Cloud Forensics Triage Framework (CFTF)

GIAC (GCFA) Gold Certification

Author: Michael Beck, mbeck.eagle@gmail.com
Advisor: *Clay Risenhoover*

Accepted: *23-June-2021*

Abstract

Digital media forensic investigations come in multiple forms and span single assets - from thumb drives, laptops, mobile phones, or a single email server to large-scale corporate incident response actions. Corporate network investigations are when analysts can become overwhelmed with the volume of internal hosts of interest, which must be forensically triaged and analyzed. The pressure to produce evidence to support or refute a case is still the same. Analysts need to deliver the evidence as quickly as possible and maintain proper evidence handling procedures. Endpoint Detection and Response (EDR) tools perform a great job identifying these systems and providing a platform to collect data. The next step of preparation and analysis of these hosts must be done and is time-consuming. This circumstance is where a Cloud Forensics Triage Framework (CFTF) can leverage cloud resources to set up a scalable and automated forensic triage framework and benefit the digital media forensic investigators. The research will explore the possibilities of using a mixture of traditional forensic media collection processes and modern cloud technologies to determine if reducing the time it takes to deliver processed media benefits the overall mean time to deliver results.

Will this reduce the time required to find the needle in the stack of needles?

1. Background

The ability of digital media forensic investigators to adapt and leverage current technology is crucial for the successful analysis of assets. The rapid adoption of cloud services creates a landscape that can be utilized to benefit investigation teams. Cloud Service Providers (CSP) are primarily Amazon AWS, Microsoft Azure, and Google Cloud Platform (GCP). For this research, Amazon AWS will be used. The CSP provides access to a user-controlled, secured, scalable, and customizable environment. These environments are scalable through size by increasing CPU, RAM, Disk, GPU and scalable by quantity, zero machines to infinite machines. Available resources determine the constraints on the scale within the CSP and budget. The benefit of this scale allows for Just in Time (JIT) availability of resources to give the investigation team the appropriate resources at the appropriate time. Budget modeling of JIT also means that the cost is only incurred when used, with significant cost savings when idle.

2. Traditional Forensics Lab

In the current cyber forensic lab, hundreds of thousands of dollars are spent on purchasing Forensic Recovery of Evidence (FRED) systems, the supporting storage arrays and network appliances, and resources for maintenance. According to the Department of Defense Cyber Crime Center (DC3) Defense Cyber Investigations Training Academy (DCITA), when asked what an investigator should use for a forensic data processing machine, the answer has been "Purchase the best equipment that you can afford." The cost of a FRED workstation typically runs between $6,000 - $25,000. In large corporate organizations, digital forensic teams consist of 6-10 examiners, each leveraging multiple FRED workstations; this capital cost can easily exceed $150,000 and be refreshed every 3-5 years. These FRED workstations are limited in their method for serial processing of source media images, meaning that each FRED processes one image at a time, only processing the following image after the former was complete.

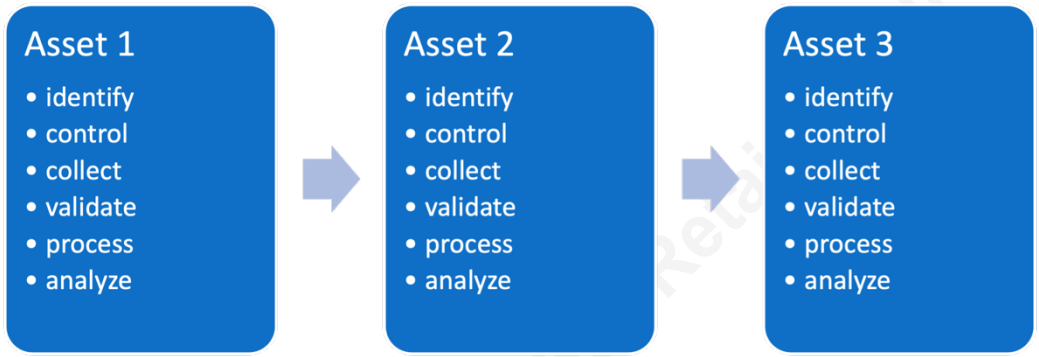

Figure 1: Serial Processing

The research will determine if there are more efficient mechanisms to collect each of the assets in a parallel manner.

2.1. Long mean time to analysis (MTTA)

In the traditional digital media forensic collection and acquisition process, there is a long Mean Time to Analysis (MTTA), which is the average amount of time it takes for a piece of digital media "asset" to be ready for analysis by an examiner. There are variables in this MTTA which we can control and shorten for large-scale investigations within a corporate environment. The variables are as follows: identification of assets, control of assets, collection of assets, validation of assets, and processing of assets. These variables constitute the MTTA and, when followed serially, continue to push the time it takes for data to be ready for analysis farther to the right of the time scale.

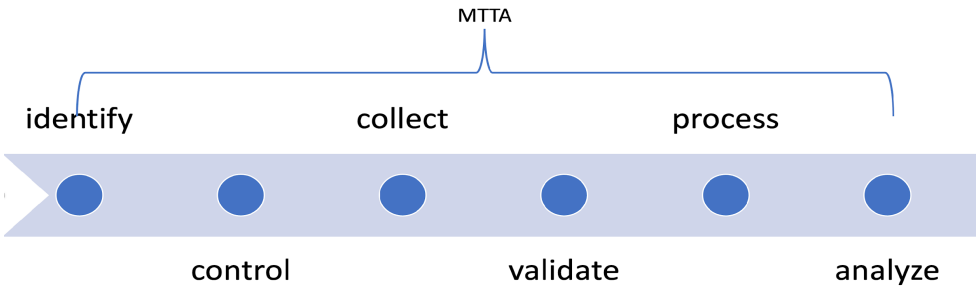

Figure 2: MTTA

The goal is to use parallel processes, cloud resources, and modern technologies to reduce the MTTA by taking portions of this process, specifically collect, validate, and

process, and moving them from serial tasks to parallel tasks. The remaining steps in the process will remain serialized.

2.2. Maintains Chain of Custody and Integrity

Maintaining the Chain of Custody and the integrity of digital media is a cornerstone of forensic analysis. If broken, it is challenging to prove facts. Chain of Custody is upheld in traditional forensic labs by having assigned digital media custodians and media handling procedures that are strictly followed. Media integrity is managed via hash value, maintaining the same hash value throughout the life of the digital media image. This concept is not new and should not change. If the hash does not match, the examiner must go back and find out what changed. The time necessary to compute the integrity hash of digital evidence is variable based on the resources available and the size of the digital evidence file. For a single piece of evidence, this could be measured in minutes. When compounded by hundreds of digital evidence files, computed in serial by a single processor, this becomes hours.

3. Problem

The traditional cyber forensics lab does not scale well with modern technology and the scope of investigations. Continuing to analyze incidents one by one slows the investigator's progress and keeps these teams under significant pressure from leadership teams to "find the answer." A solution needs to be developed to reduce the time it takes to reach the answer. Reducing the MTTA moves investigation teams closer to the end goal and "finding the answer."

4. Hypothesis

Small Incident Response teams have the responsibility of having answers quickly. These teams must provide a framework to deliver those answers quickly, reliably, and accurately and use an automated framework to accomplish these tasks. The use of Cloud Service Providers (CSP) as the Platform as a Service (PaaS) can be leveraged as this

model to model and frame an environment for use by digital forensic examiners. This framework can be used to collect digital media and forensic triage packages, validate the integrity of the data, and process the contents to be prepared for the digital forensic investigator to conduct their analysis. The research looks to hypothesize that this CFTF (Cloud Forensics Triage Framework) can provide this capability to the investigation team and be leveraged for use in any Internet-connected and centrally managed infrastructure.

The CFTF will allow forensic investigators to have data made available for analysis in a more efficient manner. The work will attempt to limit the human and technical inefficiencies that currently inhibit forensic data's expedient availability.

The CFTF is designed to be a framework to free up the investigators' critical resources to focus on the analysis of data. The collection, validation, and processing of forensic media are repeatable and must be immutable. The use of this framework has the added benefit of creating that immutability to the data being processed.

Taking advantage of technology is part of working in cyber; it is a de-facto statement that the forensics community is curious. This research and the CFTF helps to look into what can be performed using some of the basics of cloud platforms.

5. Research Method

The core capability of the CFTF will be to leverage parallel collection, validation, and analysis of digital media by simultaneously deploying collection packages to suspect systems within the corporate enterprise. The collection packages will execute pre-defined triage or digital media jobs on the suspect systems and upload the output to data repositories contained within CSP. Once the data is securely transferred to the CSP, an automated job will execute to instantiate virtual machines to run an equal number validation and processing system to the number of systems collected. For example, if 100 systems are collected, 100 validation and processing systems will be instantiated. These automated systems will tag, validate, and process each digital media collection in a unique, secured, and standardized environment. Upon completing the jobs, the output will be moved to a separate secured location within the CSP for analysis by the forensic investigation team. The 100 validation and processing systems will be terminated.

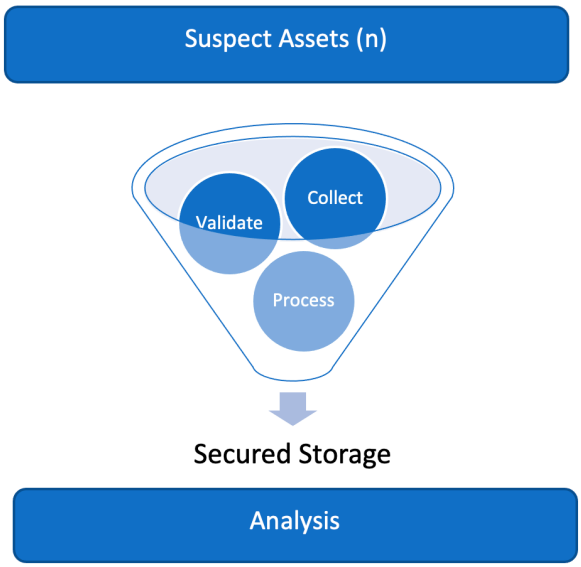

Figure 3: CFTF Framework

5.1. Prerequisites

The following prerequisites have been considered for this research, as these factors impact the deployment of tools and the ability to collect and process data effectively.

- The investigation team must have a working relationship with the network and systems security team within their environment.
- The investigation team must have working knowledge or access to a team to support services to a CSP.
- The investigation team must have a working knowledge of open-source collection and processing tools such as (KAPE, Cyber Triage, F-Response)
- The investigation team must have a working knowledge of Common of the Shelf (COTS) forensic analysis platforms such as (Magnet AXIOM, EnCase, FTK, X-Ways)

- The investigation team must have the ability to deploy collection packages to suspect systems within the corporate environment (SCCM, WinRM, Logon Scripts, EDR software)

5.1.1. Testing Procedures

A baseline physical lab was set up to capture suspect images, validate, and triage to create a control environment. The system used is a COTS Sumuri Talino forensic workstation, equipped with Intel i9-7900X 20 logical core processors, 128GB RAM, multiple Solid State Hard Disk Drive (SSD) arrays. The operating system is Windows 10 1909. Installed software includes KAPE and Magnet AXIOM Cyber. For the testing process, a pool of suspect virtual machines (VM) was configured using VMWare Workstation 15.5.7. Each of the suspect VM's represented hosts on an enterprise-connected Local Area Network (LAN). The VM's are configured as Windows 10 1909, with a shared administrative credential.

For the baseline serial collection of data, the emulation of collecting each suspect VM was conducted in order, using the current response methodology: Identify, Control, Collect, Validate, Process. The collection was executed using two different tools, KAPE for triage data and Magnet AXIOM Cyber for physical disk images. Validation of the collected data was performed using the Get-FileHash functionality of Windows PowerShell.

```
PS C:\Users\micbeck> Get-FileHash '.\                    ance.exe' -Algorithm SHA256 | fl


Algorithm : SHA256
Hash      : CAA66828909F617CD32FF93AA6021BB0B97A645BB777106BA634BB20C50C14E1
Path      : C:\Users\micbeck\                    :ance.exe
```

Figure 4: Sample Integrity Hashing

Building the baseline created the benchmark for the length of time required to perform a collection on ten suspect network-connected enterprise assets, similar to the current collection methodologies employed on the enterprise.

Parallel data collection is conducted using the same lab systems and tools, with changes to the collect, validate and process steps of the process. For this variable, the use of equal number collection, validation, and processing VM's were configured in Amazon AWS to target each of the suspect systems. Building the cloud systems creates a one-to-one relationship between the collector and suspect. This new relationship allows for improved efficiency in collecting, validation, and processing times for enterprise investigations.

Configuring collection and analysis VM's in AWS as Windows Server 2019, installed software included KAPE and Magnet AXIOM Cyber. The AWS VM's are to emulate the physical forensic workstation as close as possible. Staging of the Collector VM in AWS creates a baseline system that allows for efficient replication and repetition of the same tasks. This baseline system is configured to be immutable to preserve a clean state for each investigation and suspect collection. The baseline configuration should be set to meet the unique guidelines of each organization's requirements. This collector system is built to execute the three steps of collect, validate, and process. Once these are completed and the data is exported to a central repository for analysis, the system is terminated.

The process to perform the collection on the suspect systems is as follows for each tool.

KAPE

1. Execute pre-established KAPE package to collect triage data from the suspect system
2. The KAPE package is written to the AWS collector VM
3. The KAPE package is validated
4. The KAPE package is processed

Magnet AXIOM Cyber

1. Deploy AXIOM_collect package to collect physical disk image from the suspect system
2. The collected physical image is written to the AWS collector VM

3. The collected physical image is validated
4. The collected physical image is processed with Magnet AXIOM Cyber.

5.1.2. Tools

The tools used for this process are defined in this section and can be tailored to fit any organization. This paper is not an endorsement for any of the tools used. All credit is due to the original developers.

5.1.3. KAPE (Kroll Artifact Parser and Extractor)

Other collection platforms could be utilized to perform this action. KAPE was selected for the collection platform for ease of use, capability, repeatability, and open source. The most recent documentation for KAPE can be found at the following website: https://ericzimmerman.github.io/KapeDocs/#!index.md.
It is highly recommended to use the most recent version of KAPE; updating the core platform and modules are trivial and can be performed by running the following command:

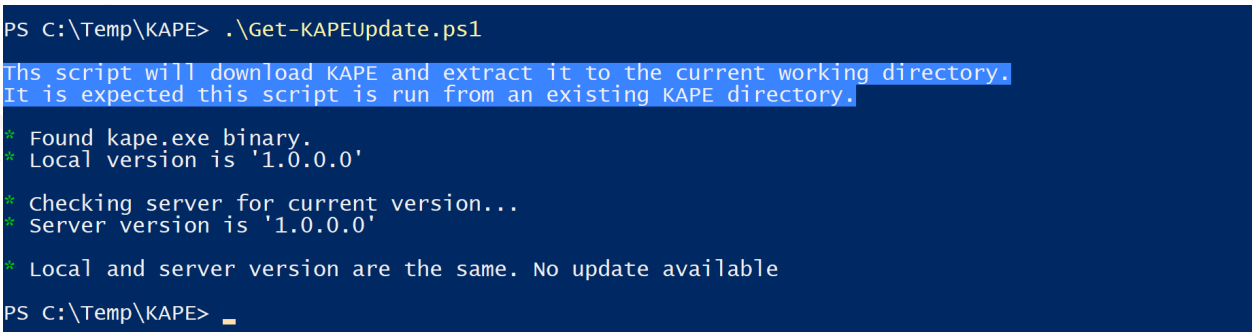

Figure 5: Updating KAPE

The configuration of KAPE is set to the following standard, allows for repeatability:

- Software version: 1.0.0.0
- Collection package: !SANS Triage
- Output package: vhdx

5.1.4. Magnet AXIOM Cyber

Magnet AXIOM Cyber is used for collection and baseline analysis. The use of a second platform for collection is required for remote acquisition of a full disk image. Magnet AXIOM Cyber is also used for the post-processing of collected data. Contact Magnet Forensics for more information on the use of AXIOM Cyber.

- Software version: 4.11.0.24063
- Collection Type: Remote Collection
- Collection Agent: Remote_Agent / Port TCP:33333
- Collection Type: Physical Drive Image

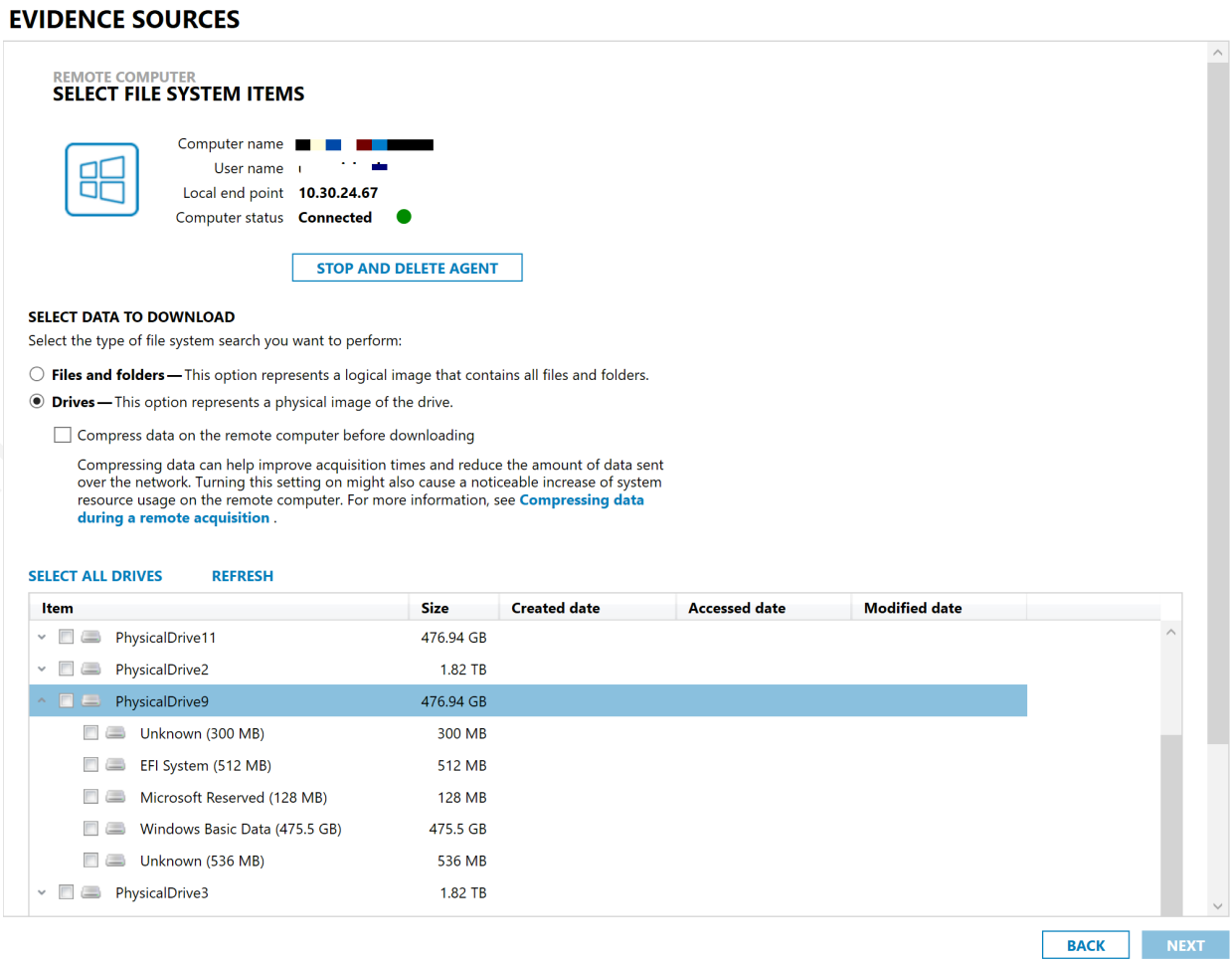

Figure 6: Magnet Cyber Remote Acquisition

5.1.5. VMware Workstation 15.5.7

Local target workstations are configured in VMware Workstation 15.5.7 to provide a sampling of configured systems to emulate a corporate environment. This virtual environment provides a method to control the systems being analyzed. The use of VMware manages some of the variables which are present in a corporate network infrastructure.

5.1.6. Amazon AWS

Amazon AWS is chosen as the Cloud Service Provider. Any of the primary CSP could be used in place of AWS. AWS is selected based on familiarity with the product and current compatibility with the chosen tools for collection and analysis. Costs and risks associated with all the CSP and due diligence will need to be performed to measure those costs and assess the risks.

The AWS instance is configured with the protection of data at the core, access to the instance is controlled via best practices to limit the source of inbound connections to only that of the required entities. It is recommended that business best practices are followed when configuring and operating in CSP environments. This section is a guide for configuring a primary AWS instance to collect and process triage data from a managed network infrastructure.

EC2 instance setup

Research note: User experience in AWS, make sure that the correct zone is used when logged in. This scenario happens between logoff/logon, and the assignment zone is not necessarily consistent between sessions.

Figure 7: Assignment zone

The instance type of t2.xlarge was selected for this research. These instance types provided sufficient resources to conduct the tests while balancing time, performance, and cost. When working on an actual enterprise investigation, it is advised to adjust the resources to fit best the applications running within the instance operating system and the

budget allotment for each project. In this case, the running cost of this instance is $0.2266/hour.

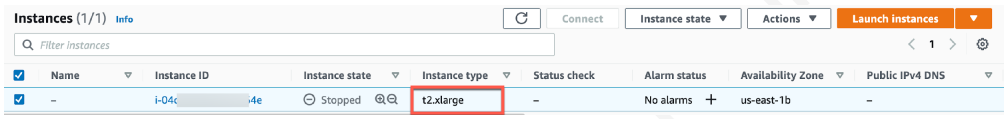

Figure 8: Instance type

Configuration of the storage arrays for this instance can be manipulated to fit the needs of the investigation. For this research, a base volume of 100GiB was selected to support the operating system and test collections. When performing analyses, it is advised to follow best practices of separating the operating system from the processing disk by adding a separate EBS volume for the case files. The management and configuration of storage nodes are conducted within the storage module of the AWS EC2 instance.

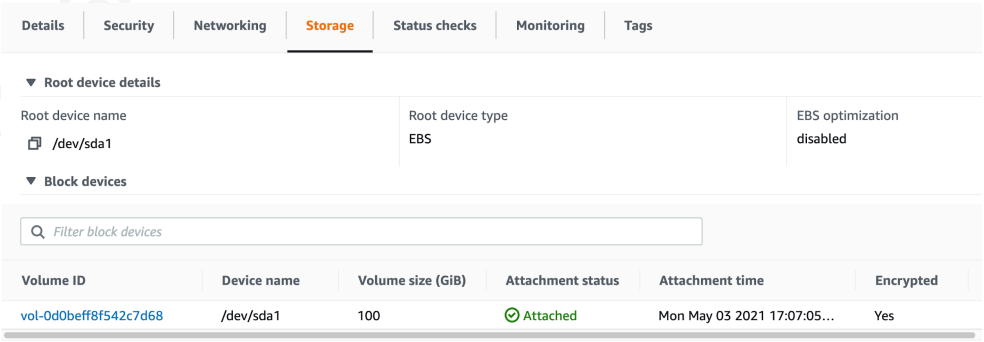

Figure 9: Storage

AWS EC2 network setup

Security is critical to the work of digital forensics. For this research project, a simple least privilege security group was configured only to allow inbound traffic from a trusted IP address. The IP address or address netblock would be a designated external IP address for enterprise environments dedicated to the incident response or forensic investigation team needed to access the cloud environment. This security group would be

configured only to allow the ports (TCP or UDP) required to perform authorized actions. All other traffic is dropped as the AWS EC2 perimeter.

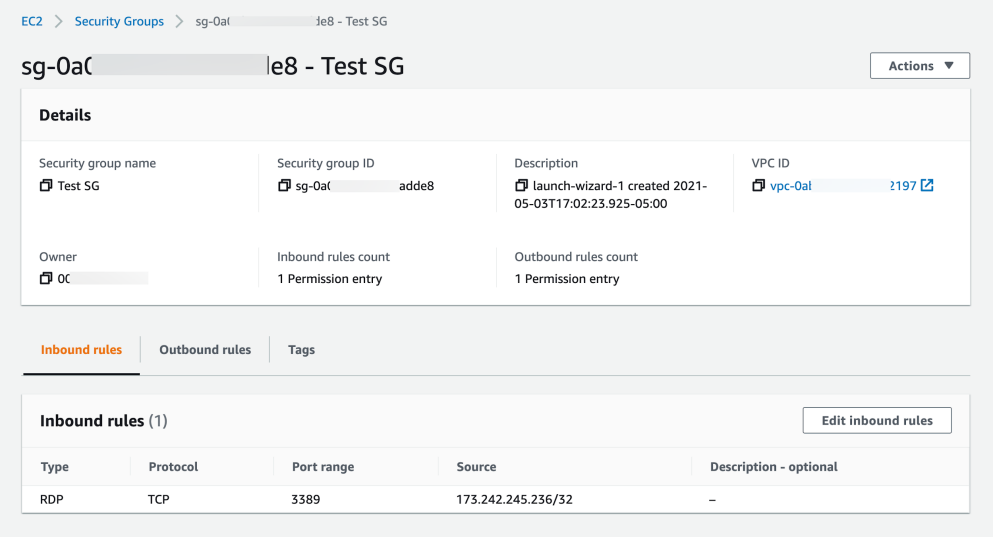

Figure 10: Security Groups

AWS EC2 access control

More refined access controls should be considered for each environment and follow the standard best practice of least privilege. When setting access controls to the environment, only allow ports, protocols, and services directly related to the actions needed to be performed. There are common mistakes that can be exploited when improper access control is allowed.

AWS S3 buckets

AWS S3 storage buckets will be leveraged for this research and the case file transport medium. The S3 storage allows for a one-way transfer of data from suspect systems into the cloud forensic processing lab. This medium allows for secure transmission and storage and a reliable mechanism to conduct repeatable and scalable actions. Key to the security of the CFTF processing lab is the integrity of the data stored within the CSP. There have been numerous cyber incidents in which the root cause was identified as misconfigured storage components of CSP. Performing due diligence to control access through the principle of least privilege is a best practice. There is an

excellent guide for AWS S3 configuration found on the Varonis blog by Jeff Petters, https://www.varonis.com/blog/how-to-use-aws-s3/.

5.1.7. Experiments

The experimentation phases of the process have been broken down into measurable segments: collect, validate, and process. These are the three phases in which working in parallel reduces the Mean Time to Analysis (MTTA).

5.1.8. Collect phase

Complete the collection by executing a pre-defined KAPE package configured to emulate a standard triage package for use during a large-scale enterprise investigation. The package is set up to gather the built-in "!SANS_Triage" compound target and create a compressed .vhdx file when complete. The output is directed to be saved on a remote system, as defined in the "—tdest variable."

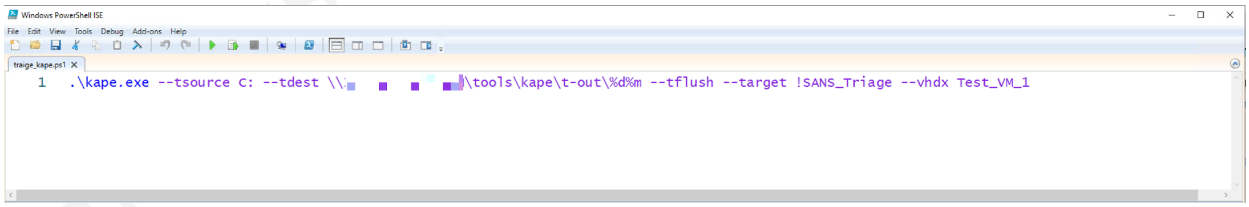

Figure 11: triage_kape.ps1 script

For the baseline test, triage_kape.ps1 was executed from the forensic lab workstation and targeted a suspect virtual machine on the network. Collection time was 83.93 seconds. The script was run against an additional ten virtual machines on the local network in serial sequence. To collect all workstations in the test pool, a total time of 1040.966 seconds is calculated. There were some outlying collection times in the entire pool that cannot be accounted for. On average, the collection times for triage_kape.ps1 was 90.65 seconds.

Collection Times			
Triage Collection - KAPE			
Runs	1	2	3
Baseline	84.0790	86.0892	81.6502
1	50.0380	521.7771	43.5433
2	48.4492	80.2298	49.0048
3	400.2894	53.3648	55.2383
4	82.3323	81.8403	82.9303
5	45.2902	59.3345	300.2238
6	84.2235	93.4421	85.0034
7	52.9336	53.0032	49.2358
8	92.0045	91.9942	87.2235
9	42.3329	294.2241	44.2373
10	57.9934	60.3343	478.2331
SUM	955.8870	1389.5444	1274.8736

Figure 12: triage_kape collection times

Modifications to the KAPE collection script are made to test the parallel collection and gather all ten virtual machines simultaneously. KAPE was configured to execute from a network share that all machines on the test network could read. The output was modified to send the data to an AWS S3 bucket for follow on validation and processing.

```
triage_kape_aws.ps1
1  \\[redacted]tools\kape\kape.exe
2  --tsource C: --tdest \[redacted]\tools\kape\t-out\%d%m --tflush
3  --target !SANS_Triage --vhdx Test_VM_1 --s3r us-east-1 --s3b CFTF --s3k ########## --s3s ##############
```

Figure 13: triage_kape_aws.ps1

This script identified some additional hurdles in an enterprise network, making sure that a clear communication path has been established to the AWS S3 bucket and PowerShell Remoting is authorized. Conducting the test run of this script on a single target showed similar times to the internal serial tests. The average collection time was slightly longer at 97.89 seconds for a single suspect VM. Targeting all 10 VM in parallel shows significant success, able to write all data on an average time of 286.12 seconds.

One of the test runs was significantly longer than the others. This anomaly is not explained and could be worth more exploration.

Full disk image collection was performed using Magnet AXIOM Cyber, and the test scenarios used in the KAPE triage process were followed. In performing the full disk image collection, it is observed that the times are much more stable, and this is believed to be due to the testing architecture of VMWare and similar disk structures across the test pool. When conducting actual collections on physical laptops over the network, this time is likely to fluctuate depending on disk use, disk health, and network bandwidth. The average collection time for the baseline and test pool was 0:13:30s. The total time to collect the test pool of 10 VM's was 2:16:20s.

During the collection phase, a discovery was made requiring a one-to-one pairing of Magnet AXIOM Cyber examiner workstations to their respective target host. This caused an issue with the attempts to automate the collection, and the process had to be adjusted to set up a pool of collection machines in AWS EC2, each one to target a single VM and run the collections in parallel.

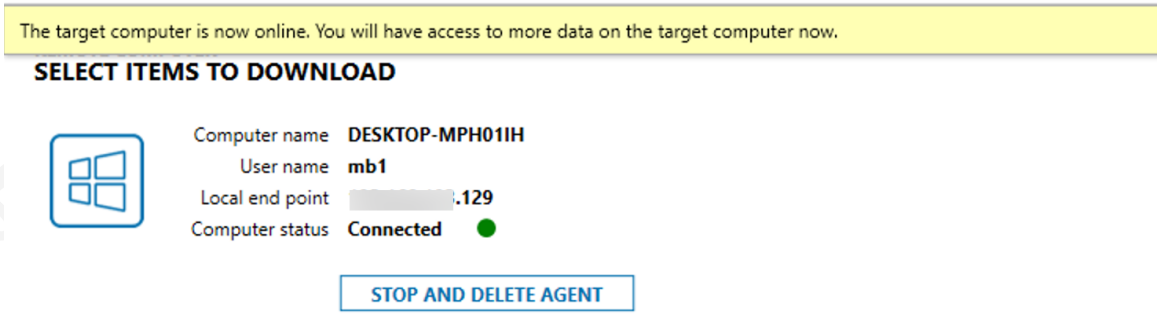

Figure 14: Magnet AXIOM Cyber - Remote collection server

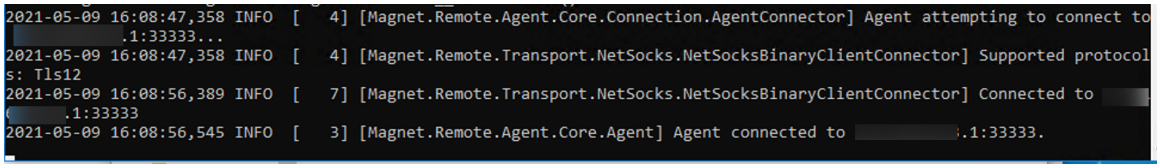

Figure 15: Magnet AXIOM Cyber - Remote collection client

To perform this action, the examiner has to execute the collection process manually on each examiner's workstation. Once executed, the suspect machines are actioned simultaneously. The first test pass shows good results, with the average collection time being 0:20:43s and the total time to collect all 10 VM's in the test pool being 0:24:46s. The delta between the average and total is due to the time taken executing the collection from each of the AWS EC2 workstations.

5.1.9. Validation Phase

Validation of collected data was expected to be improved. For the calculation of the KAPE triage packages, this did not turn out to be true. There were negligible gains in the validation phase of the collected data across the test pool. The mean time to validate the digital image files was under one second. Validation was computed using the "Get-FileHash" cmdlet included in Windows PowerShell. The time to calculate the first three .zip files was 858 Milliseconds.

```
PS M:\tools\kape\t-out> Get-ChildItem -Recurse *.zip | Get-FileHash -Algorithm SHA256 | fl


Algorithm : SHA256
Hash      : 8D06E679B02FA102109C1BEADF8AC87476610152AB2CAFB31960E46C52A236F3
Path      : M:\tools\kape\t-out\20210509T185549DESKTOP-MPH01IH\2021-05-09T185549_Test_VM_1.zip

Algorithm : SHA256
Hash      : A146F546FF29301188DA36CF4E97FF5CAD4F37E7B8FE2B3D7E9FBD72E43C1C18
Path      : M:\tools\kape\t-out\20210509T191538DESKTOP-MPH01IH\2021-05-09T191538_Test_VM_1.zip

Algorithm : SHA256
Hash      : 9AD63B2B6E12B04AF2A040B555348F88899ADB1E115B3C816B3966FF2024BDE6
Path      : M:\tools\kape\t-out\20210509T191854DESKTOP-MPH01IH\2021-05-09T191854_Test_VM_1.zip
```

Figure 16: Get-FileHash

```
PS M:\tools\kape\t-out> Measure-Command { Get-ChildItem -Recurse *.zip | Get-FileHash -Algorithm SHA256 | fl}


Days              : 0
Hours             : 0
Minutes           : 0
Seconds           : 0
Milliseconds      : 858
Ticks             : 8587465
TotalDays         : 9.93919560185185E-06
TotalHours        : 0.000238540694444444
TotalMinutes      : 0.0143124416666667
TotalSeconds      : 0.8587465
TotalMilliseconds : 858.7465
```

Figure 17: Get-FileHash time measure

Time savings were observed when calculating the physical disk image hashes. The time to calculate each hash was significant, and the ability to perform this in parallel allowed for this time to be saved. An MD5 hash is computed to validate against what is provided during the acquisition with Magnet AXIOM. To calculate the hash of all ten images on the baseline collection took 01:10:12. Compared with 00:07:02 in the AWS EC2 instance

5.1.10. Processing Phase

Processing of the KAPE triage packages is executed with KAPE in "modules mode" and targeting the previously collected .vhdx files. Two compound modules were selected for standardization of this process: "!EZ_Parser" and "MiniTimeline." The choice to use these modules was based on providing an excellent example to evaluate the collected data. Numerous modules can be leveraged and will change the processing times.

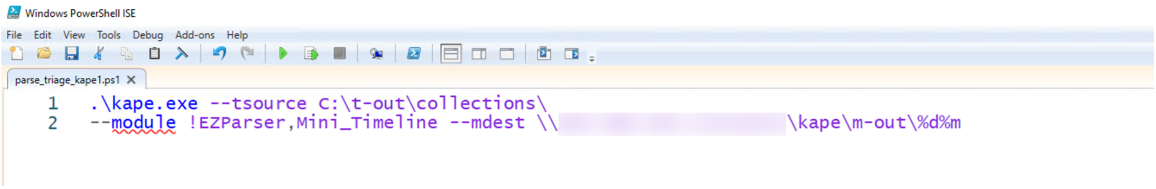

Figure 18: parse_triage_kape1.ps1

The amount of time taken to process collected .vmdk packages during the baseline sample averaged 11.50 seconds per collection and a total of 0:1:18s. Parsing within the AWS EC2 instance was consistent with the times for the baseline; the total time taken was 0:0:11s to have all ten samples processed. This shows that for KAPE processing, the time to deliver is consistent between devices, and scaling laterally will decrease the time required to process a large number of suspect assets.

Processing the full disk image with Magnet AXIOM was similar with consistent times between the baseline system and the AWS EC2 instances. On average, each case file was completed and ready for analysis in under 0:25:00s. When compiling the baseline completion time, this turned out to just over 5 hours at 05:03:35s. Within the AWS EC2 instance, the time spent on processing decreased marginally to 0:23:03 as the average time and running in parallel. The most prolonged duration of processing was 0:24:53s.

6. Findings and Discussion

Does the Cloud Forensic Triage Framework (CFTF) allow for reducing the total time spent on the collection, validation, and processing phases of investigating digital media assets? From what was discovered, yes. There is a place for leveraging this framework to gain those efficiencies in time while reducing costs. When the baseline system is reviewed to look at each phase of the process, and then the total time spent for this case, it was observed that collecting triage data, full disk image data, validating those files, and processing them took a total of 08:48:06s. This time does not include any downtime between cases, human error, breaks, computer failure, or any typical digital forensic gremlins that occur during an examination. When comparing serial processing time to the total time to perform the same function within the CFTF of 01:02:06s, the delta of over 7 hours of reduced time spent on collections.

Standard Serial Processing Top, CFTF Parallel Processing Bottom.

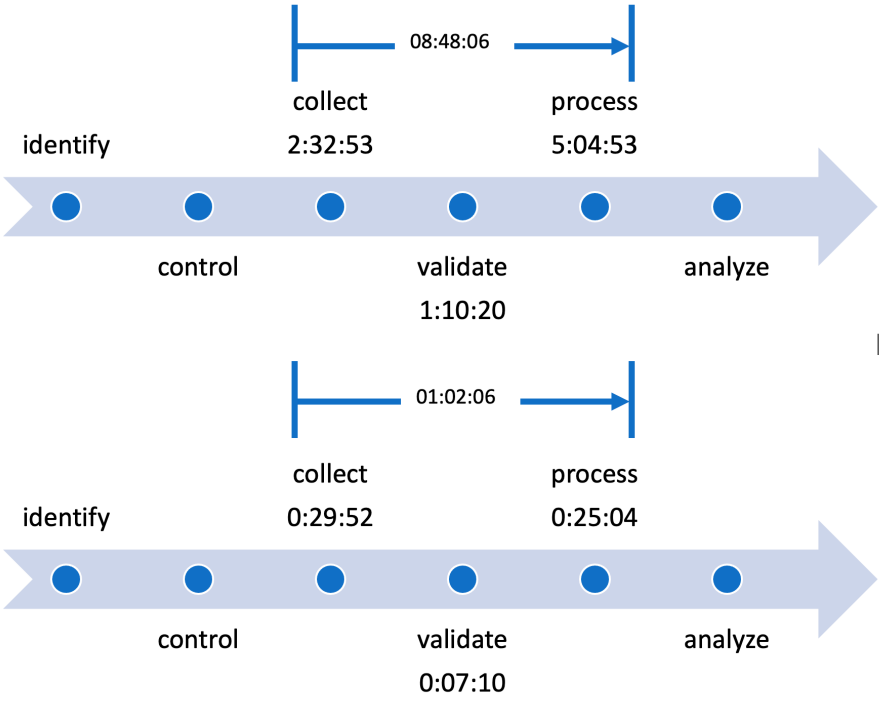

Figure 19: Time Differences

The time reduction is appreciated for the entire lifecycle of the incident response (IR) investigation. The less time taken to conduct these processes equates to a shorter mean time to analysis, allowing the investigator to spend more time conducting analysis of the data. The reduction in time taken to perform these phases impacts the rest of the IR process, the less total time to identification, containment, eradication, remediation, and lessons learned. This goes back to the principle that more time spent in the preparation of tools and techniques equates to spending less time and effort on each investigation.

6.1. Additional Findings

While conducting these tests with several assets in the AWS Cloud, these devices have a cost associated with them that has yet to be accounted for. The hypothesis was that using a CSP will reduce the collection costs, which has turned out to be true. 20 AWS instances were run for a total of 3 hours. The total time included management time, setup, installation of software, and patching. At the cost of $0.2266/hour, operating expenses totaled $13.60. This cost is an example of savings for the total operational cost

of conducting digital media investigations compared to purchasing traditional forensics hardware at $6 - $15,000 per system. It is understood that running a cloud environment must include additional best practices related to systems management, IE: turn the systems off when not in operation to save run costs.

7. Recommendations and Implications

Improving efficiencies in the digital media collection and analysis processes, which reduces resource time and human time spent conducting repeatable tasks, allows the investigator to spend more time on the items of interest. By improving this process and leveraging the CFTF opens the door to more efficient use of the investigators' primary tool, their brains, to conduct the complex analysis required. The investigator can trust that every time they need a mass collection to be performed, that collection is done in the same immutable procedure.

7.1. Recommendations for Practice

For the investigation community, the CFTF can be leveraged using currently available COTS and open-source tools. The only roadblocks to success are the limitations of the enterprise environment in which the solution is implemented. Suggestions for the community to improve are to experiment and adapt this framework to the existing procedures used by your teams currently and determine how efficiencies can be applied. The benefits to reducing the time spent on collection pay dividends on all processes which follow. Measure the time spent on collections, apply this framework, and measure the time again and determine where improvement can be made.

7.2. Future Research

Using the CFTF leads to more progress with the automation of repeatable tasks, which are time-consuming for the forensic investigation teams. More work is possible to streamline this process further and continue to reduce the time required for the forensic investigator to identify the evidence needed to prove their case, identify the guilty or innocent, and determine the root cause of a malware outbreak. Automation also leads to additional ideas of incorporating containers to do these functions; what if this whole

process was containerized in Docker and deployed directly to the suspect computer? The entire procedure could be immutable, simple to maintain, repeat, and manageable. What if the community could leverage Artificial Intelligence and Machine Learning algorithms and techniques to learn from the collected data to improve the investigative process's processing and analysis steps? Taking the information collected here and quickly comparing it to the baseline of systems in the enterprise to see what new, changed, or modified data exists and allowing the forensic investigator more time to focus on those items of interest.

8. Conclusion

The use of a Cloud Forensics Triage Framework (CFTF) has benefits for digital media investigations in the forms of resource cost: time and dollars and benefits for the investigator's work-life balance. Historically, a significant amount of time is spent on collecting and processing data. The monitoring of those tasks meant the investigator had to watch ones and zeros move, similar to watching grass grow or water boil. This time can be recaptured by moving to an updated framework to incorporate parallel collection, validation, and processing of suspect data by leveraging widely available tools and cloud resources. There are significant gains to come.

1. References

DeGrazia, M. (2019). *Triage Collection and Timeline Generation with KAPE.* Retrieved from SANS Blog: https://www.sans.org/blog/triage-collection-and-timeline-generation-with-kape/

Hartman, K. (2018). *Digital Forensic Analysis of Amazon Linux EC2 Instances.* Retrieved from SANS Reading Room: https://www.sans.org/reading-room/whitepapers/cloud/digital-forensic-analysis-amazon-linux-ec2-instances-38235

McAfee (2019). *Cloud Market Share 2019: AWS vs Azure vs Google – Who's Winning?* Retrieved from: https://www.mcafee.com/blogs/enterprise/cloud-security/microsoft-azure-closes-iaas-adoption-gap-with-amazon-aws/

Murphy, K. (2016). *Mass Triage: Retrieve Interesting Files Tool (FRAC and RIFT) Part 2.* Retrieved from SANS BLOG: https://www.sans.org/blog/mass-triage-retrieve-interesting-files-tool-frac-and-rift-part-2/

R. Montasari and R. Hill (2019). *Next-Generation Digital Forensics: Challenges and Future Paradigms.* 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3), 2019, pp. 205-212, doi: 10.1109/ICGS3.2019.8688020.

Hallman, M. (2019). *Enabling KAPE at Scale.* Retrieved from https://www.youtube.com/watch?v=YF-jDoh8BFM

Reith, M. Carr, C. Gunsch, G (2002). *An Examination of Digital Forensic Models.* International Journal of Digital Evidence, Vol.1, Issue 3.

vanBaar, R. van Beek, E. (2014). *Digital Forensics as a Service: A game changer.* Retieved from https://www.sciencedirect.com/science/article/pii/S1742287614000127

Kroll. (2020). *Kroll artifact parser and extractor – KAPE*. Retrieved from https://www.kroll.com/en/insights/publications/cyber/kroll-artifact-parser-extractor-kape

Zimmerman. (2020). *KAPE Documentation*. Retrieved from https://ericzimmerman.github.io/KapeDocs/#!index.md

Cajigas, C. (2019). *Use KAPE to collect data remotely and globally.* Retrieved from https://mashthatkey.com/2019/10/use-kape-to-collect-data-remotely-and.html

Michael Beck, mbeck.eagle@gmail.com

Magnet Forensics. (2020). *Magnet AXIOM Cyber.* Retrieved from https://www.magnetforensics.com/products/magnet-axiom-cyber/

Petters, J. (2021). *How to Use AWS S3 Securely: Setup Guide.* Retrieved from https://www.varonis.com/blog/how-to-use-aws-s3/