RAPID7

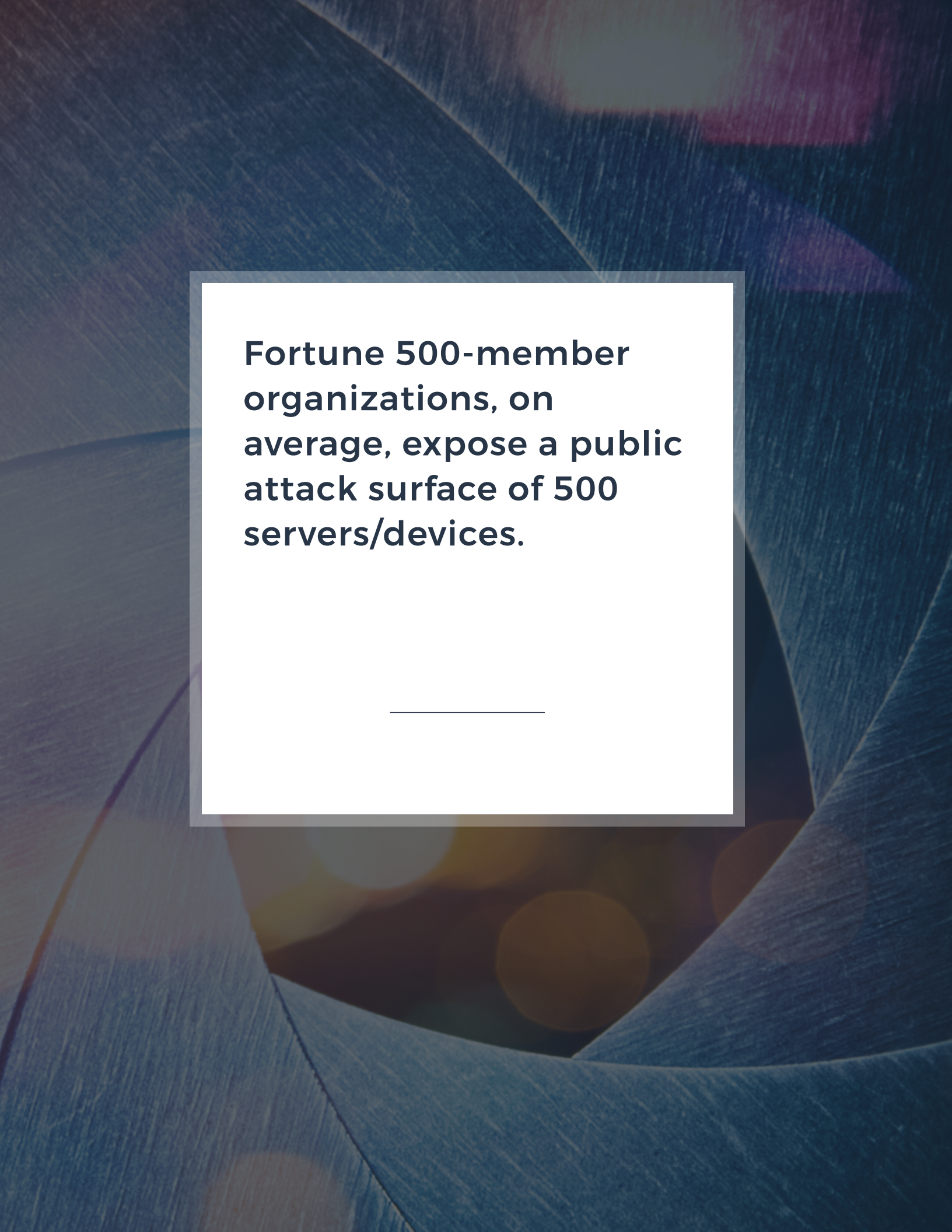
RAPID7 RESEARCH

INDUSTRY CYBER-EXPOSURE REPORT

FORTUNE 500

Rapid7 Labs | December 11, 2018

CONTENTS



Fortune 500-member organizations, on average, expose a public attack surface of 500 servers/devices.

EXECUTIVE SUMMARY

Measuring the cost and concentration of weaknesses in the public-facing configuration of internet-connected services and metadata—defined as "exposure"—is increasingly important in the face of growing cybersecurity threats. Having an accurate view of the resiliency of organizations and industry sectors to withstand malicious cyber-actions can facilitate more accurate cost models, help target efforts to reduce exposure to those that need it most, and enhance cooperative efforts between government and the private sector to better protect users and companies alike. Measurement of industry-level exposure can also inform industry-specific working groups that share cybersecurity information and threat intelligence, such as Information Sharing and Analysis Centers[1].

To understand current levels of exposure and resiliency, Rapid7 Labs measured 453[2] of the 2017 Fortune 500 List[3] for:

- Overall attack surface (the number of exposed servers/devices);
- Presence of dangerous or insecure services;
- Phishing defense posture;
- Evidence of system compromise;
- Weak public service and metadata configurations; and
- Joint third-party website dependency risks.

The Fortune 500 list is curated by a team of experts that use well-established criteria for selecting firms for inclusion. When revenues are combined, the composite list equates to approximately two-thirds of the U.S. GDP, with aggregate employment reaching nearly 29 million individuals globally. Furthermore, these organizations are incorporated in the United States, enabling the creation of a U.S.-centric view of exposure and the development of potential economic impact models as the result of malicious cyber-actions.

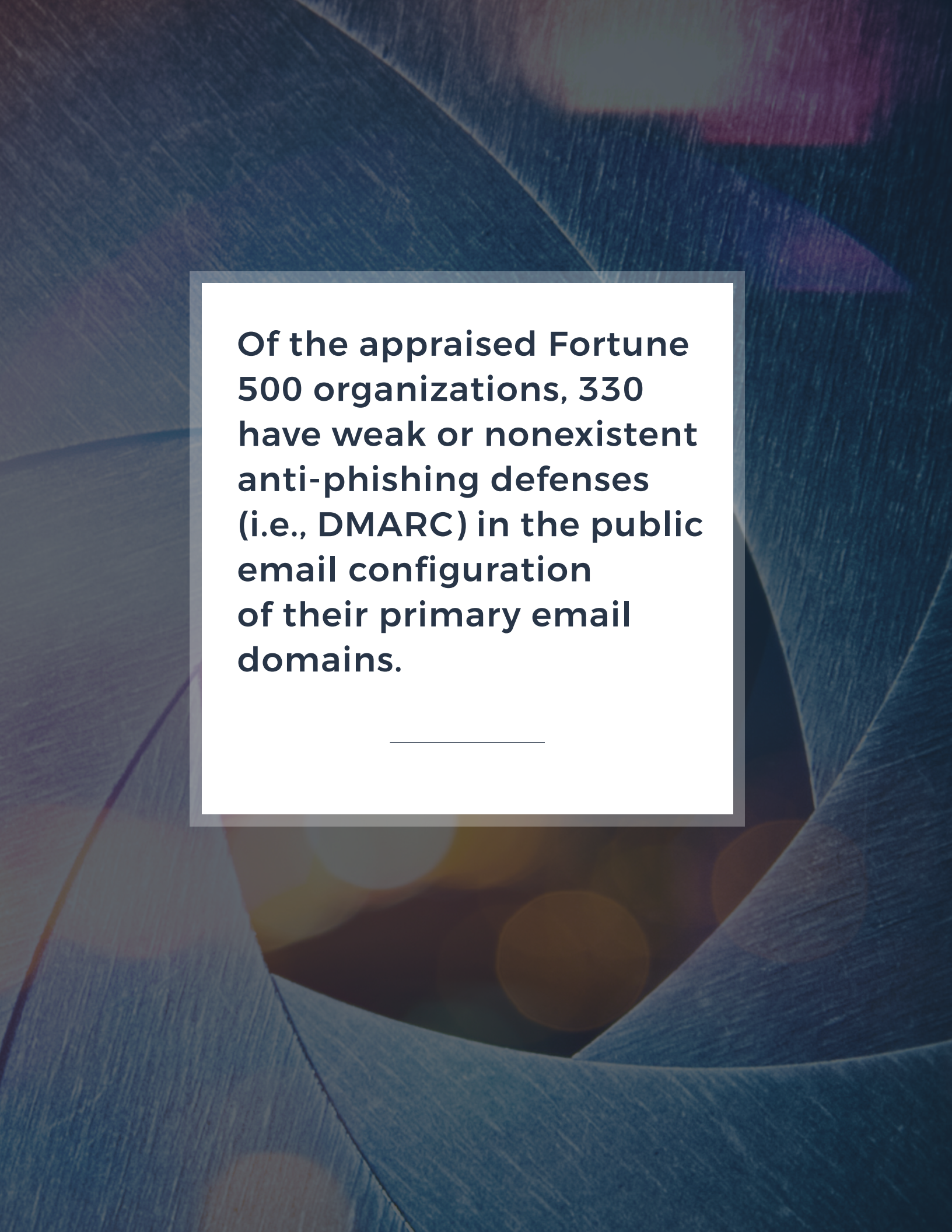
Key findings include:

- Fortune 500-member organizations, on average, expose a public attack surface of 500 systems/devices, with many companies exposing 2,500 or more systems/devices.
- Despite inherent weaknesses in Windows file-sharing and legacy Telnet servers, and known daily exploitation attempts against these vulnerable services, the average Fortune 500 organization exposes 5–10 of these services.
- Of the appraised Fortune 500 organizations, 330 have weak or nonexistent anti-phishing defenses (i.e., DMARC) in the public email configuration of their primary email domains.

[1] U.S. Dept. of Homeland Security, Information Sharing, Sep. 27, 2016, https://www.dhs.gov/topic/cybersecurity-information-sharing.

[2] This figure was chosen due to technical reasons described in Organization Internet Asset and Metadata Attribution Methodology at the end of this paper.

[3] Time, Inc., 2017 Fortune 500, http://fortune.com/fortune500/2017/ (last accessed Oct. 23, 2018).



Of the appraised Fortune 500 organizations, 330 have weak or nonexistent anti-phishing defenses (i.e., DMARC) in the public email configuration of their primary email domains.

- Every industry sector in the Fortune 500 signals how many and which cloud service providers they use in their public domain name system (DNS) metadata, with most organizations using between three and five cloud service providers, and some using 10 or more. This information can be used to craft highly effective, targeted attacks, among other actions.

- All industry sectors had members with observed malware compromises, with the Technology, Retailing, and Telecommunications sectors showing daily signs of ongoing compromise. These compromises ranged from company resources being co-opted into amplification denial-of-service attacks, to signs of EternalBlue-based campaigns similar to WannaCry and NotPetya.

The details behind these findings are presented in the remainder of the report. An important factor to consider in the context of these weaknesses is that members of the Fortune 500 list are well-resourced organizations that attract top talent in all aspects of the business, including information technology (IT) and security. The discovery of such widespread weaknesses in the exposed services of these leading organizations makes it likely that there is even greater exposure and risk in smaller organizations with fewer staff and financial resources available for securing their public internet resources. In general, companies can reduce the types of exposure identified in this report by leveraging their existing IT infrastructure, without purchasing new equipment or cybersecurity services, though reducing exposure will take time and resources.

MEASURING INDUSTRY EXPOSURE: OVERVIEW

This report documents findings regarding organizations' exposure to certain cybersecurity risks using data made available through interactions with public-facing systems over the internet. That data was then used to quantify the exposure of members of the U.S.-based Fortune 500, with results aggregated by industry sector. Measuring exposure at this level can help target cyber-risk reduction efforts, improve cybersecurity information-sharing within industry sectors, and build awareness of practices organizations can undertake to avoid future exposure.

Since 2016, Rapid7 has annually measured and reported on the exposure of specific countries to certain cybersecurity risks[4]. With this information, we engage country-level Computer Emergency Response Teams (CERT) to analyze the exposure in more detail and support action to reduce their overt exposure of critical services. To generate these reports, Rapid7 uses our internet-wide scanning platform, Project Sonar[5], and our passive sensor network, Project Heisenberg[6], to determine whether online assets are advertising vulnerable internet services or making suspicious outbound connections. We then aggregate the results at the nation-state level.

Aggregating the exposure data to the nation-state level is relatively straightforward. We use high-quality, regularly updated databases that match country location to internet addresses, with over 98% accuracy.[7] However, it takes additional effort to measure exposure at a deeper level. More robust exposure measurement of specific organizations is possible by analyzing the dedicated internet address space that those organizations own and use as part of their business operations. After matching organizations to internet addresses, exposure to certain cybersecurity risks can be quantified through publicly available data obtained with active scans and passive sensors. The Methodology section details the steps involved in:

- Attributing internet addresses and primary internet domain names to Fortune 500 organizations;
- Using Project Sonar's active scan data to identify exposure to vulnerable services and systems within the internet address space attributed to these organizations;
- Enhancing this exposure measurement by identifying the frequency and nature of interactions from this attributed internet address space with Rapid7's Project Heisenberg global passive sensor network; and
- Augmenting this direct exposure measurement with inferred exposure using "metadata" from the attributed internet address space, such as email "safety" configurations stored in internet DNS records, and detectable operating system and application version information from banners emitted by the discovered services.

The measurements can be broken down into three primary areas, each of which is covered in the following sections:

1. **Inferential measurements** using public DNS records, the most significant of which is the measurement of an organization's defenses against phishing attacks;
2. **Active measurements** using Rapid7's Project Sonar, which includes measuring both the presence of systems and services as well as the content those systems and services are exposing; and
3. **Passive measurements** using Rapid7's Project Heisenberg, which records when systems from an organization's network are contacting this honeypot collection and what actions they were trying to perform during these connections.

[4] Rapid7, National Exposure Index, Jun. 7, 2018, https://www.rapid7.com/info/national-exposure-index

[5] Rapid7, Project Sonar, https://www.rapid7.com/research/project-sonar

[6] Rapid7, Project Heisenberg, https://www.rapid7.com/research/project-heisenberg

[7] MaxMind, https://www.maxmind.com

MEASURING EXPOSURE:

INFERRING PHISHING DEFENSE CAPABILITIES

As noted in the **Methodology** section on domain name attribution (pg. 26), DNS records expose a means to identify how well an organization has configured its email service for protection from spam and phishing through the analysis of Domain-based Message Authentication, Reporting & Conformance (DMARC) records.[8] DMARC enables organizations to:

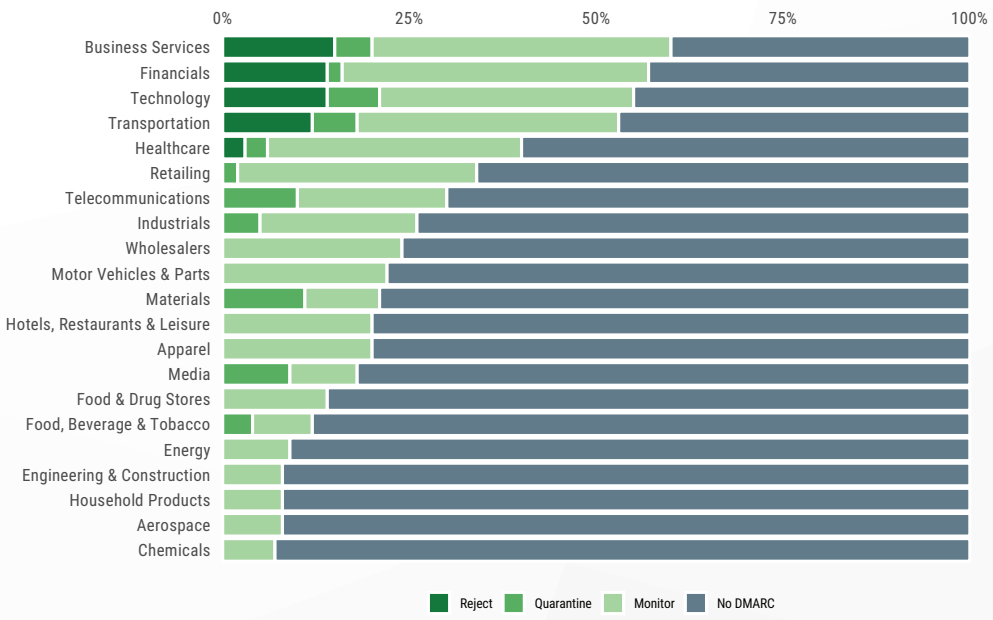
- Signal that they are using email authentication;
- Provide an email address to gather feedback about messages using their domain, legitimate or not; and
- Apply a policy to messages that fail authentication (one of "none", "report", "quarantine", or "reject").

No DMARC records—or a DMARC record of "none"—means there is no real first-line-of-defense protection from spam or phishing attacks. A properly configured DMARC record of "`monitor`" is a signal that an organization is on the path to email safety and is validating its DMARC configuration before enabling more active email defense measures.

Properly configured DMARC records with "`quarantine`" or "`reject`" have active email defense measures in place.

Figure 1 shows the percentage of DMARC adoption (by configuration category) of Fortune 500 organizations within a given sector. Green indicates that organizations within that sector are either on the path toward DMARC adoption or have fully adopted and implemented DMARC. Unfortunately, the results indicate that the majority (330) of the Fortune 500 have **not** embraced modern email safety configurations, boosting the organizations' risk of phishing attacks.

Figure 1: Email Safety Status of Fortune 500 Primary Email Domains



Since there is no direct scanning involved, DNS records are not impacted by the Project Sonar opt-out blacklist (described in the next section). Therefore, we can paint a more complete picture of the email safety configurations of the entire Fortune 500 than we can with active scanning for vulnerable services. The **Further Work** section (pg. 24) outlines additional steps that can be used to increase the scope of the examination to paint a wider picture of email safety.

[8] DMARC, https://dmarc.org (last accessed Oct. 23, 2018).

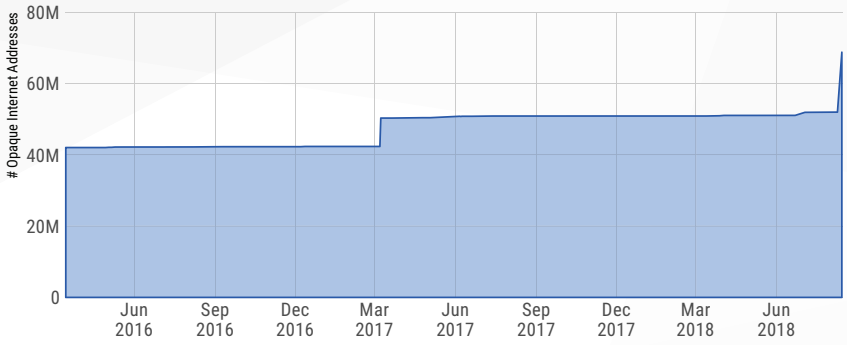
MEASURING EXPOSURE:

ACTIVE MEASUREMENTS WITH PROJECT SONAR

Project Sonar scans the internet across a wide array of services. A "service" could mean a web server, mail server, file server, database server, network equipment, or even cameras, along with many other types of servers that listen for requests over the internet. When a service on a given internet address responds positively to a probe, the positive result is recorded along with the response data. Depending on the service being scanned, this response data can include detailed version and configuration information of the scanned service.

Rapid7 adheres to the legal restrictions associated with internet scanning.[9] As a result, the probes performed by Project Sonar never involve the use of credentials, exploits for known vulnerabilities, or payloads that may cause harm to the service being probed, no matter how obvious or well-known those passwords or exploits may be.[10] While this places some limits on what we can scan and the types of service metadata that we can retrieve, there is a wide array of useful information that we can still capture.

Figure 2: Rapid7 Project Sonar Blacklist Growth



A further, self-imposed restriction comes as a result of Rapid7's "opt-out" process. Organizations may request that Rapid7 exempt specific internet address ranges from Project Sonar scans. Rapid7 obliges these requests and places the address range onto a blacklist that is restricted from the scanning process (Figure 2).

[9] Marcia Hofmann, Legal Considerations for Widespread Scanning, Rapid7, Oct. 30, 2013, https://blog.rapid7.com/2013/10/30/legal-considerations-for-widespread-scanning.

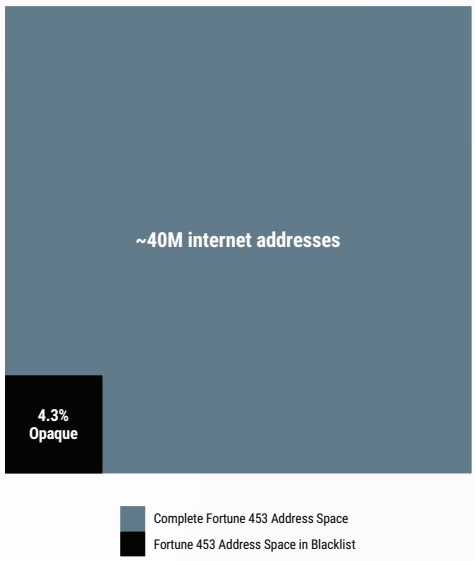
[10] Such as the very well-known passwords shipping with home routers. See Default Router Username and Password List, Router Configuration Center, https://192-168-1-1ip.mobi/default-router-passwords-list (last accessed Oct. 23, 2018).

Table 1: Number of Organizations per Sector on the Project Sonar Blacklist

INDUSTRY SECTOR	NUMBER OF ORGANIZATION WITH OPAQUE RANGES
Financials	6
Energy	4
Food & Drug Stores	2
Industrials	2
Business Services	1
Engineering & Construction	1
Retailing	1

This opt-out process had moderate impact on this exposure research. Figure 3 shows that just over 4% of possible addresses to scan for exposure are opaque to Project Sonar because of the blacklist, and Table 1 lists the impacted industry sectors. Improving the sample size to overcome this reduction in industry sector representation is addressed in the **Further Work** section (pg. 24).

Figure 3: Exposure Analysis Scan Surface Area



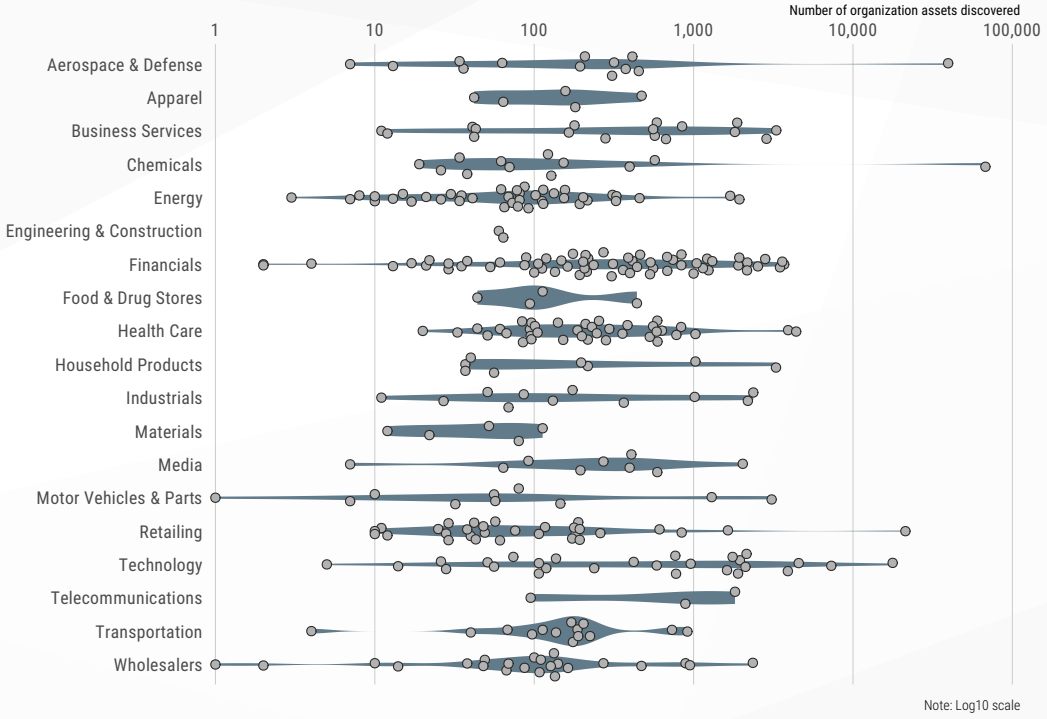
OVERVIEW OF RESULTS

Attack Surface by Sector

On average (the full distributions are below in Figure 4), organizations expose roughly 500 services that are detectable by Project Sonar. This number is neither "good" nor "bad," but each exposed node adds to the attack surface of an organization. To put it another way, each server or device has to be configured, managed, patched, and defended by an organization. There is no hard-and-fast general rule to say when this number tips the balance of risk for an organization, and how well an organization can protect these internet-exposed resources depends on a large number of factors. The fact remains that the more systems an organization exposes, the more opportunity attackers gain, regardless of the defensive capabilities of the organizations hosting these services.

It is also important to note that counts higher than 10,000 may be a sign that network equipment has been configured to respond on every port (a practice Rapid7 researchers have noted regularly in the aforementioned annual National Exposure Index reports[11]) or could be an indicator of the internet address space leasing noted above. The **Further Work** section (pg. 24) identifies potential ways of improving the accuracy of the collective list of internet address ranges.

Figure 4: Distribution of Discovered Organization Asset Totals by Sector
Each dot represents one organization; Position on axis = number assets discovered



Taking a look at Figure 4 above, we can see that industries that routinely offer higher rates of attack surface exposure include Business Services, Financials, and Technology, along with some outliers in the Aerospace, Chemicals, and Retailing industries. Organizations in these sectors may have solid asset inventories and legitimate reasons to have so many devices hanging off the internet. If your business processes do require this level of asset exposure, you must have commensurate vulnerability management, patching, and monitoring practices in place to facilitate a speedy response to discovered weaknesses or attempts by attackers to compromise your services. If your business processes are not the direct reason for this exposure and/or you do not have a well-oiled asset identification and configuration management process in place, working to reduce the surface area should be paramount, followed by plans to shore up those IT/security operational areas.

[11] Rapid7, National Exposure Index 2018, "Canary Ports," pg. 19, Jun. 7, 2018, https://www.rapid7.com/globalassets/_pdfs/research/rapid7-national-exposure-index-2018.pdf.

Critical Service Exposure: SMB and Telnet

The type of service being exposed has a direct impact on the severity of exposure (i.e., some services are less "safe" than others). One service in particular—Server Message Block (SMB)—is one of the most dangerous services for a system to expose. SMB is an all-in-one file-sharing and remote administration protocol, usually associated with Windows, that has been an attractive target for attackers and researchers alike for decades. MS03-049 in 2003, MS08-067 in 2008, and "EternalBlue" (MS17-010) in 2017 all arose from the complexity of the protocol and its central nature to Windows networking.[12] Recently, vulnerabilities in the SMB service were at the heart of the WannaCry and NotPetya attacks, which crippled networks and caused significant outages to critical business processes that cost many companies millions of dollars in lost revenue.[13]

Figure 5: Distribution of Organization Assets Exposing SMB
Each dot represents one organization; Position on axis = number assets discovered

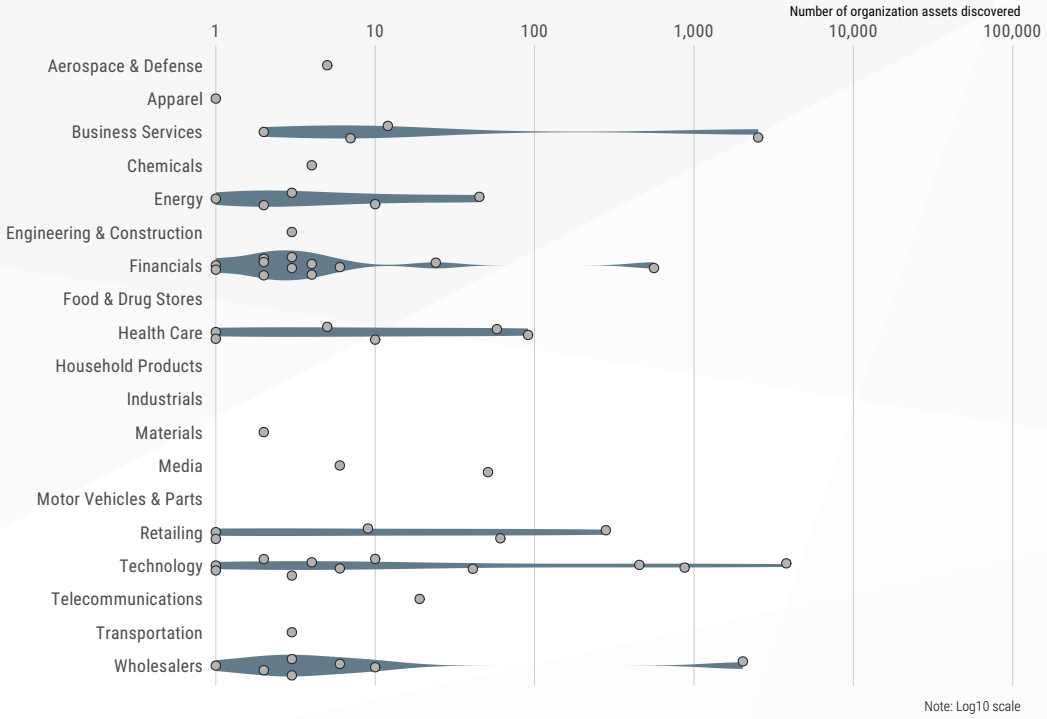


Figure 5 shows that there is still internet-based exposure to these types of attacks within the Fortune 500, with 15 out of 21 sectors having at least one member exposing SMB and an average of 10 exposed SMB nodes per organization.

While the exposure is not present throughout all the organizations in the study, there is no safe way to expose SMB services to the public internet. In light of this, Microsoft has made efforts to reduce SMB exposure for normal desktop and laptop clients; for example, current Microsoft documentation explicitly recommends blocking SMB on an internet perimeter firewall, and Windows 10 desktops automatically firewall access to port 445 by default.[14] Even exposing one asset with SMB running could end up [re-]spreading WannaCry, NotPetya, or modern variants across an entire organization.

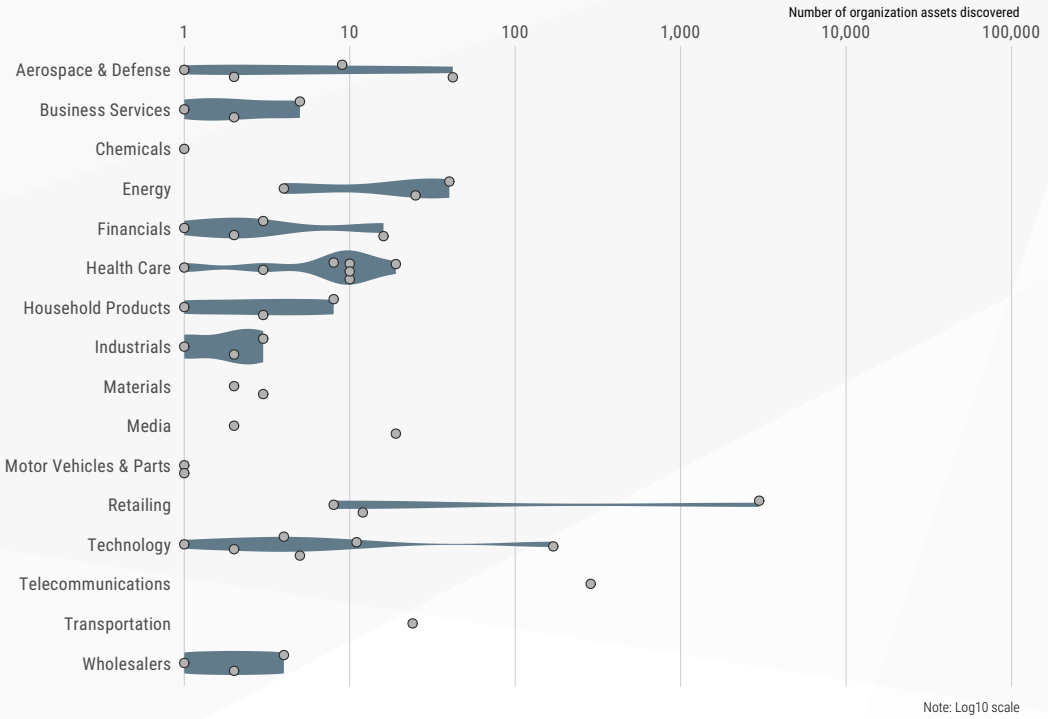
[12] Rapid7, National Exposure Index 2018, "Inappropriate Services," pg. 14, Jun. 7, 2018, https://www.rapid7.com/globalassets/_pdfs/research/rapid7-national-exposure-index-2018.pdf

[13] Bob Rudis, No More Tears? WannaCry, One Year Later, Rapid7, May 14, 2018, https://blog.rapid7.com/2018/05/14/no-more-tears-wannacry.

[14] Microsoft, Guidelines for blocking specific firewall ports to prevent SMB traffic from leaving the corporate environment, Aug. 31, 2016, https://support.microsoft.com/en-us/help/3185535/guidelines-for-blocking-specific-firewall-ports-to-prevent-smb-traffic.

Figure 6 shows the distribution of assets per organization that expose the Telnet service. Telnet exposure creates risks similar to SMB exposure. Telnet dates back to the early days of the internet, with the official "modern" standard dating back to 1983.[15] Telnet is a cleartext protocol that is used to directly log in to servers and network equipment, usually to issue commands and run scripts directly at the operating system level of the device. Telnet services have a history of vulnerabilities and exposures that put the organization at risk of credential theft, passive and active eavesdropping, and remote code execution. The cleartext nature of the protocol means that an attacker in the proper network position can read any usernames, passwords, or data being transmitted, and endpoints with weak, default, or eavesdropped passwords can be hijacked to run malicious code directly by the operating system.

Figure 6: Distribution Density of Organization Assets Exposing Telnet
Each dot represents one organization; Position on axis = number assets discovered



The nature of Telnet usage in the Fortune 500 is far from uniform. In total, 48 organizations expose Telnet and, of those, the average number of Telnet endpoints exposed is four. What they expose varies from network equipment administrative access, to direct server access to point-of-sale system access, the latter being mostly in Retailing.

There is no technical or practical justification for running a Telnet service today. It has been superseded by the Secure Shell (SSH) Transport Layer Protocol, which provides encryption-in-transport and encourages the use of digital certificates when authenticating connections.[16] If a device is truly incapable of running SSH rather than Telnet due to a lack of local computing resources, that device is simply too insecure by design to expose to the public internet, regardless of the reasoning for staying with a 40-year-old unencryptable protocol.

[15] J. Postel and J. Reynolds, Telnet Protocol Specification, Internet Engineering Task Force, May 1983, https://tools.ietf.org/html/rfc854.

[16] T. Ylonen and C. Lonvick, The Secure Shell (SSH) Transport Layer Protocol, The Internet Society, Jan. 2006, https://tools.ietf.org/html/rfc4253.

Third-Party Risk Exposure

When an organization uses third-party resources to supplement its online assets, it takes on risks associated with those third-party resources. Vulnerable third-party resources can be used as a conduit to attack the first-party organization. For example, in September 2018, security researchers noted that many sites are vulnerable to web-based credit card-skimming attacks due to their reliance on third-party content delivery networks (CDNs).[17] In another example, the Syrian Electronic Army used a compromised CDN in 2015 to take over a major news media outlet's web presence and send custom push notifications to readers.[18]

For the purposes of this study, "third-party risk" exposure is defined as being present either when:

- A measured organization is seen to be relying on resources from a third-party site when building their own websites and applications; or
- A measured organization exposes what third-party services they are actively using by leaving potentially sensitive artifacts in their published metadata.

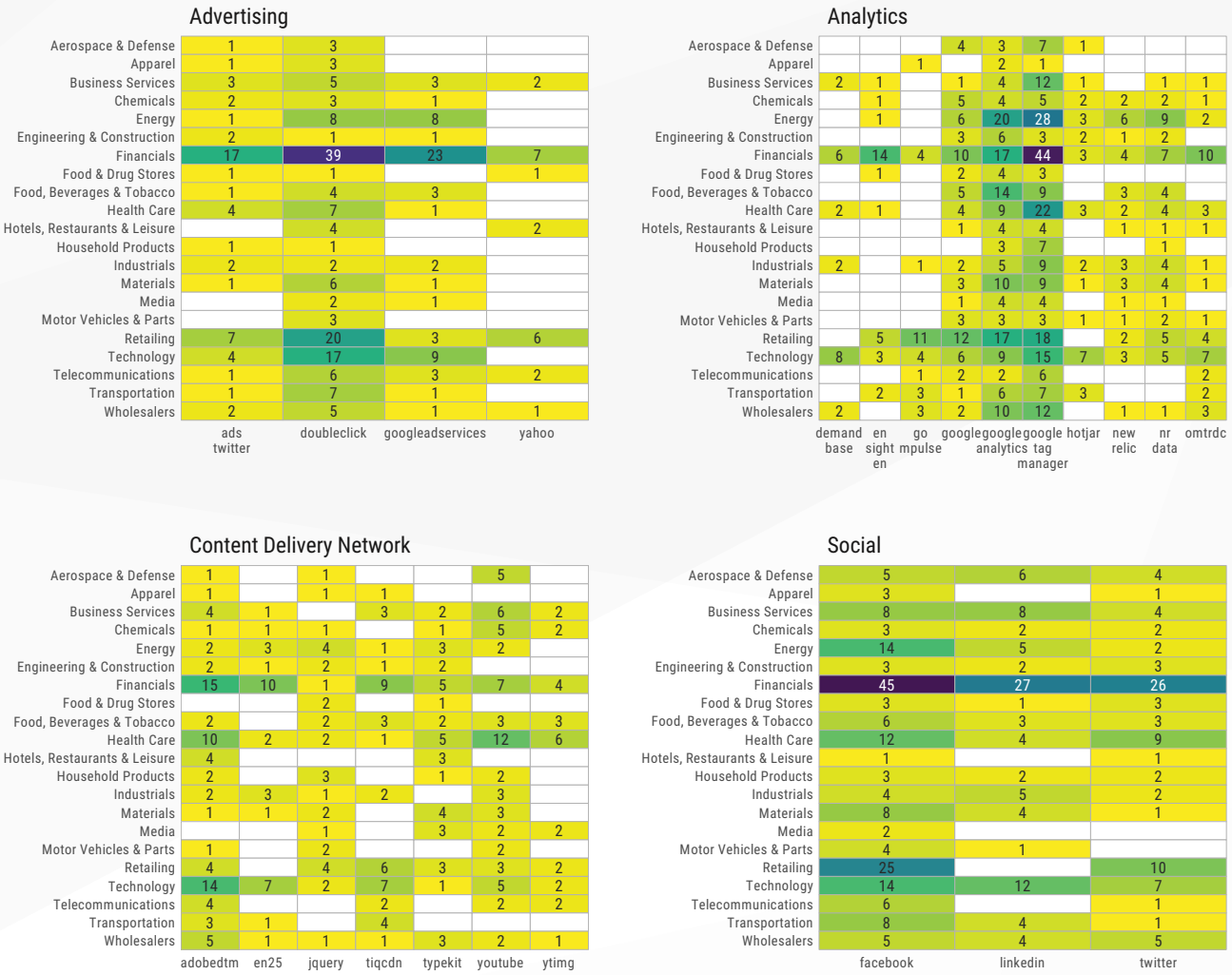
To get an idea of third-party risk when exposing web servers/service to the internet, we can examine the resources each web page loads when the page is loaded into a web browser. Project Sonar can perform this task at scale by controlling a virtual web browser, visiting the pages of the well-known domains of the organizations in the study, and capturing all the activity as each site loads resources.

---

[17] Kevin Beaumont, Magecart—new tactics leading to massive unreported fraud, DoublePulsar, Sep. 19, 2018, https://doublepulsar.com/magecart-new-tactics-leading-to-massive-unreported-fraud-5211c9883dea.

[18] Thu Pham, Malicious Hackers Take Over Media Sites via Content Delivery Network Providers, Duo Security, May 19, 2015, https://duo.com/blog/malicious-hackers-take-over-media-sites-via-content-delivery-providers.

Figure 7: Third-Party JavaScript Execution by Sector and Service Type



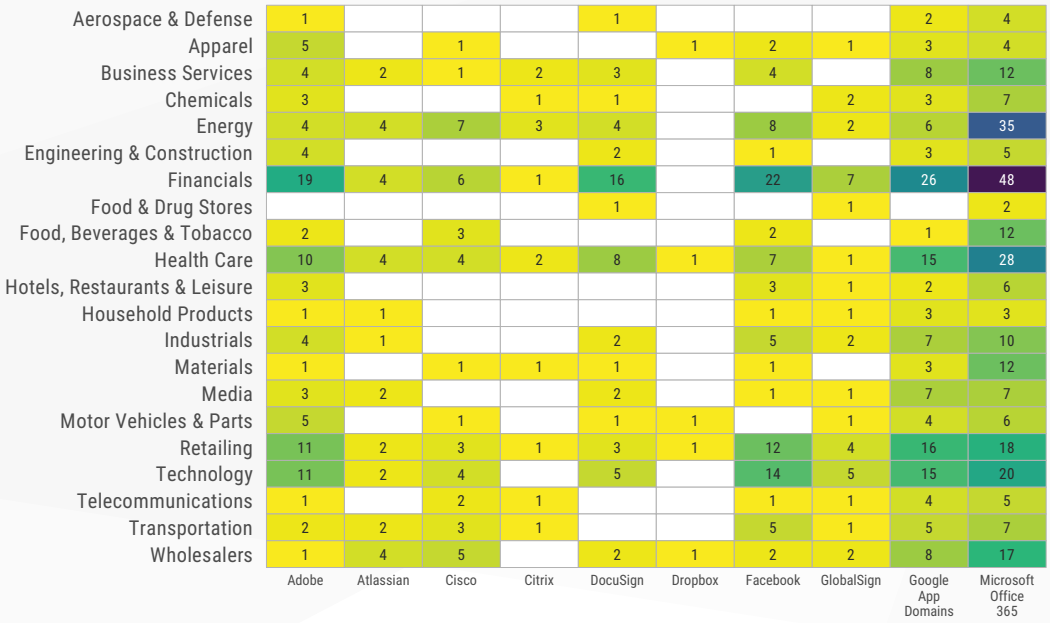
These websites load a great quantity of third-party resources, so the complete list would be difficult to visualize and comprehend. The resultant list was pared down to only the most prevalent third-party resources used across the target study list. Figure 7 shows the breakdown by general categories of resource: advertising-oriented, site analytics, sourced from a CDN, or incorporating code from social network sites such as Facebook and Twitter.

Some of these third-party services are likely resilient to cyberattacks and do not meaningfully contribute to the first-party organization's degree of exposure. For example, it is unlikely that Google would be sufficiently breached as to be an unwitting conduit for malicious activity to client organizations. However, other names in the chart might not be as resilient, and the third-party resources left off the chart are definitely not in the same class as some of these more recognizable ones.

Figure 8 focuses attention on the latter component of third-party exposure: detecting use of vendor application/cloud service.

In addition to providing the connection address for names such as www.rapid7.com, DNS records can identify secure email configurations (as detailed in Measuring Email "Safety" below). DNS records can also reveal which third-party providers an organization uses for everything from application development to cloud hosting environments to file-sharing and more.

Figure 8: Third-Party App/Cloud Usage Exposure via DNS Metadata



One way these services are exposed is through the use of verification information stored in free-form TXT records. To illustrate, Table 2 shows a sample of DNS TXT records for rapid7.com:

Table 2: Rapid7 DNS TXT Records Sample

DNS RECORD KEY	DNS TXT RECORD VALUE
rapid7.com.	smartsheet-site-validation.rapid7.com=wfJFw8OnJ0WwBCBDP7NuqH
rapid7.com.	MS=ms93061892
rapid7.com.	atlassian-domain-verification=+Mx+hFjC77glTvA7K9Tp/5x7LvbyawRYOeZpkXhE/Xys/xciI66aaIgyQQAD88E7
rapid7.com.	citrix-verification-code=3d0b3642-a1b3-4cf3-8616-c9fb8cd0c2da

- **"smartsheet-site-validation"** signals that Rapid7 uses SmartSheet, a cloud spreadsheet service.
- **"atlassian-domain-verification"** signals that Rapid7 uses cloud-based services by Atlassian, a provider of popular software development tools and platforms.
- **"citrix-verification-code"** signals that Rapid7 uses services offered by Citrix.

This may not seem like a major risk when looked at individually. In truth, many of these "validation" records are only required once and can be removed after the initial validation has occurred. These records prove that one is the true owner of a given domain since in theory, only the true owner can add, modify, or delete DNS entries.

If one were to look at those records in aggregate, it might be possible to find a common, shared third-party service in use by a large number of organizations or a boutique service provider used by only a handful of organizations. These may be high-value targets for malicious actors that seek to compromise multiple organizations, making resiliency of these third-party services all the more important..

Rapid7 researchers used Project Sonar DNS collection data to examine the TXT records of the Fortune 453 organizations in this study. Only well-known domain names were used (expanding on this effort to use additional domains is covered in the Further Work section), and Figure 8 only focuses on the most prevalent or well-known third-party services.

It may come as no surprise that virtually every industry sector uses Microsoft Office 365, and it is highly unlikely that Microsoft is going to fall prey to an attack that would enable Office 365 to be a malicious gateway into organizations. However, many other exposed (and likely less resilient) third-party providers are shared across the industry sectors, especially the Financials sector, boosting the risk that a capable attacker can use a third party to gain access to other organizations—especially when they can make a list of these common resources simply by making a DNS query.

MEASURING EXPOSURE:

CONDUCTING PASSIVE MEASUREMENTS WITH PROJECT HEISENBERG

Rapid7's Project Heisenberg is, at heart, several dozen unadvertised systems hosting a variety of fake services, such as HTTP, SMB, SSH, and many others. These honeypots are closely monitored for unsolicited connections but do nothing to attract or entice those connections. Other than internet-wide scanning research, there are no legitimate reasons for an organization to connect with the Heisenberg sensor network, so any recorded activity in Heisenberg is a high-quality indicator that an organization does not have control of its outbound connections—which further suggests either malicious activity or misconfigured service traffic coming from the organization. In essence, if there is any contact with Heisenberg by an organization, there is some type of exposure occurring in that organization. The passive connection and activity recording of Project Heisenberg is not impacted by the Project Sonar opt-out blacklist. Rapid7 does not proactively block connections originating from blacklisted IP address ranges—we merely skip actively scanning those ranges, so the entire Fortune 500 is considered in our passive measurements, not just the subset that haven't blacklisted Sonar scans.

Figure 9: Daily Time-Series of Unique Connections to Heisenberg by Industry Sector

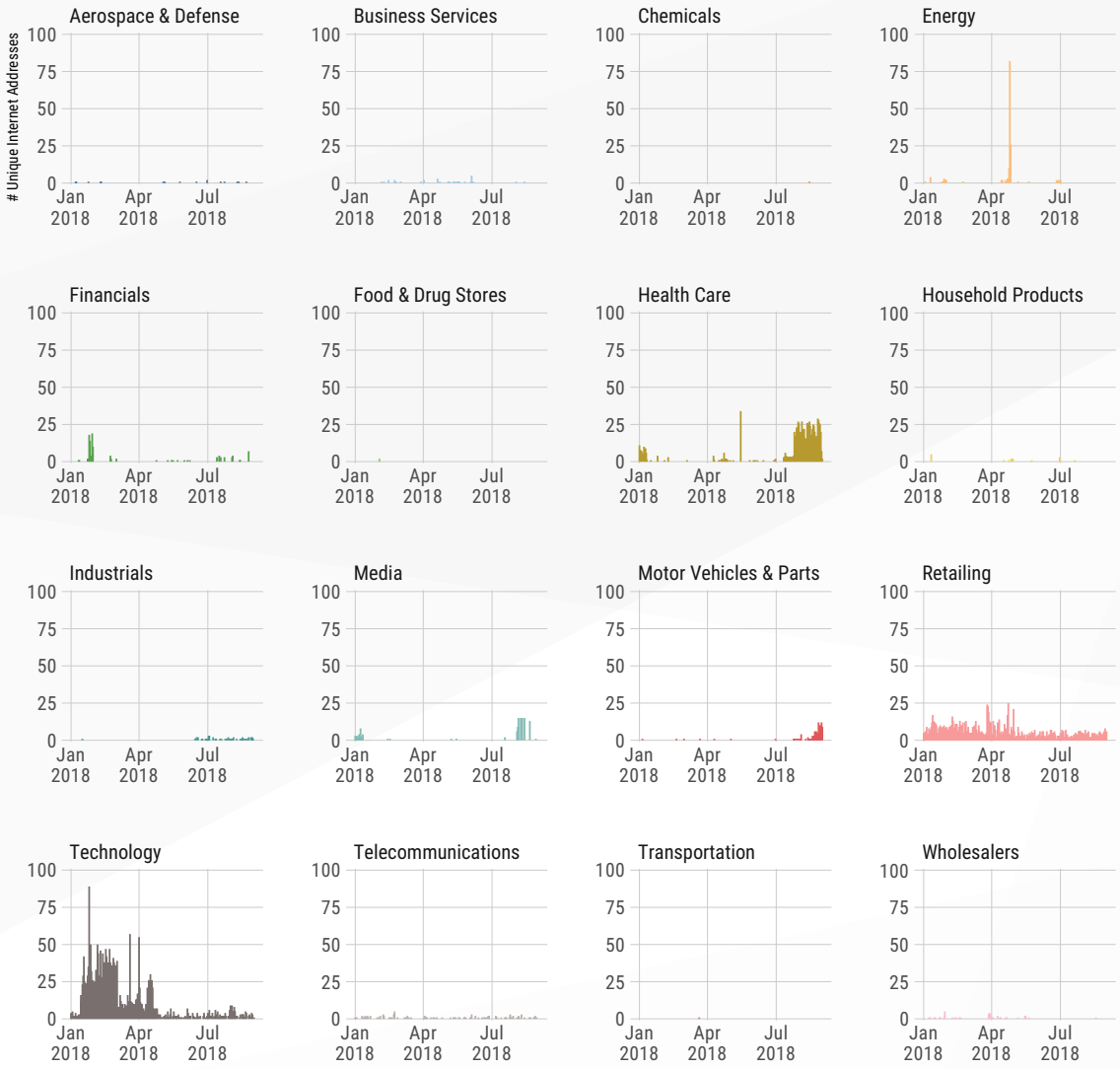


Figure 9 shows the unique, daily connections with the Heisenberg sensor network for all organizations in a given sector. Ideally, this chart should be blank. However, the chart shows lapses in control across every sector in this data set. Some sectors, such as Health Care, Retailing, and Technology, appear to have slightly higher systemic rates of control failures, but this view does not tell the whole story, since many modern networks sit behind a single internet address through which hundreds to thousands of employees, contractors, and devices communicate. This chart is handy to show presence, but we need another view to show volume.

Figure 10: Daily Time-Series of Total Connections to Heisenberg by Industry Sector

NOTE: Free Y Scales

In contrast to the unique connection footprint view, Figure 10 shows the total daily connections to Project Heisenberg across organizations in the measured industry sectors. **Note that the Y-axis is not uniform across the panels**. This free scale lets us "zoom in" on each industry and and more easily distinguish potential patterns and problems.

We see that just because an industry has a small number of unique nodes connecting to Heisenberg sensors does not mean they are inactive. Larger volumes in this view could indicate a mass malware infection internal to an organization (i.e., dozens, hundreds, or thousands of infected systems reaching out to the internet) or may be indicative of a few systems being co-opted

into denial-of-service campaigns.

To further compare industries, we can combine the data from the previous two charts to complete the macro exposure picture. Figure 11 shows the distribution of connection ratios (total connections in a day / unique sources in a day) by industry sector. Sectors with a greater number of points have organizations with more frequent gaps in configuration control or malware containment. Sectors with points further out on the axis also have gaps in monitoring as well as containment. Health Care, Retailing, and Technology organizations in this data sample appear to be lagging behind their peers in other sectors.

Some connections are more serious than others, and four of the top connection types to Heisenberg from organizations in this study are especially bad. As Figure 12 shows, throughout the first half of 2018, Heisenberg recorded daily connections that indicate multiple organizations were impacted by:

- Malware associated with SMB (i.e., WannaCry/ WannaMine/NotPetya);
- DNS denial-of-service attacks;
- Telnet cleartext credential brute-forcing; and
- SSH encrypted credential brute-forcing.

Figure 11: 2018H1 + Aug 2018 Heisenberg Connection Ratios by Industry Sector

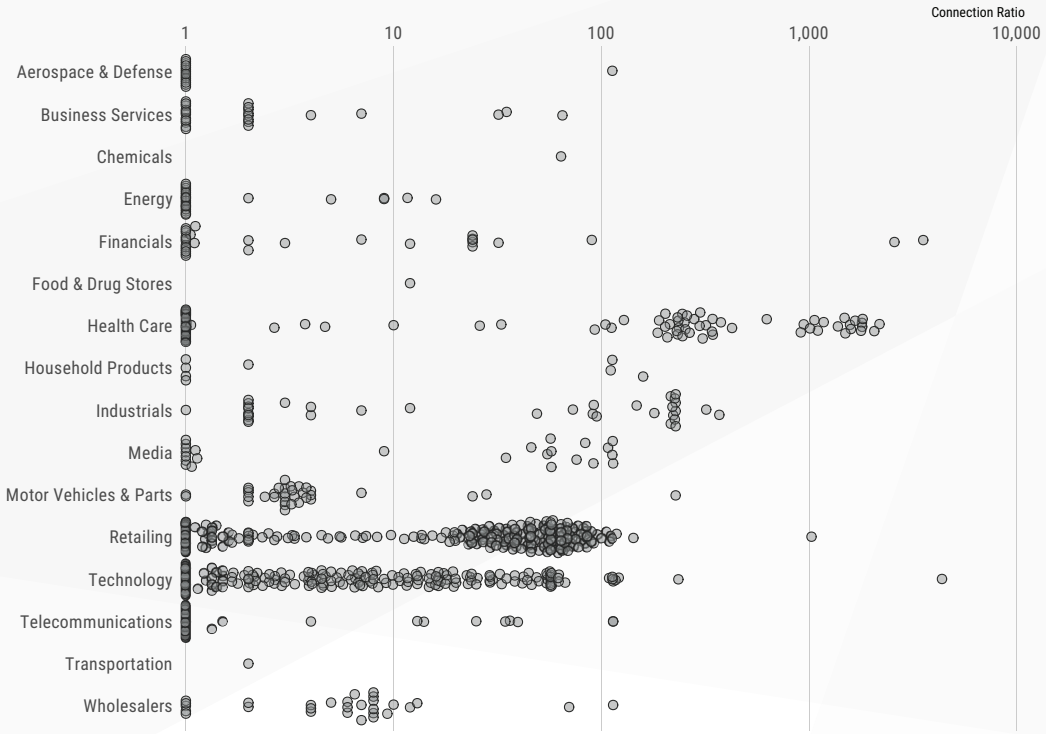
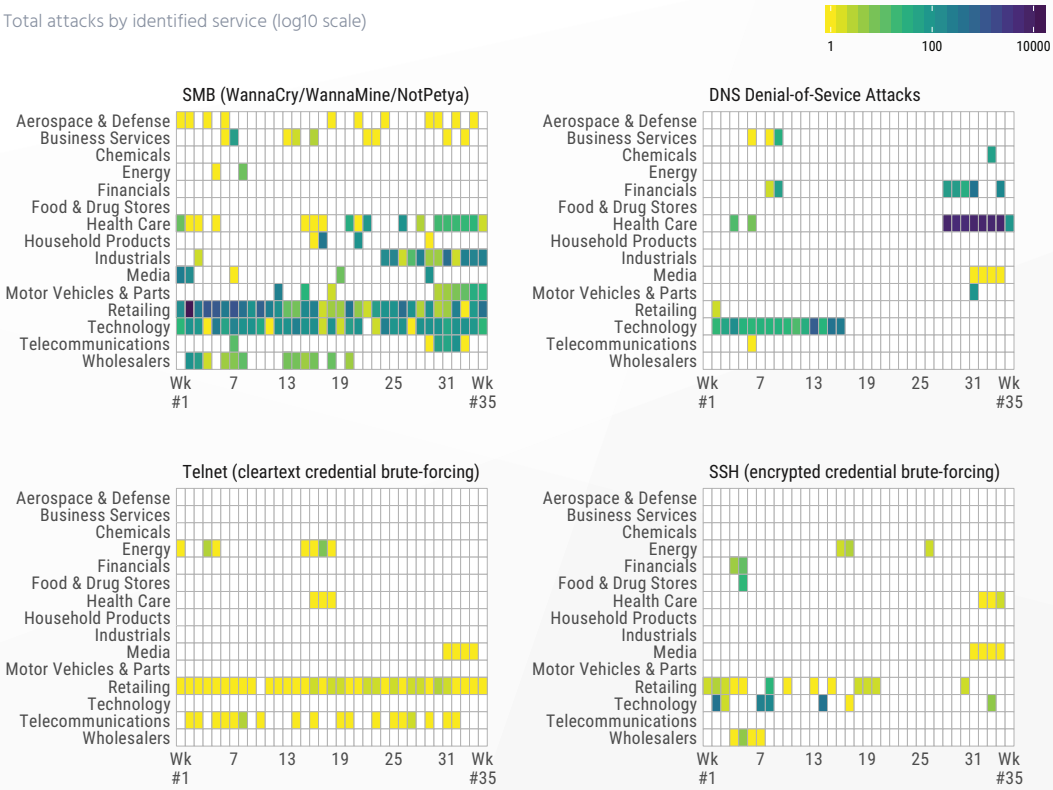


Figure 12: Signs of Serious Malicious Activity per Sector

Total attacks by identified service (log10 scale)



CONCLUSION

The methodology outlined in this report describes several ways, based on openly available internet connections, to measure the exposure of specific organizations and industry sectors to certain cybersecurity risks. While far from a complete picture of the organizations' overall cybersecurity posture, the results of this research indicate significant levels of exposure among Fortune 500 companies:

- The majority (330) of Fortune 500 organizations do not use enhanced email safety configurations, creating greater risk of phishing attacks. To reduce exposure, organizations should evaluate and strengthen their DMARC configuration settings.
- Companies in the Fortune 500 expose an average of 500 internet-accessible services each. Sixteen companies expose 5,000 or more such services, dramatically increasing their effective attack surface area from internet-based threats compared to the average Fortune 500 enterprise.
- Fifty-nine organizations expose an average of four SMB services, each creating a greater risk of susceptibility to exploitation of SMB vulnerabilities. To reduce exposure, organizations should close port 445 whenever possible.
- Forty-nine organizations expose an average of four Telnet services, each creating a greater risk of credential theft and eavesdropping as Telnet transmits information in cleartext. To reduce exposure, organizations should consider switching from Telnet to Secure Shell protocol whenever possible.
- Dozens of exposed third-party services are shared among the Fortune 500 organizations, creating a greater risk that a vulnerability in a shared third party can lead to compromise of multiple organizations. Organizations should ensure their third-party service providers are taking appropriate steps to strengthen their own security, as well as use tools such as subresource integrity signatures when sourcing these services to help reduce the likelihood of shared compromise.

Because the Fortune 500 organizations have disproportionately strong access to resources and technical expertise, the findings suggest that the severity of exposure may be greater for organizations outside the Fortune 500. An ongoing conversation with key stakeholders on the reasons for this continuing exposure, and steps to take to mitigate the cybersecurity risks posed by the exposure, may be broadly beneficial to the digital ecosystem.

FURTHER WORK

The processes and procedures used for the exposure analyses in this report are the initial steps at communicating the overall "cyber-health" of industries based on a subset of possible internet telemetry measurements. In each section, possible measurement deficiencies have been identified and will be addressed.

Improving Entity Internet Asset Attribution

The most common Internet Protocol (IP) address space (version 4, IPv4) is fully exhausted, meaning there are no "spare" blocks of IP address to assign an entity. However, organizations that currently do own IPv4 address space are not utilizing said space to capacity. The scarcity of this finite resource has resulted in the creation of a marketplace in which IPv4 space can be bought and sold.[19] While some long-standing organizations have sold portions of their IPv4 address space blocks to other parties, some retain ownership and manage the leasing of this space to others on their own. This practice results in attribution errors, which are especially vexing when corporate address space is leased in a non-attributable way to third-party hosting providers and/or cloud providers.

For this report, Rapid7 researchers manually identified attribution anomalies by comparing address space utilization and service composition with that of known hosting and cloud service providers. Further work will be performed to automate this classification, which will enable filtering out hosting and cloud service provider blocks at scale.

The initial attributed list of IPv4 address blocks to Fortune 500 organizations was performed by traditional and machine learning-based approaches to entity name mapping to IP WHOIS data. While these matches worked well, the process failed to attribute all known IPv4 space to an organization that has potentially merged with or acquired other organizations, since maintenance of WHOIS records is not often a priority when it comes to mergers and acquisitions (M&A) activities. Incorporating historical M&A activity records within the entity-matching processes will further enhance the scope and completeness of the attribution process.

Furthermore, only attributed IPv4 address blocks were used for the exposure analyses in this report. Since many organizations use third-party hosting providers and cloud services, applications, devices, and services hosted by those providers were not included in the analyses. Rapid7 has access to billions of DNS addresses, and organizations must use those addresses to claim temporary ownership of IPv4 space in hosting and cloud service providers. Further work will be done to include these separate DNS-identified resources in the point-in-time exposure analyses.

Avoiding Opt-Out Opacity

Research work like this paper depends on continuous, light-touch scanning like the kind provided by Rapid7's Project Sonar, so if enough organizations decide to opt-out of these scans, the internet research community will undoubtedly suffer. There are two future paths that can reduce the impact of the Project Sonar "opt-out" list opacity issue. As a responsible internet citizen, Rapid7 keeps the opt-out list process in place, but it may be possible to augment current processes and have the opt-out be an annual process whereby organizations may re-acknowledge their desire to have their IPv4 space remain on the opt-out list. This would provide an opportunity to restate the advantages of allowing Project Sonar scans and reduce the size of the opt-out list and preserve the statistical integrity of the surveys.

[19] IPv4 Brokers, ARIN IPv4 Market Prices & Transfer Statistics, https://ipv4brokers.net/arin-ipv4-prices-transfer-statistics/ (last accessed Oct. 23, 2018).

The second path is to just expand the sample size. There are other, notable organization lists—e.g., Inc. 5000, S&P 500, FTSE 100—that can significantly expand the sample sizes in each industry and reduce the size of the opaque regions to (perhaps) less than 1%. The previously noted attribution accuracy and expansion enhancements are key components to ensuring the validity and efficacy of this expansion process.

Finding More DNS Records

The initial exposure and email safety analyses utilized the well-known DNS domains of the organizations on the Fortune 500 list. Most of those corporations have many subsidiaries and brands, each with their own set of DNS domains. Further work will be performed to develop a machine learning-based web-crawling and data-mining process to identify these additional domains and incorporate the data associated with them into the analysis framework used in this report.

Expanding Resource Safety Evaluation

The further work to discover additional domain names will have a direct impact on the email safety analyses used for this report. Furthermore, this report only looked at one aspect of email safety (DMARC). There are additional resource records that describe other email safety configurations, such as Sender Policy Framework (SPF), which further helps reduce spam and prevents misuse of email domains by attackers. This will be included in future analyses.

Other types of DNS records (i.e., non-email-related ones) also communicate other types of both exposure and safety, and that information will also be explored for inclusion in future analyses.

Third-Party Dependency/Risk Analyses

Finally, by analyzing the overall configuration of an organization's DNS records, discovering how an organization's IPv4 networks are routed on the internet, enumerating which third-party resources an organization relies upon in its web and web application services and other indirect, public measurements, it is possible to report on both the potential fragility of an organization's overall internet presence and provide exposure views of third-party dependencies across all organizations.

STUDY METHODOLOGY

Why the Fortune 500

Aggregating exposure for specific U.S. industry sectors poses a unique problem. First, IP address space is fairly expansive. IPv4 alone supports over 4.2 billion addresses (a portion of which are not assignable), without taking into consideration the exponentially more massive IPv6 space. These addresses are assigned to various governments, companies, and service providers around the world. Second, with the onset of dynamic infrastructure ("the cloud"), it is increasingly common for companies to lease IP address space from other companies to host their services. This makes traditional methods of attributing IP addresses to particular organizations (such as by using the WHOIS lookup tool) incomplete, since the owner of the IP address may not be the owner of the service evaluated for exposure.[20]

Instead of attributing IP addresses to companies and filtering by U.S. industries, we focus on the 2017 Fortune 500 as a representative sample, from which we attribute and filter global IP address space and services hosted on dynamic infrastructure.

The 2017 Fortune 500 List was chosen for many reasons. First, it is a diverse (see Table 3) list curated by a team of experts that use well-established criteria for selecting firms for inclusion.[21] When revenues are combined, the composite list equates to approximately two-thirds of the U.S. GDP, with aggregate employment reaching nearly 29 million individuals globally. Furthermore, these organizations are incorporated in the United States, enabling the creation of a U.S.-centric view of exposure and the development of potential economic impact models.

Table 3: Fortune 500 Sector Counts

INDUSTRY SECTOR	NUMBER OF ORGANIZATIONS
Financials	88
Energy	59
Retailing	46
Health Care	40
Technology	39
Wholesalers	26
Food, Beverages & Tobacco	24
Business Services	20
Materials	19
Industrials	18
Transportation	18
Chemicals	14
Aerospace & Defense	13
Engineering & Construction	12
Household Products	12
Hotels, Restaurants & Leisure	11
Media	11
Motor Vehicles & Parts	11
Telecommunications	8
Food & Drug Stores	6
Apparel	5

A large number of these organizations have been incorporated for over 20 years and were early adopters of internet technologies. As such, the vast majority of them do own and manage significant portions of internet address space that facilitates attribution and measurement.

[20] ICANN WHOIS, https://whois.icann.org/en (last accessed Oct. 23, 2018)

[21] Time, Inc., Fortune 500, Methodology and Credits, http://fortune.com/fortune500/ (last accessed Nov. 27, 2018).

Finally, Fortune 500 member organizations attract and employ top talent at every level. This includes internal and external network and systems management personnel, as well as highly skilled and experienced application development and operations staff. Many of these organizations have representatives on committees who provide leadership and governance of groups that develop IT and internet standards. In other words, if there are exposure issues in this group of organizations, it may be a signal that exposure conditions are even more substantial in companies that do not have similar stature.

Organization Internet Asset and Metadata Attribution Methodology

The Internet Assigned Numbers Authority (IANA) coordinates the governance of key elements that enable smooth operation of the internet.[22] Two key governance elements relevant to the process of attribution include internet address space (or "IP" addresses) and domain names (the system that helps turns web addresses such as `http://www.example.com/` into internet addresses so systems can connect to internet resources).

Attributing Internet Address Space to an Organization

IANA delegates the management of internet address space to a small number of global, regional internet registries. These registries further delegate blocks of internet addresses and coordinate the metadata associated with these assignments to national and "local" registries that ultimately coordinate with internet service providers (ISPs), which assign internet addresses to users and organizations.

Table 4: Rapid7 WHOIS Record Summary

INTERNET ADDRESS ASSIGNMENT	WHOIS ATTRIBUTION
71.6.233.0/24	Rapid7 Labs. Traffic originating from this network is expected and part of Rapid7 Labs Project Sonar sonar.labs.rapid7.com (C07045996)
208.118.237.0/24	Rapid7 LLC (C02934565)

The metadata associated with these internet address assignments, such as the organization names, location information, points of contact ,and potentially the parent internet service provider, is stored in a distributed set of databases called the WHOIS service. The WHOIS service is a public resource that allows a user to retrieve information about IP number, including the organization that owns the internet address and the organization's point of contact. Each registry maintains its own WHOIS database. Individuals can use WHOIS to make interactive queries to these systems, and bulk copies of WHOIS database information are made available to organizations that will use the data for technical research purposes.

When an organization wishes to manage its own internet connection resources, it makes a request to a local ISP or local registry and is assigned one or more contiguous sets of addresses to use. This attribution metadata is stored in the appropriate WHOIS service. To illustrate what this looks like, Table 4 shows the internet address block assignments for Rapid7.

There is no data format standard for WHOIS attribution. The records are little more than free-form text. There is no required structure for each field and no requirement to use and update legal company names as they are registered or change. However, it is possible to perform manual and automated mapping of organization names to these records. Rapid7 researchers used a combination of manual inspection and a proprietary machine learning-based algorithm to identify internet address space assignments of 460 members of the Fortune 500. Of these 460, Rapid7 researchers removed seven organizations, since it was unclear how to separate the internet address space used for an organization's own "core business" purposes from the space an organization leased to customers for their business purposes. These organizations tend to fall into the categories of internet service providers or cloud computing resource providers.

[22] Internet Assigned Numbers Authority, https://www.iana.org/ (last accessed Oct. 23, 2018).

Table 5: Fortune 500 Sector Study Representation Factoring In Project Sonar Blacklist

SECTOR	FINAL COUNT	DIFFERENCE FROM ORIGINAL F500 LIST	DIFFERENCE (%)
Financials	85	-3	3.4%
Energy	53	-6	10.2%
Health Care	39	-1	2.5%
Retailing	36	-10	21.7%
Technology	33	-6	15.4%
Wholesalers	23	-3	11.5%
Food, Beverages & Tobacco	22	-2	8.3%
Business Services	20	0	0.0%
Transportation	18	0	0.0%
Industrials	17	-1	5.6%
Materials	14	-5	26.3%
Aerospace & Defense	13	0	0.0%
Chemicals	12	-2	14.3%
Household Products	12	0	0.0%
Hotels, Restaurants & Leisure	11	0	0.0%
Media	11	0	0.0%
Engineering & Construction	10	-2	16.7%
Motor Vehicles & Parts	10	-1	9.1%
Food & Drug Stores	6	0	0.0%
Apparel	5	0	0.0%
Telecommunications	3	-5	62.5%

Table 5 provides the final industry sector breakdown for the final organization list used in this study.

Further care has been taken to attempt to identify internet address space that has been leased to service providers by these organizations. Internet address space is at a premium and is a valuable commodity, making it advantageous for an organization to lease a set of internet address ranges rather than just selling it outright. However, in these lease situations, attribution information will typically identify the organization that owns the internet address space, rather than the organization that is actually using the space. Rapid7 researchers acknowledge that they may not have identified all leased internet address space, which may have increased counts in various exposure categories. Possible ways to improve leased attribution are discussed below in Further Work.

Attributing DNS Records to an Organization

A similar WHOIS registration and database service exists for DNS assignment, except this is a far more distributed service that places direct control of all the underlying records for a domain into the hands of an organization. Once assigned a domain name (e.g. "rapid7.com"), an organization sets up its own DNS server (or uses one from a DNS service provider or cloud provider), then publishes and maintains records that map DNS names to a wide array of record types and values. Organizations can add/change/delete records at will.

DNS "A" (address) records map names to internet addresses (e.g., `www.rapid7.com` currently maps to `13.33.37.212`), but it is also possible to associate other types of information with an internet name.

DNS "TXT" (text) records facilitate storing arbitrary text strings with internet names. A number of formal standards exist that provide rules for crafting specially formatted text records to convey additional metadata about that internet name resource or the domain name owner proper.

Two TXT records that are key for inferring the "safety" of an organization's email configuration are DMARC[23] and the SPF[24]. These standards enable an organization to communicate which systems are authorized to send mail on its behalf and what should be done with forged email sent by attackers or spammers. Missing, improperly configured, or overly permissive configurations of these records put organizations at risk for both increased spam and phishing attacks. Since phishing attacks have been the primary means of attackers gaining a foothold within an organization for the past few years, lack of care and attention to appropriate DMARC and SPF configuration significantly increases the likelihood of successful attacks against that organization.

Anyone can query the DNS for these and other records. As part of our research efforts into ecosystem-wide cybersecurity, Rapid7 performs millions of DNS lookups every month and stores the time-stamped record results in a large, historical database, which makes it possible to perform large-scale queries and track changes over time.

The 2017 Fortune 500 list includes the primary, well-known domain names of the members of the list. For example, "apple.com" is the well-known domain for Apple Inc., and while "acdn1.com" is also owned by Apple, this domain isn't expected to handle email, nor does it have a valid email configuration. These sites were systematically scanned by Project Sonar, and the associated DNS names for the attributed organizations were used to determine the presence of DMARC and SPF.

---

[23] The DMARC Standard, https://dmarc.org/ (last accessed Oct. 23, 2018).

[24] The SPF Standard, Apr. 26, 2014, http://www.openspf.org/.

APPENDIX:

INDUSTRY SECTOR BREAKOUT

Table 3: Fortune 500 Sector Counts

SECTOR	INDUSTRY	NUMBER IN FORTUNE 500
Aerospace & Defense	Aerospace and Defense	13
Apparel	Apparel	5
Business Services	Advertising, marketing	2
Business Services	Diversified Outsourcing Services	5
Business Services	Financial Data Services	9
Business Services	Miscellaneous	1
Business Services	Temporary Help	1
Business Services	Waste Management	2
Chemicals	Chemicals	14
Energy	Energy	8
Energy	Mining, Crude-Oil Production	11
Energy	Oil and Gas Equipment, Services	2
Energy	Petroleum Refining	9
Energy	Pipelines	7
Energy	Utilities: Gas and Electric	22
Engineering & Construction	Engineering, Construction	7
Engineering & Construction	Homebuilders	5
Financials	Commercial Banks	20
Financials	Diversified Financials	13
Financials	Insurance: Life, Health (Mutual)	7
Financials	Insurance: Life, Health (Stock)	11
Financials	Insurance: Property and Casualty (Mutual)	5
Financials	Insurance: Property and Casualty (Stock)	20
Financials	Real Estate	5
Financials	Securities	7
Food & Drug Stores	Food and Drug Stores	6
Food, Beverages & Tobacco	Beverages	4

SECTOR	INDUSTRY	NUMBER IN FORTUNE 500
Food, Beverages & Tobacco	Food Consumer Products	13
Food, Beverages & Tobacco	Food Production	5
Food, Beverages & Tobacco	Tobacco	2
Health Care	Health Care: Insurance and Managed Care	9
Health Care	Health Care: Medical Facilities	7
Health Care	Health Care: Pharmacy and Other Services	6
Health Care	Medical Products and Equipment	7
Health Care	Pharmaceuticals	11
Hotels, Restaurants & Leisure	Food Services	5
Hotels, Restaurants & Leisure	Hotels, Casinos, Resorts	6
Household Products	Home Equipment, Furnishings	4
Household Products	Household and Personal Products	8
Industrials	Construction and Farm Machinery	6
Industrials	Electronics, Electrical Equipment	4
Industrials	Industrial Machinery	7
Industrials	Miscellaneous	1
Materials	Building Materials, Glass	2
Materials	Forest and Paper Products	1
Materials	Metals	6
Materials	Miscellaneous	1
Materials	Packaging, Containers	9
Media	Entertainment	9
Media	Publishing, Printing	2
Motor Vehicles & Parts	Motor Vehicles and Parts	11
Retailing	Automotive Retailing, Services	9
Retailing	General Merchandisers	9
Retailing	Internet Services and Retailing	3
Retailing	Specialty Retailers: Apparel	7
Retailing	Specialty Retailers: Other	18
Technology	Computer Software	4
Technology	Computers, Office Equipment	8
Technology	Entertainment	1
Technology	Information Technology Services	7
Technology	Internet Services and Retailing	5

SECTOR	INDUSTRY	NUMBER IN FORTUNE 500
Technology	Network and Other Communications Equipment	3
Technology	Scientific, Photographic, and Control Equipment	1
Technology	Semiconductors and Other Electronic Components	10
Telecommunications	Telecommunications	8
Transportation	Airlines	6
Transportation	Mail, Package, and Freight Delivery	2
Transportation	Railroads	3
Transportation	Transportation and Logistics	3
Transportation	Transportation Equipment	2
Transportation	Trucking, Truck Leasing	2
Wholesalers	Wholesalers: Diversified	9
Wholesalers	Wholesalers: Electronics and Office Equipment	5
Wholesalers	Wholesalers: Food and Grocery	6
Wholesalers	Wholesalers: Health Care	6

ABOUT RAPID7

Rapid7 powers the practice of SecOps by delivering shared visibility, analytics, and automation that unites security, IT, and DevOps teams. The Rapid7 Insight platform empowers these teams to jointly manage and reduce risk, detect and contain attackers, and analyze and optimize operations. Rapid7 technology, services, and research drive vulnerability management, application security, incident detection and response, and log management for organizations around the globe. To learn more about Rapid7 or get involved in our threat research, visit **www.rapid7.com**.

QUESTIONS

Reach us at research@rapid7.com