Volatile / Live Data Acquisition in the Cloud: Disgruntled Employee Colludes with a Foreign APT

Benjamin Rader, Ahmed Mohammed, Adeolu Adeyemi

CIT 56200 - Mobile & Network Forensics

IUPUI

Dr. Connie Justice

Date: 4/25/2023

Contents

Executive Summary

In 2031, we were contacted by BrainArray to investigate a potential data breach which could impact business operations within their cloud systems and potentially jeopardize the business as a whole. The mission was to acquire and analyze both volatile and non-volatile data from BrainArray's Cloud infrastructure which is hosted on CloudNova. The focus was on prioritizing the acquisition of volatile data and ensuring that we work closely with relevant authorities to maintain compliance. We meticulously documented the data acquisition process, used hashing techniques, and followed strict policies regarding evidence handling and chain of custody. This allowed for a comprehensive investigation into the potential presence of malicious insiders, foreign APT's, systems flaws, and other threats in BrainArray's infrastructure.

Overview/Case

Overview

Introduction

As a digital forensics analyst working for "Forenzotica," I was contracted by BrainArray, an AI company that develops advanced AI models and tools, to investigate an incident involving suspicious activities within their cloud systems. These activities along with the large-scale degradation of their AI services and outright halting of their user applications, reveal the possibility of a data breach and intentional destruction of business operations. BrainArray's infrastructure was hosted on CloudNova, a cloud provider known for its strict service level agreements (SLAs), legal requirements, and processes related to data acquisitions.

Context and the AI Race

As of the time of acquiring data, it is the year 2031. With the ongoing AI race, many countries are competing to develop AI technology that will take over the stock market, develop the next cure for a cancer, or improve fusion reaction. This has led to cyberwar, increased espionage, and constant attempts of theft of proprietary AI models and IP (intellectual property). BrainArray, being at the forefront of the U.S's AI development, has to comply with the federal government and several agencies such as Homeland Security and the NSA. Not to mention, numerous regulatory bodies have arisen out of the federal and state governments to enforce laws around AI development and the use of personal data. The nature of the breach and the strategic significance of BrainArray's recent work in future ground-breaking technologies led the company to suspect that a foreign APT (advanced persistent threat) might be involved.

Initial Notice and Alarms

The analytics team at BrainArray was initially flooded with user complaints through their feedback system, and there was a sudden drop in model performance across the board, indicating that the AI models were producing completely unexpected or anomalous results.

The Task of Data Acquisition and Potential Threats

I was specifically tasked with acquiring both volatile and non-volatile forms of data from BrainArray's affected cloud infrastructure to facilitate further investigation into the potential presence or artifacts associated with an APT or other form of threat. These threats could be a malicious insider, critical system flaw, a malicious or colluding cloud provider (CloudNova), or even a not-so-advanced threat. This required me to consider various aspects, such as cloud provenance technology to prove data was not tampered with, legal and privacy concerns, the use of specialized tools applicable to CloudNova systems, and general knowledge of cloud services and architectures.

Case

General Cloud Components

BrainArray's cloud infrastructure consisted of various VPCs (virtual private clouds). It gives BrainArray's engineers an environment that has access control, network configurations, IP ranges, subnets, and customization and composability that isn't doable in the shared public cloud. However, VPCs are merely logical constructs, so there were still multi-tenancy issues which had been approached. Multi-tenancy is the idea that each physical server with all its hardware is virtualized into multiple virtualized machines. This means that the hardware could be shared on each "bare metal" or physical computer by multiple companies other than BrainArray itself.

VPC Controls and Security

Each VPC contained a mix of application instances, network security controls, load balancers, databases, and other storage systems. The security measures in place were relatively comprehensive too. They included consistent network segmentation, role-based access control for developers and engineers, IDS (intrusion detection systems), and MFA (multi-factor authentication) to complement access to the machines in the logical network. The AI orchestration VPC, which was the one most in question and with the most investigative importance, also included a hash-chaining-based cloud provenance mechanism that maintained a comprehensive audit trail for all data interactions. This included user IDs, connected hosts, timestamps, and many more attributes of relevance to an investigation.

Determining the Blast Radius and Defining Acquisition Targets

We could afford to have multiple investigators on-site during each acquisition. This was due to the sheer amount of money BrainArray was paying for the job. However, we were not allowed to contract any more employees due to the sensitivity of the investigation. As such, our manager organized a streamlined approach for what would be investigated first and what would be the order for acquisition with devices. Senior management of Forenzotica coordinated with CloudNova and BrainArray to define the scope of the acquisition phase.

Acquisition Procedure

The data acquisition process was extensive and involved imaging various infrastructure components. Namely, there were two VPCs which were thoroughly imaged and investigated with priority going towards the AI orchestration network. "Prioritization is consistently identified as a core tenet of security incident handling in numerous studies. A lack of

prioritization can result in security data fatigue, analyst burnout, and ineffective or insufficient incident response. (McRee, 2022)" Forenzotica supervisors made sure that there was an understood and pragmatic approach to prioritizing certain systems for evidence collection and data acquisition. The general priority for hardware went as such: CPU cache, RAM (random access memory), Swap space, system and user configurations (registry keys or environment variables – if not already logged via another application), then live data acquisition for things like network connections, running processes, etc. Static data acquisition such as hard-disk, SSD, and forms of data that can survive without power were the last on the list.

Volatile forms of data such as with the CPU cache and RAM were of the utmost priority. As the name suggests, volatile data can quickly be overwritten or lost. Additionally, the virtualized nature of CloudNova instances could prove to be problematic if the virtualized instance changes its hardware parameters all of the sudden or another tenant takes up RAM that has evidence on it. APTs are more likely to employ evasion methods, "hook" into running processes, or use binaries on the computer that are trusted ("LoLBins"). Therefore, imaging RAM was crucial for any device that could've been utilized for command and control (C2).

Specialized tools were used to image hardware including Volatility for RAM and EnCase for capturing disk images. These were rapidly applied to as many machines as possible and each case of acquisition required that two forensic analysts be present and one representative from CloudNova, along with authorities, which will be discussed in the next section.

It is important to note that some cloud systems do have methods for acquiring the data remotely. One example includes "watermarking" the data and using attribute-based encryption to ensure integrity during transmission (Liu et al., 2019). However, most cloud providers seem to

have limited access and complicated service level agreements that make this hard to pull off without an existing native solution from the provider.

One more issue that slowed down the acquisition process dramatically was the inherent issue of obtaining logs and images where elevated permissions were necessary. Often, this involved numerous phone calls and basic communication bottlenecks that most organizations run into on a daily basis (Prasad Purnaye & Kulkarni, 2022).

There were cases where it was possible to export virtual RAM images via copy and paste approaches. These took advantage of bi-directional copy and paste features with remote servers. This data was then fed into files running in some of our secured storage applications (Lutui & Cusack, 2021).

Privacy Regulations and Jurisdictional Requirements

Multi-tenancy posed many challenges, as CloudNova's SLAs restricted the acquisition of volatile data to avoid affecting other customers sharing the hardware. Privacy regulations and legal concerns also presented themselves with the possibility that a full RAM image would expose other cloud user's data (other companies using CloudNova). Close collaboration and negotiation took place between BrainArray, Forenzotica, CloudNova, and the related authorities and agencies. Each team reviewed the SLAs and relevant jurisdictional legal requirements to determine the extent of CloudNova's obligations and the permissible methods of data acquisition. Luckily, the cloud provenance controls in place allowed for some leeway because the data in that system could be trusted and tamper-proof due to the hash-chaining technology involved (Zawoad et al., 2018). Obtaining the RAM images from multi-tenant devices required

coordination with law enforcement and a present representative from the other affected organization.

However, one of the data centers was in California, which meant compliance with California Consumer Privacy Act (CCPA) regulations regarding user data. We had to prove that user data had been anonymized before gaining access to certain data. This was an elementary step, and our privacy team quickly gave our forensics team relevant documentation of the systems involved. Additionally, the Health Insurance Portability and Accountability Act (HIPAA) became relevant with some of our medical applications. To comply with these sorts of regulations, I coordinated with legal teams of both BrainArray and CloudNova to obtain a court order granting access to systems that housed sensitive personally identifiable information (PII) and systems with personal health information (PHI).

Lastly, there were those federal agencies such as Homeland Security and the NSA which followed strict processes when doing forensic imaging of some of the related devices. Reps from these organizations were consistently present during any briefings or meeting where data was being acquired from AI systems that were applicable to military and defense technology.

Secure Acquisition and Storage Prior to Analysis, Chain of Custody

To maintain a proper chain of custody and ensure data integrity, I documented nearly every piece of the data acquisition process, including the use of specialized tools, dates, and times of data collection, and any individuals involved. Our policy in high-risk investigations such as this one meant that no one could be trusted: the cloud provider, investigators, BrainArray, and even law enforcement. Documentation included screenshots and pictures during

the process, lists of tools, dates and times of collection, and all involved individuals with signatures.

As mentioned before, there were two investigators for each instance of acquisition. Anything that could be hashed was hashed during every step. For instance, taking an image of RAM meant that two investigators would image the device separately and store the images onto WORM (write once, read many) devices, so that they could not be tampered with. WORM devices are just one example of the few recent secure logging solutions for mitigating anti-forensics activities (Paccagnella, 2019). Hashes were put into a read-only database application and also written down onto a sheet that would be stored in secure storage in our transportation vehicle. Law enforcement, CloudNova, and affected individuals were also present during the data acquisition process. We combined all of this with our current policies and procedures around transporting the evidence onto our site and forensics lab. Every piece of evidence was also separately tagged to make it easy to obtain and organize into our forensics systems once back at the lab.

By combining technical expertise, legal knowledge, and specialized tools, I was able to navigate the challenges presented by CloudNova's SLAs, privacy regulations, and while utilizing proprietary cloud provenance technology to successfully acquire the volatile and non-volatile data necessary for the investigation.

Objectives

Primary objectives

Our primary objective of this investigation was to determine the reason behind a series of suspicious activities that were causing the degradation and destruction of infrastructure and

services. We are to investigate who was responsible for the suspicious activities, including when it all started and how long it took to discover these malicious activities. These activities included unusual data corruption, where BrainArray observed an unexpected and widespread denial of service of their AI models, which could not be attributed to normal system or maintenance errors. The company discovered that numerous databases, including recent AI and user-focused models, had been deleted. There were also indications of unauthorized access, fake user accounts, network anomalies, and malware, suggesting possible links to Advanced Persistent Threats (APTs). BrainArray detected unusual internal accounts with legitimate access to the systems, but the amount of access presented too much power over critical operation. This was true for various groups of users and the number of users is not small.

Anticipating Anti-Forensics

"People use anti-forensics to demonstrate how vulnerable and unreliable computer data can be" (Karie & Venter, 2015). This is especially true with advanced, potentially state-sponsored threats. Evidence analysis and conclusions should be made with the assumption that a threat actor attempted to destroy or corrupt the data. This is why encryption-based tamper evident methods are crucial such as with the proprietary cloud provenance system utilized by CloudNova. Unfortunately, only high-value data was logged into the secure logging system. In other words, it was not possible to continuously image RAM into cloud provenance storage, so RAM images aren't stored in these tamper-proof systems. We will not always be able to ensure that RAM had not be tampered with before the imaging took place. For all we know, the cloud providers themselves may have been involved. No data should be trusted unless it can be cross-referenced or backed up by provable tamper-proof mechanisms.

Cloud provenance System

CloudNova's cloud provenance system played a crucial role in the investigation, this system stores logs of server activities and uses hash chaining technology. Hash chaining is a technique often used in blockchain technology; it works by linking each log entry to its predecessor through cryptographic hash function (Imran & Hlavacs, 2015). This approach method made it easier to trust the logs during the investigation and ensured that the data has not been manipulated since it was initially recorded. This included user ID's, connected hosts, and timestamps that were relevant to the investigation. By utilizing the technology, our team was able to successfully acquire the necessary data and maintain its integrity throughout the entire process.

Hypothesis

With an investigation such as this one, the possibilities are wide-ranging for why numerous systems simultaneously failed. Even worse, if the threat was not an entity, but rather a system owned by BrainArray, then the expensive investigation will have mostly been a waste of money. This is all complicated by the numerous issues that come with the combination of a potential APT and data acquisition in the cloud: 1) if the threat is an APT, then there is a high probability for anti-forensics techniques, 2) volatile data is limited, time-sensitive, and most APTs utilize process-level exploits and zero-days that are difficult to understand or notice without imaged RAM, 3) the potential of CloudNova colluding with an APT makes the ramifications much worse with the possibility of direct hardware manipulation, and 4) the web of legal hurdles, SLAs, privacy legislation, and government agency involvement have made the investigation one that no one is prepared for. From our vantage point as forensics firm and given the current political and cultural climate with the AI race, it seems highly probable that the threat

actors are indeed advanced persistent threats that were looking to put the United States behind in the AI race.

Suspects and points of interest.

The primary suspects include people, inputs into the AI orchestration network (technically a VPC), and any systems that could potentially have unsanitized or unsafe inputs into the affected systems.

In terms of people, our curated lists and databases include any users with direct and indirect access. Executives and management, and those who don't have access were the information that they needed to know and nothing more. Additionally, contracts were signed to ensure that management would not let suspects know that the investigation was taking place. Some examples of suspects include AI orchestration managers, AI analysts, system administrators, cloud provider employees, developers for the AI orchestration applications, and database managers and analysts. Additionally, any employees in the same environment as the suspects were also to be investigated to a certain degree.

Search Warrants

Due to the high-risk nature of the situation and the potential of military reactions, the involved government agencies are putting many resources into getting quick warrants and access into other cloud systems which we would normally not have access to. Granted, they must still follow certain laws and regulations. However, the amount of government leaders and agencies on top of the investigation and the amount of money being thrown at the problem will make it easy to obtain warrants for virtually any of the companies involved. Not to mention, most of them are obtainable via. agreements with the National Security Agency.

Evidence

Evidence Overview

Key pieces of evidence obtained during the data acquisition process included volatile which includes, CPU cache and RAM images from the cloud infrastructure. Non-volatile data which included swap space, system and user configurations, and other forms of data. Other data included network connections, running processes, and user interactions which included complaints and feedback. Along with security logs, and records VPC's and network security controls, and documentation of the overall process.

User Access Surface for Affected Systems

Sorted by the degree of access or risk to the systems:

High Risk

1. Elevated Access to Servers
    1.1. Sys Admin - have access to change server and infrastructure configurations, credentials to access most servers directly, etc.
        1.1.1. James Carter
        1.1.2. Lisa Martin
        1.1.3. Trey Valentine
        1.1.4. Maurice Smith
2. AI Orchestration Manager - access to change AI orchestration systems which include: model deployment, model transformation, model data access, user data storage systems, etc.
    2.1. Michael Thompson
    2.2. Rebecca Johnson

3. Network Security Engineers - access to network configurations of DMZ and Cloud Nova VPCs.
    3.1. Laura Brown
    3.2. David Garcia

Medium Risk

1. Application Developers
    1.1. Susan Anderson
    1.2. Robert Lee
    1.3. Mary Martinez
2. Database Administrators
    2.1. Karen Harris
    2.2. Paul Clark
3. CloudNova Maintenance and Engineers - direct access to hardware, control over tenancy on hardware, and control over log storage.
    3.1. Traceable down to the day. CloudNova has camera footage and access logs that can verify employees or third parties which have accessed the facility.

Low Risk

1. General Employees with Cloud Access – limited access to change the data, configurations, or input data. These employees have enough access to read system configurations and utilize the intentional functionality.
    1.1. Emily Lewis
    1.2. Daniel Walker

1.3. Angela Young

1.4. Joshua Hall

List of Digital Evidence

An exhaustive nested list of the evidence procured during the acquisition phase of the investigation:

- 1
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Primary AI Orchestration Servers
  - **Make & Model**: CloudNova-CN-AIOrch-001
  - **Serial Number**: CN-OA-001-123
  - **Acquisition Date and Time**: 8/17/2031 10:15
  - **Condition**: Good
  - **Hash Value**: 8a5b423a5c6c5d5e5f6c7g8h9j1k2l3m4n5o6p7q8r
  - **Original Custodians**: Trey Valentine, Maurice Smith
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: RAM image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 2
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Primary AI Orchestration Servers
  - **Make & Model**: CloudNova-CN-AIOrch-001
  - **Serial Number**: CN-OA-001-123
  - **Acquisition Date and Time**: 8/17/2031 10:15
  - **Condition**: Good
  - **Hash Value**: 8a5b423a5c6c5d5e5f6c7g8h9j1k2l3m4n5o6p7q8r
  - **Original Custodians**: Trey Valentine, Maurice Smith
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Storage Image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 3
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Primary AI Orchestration Servers
  - **Make & Model**: CloudNova-CN-AIOrch-002
  - **Serial Number**: CN-OA-002-456
  - **Acquisition Date and Time**: 8/17/2031 10:45

    - **Condition**: Good
    - **Hash Value**: 9b6c5d4e3f2g1h0i1j2k3l4m5n6o7p8q9r0s1t2u3v
    - **Original Custodians**: Lisa Martin
    - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
    - **Components**: RAM image
    - **Chain of Custody**: Forenzotica (2 analysts), CloudNova Rep, Law Enforcement, Federal Agency Rep
- 4
    - **Evidence Type**: Image (Drive)
    - **Device Purpose**: Primary AI Orchestration Servers
    - **Make & Model**: CloudNova-CN-AIOrch-002
    - **Serial Number**: CN-OA-002-456
    - **Acquisition Date and Time**: 8/17/2031 10:45
    - **Condition**: Good
    - **Hash Value**: 9b6c5d4e3f2g1h0i1j2k3l4m5n6o7p8q9r0s1t2u3v
    - **Original Custodians**: Lisa Martin
    - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
    - **Components**: Storage Image
    - **Chain of Custody**: Forenzotica (2 analysts), CloudNova Rep, Law Enforcement, Federal Agency Rep
- 5
    - **Evidence Type**: Live Acquisition Data
    - **Device Purpose**: Primary AI Orchestration Servers
    - **Make & Model**: CloudNova-CN-AIOrch-001
    - **Serial Number**: CN-OA-001-123
    - **Acquisition Date and Time**: 8/17/2031 10:15
    - **Condition**: Good
    - **Hash Value**: 8a5b423a5c6c5d5e5f6c7g8h9j1k2l3m4n5o6p7q8r
    - **Original Custodians**: Trey Valentine, Maurice Smith
    - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
    - **Components**: Running Processes, Network Connections, Registry Keys, File System Metadata
    - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 6
    - **Evidence Type**: Live Acquisition Data
    - **Device Purpose**: Primary AI Orchestration Servers
    - **Make & Model**: CloudNova-CN-AIOrch-002
    - **Serial Number**: CN-OA-002-456
    - **Acquisition Date and Time**: 8/17/2031 10:45

- **Condition**: Good
- **Hash Value**: 9b6c5d4e3f2g1h0i1j2k3l4m5n6o7p8q9r0s1t2u3v
- **Original Custodians**: Lisa Martin
- **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
- **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
- **Components**: Running Processes, Network Connections, Registry Keys, File System Metadata
- **Chain of Custody**: Forenzotica (2 analysts), CloudNova Rep, Law Enforcement, Federal Agency Rep

- 7
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Quantum Application Server
  - **Make & Model**: CloudNova-QCN-AppVPC-001
  - **Serial Number**: QCN-AV-001-789
  - **Acquisition Date and Time**: 8/17/2031 11:30
  - **Condition**: Good
  - **Hash Value**: 0c1d2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v
  - **Original Custodians**: James Carter
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: RAM image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 8
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Quantum Application Server
  - **Make & Model**: CloudNova-QCN-AppVPC-001
  - **Serial Number**: QCN-AV-001-789
  - **Acquisition Date and Time**: 8/17/2031 11:30
  - **Condition**: Good
  - **Hash Value**: 0c1d2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v
  - **Original Custodians**: James Carter
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Storage Image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 9
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Quantum Application Server
  - **Make & Model**: CloudNova-QCN-AppVPC-002
  - **Serial Number**: QCN-AV-002-012
  - **Acquisition Date and Time**: 8/17/2031 12:05
  - **Condition**: Good

- **Hash Value**: 1d2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w
- **Original Custodians**: Michael Thompson
- **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
- **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
- **Components**: RAM image
- **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep

- 10
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Quantum Application Server
  - **Make & Model**: CloudNova-QCN-AppVPC-002
  - **Serial Number**: QCN-AV-002-012
  - **Acquisition Date and Time**: 8/17/2031 12:05
  - **Condition**: Good
  - **Hash Value**: 1d2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w
  - **Original Custodians**: Michael Thompson
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Storage Image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep

- 11
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Holographic Data Storage (Server)
  - **Make & Model**: HoloStor-HS5000X-12TB
  - **Serial Number**: HS-5000X-345
  - **Acquisition Date and Time**: 8/17/2031 13:15
  - **Condition**: Good
  - **Hash Value**: 2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x
  - **Original Custodians**: Laura Brown
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: RAM image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep

- 12
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Holographic Data Storage (Server)
  - **Make & Model**: HoloStor-HS5000X-12TB
  - **Serial Number**: HS-5000X-345
  - **Acquisition Date and Time**: 8/17/2031 13:15
  - **Condition**: Good
  - **Hash Value**: 2e3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x

- **Original Custodians**: Laura Brown
- **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
- **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
- **Components**: Storage Image
- **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep

- 13
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Holographic Data Storage (Server)
  - **Make & Model**: QuantumDrive-QD6000Z-10TB
  - **Serial Number**: QD-6000Z-678
  - **Acquisition Date and Time**: 8/17/2031 13:55
  - **Condition**: Good
  - **Hash Value**: 3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y
  - **Original Custodians**: Paul Clark
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: RAM image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 14
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Holographic Data Storage (Server)
  - **Make & Model**: QuantumDrive-QD6000Z-10TB
  - **Serial Number**: QD-6000Z-678
  - **Acquisition Date and Time**: 8/17/2031 13:55
  - **Condition**: Good
  - **Hash Value**: 3f4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y
  - **Original Custodians**: Paul Clark
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Storage Image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 15
  - **Evidence Type**: Image (RAM)
  - **Device Purpose**: Quantum Database Storage
  - **Make & Model**: Oracle QuantumX10
  - **Serial Number**: OQX10-901-234
  - **Acquisition Date and Time**: 8/17/2031 15:10
  - **Condition**: Good
  - **Hash Value**: 4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z

- **Original Custodians**: Karen Harris
- **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
- **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
- **Components**: RAM image
- **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep

- 16
  - **Evidence Type**: Image (Drive)
  - **Device Purpose**: Quantum Database Storage
  - **Make & Model**: Oracle QuantumX10
  - **Serial Number**: OQX10-901-234
  - **Acquisition Date and Time**: 8/17/2031 15:10
  - **Condition**: Good
  - **Hash Value**: 4g5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z
  - **Original Custodians**: Karen Harris
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Storage Image
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 17
  - **Evidence Type**: Router Logs and Hardware Image
  - **Device Purpose**: Photon Router
  - **Make & Model**: Cisco LightSpeed 3000
  - **Serial Number**: LS-3000-567-890
  - **Acquisition Date and Time**: 8/17/2031 15:45
  - **Condition**: Good
  - **Hash Value**: 5h6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a
  - **Original Custodians**: David Garcia
  - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
  - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
  - **Components**: Configuration History and Image of the Memory
  - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 18
  - **Evidence Type**: Firewall Logs and Hardware Image
  - **Device Purpose**: AI-Powered Firewall
  - **Make & Model**: Fortinet FortiGuardian 10G
  - **Serial Number**: FG10G-456-789
  - **Acquisition Date and Time**: 8/17/2031 16:30
  - **Condition**: Good
  - **Hash Value**: 6i7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a5b
  - **Original Custodians**: Rebecca Johnson

    - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony or witnesses from CloudNova
    - **Components**: Configuration History and Image of the Memory
    - **Chain of Custody**: Forenzotica (2 analysts), Law Enforcement, CloudNova Rep
- 19
    - **Evidence Type**: CloudNova VPC Logs (with provenance mechanism)
    - **Device Purpose**: Tamper Proof Log Storage for BrainArray's Complete VPC
    - **Make & Model**: Proprietary CloudNova Distributed Hash-Chaining Based Storage
    - **Serial Number**: N/A
    - **Acquisition Date and Time**: 8/17/2031 14:10
    - **Condition**: Good
    - **Hash Value**: 4g56i7j8k9l0m1i4f0fksda09fgr6s7t8u9v0w1x2y3z
    - **Original Custodians**: N/A
    - **Suspect or Employee Type**: BrainArray Employee, CloudNova Employee
    - **Strengths / Weaknesses / Caveats**: High admissibility in court due to provable tamper-proof nature of logs
    - **Components**: IAM logs, Network logs, Configuration monitoring and changes, security logs, and DNS for endpoints, etc. This is a distributed tamper-proof storage system that stores all of the CloudNova Logs in one place.
    - **Chain of Custody**: Forenzotica (2 analysts), CloudNova Rep
- 20
    - **Evidence Type**: Mobile Device (iOS)
    - **Device Purpose**: Mobile Phone
    - **Make & Model**: Apple iPhone 20 Pro
    - **Serial Number**: HJ1KD2LF3MG
    - **Acquisition Date and Time**: 8/22/2031 10:15
    - **Condition**: Good
    - **Hash Value**: 7j8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a5b6c
    - **Original Custodians**: Susan Anderson
    - **Suspect or Employee Type**: BrainArray Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony
    - **Components**: Everything in the Mobile Phone, RAM image, Storage Image
    - **Chain of Custody**: Forenzotica (2 analysts)
- 21
    - **Evidence Type**: Mobile Device (iOS)
    - **Device Purpose**: Mobile Phone
    - **Make & Model**: Apple iPhone 20
    - **Serial Number**: JH2LE3MG4NI
    - **Acquisition Date and Time**: 8/22/2031 14:15
    - **Condition**: Good

    - **Hash Value**: 8k9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a5b6c7d
    - **Original Custodians**: Robert Lee
    - **Suspect or Employee Type**: BrainArray Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony
    - **Components**: Everything in the Mobile Phone, RAM image, Storage Image
    - **Chain of Custody**: Forenzotica (2 analysts)
- 22
    - **Evidence Type**: Mobile Device (Android)
    - **Device Purpose**: Mobile Phone
    - **Make & Model**: Samsung Galaxy S35
    - **Serial Number**: WQ9R8S7T6U5
    - **Acquisition Date and Time**: 8/22/2031 15:30
    - **Condition**: Good
    - **Hash Value**: 9l0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a5b6c7d8e
    - **Original Custodians**: Trey Valentine
    - **Suspect or Employee Type**: BrainArray Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony
    - **Components**: Everything in the Mobile Phone, RAM image, Storage Image
    - **Chain of Custody**: Forenzotica (2 analysts)
- 23
    - **Evidence Type**: Mobile Device (Android)
    - **Device Purpose**: Mobile Phone
    - **Make & Model**: Google Pixel 12
    - **Serial Number**: XQ8S7T6U5V4R
    - **Acquisition Date and Time**: 8/22/2031 16:45
    - **Condition**: Good
    - **Hash Value**: 0m1n2o3p4q5r6s7t8u9v0w1x2y3z4a5b6c7d8e9f
    - **Original Custodians**: Maurice Smith
    - **Suspect or Employee Type**: BrainArray Employee
    - **Strengths / Weaknesses / Caveats**: May require additional expert testimony
    - **Components**: Everything in the Mobile Phone, RAM image, Storage Image
    - **Chain of Custody**: Forenzotica (2 analysts)

Analysis (Steps Taken)

False Leads and Unrelated Evidence

Unusual RF (radio) activity - RF scanning

The security team finds multiple circuit boards around the BrainArray datacenter, after further inspections using monitoring equipment It appears that the devices are using RF technology. This technology could be used to intercept communications or wireless systems data. Turns out that upon further inspection, these were simply part of side initiatives in the organization which did not fall under the purview of the security team and presented no malicious purpose.

Drone Surveillance

Employees and security teams notice new sightings of drones, one of the drones are shot down and are investigated and evidence is found of the drones capturing images and videos and using a built in visual sensors to maneuver around the premises and send data back to a local receiver. This turns out to be a part of a "fan" YouTube channel and presented no threat. This individual was found to not be a threat and was given express permission and instructions on how they could drone around the premises.

Suspicious Local Businesses

Suspicious restaurants and businesses in nearby city- Open Source Intelligence teams discover three local businesses near the datacenter which have suspicious activities and connections with foreign intelligence operatives. Obtained security footage showed these businesses operate within unusual hours of the day, appear to have repetitive customer visits, and appear to deal with only cash transactions. Later, this connection was found to be a dead end. Although, it could prove to be a local meetup for spies and other espionage. These establishments are being closely monitored by government agencies. However, they don't seem to directly relate to the recent attack.

Satellite Imagery

After working with government agencies investigators were able to obtain Satellite Imagery and discover unusual vehicles with altered license plates. This was also shown to be a dead end in terms of the current investigation. This is the conclusion of our teams work which found that the affected facilities were unrelated and disconnected from the affected systems. Therefore, these happenings could not be related to the recent incident.

Deletion of AI Infrastructure and Data

Initial RAM Image Analysis - 'CloudNova-CN-AIOrch-001' and 'CloudNova-CN-AIOrch-002'

First, we used Volatility to look through the RAM images for the primary AI orchestration servers. These were the only servers with access to the data storage servers (HoloStor-HS5000X-12TB, QuantumDrive-QD6000Z-10TB, Oracle QuantumX10) or at least in regards to controlling them. These data storage servers take API calls from the primary orchestration servers. In technical terms, all storage changes can be tracked down to processes which are directly controlled by the primary orchestration servers.

When using Volatility, we looked at the running processes and compared them with periodic CloudNova VPC Logs data. The VPC logs had proven-to-be untampered process list logs. Ultimately, the running processes were quite normal and that none of them were untrusted. However, we looked at the API calls from the data storage servers and directly correlated ID values with a particular running process. This was an update script that ran on the orchestration servers once a day. It was apparent that this was the culprit. The forensics team collaborated with the AI development teams to find that the script had no direct access to data storage and that only some of the programs running for the consumer AI services did. Therefore, our team saw

this as potential for APT-like techniques and potential process injection or hooking. The forensics team moved forward with doing further analysis on the RAM images alongside data from the VPC logs.

Correlating CloudNova IAM Logs with Memory Artifacts

As stated numerous times, the CloudNova VPC logs use tamper-proof hash chaining which makes them a trustworthy source for log data. Using the IAM (identity access management) logs, it was confirmed that no user was present or at least active on the systems during the time of deletion. Network connection also showed that there were no obvious signs of command and control from within the servers. Therefore, we looked deeper in the RAM images for advanced techniques. We used Volatility some more along with a plugin called "MalFind" to look for signs of injected code within process memory. It does this by comparing regions of memory to the expected behavior of normal processes. It also uses characteristics like VAD tags and page permissions. To enhance detection, we also used stack and VAD information together, as suggested by research, to help identify code injection attacks even when VAD information may have been modified by the attackers or protections may have been changed (Srivastava & Jones, 2017). Luckily the CloudNova VPC also had process monitoring enabled on the servers (by default) since these servers sometimes needed complex debugging and troubleshooting during development. As such, system calls (although remote) were traced down to that same specific process and cross-referenced with the Volatility results for "malfind." The forensics team then did a timeline analysis from the time when abnormal system calls and network connections started up over time. It was found that only one user had ever been present in the filesystem of the servers during this period where the process injection could've occurred. This was confirmed with the VPC logs which were tamper proof, so it could not have been possible

that the APT could've tampered to frame the user. Additional event logs for the server confirmed any potential alternatives for this explanation.

The Fake User

It was found that the applicable user (Darien Dashiell) in question was not a legitimate user. This user had privileged access to systems and yet had somehow only been created recently. This user had also been created by a parent user (Hector Zula) which did not have enough permissions to create users. This was backed up by the CloudNova VPC logs. Evidence could not be found on Hector's workstation either. Additionally, this new user had been created from the location of the AI orchestration server. Ultimately, there were only two system administrators which had credentials to access these systems or create users: Maurice Smith and Trey Valentine. It should be noted that Trey Valentine had been reported during employee interviews to have been "disgruntled" with BrainArray. In their words, disgruntled is an understatement. Trey is now the prime suspect along with the potential for Maurice having been a malicious insider.

Stealing of AI Models and Data

Memory analysis of AI infrastructure devices

The investigation team uses memory artifact analysis tools, which are used to examine the memory of the AI infrastructure for signs of intrusion. Windbg is a source-level debugging tool which allows analysts to identify exactly what happened during execution given the memory dump with source-level referring to how it can step through code or executed instructions. Windbg analyzes the "debug symbols format" included in memory dump files which maps strings in the dump file to corresponding objects in a program's source code, allowing for a

detailed view of program execution (Garcia, 2007). Using Volatility, WinDbg, and process monitoring logs from the CloudProvenance system logs from the VPC we discovered system calls that appeared to be using network adapters to communicate with whitelisted IP's. After further investigation it was determined that the applicable devices were not part of the AI infrastructure. This suggests that an external actor injected these malicious processes into employee workstations (Hector Zula) to gain unauthorized access to the AI models.

Network traffic monitoring.

BrainArray monitors the network traffic to and from the AI infrastructure devices to look for unusual activity. We found several instances of unusual patterns of large data transfer to external IP addresses, which match with the system calls found in the earlier stages of the study. The investigators use a firewall analyzer to examine the network firewalls and VPC configuration and discover that several IP addresses have been whitelisted. This shocks the security team because of the privileges and necessary permissions needed to allow this action, could only suggest a bad insider.

User Account Analysis

The investigators analyze the user accounts responsible for whitelisting the suspicious IP addresses, and discover they've just been recently created. The investigation goes further to determine whether they were created for legitimate reasons or malicious activity. Using the creation dates, permissions, and activity logs, these accounts are cross-referenced with other known users and events, and it is found that the accounts were created shortly before the whitelisting of the IP addresses and system calls. This suggests that they may have been part of the coordinated effort to steal the AI models.

Employee interviews and Background checks:

The investigation team interviews employees with access to the AI infrastructure and conducts background checks to identify workers with a motive to participate in theft. They look for any connections between the employees and the suspicious accounts or any other involvement in the authorized access of the infrastructure. The team reviews social media and communication activities of the employees, looking for any signs of contact with foreign adversaries. Lastly, they look into the access logs and badge swipes for the AI facilities to look for suspicious patterns and entries. The information collected is used to narrow down the list of potential insiders.

AI Model Corruption

Analyzing the AI Model and RAM

We captured and analyzed the memory dump by searching the AI model parameters, intermediate computation, and all input and output data. We used Volatility and WinDbg which are forensic tools to analyze memory dumps and discovered that there were inconsistencies in the parameters which resulted in data corruption. The model parameters have been modified which causes the output of the model to be poor and full of discrepancies. We found corruption evidence in the various forms below

I. **Running processes**- We analyze the processes and found malicious names, process IDs, and memory locations.

II. **Network Connection**- We found IP addresses that are outside of the network and open ports.

III. **User Credentials**- We extracted users' credentials such as usernames and passwords and we identified malicious credentials that we believed were used in corrupting the system.

IV. **Malware Artifacts**- It was discovered that malware artifacts such as malicious processes, network connections, and code injected into legitimate processes were found in the RAM. We use this to identify the type of malware, its behavior, and its damage to the system.

V. **File System Artifacts**- Some file system artifacts were found in RAM dumps, such as open files, registry keys, and file handles. This information was used to identify files that were accessed or modified, and by which processes.

With the evidence, we found how they had corrupted the model in two ways. The first finding was of what would be called an "AI trojan." An AI trojan is designed to work accurately and as normal on normal expected data, but it can behave maliciously on data samples with certain "triggers" (Xu et al., 2021). It seemed that the AI model had been modified early on to run additional code and process images with certain timestamps in a malicious manner. Specifically, the code would ingest malicious code and slowly build up a collection of malicious code to be later executed. It was a stealthy way of ingesting code via steganography mechanisms. The second finding was that the attackers had injected code to drastically change the model parameters of various CloudNova-based BrainArray AI models. This attack is not new and extremely effective when certain measures are not taken to prevent rather than simply detect it (Ramirez et al., 2022). To summarize the meaning of "model parameters", neural networks use linear algebra and networks of mathematical transformations to obtain the ultimate outputs. In this case, model parameters refers to the values used in each "neuron" of the neural network to

control the output along the way. If these are changed across the neural network, then the model quickly starts breaking down and becoming useless and inaccurate. The adversary's code was made to slowly deteriorate the model's parameters. Fortunately, the malware was not designed to cripple BrainArray's alarms that went off during this deterioration phase. Therefore, this was the first sign to BrainArray that something catastrophic was occurring with their systems.

Investigation into Trey Valentine

Cryptocurrency Transactions

The investigation did not take much longer to conclude after finding Trey and Maurice as highly suspicious. BrainArray and its work which utilize CloudNova infrastructure are under the control and guidance of multiple government agencies, all of which have wide-ranging access and resources into various parts of the internet which people normally do not have visibility. As of 2030, cryptocurrency and its use with crime have an adversarial relationship with government agencies. These agencies have better access to some cryptocurrency markets more than others, and they do a lot of work to analyze the blockchain with vast computing infrastructure and top researchers. Fortunately, interviews with Trey's subordinates showed that he talked about "crypto" quite often. The government agencies saw this as an easy opportunity. Evidence showed that Trey did start making more purchases than usual as if to be more confident in his earning and job security. Warrants with search engines and social media companies also revealed an increase in vacation option searches and accommodations overseas. This was a red flag. The government agency immediately got a warrant to look through his phone and found usernames and attributes which could be traced back to certain blockchain technologies. Cryptocurrency is not completely traceable but government agencies put a lot of work into being able to obtain data when necessary from crypto wallets such as transaction IDs, timestamp info, email addresses,

and even OAuth tokens (Chang et al., 2022). Soon enough, the government did analysis and found evidence of a 50,000-dollar worth cryptocurrency deposit into Trey's wallet. This evidence was admissible enough to charge Trey with a crime for colluding with an APT to destroy and steal BrainArray technology for monetary compensation from a foreign threat. Applicable data from Trey's wallet and information pointing to the adversary have been used alongside analysis methods such as with "heuristic clustering" to confirm that the APT can now be accurately identified (Fröwis et al., 2020).

Relevant Findings

The Initial Hypothesis Revisited

The political and systematic climate related artificial intelligence science and technologies has been a high stakes area of business in the U.S. ever since large AI models became more common. The AI race makes BrainArray an attractive target for numerous foreign and domestic threats – advanced and script kiddies alike. Before the acquisition of digital evidence, it was hard to tell what the motivation of the attack was because the AI industry is a heterogenous space when it comes to motivations and interests. In other words, it's hard to tell immediately why BrainArray was attacked. However, as will be explained, even the early analysis predominantly showed that a foreign likely state-sponsored APT was behind the wide-scale failure and exploitation of BrainArray's systems, data, and operations. Our initial hypothesis pointed to an APT because of the sophistication and practices behind BrainArray's security. During analysis, it seemed likely that only an insider with high-level privileges or an APT with zero-days could've pulled off the attack. BrainArray's systems are purposefully simple when it comes to the overall topology and architecture. This minimizes their attack surface and makes it easy to monitor and integrate with security initiatives. You can see this in the fact that

BrainArray only has a few orchestrator systems that use a singular VPC. Now, these models are deployed outwards to the "edge" (wasn't discussed earlier). However, theoretical security can be used to prove that these models are perfectly secure at the edge using extreme levels of encryption and key lengths. Therefore, the only way is to gain access directly to the VPC that hosts the AI infrastructure. Only the most advanced or lucky adversaries could've done so or an adversary which has a connection to the inside.

The Exploitation of BrainArray Systems

Cloud Memory Forensics – Deletion of AI Infrastructure/Data

It was obvious from the data obtained during this part of the analysis, that BrainArray had been attacked by a formidable threat. The initial analysis and correlation of the RAM images with the VPC logs (included process monitoring logs) showed that the attacker had injected binaries directly into periodic processes that interacted with the databases that stored BrainArray's AI models. They effectively repurposed a script that stores the AI model and metadata into a script that destroys them. Such an understanding of the systems requires vast resources, AI scientists who can understand the systems, and potentially an individual who already is familiar with them – an insider.

User Logs Analysis – Fake User Creation & Evidence of Utilization of Existing Accounts

It was shown that a process was hijacked with a malicious binary when we utilized memory forensics techniques and correlated that evidence with tamper-proof process logs. Additionally, we found that a solitary user had been present on the system in question during that time. Meaning, this user was the one who had conducted the malicious activity. Our team found that this user (Darien Dashiell) was not a legitimate user or even an employee of BrainArray. This user had been created recently by a Hector Zula, and yet, Hector did not have the access to

do so. Therefore, the trail led to the only users who could've gained access to Hector's account which was one of two system administrators: Maurice Smith or Trey Valentine – one of which had been previously reported as "disgruntled" with the ethics and management of BrainArray.

Network Device and Logs Analysis – Unauthorized & Malicious Communications

Using expert supported tools, we found proof showing manipulation of employee workstation devices to allow for unauthorized communications into BrainArray's AI systems. These workstation devices were being used as "proxies" or middlemen to gain access into systems using machines and accounts that already had existing access. The found communications on Hector Zula's machine showed transfer of BrainArray's extremely sensitive AI model data along with other communications necessary to command and control malicious operations inside BrainArray's AI systems.

Evidence of Malicious and External IP Addresses Being Added to Company Whitelists

Without being on a whitelist, a user cannot communicate over BrainArray's network. Whitelists are even stronger than blacklists because they disallow anything that's not on the list like a bouncer. This is contrary to a blacklist which operates more like the a security line that only doesn't let you in if you're on a watchlist.

IAM Logs Analysis – Fake User Creation & Evidence of Utilization of Existing Accounts

Luckily, when such an action is taken, BrainArrays stores tamper-proof logs of it in the cloud and in backup drives. Unfortunately, it was found once again that a new user had been created to do this on Hector Zula's machine. The only account that couldn't accessed his machine are one of the system administrator's accounts which points to Maurice Smith of Trey Valentine.

Network Activity - Evidence of Stealing AI Model Data

BrainArray's AI models are very large. In fact, the data storage devices which we investigated were sometimes distributed across multiple sets of drives that amounted to thousands of terabytes. Forenzotica investigated the activities associated with the malicious users, external IP addresses, and middleman devices that were being "zombified" for malicious use on the network. It was found that intermittent communication also consisted of large data transfers on the scale of terabytes. We cross-referenced this data further with tamper-proof logs (CloudProvenance) which showed access to the AI model databases which matched the timeline of these activities. Therefore, we can conclude without a doubt that the adversaries were stealing the AI models.

RAM Analysis of AI Orchestrators and Databases – Corruption of AI Models

BrainArray's AI scientists and Forenzotica collaborated to analyze the affected AI models from the incident. As previously mentioned, BrainArray was first alerted during the incident by users who were experiencing drastic corruption and functionality loss with their AI models. Forenzotica saw this as an easy piece of evidence that could be analyzed since the AI models could only be manipulated in a few ways – according to BrainArray scientists. Forenzotica analyzed the RAM images and tamper-proof process data of the applicable AI systems and found evidence of tampering with the models parameters. Think of the model parameters like memories. If someone manipulated your memories then it could drastically change how you ultimately behave. The same goes for AI models. It was immediately apparent that the threat actors had changed these "memories" in easy to notice ways. In this case, they simply changed most of the parameters with unnatural values that could not have occurred

during training. These attackers could have taken a more stealthy approach with corrupting the AI models. However, it seems that this model corruption had not been their primary object.

Evidence of Trey Valentine Colluding with Foreign APTs

Reasonable Proof of Foreign APT Involvement

BrainArray is involved with numerous government agencies which cannot be directly disclosed. However, we have evidence which proves through various indicators of compromise, malware analysis of the injected binaries, and active cyber warfare intelligence that points directly to a foreign APT. Additionally, there are various artifacts from our investigation that suggest this which we can discuss. Firstly, the particular method used to establish a foothold or command-and-control point into our network show advanced methods which have only been seen in classified incidents with government agencies. Namely, the threat actors used the AI systems as a guise and used the images being ingested by these systems to upload the malware. Such a method has only been seen in classified operations from one foreign state-sponsored APT.

Network communications also suggest a foreign APT as some of the metadata was not using English. There were other places this was seen too. RAM image analysis and malware analysis suggested that the threat actors were from a specific foreign threat. Certain language use along with the code and binaries suggests that a domestic threat is unlikely. Additionally, AI technology was used with open-source and popular knowledge graphs to calculate the connections to specific known APTs (Ren et al., 2022).

The actual malware and associated metadata of devices and all communications were also utilized in trying to attribute the APT. Things such as the programming languages used for

certain identified malicious source code, the sophistication of the design of the malware, assumed compilers, grammar and language mistakes in communication, and the overall approach to the attack were all analyzed deeply to find potential flags (Skopik & Pahi, 2020). Language seemed to be quite useful in determining the likelihood of certain APTs. However, investigators anticipated that many of these aspects can be red herrings or fabrications depending on the level of sophistication with the APT.

The external IP addresses were an easy target for forensics and BrainArray to look into. However, IP addresses are not trustworthy when it comes to adjudicating the location of or the identity of the threat because IP addresses are typically not trustworthy. However, government agencies involved with BrainArray conducted attacks on these IPs to attempt to adjudicate the locations of the threat actors along with attacks via. known forums that they use on the dark web. The identity of the involved APT was confirmed in this manner.

Additionally, the seemingly primary objective of this attack was to steal the AI model data and metadata. The combination of the current political climate and the identity of the adversarial group suggests that this was a direct attack from a state-sponsored entity which wanted to exploit BrainArray's system so that they could ultimately utilize their intellectual property and functional AI models.

Trey Valentine Marked as a Primary Suspect

The involved government agencies were given the green light to investigate Trey Valentine after finding out: 1) that the fake users could only have likely been created by one of two administrators (one of which was him), and 2) that Trey had a history of disagreement with BrainArray. They found social media evidence, Discord evidence, and numerous other public

and private communications which proved Trey's dislike for BrainArray and even talks about subverting BrainArray.

The most troubling evidence was his recent interest in foreign countries and their accommodations. This is common activity is a malicious insider is being paid in large sums of money. Therefore, government agencies and Forenzotica would move forward with an aggressive investigation of Trey Valentine, because this meant that they had a timeline to find evidence and apprehend Trey.

Cryptocurrency Transactions – Proof of Trey's Involvement

Forenzotica and government went forward with a deep investigation of Trey Valentine. This included unlimited access to his accounts through various companies, numerous warrants, and use of classified technologies to analyze his data. In the end, a specific classified unit in one of the government agencies analyzed some cryptocurrency markets to ultimately find two transactions into Mr. Valentine's cryptocurrency wallets. One transaction amounted to about fifty thousand US dollars, and the other amounted to about thirty million dollars.

Conclusion/Recommendations

Conclusion of Evidence

The investigation had revealed that the Brain Array incident was the results of a coordinated attack by a foreign APT with the assistance of an insider. Trey Valentine helped the APT orchestrate an attack which led to theft, deletion, and corruption of Cloud Arrays AI infrastructure. Our investigation used a multi-faceted approach to digital forensics, which proved to be effective in finding the culprit. Using process monitoring, cloud provenance systems, and the data integration from multiple sources. Process monitoring allowed the team to identify the

discrepancies that were running on the processes and helped us uncover the injected malware. Tools such as Volatility and WinDbg were instrumental in discovering the compromised processes and ultimately identifying Valentine as a malicious insider. Tamper log stores used in the investigation utilized cryptographic methods like hash chaining, which helped maintain the integrity and reliability of the evidence collected. The investigators also collected from multiple resources, such as employee interviews, digital evidence, network traffic logs, and memory forensics.

The investigation also delved into the financial and communication aspect of the case, tracing cryptocurrency to uncover the monetary motivation behind the attack. Obtaining a search warrant helped investigators look through search engine and social media data, which helped us discover Trey's sudden increase in overseas vacation searches, which suggested that Trey had confidence in his financial standing. Monitoring and analyzing communication channels which included email exchanges, instances messages, and social media activities for signs of contact between the APT and Trey helped provide valuable information to the planning and execution of the attack.

This investigation highlighted the need to for the organization to invest into robust cybersecurity measures and advanced forensic tools, and techniques. Moving forward Brain Array needs to improve its process for vetting employees who will have access to critical systems and data. Strengthening user account management policies, which includes monitoring account creation, escalation, and rights to delete and tamper with data. Proving ongoing security awareness and training for all employees, highlights the risks and consciousness of insider threats, and encourages a security conscious environment within the organization. Lastly the organization needs to set up a procedure in place that is regularly updated and revised. This

ensures the plan is comprehensive and provides clear procedures for detecting, containing, eradicating, and recovering from future attacks.

Conclusion of Investigative Process

Overall Steps of Investigation

The investigation consisted of prioritizing the acquisition of volatile data to prevent data loss and address multi-tenancy issues. We utilized tools such as Volatility and EnCase for imaging hardware components, and we worked with law enforcement, regulatory agencies, and other legal teams to ensure compliance with privacy regulations.

This investigation successfully obtained relevant evidence from BrainArray's affected infrastructure, which was to analyze and identify potential threats. These threats included vulnerabilities and high stakes in the event of the exploitation of high-level privileges. The threat actors included a malicious insider (Trey Valentine) and a foreign advanced persistent threat which colluded for the combined interest of making money and destroying BrainArray. This investigation was compiled in conjunction with several legal, privacy, and regulatory requirements throughout the entire collection and analysis process.

Recommendations

Moving forward we recommend focused security improvements relevant to this investigation which include: continuously updating security measures and protocols, simplifying architectures as much as possible with zero-trust approaches to their design, monitoring processes on assets in the cloud, and utilizing upcoming cloud provenance mechanism which are tamper-proof in theory and in practice. Additionally, having business continuity and disaster recovery (BCDR) baked into the design of applications and systems can mitigate expensive

breaches and destruction. If using cloud technologies for your implementations and business solutions, then be sure that an incident response plan and standard operating procedure (SOP) are in place for those systems. Lastly, having a plan for insider threat detection could be beneficial. The approach for this will differ org-to-org, but in the case of Trey Valentine, a lot could have been avoided if Trey had been somewhat "scared" from a small conversation with HR. In other words, if Trey knew that BrainArray knew about his disagreements, then he would be less likely to perform the malicious actions.

CONTINUE NEXT PAGE

Exhibits

1. 

   i. Collection of Forenzotica hard drives and WORM (write once read many) type drives with various organized redundant digital evidence and integrity protections:

      1. Malware Artifacts.
      2. File System Artifacts.
      3. Cryptocurrency transactions record.
      4. Fake user information and artifacts
      5. AI model data and parameters
      6. Unauthorized access history to infrastructure facility.
      7. Suspicious network activity logs and evidence.

8. Proprietary CloudNova Distributed Hash-Chaining Based Storage logs (CloudProvenance logs).
9. Search history of vacation options and accommodation history which includes expenses that are far above what Trey can afford based on her income.
10. Trey's personal activity on social media accounts and online forums
11. Data was also stored from RAM images and drive images from the following devices:
    a. Oracle QuantumX10 Serial Number: OQX10-901-234.
    b. QuantumDrive-QD6000Z-10TB Serial Number: QD-6000Z-678.
    c. HoloStor-HS5000X-12TB Serial Number: HS-5000X-345.
    d. CloudNova-QCN-AppVPC-002 Serial Number: QCN-AV-002-012.
    e. CloudNova-QCN-AppVPC-001 Serial Number: QCN-AV-001-789.
    f. CloudNova-CN-AIOrch-002 Serial Number: CN-OA-002-456.
    g. CloudNova-CN-AIOrch-001 Serial Number: CN-OA-001-123.
    h. CloudNova-CN-AIOrch-002 Serial Number: CN-OA-442-446.

2. 

   i. Google Pixel 12, serial number XQ8S7T6U5V4R.

   ii. Samsung Galaxy S35 serial number WQ9R8S7T6U5.

3. 

   i. Fortinet FortiGuardian 10G Serial number QT-3450-589-899.

4. 

i. Cisco LightSpeed 3000 serial number LS-3000-567-890.

References

Chang, E., Darcy, P., Kim-Kwang Raymond Choo, & Nhien-An Le-Khac. (2022). *Forensic artefact discovery and attribution from android cryptocurrency wallet applications*.

Fröwis, M., Gottschalk, T., Bernhard Haslhofer, Rückert, C., & Pesch, P. (2020). Safeguarding the evidential value of forensic cryptocurrency investigations. *Forensic Science International: Digital Investigation*, *33*, 200902. https://doi.org/10.1016/j.fsidi.2019.200902

Garcia, G. L. (2007). Forensic physical memory analysis: an overview of tools and techniques. *TKK Helsinki*, *207*, 305–320.

Imran, M., & Hlavacs, H. (2015). *Provenance in the Cloud: Why and How?* http://personales.upv.es/thinkmind/dl/conferences/cloudcomputing/cloud_computing_2012/cloud_computing_2012_5_20_20114.pdf

Karie, N. M., & Venter, H. S. (2015). Taxonomy of challenges for digital forensics. *Journal of Forensic Sciences*, *60*, 885–893. https://doi.org/10.1111/1556-4029.12809

Liu, A., Fu, H., Hong, Y., Liu, J., & Li, Y. (2019). LiveForen : Ensuring live forensic integrity in the cloud. *IEEE Transactions on Information Forensics and Security*, *14*, 2749–2764. https://doi.org/10.1109/TIFS.2019.2898841

Lutui, R., & Cusack, B. (2021). *Data acquisition from cloud network storage*. 1–6. https://doi.org/10.1109/ITNAC53136.2021.9652168

McRee, G. R. (2022). Improved detection and response via optimized alerts: Usability study. *Journal of Cybersecurity and Privacy*, *2*, 379–401. https://doi.org/10.3390/jcp2020020

Paccagnella, R. (2019). *Towards trustworthy foundations for operating system Forensics*.

Prasad Purnaye, & Kulkarni, V. (2022). BiSHM: Evidence detection and preservation model for cloud forensics. *Open Computer Science*, *12*, 154–170. https://doi.org/10.1515/comp-2022-0241

Ramirez, M. A., Kim, S.-K., Hussam Al Hamadi, Damiani, E., Byon, Y.-J., Kim, T.-Y., Cho, C.-S., & Chan Yeob Yeun. (2022). *Poisoning attacks and defenses on artificial intelligence: A survey*.

Ren, Y., Xiao, Y., Zhou, Y., Zhang, Z., & Tian, Z. (2022). CSKG4APT: A cybersecurity knowledge graph for advanced persistent threat organization attribution. *IEEE Transactions on Knowledge and Data Engineering*, 1–15. https://doi.org/10.1109/TKDE.2022.3175719

Skopik, F., & Pahi, T. (2020). Under false flag: using technical artifacts for cyber attack attribution. *Cybersecurity*, *3*(1), 8. https://doi.org/10.1186/s42400020000484

Srivastava, A., & Jones, J. H. (2017). *Detecting code injection by cross-validating stack and VAD information in windows physical memory*. 83–89. https://doi.org/10.1109/ICOS.2017.8280279

Xu, X., Wang, Q., Li, H., Borisov, N., Gunter, C. A., & Li, B. (2021). *Detecting AI trojans using meta neural analysis*. 103–120. https://doi.org/10.1109/SP40001.2021.00034

Zawoad, S., Hasan, R., & Islam, K. (2018). *SECProv: Trustworthy and efficient provenance management in the cloud*. 1241–1249. https://doi.org/10.1109/INFOCOM.2018.8485824