2024

Global Threat

Year-in-Review

sysdig

Table of Contents

Introduction

In the 2023 Global Cloud Threat Report, the Sysdig Threat Research Team (TRT) detailed how attackers were evolving and leveraging cloud services in their attacks. The team also reported on supply chain security, in addition to an obvious prioritization and targeting of telecommunications and financial sectors operating in the cloud. Most importantly, the 2023 report proved that cloud attacks often happen in 10 minutes or less.

Cybersecurity is an ever-evolving landscape and in the cyber arms race, attackers are relentless. Sysdig TRT observed significant evolutions in cyberattacks throughout 2024, characterized by the increasingly frequent use of automation, the growing scale of attacks, and the continued shift toward resource-based motivations. Harnessing automation, open source tools, and cutting-edge technologies, this year attackers have focused on maximizing profits and leveraging cloud services for both resource exploitation and novel attack vectors. The fast-paced nature of cloud environments, though, remains a central theme of Sysdig TRT's findings.

Many of the attacks Sysdig TRT captured this year were motivated by income generation and free access to otherwise expensive resources. Threat actors continue to use or sell costly resources, such as LLM access or cloud accounts, for profit. Attacks in the cloud have increased 154% year-over-year according to Check Point which, due to the ease of scalability in the cloud, increases the opportunity to make a big profit.

2024 was a year of financial strain. Not only are breaches becoming more expensive year over year, the costs that attackers put on victims are astronomical due to the scalability and speed of modern attacks.

The following report explores what the Sysdig TRT took away from 2024's most novel threat actors and attack campaigns.

Attacks happen in < 10 minutes

Learn about outpacing cloud attacks with the 555 Benchmark for Cloud Detection and Response

READ NOW →

From the exploitation of large language models (LLMs) to the increased use of automation, growing scale of attacks, and weaponization of open source tools, the 2024 edition of the Sysdig Threat Report delves into the expanding attack surface and financial strain that cloud-based organizations face. While some attacks, such as ransomware, deliver swift, high-impact financial blows, many attackers operate with more subtlety, siphoning resources in a manner that culminates in significant loss over an extended period of time. Using novel threats reported by the Sysdig Threat Research Team (TRT) throughout the year, trends that point back to the need for real-time threat detection and rapid response have emerged.

DRAIN ON RESOURCES

PLAYING THE ODDS

The newest AI drain on cloud resources

First identified by the Sysdig TRT in May 2024, LLMjacking — the theft of cloud accounts to exploit large language models — originally cost victims up to $46,000 per day. In the past six months, however, model advancements and price increases can run victims over $100,000 daily. Mirroring earlier cryptojacking and proxyjacking techniques, LLMJacking presents an even more formidable financial threat and marks a significant evolution in resource-jacking.

TURNING THE TABLES

Using good tools for evil

Sysdig TRT observed the weaponization of multiple open source tools this year, notably SSH-Snake, which facilitated credential theft and expanded attacks across the US, China, and beyond. Less than a month after the release of the open source SSH-Snake tool, the CRYSTALRAY threat group leveraged the newly created pen testing tool to steal more than 1,500 victims' credentials.

USED FOR EVIL

PERSISTENCE

THE LONG GAME

Stealth and persistence

Distributed Denial-of-Service (DDoS) attacks are often overlooked because their impacts are underestimated and they're typically considered to be easy to detect and remediate. RebirthLtd, a DDoS-as-a-Service botnet, has been marketed for video game servers. Despite its seemingly benign focus, RebirthLtd's widespread access poses risks of data exfiltration and espionage if misused — there is no telling what else could be done with its capabilities. On the other hand, RUBYCARP, a more stealthy botnet group, prioritized defense evasion and remained undetected for over a decade before Sysdig TRT's discovery.

RISKY BETS

Gambling on market trends and attacking new targets

Forward-thinking and financially motivated attackers, with eyes on market trends and cryptocurrency projections, were found exploiting legitimate tools and stolen account access to mine Meson Network (MSN) crypto prior to its token unlock a few months later. These attackers made a risky bet; if the worth of MSN went up, the mining campaign would be a jackpot. If not, it would be all risk and no reward.

NEW TARGETS

A new frontier in resource exploitation

LLMjacking

Cloud access and cloud-hosted resources can get really expensive, really fast. For this reason, they are a prime target for attackers looking to gain or sell access for a profit. Sysdig TRT has previously reported on resource-jacking — from free account trials to proxy services and now, to LLMs. Beyond financial motivations, Sysdig TRT has also observed LLMjacking attacks from nations with restricted internet access or sanctions that block access entirely. In these cases, malicious access and stolen resources become a way for these residents to bypass restrictions and connect to global internet resources.

THREAT DETAILS

Large language models (LLMs) are revolutionizing nearly every industry but come at a steep cost. In a first-of-its-kind report, Sysdig TRT exposed the exploitation of LLMs in May 2024 — **LLMjacking**. Once the attacker gained access to a victim environment via stolen cloud credentials, they identified access to cloud-hosted LLMs.

These attacks may be motivated by financial gain, but more importantly, they provided free LLM resource access.

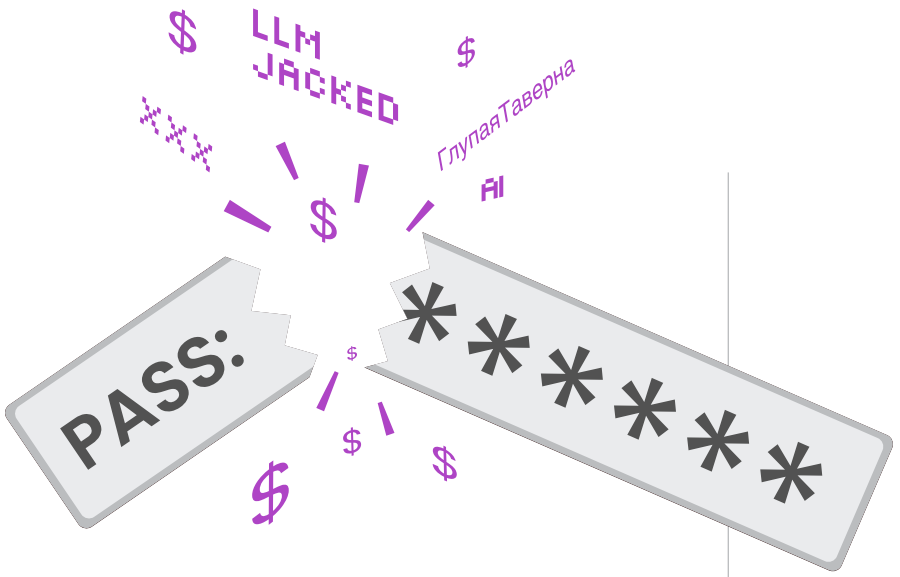
LLMjacking is similar to proxyjacking and freejacking, where the attackers look to gain access to resources that are often otherwise costly. By stealing access, they have free access to the tools or profiles they otherwise may not be able to afford and not have access to due to blocking or sanctions. Some LLM resources, such as Bedrock, Anthropic, and OpenAI, have been sanctioned and are inaccessible in countries like Russia. In addition, some website access is restricted in China and North Korea, which might encourage people in these countries to obtain access through malicious means.

Freejacking

- Sysdig TRT uncovers massive cryptomining operation leveraging GitHub Actions
- Google's Vertex AI platform gets freejacked
- AWS's hidden threat: AMBERSQUID cloud-native cryptojacking operation

Proxyjacking

- Proxyjacking has entered the chat
- LABRAT: Stealthy cryptojacking and proxyjacking campaign targeting GitLab



Resources and tools, like LLMs and proxy services, can be expensive. The stolen enterprise access in the first LLMjacking attack was a local Anthropic Claude 2. x model that could cost victims up to $46,000 per day in consumption costs. These daily costs for the newer Claude 3.5 Opus version could double or triple the daily cost.

Until the attacker's access was removed, the illicit LLM usage never stopped. A single LLM user, for example, can make 500 - 1,000 calls per day, or more. However, in July, under the presumption that LLM access was shared or sold and with the use of automation, Sysdig TRT witnessed a burst of 80,000 calls in three hours. This equates to a bill of approximately $24,000 - $30,000 for the victim — in just **three hours**!

With cryptomining, an increase in CPU resource consumption is easy to identify based on specific behaviors and will trigger an immediate alert. LLM usage, however, cannot be detected this way since there is only one behavior — a call to the LLM. LLM resource consumption will vary greatly across individual users and, therefore, it is difficult to differentiate between legitimate and malicious use.

For example, if AWS Bedrock is not usually seen in your enterprise environment, any usage will be suspicious. If you only use it in one region but then see it in another, that would be suspicious. If usage increases 10x (or some multiple), that could also be suspicious and warrant an investigation. Knowing what is normal behavior and establishing baselines for your enterprise cloud account LLM usage is critical so security teams can readily identify anomalous spikes in usage.

STOLEN ACCESS

An LLMjacking attack could cost victims

$100,000 per day

TOOLS USED

Keychecker
Verifies credentials for use with specified LLMs

OAI reverse proxy
Provides access using the stolen credentials and LLMs while maintaining control of the credentials

Hackers harness OSS with malicious intent

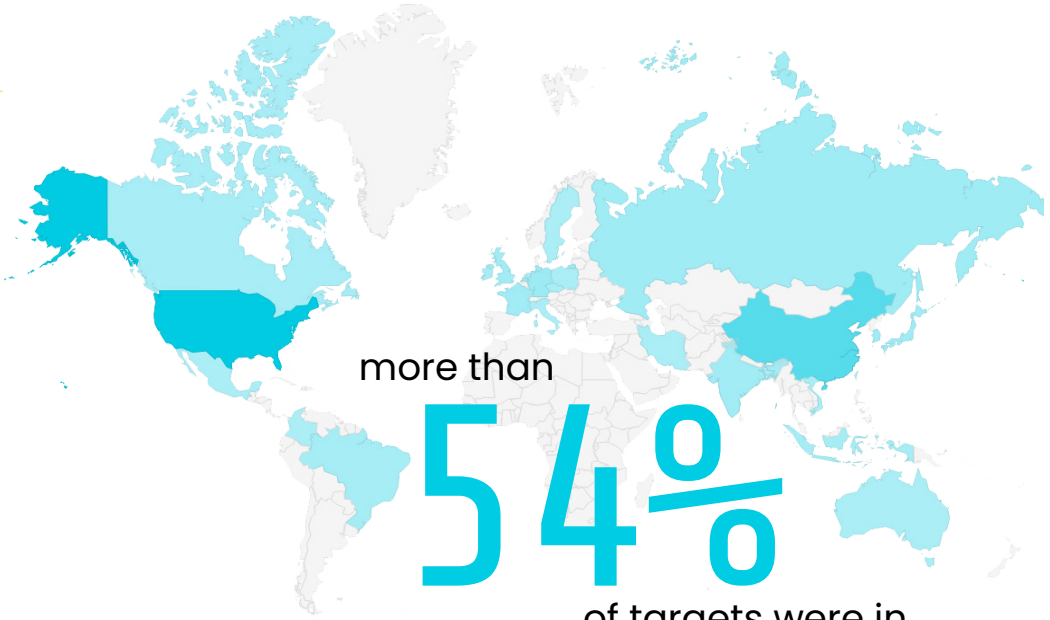
CRYSTALRAY

Open source software (OSS) security tools are free and readily available to anyone in the world with an internet connection. OSS security tools are meant to make a defender's job faster, easier, and more accurate. However, attackers are also using them with malicious intent. A defender may not think twice about seeing an enterprise user download and use a reputable OSS tool — this is the perfect cover for an attacker.

THREAT DETAILS

Sysdig TRT first identified the **CRYSTALRAY** threat actor in February 2024 using the newly developed OSS penetration testing tool SSH-Snake and continued to watch the threat actor enhance their campaign over the following months. Sysdig TRT published an **updated blog** on the actor's entire tool suite in July.

From February to July, the number of CRYSTALRAY victims grew tenfold to over 1,500 victims, with more than 54% of targeted IPs falling in the US and China. The map below visualizes the global breakout, with North America (38%), Asia (29%), and Europe (11%) being targeted most often, likely due to the number of open ports and known vulnerabilities in these regions.



Targeted countries

USA	36.4	France	2.5	Canada	1.2	Mexico	0.6	Vietnam	0.5
China	17.9	India	2.4	Ireland	0.9	Taiwan	0.6	Colombia	0.4
Germany	3.5	South Korea	2.4	Australia	0.8	Czechia	0.6	Italy	0.4
Singapore	3.5	UK	2.1	Iran	0.8	Poland	0.5	Bangladesh	0.3
Russia	2.7	Brazil	1.6	Netherlands	0.7	Sweden	0.5	Other	5.6
Japan	2.5	Indonesia	1.4						

The sophistication of the CRYSTALRAY campaign is a novelty in the collection and sale of credentials on the black market. Through the exploitation of a full OSS tool suite, this threat actor began autonomously scaling their game. Sysdig TRT identified credentials for sale — potentially related to this campaign — that were listed for $20 each. The victim impact, however, likely goes much further. There is no telling what else might happen to victims following the sale of their credentials.

Likely due to Sysdig TRT's public identification of the CRYSTALRAY threat actor, the group shut down its operations following the publication of the blog in July. As of this report, Sysdig TRT has not identified a new CRYSTALRAY campaign. Like CRYSTALRAY, some threat groups are intimidated by the threat of public spectacle and the potential for legal reprimand. In threat research, exposure is one of the greatest tools in curbing further malicious action.

The sophistication of the CRYSTALRAY campaign is a novelty in the collection and sale of credentials on the black market.

TOOLS USED

pdtm
Manage and maintain the ProjectDiscovery tool suite

SSH-Snake
SSH-based worm to map networks and collect credentials

ASN
Internet-wide network data reconnaissance

zmap
Network reconnaissance and service discovery tool

httpx
High performance HTTP-based scanner for mapping web servers

nuclei
Open source vulnerability scanning tool

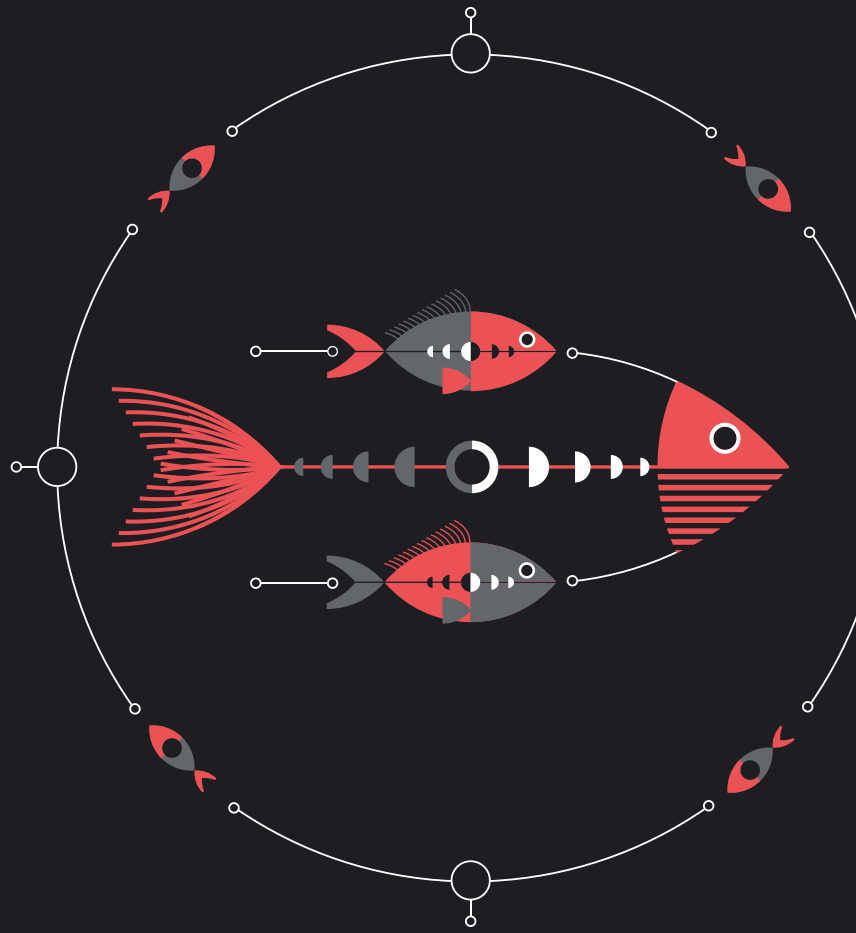
Sliver
Persistence and remote access tool similar to Cobalt Strike

Platypus
Persistence and victim access management

Greater sophistication, sustained success

RUBYCARP

Defense evasion, persistence, and stealth: near-synonyms that lay at the cornerstone of any successful attack. The longer an attacker can hide, the more money and intelligence they have the opportunity to collect. Threat actors that can successfully keep their hubris in check and evolve their tools, stealth, and capabilities can continue making money unscathed for a greater period of time. Stealthy advanced persistent threats (APTs) have been known to lie in wait and collect or exfiltrate data for years at a time. For many attackers, supplementing their income with a few hundred dollars a month is worth the tempered approach.



RUBYCARP

In a shocking discovery in April 2024, Sysdig TRT unveiled a botnet group, dubbed **RUBYCARP**, that had been operating seemingly untouched for at least 10 years. Why has there been no legal recourse against the group in the last decade? Likely, it's simply because open source attribution is very difficult. There are hundreds of other reports about the shellbot perl script Sysdig TRT identified, but nothing else about those behind the tool.

In recent years, however, there have been reports of a **nine-year-old Romanian botnet** that ended in arrest and another **newer Romanian botnet**, only in operation for three years. Perhaps this particular group is just that good at evading detection. After all, the primary domain used to store its cryptomining setup logs was the subdomain physics. uctm[.]edu for the University of Chemical Technology and Metallurgy in Bulgaria.

Rather than build out an extensive suite of known tools for which they can assume defenders have detections, RUBYCARP members reduced their risk of being detected by customizing and regularly updating many of its kill chain necessities over the decade.

Furthermore, the RUBYCARP botnet group targeted several different vulnerabilities to expand its potential victim pool and minimize detection of its operations. In addition to the originally reported targeting of vulnerable Laravel applications (CVE-2021-3129), they also targeted Alibaba Cloud and vulnerable GitLab Community and Enterprise Edition servers (CVE-2021-22205).

RUBYCARP is a successful financially motivated botnet group. One member made $360, or over 1,600 Romanian lei (RON), in just one day. Sysdig TRT also identified a member's wallet that was active for two years and amounted to over $22,800 (100,000 RON) during that time. The team is confident that there are many other wallets with gains such as these, and a little bit now can add up to a lot later — especially considering the cost of living in Romania.

TOOLS USED

Banner
Port scanner tool

Masscan
Common network discovery tool

X kernel module
Used to hide files and processes

Brute
Custom SSH brute force tool

Shellbot
Perl script to join IRC servers

IRC servers
Used for overall team communications and a a working folder to store details for each campaign — to keep them separated

C3Bash
Custom command line miner setup

Operating seemingly untouched for **at least**

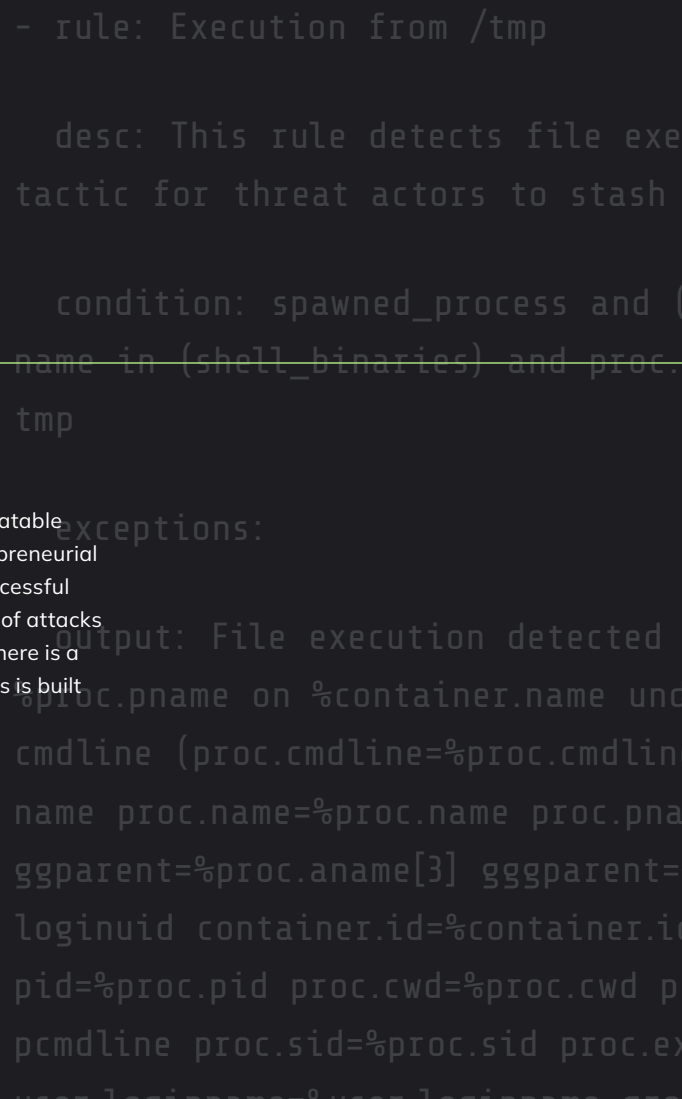
10 years

Why has there been no legal recourse against the group in the last decade?

Attackers at your service

REBIRTH

Any successful enterprise will tell you that a streamlined, repeatable sales motion is the key to success — the same is true for entrepreneurial threat actors. When threat actors can build and automate successful attacks, like DDoS, they can keep delivering a greater number of attacks or services and grow a strong Rolodex of repeat customers. There is a marketing engine behind some attack groups and their success is built upon intentional efforts.



THREAT DETAILS

Distributed Denial-of-Service (DDoS) attacks are often overlooked because they are overhyped and typically easy to remediate. However, although they may not appear to be inflicting pain on your organization once the DDoS is over, the attacker may still have access. These types of attacks are a good distraction from other malicious activities, which could leave victims at risk for data exfiltration, espionage, or worse. Attackers won't necessarily refine their techniques unless they need to, because changing methods can be hard and time-consuming. DDoS is still an effective attack technique, as observed this year when **Microsoft** Azure's global infrastructure was down for several hours.

In March 2024, Sysdig TRT uncovered a DDoS-as-a-Service botnet called **RebirthLtd**, developed from Mirai malware source code. While the group markets its services for DDoSing video game servers, there is nothing to say it can't be used beyond gaming targets.

Once the RebirthLtd botnet is sold, there is no telling what could be done with its access.

BOTNET FOR SALE

Once sold, the access can be used for anything. More importantly, evaluating, understanding, and defending against the RebirthLtd botnet operations will improve defensive security against other, more destructive DDoS attacks.

The team discovered the development of the service offering from its early testing phases in 2019, to the present advertisement and sale starting in January of this year. The group charges $15 - $53 for use of the botnet, with varying levels of access, support, and capability at different costs, providing what Sysdig TRT assumes is a modest supplementary income for the RebirthLtd threat actors since its Telegram marketing efforts began early this year.

Telegram message following the March publication of the RebirthLtd article by Sysdig TRT.

TOOLS USED

Mirai
Malware source code used to build this botnet

QBot and STDBot
Malware trojans used for establishing backdoors

Various exploits
NETIS Router for backdoors, user-agents lists for http attacks, IP scanner for Telnet/SSH, and brute forcing

tcpflood and udpflood
For DDoS capabilities and geoblocking restrictions

Market-driven attacks

Meson Crypto CDN

Money-motivated threat actors are invested in reaping the greatest rewards for their malicious gambles. Cloud processes and automation enhance an attacker's ability to skate to the puck, so to speak. The faster they can work and the easier they're able to scale, the more money they stand to make. Financially motivated attackers are automating their tedious processes to get more bang for their buck.

THREAT DETAILS

Hackers are quick to adopt new technologies and target new opportunities, but in February 2024, **one particular attacker** ramped up their operation in anticipation of a new crypto token unlock — before a price was even placed on the tokens. Meson Network (MSN) is a blockchain project on Web3 that is meant to replace more expensive traditional cloud storage solutions, such as Google Drive or Amazon S3.

Within minutes of obtaining access to the victim's environment, the attacker attempted to create 6,000 nodes using the compromised cloud account. This process was automated, taking approximately 20 seconds to launch each batch of 500 micro-sized EC2 instances per region. With micro-sized nodes, 6,000 nodes could cost the victim $2,000 per day, but with public IP addresses that goes up to $22,000 per day. Unless enterprise victims limit or restrict the number of nodes that can be created, attackers will continue to churn out as many as possible. The victim costs beyond these estimations could be astronomical.

The attacker was taking a gamble on their potential financial gain by preparing to mine a locked cryptocurrency. The attacker's financial outcome for this operation is difficult to compute because MSN coins are not mined based on CPU. Instead, the formula is based on several scores. Additionally, the price of MSN during this attack in February was approximately $4.60, but dropped considerably when the crypto token was launched in May; the MSN price has not been valued over $1 since. In this situation, planning for a future windfall didn't pay off.

TOOLS USED

There were no tools used in this attack, other than the MSN mining software downloaded and run directly from the official **website**.

Hackers could cost victims

$22,000

per day

Takeaways and predictions

Cybersecurity is constantly evolving — and nowhere is that more true than in the cloud. As we've reported over the last two years, threat actors and cybercriminals are just as dedicated to innovation and their operational success as we are to defending against them.

In 2025, we expect attackers will continue to automate tasks to expedite their kill chain for faster, large-scale data exfiltration, intelligence collection, and financial gain. We also anticipate attackers both targeting and using new, innovative technologies and tools for these same reasons, likely increasing the success rate of their attacks. AI is the future — we witnessed the early stages of the AI-cybersecurity struggle in 2024 and there will be a lot more to come in 2025. LLMjacking is only the beginning of the new threats we should expect against AI, and we also know that AI will be used by threat actors to strengthen their attacks over the next 12 months.

1 SCALABILITY

Due to the ease of scaling in cloud environments, the rate of DDoS attacks will increase in 2025. Denial of service causes mass panic at an enterprise scale and can divert eyes while attackers go deeper.

2 ATTACK SURFACE

The attack surface will continue to grow over the next year, especially with the use of LLMs in every sector. Data that was once compartmentalized will continue to be centralized and fed into LLMs in the hopes of higher productivity. Inadvertently, by pushing massive amounts of data into LLMs, we are, in some cases, creating a new concentration risk and increasing the attack surface and opportunity for attackers.

3 AUTOMATION

The rise of LLMs will contribute to the success of attacks. Attackers will use audio and visual clones to successfully target MFA in 2025. On the other hand, we expect an increase in prompt engineering attacks as attackers continue to get the lay of the land and understand the inner workings of LLMs.

4 COST

We expect the enterprise victim cost of attacks to increase. **According to IBM**, the average cost of a breach in 2024 is $4.68M. However, for public cloud breaches, that figure increases to $5.17M. The US alone had over 1,500 **reported breaches** in the first half of 2024. Considering these projections, we expect that global cyberattacks in 2025 will cost over $100B.

> **Preventing attacks is simply insufficient as attackers' means of defense evasion continue to mature.**

Michael Clark
Senior Director of Threat Research

Proactive security programs should always assume compromise. Cyberattacks will continue, likely at a greater frequency, and preventing attacks is simply insufficient as attackers' means of defense evasion continue to mature. Powerful real-time detection and rapid response actions help defenders identify and stop the unknown. Resilience following a cyberattack will keep the business moving.

Cloud attacks will continue to become faster, more sophisticated, and more expensive year over year.

The Sysdig Threat Research Team

The Sysdig Threat Research Team (TRT) is on the cutting edge of threat research with a special focus on cloud and container environments, runtime, and threat detection. They are a well-known group of skilled researchers who first reported the 10-minute time for cloud attacks, set the **555 Benchmark for Cloud Threat Detection and Response**, and uncovered novel threats, such as **PURPLEURCHIN** and **SCARLETEEL**. They have also discovered several large cryptomining campaigns and botnets weaponizing defender tools, such as DockerHub, GitLab, AWS, and more.

Sysdig TRT is responsible for tracking the cloud and container threat landscape, developing and improving Sysdig and Falco detection analytics, and producing and delivering content that educates the security community about their research findings. Since the team was established in 2021, Sysdig TRT has written over 500 detection rules for the open source Falco community. In addition to threat intelligence feeds and customized honeypots, the team uses ML and AI algorithms to refine and improve Sysdig's threat detection rules and capabilities.

Sysdig TRT's highly skilled security experts are dispersed across the globe. The team possesses diverse experiences in governmental, commercial, and academic arenas, and their expertise includes offensive and defensive security operations, computer network operations, malware analysis, and more. You'll catch members of this team regularly at events, big and small, around the world, such as BlackHat, RSA, fwd:cloudsec, and KubeCon.

What's new from TRT?

To stay up to date on the latest cloud threat research, trends, and best practices, visit the Sysdig Threat Research Team's resource center.

LEARN MORE →

Falco

Falco is like a network of security cameras that can be deployed across your distributed infrastructure. It provides real-time detection, monitoring, and observability capabilities for any environment based on runtime security. Falco offers a rich set of out-of-the-box security rules specifically built for Kubernetes, Linux, and the cloud that are open source and community-driven. Since 2021, Sysdig TRT has written over 500 detection rules for the Falco community.

About Sysdig

In the cloud, every second counts. Attacks move at warp speed, and security teams must protect the business without slowing it down. Sysdig stops cloud attacks in real time, instantly detecting changes in risk with runtime insights and open source Falco. We correlate signals across cloud workloads, identities, and services to uncover hidden attack paths and prioritize real risk. From prevention to defense, Sysdig helps enterprises focus on what matters: innovation.

LEARN MORE →

RP-010 REV. A 10/24