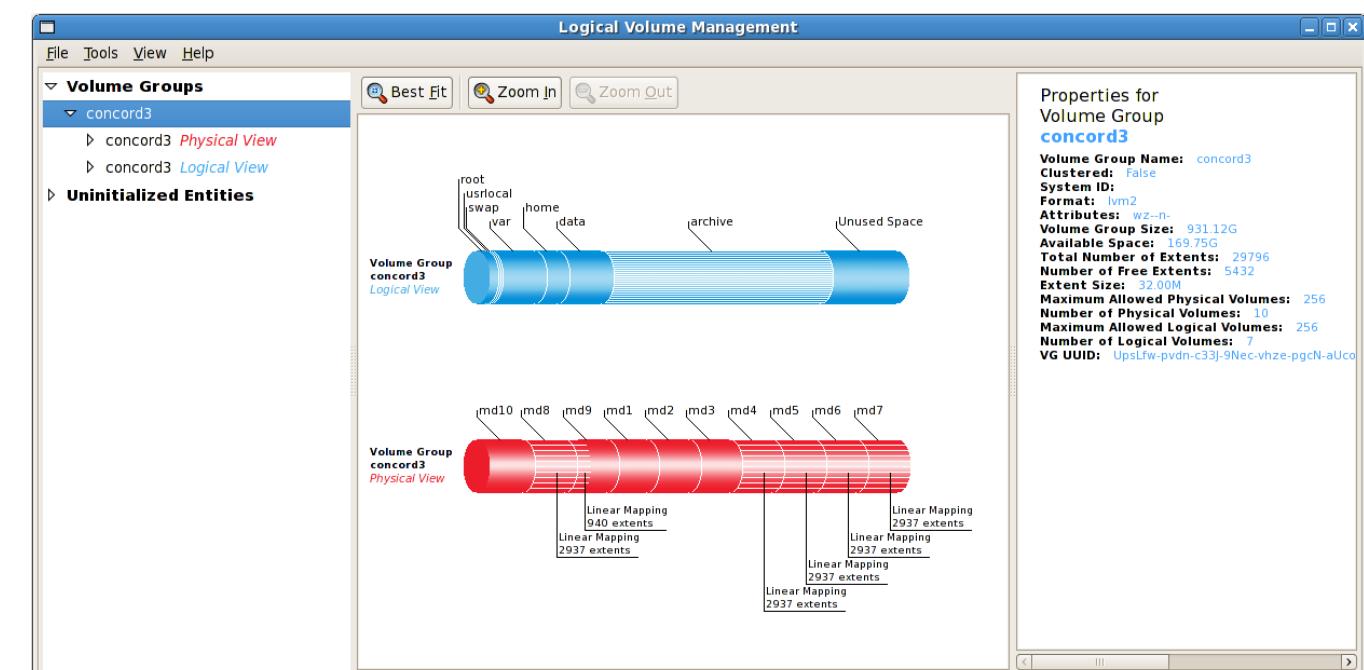
KVPM – KDE interface `apt-get install kvpm`

system-config-lvm – Gnome interface `apt-get install system-config-lvm`

KVPM



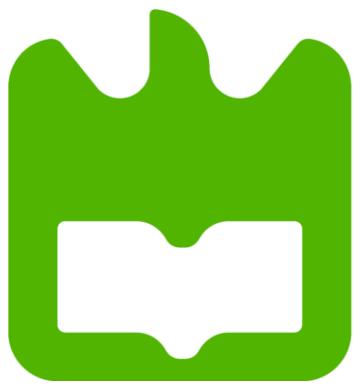
system-config-lvm



Exercises

07-Lab 1 – Identify partitions types with different tools





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

RAM Analysis

Artur Varanda

School Year 2021-2022

If evidence of compromise is never written to a hard drive, we cannot rely on disk forensics!

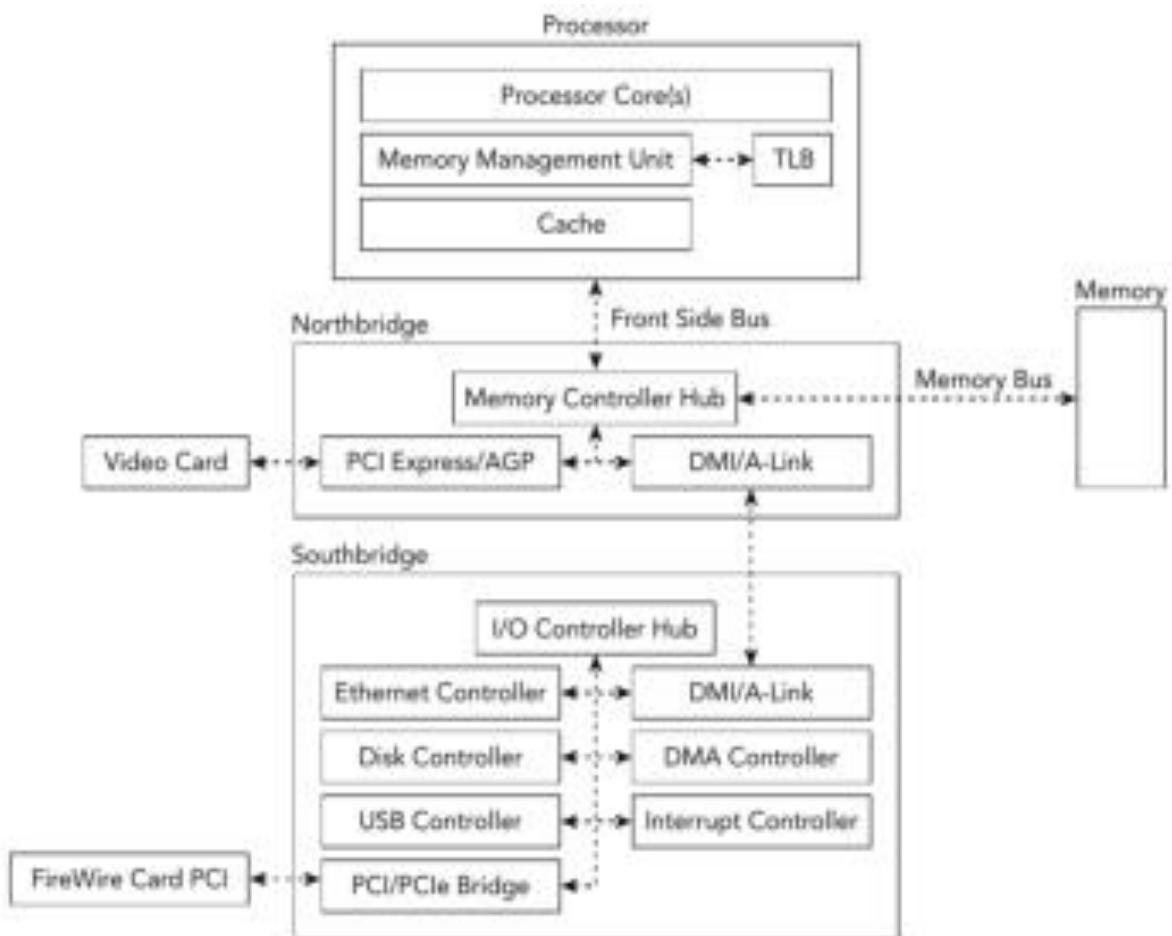
Volatile memory has a high potential to contain:

- malicious code from an infection, in whole or in part, because it must be loaded in memory to execute
- evidence that system resources were allocated by the malicious code
- encryption keys and passwords, or the plain-text contents of files before they were encrypted

Architecture

- CPU – accesses main memory to obtain its instructions and then executes them
- RAM – volatile memory that is much slower than CPU
- Cache – faster than RAM, but still slower than CPU
- MMU – Memory Management Unit to help find where the data is stored (RAM or Cache)
- TLB – Translation Lookaside Buffer is a special cache for MMU to translate memory addresses

Figure: Physical organization of a modern system

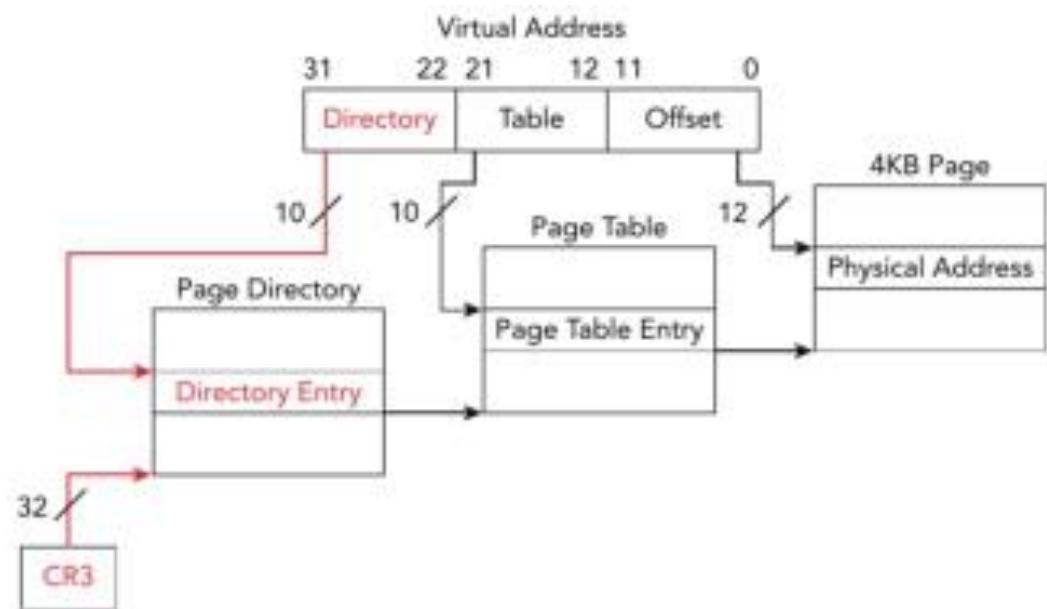


Direct Memory Access (DMA)

- provide I/O devices the capability to directly transfer data stored in system memory without processor intervention
- improves performance
- CPU initiates a data transfer and DMA controller manages the data transfer

- provides the ability to virtualize a linear address space
- creates an execution environment in which a large linear address space is simulated with a modest amount of physical memory and disk storage
- typical page size is 4 kB
- different paging structures are used for different processes
 - ✓ the OS provide each process the appearance of a single-programmed environment through a virtualized linear address space

Figure: Address translation to a 4 kB page using 32-bit paging



Impact on memory forensics

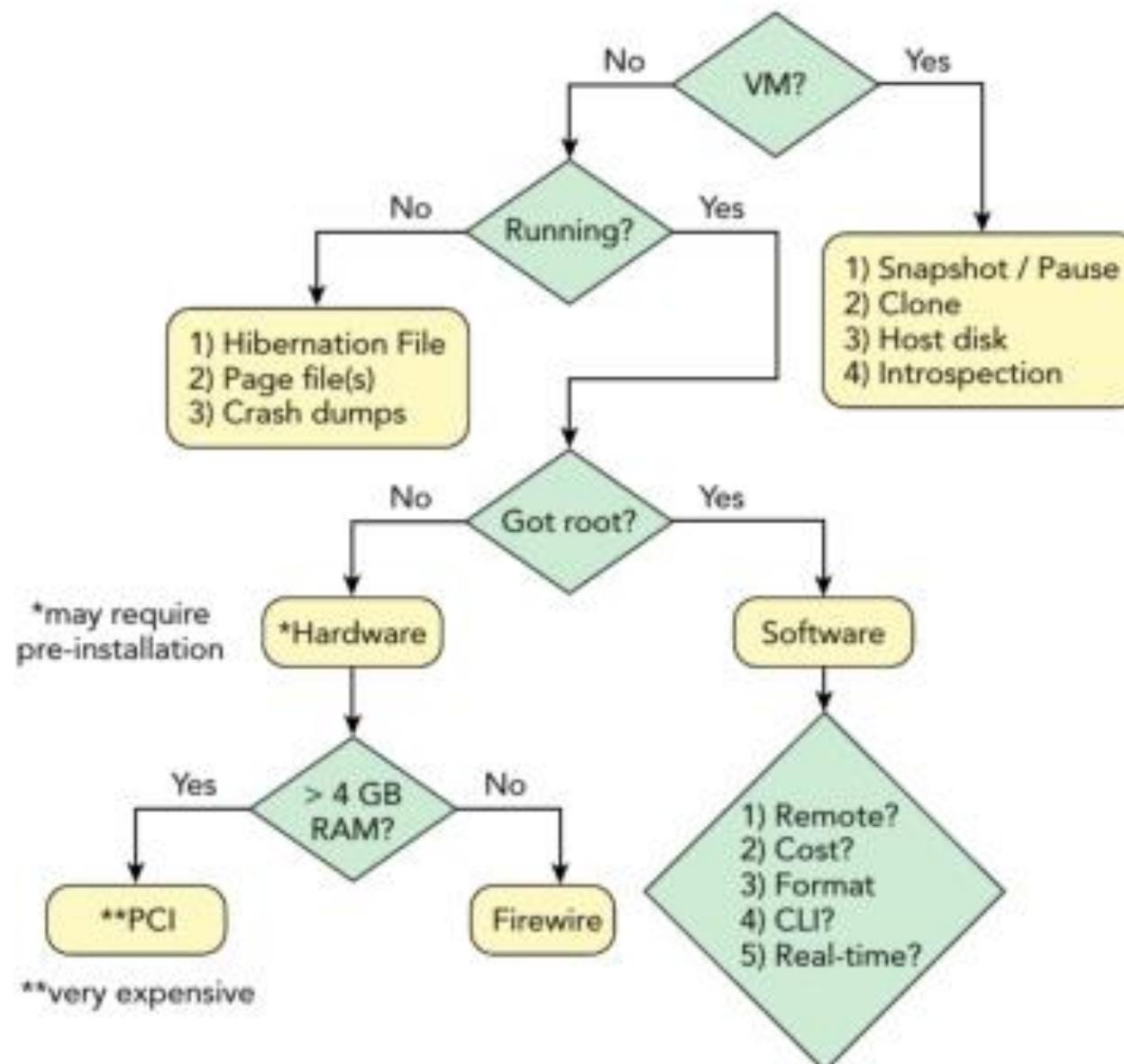
Forensics software must emulate the virtual address space and transparently handle virtual-to-physical-address translation

Memory Acquisition

Memory acquisition (also know as dumping, capturing, sampling)

- copy the contents of the volatile memory to a non-volatile storage
- an important source to get a better understanding of what happened
- decision must be made about which data to collect and the best method for doing so
 - ✓ methods and tools depend on the goals of the investigation and the characteristics of the system
 - ✓ choosing a proper tool is important to avoid corrupt memory images, destroyed evidence, and limited, if any, analysis capabilities

Memory acquisition decision tree



Decisions to make:

- remote or local – do you have physical access to the target system? Is it a server with no keyboard or monitor attached?
- cost – do you have budget restrictions on the acquisition software you can buy?
- file format – does your analysis tool support the file format of the acquisition tool?
- CLI or GUI – do you prefer command-line or graphical user interface tools? A CLI tool might have a smaller footprint, besides you might not have graphical engine running acquisition or runtime
- interrogation – Do you need a full physical memory dump or just the ability to determine the running processes, network connections?

Before you acquire physical memory, you should always consider the risks

- most OSs do not provide a supported native mechanism for acquiring physical memory
- memory acquisition tools might leave the system unstable
- poorly written malware can be unstable and may behave in an unpredictable way
- is the target a mission-critical system that can be shut down or rebooted only in extreme circumstances?

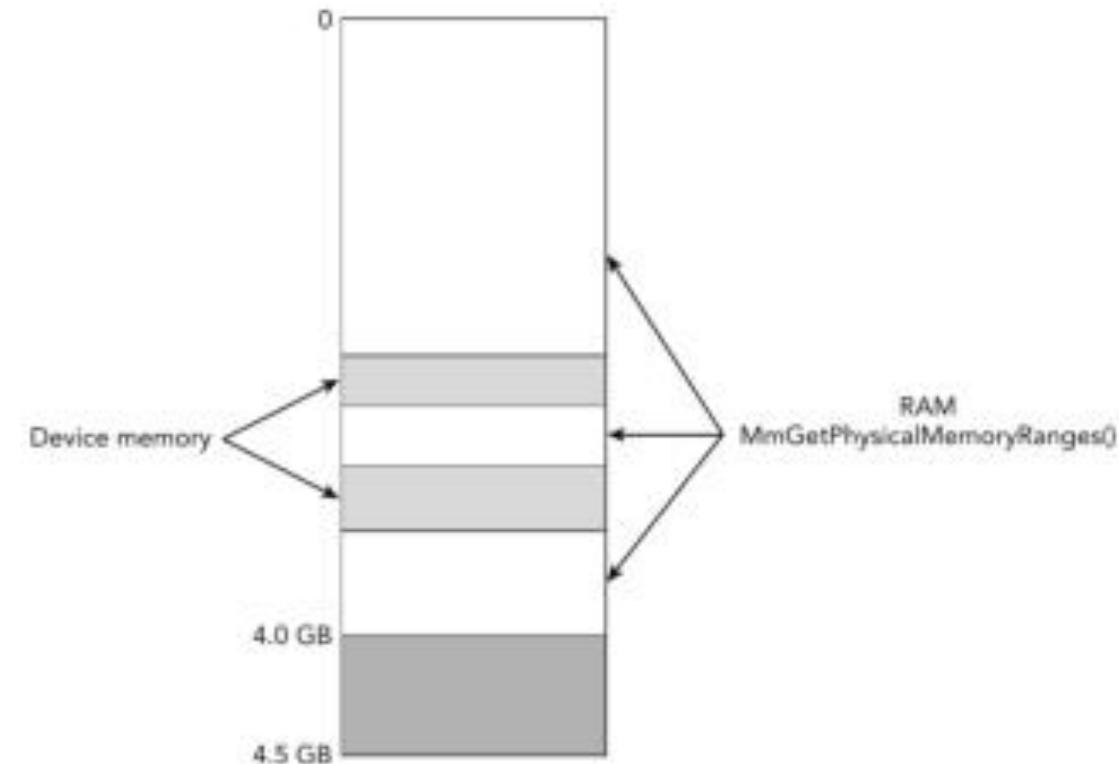
There might be circumstances in which the consequences (i.e., death, environmental damage) of destabilizing a system are never worth the risk.

Why memory acquisition can lead to system instabilities and evidence corruption?

- atomicity – memory acquisition is not an atomic operation and the contents of RAM are constantly changing. During acquisition, other processes are writing memory, the kernel is adding/removing linked-list elements, network connections are being initiated or torn down, and so on
- cache coherency – processors were not designed to accommodate the simultaneous mapping of the same physical address with multiple cache attributes (non-cached, cached, write-combined). A poorly written acquisition tool can easily invalidate the very memory being acquired.

- device memory – there are physical memory regions reserved for use by the firmware, by the ISA or PCI busses, or by various motherboard devices. Reading from one of these regions may alter the state of the device you are accessing.
- ✓ few tools are able to acquire these regions with reliability and accuracy

Figure: Physical memory layout



Choosing the proper time to depends on a number of factors.

List of suggestions:

- plan the acquisition when the suspect is online (or at least logged in), which can give you access to:
 - ✓ the suspect's logon session, information about cloud services or remote storage
 - ✓ and any encrypted documents that the suspect might have been viewing
- avoid the most active periods:
 - ✓ so that the suspect doesn't detect your activity
 - ✓ to minimize the number of anomalies you encounter when you analyze the evidence

Local acquisition to removable media:

- never dump memory to the target system's local drives, such as the C: partition
- dumping memory to an external USB, ESATA, or Firewire drive
- the file system of the external drive must support file sizes equal to the amount of RAM (FAT32 has a limit of 4 GB)
- advices:
 - ✓ removable media should be used only on one computer to avoid spreading malware
 - ✓ wipe removable media before using (or re-using) it to acquire evidence
 - ✓ do not plug possibly infected removable media directly into your forensic workstation, inspect it on another computer, then copy the evidence over an isolated network

Runtime interrogation:

- use automated tools that log all the preformed steps

Remote acquisition

- typically, the acquisition tool is pushed over the network to the target system
- the tool can run through a scheduled task or service
- the dump can be stored on a network share (last resort) or via a stream with netcat, but this method has some risks
 - ✓ administrator credentials and the contents of the target system's RAM may be exposed
 - ✓ create a temporary admin account and use an encrypted channel with a tool that supports TLS
 - ✓ configure the firewall to limit the traffic between the target and the remote acquisition system
- the use of compression is recommended

Acquisition tools

All software-based acquisition tools follow a similar method:

- load a kernel module that maps the desired physical addresses into the virtual address space of a process
- access the data from the virtual address space
- write the contents to the requested non-volatile storage
- most tools avoid acquiring device memory regions
 - ✓ the acquisition process is more stable
 - ✓ but might miss important evidence, such as sophisticated rootkits

There are many tools, just to name a few:

- Windows
 - ✓ AccessData FTK Imager (free)
 - ✓ EnCase (comercial)
 - ✓ Winpmem (open source)
- Mac
 - ✓ OSXPmem (open source)
- Linux
 - ✓ LiME - Linux Memory Extractor (open source)
 - ✓ AVML - Acquire Volatile Memory for Linux (free)

There are a few tools, but only some are open source:

- volatility – available for Windows, Mac and Linux
- rekall – fork from volatility

```
volatility -f memdump.raw imageinfo  
  
volatility -f memdump.raw --profile=WinXPSP3x86 pslist  
  
volatility -f memdump.raw --profile=WinXPSP3x86 consoles  
  
volatility -f memdump.raw --profile=WinXPSP3x86 -p 1564 memdump -D .  
  
volatility -f memdump.raw --profile=WinXPSP3x86 sockets  
  
volatility -f memdump.raw --profile=WinXPSP3x86 connections  
  
# hiberfil.sys is a dump made by the OS in compressed format  
  
# it is necessary to decompress to improve the analysis:  
  
volatility -f hiberfil.sys --profile=WinXPSP3x86 imagecopy -O hiberfil.raw
```

Windows

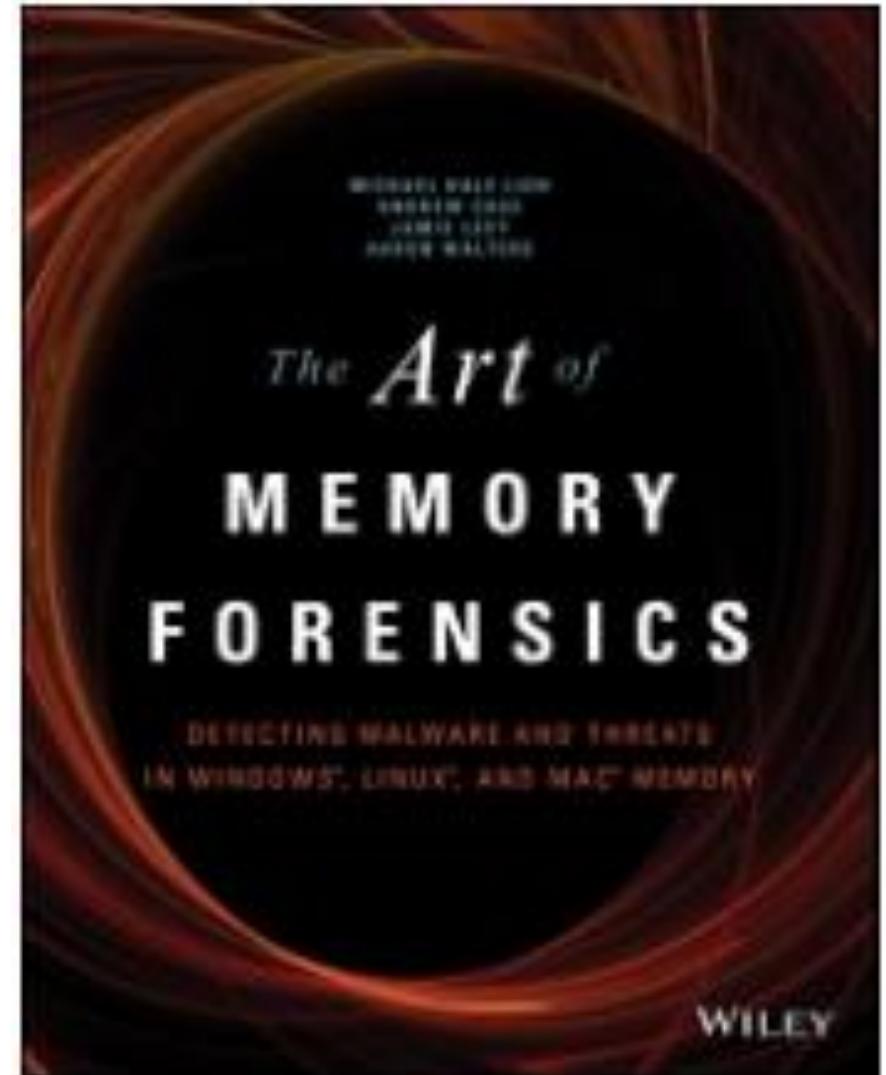
<https://github.com/volatilityfoundation/volatility/wiki/Command-Reference>

Linux:

<https://github.com/volatilityfoundation/volatility/wiki/Linux-Command-Reference>

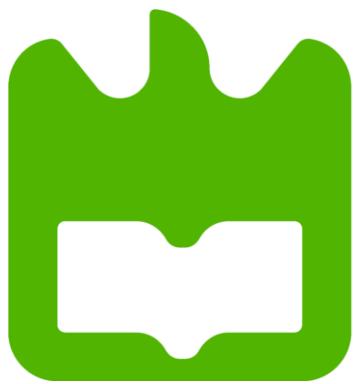
Book

- Title: The Art of Memory Forensics: detecting malware and threats in windows, linux, and Mac memory
- Authors: Ligh, M. H., Case, A., Levy, J., & Walters, A.
- Publisher: John Wiley & Sons
- Date: July 28, 2014
- ISBN: 978-1118825099
- <https://www.amazon.com/exec/obidos/ASIN/0321268172/>



08-Lab01 – Analysis of a memory dump with Volatility v2.6





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Mobile Forensics

Artur Varanda

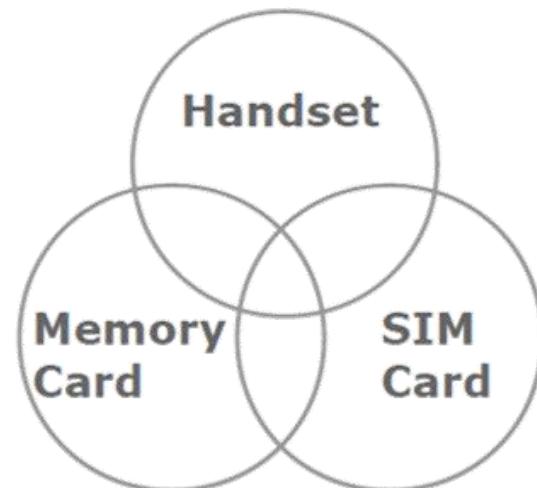
School Year 2021-2022

Phones, especially smartphones, have a huge potential of providing evidences

- are part of our everyday life
 - ✓ screen checks/day and h/day usage of smartphones
 - ✓ they store a huge amount of diverse information:
 - ✓ logs of calls, messages, GPS, network connections contents of messages, emails, multimedia (photos and video), social networks, etc...
- sales of smartphones surpassed PCs by the end of 2011

Where is data located in phones?

- data can be physically stored in 3 different locations:
 - ✓ handset, memory card and SIM card
- some types of data may be found in more than one location:
 - ✓ contacts on SIM and handset
 - ✓ pictures on handset and memory card

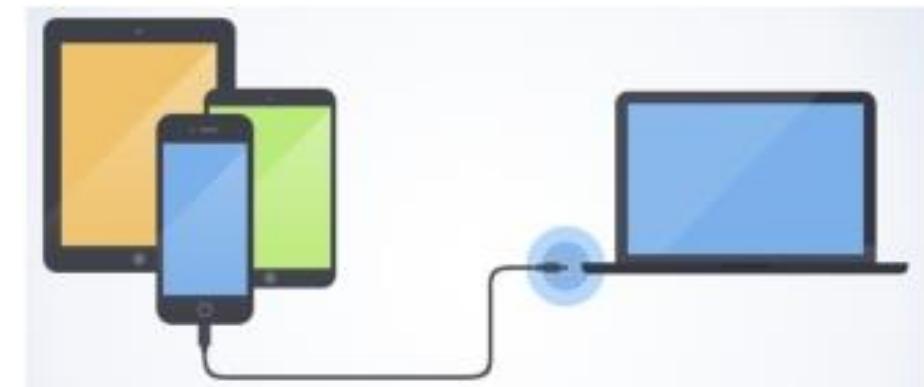
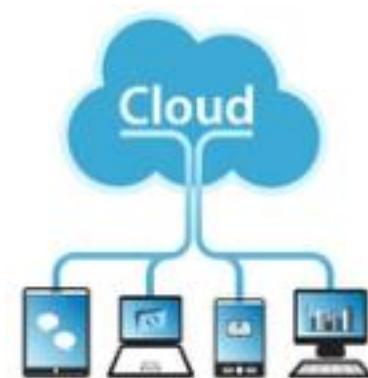


Retrieval approach:

- Examine every area (handset, memory card and SIM) independently
 - ✓ to be sure of capturing all the information you can

Can data be stored anywhere else?

- Service providers → requires additional legal procedures
- Cloud services → might require additional legal procedures
- Handset backups → more common in iOS devices



Disambiguation

- UICC (Universal Integrated Circuit Card) – is the technical name of the physical part of the smart card
- SIM (Subscriber Identity Module) – is a logical module stored inside the smart card
 - ✓ in the early stages a SIM consisted of the hardware and the software

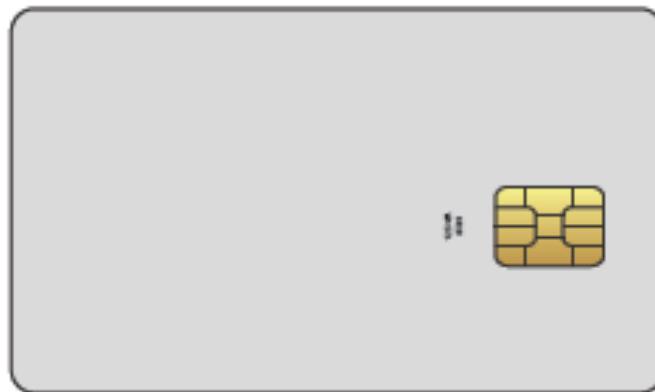
A given card can contain multiple SIMs

This would allow multiple phone numbers or accounts to be accessed by a single UICC.



12-in-1 UICC : https://multi-com.eu/details.id_pr,2769,key,sim-max-12-in-1-card.html

How many sizes/formats exist?

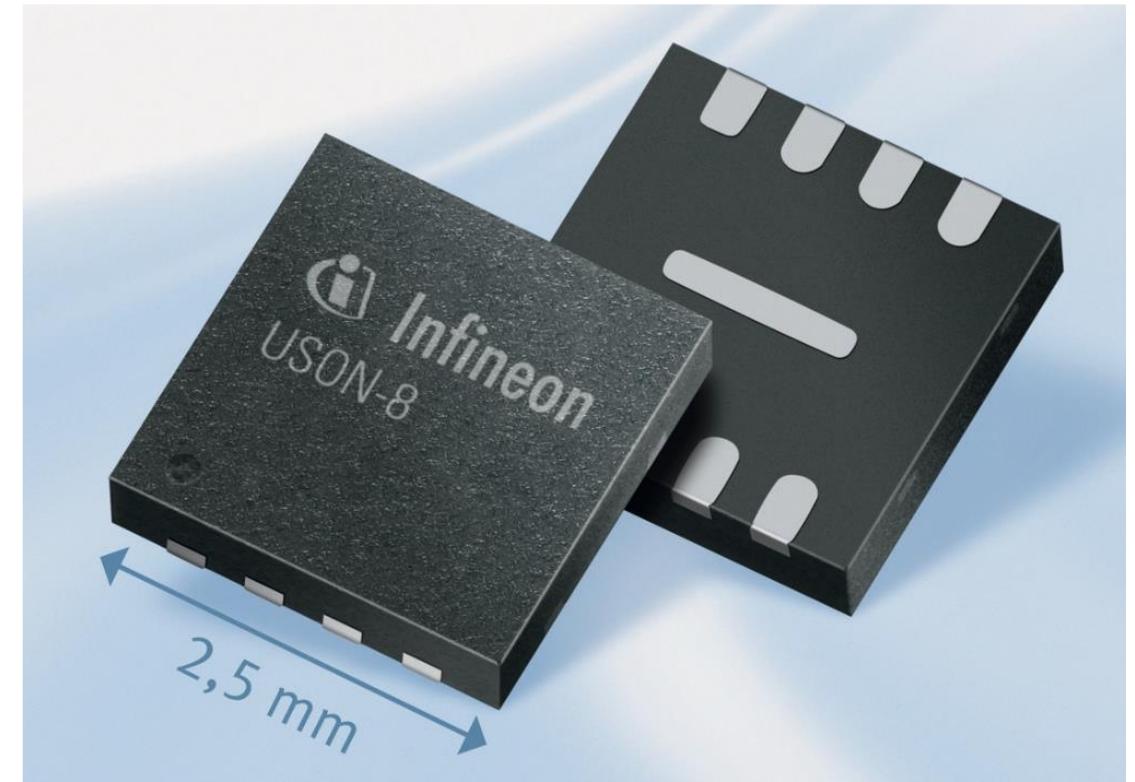


Variant	1FF	2FF ("Mini SIM")	3FF ("Micro SIM")	4FF ("Nano SIM")
Year of launch	1991	1996	2003	2012
Dimensions (mm)	85.6 x 53.98	25.0 x 15.0	15.0 x 12.0	12.3 x 8.8

These are user replaceable

Embedded UICC (also know as eSIM)

- permanently embedding into devices used in machine-to-machine (M2M) applications
- not replaceable by a regular user
- 2 formats MFF1 and MFF2 { both have the same size
 - ✓ MFF1 is socketable (replaceable with special tools)
 - ✓ MFF2 is soldered



These are non user replaceable

Main characteristics

- processor
- storage
 - ✓ memory to store text based user data e. g. SMS, contacts and calls
 - ✓ traditionally held just 16 to 64 KB, but there are some with 1 GB

UICC are also known as "SIM cards"

- mandatory in GSM networks
- standardized by 3GPP: <https://www.3gpp.org/>



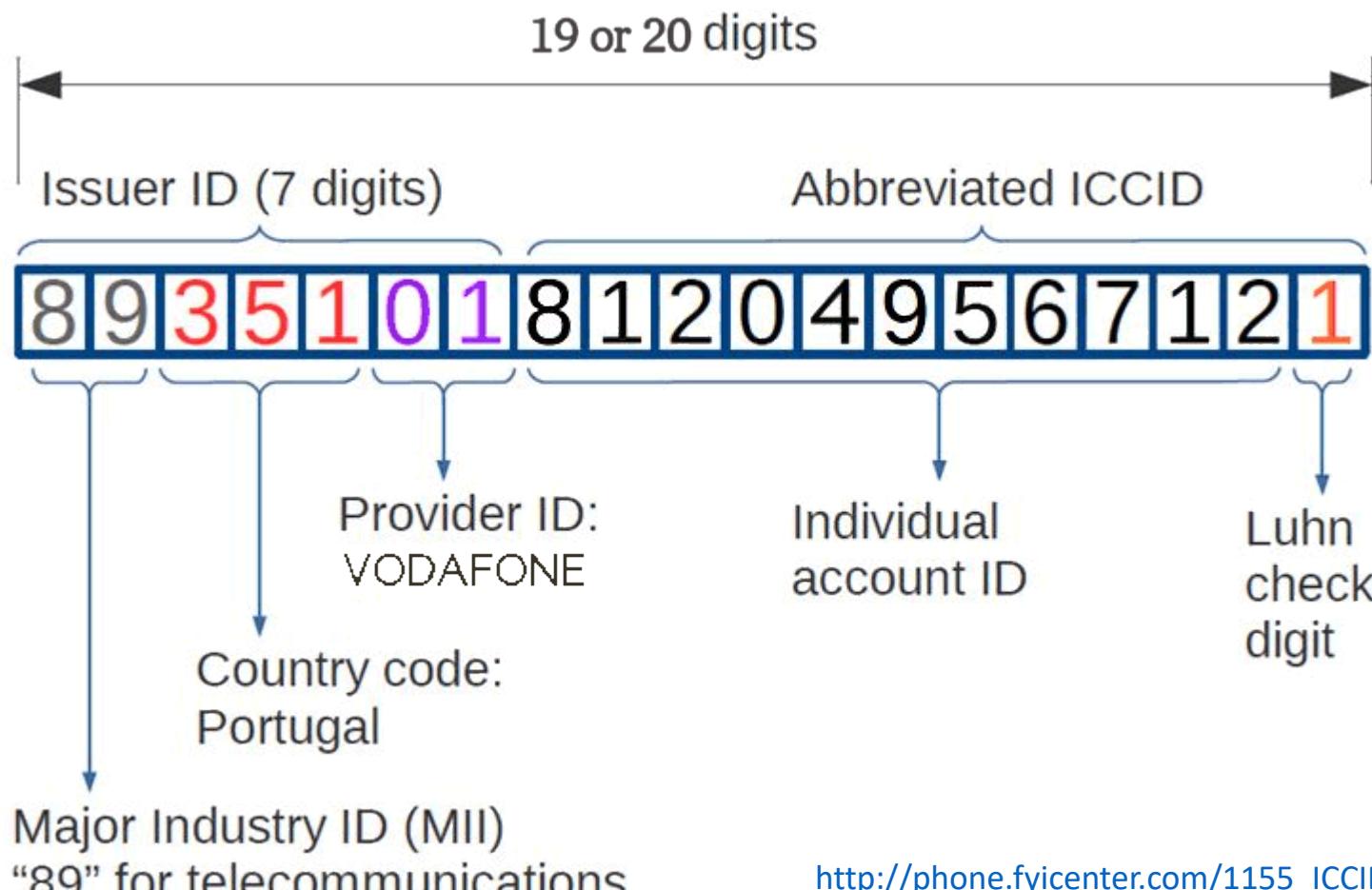
Integrated Circuit Card Identifier (ICCID)

- uniquely identifies the card
- 19 or 20 digits in length
- often printed on the outside (may be abbreviated)
- always stored digitally in the card



ICCID identifies issuing service provider and country

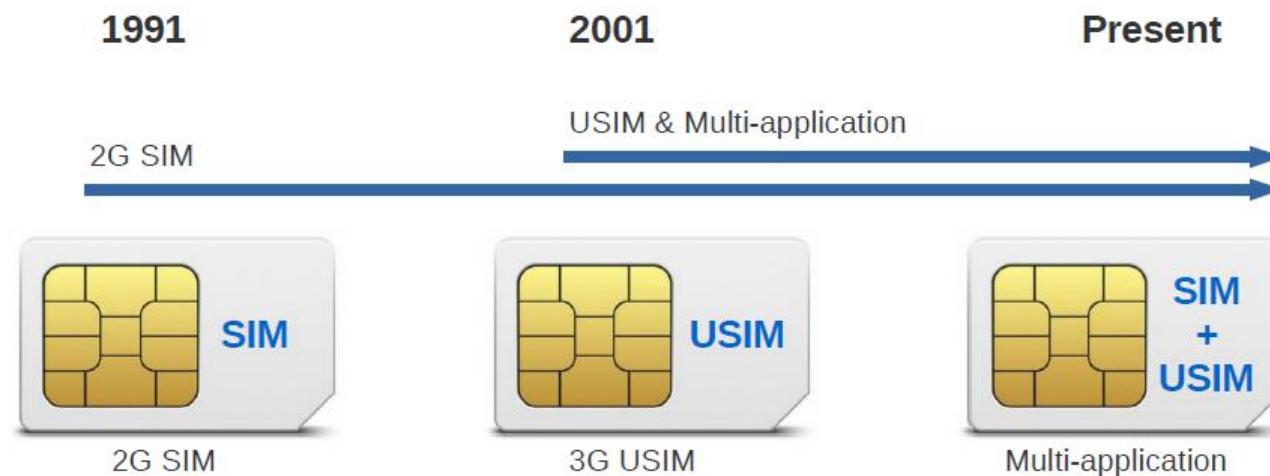
Integrated Circuit Card Identifier (ICCID)



Role of the SIM

- **Authentication** - the mobile network uses a challenge/response security mechanism to allow access to the network;
- **Accountability** - the SIM contains a unique reference number that identifies both the card and the subscriber to ensure that associated costs are allocated correctly;





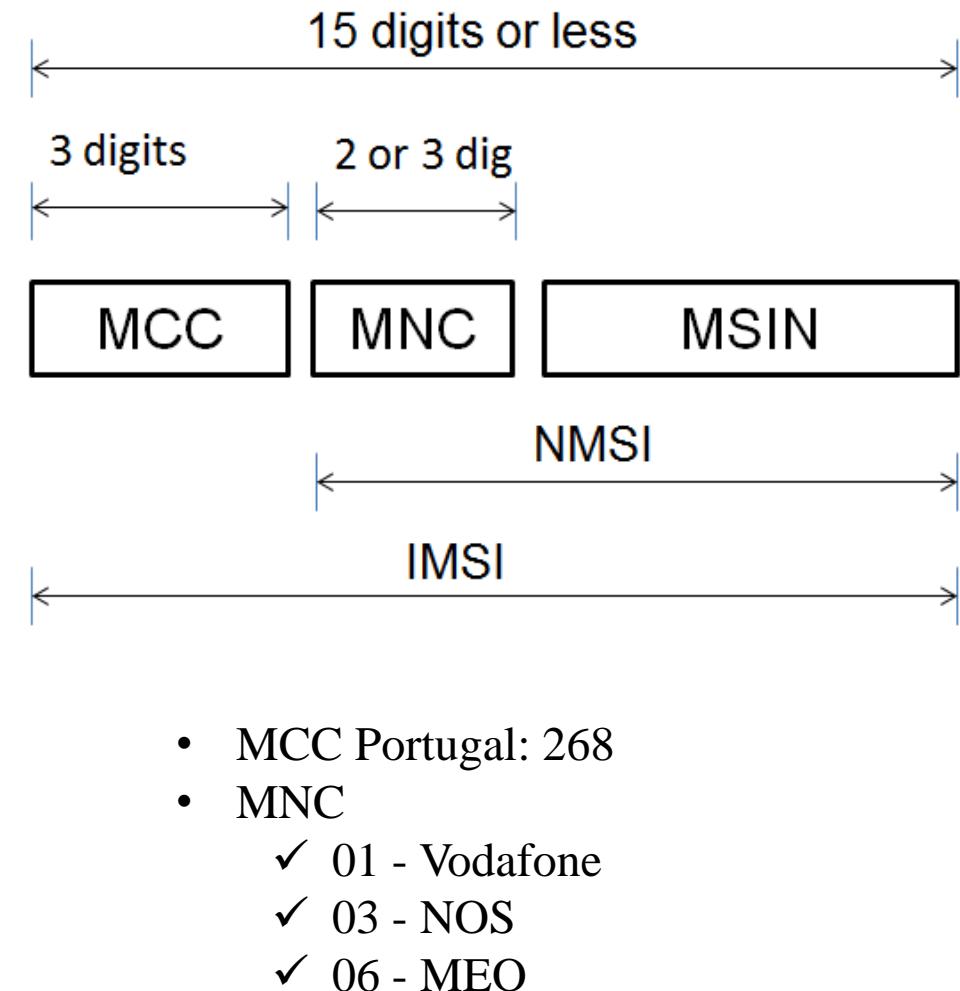
USIM - Universal Subscriber Identity Module

- for 3G and newer networks
- compared with SIM:
 - ✓ higher security, bigger and improved phonebook, can run small applications

Multi-application cards have 2 partitions: SIM + USIM

IMSI

- uniquely identifies the subscriber
- stored digitally in the card
- cannot be changed in a normal card
- can also identify issuing service provider and country
- usually not known by the owner
- composed by:
 - ✓ Mobile Country Code (MCC)
 - ✓ Mobile Network Code (MNC)
 - ✓ Mobile Subscription Identification Number (MSIN)

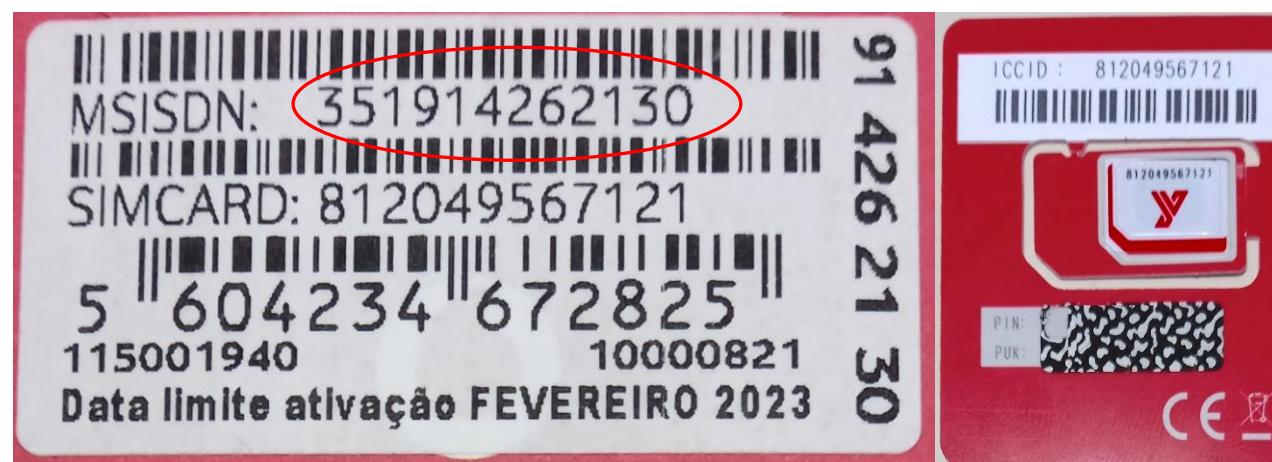


<https://www.msisdn.net/list-of-mcc-mnc/>

MSISDN

- just like the IMSI, the DSISDN is also an important number used for identifying a mobile subscriber
- used for routing calls to the subscriber
- it is the number normally dialed to connect a call to the mobile phone
- The ITU-T recommendation E.164 limits the maximum length of an MSISDN to 15 digits. 1-3 digits are reserved for country code.

MSISDN = Country Code + Subscriber Number



Phones vary enormously

Huge variation between different handsets

- shape, keyboard, connectivity, features, memory, etc

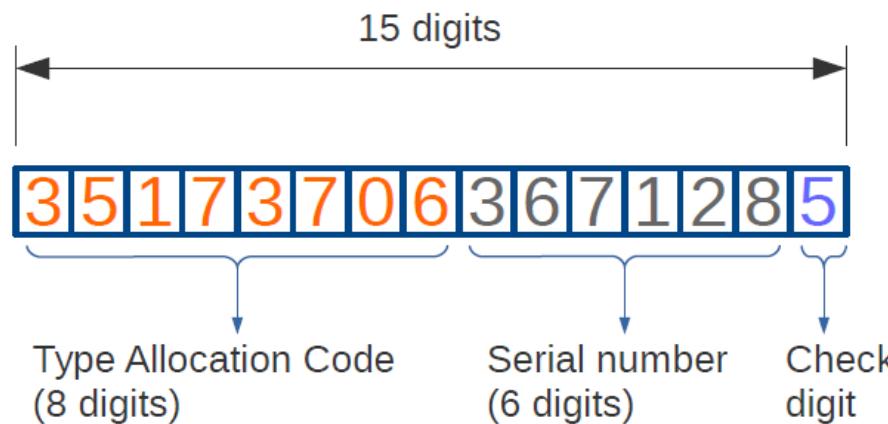


What they have in common?

International Mobile Equipment Identifier (IMEI)
All handsets must be uniquely identified by IMEI

IMEI

- printed under the battery, or the back of the device
- stored digitally in the handset
 - ✓ can be displayed on most phones by entering *#06#
- first 8 digits identify manufacturer and model



*#06#

IMEI

355995057333900 / 01

355996057333908 / 01

<https://www.imei.info/>

XRY Software

Supports device search by the first digits of TAC

https://en.wikipedia.org/wiki/International_Mobile_Equipment_Identity

[\(Manufacturer and model\)](https://en.wikipedia.org/wiki>Type_Allocation_Code)

[\(GSMA-approved group that allocated the TAC\)](https://en.wikipedia.org/wiki/Reporting_Body_Identifier)

[\(Check digit computation\)](https://en.wikipedia.org/wiki/Luhn_algorithm)

What data can potentially be an evidence on the (U)SIM?

- ICCID and IMSI
- phonebook / contacts
- SMS messages, including deleted (if not overwritten)
- call information

	SIM	USIM
Dialled	Yes (no time, date or duration)	Yes (optional: time, date and duration)
Received	No	Yes (optional: time, date and duration)
Missed	No	Yes (optional: time, date and duration)

What happens when a SMS on the (U)SIM is deleted?

- the message status is changed to indicate message no longer required
- however, the content of the message is typically left intact
- message is only overwritten when space is required for a new SMS
- So deleted SMS may be retrieved by accessing SIM via a card reader

Each SMS message has a maximum length

- 160 characters (for the GSM Latin alphabet)
- less characters per message for Arabic, Hebrew, etc
- long SMS are split into 2 or more separate messages



Network data on SIM card

- location area information can be retrieved from SIM card but its value is limited
 - ✓ likely to reflect location of seizure, but could be used to ensure the handset hasn't been switched on after seizure
- list of allowed or forbidden networks
- other network information can also be retrieved, e. g. TMSI, Kc, etc
 - ✓ may not be relevant to many investigations
 - ✓ more details: Forensics Wiki (https://forensicswiki.xyz/wiki/index.php?title=SIM_Cards) and (https://en.wikipedia.org/wiki/SIM_card)



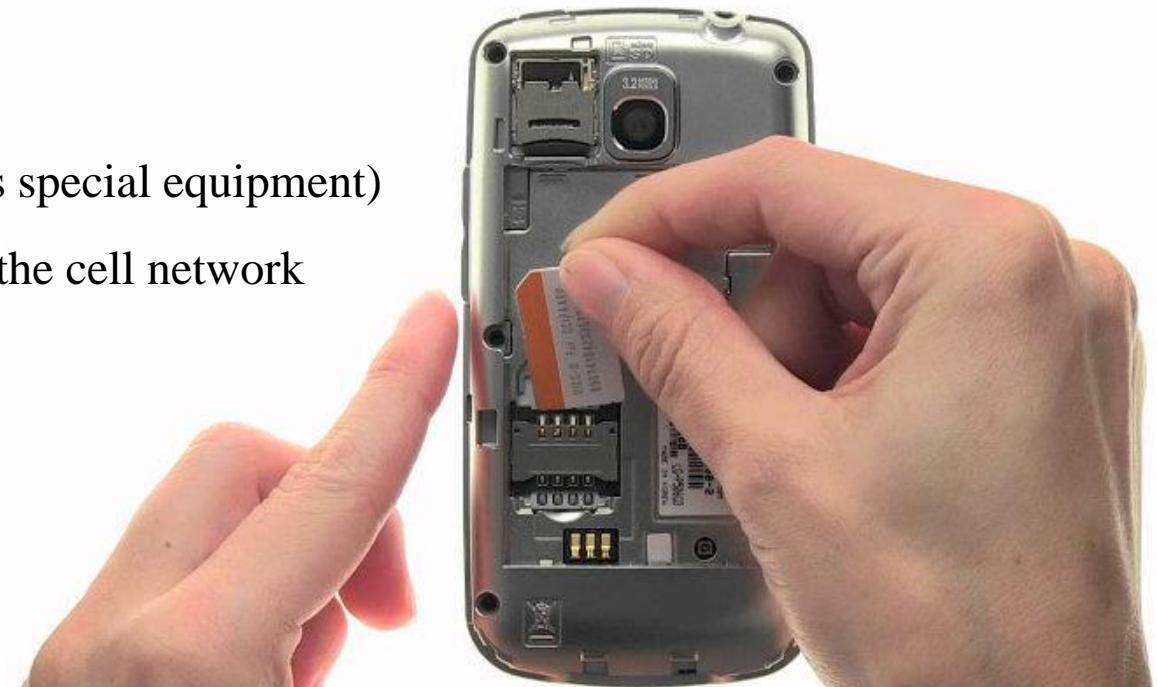
Warning

Some feature phones only work with a SIM card inserted.

However, if the handset detects that the SIM card has changed (based on IMSI, ICCID or both) it could potentially delete the call register entries

Solution:

- when available, clone SIM card ICCID and IMSI (requires special equipment)
 - ✓ SIM with cloned ICCID and IMSI cannot connect to the cell network

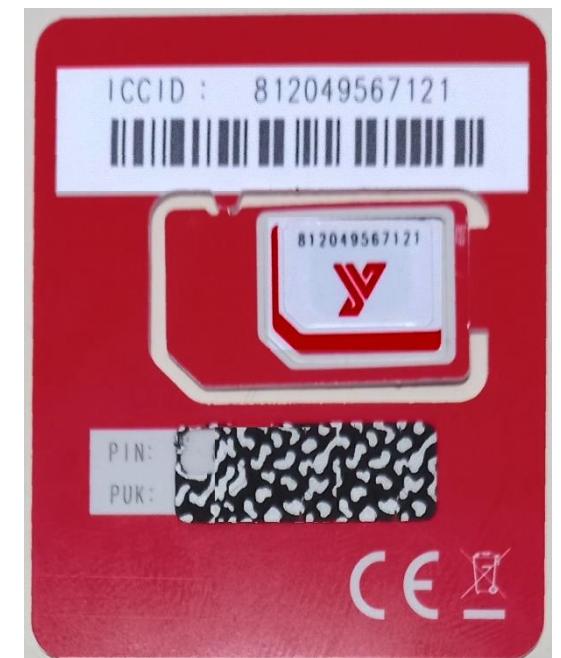


SIM card security

- PIN (personal identification number) - if the PIN is enabled and entered incorrectly 3 times in a row, the SIM card will be blocked
- PUK (PIN unlock key) - if PUK is entered 10 times incorrectly, the SIM card will become **permanently blocked and unrecoverable**
 - ✓ PUK is a SIM-specific code assigned by the service provider
 - ✓ cannot be changed by the user

Without the PIN

- only ICCID can be read from the SIM card
- with the ICCID ask the service provider for the PUK ⇔ requires legal procedures



Mobile Station International Subscriber Directory Number (MSISDN)

- SIM cards can store one or more phone numbers (MSISDNs)
- but entries are **unreliable**
 - ✓ the number may never have been used
 - ✓ the MSISDN may have been ported over from a previous SIM
 - ✓ in older handsets the number may be missing or **edited by owner**
 - ✓ MSISDN must be confirmed with the service provider

Data Acquisition

Interfaces to acquire data

- Cable – fast and secure



- Bluetooth – slower than cable, leaves footprints in PC and handset



XRY Device Manual details the preferred connection

- same handset with different connection interface may produce different results
- “Manually find devices and apps” and read info

Ensure network isolation before data extraction

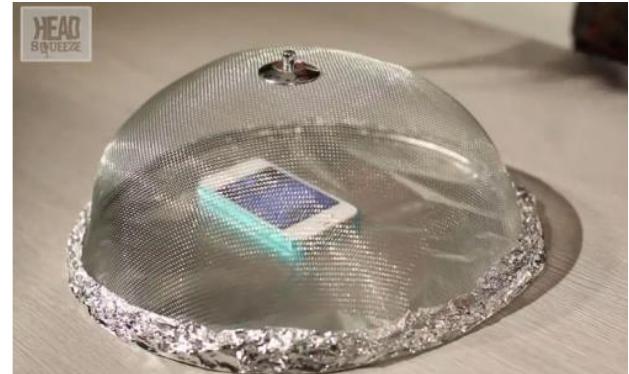
- airplane mode ON - many devices allow it in the block screen

or

- remove the SIM card and turn off Wi-Fi
- use a faraday cage
- signal jammers (might require special permission)

Importance of network isolation

- to avoid data changes (SMS, chat, calls, etc)
- to avoid remote locks or wipes

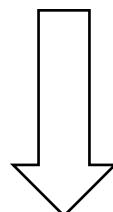


How it works?

- acquisition software asks handset what data is available
- handset may or may not provide data
 - ✓ usually, it's not possible to get deleted data
- different protocols are used for
 - ✓ different handsets and OS
 - ✓ different data types

Different types of logical acquisition

- full acquisition (through an agent app)
- ask the smartphone to do a backup



What data can be retrieved in a logical acquisition?



SIM Card



Feature phone



Smartphone

Live data
(undeleted)

Live SIM data can
be retrieved

Live handset data
can be retrieved

Live handset data
can be retrieved

Deleted data

only SMS with
card reader

deleted handset
data cannot be
retrieved

deleted handset
data may be
retrieved

Logical acquisition requires access to the OS.

How to beat the security code?

- ask the owner!
- XRY can get the security codes of some devices, check the “Device Manual”
- smudged swipe pattern
- manufacturers defaults (check user manual)
- on some devices XRY can do a physical acquisition without the security code



How it works?

- data is recovered in raw form
 - ✓ copy bit by bit
 - ✓ provides a lot of data, including deleted data (not overwritten)
 - ✓ requires decoding of the raw data
 - CPU intensive
 - software may not be able to decode everything
 - cannot be done if device is encrypted

What data can be retrieved in a physical acquisition?



SIM Card



Feature phone



Smartphone

All data is extracted, bit by bit

(there is no distinction between deleted and undeleted data)

N/A

Data can be decoded:
previous IMSI/ICCID details + bluetooth pairings + some security codes + apps data + etc

deleted data may be recoverable

encrypted devices cannot be decoded

Single hash for repeatability is not feasible

In mobile data acquisition hashing cannot be treated the same way as in HDD or SSD

- each time the handset is turned ON something changes
 - it is not feasible to have a single hash value for the acquired image
 - it is very important to hash ALL content individually (every file or digital object)
- ✓ this is done automatically by acquisition software like XRY

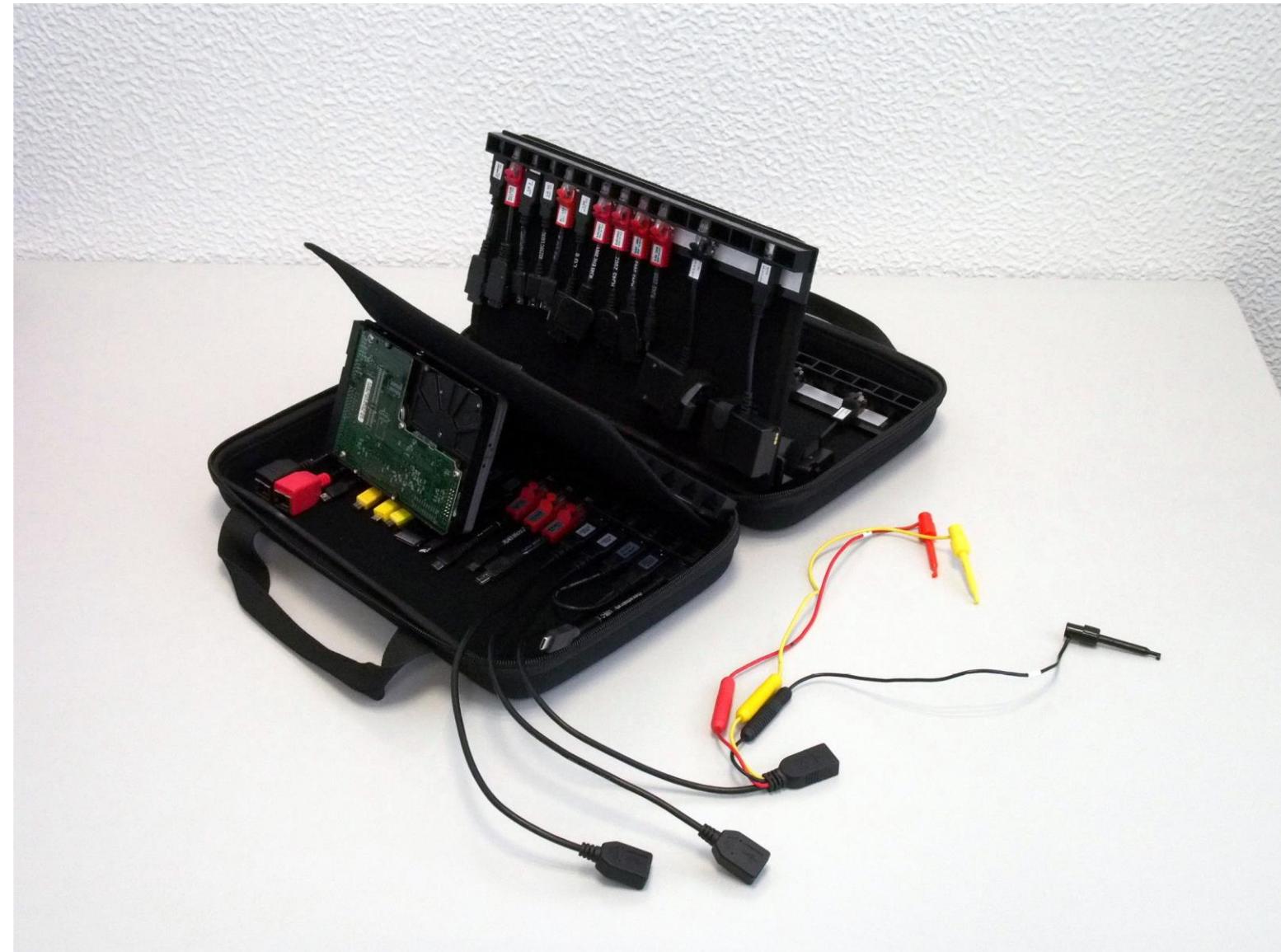


79054025
255fb1a2
6e4bc422
aef54eb4

Some software tools:

- **XRY** - to acquire data on mobile devices
- **FTK Imager** - to acquire data on memory cards
- **Autopsy** - to analyze acquisitions, some support for Android
- **SQLiteBrowser** - to see browsers caches and mobile apps databases
- **file** - to identify file types regardless of extension
- **strings** - to extract strings (ascii, utf-8 or utf-16)
- scripting to speed up repetitive tasks
- phones specifications: <https://phonescoop.com> and <https://gsmarena.com>

Hardware tools



Hardware tools



Service providers can give additional information:

- subscriber details (name, address, payment details, etc)
- calls made/received
- SMS and MMS logs
- voicemail
- location information
- ICCID / IMSI / IMEI / phone number (MSISDN)
- PUK code



Open the file MAVEN.xrycase with XAMN Launcher

1- open the SIM acquisition

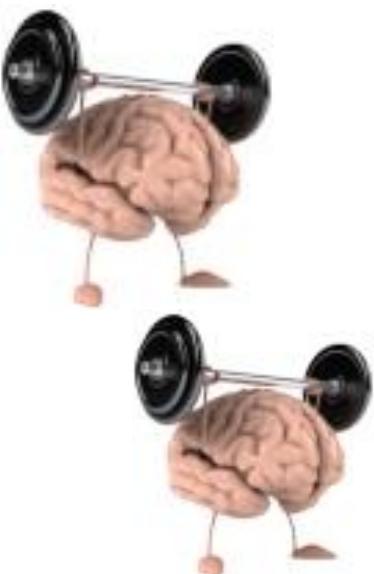
- what is the card ID (ICCID)? Who issued this SIM card? what is the network operator?
- what is the phone number?
- click on Device, a new tab will open
- what was the last network the device was connected?

2 - open the Agent acquisition

- what is the phone brand and model?
- what is the device timezone?
- click on ALL, a new tab will open
- what is the phone number of Kara Thrace?
- what number was dialed on 16-06-2016 21:21:11 UTC?

3 - Look for MMS, choose MMS ID 1

- “launch” the image file and look at its properties (exif info)
- find the GPS coordinates, what is the address where the photo was taken?
- the photo on the MMS ID 2 was in the same area?



Still in the file case MAVEN.xrycase, open the Backup acquisition

1 - add a word list filter with Kik (a messaging app)

- change to File tree view pane

2 - profile picture

- /data/data/kik.android/8c66b2ac[...]/cache/profPics/
- can you see the profile picture?
- can you tell the username?

3 - look for the app databases

- directory /data/data/kik.android/databases/
- save this file 8c66b2ac[...]37.**kikDatabase.db**

4 - with SQLiteBrowser

- open database kikDatabase.db
- see contents of table KIKcontactsTable
- what is the photo time stamp of user funnyordie?

5 - with MFT Stampede

- convert the photo-timestamp to readable format





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Open Source INTelligence

Artur Varanda

School Year 2021-2022

History

HUMINT Human Intelligence – intelligence gathered by means of interpersonal contact, typical activities consist of interrogations and conversations with persons having access to information

SIGINT Signals Intelligence – intelligence-gathering by interception of signals, whether communications between people or from electronic signals not directly used in communication

SIGINT (Signals Intelligence) can be subdivided into:

COMINT Communications Intelligence – deals with messages or voice information derived from the interception of communications between people

ELINT Electronic signals intelligence – intelligence-gathering by use of electronic sensors

FISINT Foreign instrumentation signals intelligence – telemetry, tracking systems, video data links, and arms control

IMINT Imagery intelligence – collects information via satellite and aerial photography

MASINT Measurement and Signature Intelligence – analysis of data obtained from sensing instruments for the purpose of identifying any distinctive features associated with the source, emitter or sender, to facilitate the latter's measurement and identification. This often includes radar intelligence, acoustic intelligence, nuclear intelligence, and chemical and biological intelligence

- Since the 1930s, at the University of Princeton, the Foreign Broadcast Information Service (FBIS) collected information serving as an intelligence source in World War II and the Cold War;
- OSINT terminology was introduced by Americans in the mid-1970s and in Europe in the 1980s given the importance of criminal analysts in combating crime, including organized crime;
- In 2005, after the failure of September 11, CIA opened an Open-Source Center (OSC). However, it was discredited, because it was believed that non confidential information was not valuable.

Information Sources

Origin

- Primary information source is a place where original information resides (usually true)
- Secondary information source is a document or record that relates or discusses information originally presented in primary information sources (needs to be verified)
- Tertiary information sources are made up of lists or summaries of primary or secondary sources, such as bibliographies, lists of readings, or articles on research

Authority

- Closed Source Information
- Open Source Information

- A closed source involves obtaining a judicial authorization or formal explicit or implicit authorization through the delegation of powers
- Is outside the legal limits of collecting information on its own initiative
 - to maintain information privacy
 - to avoid the inadmissibility of its use
 - to avoid civil, disciplinary and criminal liability over the expert or his organization

Institute of Registries and Notaries (IRN)

- civil identification
- car registration
- collective entities
- criminal record
- land registry
- visa applications
- trademark registration
- registration of religious entities
- foundations and associations
- judicial power of attorney certificates
- marriages, divorces and sex changes
- nationality attributions and losses
- wills and public scriptures
- *etc*

Tax Authority (AT)

- tax registration
- taxable assets
- household
- taxed IRS
- billing
- taxed IUC, etc

Institute of Mobility and Transports (IMTT)

- driver identification
- automotive inspections
- automotive booklets
- etc

Open vs. Closed sources

	Closed	Open	
Primary	Radiography (X-Ray)	Fact	
Secondary	Radiological Report	News in a Newspaper	Legal Authorization Required

Open source information is said to be open if:

- it is fully accessible by third parties
- can be of individual or collective origin
- can be collected and processed automatically or manually

Open source categories

- **Traditional Media** newspapers, magazines, radio, and television
- **Internet** communities and user-generated content (social networks, video sharing networks, wikis, chats and blogs)
- **Public Data** official data from governmental and other organizations, conferences, speeches, from companies
- **Observation and Reporting** data collected by specialized citizens
- **Pictures, Videos and Sound** maps, satellite images
- **Professional and Academic** greyliterature, reports and articles

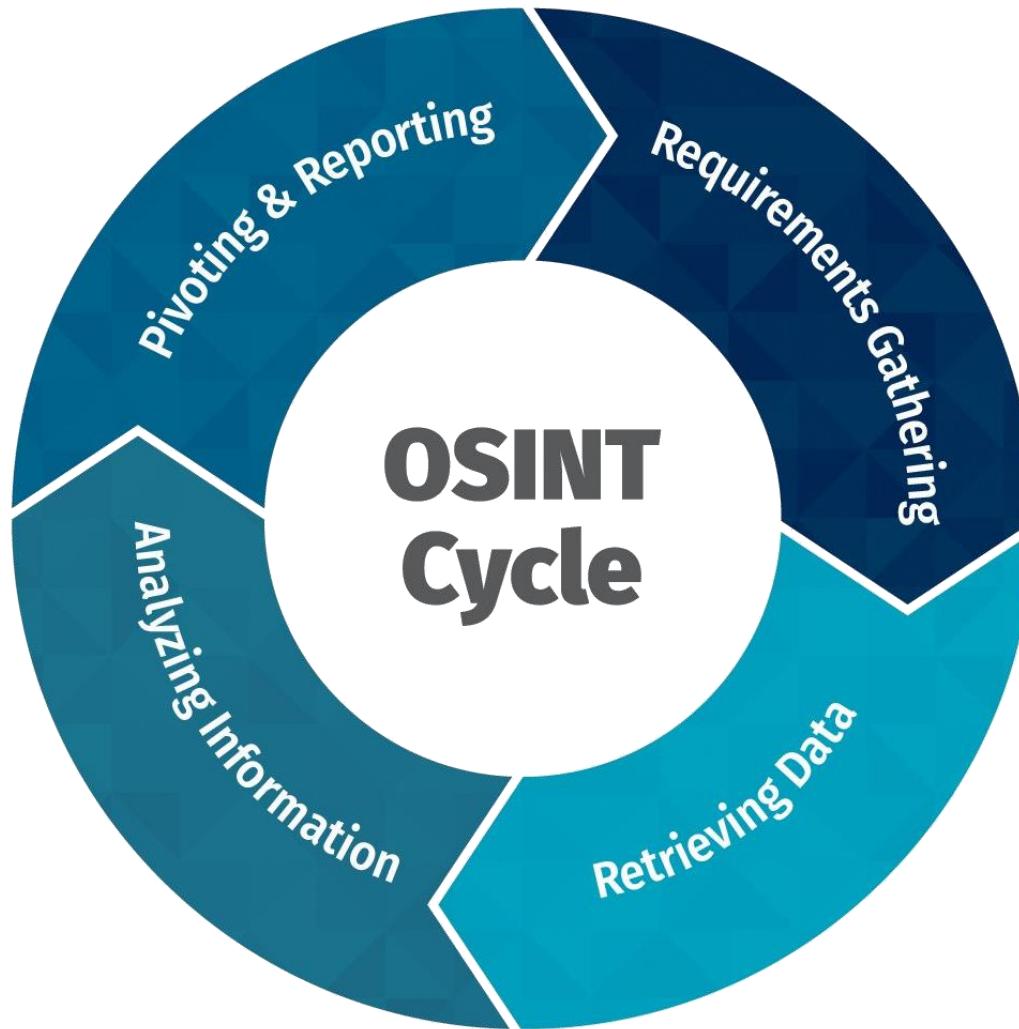
Advantages

- open sources are less costly and quick to access
- open sources provide information beyond closed sources and in greater quantities
- open sources can reduce the need for the production of classified intelligence and complement the latter
- does not compromises the purposes of the investigation
- open sources are a contextual element for classified intelligence operations

Disadvantages

- open source information does not replace classified information due to its intentionally secretive nature
- can be subject to misinformation or arbitrariness and therefore generally require a validation process
- need experts in various domains, this is more relevant if foreign languages are involved
- the amount of information available on the Internet is usually excessive, this implies the use of specific tools

Information to Intelligence Cycle



Advantages of Open Sources at the strategic level

- can help identify strengths, threats, risks and opportunities for sustainability and long-term e.g. growth in the number of refugees in Europe
- important in the cultural, demographic and geographical components in assessing opportunities and alliances
- essential to contextualize internal strategic information by comparison and benchmarking

Advantages of Open Sources at the operational level

- important to define and subsequently contextualize operations at the human and geographical level, such as distribution
 - e. g. Google Maps, Bing Maps
- very important in coordinating joint commercial or other international operations where classified information does not exist
 - e. g. hotel information, habits, etc

Advantages of Open Sources at the tactical level

- enabling current and imminent information is an important resource for addressing medium-term priorities and assessing the individual reliability of forces
 - e. g. new products, new EU incentives, etc
- can provide good information about the human context for medium term operations
 - e. g. current demographic, cultural, etc

Advantages of Open Sources at the technical level

- it's an important resource for assessing the effectiveness of forces at the technical level
e.g. information about computer resources, level of automation, etc
- it allows to collect information about the systems, transport, communications and even financial context

Skills of the Analyst

Discipline in collecting due to excessive information

Time	Task	Description
15 minutes	Requirements definition	Ensure an understanding of commander's intent
30 minutes	Internet Collection	Use search tools, rapidly identify top ten sites and review
15 minutes	Resources' Table	Create Resources' Table for future use and for customer's reference
60 minutes	Commercial Collection	Use fee sources, identify top 20 items for exploitation
60 minutes	Analysis	Read, understand, evaluate, and structure collected information
60 minutes	Production	Carefully create an analytical summary, table of contents, and slides

4 hours – Total time to produce an open source analysis report using only internal sources

Source: NATO OSINT manual

OSINT process

- 1. know who knows** have in-depth knowledge of the available sources' characteristics
- 2. know what's what** ability to evaluate and assess the validity, scope, degree of accuracy and timeliness of the requirements
- 3. know what's hot** ability to distinguish what is important and relevant
- 4. know who's who** ability to distinguish between facts and speculation and avoid cognitive bias from the sources

Source: NATO OSINT manual

Cognitive deviations (bias)

- systematic error in thinking that affects the decisions and judgments that people make
- individuals create their own *subjective social reality* from their perception of the input
- may lead to perceptual distortion, inaccurate judgment, illogical interpretation, or what is broadly called irrationality

Important

The analyst should be able to avoid cognitive bias from the sources and from his own education, origin, religion, culture and profession

Open Source Possibilities

Traditional Media Sources

- allow to establish chronologies
- contextualize other sources of information
- update other information
- requires validation

Internet Sources

- provide personal information
- allow in-depth research
- can be worked automatically
- allows data mining and pattern identification
- requires validation

Public Data Sources

- provide specific information about organizations
- may provide important strategic information
- important at the strategic level
- does not require validation

Report Analysis Sources

- provide very specific information about organizations
- fill in gaps from other sources
- requires specific knowledge
- does not require validation

Professional and Academic Sources

- provide important technical and scientific information
- essential at the technical level
- allow validation of other sources
- requires specific knowledge

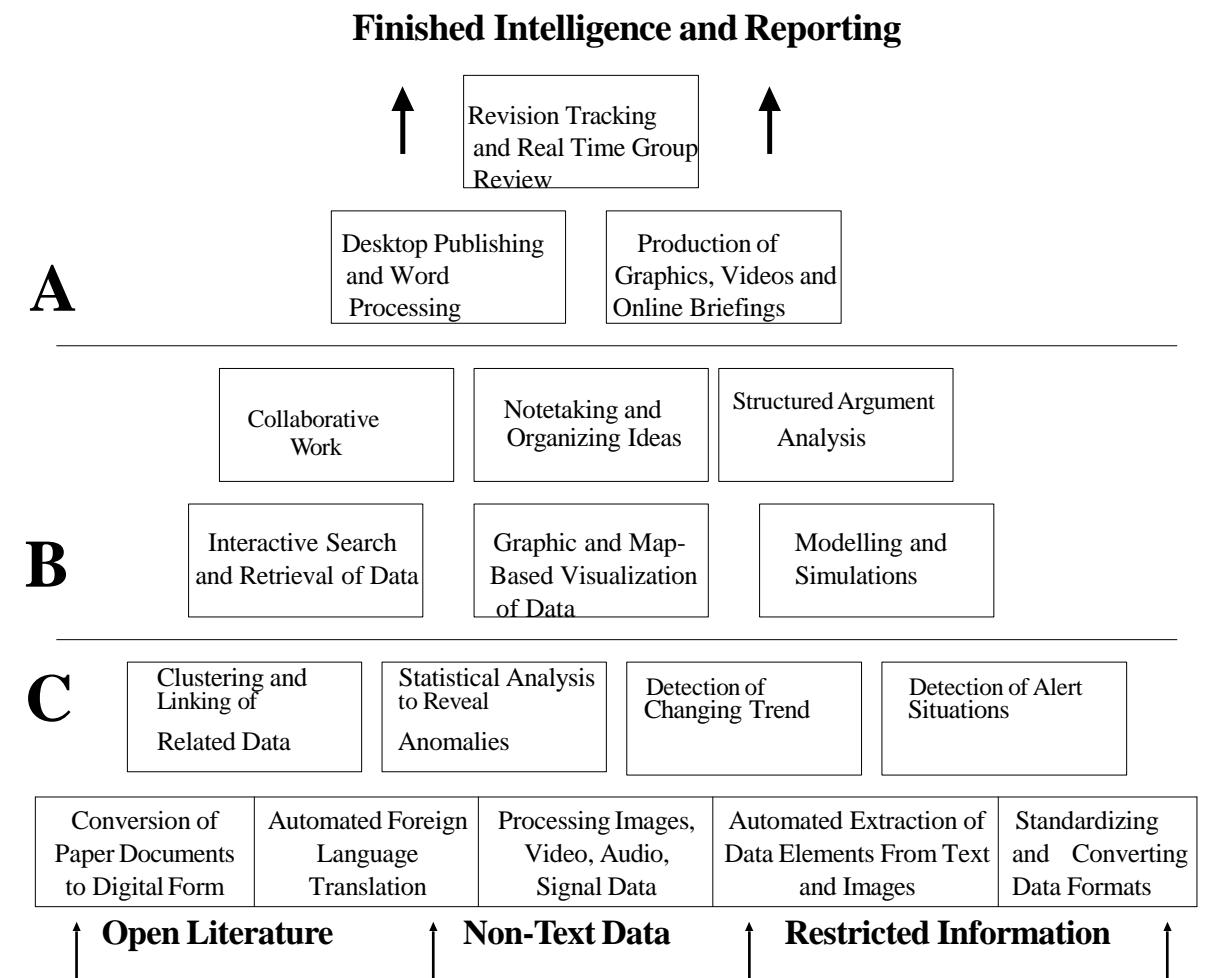
Automated Processing

Software functionalities for a optimal OSINT toolkit:

A Dissemination – publishing and production management functionalities

B Analysis – combines collaborative work tools with data visualization and manipulation tools with thinking tools

C Automated pre-processing



		Method	
		Non Intrusive	Intrusive
Type of Source	Open	OSINT	Illegal use for Crime
	Closed	Social engineering Crime	Hacking Crime

Resources

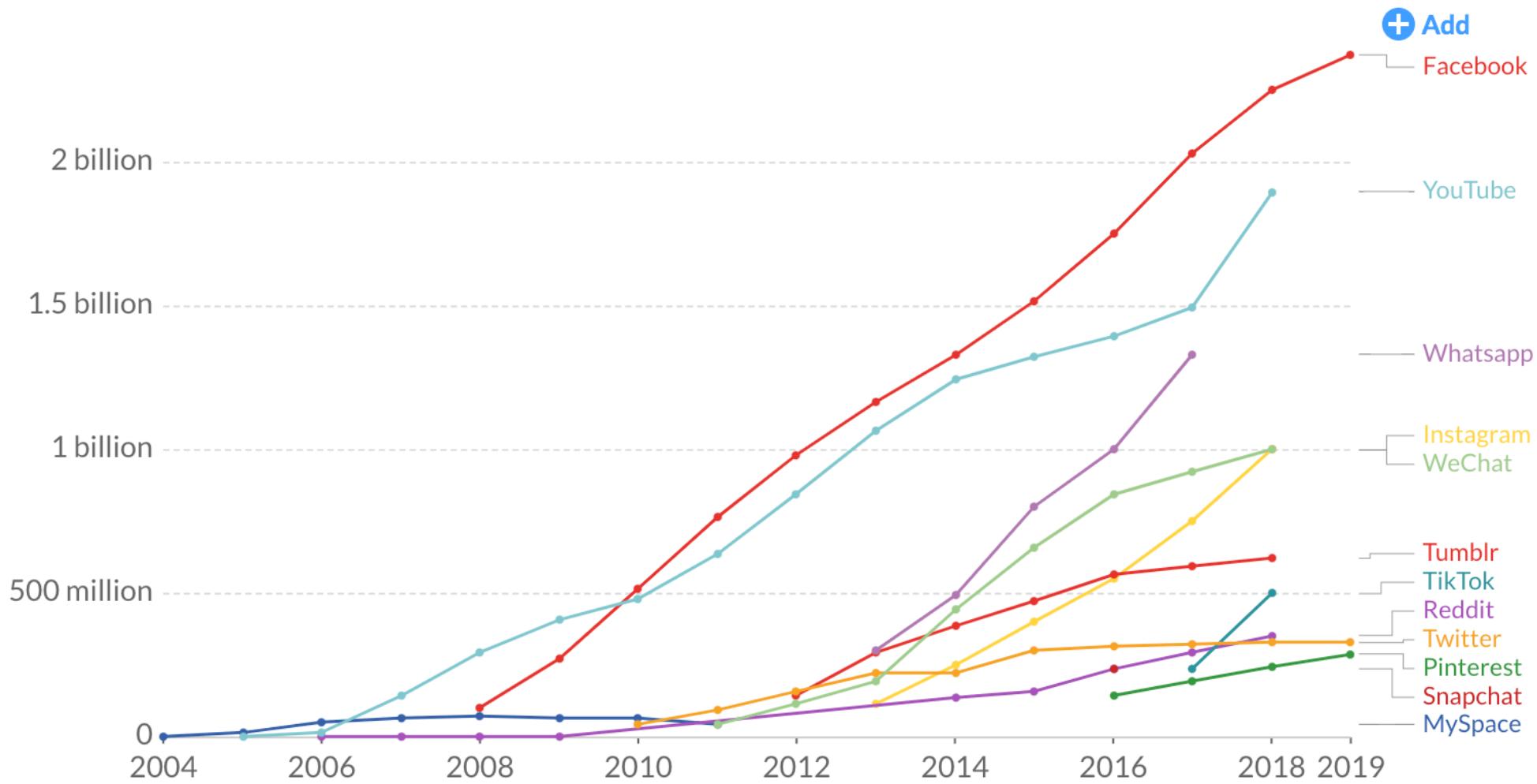
- crawlers
- analytics
- social media searches
 - Facebook, Twitter, YouTube, Instagram, FourSquare, etc
- digital footprint, anonymization, IP address as ID
- Application Programming Interface (API)
- darkweb



Facebook

Number of people using social media platforms

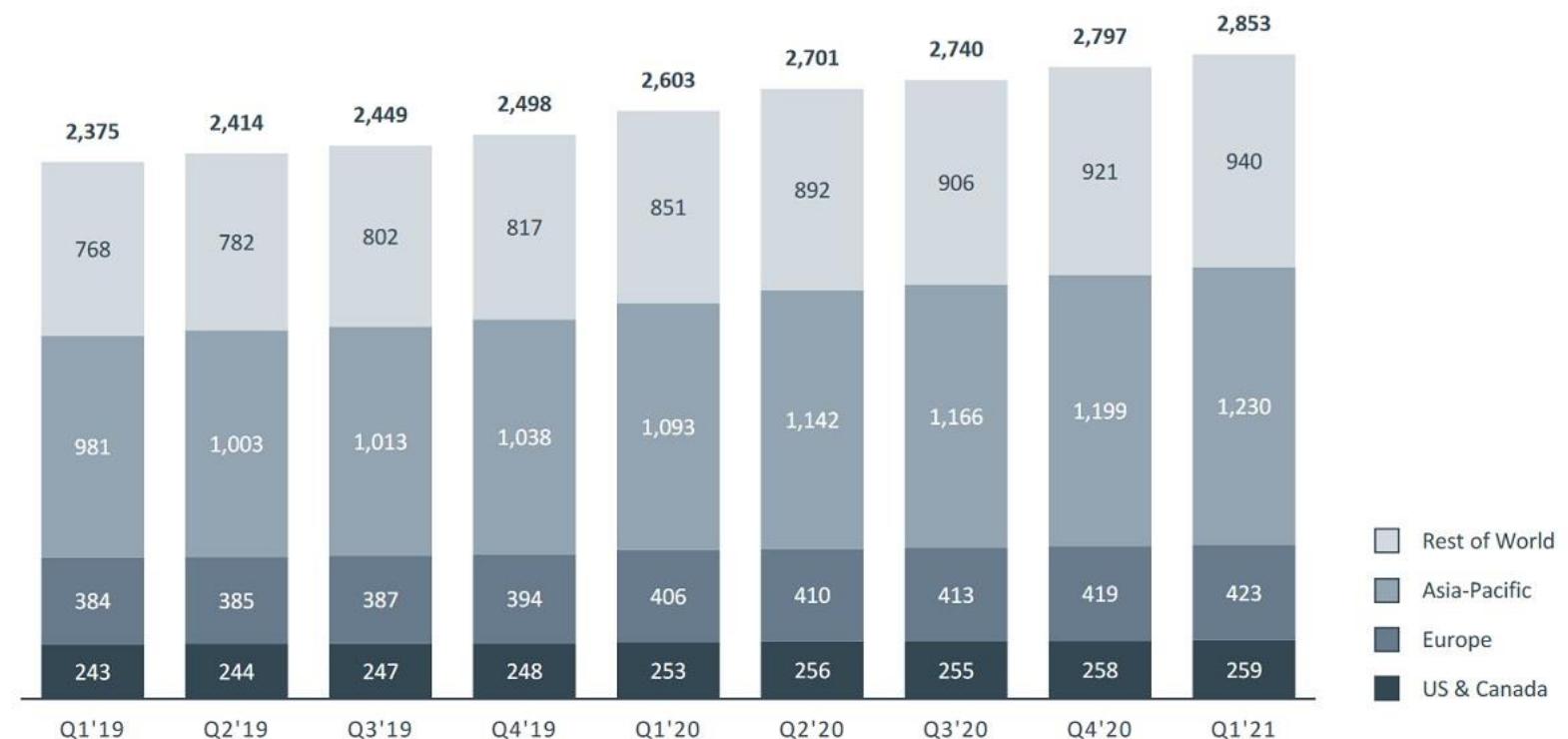
Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.



Facebook Monthly Active Users (MAUs)

FACEBOOK

In Millions



Please see Facebook's most recent quarterly or annual report filed with the SEC for definitions of user activity used to determine the number of our Facebook DAUs and MAUs. The numbers for DAUs and MAUs do not include users on Instagram, WhatsApp, or our other products unless they would otherwise qualify as DAUs or MAUs, respectively, based on their other activities on Facebook.



- more than 2300 million Monthly Active Users (MAUs) worldwide
- countries with more users: United States, India and Brazil

The Graph API

Is the interface that allows any computer program to query Facebook data



- Nodes represent “things”: users, photos, pages, comments, *etc*
- Links represent relationships between “things”
- Fields represent “things” attributes: birthday, name, *etc*

Graph API Documentation



- Facebook for developers
<https://developers.facebook.com/>
- Graph API Explorer
<https://developers.facebook.com/tools/explorer/>

Graph API examples



- `curl -i -X GET \
"https://graph.facebook.com/your-facebook-user-id?
fields=name&access_token=your-access-token"`
- `curl -i -X GET \
"https://graph.facebook.com/your-facebook-user-id/photos?
access_token=your-access-token"`
- `curl -i -X POST \
"https://graph.facebook.com/your-facebook-user-id?
email=you@your-email.com&access_token=your-access-token"`

Facebook Graph API Fields

- id
- about
- address
- age_range
- bio
- birthday
- currency
- devices
- education
- first_name
- gender
- hometown
- inspirational_people
- interested_in
- languages
- location
- meeting_for
- middle_name
- name
- political
- relationship_status
- religion
- significant_other
- sports
- quotes
- timezone
- work
- website
- cover
- ...

Facebook Graph API Links

- accounts
- activities
- albums
- books
- events
- friendlists
- games
- groups
- likes
- movies
- music
- statuses
- television
- videos
- family
- friends
- mutualfriends
- subscribers
- subscribedto
- ...

Twitter



Two ways to use Twitter API

- REST API – provides access to read and write data on Twitter
- Streaming API – provides access to read large – scale and real-time Twitter data globally



Twitter API documentation

<https://developer.twitter.com/en/docs>

Twitter API example

```
https://api.twitter.com/2/users/84092683?user.fields=location
-H "Authorization: Bearer $ACCESS_TOKEN
```

Twitter API – REST interface

GET [https://api.twitter.com/1.1/users/show.json?screen_name=\[user_name\]](https://api.twitter.com/1.1/users/show.json?screen_name=[user_name])

- 1. created_at
- 2. default_profile_image
- 3. description
- 4. expanded_url
- 5. favourites_count
- 6. followers_count
- 7. friends_count
- 8. id
- 9. lang
- 10. location
- 11. name
- 12. profile_background_image
- 13. screen_name
- 14. statuses_count
- 15. time_zone
- 16. verified
- 17. geo
- 18. in_reply_to
- 19. in_reply_to_status_id
- 20. status

Twitter Online Tools

Description	URL
Twitter advanced search	https://twitter.com/search-advanced
Twitter location search (enter GPS)	https://twitter.com/search?q=geocode%3A40.6306263%2C-8.6588713%2C1km&src=typed_query
Twitter time search (keyword + date)	https://twitter.com/search?q=Universidade%20de%20Aveiro%20since%3A2021-12-01%20until%3A2021-12-12&src=typed_query
All my Tweets (entire archive)	https://www.allmytweets.net
Closest friends	https://mentionmapp.com/
Twitter analytics	https://foller.me
OmniSci (mapped Tweets)	https://www.omnisci.com/demos/tweetmap
OMTM (mapped Tweets)	https://onemilliontweetmap.com
Real time monitoring	https://tweetdeck.twitter.com
Real time search	https://twitterfall.com
Twitter analytics	https://www.twitonomy.com
Identifies fake accounts	https://www.socialbakers.com/feature/fake-influencers-detection
Analyse associates	https://followerwonk.com/analyze
Analyse users	https://followerwonk.com/compare
Locate profiles by interest	https://followerwonk.com/bio
Topic explorer	https://tweettopicexplorer.neoformix.com

Google

Functionalities

Description	Query or tool url
Calculator and converter	<p>“how much is 20% of 135”</p> <p>“cos(3x)+sin(x),cos(7x)+sin(x)”</p> <p>“1 light year to au”</p> <p>“what is the volume of a cylinder with radius 4cm and height 8cm”</p>
Advanced search	https://www.google.pt/advanced_search
Translator	https://translate.google.com
Image search	https://images.google.com
News	https://news.google.com
Google Drive	https://drive.google.com
Book search	https://books.google.com
Video search	https://www.google.com/videohp
Academia	https://scholar.google.com
Finance	https://www.google.com/finance/
Trends	https://trends.google.com
Ngrams	https://books.google.com/ngrams



Google Hacking – Possibilities



- more assertive research
- less false positives
- access to devices connected to the network configured by default
- access to contente not available through normal search
- *etc*

Database of examples: <https://www.exploit-db.com/google-hacking-database>



Google Hacking – Examples

- [intext:"password" site:www.domain.com](#) ([example](#))
- ["contrato" filetype:pdf](#) ([example](#))
- ["text" filetype:doc | filetype:docx](#) ([example](#))
- [allintext:"*.@gmail.com" OR "password" OR "username" filetype:xlsx](#) ([example](#))
- [inurl:view/view.shtml](#) ([example](#))
- [parent directory Index of mp3](#) ([example](#))
 - mp4,mov,jpg,jpeg,...

Google Hacking – Examples



- [ext:txt intext:@yahoo.com intext:password](#) ([example](#))
- [password console-password ext:cfg -git](#) ([example](#))
- [intitle:"index of " "*passwords.txt"](#) ([example](#))
- [inurl:login.txt filetype:txt](#) ([example](#))
- [Index of /backup](#) ([example](#))
- [inurl:/intranet/login.php](#) ([example](#))

Google Hacking – Examples



- "parent directory" /appz/ -xxx -html -htm -php -shtml -opendivx -md5 -md5sums
- "parent directory" DVDRip -xxx -html -htm -php -shtml -opendivx -md5 -md5sums
- "parent directory" Xvid -xxx -html -htm -php -shtml -opendivx -md5 -md5sums
- "parent directory" Gamez -xxx -html -htm -php -shtml -opendivx -md5 -md5sums
- "parent directory" MP3 -xxx -html -htm -php -shtml -opendivx -md5 -md5sums
- "parent directory" Name of Singer or album -xxx -html -htm -php -shtml -opendivx -md5 -md5sums

Other Open Sources

- Allow searches by address
- Work by layers
- Preferred XML as information exchange format
- Allow detailed view
- Area measurement tools
- Allow interactive presentations

[Google Maps](#)

[Google Earth](#)

[Bing Maps](#)



Foursquare

- allows to search public places and goers
- allows to search people and places visited
- <https://pt.foursquare.com/city-guide>



YouTube

- Allows you to search for videos by themes
- Has a system of recommendations and alerts
- Allows you to search for users/commenters
- Allows you to view the entire comment history
- <https://www.youtube.com>

- “Whois” online
<https://whois.domaintools.com/>
- Way Back Machine – Internet Archive
<https://web.archive.org/>
- Google Cache
<https://cachedview.com/>
- Email Header Analyzers
<https://toolbox.googleapps.com/apps/messageheader>
<https://www.whatismyip.com/email-header-analyzer>
<https://www.gaijin.at/en/tools/e-mail-header-analyzer>

Utilities(1)

Utilities(2)

- Profile pictures

<https://www.picuki.com>
<https://gramhir.com>

- Reverse image

<https://tineye.com>
<https://www.google.com/imghp?hl=en>

- Fake photos

<https://fotoforensics.com>

DarkNet

Visible Web

- also called the Surface Web, Indexed Web, Indexable Web or Lightnet
- portion of the World Wide Web that is available to the general public
- search able with standard web search engines

Visible Web

versus

Deep Web

Deep Web

- has contentes that can not be found or directly accessed via search engines
- sites that are purpose fully designed to keep search crawlers out
- however, doesn't requires special browsers
- instead a direct link to access is needed

Dark Net

- is na overlay network (a network built on top of the Internet)
- was designed specifically for anonymity
- *e.g. Tor, I2P, Freenet, DN42 etc*

Dark Net

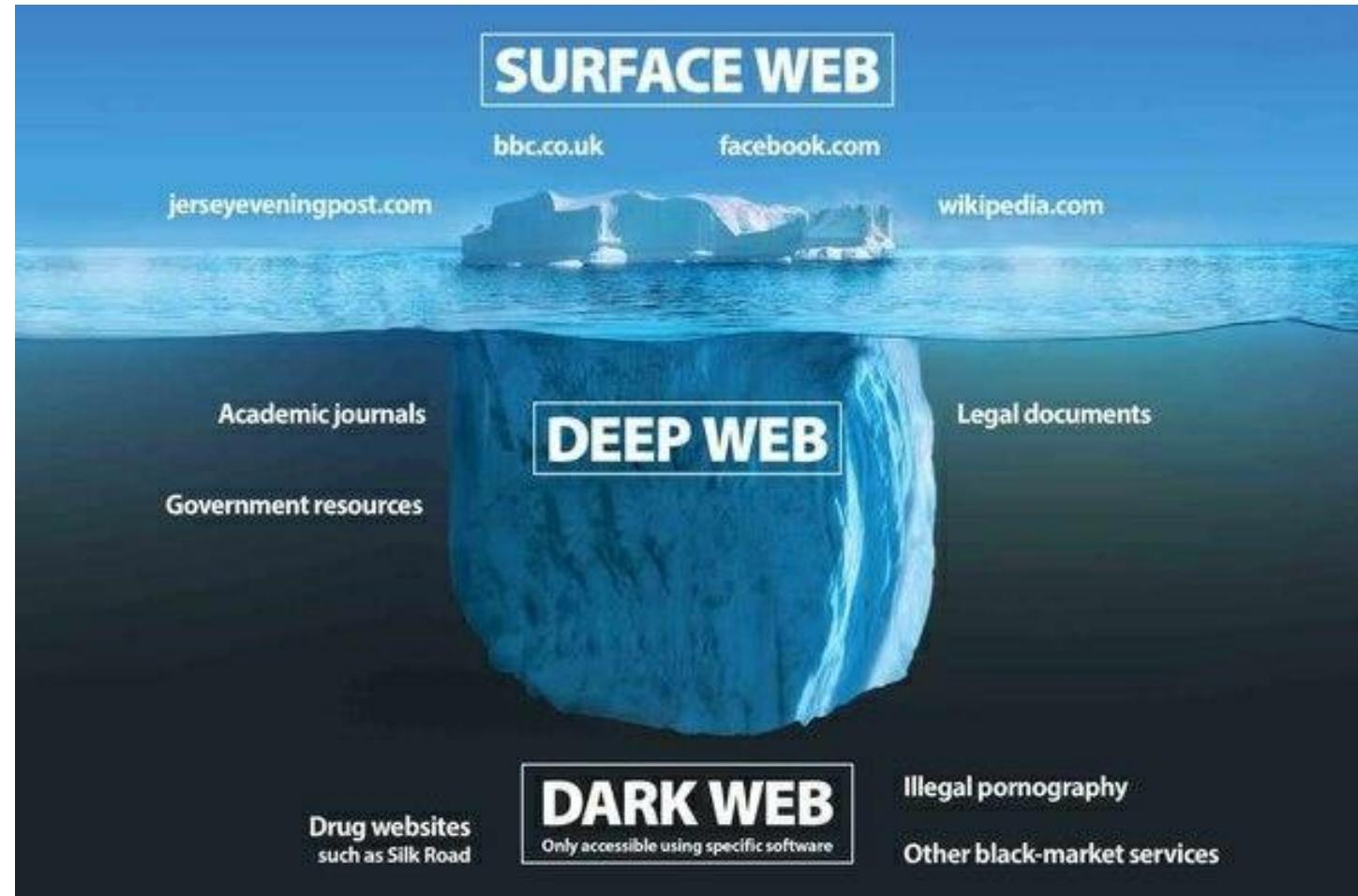
versus

Dark Web

Dark Web

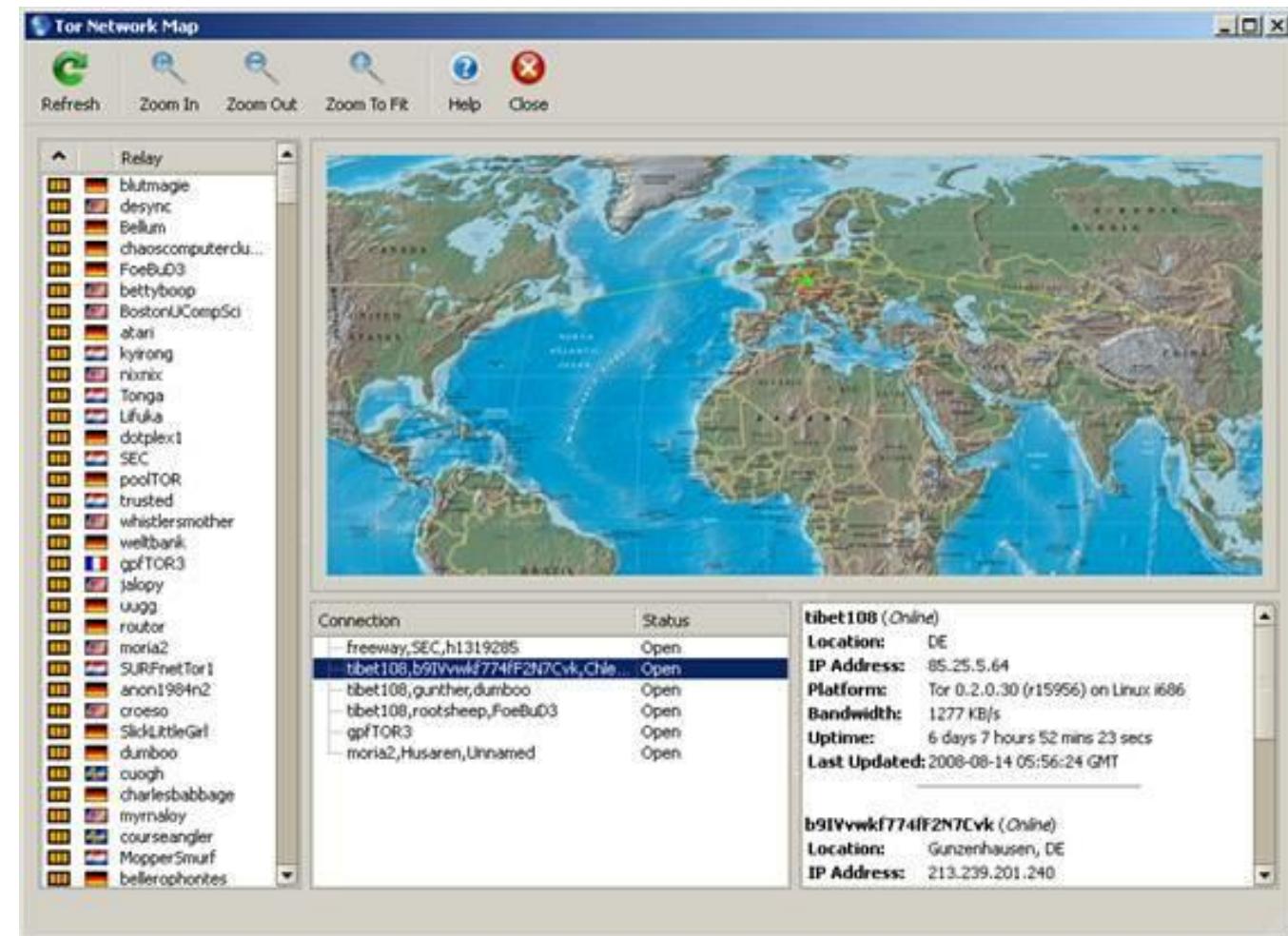
- Refers to websites on a Dark Net
- Market Places, Dark Net Markets, *etc*

Perspective



The Onion Router (TOR)

<https://www.torproject.org>



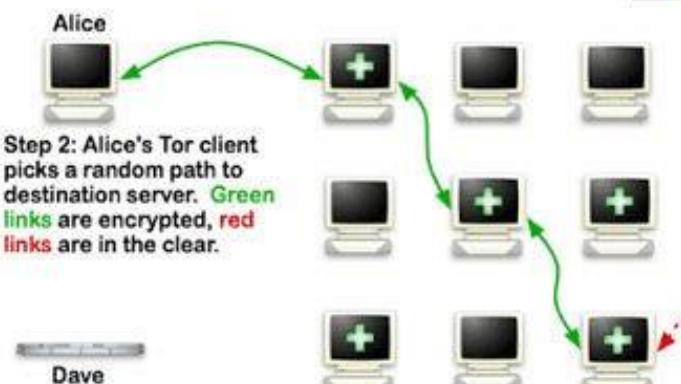


How Tor Works: 1

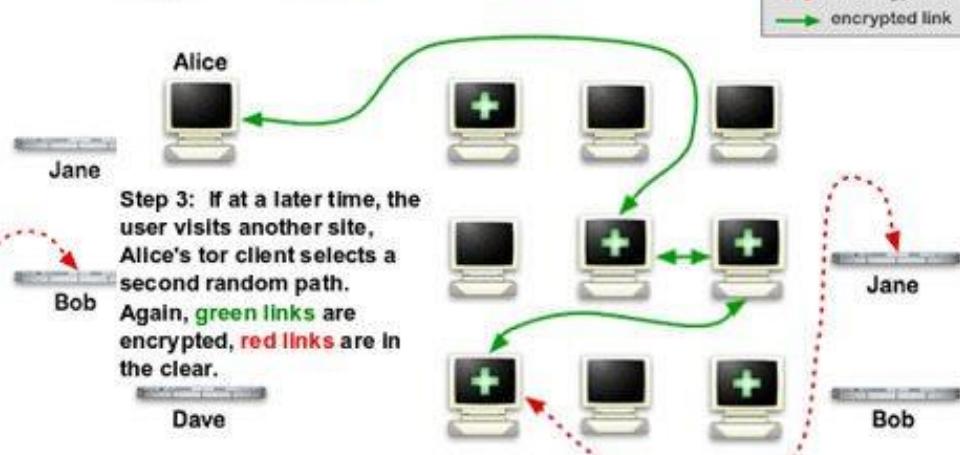


Tor node
... unencrypted link
— encrypted link

How Tor Works: 2



How Tor Works: 3

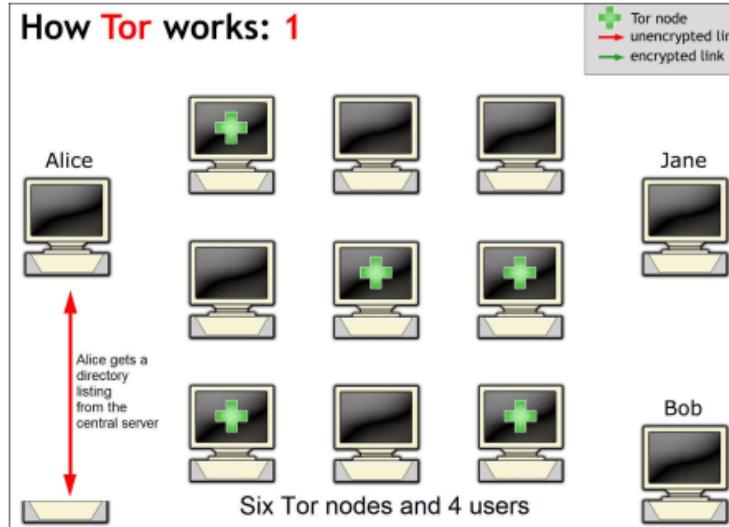


TOR Connections

TOR Connections

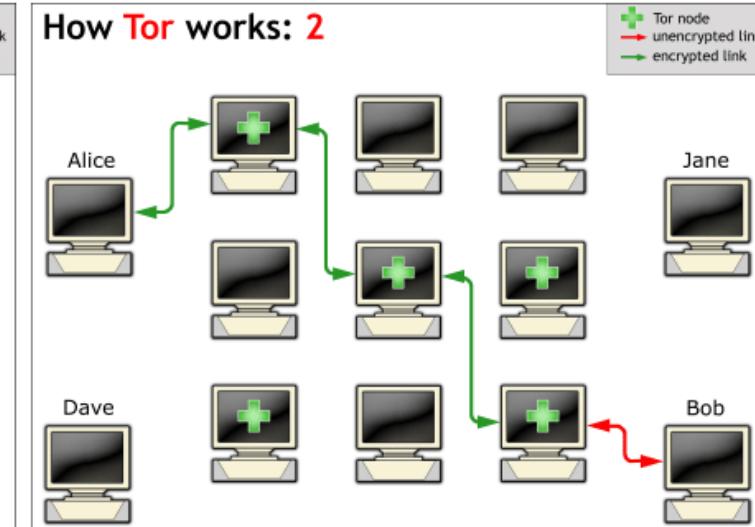
Connection set up

How Tor works: 1



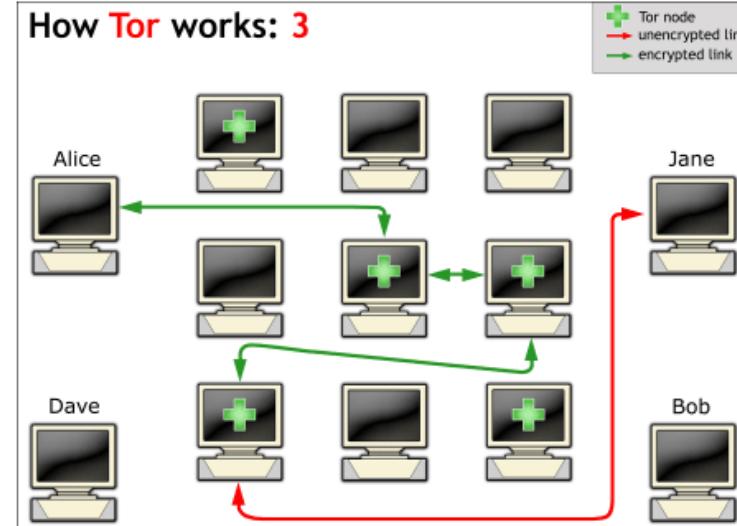
Connection

How Tor works: 2



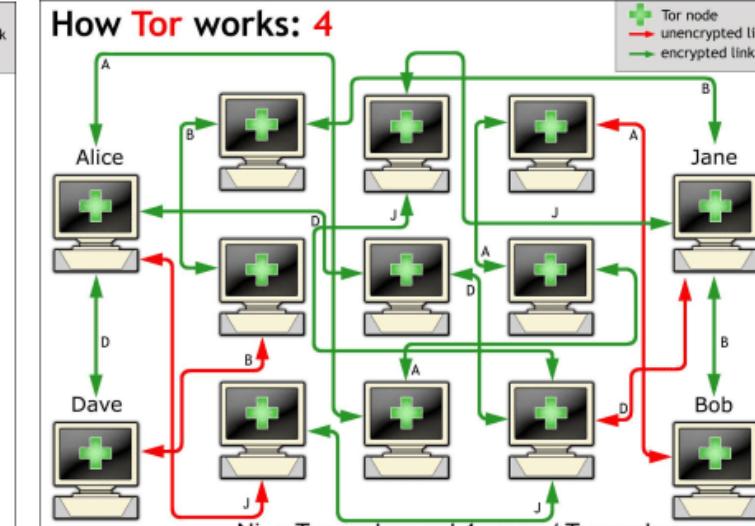
Connection Timeout - entry node change

How Tor works: 3



A real scenario - multi purpose node

How Tor works: 4



Useful Links:

The Hidden Wiki

http://zqktlwiuavvvqqt4ybvgvi7tyo4hjl5xgfuvpdf6otjiycgwqbym2qad.onion/wiki/index.php/Main_Page

OnionLinks

<http://jaz45aabn5vkemy4jkg4mi4syheisqn2wn2n4fsuitpccdackjwxplad.onion>

TorLinks

<http://torlinksge6enmcyyuxjpjkoouw4oorgdgeo7ftnq3zodj7g2zxi3kyd.onion>

Ahmia

<http://juhanurmihxlp77nkq76byazcldy2hlmovfu2epvl5ankdibsot4csyd.onion>

References

Websites

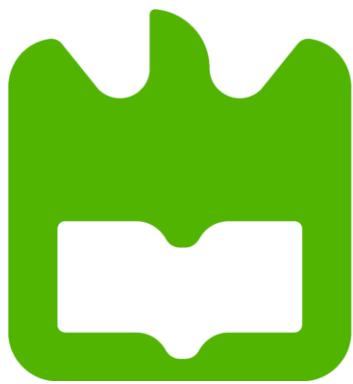
- <https://www.inteltechniques.com/links.html>
- <http://rr.reuser.biz/>
- <https://www.exploit-db.com/google-hacking-database>

- Michael Bazzell, “Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information”, 6th edition, 2018
ISBN-13978-1984201577
- NATO, “NATO Open Source Intelligence Handbook”, 2001

Books

- The Open Source Manual for the Polícia Judiciária, by António Jorge Filipe Fonseca (Phd) 2015
- Mário Antunes, Baltazar Rodrigues, “Introdução à Cibersegurança: A Internet, os Aspetos Legais e a Análise Digital Forense”, 1st edition, 2018
ISBN-13978-9727228614





universidade
de aveiro

Computer Systems Forensic Analysis

AFSC

Documentation and Reporting

Artur Varanda

School Year 2021-2022

All computer and digital media examinations are different

- consider all circumstances as you proceed
- some recommendations may not apply to every situations
- examiners may need to adjust to unusual or unexpected conditions in the field take into account that computers and other electronic devices are borderless
- ✓ multiple jurisdictions and agencies may be involved

About these guidelines

They just outline general principles; examiners may need to adjust to unusual or unexpected conditions

Electronic devices

In the context of these guidelines “Electronic devices” are all computer system components and other electronic devices, including digital and electronic media.

All electronic devices should be:

- examined physically
 - ✓ include a physical description and detailed notation of any irregularities, peculiarities, identifying markings, and numberings
- catalogued with photos and all details that allow its unique identification
- listed in a inventory
 - ✓ apply the tag system as mentioned in previous classes

When examining a computer:

- clock and time zone:
 - ✓ the system date and time should be collected, preferably from the BIOS
 - ✓ date and time should be compared to a reliable known time source and any differences noted
 - ✓ in the triage phase read the OS configured time zone. By default:
 - Windows OS configures the BIOS' clock to local time
 - Linux and Mac OS configure the BIOS' clock to UTC
- if the BIOS is accessible then
 - ✓ take note of the drives parameters and boot order
 - ✓ if present, take note also of system serial numbers, component serial numbers, hardware component hashes, etc

BIOS Access – Warning

Before trying to access the BIOS, first unplug all storage media (SSD, HDD, memory cards, etc) from the computer.

Examination of media should be conducted in a forensic environment

- one which is completely under the control of the examiner
- no actions are taken without the examiner permit them to happen
- when the examiner permits or causes an action, he must be able to predict with reasonable certainty the outcome of the action
- use a forensically sound operating system
- physical write-blocking devices are mandatory in OS that are not forensically sound
- avoid conducting an examination on the original evidence media, always use forensic copies

Digital evidence is said to be **forensically sound** if it was collected, analyzed, handled and stored in a manner that is acceptable by the law, and there is reasonable evidence to prove so.

Examples of things to write in the report

- number and type of partitions or volumes (in SSD, HDD, or other large writable media)
- number of sessions in optical discs
- file system type
- installed operating systems
- in some cases, *e. g.* illegal content, pedophile images, *etc*, include folder structure, filenames, date/time stamps, logical file sizes, hash values (MD5 and SHA256)
- files created by the OS including, but not limited to
 - ✓ boot files, registry files, swap files, temporary files, cache files, history files and log files
- list of installed applications
- user created files should be examined using native applications, file viewers or hex viewers
 - ✓ this includes text documents, spreadsheets, databases, financial data, electronic mail, digital photographs, sound, other multimedia files, etc

(continuation)

- report unused and unallocated space on each volume
 - ✓ search for previously deleted data, deleted folders
 - ✓ slack space data and data placed there by the user with the intent to hide it
 - ✓ deleted filenames of apparent evidentiary value
- report any irregularities or peculiarities in the system area of the volume (*i. e.* FAT, MFT, *etc*)
- report any hidden areas of the media, such as HPA
- report any recovered data and the process used
- forensic tools used
 - ✓ name and version of the tool
 - ✓ reference any validation test performed by examiner, the examiner's agency, or other reputable organization

Windows OS

Windows XP

C:\Documents and Settings

Windows 7 and newer versions

C:\Users

Inside users directory (some of these names may be translated accordingly to the configured local)

Desktop

AppData or Application Data Contacts, Favorites, Downloads

Documents, Pictures, Music, Videos, Virtual Machines, etc

From Windows Vista to newer versions: C:\\$Recycle.Bin

- it's a special directory, hidden and protected by SO
- stores files before their complete deletion
- has one subdir for each user with his SID *e.g.*

S-1-5-21-743533327-1274932866-229777416-1000

- ✓ in Windows always begin with S-1-5-21
- ✓ S = SID, 1 = version, 5 = Identifier authority
- ✓ 21-743533327-1274932866-229777416 = domain identifier
- ✓ 1000 = user account relative identifier (RID),
- ✓ RID can be paired with user account name in the systems' registry
 - HKLM\SAM\Domains\Account\Users – not accessible through the normal Registry view on a live system
- ✓ Administrator account RID is 500 by default
- ✓ missing RID in the registry means the account was deleted

AppData **replaces** Application Data from Windows XP

- stores data from applications, because they don't have security permissions to write data to the User's primary directories
- There are 3 sub-directories: Local, LocalLow and Roaming

AppData\Local

- stores user specific application data, e. g. temporary files and cache files
 - ✓ **Internet Explorer:** C:\Users\<User>\AppData\Local\Microsoft\Windows
 - ✓ **Firefox:** C:\Users\<User>\AppData\Local\Mozilla\Firefox
 - ✓ **Windows 8 and 10**
 - C:\Users\<User>\AppData\Local\Packages

AppData\LocalLow

- similar to the Local folder
- only to write low integrity data, e. g. Internet Explorer add-ons

AppData\Roaming

- specific data which are computer independent
- should roam with the user's profile
- ✓ data stored in this directory should be accessible when the user logs into another computer from the same network
- there maybe subfolders related to a variety of installed application

The Windows Registry is a central hierarchical database used to store information necessary to configure the system for one or more users, applications and hardware devices.

Windows Registry can be an excellent source of potential evidence

- there are two distinct types of registries
 - ✓ 95 / 98 / 98SE / ME
 - ✓ NT / 2000 / XP / VISTA / WINDOWS 7 / WINDOWS 8 / WINDOWS 10

95 / 98 / 98SE / ME

- <windir>/SYSTEM.DAT
- <windir>/USER.DAT
- <profiles>/<username>/USER.DAT – **Zero or more User files**

NT / 2000 / XP / VISTA / WINDOWS 7 / WINDOWS 8 / WINDOWS 10

- <windir>/system32/config/SYSTEM
- <windir>/system32/config/SOFTWARE
- <windir>/system32/config/SECURITY
- <windir>/system32/config/SAM
- <windir>/system32/config/systemprofile/NTUSER.DAT
- <profiles>/<username>/NTUSER.DAT – **One or more User files**
- <profiles>/<username>/<local appdata>/Microsoft/Windows/UsrClass.DAT – **User's Class File**

The Registry contains 3 major categories:

- User Specific Information
 - ✓ *e. g.* Desktop Preferences, Typed URLs, Messenger Contacts, Most Recently Used (MRU) Lists and Passwords
- System Specific Information
 - ✓ *e. g.* Network Settings, Time Zone Information, Registered Owner details, Last Shutdown Date/Time and Hardware information
- Application Specific Information
 - ✓ *e. g.* File Associations, Application Registration Information, *etc*

There are typically 5 root level keys in a Windows registry

- HKEY LOCAL MACHINE (HKLM) – is by far the most significant key in the registry
 - ✓ has subkeys named Software, System, Security, Sam and Hardware
 - ✓ stored in separate files: SOFTWARE, SYSTEM, SECURITY, SAM
- HKEY USERS – stored in one NTUSER.DAT file per user
- HKEY CLASSES ROOT – can be ignored, it is merely an alias for HKLM/Software/Classes
- HKEY CURRENT USER – is an alias for HKEY USERS for the user currently logged on
- HKEY CURRENT CONFIG – is an alias to the current hardware profile stored at
HKLM/System/CurrentControlSet/Hardware Profiles/Current

Tools to analyze the Windows Registry

regedit

%windir%\regedit.exe

- ✓ included in all Windows installations
- ✓ hides some registry entries

Registry Browser

https://lockandcode.com/software/registry_browser

- ✓ nice GUI, shows registry with similar structure as regedit
- ✓ shows all registry entries
- ✓ has a “device manager” like view
- ✓ has a nice report tool

Registry Report

<https://www.gaijin.at/en/files?dir=old-software>

- ✓ easy to use GUI, but requires some knowledge about the registry structure
- ✓ produces detailed reports

Reg Ripper

<https://github.com/kireyn/RegRipper2.8>

- ✓ powerful command line tool, very good to automate tasks
- ✓ requires good knowledge about the registry structure

Forensic Report

The forensic report should contain:

- Preamble – pages with roman numeration
 - ✓ Declaration of honor – sometimes it's a separate document
 - ✓ List of acronyms in alphabetical order
 - ✓ List of contents
- Body – start the Arabic numeration of the pages
 - ✓ Introduction
 - ✓ Several analysis chapters, usually one per device
 - ✓ Conclusions
- Epilogue
 - ✓ Bibliography – not always required
 - ✓ Appendixes – not always required
 - ✓ Glossary – recommended

Declaration of honor

- usually it's a separated document
- but can be included in the report

Example

I, (expert full name), holder of the identity card number XXXXXXXX, valid until 20XX-XX-XX, with professional domicile at the University of Aveiro, located at Campos Universitário de Santiago, Gloria, 3810 - 193 Aveiro, declare on a commitment to honor the veracity of the information provided in this report, of which this declaration is a part.

Introduction

- what are you looking for and why
- list of the devices being analyzed and their IDs
- explain the structure of the report

Example

The authors of this report were appointed computer experts in the NUIPC XXXXXXXX process under the protocol between the Aveiro University and the Office of Cybercrime of the Generals' Attorney Office. In this context, XX devices of the defendant (full name) were delivered for analysis. The devices are described in Table X with their assigned identification code (ID). [...]

In this case, the following diligences were requested:

- bla, bla, . . .

Analysis

- one chapter per analysis, *e. g.* for each device, or DNS analysis, *etc*
- for each chapter:
 - ✓ detail the device characteristics and ID
 - ✓ the procedures you made on that device, *e. g.* forensic copy, run anti-virus, *etc*
 - ✓ explain clearly what you found
 - ✓ anti-virus results
 - ✓ ...

Conclusions

- reconstruct the events based on the evidences found
 - ✓ and include a reference to the chapter and section in the report where you detailed how you found the evidence
- report all the found evidences, either incriminating or exculpatory
- always use clear text and avoid complex technical terms when possible, if needed reference to a glossary explaining the technical terms

Example

Under the NUIPC XXXXXX process, a total of XX devices of the defendant (full name) were submitted for analysis. In YY devices, no relevant information was found for the process (list the IDs).

Only ZZ devices contain relevant information (list the IDs). In this report it was possible to establish the following:

- bla, bla, . . .

Bibliography

- citations help to increase the credibility of the report and the expert
- cite reference books in the field, or other medium with high reputation in the field being analysed, *e. g.* RFC
- there are many bibliography styles, *e. g.* APA, Chicago, IEEE, Harvard, . . .
 - ✓ choose one and be consistent through the report
 - ✓ use tools to help format the references, *e. g.* JabRef, Mendeley, MS Word Reference tool

Example

- 1 Y. Rekhter et al. Address Allocation for Private Internets. RFC 1918 (Best Current Practice). RFC. Updated by RFC 6761. Fremont, CA, USA: RFC Editor, fev. de 1996. doi: 10.17487/RFC1918. url:<https://www.rfc-editor.org/rfc/rfc1918.txt>.
- 2 J. Postel. Simple Mail Transfer Protocol. RFC 821 (Internet Standard). RFC. Obsoleted by RFC 2821. Fremont, CA, USA: RFC Editor, ago. de 1982. doi:10.17487/RFC0821. url:<https://www.rfc-editor.org/rfc/rfc821.txt>.

Appendices

The Appendices section should be used to include information that helps to demonstrate, or complements, the expert conclusions.

- include documents **relevant** to the case being analysed:
 - ✓ reports generated by used tools
 - ✓ technical specifications from hardware vendors
 - ✓ reports, or parts of a report, produced by someone else
 - ✓ ...

Glossary

- explain technical terms in lay language
- this section is important for the non technical staff that must read the report

Example

Phishing – Phishing is the attempt to acquire sensitive information such as usernames, passwords, and credit card details (and sometimes, indirectly, money), often for malicious reasons, by masquerading as a trustworthy entity in an electronic communication. The word is a neologism created as a homophone of fishing due to the similarity of using a bait in an attempt to catch a victim. Phishing is typically carried out by email spoofing or instant messaging, and it often directs users to enter details at a fake website whose look and feel are almost identical to the legitimate one.

