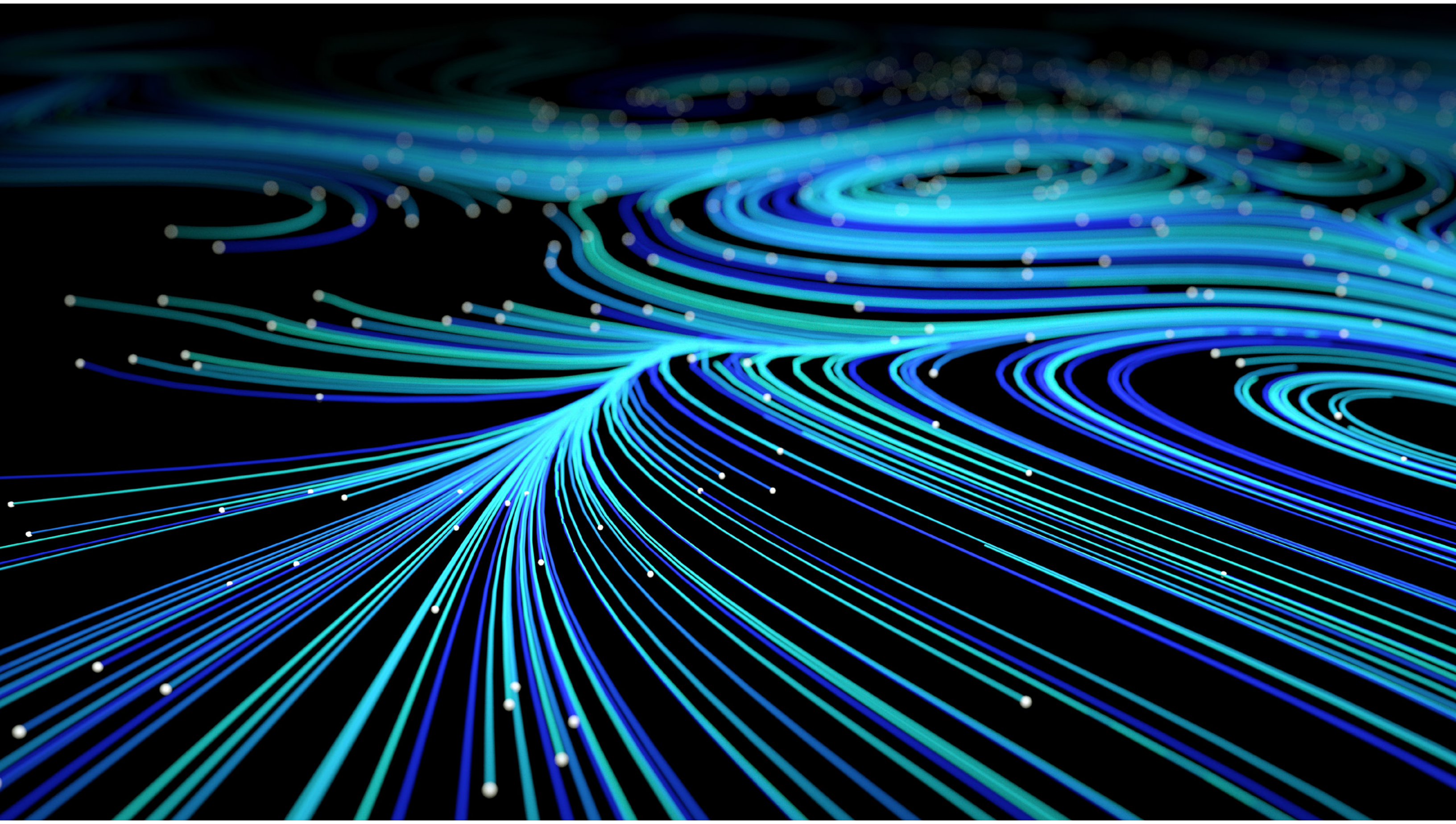
Google Cloud
Security

Threat Horizons

H2 2024 Threat Horizons Report

Office of the CISO

Table of Contents



Mission Statement

The Google Cloud Threat Horizons Report provides decision-makers with strategic intelligence on threats to not just Google Cloud, but all providers. The report focuses on recommendations for mitigating risks and improving cloud security for cloud security leaders and practitioners. The report is informed by Google's Threat Analysis Group (TAG), Mandiant, Google Cloud's Office of the CISO, Product Security Engineering, and various Google Cloud intelligence, security, and product teams.

Executive Summary

Arming Cloud Defenders with Security Mitigations from the Serverless Frontlines

Serverless computing has emerged as a transformative approach to application development, promising scalability, reduced operational overhead, and faster time-to-market.

Serverless products also create opportunities for threat actors in cloud providers from potential security misconfigurations in customer environments. What does this mean for cloud security professionals?

Based on recent serverless cloud threats that our security and intelligence teams are seeing, the following are three key considerations to prioritize when developing your cloud security strategy:

- **Compromised credentials:** Threat actors continue exploiting weak passwords to gain unauthorized access to Google Cloud projects. At the same time, serverless computing may make cryptomining an even more attractive target for some threat actors, underscoring the importance of efforts to identify suspicious activity in cloud environments.
- **Exploited misconfigurations:** Our detection and response investigations indicate ensuring adequate serverless security best practices is necessary to help defend against threat actors seeking to exploit misconfigurations.
- **Distribution of malware:** Threat actors are leveraging serverless technology and adjusting tactics in response to previous detection by network defenders.

The Google Cloud Cybersecurity Forecast 2024 report predicted that *"cyber criminals and nation-state cyber operators will more heavily leverage serverless technologies within the cloud because it offers greater scalability, flexibility, and can be deployed using automated tools."*

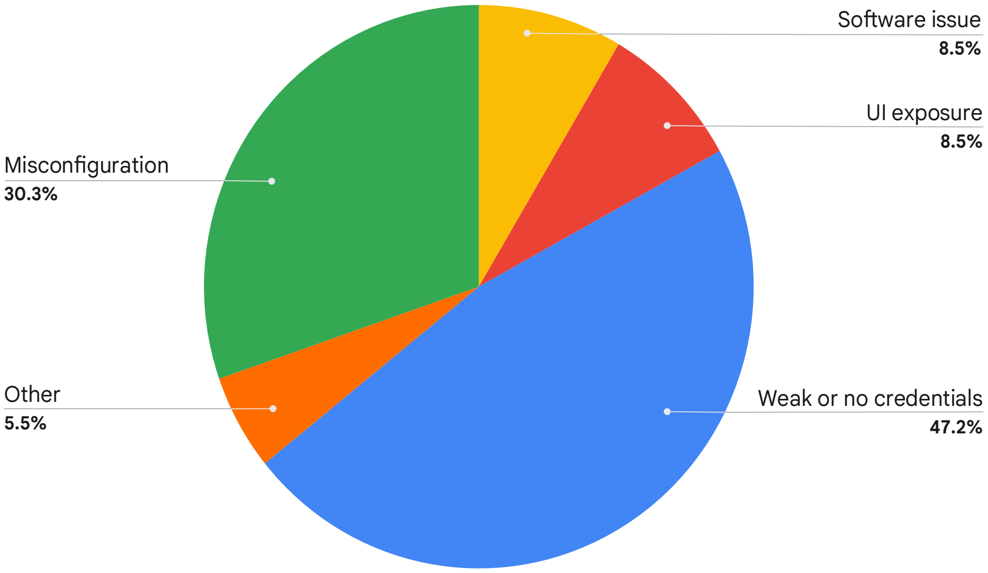
We have seen threat actors live up to that prediction by exploiting serverless computing security hygiene gaps. The following sections dive deeper into the key takeaways from these threats to serverless computing to better enable cloud security defenses.

By The Numbers: Identity Challenges Continue to Pose Risk to Serverless Environments

As part of Google Cloud's continued commitment to security, the Office of the Cloud CISO monitors incident activity and trends associated with how threat actors are gaining unauthorized access to cloud environments and their objectives once inside. This data, along with new insights derived from the Google Security Operations platform (formerly Chronicle), can be found below.

Google Cloud investigated initial access vectors across multiple sources for H1 2024, looking at both successful intrusions into customer environments as well as potential vulnerabilities or gaps found in anonymized Google Security Operations data across a large customer base. This approach allowed us to not only assess how threat actors broke into customer cloud environments in H1, but also determine which areas have the greatest potential for security growth for organizations in H2.

Initial Access Vectors of Concern (H1 2024)



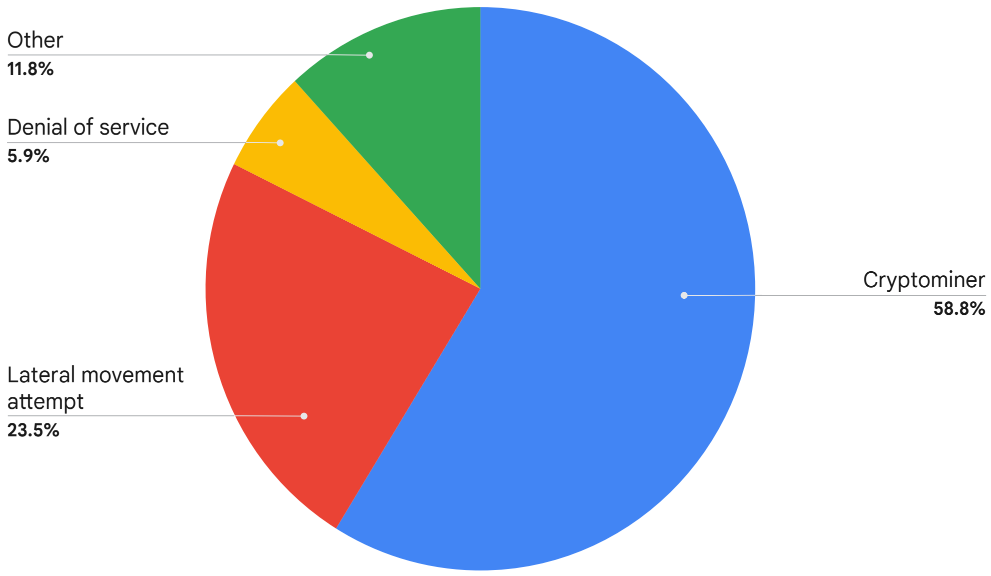
Weak or no credentials remained a key driver of initial access, accounting for the most frequent successful vector and the second most commonly seen trigger for detection rules. Misconfiguration, however, jumped to over 30%, largely due to the high volume of detections of misconfigured or poorly configured environmental factors.

While these misconfigurations were not always exploited by threat actors, they remain an open door for potential malicious activity. An example of a common misconfiguration issue would be service account keys being either overly permissioned or having insufficient preventative controls from malicious use. The risk posed by misconfiguration highlights one key benefit of serverless computing minimizing the configuration oversight required for server maintenance of critical processes.

Additionally, these findings support the relevance of serverless architecture as part of a broader defense in depth strategy, as a preventative control alongside other detective controls in place across the length of a potential intrusion to find and stop attackers at multiple points in the process. The 'Other' category included a host of suspicious detections, such as penetration testing tools successfully infiltrating instances and attempted DNS tunneling efforts.

End goals of intrusions largely remained the same over H1 2024, as nearly 59% of intrusions were motivated by cryptomining efforts, which is slightly lower than our observations from H2 2023 (65%).

Observed Impact of Intrusion (H1 2024)



Mitigations

- Many scenarios that use service account keys can be accomplished with more secure authentication methods that don't rely on downloading and distributing key files. Additionally, Google Cloud uses organizational policy defaults to reduce the risk posed by service account key threats as part of its secure by default architecture. We recommend that you assess and reduce unnecessary service account key usage with the guidance found here.
- Ensure full adoption of multifactor authentication (MFA) for administrative access to serverless web apps as well as other Google Cloud instances.
- Penetration testing is necessary to prevent threat actors from using basic offensive security tools to access your environment
- Leverage Google Security Command Center's (SCC) Event Threat Detection to identify suspicious activity within your organization's cloud environment, such as inappropriate token generation or anomalous geolocation observations. Take advantage of Google SCC's cryptomining protection program for eligible organizations.

Threats to Serverless Functions and Backend Services

Serverless computing offers undeniable advantages, but security must be integrated from the start. By understanding the unique threat landscape and implementing robust mitigations, organizations can leverage the strengths of serverless while protecting applications, data, and their cloud infrastructure.

Throughout the course of incident response and proactive engagements during the last two years, Mandiant has observed a multitude of threats to serverless architecture across all cloud providers. The following threats should be top-of-mind when deploying or operating serverless architecture:

- Hard-coded and clear-text secrets
- Attackers utilizing serverless infrastructure for malicious purposes
- Insecure architecture and development practices
- Misconfigured backend services

Hard-coded and Clear-text Secrets

The practice of embedding secrets, such as API keys and database credentials, directly within serverless function code or environment variables should be avoided at all costs. Unfortunately, this practice remains widespread across all cloud platforms, and clear-text secrets are commonly identified by Mandiant during both incident response and proactive engagements with clients. Some of the main risks include:

- **Exposure:** If your code is ever exposed (leaked repository, misconfigured permissions, compromised hosting environment, etc.), attackers could gain access to the clear-text credentials. In addition, if an attacker is able to gain read-only access to cloud resources, they could access clear-text credentials stored in function code or variables. In both cases, this could allow escalation of privileges within the cloud environment or the ability to move laterally to additional platforms or services.
- **Version Control:** Secrets in code or environment variables are often committed to version control, creating a long-term risk even if the initial exposure is fixed.
- **Credential Rotation:** Hard-coded secrets make it challenging to rotate credentials regularly. Credential rotation helps limit the potential damage if a secret is compromised. However, with hardcoded secrets, rotating credentials would require modifying and redeploying the entire function, introducing operational overhead and increasing the risk of errors.

Mitigations and Best Practices

- **Secret Manager:** Utilize Google Cloud Secret Manager to securely store and manage your secrets. Cloud Run integrates with Secret Manager to allow you to mount secrets as environment variables or files.
- **Never Store Secrets Directly in Environment Variables:** Secrets stored directly in environment variables are not encrypted and can be easily accessed. Cloud Run proactively creates recommendations if it detects environment variables that could be passwords, API keys or Google application credentials.
- **Principle of Least Privilege:** Follow the principle of least privilege by granting your Cloud Functions or Cloud Run services only the permissions they need to access the required resources. This minimizes the potential damage if your code or credentials are compromised.
- **Security Scanning:** Regularly scan your code, dependencies, and cloud resources for potential exposures of secrets and credentials. These scans can be conducted using open source tools such as trufflehog and detect-secrets, or using cloud provider tooling such as Sensitive Data Protection in Security Command Center.

Attackers Utilizing Serverless Infrastructure for Malicious Purposes

Over the past few years, Mandiant has observed threat actors such as UNC2465, UNC4713, and APT41 leveraging serverless infrastructure for malware distribution or Command and Control (C2) communication. Threat actors utilize serverless runtime environments by employing them either as a proxy for traffic destined to an adversary-controlled infrastructure or by directing traffic directly to the compromised machine[1]. This enables threat actors to conceal their malicious traffic more effectively, facilitated by the communications being transmitted to and from subdomains of the cloud provider.

Threat actors have the ability to manipulate functions in such a way that they only accept requests that adhere to specific criteria, such as user-agent, URI paths, headers, or query parameters. In the event that a request does not meet one or more of these requirements, the threat actors have the capacity to redirect the traffic to a benign website or, in the case of an existing function being utilized, allow the function to execute as originally intended. The next article in this report expands on this topic and details on how threat actors are using serverless cloud services to distribute malware.

Mitigations and Best Practices

- Restrict egress traffic from all resources (cloud and on-premise) except where explicitly required. Monitor traffic for communication with unauthorized cloud services. If outbound connection is required, Google Cloud Secure Web Proxy can help to monitor and secure outbound traffic from VMs, containers, and serverless environments.
- Ensure serverless functions and services are behind an API Gateway and Application Load balancer which allows additional security benefit, such as:
  - » **Web Application Firewall (WAF)** integration to filter out malicious traffic based upon common web-based attacks
  - » **Identity Integration or API keys** to control access for authentication and authorization
  - » **HTTPS Enforcement** for all incoming requests to ensure encryption is implemented in transit to and from serverless functions
  - » **Enhanced Logging and Monitoring** to provide detailed logs of API calls, error, track API performance, and anomalies
- Review and remove any unnecessary permissions granted to IAM users or roles that allow them to create, modify, or execute serverless resources. IAM recommender can help to identify and remove excess permissions from principals in Google Cloud.
- Ensure principles of least privilege are implemented for the function or service, refer to the subsequent section for precise guidance.

Read and understand Cloud Run security design, which also applies to Cloud Functions. Note that Cloud Run and Cloud Functions execute by default in an isolated and sandboxed environment.

Insecure Architecture and Development Practices

In a serverless architecture, the code is executed in short-lived containers. This means that there is no persistent infrastructure to attack, making it harder for threat actors to gain a foothold in the cloud environment. However, since the code itself is the core of a serverless function, any vulnerabilities within it can be exploitable. This includes injection flaws (e.g., SQL Injection, XSS), insecure dependencies, and logic errors. The risk is that an attacker can use weaknesses in serverless resources to move laterally to other cloud infrastructure where they can gain a deeper foothold or access data.

For example, an attacker may be able to leverage a vulnerable function to access its service account credentials. Cloud Functions in Google Cloud uses a default service account for function execution, which is provisioned with the editor role. Compromising the service account token would allow the attacker broad permissions across the project, including listing all the cloud storage buckets and retrieving the objects within them.

Mitigations and Best Practices

- **Secure Coding:** Adhere to secure coding principles, use static and dynamic analysis tools, and keep dependencies updated to minimize vulnerabilities. Beyond adhering to principles, leverage OWASP's checklist (e.g., input validation, output encoding, error handling) for specific guidance. In Google Cloud, Artifact Analysis can provide vulnerability information for the container images stored in Artifact Registry.
- **Principle of Least Privilege:** Grant serverless workloads only the permissions absolutely necessary for their operation. In Google Cloud, we recommend creating a unique service account for each serverless resource and granting it the minimum IAM role necessary. Organization policies should be used to prevent automatically granting the Editor role to default service accounts in new projects. This policy is now enforced by default on all new customer organizations.
- **Logging and Detection:** Leverage admin activity audit logs to identify service account usage outside of expected activity. For example, develop detections to alert on usage of a function's service account from unexpected IP ranges or accessing unexpected resources.

Misconfigured Backend Services

Organizations that use serverless Backend as a Service (BaaS) providers entrust them with the storage and management of their application's data. However, misconfigured security measures when implementing the BaaS resource can expose the data to unauthorized access or leaks.

- **Publicly Exposed API Endpoints:** When API endpoints are accessible without adequate authentication or authorization, they become susceptible to exploitation. For example, unauthenticated access to these endpoints enables attackers to probe for vulnerabilities, exfiltrate sensitive data, and manipulate the application's functionality.
- **Insecure APIs:** Even with authentication in place, APIs can remain vulnerable if they fail to adhere to security best practices. For example, insufficient input validation exposes the application to injection attacks, improper error handling can result in information leakage, and inadequate rate limiting facilitates brute-force attacks.
- **Misconfigurations:** BaaS providers offer significant flexibility in configuration, but this can inadvertently lead to misconfigurations that compromise the data's security. For example, overly permissive access controls and misconfigured storage settings can all contribute to data exposure.

Mitigations and Best Practices

- **Automation:** Treat BaaS configurations like software code. Use Infrastructure as Code (IaC) tools to define and manage configurations. This allows you to version control, test, and automate changes, reducing the risk of human error. Prior to deploying resources using IaC, utilize a scanning tool to identify misconfigurations and secrets.
- **Configuration Baselines:** Establish and maintain security configuration baselines for your BaaS platform. These baselines should define secure default settings, access controls, encryption requirements, and other security parameters.
- **Security Review:** Regularly review BaaS configurations to identify and address misconfigurations promptly. Automated configuration scanning tools can significantly streamline the review process. These tools can scan BaaS configurations for common misconfigurations, vulnerabilities, and deviations from security best practices.

Threat Actors Experimenting with Serverless Cloud Services to Distribute Malware

Serverless architectures are attractive to developers and enterprises for their flexibility, cost effectiveness, and ease of use. These same features make serverless computing services for all cloud providers attractive to threat actors, who use them to deliver and communicate with their malware, host and direct users to phishing pages, and to run malware and execute malicious scripts specifically tailored to run in a serverless environment. The security research community has uncovered a wide range of abuse of legitimate serverless infrastructure by malicious actors. This abuse affects all cloud service providers, including Google Cloud, AWS, Azure, CloudFlare, and others.

Google's Threat Analysis Group's (TAG) mission is to track, monitor and counter serious threats against Google and our users. In 2023, TAG detected financially motivated actors abusing Google Cloud's serverless compute products, Cloud Run and Cloud Functions, to distribute malware and host phishing pages.

In response, teams across Google worked together to disrupt the abuse by hunting for malicious instances, updating detections in Safe Browsing, and adding product-level security improvements to prevent future threat activity. As described in the case study below, our intervention reduced one malware campaign by 99% compared to its peak levels.

Google Cloud Run and Cloud Functions are services provided by Google for building and deploying web services. Some threat actors take advantage of the platform's flexibility and ease of deployment, which is intended to ensure users' experience is favorable. The platform's administrative panels provide detailed information about requests and performance metrics. This is a familiar interface to malware distributors as it resembles the Traffic Distribution Systems (TDS) they commonly use to determine campaign success metrics.

Case Studies

Google security teams actively hunt for and disrupt threat activity attempting to use Google Cloud surreptitiously to distribute malware. These case studies from the last year illustrate Google Cloud's proactive approach to detecting and countering abuse of our serverless computing products and highlight our continuous efforts to implement countermeasures that keep users safe and ensure our platforms are secure and trustworthy.

In both cases, financially motivated threat actors used Google Cloud container URLs and legitimate Google Cloud domains such as `cloudfunctions.net` to distribute infostealer malware and host credential phishing pages.

Astaroth infostealer Distribution on Cloud Run and Cloud Functions

Over the years, the distributors of the Astaroth infostealer have abused a wide array of legitimate online services and cloud service providers to distribute their malware to users. These threat actors have experimented with a number of cloud platforms, including Google Cloud, Amazon AWS, Microsoft Azure and others.

Their abuse of serverless computing resources dates back to at least 2019 when security researchers observed them using Cloudflare Workers to create randomized URLs to prevent automated analysis and deliver a malicious payload. Based in Latin America, the distributors of the Astaroth infostealer primarily target users in Brazil, and are well known for their ability to quickly update their malware and distribution techniques to evade detection.

In mid-2023, TAG and Safe Browsing detected Google Cloud abuse by actors we track as PINEAPPLE, who leveraged Cloud Run and Cloud Functions to distribute the Astaroth infostealer. PINEAPPLE used compromised Google Cloud instances and Google Cloud projects they created themselves to create container URLs on legitimate Google Cloud serverless domains such as **`cloudfunctions.net`** and **`run.app`**. The URLs hosted landing pages redirecting targets to malicious infrastructure that dropped Astaroth. If clicked, the Cloud Run and Cloud Function URLs redirected to a Google Cloud storage bucket hosting a ZIP archive that contained a malicious Microsoft Installer (MSI) file.

PINEAPPLE attempted to deliver the malicious URLs in spam campaigns using tax and finance themed lures to convince users to click on the link. The overwhelming majority of these email campaigns were blocked on arrival for Gmail and Workspace users. Over a 14 day period in June 2024, 95% of the emails were blocked. Several of the campaigns masqueraded as Brazil's revenue service, Receita Federal do Brasil, while others impersonated messages from WhatsApp.

PINEAPPLE varied their techniques to convince email gateways their emails were authentic - for example, using mail forwarding services, which do not drop messages with failed SPF records, or placing unexpected data in the SMTP **`Return-Path`** field to trigger a DNS request timeout and cause SPF email authentication checks to fail.

When we detected PINEAPPLE abusing Cloud Run and Cloud Functions, teams across Google worked together to hunt and disrupt related activity. We updated detection signatures and implemented mitigation measures that significantly reduced the volume of the Astaroth campaigns by 99% compared to the campaign's peak.

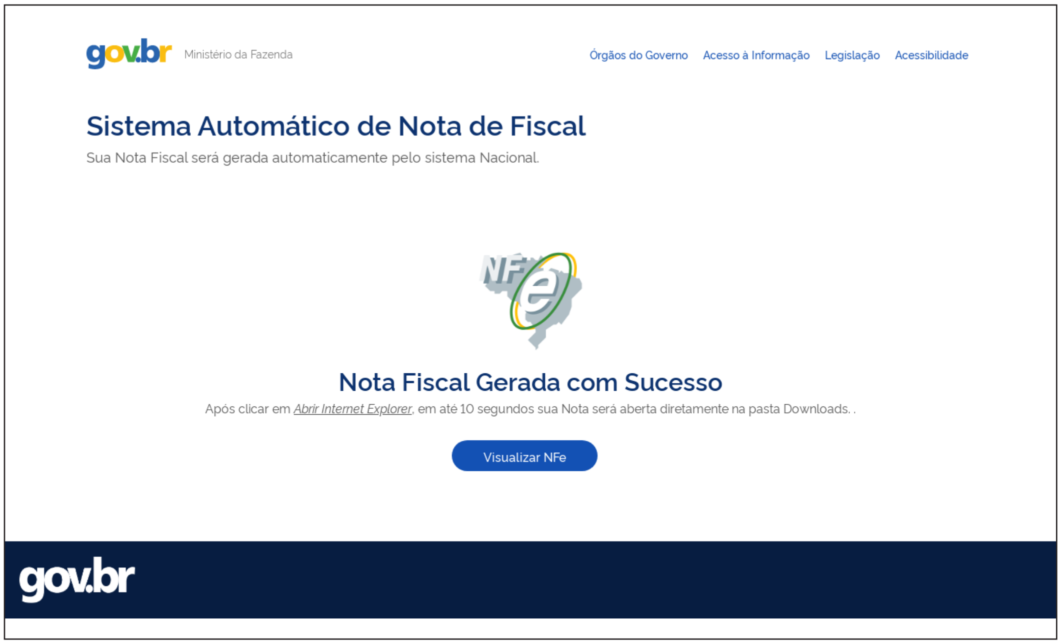
As our teams discovered new abuse attempts, Safe Browsing and TAG updated signatures and created tailored detections to identify and block the campaigns. We also added malicious URLs to the Safe Browsing blocklist. Google disabled the malicious Cloud Run sites and suspended the associated Google Cloud project. We also implemented product level security improvements to significantly increase the difficulty of our platforms being used by this threat actor.

PINEAPPLE reacts quickly and iteratively adapts their tactics, techniques, and procedures (TTPs) in response to new detections. Following Google's disruption of their scaled abuse campaigns, they attempted to continue abusing Cloud Run intermittently at lower volumes.

In one recent campaign blocked by Gmail, PINEAPPLE's spam emails impersonated Brazil's finance ministry and directed recipients to a social engineering page mimicking the Brazilian government's electronic tax document system (Portal da Nota Fiscal Eletrônica). The site directed visitors to click a button to view an electronic tax document generated by the system.

If clicked, the link directed users to an LNK payload hosted on an attacker-controlled IP address. In a likely effort to evade detection, the attackers incorporated multiple legitimate services into the campaign. Links on the social engineering site used the `ms-search://` protocol to direct users to the attackers' IP address, and threat actors hosted their site on Google's Cloud Run. Google disabled the malicious Cloud Run site and suspended the associated Google Cloud project.

In March 2024, PINEAPPLE campaigns temporarily updated their distribution mechanism to use Google Compute Engine (GCE) instances with static public IPs. Similar to their past activity, the campaigns



Sistema Automático de Nota de Fiscal

Sua Nota Fiscal será gerada automaticamente pelo sistema Nacional.

Nota Fiscal Gerada com Sucesso

Após clicar em *Abrir Internet Explorer*, em até 10 segundos sua Nota será aberta diretamente na pasta Downloads. .

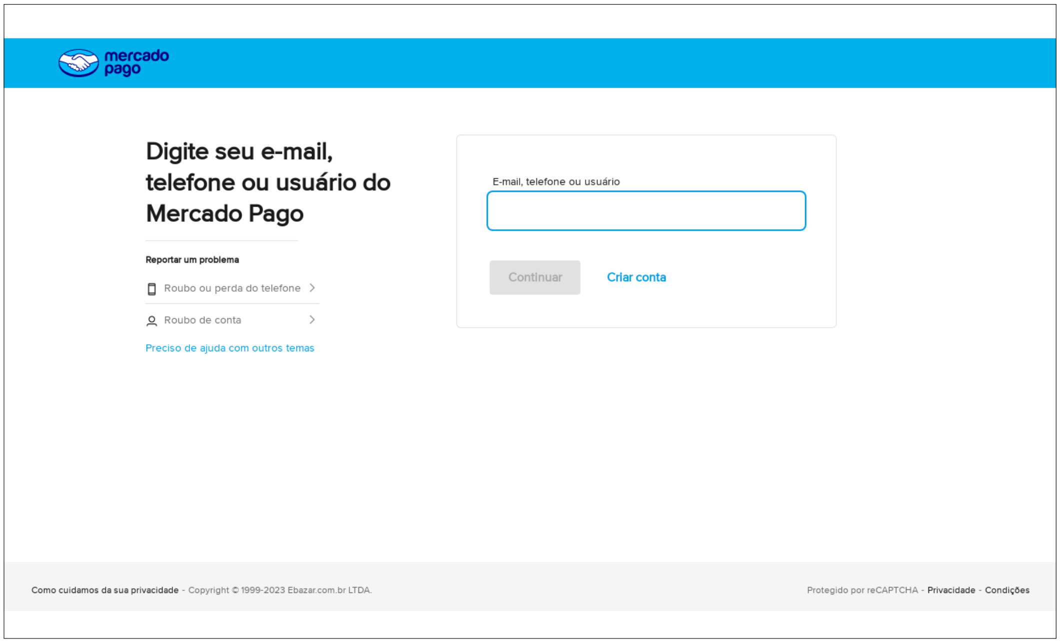
Social engineering page impersonating the Brazilian government's electronic tax document system distributed malicious links via email. The GCE links served an unencrypted archive containing a ZIP or LNK file. PINEAPPLE varied the type of archive they used, including ones they had not used in the past such as `.xz` and `.bz2`. In some cases, the archive contained HTM, HTML, or MSI files instead of an LNK.

Within days of attempting to abuse GCE in their campaigns, PINEAPPLE also experimented with other cloud platforms. In late March 2024, we observed them incorporate Azure Cloud Services and Tencent Cloud into their campaigns.

Shortly after, in campaigns in May and June 2024, they continued sending spam spoofing Brazilian federal agencies. The malicious emails contained links to landing pages on dedicated virtual servers created through GoDaddy's reverse IP hostname service. We continue to monitor their campaigns and regularly update Google's protections to ensure users are protected.

Phishing Serverless Projects

Another Latin America-based financially motivated actor, FLUXROOT, has experimented with Google Cloud containers and tested detection rates for Google Cloud URLs in VirusTotal. FLUXROOT is known publicly for distributing Grandoreiro banking malware. In 2023, TAG identified multiple Google Cloud serverless projects being used to harvest credentials for one of Latin America's largest online payment platforms. Upon discovering the FLUXROOT sites, TAG and Safe Browsing updated detection signatures and added the sites to the Safe Browsing blocklist.



Digite seu e-mail, telefone ou usuário do Mercado Pago

Reportar um problema

E-mail, telefone ou usuário

Como cuidamos da sua privacidade - Copyright © 1999-2023 Ebazar.com.br LTDA.

Protegido por reCAPTCHA - Privacidade - Condições

Credential harvesting page hosted on Google Cloud serverless project

Google Cloud Trust & Safety suspended the associated Google Cloud projects, and updated our detections against similar abuse. More recently, FLUXROOT has continued distributing Grandoreiro, using cloud services such as Azure and Dropbox to serve the malware.

Impact

These case studies point to a growing concern: the abuse of serverless computing for malicious purposes. Threat actors take advantage of the flexibility and ease of deployment of serverless platforms to distribute malware and host phishing pages. Threat actors abusing cloud services shift their tactics in response to defenders' detection and mitigation measures. PINEAPPLE threat actors, for example, have repeatedly evolved their TTPs and experimented with different cloud services in their attempts to evade detection and continue to distribute Astaroth.

Mitigations

Security teams across Google monitor continuously for threats to our users and attempts to abuse our products. Safe Browsing and TAG regularly update detection signatures and add malicious domains and URLs to the Safe Browsing blocklist. Google Cloud Trust & Safety routinely monitors for abuse of Google Cloud services and suspends attacker-operated Google Cloud projects, and Google Cloud's Product Security Engineering team identifies security gaps and mitigations that help drive product-level security improvements that make it increasingly difficult for threat actors to abuse our services.

We also recommend the following approaches for Google Cloud customers to help prevent malware in serverless computing:

- For identities and permissions, closely manage accounts with high privilege and administrator access and apply least privilege principles to ensure each user has the minimum required permissions.
- Incorporate monitoring and controls to detect malware, unwanted software, exploits, and other host-based threats by leveraging Applied Threat Intelligence in Google Security Operations and Google Threat Intelligence. Public and private sector cloud defenders can also collaborate with the U.S. Dept. of Homeland Security, Cybersecurity and Infrastructure Security Agency's Malware Analysis Service.
- Use Workspace alerts for leaked passwords to monitor for compromised credentials, which are often stolen by infostealer malware. Implement a playbook resetting user credentials and checking affected hosts for signs of malware. Mandiant's Digital Threat Monitoring provides additional, advanced protection for monitoring underground marketplaces, paste sites, blogs, forums, and malware repositories to detect unknown data and credentials leaks.
- If using Google Cloud Run, from a back-end services perspective, containerized workload risk mitigations include incorporating Google Security Command Center's Container Threat Detection and refraining from downloading untrusted containers.
- Configure Cloud Functions network settings and Cloud Run network settings to enable control of network ingress and egress to and from individual functions.

Contributors

Cris Brafman Kittner

Charles DeBeck

Kristen Dennesen

Dmitrij Lenz

Crystal Lister

Daniel Medina

Ashik Saji

Will Silverstone

Nader Zaveri

Google Cloud