Coalition®

coalitioninc.com

Cyber Threat Index 2023

Insights on internet security, cyber risk, and the security trends organizations faced over the last year

Table of Contents

3 Executive Summary

4 Key Takeaways

6 SECTION 1
Vulnerability Trends & Predictions

12 SECTION 2
Internet-exposed services

27 SECTION 3
Industries

35 SECTION 4
Attacker Behaviors

45 Conclusion

45 About Coalition

46 Methodology

Executive Summary

Ransomware groups continued to attack thousands of organizations in 2022. As in prior years, many organizations desperate to retake control of their systems continued to pay these cyber hoodlums large ransoms. Whether ransoms were paid or not, attackers prevented hospitals from treating patients, blocked municipal governments from providing services, and even circulated private records snatched from school districts.

And, of course, Log4Shell, described by some as the single biggest vulnerability ever found, was discovered in late 2021 but caused plenty of trouble in 2022.

Yet the headlines tell only a small part of the digital risk story. Thousands of less-publicized vulnerabilities are identified every month. Coalition, Inc. ("Coalition") predicts that we can expect to see a 13% increase in average monthly critical Common Vulnerabilities and Exposures (CVEs) over 2022; however, we also believe the Common Vulnerability Scoring System (CVSS) to be a flawed system, which led to the creation of our own scoring mechanism.

To help decision-makers in the insurance and security sectors understand the digital risk landscape, Coalition compiled this report based on critical information gathered from our underwriting and claims practices, as well as from internet scans of 5.2 billion IP addresses — an unprecedented data set that comprises the entire IPv4 address space and relevant IPv6 addresses.

We also maintain a global network of honeypots (sensors) to observe attacks from the inside and understand more deeply what techniques attackers use. We combined these scans, sensors, and data sets into this research to provide key insights into the prevalence of vulnerabilities across the internet and how threat actors are exploiting them.

Our report this year covers four key areas:

- What we uncovered in our internet-wide scans
- Common anomalies we detected
- How attack surfaces vary between different industries
- Our analysis of the vulnerabilities disclosed in 2022

Additionally, through our sensors, we discovered variants of existing vulnerabilities, how attackers use these vulnerabilities to carry out cybercrimes, and, most importantly, *exploits yet to be publicly identified*.

What this means:

- If you are a security analyst or cybersecurity practitioner, this report provides valuable context about the threat landscape and the most dangerous vulnerabilities.
- If you are considering cyber insurance or are responsible for a cyber insurance policy, this report will help you assess the risk of various vectors and vulnerabilities.
- If you are a security executive or IT chief, this report will inform your decisions about the role security plays in your operations and infrastructure — and help your team make wiser decisions.

Key Takeaways

Our internet reconnaissance in 2022 yielded information that generates a three-dimensional view of the existing digital risk landscape.

Coalition found enormous volumes of CVEs that security analysts needed to respond to throughout the year. Moreover, each month, thousands of new vulnerabilities were discovered.

Keeping up with every one of those is obviously impossible, which is why cybersecurity executives and practitioners must find ways to prioritize intelligently. This is critical since we know attackers also read CVE reports, and the publication of a CVE can expose organizations to an attack — sometimes within days. Sometimes even sooner.

For example, within a day following the December 10, 2021, publication of Log4Shell, a pitched battle ensued online. As scores of threat actors scurried to exploit the vulnerability, cybersecurity personnel across the globe hurried to create patches and thwart them.

While patching is a continuous effort within organizations, security teams struggle to keep up. In fact, we found that 94% of organizations scanned in the last year have at least one unencrypted service exposed to the internet. Additional key highlights of our report include:

Median Time to Exploit
The time to exploit a CVE varies but is a critical component of digital risk. While CVEs such as CVE-2022-0543, which impacted Redis, garnered exploit-seeking traffic within three days of disclosure, others, like CVE-2022-40684, which impacted Fortinet, did not show up in scans until a month later.

Ransomware Wreaks Havoc
Ransomware remains an enormous problem. Elasticsearch and MongoDB databases have a high rate of compromise, with signals showing that a large number have been captured by ransomware attacks.

Remote Protocols Remain Vulnerable
Remote Desktop Protocol (RDP) remains by far the most common remote-scanning protocol by attackers — and predictably, RDP scanning traffic is very high. This means attackers are still leveraging old protocols with new vulnerabilities like RDP to gain access to systems, which is why quickly patching these is of paramount importance.

Predicting 2023 CVEs
We expect to see more than 1,900 new CVEs per month in 2023, including 270 high-severity and 155 critical-severity CVEs using CVSS.

Cybersecurity leaders and practitioners must be more vigilant than ever to the vulnerabilities that already exist among their networks and assets. They must remain alert to newly published CVEs and respond quickly to close those security gaps. They need an effective way to prioritize the enormous volume of vulnerabilities announced each month.

Ranking the Risk and Urgency of Vulnerabilities

With the overwhelming volume of CVEs, cybersecurity experts need a way to evaluate the risk of each. Traditionally, they do this by estimating the probability that a software vulnerability will be exploited, often using the Exploit Prediction Scoring System (EPSS) and the CVSS. However, when new CVEs are disclosed, important information about them, such as CVSS scores and criticality levels issued by the National Vulnerability Database (NVD), may not be immediately available.

To combat this, Coalition developed the Coalition Exploit Scoring System (CESS), which serves as a CVE prediction and scoring system to help security teams calculate the risk and impact of newly published CVEs. The system features:

1. **CVSS Predictor**: A combination of eight deep learning models that can predict the CVSS score for a vulnerability, given its description.
2. **CESS Exploit Availability Predictor**: Predicts the likelihood of exploit availability in the near future by modeling past exploit availability for CVEs.
3. **CESS Exploit Usage Predictor**: Predicts the likelihood of exploit usage against Coalition policyholders by modeling past attacks.

This system, inspired by EPSS and CVSS, delivers custom-built information to assist cyber insurance underwriting by measuring how likely attackers will actually exploit a CVE.

Powered by machine learning, CESS is an industry standard for assessing the severity of computer vulnerabilities. CESS assigns severity scores to vulnerabilities enabling responders to prioritize responses and resources according to their threat level. The range is 0 to 10, with 10 being the most severe.

Core to the system is the ability to provide security researchers and underwriters with two key pieces of information: the likelihood of exploit availability and the likelihood of exploit usage. We then assign a score based on these two components in a percentile scale, with a 1.0 being 100%.

SCANNING THE INTERNET

At the heart of Coalition's intelligence gathering are its internet scans.

Coalition Control, our proprietary scanning platform, continuously scans the entire IPv4 space, comprising 4.2 billion IPv4 addresses, as well as over 1 billion IPv6 addresses. The scanning infrastructure is globally distributed across multiple countries and providers and collects data from more than 220 ports every 30 days.

This process includes gathering extra data like screenshots for RDP or other types of data enrichment for services running on different ports. For example, when it comes to the Secure Shell protocol (SSH), Control collects all SSH keys, algorithms, and ciphers supported for all SSH servers we find running, versus collecting only the version of the SSH server.

Coalition's data collection also goes beyond internet scans. Our extensive network of honeypots spans multiple locations around the globe. As threat actors conduct their own scans, these sensors appear as enticing targets that lack protection against multiple known vulnerabilities and run outdated software and appliances. Operating these sensors gives us insight into the behaviors of potential attackers and provides important clues about what they are scanning on the internet and the weaknesses they are finding.

While we check more than 5 billion IP addresses, only a subset have any type of service operating on them. To ensure privacy and anonymity, we maintain an ongoing blocklist of people or groups that have requested not to be scanned. However, these represent a small percentage of IP addresses and do not impact our insights. To gain insight and data into your organization's risk landscape, sign up for Coalition Control at control.coalitioninc.com.

SECTION 1

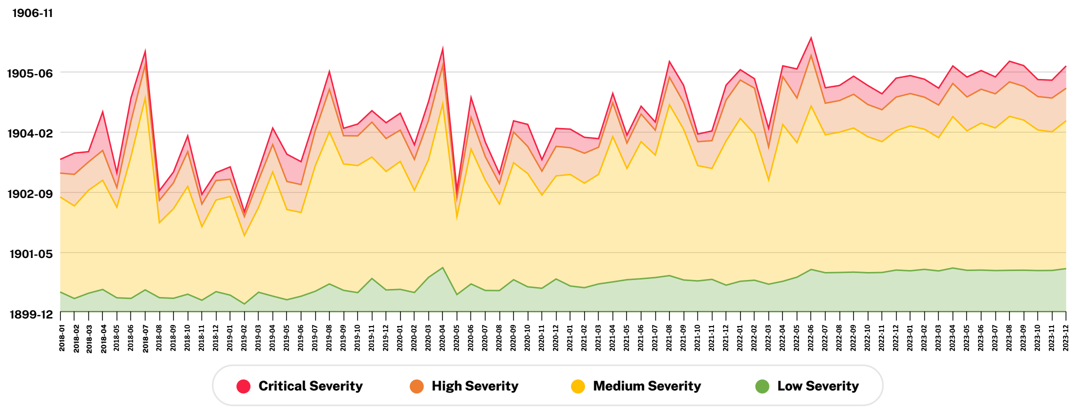
Vulnerability Trends & Predictions

The CVE Evolution

In 2000, there were about 1,000 disclosed vulnerabilities, according to CVEDetails (Figure 1.1). IT and security teams could review and remediate such a low volume without much issue. These systems were also less complex and could be easily siloed compared to today, meaning vulnerabilities had a lesser impact on entire organizations.

CVEDetails also found that the number of disclosed vulnerabilities exploded to over 23,000 in 2022, a 2,200% increase in the last 22 years (Figure 1.1).

Number of Vulnerabilities 2018-2022 *(Figure 1.1)*



As the number of CVEs exploded, we also saw significant growth in digital risk related to CVEs over the last five years. The incredible volume makes tracking increasingly difficult, and new ones are added to the database at an alarming rate.

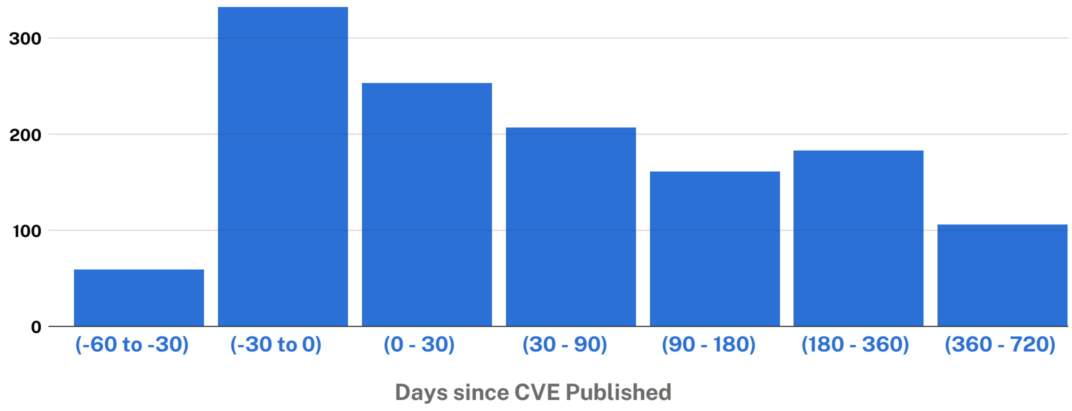
Time to CVE Exploit & Availability

Not all CVEs are exploitable, and there are varying degrees of difficulty in creating exploits for CVEs. Coalition monitors CVE exploit availability using sources such as GitHub (Figure 1.2) and Exploit-DB (Figure 1.3).

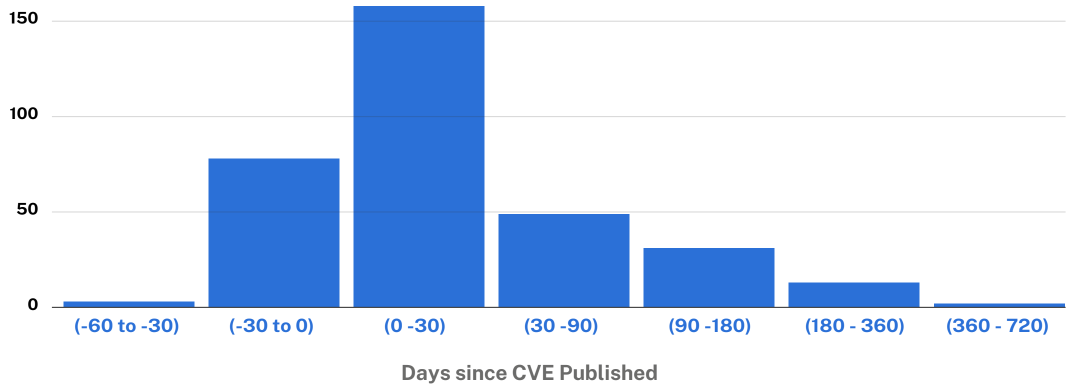
On average, in 2022, verified exploits were published on Exploit-DB after 30 days of CVE disclosure, and we found evidence of potential exploits in GitHub repositories 58 days after disclosure.

The figures below show counts of CVEs that have exploits available, partitioned by the number of days since the CVE was published. We can see that, for most CVEs, the time to exploit is within 90 days of CVE public disclosure, with the majority exploited within the first 30 days.

Number of CVEs with Exploits found on GitHub *(Figure 1.2)*



Number of CVEs with Exploits found on Exploit-DB *(Figure 1.3)*

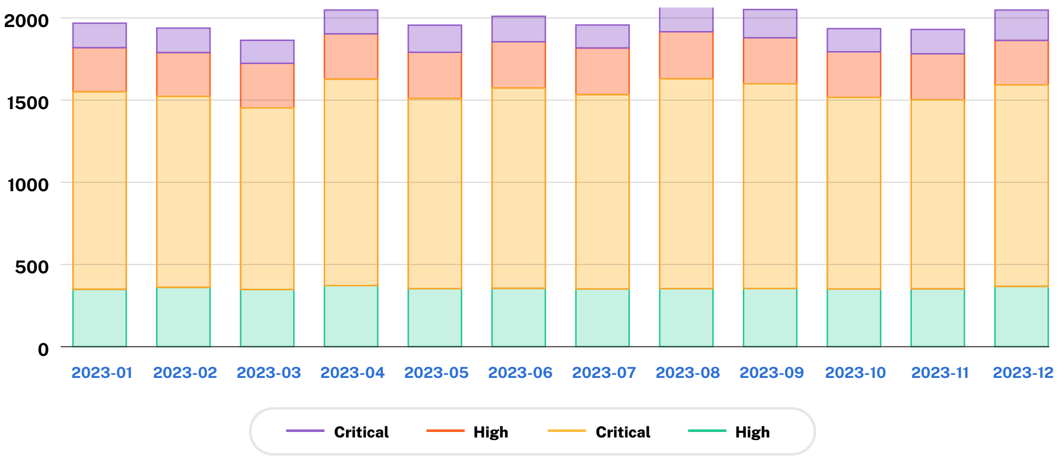


For many CVEs, exploit proof-of-concepts (POCs) were made publicly available before the CVEs were officially "published" by the NVD. This highlights the importance of accelerated patching. Security teams should prioritize applying updates on public-facing infrastructure within 30 days of patch release.

CVE Predictions for 2023

Using CVE disclosure data from the last 10 years, we built a Seasonal ARIMA forecast model to predict the number of CVEs that could be published in 2023 (Figure 1.4). The model predicts that in 2023 we can expect to see more than 1,900 new CVEs per month, including 270 high-severity and 155 critical-severity CVEs, a 13% increase in average monthly critical CVEs from 2022.

CVE 2023 Forecast *(Figure 1.4)*



This can be an overwhelming number of CVEs for IT and security professionals to analyze meaningfully. Therefore, it is essential to understand which CVEs are primed to be exploited and can dramatically impact your business.

Analyzing 2022 CVEs using CESS

We processed all 2022 CVEs through CESS as of November 9, 2022. CESS successfully scored about **17,500** because CVEs with a status of “RESERVED,” “DISPUTED,” or “REJECT,” as defined by the NVD, were not scored.

Of these, **13,000** CVEs had *ATTACK VECTOR = NETWORK*, meaning they were relevant for internet security.

Of these relevant CVEs, **9,500** affected a software *Application*. The remaining were vulnerabilities for either *Hardware* or *Operating Systems*. This data filtering was performed using Common Platform Enumeration (CPE) strings associated with the CVEs.

Below are the CVSS scores of the 9,500 *software application* CVEs (Table 1.1):

Top Vulnerable Products by CVSS Score *(Table 1.1)*

Vendor_Product	Num CVEs	Mean CESS	Mean EPSS	Mean CVSS
cybelsoft_thinvnc	1	0.3536	0.3111	10.0
deno	1	0.3562	0.3212	10.0
linuxfoundation_loopback_connector_postgresql	1	0.4968	0.3718	10.0
microsoft_azure_arc_enabled_kubernetes	1	0.3037	0.3880	10.0
microsoft_azure_stack_edge	1	0.3037	0.3880	10.0

oracle_communications_cloud_native_communication_proxy	1	0.7231	0.5097	10.0
sap_content_server	1	0.3270	0.3625	10.0
squirrel_lang	1	0.5266	0.4851	10.0
vm2_project	1	0.7270	0.3627	10.0
vmware_spring_cloud_gateway	1	0.7231	0.5097	10.0

From the above list, four of the top 10 CVEs, according to their CVSS scores, were either end-user applications or components of a software stack with limited capabilities.

When we consider CVEs that have a high chance of exploit usage using CESS scores, we get a completely different view. Only one entry in this list is an end-user application (Table 1.2); most others were deployed in internet-facing servers and services, thereby providing an attack vector for malicious actors.

Table 1.2

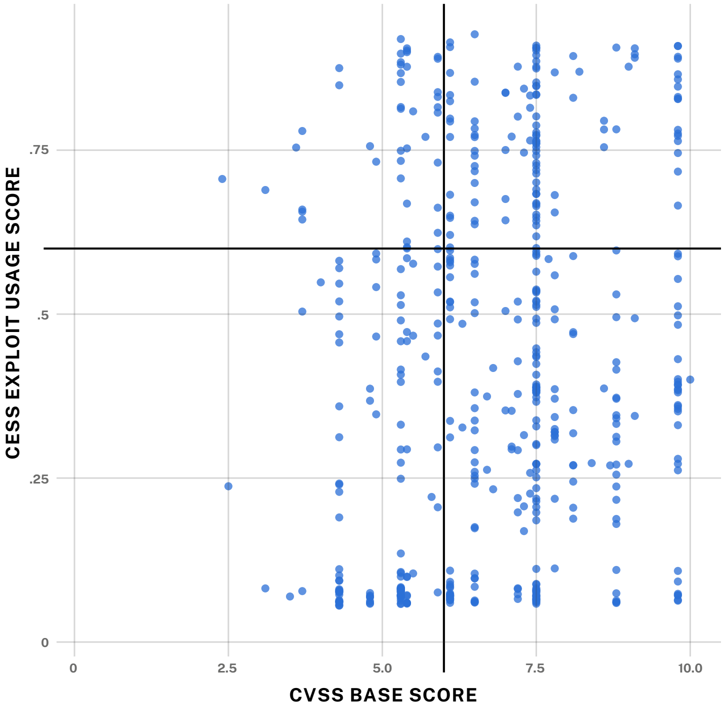
Vendor_Product	Num CVEs	Mean CESS	Mean EPSS
oracle_zfs_storage_appliance_kit	14	0.7866	0.1717
oracle_http_server	13	0.7866	0.1717
apache_http_server	11	0.7866	0.0524
oracle_enterprise_manager_ops_center	3	0.7634	0.3688
netapp_clustered_data_ontap_antivirus_connector	8	0.7622	0.3688
netapp_santricity_smi_s_provider	5	0.7622	0.3688
netapp_storagegrid	8	0.7622	0.3546
openssl	11	0.7622	0.5375
tenable_nessus	14	0.7622	0.3546
netapp_clustered_data_ontap	22	0.7622	0.5331

Oracle tops the list with multiple products that had multiple vulnerabilities this year. In fact, Oracle had 21 vulnerabilities known to be exploited in 2022 alone, per the Cybersecurity and Infrastructure Security Agency (CISA).

By using CVSS Predictor to understand the severity of a CVE and CESS to calculate the likelihood of an exploit, these two systems help security teams decide which CVEs to prioritize in their research.

From the scatter plot of the two scores below (Figure 1.5), CVEs in the top right quadrant have the highest CVSS and CESS scores and represent the greatest digital risk. Based on combining these two scores, these CVEs should be the focus of security teams.

CVSS vs CESS Scores 2020-2022 Scatterplot *(Figure 1.5)*



Exposure for Top CVEs

The table below contains only the top 20 CVEs by CESS Usage Score and the number of exposed targets (Table 1.3).

Table 1.3

CVE	CVSS Score (v3)	CESS Exploit Availability Probability	CESS Exploit Usage Probability	Num Targets
CVE-2022-22719	7.5	0.49351814	0.786557	4677238
CVE-2022-22721	9.1	0.5627909	0.76335865	4677238
CVE-2022-22720	9.8	0.6414925	0.74427074	4677238
CVE-2022-28614	5.3	0.4225313	0.73841906	4677238
CVE-2022-28615	9.1	0.5279814	0.7264665	4677238
CVE-2022-31813	9.8	0.6448038	0.71925247	4677238
CVE-2022-26377	7.5	0.35287416	0.71757317	4677238
CVE-2022-30556	7.5	0.57295173	0.69861144	4677238
CVE-2022-29404	7.5	0.34253964	0.68992347	4677238
CVE-2022-23943	9.8	0.68455976	0.6566089	4677238
CVE-2022-0778	7.5	0.44951636	0.7621613	1971558
CVE-2022-1292	9.8	0.69729155	0.71639705	1163529
CVE-2022-2097	5.3	0.35500753	0.68924874	1163529
CVE-2022-1473	7.5	0.3828603	0.6751895	1163529

CVE-2022-3602	9.8	0.6993933	0.6241492	1163529
CVE-2022-2068	9.8	0.7449148	0.6173774	1163529
CVE-2022-30522	7.5	0.34253964	0.68992347	88433
CVE-2022-0391	7.5	0.38132235	0.653848	6319
CVE-2022-2048	7.5	0.36651352	0.61484987	2085

The table below explains what vulnerable configurations (CPEs) contributed to these numbers (Table 1.4):

Table 1.4

CPE	CVEs	num_targets
cpe:/a:apache:http_server	['CVE-2022-22719' 'CVE-2022-22720' 'CVE-2022-22721' 'CVE-2022-23943' ‘CVE-2022-26377’ ‘CVE-2022-28614’ ‘CVE-2022-28615’ ‘CVE-2022-29404’ ‘CVE-2022-30556’ ‘CVE-2022-31813’]	4677238
cpe:/a:openssl:openssl	[‘CVE-2022-0778’ ‘CVE-2022-1292’ ‘CVE-2022-1473’ ‘CVE-2022-2068’ ‘CVE-2022-2097’ ‘CVE-2022-3602’]	1163529
cpe:/a:nodejs:node.js	['CVE-2022-0778']	442910
cpe:/a:mariadb:mariadb	['CVE-2022-0778']	365085
cpe:/a:apache:http_server:2.4.53	['CVE-2022-30522']	88433
cpe:/a:python:python	['CVE-2022-0391']	6319
cpe:/a:eclipse:jetty	['CVE-2022-2048']	2085

Looking at the vulnerable configurations above, the top 10 CVEs with very high CESS usage scores impact Apache HTTP Server, followed by OpenSSL and NodeJS.

SECTION 2

Internet-Exposed Services

In this section, we'll take an in-depth look at multiple services and ports using data collected from our internet-wide scans. We focused this analysis on the main services used and the types of security configurations and protocols used by said services.

Web Services

Web services are some of the most commonly used technologies on the internet. These services are some of the most important pieces of software on the internet, serving web applications and websites to users.

HTTP vs. HTTPS

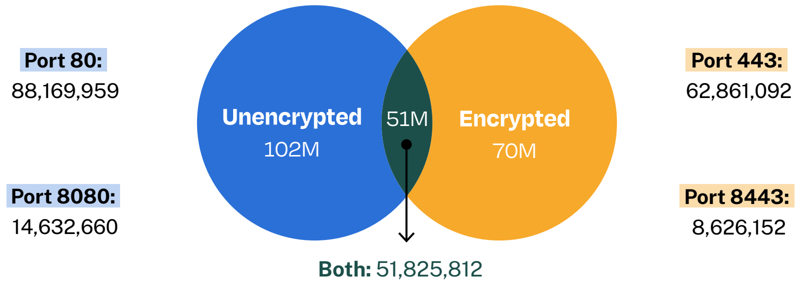
HTTP (HyperText Transfer Protocol) offers a set of rules and standards that govern how information can be transmitted on the World Wide Web (WWW). HTTP is a plaintext protocol susceptible to Man-in-the-Middle attacks where a malicious actor can intercept and view all communication between a server and a user.

HTTPS is an extension of the HTTP protocol that aims to secure connections with a set of cryptographic keys to encrypt and validate data. HTTPS has become the de facto standard for websites. In addition to providing security, it also establishes trust between a website and its users. The most common way for websites to use HTTPS and have a secure connection is by obtaining a Secure Sockets Layer (SSL) or Transport Layer Security (TLS) certificate. TLS is preferred over SSL as it provides additional checks to ensure message integrity.

Our first question when looking at our internet scanning data is: how many IP addresses are running on HTTP vs. HTTPS?

We focused on the main ports, specifically port 80, 8080 for HTTP and 443, 8443 for HTTPS (Figure 2.1). It is important to ask this question because redirecting unencrypted services to encrypted ones is one of the more straightforward security controls to implement. So, understanding if organizations are applying the basics gives us a preview of how the rest of the services might be configured.

IP Addresses Running HTTP or HTTPS Based Services *(Figure 2.1)*



There still exists a higher number of servers that only have port 80 (88 million) running compared to 443 (62 million). However, we remain optimistic as the gap between 80 and 443 has been slowly closing over time. We found that close to 12.3 million IP addresses enforce redirects from an unencrypted port to one configured with TLS.

We strongly recommend switching to HTTPS using TLS and implementing HTTP to HTTPS redirection for services that end users connect to. Refer to Coalition's Help Center for steps outlining how to make this upgrade.

Top Web Servers Used Today

The following table shows the top web servers used (Table 2.1).

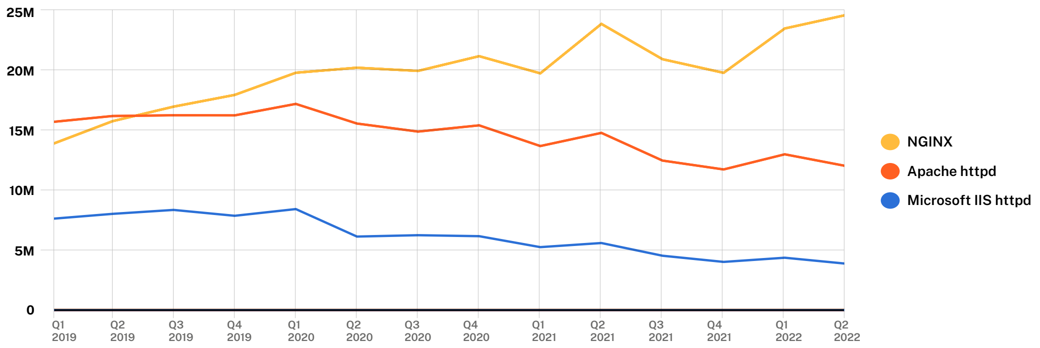
Table 2.1

APP NAME		TARGET IP
1	Nginx	20,936,101
2	Apache	15,250,552
3	IIS	4,862,188
4	Microsoft HTTPAPI	2,468,111
5	lighttpd	1,388,629
6	Express	1,087,213
7	OpenResty	673,931
8	Kestrel	477,953
9	LiteSpeed	472,964
10	Tengine	419,412

NGINX and Apache are still the most widely used web servers, followed by Microsoft Internet Information Services (IIS).

NGINX and Apache are the leading web servers due to their reliability, ability to scale, ability to handle high volumes of traffic, and because they are both open-source. NGINX is designed to handle 10,000 simultaneous client connections (C10K) by using asynchronous event-driven architecture. Data from W3Techs shows similar results, with NGINX at 34% and Apache at 32%.

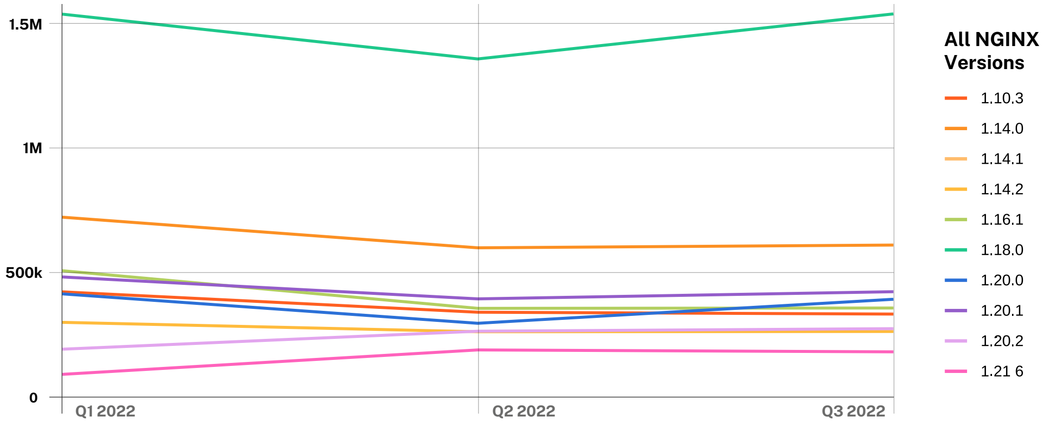
Top Web Servers 2019-2022 *(Figure 2.2)*



This has been the state of web server usage since the middle of the second quarter in 2019, when NGINX overtook Apache as the most popular web server (Figure 2.2).

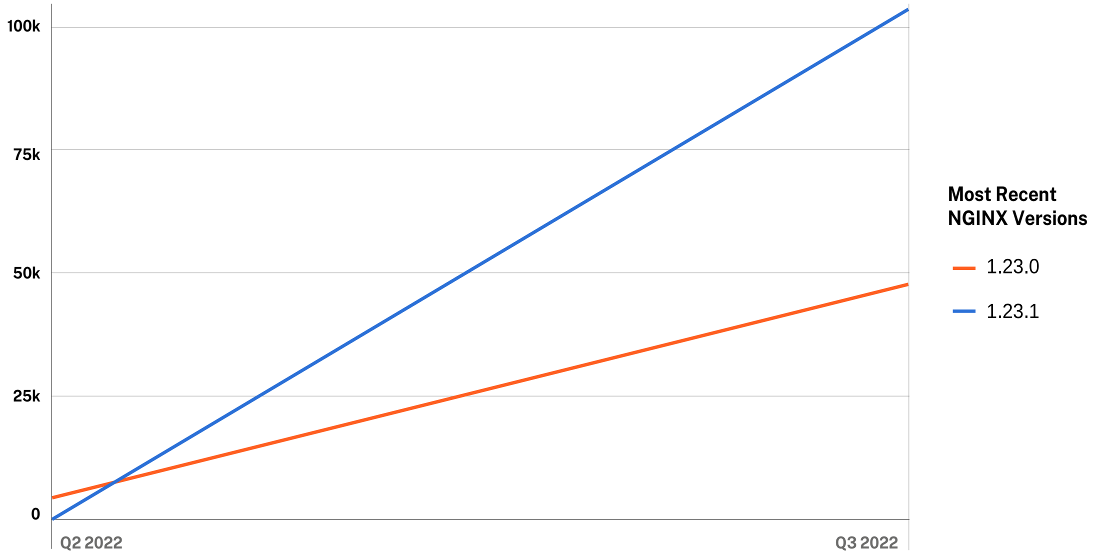
On NGINX, the dominant version on the internet is 1.18.0, which was released on April 21, 2020 (Figure 2.3). This version is over two years old and four updates behind the latest release. While it is not uncommon to run older versions, as evident from the graph, we strongly urge businesses to follow regular upgrade cycles and patch cadence to mitigate any security vulnerabilities in older software.

Figure 2.3



The most updated version observed at the time of this writing was 1.23.0 to 1.23.2; however, the volume of installations for these versions is far smaller, which is normal since they are newer versions (Figure 2.4).

Figure 2.4

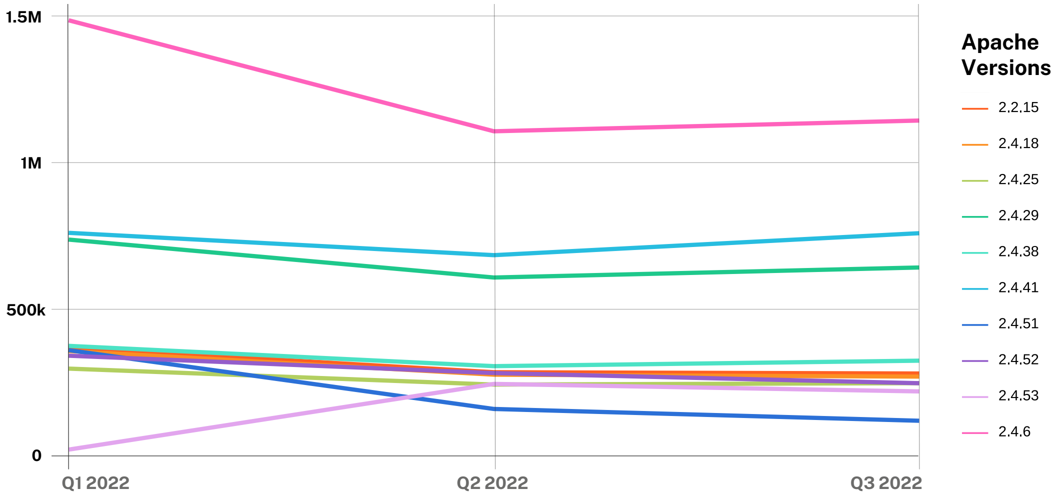


For Apache, the top observed version is 2.4.6, which was released on July 22, 2013 (Figure 2.5). It is concerning to see this 9-year-old product version being actively used today. Multiple bypass vulnerabilities have been found for this version, including:

1. CVE-2013-5704
2. CVE-2014-8109
3. CVE-2015-3185

All of these vulnerabilities result in the bypass of security controls or restrictions, providing attackers with a set of possibilities for potential entry points to systems.

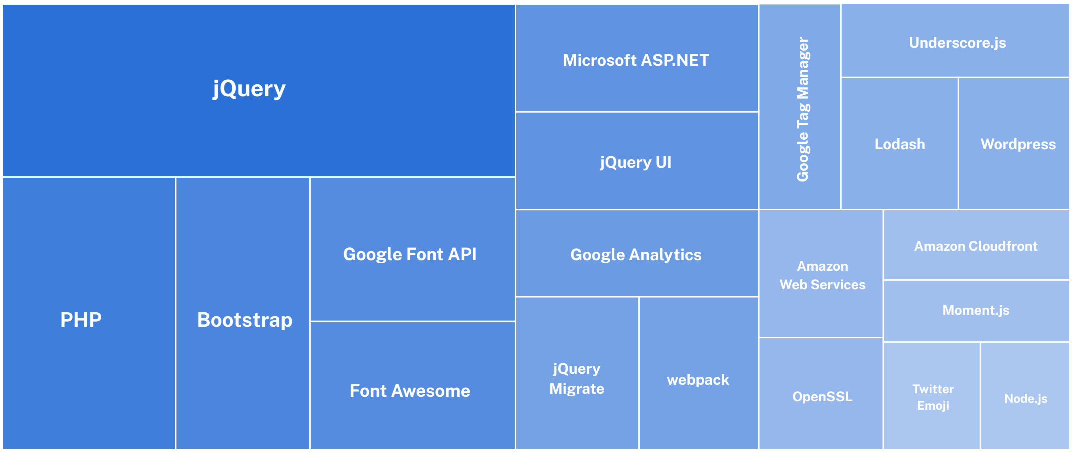
Figure 2.5



Top Web Application Technologies

The web server is only one piece of the technology stack for serving web applications (Figure 2.6). What about web technologies being used to develop web applications?

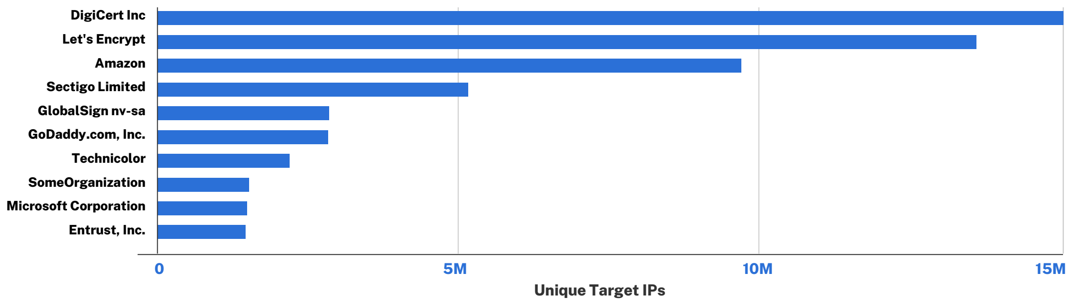
Figure 2.6



jQuery is a widely deployed JavaScript library that makes using JavaScript on your website easier. jQuery takes a lot of common tasks that require many lines of JavaScript code to accomplish and wraps them into methods that you can call with a single line of code. It is implemented by 77% of the 10 million most popular websites. PHP is second, followed by Bootstrap.

SSL certificates used in HTTPS services also have interesting characteristics worth looking into. The top certificate issuer is still DigiCert (Figure 2.7). Remember that our scanning is based on IP addresses. We believe if we observed via domain, this might look different, and Let's Encrypt might take the first spot.

Top SSL Certificate Issuers *(Figure 2.7)*



Overall, the takeaway from the web server data is that organizations are not only choosing to use unencrypted services in a higher volume than encrypted, but they also seem to be struggling with patching their web servers. From our internet-wide scan data, we estimate that 94% of organizations scanned in the last year have at least one unencrypted service exposed to the internet.

▷ Did you know?

Patch cadence is a signal used when evaluating risk for cyber-insurance, but you should consider the possibility of always having your software up to date and running on the latest stable versions.

Telnet/SSH

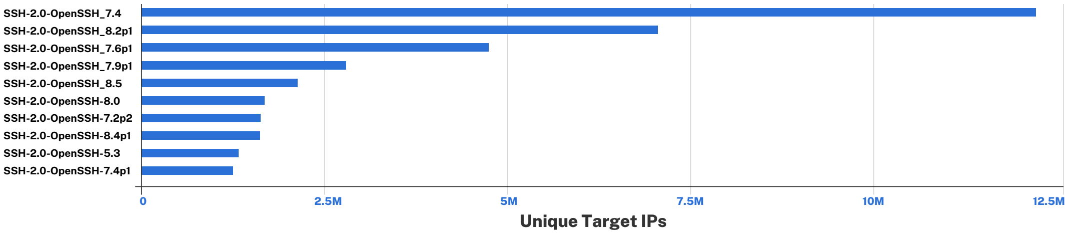
SSH and Telnet (teletype network) are both network protocols; the primary difference between them is that SSH provides users with an encrypted connection, while the data transmitted through Telnet is clear-text.

It is important to note that Telnet suffers from the same problem as HTTP: if you connect with Telnet while on a shared wireless network with someone else, all the credentials will be visible to hackers in that network. SSH, however, provides a secure, encrypted channel between a client and a server.

Commonly used to manage remote machines and transfer files in a secure form, SSH is usually found on Transmission Control Protocol (TCP) port 22, while Telnet is found on port 23. Our study will analyze this service's use and exposure and also extract other data, such as the encryption algorithms and keys used.

In 2022, we found **5,512,507 IP** addresses running a Telnet service. However, unlike in the web server world, we found **47,383,594 IP** addresses running SSH services, which is much higher than its insecure counterpart.

SSH Versions *(Figure 2.8)*

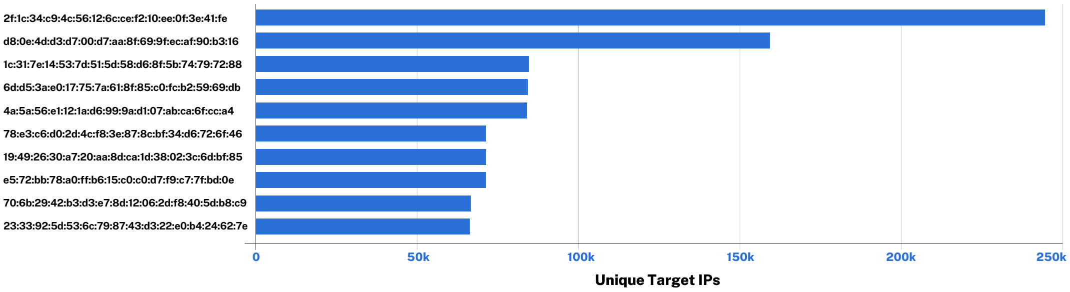


Simply saying SSH is "the secure version" of Telnet would be misleading. SSH requires software that often needs patching and must be configured correctly to be secure. Therefore, ensuring that you are using the right SSH services is important.

For example, the most-used versions of SSH as seen in Figure 2.8 are quite robust: OpenSSH 7.4 was released on December 19, 2016, followed by 8.2p1 on February 14, 2020. Fortunately, no major vulnerabilities have been found for these versions. SSH is a battle-tested protocol, hence why it is also widely used and the standard for remote management and file transfer on Unix servers.

SSH has a few different properties that are interesting to explore further. We will start by looking at the ciphers (algorithms that perform data encryption) used by SSH and the fingerprints (unique properties) of those ciphers.

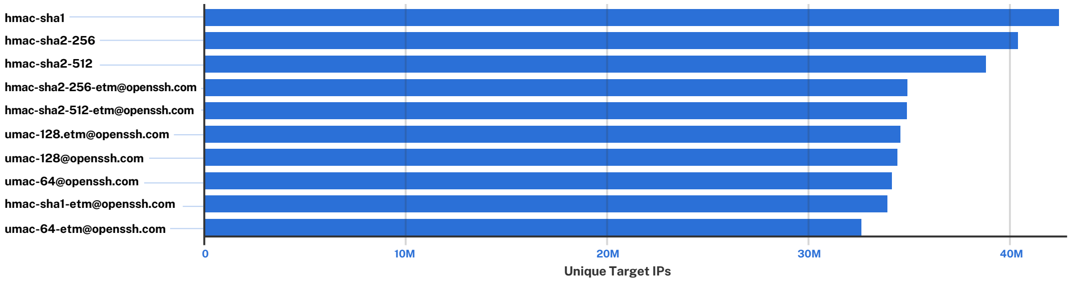
Top SSH Ciphers *(Figure 2.9)*



We can see that the first two fingerprints have a high volume of IP addresses as seen in Figure 2.9. Looking into the providers where these IP addresses were detected, we see they are mostly internet service providers (ISPs) geographically distributed in multiple locations. This means these devices may be provided by the ISP, like a router or modem that is white-labeled and resold.

Moving to another property in SSH, Message Authentication Code (MAC) algorithms are used to verify packet integrity. Looking at the top MAC usage, we see there is still a strong prevalence of hmac-sha1, which uses the SHA-1 algorithm that has been deprecated by the National Institute of Standards and Technology (NIST) (Figure 2.10). Although it is still secure to use for HMAC as it does not depend on the underlying hash function being resistant to collisions, it is best practice to always update to a stronger and better algorithm.

Top MAC Usage *(Figure 2.10)*

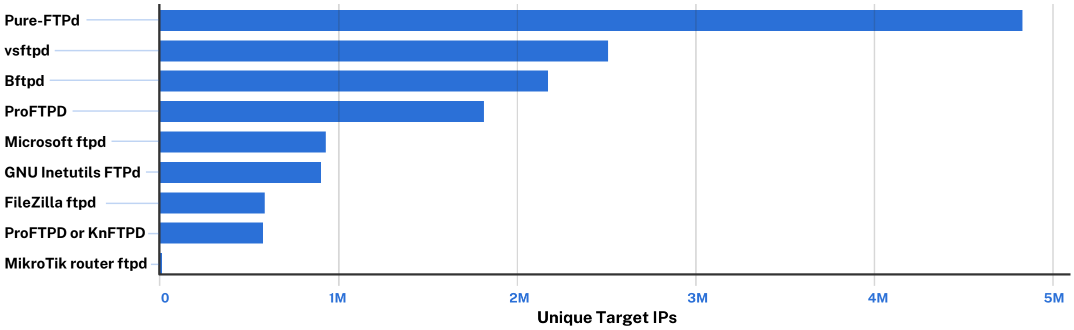


Although many IP addresses (SSH) choose the most secure communication service, we can still see a lack of correct security configurations being enforced. State-backed actors have the resources and know-how to break encryption and have evolved rapidly during the last few years. That is why system administrators must deploy the most secure configurations possible.

FTP/FTPS

The File Transfer Protocol (FTP) is a network protocol used to transfer files between a client and a server. FTPS is the same as FTP but with a TLS/SSL tunnel. FTP is still a widely used protocol for file transfer. In total, we found **19,522,191** IP addresses with either FTP or FTPs service running. Out of the **19,469,543** FTP servers we found running on the internet, ProFTPD is the most used by a wide margin (Figure 2.11).

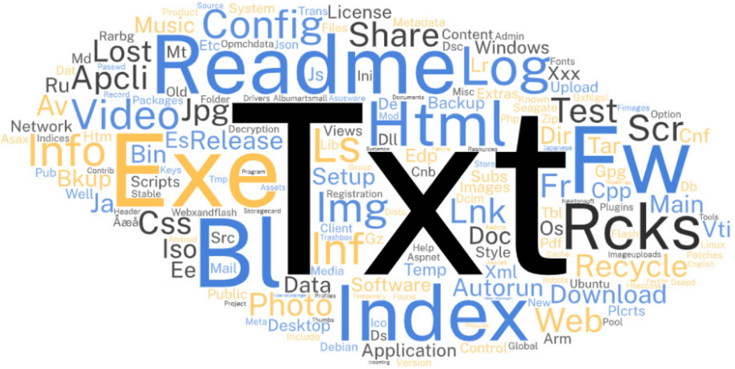
Top FTP Server Software *(Figure 2.11)*



One interesting point regarding FTP is that there is still a considerable number of IP addresses that allow anonymous (default username/password per Request for Proposal (RFP) access to their content. We found a total of **349,015** servers exposing this type of access and with different types of content ranging from network storage appliances to companies that had their documents exposed (clearly by mistake).

Based on the most common words in the filenames, we generated a word cloud to show some of the most common keywords to understand what type of data is exposed most commonly (Figure 2.12).

Figure 2.12



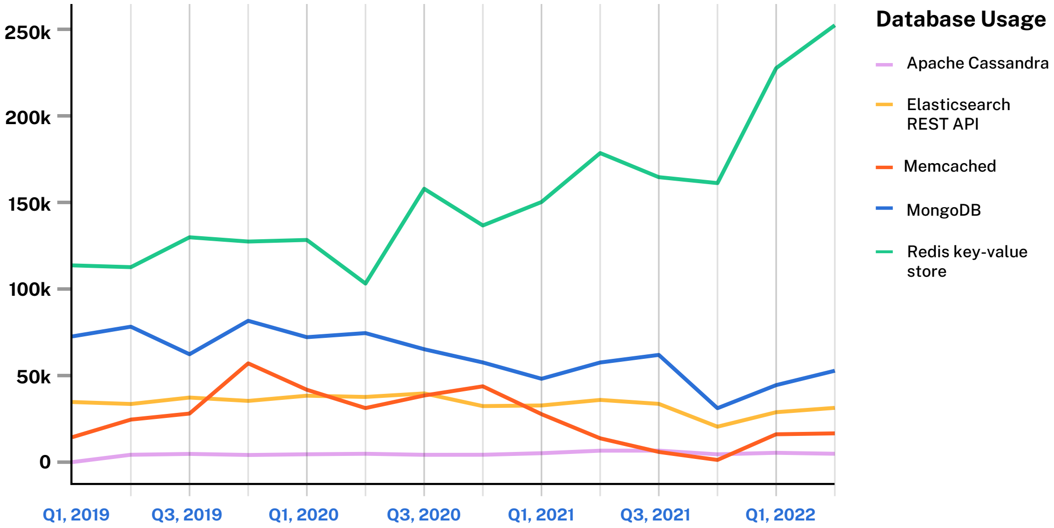
We can see some interesting words that reveal potential exposures or data leaks, such as "Config," which typically refers to folders with configuration files that may have usernames and passwords. "Backups" is typically data with access restricted to a select group.

Data Storage

Databases exposed to the internet are the origin of many of today's data leaks. We will focus on a subset of specific technologies: Elasticsearch, Redis, MongoDB, Memcached, and Apache Cassandra for unauthenticated databases. Then we will look at PostgreSQL, MySQL, and Microsoft SQL, which have authentication by default.

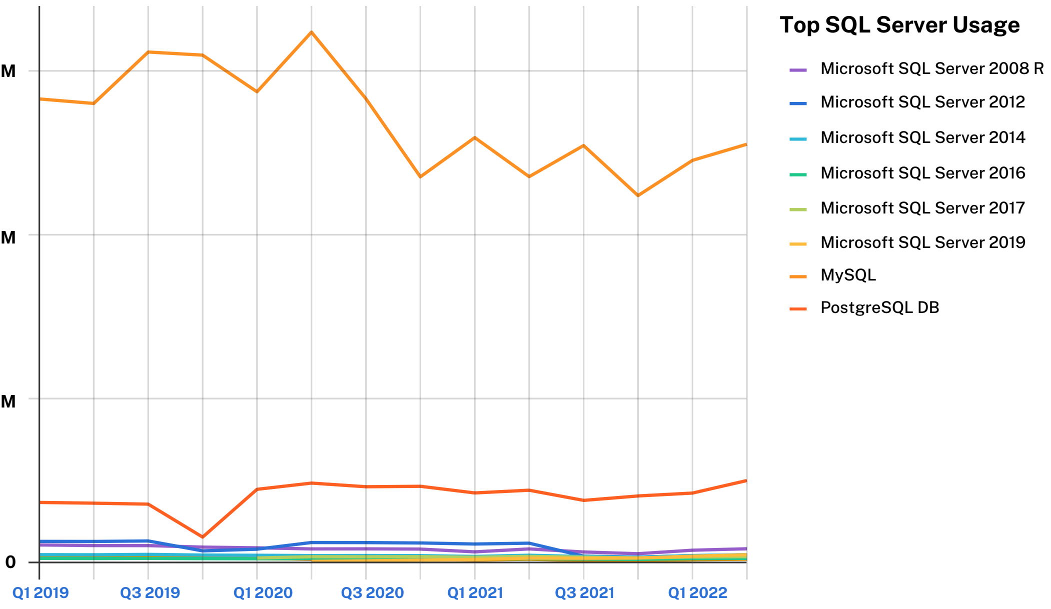
Looking at the below image, we can observe that these databases are still widely used and have also grown in use over time (Figure 2.13) . Specifically, Redis has shot up from around **110,000** instances in 2019 to **250,000** in 2022.

Figure 2.13



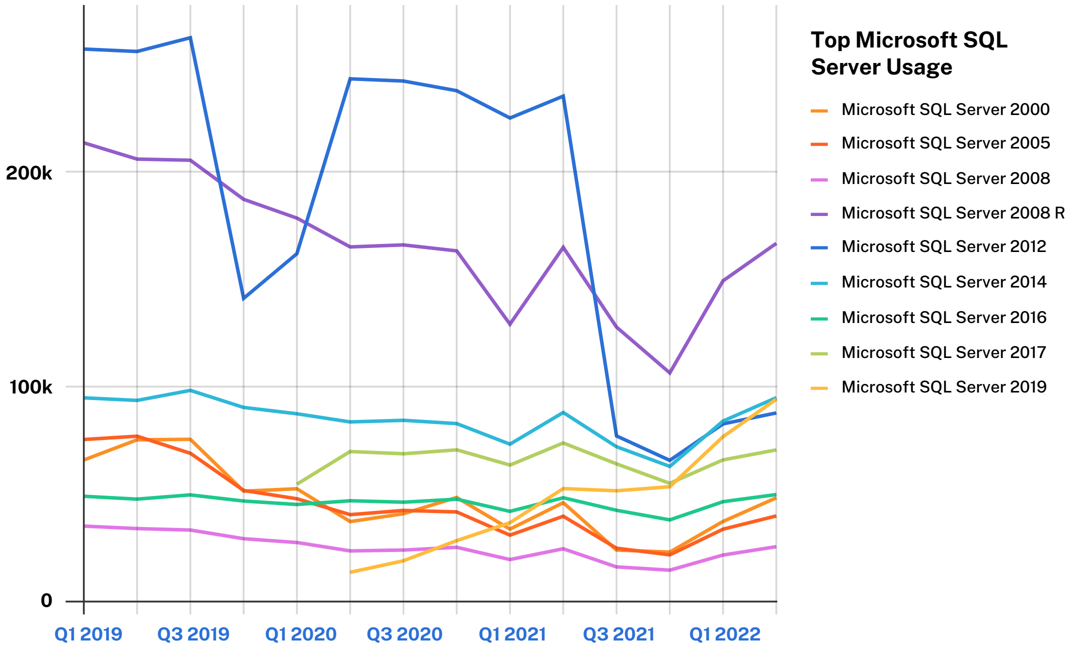
As for authenticated databases, MySQL is the longtime leader.

Figure 2.14



We can see, however, that there are still plenty of old installations of Microsoft SQL. If we drill into just Microsoft SQL, we see patching cycles where certain versions disappear and newer versions appear, like in Q2 2021 (Figure 2.14) when Microsoft SQL 2019 started to overtake older versions. However, it is also clear that Microsoft SQL 2008 R2 is extremely prevalent despite being end-of-life since July 9, 2019 (Figure 2.15).

Figure 2.15

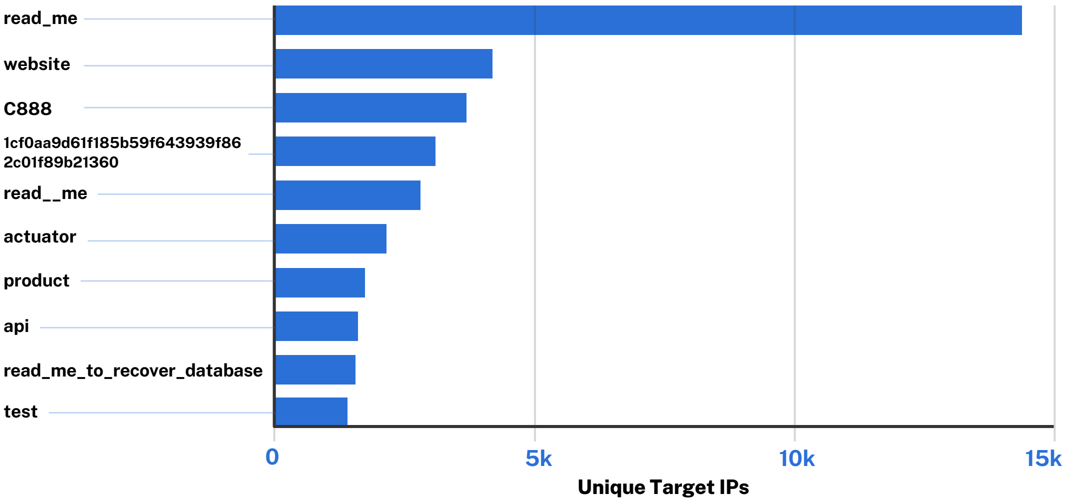


Elasticsearch

Over 2022, we detected a total of 4,962,164 IP addresses running Elasticsearch. Out of those, we found a total of 22,846 databases had been ransomed throughout the year. A total of 140 Terabytes of data were exposed to the internet with no authentication consisting of 178,902,591,446 documents.

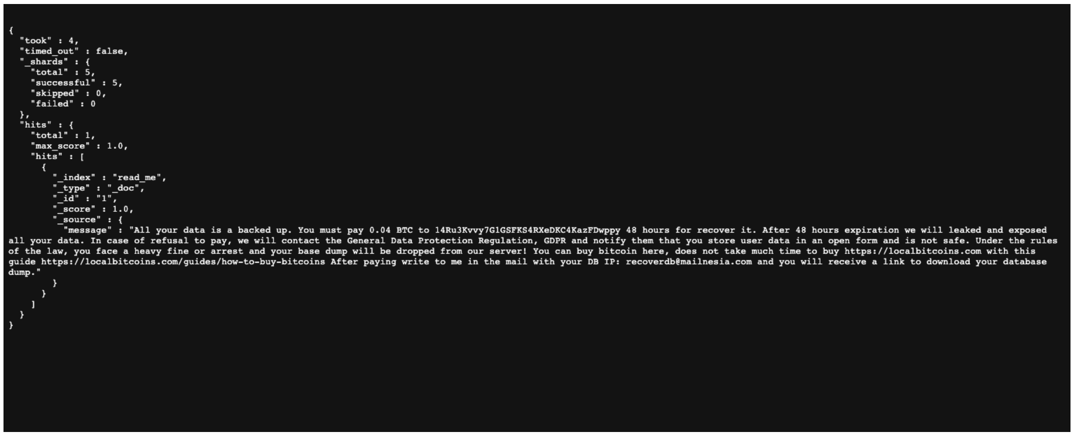
By looking at the below index names (Figure 2.16), we see the signal that shows the high volume of databases targeted for ransom.

Elasticsearch Index Names *(Figure 2.16)*



All the read_me's are essentially indexes renamed by threat actors who ransomed these databases, as seen in the following image when we look at the content of an index (Figure 2.17).

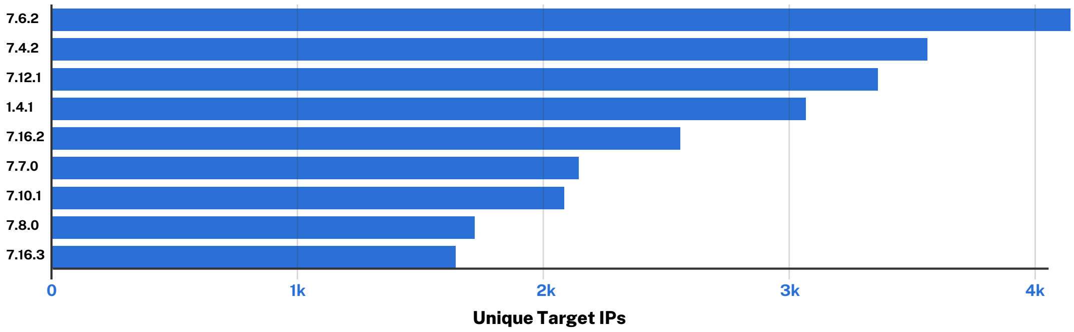
Figure 2.17



```
{
  "took" : 4,
  "timed_out" : false,
  "_shards" : {
    "total" : 5,
    "successful" : 5,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : 1,
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "read_me",
        "_type" : "_doc",
        "_id" : "1",
        "_score" : 1.0,
        "_source" : {
          "message" : "All your data is a backed up. You must pay 0.04 BTC to 14Ru3Kvvy7G1GSFKS4RXeDKC4KazFDwppy 48 hours for recover it. After 48 hours expiration we will leaked and exposed
all your data. In case of refusal to pay, we will contact the General Data Protection Regulation, GDPR and notify them that you store user data in an open form and is not safe. Under the rules
of the law, you face a heavy fine or arrest and your base dump will be dropped from our server! You can buy bitcoin here, does not take much time to buy https://localbitcoins.com with this
guide https://localbitcoins.com/guides/how-to-buy-bitcoins After paying write to me in the mail with your DB IP: recoverdb@mailnesia.com and you will receive a link to download your database
dump."
        }
      }
    ]
  }
}
```

In Elasticsearch scans, we continue to see the pattern we observed in previous sections of this report, where organizations are struggling with patching. The most used version (7.6.2) was released on March 31, 2020 (Figure 2.18).

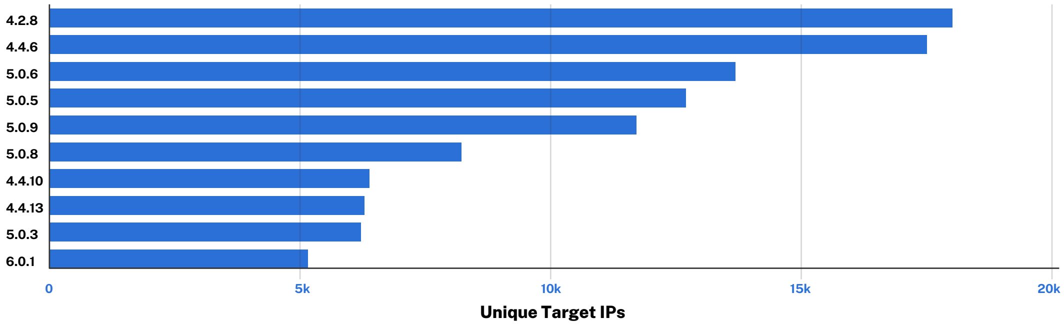
Elasticsearch Top Used Versions *(Figure 2.18)*



MongoDB

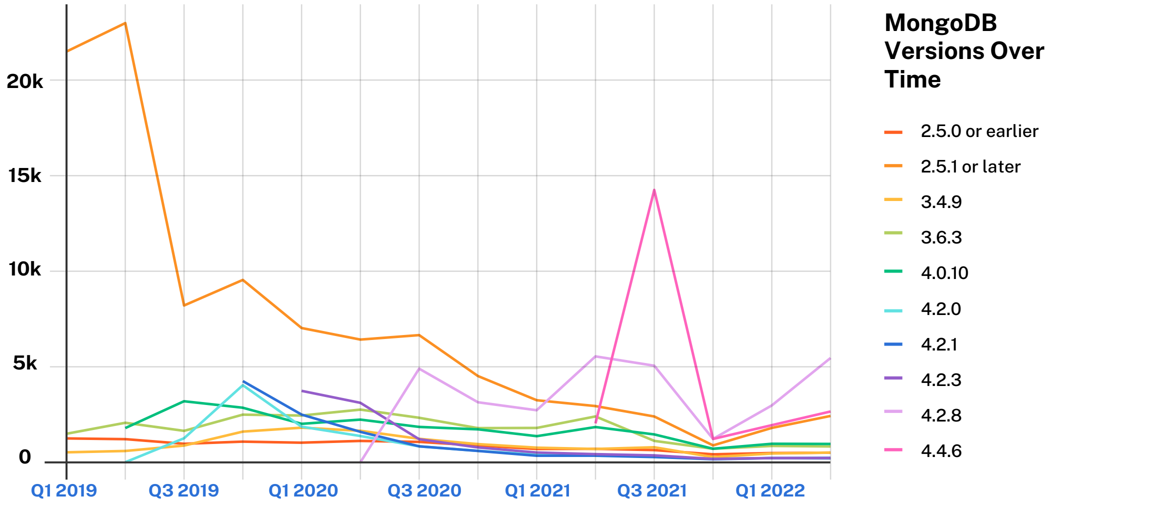
MongoDB, a NoSQL database, has seen high adoption throughout the years. It is easy to spin up and start using in a project. A total of 264,408 IP addresses running MongoDB instances were found in 2022. The most used version is 4.2.8, released on June 15, 2020 (Figure 2.19).

MongoDB Top Used Versions *(Figure 2.19)*



From a trends perspective, it is interesting to see the decline in 2.5.1 but also that 4.2.8 continues to grow despite being more than two years old (Figure 2.20).

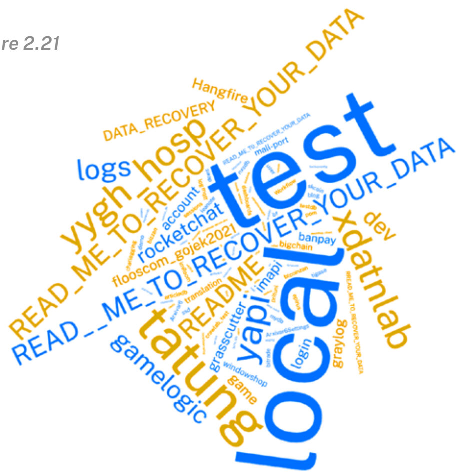
Figure 2.20



MongoDB databases were actively targeted throughout 2022, with a total of 68,423 of them hacked. That is, 26% of all installations found during the year were compromised. In total, we found 9.7TB of data exposed in MongoDB without authentication.

Looking at the database names in MongoDB exposures, the story tells itself, as the top names are essentially read_me_to_recover_your_data, which is the naming that threat actors use when they ransom a MongoDB installation (Figure 2.21).

Figure 2.21



We also found geopolitical impacts in cyber. Our searches turned up five databases renamed to SLAVA_UKRAINI, which stands for "Glory to Ukraine!" — a national salute used worldwide after the invasion of Ukraine by Russia. Interestingly, most, if not all, of the hacked databases were hosted in Russia.

Cassandra

Cassandra is a NoSQL database originally designed at Facebook (now Meta). Since 2019, from our scans, we have seen approximately 3,000 Cassandra installations exposed to the internet. This number has remained relatively stable.

The most used version, 2.0.15, is no longer maintained, and the 3.x branch will soon no longer be maintained, so the pattern continues where we see organizations not patching their software (Figure 2.22).

Version 2.x of Cassandra does not receive any security updates to patch existing vulnerabilities, which could allow attackers to access these databases. Looking at these exposed databases, we can see the type of data stored and what may have been stolen, everything from tokens that verify email registrations to customer and user data (including PII and PHI). It is exactly the kind of data you do not want to be leaked (Figure 2.23).

Upgrading and patching internet-facing software is critical, and organizations should implement a regular patching process.

Cassandra Top Used Versions *(Figure 2.22)*

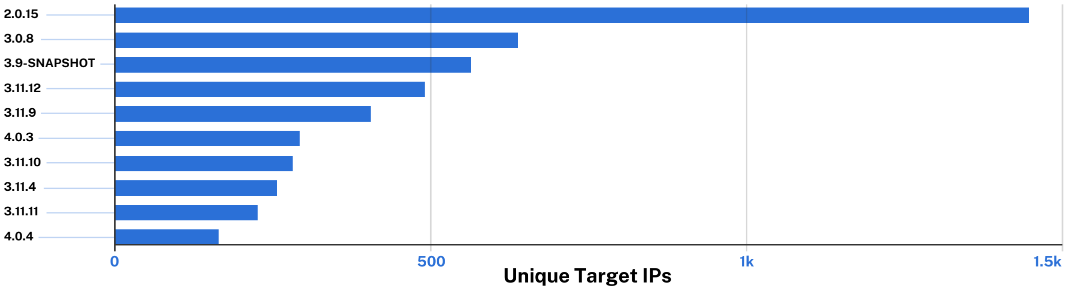
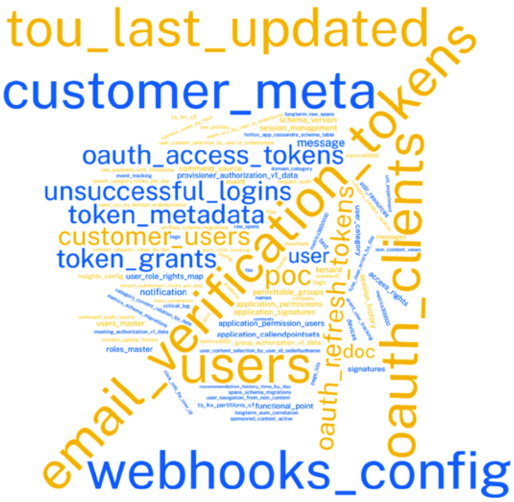


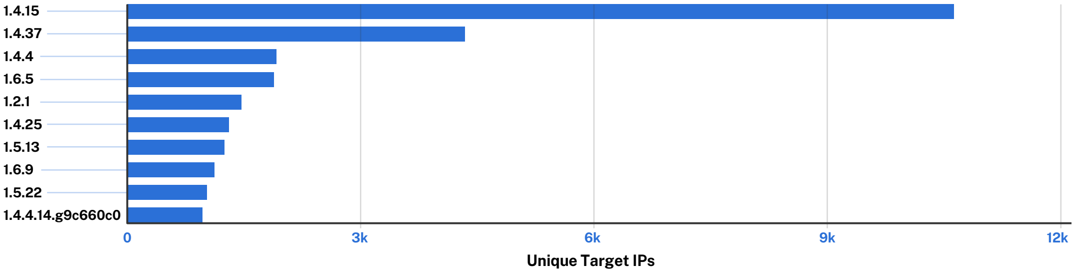
Table Names Found in Exposed Cassandra *(Figure 2.23)*



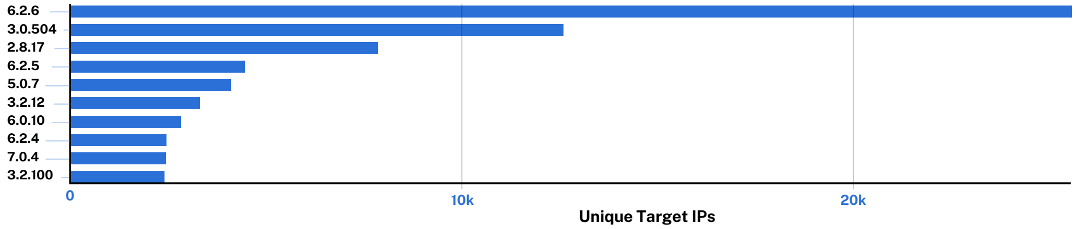
Redis and Memcached

Both of these are in-memory data storage technologies typically used to store ephemeral data. We found 34,341 Memcached instances and 102,823 Redis instance. We found these instances were leaking more than four Terabytes of data across them. In terms of versions, both the top versions for each are old versions; however, it is much worse for Memcached, where 1.4.15 was originally released in 2012 (Figure 2.24), and Redis 6.2.6 was released in 2021 (Figure 2.25).

Memcached Top Used Versions *(Figure 2.24)*



Redis Top Used Versions *(Figure 2.25)*

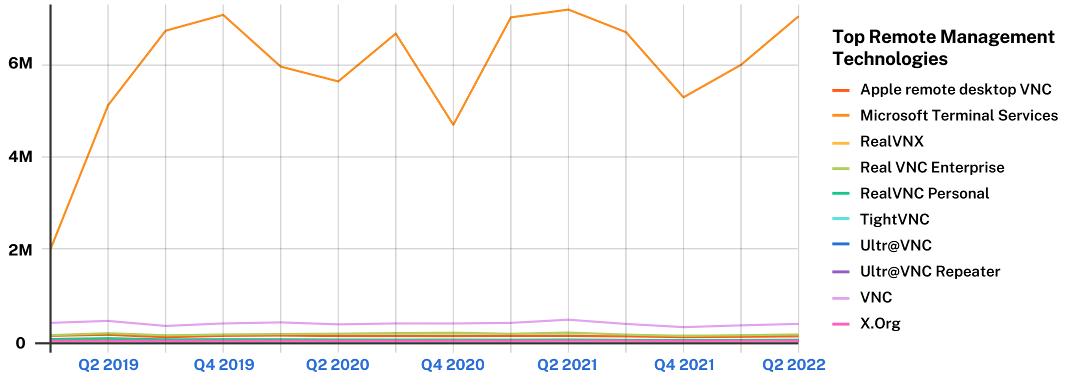


Remote Management Services

This section will focus on three remote management services: RDP, X11, and Virtual Network Computing (VNC). We focus on these three because they are the main services used with internet exposure and are also often seen as the initial root causes of ransomware cases.

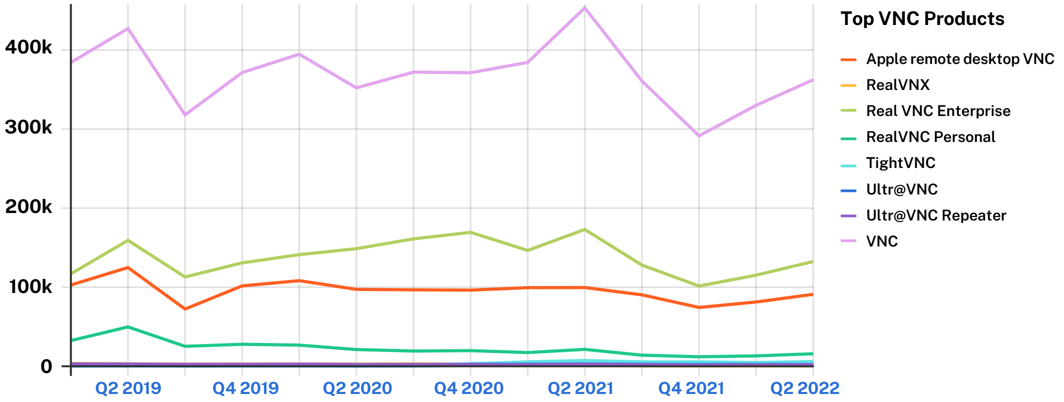
If we look into historical trends, we can see that RDP has been around for a while in high volume and will most likely continue to be (Figure 2.26).

Figure 2.26



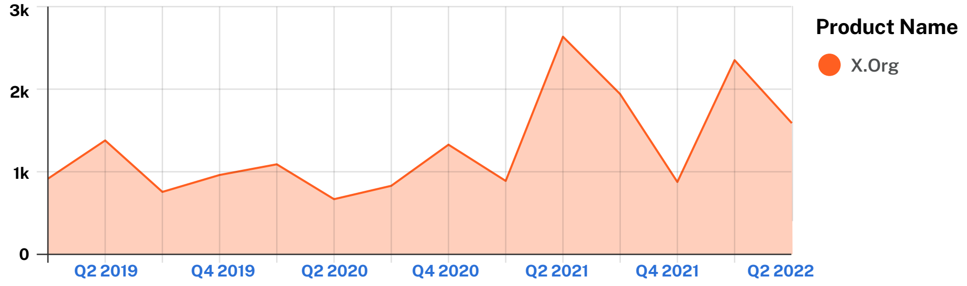
Sitting atop the hill of remote management services, RDP is king, with a presence of over 7 million IP addresses. VNC has multiple versions; in the following image, we remove RDP and X11 to focus only on the distribution and usage of VNC servers (Figure 2.27).

Figure 2.27



Although there was a slight dip in usage, VNC is still very much present on the internet. Last but not least, X11 has a much more reduced footprint of fewer than 2000 servers exposing this service to the internet (Figure 2.28).

X.Org Usage Over Time *(Figure 2.28)*



For each of these three services, we take a screenshot anytime we find one open. Here are the number of screenshots we took across these three services for this year alone (Table 2.2).

Table 2.2

RDP IMAGES	VNC IMAGES	X11 IMAGES
3,143,042	16,120	6,735

Why is there a gap between the number of screenshots and services found? For example, on RDP, organizations sometimes activate a security feature called network layer authentication (NLA), which does not show the RDP login screen before a client authenticates with a certificate.

However, even today, it is still easy to find critical systems or smart home devices exposed to the internet.

Overall, it is clear that companies are struggling to update their systems since a large majority of database servers are running older versions of software, potentially making them prime targets for attackers who ransom or steal data.

Organizations continue to expose databases without authentication to the internet, and many are automatically ransomed by attackers continuously scanning the internet for these technologies.

RDP continues to be a widely used technology that organizations should ensure they are using securely as attackers also continue to target it heavily. We will go into more detail on this in our Attacker Behaviors section.

SECTION 3

Industries

Below, we will dive into the current industry trends using a sample of Coalition's current cyber insurance policyholders.

We have selected 100 enforcement policies and insured companies belonging to each of the following industries:

- Consumer Services
- Financial Services
- Healthcare
- Professional Services
- Real Estate
- Technology

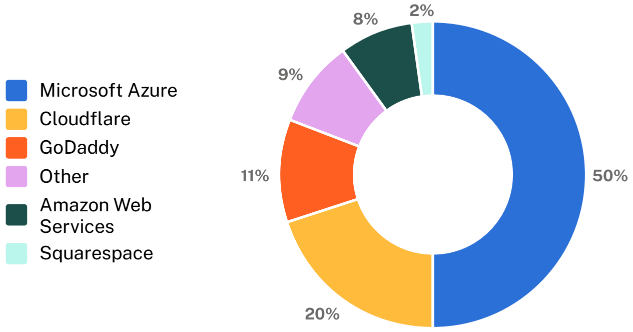
Each industry section covers the top tech choices, email providers, cloud providers, security check findings, asset breakdowns, and data leaks. The aggregations stem entirely from the underwriting scans run on these companies in the quoting process before binding, and all calculations and plots are over the entire industry.

INDUSTRIES

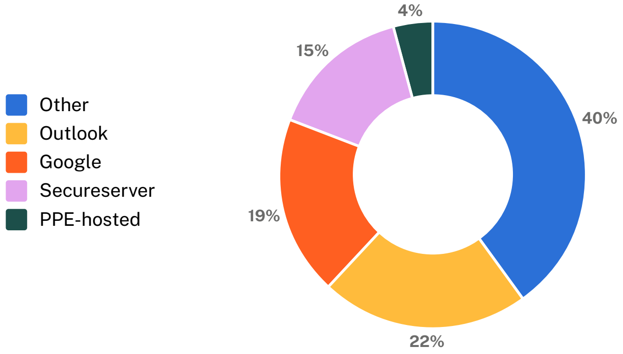
Consumer Services

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
30.70	43.52%	2.29%	9.36 OUT OF 10	9.45

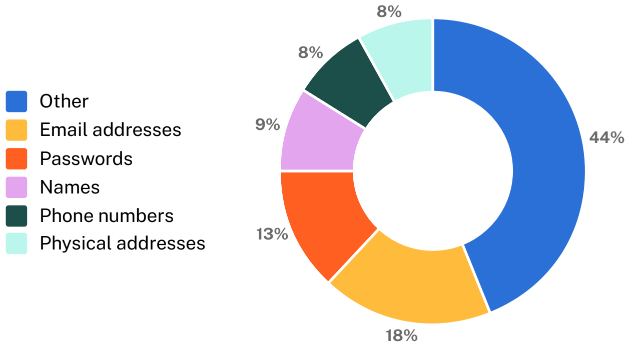
Cloud Providers



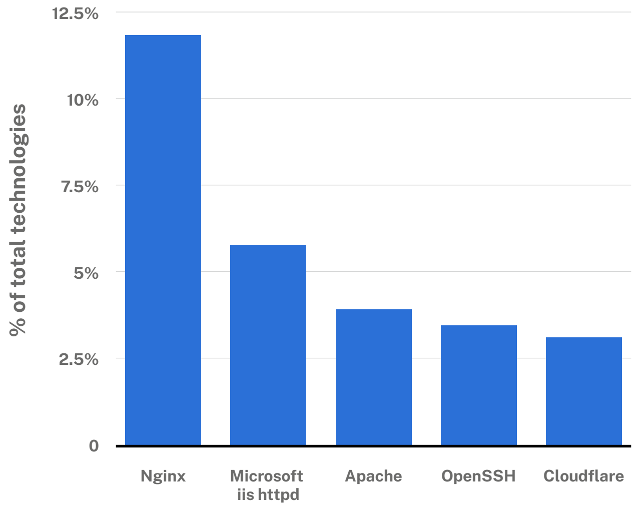
Email Providers



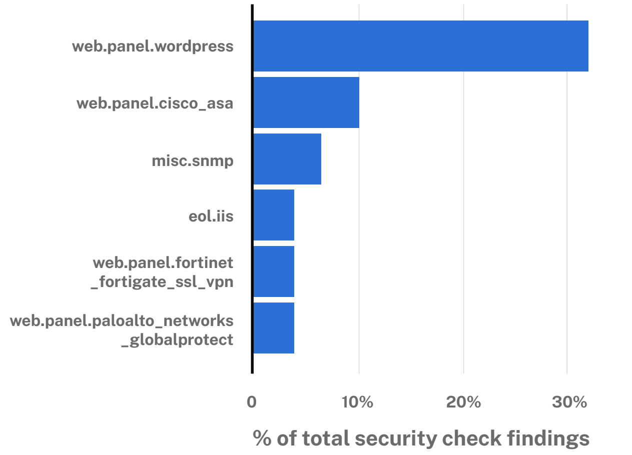
Type of Data Leaked



Technologies



Security Check Findings

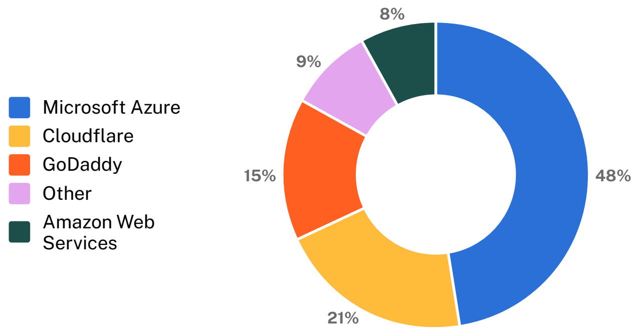


INDUSTRIES

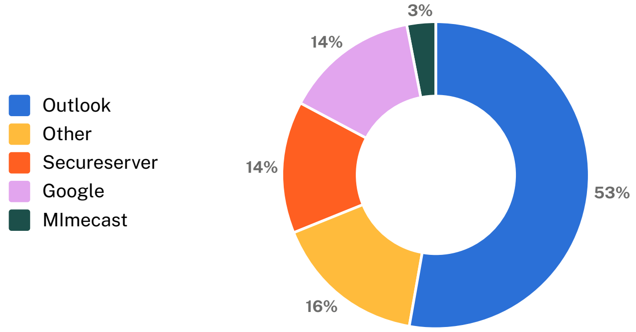
Financial Services

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
27.70	35.17%	4.25%	9.24 OUT OF 10	7.43

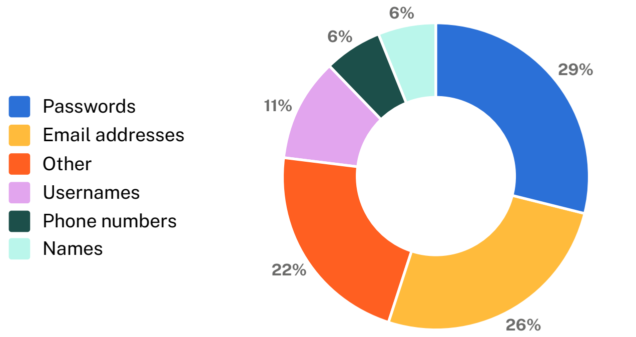
Cloud Providers



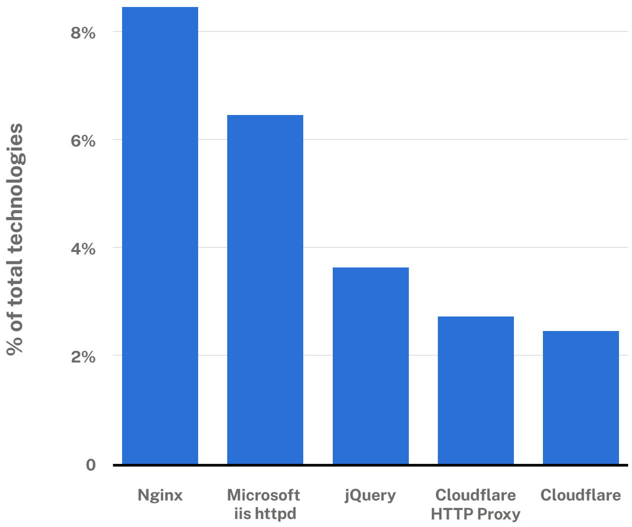
Email Providers



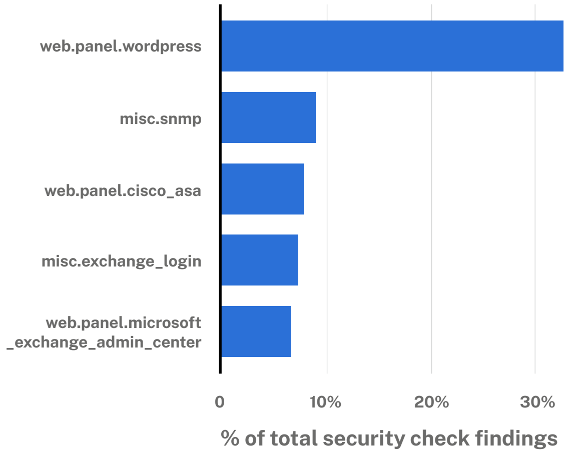
Type of Data Leaked



Technologies



Security Check Findings

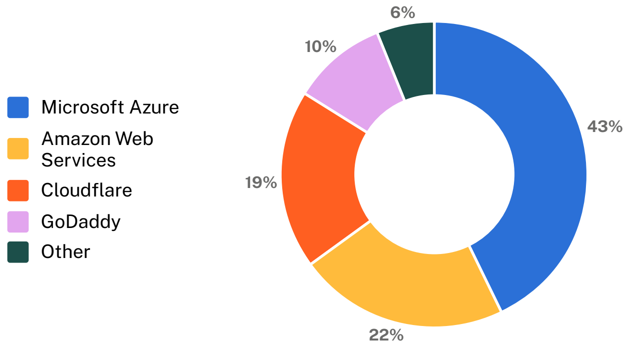


INDUSTRIES

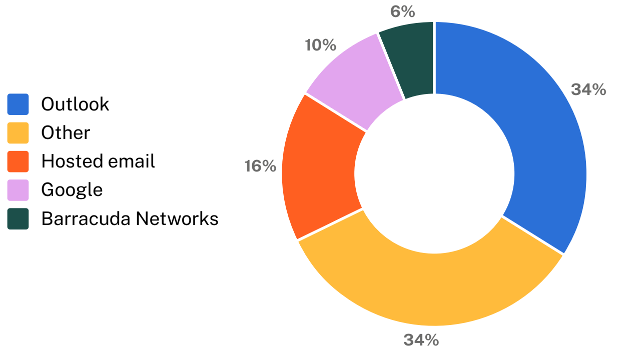
Healthcare

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
26.90	31.61%	5.29%	8.91 OUT OF 10	5.91

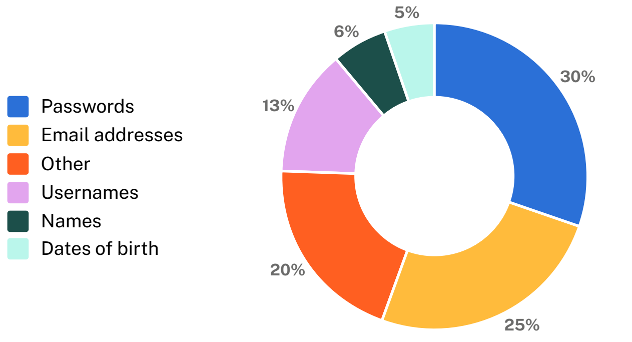
Cloud Providers



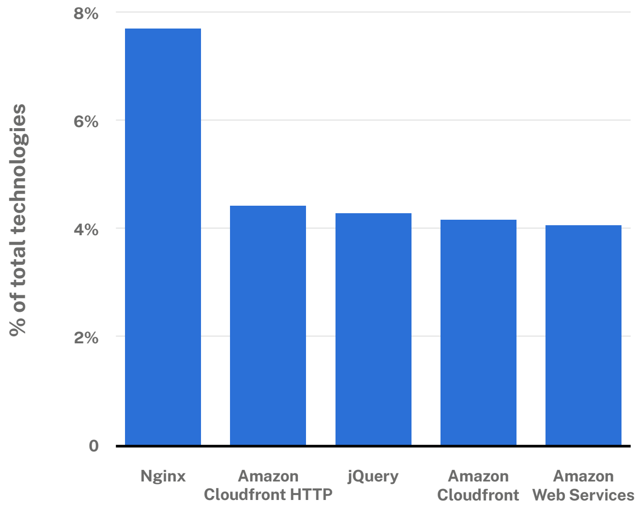
Email Providers



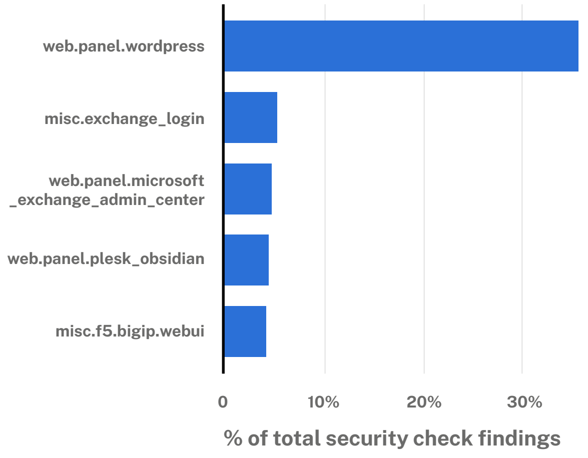
Type of Data Leaked



Technologies



Security Check Findings

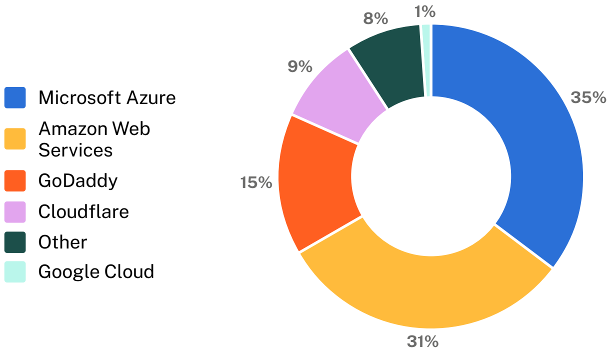


INDUSTRIES

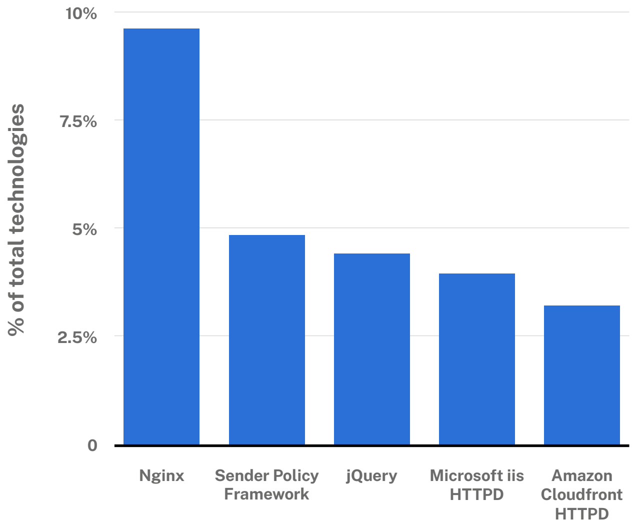
Professional Services

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
40.80	40.25%	2.57%	9.42 OUT OF 10	7.67

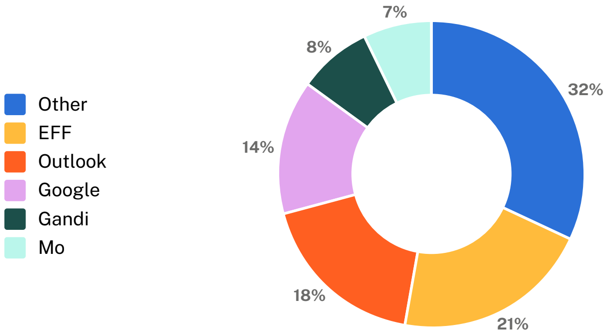
Cloud Providers



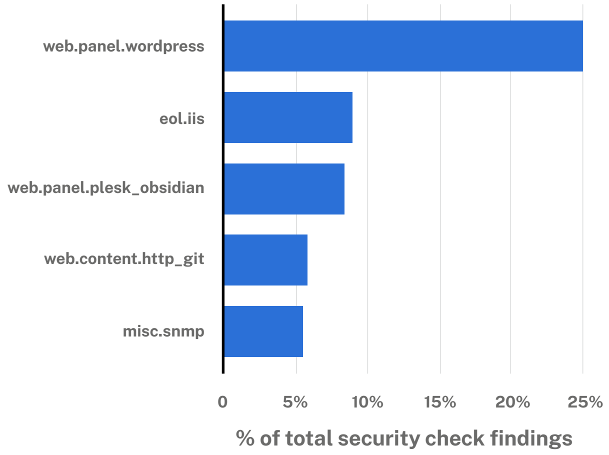
Technologies



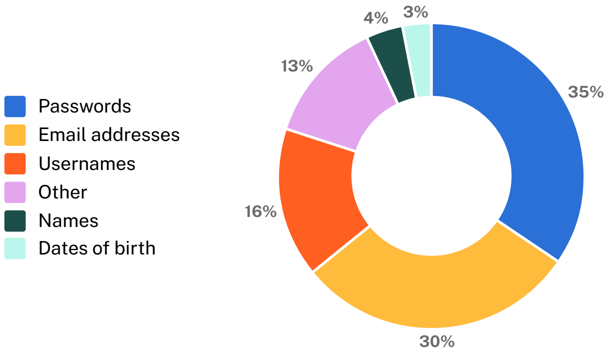
Email Providers



Security Check Findings



Type of Data Leaked

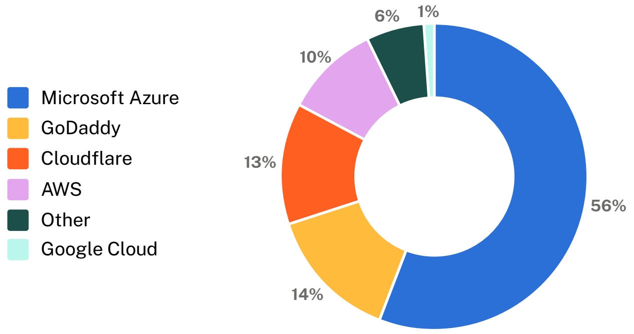


INDUSTRIES

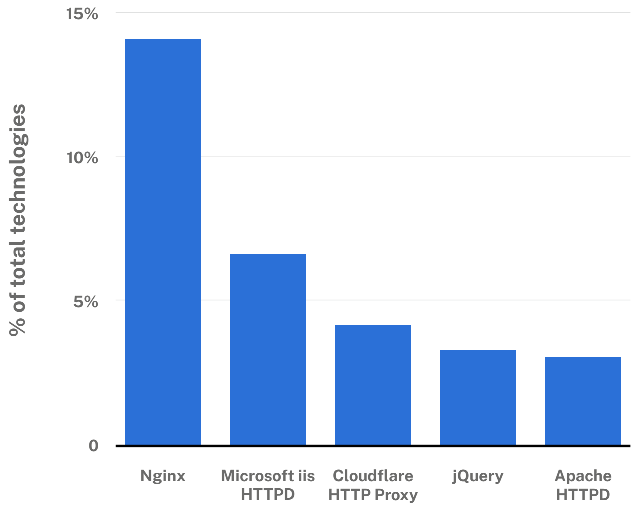
Real Estate

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
24.30	41.8%	4.03%	7.78 OUT OF 10	9.30

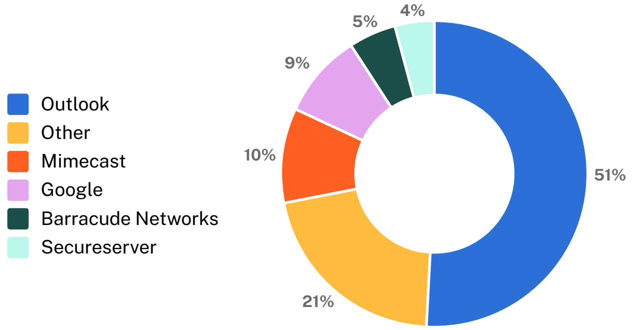
Cloud Providers



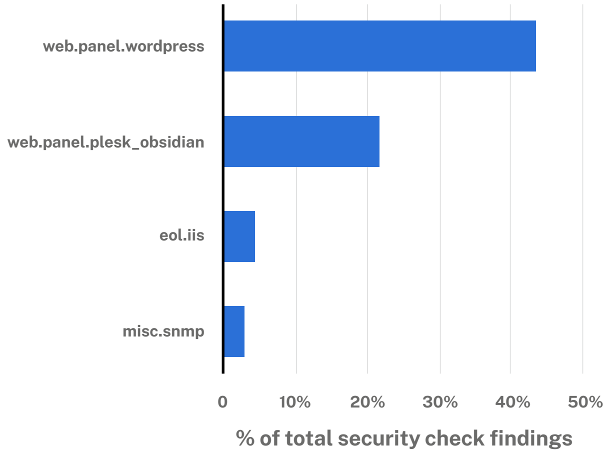
Technologies



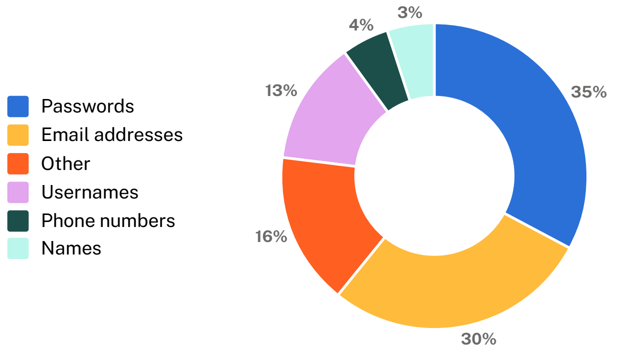
Email Providers



Security Check Findings



Type of Data Leaked

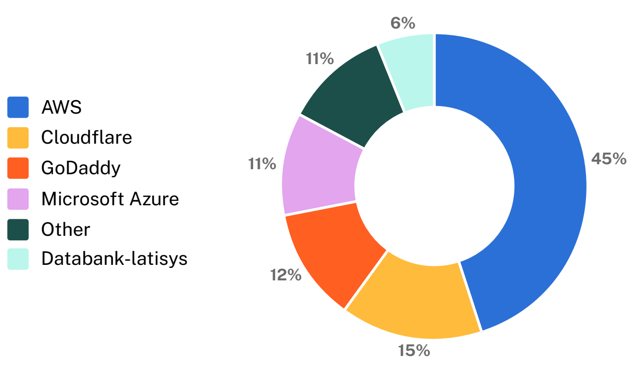


INDUSTRIES

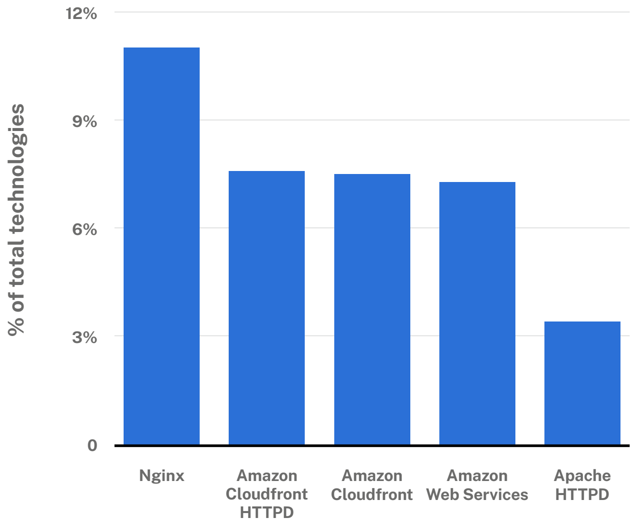
Technology

Distinct Tech Count (per company)	Cloud Hosted Asset Ratio	Security Check Findings Frequency	Average CVE Criticality	Distinct Data Leaks (per company)
55.60	37.59%	1.21%	9.29 OUT OF 10	5.59

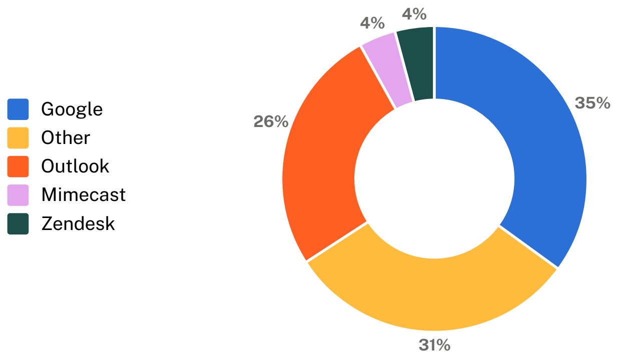
Cloud Providers



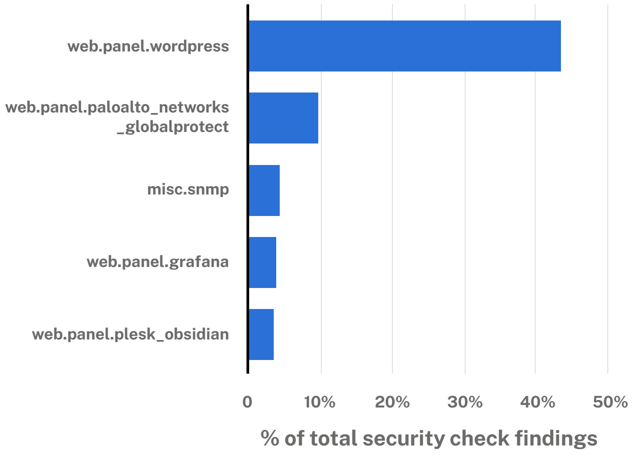
Technologies



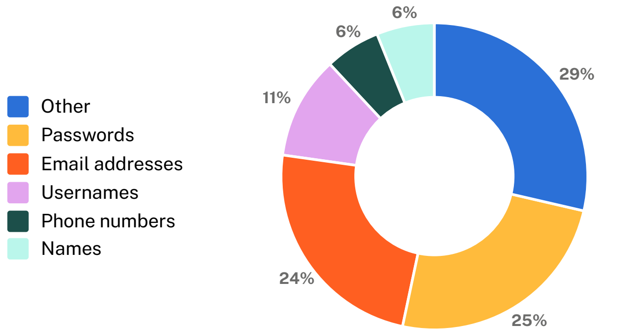
Email Providers



Security Check Findings



Type of Data Leaked



Industries Overview

By looking at the data across industries, we observe quite a few interesting trends.

Technology companies have the most complex stacks

Focusing on a company's tech stack, it is not surprising that the technology industry uses the largest number of unique technologies. Only the professional services industry comes close. The most popular choices among developers center around jQuery, Microsoft IIS, and Cloudflare, with NGINX holding the top spot across every industry. While other industries utilize Microsoft IIS quite a bit, the tech sector focuses more on Amazon products.

Most industries favor Microsoft cloud services; technology prefers AWS

Cloud usage across industries follows a similar trend. While other industries favor Microsoft with a high rate of Azure usage, tech companies tend to favor Amazon Web Services (AWS). Google Cloud Platform barely pops up in any of the breakdowns, despite accounting for 11% of the cloud marketplace in 2022. When examining asset breakdowns, most industries have between 35-45% of their total assets on the cloud, except for healthcare (30%).

It is surprising that the consumer services industry stores the highest percentage of assets in the cloud because the industry processes and stores customers' personally identifiable information (PII). Storing PII in the cloud is a significant security risk because, if not done correctly, it can be exposed for anyone to view and download.

Office 365 has the lion's share of email usage

Among email providers, Microsoft reigns supreme. While Gmail thrives in the tech industry, O365 has a greater market share across industries. It is interesting to note that healthcare, real estate, and financial services tend to use O365 far more frequently than Gmail. Meanwhile, among technology, professional services, and consumer services industries, Gmail and O365 are used at a similar rate.

Healthcare and real estate tend to have less-serious CVEs

While it is unfortunate that healthcare and real estate tend to have more security findings detected per asset, it is positive that they seem to be targeted with less harmful CVEs. The opposite is true within other industries. On average, many see fewer security findings, but the discovered CVEs are often more critical. Across all industries, WordPress is by far the most commonly detected security issue on scanned assets.

Emails and passwords top the list of compromised information

When looking at the industry breakdown of the kind of information lost due to a data breach, email addresses and passwords remain consistently on top. The only industry that stands apart is the consumer services industry. Most of the data lost from this sector are email addresses.

Other industries typically lose passwords, with emails following closely behind. It is also interesting that healthcare is one of the leaders across industries, with one of the lowest amounts of distinct breaches on average. That is a good sign, as these companies are more likely to hold highly sensitive patient care data.

SECTION 4

Attacker Behaviors

This section will examine some of the traffic collected by our honeypots and delve deep into the revelations the data provides about attacker behavior during the past year.

The web traffic received by our honeypots is divided into three categories: benign, malicious, and unknown. Benign traffic is web traffic not intended to cause harm and is usually generated by internet scans performed by security companies, search engine crawlers, or universities. Their scans are usually motivated by research or commercial reasons.

In contrast, malicious traffic is intended to cause harm. Often this classification results from attackers trying to exploit specific vulnerabilities or traffic we receive from known malicious actors like botnets.

Other types of honeypot traffic do not easily fit into the benign or malicious categories. Often, this can be attributed to the traffic being a novelty. This data is still collected, and in the future — when we categorize this traffic — we will look back and review how this type of traffic evolved.

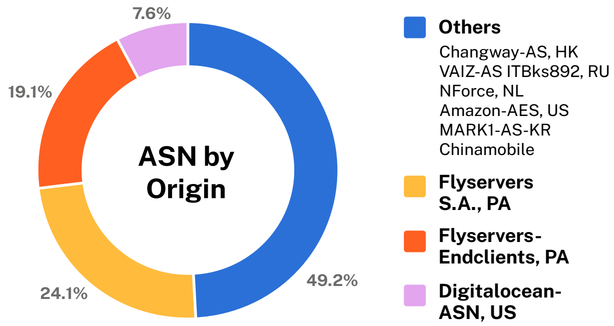
In the following sections, we will look first at the geographic distribution and types of protocols used by attackers. Then, we will share insights into the most frequently exploited CVEs and a pair of particularly significant vulnerabilities in Redis and Fortinet. Finally, we will look at the HTTP paths attackers seek to exploit.

Geographic Trends

We distribute our honeypots in countries worldwide. This allows us to have an accurate representation of the geographical trends because many attackers often target specific countries, and many entities block traffic completely from specific countries. Because we possess the geolocation origin of every payload sent to our honeypots, we can analyze the top countries, regions, and the autonomous system numbers (ASNs) attackers attempt to exploit.

Looking at the top ASNs of the traffic's origin (Figure 4.1), we see various public-cloud providers such as DigitalOcean and Amazon AWS.

Figure 4.1



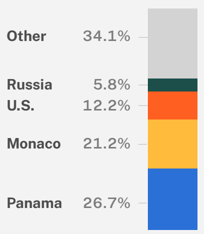
Knowing both where our honeypots are located as well as the origin of the attacks, we can break down traffic country by country. The top origin for attacks targeting the United States (Figure 4.2) is Panama, corresponding to the top ASN we saw above. Nearly all attacks on China originated from

Russia (Figure 4.3). Perhaps our most interesting observation is that nearly half of the attacks on Ukraine originated from within its own country (Figure 4.4).

Figure 4.2

United States Traffic

TOP ATTACKER COUNTRIES

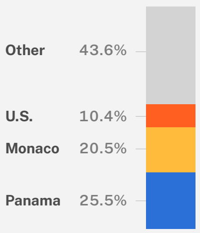


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 64.225.51.53	250,335,405
2. 138.99.216.81	222,116,963
3. 194.165.16.76	201,175,917
4. 194.165.16.11	197,599,853
5. 45.227.254.55	195,616,392
6. 194.165.16.77	195,437,376

United Kingdom Traffic

TOP ATTACKER COUNTRIES

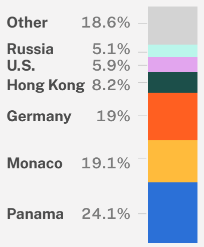


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 194.165.16.76	34,595,712
2. 194.165.16.11	34,051,707
3. 45.227.254.55	33,618,051
4. 44.196.108.12	33,573,874
5. 194.165.16.77	33,446,625
6. 194.165.16.10	33,419,025

Canada Traffic

TOP ATTACKER COUNTRIES

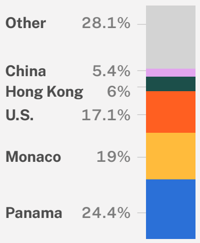


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 193.142.146.204	289,210,160
2. 92.255.85.195	63,565,766
3. 194.165.16.76	47,412,279
4. 138.99.216.81	47,085,115
5. 194.165.16.11	46,713,476
6. 45.227.254.55	46,426,007

Japan Traffic

TOP ATTACKER COUNTRIES



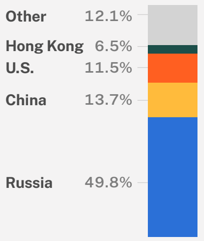
TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 64.225.51.53	124,748,034
2. 223.83.209.102	51,993,828
3. 138.99.216.81	46,314,660
4. 194.165.16.76	37,266,890
5. 194.165.16.11	36,866,981
6. 45.227.254.55	36,706,499

Figure 4.3

China Traffic

TOP ATTACKER COUNTRIES

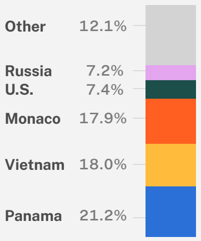


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 185.156.72.37	958,960,435
2. 64.225.51.53	181,904,405
3. 223.83.209.102	96,816,081
4. 52.131.45.108	68,433,952
5. 185.70.104.80	61,212,588
6. 92.255.85.195	47,847,831

Russia Traffic

TOP ATTACKER COUNTRIES

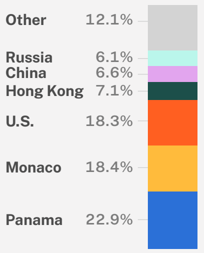


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 37.157.31.143	4,399,346
2. 103.125.190.35	3,938,953
3. 103.89.89.81	3,386,565
4. 194.165.16.76	2,703,287
5. 194.165.16.73	2,669,657
6. 45.227.254.55	2,638,079

Hong Kong Traffic

TOP ATTACKER COUNTRIES



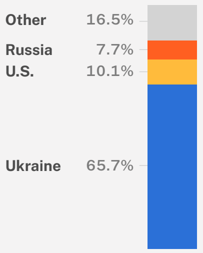
TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 64.225.51.53	124,807,881
2. 223.83.209.102	40,066,427
3. 194.165.16.71	28,167,816
4. 194.165.16.11	27,725,822
5. 138.99.216.81	27,606,311
6. 45.227.254.55	27,405,523

Figure 4.4

Ukraine Traffic

TOP ATTACKER COUNTRIES

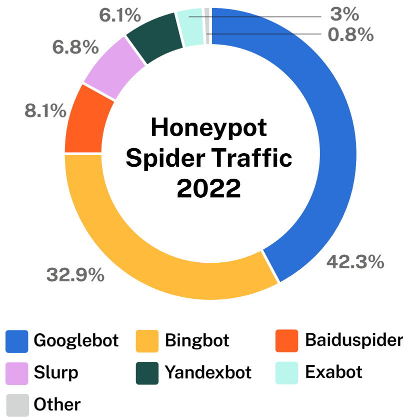


TOP ATTACKER IPs

ORIGIN IP	RECORD COUNT
1. 185.156.42.225	35,993,103
2. 185.156.41.181	1,707,703
3. 185.67.2.19	1,299,411
4. 165.22.31.110	1,069,133
5. 185.67.2.56	997,120
6. 194.28.84.86	953,281

Search Engine Spiders: Real and Fake

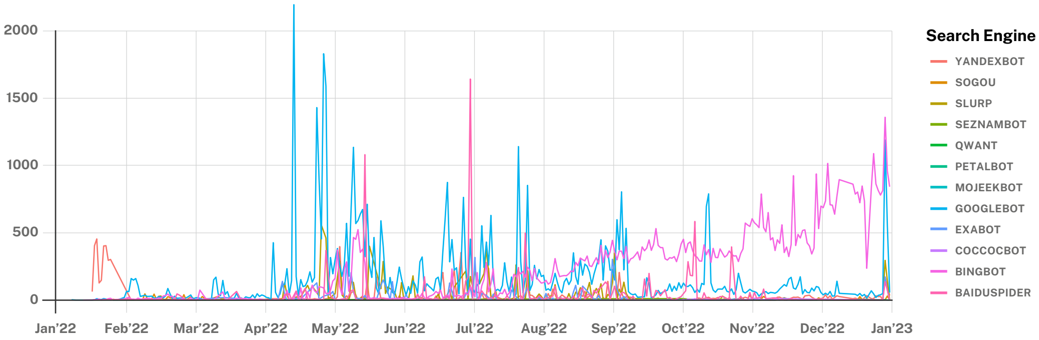
Figure 4.5



Attackers increasingly use search engine crawlers to hide their activity and evade security protocols. Search engine crawlers, or "spiders," generate a lot of the benign traffic picked up by our honeypots. Companies operating search engines scan the internet to index the content of websites so that those websites can be displayed in search results. In 2022, our honeypots logged traffic from many different spiders (Figure 4.5). The two that generated the most traffic on our honeypots are Google's Googlebot and Bing's Bingbot. These accounted for traffic well above the rest, representing 42.3% and 32.9%, respectively, of the total spider traffic we received.

Search engine crawlers use the User-Agent HTTP header, a characteristic string that identifies the application, operating system, vendor, or version of the requesting user, to identify themselves. This makes it easy for our honeypots to identify them. However, it also makes it very easy for attackers to hide behind the identity of a search engine crawler. During the past year, we have seen traffic spikes in this area, likely due to attackers masquerading as spiders (Figure 4.6).

Honeypot Traffic by Crawler *(Figure 4.6)*



If we look at the spike in Googlebot traffic on April 15, 2022, we see the IP address that is the source of this spike is 106.75.15.80. Looking through our data, we see that this IP address has previously attempted to conceal its identity by posing as other search engine crawlers. We can also see that the operators of this address have tried to exploit CVE-2020-5902, a remote code execution vulnerability found in F5 BIG-IP devices.

Modifying user-agents is common practice among malicious actors and something to be aware of. Analyzing the traffic, we see that the traffic from our top five highest-traffic crawlers spread across 72 unique user-agents (Table 4.1).

Table 4.1

Unique User-Agents by Crawler

	SEARCH ENGINE CRAWLER	USER-AGENT
1	GOOGLEBOT	35
2	BINGBOT	7
3	BAIDUSPIDER	24
4	SLURP	4
5	YANDEXBOT	2

An easy way to combat this is to perform Domain Name System (DNS) and reverse DNS lookups when an IP address identifies itself as a search engine crawler. This allows you to see what domain the IP address belongs to and then verify if that IP address is listed under the domain's DNS records.

Top Attack Types

All traffic we see on our honeypots receives one or more tags as a way to organize and classify the traffic. The types of tags we use range from the specific name of the technology or CVE the attacker is trying to exploit to generic "scanner" traffic like HTTP or SSH. Below are the top 10 tags for 2022, which show the top types of protocols attackers seek to exploit (Table 4.2).

Table 4.2

1	RDP_SCANNER	37.67%	Scanning for Remote Desktop Protocols
2	SSH_SCANNER	10.11%	Valid SSH connections
3	ICMP_ECHO_REQUEST	1.53%	Ping event
4	HTTP_SCANNER	1.06%	Scanning for HNAP routers
5	SSL_SCANNER	0.64%	Valid SSL Connections
6	SMB_SCANNER	0.18%	Scanner for SMB Protocol often affiliated with the exploitation of Microsoft Windows
7	HTTP_REFLECTION	0.14%	The source of the event tried to make one of our sensors access something from a third party, a potential DDoS attack
8	PROXY_SCANNER	0.14%	Scanning for open proxies
9	UPNP_SCANNER	0.11%	Scanner for UPNP protocol
10	VOIP_SCANNER	0.09%	Scanner for VOIP protocol

Much of our traffic is some kind of "scanner" traffic used to denote when a potential attacker attempts to identify what services are running on each IP as they broadly scan large portions of the public IP space. Although this type of traffic may attempt to identify specific services they wish to exploit, it is important to separate this from an actual exploitation.

On the one hand, we have the benign tag for when we determine that the source of the request is a harmless actor, such as a university or search engine web crawler. On the other hand, we have CVE tags that denote traffic targeting specific exploits.

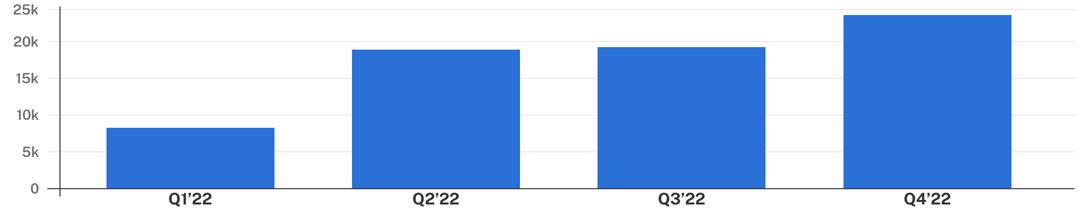
RDP Traffic Remains Huge

Over 37% of all traffic we see is RDP-related. Below is a graph showing the number of unique origin IPs for RDP traffic, along with a running average. Early in the year, we saw a sharp increase in traffic leading up to May (Figure 4.7).

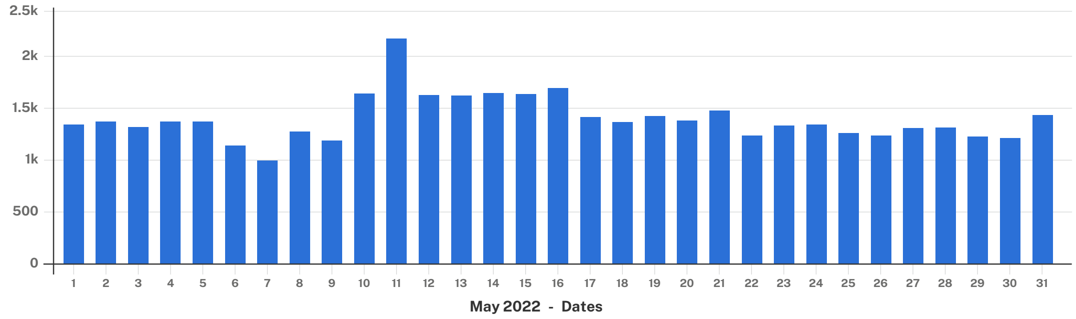
On May 11, 2022, Microsoft announced a patch for 75 vulnerabilities, including three that were zero days (Figure 4.8). Among these zero-day vulnerabilities was CVE-2022-22017 (published in the NVD on May 10, 2022), a Remote Code Execution (RCE) vulnerability in RDP for Windows 11 and Windows Server 2022. While RDP traffic rose sharply in the weeks leading up to the day this CVE was published, it continued to rise throughout the year, albeit at a slower pace.

When a new CVE is published, attackers typically begin scanning the internet to create an inventory of potentially vulnerable targets. Since RDP is one of the most commonly exploited technologies, it is unsurprising that this CVE caused a significant increase in related traffic.

RDP Scanning 2022 Quarter over Quarter *(Figure 4.7)*



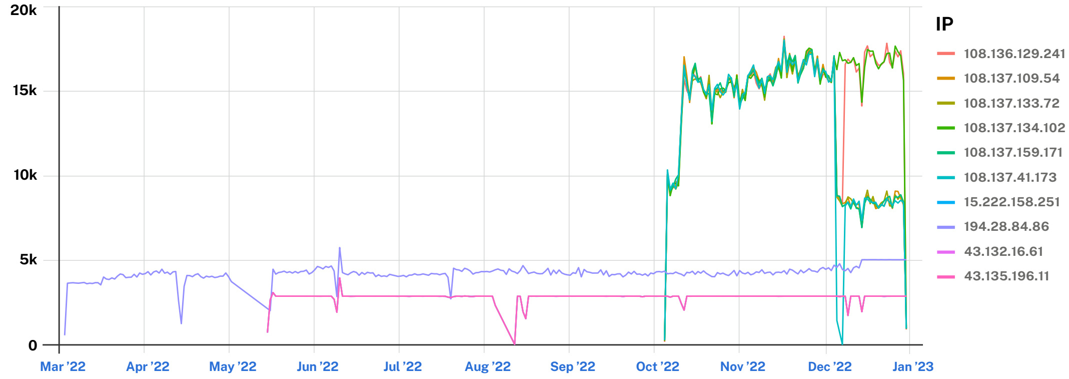
CVE-2022-22017 Exploitation over May 2022 *(Figure 4.8)*



ICMP Traffic Continues

We see a large number of Internet Control Message Protocol (ICMP) echo requests daily. These requests are typically known for one of two things. The first is a simple ping. Many IPs we see making these requests make exactly 2,880 requests each day. Some quick math will tell you these IPs are making a request every 30 seconds, likely meaning they are trying to determine if they can make a healthy connection to our servers. ICMP echo requests are also a common method of Distributed Denial of Service (DDoS) attacks. We saw increased request frequency from a handful of IPs, one originating from Ukraine (Figure 4.9).

Figure 4.9



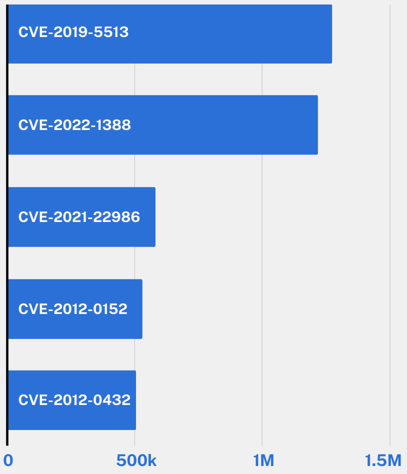
The Five Most-Targeted Vulnerabilities

Our honeypots are configured with many vulnerabilities that have known exploits so that we can observe attacker behaviors. We tag any traffic attempting to exploit those CVEs. Throughout the year, we saw spikes in such traffic, typically correlated with discoveries or disclosures related to CVEs. There were also some CVEs frequently exploited throughout the entire year.

Below are the top five CVEs we saw the most traffic for in 2022. These are the vulnerabilities that we have observed attackers targeting the most often (Figure 4.10).

Figure 4.10

Top Five CVEs at a Glance



1. **CVE-2022-1388**
    a. F5 BIG-IP iControl
    b. Remote Code Execution (RCE) vulnerability with authentication bypass
    c. Base Score (CVSS v3.x): 9.8 Critical

2. **CVE-2021-22986**
    a. F5 BIG-IP iControl
    b. Remote Code Execution (RCE) vulnerability with authentication bypass
    c. Base Score (CVSS v3.x): 9.8 Critical

3. **CVE-2019-5513**
    a. VMWare Horizon Connection Server
    b. Information disclosure of internal domain names, the Connection Server's internal name, or the gateway's internal IP address
    c. Base Score (CVSS v3.x): 5.3 Medium

4. **CVE-2012-0152**
    a. Remote Desktop Protocol (RDP) Microsoft Windows Server 2008 R2 and R2 SP1 and Windows 7 Gold and SP1
    b. Terminal Server Denial of Service Vulnerability
    c. Base Score (CVSS v2.0): 4.3 Medium

5. **CVE-2012-0432**
    a. Novell NCP implementation in NetIQ eDirectory
    b. Stack-based buffer overflow allows remote attackers to have an unspecified impact via unknown vectors
    c. Base Score (CVSS v2.0): 10.0 High

F5 BIG-IP iControl: The Most-Targeted Vulnerability

The first observation we make is that the top two CVEs are the same service, F5 BIG-IP iControl. CVE-2021-22986 was published in March 2021 and affected versions, including 16.0.x before 16.0.1.1. CVE-2022-1388 was published in May 2022 and affected versions, including 16.1.x before 16.1.2.2.

The PoCs referenced in these two CVEs are nearly identical, likely meaning the vulnerability was reintroduced with the 16.1 update. The fact that even software vendors can make mistakes and create new vulnerabilities underscores the importance of keeping up with patch cadence cycles and following vendor recommendations.

RDP for Windows Server 2008: The Oldest High-Traffic Vulnerability

Despite being published 10 years ago, CVE-2012-0152 continues to be targeted at a high rate. As we saw earlier in Top Tags, RDP dominates our honeypot traffic.

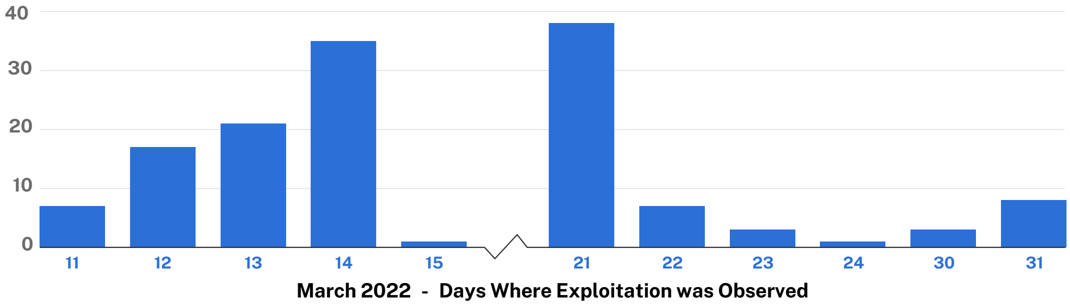
Redis - The Most Severe Vulnerability

On February 18, 2022, NIST published CVE-2022-0543, a remote code execution vulnerability in Redis, an open-source (BSD licensed), in-memory database. It received a CVSS 3.x score of 10.0, the highest severity as an attacker would get full control over a Redis instance by exploiting this vulnerability.

On March 11, 2022, we began seeing attempts to exploit this vulnerability on our honeypots, indicating this had become an actively exploited vulnerability (Figure 4.11). After our team determined the susceptible versions of Redis, we began identifying all affected policyholders.

Although this is a remote code execution vulnerability, we primarily saw two types of exploits. The first was simply an attempt to expose the system password (`cat /etc/passwd`). The second exploit was an attempt to download a file called `russia.sh` from a specific IP address.

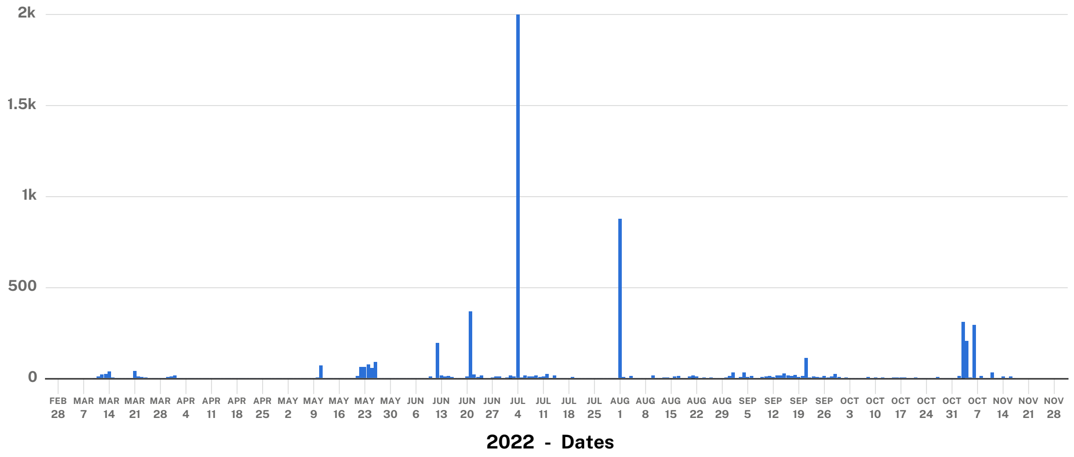
Figure 4.11



Although this CVE did not account for a large percentage of our CVE-related traffic for the year, we rushed to evaluate its potential negative impact on our policyholders, especially due to its high severity.

Throughout the remainder of the year, attacks for this CVE were sporadic, with the biggest spike coming on July 4 (Figure 4.12). Breaking down this data further, we see many spikes were from lone IPs. On July 4, all traffic came from 138.2.67.235, originating from Singapore. On August 1, all traffic came from 52.198.49.137, originating from Japan.

Exploitation Activity *(Figure 4.12)*

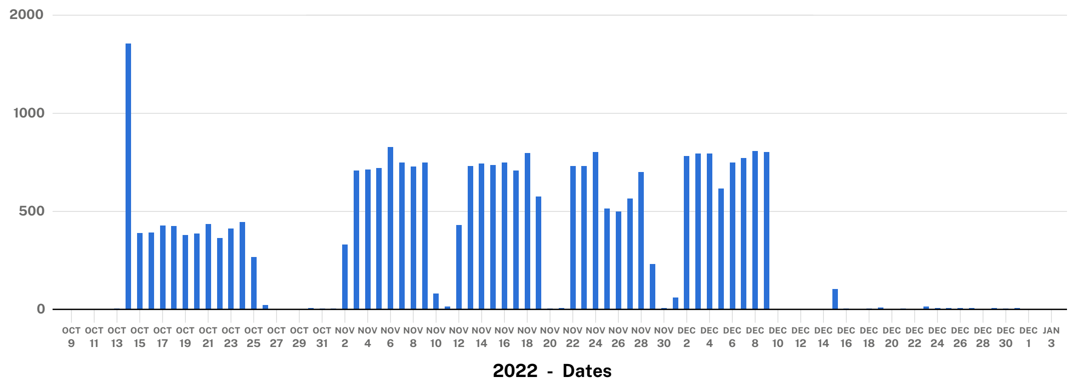


Fortinet - The Quickest to Be Exploited

On October 10, 2022, Fortinet announced an authentication bypass vulnerability, CVE-2022-40684, for three of its products, FortiOS, FortiProxy, and FortiSwitchManager. This vulnerability allowed unauthenticated attackers to update the SSH key for the admin user via specifically crafted HTTP or HTTPS requests. Once the SSH key was updated, the attacker could SSH into the vulnerable machine and perform administrative operations.

Just two days after the announcement, on October 12, we began to see an increase in scanning for generic Fortinet appliances on our honeypots (Figure 4.13). On October 13, at approximately 4:52 PM UTC, the PoC exploit code was released. Less than seven hours later, at 11:42 PM UTC, we began to see attempted exploitation on our honeypots.

Exploitation Activity *(Figure 4.13)*



Over the next 14 days, a steady stream of traffic attempted to exploit this vulnerability. Today, our honeypots have logged over 22,000 total events and approximately 50 unique SSH keys (Table 4.3). The traffic has been spread across 208 unique IP addresses, with most of the traffic coming from Germany, the United States, the United Kingdom, Singapore, and the Netherlands.

While a majority of attackers have tried to modify the SSH key for the admin user, we have also logged requests with the username set to *fgate* and *fortinet-tech-support*.

Table 4.3

	Attacker Country	# of Events
1	Germany	8,740
2	United States	4,272
3	United Kingdom	3,818
4	Singapore	2,884
5	Netherlands	1,413
6	Canada	881
7	India	331
8	Australia	71
9	Italy	62
10	Nicaragua	36

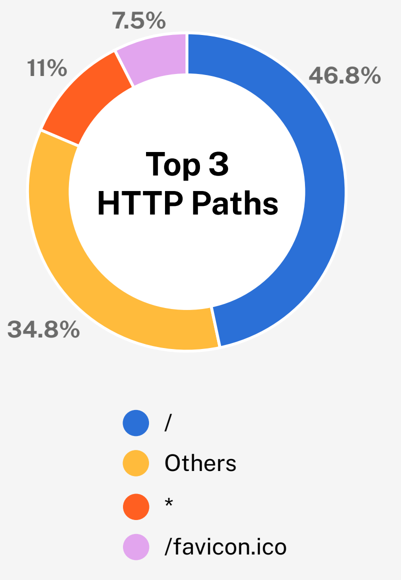
One SSH key was sent by two separate IP addresses. This leads us to believe that these IP addresses belong to the same attacker. If we look at the activity of these IP addresses across the last year, we can see that both IP addresses were only active on one day, November 6, 2022. Even though they were only active on one day, they both managed to trigger 35 unique tags across all of our honeypots.

Top HTTP Paths

HTTP paths specify which resource on a host a client is trying to access. When an individual is trying to interact with one of our honeypots, one of the fields that we pay close attention to is the HTTP path. HTTP paths can give us different insights into the motives of the client or attacker. Over this past year alone, our honeypots have seen **305,694** unique paths.

Some common HTTP paths represent most of the traffic we receive on our honeypots. However, HTTP paths can often be unique to a specific product or vulnerability. Combining unique HTTP paths with other signals, we can gain more insight into the motivations behind the web traffic.

Top 3 HTTP Paths (%) *(Figure 4.14)*



The top three HTTP paths that our honeypots logged were (Figure 4.14):

- /
- *
- /favicon.ico

These three HTTP paths combined accounted for approximately 65% of the HTTP traffic we received, approximately 47% for / (root), 11% for *, and 7.5% for ***/favicon.ico***.

Top HTTP Paths *(Figure 4.15)*

	HTTP Path	# of Events
1	/portal/favicon.ico	5,141,231
2	/ws/v1/cluster/apps/new-application	3,316,211
3	/ws/v1/cluster/apps	3,216,693
4	/robots.txt	3,018,171
5	/c/version.js	2,455,862
6	/system_api.php	2,455,510
7	/stalker_portal/c/version.js	2,454,988
8	/stream/live.php	2,454,779
9	/streaming/clients_live.php	2,454,672
10	/flu/403.html	2,453,989

The remaining top 10 HTTP paths, representing approximately 35% of HTTP traffic, can give us some more insight into the landscape of attacker activity (Figure 4.15).

In spots two and three, we find the first HTTP paths on our list that are associated with a vulnerability. ***/ws/v1/cluster/apps/new-application*** and ***/ws/v1/cluster/apps*** are paths that can be used to exploit a Hadoop YARN REST Application Programming Interface (API) vulnerability allowing an attacker to remotely execute code without authorization.

Next on the list is ***/robots.txt***. This path is used as a resource for hosts to communicate with web crawlers and other web robots. Out of the approximately 5,600 IP addresses we have flagged as search engine crawlers, only about 2,000 IP addresses have attempted to access the /robots.txt file on our honeypots.

/manager/html is a well-known Apache Tomcat path where attackers can upload files to perform remote code execution. This path is protected by authentication; however, most payloads associated with this path attempt to sign in using some of the most common credentials.

/mgmt/tm/util/bash is a path attackers can use to exploit CVE-2022-1388. Through this path, attackers can send a carefully crafted POST request to bypass iControl Rest Authentication on F5 BIG-IP devices and gain admin access to the device.

/tmui/tmui/system/settings/redirect.jsp is a unique path where attackers can exploit a remote code execution vulnerability, CVE-2020-5902, on F5 BIG-IP devices through the Traffic Management User Interface (TMUI).

Conclusion

Cybersecurity has never been easy. Sometimes it may seem like an endless, high-stakes game of cat-and-mouse between threat actors and organizations, a constant battle to see who can stay one step ahead.

Behind the headlines, thousands of CVEs go unnoticed by the public — but not by threat actors. Some are inconsequential, others serious, and a few represent a critical threat and increased risk for many organizations.

With so many vulnerabilities to address, systems often go unpatched for years, leaving huge swaths of the internet unprotected.

What all this means is that leaders responsible for protecting network security need the most accurate and insightful information to act upon — and they need an effective way to prioritize which CVEs to respond to. We have attempted to provide that necessary context and the CVSS/CESS framework to help cybersecurity leaders and practitioners make informed decisions about their digital risk and react quickly to harmful vulnerabilities.

They need the best solutions and tools to help track and defend against sophisticated threat actors and ransomware thieves. Most of all, we all need expert assistance to mitigate, actively assess, and, if need be, respond to the growing risk.

About Coalition

Coalition is the world's first Active Insurance provider designed to prevent digital risk before it strikes. By combining comprehensive insurance coverage and cybersecurity tools, Coalition helps businesses manage and mitigate digital risks.

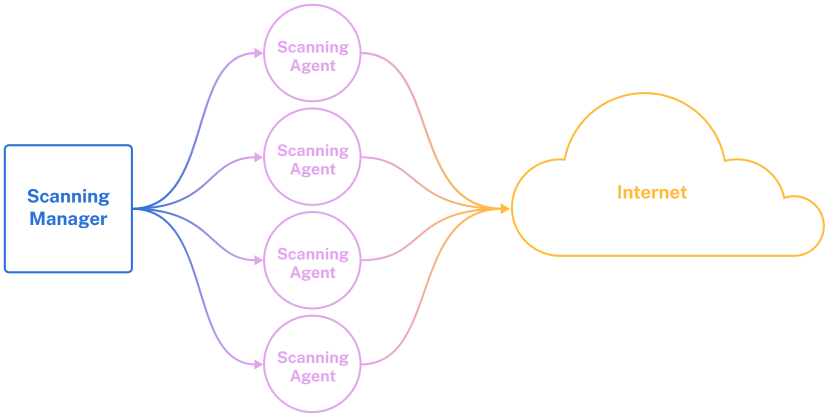
Through its partnerships with leading global insurers, including Arch Insurance North America, Allianz, Ascot Group, Lloyd's of London, Swiss Re Corporate Solutions, and Vantage, Coalition offers its Active Insurance products on behalf of its carrier partners in the U.S., U.K., and Canada, and its security products to organizations worldwide. Coalition's Active Risk Platform provides automated security alerts, threat intelligence, expert guidance, and cybersecurity tools to help businesses remain resilient in the face of cyberattacks. Headquartered in San Francisco, Coalition is a distributed company with a global workforce that collaborates both digitally and in office hubs across the globe.

Methodology

Scanning the Internet

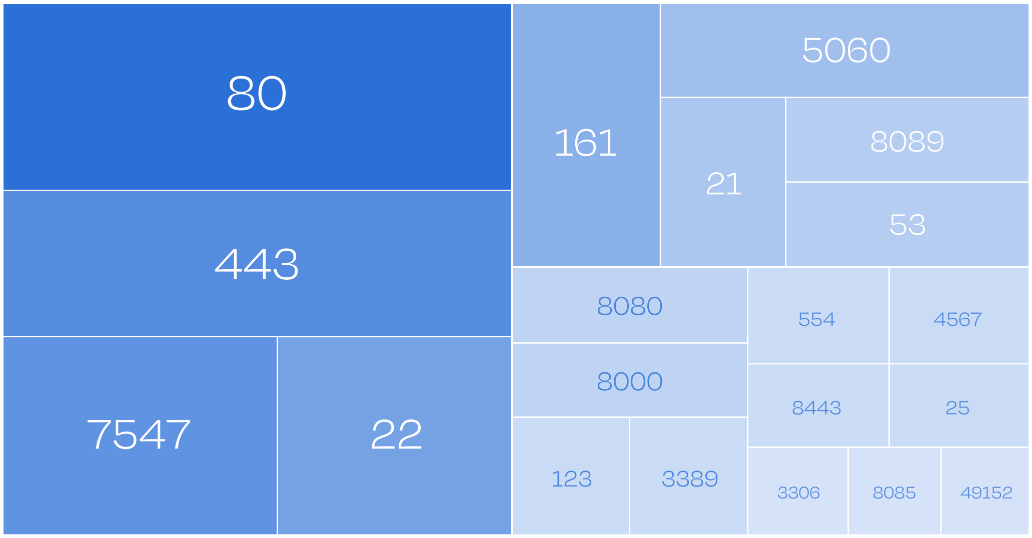
When scanning, we used our internet-scanning platform called Coalition Control. This platform continuously scans the entire IPv4 space and parts of the IPv6 multiple ports per month. Our platform first starts a round of TCP-SYN scanning across all IP addresses, followed by service identification and protocol enrichment scanning depending on the port or service being scanned (Figure 5.1).

Figure 5.1



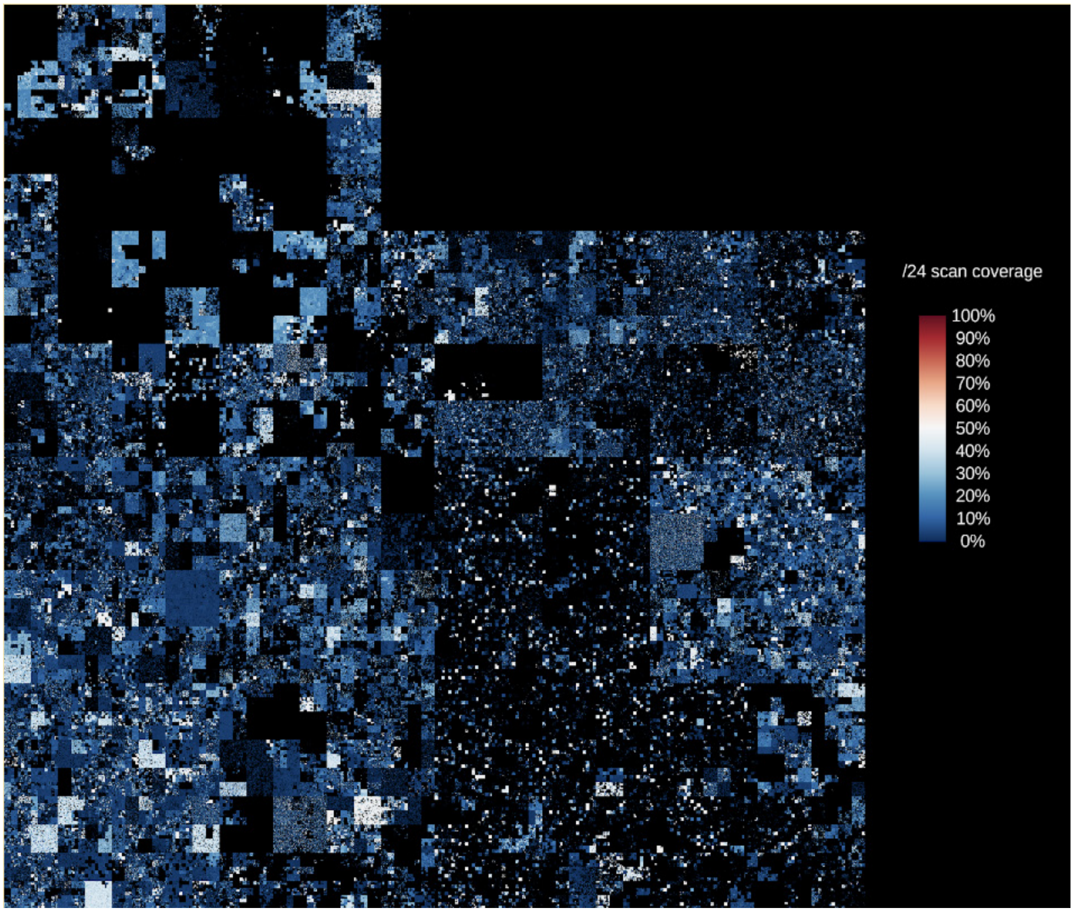
Our scanning infrastructure is geo-distributed across multiple countries and providers and uses custom task distribution and scanning modules built in-house. We collect data from more than 220 ports every 30 days, including all protocol enrichments for services running on different ports (for example, we collect all SSH keys, algorithms, and ciphers supported for all SSH servers we find running vs. just collecting the version of the SSH server). The top ports identified with services were **80**, **443**, **7547**, **22**, and **161** (Figure 5.2).

Top 20 Ports Found with Services Open to the Internet *(Figure 5.2)*



One interesting point about scanning the internet is that although we have to check approximately 5.2 billion IP addresses, only a subset of those have any type of services running on them, as seen by the following heatmap that represents the entire internet (Figure 5.3). All the networks that appear "lit up" are ones that had IP addresses that replied to our scanning at least once.

Figure 5.3



Another important point is that we also have an ongoing blocklist of IP addresses that request not to be scanned—this is a small list, however, and does not impact our results.

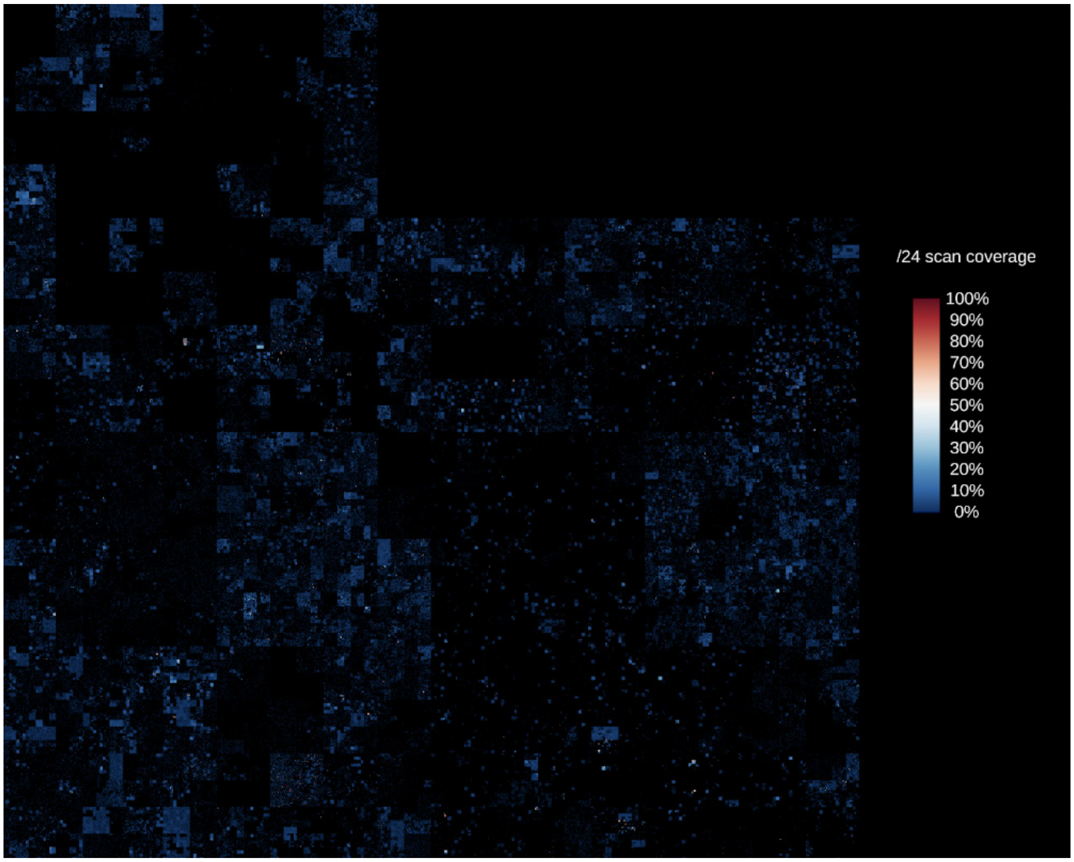
Of the over 5.2 billion IP addresses on the internet, only **441,760,930** were found running at least one service on the ports we scanned.

Sensors or Listening to the Internet

It is important to note that at Coalition, we do not just scan the internet. We have set up an extensive network of sensors that are geo-distributed across multiple locations and providers. Our sensors act as machines that appear unprotected against multiple known vulnerabilities or are running outdated software and appliances. Running these sensors gives us an idea of what attackers are doing and what is being scanned on the internet.

Additionally, our sensors have helped us discover new vulnerabilities that had not yet been publicly announced, variants of existing vulnerabilities, and how attackers use said vulnerabilities to carry out cybercrimes (Figure 5.4).

Figure 5.4



Coalition®

coalitioninc.com

55 2ND STREET, SUITE 2500
SAN FRANCISCO, CA 94105

Any report or similar presentation posted on info.coalitioninc.com/cyber-threat-index-report-2023 or elsewhere has been prepared by Coalition Inc. (defined below) solely for information purposes. You are advised to read this disclaimer carefully before accessing, reading, or making any other use of a report or similar presentation. The information presented as part of this report or similar presentation and any material published in relation to this report or similar presentation are provided for general purposes only and should not be construed as advice of any kind or the rendering of consulting, financial, legal, tax, or other professional services from Coalition Inc. or any of its subsidiaries, affiliates, associates, employees or representatives (collectively, "Coalition"). Coalition makes no representation or warranties about the accuracy or suitability of information provided in report or similar presentation and related materials (such as handouts, presentation documents and recordings). The content of this report may not apply directly to specific circumstances. Professional advice should be sought before any action is taken in relation to information disseminated during the report or similar presentation. This report or similar presentation does not constitute a contract for consultancy or advice and this report or similar presentation does not establish any contractual relationship with Coalition. You should consult with an attorney, consultant or other professional to determine what may be best for your individual needs. The opinions expressed in a report or similar presentation are those of the speaker(s) only and may not necessarily represent the views of Coalition. To the extent permitted by law Coalition exclude any and all liability for any loss claim or damage, cost or expense including any indirect or consequential damages or lost profit whether arising in negligence or otherwise suffered in connection with the access to participation in or use of the report or similar presentation by you or any other person. By acceding to this report or similar presentation you acknowledge that the information and materials contained in the report or similar presentation may contain inaccuracies or errors. The content of the information provided in the report or similar presentation is for your general information and use only. It is subject to change at any time and without notice. The report and similar presentation may include links to other resources or websites. These links are provided for your convenience only and do not signify that Coalition endorses, approves or makes any representation or claim regarding the accuracy of copyright compliance legality or any other aspects of the resource(s) or website(s). Further you are advised not to send any confidential or sensitive information to Coalition in the context of a report or similar presentation. Copyright © 2023. All rights reserved. Coalition and the Coalition logo are trademarks of Coalition, Inc. or its affiliates.