International Journal of Safety and Security Engineering
Vol. 14, No. 2, April, 2024, pp. 421-433
Journal homepage: http://iieta.org/journals/ijsse

# Micro Cloud Services Forensics as a Framework

Abubakr Shehata[1,2], Heba K. Aslan[3,4], Young-Im Cho[6*], Mohamed S. Abdallah[4,5,6]

[1] Information Security, Faculty of Information Technology and Computer Science, Nile University, Cairo 12677, Egypt
[2] Senior IT Manager, Gulf Agency Company (Egypt), Cairo 11771, UAE
[3] Center of Informatics Science, Faculty of Information Technology and Computer Science, Nile University, Cairo 12677, Egypt
[4] Informatics Department, Electronics Research Institute (ERI), Cairo 11843, Egypt
[5] AI Lab, DeltaX Co., Ltd., 3F, 24 Namdaemun-ro 9-gil, Jung-gu, Seoul 04522, Republic of Korea
[6] Department of Computer Engineering, Gachon University, Seongnam 13415, Republic of Korea

Corresponding Author Email: yicho@gachon.ac.kr

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/ijsse.140210

**Received:** 11 January 2024
**Revised:** 5 March 2024
**Accepted:** 12 March 2024
**Available online:** 26 April 2024

***Keywords:***
*Digital Forensic Research Workshop (DFRWS), digital forensics, events reconstruction, forensics-as-a-service, forensics investigation, microservices*

## ABSTRACT

Investigating digital crimes in cloud service environments is complex due to the decentralized nature of these services, posing challenges in data collection and presenting credible evidence in court. While existing research focuses more on external investigators, Cloud Service Providers (CSPs) have less responsibilities. To address this gap, a new framework named Microservices Forensics as a Service (MsFaaS) is introduced, aiming to ensure the reliable presentation of evidence. MsFaaS integrates international law enforcement, assigning responsibility to CSPs validated by local authorities where incidents occur. The framework consolidates existing literature, tackling unresolved challenges like legality, standardization, and data collection through the collection of diverse data types and the use of event reconstruction techniques to construct a comprehensive crime scene in both real-time and postmortem scenarios. Blockchain secures collected data against tampering, while hash functions and public key cryptography validate Microservices workflows against man-in-the-middle attacks. Machine learning enables proactive response actions to incidents. Moreover, MsFaaS facilitates auditing and recording of both internal and external cloud traffic, producing evidence reports certified by local authorities. By addressing the limitations of traditional digital forensics, MsFaaS enhances investigation reliability and effectiveness, offering services for internal CSP auditing and maintaining Chain of Custody integrity critical for trial decision-making.

## 1. INTRODUCTION

Global data centers are served by a vast array of services offered by cloud service providers. Geographically speaking, multi-tenancy and nodes are dispersed throughout various regions in several nations. Compared to traditional monolithic architecture, Software as a Service (SaaS) microservices have a heterogeneous architecture that makes them more susceptible to attacks. On the other hand, because all the features are consolidated into one location, monolithic architecture is effectively managed by firewalls and monitoring tools. This leads to centralized evidence and enhanced durability for forensic missions. However, the microservices environment exhibits a dynamic nature with particular features, such as multiple interfaces, geographically distributed components and logs, and the capacity to launch individual microservices (MS) on different platforms like virtual machines, containers, or directly on a hypervisor node [1]. These MS entities can communicate with one another via a self-contained, loosely coupled, and scalable network [2], which presents various security risks, especially for microservices [3]. The difficulties in digital forensics make it difficult for the Digital Forensics Investigator (DFI) to gather reliable evidence, which shatters the Chain of Custody (CoC) and causes the court to reject the digital evidence. The possibility of forgery and the absence of certification from an accredited body are the main causes of rejection.

Three main sectors have been identified in literature as areas where the challenges in digital forensics are being addressed. First, in June 2014, the National Institute of Standards and Technology (NIST) released a report titled NISTIR 8006 [4], which was updated in August 2020. NIST took on the task of identifying these challenges. The difficulties that prevented the Digital Forensics Investigator (DFI) from providing solid evidence in court were divided into nine categories in this report. Second, a six-phase cloud forensics methodology based on best practices was established by the Digital Forensic Research Workshop (DFRWS) [5]. Thirdly, several approaches were put forth in the literature to address the issues mentioned in the NIST report. Instead of offering a comprehensive solution that addresses every challenge, researchers concentrated on addressing particular categories of challenges. A central forensics server, external third-party interactions, and data encryption to guarantee security and integrity were among the suggested solutions. While some methods achieved data integrity by utilizing blockchain and machine learning, they were unable to maintain the Chain of Custody (CoC), an essential prerequisite for providing reliable evidence. As a result, architecture, first response, anti-forensics, analysis, training, and role management—six of the nine challenges—have been tackled. However, challenges related to data collection, legality, and standardization remain unaddressed. The reason these issues are not addressed is that researchers neglected to consider the legal implications. Governments and international organizations are required by law to safeguard and maintain data that aids in the investigation of digital crimes. For example, adherence to the General Data Protection Regulation (GDPR) of the European Union (EU) guarantees the safeguarding of personal information [6]. When any pertinent information was omitted, the Chain of Custody was unavoidably jeopardized [7]. As a result, the existing literature falls short of providing a comprehensive solution for presenting robust Electronic Evidence (EE) in court.

The suggested solution in this paper is the Microservices Forensics as a Service (MsFaaS) framework. The issues of data collection, legality, and standardization are all addressed by our framework. It presents solutions to address these issues and integrates insightful information from the body of current literature. Using a variety of data sources, including network traffic, swap files [8], VM dumps [9], containers, firewall logs, microservices activity, and registered user data, MsFaaS is intended to be deployed and managed by the cloud provider. We use event reconstruction techniques [10] to create a complete picture of the crime scene before it happens, in both post-mortem and live situations [11]. The solution flow, which includes data collection, segregation, attribute augmentation, normalization, examination, analysis, and court presentation, is in line with the DFRWS methodology. By using blockchain technology [12] and storing data in hashed format [13], we can guarantee data integrity. To add to the trust factor, a third-party server [14] housed inside a government agency is included. Hashed values are replicated to this server at the local government authority, which produces a certified report that validates the entire Chain of Custody (CoC). As such, MsFaaS functions as a security and forensic tool. The framework offers evidence with a full CoC that is admissible in court from a forensic standpoint. It also makes it easier to monitor and document internal or external threats within the cloud environment. Regarding security, MsFaaS provides a validation mechanism against Man-In-the-Middle (MITM) attacks on microservices and can also initiate an incident response in the event of hacking activity or VM malware infections. By turning distributed metadata that is intangible into concrete EE components, it can be presented in court as a certified report.

The main contributions of this paper are as follows:

- Design a framework that provides the forensic as a service for cloud microservices.
- Data collection, legality, and standardization—three previously identified and unresolved NIST challenges—are resolved by the MsFaaS framework.
- Design a framework that maintains the CoC to be able to present a robust electronic evidence in court.
- Services like a forensics-certified report, MSs security validation, and CSP auditing are provided by the MsFaaS framework.
- Machine learning techniques are integrated into the MsFaaS framework to improve analysis.
- A validation service for eCommerce is offered by the MsFaaS framework as a use case to protect against Man-in-the-Middle (MITM) attacks.

The remainder of this paper is organized as follows: Section 2 provides background information and a literature review. Section 3 presents the detailed description of our proposed framework, MsFaaS. Finally, Section 4 concludes the paper.

## 2. BACKGROUND AND LITERATURE REVIEW

Extensive scholarly research has been done on cloud forensics because of its global importance. Nevertheless, these studies have only addressed a subset of the challenges, and a comprehensive solution has yet to be achieved. To establish a well-structured framework for cloud forensics, the National Institute of Standards and Technology (NIST) has generated a report outlining the challenges in cloud forensics, arranged conferences, and developed stages to enhance forensic methodologies. We go over NIST's contributions and highlight some of the solutions found in the literature in the subsequent subsections.

### 2.1 Background

The efforts made by NIST and DFRWS to address the difficulties involved in carrying out a secure digital forensic investigation, along with the crucial steps needed to complete it, were described in this subsection.

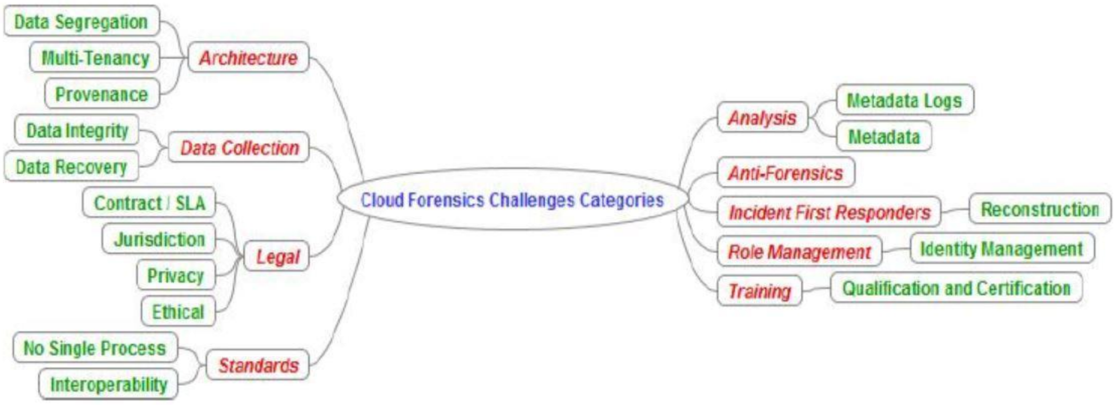

**Figure 1.** Mind map challenges categories (NIST) [4]

### 2.1.1 Challenges

The NIST report classifies cloud forensics challenges into subcategories and categories. Technical issues take precedence, followed by legality and organizational concerns. The specific technical challenges vary depending on the type of cloud computing services, such as Platform as a Service (PaaS), Infrastructure as a Service (IaaS), and Software as a Service (SaaS) [15]. Scalability and development by various teams and platforms are characteristics of the SaaS microservice architecture. The dispersion of cloud services across international borders and legal jurisdictions gives rise to organizational and legal challenges. In Figure 1, the primary categories that were taken from the NIST report [4] are shown in red text, and the subcategories are shown in green text to give specific details about each challenge.

The following sums up the challenges: The architecture's complexity stems from the use of microservices and multi-tenancy. Data collection presents challenges to investigators because it requires assembling resources that are dispersed among several nodes. They are unable to confirm whether data traffic coming from reliable microservices is authentic or if man-in-the-middle hackers could be manipulating it. Another layer of complexity is introduced by the volatility of data, since any microservice or virtual machine—along with any related logs and swap files (vRAM and vHD)—can be deleted. In postmortem situations, legal concerns about integrity, validity, and privacy prevent investigators from gathering important information. Furthermore, the regular contract or Service Level Agreement (SLA) often lacks provisions for data retention after service deletion, leading to a compromised CoC during trial proceedings. The standardization challenge is exacerbated by the proliferation of diverse forensics tools and practices. Despite the existence of ISO 27037, NIST acknowledges the absence of System Operational Procedures (SOPs), testing procedures, documented tool usage, and interoperability among Cloud Service Providers (CSPs). Correlating various file versions and formats with events reconstructed in a metadata format is a major challenge in analysis. Additionally, due to variations in time synchronization across various geographical locations, the gathered data may display a variety of timestamps.

Techniques used to thwart or deceive forensic analysis are included in the concept of anti-forensics. Techniques like malware, obfuscation, data hiding, and other strategies that jeopardize the integrity of evidence are used to avoid forensic tools. The efficacy of data collection is greatly impacted by the incident first responder, who symbolizes the CSP's obligation to act quickly to prevent data loss. The difficulty of identifying data owners and their true identities—which might not be directly related to their authenticated credentials—represents the role management challenge. Finally, the training challenge relates to the specific knowledge and abilities needed by investigators in cloud forensics, such as proficiency in mobile, network, database, IoT, microservices, and virtual machine environments. Bringing these issues under control requires assembling a group of specialists.

### 2.1.2 Digital forensics research workshop

DFRWS is an annual conference that was first held in 2001. It serves as a platform for academic and practitioner collaboration to develop the latest advancements in digital investigations, including cloud forensics. The workshop aims to establish best practices and methodologies in the field. The investigation process consists of six phases, as described below.

Phase one, Pre-Process, involves the creation of an action plan that outlines the investigation roadmap. This plan includes defining the scope of the incident, identifying the affected services, determining the nature of the damage, and assessing the available data for operational and evidentiary purposes. Phase two, Identification, focuses on detecting anomalous patterns and identifying the sources of valuable information. This phase aims to distinguish between useful and irrelevant evidence, with the selection criteria based on their relevance to the incident type or specific attacks. Phase three, Acquisition & Preservation, is concerned with ensuring the availability and integrity of evidence in live scenarios. It involves preserving the crime scene without making changes whenever possible. This can be achieved by safeguarding valuable data from being lost. For instance, it may involve storing network traffic metadata in a meaningful format within a database, collecting data from microservices, extracting dumps from virtual machines, and saving swap files throughout the incident securely. Preliminary findings based on the kind or source of the digital evidence are presented in Phase 4, Analysis. It entails putting pieces of the investigation together to figure out what happened and how the digital crime was committed. The logical analysis of every piece of evidence produced during the analysis phase is the task of phase five, evaluation. This phase involves assessing the impact of the evidence on the incident, validating evidence patterns, and attempting to recover hidden or encrypted data. The goal of phase six, presentation and post-processing, is to bring the investigation to a close. This can be accomplished by choosing to close the case owing to insufficient evidence or by presenting the evidence in a CoC form. The DFRWS conference plays a vital role in advancing digital investigations by providing a framework that encompasses these six phases.

## 2.2 Literature review

Based on the phases proposed by DFRWS, researchers have published numerous papers that have made significant progress in resolving six out of the nine challenges identified in the NIST report. We have classified the findings from the literature into four main deliverables. Firstly, the utilization of a central server for aggregating artifacts in a single location has proven effective. Secondly, adopting a forensic-as-a-service approach has shown promise in handling incidents more efficiently. Thirdly, leveraging machine learning and AI techniques has enhanced the analysis process. Lastly, encryption methods have been employed to ensure data confidentiality. These deliverables have successfully addressed challenges related to architecture, anti-forensics, role management, first response, training, and analysis. However, challenges related to data collection, standards, and legality remain unresolved. Table 1 provides a comparison between several frameworks, outlining the challenges they have successfully addressed and their corresponding subcategories. In addition to the subcategories identified by NIST, we have introduced additional unsolved subcategories that, from our perspective, contribute to maintaining the CoC and ensuring the production of trusted evidence that can be presented in court. These subcategories include: data collection (full data evidence), standardization (API gateway, network metadata, PaaS information, SaaS MS, IaaS VM), legality (chain of custody), anti-forensics (encryption and decryption), first response (postmortem, live, security), and finally, analysis (normalization and machine learning).

Numerous solutions have been presented by researchers [16-19] for utilizing a central server, effectively addressing the challenges related to architecture and role management. The central server idea is employed by the authors of a schema named "Secure Log" that they introduced by Joshi and Chillarge [20]. It is only compatible with systems that have logs available. A different suggested approach separates gathered data before keeping it in one place, which solves the role management problem by giving the CSP responsibility for the forensics mission rather than an outside investigator and enabling efficient utilization by all forensic activities [6]. Furthermore, a solution proposed by Radha Rani and Geethakumari [21] presents an algorithm that uses Deep Learning Modified Neural Network (DLMNN) classifiers and Modified Elliptic Curve Cryptography (MECC) to thwart anti-forensics attacks. This algorithm looks through transferred data to find elements that have been compromised. In addition, the CSP's involvement of highly skilled security teams to supervise the forensic server helps to address the training challenge [21]. Lastly, despite the restrictions on resource accessibility, Ali et al. [22] recommend situating the forensic server on the ISP side.

The first response challenge is efficiently managed because of the CSP's accountability and the evidence's accessibility on a central server. First response actions are triggered either when an incident occurs or when it is reported [14]. This involves securing systems once a threat is detected in real-time scenarios and providing valuable information in postmortem scenarios. The basis for developing significant evidence for forensic investigations as well as first response capabilities is event reconstruction based on timeline methodology, as covered in the study of Raju and Geethakumari [23]. The idea of keeping virtual machine snapshots is also presented by Raju and Geethakumari [24], which can be taken and kept in a safe place prior to incidents happening. However, this approach requires substantial effort and costs, as well as unlimited storage to accommodate various VM versions.

Blockchain technology has been employed to ensure data integrity, transparency, and auditability. Dasaklis et al. [12] proposed the Internet-of-Forensics (IoF) solution, which utilizes Blockchain characteristics to maintain the CoC for multiple devices/systems [25]. It should be noted that one block of evidence may be insufficient in court. Therefore, Blockchain is employed to maintain the integrity of all data, as a comprehensive representation of the crime scene.

New technologies have greatly facilitated and improved analysis results. To improve forensics prediction in an Internet of Things (IoT) setting, for instance, Koroniotis et al. [26] used deep learning (DL). Data mining methods and machine learning (ML) approaches have been employed by other researchers to discern between legitimate and malicious data [18]. Additionally, it has been possible to detect compromised authentication using Artificial Intelligence (AI) [27].

In addressing the standardization challenge, researchers have collected data from insufficient related sources. However, this has resulted in a broken CoC as only a few sources have been considered. These sources include, for instance, network metadata, API Gateways, PaaS, and IaaS through the extraction of snapshot dumps for proof. For example, in the study of Hemdan and Manjaiah [9], network metadata and virtual machine snapshots are used to gather data. To monitor CPU and RAM usage and identify CPU-Miner and HTTP-flood attacks, Sharma et al. [28] use Variational Autoencoders (VAEs) techniques. There is a gap in addressing the Standardization challenge as shown in Table 1. None of the current frameworks address the operational-level data of microservices. In the present paper, we will incorporate microservices data along with other data sources.

**Table 1.** Literature frameworks comparison

| Cat. | Data Collection | | | Standardization | | | | | Legal | | | Architect | | Anti-forensics | | Role Mng | First response | | | | Train | Analysis | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Subcategories | Data Integrity | Data Available | Full Required Data | API G. W | NT Metadate | PaaS Info | SaaS MS | IaaS VM | Data Privacy | Law Enforcement | Chain of custody | Data Provenance | Data Segregation | Encrypt Evidence | Decrypt Info | Responsible | Evident construction | Postmortem | Live | Security | Qualifications / Tools | Meta data | Normalize | ML/AI |
| [6] | ✓ | - | - | - | - | - | - | - | - | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [13] | ✓ | ✓ | - | - | ✓ | - | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [15] | ✓ | ✓ | - | ✓ | ✓ | - | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [16] | - | ✓ | - | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | = | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| [17] | - | ✓ | - | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| [18] | - | ✓ | - | ✓ | - | - | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ |
| [19] | - | ✓ | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | - | - | - | - |
| [20] | - | ✓ | - | ✓ | ✓ | - | - | - | - | - | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ |
| [21] | ✓ | ✓ | - | - | ✓ | - | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | - |
| [23] | - | ✓ | - | - | ✓ | - | - | ✓ | - | - | - | - | - | - | - | - | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [24] | ✓ | ✓ | - | ✓ | ✓ | ✓ | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | - |
| [25] | - | ✓ | - | ✓ | ✓ | - | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | ✓ | ✓ |
| [28] | ✓ | ✓ | - | - | - | - | - | - | ✓ | ✓ | - | ✓ | ✓ | - | - | - | - | - | - | - | - | - | - | - |

The legal challenge remains unresolved. If the court does not accept the CoC and use its jurisdiction during the trial, all efforts are rendered ineffective. Therefore, Ali et al. [22] discusses the application of law in cases involving digital crime, emphasizing the collaboration between technical, organizational, and legal aspects. Cloud forensics presents three primary legal challenges: data ownership (credentials and accessibility), loss of location (distributed artifacts), and the confiscation process (data acquisition). The literature suggests a technical framework that "keeps gathered data on a different forensic server that is under law enforcement's control" [22]. Additionally, in 2013 Microsoft contested a warrant issued by the US federal government to obtain data from an account located on servers in Ireland. This led to a legal battle in the USA known as "Microsoft Ireland" [29]. With the passage of the Clarifying Lawful Overseas Use of Data (CLOUD) Act in 2018, US-based technology companies such as Microsoft were able to compel federal law enforcement agencies to produce artifacts, regardless of the physical location of the data.

Although various solutions have been proposed in the literature, addressing multiple challenges, a comprehensive solution that resolves all challenges raised by NIST is still lacking. In the following section, we provide a detailed description of the proposed framework.

## 3. MICROSERVICES FORENSICS AS A SERVICE (MSFAAS) FRAMEWORK

The proposed framework consists of a segmented service operated by the internal cloud security team, addressing common scenarios of Live and Postmortem investigations. The framework is divided into two sections: Section one involves a service that can be offered to end customers through contracted monthly fees with SLAs. This service includes the following components:

- Provision of an incident forensics report certified by a government authority. This report upholds the CoC and plays a crucial role in decision-making during trials.
- Monitoring and control of end-user activities on a day-to-day basis.
- Validation service to safeguard critical Microservices against Man-in-the-Middle (MITM) attacks.

Section two encompasses a forensics tool that aids in internal CSP auditing and control. The features of this tool include:

- Facilitating the DFI mission by constructing a comprehensive crime scene.
- Monitoring and controlling breaches and threats originating from within the CSP.
- Providing an auditing tool to support ISO 27xxx and PCI compliance.

By offering these services, the MsFaaS framework aims to address the limitations of conventional digital forensics methodologies and enhance the reliability and effectiveness of investigations.

The MsFaaS investigation framework draws inspiration from the phases of DFRWS and incorporates the stages of event reconstruction. It represents a collaborative effort involving technical, organizational, and legal aspects.

As such, MsFaaS has been designed as a solution that uses organizational processes and contracting methods. It also incorporates a third-party governmental reference, mainly from the "Microsoft Ireland" case [29]. Our method focuses on law enforcement cases in which the CSP headquarters or the evidence server is under the same court's legal jurisdiction, regardless of the location of the CSP's branches (being located abroad).

Our methodology aims to consolidate existing ideas from literature into a unified framework and address unresolved challenges, including legality, standardization, and data collection. Specifically, we use the idea of a central server as described in the previous studies for gathering evidence [16-19]. Taking inspiration from a client service named LAUXUS that was proposed by Desausoi [6], we provide CSP forensics as a service. In addition, we model the system suggested in TamForen [14] by ntegrating the participation of a third party to enhance the validity of the evidence. The framework encompasses the collection of various data types from different sources, including firewall logs, network traffic, output values from Machine Learning (ML) models, streamed data from API gateways, hashed data from microservices that are sent for validation, and registered user information. Finally, the collected data is secured using blockchain technology, and critical values are hashed to meet privacy requirements.

### 3.1 MSFaaS framework

The MsFaaS framework comprises hardware, software, and a set of processes organized within a specific workflow. The hardware component consists of a server, which can be either physical or virtual, equipped with storage capacity to host two databases: a relational database and a non-relational database. The server is connected remotely to the Cloud Service Provider's (CSP) internal network, routers, switches, firewalls, and IDS/IPS devices. Additionally, the non-relational database is connected remotely to a third-party server via an IPSec Virtual Private Network (VPN) to facilitate data storage for replication purposes. The MsFaaS framework is built based on the principles of DFRWS and the Event Reconstruction methodology, as depicted in Figure 2.

The MsFaaS framework's workflow is divided into multiple phases. Using the most recent cloud threat scenarios, Phase 1 identifies all pertinent data sources associated with the incident. Planning suitable reactions for every kind of incident, including zero-day attacks, is another aspect of this phase. Data collection (Step A) is the first step in the event reconstruction process, which starts in Phase 1. Phase 2 involves the identification, filtration, and segregation of various data types according to their sources and types. To carry out the first reaction, either as a security countermeasure or to move forward with event construction, anomaly detection techniques are utilized. Data segregation is the step that is referred to as Step B. Phase 3, the core phase, is concerned with protecting the data while upholding its CoC. Using Steps C: Extra Attribute and D: Normalization, the gathered metadata is converted from its unprocessed state into one that is readable by humans. Then, blockchain technology is used to complete Step E: Data Integrity. Furthermore, the government designates a third-party location for the hashed replication of the data. Using deep learning techniques, analysis reports are produced in Phase 4, offering coherent and pertinent pieces of evidence. Phase 5 assesses the overall picture of the crime and looks at the logical relationship between all incident elements. Phase 6 concludes with the generation of a forensics report, which can be certified by comparing the report's hashed values with a counterpart that is kept on a replicated third-party server outside of the CSP. All legal requirements are met by the certified report.

The MsFaaS framework provides two main services: the Customer Forensics Service and the CSP Internal Auditing Tool. Both services follow the same workflow, but their outputs vary depending on the specific incident types. The Customer Forensics Service aims to provide customers with trustworthy evidence accompanied by valid CoC data in the event of hacking or other violations. This service involves a monthly subscription and SLA to monitor day-to-day activities and network traffic, distinguishing between malicious and benign activities. The validation service ensures the integrity of important JSON objects transferred between Microservices (MSs). For example, in online sales, JSON objects such as "Sales Order," "Payment," and "Store transfer" are stored for future validation. Hashes of these objects are generated by the sender MS and stored in the MsFaaS database. The receiver MS will only proceed if the object is validated based on the stored hashes.

The CSP Internal Auditing Tool focuses on internally monitoring all traffic within the CSP's network. It encompasses various data types involved in transactions within the CSP's local area network, including monitoring internal employees' activities using single-sign-on authentication. This service helps manage internal threats against CSP assets and helps achieve compliance with standards such as ISO 27037 or PCI. For instance, if the service detects a credit card number in sniffed network traffic, it triggers an incident for validation or potential data leakage.

All recorded data is replicated to a third-party governmental server. The hashed records in the forensic report are compared to those that were previously replicated to the third-party server to authenticate and verify the report. This way, the third party can confirm the report without needing to access the data in clear text format. For example, in Egypt, the National Telecommunications Regulatory Authority (NTRA) is required to store the necessary data for forensic purposes according to the "anti-digital data crimes" law. The need for the governmental party is to have a certified report. However, MsFaaS report can be verified by any third-party. The subsequent sub-sections provide description for the required pre-processing, reconstructing the event, and presenting reliable evidence in court.

### 3.1.1 Pre-process phase

The CSP establishes a set of requirements to guarantee that the reconstruction of events satisfies the needs for handling legal challenges, standardization, and data collection. Data integrity, customer data privacy within the CSP, data retention policies, data source types, preserved data format after normalization, and the security of transferred data to a third party are some of these challenges. The legality of multi-jurisdictional cases and multi-tenant environments is supported by these organizational and technical requirements [30].

The framework includes a set of legal guidelines that support both technical and organizational features. Organizational requirements are met by a retention policy that preserves the history of artifacts. System Operational Procedures (SOP) and the client's SLA contracts with local authorities under international law serve as the guiding principles for this policy [29]. From the technical perspective, a third-party server at the local authorities is used to verify the integrity of the evidence to prevent forgeries, and a local forensics server is implemented as an evidence store. A secure site-to-site VPN tunnel is used to safeguard the SOPs and makes it easier to move replicated hashed data from the CSP's local environment to the third-party.

The goal of the standardization process is to create a single, common SQL format from data gathered from various cloud platforms. This involves identifying metadata, segregating, normalizing data, and conducting an examination before analyzing the presented evidence. For instance, meaningful data is created from artifacts gathered from IaaS, SaaS, and PaaS platforms.

The goal of the data collection process is to locate the required artifacts. It completes "Step A-Data Collection" for the Event Reconstruction (Figure 2). When conducting live incident investigations and the crime scene has not been altered, this procedure is used. However, in postmortem situations, the evidence might have disappeared. Either the logs could be changed, or the targeted virtual machine or container could be erased. We gathered information based on a list of potential cloud attack scenarios described by Raju and Geethakumari [24] to establish an effective data retention policy. The following data types must be gathered:

- Sniffed raw metadata packets from both external and internal networks.
- User activity logs, firewall logs, and sessions that are unknown.
- User identity for cloud services, including name, location, payment information (but not credit card numbers, passwords, or secrets), hardware MAC address, used protocols, credentials, and used ports.
- A stream of events produced by the API getaway.

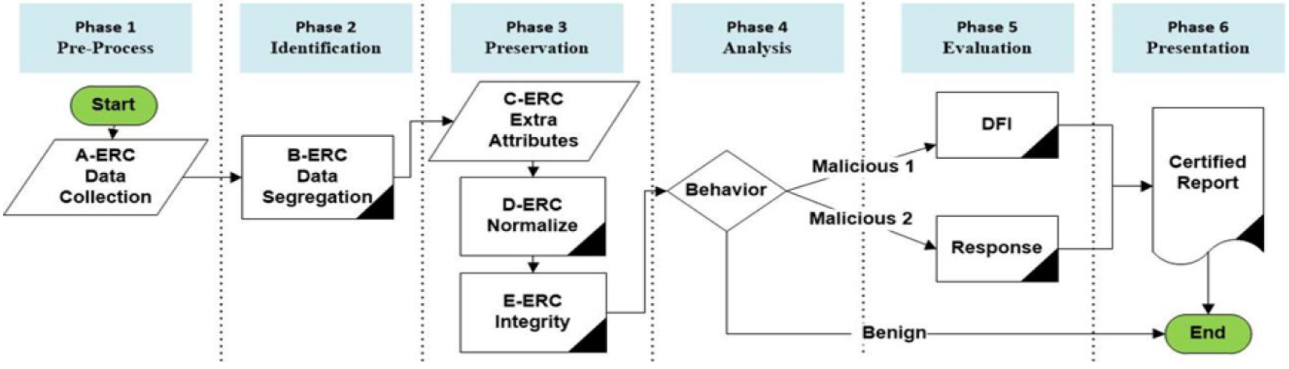

**Figure 2.** MsFaaS framework workflow

- Metadata traffic from virtual machines and dump images (like vHD, swap files, and vRAM) in case of incident handling.
- Microservices stored hashed data for a future validation period by subscribing to the MsFaaS "Validation Service."
- HTTP flow in containers at the API interfaces, RAM, CPU, and hardware performance statistics.
- Values and patterns produced by past machine learning forecasts.

By gathering all known data types from all accessible sources, we greatly reduce the difficulty of gathering data and create a robust data correlation. This guarantees that all data sources and bits of evidence from IaaS, PaaS, and SaaS platforms go toward creating a complete picture of the crime scene.

### 3.1.2 Identification phase

The main goal of this stage is to identify irrelevant and valuable data types. It corresponds to "Step B-Data Segregation" in Event Reconstruction (Figure 2). Depending on its source and how it correlates with other data, the collected data must be separated and classified during this phase. It is necessary to process and convert raw metadata into distinct records from big data streams that are obtained from network nodes by the Intrusion Detection System (IDS) or events that are taken out of API gateways. For instance, Instance_uuid and the instance's unique ID [6], distinguish data related to a particular tenant and links it to the customer's identity. Different microservices within a tenant run independently and could be dangerous. Before they can be removed or moved to another node in a different location, we use UUIDs, firewall logs, and user data to create correlations between already-existing VMs and containers. This is achieved by classifying massive metadata streams into distinct records according to predefined categories using anomaly-based or supervised machine-learning classification algorithms. Even though each set of identified logs might not tell the whole story, they are helpful for analysis and future normalization procedures.

### 3.1.3 Preservation phase

This phase is the most important since it is building the entire image while maintaining its integrity. It comprises writing thorough documentation for the procedures that support the issue of a certified report and contribute to a robust CoC. This phase uses three dedicated storage systems: C-Extra Attributes, D-Data Normalization, and E-Data Integrity. It consists of steps for event reconstruction (Figure 2). Three databases hold the data: one is a relational database housed within the CSP, another is a non-relational database that contains hashed copies of all the data within the CSP, and a third-party database that has been granted permission by the government to validate the hashed values. This is how the process goes.

Step C-extra attributes. Records are gathered as evidence based on the type of threat. For example, in a DDoS attack, additional connections are added, and multiple IP addresses contribute to the flow of HTTP packets. There is a service outage as a result of the CPU, RAM, and API interface hardware performance reaching their thresholds. Added features like data flow direction, timestamp, verified user login credentials, payment identity (not sensitive data), used tokens, and protocols are useful in these kinds of situations. These additional features help distinguish between an internal or unknown external source for the attack, as well as whether it is the consequence of malware inside the VM as a result of incorrect configuration (lack of antivirus), or a hacking attempt. Although some of these features might not be required, this procedure produces strong proof for the normalization stage.

Step D-data normalization. Using the Event Reconstruction methodology, the main goal of this step is to convert raw metadata into meaningful information [23]. Furthermore, the transformed data is organized on a timeline to construct a coherent and user-friendly narrative [23].

As stated in the study of Raju and Geethakumari [23], the purpose of this step is to link each piece of evidence to a specific timeline. This involves transforming the separated rows of raw metadata into a comprehensible human-readable format. A relational database is used to store and organize the data, displaying behavior in a comprehensible way. As a result, its evaluation is based on severity and risks. For example, the security category includes events pertaining to login and logout, whereas the user behavior categories contain information about file names, protocols, and IP addresses that are targeted or websites that fall into the categories of governmental, educational, hacking, and proxy. This facilitates an understanding of the risk associated with each behavior and allows for appropriate categorization. Thus, a unified timeline based on Greenwich Mean Time (GMT) is used to represent the gathered artifacts to procedure meaningful records in the format <TimeDescription>. GMT is used to account for time differences across the globe and reduces the risk associated with performing actions from different locations. As shown in Figure 3, the data gathered from various sources is kept in two formats. First, the CSP stores the normalized data in a relational database for use in machine learning and internal monitoring. Second, a non-relational database that has been replicated at the third-party location houses the hashed version. A few benefits of using hashed values are protection of data privacy, decreased storage volume, and the third party's ability to provide validation certificates based on the hashed values.

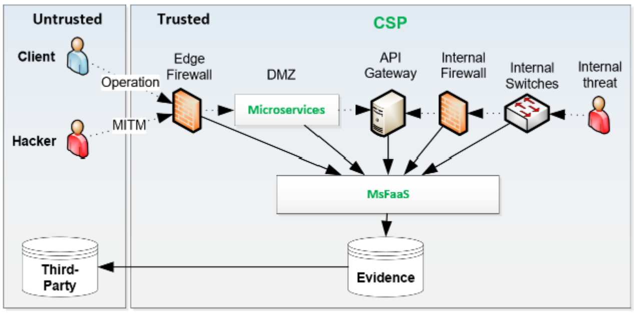

**Figure 3.** Collected data sources

Step E-data integrity. Data integrity is maintained through two mechanisms. Firstly, Blockchain is employed for data storage in the non-relational database. Additionally, hashed values are replicated on a daily basis to a database located at a local third-party. The immutability of the hashes is guaranteed by the blockchain as they are stored in blocks. Each block header holds the previous block's hash value, a timestamp, a nonce, the Merkle hash root (shown in Figure 4), and other pertinent data. As a result, the Blockchain—a chain of hash blocks—is created. Using the hash value from the previous block makes it simple to identify any small changes made to a block. The second method entails creating a site-to-site VPN tunnel over IPsec in order to safely replicate the hashed values to a third-party. A local government agency that has been designated as the third party is able to verify the copied data upon request.

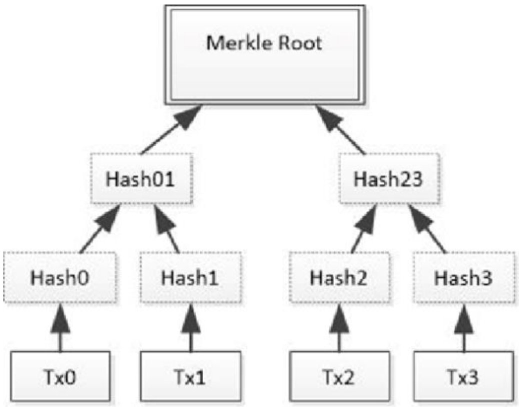

**Figure 4.** Merkle root

MsFaaS provides a further way to guarantee data integrity as a service, allowing users to submit a hashed copy of any data that needs to be validated later. For example, MsFaaS offers a validation technique in the context of microservices in an eCommerce application. The hash values are kept in a database under certain headings (like Purchase Order), which can subsequently be verified when the store is transferring the items to the customer (more information on this is provided in the use-case subsection). The coherence of the entire crime story is not ensured by the integrity of individual pieces of information alone, which could lead to a breach of the CoC and possible rejection during the trial. To address this, MsFaaS employs a strategy of replicating the entire related dataset (including SaaS, PaaS, and IaaS) to guarantee the availability of governmental authorities' validation and certification upon request. Two examples—one involving the National Telecom Regulation Authority (NTRA) and Egyptian law, and the other involving US federal law—will be discussed during the presentation phase to support this idea.

3.1.4 Analysis phase

The DFI uses the analysis phase as its first point of contact when gathering evidence. It is an essential workspace for figuring out when, what, and how the incident happens. This helps to ensure that the outputs provided during the presentation phase are accurate. Analysis is the first step in decision-making and incident response in the context of MsFaaS. To facilitate this stage, we leveraged a diverse range of data from various sources to determine the necessary actions. We used approaches from literature as well as state-of-the-art technologies, like supervised machine learning algorithms that produce accurate predictions, considering the range of data types involved. Our contribution entails an anomaly-based detection system consisting of four steps: dataset preparation, correlation feature engineering, model training preparation, and classification. These steps help distinguish between patterns that are suspicious and those that are benign. CSE-CIC-IDS-2018 dataset [31], made available by the Canadian Institute for Cybersecurity, was used in our research. This dataset was created by recording network traffic on Amazon Web Service (AWS) over a ten-day period in a controlled network environment. It includes the signatures of different types of security breaches. Excel files are used to store data that has been recorded in raw format. There are 80 features (data types) in the dataset.

In the initial phase of the dataset preparation process, we extracted data about protocols, timestamps, flow durations, ports, and packet sizes. This information encompasses multiple records, depicting both benign traffic patterns and various attack types, including DoS, DDoS, brute force, bot, and web attacks. Subsequently, the data was prepared by removing unnecessary features.

We carried out a comparative analysis in the second step, correlation feature engineering. This analysis resulted in the creation of a heatmap diagram (Figure 5) illustrating the relationships among different pieces of information.

We used a filter in the third step, which is the preparation of the model for training, to identify the 32 most pertinent features based on their lowest variance. The output of this filtering procedure included the 32 most correlated features, which create unique patterns.

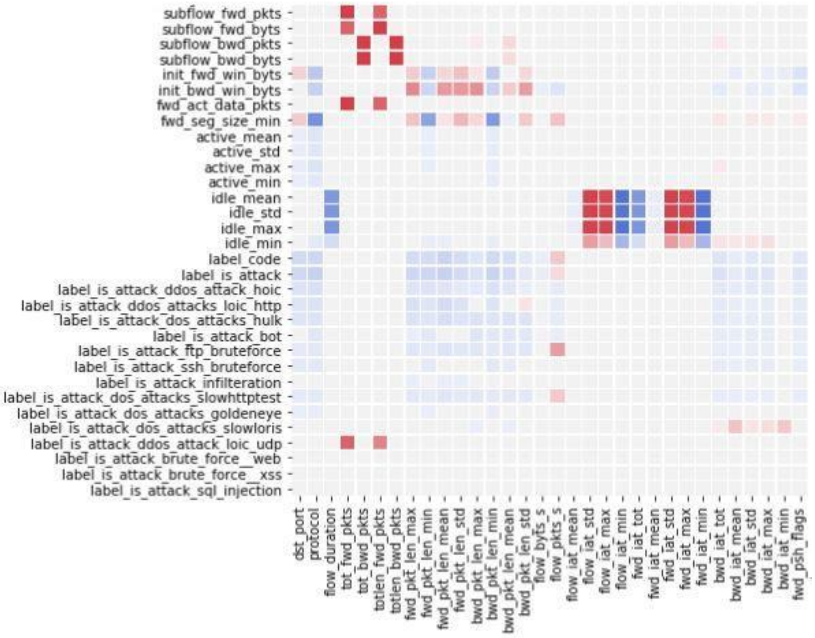

**Figure 5.** ML process 2- features correlation – heatmap

Then, the fourth process which involves the classification step. When deciding whether a traffic pattern is malicious or benign, the decision-making process depends on certain factual considerations. An incident response, for example, is triggered by attacks like DoS, DDos, Brute Force, Heartbleed, Infiltration, and Botnets. On the other hand, regular traffic is captured for potential future forensic incidents along with additional data from multiple sources.

An extra detection method built into the model is the Event Sequence Graph (ESG), which is defined in the study of De et al. [32] and records the actions of various interactive systems. The Variational AutoEncoders (VAEs) technique is used to monitor CPU and RAM usage in order to detect CPU-Miner and HTTP-flood attacks.

### 3.1.5 Evaluation phase

The main responsibilities of the DFI during this phase are upholding the CoC and guaranteeing the orderly flow of events. The investigator makes sure that all evidence is gathered in accordance with legal protocols and confirms the data's logical coherence from the analysis phase. Putting the right logs in the right order on a timeline is essential to creating a coherent story of events. For example, actions taken with the same credentials after logging into a cloud account should be examined. When data encryption occurs, the DFI can use the Forensics Tool Kit (FTK) [33] to unlock the evidence that has been encrypted. Each piece of evidence must also be linked by the investigator to the relevant legal process, whether those procedures are governed by contracts with local authorities, SLA, or international law. This mission is greatly aided by MsFaaS, which offers predefined scenarios backed by thorough evidence from a variety of data sources.

### 3.1.6 Presentation phase

Providing the investigator with a thorough analysis report that can be used as evidence in court is the main goal of the MsFaaS service. Law enforcement is therefore involved. Karagiannis and Vergidis [29] claims that "Microsoft Ireland" was forced to reveal data kept on several servers, some of which are based in the US and others of which are housed in foreign datacenters in other nations, by the US federal law known as the CLOUD Act.

In the study of Karagiannis and Vergidis [29], the authors identified five legal circumstances in which the law can be enforced when the crime and the court are in the same location of the court: Physical location of CSP headquarters, physical location of datacenter hosting the digital evidence, the end-user location, the direct consequences of the crime, and the nationality of the perpetrator or victim, regardless of the crime's location. Similarly, organizations in Egypt are required by national law to take the necessary steps to combat digital crimes. For the analysis report to be admitted as evidence in a trial, official validation is therefore necessary. Certification is obtained after the report being verified using a hashed copy which guarantees that the analysis report has not been altered. The report may, in the worst case, be used as additional evidence in the trial. This is shown in Figure 6. The investigator can present the court with a certified report by handling the privacy-related legal challenges and maintaining the CoC.

## 3.2 Use case (e-commerce)

This subsection demonstrates the utilization of a customer service that incorporates an End-User SLA to provide a validation method for any workflow. The service is specifically concerned with verifying important tasks in a workflow process. In our example, we model an eCommerce workflow and focus on the delivery and payment procedures as MSs within a SaaS application where all MSs are encapsulated in containers. The service ensures the integrity of data (referred to as purchasing 'Order') transmitted between the order MS and the delivery MS, guaranteeing that it remains uncompromised and unaltered by any means. The 'Order' is represented as a JSON object [34], containing pertinent details such as the payment amount and the items to be sent. Therefore, the SLA should make sure that the delivery items' quantities match those stated in the payment MS. The service verifies that the data has not been altered or tampered with, either internally or externally. MsFaaS keeps the data needed to reconstruct the entire crime scene in the event of a hacking incident.

Selected sales workflow comprises three MSs represented as containers with APIs: OrderMS, Delivery-MS, and MsFaaS-MS. These APIs serve as the backend components, as depicted in Figure 7, and are developed using ASP.Net Core version 6 and utilize a SQL standard database. We used 'Postman' testing tool to emulate hacker and client-side activities for the frontend simulation.

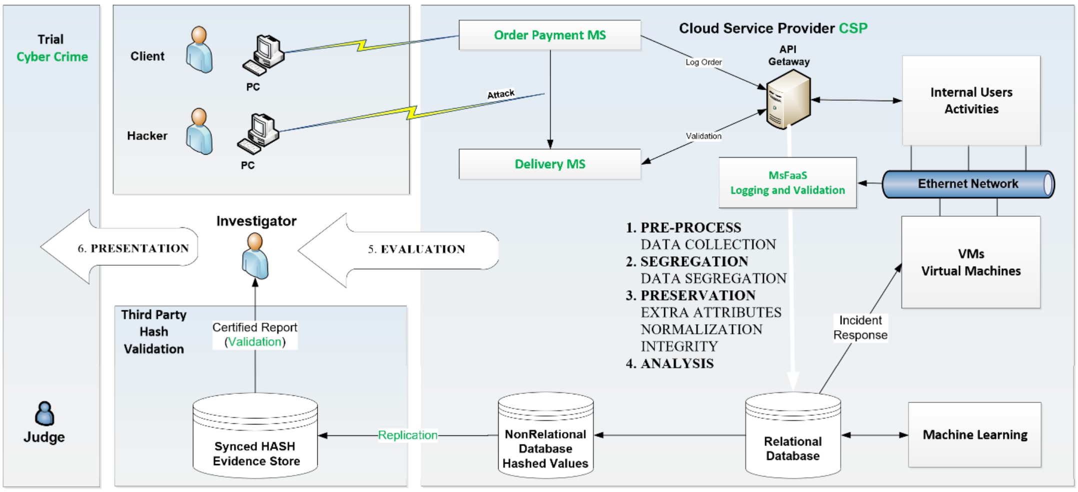

**Figure 6.** MsFaaS framework

In our test scenario, we are exposed to two types of orders: normal orders (indicated in green in Figure 7) that represent regular operations, and modified orders (indicated in red in Figure 7) that have been tampered with by a hacker as shown in Figure 7. The 'Postman' tool acts as a Man-in-the-Middle (MITM) hacker [35], intercepting traffic and modifying the order object before sending it to the Delivery-MS for processing. The goal of the modification is to increase the number of items to be sent, which does not correspond to the paid amount. The "suite prep" tool and Kali Linux are two examples of tools that could be used for interception in practical scenarios. To change the payload (JSON object) before it reaches the recipient for additional processing of the manipulated data, the attacker impersonates the MS sender's IP address, intercepts the sessions, and then sniffs the traffic.

Considering the six phases encompassing the sales workflow, the test is conducted in two scenarios, as depicted in Figure 7: The normal workflow and the workflow with the MITM attack.

### 3.2.1 Pre-process phase

In this phase, the customer subscribes to the MsFaaS validation service, and all operational data is legally collected and hashed according to the SLA. Additionally, other necessary information and metadata, such as login credentials, source IP (whether from outside or inside the CSP), and payment information (excluding credit card numbers), are collected in accordance with the CSP's SOP. The test proceeds as follows:

Step 1: The Postman tool sends an order to the Order-MS. This order includes the items to be purchased and the corresponding payment value. In our use case, we assume an order consisting of two items with a payment value of $2k, as depicted in Figure 7. The order, represented as a JSON object, will be converted into a hash value.

Step 2: The calculated hash is then sent to MsFaaS.

Step 3: Order-MS forwards the normal order (represented by the green color in Figure 7) to the Delivery-MS for processing the delivery operation.

### 3.2.2 Identification phase

In this phase, the data is distinguished based on its source and type. The hashed values are sourced from the MSs, which originate from specific workplaces and possess unique UUIDs. On the other hand, client (buyer) information is collected from registered clients. Additionally, the firewalls are logging traffic during this phase.

### 3.2.3 Preservation phase

All received data are saved in their designated locations based on their types. A non-relational database, which contains the customer-specific orders category, stores the hash value. Network traffic and client-registered information are stored in the relational database. Furthermore, all mentioned records are hashed and secured using Blockchain technology and replicated to third-party storage for future validation.

Step 4: The hashed value of the Order is stored along with references to the customer and sales client information. Additionally, a separate record is created to store all metadata related to IP sources and UUID, along with any additional data such as timestamps.

### 3.2.4 Analysis phase

This phase involves decision-making based on the utilization of ML to detect the behavior of the operation. The objective is to determine whether the operation is benign, exhibiting a correct correlation between all types of information, or if the traffic shows malicious behavior, which can be identified through anomaly detection or hash incompatibility.

Step 5: Before starting the delivery process, the Delivery-MS compares the order's hash with the records kept in MsFaaS.

### 3.2.5 Evaluation phase

Based on the results obtained during the analysis phase, the sales operation proceeds normally in case of benign traffic, while triggering an alert in case of any violation. This information serves as support for constructing a logical narrative, which can be considered as a complete crime scene during future forensic investigations by the digital forensic investigator (DFI).

Step 6: Upon successful validation of the order, it is sent to the customer. However, if the order is compromised and modified to include four items instead of two (with a total amount of $2k), the workflow will be blocked, and appropriate alerting actions will be initiated.

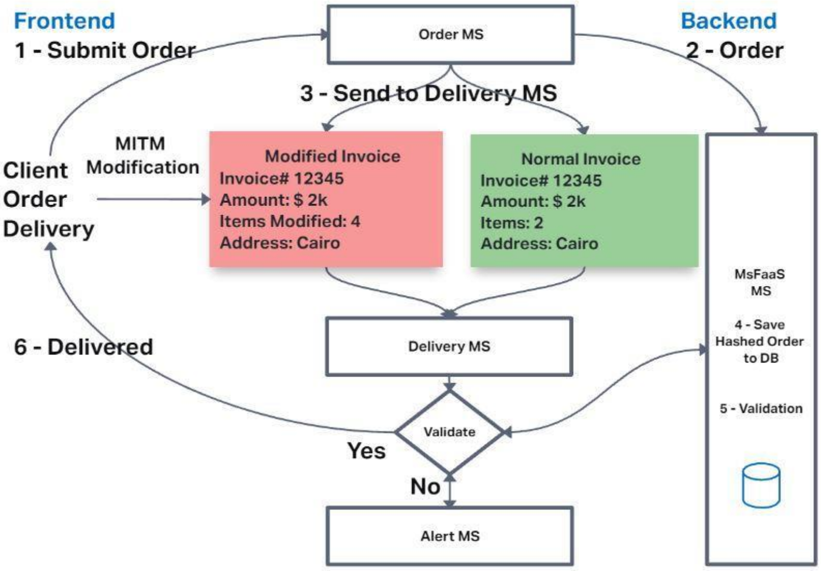

**Figure 7.** POC e-commerce validation

**Table 2.** Literature compared to MsFaaS frameworks

| Cat. | Data Collection | | | Standardization | | | | | Legal | | | Architect | | Anti-forensics | | Role Mng | First response | | | | Train | Analysis | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Subcategories | Data Integrity | Data Available | Full Required Data | API G. W | NT Metadate | PaaS Info | SaaS MS | IaaS VM | Data Privacy | Law Enforcement | Chain of custody | Data Provenance | Data Segregation | Encrypt Evidence | Decrypt Info | Responsible | Evident construction | Postmortem | Live | Security | Qualifications / Tools | Meta data | Normalize | ML/AI |
| [6] | ✓ | - | - | - | - | - | - | - | - | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [13] | ✓ | ✓ | - | - | ✓ | - | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [15] | ✓ | ✓ | - | ✓ | ✓ | - | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [16] | - | ✓ | - | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | = | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| [17] | - | ✓ | - | ✓ | ✓ | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| [18] | - | ✓ | - | ✓ | - | - | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ |
| [19] | - | ✓ | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | - | - | - | - |
| [20] | - | ✓ | - | ✓ | ✓ | - | - | - | - | - | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ |
| [21] | ✓ | ✓ | - | - | ✓ | - | - | ✓ | ✓ | - | - | ✓ | ✓ | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | - |
| [23] | - | ✓ | - | - | ✓ | - | - | ✓ | - | - | - | - | - | - | - | - | - | ✓ | ✓ | - | ✓ | ✓ | - | - |
| [24] | ✓ | ✓ | - | ✓ | ✓ | ✓ | - | - | ✓ | - | - | ✓ | - | ✓ | - | ✓ | - | ✓ | ✓ | ✓ | ✓ | ✓ | - | - |
| [25] | - | ✓ | - | ✓ | ✓ | - | - | - | - | - | - | ✓ | ✓ | - | - | ✓ | - | ✓ | ✓ | - | ✓ | ✓ | ✓ | ✓ |

3.2.6 Presentation phase

The presentation phase represents the final stage, wherein historical records are presented to the customer or valid certified evidence is provided to the court in the event of a violation.

The test successfully demonstrates one case among numerous attacking scenarios that can be executed in MSs and cloud environments. It emphasizes how useful it is to use important stored data (like in this instance, the hash value) that can be connected to different kinds of attacks or malicious activity. Consequently, this outcome implies that multiple validation applications can utilize the same service for maintaining data integrity.

## 3.3 MsFaaS impact

By providing reliable and trustworthy evidence, MsFaaS framework significantly impacts the quality of investigations. To ensure a comprehensive forensic report, the collected information in the microservice validation case must be supplemented with relevant data from all interconnected cloud environments such as firewalls and the API gateway which the container uses. To address this, we have formulated a series of inquiries that are included in the forensics report:

- Who is responsible for the incident? Is he/she a registered user or an unknown entity?
- What is the type of collected data to establish a comprehensive image? (e.g., user credentials, firewall traffic, MS hashed values, container UUID, MAC addresses, IPs, and timeline).
- How did the perpetrator carry out the incident? This entails explaining the technical methods employed and mentioning the attack type.
- When did the incident occur? This provides insight into the logical coherence of the incident.
- What is the impact of the incident? This elucidates the effects on both the criminal and the victim.
- Can the incident be replicated or forged by others? This examines the chance of data forgery and ensures evidence integrity.
- Were the evidence and data collected in a lawful manner? This addresses the legal aspects of the entire system's processes including the types of data collected to establish a comprehensive image (e.g., user credentials, firewall traffic, MS hashed values, container UUID, and timeline).

MsFaaS supports the postmortem forensic scenario and provides an effective answer to these questions. Reconstructing the crime scene while preserving its integrity is one of the most challenging phases of cloud forensics. By addressing these inquiries, the framework provides effective internal monitoring and control of activities in both the untrusted customer environment and the trusted environment of the CSP. This helps to establish a cloud environment that is secure and well-managed. Furthermore, MsFaaS effectively resolves previously uncovered NIST challenges related to legality, standardization, and data collection (as shown in Table 2). As a result, MsFaaS serves as a foundation for establishing an instant response service aimed at preventing security breaches. Table 2 shows a comparison of MsFaaS with other frameworks proposed in literature.

## 4. CONCLUSIONS AND FUTURE WORK

Currently, a full forensic service has not yet been achieved due to the inherent complexity of cloud forensics. In this context, we have presented the hypothetical framework of MsFaaS to demonstrate the viability of forensics-as-a-service as an achievable goal. The implementation of this solution as a service by the cloud service provider (CSP) yields tangible results. As a prerequisite in our framework, the CSP can establish a set of rules and regulations to address the challenges outlined by NIST. By implementing the proposed framework, MsFaaS ensures the preservation of chain of custody, providing trusted evidence that can be presented in a court of law. The MsFaaS framework offers effective technical solutions by leveraging previous research and innovative ideas. It involves the collection of diverse data types, their normalization through correlation and classification using machine learning algorithms. The framework significantly aids digital investigators in the SaaS microservices environment, particularly in postmortem investigations. Additionally, it provides a validation method to safeguard microservices against man-in-the-middle attacks. Thus, the comprehensive framework effectively addresses challenges such as Cloud Architecture, Data Collection, Standards, Training, Legal Considerations, Anti-forensics, and Incident Response. Under the sponsorship of the CSP, the proposed framework successfully addresses a set of questions posed in forensic reports. The framework offered services for internal CSP auditing, and producing reports that uphold the Chain of Custody (CoC) which play a crucial role in decision-making during trials. By offering these services, the MsFaaS framework aims to address the limitations of conventional digital forensics methodologies by resolving uncovered challenges including legality, standardization, and data collection which enhances the reliability and effectiveness of investigations.

Future work should focus on two key directions. Firstly, law enforcement agencies should mandate CSPs to implement regulations and roles utilizing such innovative solutions. Secondly, the development of multiple response actions tailored to different threats and scenarios is crucial. Attacks such as Brute Force, DoS, DDoS, Heartbleed, Web Attacks, Infiltration, and Botnets are already recognized, and corresponding best response practices exist. MsFaaS could be further developed into a comprehensive security service, incorporating emerging technologies such as deep learning and artificial intelligence.

## ACKNOWLEDGMENT

This study was funded by Korea Agency for Technology and Standards in 2022, project numbers are 1415181629(20022340, Development of International Standard Technologies based on AI Model Lightweighting Technologies) and 1415180835(20020734, Development of International Standard Technologies based on AI Learning and Inference Technologies).

## REFERENCES

[1] Bhardwaj, A., Rama Krishna, C. (2021). Virtualization in cloud computing: Moving from hypervisor to containerization—a survey. Arabian Journal for Science and Engineering, 46: 8585-8601. https://doi.org/10.1007/s13369-021-05553-3

[2] Bushong, V., Abdelfattah, A.S., Maruf, A.A., Das, D., Lehman, A., Jaroszewski, E., Coffey, M., Cerny, T., Frajtak, K., Tisnovsky, P., Bures, M. (2021). On microservice analysis and architecture evolution: A systematic mapping study. Applied Science, 11(17): 7856. https://doi.org/10.3390/app11177856

[3] Berardi1, D., Giallorenzo, S., Mauro, J., Melis, A., Montesi, F., Prandini, M. (2022). Microservice security: A systematic literature review. PeerJ Computer Science, 7(3): e779. https://doi.org/10.7717/peerj-cs.779

[4] Herman, M., Iorga, M., Salim, A., Jackson, R., Hurst, M., Leo, R., Leo, R., Lee, R., M. Landreville, N., Mishra, A., Wang, Y., Sardinas, R. (2020). NIST cloud computing forensic science challenges. U.S. Department of Commerce, National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8006.pdf.

[5] Gómez, J.M.C., Mondéjar, J.C., Gómez, J.R., Martínez, J.M. (2021). Developing an IoT forensic methodology: A concept proposal. Forensic Science International: Digital Investigation, 36: 301114. https://doi.org/10.1016/j.fsidi.2021.301114

[6] Desausoi, L. (2020). Building a secure and auditable personal cloud. Marster dissertation. Université Catholique de Louvain, Leuven, Flemish Braban, Belguim. http://hdl.handle.net/2078.1/thesis:25194.

[7] D'Anna, T., Puntarello, M., Cannella, G., Scalzo, G., Buscemi, R., Zerbo, S., Argo, A. (2023). The chain of custody in the era of modern forensics: From the classic procedures for gathering evidence to the new challenges related to digital data. Healthcare, 11(5): 634. https://doi.org/10.3390/healthcare11050634

[8] Purnaye, P., Kulkarni, V. (2022). Information retrieval for cloud forensics. In: Satapathy, S.C., Peer, P., Tang, J., Bhateja, V., Ghosh, A. (eds) Intelligent Data Engineering and Analytics. Smart Innovation, Systems and Technologies, vol 266. Springer, Singapore, 11-18. https://doi.org/10.1007/978-981-16-6624-7_2

[9] Hemdan, E.E.D., Manjaiah, D.H. (2021). An efficient digital forensic model for cybercrimes investigation in cloud computing. Multimedia Tools and Applications, 80: 14255-14282. https://doi.org/10.1007/s11042-020-10358-x

[10] Araz, J., Spannowsky, M. (2021). Combine and conquer event reconstruction with Bayesian ensemble neural networks. Journal of High Energy Physics, 2021: 296. https://doi.org/10.1007/JHEP04(2021)296

[11] Baror, S.O. Venter, H.S., Adeyemi, R. (2020). A natural human language framework for digital forensic readiness in the public cloud. Australian Journal of Forensic Sciences, 53(5): 566-591. https://doi.org/10.1080/00450618.2020.1789742

[12] Dasaklis, T., Casino, F., Patsakis, C. (2021). SoK: Blockchain solutions for forensics. In: Akhgar, B., Kavallieros, D., Sdongos, E. (eds) Technology Development for Security Practitioners, 21-40. https://doi.org/10.1007/978-3-030-69460-9_2

[13] Sachdeva, R., Gupta, S. (2021). A novel focused crawler with anti-spamming approach & fast query retrieval. In: Smys, S., Balas, V.E., Kamel, K.A., Lafata, P. (eds) Inventive Computation and Information Technologies. Lecture Notes in Networks and Systems, vol 173. Springer, Singapore, 315-331. https://doi.org/10.1007/978-981-33-4305-4_25

[14] Ye, F., Zheng, Y.Z., Fu, X., Luo, B., Du, X.J., Guizani, M. (2020). Tamforen: A tamper-proof cloud forensic framework. Transactions on Emerging Telecommunications Technologies, 33(4). https://doi.org/10.1002/ett.4178

[15] Khan, Y., Varma, S. (2020). An efficient cloud forensic approach for IaaS, SaaS and PaaS model. In 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, pp. 1-6. https://doi.org/10.1109/IDEA49133.2020.9170707

[16] Awuson-David, K., Al-Hadhrami, T., Alazab, M., Shah, N., Shalaginov, A. (2021). BCFL logging: An approach to acquire and preserve admissible digital forensics

evidence in cloud ecosystem. Future Generation Computer Systems, 122: 1-13. https://doi.org/10.1016/j.future.2021.03.001

[17] Achar, S. (2022). Cloud computing forensics. International Journal of Computer Engineering and Technology, 13(3): 1-10. https://www.doi.org/10.17605/OSF.IO/9N64K

[18] Peng, L.W., Luo, J, Li, J. (2020). Information fusion-based digital forensics framework in cloud environment. In 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp. 279-283. https://doi.org/10.1109/ICAIBD49809.2020.9137434

[19] Razaque, A., Aloqaily, M., Almiani, M., Jararweh, Y., Srivastava, G. (2021). Efficient and reliable forensics using intelligent edge computing. Future Generation Computer Systems, 118: 230-239. https://doi.org/10.1016/j.future.2021.01.012

[20] Joshi, S.N., Chillarge, G.R. (2020). Secure log scheme for cloud forensics. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, pp. 188-193. https://doi.org/10.1109/I-SMAC49090.2020.9243428

[21] Radha Rani, D., Geethakumari, G. (2020). Secure data transmission and detection of anti-forensic attacks in cloud environment using MECC and DLMNN. Computer Communications, 150: 799-810. https://doi.org/10.1016/j.comcom.2019.11.048

[22] Ali, S.A., Memon S., Dhomeja, L., Djokic D., Sahito, F. (2022). Cloud forensics framework for law enforcement agency. Journal of Southwest Jiaotong University, 57(2): 83-96. https://doi.org/10.35741/issn.0258-2724.57.2.8

[23] Raju, B.K.S.P.K., Geethakumari, G. (2017). Timeline-based cloud event reconstruction framework for virtual machine artifacts. In: Sa, P., Sahoo, M., Murugappan, M., Wu, Y., Majhi, B. (eds) Progress in Intelligent Computing Techniques: Theory, Practice, and Applications. Advances in Intelligent Systems and Computing, vol 719. Springer, Singapore, 31-42. https://doi.org/10.1007/978-981-10-3376-6_4

[24] Raju, B.K., Geethakumari, G. (2019). SNAPS: Towards building snapshot based provenance system for virtual machines in the cloud environment. Computers & Security, 86: 92-111. https://doi.org/10.1016/j.cose.2019.05.020

[25] Kumar, G., Saha, R., Lal, C., Conti, M. (2021). Internet-of-forensic (IoF): A blockchain based digital forensics framework for IoT applications. Future Generation Computer Systems, 120: 13-25. https://doi.org/10.1016/j.future.2021.02.016

[26] Koroniotis, N., Moustafa, N., Sitnikova, E. (2020). A new network forensic framework based on deep learning for internet of things networks: A particle deep framework. Future Generation Computer Systems, 110: 91-106. https://doi.org/10.1016/j.future.2020.03.042

[27] Shaikh, A.H., Meshram, B.B. (2022). Cloud attacks and defense mechanism for SaaS: A survey. In: Balas, V.E., Semwal, V.B., Khandare, A. (eds) Intelligent Computing and Networking. Lecture Notes in Networks and Systems, vol 301. Springer, Singapore, 43-52. https://doi.org/10.1007/978-981-16-4863-2_4

[28] Sharma, P., Porras, P. Cheung, S., Carpenter, J., Yegneswaran, V. (2021). Scalable microservice forensics and stability assessment using variational autoencoders. arXiv, 13193. https://doi.org/10.48550/arXiv.2104.13193

[29] Karagiannis, C., Vergidis, K. (2021). Digital evidence and cloud forensics: Contemporary legal challenges and the power of disposal. Information, 12(5): 181. https://doi.org/10.3390/info12050181

[30] Fernandes, R., Colaco, R.M., Shetty, S., Moorthy R. (2020). A new era of digital forensics in the form of cloud forensics: A review. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, pp. 422-427. https://doi.org/10.1109/ICIRCA48905.2020.9182938

[31] CSE-CIC-IDS-2018 on AWS. Canadian institute for Cybersecurity (2018). https://drive.google.com/drive/folders/1HrTPh0YRSZ4T9DLa_c47lubheKUcPl0r.

[32] De, S., Barik, M.S., Banerjee, I. (2020). A digital forensic process model for cloud computing. In 2020 IEEE Calcutta Conference (CALCON), Kolkata, India, pp. 106-110. https//doi.org/10.1109/CALCON49167.2020.9106500

[33] Bhagat, S.P., Meshram, B.B. (2021). Digital forensic tools for cloud computing environment. In: Senjyu, T., Mahalle, P.N., Perumal, T., Joshi, A. (eds) ICT with Intelligent Applications. Smart Innovation, Systems and Technologies, vol 248. Springer, Singapore, 49-57. https://doi.org/10.1007/978-981-16-4177-0_7

[34] Truica, C.O., Apostol, E.S., Darmont, J., Pedersen, T.B. (2021). The forgotten document-oriented database management systems: An overview and benchmark of native XML DODBMSes in comparison with JSON DODBMSes. Big Data Research, 25: 100205. https://doi.org/10.1016/j.bdr.2021.100205

[35] Pingle, B., Mairaj, A., Javaid, A.Y. (2018). Real-world Man-In-The-Middle (MITM) attack implementation using open-source tools for instructional use. In 2018 IEEE International Conference on Electro/Information Technology (EIT), Rochester, MI, USA, pp. 192-197. https://doi.org/10.1109/EIT.2018.8500082