ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse247942020.pdf
https://doi.org/10.30534/ijatcse/2020/247942020

High performance computer analysis for indexing emails in digital forensic analysis (Computer Forensics)

Enrique Lee Huamaní [1], Lizbardo Orellano Benancio[2], Danny Lezama Vega[3], Alexi Delgado[4]
[1]Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima-Perú, ehuamaniu@uch.edu.pe
[2]Oficina de Peritajes del Ministerio Público, Lima-Perú, orellano01@yahoo.com
[3]Oficina de Peritajes del Ministerio Público, Lima-Perú, dannylezamavega@hotmail.com
[4]Mining Engineering Section, Pontificia Universidad Católica del Perú, Lima-Perú, kdelgadov@pucp.edu

ABSTRACT

The fiscal investigation requires as scientific support the expert work carried out by specialized experts using equipment, hardware and software under certain methodologies, being the limitation for the analysis and presentation of the requested information the amount of high performance computer equipment, reason in this investigation the comparative analysis was carried out taking as sample a report generated with more than 22, 000 e-mails with FRED SR WSC621E equipment and MAGNET AXIOM forensic software, in order to demonstrate the computer performance for the generation of the data for the presentation of the expert report and its report in PST format of e-mails according to a requested object of study, for tax investigation, in front of commercial computers including laptops

Key words: Expertise, digital forensics, computer forensics, mail electronic.

1. INTRODUCTION

With the advance of technology, e-mail communication media are part of different licit and illicit activities, being the illicit ones the ones that deserve an investigation, that through an expertise the technical scientific support is achieved for the clarification of a certain fact, being the problem the great capacities of the storage media, the e-mails and their attached files, requiring high performance computational processes and the use of methodologies and forensic tools for the accomplishment of the expert work according to a certain object of study.

The expert's office of the Public Prosecutor's Office receives different requests for the analysis and extraction of emails from evidence contained in image copies, which have previously been secured by means of the HASH code, The problem is the limitation of high-performance computer equipment, so the objective is to evaluate the performance of a high-performance computer versus conventional computers, using as data the indexing of a half year of data from an image copy with more than 22,000 emails using the AXION MAGNET forensic software, being part of the problem the number of emails and their high computer demand for both analysis and reporting.

Electronic mail is the most widely used means of communication, shortening the time and distance between people and allowing for the exchange of multiple pieces of information. It is of legal interest as an evidentiary document in a trial [1].

2. METHODOLOGY

The methodology in digital forensic analysis or computer forensics focuses on the treatment of evidence at the crime scene that is related to securing the scene, the identification of evidence and its capture up to the activities in the forensic laboratory that are related to the preservation of evidence its analysis and the presentation of results, being the methodology the one related to the forensic laboratory.

2.1 Collection and/or identification

It is the process of collecting equipment, devices among others and documentation of seizure and abduction, which may contain the evidence or acquisition or forensic copy of the information. The identification of the digital evidence and its container means is carried out, being the compilation and acquisition processes previous to the present expert work. The chain of custody is verified to ensure the traceability and continuity of the evidence, as well as the identification of the physical evidence through the description and details such as brand, capacity, model and series of the device containing the image copy, verifying its correspondence as an original copy through the HASH codes, which for this expert work was carried out with the AccesData FTK Imager 4.2.1.4 software that guarantees the integrity of the information, which is detailed in the following image, Figure 1 shows the verification of the HASH codes.

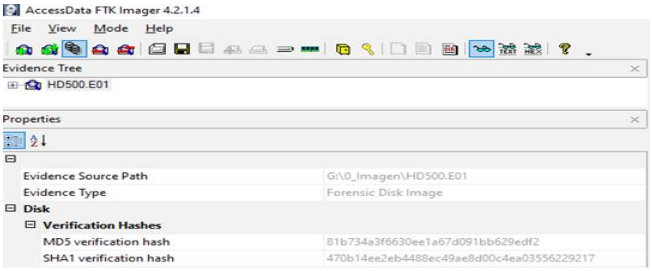

Figure 1: Verification of the HASH Codes that guarantee the integrity of the information in the image copy

2.2 Conservation/Preservation

The evidence is preserved through image copying with its corresponding HASH code that guarantees its integrity, safeguard and preservation as digital evidence, i.e., its originality for its subsequent admissibility as evidence for the investigation, i.e., its evidential value is preserved. Digital preservation is about safeguarding digital material in the short and long term. The evidence is preserved on the high-performance computer (Forensic Recovery of Evidence) FRED SR WSC621E detailed in the following image [2].

Figure 2: High Performance Computer: Forensic Recovery of Evidence

2.3 Analysis

The image copy is analyzed according to the object of study requested for the investigation and the analysis and interpretation of the digital evidence is carried out with the MAGNET AXIOM forensic software, with 665,771 elements indexed related to emails, which is detailed in Figure 3 and Figure 4.

Figure 3: Case summary with MAGNET AXIOM V3.9.0.18130 forensic software

Figure 4: 958,205 items found with MAGNET AXIOM V3.9.0.18130 forensic software

The object of this study is the indexation of the semester of the year 2014 that corresponds to 22,438 elements of e-mails that are detailed in Figure 5.

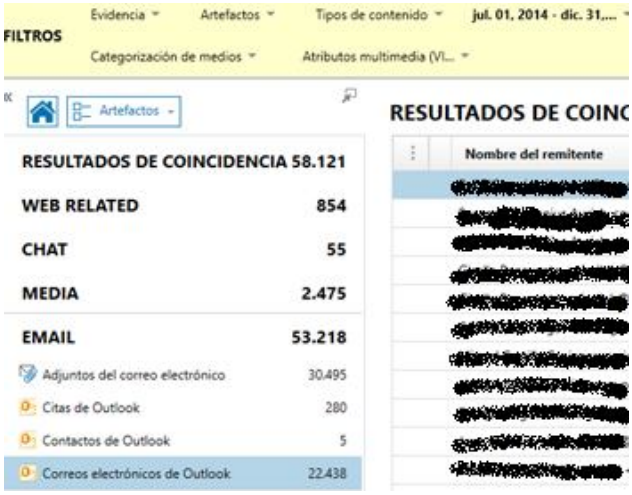

Figure 5: 22,438 emails found in one half year with MAGNET AXIOM V3.9.0.18130 forensic software

2.4 Presentation of results

The report of findings is made by means of an expert's report annexing the reports according to the object of study requested, selecting from the multiple export options the PST format detailed in Figure 6.

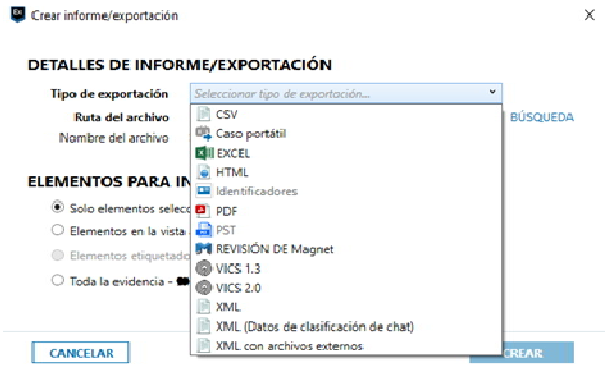

Figure 6: Different options for generating reports with MAGNET AXIOM V3.9.0.18130

PST format files contain elements of e-mail messages, appointments, task notes, and are used as current or archived mailboxes, also known as Personal Folder File (PFF) format [3] [4].

Being the email extraction reports for semesters of 4 years. The export is done in PST format, with their respective HASH codes, being the file with the name "Export.pst" for the comparison of the performance of the computer equipment the extraction report of a semester with a weight of 4,896 MB that is detailed in the following image.

Figure 7: Export of a semester in PST format for computer performance comparison

The expert's report is sent with its respective chain of custody and the evidence and digital reports are attached. Figure 8 shows part of the expert's report.

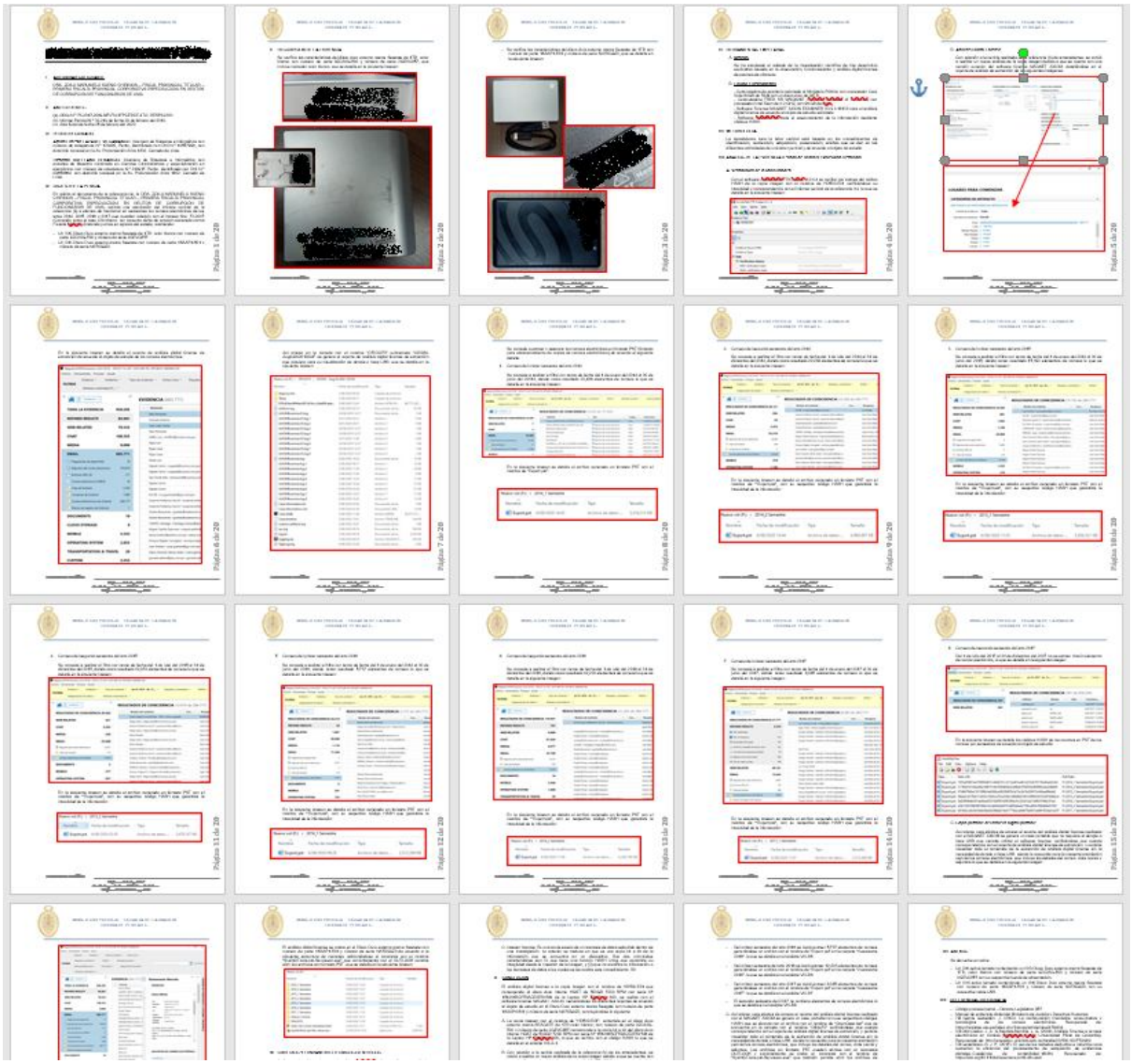

Figure 8: Part of the expert's report that details the presentation of the expert's work

Reports that can be analyzed, reviewed or any action considered pertinent within the investigation phase are made available for tax investigation by professionals other than an analyst or forensic expert [5].

3. STANDARDS

3.1 ISO/IEC 27001:2013 standard

ISO/IEC 07001:2013 is an international standard issued by the ISO standardization organization and is based on information security management and the implementation of security measures. The standards related to this research work are ISO 27037 and ISO 27042 [6].

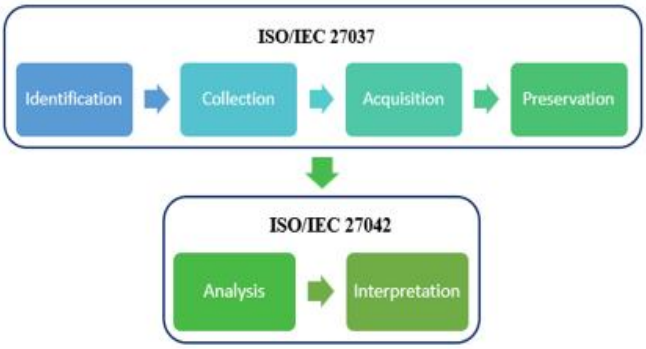

Figure 9: ISO 27037 and 27042 processes

3.2 Standard ISO/IEC 27037:2012

The standard ISO/IEC 27037:2012 "Information technology - Security techniques - Guidelines for identification, collection, acquisition and preservation of digital evidence" [7], defines the guidelines in information technology - security techniques - guidelines for cybersecurity. Being the specific activities related to the identification, collection, acquisition and preservation of digital evidence, whose basic principles are oriented to the processes of auditable, reproducible, defensible.

3.3 Standard ISO/IEC 27042:2015

ISO/IEC 27042:2015 "Information technology - Security techniques - Guidelines for the analysis and interpretation of digital evidence" [8], defines the guidelines for the analysis and interpretation of digital evidence. It defines the best practices for the selection, design and implementation of analysis and records of technical and scientific information so that the processes can be subjected to scrutiny when necessary.

4. RESULTS

From the analysis made to the secured image copy with its respective HASH code made with the forensic software MAGNET AXIOM 958,205 files were extracted being the total of extracted emails object of study 665,771 that were classified by semesters, being one of these semesters, the sample for the performance comparison analysis between the high performance computer (Forensic Recovery of Evidence) FRED SR WSC621E and a desktop and 3 laptop computers detailed in Table 1.

Table 1: Attributes of Cleveland dataset

Item	Feature	Processor	Memory	Operating System
HP I7	Laptop HP 250G6 modelo RTL8723DE	Intel Core I7 7500U 2.7HZ – 4 núcleos	8 GB	Windows 10
Laptop FRED	Laptop FRED-L forensic Laptop modelo P750TM1	Intel Core I7 9700K 3.6GHZ – 8 núcleos	64 GB	Windows 10
PC I9	Computadora Ensamblada I-9	Intel Core I9-9900K de 3.6GHz - 16 núcleos	64 GB	Windows 10
HP ProBook 4530s	Laptop HP ProBook modelo 4530s	Intel Core I7 -2620M – 4 núcleos	8 GB	Windows 10
FRED PC	FRED SR WSC621E	Intel Xeon Silver 4114 202 GHZ - 40 núcleos	96 GB	Windows 10

The sample taken for the performance comparison named "Export.pst" of size 4,968,697KB is a report generated that

took a time of 185 minutes in the FRED SR WSC621E, replicating the same report generation procedure in the FRED laptop that had an additional delay of 10 minutes, in the PC I9 of 31 minutes, in the HP I9 laptop 35 minutes and in the HP proobook 89 minutes, which is detailed in Figure 10.

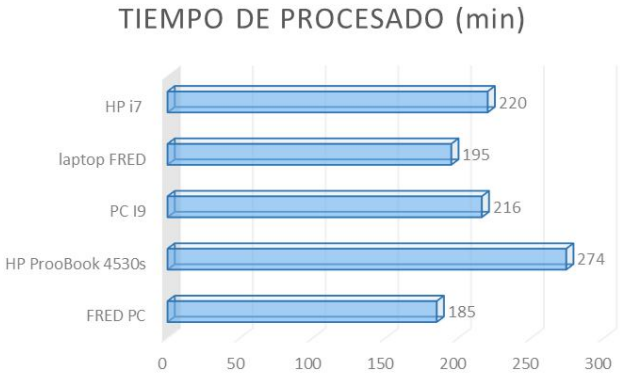

Figure 10: Sample processing time

5. DISCUSSIONS AND CONCLUSIONS

In the different expert tasks there is the inconvenience of estimating the processing time, analysis, indexing or report generation with different forensic software that includes MAGNET AXIOM. This research study is very important for the different institutions that perform expert tasks to estimate the time in report generation and the alternative of using commercial computer equipment. It has also allowed us in the area of digital forensic analysis, to estimate the processing load with different computer equipment that allows decision making in the face of the great burden of expert work and the fact of having a single high performance computer. As a proposal for the future, the implementation of custom-assembled computer equipment is planned, which will allow high performance computing with commercial hardware components to minimize costs.

We can conclude that the report generation time of the FRED computer was 3 hours and 5 minutes, while the FRED laptop had an additional delay time of 5% (3 hours and 15 minutes), the I9 PC had an additional delay time of 17% (3 hours and 36 minutes), The HP I9 laptop has a delay time of 48% (3 hours and 40 minutes) and the HP proobook (4 hours and 34 minutes), so it is feasible to use commercial computers as an alternative for the generation of forensic software reports such as MAGNET AXIOM in front of high cost and performance computers.

REFERENCES

1. B. P. De Gallo et al.**Ontología para el Análisis Forense de Correo Electrónico**, no. October 2016, 2014.
2. A. Salvador Benítez and A. Ruiz Rodríguez. **Metadatos para la preservación de colecciones digitales**, Cuad. Doc. Multimed., vol. 16, no. 16, pp. 48–60, 2005.
3. J. Andrés, O. Castro, L. Gisell, and B. Montilla. **Análisis Forense a correos electrónicos en Outlook**".
4. E. L. Huamaní, P. Condori, and A. Roman-Gonzalez. **Implementation of a Beowulf Cluster and Analysis of its Performance in Applications with Parallel Programming**, Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 8, pp. 522–527, 2019.
https://doi.org/10.14569/IJACSA.2019.0100869
5. J. Castellanos, **Metodo Deductivo E Inductivo**, Pontif. Univ. Javeriana, 2017.
6. K. Peciña, R. Estremera, A. Bilbao, and E. Bilbao. **Physical and Logical Security management organization model based on ISO 31000 and ISO 27001**, Proc. - Int. Carnahan Conf. Secur. Technol., pp. 1–5, 2011.
https://doi.org/10.1109/CCST.2011.6095894
7. D.-Y. Kao and G.-J. Wu. **A Digital Triage Forensics Framework of Window Malware Forensic Toolkit**. 49Th Annu. Ieee Int. Carnahan Conf. Secur. Technol., pp. 217–222, 2015.
8. W. Ni, Z. Guo, G. Sun, and H. Chi. **Investigation of forest height retrieval using SRTM-DEM and ASTER-GDEM**, Int. Geosci. Remote Sens. Symp., pp. 2111–2114, 2010.