JupiterOne

THE

State of Cyber Assets Report

2023

JupiterOne Research

Contents

Contents

Scaling security to a fragmented attack surface

Goal

Understand cyber assets, liabilities, attack surfaces, and their relationships to each other in the modern enterprise.

Summary

Not only is the attack surface growing, but the scale of the problem is now untenable.

“By redefining the cybersecurity control plane, we can better adapt to our environments' growing complexity.”

Understand cyber assets, liabilities, attack surfaces, and their relationships to each other in the modern enterprise.”

The evolving state of the modern cyber attack surface is the reason we created The State of Cyber Assets Report (SCAR). It's one of many annual reports to understand cyber assets, liabilities, attack surfaces, and their relationships to each other in the modern enterprise.

In the era of increasingly destructive and disruptive cyber threats, centralized cyber assets are a business liability. Threats to the confidentiality, integrity, and availability of organizations have made it necessary to adapt by decentralizing our cyber assets across a growing number of cloud service providers, environments, and services.

Not only is the attack surface growing, but the scale of the problem is now untenable. That's why we've set out to conduct and write this research.

Cybersecurity practitioners are grappling with an unprecedented amount of complexity in 2023. Continuous integration and deployment (CI/CD) pipelines result in a steady stream of changes that can each introduce new possibilities for misconfigurations, policy exceptions, or human error. Security teams need context to scale security policy and enforcement to the distributed evolving attack surface.

Striving for reduced complexity is not possible for cybersecurity teams. Instead, we must learn to accept our increasingly complex environments by rethinking our definition of the cybersecurity control plane. Achieving consistent situational awareness across new asset types and environments requires a shift toward unified cyber insights. Its flexibility is especially suitable for increasingly modular approaches consistent with hybrid multicloud architectures. CSMA enables a more composable, flexible and resilient security ecosystem. Rather than every security tool running in a silo, a cybersecurity mesh enables tools to interoperate through several supportive layers, such as consolidated policy management, security intelligence and identity fabric.

Jasmine Henry
Lead Researcher
Senior Director of Data Security and Privacy

Research questions

The following questions formed the basis of the research conducted for the 2023 State of Cyber Assets Report. All of the findings in this report derive from these core questions.

01 What is the composition of cyber asset inventories?

- What are the average number of **APPLICATIONS**, **DATA**, **DEVICES**, **NETWORKS**, and **USERS**?
- How many assets and accounts are in an AWS, GCP, or Azure environment?
- Are security practitioners inventorying all types of cyber assets?
- What is the value of a cyber asset?

02 How do security practitioners interact with security data?

- What is the composition and volume of security data?
- What are the most leveraged types of security technologies?
- How many security data sources are being correlated and aggregated?
- Do security teams have comprehensive, data-driven visibility across the attack surface?

Research questions

The following questions formed the basis of the research conducted for the 2023 State of Cyber Assets Report. All of the findings in this report derive from these core questions.

03 Which assets tend to have more liabilities (or vulnerabilities)?

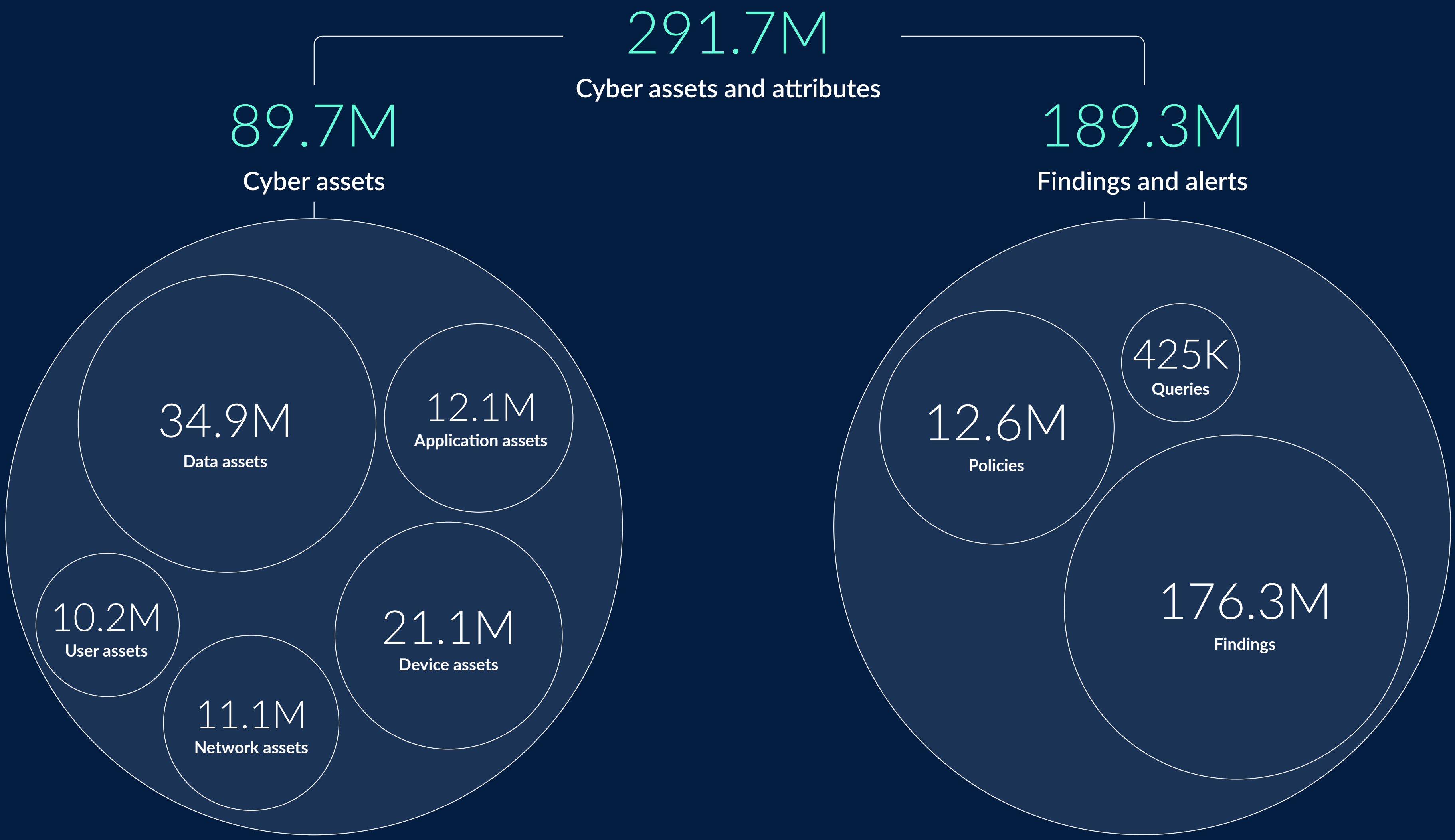
- What tools are security practitioners using to identify and detect vulnerabilities?
- What is the ratio of vulnerabilities to assets, and which types of assets have the most vulnerabilities?
- Which assets are the most critically vulnerable?

04 How do security practitioners navigate their attack surface and data sources?

- What assets do security practitioners query?
- Which assets are most related?

Executive summary and key findings

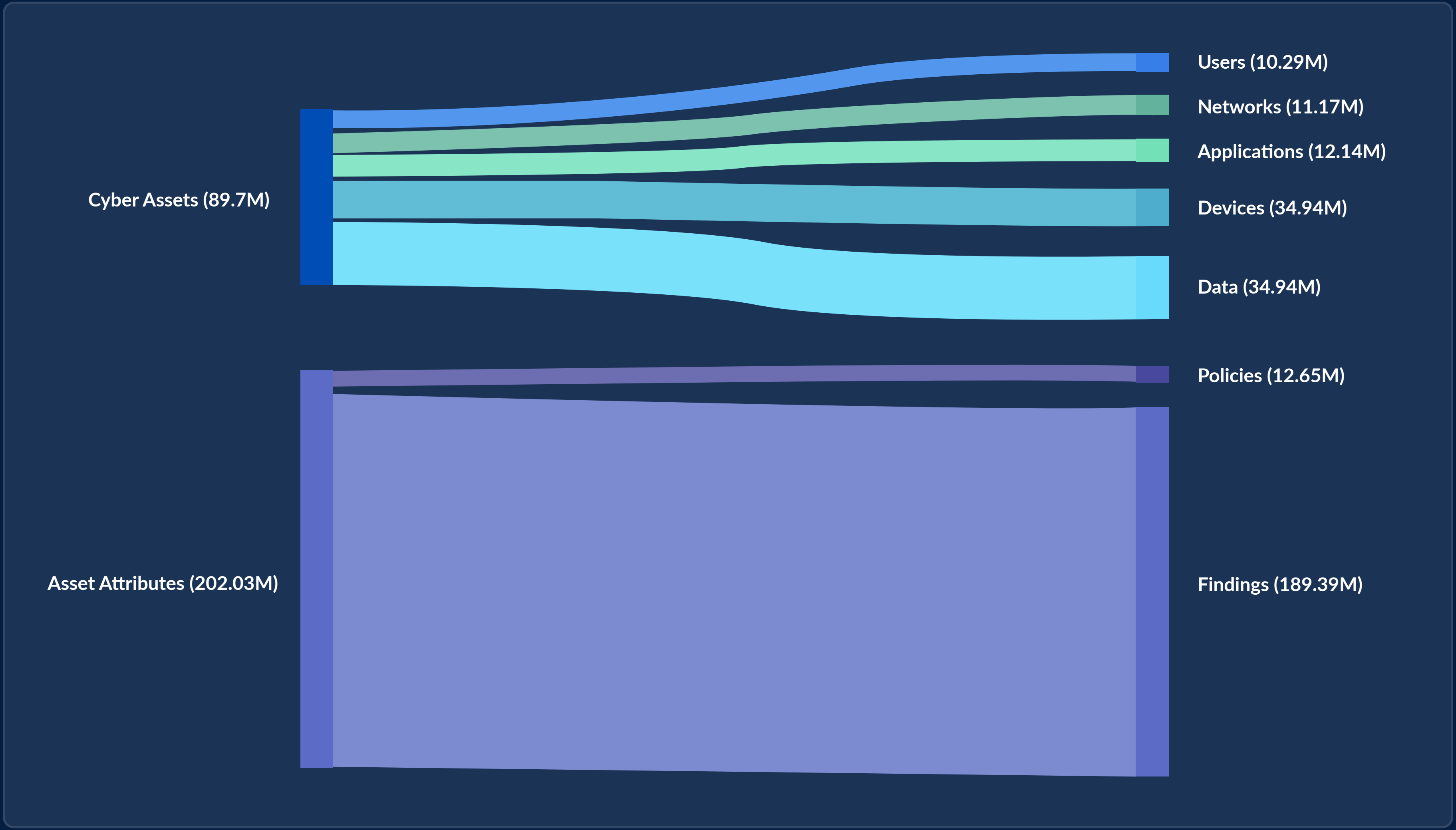
Our research included an analysis of:



Introduction

Overview of SCAR data

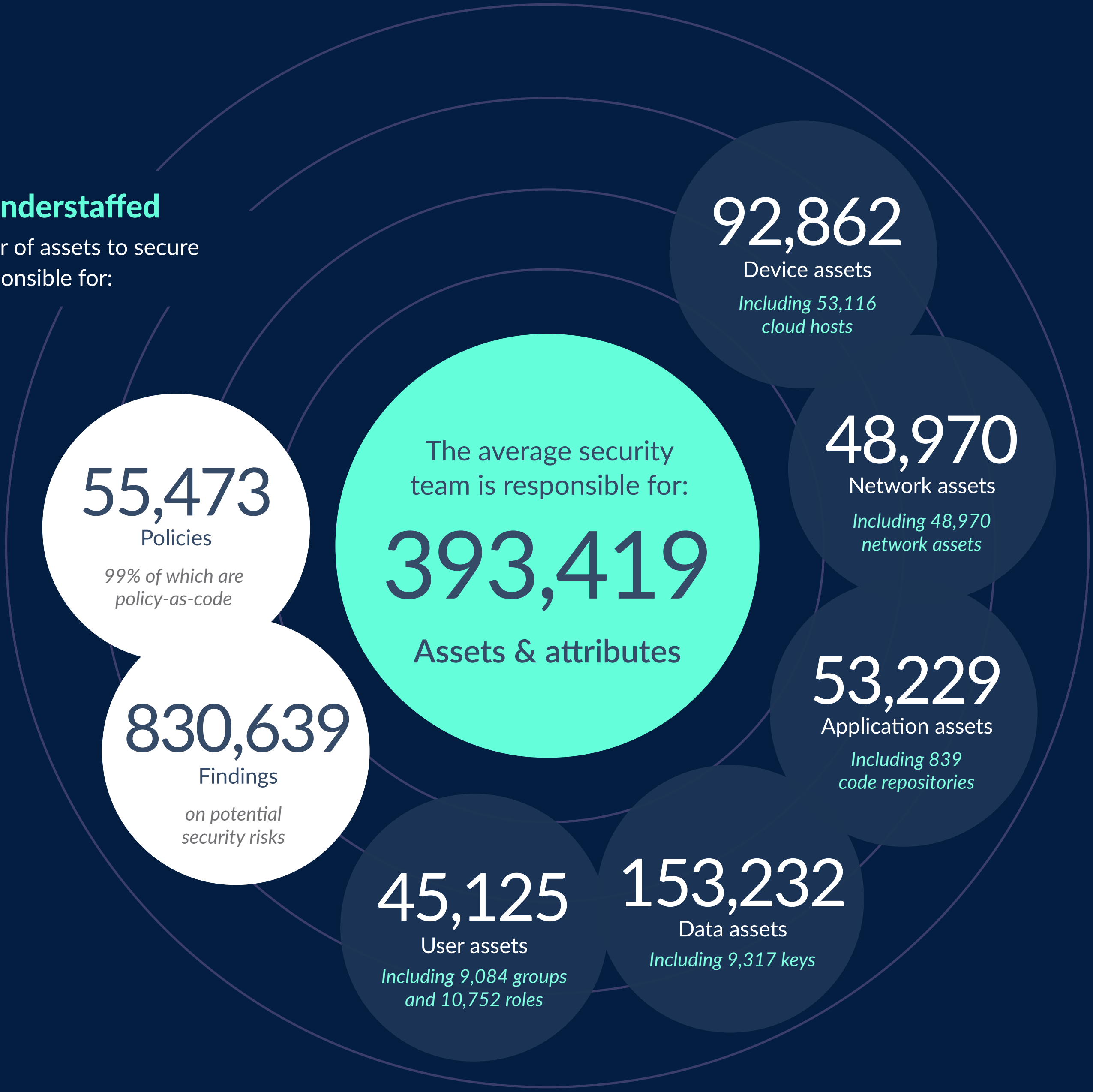
Executive summary chart 1: ratio of cyber assets to attributes



On average

Security teams are fatigued and understaffed

Security teams have an unprecedented number of assets to secure and manage. The average security team is responsible for:



The average security team is responsible for:

393,419

Assets & attributes

92,862
Device assets
Including 53,116 cloud hosts

48,970
Network assets
Including 48,970 network assets

53,229
Application assets
Including 839 code repositories

153,232
Data assets
Including 9,317 keys

45,125
User assets
Including 9,084 groups and 10,752 roles

55,473
Policies
99% of which are policy-as-code

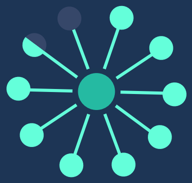
830,639
Findings
on potential security risks

Year-over-year, the average security organization has experienced a **132.86%** increase in cyber assets and a **588.98%** increase in security findings.

The mean value of a cyber asset is **$17,711**, a staggering number considering the volume of both assets and findings (or liabilities) that security practitioners must oversee.

The state of unified cyber insights

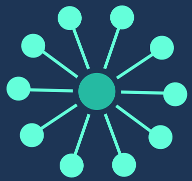
Security teams at today's organizations work with a broad range of data coming from across the environment. Understanding the number and variety of data sources is foundational to understanding the complexity these teams must navigate.



8.67

The average (mean) security team is correlating 8.67 security data sources.

Organizations of all sizes have the greatest adoption of unified data sources for **APPLICATIONS** (100%), **DEVICES** (92.22%), and **USERS** (82.22%). The asset classes with the fewest data sources are **NETWORKS** (62.72%) and **DATA** (49.29%).

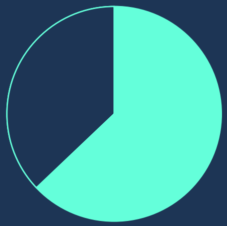


10.11

Mid-sized organizations (50-499 employees) have the most data sources with an average of 10.11 sources.

2/3

Nearly two-thirds of security data points originate from technologies used to secure applications.



61%

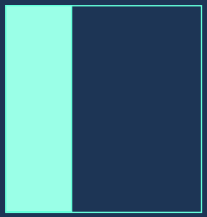
Of all data utilized by practitioners comes from non-security systems, including CSPs, HRIS, task management, and DevOps tools.

The security findings dilemma

Growing asset volumes are concerning, but it is the opportunities these assets present to potential attackers that pose the real dilemma for security teams. Security findings tell an even deeper story than just asset information alone.

2.11

The average organization has 2.11 findings (or vulnerabilities) for every asset.



1/3

DEVICES, especially cloud hosts, are linked to over 1/3 of security findings (36.84% of findings), but represent 96.1% of critical findings.

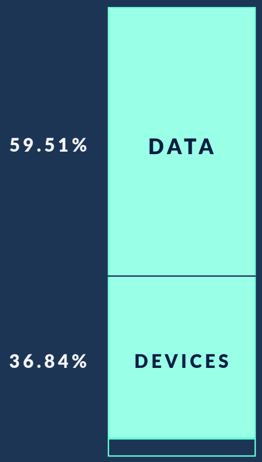
The SCAR includes analysis of a total of

189.39M

⚠ Security findings

89.7M

Cyber assets



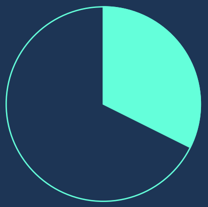
96.35%

The two most vulnerable superclasses, **DATA** and **DEVICES**, collectively represent 96.35% of all unresolved security findings.

- **DATA** is the most vulnerable superclass of assets with 59.51% of all security findings.
- **DEVICES** is a distant second-place with 36.84% of unresolved findings.
- Drilling down further, just two types of cloud assets – hosts and images – are responsible for an incredible 93.18% of all security findings.

The cloud-native attack surface expands

Cloud adoption continues to grow for businesses of all sizes, making cloud providers and services a tempting target for attackers. In many cases, organizations are using more than one cloud service provider and must manage many accounts across these providers.



31%

Of orgs are using 3 CSPs simultaneously, and in those cases, they are using AWS, GCP, and Azure.

60%

Of the 89.7 million assets included in this analysis originate from a cloud service provider (CSP) environment, which doesn't include the countless cloud-based apps, services, and network assets in today's enterprise.

225

The average (mean) resources that security teams at large organizations are tasked with securing are nearly 225 total AWS accounts, GCP projects, and Azure subscriptions.

- Security teams at small organizations have 171.05 total environments to secure.
- Mid-sized organizations are responsible for securing 559.24 unique accounts, projects, and subscriptions across major CSPs.
- Organizations with fewer than 500 employees have more unique AWS accounts, GCP projects, and Azure subscriptions than employees, on average.

Section

1

The State of Cyber Assets:

An overview of asset classes, attributes, & trends

Assets per employee by organization size

Chart 1.1: Composition of cyber assets by superclass

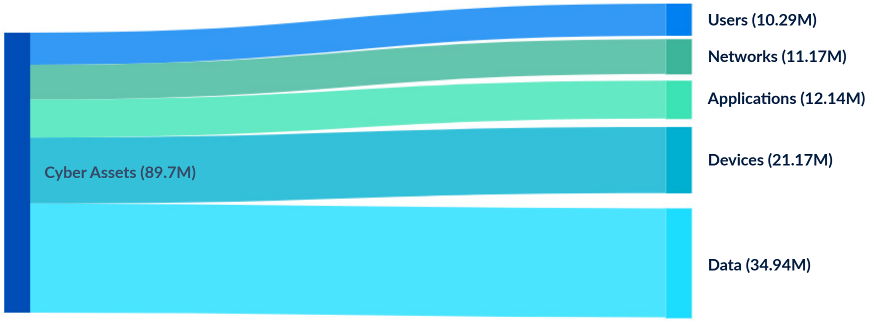


Chart 1.2: Average assets per employee by organization size

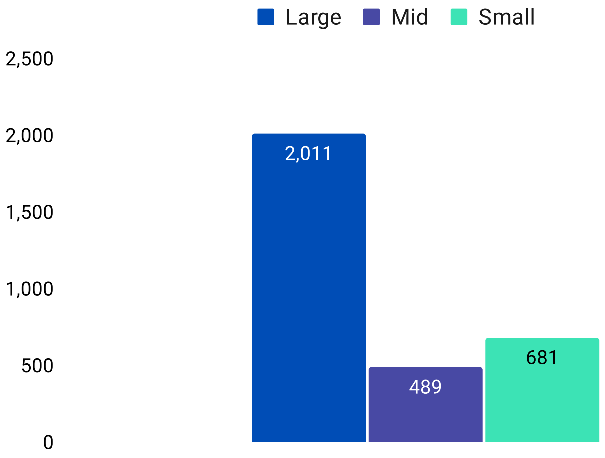


Table 1.1: Composition of cyber assets by superclass

Large	Mid	Small
2,011	489	681

Table 1.2: Average assets per employee by organization size

Superclass	Number	Relative percent
APPLICATIONS	12,136,257	13.53%
DATA	34,937,076	38.95%
DEVICES	21,172,433	23.60%
NETWORKS	11,165,333	12.45%
USERS	10,288,505	11.47%

Average assets per employee by industry

Chart 1.3: Average assets per employee by industry

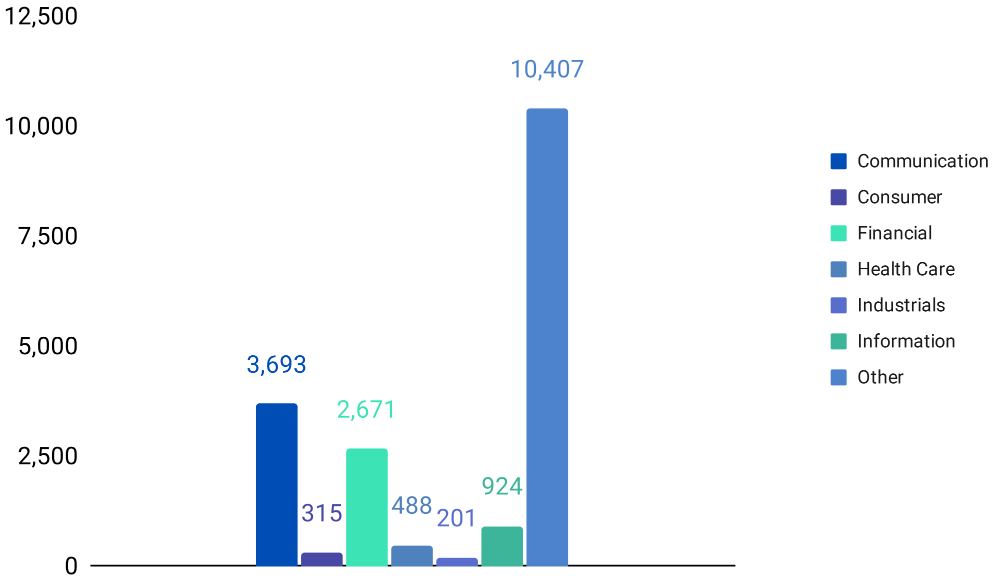


Table 1.3: Average assets per employee by industry

Communication[1]	Consumer[2]	Financial	Health care	Industrials[3]	Information	Other[4]
3,693	315	2,671	488	201	924	10,407

[1] Telecommunication Services, Media & Entertainment
[2] Consumer Discretionary, Consumer Staples
[3] Capital goods, commercial & professional services, transportation
[4] Energy, materials, utilities, and real estate.

What's the value of a cyber asset?

Cyber assets are fundamental to the mission and operations of modern organizations. Hard, or physical, assets no longer make up the majority of an organization's balance sheet. The value of a digital asset is determined both within and outside of the business context by its ability to create present and future value.

The mean value of a cyber asset – calculated as the total number of assets divided by market capitalization – is $17,711, a staggering number considering the volume of both assets and findings (or liabilities) that security practitioners must oversee.

Total number of assets ÷ Market capitalization

= Asset Value of $17,711

Section

2

An overview of superclasses

Asset superclasses

Assets and attributes are grouped into superclasses for the purpose of visualization and analysis. These asset superclasses are largely drawn from Sounil Yu's Cyber Defense Matrix, with the addition of two superclasses to describe asset attributes: `FINDINGS` and `POLICIES`.

The definitions and inclusions for each category are included below, or described further in JupiterOne's data classification model. This report will utilize `ALL CAPS` for these classification patterns to easily identify the superclasses by name.

Devices

The `DEVICES` superclass consists of workstations, servers, phones, tablets, containers, hosts, peripherals, storage devices, network devices*, web cameras, infrastructure, and more. It also includes operating systems, firmware, and any other software native to a device.

Chart 2.1: Composition of the `DEVICES` *superclass*

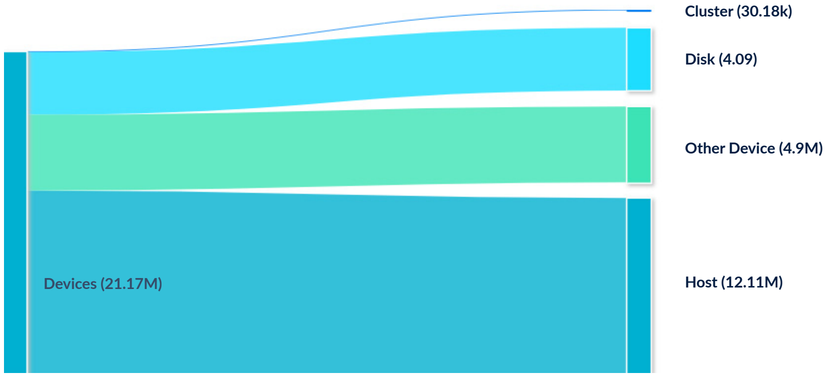


Table 2.1: Composition of the `DEVICES` *superclass by number and relative percent*

Class	Number	Relative percent
Cluster	30,179	0.14%
Other Device	4,944,674	23.35%
Disk	4,087,006	19.30%
Host	12,110,574	57.20%

[2] Hardware networking devices (like switches and routers) are included here because those devices are separate from the network communication pathways they create.

Asset superclasses

Networks

NETWORKS are communications channels, connections, and protocols that enable traffic to flow among DEVICES and APPLICATIONS, including both physical and virtual networking systems such as cloud firewalls. This superclass also includes Domain Name Systems (DNS), Border Gateway Protocol (BGP), Virtual Private Clouds (VPCs), Virtual Private Networks (VPNs), Content Delivery Networks (CDNs), and certificates.

Chart 2.2: Composition of the NETWORKS *superclass*

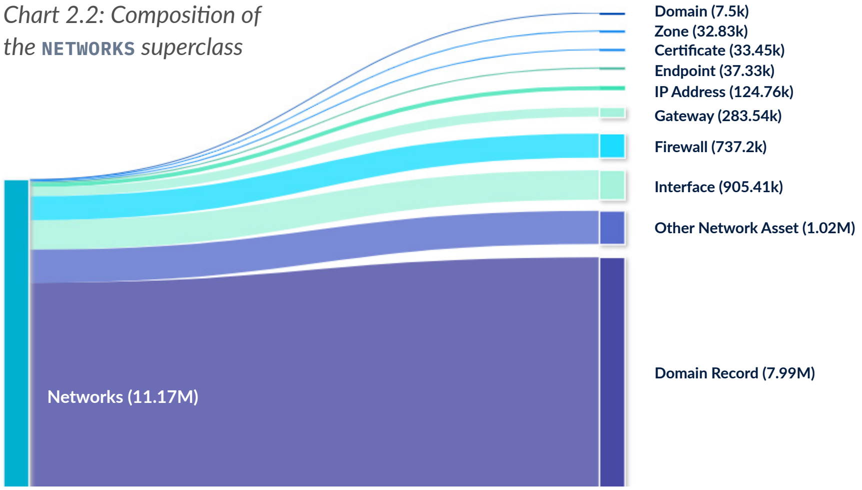


Table 2.2: Composition of the NETWORKS *superclass by number and relative percent*

Class	Number	Relative percent
Certificate	33,451	0.30%
Domain	7,495	0.07%
Endpoint	37,333	0.33%
Firewall	737,200	6.60%
Gateway	283,543	2.54%
Interface	905,415	8.11%
IP Address	124,761	1.12%
Other Network	1,017,207	9.11%
Domain Record	7,986,103	71.53%
Zone	32,825	0.29%

Asset superclasses

Applications

APPLICATIONS are software code and applications on the devices, separate from the operating system/firmware. This class includes serverless functions, APIs, and microservices.

Chart 2.3: Composition of the APPLICATIONS *superclass*

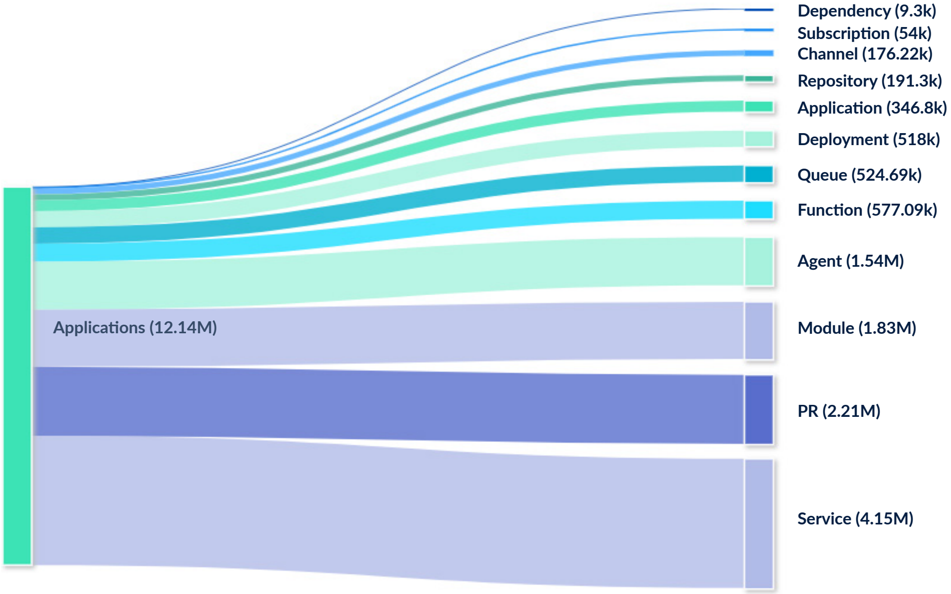


Table 2.3: Composition of the APPLICATIONS *superclass by number and relative percent*

Class	Number	Relative percent
Agent	1,542,582	12.71%
Application	346,798	2.86%
Channel	176,220	1.45%
Dependency	9,325	0.08%
Deployment	518,267	4.27%
Function	577,086	4.76%
Module	1,834,693	15.12%
PR	2,214,178	18.24%
Queue	524,693	4.32%
Repository	191,340	1.58%
Service	4,147,037	34.17%
Subscription	54,038	0.45%

Asset superclasses

Data

DATA includes data-at-rest, data-in-motion, and data-in-use. This superclass includes databases, S3 buckets, storage blobs, and files. Also, the **DATA** superclass includes logs, records of changes, tasks, and notification channels. Secrets are also grouped with data, including encryption keys, key pairs, and vaults.

*Chart 2.4: Composition of the **DATA** superclass*

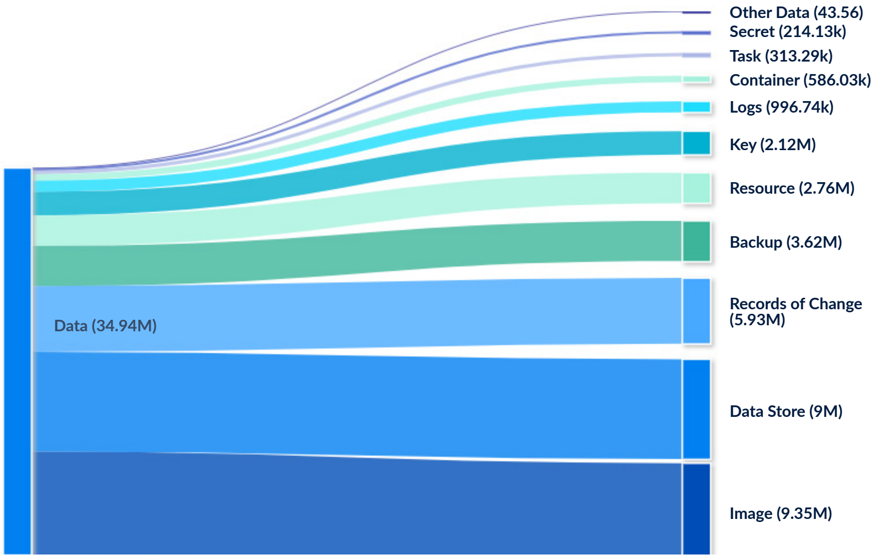


*Table 2.4: Composition of the **DATA** superclass by number and relative percent*

Class	Number	Relative percent
Backup	3,623,434	10.37%
Container	586,032	1.68%
Data Store	8,995,971	25.75%
Image	9,345,869	26.75%
Key	2,124,354	6.08%
Logs	996,736	2.85%
Other Data	43,560	0.12%
Records of Change	5,934,821	16.99%
Resource	2,758,872	7.90%
Secret	214,133	0.61%
Task	313,294	0.90%

Asset superclasses

Users

USERS are people and who use the resources in other asset superclasses, and the identities associated with these users. The **USERS** superclass also includes groupings of users, including teams, organizations, and sites.

*Chart 2.5: Composition of the **USERS** superclass*

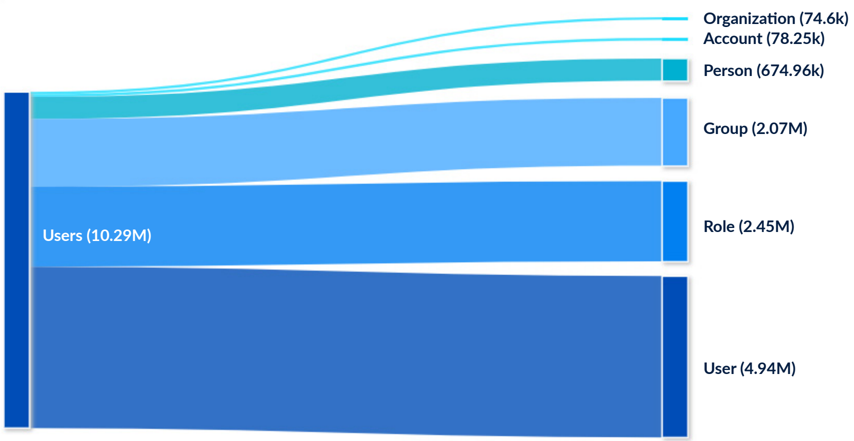


*Table 2.5: Composition of the **USERS** superclass by number and relative percent*

Class	Number	Relative percent
Account	78,253	0.76%
Group	2,071,141	20.13%
Organization	74,678	0.73%
Person	674,955	6.56%
Role	2,451,556	23.83%
Other User*	4,937,922	47.99%

*The 'other user' category is defined as a set of credentials for an application.

Attribute superclasses

Findings

The FINDINGS category consists of alerts and results, incidents data, monitoring trails, threat intel, and vulnerabilities from both human and non-human sources. FINDINGS are part of the complex graph of relationships and dependencies in security organizations, but they're not an asset. Many FINDINGS are liabilities, particularly if they are related to a critical asset.

Chart 2.6: Composition of the **FINDINGS** *superclass*

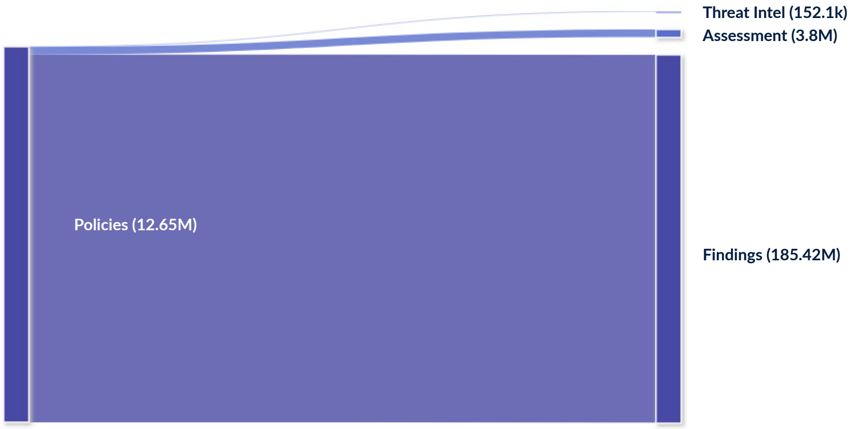


Table 2.6: Composition of the **FINDINGS** *superclass by number and relative percent*

Class	Number	Relative percent
Findings	185,418,521	97.91%
Assessment	3,815,165	2.01%
Threat intel	152,119	0.08%

Attribute superclasses

Policy

Much like `FINDINGS`, `POLICY` falls outside the classification of traditional assets and should be considered an attribute of cyber assets in the sense they act as guardrails to protect assets.

IAM policies, control policies, configurations, requirements, and rulesets all fall within our classification model for `POLICY`. Human-generated policy and procedure documents are here, too, though they're a negligible percentage compared to other forms of `POLICY`.

Chart 2.7: Composition of the `POLICY` *superclass*

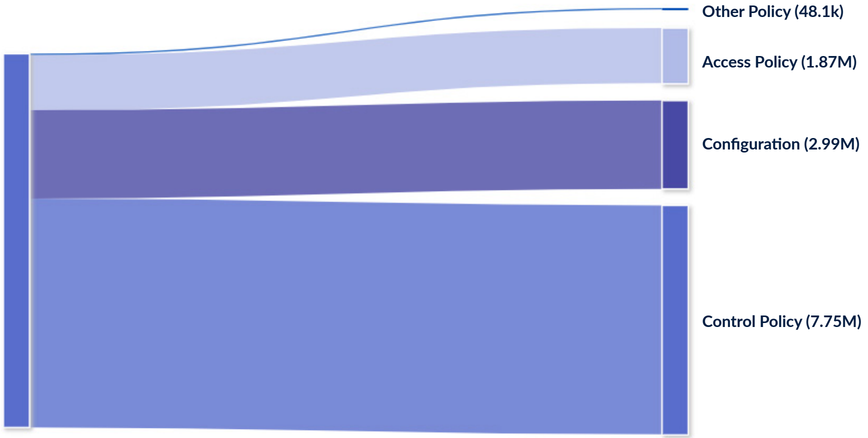


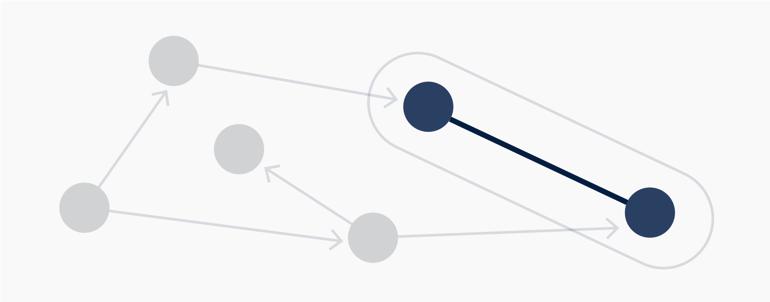
Table 2.7: Composition of the `POLICY` *superclass by number and relative percent*

Class	Number	Relative percent
Access policy	1,867,017	14.76%
Configuration	2,985,453	23.60%
Control policy	7,747,303	61.25%
Other	48,166	0.38%

An overview of superclasses

Additional terminology

A TRIPLET IS A UNIQUE OCCURRENCE OF 2 NODES & 1 EDGE



Additional information on the underlying graph data model can be found in this GitHub repo or docs.

WHERE CRITICAL ASSETS FIT IN THIS ANALYSIS

The definition of a critical asset varies between attackers, defenders, and even security practitioners. JupiterOne's graph data model prioritizes the following production asset classes with a 'critical' tag:

Superclass	Asset class
APPLICATIONS	Applications Code repos Functions
DATA	Data stores
DEVICES	Hosts

The Graph Data Model

The underlying data model for this analysis is based on graph theory; specifically, a reference model used to describe cyber assets and their complex interactions in a modern organization. The data model is defined by a set of entities and their relationships, or nodes and edges:

- An entity is a node or vertex in the graph that represents a cyber asset.
- A relationship is the edge between two entity nodes in the graph.
- A triplet is a set of two nodes and one edge, or a single cyber asset relationship.

Cyber asset relationship

A relationship is the connection between two or more cyber assets. Assets in isolation don't tell us the entire picture – it's how they connect and work together to provide value or risk to a business and processes.

Context or 'cyber asset context'

Context is metadata, data, or information that provides added perspective or attributes of any cyber asset(s). Context provides a broader understanding into a cyber asset, its relationships, and how they relate to one another in the broader system or environment.

Critical assets

Critical assets, as defined by CISA, are cyber assets that are essential to maintaining operations and achieving the organization's mission. If the confidentiality, integrity, or availability of a critical asset is breached, there are generally significant business consequences.

While critical assets often have common characteristics, the designation of a cyber asset can vary significantly between organizations. Criticality is largely dependent on the business environment, processes, risk appetite, and policy.

Section

3

Security data sources

On average

Security teams are, arguably, only as good as their data sources. *More* data, however, isn't necessarily *better*. Security data sources that cover a wider range of asset classes and security functions may create visibility, but it's not a perfect indicator of maturity.

A security team may have a lot of data, but lack comprehensive visibility or authority to enforce a consistent security posture across cloud providers, environments, or asset types.

The security data sources analyzed in the SCAR are probably not a comprehensive list of all the data that security teams monitor, investigate, or reconcile. These data sources are effectively limited by the ~180 integrations that JupiterOne offers, although a significant number of organizations in this analysis have built custom integrations to either homegrown, on-premises, or vendor-supported systems.

On average

The average (mean) security team is correlating 8.67 security data sources. Mid-sized organizations (50-499 employees) have the most data sources with an average of 10.11 sources.

Further research is needed to understand why mid-sized organizations have more integrations, since the answer is probably more complex than just adoption. Mid-sized organizations may have fewer barriers to integrating data sources compared to their peers at larger organizations, who must contend with many silos, policies, and system owners before gaining data access.

The number of data sources that security practitioners actively correlate is likely a subset of all available data sources. While little publicly available information exists on the number of systems used by security practitioners in 2023, it is likely that the correlated data sources represent half – or fewer – of all the tools used by security teams.

The total number of technologies and potential data sources could be much higher when including all homegrown solutions, open source software, and solutions that are shared between security and other teams. One recent Panaseer study of the largest enterprises with 5,000 or more employees revealed teams are managing 76 discrete security tools on average.

8.67

The average (mean) security team is correlating 8.67 security data sources.

76

One recent Panaseer study of the largest enterprises with 5,000 or more employees revealed teams are managing 76 discrete security tools on average.

On average

Chart 3.1: Average security data sources by industry and organization size

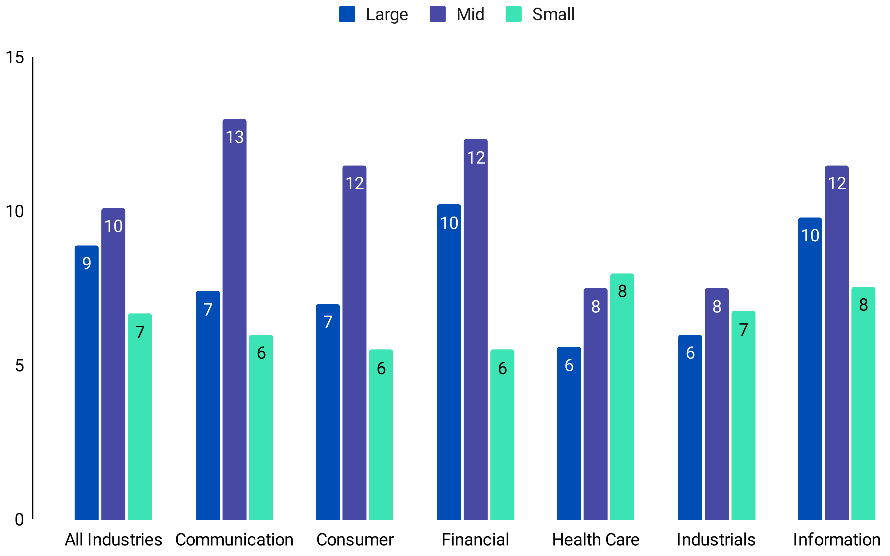


Table 3.1: Average security data sources by industry and organization size

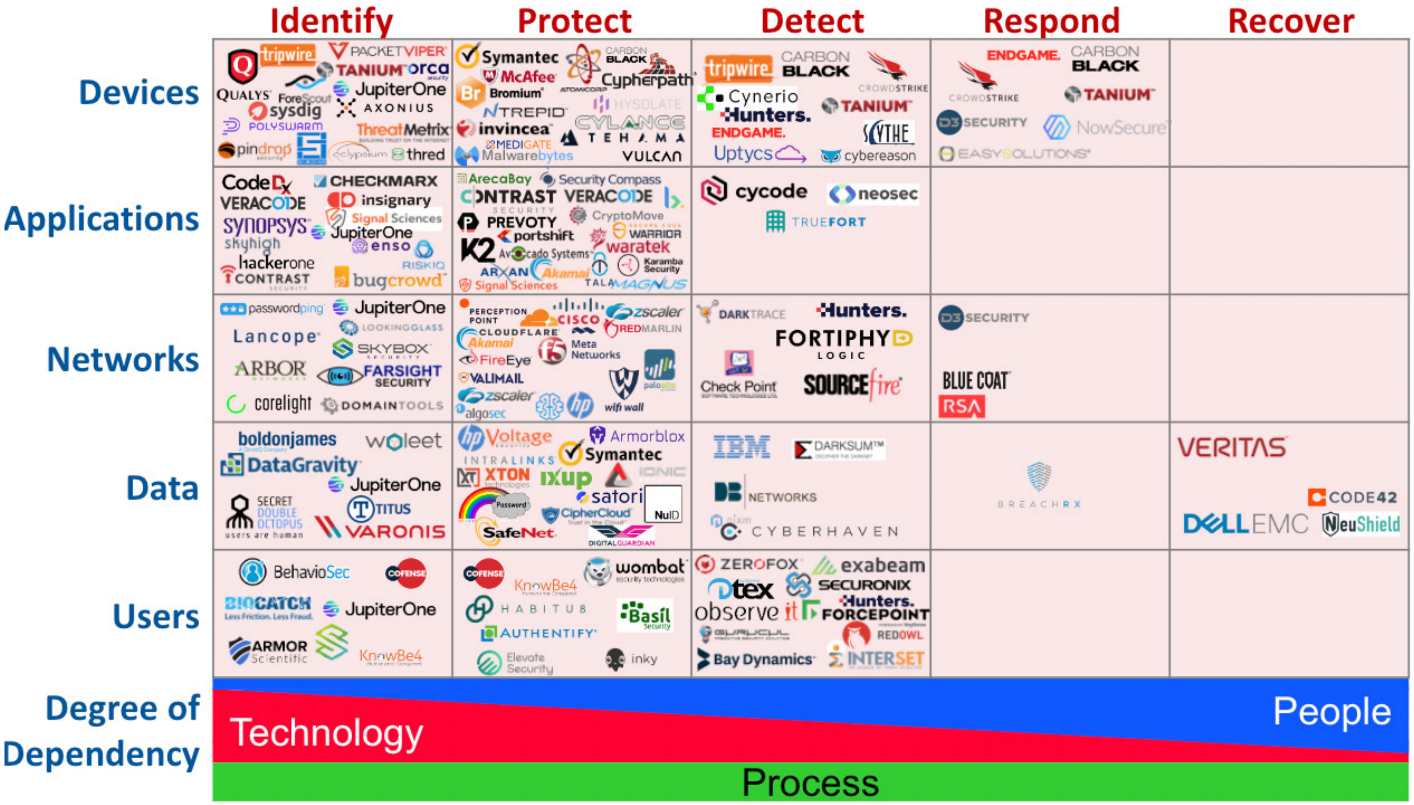
	Large	Mid	Small
All industries	8.90	10.11	6.67
Communication	7.40	13.00	6.00
Consumer	7.00	11.50	5.50
Financial	10.21	12.33	5.50
Health care	5.60	7.50	8.00
Industrials	6.00	7.50	6.75
Information	9.81	11.50	7.56

Most common types of security data

To understand how security data sources compare, it is necessary to use a framework that is mutually exclusive and collectively exhaustive, such as Sounil Yu's Cyber Defense Matrix. It combines the five distinct superclasses of cyber assets in this report with the five distinct cybersecurity functions of the NIST cybersecurity framework.

Cyber Defense Matrix is a broadly applicable map for understanding the complex cybersecurity technology landscape. The data sources included in this analysis should provide enough visibility into the 'security data stack' to offer useful context to readers on how practitioners are approaching unified cyber insights.

Security technologies by asset classes & operational functions



Security data sources

More data doesn’t imply maturity

It is unreasonable to assume that a large number of data sources is a primary factor contributing to an organization’s security maturity. An organization may have dozens of systems to manage laptops and user identities, but little oversight into cloud security. A variety of data sources across different asset superclasses and operational functions is much more valuable than sheer amount of data.

Instead, organizations that are able to correlate and find cross-system relationships from a wide variety of data sources that cover all asset types and security functions are best able to navigate security requirements both left and right of boom. The most forward-looking cybersecurity teams are building multi-year strategies for unified cyber insight.

Observant readers will notice that this heatmap only includes one dimension of Cyber Defense Matrix, the asset classes, while excluding the cybersecurity operational function. That decision was primarily made for the sake of simplicity – analyzing data source adoption by class is an easier visualization that better displays areas where security teams may be struggling to integrate data sources.

Mapping data source adoption by security function is also complicated for several reasons. Many cybersecurity vendors cover multiple functions. An Extended Detection and Response (XDR) solution, for example, could be used to **Identify**, **Protect**, and **Detect** the security of DEVICES. Cybersecurity graphs are also a tool designed to help defenders **Identify** relationships between their cyber assets. A cybersecurity data source in a graph will generally help practitioners identify assets, even if the source has a different primary function like **Protect**, **Detect**, or **Respond**.

Heatmap 3.1: Percent adoption of data source by asset superclass & size

	Large	Mid	Small	All
DEVICES	100%	100%	76.67%	92.22%
APPLICATIONS	100%	100%	100%	100%
NETWORKS	62.07%	92.11%	40.00%	64.72%
DATA	41.38%	63.16%	43.33%	49.29%
USERS	100%	100%	46.67%	82.22%

Key 🟥 +1.00 🟧 +0.40 🟨 +0.00

More data doesn't imply maturity

DEVICES asset security sources

DEVICES security data sources have the second-highest aggregated adoption rate across organizations of all sizes. Organizations with 50 or more employees in the large and mid-sized categories have a 100% adoption rate of at least one device data source. The most common security data sources in this category are *Endpoint Detection and Response* (EDR/XDR), *Unified Endpoint Management* (UEM), *IT Asset Management* (ITAM), and *Managed Detection and Response* (MDR).

Addigy | AUTOMOX
CROWDSTRIKE | jamf
Qualys. | SNIPE-IT OPEN SOURCE ASSET MANAGEMENT

APPLICATIONS asset security sources

All analyzed organizations are using at least one data source to secure APPLICATIONS, including third-party apps and homegrown code. *SAST, DAST*, and *software asset management* are the most common security data sources in this category. There was also significant adoption of *web application firewall* (WAF), *Runtime Application Self-Protection* (RASP), and *source code compromise* tools.

Checkmarx | GitLab
Cobalt | hackerone
GitHub | snyk

NETWORKS asset security sources

On average, more than half of organizations of any size are utilizing one or more security data sources for insight into network security, though adoption varies significantly by size. Mid-sized firms with 50-499 employees had over 92% adoption of network security data, compared to just 62% at their larger counterparts and 40% at small firms. Network security sources include *Content Delivery networks*, *Load Balancing services*, *Secure Access Service Edge* (SASE) vendors, *Domain Registry* services, and *Network Intrusion Prevention* (NIPS or NIDS).

CISCO | GoDaddy
CLOUDFLARE | RAPID7
fastly | netskope

More data doesn't imply maturity

DATA asset security sources

Security practitioners have integrated the fewest sources of insight into **DATA** assets, with just under 50% adoption across organizations of all sizes. Mid-sized organizations are the most likely to be utilizing a data asset source for insight into their attack surface with over 63% adoption, compared to 41% and 43% adoption by large and small organizations, respectively. The most frequently adopted **DATA** integrations include *Data Classification* technologies, *Secrets Discovery*, and *Data Loss Prevention* (DLP).

While **DATA** is likely an asset class where security teams are struggling to gain unified insight, we hypothesize that security practitioners have data on the security of these assets, it's just not integrated with other sources. Many organizations of all sizes utilize open source technologies, homegrown systems, and policy-based enforcement to secure **DATA**.

DATADOG

databricks

sysdig

DATASTAX

USERS asset security sources

All mid-sized and large organizations, consisting of all firms with 50 or more employees, are utilizing at least one source of data on **USERS**. Commonly adopted data sources on user security include *Identity Providers* (IDP), *Multi-Factor Authentication* (MFA) vendors, *Identity and Access Management* (IAM) technologies, *Credential Management*, and *Security Awareness Training*.

While just 46% of small organizations with 49 or fewer employees are using a dedicated security source for user security data, this does not imply a lack of capability or adoption. Instead, smaller firms are probably more likely to rely on multi-purpose tools, such as *Enterprise Workspace* technologies or *HR Information Systems* (HRIS), to identify and secure their users.

DUO

KnowBe4

jumpcloud

okta

Other data sources

Not all security data comes from security technologies. All of the organizations included in this analysis had integrated at least one data source from a system that falls outside the realm of what is traditionally owned or administered by security. Most commonly, these data sources include:

- Cloud Service Providers
- Source Code Management (SCM)
- Work Management & Project Tracking
- Human Resources Information Systems (HRIS)
- Collaboration and Communication Technologies
- Application Error Tracking
- DevOps & Site Reliability Engineering (SRE) Tools

Security is an inherently collaborative and interdependent function that is tasked with auditing, governing, using, and testing the security of business systems. Integrating non-security data sources is necessary to understand the attack surface, pass audits, and perform investigations with any degree of efficiency and accuracy. Tomorrow's teams are likely to pull from an increasingly broad range of data sources to systematically uncover unified cyber insights.

Security data

Are security teams correlating the right data?

Integrating a data source does not necessarily signify security maturity. But there is growing evidence to suggest that the most effective security teams are proactive about creating a security fabric. A higher number of data sources integrated into a single place enables security teams to understand complex cross-system dependencies and events.

All of the security teams analyzed have adopted at least one data source used to identify `DEVICES` and `APPLICATIONS`. The majority have visibility into network assets, while over 40% have integrated a system for user identities. The adoption of data sources used to identify or protect data is much lower and bears further investigation in future research efforts.

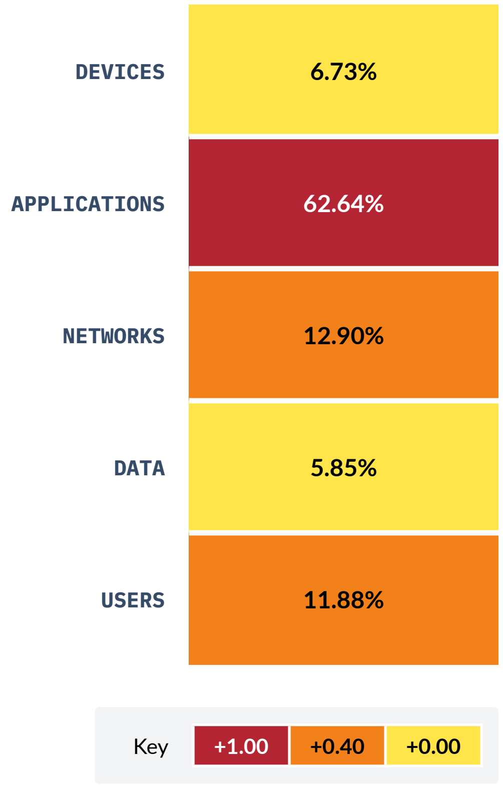
The loudest sources of security data by size and function

Not all security data sources are created equal. Some security data sources contribute more vulnerabilities and alerts since they are high volume. The security data sources that scream the loudest aren't necessarily the most valuable or worthy of attention. Instead, the security sources that generate the highest volume of data may demand more context in order to matter – the loudest sources generally need to be correlated with multiple sources of intelligence for actionable insight.

Our analysis revealed that 61% of all data utilized by practitioners comes from non-security systems, including CSPs, HRIS, task management, and DevOps tools. It is likely that security's need for data sources owned by DevOps, HR, and product teams will continue to increase. After all, "identity and context have become the ultimate control plane in a distributed environment that supports assets and access from everywhere," writes Felix Gaehtgens, James Hoover, Henrique Teixeira, et al., Gartner.

The remaining 39% of security data originates from security technologies. These sources are visualized across the asset superclass and cybersecurity functions below. While this analysis is not necessarily comprehensive of all security data sources, it's an interesting glimpse into the most high-volume data sources at over 200 organizations with security teams.

Heatmap 3.2: Relative percent of security data by sources



DEVICES	6.73%
APPLICATIONS	62.64%
NETWORKS	12.90%
DATA	5.85%
USERS	11.88%

Key	🟥	🟧	🟨
	+1.00	+0.40	+0.00

Loudest sources of security data by size and function

Relative volume of DEVICES asset data

Device data sources contribute just 6.73% of security data points, making them the second-quietest set of systems by asset superclass. While vulnerability scanners and device ID/IPS technologies are not particularly quiet, XDR and UEM solutions for physical endpoint devices generate a relatively low volume of data points on a daily basis.

Relative volume of APPLICATIONS asset data

Few readers should feel surprised that nearly 63% of security data points trickle from systems used to secure APPLICATIONS. It's a notion that tracks, particularly for individuals who have spent any time in the vicinity of SAST, DAST, WAF, and application vulnerability scanners since these technologies can produce a lot of volume. Application data sources generate a higher number of security data points than all other application superclasses combined, including the aggregate of data points on DEVICES, NETWORKS, DATA, and USERS.

Relative volume of NETWORKS asset data

Network data sources are neither loud nor quiet compared to other data sources, only contributing 12.9% of security data points that practitioners must correlate and prioritize. These data points are generated by domain registers, ID/IPS, and SASE technologies.

Relative volume of DATA asset data

Security practitioners have the fewest data points on the security of their DATA assets at just 5.85% of total. DATA is the asset class with the fewest data sources and the quietest data sources, which could indicate a lack of unified visibility at some organizations. Most of the volume in the DATA assets category is generated by secrets discovery and DLP systems.

Relative volume of USER asset data

USERS are arguably the most unpredictable of all asset classes, since it's the grouping that contains human users. Still, just 11.88% of security data points are generated by systems dedicated to user security, such as IAM, IDP, HRIS, and training technologies. Since security practitioners have limited ability to scan people for vulnerabilities, some degree of relative silence can be expected in this asset class. In addition, policies and procedures to protect employee privacy may limit the number of user security data sources that organizations integrate. HRIS, User Behavioral Analytics, and Learning Management Systems (LMS) may limit access to a very small number of trusted users and applications to protect PII.

11.8%

of security data points are generated by systems dedicated to user security data.

6.73%

Device data sources contribute just 6.73% of security data points.

63%

of security data points trickle from systems used to secure **APPLICATIONS**.

Section

4

Findings

What it means for security

Security teams, collectively, have more than twice the number of FINDINGS as assets. This year's SCAR analysis included 89.7 million APPLICATION, USER, NETWORK, DEVICE, and DATA assets, and an incredible 189.39 million security FINDINGS. More than 99% of the security teams analyzed have a higher number of FINDINGS, or liabilities, than assets.

Are organizations becoming more vulnerable year-over-year? The answer is most likely yes, but it begs the question of how to measure vulnerability year-over-year or quarter-over-quarter. Even with an agreed-upon definition of 'vulnerable' as a measure of weaknesses in a target, there are multiple ways that weaknesses could increase:

- Size of potential exposed attack surface
- Length and number of attack paths
- Number of critical assets with vulnerabilities
- Number of assets with exploitable vulnerabilities.

Numeric data on FINDINGS is not a meaningful indicator of the degree to which an organization or application is vulnerable, particularly not in isolation. Security practitioners must learn to live with an incredible number of vulnerabilities, and even more uncomfortably, defend this new reality to their colleagues in other business domains.

The most vulnerable asset superclasses

Asset relationships are a crucial source of context. Graphs provide security practitioners with a contextually rich understanding of assets, attributes, and the relationships between these nodes. Security data sources do not reveal actionable insights in isolation.

A security organization's FINDINGS by sheer number is not a meaningful indicator of posture, particularly not without any context. FINDINGS demand context from relationships that span systems, particularly the relationships between findings, assets, and policies:

- Are there internet-facing EC2 instances that are allowed access to non-public S3 buckets?
- Are there cross-account IAM trust relationships to external or vendor accounts?
- Which PRs or developer updates introduced new vulnerability findings this past week?
- Are there orphaned DNS records pointing to non-existent resources?
- Are services with one or more critical vulnerabilities exposed to the internet?

Analyzing the first-degree relationships between FINDINGS and asset superclasses can reveal more than vulnerability severity or source alone.

What it means for security

The realities of cloud-native security have created an upside-down reality when it comes to maturity and visibility. A high number of unresolved vulnerabilities does not necessarily indicate exposed weaknesses.

Instead, it could be a characteristic of organizations with a greater number of mechanisms for vulnerability discovery or companies with greater maturity in security data collection that are working to defend with graphs. Their FINDINGS are many, and come from many different sources.

The most vulnerable asset superclasses

Chart 4.1: Composition of FINDINGS by asset superclass

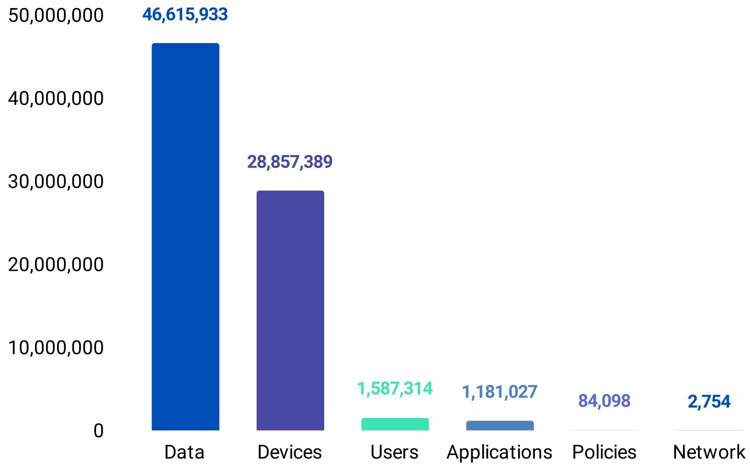


Table 4.1: Composition of FINDINGS by number and relative percent

Asset superclass	Number	Relative percent
DATA	46,615,933	59.51%
DEVICES	28,857,389	36.84%
USERS	1,587,314	2.03%
APPLICATIONS	1,181,027	1.51%
POLICIES	84,098	0.11%
NETWORK	2,754	0.00%*

*Represents fewer than 0.01% of all FINDINGS

FREQUENCY

DATA is the most vulnerable superclass of asset, at least on the basis of sheer frequency, comprising 59.51% of all FINDINGS. DEVICES is a distant second-place, but still comprises over one-third of all FINDINGS by superclass at 36.84%. The two most vulnerable superclasses, DATA and DEVICES, collectively represent 96.35% of all security FINDINGS, while the remaining 3.65% is distributed among USERS, APPLICATIONS, NETWORKS, and POLICIES.

The most vulnerable asset superclasses

DATA FINDINGS

Images, records, and containers account for 87% of the 46.62 million findings in the **DATA** superclass. A smaller, but still significant, percentage of findings in this superclass is attributable to records, datastores, and keys.

APPLICATIONS, POLICIES, AND NETWORKS FINDINGS

The three superclasses with the least findings – applications, policies, and networks – collectively represent just 1.62% of the security findings analysis. Configuration and control policy exceptions are most commonly linked to policy findings, followed by firewall and gateway findings.

DEVICES FINDINGS

Over 99% of the 28.86 million findings within the devices superclass are linked to cloud hosts.

USERS FINDINGS

While the 1.59 million findings linked to the user superclass is a relatively small percentage of the findings data set, the majority of these originate from accounts and access roles.

The most vulnerable asset superclasses

The most vulnerable asset classes

Just ten asset classes from four superclasses represent 77.64 million vulnerabilities, or more than 95% of the vulnerability triplets included in this analysis. While sheer number of vulnerabilities alone is not an exclusively reliable indicator of the degree to which an asset class has exposed weaknesses, it still bears further examination. Just a few asset classes are responsible for the vast majority of noise that security practitioners must contend with.

Focused analysis of the most vulnerable assets within the superclasses further highlights the fact that the majority of findings come from a tiny minority of asset classes, particularly Images and Hosts.

Chart 4.2: Top 10 most vulnerable asset classes by superclass

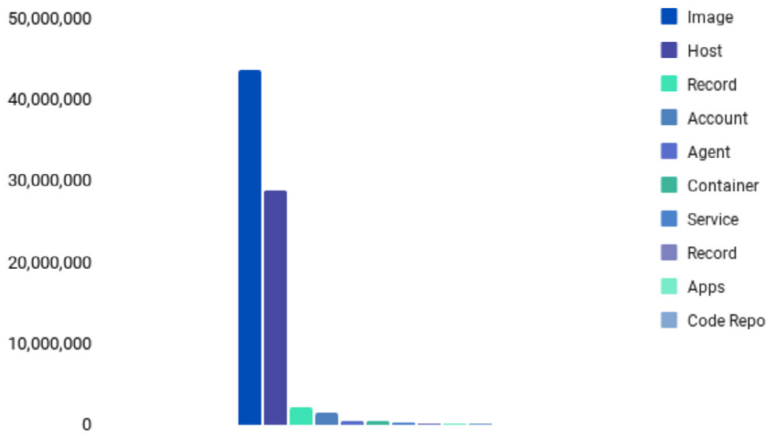


Table 4.2: Top 10 most vulnerable asset classes by number & relative percent

Class	Number	Relative percent	Superclass
Image	43,613,515	56.17%	DATA
Host	28,734,120	37.01%	DEVICES
Record	2,120,711	2.73%	DATA
Account	1,437,973	1.85%	USERS
Agent	487,832	0.63%	APPLICATIONS
Container	456,923	0.59%	DATA
Service	264,088	0.34%	APPLICATIONS
Record	211,053	0.27%	DATA
Application	206,249	0.27%	APPLICATIONS
CodeRepo	109,763	0.14%	APPLICATIONS

Findings by severity

Not all FINDINGS are created equal, at least not from a security practitioner's perspective. Security practitioners are generally significantly more interested in "Critical" and "High" alerts than those labeled as "Informational," particularly if those high-urgency alerts are attached to a critical asset or coming from multiple sources at the same time.

While first-degree relationships involving FINDINGS are important to security practitioners since they show which assets are connected to a vulnerability, they're still only one small part of the equation. The true criticality of the vulnerability depends on the criticality of the asset, the policies, network configurations, and countless other factors. By reconciling vulnerability classifications by many different vendor approaches, one can tell the true industry distribution of FINDINGS by severity.

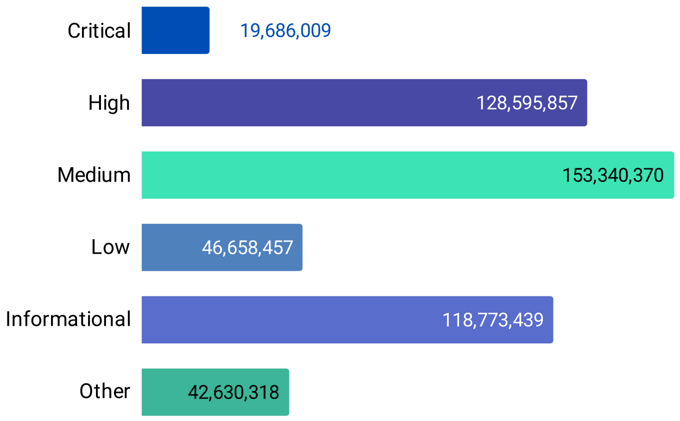
While severity classification is just one part of the contextual puzzle that security teams must navigate, it's an undeniably important component of understanding the potential exploitability of security FINDINGS. Vulnerabilities categorized as "Low" and "Informational" collectively comprise nearly one-third (32.4%) of the FINDINGS. In contrast, "Critical" and "High" vulnerabilities represent 29.1% total, at 3.9% and 25.2%, respectively.

Table 4.3: Composition of **FINDINGS** *by severity classification, number, and relative percent*

Severity classification	Number	Relative percent
Critical	19,686,009	3.9%
High	128,595,857	25.2%
Medium	153,340,370	30.1%
Low	46,658,457	9.2%
Informational	118,773,439	23.3%
Other*	42,630,318	8.4%

**The 'Other' classification consists primarily of FINDINGS from homegrown systems or integrations without available severity classification, or proprietary vendor classifications that cannot easily be reconciled with SCAR severity classifications.*

Chart 4.3: Composition of **FINDINGS** *by severity classification*



Findings by severity and organizational size

Mid-sized and small organizations

Mid-sized and small organizations have fewer critical vulnerabilities than their larger peers. Organizations with 500 or fewer employees had a significantly higher number of informational vulnerabilities that generally don't require an urgent remediation effort.

Large organizations

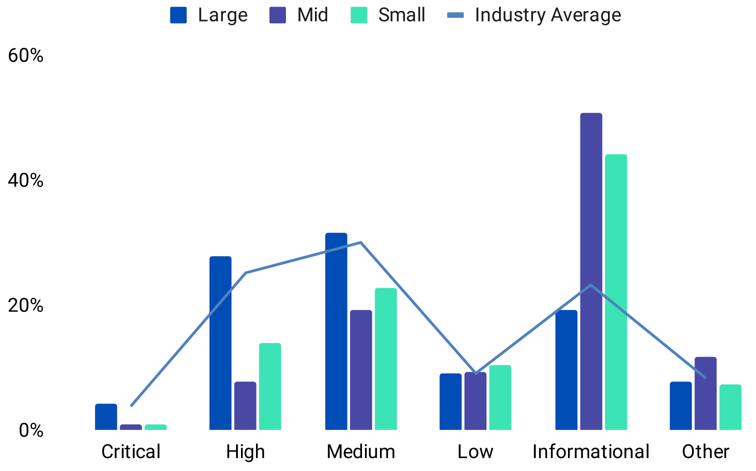
Large organizations with 500 or more employees have far more Critical, High, and Medium severity vulnerabilities compared to their smaller counterparts. Do larger organizations have more critical vulnerabilities since these vulnerabilities are longer-lived? If so, why?

We hypothesize that organization size or employee count has very little to do with vulnerability severity. Instead, large organizations may have more complicated infrastructure and legacy systems that are harder to patch, update, and protect. End-of-life, legacy systems probably result in longer-lived vulnerabilities, not size.

Table 4.4: Relative proportion of **FINDINGS** *by severity and organizational size*

Class	Large	Mid	Small	Industry average
Critical	4.29%	1.05%	0.96%	3.9%
High	27.81%	7.72%	14.07%	25.2%
Medium	31.65%	19.29%	22.82%	30.1%
Low	9.11%	9.31%	10.44%	9.2%
Informational	19.25%	50.80%	44.24%	23.3%
Other	7.89%	11.84%	7.46%	8.4%

Chart 4.4: Relative proportion of **FINDINGS** *by severity and organizational size*



Findings by severity and and industry

What does industry have to do with the severity of an organization's vulnerabilities? According to this initial analysis, very little. While there is variation among industries, it doesn't necessarily follow the patterns one might expect. We hypothesized that organizations in highly regulated industries such as finance and healthcare may have fewer Critical or High vulnerabilities due to stricter regulatory and compliance requirements, but the data did not support this hypothesis.

Chart 4.5: Relative proportion of **`FINDINGS`** *by severity and industry*

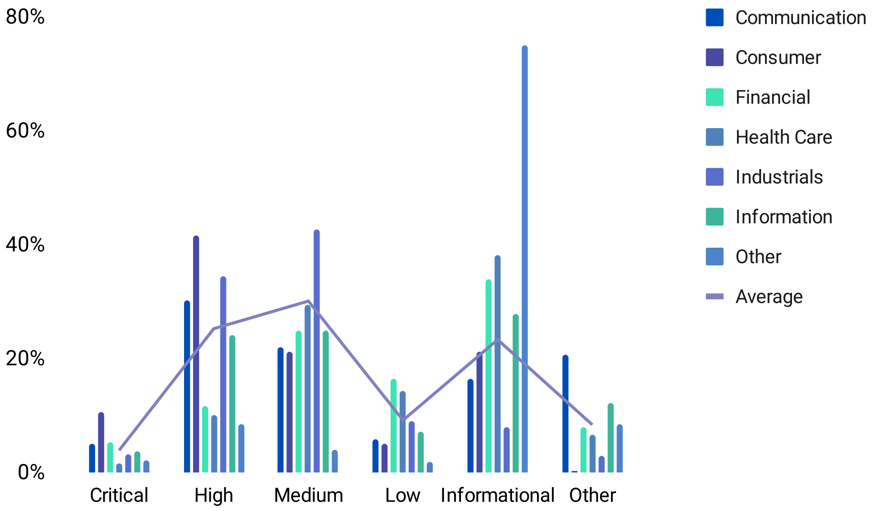


Table 4.5: Relative proportion of **`FINDINGS`** *by severity and industry*

Severity	Communication	Consumer	Financial	Health care	Industrials	Information	Other	All industries (Average)
Critical	5.11%	10.46%	5.26%	1.63%	3.13%	3.77%	1.98%	3.9%
High	30.21%	41.69%	11.62%	9.99%	34.41%	24.00%	8.58%	25.2%
Medium	22.02%	21.29%	24.80%	29.29%	42.56%	24.95%	4.09%	30.1%
Low	5.76%	5.15%	16.44%	14.41%	9.03%	7.11%	1.89%	9.2%
Informational	16.34%	21.07%	33.84%	38.18%	7.91%	27.93%	75.09%	23.3%
Other	20.57%	0.35%	8.03%	6.50%	2.97%	12.24%	8.38%	8.4%

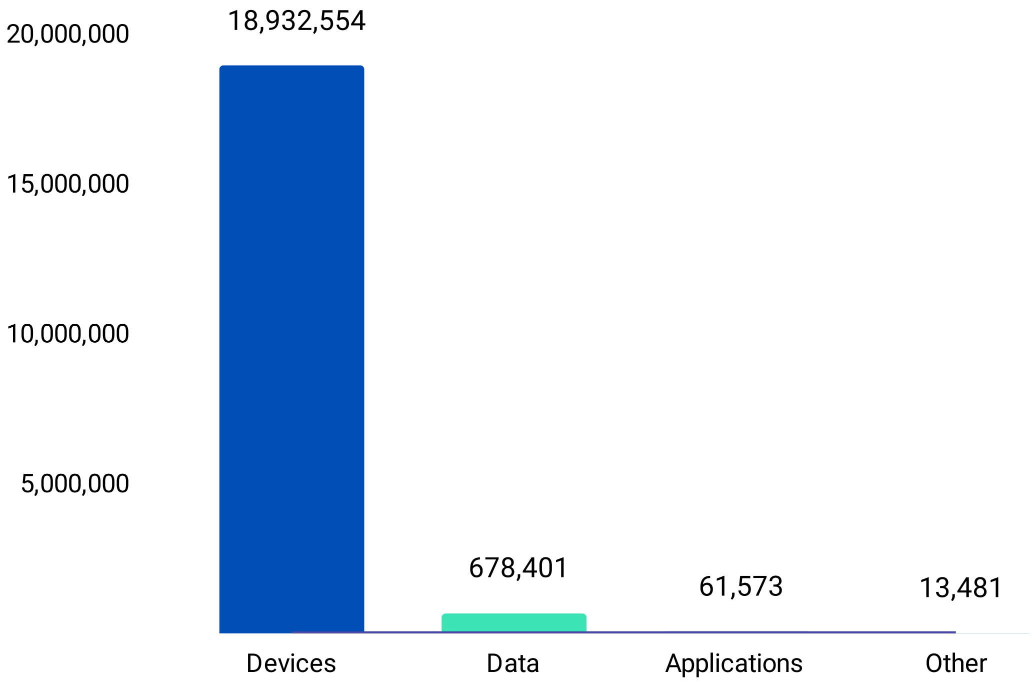
The source of critical findings

Prior analysis in this report has shown that a small handful of superclasses are linked to the vast majority of FINDINGS. DATA, DEVICES, and USERS are linked to 98.38% of all security FINDINGS, while NETWORKS, POLICIES, and APPLICATIONS are linked to the remaining 1.62% of all security FINDINGS.

The seven superclasses in this analysis represent the aggregation and clustering of a total of 125 unique asset classes from the asset-relationship graph data model. Within those 125 classes, just two classes – Hosts and Images – are responsible for an incredible 93.18% of all security FINDINGS.

Given the fact that Hosts and Images yield a disproportionate number of FINDINGS, are they also responsible for the majority of critical vulnerabilities? The answer is complicated.

Chart 4.6: Composition of critical security **FINDINGS** *by superclass*



The source of critical findings

DEVICES – especially cloud hosts – are linked to just over one-third of security FINDINGS (36.84% of FINDINGS), but represent 96.1% of critical FINDINGS.

While assets in the DATA superclass represent a significant 59.51% of all security findings, they represent just 3.45% of all security findings classified as critical. This likely indicates that cloud hosts are the source of the majority of critical security findings, based on the potential severity and/or exploitability of the exposed security weaknesses.

Why do cloud devices represent such an incredibly disproportionate number of critical findings? Data shows that cloud host vulnerabilities are skewed toward more critical classifications than other asset classes. This could be due to the potential for cloud host vulnerabilities to have a greater impact compared to other assets. It also could be due to the inherent complexity of cloud host security, including factors of access, network connectivity, and shared responsibility with CSPs.

In particular, it is worth examining whether cloud host vulnerabilities are more frequently downgraded to a lower severity classification or if these vulnerabilities have much longer lifespans than other asset classes.

Table 4.6: Composition of critical security FINDINGS by superclass

Superclass	Critical FINDINGS	Relative percent of critical FINDINGS	Relative percent of all classifications of FINDINGS	Proportional difference in critical FINDINGS
DATA	18,932,554	96.17%	36.84%	(-59.33%)
DEVICES	678,401	3.45%	59.51%	56.06%
APPLICATIONS	61,573	0.31%	1.51%	1.21%
Other	13,481	0.07%	2.14%	2.07%

**Comprised of the USERS, NETWORKS, and POLICIES superclasses*

Where vulnerabilities come from

An in-depth look at where vulnerabilities come from

`FINDINGS` represent a disproportionate share of the data in the SCAR analysis, and they also demand a disproportionate share of security practitioners' attention on a daily basis. To better understand the source of security findings, we analyzed the types of security technologies that generate this superclass of data, as well as the average number of findings per source.

Chart 4.7: Average (mean) number of security findings, top 10 data sources

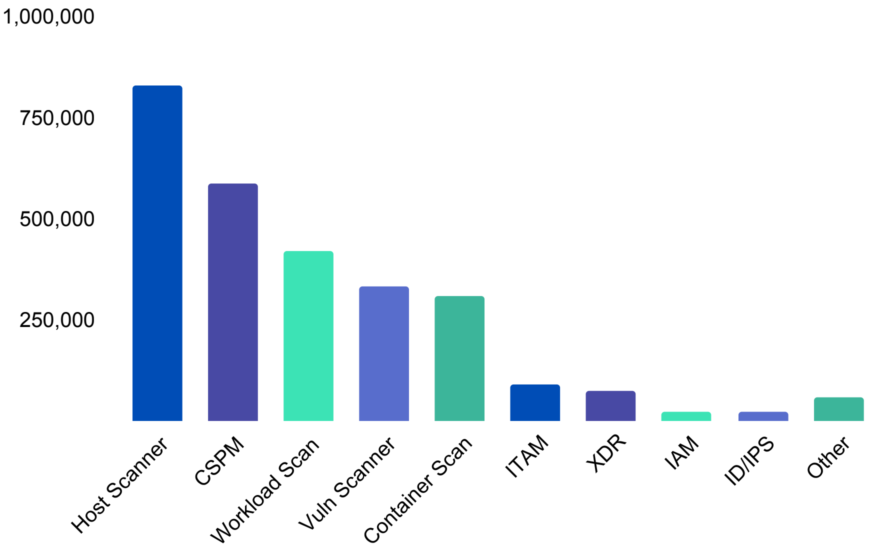


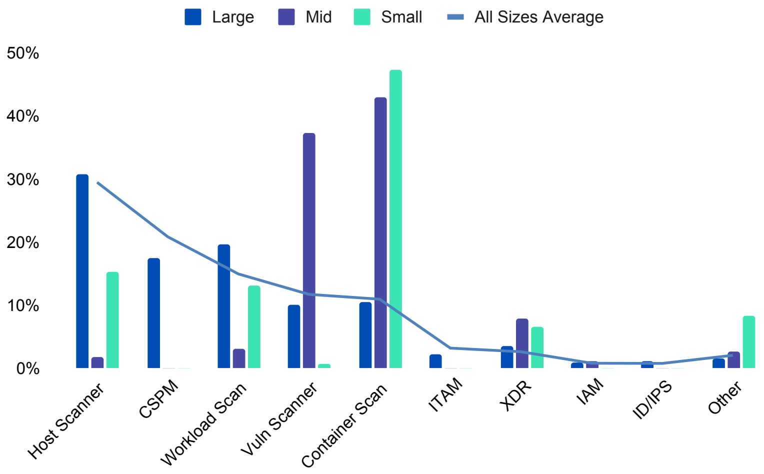
Table 4.7: Average (mean) number of security findings, top 10 data sources

Findings source	Average number	Relative percent
Host scanner	830,042.32	30.16%
CSPM	587,400.07	21.34%
Workload scan	421,750.43	15.33%
Vuln scanner	330,797.67	12.02%
Container scan	309,322.55	11.24%
ITAM	91,256.09	3.32%
XDR	75,125.84	2.73%
IAM	23,912.33	0.87%
ID/IPS	23,128.61	0.84%
Other*	59,241.95	2.15%

**Includes threat feeds, secrets discovery technologies, DAST, SAST, SSPM, IAC, and more.*

Where vulnerabilities come from

Chart 4.8: Relative frequency of security findings by organizational size, top 10 data sources



When the top ten sources of security findings are analyzed by organizational size, it reveals that the number and variety of data sources are likely positively correlated with the number of employees. Large organizations have far more findings originating from Cloud Security Posture Management (CSPM), Information Technology Asset Management (ITAM), and ID/IPS than their smaller-sized counterparts.

Table 4.8: Average (mean) number of security findings by top data sources & size

	Large	Mid	Small	All sizes (Average)
Host scanner	1,218,343.85	11,379.50	15,495.36	830,042.32
CSPM	688,613.53	589.00	-	587,400.07
Workload scan	777,755.45	18,678.61	13,364.53	421,750.43
Vuln scanner	403,046.19	227,691.59	671.47	330,797.67
Container scan	416,192.24	262,985.47	48,234.86	309,322.55
ITAM	91,256.09	-	-	91,256.09
XDR	139,234.38	48,693.19	6,791.78	75,125.84
IAM	41,380.82	7,778.15	118.00	23,912.33
ID/IPS	48,717.50	538.67	28.12	23,128.61
Other*	62,700.53	16,376.18	8,504.66	59,241.95

**Includes threat feeds, secrets discovery technologies, DAST, SAST, SSPM, IAC, and more.*

Section

5

Securing the cloud attack surface

CSP adoption among scar organizations

60% of the assets included in this analysis originate from a cloud service provider (CSP) environment.

This means that a staggering percentage of the data that security teams utilize originates from their CSP integration, which includes a mixture of cloud resources such as cloud devices and data, as well as CSP services used for security functions such as intrusion detection and prevention (ID/IPS), identity and access management (IAM), and policy.

Security teams in 2023 are not tasked with governing a single AWS account or GCP project. Instead, organizations of all sizes have a growing number of environments across CSPs, representing segmentation of environment by purpose (develop, test, production, archive), product line, or customer.

CSP adoption among SCAR organizations

All organizations included in the SCAR analysis have adopted the cloud to a great extent, although some organizations are 'cloud-native' while others are cloud adopters due to the longevity of the organization, business model, or countless other reasons. SCAR organizations are all organizations that have adopted a cyber asset relationship graph for security, which may have once implied a greater sophistication of cloud adoption than industry averages. But, research from reliable third-party sources such as O'Reilly and Fortinet shows the gap is closing and that premises-based organizations are rapidly approaching extinction.

While 100% of SCAR organizations use the cloud, O'Reilly Cloud Adoption report data shows that 90% of organizations are utilizing the cloud across industries, with 77% using or pursuing a cloud-native strategy. Among the largest organizations with 1,000 or more employees, cloud adoption is 94%.

SCAR organizations have a bit deeper cloud adoption than true industry-wide averages, and they're also slightly more likely than industry average to be multi-cloud. 31% of the organizations included in the SCAR analysis are using three CSPs, with a presence in AWS, GCP, and Azure, compared to 26% of organizations using all CSPs in a recent VMWare study.

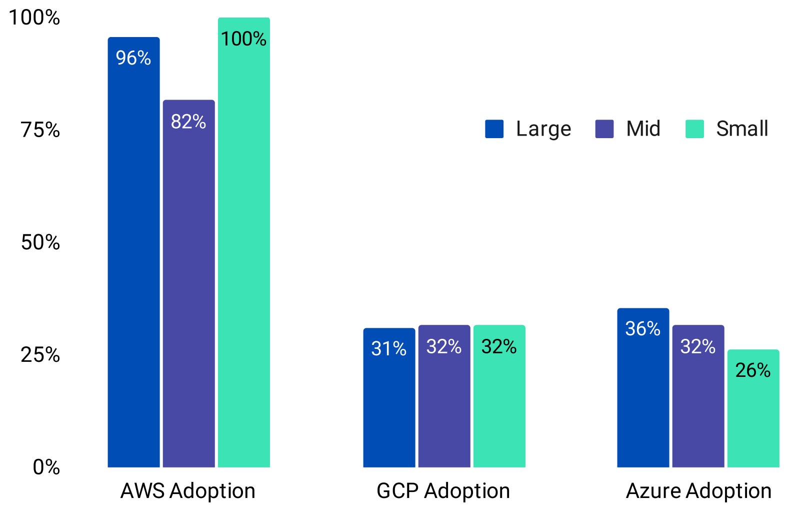
CSP adoption among SCAR organizations

AWS is the dominant cloud provider among SCAR organizations and in private industry, with an average of 92.44% adoption across organizations of all sizes included in this analysis. Large organizations are more likely than average to use Azure, with 35.56% adoption rates. Mid-sized organizations had the lowest AWS adoption rates at just 81.58% total. Small organizations kept pace with larger organizations on GCP adoption (31.58% adoption versus 31.42% median adoption across all sizes) and barely lagged on Azure adoption (26.32% adoption versus 31.15%), with 100% adoption of AWS.

Table 5.1: Average (mean) CSP adoption by organization size

Values	Large	Mid	Small	All sizes (Average)
AWS adoption	95.74%	81.58%	100.00%	92.44%
GCP adoption	31.11%	31.58%	31.58%	31.42%
Azure adoption	35.56%	31.58%	26.32%	31.15%

Chart 5.1: Average (mean) CSP adoption by organization size



Average AWS, GCP, & Azure environments by size

The average, or mean, resources that security teams at large organizations are tasked with securing are nearly 225 total AWS accounts, GCP projects, and Azure subscriptions.

Security teams at small organizations have 171.05 total environments to secure, while mid-sized organizations are responsible for securing 559.24 unique accounts, projects, and subscriptions across major CSPs. Organizations with fewer than 500 employees have more unique AWS accounts, GCP projects, and Azure subscriptions than employees, on average, further illustrating the significant burden on security teams and the expanding attack surface.

Further research is necessary to understand whether the increase in CSP accounts represents a continued trend toward greater segregation in cloud environments and distributed architecture patterns. Just as importantly, research is needed to understand the motives behind a greater number of projects – is this driven by a need for isolation architecture, resource segmentation, or other reasons?

Table 5.2: Average (mean) unique CSP accounts by organization size

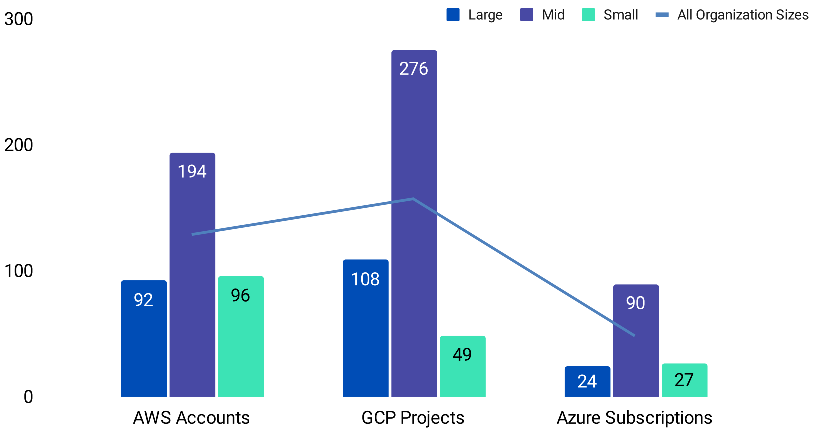
Values	Large	Mid	Small	All sizes (Average)
AWS Adoption	91.98	194.03	95.63	128.74
GCP Adoption	108.47	275.66	48.58	157.104
Azure Adoption	24.49	89.55	26.84	48.25
Total CSP Accounts	224.94	559.24	171.05	334.09

Distributed architecture and the DIE triad

The DIE Triad is an architectural model by Sounil Yu, which utilizes Distributed, Immutable, and Ephemeral solutions to directly support the organization's capacity for recovery and resilience.

A higher number of CSP accounts is one way for organizations to adopt distributed architectural techniques. A higher number of CSP accounts is likely to reduce the number of cyber assets and critical assets in each account, project, or subscription. Distributed architecture can improve resiliency by reducing dependence on single systems. It can also reduce the blast radius of potential incidents.

Chart 5.2: Average (mean) unique CSP accounts by organization size



Average AWS, GCP, & Azure environments by industry

Industry-based analysis of CSP environments yielded results that may challenge many preconceptions about technological differences between industries and sectors. The information sector doesn't necessarily outpace its peers when it comes to the total number of CSP accounts, projects, and subscriptions.

The Communication industry – consisting of organizations in telecommunication services, media, and entertainment – had the highest number of CSP environments by a significant margin with a particularly high number of GCP projects. The Industrials industry, which represents capital goods, commercial & professional services, and transportation, also had a significantly higher-than-median number of AWS Accounts, GCP projects, and Azure subscriptions. Organizations in the Information sector also out-paced the mean, but had fewer CSP accounts than peers in Industrials and Communication.

Further research is needed to better understand how motives and benefits for more distributed CSP architecture varies across industries, and the difference in perceived benefits between consumer and business-serving organizations. In particular, future research should examine the role being in a 'highly regulated' industry can play in an organization's decision to consolidate its CSP accounts instead of greater segmentation, and whether health care regulations have a greater impact than financial regulations in this area.

What it means for security

An estimated 90% of organizations are vulnerable to cross-account security vulnerabilities. Unlike other CVEs, cross-connect cloud security vulnerabilities aren't always reported or tracked in a central location by vendors. Instead, cross-account vulnerabilities are generally the result of toxic combinations of misconfigurations, privileges, and absent controls.

Isolation architecture is an increasingly important part of securing the cloud. Security teams have spent 2022 striving to minimize the amount of trust between cloud resources, while increasing the number of tight perimeters. Deploying additional CSP accounts is not the only isolation architecture approach that security teams used in 2022, but it's one that stood out in our SCAR analysis.

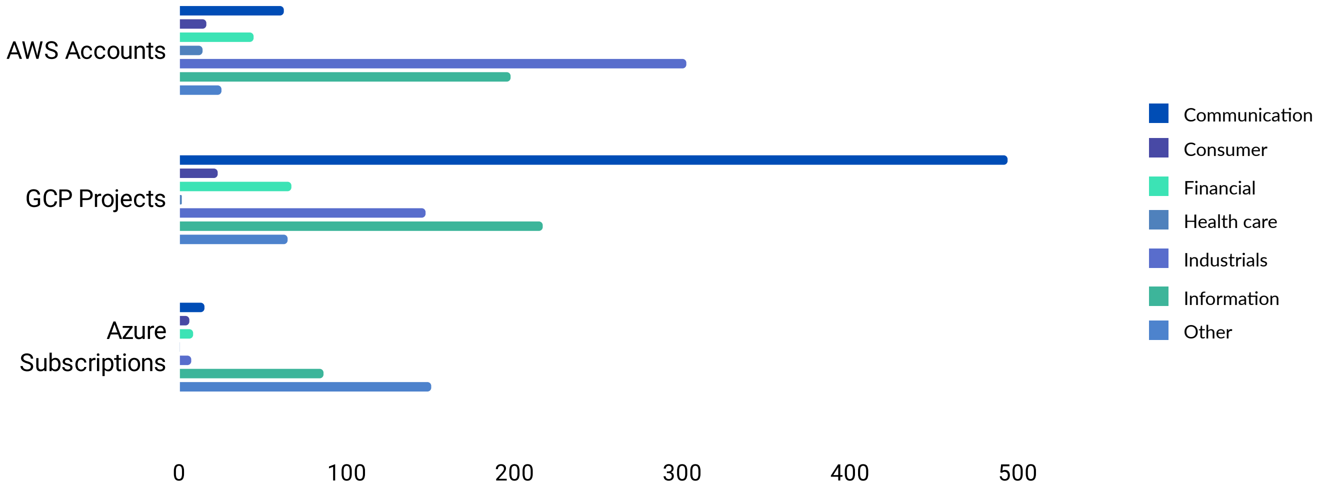
Security teams spent the past year deploying unprecedented numbers of AWS accounts, GCP projects,and Azure subscriptions as one way to create secure perimeters in the cloud.

Average AWS, GCP, and Azure environments by industry

Table 5.3: Average (mean) number of CSP environments by industry

Values	Communication	Consumer	Financial	Health care	Industrials	Information	Other
AWS Adoption	62.71	16.00	44.47	14.31	302.22	197.93	25.40
GCP Adoption	493.71	23.50	67.37	2.08	146.44	216.33	64.60
Azure Adoption	15.57	6.33	8.47	0.38	7.11	85.54	150.80
Total	572.00	45.83	120.32	16.77	455.78	499.80	240.80

Chart 5.3: Average (mean) number of CSP environments by industry



Section

6

Relationships and queries

Asset relationships

A relationship is the connection between two or more cyber assets. Assets in isolation don't tell the complete story – it's how they interoperate and work together that provides value.

As companies move more of their assets and activities to the digital environment, relationships have become more complex and understanding the connections between assets more important. As a result, when a security incident happens, the data you need to assess often lives in unrelated systems and tooling.

Relationships between cyber assets matter a lot, especially when considering the following scenarios:

- An admin account probably needs attention if it's owned by a former employee.
- An application generally shouldn't be related to dozens of privileged users.
- A system should almost always be related to one system owner.
- A policy's relationships to critical assets or privileged users can determine the immediate blast radius potential for an incident.

Threat actors have long recognized the importance of relationships. The relationship between an over-privileged user and sensitive assets is how and why social engineering and account takeover are highly successful tactics for threat actors. The blast radius of a realized risk is the product of relationships around a vulnerable asset and includes user permissions, configurations, and integrations.

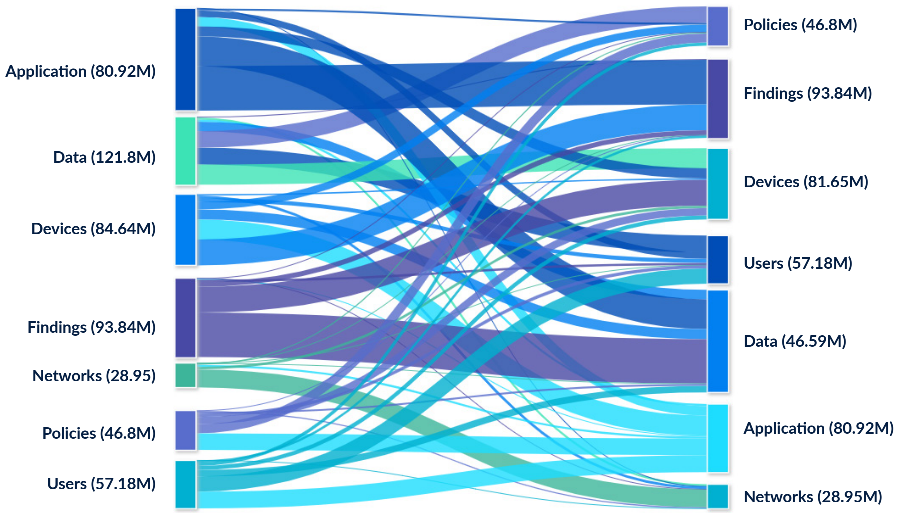
Asset relationships

For the purpose of this analysis, we examined the most common types of relationships between all seven superclasses of cyber assets.

Table 6.1: Frequency and relative percent of asset relationships by superclass

Superclass	Relationship frequency	Relative percent
APPLICATIONS	80,922,345	15.86%
DATA	121,801,260	23.87%
DEVICES	84,643,632	16.59%
FINDINGS	93,843,511	18.39%
NETWORKS	28,953,130	5.67%
POLICIES	46,796,079	9.17%
USERS	5,340,0481	10.46%

Chart 6.1: Relationships between assets and attributes by superclass



Triplets, or first-degree relationships

Triplets represent first-degree relationships between assets. A user and their laptop are a first-degree relationship between a user and device.

Chart 6.2: Frequency of triplets by superclass

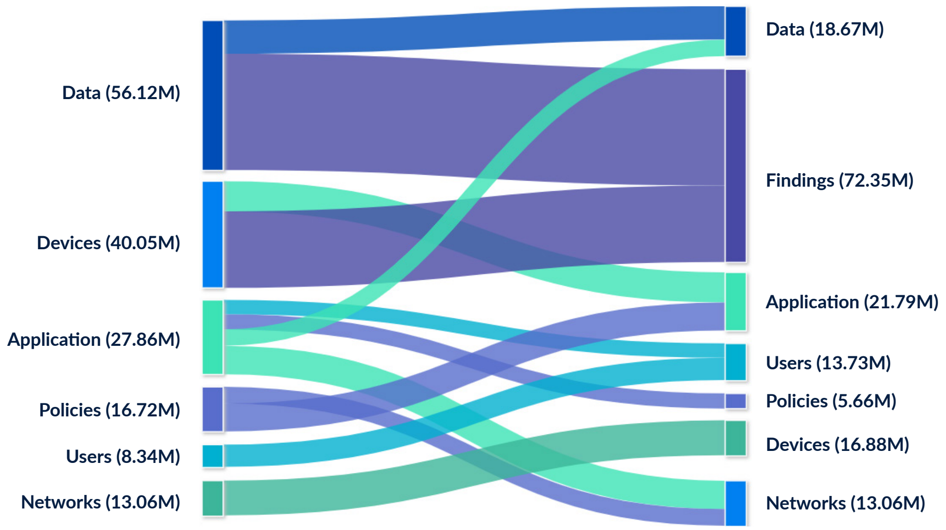


Table 6.2: Frequency of triplets by superclass

Superclasses	Frequency
DATA & FINDINGS	43,613,515
DEVICES & FINDINGS	28,734,120
APPLICATIONS & DEVICES	21,957,891
APPLICATIONS & POLICIES	16,146,341
NETWORKS & NETWORKS	13,061,384

Superclasses	Frequency
DATA & FINDINGS	12,507,299
DEVICES & FINDINGS	8,344,929
APPLICATIONS & DEVICES	6,237,957
APPLICATIONS & POLICIES	6,164,793
NETWORKS & NETWORKS	5,380,903

A short note on directed vs. undirected graphs

Chances are, a small section of our most graph-obsessed SCAR readers are asking, “But what about directed relationships?” Great question.

Graph relationships come in two primary varieties, which are often called “directed” and “undirected.”

- **Undirected** graphs refer to two-way relationships.
- **Directed** graphs refer to one-way relationships.

Both directed and undirected relationships exist in cyber asset relationship graphs, although directed one-way relationships are significantly more common. The visualization of relationships does not account for directed relationships, however, since we had to draw the line somewhere for simplicity of analysis.

The complexity of security queries

Graph databases allow security practitioners to ask complex questions. But, how complex are the questions that practitioners are asking?

One possible measure is the number of asset classes included in a single query. Queries that include two or three asset and attribute superclasses can indicate a greater level of sophistication since they pertain to cross-system relationships.

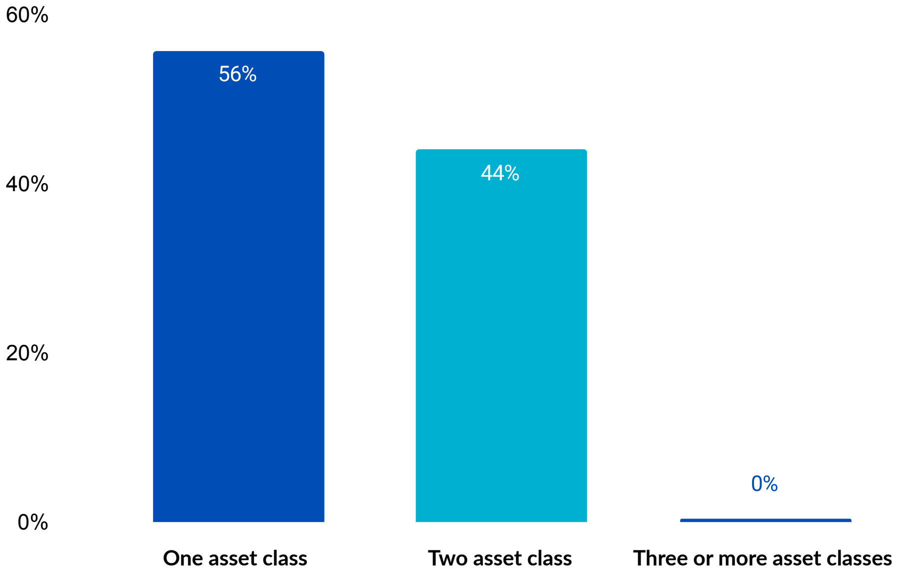
- A query for all findings concerns one superclass, and is likely to yield an overwhelming number of results for security practitioners.
- A query for all critical assets with findings effectively involves two classes, and will yield slightly more actionable results.
- Querying for all critical assets with a critical finding and policy exception concerns three superclasses, and is likely to provide the smallest number of true positives by surfacing the most important assets with security vulnerabilities and potential misconfigurations.

While query complexity can provide some insight into security practitioner behavior and how they interact with their data sources, it's not a perfect measure of sophistication. The number of asset classes involved in a query is just one measure of how security practitioners ask questions or interact with their cyber assets. Query parameters are another method practitioners can use to surface the most important results, by specifying whether findings should be open, the period of time in which an asset was created, or the team or individual who should own an asset.

Table 6.3: Number of asset classes in queries by relative percent

One asset class	55.70%
Two Asset classes	43.98%
Three or more asset classes	0.32%

Chart 6.3: Number of asset classes in queries by relative percent



Relationships and queries

The most frequently queried superclasses of cyber asset

Not all asset classes receive equal attention from security practitioners, at least as not when it comes to querying.

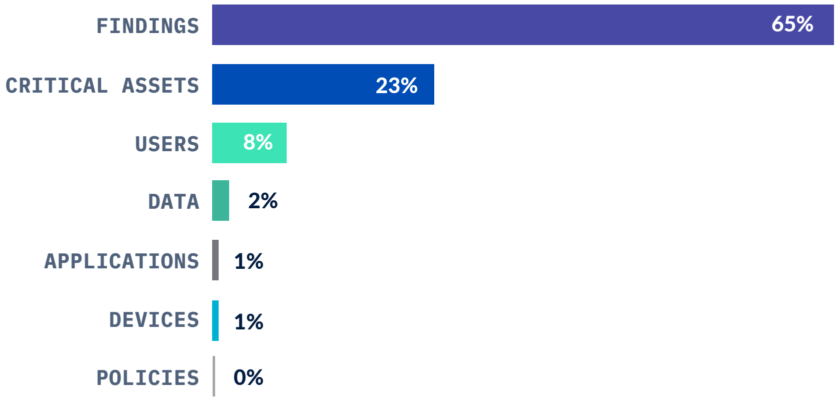
- FINDINGS receive a dominant share of attention at nearly two-thirds of total practitioner queries (65.35%).
- Critical assets, a mission-critical combination of APPLICATIONS, DATA, and DEVICES , receive less than one-quarter of attention from all security practitioner queries.
- USERS are the focus of nearly 8% (7.95%) of security practitioner queries
- Less than 4% of total queries focus on FINDINGS, DATA, DEVICES, and POLICIES.
- NETWORKS are very infrequently queried by security practitioners.

Table 6.4: Relative frequency of queries by asset superclass

APPLICATIONS	0.80%
Critical Assets	23.42%
DATA	1.83%
DEVICES	0.59%
FINDINGS	65.35%
POLICIES	0.05%
USERS	7.95%

**Critical asset definitions vary by security organization, but typically span multiple asset classes including APPLICATIONS, DATA, and DEVICES, or applications, code repos, data stores, functions, hosts, and logs.*

Chart 6.4: Relative frequency of queries by asset superclass



Section

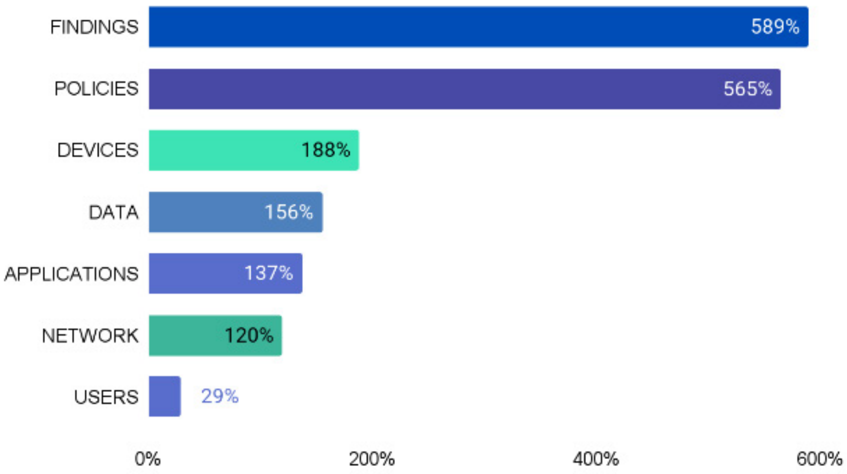
7

Year-over-year change

Table 7.1: Average (mean) number of cyber assets, year-over-year change

Superclass	2022 average	2023 average	Change	Percent change
APPLICATIONS assets	165,633	393,419	227,786	+ 137.36%
DEVICES assets	32,190	92,862	60,672	+ 188.48%
NETWORK assets	22,277	48,970	26,693	+ 119.82%
DATA assets	59,971	153,232	93,261	+ 155.51%
USERS assets	35,018	45,125	10,107	+ 28.86%
FINDINGS assets	120,561	830,639	710,078	+ 588.98%
POLICIES assets	8,345	55,473	47,128	+ 564.75%

Chart 7.1: Average (mean) number of cyber assets, year-over-year change



Section

8

In conclusion

In conclusion

By analyzing millions of data points to summarize the state of cyber asset inventories each year, we learn a lot about where security practitioners are focused.

Over the past 12 months, there's been an incredible – and almost certainly unprecedented – growth in the security practitioners' inventory of cyber assets. The average backlog of security findings has grown over three times faster than asset inventories, with nearly 600% year-over-year.

In conclusion, we want to surface and summarize the most interesting discoveries from this journey – not because we have all the answers, but because we have millions of data points on how little we truly know:

1. Unified cyber insights are critical

Security practitioners aren't omniscient. Visibility into cross-system relationships is only as good as the integration and correlation across data sets. **Unified cyber insights** matter a lot if anyone wants to effectively defend the cloud-native attack surface, but teams may be struggling to make a case for data access to systems owned or administered by other teams.

2. Cyber assets are business assets

Everyone knows that modern businesses cannot function, let alone succeed, without their cyber assets in both cloud and physical environments. Still, security teams have long struggled to convince business leaders just how much cyber assets are worth. Understanding that the average asset is worth $17,711 in 2023 may not help security teams get enough budget, but it feels like a start towards quantifying the value of cyber assets. Ideally, future SCAR research will include richer quantitative measures of critical asset value or the potential liabilities of security findings.

3. We have a cloud host and image problem

The discovery that 96.1% of security findings are linked to cloud hosts and images is unsettling and demands greater attention towards the ways in which we detect, prioritize, and remediate vulnerabilities across asset classes. Are some assets really hyper-vulnerable, or are those assets unreasonably likely to collect tons of critical findings?

4. The modern attack surface is distributed

Security practitioners are frequently responsible for hundreds of unique CSP accounts each in 2023. Distributed architecture is necessary to isolate our sensitive data and reduce the potential impact of destructive attacks. But, this hyper-growth in distributed architecture techniques has introduced a new era of complexity for multi-cloud security teams that must contend with less standardization than ever before and create a centralized control plane for common policy.

Appendix A: Firmographics

Firmographics (sometimes referred to as emporographics or firm demographics) are characteristics of businesses used for segmentation and research, including company size, location, and industry.

Firmographics were available for a subset of the total SCAR data population, which includes several thousands organizations. This was further narrowed to a total of 228 organizations with enterprise JupiterOne accounts, with much analysis focusing on a subset population whose integrations and data ingestion represented the most significant usage patterns, yielding a richer understanding of cyber assets, asset attributes, and relationships across security data sources.

The deeper focus on firmographics and variation across organizations by size and industry is a new feature of the 2023 SCAR report to provide our readers with a better understanding of how they stack up against similar peers.

Firmographic data enrichment was performed prior to analysis to protect the anonymity of research subjects during the analysis stage of SCAR research. Firmographic characteristics were pulled from reliable sources of OSINT and private data repositories such as ZoomInfo. Manual validation was performed on a randomly selected sample of enriched data to ensure accuracy of industry classification, employee counts, and market capitalization data.

Appendix A: Firmographics

SCAR firmographics by size

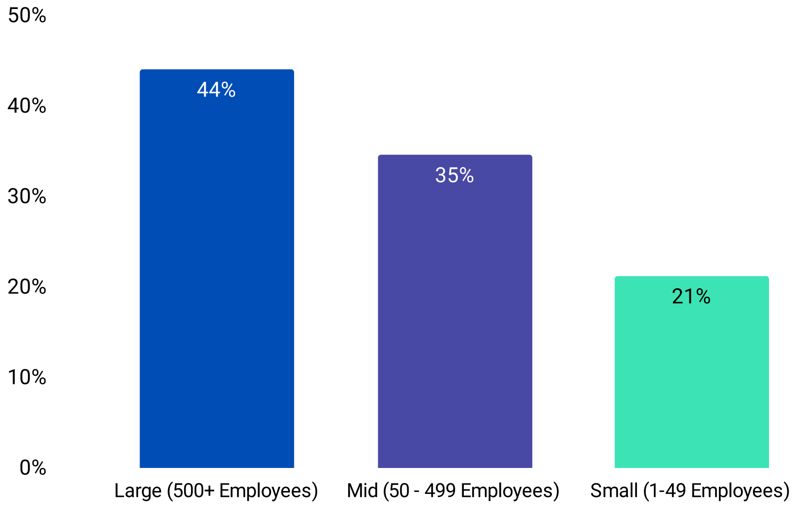
The size classification in the SCAR report is based on the number of employees, using the category definitions from the ADP Employment Report® developed by the ADP Research Institute in collaboration with Stanford Digital Economy Lab.

The relative proportions of large, mid-sized, and small organizations in the SCAR data set does not necessarily reflect the size ratios of the US economy. Small organizations with 49 or fewer employers represent the largest proportion of US employers, according to NAICS, meaning the SCAR dataset has an overrepresentation of large and mid-sized organizations. This dataset is most likely a closer reflection of US organizational firmographics among businesses that employ a security team, according to data on cybersecurity hiring from Zippia.

Table a.1: SCAR organization sizes by relative percent

Size category	Relative percent
Large (500+ Employees)	44.09%
Mid (50 - 499 Employees)	34.65%
Small (1 - 49 Employees)	21.26%

Chart a.1: SCAR organization sizes by relative percent



Appendix A: Firmographics

Firmographics by industry

The industry groups used in the SCAR are based on the sectors utilized in the MSCI's Global Industry Classification Standard (GICS®). Some GICS® were clustered, or combined to ensure greater anonymity for included organizations and more equivalent proportional representation among industry groups.

'Consumer' industry definition

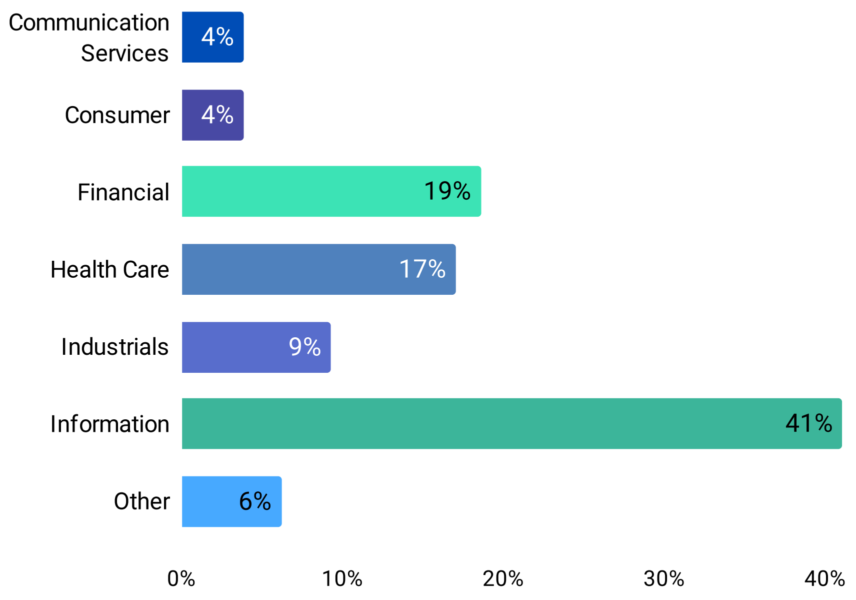
The SCAR analysis has grouped the consumer discretionary and consumer staples category into a single 'consumer' category that includes a broad mix of retailers, consumer services, food & beverage, and household products.

'Other' industry definition

Organizations from the real estate, utilities, and energy sectors have been grouped into a single category labeled 'Other' due to relatively low proportional representation from each of these three GICS® sectors.

It is worth noting that the SCAR uses 'industry' in the same way that the GICS® standard uses 'sector'. Within the original, reference classification standard, the 11 GICS® sectors are further divided into 24 sub-classifications of industry groups, 69 industries, and 158 sub-industries. The reference table below lists the standard's sub-categories to help readers better understand the types of organizations and business models represented in each category of SCAR industry.

Chart a.2: SCAR organizational industry by relative percent



Appendix A: Firmographics

Table a.2: SCAR organizational industry by relative percent

Industry	Equivalent GICS® sector(s)	SCAR relative percent
Communication	Communication services	3.88%
Consumer	Consumer discretionary, consumer staples	3.88%
Financial	Financials	18.60%
Health Care	Health care	17.05%
Industrials	Industrials	9.30%
Information	Information technology	41.09%
Other	Energy, Utilities, Real Estate	6.20%

Table a.3: How SCAR industries map to GICS® sector & industry

SCAR industry	GICS® sector(s)	GICS® industry groups
Communication	Communication services	• Telecommunication Services • Media & Entertainment
Consumer	Consumer discretionary	• Automobiles & Components • Consumer Durables & Apparel • Consumer Services • Retailing
	Consumer staples	• Food & Staples Retailing • Food, Beverage & Tobacco • Household & Personal Products
Financial	Financial	• Banks • Diversified Financials • Insurance
Health care	Health care	• Health Care Equipment & Services • Pharmaceuticals, Biotechnology & Life Sciences
Industrials	Industrials	• Capital Goods • Commercial & Professional Services • Transportation
Information	Information technology	• TSoftware & Services • Technology Hardware & Equipment • Semiconductors & Semiconductor Equipment
Other	Energy	• Energy
	Materials	• Materials
	Utilities	• Utilities
	Real estate	• Real Estate

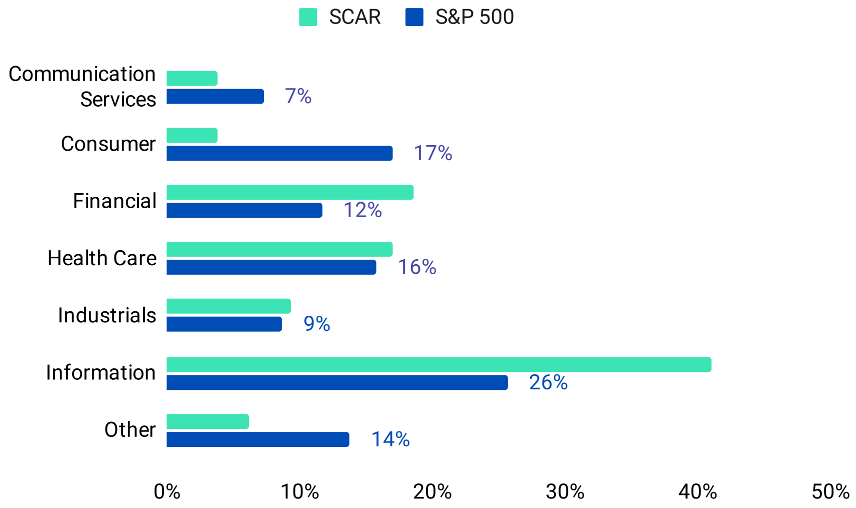
Appendix A: Firmographics

Relative proportion of SCAR industries vs. S&P 500

The SCAR's relative industry weightings is likely quite different from actual representation in the US private sector.

When comparing the relative percentages of SCAR organizations by industry compared to the sector's relative weight in the S&P 500 index, it is clear that the SCAR has a significant proportional overrepresentation of organizations from the Information and Financial industries. The SCAR is significantly relatively underweight in Communications, Consumer, and Other industry categories.

Chart a.3: SCAR industry relative percent vs. S&P 500 sector weighting*



Notably, the S&P 500 index is a subset of the US private sector, and accordingly, just one way to analyze SCAR industry firmographics. Further inquiry is important to better understand how SCAR industries and S&P 500 sectors represent actual private sector organizations with security teams and attack surface management technologies.

Table a.4: SCAR industry relative percent vs. S&P 500 sector weighting

Industry /sector	SCAR	S&P 500*	Difference
Communication Services	3.9%	7.3%	(-3.4%)
Consumer	3.9%	17.0%	(-13.1%)
Financial	18.6%	11.7%	6.9%
Health Care	17.1%	15.8%	1.3%
Industrials	9.3%	8.7%	0.6%
Information	41.1%	25.7%	15.4%
Other	6.2%	13.8%	(-7.6%)

**S&P 500 sector weighting data via S&P 500*

Appendix B: Methodology

Ultimately, the SCAR intends to be a rigorous resource by security practitioners and researchers, for other practitioners. This report strives for an exceptional commitment to academically rigorous research methods, a strong attention to detail, and unbiased presentation of research findings.

Approach to error and corrections

We acknowledge that the authors and reviewers of this report are human, and accordingly, imperfect. Despite significant review, it is entirely possible that mistakes could happen due to human error. Any errors discovered will be promptly corrected in the report, and noted in the Appendices in the future in a section labeled "Appendix D: Corrections."

We believe that peer review is crucial, which is why most SCAR analyses include a data table for the greatest possible transparency into data inputs, while still maintaining anonymity over the organizations that are part of this analysis. We encourage you to check our analyses or find your own stories within the data for your own research and reporting, while providing credit to JupiterOne as your data source.

Disclaimer

No claims are made that the research in this report is completely representative of all organizations worldwide. Claiming to have a perfect understanding of cyber assets at organizations of all sizes and industries would be unscientific and unreasonable. Instead, we simply claim to have analyzed a large data sample, and have provided as much visibility into firmographic factors as reasonably possible.

We believe that many of the findings are appropriate for general application to other organizations, especially cloud-native organizations, but we do not claim total global representation or a complete absence of bias. Readers of this report (and all other reports) are encouraged to be objective and critical of methodology applied.

While we have applied many possible controls against bias, such as rigorous review of this research by our peers and transparency into our data, bias nearly always exists and it's healthy to acknowledge this fact.

Appendix B: Methodology

Acknowledgment of selection bias

The data sample analyzed for SCAR is sourced from the organizations that use JupiterOne's Cyber Asset and Attack Surface Management (CAASM) product. Accordingly, selection bias is possible based on the organizations that find value in JupiterOne's product, which is discussed significantly in Appendix A: Firmographics in analyses of organization size, industry, and how these stack up to widely adopted classification methods and real-world weightings.

Our customers are generally cloud-native, which means the organizations analyzed for the SCAR could have a lower number of legacy systems and on-premises-based deployments than the true, global mean. This is due to the fact that CAASM products such as JupiterOne are generally designed for the asset and attack surface management requirements of cloud-native architectures.

Notably, the data in the SCAR sample could be further limited based on the customer's chosen integrations with JupiterOne's CAASM product since not all customers integrate all of their systems. Integrations typically include all IaaS and PaaS products, but not necessarily all SaaS, especially not SaaS that fall outside the administration of the security team (e.g., our customers may not always integrate their CRM or marketing SaaS tools with JupiterOne). Also, not all customers choose to create integrations between JupiterOne and homegrown systems, or integrations with legacy systems.

Some customer data on findings may be omitted as well, since JupiterOne offers customers the option to ingest findings for severity medium or higher. Other sources of findings, such as AWS Guard Duty or Inspector, are not similarly filtered. Future editions of the SCAR will attempt to explore the distribution of findings by severity, to better understand the proportion of critical, high, medium, low, and informational findings.

In short, the data reflected here is likely impacted to some extent by selection bias, and the result is likely a more cloud-native mixture of cyber assets than the reality of cyber assets at organizations worldwide. We hope to better understand, articulate, and control for possible selection bias in future iterations of the SCAR.

Appendix B: Methodology

The SCAR Process

The collection and conversion method consisted of the direct recording of cyber asset inventory data for customers of the JupiterOne CAASM solution. The majority of asset data was captured between September and December 2022, while the query data set represents a single point-in-time capture for a one-week period in the 4th quarter of 2022.

Significant and reasonable efforts were made to protect customer anonymity and avoid exposure of critical data to our data science and analyst teams, by ensuring that only sanitized data was included in the data used for analysis. No data that reveals customer sensitive information was included, with "critical" data made unavailable via access controls. JupiterOne's approach to classifying, managing, and protecting the confidentiality of customer data is described in the following documents:

- The JupiterOne Data Model
- JupiterOne Data Management Policy
- JupiterOne Data Protection

All contributors to the SCAR report were required to accept these policies. While IAM controls barred critical data from exposure to analysts, analysts were further instructed to be conservative with what data was used in this report. As such, we may have omitted the analysis of some data that probably could have yielded more interesting insights.

Source data was aggregated into a report by querying the data lake, and then transferred into a spreadsheet for greater consistency and control over granular analysis by asset class. Several analysts worked collaboratively to ensure consistency of analysis efforts, including clustering and grouping of sub-classes within each super class.

JupiterOne's data model allows a single cyber asset to fit multiple asset classes. Similarly, a single cyber asset may enter JupiterOne's graph multiple times due to redundancy within integrations. For example, a laptop may be reflected in both a device management integration and an endpoint detection and response integration. Care was taken to deduplicate all assets with multiple classifications and ensure they were only included once in the final analysis. Several data analysts and security practitioners worked collaboratively to review all groupings of assets with multiple classes, and categorize them into superclasses once according to best fit.

After the creation of data tables and basic graphs, the SCAR analysis was subject to interpretation by the authors, who shared a common lens as long-term security practitioners. The resulting graphs, tables, and written analysis were subjected to several weeks of rigorous peer review by analytics experts, data scientists, security practitioners, and engineers to ensure the data and interpretations were fair.

Appendix C: Acknowledgments

The authors would like to express their sincere gratitude to the many people who contributed to this research or assisted in the review process. This report is only possible with the efforts of many people and we are grateful for their support.

SCAR contributors

Significant contributors to this report include the individuals responsible for the design, editing, promotions, and ideation:

- Jasmine Henry, Lead Researcher and Author
- Scott McFarlane, Editor
- Sarah Hartland, Director of Demand Generation & Digital Marketing
- Corey Tomlinson, Senior Content Marketing Manager
- Valerie Zargarpur, Director of Global Events
- Tanvi Tapadia, Product Marketing Associate
- Benton Walker, Creative Technologist & Marketing Developer
- Chum Wongrassamee, Head of Design
- Dave Moy, Senior Designer
- Josephine Wong, Graphic Designer

SCAR reviewers

Also, we would like to acknowledge the following people for their review, feedback, and encouragement:

- Chasen Bettinger, Product Manager
- David Affonso, Account Executive
- Guillaume Ross, Deputy CISO
- Kenneth Kaye, Infrastructure Security Automation Engineer
- Sounil Yu, CISO
- Ashleigh Lee, Sr. Manager, Community and Customer Marketing

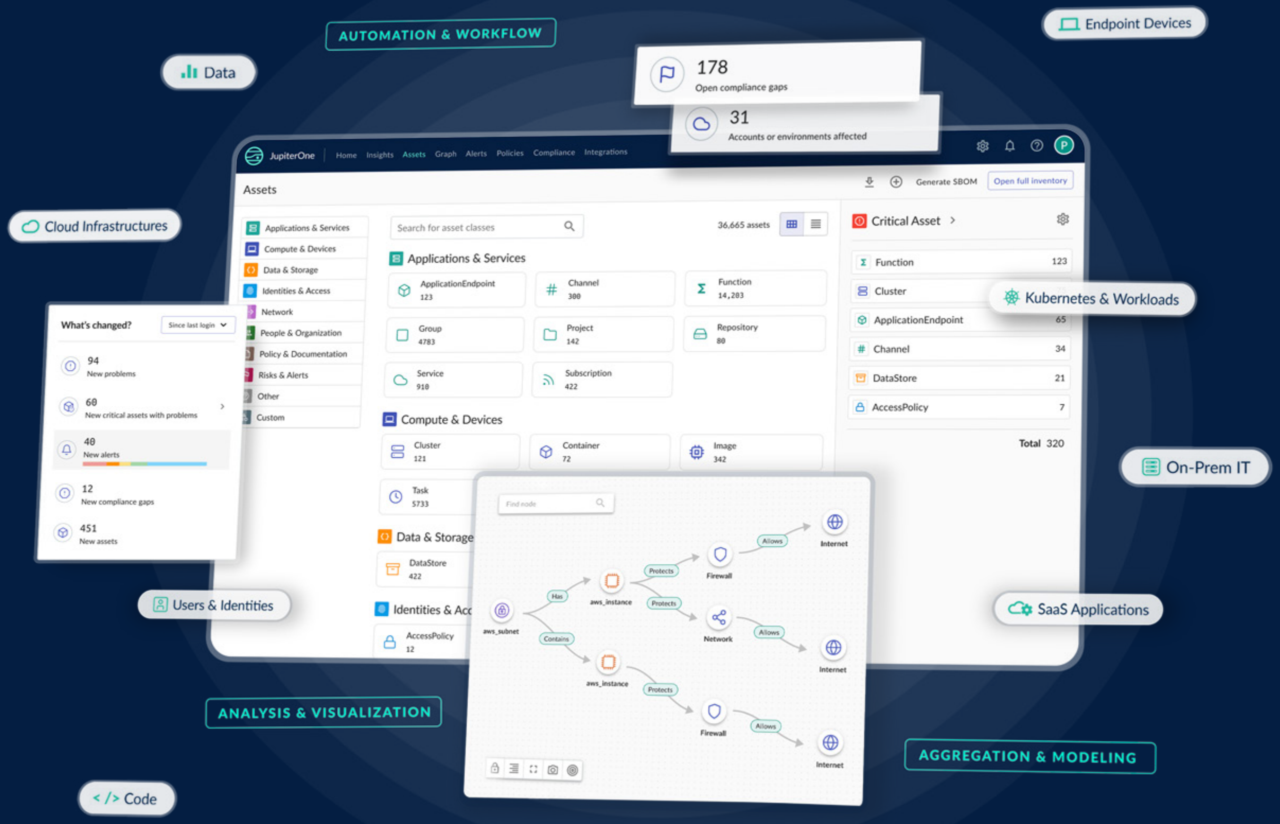
SCAR readers (that's you!)

This report highlights the incredible dedication of the worldwide community of security practitioners. We are grateful to be part of such an exceptional, dedicated community. Our goal is to help our security peers advocate for necessary resources and recognition, and move toward a healthier and more secure future for all of us.

Thank you for reading this report, citing it, and sharing your feedback on how we can improve future editions of this research. We are listening, receptive to your thoughts, and you are encouraged to reach us at research@jupiterone.com.

Cite the SCAR

You are permitted to use statistics, figures, and other information from this report, provided that you cite the source as JupiterOne 2023 State of Cyber Assets Report (or 2023 JupiterOne SCAR) and do not modify the content in any way. Exact quotes are permitted. If you would like to link to the report, we ask that you link to JupiterOne.com/SCAR.


AUTOMATION & WORKFLOW
Endpoint Devices
Data
178
Open compliance gaps
31
Accounts or environments affected
Cloud Infrastructures
JupiterOne
Home
Insights
Assets
Graph
Alerts
Policies
Compliance
Integrations
Assets
Generate SBOM
Open full inventory
Applications & Services
Compute & Devices
Data & Storage
Identities & Access
Network
People & Organization
Policy & Documentation
Risks & Alerts
Other
Custom
Search for asset classes
36,665 assets
Applications & Services
ApplicationEndpoint 123
Channel 300
Function 14,203
Group 4783
Project 142
Repository 80
Service 910
Subscription 422
Compute & Devices
Cluster 121
Container 72
Image 342
Task 5733
Data & Storage
DataStore 422
Identities & Acc
AccessPolicy 12
Critical Asset
Function 123
Cluster
ApplicationEndpoint 65
Channel 34
DataStore 21
AccessPolicy 7
Total 320
Kubernetes & Workloads
What's changed?
Since last login
94 New problems
60 New critical assets with problems
40 New alerts
12 New compliance gaps
451 New assets
On-Prem IT
Find node
Internet
Allows
Firewall
Protects
Has
aws_instance
aws_subnet
Contains
Network
Users & Identities
SaaS Applications
ANALYSIS & VISUALIZATION
AGGREGATION & MODELING
</> Code

JupiterOne

THE

State of Cyber Assets Report

2023

JupiterOne.com

JupiterOne Research