



2022 State of the Internet Report

Introduction

The Internet has revolutionized how we communicate, share information, and conduct business. Remote work arrangements, cloud adoption, and zero trust deployments have increased the Internet exposure of most organizations. While organizations' technological footprint once resided behind the closed doors of corporate networks, shifts in the way we use technology and increased connectivity have ushered in an era of Internet exposure.

The use of services and devices not approved by IT teams, or shadow IT, has increased over the last 2 years, driven in large part by new remote work demands. This increased connectivity beyond the purview of IT and Security teams now poses additional risk to organizations in the form of improperly managed devices and services connected to the Internet.

Censys maintains the most comprehensive view of assets on the Internet by continuously scanning the public IPv4 address space across the 3,500 most popular ports. This data is freely available via [Censys Search](#), and we provide data for [researchers](#) and [enterprise security teams](#). This Internet-wide scan data also powers our Attack Surface Management (ASM) product, which comprehensively maps organizations' Internet exposure. This data allows us to understand broader trends in Internet security and how organizations are exposed across the Internet.

As researchers and security practitioners ourselves, we wrote this report with the goal of sharing our visibility with the wider security community. We're eager to share what we've learned, and we hope it's useful and informative to security practitioners, executives, and enthusiasts alike.

Executive Summary

- **Misconfigurations**—including unencrypted services, weak or missing security controls (Content Security Policy (CSP), etc.), and self-signed certificates—**make up roughly 60% of the risks we observe across the Internet.** Exposures of services, devices, and information represent **28%** of observed risks in our data, and **Software Vulnerabilities** represent **12%** of risks observed in 2022.
- With so many organizations migrating services to the cloud, there's a lot of attention on cloud security and exposure. However, there's still significant exposure risk for on-premises infrastructure. **The majority of the Internet hosts and services do not run on a major cloud provider, but rather are hosted on-premises or in a conventional datacenter.** Despite increasing cloud adoption, Internet exposure isn't just a cloud problem.
- Vulnerability management continues to pose challenges. [Research suggests](#) that generally, it takes over 200 days to patch severe vulnerabilities, and we observed **three distinct types of behavior in response to vulnerability disclosures:** near-immediate upgrading (Log4j), upgrading only after the vulnerability is being actively and widely exploited (GitLab), and near-immediate response in the form of taking the vulnerable instance off the Internet entirely, or in other cases, patching (Confluence).
- Organizations have an average of **44 different domain registrars and presence in 17 different hosting providers** (including cloud, datacenter, and on-premises equipment). A reported [59% increase](#) in shadow IT, driven by remote work demands over the last 2 years, has likely contributed to this sprawl. Organizations must continue enabling their employees, but this can lead to visibility issues when IT and Security teams are left out of the conversation.

What's in the Report?

The Internet as a Whole:

This section serves as a macro view of the entire Internet. We examine popular services, the standard and non-standard ports where they run, and autonomous systems where they're hosted.

What's Out There?

We explore some of the major services that run on the Internet—HTTP, SSH, and FTP—at a high level.

Where Do Services Run?

Most of the Internet is still not in one of the major clouds we studied. We examine where hosts and services run on the Internet beyond these clouds.

The Attack Surface of the Internet:

Using our Internet-wide scan data and risk fingerprints that power our Attack Surface Management (ASM) platform, we dive into risks and vulnerabilities on the Internet.

Risk and Vulnerability Overview

Using our risk detection fingerprints, we examine the most commonly observed Censys-visible exposures, misconfigurations, and vulnerabilities across two point-in-time samples of over 4 million hosts.

The Internet's Response to Major Vulnerabilities

Profiles of how the Internet responded to three major vulnerabilities from the last year and a half. We look at how quickly vulnerable services were patched and upgraded, how many vulnerable instances remain, and instances of services being taken off of the public Internet altogether in response to vulnerability disclosure.

Attack Surfaces of Organizations:

We used our ASM platform to generate attack surfaces for 37 medium and large organizations to better understand companies' public-facing Internet footprints.

How to Read This Report

This report does not need to be read in order: each section stands on its own and can be read independently of the others. At the beginning of each section, you'll find a box like the one below with high-level summary statistics. **At the end of each section, you'll see takeaways and suggestions for action items for IT and Security practitioners.** You can skim these to obtain a high-level summary, or you can read on to dig into the details. Either way, we've got you covered.



We recommend readers start with the Introduction and Glossary, but beyond that, this is a bit "choose your own adventure."

While we hope that readers find all the contents of the report useful, we recognize that certain sections may be of particular interest to different audiences. Below are some suggestions on where to begin, depending on your role or area of focus:

CISOs and other executives

Attack Surfaces of Organizations provides insight into the Internet-facing attack surfaces of 37 medium and large companies.

Security engineers, researchers, analysts, and other practitioners

Risk and Vulnerability Overview examines the top risks and vulnerabilities on the Internet. If you feel overwhelmed with **gestures wildly** in security and are unsure of how to prioritize patching or security-related maintenance, this section may help inform those efforts.

The Internet's Response to Major Vulnerabilities provides our perspective on how responders addressed several major vulnerabilities over the last year and a half.

Glossary

Attack Surface:

[NIST defines](#) an attack surface as “the set of points on the boundary of a system, a system element, or an environment where an attacker can try to enter, cause an effect on, or extract data from.”

Autonomous System (AS):

An AS is a group of hosts with the same routing policy, managed by one network operator. ASes help route traffic across the Internet. Each AS receives an Autonomous System Number (ASN) as an identifier.

CVE:

[Common Vulnerabilities and Exposures](#). This is a list of publicly disclosed cybersecurity vulnerabilities. Each vulnerability receives a CVE ID in the following format: CVE-YYYY-NNNN, where YYYY is the year of initial request for disclosure of the vulnerability, and NNNN is the number it was assigned.

Host:

A computer or other device connected to the Internet.

Service:

The system running on a host that can receive and communicate on the Internet. Services are usually identified by the OSI-model L7 (application) protocol that they use for communication, although Censys identifies some specific services that run on top of HTTP (e.g., Elasticsearch, CWMP, etc.). A few specific services we'll discuss include:

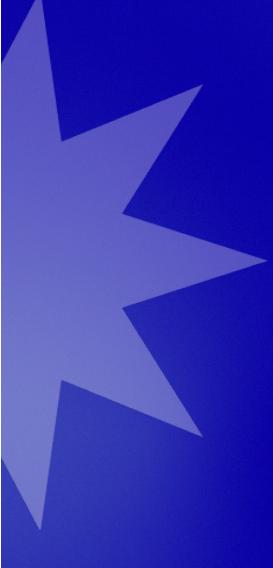
- **HTTP**, or Hypertext Transfer Protocol, is the service used for data transfer between web hosts/websites and web servers.
- **SSH**, or Secure Socket Shell, is a secure network protocol that enables secure remote access and file transfer between systems on a network.
- **FTP**, or File Transfer Protocol, is used to transfer files from a server to a client on a network.

Universal Internet Dataset (UIDS):

Our Internet-wide scan dataset. It is derived from continuous scanning of the entire IPv4 space on over 3,592 ports, as well as:

- **More Frequent Global Scan of Popular Ports.** We scan the whole IPv4 space on 137 ports with IANA-assigned services every day.
- **Cloud Provider Scans.** Since many cloud hosts are ephemeral, we scan the 1,440 most popular ports on Amazon, Google, and Azure hosts every day.
- **Global Scan of Less Popular Ports.** We scan the whole IPv4 space on 3,455 additional ports on a regular basis, completing a walk every 10 days.
- **Global Scan of Every Other Port Number.** We scan the entire IPv4 address space across ALL ports (65535) at a low background rate.

The Internet as a Whole



HTTP, SSH, SMTP, and FTP represent 87% of the services on the Internet; HTTP alone represents 81%. But HTTP isn't just public websites, it includes APIs, web-based control panels for Internet-connected devices, and more.

We note that an interesting FTP configuration (running on port 40029) and prevalence of the 3DES-CBC SSH cipher seems to be concentrated among Asian ASes. This suggests regional differences in common configuration practices.

Over 10 million SSH services continue to support ciphers with known vulnerabilities like 3DES.

We begin by taking a broad look at the Internet. Using a single daily snapshot per month from June 2021 to March 2022—an average of 220,763,081 hosts per snapshot—we start analyzing Internet-wide trends. Specifically, we examine which ports, services, and software are most prevalent on the Internet, and the autonomous systems and regions where they run.

What's Out There?

HTTP

HTTP, or Hypertext Transfer Protocol, is the protocol used for data transfer between clients, like web browsers, and servers.

HTTP is overwhelmingly the most common service observed across the Internet. HTTP is found on 222 million hosts and makes up 81% of Internet services. In addition to websites and web servers, HTTP includes APIs, caches, proxies, and web-based control panels for Internet-connected devices.

THE INTERNET AS A WHOLE

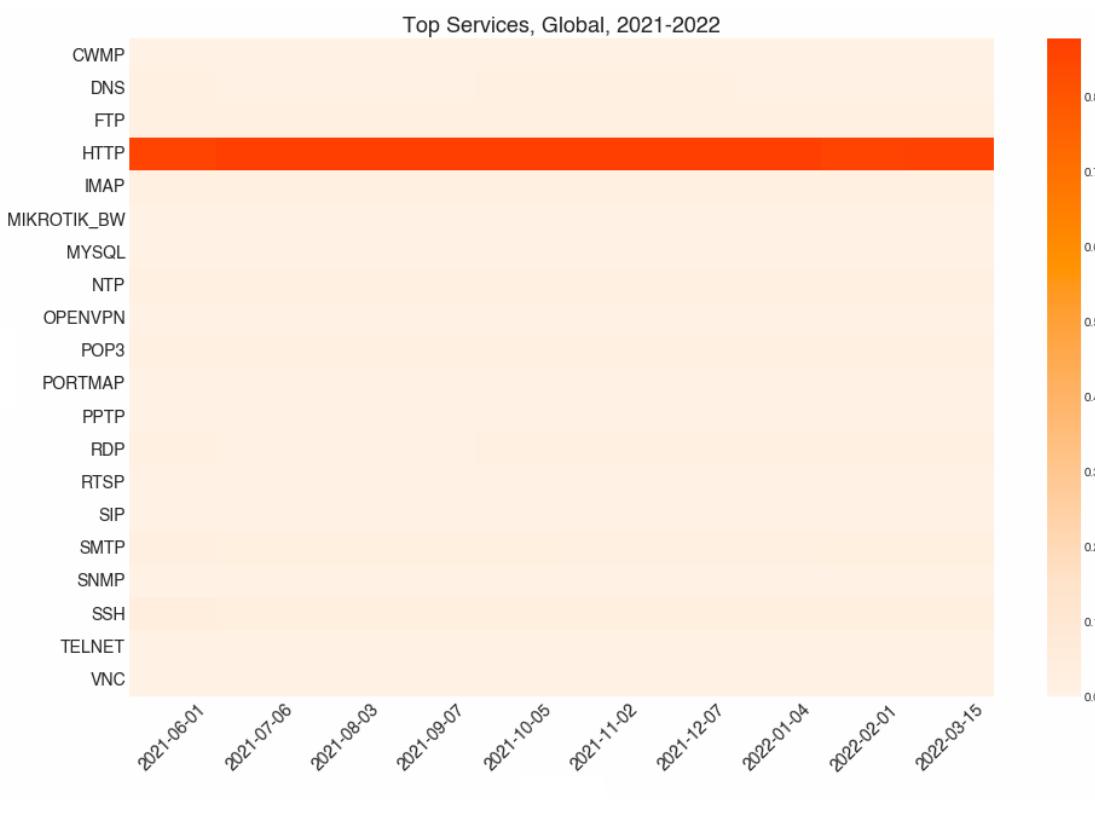


Figure 1: Breakdown of top services on the Internet across 222 million hosts. This is not a bar chart, but rather a heatmap over 10 monthly snapshots of Internet-wide scan data. HTTP represents, by far, the most services on the Internet.

While HTTP is most commonly associated with TCP/UDP ports 80 and 443, **the vast majority of HTTP doesn't run on standard ports**—we observe it running across the widest range of ports of any service. Only 7% of HTTP services observed by Censys run on port 80, while 5% run on port 443.

The next most commonly observed ports running HTTP services are 7547 (2%) and 30005 (1%). These percentages may seem low, but 1% and 2% of millions of services still represent a substantial amount of HTTP.

While 80, 443, 7547, and 30005 are the most popular ports for HTTP, 85% of HTTP services on the Internet do not run on one of these ports. Scanning only a few ports to search for HTTP, even if scanning the most popular ones listed above, results in omission of the bulk of HTTP services on the Internet.

Many services have an [IANA-assigned default port](#), though services are often set to run on non-standard ports.

While running services on non-standard ports is not in itself a risk, [prior work has shown](#) that services on non-standard ports tend to be less secure in practice.

Moreover, running services on non-standard ports can provide a false sense of security, especially if the service owner is relying on [security through obscurity](#) to protect their assets.

[CPE WAN Management Protocol \(CWMP/TR-069\)](#), which uses HTTP, often runs on ports 7547 and 30005. This protocol is often used by Internet Service Providers (ISPs) to configure consumer routers. Exposure of CWMP to the Internet can be problematic, as [several vulnerabilities](#) and misconfigurations in CWMP have been exploited in the past.

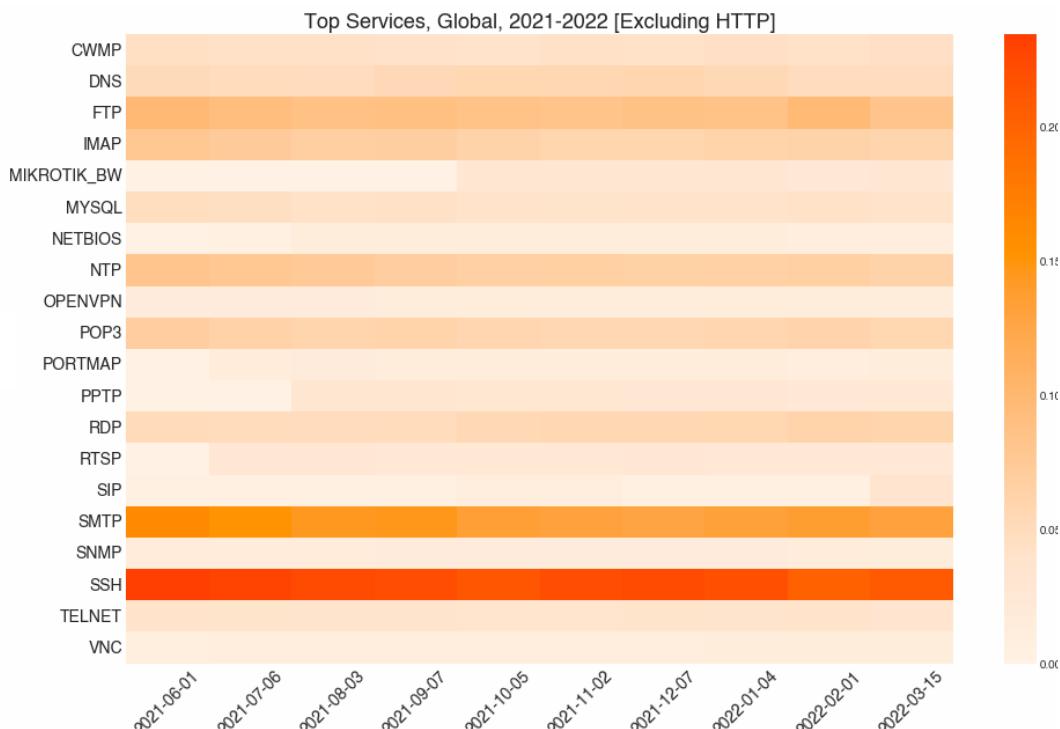


Figure 2: Breakdown of top services on the Internet across 222 million hosts, sans HTTP.

SSH

SSH (Secure Socket Shell), is a secure network protocol that enables secure remote access and file transfer between systems on a network.

In contrast to the distribution of HTTP services across many ports, 75% of SSH services run on the IANA-assigned port 22. 25% of SSH services run on a non-standard port.

75% of SSH services observed by Censys support AES or ChaCha20 ciphers (most of which are shown above the gray horizontal line in the graph below), which are currently recommended secure cipher options. However, as of March 15, 2022, we observe over 10 million SSH services using 3DES-CBC (5% of all observed SSH services), which has been recommended against by [NIST since 2017](#).

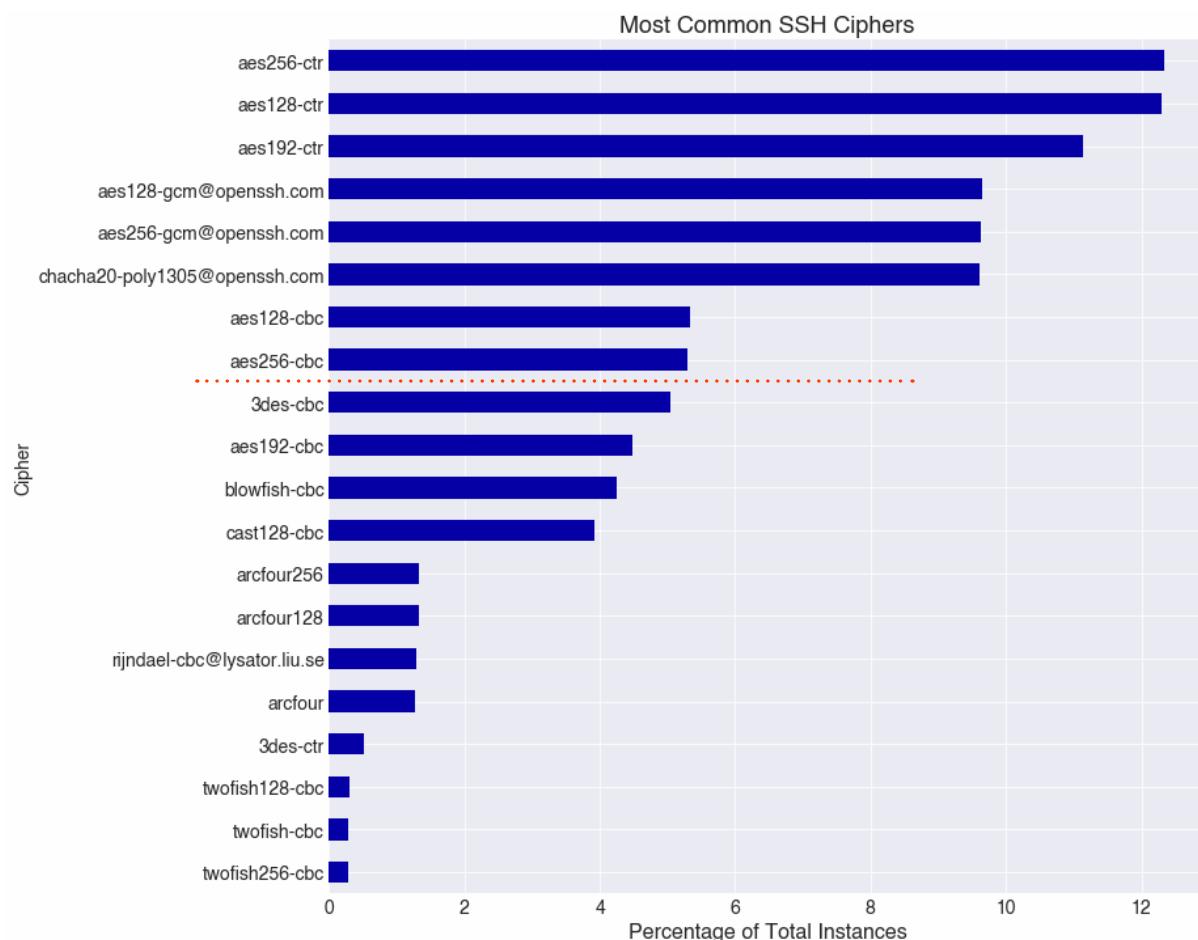


Figure 3: Most common SSH ciphers. Ciphers above the dotted orange line represent 75% of the versions we see.

SSH servers in Asia disproportionately rely on 3DES ciphers. Amazon accounts for the bulk of hosts, but CHINANET-BACKBONE, Korea Telecom, CHINA169, and ALIBABA also have strong 3DES presence.

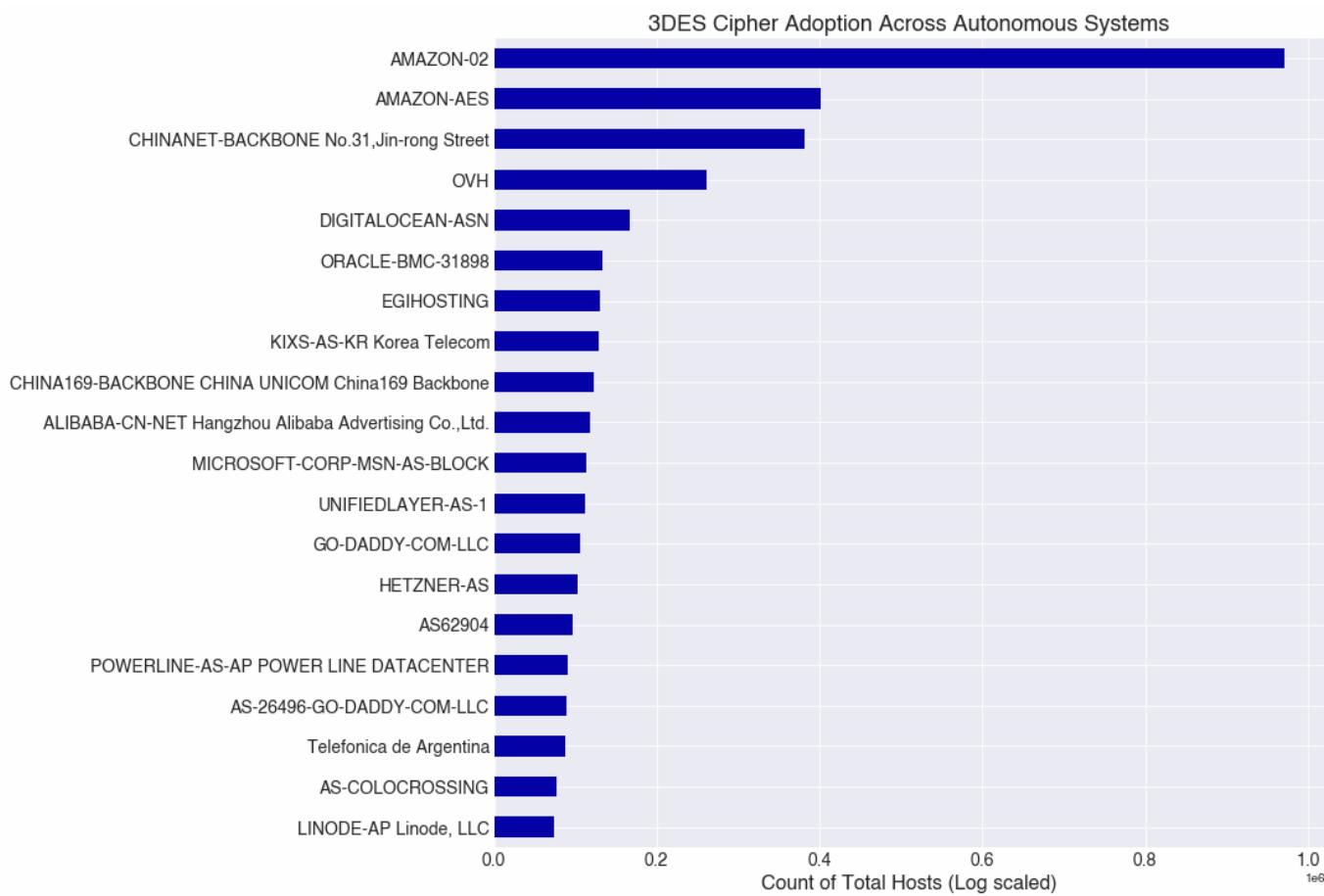


Figure 4: Host counts for top Autonomous Systems using 3DES-CBC SSH ciphers.

FTP

FTP, or File Transfer Protocol, is used to transfer files from a server to a client on a network.

When examining FTP, we see that 84% of FTP services run on IANA-assigned port 21. The next most commonly observed port running FTP is 40029, which runs 3% of all FTP services Censys sees. This was a bit surprising to us, as 40029 doesn't have any [IANA-assigned service](#), and 40029 does not appear in any [FTP-related IANA assignments](#).

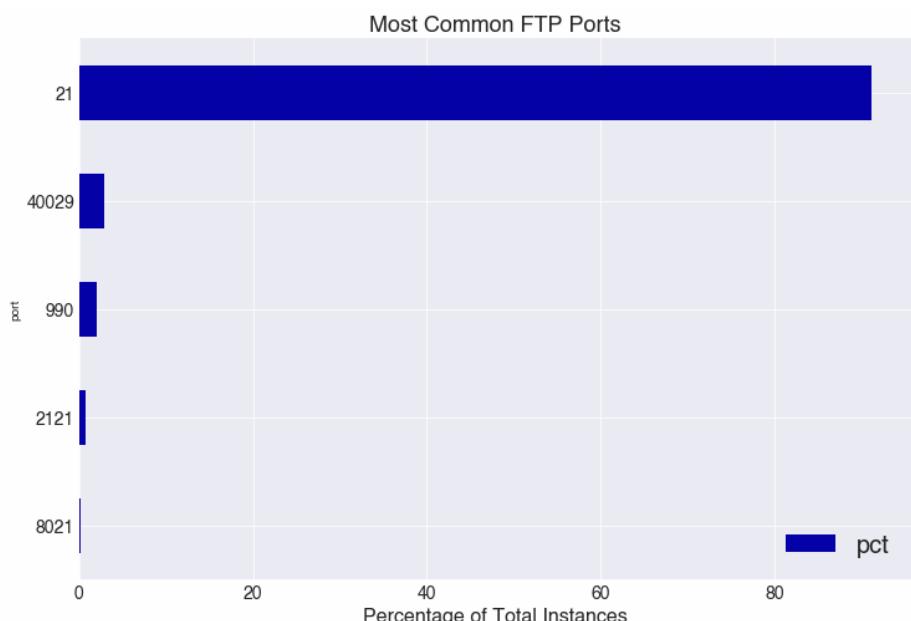
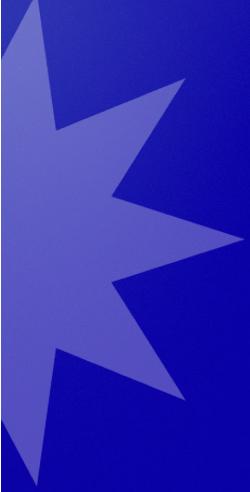


Figure 5:
Top 5 FTP ports.

The 40029 configuration appears to be related to [Alibaba Cloud instances](#). It's unclear whether this uncommon FTP configuration is related to a base image provided by Alibaba to its customers.

Our findings of this unorthodox FTP configuration and the use of 3DES SSH ciphers primarily on Alibaba Cloud instances may speak to regional differences in configuration practices and is something we hope to explore in future research.

Where Do Things Run?



In spite of increasing cloud adoption over the last few years, most of the Internet does not run in a major cloud. Only 9% of hosts with services run in one of the four major clouds we studied.

When examining hosts across four major US-based cloud providers, we see wide adoption of the region corresponding to the Eastern US, primarily driven by Amazon. Western Europe has strong Azure and Google presence. The Central US region is almost entirely driven by Google.

Now that we've looked broadly at popular services on the Internet, we shift to studying where they run on the Internet.

While the Internet has become increasingly reliant on several large cloud providers, Amazon, Microsoft (Azure), Google, and Oracle make up only 9% of all hosts with services on the Internet. This may initially seem surprising, but the sense that "everything is in the cloud now" is likely primarily driven by the value we assign to services that are in these clouds. If a major cloud experiences an outage, it's not just one business that's also experiencing an outage, but hundreds or thousands across many industries.

THE INTERNET AS A WHOLE

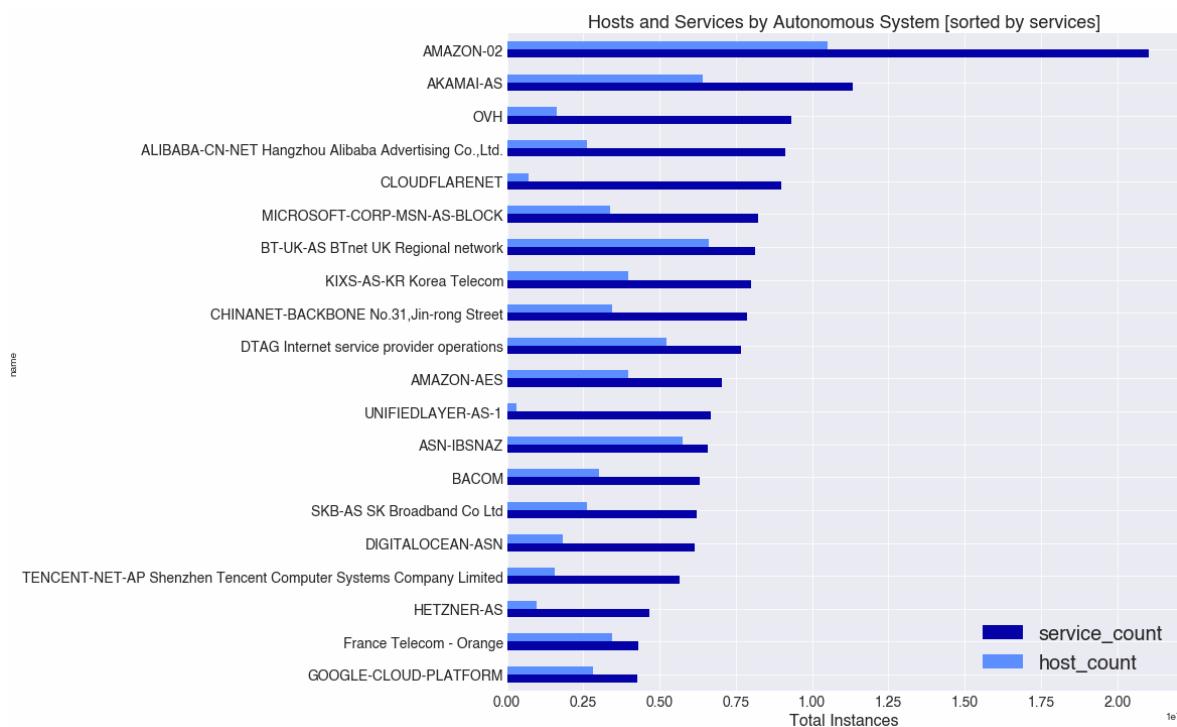


Figure 6: Hosts and services by Autonomous System, sorted by number of services.

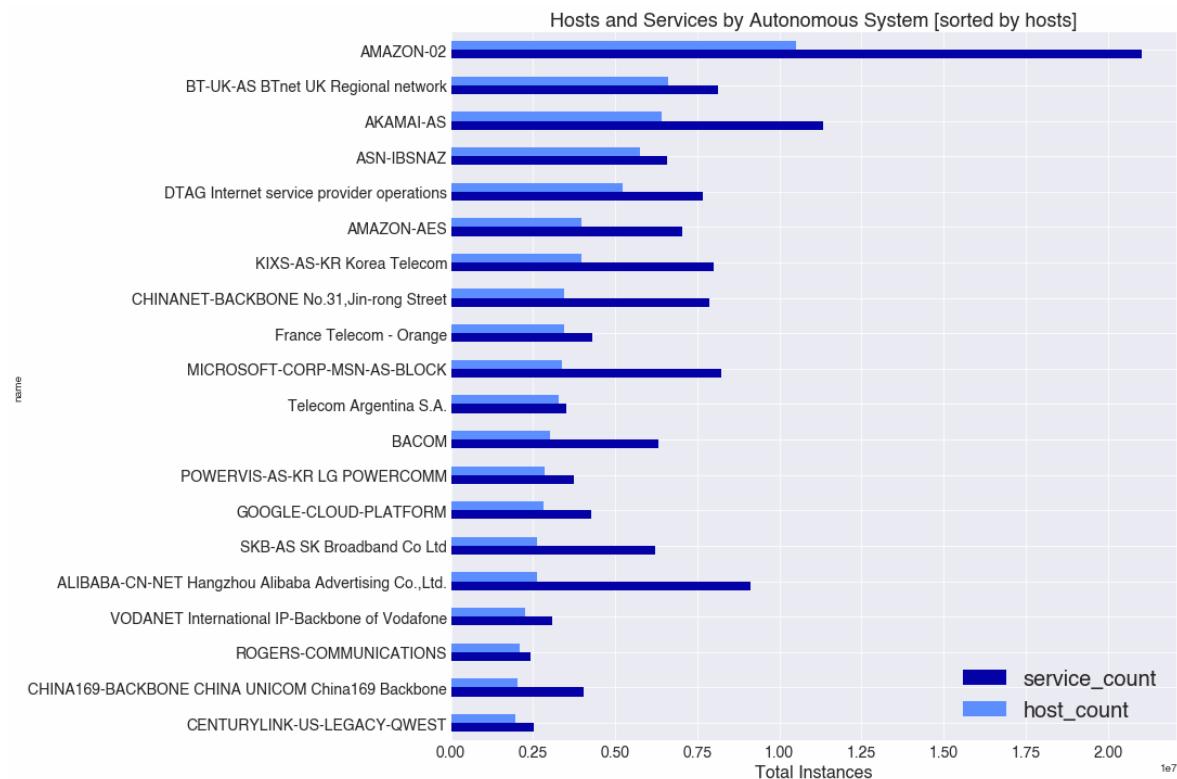


Figure 7: Hosts and services by Autonomous System, sorted by number of hosts.

We examined hosts across four US-based cloud providers and observed that the eastern US-based region tends to be the most popular region across these clouds, though that's driven by Amazon's us-east regions. Google, for instance, houses many hosts in their central US and western Europe regions. In addition to an eastern US region, Azure also has strong presence in western Europe.

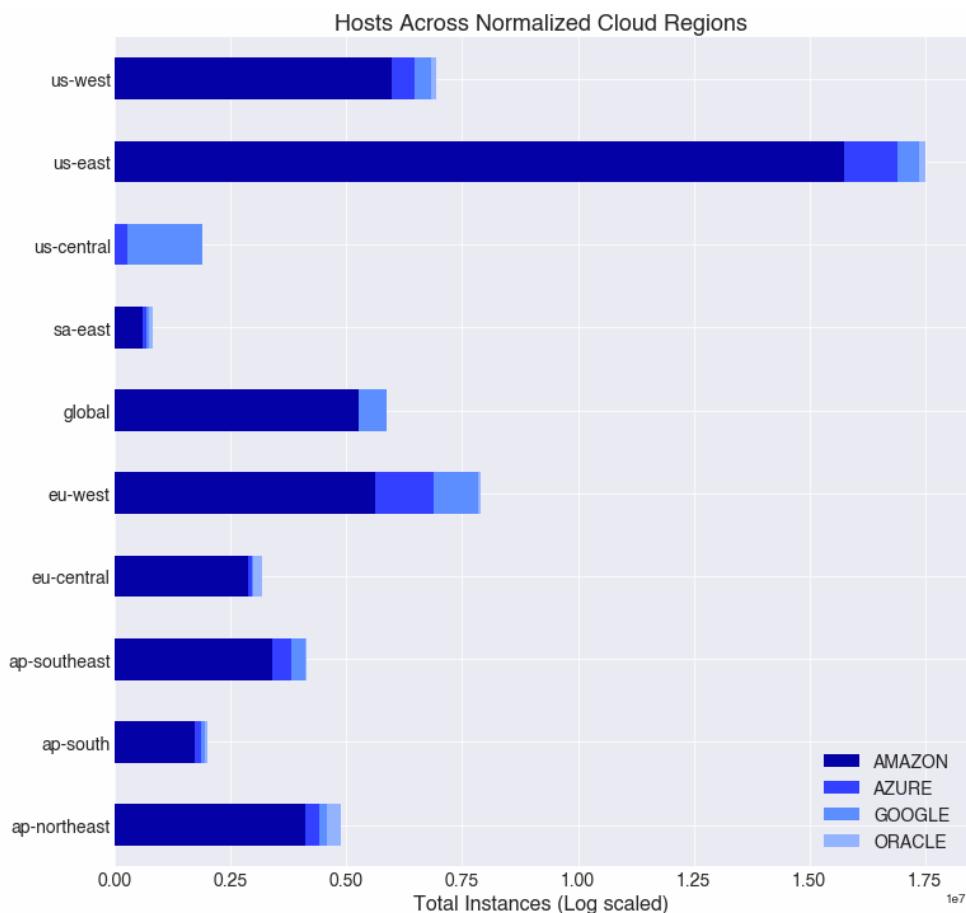


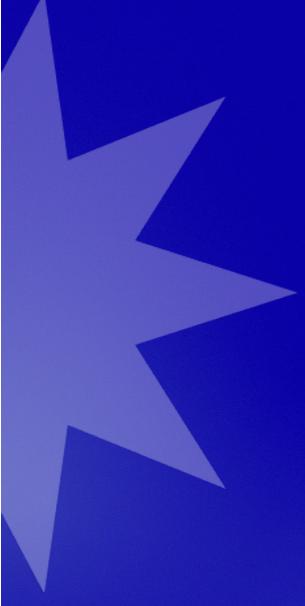
Figure 8: Presence of hosts across normalized regions for four major cloud providers.

Using cloud resources can make scaling and growth much easier for organizations large and small. A multi-region approach can help mitigate availability issues when an entire region goes down.

Takeaways

- Let's get this one out of the way: It's 2022. Security through obscurity is not a valid strategy for protecting your assets. Running services on non-standard ports won't keep threat actors from finding them.
- **If you've recently inherited infrastructure through a merger or acquisition, or perhaps haven't taken a look at the Internet-facing assets in your organization in a while, there's no time like the present.** Defenders, put on your offensive security hat and do some reconnaissance of your organization's Internet-facing footprint with the [Censys Search](#) interface. If you're concerned there are additional assets you don't know about (spoiler: there probably are), Censys's Attack Surface Management platform can help identify these assets so you can lock them down and get your sensitive services off the public Internet.
- **If it's possible, someone will do it.** Using deprecated SSH ciphers or exposing sensitive services to the Internet can be easy to do, especially without guard rails in place. **Find ways to make the "secure" option an easy default in your organization.**
- While having assets in multiple clouds or hosting providers can complicate your attack surface (as we'll see in [Attack Surfaces of Organizations](#)), a multi region cloud or datacenter strategy can be helpful in the case of an outage. Being able to failover to a functioning region helps ensure your services remain available for your employees and customers. Keeping track of that footprint can be tricky, and **it's important to regularly assess any new hosts or services that pop up in your infrastructure.**

The Attack Surface of the Internet



Misconfigurations and Exposures represent 88% of the risks and vulnerabilities Censys observes across the Internet. While CVEs and advanced exploits often make headlines, they represent just 12% of risks we observe on the Internet.

The Internet as a whole responds to different major vulnerabilities in varying ways. The Log4j vulnerability of December 2021 saw quick, widespread remediation. In contrast, remediation for the GitLab remote code execution vulnerability (CVE-2021-22205) announced in May 2021 didn't catch on until about 6 months later, when it was discovered that a botnet was exploiting the RCE.

As devices increasingly move online and digital infrastructure scales, so too does the number of risks that make up the Internet's attack surface. We examine the Internet's attack surface from two perspectives: risks and vulnerabilities across the Internet and the Internet's response to several high-profile vulnerabilities from the prior year.

Risks encompass settings or conditions (including vulnerabilities) that increase the potential for data breaches, information leaks, or destruction of assets.

While Censys has visibility into hundreds of risks and vulnerabilities, the reader should note that all data here represents 1) what our scanner can see, and 2) what we have detection fingerprints written for. We do not claim that these are "all the risks on the Internet" but rather the public Internet-facing risks and vulnerabilities visible to a non-intrusive network scanner like Censys. Lastly, the reader should also note that the Censys scanners never attempt an actual exploit against a resource, and as such, all information in this report is derived from our non-invasive scan data.

This data is biased toward risks and vulnerabilities that have some public Internet-facing artifact (e.g., a specific value in a banner message), but these are often the first things a threat actor will observe when doing reconnaissance on an organization.

Censys-visible Risks and Vulnerabilities on the Internet

We evaluated the presence of various risks and vulnerabilities across random samples of 2.2 million hosts from November 30, 2021, and 2 million hosts from roughly half a year later on June 10, 2022, all drawn from UIDS. To perform sample selection for each date, we joined UIDS with ASdb, a dataset that maps public autonomous systems (identified by ASN) to organizations and industry types. We then randomly selected 1% of hosts from each [ASdb](#) industry categorization to ensure representation of hosts across a variety of industries.

Notably, there was minimal change in overall observations across the two dates.

ACROSS ALL INDUSTRIES

At the time of this analysis, Censys had over 250 risk and vulnerability detection fingerprints. The graph below illustrates the spread of distinct risks across hosts in the various ASdb industries, not the amount of risks in each industry.

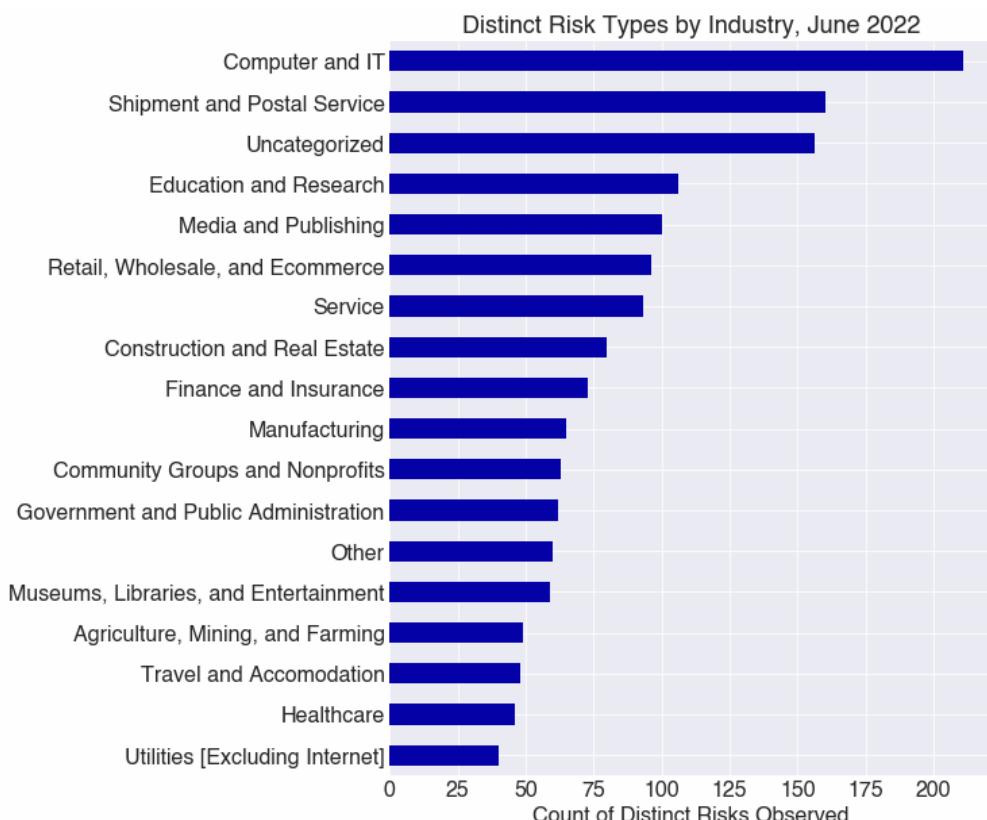


Figure 9: Presence of hosts across normalized regions for four major cloud providers.

When examining distinct risk types (i.e., widest spread of different risks), we see that the Computer and IT industry has the highest variety of risks, which isn't particularly surprising given the composition of this industry (ISPs, telecom providers, cloud providers). Shipment and Postal Services may seem out of place as the industry with the second-most varied range of risks, but ASdb categorizes two major Amazon ASes as Shipment and Postal Services, driving much of the risk variety here.

Across our 2021 and 2022 samples, misconfigurations make up roughly 60% of Censys-visible risks. For our purposes, 'misconfiguration' includes risks such as unencrypted services, weak or missing security controls (Content Security Policy, etc.), and self-signed certificates.

Exposures of services, devices, and information represent 28% of observed risks in our 2021 and 2022 data – this includes things like unintentional database exposures, exposed storage, IoT devices, exposed credentials, or API keys.

In contrast, vulnerabilities represent 12% of observed risks in our 2021 and 2022 snapshots. Vulnerabilities include end-of-life or outdated software and CVEs.

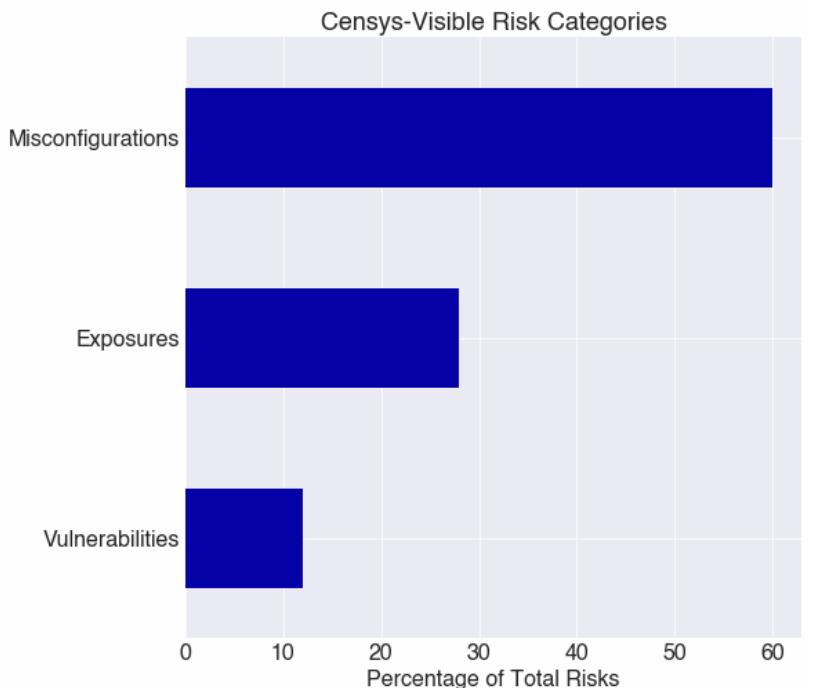


Figure 10:
Percentages of Censys-visible risk categories.

While CVEs, zero-days, and other advanced exploits dominate the security news cycle, the risks we most commonly observe are less provocative. Misconfigurations

and exposures collectively represent 88% of observed risks and are often best remedied through good security hygiene.

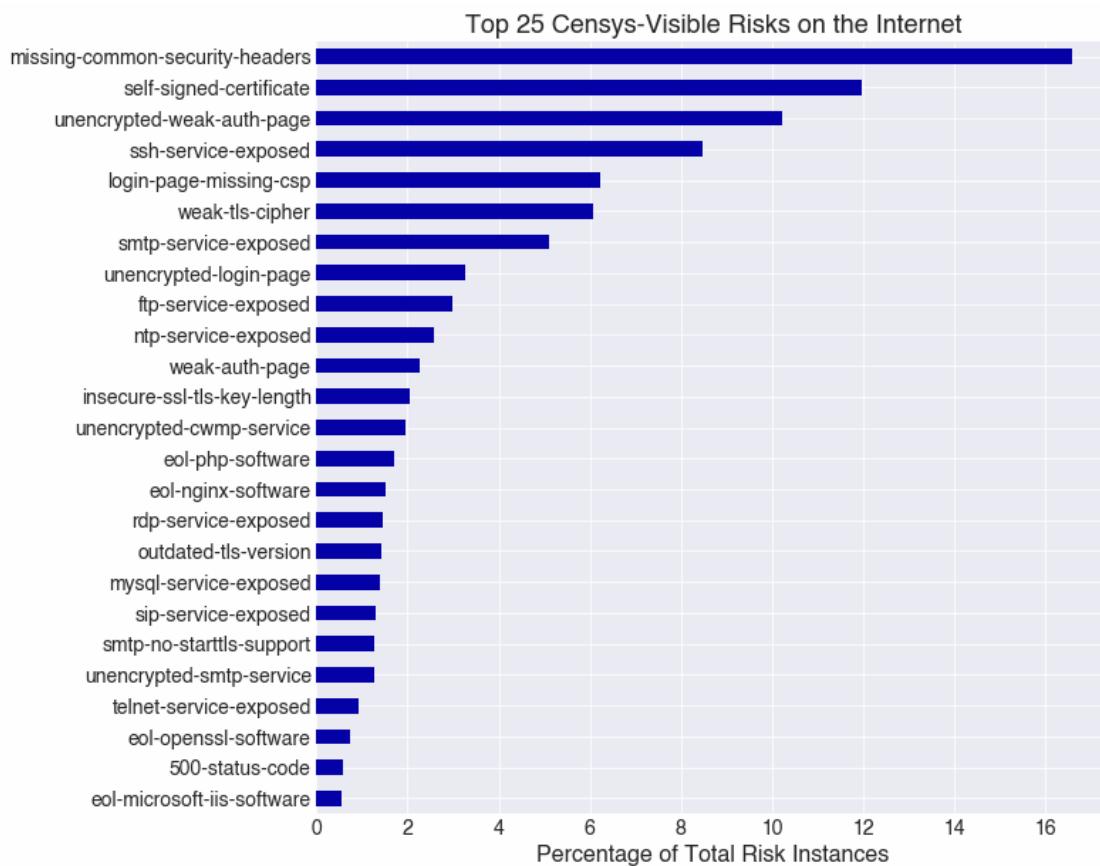


Figure 11: Percentages of Censys-visible risks across 2 million hosts, June 2022.

The top 3 risks observed across both snapshot dates (June data shown above) are **missing common security headers**, **self signed certificate**, and **unencrypted weak authentication page**, each of which we categorize as Misconfiguration.

Missing common security headers indicates that we did not detect any common security headers, such as [Content Security Policy](#) (CSP), [Cross-Origin Resource Sharing](#) (CORS), or [Strict Transport Security](#) (STS), on a service. Lack of these headers can make affected services a target for XSS or data injection attacks.

Self-signed certificate indicates that we discovered a certificate that was signed by its own private key instead of a trusted Certificate Authority. Services without identity verification are a target for man-in-the-middle attacks and phishing campaigns.

We categorize both of the above as low-severity risks, meaning that while exploitation of them may not lead a threat actor directly to an organization's crown jewels, they could be weaponized as part of an exploit chain or used to gather additional information about the organization.

Unencrypted weak authentication page represents just over 10% of the risks we observe in our Internet-wide sample. These authentication or login pages use basic or digest authentication without [TLS](#), making submitted credentials susceptible to interception and hash cracking techniques.

We categorize this as a high severity risk, as it can easily lead to credential theft. Moreover, Verizon's 2022 Data Breach Investigations Report indicates that "Use of stolen creds (Hacking)" is the top distinct Action variety (i.e., tactic) observed across their incidents and breaches dataset (p. 15). While credential theft is by no means a new tactic, it remains effective for threat actors.

Beyond these, many observed risks and vulnerabilities are exposures and end-of-life or deprecated software.

FINANCE AND INSURANCE (JUNE 2022)

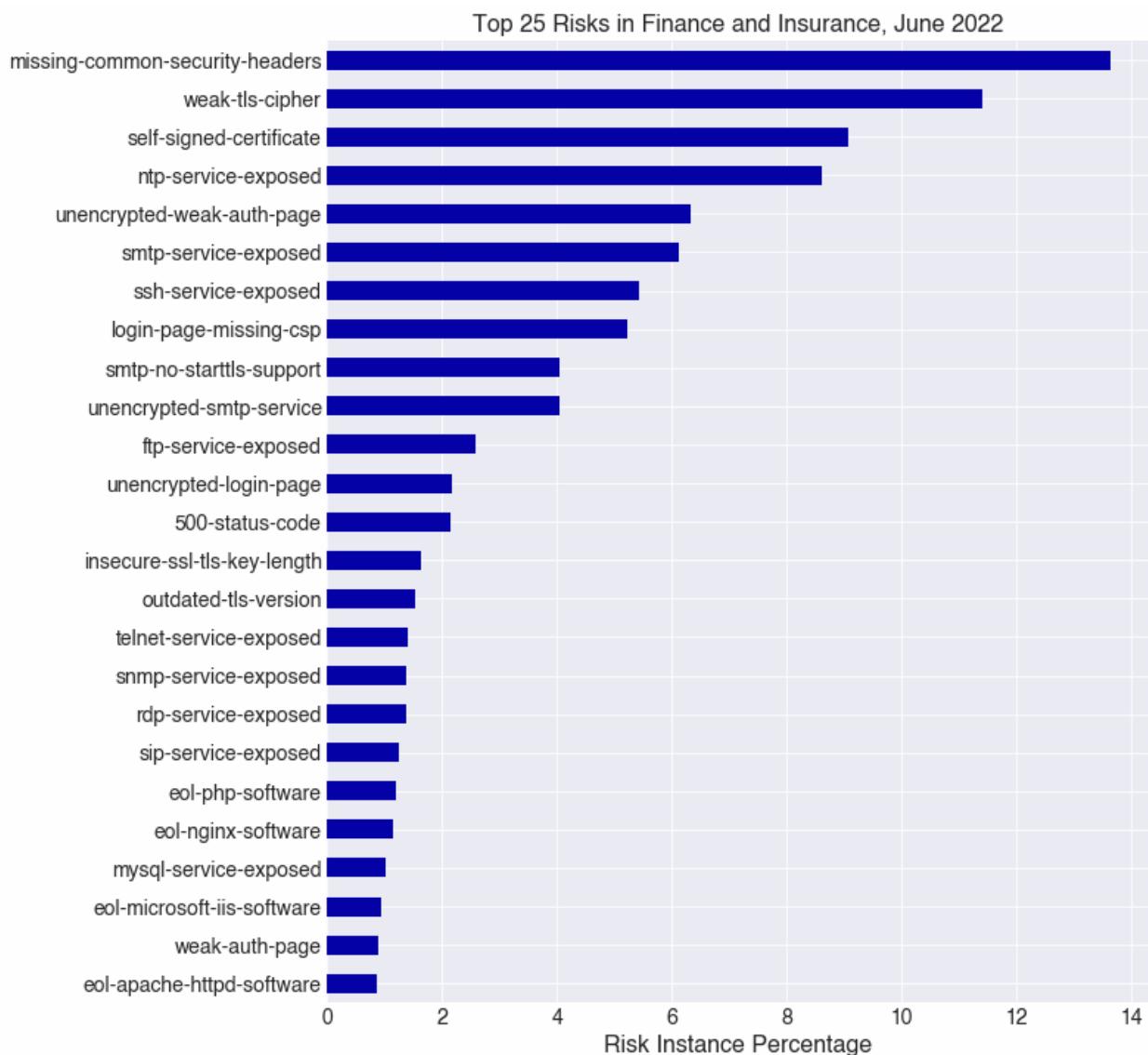


Figure 12a: Percentages of Censys-visible risks across hosts in the Finance and Insurance industry per ASdb, June 2022.

FINANCE AND INSURANCE (NOVEMBER 2021)

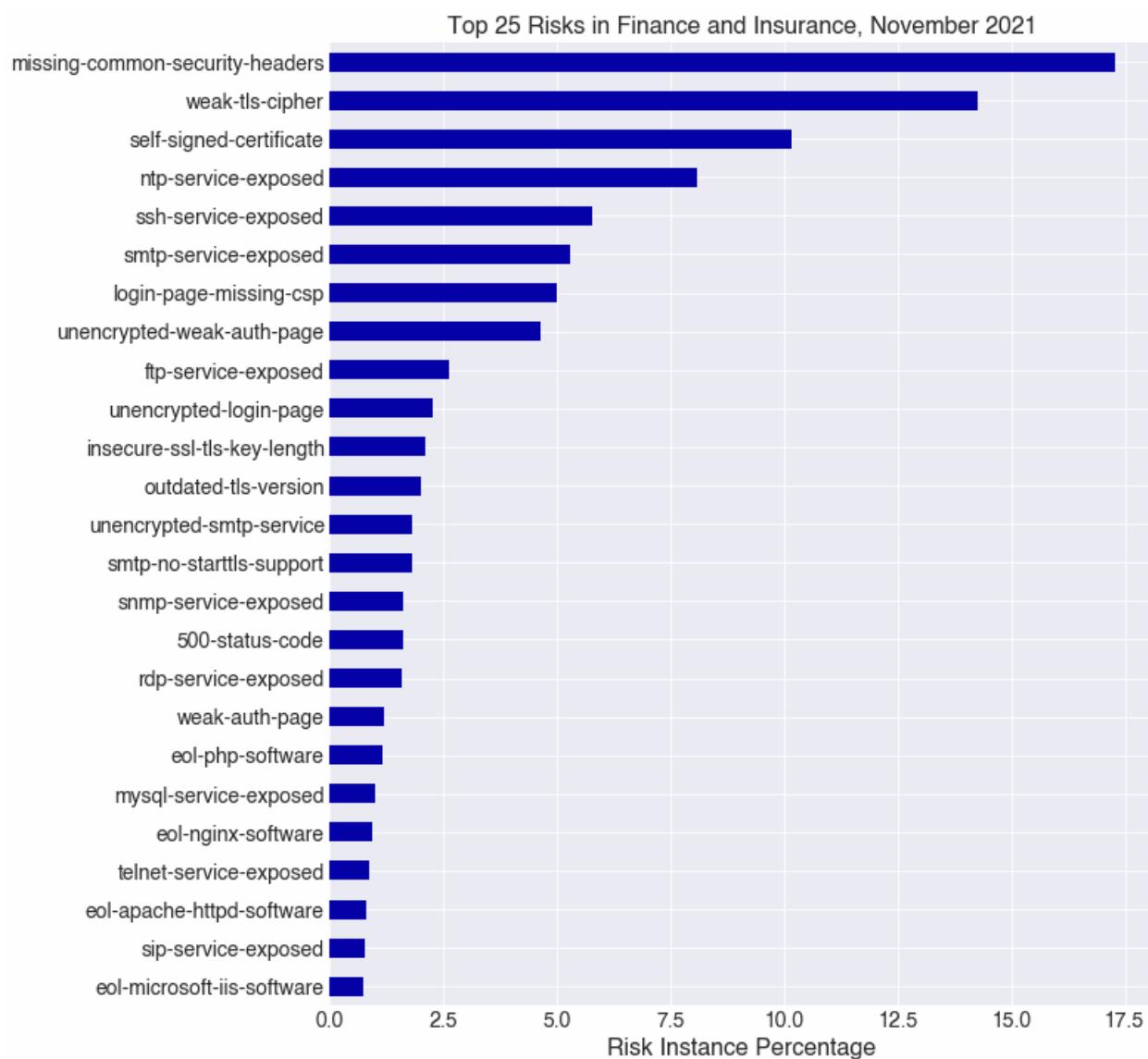


Figure 12b: Percentages of Censys-visible risks across hosts in the Finance and Insurance industry per ASdb, November 2021.

RETAIL, WHOLESALE, AND ECOMMERCE (JUNE 2022)

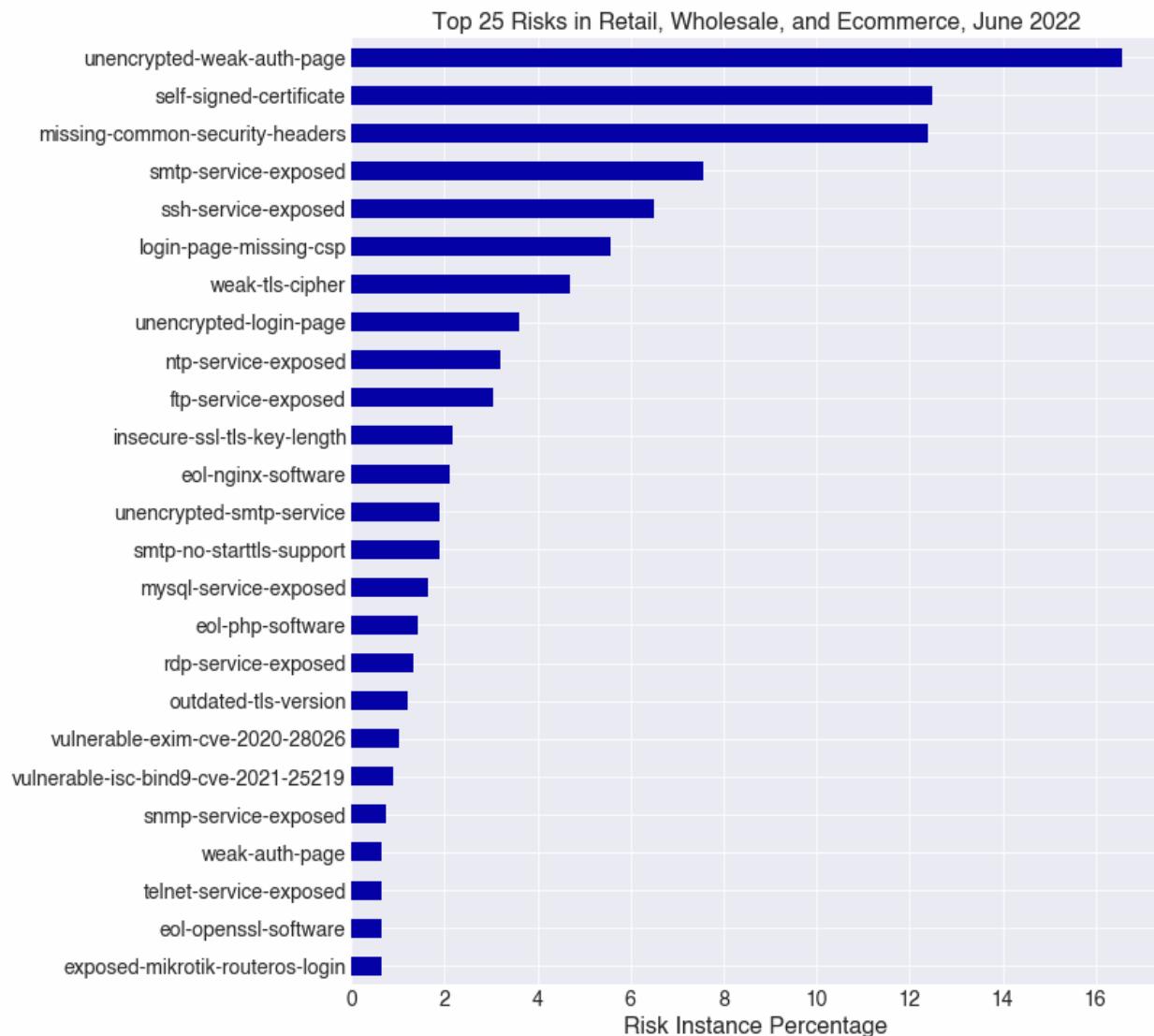


Figure 13a: Percentages of Censys-visible risks across hosts in the Retail, Wholesale, and Ecommerce industry per ASdb, June 2022.

RETAIL, WHOLESALE, AND ECOMMERCE (NOVEMBER 2021)

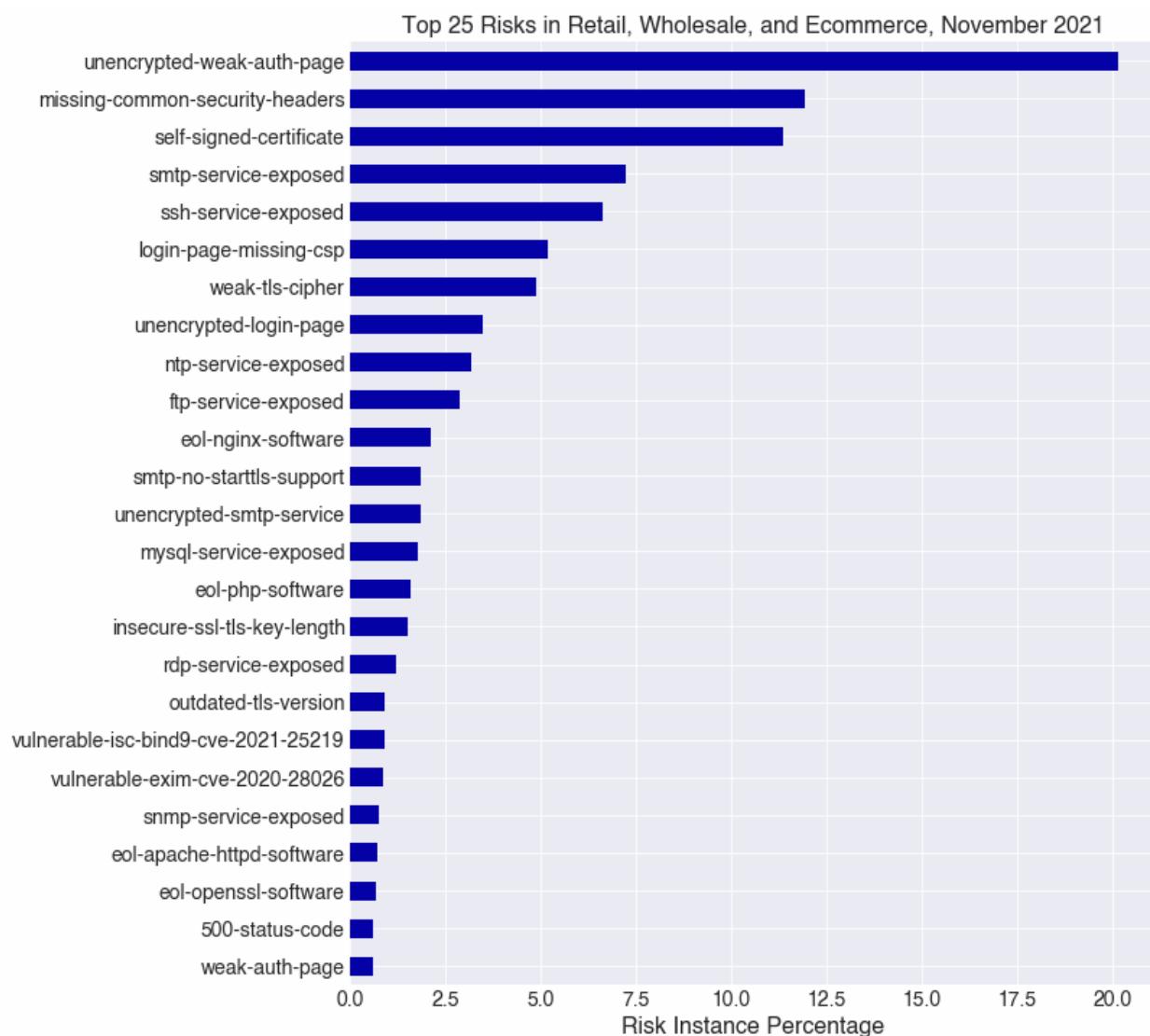


Figure 13b: Percentages of Censys-visible risks across hosts in the Retail, Wholesale, and Ecommerce industry per ASdb, November 2021.

UTILITIES [EXCLUDING INTERNET] (JUNE 2022)

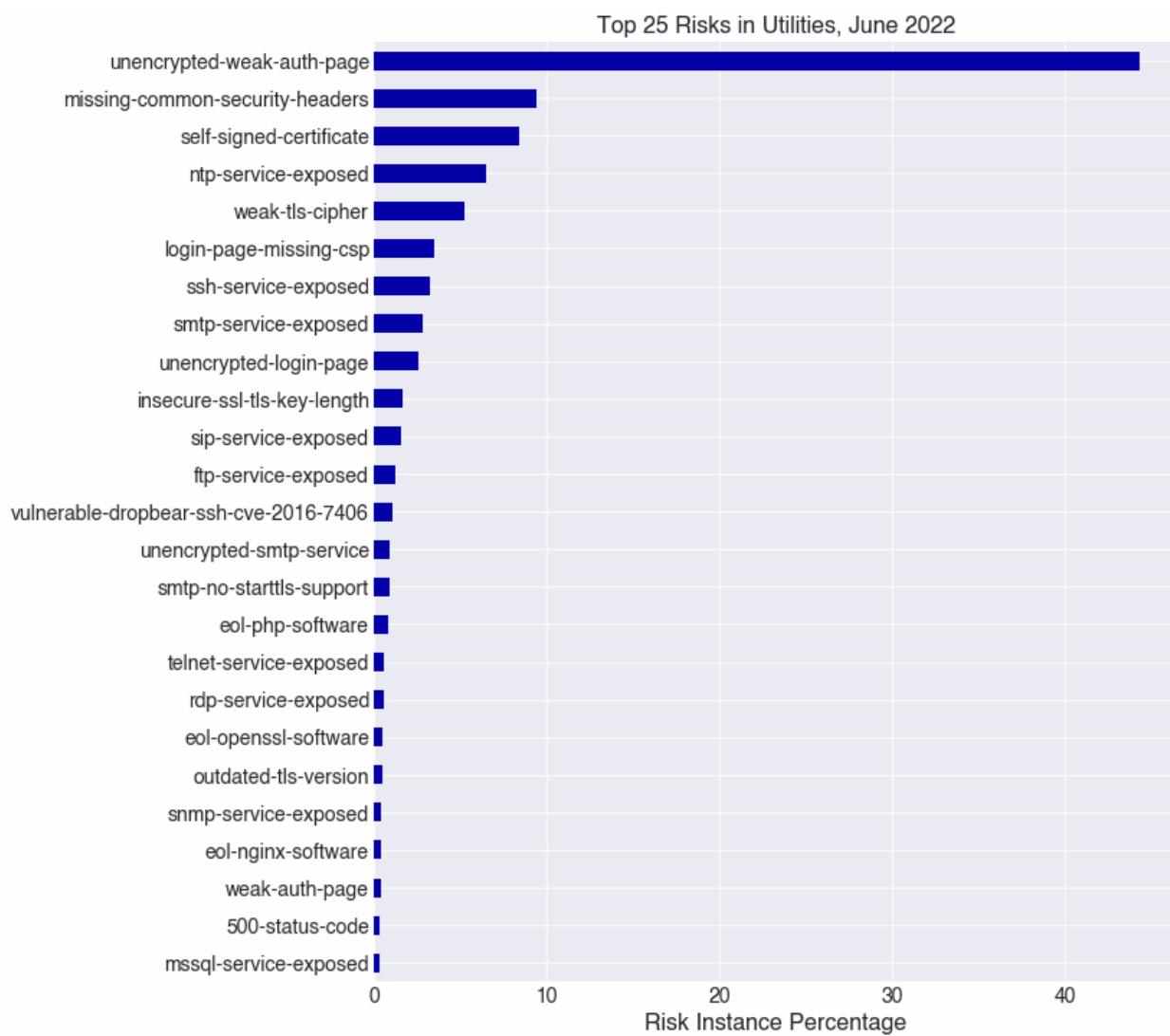


Figure 14a: Percentages of Censys-visible risks across hosts in the Utilities (excluding Internet) industry per ASdb, June 2022.

UTILITIES [EXCLUDING INTERNET] (NOVEMBER 2021)

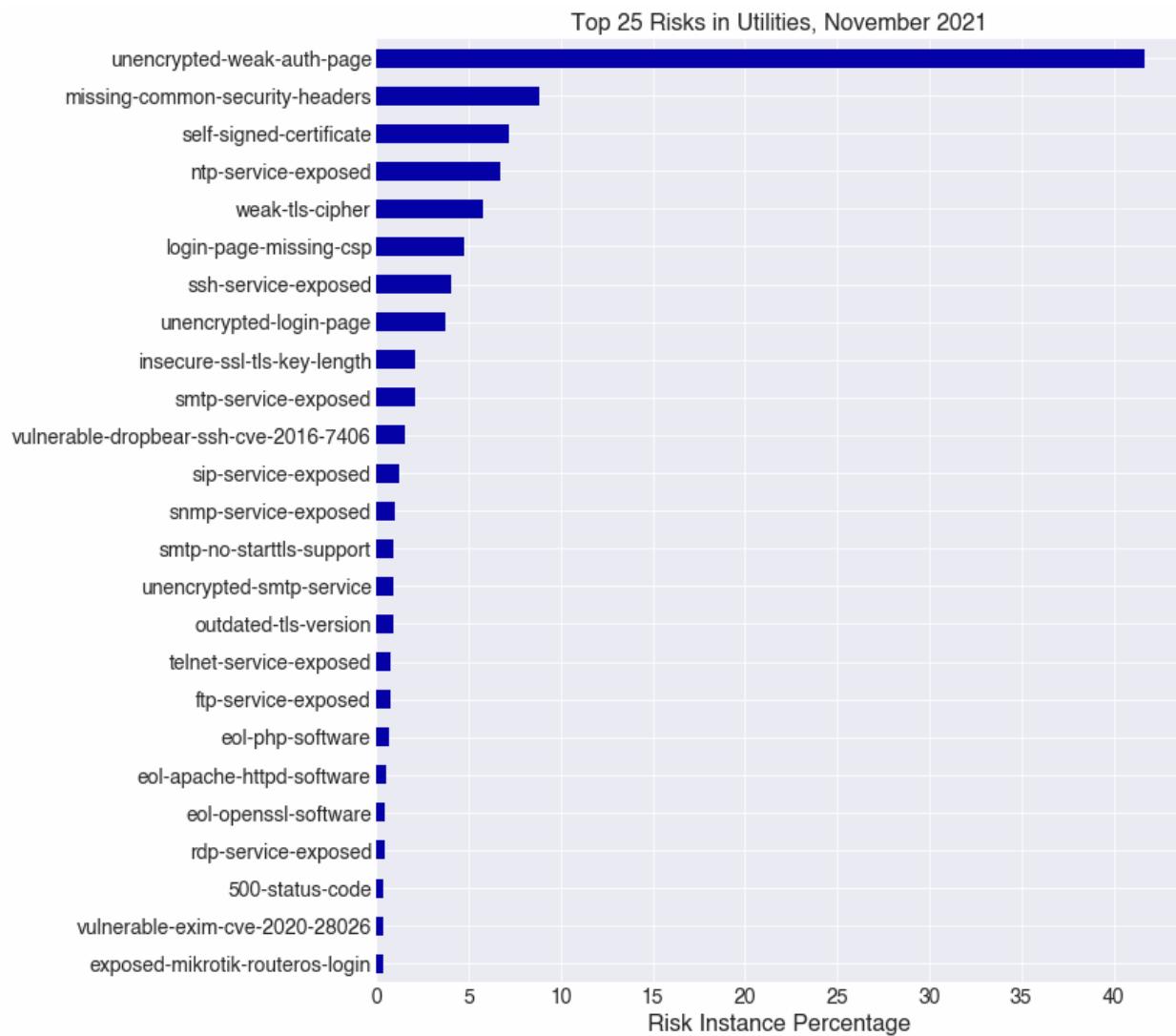


Figure 14a: Percentages of Censys-visible risks across hosts in the Utilities (excluding Internet) industry per ASdb, November 2021.

The risk profile of the Utilities industry stands out because so much of it is driven by unencrypted weak authentication pages. While unencrypted weak authentication page is one of the top three risks we observe overall, it represents over half of the observed risks for this industry—driven primarily by a US-based electric utility. With [increasing concern](#) over potential cyberattacks targeting Utilities, this particular risk could offer threat actors a relatively easy way into Utility networks unless remediated.

The Internet's Response to Major Vulnerabilities

New vulnerabilities are reported daily, and there is often extensive coverage of their mechanics—how the exploit works and how to detect and defend against it. Less widely understood is how the Internet responds to these vulnerabilities. How long does it take for vulnerable devices to be patched or upgraded? Do devices get patched or simply taken offline?

Here, we explore the Internet’s response to the Log4j remote code execution (RCE) vulnerability, the [GitLab RCE](#) vulnerability and botnet (CVE-2021-22205), and the Confluence [OGNL injection](#) vulnerability (CVE-2021-26084). [Per the Cybersecurity and Infrastructure Security Agency \(CISA\)](#), the Confluence and Log4j vulnerabilities were among the top routinely exploited vulnerabilities in 2021.

LOG4J

On December 9, 2021, a severe remote code execution (RCE) vulnerability, “Log4Shell,” was disclosed in the Log4j logging library maintained by the Apache Foundation and used by countless Java applications. The widespread library adoption across many applications made this vulnerability harder to identify and more dangerous than most. A threat actor could create a malicious payload that tricks a server into loading executable code from an threat actor-controlled location, resulting in remote code execution (RCE) with the permission levels of the user running the service.

While researchers identified hundreds of projects as vulnerable to this attack, we focused our analysis on a select number of widely used and deployed products, specifically devices and products in which Censys was able to derive a version number: Unifi devices, Metabase instances, Rundeck instances, and Neo4j.

Censys observed 105,497 total services running what would become a Log4j vulnerability target. Of those, 102,060 services were vulnerable to this attack. By March 2022, only 36% of services were left vulnerable.

The response to the Log4j vulnerability was swift. The sharp decrease in vulnerable devices in Figure 15 is a testament to the criticality of the vulnerability and the widespread coverage it received.

THE ATTACK SURFACE OF THE INTERNET

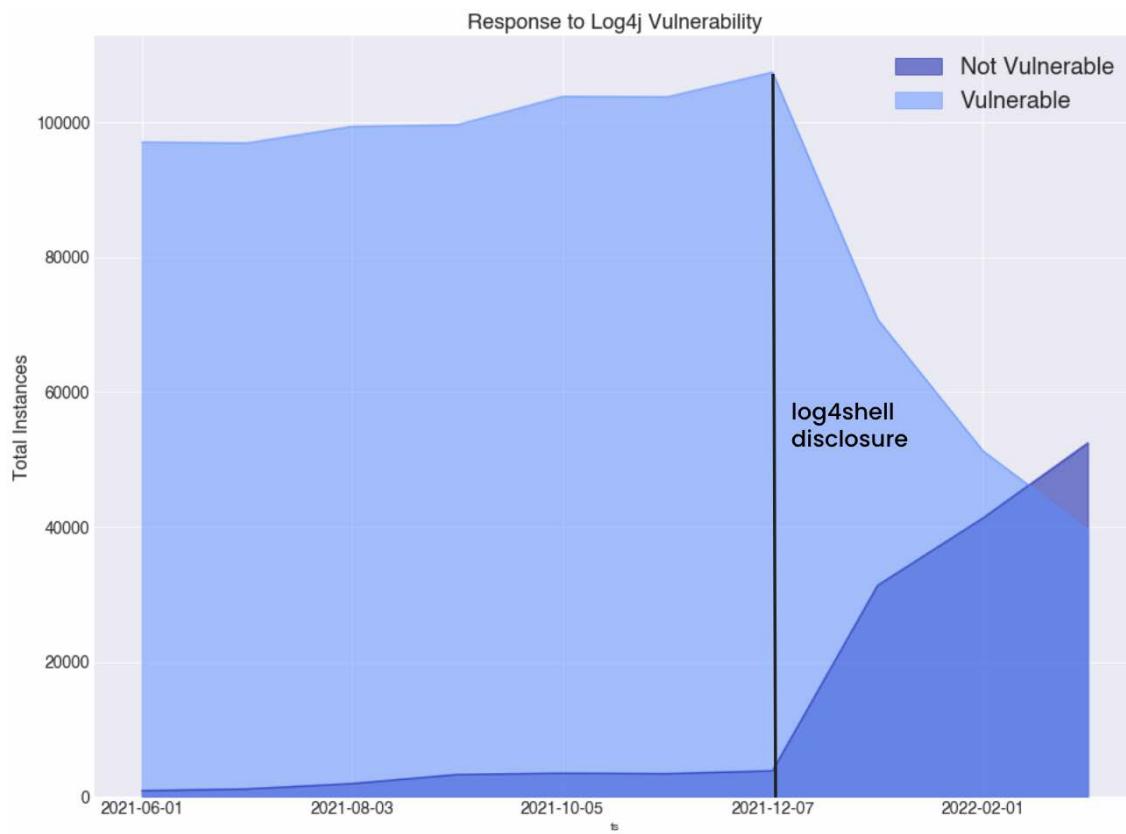


Figure 15: Vulnerable Log4j instances over time, June 2021–March 2022.

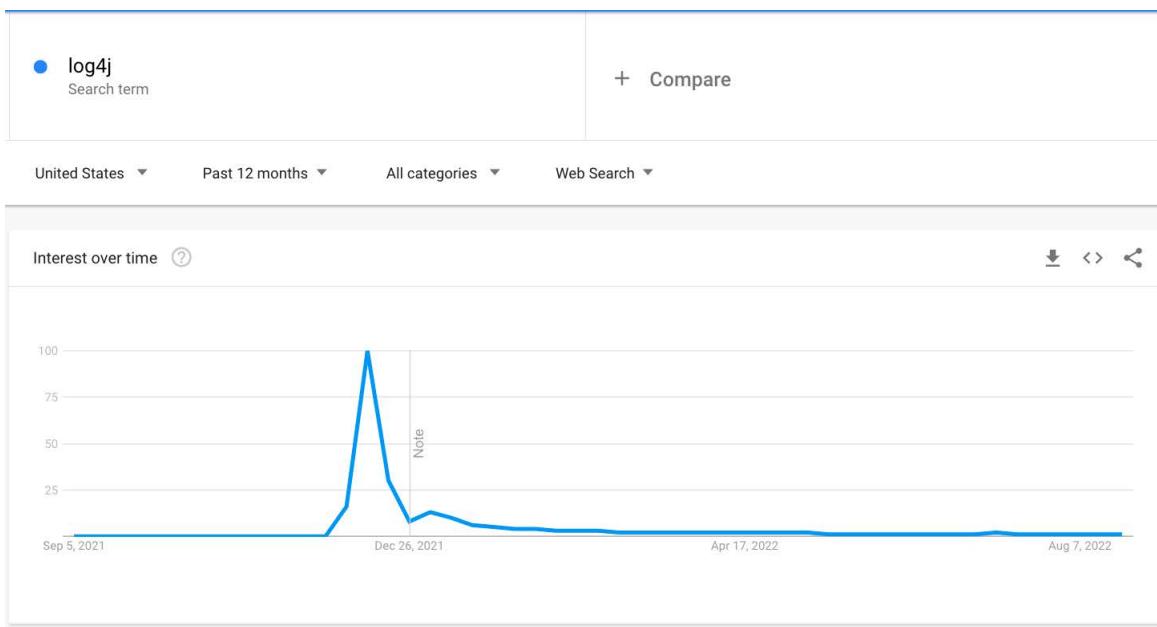
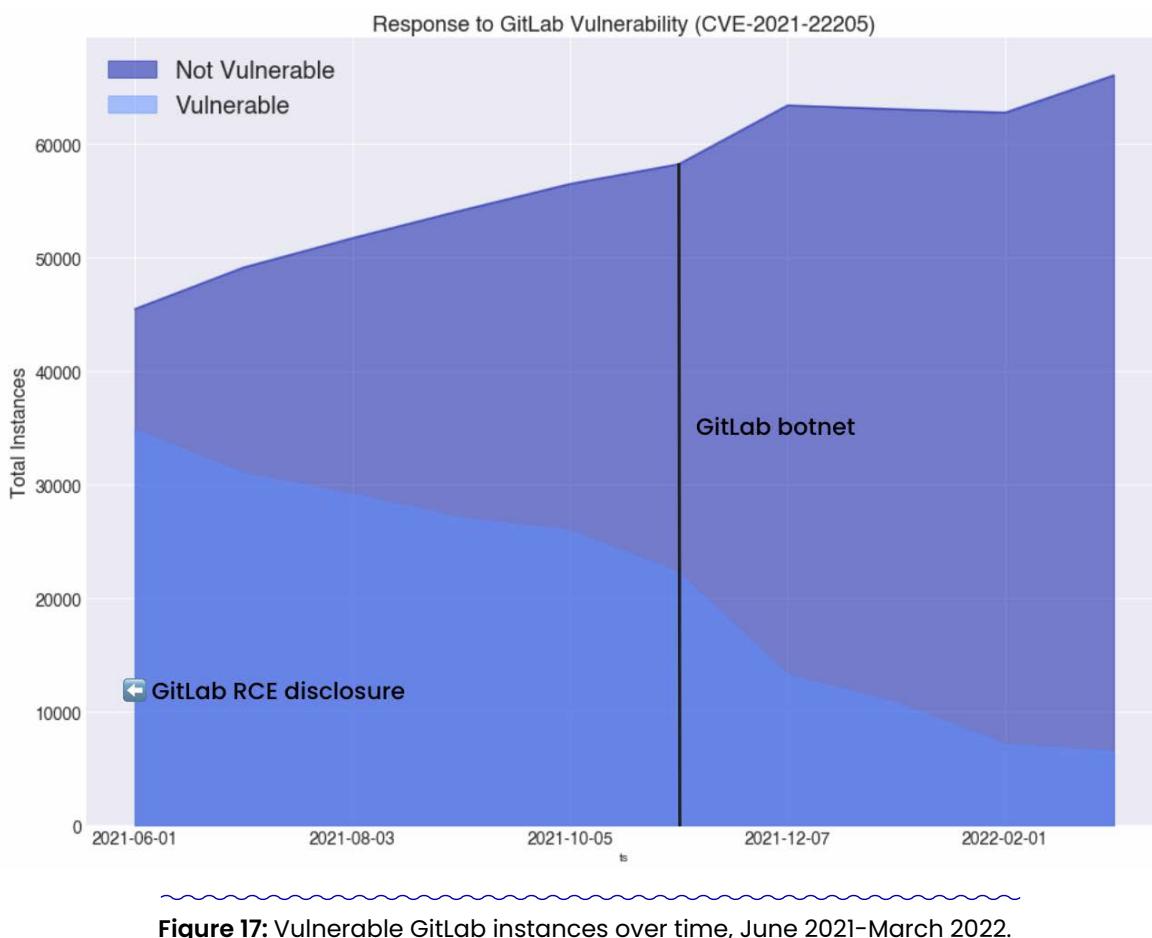


Figure 16: Google Trends search for “log4j” from early August 2021 to late July 2022. Interest spikes in late November 2021.

GITLAB

In early May 2021, a critical vulnerability in GitLab Server (CVE-2021-22205), which could result in a threat actor executing arbitrary code on the target host, made its way into the public eye. At the time, Censys observed over 78,000 services running this software, 48% of which were vulnerable to this particular attack. While vulnerable versions declined for the next six months, Censys observed a 56% drop in vulnerable versions between May and September 2021.



In contrast to the Log4j vulnerability response (Figure 15), the reaction to the GitLab vulnerability (Figure 17) proceeded more gradually until [researchers discovered a botnet](#) composed of thousands of compromised GitLab servers participating in DDoS campaigns capable of generating over one terabit of network traffic per second.

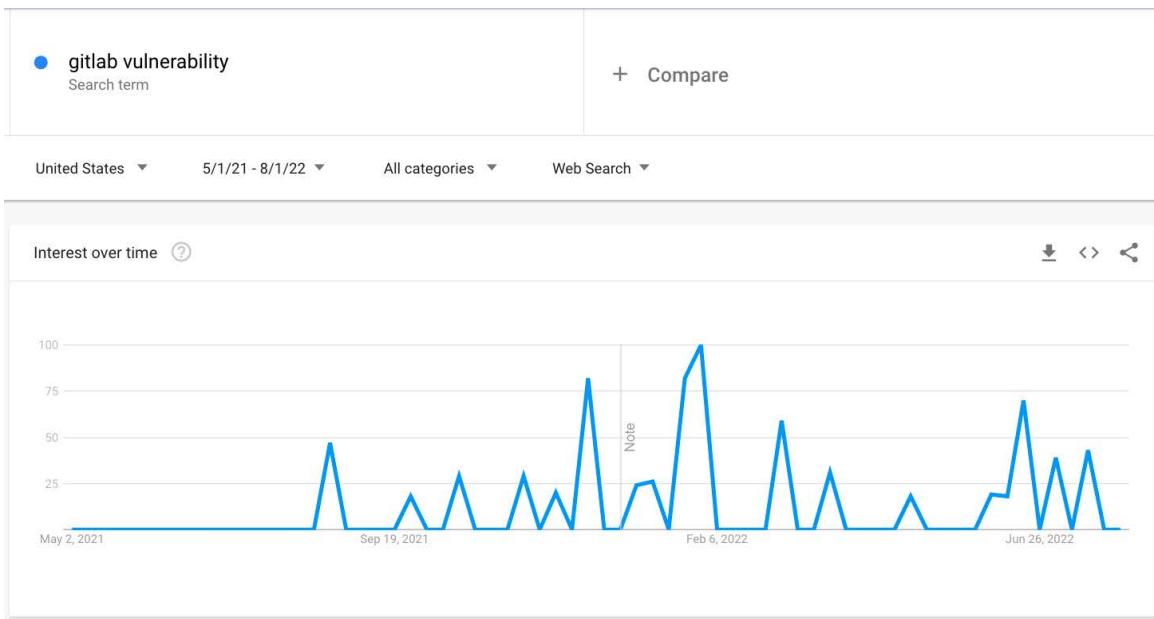


Figure 18: Google Trends search for “gitlab vulnerability” from May 1, 2021 through July 27, 2022. Interest picks up October 31–November 6, 2021, and peak “interest” in the search term takes place from December 12–18, 2021.

It was only then that this vulnerability gained widespread attention both on social media and in the news, and for the next four months, Censys observed over a 70% decline in the number of vulnerable GitLab services online.

CONFLUENCE

In August 2021, Censys observed over 14,000 Confluence services online, only 45 of which were not vulnerable to an OGNL injection that resulted in remote code execution. Once more information was made available about this specific vulnerability, Censys observed a massive effort to fix these systems.

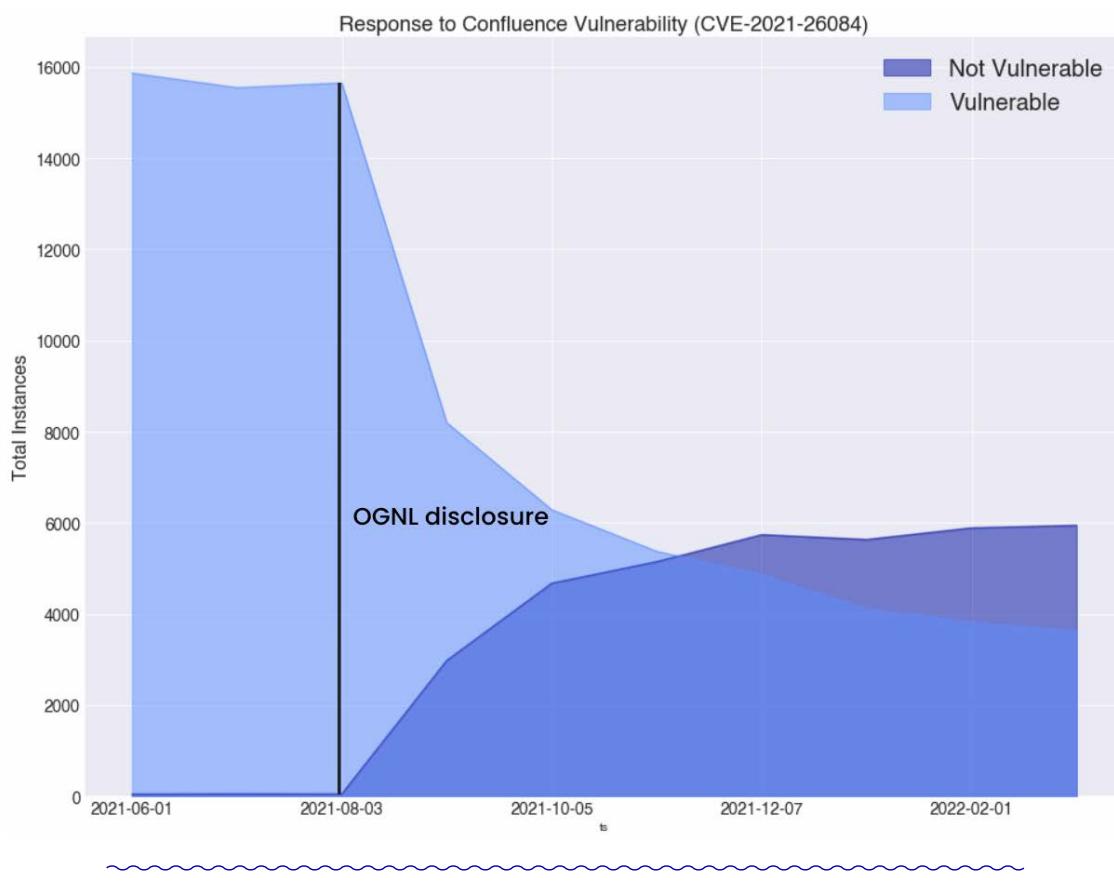


Figure 19: Vulnerable Confluence instances over time, June 2021–March 2022.

In a departure from either of the patch curve patterns observed in the GitLab or Log4j vulnerabilities, Censys observed that **many of these systems disappeared entirely from the public Internet**. 3,363 vulnerable Confluence servers existed six months later, while only 5,253 had signs of being patched. With 8,766 total Confluence servers online, it seems that **many of these instances may have been unused or forgotten resources which were only removed once this significant vulnerability had been made public**. It's also possible that users shifted to Confluence Cloud and, while they're still using Confluence, it's no longer visible on the public Internet in the same way that these instances were.

In comparing these three major vulnerabilities, we see three distinct patterns of response:

- **Quick upgrading upon disclosure:** Log4j response was exemplary in speed (but not so much in the chaos and stress it caused for Security teams everywhere).
- **Vulnerability required wide-scale exploitation before remediation:** Gitlab took a botnet to get traction (lots of things fly under the radar until they...don't).
- **Quick upgrading, and removing instances from the public-facing Internet:** Response to Confluence was quick but much of the remediation was taking things off the public Internet, rather than upgrading as seen in Log4j or Gitlab.

While most vulnerabilities are nowhere near as severe as the Log4j vulnerability (to every responder's relief), reducing the time between vulnerability disclosure to upgrade for even medium and lower risk vulnerabilities could improve organizations' overall security posture.

ONGOING: DEADBOLT RANSOMWARE

In [January 2022](#), a group calling themselves Deadbolt targeted a series of QNAP NAS devices made for consumers and small businesses that run the QNAP QTS (Linux-based) operating system, infecting the devices with ransomware.

Instead of encrypting the entire device like many ransomware variants, this ransomware targets specific backup directories for encryption and vandalizes the web administration interface with an informational message explaining how to remove the infection. This enabled our research team to identify and track infected devices.

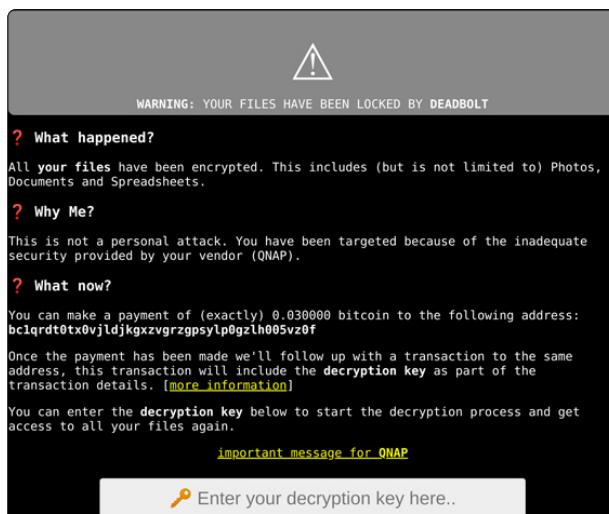


Figure 20:
Deadbolt ransom message.

In the initial January campaign, we observed 4,988 QNAP devices locked by Deadbolt. Along with the self-explanatory HTML title, “[ALL YOUR FILES HAVE BEEN LOCKED BY DEADBOLT](#),” the HTTP response body includes a unique Bitcoin address where the victim is urged to send 0.03BTC (equivalent at the time to USD 1,100) to unlock their newly hacked device. If the actor had received a 100% return from this attack, that would net them a prize of \$4,484,700 US dollars.

By May, we observed [two additional](#) Deadbolt attacks. With each campaign, along with general information about what hosts were infected with Deadbolt, we could also obtain and track every unique Bitcoin wallet address used as a ransom drop.

We teamed up with [Concinnity Risks](#) to determine the exact amount of money involved in this attack by tracking the Bitcoin wallet transactions associated with an infection; as of April, we concluded the following.

Number of ransoms paid	132
Total BTC	3.96
Total USD	\$187,665

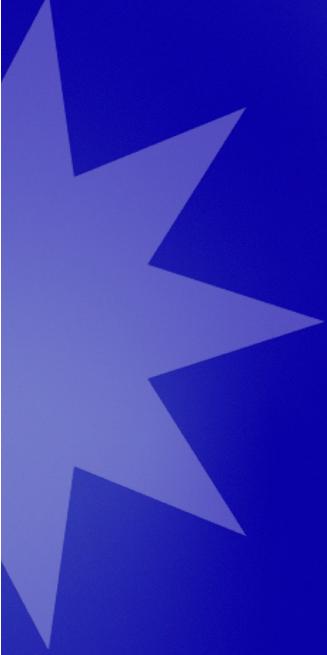
Note that this does not include the most recent set of infections but gives us good insight into the inner workings of a ransomware campaign.

We continue to monitor the Internet for signs of QNAP Deadbolt infections, and you can too, using our [Deadbolt ransomware dashboard](#).

Takeaways

- Zero days, CVEs, and interesting technical exploits are often trending topics on infosec Twitter, but they don't represent the majority of risks we observe across the Internet. Admittedly, part of this may be due to the way we obtain this data—i.e., requiring a fingerprint based on a public Internet-facing artifact. However, identification of misconfigurations and exposures can be among the first observations a threat actor makes when performing initial reconnaissance on an organization. **Good security hygiene that addresses misconfigurations and exposures may not be as exciting as a zero day, but it's a critical piece of defense in depth for any security program.**
- When the GitLab vulnerability was disclosed, it gathered widespread attention only after being leveraged in a botnet 6 months after the initial disclosure. New vulnerabilities are disclosed (almost) daily and it can be challenging to keep track of all of them, let alone those specifically that are relevant to your organization. [CVETrends](#), [CISA alerts](#), and [r/netsec](#) are useful resources for learning about newly disclosed vulnerabilities and exploits. **Knowing is half the battle, but it's also important to ensure your organization has a vulnerability management process and engages in regular patching across all assets.**

Attack Surfaces of Organizations



AWS is the dominant hosting provider for the medium to large organizations we studied, representing where we see 80% of hosts. AWS is also where we observe 75% of risks and vulnerabilities for these organizations—not surprising, given the large AWS footprint for these organizations.

These organizations have an average of 44 domain registrars and 17 hosting providers (including cloud, datacenters, and on premises servers). This sprawl can make inventory and defense of assets particularly difficult for Security teams.

Misconfigurations represent 70% of the risks observed across these organizations' attack surfaces, while Exposures represent 16%. This is a shift from what we see across the Internet, where Misconfigurations represent 60% and Exposures make up 28%.

Given the widespread migration of daily business functions to digital platforms, organizations now have to have some kind of online presence. But what does that mean? Companies may perceive their online presence to consist solely of their public websites and web servers. But in practice, that digital footprint often looks different than expected.

We used Censys' attribution algorithm to generate attack surfaces for 37 medium to large companies of varying industries, shown on the next page.

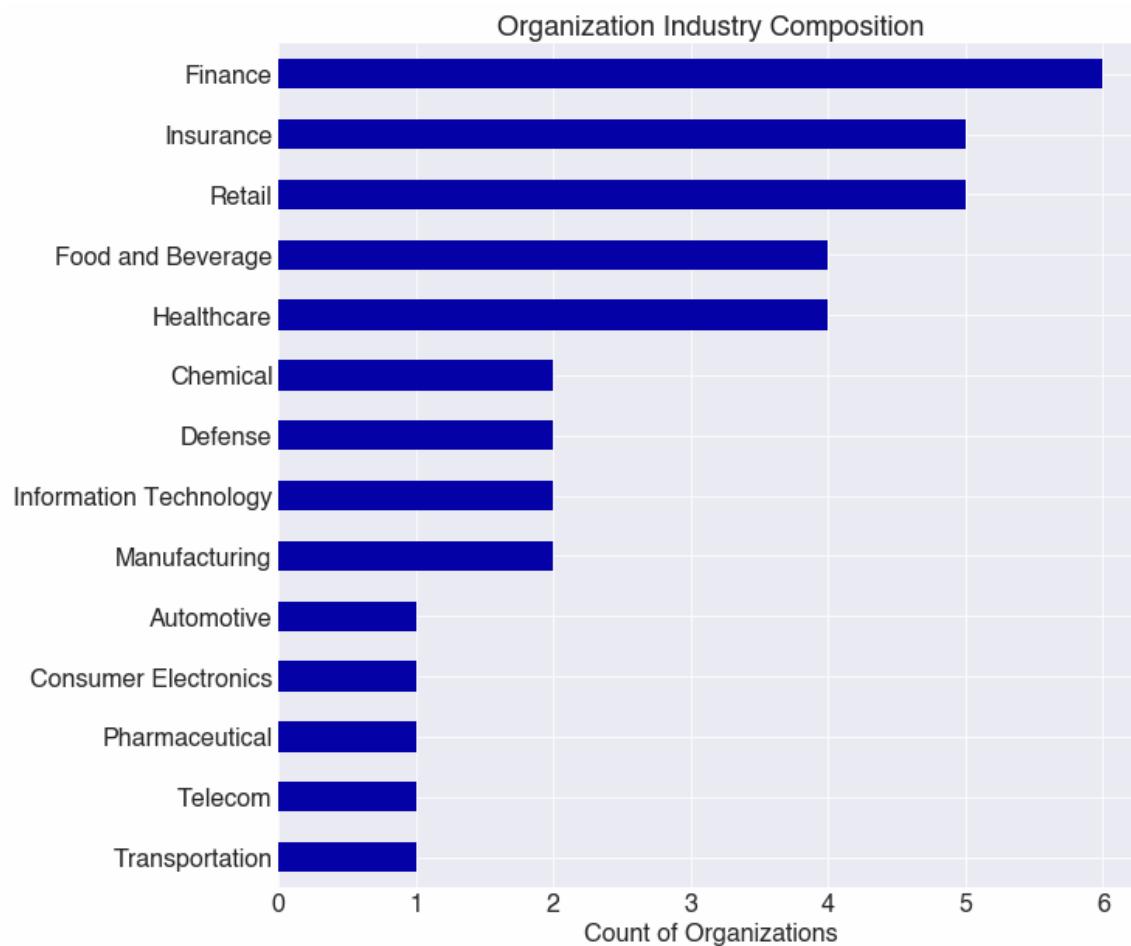


Figure 21: Industry composition of organizations sample.

Overview

Collectively, 80% of the hosts for our sample organizations are in AWS, though it is important to note that this is an aggregate figure. This doesn't mean that each of these organizations use AWS as their primary cloud (or even at all), but rather when aggregating all of the hosts in our sample and examining where they're hosted, AWS comes out on top.

75% of observed risks and vulnerabilities across these organizations are in AWS, which isn't particularly surprising, given the heavy AWS presence of the sample.

The 37 organizations have hosts in an average of 17 different hosting provider locations (median 14)—including cloud, datacenter, and on-premises equipment. These organizations also have, on average, 44 domain registrars (median 30).

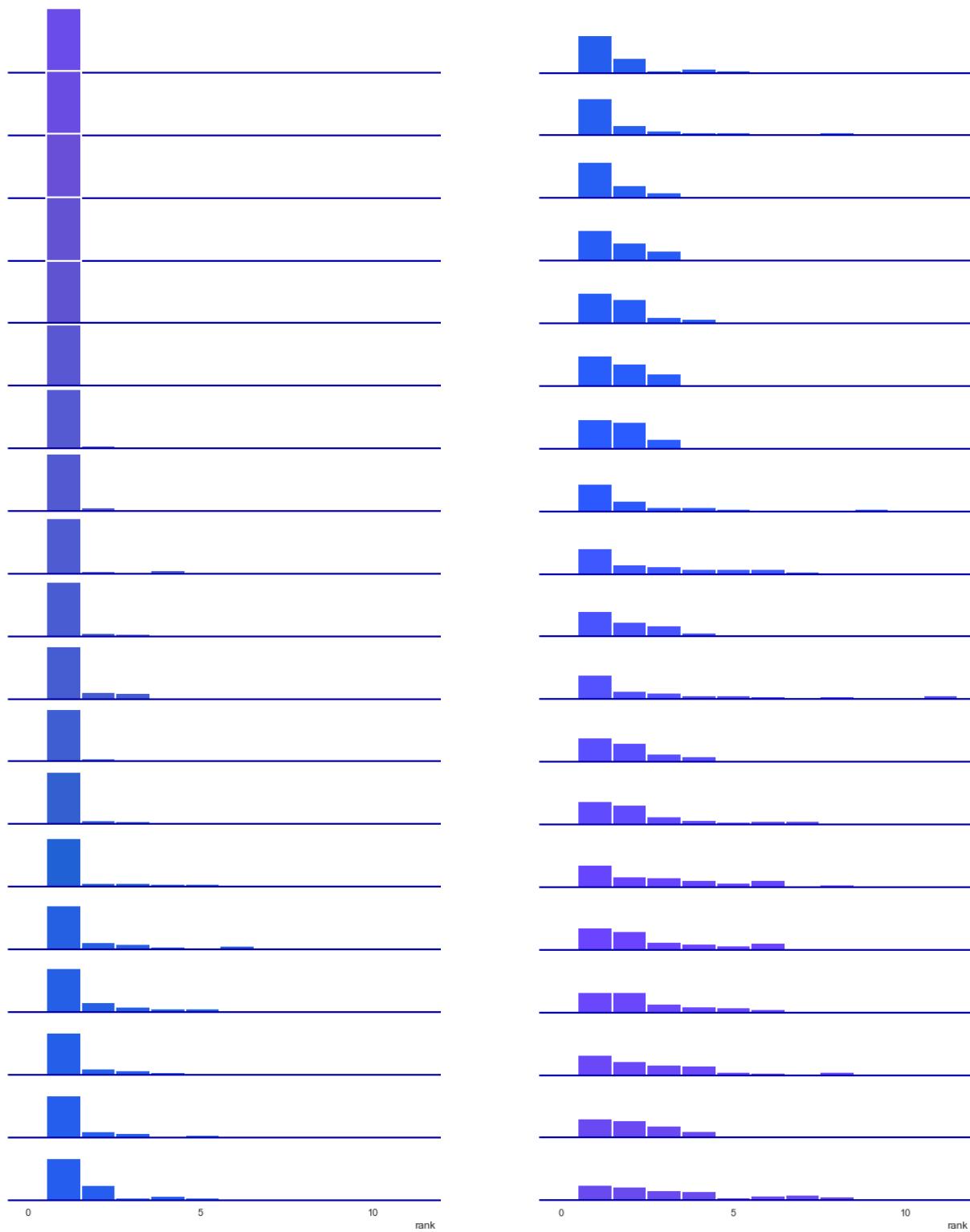


Figure 22: Hosting provider sprawl distributions. Each plot represents an organization in our sample.

While we see an average of 17 different hosting provider locations among the 37 organizations studied, there is variety in the spread of assets across these providers. In the graph above, each plot represents the distribution of hosting providers of one of the 37 organizations studied. Starting with the plot on the top left, organizations with a presence primarily in one hosting provider are shown, and as we move to the right column of plots, we can see organizations with greater amounts of sprawl across hosting locations.

Many of these organizations also rely on a [CDN](#), and Akamai is by far the most popular CDN choice among our sample organizations. Given Akamai's enterprise-focused offerings, this is not surprising, as our sample consists of medium to large organizations.

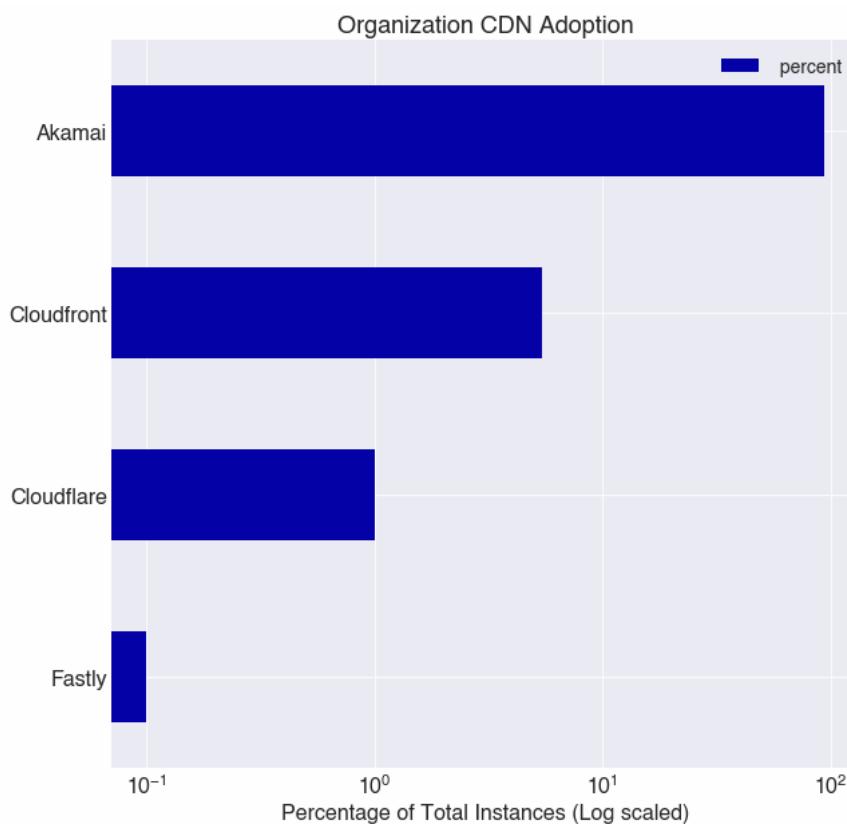


Figure 23:

CDN presence among the 37 organizations.

Risk Profile

Observed Vulnerabilities percentages are nearly the same across the Internet and our company sample, representing 13% in our sample and 12% Internet-wide.

However, Misconfigurations represent 70% of the risks observed across these organizations' attack surfaces, while Exposures represent 16%. This differs from overall Internet risk trends, where 60% are Misconfigurations, and 28% are Exposures.

One possible explanation for the lower Exposure percentage for our organization sample could be firewall rules enabled by the organizations studied.

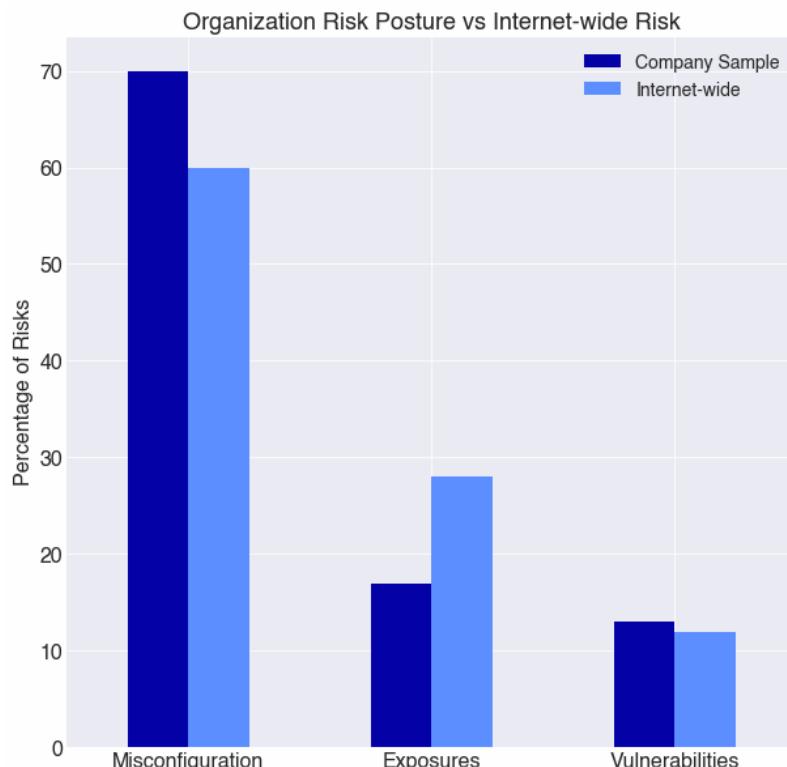


Figure 24:

Comparison of the 37 organizations' risk footprint vs risks on the Internet as a whole.

The top risks among the 37 organizations are a departure from the top risks seen across the Internet at large. We categorize the `insecure-ssl-tls-key-length` risk as a misconfiguration, which likely drives some of the breakdown that we see in Figure 24. However, it's worth noting that, while `unencrypted-weak-auth-page` was one of the top (and higher severity) risks in our Internet-wide sample, it nearly falls out of the top 20 risks in our companies sample.

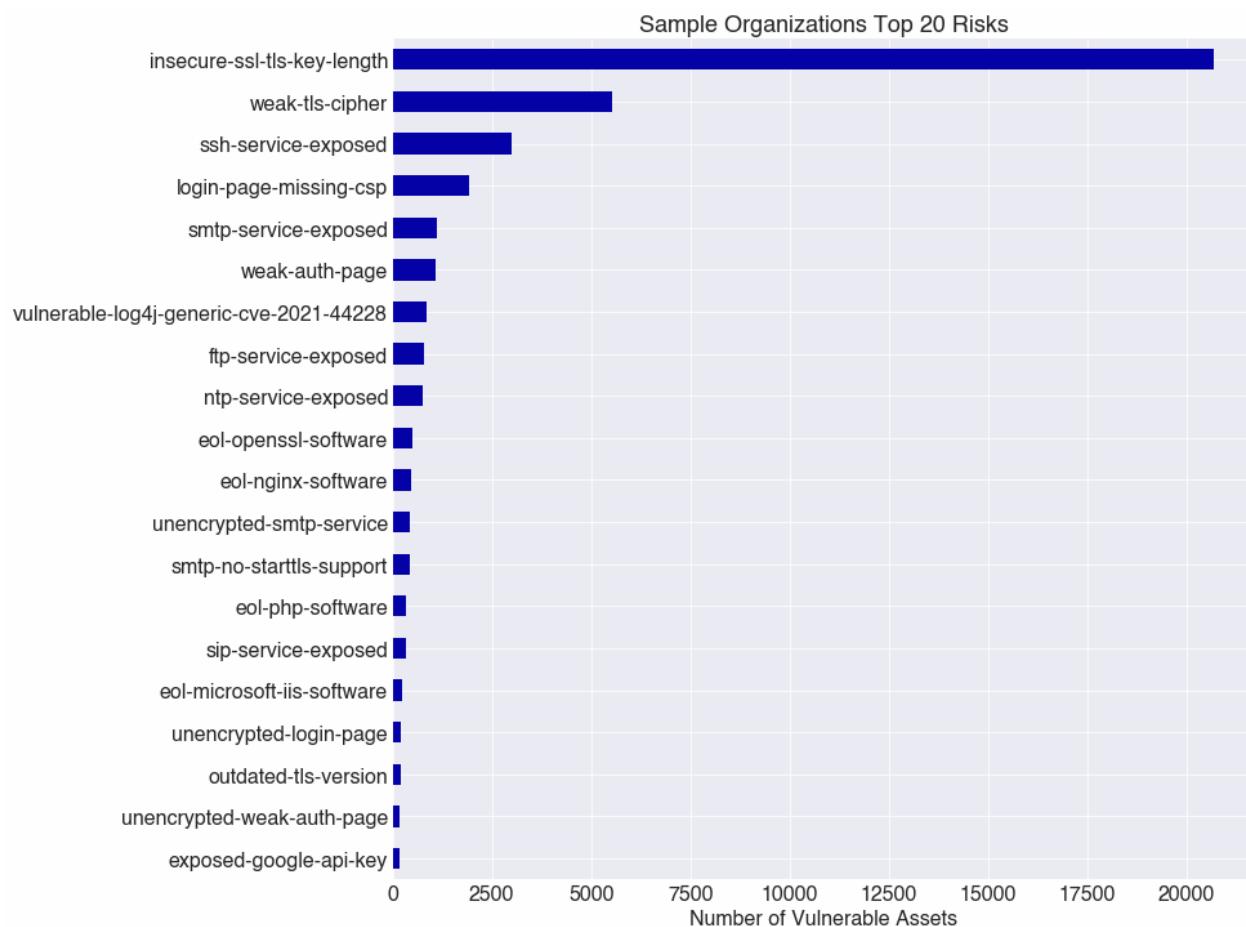


Figure 25: Top 20 risks observed among the attack surfaces of our sample organizations.

This may suggest that these medium and large organizations have the resources to implement better security hygiene practices than the Internet at large, and also perhaps that there are firewalls protecting the perimeter of these organizations.

While we hesitate to make wider generalizations from a sample of 37 organizations, there are things we can learn from observing real companies' public-facing Internet footprints. Namely, we recognize that the actual public-facing Internet presence may look very different than their Security teams anticipate.

Takeaways

- Our data suggests organizations often adopt a multi-provider strategy, but do you really know all the clouds and data centers where your assets live? If you're unsure, you're not alone—many companies use multiple hosting providers and face similar challenges. This can pose difficulties for IT and Security teams, as **it's impossible to secure assets you don't even know you own.**
- While it may be expected for a company to use several domain registrars, an average of 44 different registrars suggests that there are likely domains purchased by organizations that are beyond the view of IT and Security teams. **If these domains expire and are purchased by threat actors, they can be weaponized for phishing, brand impersonation, and other attacks.**

Conclusion

The Internet is constantly evolving as new technologies emerge, new vulnerabilities are discovered, and organizations grow their online presence. Moreover, the recent increase in remote work and the need to enable employees has exacerbated existing challenges with shadow IT and asset visibility.

While CVEs and advanced exploits make the news, undetected misconfigurations and service exposures represent the majority of risks we observe. Moreover, response to CVEs and exploits is highly variable, depending on factors such as severity, observation of exploitation in the wild, and how widely they are discussed in the news and social media.

Things change rapidly on the Internet, but security begins with visibility. It is our hope that sharing our visibility into these trends will inform security practitioners and executives alike, and will help us all work together toward a safer Internet.