SANS

Analyst Program

Whitepaper

A SANS 2021 DFIR Cloud Report: Partly Cloudy with a Bunch of DFIR

Written by **Domenica Lee Crognale** and **Heather Mahalik**

October 2021

Uptycs

Executive Summary

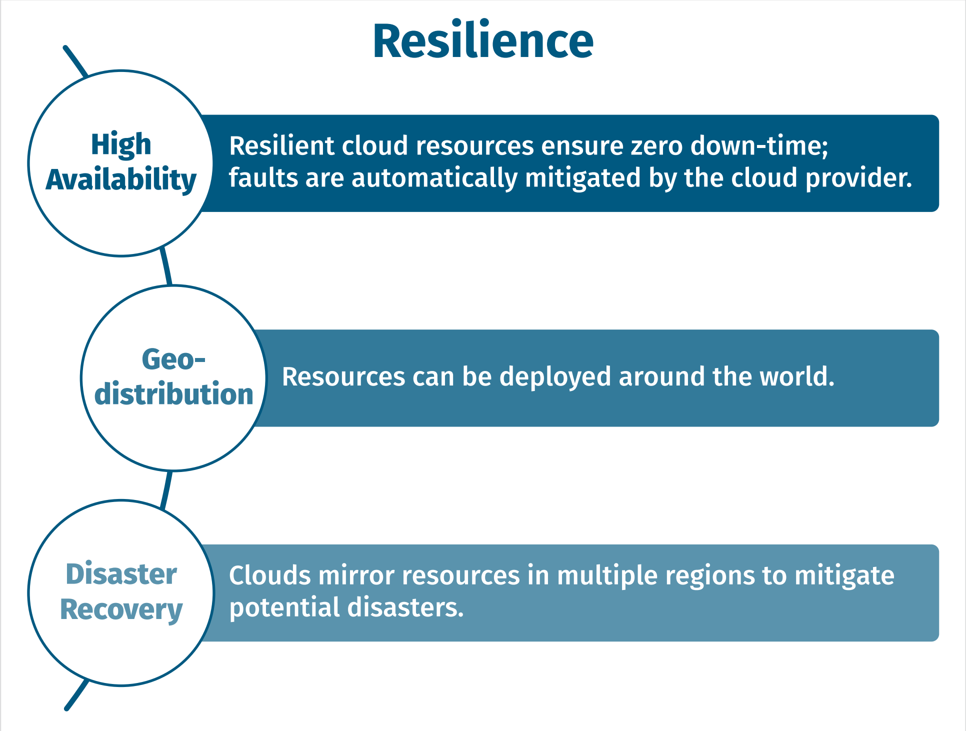
More and more individuals and organizations now leverage the cloud for their storage, processing, and computing needs, and the data stored in the cloud is equally diverse. This abundance of data provides an attractive target for attackers. For these reasons, we see a rapid increase in the need for cybersecurity professionals who can both protect the stored data and also investigate breaches that occur. While the approach may differ slightly from a typical digital forensic and incident response (DFIR) investigation, organizations gain the added benefits of increased logging, which can aid practitioners in their analysis. This paper explores the major cloud providers, models and deployments, reported attacks on the cloud, the challenges and processes surrounding collection, preservation, and analyzing cloud data, and some best practices moving forward.

Cloud Computing: Everybody's Doing It

Most everyone with a digital device has utilized a cloud resource in 2021. Whether you check out new trends on TikTok, share first-day-of-school photo albums with your family via Apple's iCloud or Google Photos, or work from home due to COVID-19 restrictions and use Slack, Zoom, or Microsoft Teams to communicate with co-workers, the cloud stays hard at work behind the scenes.

The term *cloud computing* refers to the process of offloading some or all of your typical computing functions and IT resources so that you can perform them over the internet without maintaining hardware, software, storage, or other physical components in-house.[1]

Cloud adoption has grown year over year, with more than 90% of businesses using some type of cloud capabilities in 2021.[2] More companies are moving in this direction because of the many benefits to entrusting certain services to firms that specialize in the various IT offerings. Cloud computing provides an attractive alternative for many organizations, due to its capacity to provide a resilient, responsive approach to traditional computing environments. See Figures 1 and 2.



Resilient cloud resources ensure zero down-time; faults are automatically mitigated by the cloud provider.

Resources can be deployed around the world.

Clouds mirror resources in multiple regions to mitigate potential disasters.

Figure 1. Cloud Resilience Characteristics[3]

[1] https://aws.amazon.com/what-is-cloud-computing/

[2] "Five Reasons More Businesses Are Choosing Cloud," www.forbes.com/sites/forbestechcouncil/2020/10/14/five-reasons-more-businesses-are-choosing-cloud/?sh=7f70dabe33d9

[3] SANS FOR509 Enterprise Cloud Forensics and Incident Response, www.sans.org/cyber-security-courses/enterprise-cloud-forensics-incident-response/

These benefits ultimately lead to an overall decrease in operating costs for the business in the long term and an increase in consumer confidence.

Cloud Computing Environments

The diverse cloud ecosystem offers varying combinations of models and deployments molded to fit a particular customer's needs. Cloud services often follow these three models: software-as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS).

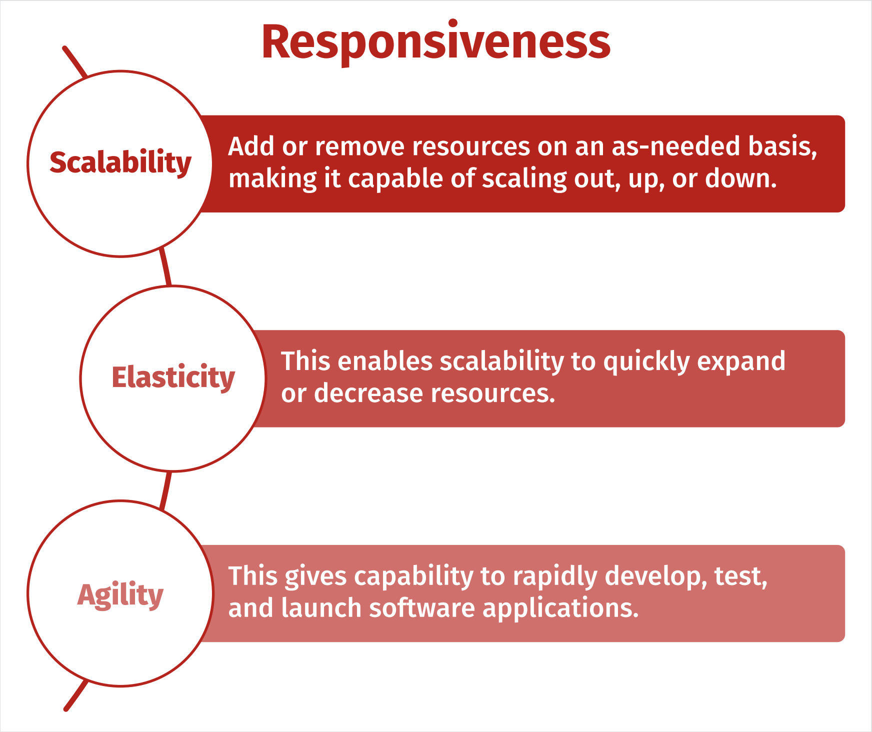

Responsiveness
Scalability
Add or remove resources on an as-needed basis, making it capable of scaling out, up, or down.
Elasticity
This enables scalability to quickly expand or decrease resources.
Agility
This gives capability to rapidly develop, test, and launch software applications.

Figure 2. Cloud Responsiveness Characteristics[4]

SaaS is one of the more commonly used cloud services. This typically means paying for a service to access software capabilities that require no deployment, configuration, or maintenance by your own personnel. Clients who make use of SaaS often want an application or a suite of applications to handle tasks that they do not desire to manage or cannot manage on premises. SaaS applications operate analogous to desktop applications and are often transparent to users. Dropbox represents a good example of an application that can handle excess storage needs for your business if you do not want to invest money in costly network storage hardware and the associated maintenance. HR applications often present a logical choice for businesses looking for ways to track billable hours, training, expenses, benefits, and other important items for their workforces.[5] Many businesses have switched from maintaining their own mail servers to using a service such as Gmail and have replaced their physical Microsoft Office Suite of software with Office 365, because they no longer have to maintain equipment, licenses, and employ personnel to keep these applications up to date.[6]

PaaS provides another option for businesses. This cloud service offers both the infrastructure and the operating system and usually some additional resources such as database support and Java. Although not necessarily geared toward end users, the option provides an ideal solution for development environments. With PaaS, developers need not focus on maintaining servers and operating systems, but instead can focus their energies on building and deploying new applications. Google's App Engine, an example of PaaS, gives developers access to Node.js, Java, Ruby, C#, Python, Go, and PHP in order to allow developers to effectively write code.[7]

[4] SANS FOR509 Enterprise Cloud Forensics and Incident Response, www.sans.org/cyber-security-courses/enterprise-cloud-forensics-incident-response/

[5] www.comptia.org/content/articles/what-is-saas

[6] www.cloudendure.com/blog/cloud-technology-iaas-paas-saas

[7] https://azure.microsoft.com/en-us/overview/what-is-paas

IaaS transfers the most control to its users. Subscribers to this service often need access to one or more of the resources on offer and choose IaaS as the least costly method of attaining it. This option allows for access to storage, network resources, and additional computing power, all with the ability to interact, scale, and monitor as necessary. Examples of IaaS include Microsoft's Azure, Google Cloud Platform (GCP), and Amazon Web Services (AWS). Each model provides its own benefits to the end consumer, but as more of the overhead moves to the cloud, the less oversight an entity may have over the associated logs.

Cloud Deployments

In addition to the multiple service models available, organizations have a mix of deployment options to choose from in the cloud—namely public, private, and hybrid.

Public clouds make resources such as compute, storage, infrastructure, networking, hardware, and software available via the internet to organizations, commonly referred to as tenants. These resources are shared among many *tenants*, and so a public cloud represents an attractive offering due to its zero maintenance, high availability, and capability to scale as needed.[8] Some of the largest public cloud service providers include Azure, AWS, and GCP.

In contrast, organizations maintain a **private** cloud on a private network accessible only to the organization, with the aforementioned resources not shared outside of the organization. You can have the infrastructure completely on premises or utilize third-party data centers, or the organization can rely on a mixture of the two. This type of deployment proves ideal for businesses that require a certain amount of customization, and government agencies or entities operating under strict regulations often choose this method. Some of the leaders in the private cloud space include Dell, IBM, and HPE, but the major public cloud providers offer private cloud services as well.

Hybrid clouds offer a little bit of both: security for those parts of your business that must remain private and that you must keep in your own data center, for example, and the flexibility to scale up using public cloud resources when the business demands it.

A **multicloud** approach, a growing trend with organizations, utilizes two different cloud providers for redundancy, or perhaps the organization uses one major provider and augments that with a regional offering such as Alibaba, Huawei, or Tencent Cloud.

[8] https://azure.microsoft.com/en-us/overview/what-are-private-public-hybrid-clouds/#private-cloud

Trends in Cloud Computing

Surge in Remote Work

With the uptick in remote work, more data is scattered globally than ever before. This may initially pose some hurdles for investigators who need quick, timely access to data for analysis. The old-school methods of physically connecting to and acquiring drive images do not prove feasible in many remote environments. We need the ability to collect data from systems regardless of their geographic location in a matter of minutes/hours and not days/weeks. While we can do this in most cloud environments, the effectiveness relies on the knowledge and expertise of your in-house or third-party forensics team. Much of the setup for effective remote access relies on having account access on a particular cloud network as well as having resources available in a particular region. Improper setup in this phase could result in significant fees for the client.

Increased Focus on Serverless and Container Technology

Virtualization technology continues to improve, and the ability to spin up a small bit of code on demand or utilize a sandboxed container application only when needed to keep overhead costs manageable has a big appeal for customers. The big three providers refer to their serverless offerings as *functions* or *lambda functions*, and you're probably familiar with container technologies such as Kubernetes and Docker. While customers see measurable benefits, the nature of these technologies can pose a problem for DFIR professionals. Serverless architecture lets you run a small bit of code based on a triggering event, running for mere moments, often leaving little trace evidence behind. Similarly, containers can be spun up and down at will, purging trace resources and log data while doing so, which makes the ephemeral nature of these technologies a sought-after technique for attackers trying to cover their tracks in your environment.

Volume of Data Continues to Expand

Unsurprisingly, the volume of data that investigators must process and analyze continues to grow each year, and the increase in volume leads to longer lead times for many of the forensic processes that need to be carried out as part of the investigation. It's no longer efficient to grab and process data from only one computer/server at a time when so much data requires analysis. Practitioners need the ability to simultaneously process and analyze multiple machines at the same time to provide timely, actionable intelligence to their customers. This may require investigators to process and analyze in the cloud, where the investigation can also take advantage of highly available, scalable resources that remain ready at a moment's notice.

Rise in Cloud Attacks

As with any new technology, as more people adopt it, the cyberattacks soon follow, and because cloud computing grows exponentially year over year, security teams feel the burden. Teams find themselves pivoting from host-based forensic techniques in favor of investigation practices tailored toward the cloud, but this doesn't come without a lot of time, investment, and knowledge. They must also orchestrate all this while carrying out an investigation cycle that itself seems to grow each year.

Attacks on Cloud

As individuals and businesses consume more cloud resources, attackers have been quick to set their sights on a new platform to target. In fact, attackers don't care much where that data is stored; they just want to access it so that they can disrupt, destroy, or exploit it for gain or profit. MITRE ATT&CK™ Matrix does a good job at documenting attacks on the cloud and includes information targeted to Azure AD, Office 365, Google Workspace, SaaS, and IaaS (which includes AWS) cloud-based platforms.[9]

Attack targets for cloud providers include:

- Business email compromise (BEC)
- Password spraying
- Credential stuffing
- Obtaining/validating of credentials with legacy protocols
- Data exfiltration (e.g., Graph API)
- Bucket attacks by exploiting misconfiguration (e.g., GCPBucketBrute and similar toolkits for AWS S3 buckets)
- Exploiting OAuth tokens
- Vulnerable/outdated preconfigured containers
- Abuse of cloud resources (e.g., VM utilization for cryptocurrency mining)

The following represent just a few real-world examples of attacks carried out against data and resources residing in the cloud.

Stolen Zoom Credentials Sold on the Dark Web

Half a million Zoom credentials ended up for sale on the dark web for mere pennies, or in some cases free—and this wasn't the first find of its kind. Attackers have targeted both personal and business accounts, with email addresses, passwords, host keys, meeting IDs, and Zoom account types among the list of items for sale.[10] Credential stuffing, an attack that automates attempted usage of legitimate username/password combos into login forms, relies on the reuse of credentials across multiple sites.[11]

[9] https://attack.mitre.org/matrices/enterprise/cloud/

[10] https://attack.mitre.org/matrices/enterprise/cloud/

[11] "Credential stuffing," https://owasp.org/www-community/attacks/Credential_stuffing

Business Email Compromise

The Russian attack group known as Nobelium—attributed to last year's infamous SolarWinds hack—recently spun up a new campaign targeting email. More than 3,000 emails were distributed to more than 150 organizations across 24 countries, with most recipients identified as government agencies or entities with ties to the government. According to Tom Burt, corporate vice president of Customer Security and Trust at Microsoft, phishing emails harboring the NativeZone backdoor, capable of "a wide range of activities from stealing data to infecting other computers on a network," were distributed from USAID's Constant Contact account.[12]

Data Exfiltration Using MS Graph APIs

One of the most notable cloud attacks in recent history occurred when purported Russian attackers hid the SUNBURST malware in SolarWinds's legitimate Orion software, and SolarWinds delivered it unknowingly to customers around the world as a normal software update. After initially breaching the systems, attackers leveraged Microsoft Graph APIs to carry out further attacks on their victims.[13]

Microsoft describes the MS Graph API as functionality that "exposes REST APIs and client libraries to access data on Microsoft cloud services" to include resources such as Microsoft 365, OneNote, Outlook/Exchange, SharePoint, Teams, Azure Active Directory, Windows 10 services, and more, as shown in Figure 3.[14]

Attackers targeted high-traffic applications including the Mail app to disguise their activity and added new credentials to these applications to facilitate data theft. The credentials allowed them to add additional permissions to applications, authenticate to Azure AD, and move laterally through the systems undetected. The same MS Graph APIs also allowed attackers to exfiltrate sensitive data from within the compromised applications.[15]

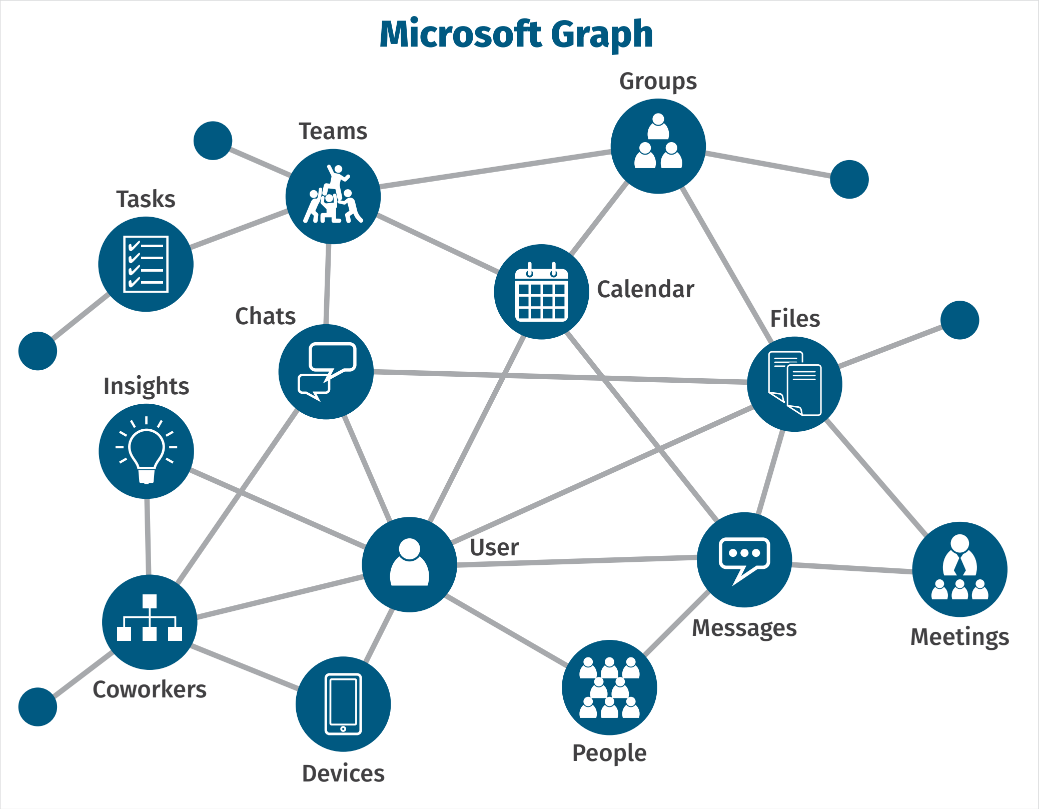


Figure 3. Microsoft Graph API[14]

[12] "SolarWinds attacker Nobelium targets almost 3,000 emails," www.arnnet.com.au/article/688722/almost-3-000-emails-targeted-by-nobelium-attack/

[13] "The US is readying sanctions against Russia over the SolarWinds cyber attack. Here's a simple explanation of how the massive hack happened and why it's such a big deal," www.businessinsider.com/solarwinds-hack-explained-government-agencies-cyber-security-2020-12

[14] Image adapted from https://docs.microsoft.com/en-us/graph/overview

[15] "Behind the Scenes of the SunBurst Attack," https://thenewstack.io/behind-the-scenes-of-the-sunburst-attack/

Crypto-mining in Containers

Crypto-mining, the process of surreptitiously using a computer's resources to mine digital currency, is an increasingly popular trend among attackers, and cloud resources represent the ideal opportunistic target.[16] In August 2020, the attack group Team TNT not only utilized standard crypto-mining modules but also employed the first worm of its kind that directly targeted both vulnerable Docker and Kubernetes systems. After initially infecting the server, the malware conducts a port scan using masscan to look for unsecured AWS configuration and credential files for exfiltration to their command and control servers.[17] This proves especially alarming as more consumers take advantage of containerization functionality available within the various cloud platforms.

DFIR Challenges

While threat detection capabilities continue to evolve, many challenges remain for those working in DFIR, with the myriad cloud vendors, service models, and deployment types only adding complexity to the mix.

Lack of Cloud Expertise

Cloud self-computing can be challenging—and, just as investigators carved out a place for themselves in host-based forensics by picking a favorite operating system, we can say the same for specializing in artifacts from one of the major cloud providers. Although you will find many of the constructs and offerings similar, you might find yourself quickly confused due to varying terminology. Becoming an expert in *cloud vernacular* is often the first step to getting what you need from your security team. If you cannot speak the language, you may not see the full picture.

Logging

Each cloud provider stores artifacts differently. Will the service model affect your access to logs? Will you need to tweak configurations prior to an incident to ensure that you get as many logs as possible? The answer is yes. Although SIEM solutions will record events, the log data—if properly configured—will contain the most crucial information. Most major cloud providers disable many essential logs by default, so knowing how to *properly configure your cloud services*, such as enabling verbose logging, will result in more data for analysis. To add additional complexity to the mix, most consumers adopt a *multicloud approach*, which utilizes two or sometimes multiple cloud providers, further complicating the gathering, normalizing, and analysis of artifacts in an efficient manner.

[16] www.malwarebytes.com/cryptojacking

[17] "Cryptojacking worm steals AWS credentials from Docker systems," www.bleepingcomputer.com/news/security/cryptojacking-worm-steals-aws-credentials-from-docker-systems/

Access to Data

When the data and artifacts required for an investigation reside in the cloud, it often eliminates the need for *access to the physical media* for acquisition and preservation purposes. However, certain geographic hurdles still present themselves as part of cloud data acquisitions. It's important to stand up resources in the region where the incident occurred to make efficient and cost-effective choices for acquiring logs or snapshots of affected systems. This requires preparation and forethought from your response teams so that they find systems, tools, and resources in place at the exact moment they need them.

This goes hand in hand with *access requirements*, as you will find these often a prerequisite for getting access to the data needed for an investigation. In-house IR teams can more easily achieve this, but for outsourced cloud responders, this represents one aspect of the investigation that they must handle up front and with immediacy to facilitate a rapid response. Any lag in response time could lead to the inability to collect data. Some logs, even if turned on by default, remain available only for a finite period, so collection must happen as soon as possible upon attack detection.

Traditional DFIR Methods Were Not Designed for the Cloud

Traditional DFIR methods were not developed for cloud-based incidents (as apparent with every step of the process from collection to analysis and with the level of difficulty increases in multicloud environments). As recently as 2020, the AWS Online Tech Talks: AWS Digital Forensics Automation at Goldman Sachs by Ryan Tick, Vaishnav Murthy, and Logan Bair outlined an example, detailing the 21-step manual process for the collection of forensic copies from AWS systems required before the start of any investigation.[18]

Organizations could and should automate many of these processes by using supplied templates and built-in functions invoked by a triggering event or indicator of compromise (IoC). The function kicks off the simultaneous collection of snapshots and memory images of the affected systems at the exact moment the detection is alerted, so no downtime occurs with any manual processes.[19] Note that all the major cloud providers and modern DFIR platforms have similarly tailored offerings that can automate these manual processes.

Time and Cost Constraints

With DFIR in the cloud, we often face a race against time. Getting access to incident evidence, such as logs, before it is purged can be the difference between a fruitful or an unsuccessful forensic effort. Organizations should automate as many processes as possible so that analysts can begin their investigation immediately. Processing data in the cloud can significantly cut down on the time that it takes to access, download, transfer, and import images into traditional forensic workstations for forensic processing and analysis.

[18] "AWS re:Invent 2020: Automated forensic artifact collection on AWS with Goldman Sachs," www.youtube.com/watch?v=W4Ih9zvuBa4

[19] "How to automate forensic disk collection in AWS," https://aws.amazon.com/blogs/security/how-to-automate-forensic-disk-collection-in-aws/

Tool Limitations

As cloud forensics evolves, you'll likely always find new tools available to accomplish the job. In fact, many of the major cloud providers offer free and paid built-in solutions that can aid in DFIR investigations by providing detection and threat-hunting capabilities. You'll find it beneficial to weigh the positives and negatives to using this approach over others before you find yourself in the middle of an investigation. Turning on alerting, logging, and other detection systems that could assist in analysis comes at a cost. Turning on and storing logs that are otherwise purged also increases fees for the customer; while these methods offer an overview of activity, do not mistake them for a thorough forensic investigation.

In contrast, some of the available paid solutions from cloud providers—Amazon Detective (holistic threat hunting and searching through available log sources) or Amazon GuardDuty (continuous monitoring) and Azure Sentinel (a cross-platform SIEM/SOAR solution that uses AI for threat hunting), for example—are quite robust, but they can create substantial price hikes quickly. Some of the cloud solutions can provide a quick overview assessment of the situation, but they do not rely on them to be as thorough as a full DFIR investigation. Because the nature of cloud investigations often spans multiple users, systems, regions, and cloud platforms, finding the right tool(s) for the job often proves daunting, and as more data moves to the cloud, this issue will continue to grow. Currently, examiners must often utilize multiple processes, techniques, parsers, and analytical tools to achieve results—and to compound the problem further, cloud providers can change their processes or output on a whim. This can mean that the tools you used to utilize for parsing the data no longer work the same way and also require updates. Adaptability in this field on the part of the analyst and the solutions vendors deliver is paramount.

Moving Forward

Evaluate of *all* your data sources. An investigation that encompasses every data source available produces a better result, and the ability to analyze this data in aggregate can help draw beneficial conclusions and provide a better overall picture of the attack. Analyzing a combination of cloud logs, host systems, and system memory gives insight into movement throughout the network, staged attacks, and possible infection vectors such as fileless malware, rootkits, and process hollowing, which have become more prevalent as attackers attempt to leave little to no footprints behind.

Automate, automate, automate. Cloud platforms can automate processes already built in to the environment. Have a plan, make a template, and utilize built-in functions to automate some of the repeatable processes that would otherwise slow down your collection and processing efforts. The quicker you can arm your analysts with data to investigate, the more efficiently they will perform their roles.

Invoke the power of the cloud. Processing your data in the cloud vs. on premises using traditional forensic methodologies has its advantages. Leveraging the cloud can drastically reduce the amount of time required to investigate an incident. While shifting your investigation to the cloud may incur additional costs, you must weigh this against the benefits of saved time and reduced risk. Assess your situation and take advantage of the power of the cloud when it makes sense.

Time is not on our side; the early bird gets the W0RM . . . or whichever potentially destructive malware comes our way. The dynamic nature of the cloud means that time is of the essence with every investigation. Using events to trigger automation, leveraging the cloud for processing, and integrating forensic capabilities with your SIEM/SOAR or EDR/XDR solutions will give you a jump on the investigation the moment you detect malicious activity.

Good preparation is essential to the investigation. As the cloud continues to evolve, it requires a constant shift in your processes. Technology and procedures can become obsolete quickly, so investing time in testing, assessments, and purple team exercises can keep your organization ahead of the curve. Similar to a good disaster recovery plan, organizations should regularly exercise, document, and debrief procedures so as to improve them. Optimally, you want to minimize disruption and get your business back up and running with the least impact to your company or your customers.

Investment in people is paramount. While many organizations have in-house incident responders, this doesn't mean that they necessarily are experts in cloud. Tapping the right people for the job is essential in any successful cloud incident. Many reputable incident response firms excel in cloud analysis. If your organization does not have experts in-house, you should best leave this job to an outside firm. They can be available at the first detection of a breach but also periodically assess your environment to ensure its proper configuration for detection and remediation. For those companies with limited resources, consider one of the widely available incident response playbooks for the cloud. Playbooks can provide tactics, techniques, and procedures for responding to an incident, but they will also document best practices for normal activities. Many firms offer free templates to get started.

Do your homework. It takes knowledge, practice, and time in the field to become a DFIR cloud expert, and because of the uniqueness of each cloud platform, this represents no small feat. You can find no shortage of excellent training, certifications, webcasts, and blog posts currently shared within the community. Consider platform-based certifications in security or engineering, vendor-led instruction, or tool-/platform-agnostic training to augment your existing forensic skills. Cloud computing is evolving fast, so prepare to roll with the punches.

About the Authors

Domenica Lee Crognale

SANS Certified Instructor **Domenica Lee Crognale** is a co-author and instructor of SANS FOR585: Advanced Smartphone Forensics. She has 15 years of experience in cybersecurity, with a focus on digital forensics and mobile device security. Lee developed a love for mobile device forensics while working in both the law enforcement and intel communities, where she was fortunate to work on many high-profile cases. She has provided specialized training to military special forces, the U.S. Coast Guard, and other government agencies, and has tested and validated various forensics utilities, researched artifacts associated with mobile operating systems and numerous third-party apps, and provided security assessments for many mobile applications.

Heather Mahalik

Heather Mahalik is a SANS senior instructor and course lead for FOR585: Smartphone Forensic Analysis In-Depth. As the senior director of digital intelligence at Cellebrite, Heather focuses on forensic research and making the community smarter with regard to all aspects of digital intelligence. Her background in digital forensics and e-discovery covers smartphone, mobile device, and Mac and Windows forensics, including acquisition, analysis, advanced exploitation, vulnerability discovery, malware analysis, application reverse engineering, and manual decoding. Prior to joining Cellebrite, Heather focused on mobile device forensics in support of the federal government and served as a technical lead performing forensic examinations for high-profile cases. Heather maintains www.smarterforensics.com, where she blogs and shares presentations.

Sponsor

SANS would like to thank this paper's sponsor:

Uptycs