Microsoft

Microsoft Digital Defense Report

September 2020

Table of contents

Introduction

TOM BURT, CORPORATE VICE PRESIDENT, CUSTOMER SECURITY AND TRUST

The global events of the past 12 months have brought unprecedented change to the physical and digital worlds. Cybercrime, however, is a constant. We've seen that cybercriminals continue—and sometimes escalate—their activity in times of crisis. Defending against cybercriminals is a complex, ever-evolving, and never-ending challenge.

But knowledge is power. For security professionals to create successful defense strategies, they need more diverse and timelier insights into the threats they are defending against. We're proud to provide the global community with the latest in a long series of security intelligence reports. The Microsoft Digital Defense Report is a reimagining of Microsoft's Security Intelligence Report (SIR), first published in 2005, and it brings together more insights, from more teams, across more areas of Microsoft than ever before targeting a broader audience for consumption.

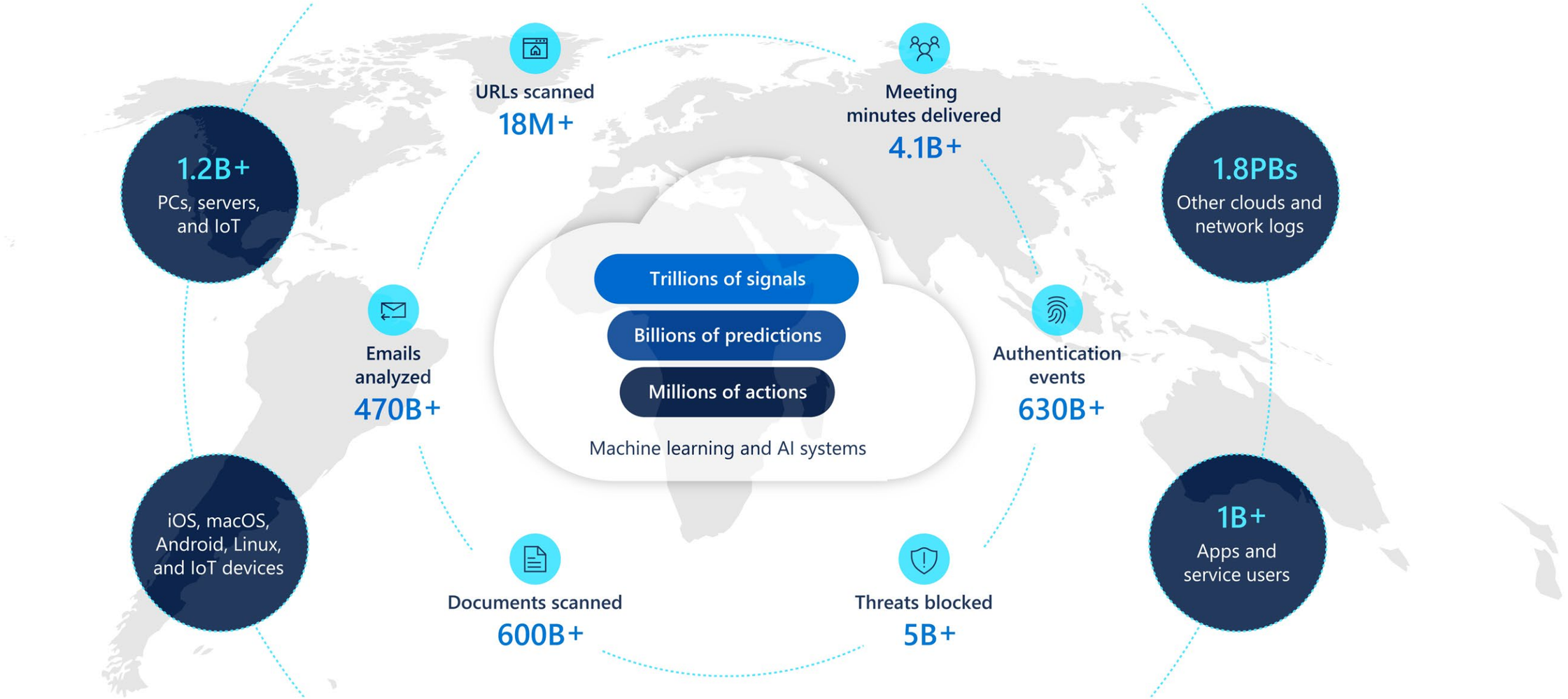
Microsoft serves billions of customers globally, allowing us to aggregate security data from a broad and diverse spectrum of companies, organizations, and consumers. Our unique position helps us generate a high-fidelity picture of the current state of cybersecurity, including indicators to help us predict what attackers will do next. This picture is informed by over 8 trillion security signals per day.

There are thousands of security experts across 77 countries working at Microsoft, interpreting and contributing to the insights gained from our advanced engineering and telemetry. Our security experts include analysts, researchers, responders, engineers, and data scientists. This report draws on insights, data, and signals from across Microsoft, including the cloud, endpoints, and the intelligent edge.[1] We also share lessons learned from customers transitioning to a fully remote workforce and frontline stories from our incident responders. Of course, there's a great deal of malign activity we don't see, some of which is reported on by others in the industry. While the defender community at Microsoft works hard to identify threats and keep our customers informed, the bad actors are skilled and relentless. By continually sharing insights that we and others in the industry derive from the work we do, we hope to empower everyone to defend the online ecosystem more effectively.

[1] *These signals are collected with customer privacy in mind. The data we collect depends on the context of your interactions with Microsoft and the choices you make, including your privacy settings and the products and features you use.*

Unique insights informed by trillions of signals

Monthly volume and diversity of signals used by Microsoft security operations



Microsoft invests significantly to increase and improve the knowledge we derive from our telemetry. These investments deliver the highly synthesized and integrated insights that we share here. The goal of this report is to help organizations understand how cybercriminals are shifting their modes of attack and the best ways to combat those attacks. We've approached the writing and sharing of this report in the spirit of enabling the community to benefit from the insights, observations, and transparency generated by our unique mission and vantage point.

Digital defense: Our 2020 focus areas

2020 has brought major disruptions to both the physical and digital worlds, and these changes are also evident in the cyberthreat landscape. Certain types of attacks have escalated as cybercriminals change tactics, leveraging current events to take advantage of vulnerable targets and advance their activity through new channels. Change brings opportunity, for both attackers and defenders, and this report will focus on the threats that are most novel and relevant to the community in this moment.

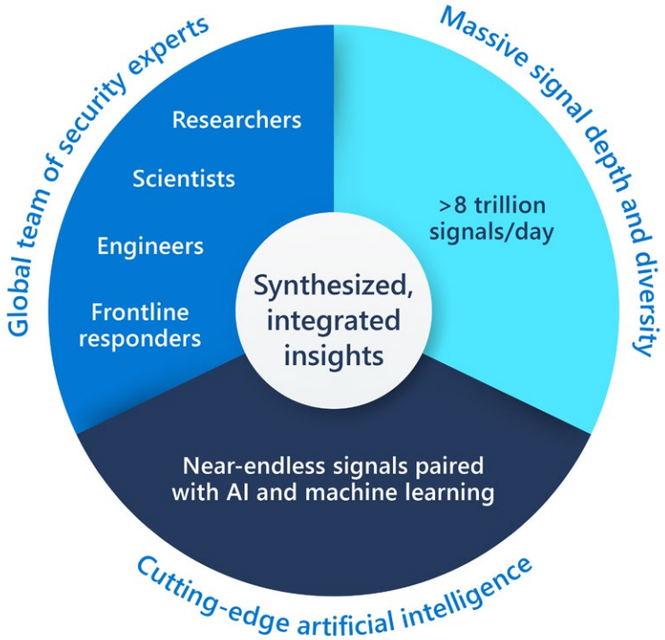
Looking at the data and signals from the cross-company teams, three top-level areas came into the sharpest focus: cybercrime, nation state threats, and the remote workforce.

The state of cybercrime

Cybercrime is a business. Like any other business, there's a need to innovate to be profitable and successful. Some types of cybercrime persist independent of economic, political, or social changes, while certain types are fueled by these changes. Cybercriminals can be found globally and have different skill sets and motivations. In this section, we look at the complexity that Microsoft and the industry face when defending against an extremely diverse set of criminal actors and their ever-evolving tactics and techniques.

A key area we address in this section is the opportunistic nature of cybercriminals as they capitalized on interest and fear related to the COVID-19 pandemic and other disruptive events. As the virus spread globally, cybercriminals pivoted their lures to imitate trusted sources like the World Health Organization (WHO) and other national health organizations, in an effort to get users to click on malicious links and attachments.

Synthesized insights drawing from security and threat intel teams across Microsoft and over 8 trillion security signals per day.



Email phishing in the enterprise context continues to grow and has become a dominant vector. Given the increase in available information regarding these schemes and technical advancements in detection, the criminals behind these attacks are now spending significant time, money, and effort to develop scams that are sufficiently sophisticated to victimize even savvy professionals. Attack techniques in phishing and business email compromise (BEC) are evolving quickly. Previously, cybercriminals focused their efforts on malware attacks, but they've shifted their focus to ransomware, as well as phishing attacks with the goal of harvesting user credentials. Human-operated ransomware gangs are performing massive, wide-ranging sweeps of the internet, searching for vulnerable entry points, as they "bank" access, waiting for a time that's advantageous to their purpose.

We share leading indicators of where attacks might be headed next, as we provide a look into adversarial machine learning (ML), attacks on ML systems, and why it's so important for organizations to take steps to secure them. We propose democratizing ML for customized anomaly detections and a scalable effort, making it ubiquitous across the variety of data types.

Modern business systems are driven by a complex, global supply chain. In this section, we focus our observations and recommendations for supply chain security on third-party services, open-source software, and Internet of Things (IoT) hardware, concluding with a look at changes to the regulatory landscape.

Nation state threats

Nation state actors are well-funded, well-trained, and have more patience to play the "long game," which can make identification of anomalous activity more difficult. Like cybercriminals, they watch their targets and change techniques to increase their effectiveness. To protect our customers, Microsoft spends significant resources monitoring and disrupting nation state attacks attempted on our platform. In this section, we explain the four main approaches Microsoft employs to thwart nation state actors: technology, operations, legal action, and policy.

We also provide our analysis of the intent behind nation state threats and how to defend against them. We look at top-level trends in country-of-activity origin, targeted geographic regions, and the top nation state activity groups detected. Interestingly, nation state activity is significantly more likely to target organizations outside of the critical infrastructure sectors. The most frequently targeted sector has been non-governmental organizations (NGOs), such as advocacy groups, human rights organizations, nonprofit organizations, and think tanks focused on public policy, international affairs, or security.

Whatever the strategic objectives behind the activity, nation state actors have these common operational aims: espionage, disruption or destruction of data, and disruption or destruction of physical assets.

Finally, we examine some of the most common attack techniques used by nation state actors in the past year: reconnaissance, credential harvesting, malware, and virtual private network (VPN) exploits. As an example, we see advanced adversaries investing heavily in development of unique malware, in addition to using openly available malicious code.

Security and the remote workforce

Almost overnight, the workforce of thousands of organizations around the world became entirely remote. School closures forced millions of students to rapidly transition to learning from home—and added significant challenges for parents and caretakers. Although workforces around the world, regardless of size, have been trending toward mobility in some aspects of their operations, few companies and learning institutions were set up to operate 100% remotely. Operational tasks like software or device patching and updates had previously been accomplished when mobile workers routinely returned to the office, but after the COVID-19 outbreak, this option temporarily disappeared. At Microsoft, COVID-19 became a catalyst for managing a remote workforce with immediacy and at scale.

In this chapter, we take a closer look at three important areas of consideration for an at-scale remote workforce: infrastructure, data, and people. Architecture and infrastructure designed to support an on-premises workforce isn't well-equipped to support fully remote workers. We explain how organizations can support a secure, remote workforce through VPN architecture and the principles of Zero Trust, and how Microsoft used Zero Trust principles to enable its own transition to remote work. We discuss some of the largest availability and security concerns facing the remote workforce, such as distributed denial of service (DDoS) attacks, and we explore what this means for organizational privacy and security.

Data protection practices continue to increase, as workforces become remote and teams collaborate on vital assets without being physically together. Correspondingly, we observe the continued increase in usage of information rights management to enforce policies aimed at protecting confidential information and intellectual property.

Finally, we reflect on our enterprise-scale exercise in resilience and lessons learned as the world moves through a global pandemic and billions of individuals adapt to working, learning, and socializing from their home environments. To some degree, the enterprise response to COVID-19 changed not only our operational procedures but also the very vocabulary that we use to describe our reactive measures. In this section, we consider "the extended security boundary," "pandemic resilience," "human infrastructure," and ways in which traditionally less scrutinized areas have become critical to enterprise resilience.

How to read this report

To help you get the most out of this report, we've incorporated navigational elements throughout to increase readability.

Telemetry icons

These guide you to areas that show you **what we're seeing.**

Best practice icons

These guide you to our recommended controls.

Actionable learnings

Located at the end of the report, these provide a summary based on industry guidance and what we've seen work best for Microsoft and our customers.

Glossary

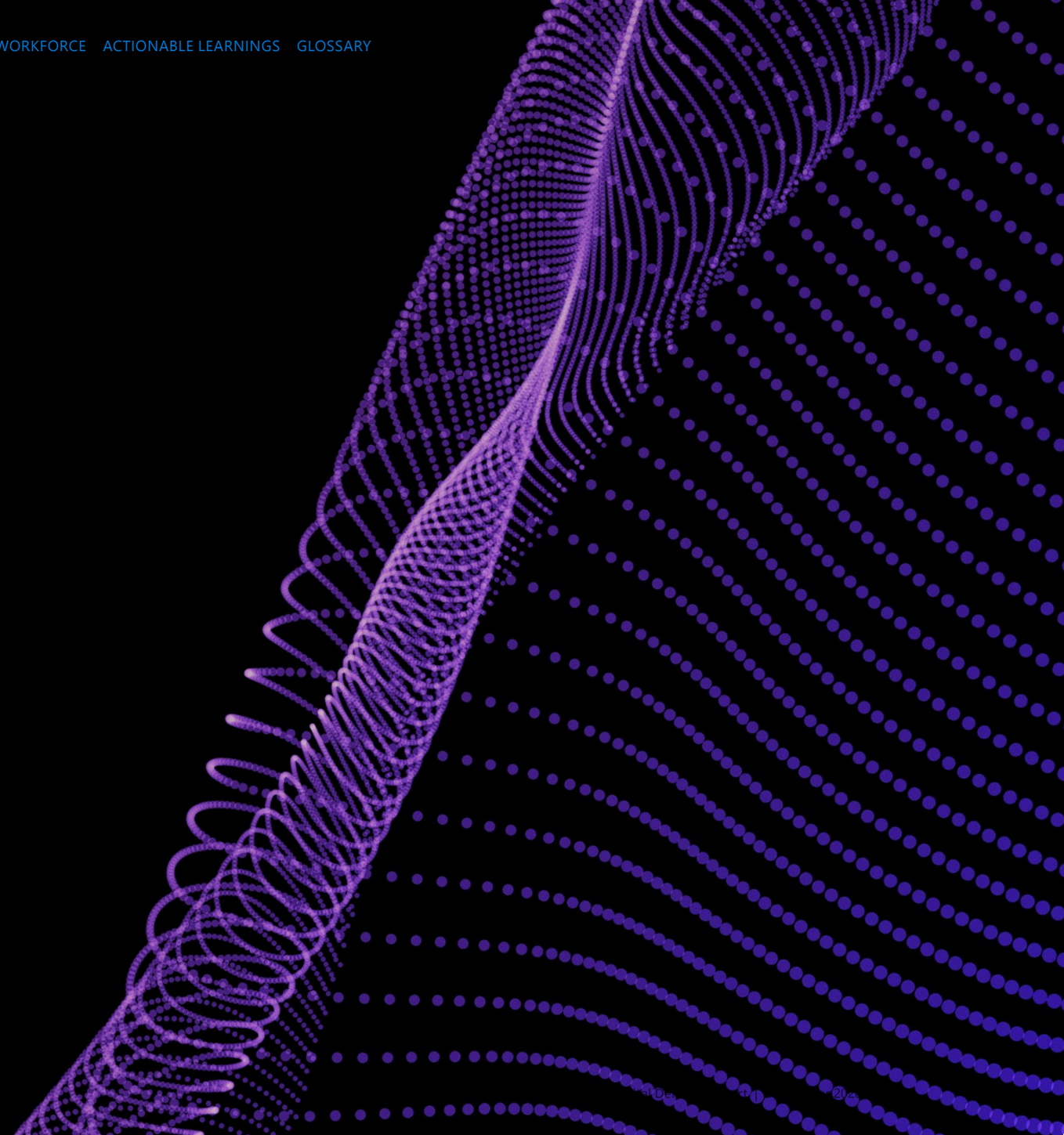
A clarification of acronyms and terminology, located at the end of the report.

Learn more

Resources and references for deeper technical information and further reading are listed in each chapter.

1

The state of cybercrime



Introduction

AMY HOGAN-BURNEY, GENERAL MANAGER, DIGITAL CRIMES UNIT

Cybercrime is an ongoing and escalating challenge for both the public and private sectors around the globe. It's no longer effective to view cybercrime solely through the lens of a specific criminal or type of crime. Rather, cybercrime is now a large and diverse enterprise, often available for sale, to commit a variety of attacks, such as fraud, theft, and spying. It could be financially motivated, or nation state supported, or both.

Activity like this is possible because of a global network of actors and infrastructure used to facilitate the crime. Arrests and prosecution can be elusive owing to the international nature and the often-differing locations of the criminal, crime, and victim. The actions of cybercriminals are often disguised, leveraging a crisis or major global event and the daily news cycle and relying on personal behaviors to perpetrate a criminal act. As varied as the motivations are, so too are the number and types of victims: individuals, governments, industries, or corporations can be targeted.

Despite sophistication and diversity of the attacks, the methodology is often the same, whether the actors use large-scale attacks for financial gain or targeted attacks to support geopolitical interests. A phishing email can be a massive campaign targeting millions of users or a single, targeted email that represents a socially engineered marvel many months in the making. Homoglyphs, or spoofed domains, can be used to trick victims.

An example of this usage is "Microsoft.com" and "Micr0soft.com," where the first "o" is replaced by a zero and can be easily overlooked by human readers. This malicious domain, Micr0soft.com, can then be used to distribute malware, steal credentials, or support a fraudulent website. That malware then can be used to create a botnet to facilitate a DDoS attack against a bank, distribute ransomware, or steal sensitive information about a nation's critical infrastructure. We look for commonalities across various environments and ecosystems to understand and disrupt these methods. We work to dismantle the criminals' infrastructure, sharing information gathered through the course of our investigations as appropriate. We can then offer these additional insights to our product teams who use this information to provide greater protection to our customers.

In this chapter, we present what we're seeing as the state of cybercrime through evolving techniques in phishing and BEC, ransomware, and malware. We also focus our observations and recommendations on supply chain security. Finally, we include a deeper look into four trends in ML.

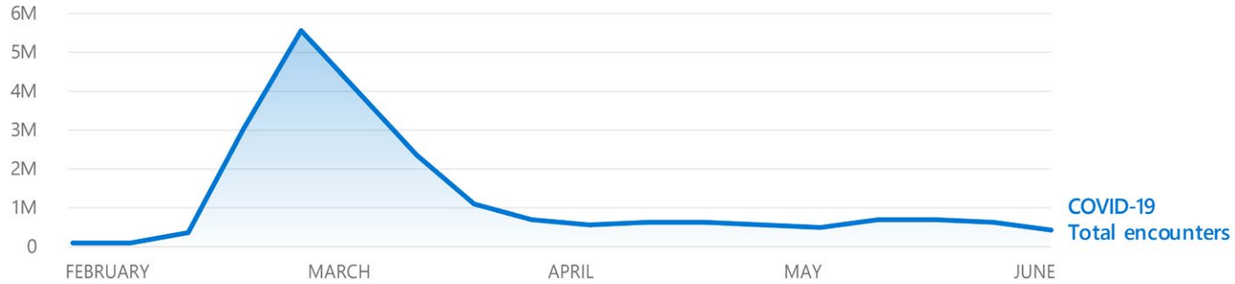
What we're seeing

Cybercrime follows the contemporary issues of the day

The threat environment we face continues to evolve. Cybercriminals are creative, well-resourced, well-organized, and innovative. They move quickly to discover new threat vectors, use new exploits, and respond to new defenses.

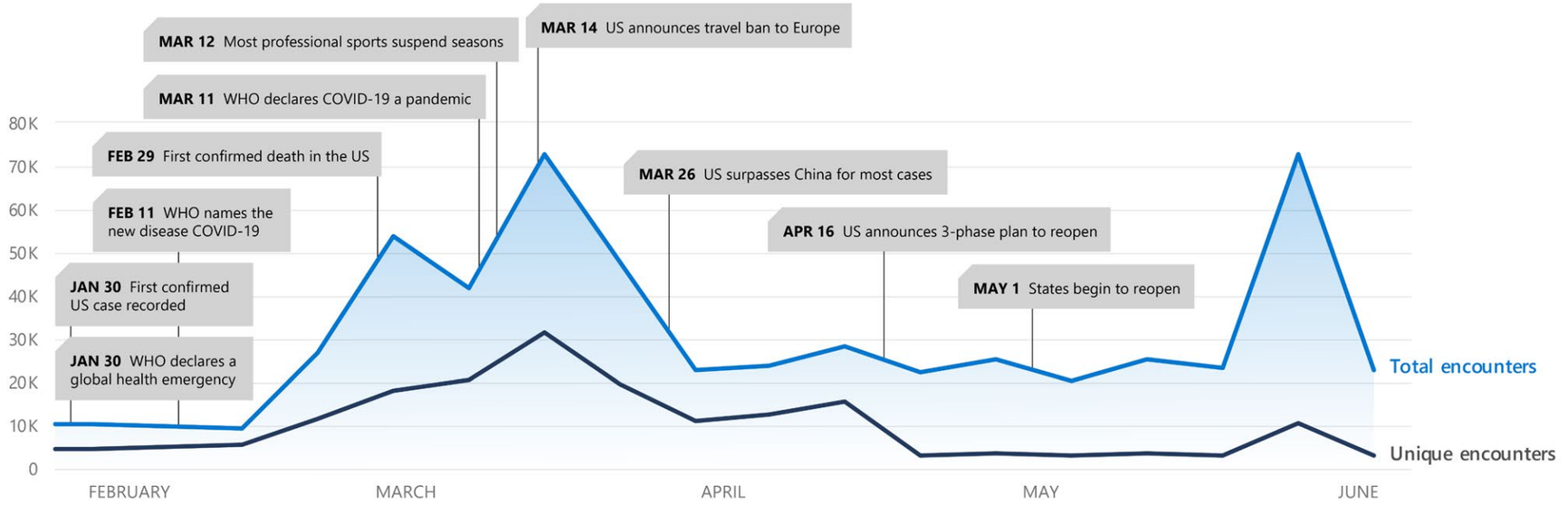
Attackers are opportunistic and will even switch lure themes daily in accordance with news cycles, as seen in cybercriminals' use of the global COVID-19 pandemic to broadly target consumers, as well as to specifically target hospitals and healthcare providers.

COVID-19-themed attacks: United States

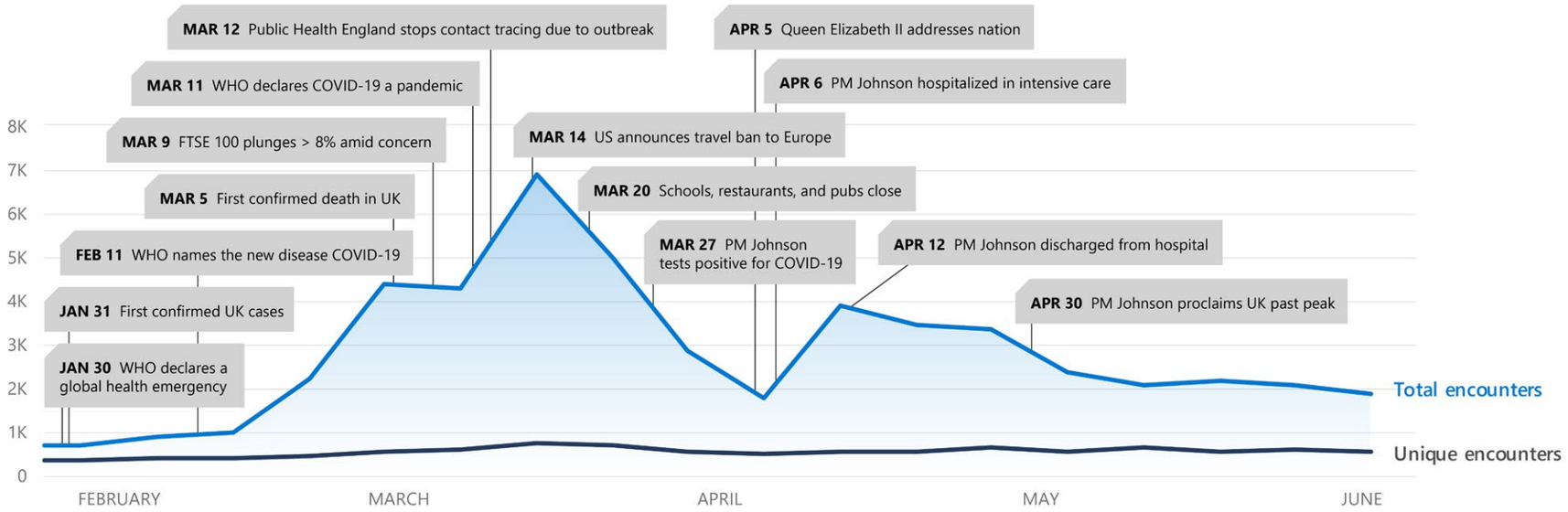


The lack of basic security hygiene in any given ecosystem continues to enable cybercriminals to use well-known vulnerabilities—or new variants of them—to exploit their environments. They also were observed to leverage the fear and uncertainty associated with COVID-19 with great success. While the COVID-19-themed attacks represent a small percentage of the total of malware we observed, our tracking of these themed attacks shows how rapidly cybercriminals move to adapt their lures to the topics of the day.

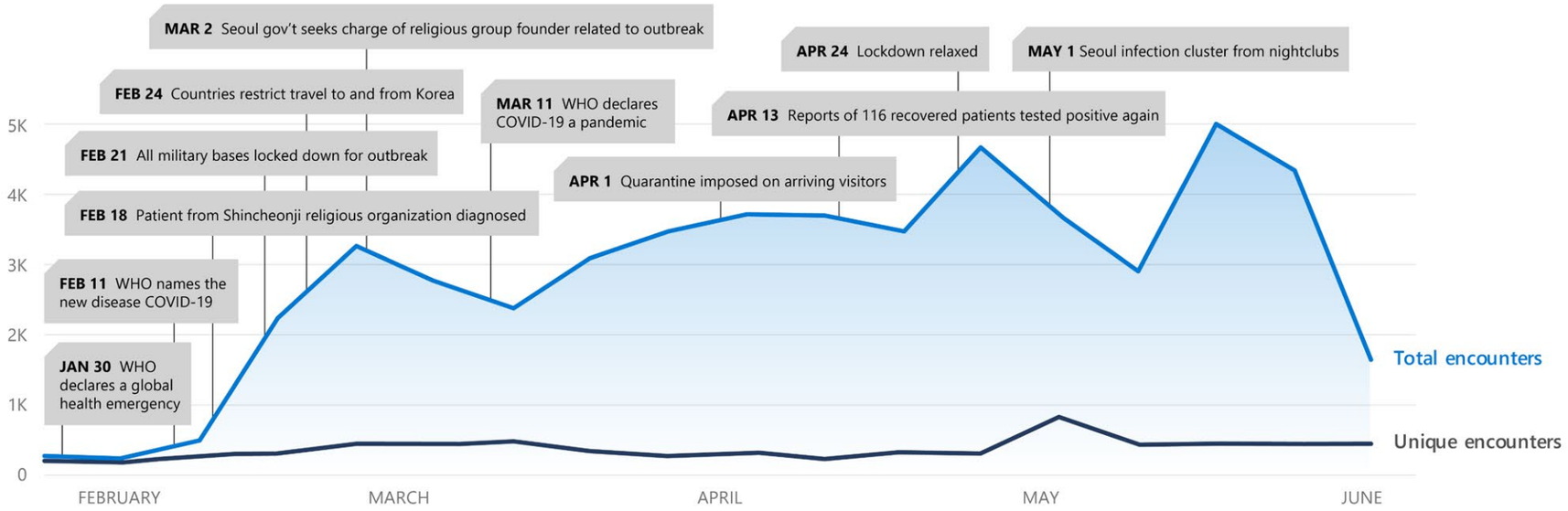
Instances of unique and total malware encounters in relation to local news events of the day, as seen in the United States, the United Kingdom, and South Korea



COVID-19-themed attacks: United States



COVID-19-themed attacks: United Kingdom

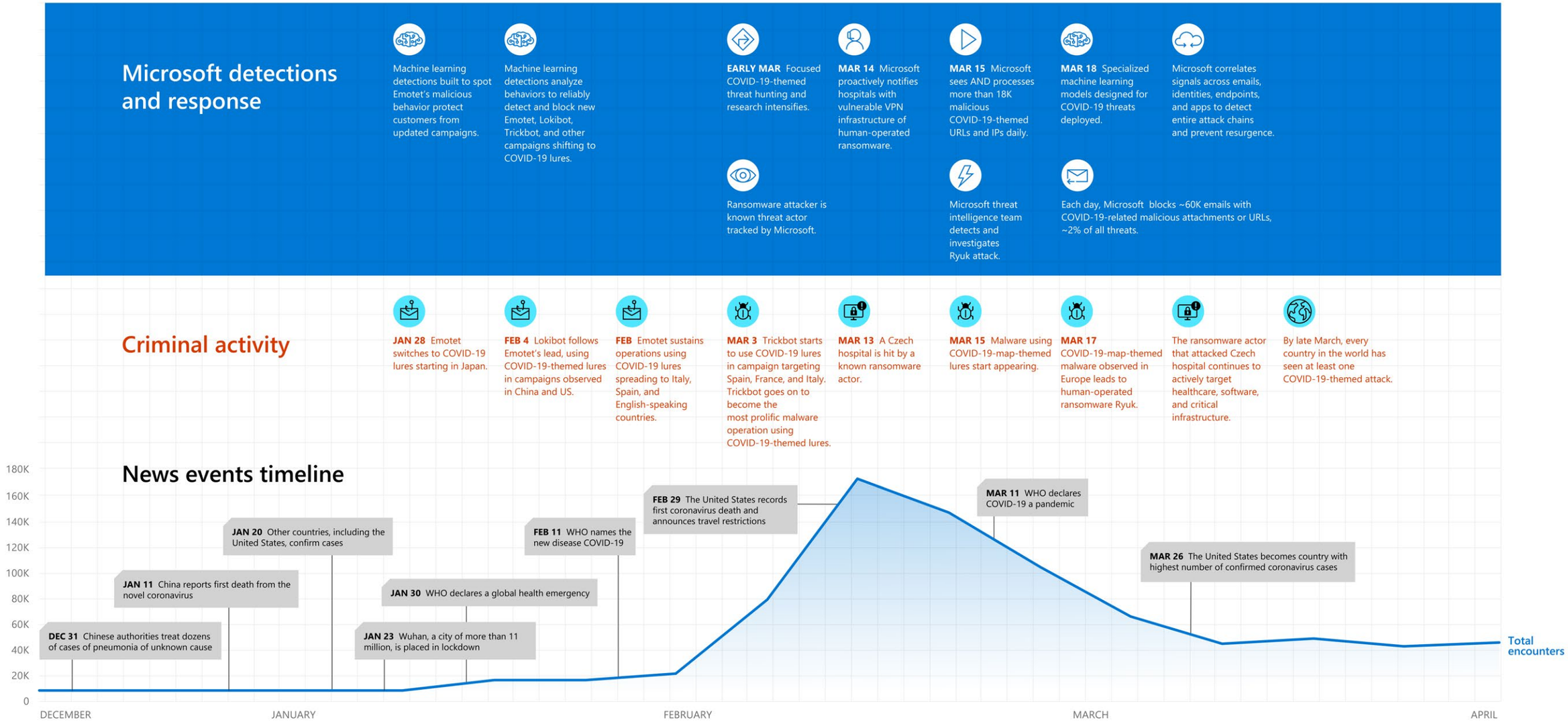


COVID-19-themed attacks: South Korea

As our threat intelligence teams at Microsoft actively monitor and respond to these shifts in focus, our telemetry shows that, while the overall volume of malware remains largely the same, adversaries used the COVID-19 theme to socially engineer lures around the anxiety and the flood of information associated with the pandemic.

Attackers also exploited the crisis to reduce their dwell time within a victim's system—compromising, exfiltrating data, and in some cases ransoming quickly—apparently believing that there would be an increased willingness to pay as a result of the outbreak. It's a common understanding among defenders to "never waste a crisis" and to use the information from the crisis to help inform the investments needed. Cybercriminals share the same philosophy; they seek to blend their well-established tactics and malware with human curiosity and our need for information.

News events, criminal activity, and Microsoft's response



Our intelligence shows that these attacks settled into a rhythm that's the normal ebb and flow of the threat environment. Our telemetry shows that China, the United States, and Russia were hit the hardest, but every country in the world saw at least one COVID-19-themed attack, with the volume of successful attacks in outbreak-hit countries increasing as fear and the desire for information grew.

Of course, COVID-19 was just one of the themes exploited by cybercriminals. Microsoft's Digital Crimes Unit (DCU) continues to work with partners and proper authorities in more than 50 countries to take coordinated legal and technical steps to take down malicious domains and URLs, and, where possible, to prosecute the individuals behind them.

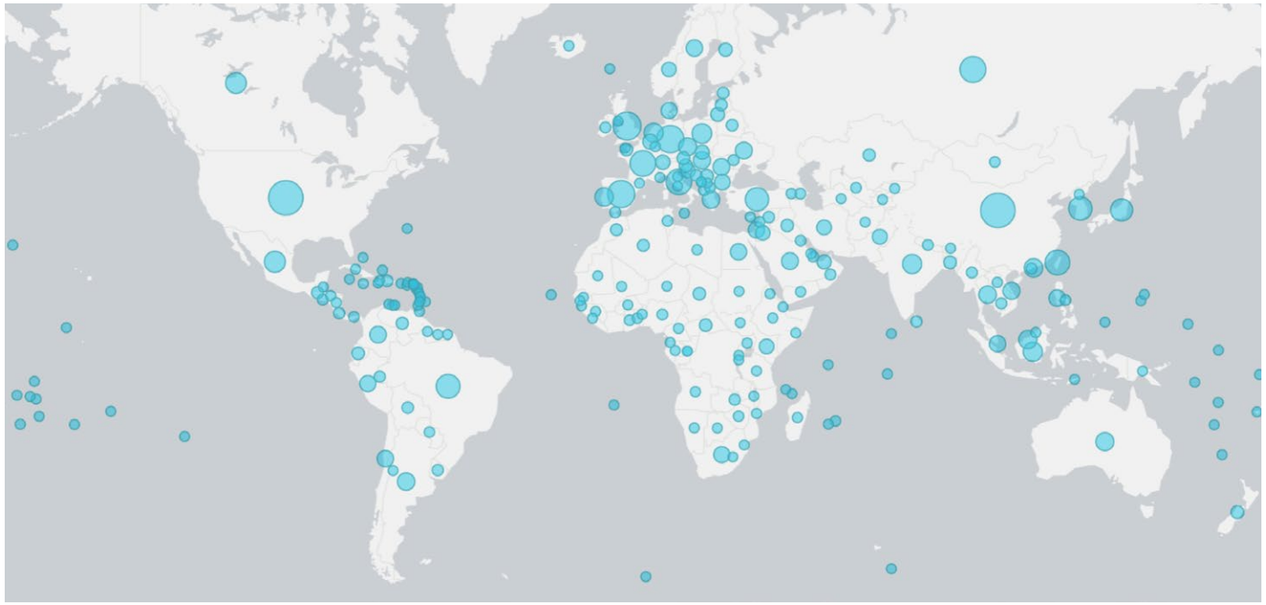
For more information on COVID-19-related cybercrime, see the Nation state threats chapter of this report.

Learn more:

Microsoft takes legal action against COVID-19-related cybercrime 2020/07/07

Microsoft shares new threat intelligence, security guidance during global crisis 2020/04/08

Relative impact of COVID-19-themed attacks across the world by file count (as of July 1, 2020)



Phishing and business email compromise

Email phishing in the enterprise context continues to grow and has become a dominant vector. Essentially, phishing occurs when individuals or organizations receive a fraudulent email encouraging them to click on a link, giving the cybercriminal access to a device or personal information.

From our view in Office 365 telemetry, we see customers facing three main forms of phishing:

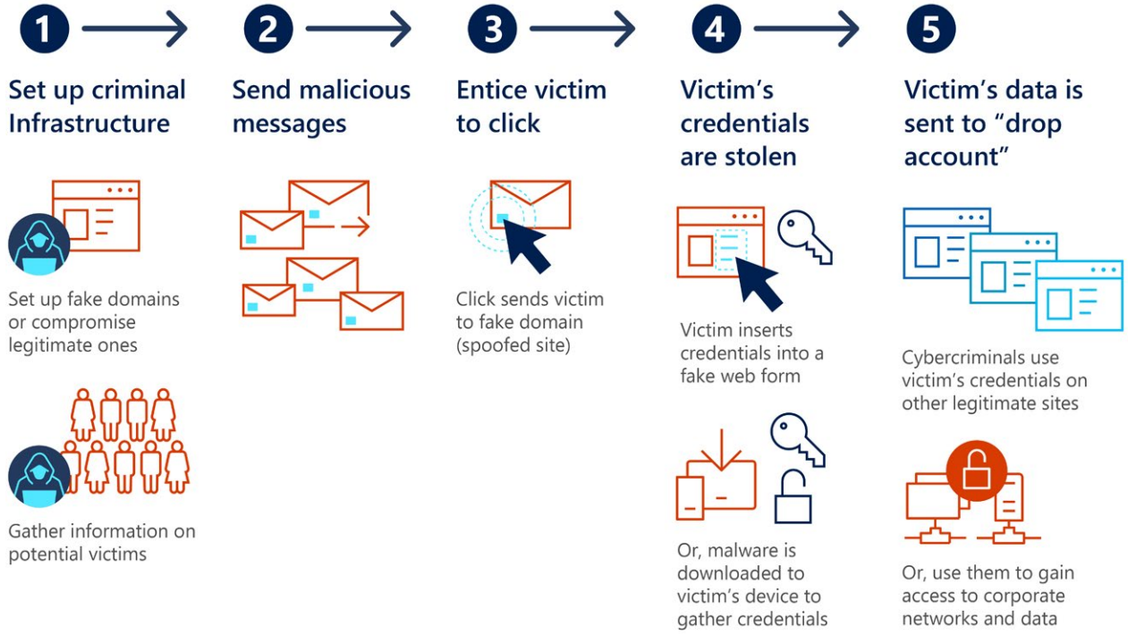
1. Credential phishing

The cybercriminal attempts to pose as a well-known service in the email template (typically Microsoft and other established enterprise SaaS services). They then try to lure the user into clicking on a link, which, through a series of misdirections, will present a fake login page to the user. Once a credential has been compromised, it can be used to launch different kill chains to build further persistence inside an organization by using cloud-only APIs and systems, and to move around laterally to steal data, money, or otherwise breach the organization.

Example of credential phishing:
The cybercriminals begin by setting up a criminal infrastructure designed to steal an individual's credentials (phishing kits often contain everything they need for this). The cybercriminals send malicious email to the unsuspecting individual, who then clicks on a link within the email. The individual might then be taken to a fake web form to enter their credentials, or the site might contain malware that's automatically downloaded to their device, capturing credentials stored on the device or in the browser memory. The victim's credentials are then collected by the cybercriminals who use the credentials to gain access to legitimate websites or even the victim's corporate network.

Reduce risk of security breaches with strong authentication.[2] Strong authentication can protect your users from the vast majority of identity attacks. Passwordless authentication options are recommended for best security and user experience. Using an authenticator app is always the preferred option over SMS/voice authentication.

Credential phishing example



[2] *Office 365 Advanced Threat Protection and Microsoft Defender Advanced Threat Protection both protect against credential phishing and BEC.*

2. Business email compromise (aka CEO fraud or vendor compromise)

Business email compromise is a type of phishing that specifically targets businesses. It's characterized by techniques used to pose as someone who the victims will likely take notice of, such as the company CEO, CFO, or the accounts receivable clerk. BEC can also involve a business-to-business transaction. For example, the cybercriminal might fraudulently access a company's system and then pose as that company to fraudulently request payment from another company.

Depending on the entity that gets impersonated or spoofed, the attack can be categorized as BEC, CEO fraud, or vendor compromise.

While the industry and vendor-specific terminology for this class of attacks has been evolving, at its core it employs three basic patterns of a cybercriminal posing as another entity:

- Cybercriminal creates an *exact spoof* of an established person or organization that the recipient has a relationship with, and then gains trust to deliver the fraud. For example, the email comes from an infrastructure controlled by the cybercriminal but appears to come from a legitimate person (sender address is identical).
- Cybercriminal creates a *look-alike impersonation* of an established person or organization that the recipient has a relationship with, and then gains trust to deliver the fraud. For example, the criminal signs up for and sends from this account: YourCompanyCEO@outlook.com.
- Cybercriminal creates a *look-alike domain* and registers it for the sole purpose of delivering a BEC attack. For example, instead of CEO@microsoft.com, the criminal sets up the domain micros0ft.com (number zero replaces letter "o") and sends email from it.

At its core, BEC is a social engineering scam. However, given the increase in available information regarding these schemes and technical advancements in detection, the criminals behind these attacks are now spending significant time, money, and effort to develop scams that are sufficiently sophisticated to victimize increasingly savvy professionals.

...the criminals behind these attacks are now spending significant time, money, and effort to develop scams that are sufficiently sophisticated to victimize increasingly savvy professionals.

3. Combination of BEC and credential phishing

We're also seeing new attack kill chains that combine the two forms to deliver more sophisticated kill chains. This attack can start with credential phishing. Once an account is compromised, the cybercriminal sets up mailbox "forwarding" rules to monitor for financial transactions. Often these forwarding rules include keywords such as "invoice," "accounts receivable," "funds," "overdue," "payroll," or "IBAN" and send all relevant email to a collection email account controlled and monitored by the cybercriminal. The cybercriminal then inserts a victim impersonation email in the middle of a communication to misdirect and steal money or information.

While credential phishing and BEC continue to be the dominant variations, we also see attacks on a user's identity and credential being attempted via password reuse and password spray attacks using legacy email protocols such as IMAP and SMTP.

Example of BEC where cybercriminal masquerades as CEO:

Depending on the level of sophistication of the attack, the cybercriminals might begin by monitoring the CEO's mail account for information such as relationships within the company, common phrases used, business activities or travel, and especially information about payments or wire transfers. The cybercriminals might set mailbox forwarding rules using keywords such as "invoice" or "accounts receivable" so that mail containing these words is directed to a drop email account that they control and monitor.

The cybercriminals can then send mail requesting sensitive information or payments to be made, using links or instructions to wire payment to a fraudulent bank account. Unsuspecting employees will comply with the requests, believing the requests come from the CEO, and wire the payment to the cybercriminal's account.

Example of BEC where cybercriminal masquerades as CEO

COVID-19-themed phishing lures

Microsoft tracks thousands of email phishing campaigns that cover millions of malicious messages every week. Phishing campaigns are more than just one targeted email at one targeted user. They include potentially hundreds or thousands of malicious emails targeting hundreds or thousands of users, which is why they can be so effective. Of the millions of targeted messages we see each day, roughly 60,000 included COVID-19-related malicious attachments or malicious URLs during April 2020. We've seen many instances of attackers impersonating established entities like the WHO, Centers for Disease Control and Prevention (CDC), and the Department of Health to get into inboxes. The trendy and pervasive Trickbot and Emotet malware families are very active in rebranding their lures to take advantage of the outbreak. As of June 30, 2020, we observed 79 threat variants globally using COVID-19-themed lures.

What we're seeing in Office 365 telemetry

In 2019 we blocked over 13 billion malicious and suspicious mails, out of which more than 1 billion were URL-based phishing threats (URLs set up for the explicit purpose of launching a phishing credential attack).[3] These URLs were set up and weaponized just in time for the attacks and had no previous malicious reputation. We're seeing approximately 2 million such URL payloads being created each month for credential harvesting, orchestrated through thousands of phishing campaigns.

Evolving techniques

Attack techniques have continued to evolve over the past year, with notable variation in campaign orchestration, URL payloads, and delivery mechanism.

Attack techniques

- **Increased usage of cloud services and compromised infrastructure**
  As detection rates improve, we're seeing cybercriminals shift tactics and use cloud services and compromised email and web hosting infrastructures to orchestrate these phishing campaigns. They aim to hide among these reputable cloud services to avoid detection. We've seen cybercriminals leverage the most popular cloud services, email sending services, and file sharing services to launch attacks.

- **Rapidly changing campaigns**
  We're also seeing attack campaigns that are being rapidly changed or morphed to avoid detection. Morphing is being used across sending domains, email addresses, content templates, and URL domains. The goal is to increase the combination of variations to evade detections. In Office 365, we're investing in sophisticated campaign clustering intelligence to enable security operations center (SOC) teams to piece together these complex campaigns from their fragments.

- **Constant evolution in payload delivery mechanisms**
  Over the last year, we saw interesting techniques used for launching attacks. We saw cybercriminals using poisoned search results and legitimate URLs that linked to those searches to deliver an attack. In another attack, we saw cybercriminals use custom 404 pages to host phishing payloads. We've also seen man-in-the-middle components used to present less suspicious sites to the targets and captcha and other evasion tools to hide detections.

As defenders, we continue to leverage this data and telemetry to add enhanced detection and intelligence capabilities in our Office 365 filtering services, across Exchange Online Protection and Office 365 Advanced Threat Protection, to help secure our customers from these modern and increasingly sophisticated attack variations.

Up until a few years ago, cybercriminals focused their efforts on malware attacks because they provided the greatest ROI.

More recently, they've shifted their focus to phishing attacks (~70%) with the goal of harvesting user credentials.

What we've seen in Office 365 Advanced Threat Protection detection in the past year:

6T
Messages scanned

~13B
Malicious emails blocked

~1.6B
URL-based email phishing threats blocked

~1.7-2M
URL payloads being created each month, orchestrated through thousands of phishing campaigns

[3] *Office 365 email gateway services across Exchange Online Protection and Advanced Threat Protection scan trillions of emails every year.*

What we're seeing in the Digital Crimes Unit

The Microsoft Digital Crimes Unit is a team of attorneys, investigators, data scientists, engineers, analysts, and business professionals who fight cybercrime globally through the innovative application of technology, forensics, civil actions, criminal referrals, and public and private partnerships, while protecting the security and privacy of our customers. Some of the DCU observations about phishing and BEC are shared here.

Email-facilitated cybercrime, such as credential phishing for cloud services and BEC, continues to expand in scope and complexity. According to the Federal Bureau of Investigation (FBI), as reported in the 2019 Internet Crime Report, the Internet Crime Complaint Center (IC3) received nearly a half a million complaints and recorded over $3.5 billion in losses to individual and business victims. The most frequently reported complaints involved email-facilitated cybercrime and the most costly involved BEC crimes. According to the IC3, BEC complaints totaled 23,775 and accounted for losses of more than $1.7 billion—representing nearly half of all financial losses owing to cybercrime.

Targeting the C-Suite

Although BEC scams have posed a significant threat to businesses and individuals for many years, these schemes continue to evolve in sophistication and reach. Early BEC schemes tended to focus on spoofing or impersonating emails from CEOs or CFOs in an attempt to redirect wire transfers to fraudulent accounts controlled by criminals. Current BEC schemes include multiple phases and usually start with reconnaissance of targeted businesses prior to initiating an attack. Research by cybercriminals during the initial phase results in identifying specific accounts to target for account compromise through techniques like credential phishing. We've observed threat actors targeting C-suite, accounting, and payroll employees in these credential phishing attacks.

Brand imitation continues to grow

Many credential phishing scams start by imitating a well-known brand, enterprise cloud service, or government entity. They require users to click on a link that leads to a fake login site and lures them into revealing their credentials. Another common tactic is to spoof trademarks from trusted brands in the phishing "lure." It's no surprise that well-known brands in banking, telecommunication, and technology are frequently spoofed to deceive recipients. Based on our Office 365 telemetry, the top five spoofed brands are Microsoft, UPS, Amazon, Apple, and Zoom. The websites that host the fake login pages are composed of phishing kits deployed by cybercriminals. For instance, there's a "Microsoft"-branded phishing kit that "advises" recipients that they've received a voicemail. Each phishing kit also includes a "drop account," or email account set up by the criminal to receive credentials provided by the unsuspecting victim. There are literally thousands of different phishing kits sold on the dark web today. DCU routinely analyzes these kits to develop insights to thwart ongoing and future credential harvesting and BEC campaigns targeting our customers.

Top 5 spoofed brands

Microsoft
UPS
Amazon
Apple
Zoom

BEC-targeted industries

Given the time dedicated to research and post-compromise reconnaissance, cybercriminals know which industries frequently wire large payments. Based on our BEC investigations, the top 10 most targeted industries for BEC attacks are accounting and consulting, wholesale distribution, IT services, real estate, education, health care, chemicals, high tech and electronics, legal services, and outsourced services. Cybercriminals also understand business and payment cycles. As a result, DCU sees phishing campaigns directed at specific industries and BEC attacks timed to increase the ROI.

Top 10 targeted industries for BEC attacks

Accounting & Consulting
Wholesale Distribution
IT Services
Real Estate
Education
Healthcare
Chemicals
High Tech & Electronics
Legal Services
Outsourced Services

Our strategy

Given the enormity and complexity of the credential phishing and BEC issue, the DCU, together with security engineering teams across Microsoft, is executing a strategy focused on four functions: disrupt, deter, strengthen, and communicate.

Disrupt: The disrupt function is to analyze and map the technical infrastructure utilized by cybercriminals to launch BEC attacks. The infrastructure includes phishing URLs, phishing kits, collection accounts, drop accounts, and homoglyph domains. DCU takes action to take down these aspects of the criminal technical infrastructure through established notice and take-down procedures or through judicial intervention. Comprehensive disruption mitigates the likelihood of successful BEC attacks and increases the cost to cybercriminals, reducing the ROI and incentive for cybercrime. The data produced by disruption efforts informs allocation of resources and other decisions to drive strategic decision-making to yield the highest impact.

Deter: The deter function focuses on identifying, investigating, and developing cases for criminal referral to law enforcement agencies throughout the world, or for civil action as appropriate. Successful enforcement actions have a powerful impact by acting as both a specific and general deterrence against future cybercrime.

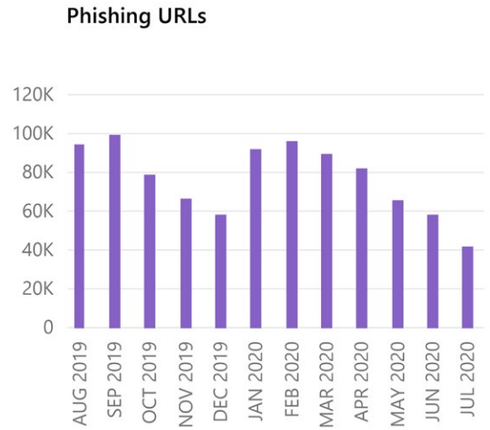
Strengthen: The evidence and insights gained from DCU case development are shared with Microsoft product security engineering teams and help inform advancements in product security and technical measures to defeat phishing and BEC.

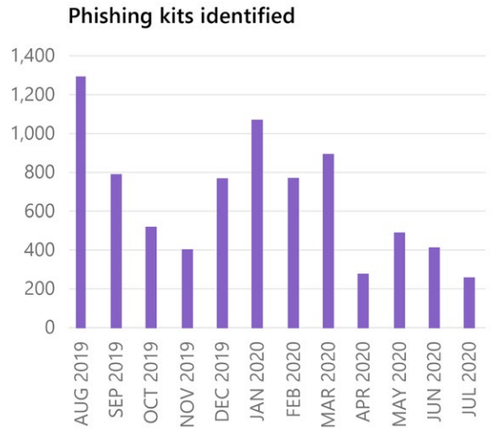
Communicate: Finally, DCU communicates our findings internally and externally to provide guidance to customers on how to avoid falling victim to BEC scams.

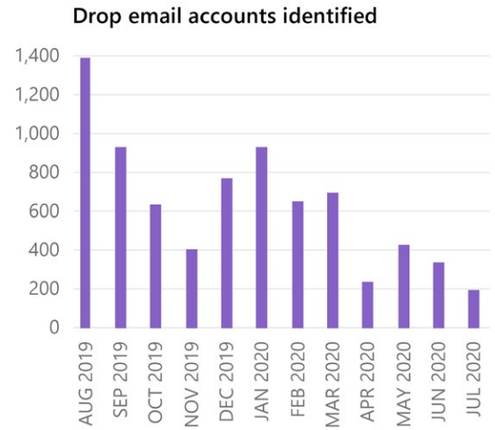
Learn more:

Top 10 ways to secure your data

Phishing URLs and phishing kits target the Microsoft brand. Drop email accounts refer to the bad actor emails to which the stolen credentials (that are the result of successful phishing) are sent.







Prevention checklist:

- Set up multi-factor authentification
- Stop auto-forwarding for email
- Train your users

Training employees to combat phishing

Every year, we provide more than 200,000 employees and external staff with the experience of being phished, along with prevention education and reporting guidance. We then follow up with users who were susceptible through quarterly simulations to help them better identify key indicators in the future. In addition, full-time employees are required to complete annual and updated Security Foundations training. These levers, combined with ongoing awareness campaigns, are what drive positive user behavior.

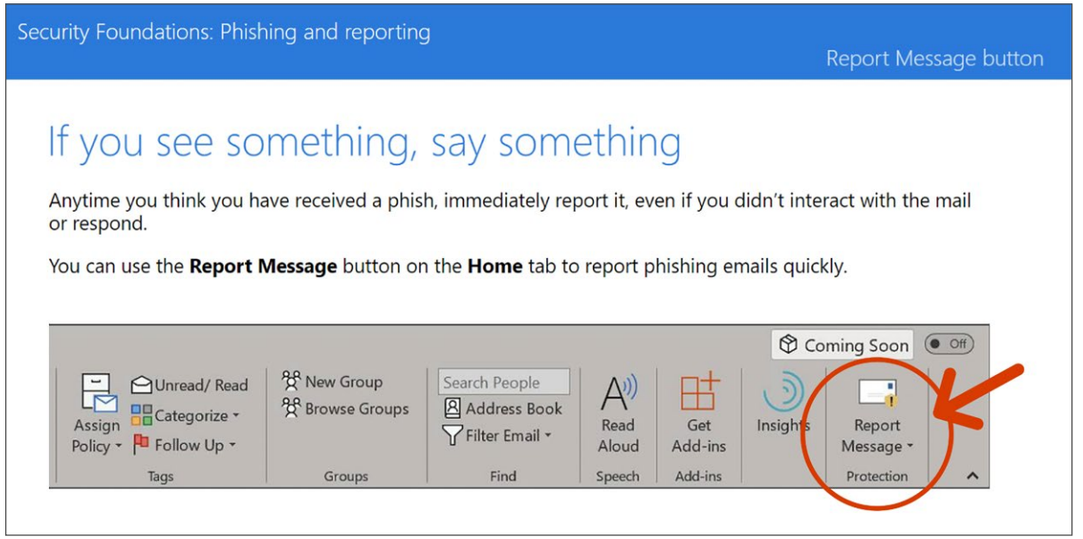
Our employees have a solid awareness of cybersecurity risk behaviors, but they aren't impervious to savvier tactics. This means our prevention education programs must honor what they already know by shifting away from general awareness content and a one-size-fits-all approach toward a more tailored and data-driven strategy. This focused emphasis ensures employees remain engaged with our content and training and that we're focused on experiences our employees will likely encounter in our environment.

Our strategy has moved away from a compliance-centric mindset and toward an adult learning and skills-building model. This model guarantees we build skills progressively over time to help us reach the desired behavioral outcomes and reduce risk. We use a series of micro-learnings, or shorter online trainings, launched throughout the year. This cadence allows employees to absorb the information in manageable pieces, keeps our employee engagement high, and helps solidify their skills with knowledge reinforcement.

Annually we're seeing continual improvements in employees' abilities to recognize and report phishing as a result of employee training at Microsoft.

Phishing is, of course, just one element of an overall security awareness program. Organizations should establish an awareness program that takes a holistic approach, utilizing multiple levers with data and telemetry at its center. Determining what areas of behavior are driven by a lack of knowledge will best be addressed with a "training first" approach. Areas where employees have the knowledge but are still not displaying desired security behaviors should be addressed through other efforts, like targeted campaigns, leadership messaging, outreach events, and a closer look at process and procedures. This is why telemetry—and knowing what to measure—is so critical to the effectiveness of a program.

Employee phishing training module



Learn more about phishing and business email compromise:

Microsoft shares new threat intelligence and security guidance during global crisis 2020/04/08

Protecting against coronavirus-themed phishing attacks 2020/03/20

Protect yourself from phishing schemes and other forms of online fraud

Ransomware: A high-impact, human-driven threat

Historically, cyberattacks were seen as a sophisticated set of actions targeting particular industries, which left the remaining industries believing they were outside the scope of cybercrime, and without context about which cybersecurity threats they should prepare for. Ransomware represents a major shift in this threat landscape, and it's made cyberattacks a very real and omnipresent danger for everyone. Encrypted and lost files and threatening ransom notes have now become the top-of-mind fear for most executive teams.

INTERPOL issued a warning to organizations at the forefront of the global response to the COVID-19 outbreak because they've also become targets of ransomware attacks designed to lock them out of their critical systems in an attempt to extort payments.[4]

Ransomware's economic model capitalizes on the misperception that a ransomware attack is solely a malware incident, whereas in reality ransomware is a breach involving human adversaries attacking a network.

For many organizations, the cost to rebuild from scratch after a ransomware incident far outweighs the original ransom demanded. With a limited understanding of the threat landscape and how ransomware operates, paying the ransom seems like the better business decision to return to operations. However, the real damage is often done when the cybercriminal exfiltrates files for release or sale, while leaving backdoors in the network for future criminal activity—and these risks persist whether or not the ransom is paid.

Paying a ransom doesn't remove the attacker

The rise in popularity of ransomware—and its subsequent rise in fame—has led to a cybersecurity narrative that focuses on the intricacies of the ransomware payload itself and the novelty of encryption methods utilized. This makes it seem as if the ransomware were appearing on networks through characteristics of the malware and that defenses should focus on preventing encryption. However, this approach often fails to address the root problem because it ignores the human actors behind the threat, the specificity of their targets, and that access to their networks might already be compromised.

Understanding and fixing the fundamental security issues that led to the compromise in the first place should be a priority for ransomware victims.

Microsoft's Detection and Response Team (DART) engages with customers around the world, helping to protect and harden against attacks before they occur, as well as investigating and remediating when an attack has occurred. DART has observed that ransomware continues to be the most common reason behind our incident response engagements from October 2019 through July 2020. More threat actors are using open-source tools like Cobalt Strike, MimiKatz, ProcessHacker, and LaZagne to initiate the attack that ultimately delivers ransomware payload. The malware used in these incidents is usually delivered through spear phishing or by exploiting public-facing applications.

Human-operated ransomware

Human-operated ransomware is sometimes referred to as "big game ransomware," a term that implies cybercriminals select specific networks for their value proposition and then hunt for entry vectors. This approach has been the exception, not the rule, in most major ransomware attacks in the past year. Cybercriminals perform massive wide-ranging sweeps of the internet, searching for vulnerable entry points. Or they enter networks via "commodity" trojans and then "bank" this access for a time and purpose that's advantageous to them. For example, cybercriminals used Dridex (a strain of banking malware that leverages macros in Microsoft Office) to gain initial access to networks, and then ransomed a subset of them with the DoppelPaymer ransomware during the 2019 Christmas holiday season. As another example, cybercriminals exploited vulnerabilities in VPN and remote access devices to gain credentials, and then saved their access to use for ransoming hospitals and medical providers during the COVID-19 pandemic. In these attacks, the cybercriminals actively make decisions as they go, controlling each attack step based on the configurations they encounter in the network. They decide which data to exfiltrate, which

[4] *https://www.interpol.int/News-and-Events/News/2020/Cybercriminals-targeting-critical-healthcare-institutions-with-ransomware*

persistence mechanisms to use for future access to the network, and ultimately, they even decide which ransomware payload to deliver. This scenario is far from a commodity or automated threat. Microsoft refers to this class of attacks as "human-operated ransomware" to reflect the current threat landscape more accurately.

Ransomware attacks are often described in terms of their payload. However, Microsoft has observed multiple ransomware payloads being deployed by cybercriminals using the same infrastructure for concurrent campaigns. The selection of which payload and tools used was largely dependent on the terrain the cybercriminals landed in; choices were based on which security tools were present, whether the network had good cybersecurity basics in place, and which data the cybercriminals wanted to exfiltrate from the network.

Payloads can also be varied by cybercriminals to avoid attribution. If a certain ransomware has been reported in the news and there's heightened awareness of it, switching to a different payload (and sometimes just ransom note) can reduce the pressure on cybercriminals and their concern about the possibility of getting disrupted. In April 2020, Microsoft observed a prolific ransomware criminal responsible for more than 100 major incidents suddenly switch to using the infamous WannaCrypt payload at the end of their typical attack pattern, after attacking hospitals and healthcare organizations during the COVID-19 crisis.

Ransomware criminals are intimately familiar with systems management concepts and the struggles IT departments face. Attack patterns demonstrate that cybercriminals know when there will be change freezes, such as holidays, that will impact an organization's ability to make changes (such as patching) to harden their networks. They're aware of when there are business needs that will make businesses more willing to pay ransoms than take downtime, such as during billing cycles in the health, finance, and legal industries. Targeting networks where critical work was needed during the COVID-19 pandemic, and also specifically attacking remote access devices during a time when unprecedented numbers of people were working remotely, are examples of this level of knowledge.

Once inside a network, this proficiency in understanding protective and detective controls continues to contribute to the success of the cybercriminals. Through reconnaissance they'll select machines with no or poorly configured antivirus software to perform most of their actions, modifying their techniques if they sense they might be detected. Unfortunately, there are also numerous examples of situations where cybercriminals simply performed their attacks as they wished, with poor cyber hygiene leaving no blocking controls in their way. This can happen in spite of multiple stages of their attack causing detections in antivirus and endpoint detection and response products, signifying cybercriminals' understanding of the challenges modern IT departments and SOCs face in rapidly triaging and responding to fast-paced attackers.

In some instances, cybercriminals went from initial entry to ransoming the entire network in less than 45 minutes.

While human-operated ransomware campaigns vary in their tools and ultimate choice of ransomware payload, all of them need to operate in a similar manner to effectively enter a network and move laterally to achieve their goals. While initial entry vectors can vary, such as through unpatched vulnerabilities, the biggest factor leading to success in moving through multiple systems and deploying ransomware is the cybercriminal's ability to gain access to highly privileged account credentials. Cybercriminals rely on off-the-shelf tools used for systems administration or security testing and built-in tools to move from machine to machine, but they need administrative credentials, such as those of a domain administrator, to gain access. To deploy ransomware across an entire organization, cybercriminals must capture a credential and a system with the rights to do this. Domain administrator accounts are often used for their ability to utilize Active Directory policies and file shares intended for software distribution to maliciously deploy devastating ransomware payloads.

Over 70% of human-operated ransomware attacks in the past year originated with Remote Desktop Protocol (RDP) brute force.[5]

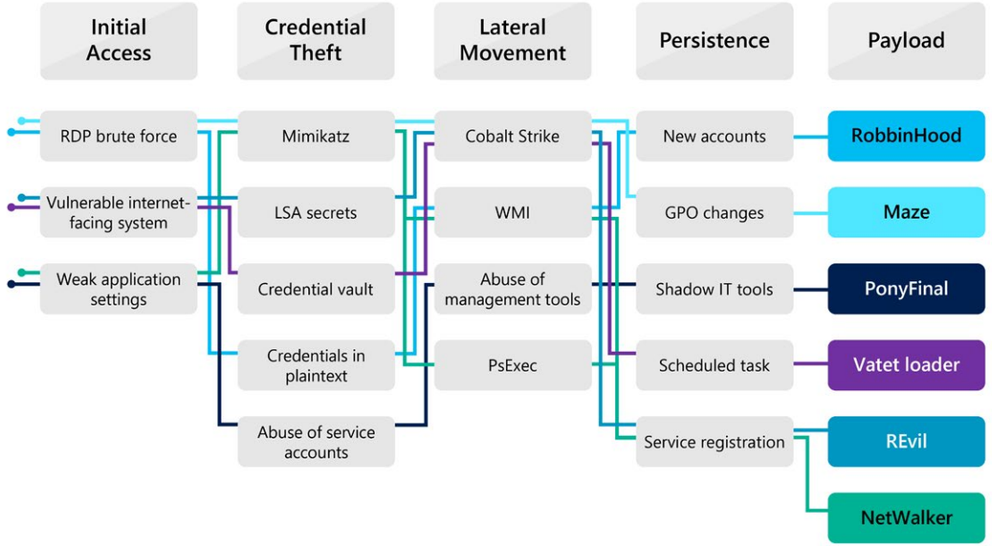
By narrowing the focus for IT departments to the overlapping techniques used by attackers regardless of which payload they deploy, defensive strategies can be moved from a reactionary detection-only option, which changes with every variation in ransomware, to simple steps that provide significant ROI across multiple classes of attacks.

Learn more:

Human-operated ransomware attacks: A preventable disaster 2020/03/05

Ransomware groups continue to target healthcare, critical services; here's how to reduce risk 2020/04/28

Ransomware operations attack pattern detail observed by Microsoft Threat Protection intelligence in early 2020. Payloads might vary, but mitigations apply across all varieties.



Two steps to follow:

Know your perimeter

RDP brute force is by far the most common entrance point

"Abandoned" systems still online and discoverable

Use the tools the attackers use: Shodan.io can help you understand your attack surface

Practice credential hygiene

Domain admin level accounts should not be used to run services, scheduled tasks, or permitted to log on to worksatations

Randomize local admin passwords

Embrace "logical segmentation"—vulnerable systems should not have credential overlap with the rest of the network

[5] *https://www.bleepingcomputer.com/news/security/fbi-says-140-million-paid-to-ransomware-offers-defense-tips/*

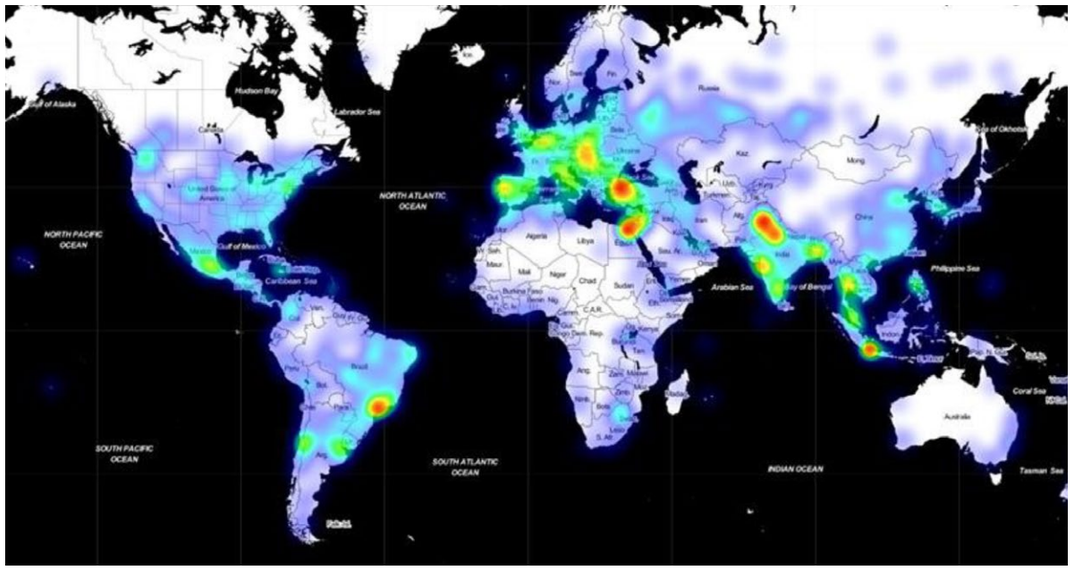
Malware: The evolution of banking trojans

For busy security operations centers, keeping up with the volume of alerts and data coming in from various security tools and monitoring platforms can be difficult. This so-called alert fatigue is a huge factor in organizations not being able to respond quickly and accurately enough to security incidents—a factor cybercriminals are acutely aware of. When faced with deciding which alerts to prioritize, many SOCs will prioritize threats perceived as targeted (such as nation state–related threats) over those which are "commodity," a term used to describe malware that isn't customized or intended for use on specific targeted networks. This assumption that commodity threats are less impactful to a business, combined with the security industry's naming convention for threats, leads to confusion that cybercriminals of all skill sets will use to their advantage. From nation state attackers utilizing off-the-shelf tools intended for Red Teams to major criminal operations pivoting their behaviors in well-known malware families, the desire to blend in and not be remediated has turned some of these assumptions upside down. Nowhere was this more impactful in the past year than in the category of banking trojans.

Judging malware by its family or detection name is a lot like judging a book by its cover—it might not tell you what the malware is capable of. Banking trojans, intended to steal credentials from online banking and finance sites, have been a consistent threat in the landscape over the past decade. The nature of many banking trojans was always modular and extensible, with additional modules and capabilities deployed or downloaded after the initial infection. In the past few years, the potential of this capability for attacks against enterprise environments, where the banking trojan malware might not be immediately triaged or removed, was realized with devastating success by several major banking trojans. The two banking trojans most infamous for adopting this new model are Emotet and Trickbot.

Originally intended to steal banking data, components in Emotet and Trickbot reporting IP addresses and DNS ranges back to the operators revealed just how many major networks their malware was installed on and checking in from. In early efforts to capitalize on this network access, team-ups of malware payloads were seen. Examples include Emotet's leading to Trickbot infections that utilized the Eternal Blue vulnerability for lateral movement and Emotet's downloading second-stage trojans with credential theft capabilities such as Qakbot or IceID. This piggybacking has led to disastrous results for organizations, as these

Trickbot banking trojan prevalence, April 2020

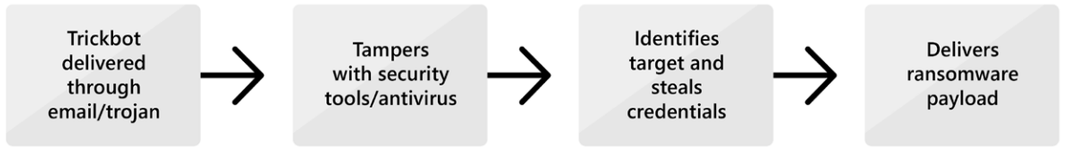


second-stage malware payloads used brute force attacks against Active Directory passwords, causing massive downtime and network outages. Trickbot reached the most devastating stage in this evolution in 2019 when the operators began to utilize the Trickbot install base to deploy the Ryuk ransomware.

Taking advantage of Trickbot's prevalent install base, operators select networks that seem valuable and advantageous to them at the time: the ones that will pay a high ransom or have interesting data. They then use the Trickbot implant to download additional tools, such as Cobalt Strike, and transform into a human-operated ransomware campaign.

Banking trojans have succeeded by being underestimated and by using their breadth of infections to their advantage, selectively targeting networks for follow-on ransomware, data exfiltration, or sale of access on the dark web. Organizations can help with alert fatigue and bombardments of similar threats by focusing on the behaviors of the malware and delivery mechanisms, and by mitigating the issues that allow entry into their networks. Most banking trojans are delivered via email attachments and especially via macro-enabled Microsoft Office documents.

Simple solutions for blocking malicious macros, such as using trusted locations, can dramatically decrease an attack surface and reduce the risk of dangerous post-infection activities.



Learn more:

New action to disrupt world's largest online criminal network 2020/03/10

An inside look at the global battle with botnets 2020/07/24

Promoting trusted information

There's so much COVID-19 commentary circulating online that even those with the keenest eye for misinformation can sometimes find it tricky to separate fact from fiction.

That's why Microsoft and LinkedIn, along with other technology companies are working to keep millions of people connected while also combating fraud and misinformation about the pandemic. We're giving prominence to authoritative content across platforms and coordinating with government healthcare agencies around the world to share critical updates.

Learn more: *Data, supplies, community: How Microsoft is supporting efforts to combat COVID-19*

How is Microsoft highlighting content from credible sources?

We're taking new steps across our services, including Bing, LinkedIn, Microsoft News, and Microsoft Advertising. On Microsoft News and LinkedIn, we're creating curated resources based on official guidance from organizations such as the WHO and the CDC. We're also aggregating data on case locations,[6] infection prevention, and more from authoritative sources and presenting it in an accessible, easy-to-use way across Microsoft News[7] and Bing.[8]

Bing users in multiple countries can access reliable health information through public service announcements for some COVID-19 queries. They can also find reliable information in answers near the top of search results pages and in sidebar windows. In the United States—and soon in other countries—information is provided on testing protocols and locations for anyone who thinks they have symptoms of the virus.

Bing is also prioritizing trusted news sources and piloting algorithmic defenses to help promote reliable information about COVID-19.

On LinkedIn, a global team of more than 65 experienced journalists is curating news and perspectives about the coronavirus pandemic from trusted sources. The team is spotlighting this coverage in a "Special Report: Coronavirus" box on the LinkedIn homepage, with relevant and accurate stories. They're also reaching audiences through the Daily Rundown news summary, which reaches 46 million people, in 96 countries, in nine languages. When users search for coronavirus-related terms or hashtags, they'll see trusted information modules at the top of the results page.

How is Microsoft helping to surface credible reporting on the crisis?

Microsoft News has COVID-19 information hubs in 39 markets across the globe. An experienced team edits content from more than 4,500 trusted news brands, including *The New York Times*, Reuters, and the *Wall Street Journal*. These hubs also contain links to official tools and information from sources like the WHO and the CDC. When people search Bing for "coronavirus updates" and related queries, it will point to these hubs in some news answers.

Learn more: *Microsoft to join White House-led consortium to fight COVID-19*

What measures is Microsoft taking in relation to advertising?

Microsoft's Sensitive Advertising policy and LinkedIn's Advertising policies prohibit ads that capitalize on the pandemic and company pages that improperly sell medical supplies and solutions. These policies allow Microsoft and LinkedIn to remove or limit advertising and company pages in response to a sensitive tragedy, disaster, death, or high-profile news event and are being applied to block ads related directly to COVID-19. In certain instances, advertising related to the pandemic will only be allowed from trusted sources. Any advertising that exploits the coronavirus crisis for commercial gain, spreads misinformation, or might pose a danger to users' safety is prohibited.

What is the role of the Digital Crimes Unit at Microsoft?

Our Digital Crimes Unit is also analyzing data regarding cyberthreats associated with malware, phishing, and fraud. Between late March and the beginning of July 2020, we've reported 13,971 potentially malicious COVID-19-themed domains and 23,123 URLs to the proper authorities so that they can be taken down and, where possible, the individuals behind them prosecuted.

Who else is Microsoft working with?

Teams within Microsoft Research are collaborating closely with Professor Jacob Shapiro, director of the Empirical Studies of Conflict Project at Princeton University, to characterize the types and extent of misinformation and disinformation narratives online related to COVID-19. This work helps our researchers, product teams, and industry partners understand the global information environment our customers are exposed to.

[6] *https://blogs.bing.com/search/march-2020/Stay-informed-on-the-coronavirus-pandemic-with-Bing-and-Microsoft-News*
[7] *https://www.msn.com/en-gb/news*
[8] *https://www.bing.com*

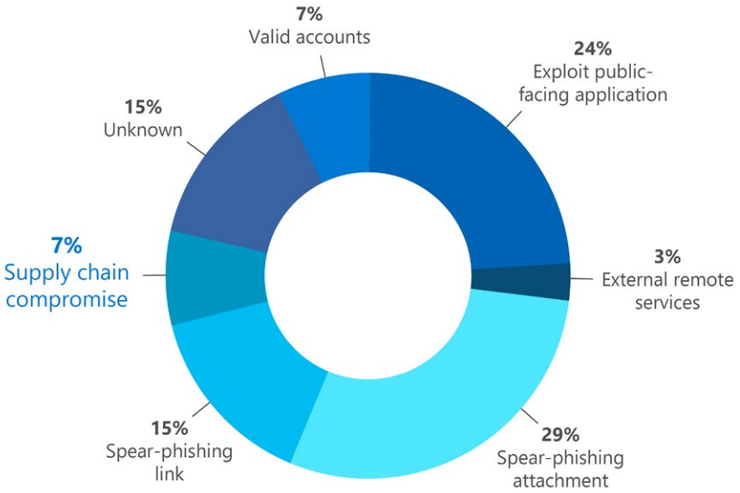
In focus: Supply chain security

For years Microsoft has been tracking threat actors who use supply chain compromise as an entry point for attacks. Supply chain attacks are particularly insidious because they take advantage of the trust that users and organizations place in the hardware, software, and third-party services they use.

The increased number of supply chain attacks over the past few years has become an important topic in many cybersecurity conversations and is a growing source of concern across the global supply chain. The past 12 months have been an unprecedented time of focus on supply chain security, given the acceleration of interdependencies resulting from changes in global remote workforces in response to COVID-19, as well as new and evolving regulations in the United States and Europe.

In this section, we focus our observations and recommendations for supply chain security on third-party services, open-source software, and IoT hardware and conclude with a look at changes to the regulatory landscape.

DART incident response initial attack vector (June 2019–March 2020)



Third-party services

Through its engagements in assisting customers who have been victims of cybersecurity intrusions, the Microsoft Detection and Response Team has observed an uptick in supply chain attacks between July 2019 and March 2020. These attacks include targeting IT service providers to get a foothold in their customers' systems. Although there was an increase, supply chain attacks represented a relatively small percentage of DART engagements overall.

As organizations grant widespread administrative or system access to external entities, managed service providers (MSPs) and IT outsourcers become high-value targets for attackers. All too often, IT outsourcers or MSPs aren't threat modeled with the same diligence as internal systems. Over the past year, there have been numerous high-profile compromises of MSPs to deliver ransomware, exfiltrate data, or provide financial gain to attackers in some way. These attacks follow a common pattern of entering the service provider's network via phishing email or through a network weakness with RDP brute force, and then moving laterally to gain administrative privileges. Once the attackers have gained full access to the service provider, they can access the customers via the same legitimate support channels and accounts used for remote support or software maintenance.

To limit the risk of attack, organizations should vet their service providers to ensure they follow best practices for least privilege access on accounts and services. Access to the network for support should be monitored and secured via multi-factor authentication (MFA) and just-in-time access.

Learn more:

Guarding against supply chain attacks—Part 1: The big picture 2019/10/16

Guarding against supply chain attacks—Part 2: Hardware risks 2020/02/03

Guarding against supply chain attacks—Part 3: How software becomes compromised 2020/03/11

Defending the power grid against supply chain attacks—Part 1: The risk defined 2020/02/18

Defending the power grid against supply chain attacks—Part 2: Securing hardware and software

Third-party and open-source software

Given the diversity of solutions and services provided by our third-party ecosystem, Microsoft deploys flexible methodologies. For example, we ensure that licensors of commercial third-party software adhere to the fundamental tenets of our Security Development Lifecycle.

We recognize that open-source software requires a different approach. We apply security controls to the software itself, vetting it by using a combination of automated tools and manual processes before allowing it to be used in a product or service.

Modern software projects are heavily dependent on open-source software, from operating systems to widgets, and from back-end data analysis to front-end user interfaces. There are well over 2 million[9] open-source components available for developers to use, and according to the GitHub State of the Octoverse report (2019), the average software project depends on more than 200 other components. Open-source software has provided great benefits, allowing developers to focus more on building software applications and less on reinventing common components.

However, this pervasiveness has attracted the attention of attackers, who in recent years have increasingly turned their focus to the open-source software supply chain. While traditional security vulnerabilities have always affected the software supply chain, attackers can now target an ecosystem impacting multiple users at once, with more damaging results. For example:

- An attacker hijacks the account of the publisher of a popular package and publishes a malicious update, such as in the case of the npm bb-builder package, which was discovered a year after it was compromised. Unfortunately, only 12.84% of GitHub accounts and 9.3% of npm maintainers use MFA to protect their accounts, which is one of the most effective ways to mitigate the risk of account takeover. Thankfully, this percentage is higher for well-established open-source projects: 52% of maintainers have enabled two-factor authentication for open-source projects with more than 100 contributors.
- An attacker publishes a package with a name similar to a popular package and waits for some developers to accidentally type the wrong package name (known as "typosquatting"), obtaining the attacker's malicious version. Examples include the python3-dateutil PyPI module and the 725 malicious packages from the RubyGems ecosystem that mined crypto-currency.
- An attacker publishes a malicious package and waits for developers to install and use it. Examples include the npm 1337qq-js and m-backdoor modules.
- An attacker finds credentials, API keys, or other secrets that were accidentally published by an open-source project author. Each month, 8% of all public GitHub repositories are found to contain a potential secret, with many of these automatically detected and revoked.

At Microsoft, we use open-source software extensively in our products and services and have built a program that uses people, process, and technology to identify and manage security risk at scale. At the center of this program is a system that automatically identifies open-source software components and performs certain checks to ensure they're free of known security defects. This is an essential first step to using open-source safely. It enables us to accurately identify the right open-source components to use and understand whether they contain vulnerabilities. It's important that this process be automated as much as possible. Fortunately, there are many tools, commercial and open source, that can help, including GitHub's Dependency Graph and Dependabot alerts. On top of this, we add processes for manually reviewing critical open-source packages and scanning millions of open-source components for signs of compromise.

Over the past year, we've found that nearly 5% of the open-source packages used at Microsoft had at least one reported security vulnerability, and about a quarter of those reported security vulnerabilities were categorized as high risk. Automation is essential to promptly alerting affected engineering teams and tracking remediation activities. According to GitHub, automating alerts in vulnerable dependencies can reduce the average time to remediate by over 75%. As open-source software is a growing and increasingly critical part of modern products and platforms, it should be held to the same standards as proprietary software.

Learn more:

Collaborating to improve open source security—How the ecosystem is stepping up - RSA Conference 2020/02/28

Microsoft Security Engineering: Open source security

Threats, Risks, and Mitigations in the Open Source Ecosystem, Open Source Security Coalition

[9] *Data for this from modulecounts.com*

IoT security insights from honeypots

There were 26.66 billion active Internet of Things (IoT) devices in 2019, and it's estimated that by 2022, there will be 50 billion consumer IoT devices worldwide.[10] While manufacturers of IoT devices have a responsibility to design secure products, this rapid proliferation has made these products appealing targets for a growing volume of cyberattacks.

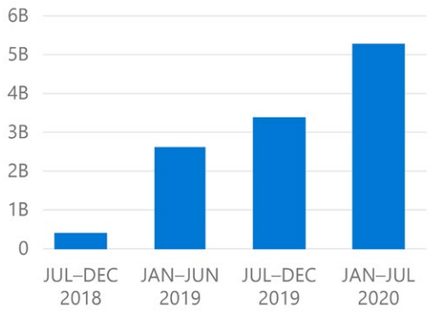
To gather intelligence about attacks on remote terminals, routers, and IoT devices, Microsoft Threat Intelligence Center (MSTIC) deploys honeypots (intended to mimic likely targets of cybercriminals) over Azure and a range of other cloud providers. The honeypots engage attackers to interact with them using various known protocols. Security teams across the company analyze this data to discover new trends and emerging threats.

The IoT security research team is one of the teams that processes the data, reverse-engineers malicious samples, and helps to draw end-to-end attack scenarios. This work helps us discover trends and better understand the risks of using and deploying IoT devices.

Summary of findings from the IoT security research team:

- IoT threats are constantly expanding and evolving. Honeypot data in the first half of 2020 indicate an approximate 35% increase in total attack volume compared to the second half of 2019.
- Vulnerability management plays a crucial role in securing IoT assets, because when a new vulnerability is discovered it's rapidly adopted by malicious actors. These actors take advantage of the time it takes to patch and update the firmware. Various scanners and passive network-based profiling technologies make this malicious capability achievable and easy to deploy.

Total attacks on honeypots



Analysis details:

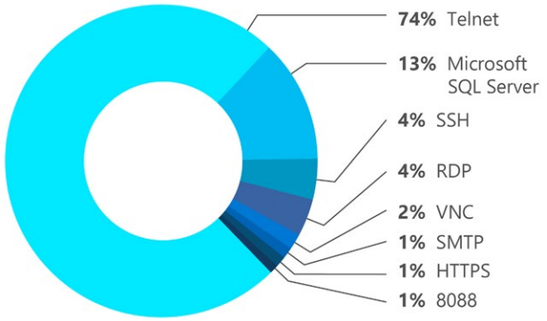
- **The most abused protocols we see are Telnet,** MS SQL, and Secure Shell (SSH), with RDP and Virtual Network Computing (VNC) also being notable.
- The most common attack scenario includes using default credentials to obtain a shell access to download and execute malware. All kinds of devices are being targeted; honeypots exposed malware for many types of architectures, including but not limited to x86, MIPS, MIPSEL, SH4, ARM4/5/6/7, and SPARC.
- The default credentials are associated with a variety of different products such as routers, remote terminals, IP cameras, and multimedia systems.
- Our analysis indicates three typical infiltration techniques:

  Misconfiguration: Use of default passwords and internet-facing listening ports to get remote shell access to the devices.

  Supply chain: In one analysis, an IP camera module manufacturer introduced remote technical support capabilities in the default firmware. These capabilities were reverse-engineered and used by attackers to gain a remote root shell on the affected devices. Because the firmware was implemented without modifications by dozens of IP camera vendors, millions of devices were potentially exposed. Honeypot data indicate a campaign that's actively trying to exploit this vulnerability.

  Vulnerability exploitation: Malware is weaponizing well-known exploits to propagate to new devices. Analyzed samples include exploits for various router vendors, IP cameras, and multimedia systems.

Abused protocols in January–March 2020



Recommendations:

Map IoT assets and apply security policies to reduce the attack surface and the potential risk of implementing IoT technologies.

Make sure to use a different network for IoT devices and be familiar with all exposed interfaces. Previous analyses at Microsoft have shown that even the IoT vendor might be unaware of remote administration capabilities that might be introduced in a module's firmware (which is often created by another manufacturer).

[10] *https://cybertechaccord.org/iot-security/*

IoT/OT Security insights from CyberX "2020 Global IoT/ICS Risk Report"

Operational technology (OT) networks used in industrial and critical infrastructure environments have traditionally been air-gapped from corporate IT networks and the internet, but digital transformation has increased both the connectivity and number of devices in these environments, leading to higher risk.

Further increasing risk, many of the legacy IoT/OT protocols and embedded devices in these environments were designed years ago—lacking modern controls such as encryption, strong authentication, and hardened software stacks—while the OT networks themselves are often flat, unsegmented, and lacking zero-trust policies.

Adversaries targeting this expanded attack surface can cause substantial corporate impact, including safety and environmental incidents, costly production downtime, and theft of sensitive intellectual property such as information about proprietary formulas and manufacturing processes.

OT cyber risk was recently highlighted in an alert (AA20-205A) from the U.S. Cybersecurity and Infrastructure Security Agency (CISA) and the National Security Agency (NSA). The alert was titled "NSA and CISA Recommend Immediate Actions to Reduce Exposure Across Operational Technologies and Control Systems." In the alert they stated that "Over recent months, cyber actors have demonstrated their continued willingness to conduct malicious cyber activity against critical infrastructure (CI) by exploiting internet-accessible operational technology (OT) assets."

CyberX is an IoT/OT cybersecurity company recently acquired by Microsoft to help clients gain broader and deeper visibility into their IoT/OT risk. In its "2020 Global IoT/ICS Risk Report," CyberX analyzed data collected from more than 1,800 production industrial control system (ICS) networks, across diverse industries including energy utilities, manufacturing, pharmaceuticals, chemicals, and oil and gas. The data was analyzed using CyberX's passive, agentless monitoring technology and proprietary Network Traffic Analysis (NTA) algorithms.[11]

Some of the key insights from this report include:

- 71% of sites have unsupported Microsoft Windows systems, such as Windows 2000, Windows XP, and Windows 7, that no longer receive regular security patches from Microsoft, making them especially vulnerable to ransomware and destructive malware. Even excluding Windows 7 systems which became unsupported in January 2020, the percentage of sites with unsupported Windows systems is still quite high at 62%.
- 64% of sites have unencrypted passwords traversing their networks, making it easy for adversaries to compromise systems simply by sniffing the network traffic.
- 66% of sites aren't automatically updating Windows systems with the latest antivirus definitions.
- 54% of sites have devices that can be remotely accessed from internal networks by using standard management protocols such as RDP, SSH, and VNC, enabling attackers to pivot undetected from initial footholds to other critical assets. For example, during the TRITON attack on the safety systems in a petrochemical facility, the adversary leveraged RDP to pivot from the IT network to the OT network in order to deploy its targeted zero-day malware.

[11] *The analysis was performed on an anonymized and aggregated set of metadata with identifying information removed. Rigorous attention was paid to preserving the confidentiality of sensitive customer information.*

Recommendations for securing IoT/OT Networks

Microsoft recommends the following actions to better secure and manage the risk associated with IoT and OT devices:

Avoid exposing IoT and OT devices directly to the internet. Avoid exposure or create custom access controls to limit exposure. This is true of incoming and outgoing network controls to protect against intrusion and exfiltration.

Map the digital terrain. Create a current map of your network and all the IoT/OT devices connected to it. Understand how devices are communicating with each other and other parts of your corporate network.

Focus on mitigating risk for your "crown jewel" processes. Use threat modeling (via automated attack simulation techniques, Red Teaming, and/or tabletop exercises) to predict the most likely paths adversaries would take to compromise your most critical or "crown jewel" processes. These are the processes in which compromise would cause major financial impact owing to production downtime, major safety incidents, or theft of intellectual property. Then focus on mitigating risk to those specific attack paths by addressing configuration errors, patching, network segmentation, or compensating controls such as continuous network security monitoring.

Implement Zero Trust IoT/OT strategies. Use a separate or segmented network for IoT and OT devices. Consider more granular segmentation policies within the OT network itself, across multiple layers of the Purdue model. Audit and revalidate identities and credentials with authorized access to IoT devices, users, and processes.

Require approval and cataloging of any IoT/OT devices. Implement a policy for all devices running in your environment.

Centralize asset/configuration/patch management. Conduct routine asset discovery and configuration/patch audits against deployed IoT and OT devices.

Remove IT/OT silos. Modern attacks often cross IT/OT boundaries. For example, in the TRITON attack on the safety systems in a petrochemical facility, the adversary compromised OT remote access credentials from a Windows system on the IT network before pivoting to the OT network. Additionally, it's not always clear who's responsible for the security of the OT network—is it the IT security organization, the OT organization, or various third-party organizations that manage and maintain automation equipment in these facilities? Ensure the lines of responsibility are clear, and also that the corporate SOC has continuous visibility into potential OT security threats—via their security information and event management (SIEM)/Security Orchestration, Automation and Response (SOAR) system—so they can quickly detect multistage IT/OT attacks and stop them before they cause harm.

Continuously monitor for unusual or unauthorized behavior. It's no longer sufficient to rely on static indicators of compromise to detect attacks. Adversaries are increasingly using "living off the land" tactics in the OT domain, just as in the IT domain. This change in strategy requires vigilant monitoring for unusual or unauthorized behavior such as new remote access connections, code or configuration changes made to programmable logic controllers, printers browsing SharePoint sites, or even new devices being connected to the network for the first time.

Plan for incident response. Define policies for isolation of IoT and OT devices, preservation of device data, ability to maintain logs of device traffic, and capture of device images for forensic investigation.

Don't forget third parties. If your devices are deployed/managed by a third party, include explicit terms in your contracts detailing security practices to be followed and audits that report security status and health of all managed devices. Ensure third parties are using secure remote access (VPNs, strong authentication, and rotating passwords) to access your network. Where possible, define service-level terms in IoT and OT device vendor contracts that set a mutually acceptable window for investigative response and forensic analysis to any compromise involving their product.

Create a manageable OT Windows upgrade schedule. IIoT/OT systems are typically more difficult to upgrade than corporate IT systems. Many IoT/OT networks run 24x7 and have limited maintenance windows. Windows systems often host legacy OT applications that would need to be extensively tested or rewritten after an upgrade. Food and Drug Administration–validated systems in the pharmaceutical industry require a new cycle of validation after being upgraded.

While creating a software and hardware upgrade schedule is more difficult for OT than it is for IT, it isn't impossible given the appropriate top-down management attention and resources. The inescapable truth is that no matter how difficult they are to upgrade, legacy Windows systems introduce risk to your organization's people, production, and profits.

If you can't update your Windows systems, ensure you're aware of all legacy Windows systems in your facilities and implement compensating controls such as continuous monitoring to quickly detect when they're being targeted.

Secure your OT Windows endpoints. In the past, automation vendors prohibited end-user organizations from deploying endpoint security controls such as antivirus protection on their systems, but this situation has changed. Ensure your OT Windows endpoints are running modern endpoint protection controls to quickly detect malware and prevent unauthorized applications from running.

Learn more:

CyberX Global 2020 IoT/ICS Risk Report

US CISA/NSA Alert 2020/07/23

Staying safe and smart in the Internet of Things era 2020/05/14

Cyber Tech Accord

Azure IoT security

Corporate IoT: A path to intrusion 2019/08/05

Microsoft logistics team gains supply chain visibility using Azure IoT Central 2020/01/11

Evolving regulations to manage supply chain risk

Generally, a good supply chain risk management strategy is a coordinated, organization-wide effort to identify, monitor, detect, and mitigate threats to supply chain continuity and profitability. Threats to the supply chain include cost volatility, material shortages, supplier financial issues and failures, and natural and manmade disasters.

An increased regulatory focus on supply chain security has led to requests for Supply Chain Risk Management Assessments and the U.S. Department of Defense (DoD) defining a new Cybersecurity Maturity Model Certification (CMMC), intended to improve supply chain security and the cybersecurity posture of the defense industrial base. Released in January 2020, the new CMMC framework and certification will require DoD contractors and subcontractors to obtain third-party certification of their cybersecurity maturity.[12]

The CMMC framework contains five maturity levels starting with basic safeguarding at Level 1, moving to broad protection of Controlled Unclassified Information at Level 3, and culminating with reducing the risk from Advanced Persistent Threats at Levels 4 and 5. Each level is cumulative and designed to provide increased assurance that a contractor can adequately protect certain types of sensitive information.

[12] *The DoD simultaneously released Appendices A-F for the CMMC and an Overview Briefing.*

Machine learning in security

As machine learning takes an increasingly central role in the operations and products of organizations around the world, it's increasingly vulnerable to malicious actors. At the same time, it's becoming an invaluable tool in combatting cybercrime. Here are four trends to consider when defending against attacks on ML systems, and in using ML to defend against alert fatigue and make threat detection more effective and efficient.

Preparing your industry for attacks on machine learning systems

Adversarial machine learning, the name given to efforts to attack ML systems, is a growing reality in the software industry. Google,[13] Microsoft,[14] and IBM[15] have signaled—separate from their commitment to securing their traditional software systems—specific initiatives to secure their ML systems.

In February 2019, Gartner, the leading industry market research firm, published its first report on adversarial machine learning,[16] advising that "Application leaders must anticipate and prepare to mitigate potential risks of data corruption, model theft, and adversarial samples."

There are good reasons for industries to consider securing their ML systems:

1. In the last three years, companies heavily invested in ML (Google, Amazon, Microsoft, Tesla) faced some degree of adversarial attacks[18-21]—a bellwether of the rise of adversarial ML.

2. Standards organizations like ISO[22] and NIST,[23] whose endorsements have been historically sought after in the industry,[24] are forming certification rubrics to assess security of ML systems. Governments are also showing signs of increased appetite for regulations on ML security, with the earliest regulatory activity likely to be in the EU and the United States. A complete checklist has been released in the EU to assess trustworthiness of ML systems, which will potentially form the basis of new regulations.[25]

3. ML is rapidly becoming core to organizations' value propositions (with a projected annual growth rate of 39% for ML investments in 2020),[26] and it's only natural that organizations invest in protecting their crown jewels.

Model poisoning is the top perceived threat against ML for business decision makers.[17]

[13] "Responsible AI Practices." [Online]. Available: https://ai.google/responsibilities/responsible-ai-practices/?category=security; [14] "Securing the Future of AI and ML at Microsoft." [Online]. Available: https://docs.microsoft.com/en-us/security/securing-artificial-intelligence-machine-learning; [15] "Adversarial Machine Learning," July 2016. [Online]. Available: https://ibm.co/36fhajg ; [16] S. A. Gartner Inc., "Anticipate Data Manipulation Security Risks to AI Pipelines." [Online]. Available: https://www.gartner.com/doc/3899783; [17] https://arxiv.org/abs/2002.05646; [18] A. Athalye, L. Engstrom, A. Ilyas, and K. Kwok, "Synthesizing robust adversarial examples," arXiv preprint arXiv:1707.07397, 2017.; [19] J. Li, S. Qu, X. Li, J. Szurley, J. Z. Kolter, and F. Metze, "Adversarial Music: Real World Audio Adversary Against Wake-word Detection System," Advances in Neural Information Processing Systems, 2019, pp. 11 908–11 918.; [20] P. L. Microsoft, "Learning from Tay's introduction," March 2016. [Online]. Available: https://blogs.microsoft.com/blog/2016/03/ 25/learning-tays-introduction/; [21] "Experimental Security Research of Tesla Autopilot," Tech. Rep. [Online]. Available: https://bit.ly/37oGdla; [22] "ISO/IEC JTC 1/SC 42 Artificial Intelligence," January 2019. [Online]. Available: https://www.iso.org/committee/6794475.html; [23] "AI Standards." [Online]. Available: https://www.nist.gov/topics/artificial-intelligence/ai-standards; [24] R. Von Solms, "Information security management: Why standards are important," Information Management & Computer Security, vol. 7, no. 1, pp. 50–58, 1999.; [25] "Ethics guidelines for trustworthy ai," November 2019. [Online]. Available: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai; [26] "2018 AI predictions-8 insights to shape business strategy," Tech. Rep. [Online]. Available: https://www.pwc.com/us/en/advisory-services/assets/ai-predictions-2018-report.pdf

Despite the compelling reasons to secure ML systems, in a survey spanning 28 different organizations, we found that most industry practitioners are yet to come to terms with adversarial machine learning.[27] Of the 28 organizations surveyed, 25 indicated that they don't have the right tools in place to secure their ML systems and are explicitly looking for guidance.

Attacks on machine learning systems

Poisoning attack	Attacker contaminates the training phase of ML systems to get intended result
Model stealing	Attacker is able to recover the model by constructing careful queries
Model inversion	Attacker recovers the secret features used in the model through careful queries

Recommendations:

- Perform threat modeling on ML systems, just as you would security software.
- Prepare your security analysts to identify and respond to such attacks using the bug bar.

Leveraging machine learning to reduce alert fatigue

ML and alert fatigue

Alert fatigue is a huge concern, as security analysts face the burden of triaging and sifting through a sea of alerts, often correlating alerts from different products manually or using a traditional correlation engine. One of the top challenges in SOCs today is finding the balance: identifying unique, advanced threats while maintaining a manageable number of incidents. With the increasing amount of data and diversity of sources collected in a SOC, the defender community must harness the power of ML to address this challenge.

Security and compliance products with ML built in[28] can help security analysts, data scientists, and engineers act more efficiently and effectively by enabling them to focus on the threats that matter. Identifying threats that fly under the radar while maintaining a low level of alert fatigue is achieved by incorporating graph-based ML and a probabilistic kill chain (likelihood of a series of activities being part of an attack flow).

ML can stitch together millions of lower fidelity anomalous signals into just tens of high-fidelity alerts. In internal evaluations and measurements with customers, we found a median 90% reduction in alert fatigue using a Microsoft technology called Fusion. In one application, Fusion evaluates tens of thousands of suspicious uses of Tor browsers and hundreds of thousands of suspicious firewall rule violations per month, fusing them together into just a dozen high-fidelity incidents of "compromised devices" that are surfaced to the SOC analyst.

Enrichment

Given the high number of investigations a SOC handles daily, it's crucial to prioritize the investigations of users with the highest impact to the organization, in case the account or device is breached. Enriching user metadata with information about the user "blast radius" based on their position in the organization and role importance provides this valuable prioritization and reduces the mean time to resolve the most critical incidents.

For example, a login by a user from a location they're not usually associated with or have never used before, would be considered suspicious and an alert would be issued. However, if the user's peers were logging in from the same location, this similar activity would indicate the event is probably benign and the associated alert would be suppressed. This enrichment reduces mean time to resolution by enabling security personnel to focus their investigations on alerts that matter. It can also reduce SecOps investigation time and level of expertise required.

Peer data: For a given user, we infer the ranked list of the user's peers and frequent collaborators.

Blast radius: For a given user, we infer the user's "blast radius," or the list of important assets that the user has access to.

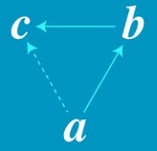
21,000 HOURS

Time spent annually by security analysts triaging false positives.[29]

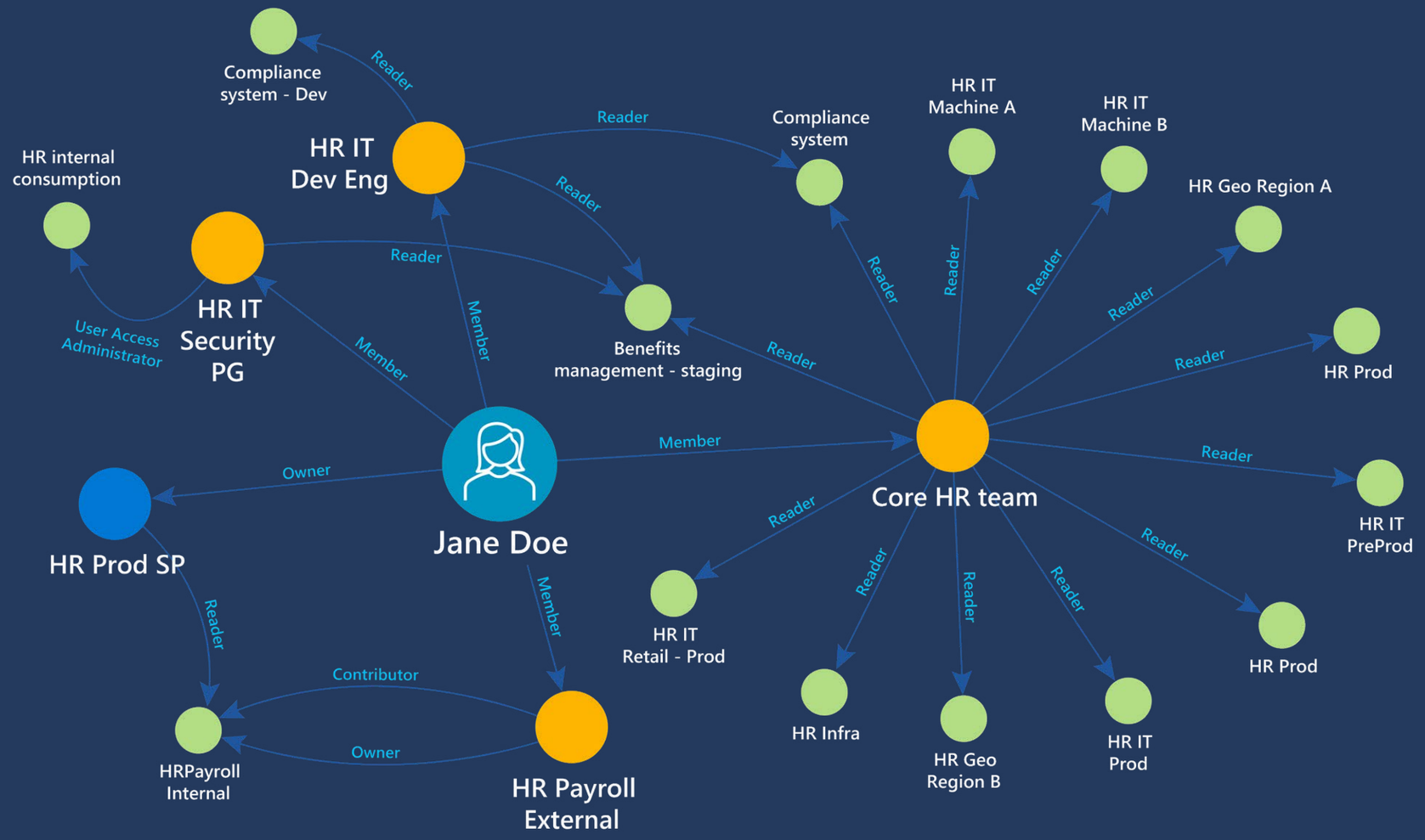
[27] Ram Shankar Siva Kumar et al., "Adversarial Machine Learning–Industry Perspectives," Deep Learning and Security Workshop, co-located with the 41st IEEE Symposium on Security and Privacy, https://arxiv.org/abs/2002.05646/2020;
[28] Azure Sentinel and Microsoft Defender Antivirus; [29] Ponemon Institute; [30] Using Azure Active Directory (AAD) role importance

What is a blast radius?

BLAST RADIUS:
The full privileges of an entity in your cloud include hidden/transitive privileges.



The full impact if an entity in your cloud is compromised.

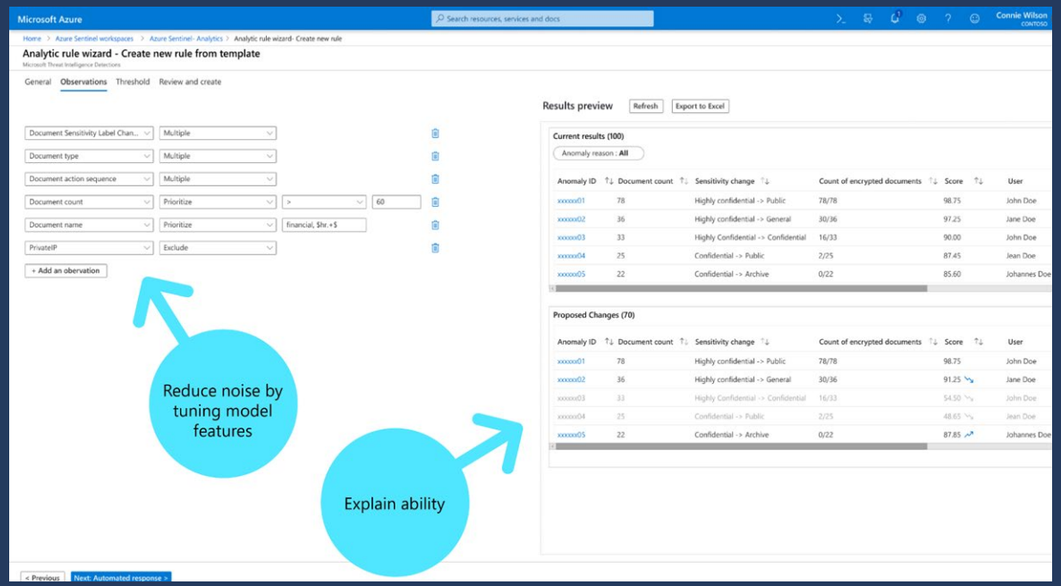


Democratizing machine learning in security with SOC-ML

Microsoft security analysts and data scientists are constantly adding new ML anomalies and successfully finding valuable incidents in customer environments. To scale this effort and make it ubiquitous across the variety of third-party and custom data types, we need the help of the security community—especially managed security service providers (MSSPs) and large organizations that have the necessary security expertise. The major challenge this community encounters is finding security analysts with data science skills needed to develop ML models.

To address this challenge, we advocate democratizing the development of these security ML models, making them accessible to security professionals with no data science background. SOC-ML is an effort to address this challenge, allowing security analysts to customize a model within a familiar user experience, fitting it to the needs of their organizations. It provides the ability to reduce alert fatigue by controlling the features used in the ML model and the thresholds of the scores produced. SOC-ML also provides explainable model scores, which has been a roadblock for wide ML adoption in security. Although democratization of ML continues to trickle into various technology areas, SOC-ML is a unique and disruptive approach in the security space.

Democratizing security ML with explainable model scores



Leveraging anomaly detection for post-breach detection

Behavioral anomaly detection has become increasingly important in enhancing post-breach attack detection. Once an attacker has penetrated the enterprise through phishing or other means, there's very little labeled data, and the attacker has a huge variety of options, reducing the applicability of supervised ML and suggesting instead unsupervised methods that don't require labeled data, such as anomaly detection.

While the methods of anomaly detection are varied, some common themes are present:

- **Behavioral changes over time** are important. Models that can encode time-varying behaviors such as time-of-day, day-of-week, and other variations over time are critical for effective use in security settings.
- **Automatic, streaming model updating** is more suitable for anomaly detection than the batch train/test process of supervised ML.
- The combination of weak anomalies (those that individually are prone to high false positive rates) to produce **strong, attack-relevant rarity** is a powerful method to deploy anomaly detection in security applications. Combining several rare events all indicating attack is much stronger than considering them independently.
- For post-breach kill chain detection, **graphs** are the topology upon which we combine anomaly scores.

We can see in the figures an example of anomaly detection and weak signals combined to produce actionable detection of the Parinacotta ransomware group for one of our customers. This case was reported through our Microsoft Threat Experts targeted attack notifications directly to the customer. The root-cause brute force followed by lateral movement and the dropping of ransomware on victim machines was discovered via probabilistic graph combination of anomaly and alert signals. This patented work was done as part of research for Microsoft Threat Protection.

Go to Actionable Learnings

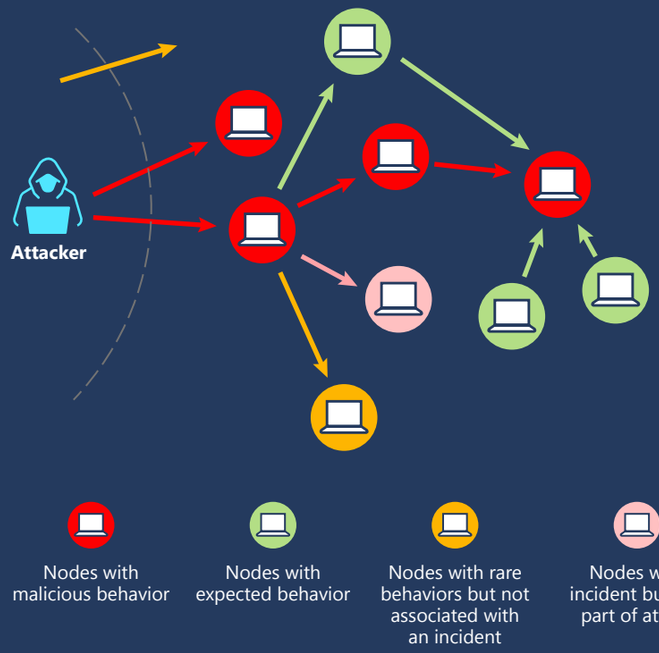
Learn more:

Azure Sentinel uncovers the real threats hidden in billions of low-fidelity signals 2020/02/20

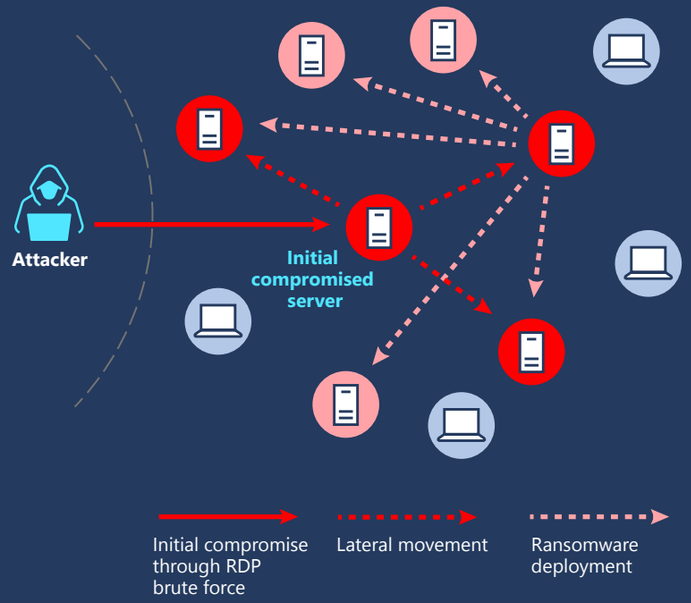
The science behind Microsoft Threat Protection: Attack modeling for finding and stopping lateral movement 2020/06/10

Advanced multistage attack detection in Azure Sentinel 2020/02/18

Actionable alerts based on weak anomalies

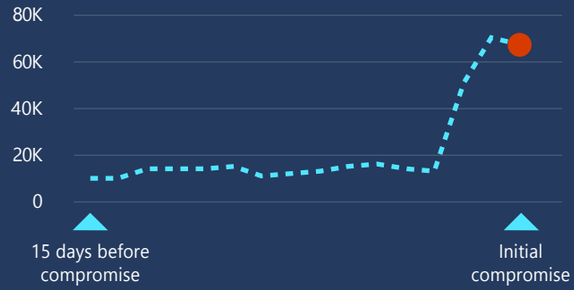


Human-operated ransomware detection



Kill chain detection of human operated ransomware drop following initial penetration two days later. This anomaly detection was achieved through combination of weak signals and alerts to form a strong incident discovery.

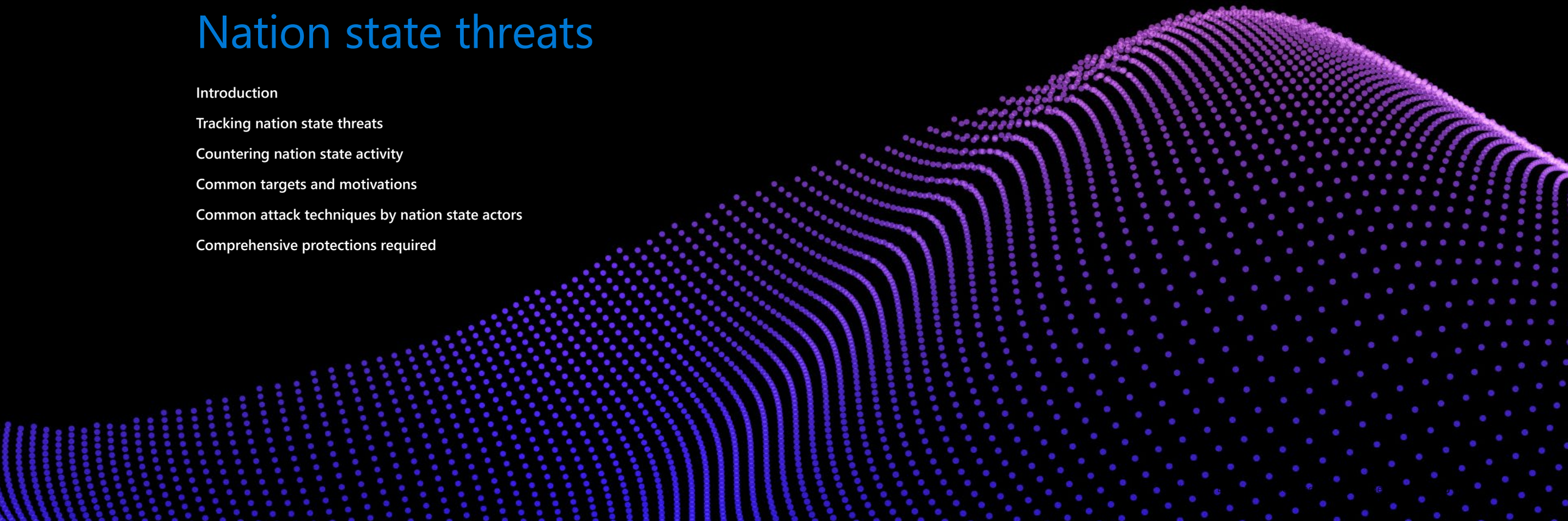
Identifying an attempt of brute force attack



Daily RDP inbound: Time series anomaly detection of brute force attempts followed by successful logon and breach.

2

Nation state threats



Introduction

JOHN LAMBERT, DISTINGUISHED ENGINEER AND GENERAL MANAGER, MICROSOFT THREAT INTELLIGENCE CENTER

Nation state threats are defined as cyber threat activity that originates in a particular country with the apparent intent of furthering national interests. These attacks represent some of the most advanced and persistent threat activity Microsoft tracks.

Nation state activity groups are focused, have the means to develop and deploy novel techniques and tactics, and are constantly working to improve their capabilities. We understand the complex nature of nation state cyberthreats and mobilize all our security analysis and products to discover, track, and defend our customers against them.

Microsoft's approach to defending against nation state threats is as multifaceted as the threats themselves. Our approach rests on a thorough understanding of the tactics and techniques these groups use, their targeting patterns, and the possible objectives driving their activity.

These insights, along with the fidelity Microsoft signals provide, allow us to better spot emerging malicious campaigns, warn customers about the activity, and implement protections against them. The Microsoft Threat Intelligence Center follows these threats, builds comprehensive profiles of the activity, and works closely with all Microsoft security teams to implement detections and mitigations to protect our customers. Each of those teams shares related indicators, detections, and research to build unparalleled visibility into these advanced threats across the full suite of Microsoft products and services.

In this chapter, we share some of Microsoft's intelligence on the tactics commonly used by nation state groups, the industry verticals and geographic regions they target most often, and assessments on the possible motivations for this malicious activity. Our hope is that sharing this knowledge will help security organizations better understand and defend against these sophisticated and consequential threats.

Tracking nation state threats

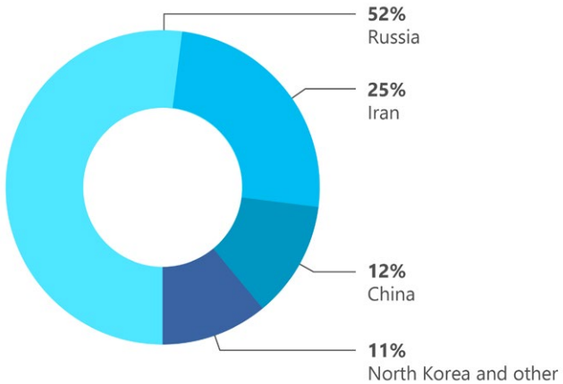
Microsoft tracks nation state activities to protect our platforms, our services, and our customers. We use a variety of metrics and sophisticated data integration techniques to better understand targeting, motivations, and customer impact. MSTIC focuses on nation state activities because these tactics, techniques, and procedures are often unique and novel, prompting downstream actors such as cybercriminals and smaller nation states to eventually copy their methods.

The use of open-source defensive toolkits to improve proactive detections of threats has increasingly become a way for nation state actors to obfuscate their activity by hiding in plain sight using weaponized open-source code.

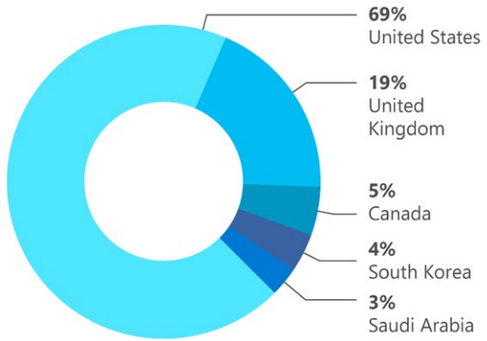
MSTIC analysts focus on nation state activities regardless of platform, targeted victim, or geographical region, and we maintain visibility and active threat hunting worldwide. The information presented here provides a snapshot of customers impacted by nation state activity. It's important to note that even if a particular industry sector or geographic region isn't represented in the information here, nation state activity spans nearly every industry sector and geographic region. In other words, protections against these tactics are critical for every organization and individual. Our intelligence is also impacted by the degree to which our products and platforms are utilized in a particular geography or for particular purposes.

Nation state notifications

When a customer (organization or individual account holder) is targeted or compromised by nation state activities that Microsoft tracks, we deliver a nation state notification (NSN) to the customer. Over the past two years, Microsoft has delivered over 13,000 NSNs. The highest percentage of NSNs represented activity originating in Russia, followed by Iran, China, North Korea, and other countries.



Country of activity origin for NSNs (July 2019–June 2020)



Top 5 targeted geographic regions based on NSNs (July 2019–June 2020)

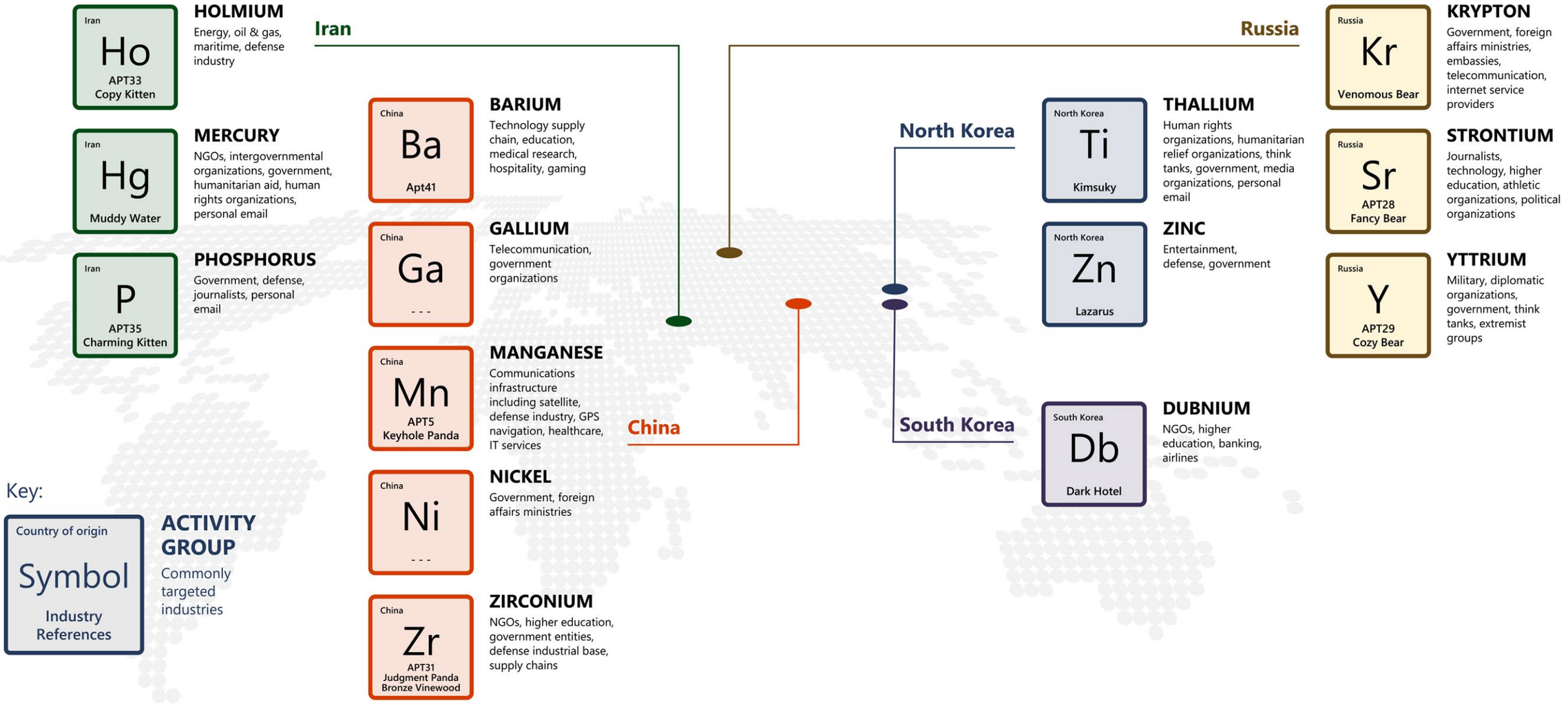
Countering nation state activity

Nation state actors are generally well-resourced and capable adversaries. Our relentless pursuit of these adversaries and our continuous development of new capabilities to detect and deter malicious activity support our commitment to customer protection.

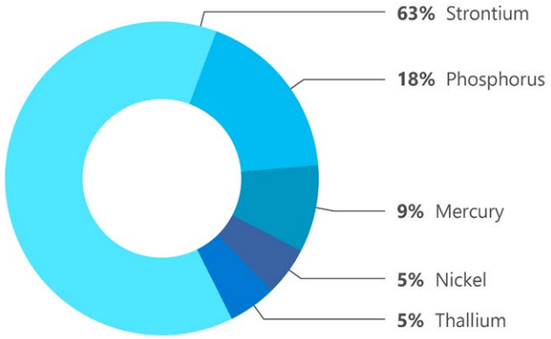
Guide to nation state actors discussed in this report

Throughout this chapter, we cite examples of nation state actors to provide a deeper view into attack targets, techniques, and analysis of motivations. Microsoft identifies nation state activities with chemical element names, just some of which are shown on the following page together with the countries from which the actors operate. This small sample of the total nation state actors tracked by Microsoft represents those that were most active in the last year (July 2019–June 2020) and made most effective use of the tactics detailed in this chapter.

Sample of nation state actors and their activities



Top 5 detected nation state activity groups targeting Microsoft customers (July 2019–June 2020)



Our approach

Microsoft uses four main approaches to disrupt nation state actors—technology, operations, legal action, and policy—and each one plays an important role in our commitment to protecting customers.

Leveraging technology:
Microsoft's cumulative knowledge of the global threat landscape enables our products and services to constantly create and update new detections that protect and defend against nation state activities at scale. These collective defenses represent the most effective method to counter nation state threats because they're informed by the extensive threat intelligence resources built into each product and enabled by world-class engineering.

Taking action against malicious operations:
From time to time, Microsoft will have sufficient information to warrant a one-time deletion or shutdown of infrastructure or assets associated with a nation state attacker. By taking proactive action against malicious infrastructure, the actor loses visibility, capability, and access across a range of assets previously under their control, forcing them to rebuild. An example of this scenario would be the THALLIUM disruption in late 2019.[31]

Leveraging legal actions:
One of Microsoft's unique resources in the fight against nation states is the Digital Crimes Unit. The DCU is responsible for lawsuits against STRONTIUM, BARIUM, PHOSPHORUS, and THALLIUM. Using litigation to seize domains and assets used by nation state actors against Microsoft customers, DCU has been instrumental in shutting down those attack vectors. These four cases have led to the takedown of hundreds of domains and the protection of thousands of customers. Lessons learned from the cases are shared with Microsoft engineering teams to help improve our operational and technical disruption capabilities.

Informing public discourse and policy:
Microsoft uses its voice to raise awareness about nation state activities when we see them, highlighting the context and impacts of the incidents. This reporting helps drive a broad discussion about what can be done to combat malicious nation state activities across governments, enterprises, academia, social organizations, and the public. Talking publicly about nation state attacks is an important part of deterrence. For example, in October 2019, Microsoft identified publicly that an account associated with a U.S. presidential campaign had been targeted by PHOSPHORUS.[32, 33] This public disclosure drove a significant dialogue about the attack and election security, with sustained media coverage, putting a spotlight on the necessity of protecting political campaigns from foreign actors.

[31] *https://blogs.microsoft.com/on-the-issues/2019/12/30/microsoft-court-action-against-nation-state-cybercrime/*
[32] *https://blogs.microsoft.com/on-the-issues/2019/10/04/recent-cyberattacks-require-us-all-to-be-vigilant/*
[33] *https://www.nytimes.com/2019/10/04/technology/iranian-campaign-hackers-microsoft.html*

Common targets and motivations

To provide a deeper analysis of some of the observed nation state activity, we look at frequently targeted sectors and the motivations of the attackers.

Critical infrastructure

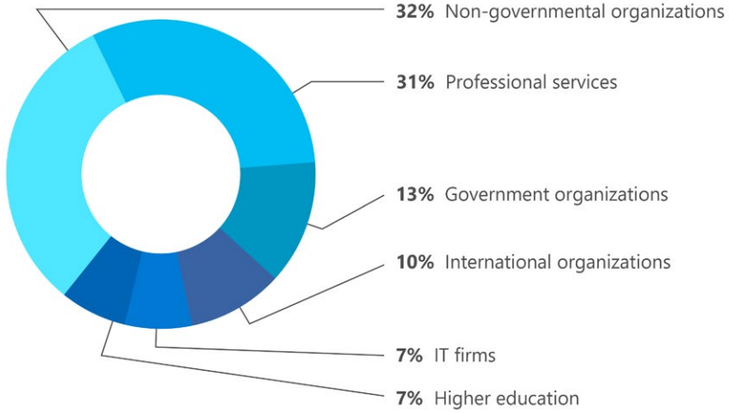
Nation state activity is more likely to target organizations outside of the critical infrastructure sectors by a significant measure, with over 90% of notifications served outside of these sectors.

Within the critical infrastructure sectors, targeting of IT organizations represents over 60% of nation state activity, followed by commercial facilities, critical manufacturing, financial services, and the defense industrial base.

Frequently targeted sectors

The sectors listed below were consistent targets of nation state activity in the period from July 2019 through June 2020. We've listed them in order of prevalence with respect to the number of NSNs delivered to each sector.

Top 6 targeted industry sectors by NSNs delivered (July 2019–June 2020)



Top 6 targeted industry sectors

32%
Non-governmental organizations

The NGO recipients of NSNs included organizations such as advocacy groups, human rights organizations, nonprofit organizations, and think tanks focused on public policy, international affairs, and security. STRONTIUM was responsible for more than three quarters of activity triggering NSNs in this sector through heavy targeting of sports associations. DUBNIUM was the second most active group, followed by ZINC, THALLIUM, and PHOSPHORUS.

31%
Professional services

Support services and trusted third parties can serve as launch points for intrusions or collection against government targets, whose systems might have more robust security protocols in place. In the last year, STRONTIUM attempted to compromise consulting firms and government contractors, specifically U.S. defense contractors and large public affairs, corporate legal, IT, media, and physical security consultancies operating in the United States, Europe, and the Middle East.

13%
Government organizations

Nation state threat actors attempted to compromise foreign ministries and equivalent diplomatic institutions in countries across the Americas, Australia, Central and Southern Africa, Europe, and the Middle East. NICKEL made an aggressive push against foreign ministries in 11 countries across Latin America and Europe during this period. Microsoft discovered YTTRIUM had gained footholds in the diplomatic missions of several European countries. Military institutions and national legislatures were also popular targets in the government sector.

10%
International organizations

Among international organizations, global and regional organizations working on governance norms and human rights were top targets. MERCURY attempted to compromise a high number of targets involved in work with refugees. THALLIUM focused its efforts on a regional organization devoted to developments in Africa. As the COVID-19 pandemic unfolded, both DUBNIUM and PHOSPHORUS directed their attention toward global health institutions.

7%
Information technology firms

As mentioned earlier, IT is the most heavily targeted of the critical infrastructure sectors. Attacks on IT products and service providers are the most prevalent in this category, followed by companies that provide internet routing, access, and connections. MANGANESE was the most active against IT sector targets, pursuing account compromise in U.S.-based companies involved in microchip technology, cybersecurity, and networking solutions. MERCURY targeted network technology providers in the Middle East.

7%
Higher education

Actors STRONTIUM, PHOSPHORUS, BARIUM, THALLIUM, and ZINC all targeted universities in this period. Victims were spread across the globe, including institutions of higher education in the United States, Asia, Europe, Latin America, and the Middle East. Colleges and universities are targeted because they often house cutting-edge research initiatives that might be of interest to nation state actors. For example, a suspected nation state group operating out of China compromised accounts at a U.S. university involved in COVID-19 vaccine research in March. Beyond scientific research itself, nation state actors originating from both North Korea and Iran targeted global universities' subject matter experts (SMEs) that influence international policy on topics ranging from international security and nuclear nonproliferation to domestic politics and human rights in their respective regions. Accessing the SMEs' personal and professional contacts could serve espionage and counterespionage purposes for the supporting nation state.

Targeting major events and other opportunities

Major current events are commonly targeted by nation state actors or leveraged as opportunities to further operational objectives. Microsoft has observed several such operations targeting events around the globe, including elections and individuals tied to political campaigns as well as the Olympic Games. More recently, several nation state actors took advantage of the global COVID-19 pandemic to craft targeted spear-phishing lures in credential theft and malware deployment efforts.

Political campaigns

In a 30-day period between August and September 2019, MSTIC observed PHOSPHORUS performing password sprays, making more than 2,700 attempts to identify consumer email accounts belonging to specific Microsoft customers, and then proceeding to attack 241 of those accounts. The targeted accounts are associated with a U.S. presidential campaign, current and former U.S. government officials, journalists covering global politics, and prominent Iranians living outside Iran. Four accounts were compromised as a result of these attempts. (These four accounts weren't associated with the U.S. presidential campaign or current and former U.S. government officials.) As we approach the 2020 U.S. general election, we're likely to see activity increase after this report was written. MSTIC has previously observed similar patterns by STRONTIUM and other actors targeting elections or politically affiliated organizations in other countries. ZIRCONIUM, operating from China, has attempted to gain intelligence on organizations associated with the upcoming U.S. presidential election. We've detected thousands of attacks from ZIRCONIUM between March 2020 and September 2020 resulting in nearly 150 compromises. As additional elections occur around the world, Microsoft will continue to monitor and notify our customers of these attacks.

2020 Olympic Games

As the world prepared for the Tokyo Summer Olympic Games in 2020, at least 16 national and international sporting and anti-doping organizations across three continents were targeted. The attacks began on September 16, 2019, just before news reports emerged about new potential action being taken by the World Anti-Doping Agency (WADA). Following patterns detected in attacks against organizations affiliated with previous Olympic Games, some of these attacks were able to successfully gain access to targeted accounts. STRONTIUM has consistently targeted these organizations using spear phishing, password spray, and both open-source and custom malware to compromise devices for further espionage.[35]

COVID-19 outbreak

Microsoft observed 16 different nation state actors either targeting customers involved in the global COVID-19 response efforts or leveraging the crisis in themed lures to expand their credential theft and malware delivery tactics. These COVID-19-themed attacks targeted prominent government healthcare organizations in efforts to perform reconnaissance on their networks or people. They targeted global medical relief and humanitarian aid organizations with spear-phishing lures paired with malicious macros in Word documents. They also crafted unique credential theft lures spoofing a popular U.S. fast food chain's COVID-19 response email and spoofed offers of online coupons. Academic and commercial organizations involved in vaccine research were also targeted.

Although these activities didn't represent significant operational shifts by these actors during the COVID-19 crisis, it's imperative that every organization affiliated with or affected by a major public event understand how their organization defends against the tactics commonly used by these actors.

Learn more :

New cyber attacks targeting US elections 2020/09/10

[34] *https://www.wired.com/story/untold-story-2018-olympics-destroyer-cyberattack/*

[35] *https://blogs.microsoft.com/on-the-issues/2019/10/28/cyberattacks-sporting-anti-doping/*

Common operational aims

Nation state actors often conduct intrusions in pursuit of larger strategic objectives, which they view as critical to the stability, prosperity, and sometimes survival of their home nations. As such, these groups are highly motivated and willing to invest significant time and resources into achieving their operational objectives. Whatever the strategic objective behind it, nation state cyberthreat activity has common operational aims: to steal information (about you, your business, personnel, customers, or partners) or to disrupt or destroy components of your operations.

Espionage

Microsoft tracks several nation state activity groups for whom espionage is their main goal. These groups operate to gather information about the intentions and movements of targets of interest. Some of the information they're looking for includes government correspondence, proprietary business information, and individual contacts.

Government correspondence

Government correspondence speaks directly to national plans and intentions. Threat actors across the globe focus on compromising foreign ministries because these organizations produce, exchange, and store voluminous content about their home nations' foreign policy and security priorities. The extensive targeting of think tanks and policy-oriented NGOs is probably also reflective of a hunt for insights into government policy. Think tanks and NGOs focus largely on government policy often without the baseline attention to cybersecurity that many government agencies have, making them attractive targets from which to collect insights into national policy.

Proprietary business information

Nation state threat actors seek to collect trade secrets and intellectual property for a number of reasons, which can include accelerating the growth of their own domestic industries, advancing a military program, or determining how advanced an adversarial nation's weapons or technology capabilities are. As the COVID-19 pandemic unfolded, China-based nation state threat actors targeted medical research institutions in the United States and Asia, highlighting competition for medical innovations as another potential motive for proprietary information theft.

Individual contacts

In addition to directly attempting to collect information about experts in their fields of interest, nation state threat actors seek information on individuals to learn about their intentions and ideas, similar to the tactics described in the Government correspondence section. They might also be seeking to improve the effectiveness of spear phishing and other social engineering efforts and to surface additional attack vectors within a target's personal or professional networks. In April and May 2020, ZINC actors sent phishing lures spoofing multiple U.S.-based aerospace firms to employees at competitor firms in Asia and Europe, probably hoping to increase collection opportunities in the sector. For more information and detailed examples of how this information is employed in operations, see the Common attack techniques by nation state actors section of this report.

Disruption or destruction

Although less common than espionage, nation state threat actors have conducted intrusions intended to disrupt or destroy data or physical assets at targeted facilities or institutions. The U.S. National Institute of Standards and Technology (NIST) defines a disruption as "an unplanned event that causes the general system or major application to be inoperable for an unacceptable length of time."[36] Disruptive operations can, for example, result in minor or extended power outages or extended network unavailability. Destructive operations involve "overwriting, erasing, or physically destroying information," equipment, or facilities.[37]

A look at past activities suggests nation state actors might engage in disruptive or destructive operations to retaliate against actions by a perceived actor, to deter or discourage continuation of certain activities, or to test new or developing cyber tools and techniques. For example, actors operating from Russia have a history of launching disruptive and potentially destructive cyberattacks in response to perceived anti-Russian actions in international sport. In advance of the Olympic Games in 2016 and 2018, respectively, suspected Russia-based threat actors stole and leaked athletes' sensitive medical data and rendered inoperable the servers comprising the IT backbone of the Olympic Games.[39]

Late in 2019, Microsoft observed STRONTIUM attempting to compromise at least 16 national and international sporting and anti-doping organizations across three continents. These attempts began when WADA announced its decision to investigate Russia's anti-doping agency in September 2019 and continued through December 2019 when WADA ultimately ruled to ban the Russian national team from participation in the Olympic Games.[40]

...nation state actors may engage in disruptive or destructive operations to retaliate against actions by a perceived actor, to deter or discourage continuation of certain activities, or to test new or developing cyber tools and techniques.

[36] *https://csrc.nist.gov/glossary/term/disruption*

[37] *https://csrc.nist.gov/glossary/term/Destruction*

[38] *"Russian hackers leak Simone Biles and Serena Williams files," BBC News, September 13, 2016, https://www.bbc.com/news/world-37352326; "WADA confirms another batch of athlete data leaked by Russian cyber hackers 'Fancy Bear,'" WADA website, September 14, 2016. [Online]. Available: https://www.wada-ama.org/en/media/news/2016-09/wada-confirms-another-batch-of-athlete-data-leaked-by-russian-cyber-hackers-fancy*

[39] *Greenberg, Andy. "The Untold Story of the 2018 Olympics Cyberattack, the Most Deceptive Hack in History," WIRED Magazine, October 17, 2019. [Online]. Available: https://www.wired.com/story/untold-story-2018-olympics-destroyer-cyberattack/*

[40] *https://blogs.microsoft.com/on-the-issues/2019/10/28/cyberattacks-sporting-anti-doping/*

Common attack techniques by nation state actors

Understanding what nation state threat groups are doing and why is one part of the story, but even more important is to know how they conduct their operations so that organizations can put the appropriate defense protections and practices in place. The sections below represent the most common tactics used by nation state actors in the last year.

Reconnaissance

Microsoft has observed a significant increase in password spray attacks by nation state actors in the last year that indicate a new level of sophistication in their reconnaissance tactics. Identifying and accessing legitimate user accounts is a critical method that an increasing number of nation state actors are using to successfully infiltrate a targeted organization and achieve operational aims.

PHOSPHORUS, HOLMIUM, and STRONTIUM are examples of actors associated with nation state operations that have successfully refined the password spray tactic for reconnaissance over the past two years. Password sprays have included known accounts across large organizations but have also included iterations of possible name formats to discover accounts, including personal accounts. PHOSPHORUS has targeted a wide variety of organizations in the past year, with particular focus on targeting personal email accounts of individuals employed or affiliated with their objective. These operations have largely focused on individuals involved in human rights organizations, humanitarian or refugee relief efforts, NGOs defining or setting policies associated with international sanctions, or personal accounts of government employees.

Personal email accounts are a particularly effective target for a reconnaissance operation since most individuals don't apply the level of security protections applied by larger organizations (such as MFA and complex passwords). Individuals in smaller organizations might also use consumer email platforms for official communications, making them high-value targets for surveillance. If a large-scale password spray or password guessing operation is launched against a set of known accounts, one of the most effective ways to proactively prevent those attacks is through the combination of MFA and strong passwords.

HOLMIUM uses a combination of tactics to gain access to networks, including socially engineered spear-phishing operations and password spray attacks. During one wave of activity over a two-day period in January 2020, HOLMIUM password spray activity attempted logins against 33,141 different tenants with an average of 2.74 accounts per tenant.

Example of name reconnaissance conducted by PHOSPHORUS: iterations of name formats including personal attacks

J.Smith@contoso.com
John.smith@contoso.com
John.m.smith@contoso.com
JohnSmith@contoso.com
johnmsmith@contoso.com

Other capabilities of HOLMIUM include use of open-source "penetration testing tools" as well as custom PowerShell-based tools. Attacks have targeted entities affiliated with the oil and gas sector and its supporting industries, including maritime transportation and logistics companies. HOLMIUM actors have been observed deploying Ruler, a self-described "tool to abuse Exchange accounts" on compromised accounts.[41] The Ruler command and control server in turn deploys a custom PowerShell-based implant identified as "Powerton."

[41] *https://github.com/sensepost/ruler*

Credential harvesting

Credential harvesting is a common tactic used by nation state actors to improve their ability to compromise a targeted organization. Microsoft tracks activity by multiple actors that leverage spear-phishing emails, carefully crafted imitation domains, and spoofed login pages to convince employees to login using their corporate credentials. These logins on maliciously crafted websites steal and then reuse the usernames and passwords recorded by unsuspecting employees.

THALLIUM is one Microsoft-identified nation state actor that uses credential harvesting as a successful tactic to gain access to NGOs, government organizations, and other professional services organizations affiliated with their areas of geopolitical interest. THALLIUM spends significant resources purchasing and using imitation domains that resemble their targeted organization's name, or using the name of an online service the target uses, as the subdomain. For example, THALLIUM registered the *client-mobile.work* domain and then created a subdomain that includes the words "login" and "outlook" (such as *login-outlook.client-mobile.work*) to raise the likelihood they will successfully convince a targeted individual to provide their credentials. THALLIUM then set up a cloned authentication page to convince the target to provide their username and password, thinking they were logging into a legitimate corporate site. It's also worth noting that use of the word "mobile" could also indicate they anticipated a user might click on this lure from a mobile phone and the screen size or difficulty in seeing the full URL on these devices might increase the likelihood a user will fall for the lure.

THALLIUM's spoofed pages include most major online service brands. The servers they use to conduct this credential theft are often compromised servers of other legitimate companies, such as a small business computer that has been compromised and then set up as infrastructure for this intrusion. THALLIUM has also been known to use inexpensive regional virtual private server (VPS) providers as infrastructure for similar operations. In the last year, MSTIC has discovered more than 500 domains that were registered by THALLIUM to conduct credential theft operations—most mimicking a well-known brand to successfully lure their targeted user.

THALLIUM typically attempts to trick victims through spear phishing. By gathering information about the targeted individuals from social media, public personnel directories from organizations the individual is involved with, and other public sources, THALLIUM is able to craft a personalized spear-phishing email in a way that gives the email credibility to the target. As seen in the sample spear-phishing email below, the content is designed to appear legitimate, but closer review shows that THALLIUM has spoofed the sender by combining the letters "r" and "n" to appear as the first letter "m" in "microsoft.com." The link in the email redirects the user to a website requesting the user's account credentials. By tricking victims into clicking on the fraudulent links and providing their credentials, THALLIUM is then able to log in to the victim's account.

Additionally, THALLIUM will often create a new mail-forwarding rule in the victim's account settings. This mail-forwarding rule will forward all new emails received by the victim to THALLIUM-controlled accounts. By using forwarding rules, THALLIUM can continue to see email received by the victim, even after the victim's account password is updated.

Credential harvesting example

Learn more:

STRONTIUM: Detecting new patterns in credential harvesting 2020/09/10

Malware

Malware represents one of the most effective and commonly used methods by nation state actors to successfully infiltrate and maintain persistence on a targeted system. Advanced adversaries invest heavily in development of unique malware, but also are known to obfuscate their activity by using openly available malicious code for mainstream online criminal activity. Even more common is the malicious use of known vulnerabilities in commercial enterprise software to take advantage of poorly secured targets.

BARIUM is a long-running actor with abundant success in compromising victims since 2006. Although the group has been observed mostly targeting information technology, software, and IT products and services companies, Microsoft has seen BARIUM target a wide range of organizations including but not limited to retail, manufacturing, NGOs, government organizations, gaming, and hospitality. BARIUM is particularly effective at improving on their own custom malware and leveraging open-source tools to limit detection.

BARIUM appears to frequently upgrade and refine their toolkit with slight modifications to their implants to stay ahead of defenders. Some of the modifications to their implants suggest that BARIUM actively develops and maintains either the primary codebase or a fork of PlugX, Strilix (Crosswalk), and Winnti malware.

The attention to detail seen in BARIUM reconnaissance, infrastructure setup, and custom malware has contributed to abundant success in compromising targets.

Summary of techniques BARIUM has used successfully in past operations:

- **Windows shortcut (.Lnk) with hidden payloads**
  Windows Shortcut.lnk with hidden payloads was observed in the summer 2016. The RAR archive BARIUM sent to the victim contained a Windows Shortcut file, Resume.docx, and an executable named Thumbs.db. When the victim ran the .lnk file, it installed the Win32/Barlaiy3F[42] implant and then displayed a legitimate Resume.docx.
- **Windows compiled html help files. (.Chm)**
  Because Microsoft HTML Help is a necessary feature in the Windows platform, HTML Help generates .CHM files legitimately that can be viewed by using the Windows HH.exe viewer. BARIUM delivers malicious .CHM files that execute when the user opens them. While the user will see what appears to be a legitimate document in the Help Viewer (commonly a résumé), the malware silently runs JavaScript that executes the malware delivery and stores it in xml.htm.
- **Microsoft PowerPoint and Word documents with macro**
  This technique required the victims to "enable content" (macros) within a malicious PowerPoint or Word file typically delivered in an email. This technique can bypass protections if users aren't careful and manually enable the macro when opening the document. The theme of these files was typically focused on résumés or business development, aligning with themes uncovered in the reconnaissance phases. When these macros were successfully run, a backdoor trojan was installed.

BARIUM also uses what's described as "supply chain" techniques as an initial infection vector against organizations globally. In these cases, BARIUM compromises software companies and modifies their source code to install malware. Modifying legitimate software and compromising update processes allows BARIUM to selectively compromise organizations of their choosing. BARIUM has embedded malware in legitimate software executables which, when deployed, beaconed over DNS to infrastructure that BARIUM controlled. By monitoring this DNS traffic, BARIUM could select their victim and customize delivery of more specialized modules to clients of their choosing.

More recently, BARIUM has shifted their techniques to start using more open-source malware tooling in their attacks. Microsoft has observed them dropping batch files that are used to create persistence for their Implants using Crosswalk/Strilix and CobaltStrike, two common open-source penetration testing tools that can easily be weaponized for attacks.

For more information on software supply chain, see the In focus: Supply chain security section of this report.

[42] *https://www.microsoft.com/en-us/wdsi/threats/malware-encyclopedia-descriptionTrojanDropper:Win32/Barlaiy.A!dha*

VPN Exploits

As the world adjusts to increased numbers of remote workers, global enterprise IT departments rely on VPNs to improve the connectivity and security of systems. Typically, this critical service is managed by third-party software deployed across our devices. Since mid-2019, Microsoft has observed nation state actors consistently targeting and frequently compromising outdated and unpatched VPN infrastructure. This activity indicates they view it as an easy and effective method for penetrating and persisting on a targeted network by using compromised credentials.

MANGANESE has been successful in compromising victims since 2010. The group targets companies in sectors such as satellite communications, defense, healthcare and public health, IT products and services, and GPS navigation. They're known to be quite stealthy, maintaining access to their victims for long periods of time, and conducting successful operations against multiple victims over several years with few updates to their malware. They frequently use social profiles from highly visited internet sites for their malware configuration strings, making it difficult for targeted organizations to detect.

MANGANESE is known for leveraging stolen credentials from victim networks to maintain long-term access. Their most common tactic for maintaining persistence on an account is by connecting to remote access locations using legitimate accounts, which reduces the chances of being detected. They also rely on heavy use of VPN providers for infrastructure and short-timespan connections to further reduce their trackable footprint.

For more information on VPN exploits, see the Security and the remote workforce chapter of this report.

Applying security patches for internet-facing systems is critical in preventing attacks. When managing VPN or VPS infrastructure, it's critical for organizations to know the current status of related security patches. In October 2019, both the National Security Agency and National Cyber Security Centre (NCSC) put out alerts on these attacks and encouraged enterprises to patch.

Web shell-based attacks on the rise

Microsoft's Detection and Response Team (has seen that web shell-based attacks are also on the rise. Often written in typical web development programming languages (such as ASP, PHP, JSP), web shells allow an attacker to execute commands and to steal data from a web server or use the server as a launch pad for further attacks against the affected organization. While multiple threat actor groups, including ZINC, KRYPTON, and GALLIUM, have been observed using web shells, their prevalence has increased, and Microsoft Defender Advanced Threat Protection now detects an average of 77,000 web shells and related artifacts on an average of 46,000 unique machines every month.

DART completed an incident response investigation and remediation to an incident that used a web shell to gain a foothold in a customer environment. The customer was a public sector organization that had their internet-facing servers misconfigured, allowing attackers to upload a web shell and gain initial entry to further compromise their network. DART's investigation revealed that the attackers uploaded multiple web shells leading to compromise of domain admin and service accounts. The attackers also installed a dynamic link library backdoor on an Outlook Web Access server. To persist on the server, the backdoor implant registered itself as a service or as an Exchange transport agent, which allowed it to access and intercept all incoming and outgoing emails, exposing sensitive information of the customer.

Comprehensive protections required

The skill and persistence of malicious nation state activity increases the difficulty of detecting and protecting against these threats. Their impact can be wide ranging and highly damaging. These adversaries are well funded, employ techniques of tremendous breadth and sophistication, and are motivated by objectives of national significance—which might lead to their compromising networks for unexpected purposes. More than other adversaries, nation state attackers target individuals specifically for access to their connections, communications, and information.

Defense-in-depth strategies against nation state adversaries should include educating employees on how to avoid being targeted themselves. Enabling MFA on all personal or work accounts is a critical security practice. This best practice for the home environment, preventing attacks including those that could result in financial loss, will also help to protect corporate assets from attacks originating from personal email addresses, file sharing services, and social media.

While nation state attacks are often sophisticated or can deploy zero-day vulnerabilities to gain access to networks, defense-in-depth strategies and proactive monitoring can greatly reduce the actor's dwell time on a network, potentially enabling disruption of their activities before they reach their goals. Above and beyond enabling MFA, IT departments should prioritize steps to mitigate lateral movement by attackers; specifically, credential hygiene and network segmentation. To limit the damage of data exfiltration, information rights management can be applied to files. Building protective controls into your network will raise the threshold for attackers, improving your organization's ability to detect anomalous activity in the environment.

Go to Actionable Learnings

Learn more:

Microsoft Threat Protection stops attack sprawl and auto-heals enterprise assets with built-in intelligence and automation 2020/02/20

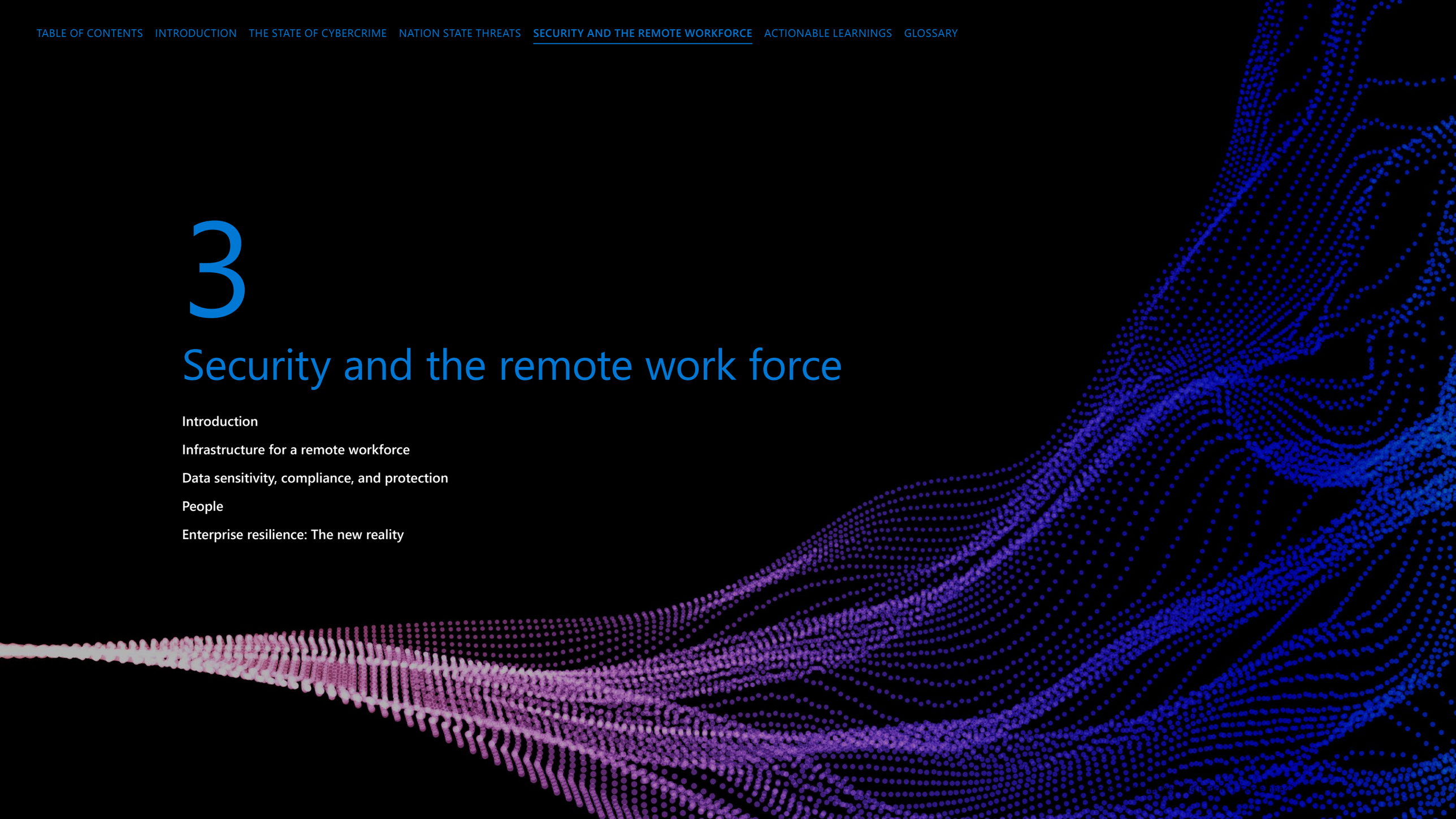
Microsoft takes court action against fourth nation-state cybercrime group 2019/12/30

MITRE ATT&CK APT 29 evaluation proves Microsoft Threat Protection provides a deeper end-to-end view of advanced threats 2020/04/21

The Verge: Microsoft has warned 10,000 people that nation-state hackers are targeting them 2019/07/18

3

Security and the remote work force



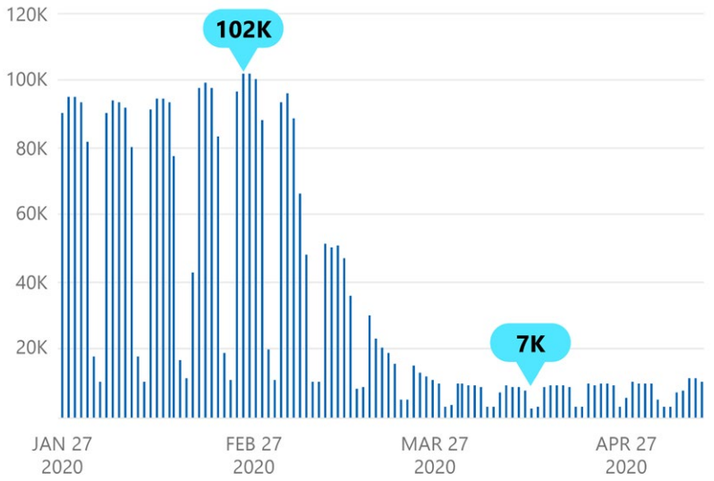
Introduction

BRET ARSENAULT, MICROSOFT CHIEF INFORMATION SECURITY OFFICER

A hallmark of the digital economy in the 21st century is the ability of employees to work from anywhere, fueled by significant innovations in infrastructure, mobile endpoints, and communication applications. Although there were increasingly more opportunities to work remotely in recent years, the COVID-19 pandemic catalyzed these efforts, forcing a significant portion of the workforce to work from home. As Microsoft CEO Satya Nadella stated, "We have seen two years' worth of digital transformation in two months."[43]

The immediacy and scale of moving to a remote workforce brought along new security challenges. When Microsoft recommended employees work from home even before governmental regulations required it, its workforce was able to respond quickly. As long as employees had secure, adequate internet access at their remote locations, they could continue being fully productive and, most critically, they could protect themselves and their families from the pandemic. We were successful in making this transition because we invested in a Zero Trust architecture, including MFA, ubiquitous device management, and conditional access enforcement. The graph on this page shows how quickly Microsoft became a remote workforce and the scale at which we had to manage the requisite infrastructure, applications, and personnel.

Microsoft employee badge scans for building access



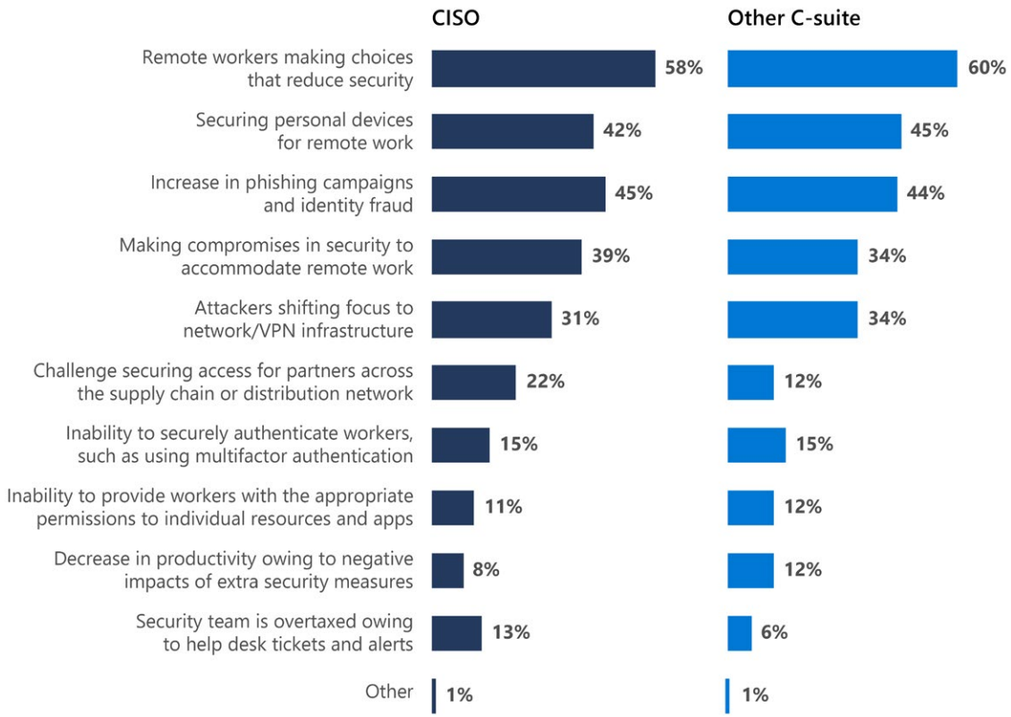
[43] *https://www.microsoft.com/en-us/microsoft-365/blog/2020/04/30/2-years-digital-transformation-2-months/*

As more organizations adapt to enabling secure remote work options, whether in the short or long term, cybersecurity provides the foundation for operational resiliency. Successful transition in any type of disruption requires a strategic combination of planning, response, and recovery. To maintain cyber resiliency, an organization should regularly evaluate its risk threshold and ability to operationally execute key processes through a combination of human efforts and technology products and services.

This chapter provides a look at three areas of risk that have become more apparent with the rapid onset of a large-scale remote workforce. First, we'll explain how organizations can support a secure, remote workforce through the principles of Zero Trust, including VPN architecture and virtualization—which helped enable Microsoft's own transition to remote work. This section includes how we address security risks associated with managing devices not owned by an organization (and often owned by employees and partners). Second, we address the human element as fundamental to a secure workforce by looking at challenges such as insider threats and social engineering by malicious actors. Third, we explore lessons learned as we engaged in an exercise in enterprise resilience during the COVID-19 outbreak.

Security and remote workforce concerns

Security decision makers in the United States perceive their most common remote workforce challenge as remote workers making choices that reduce security. Securing personal devices for remote work and the increase in phishing campaigns and identity fraud are also concerns.[44]



[44] *Based on 506 responses from U.S. security decision makers, representing organizations of 1,000+ employees (April–May 2020)*

Infrastructure for a remote workforce

Infrastructure security is a fundamental concern when scaling a remote workforce. There's a substantial difference between an infrastructure that supports employees taking laptops home to monitor email and a complete architecture for a mobile workforce. Originally built for in-office work, many infrastructures weren't initially prepared for a large scale—and immediate—transition to a remote workforce.

Zero Trust security model

By treating every access attempt as if it were originating from an untrusted network, the Zero Trust strategy helps solidify the security of VPNs for working from home. This is precisely the approach needed with today's remote workforces because they're coming from untrusted home networks and the VPN architectures used to extend corporate networks are sometimes failing.

Zero Trust defined: Instead of assuming everything behind the corporate firewall is safe, the Zero Trust model assumes a breach and verifies each request as though it originates from an open network. Regardless of where the request originates or what resource it accesses, Zero Trust teaches us to "never trust, always verify." Every access is fully authenticated, authorized, and encrypted before granting access. Micro-segmentation and least-privileged access principles are applied to minimize lateral movement. Rich intelligence and analytics are utilized to detect and respond to anomalies in real time.

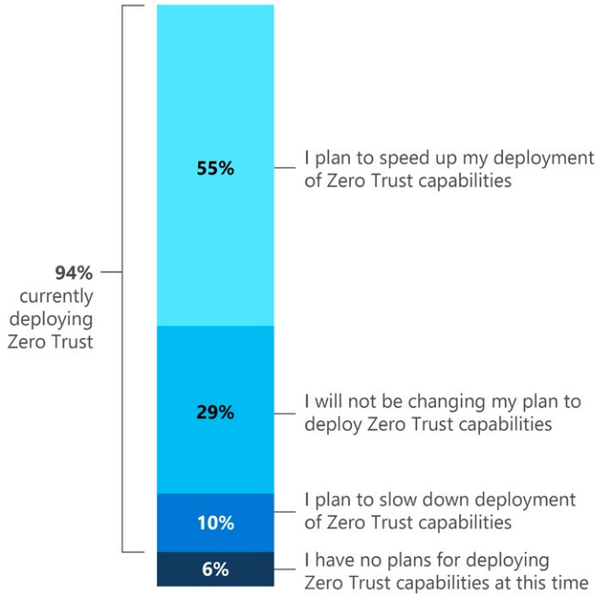
Zero Trust helps secure corporate resources by eliminating unknown and unmanaged devices and limiting lateral movement. As the COVID-19 pandemic forced the "lift and shift" of the corporate workforce and much of its workload to a distributed—or even scattered—model, the notion of the corporate security perimeter changed dramatically. Traditional security policies within the perimeter became much harder to enforce across a wider network made up of home and other private networks and unmanaged assets in the connectivity path. Corporations that had already made investments, even partial investments, in areas that contribute to Zero Trust, were able to leverage advances already made by their security teams in areas such as strong identity, enrolled device management and device health, alternate access for unmanaged devices, and application health. Even in cases where these investments weren't focused on a full Zero Trust implementation, they became the basis for improved security risk mitigation and management. They collectively helped enable the extension of the security perimeter to accommodate a physically dispersed workforce and related assets.

Microsoft's approach to Zero Trust enablement allows for staggered or progressive enablement, where the highest priority capabilities—such as Identity as a Service (IDaaS)—can be invested in earlier, with other capabilities enabled subsequently. This approach allows customers the ability to plan their spend and make capital expenditures in alignment with essential business needs and growth opportunities.

From a technical perspective, Microsoft's approach focuses on access architectures that use policy to explicitly validate the trustworthiness of users and devices. Through validation, Microsoft dynamically addresses insufficient trust via increasing assurance, limiting access, or blocking access altogether. This approach doesn't rely on network architectures but instead uses IDaaS as the decision engine based on organizational policy. The result is mobility and choice to enable productivity for end users (and remote workers). This model can be extended further to allow users to choose any device from which to work because the access architectures can now account for the variability in trust of these potentially unmanaged devices.

When we asked enterprise security decision makers about deploying Zero Trust, we learned that most have deployed it to some extent, and over half plan to speed up Zero Trust deployment.[45]

Impact of the COVID-19 pandemic on Zero Trust deployment



Learn more:

Enable a remote workforce by embracing Zero Trust security

Implementing a Zero Trust security model at Microsoft 2019/08/06

VPN architecture

Historically, VPNs have been used by organizations to extend their network to additional sites and individual users. Traffic from these sites or individual users was sent to a central site and then out to the internet. For security teams, the network was the original visibility and control plane, and for some companies, it remains a key component of their security strategy. It's because of this reliance on network architectures that security teams tended to enforce the use of full VPN tunnels to extend their network security strategy. An architecture that routes all remote traffic back to the corporate network was originally intended to provide the security team with the following capabilities:

- Prevention of unauthorized access
- Control of authorized user access
- Network protections such as intrusion detection and prevention (IDS/IPS) and DDoS mitigation
- Data loss prevention

The infrastructure required to support the VPN architecture and related security stack was often designed to support only a small and controlled percentage of the user population at any given time. When some workforces moved to working nearly 100% remotely, their architectures became overtaxed under the increased load. That load included not only day-to-day work with email and documents, but also increased use of collaboration software utilizing audio and video, which would oftentimes be degraded. The full VPN tunnel architecture created usability issues for end users, and in some cases, the VPN infrastructures failed completely as a result. We reached out to 123 customers during the initial weeks of the COVID-19 outbreak after noting some latency in their audio or video quality owing to the traffic volume. To tackle this issue, some customers doubled down on their network architectures to increase capacity, while others looked for new ways to manage risk in a more direct-to-internet network architecture.

[45] *Based on 524 responses from security decision makers in the United States, the UK, and India representing organizations of 500+ employees (May 2020)*

Controlling access to corporate resources

Personally owned, unmanaged devices that access organizational resources are a security risk associated with remote workforces, as well as third-party vendors and partners.

As part of Microsoft's Zero Trust strategy leveraging Azure AD Conditional Access, we can block direct access to Microsoft systems from unmanaged devices, but allow the employee to decide if they would like to enroll their personal device into management in order to gain access. For unmanaged devices, we provide an alternative access in the form of a rich desktop environment through Windows Virtual Desktop (WVD). Leveraging WVD and Azure Firewall, we can control who accesses information via identified devices. Moreover, through the nature of this virtualized environment, we effectively prohibit the export of information from our Microsoft-controlled environments while assuring our users' productivity.

An additional risk, which is related to supply chain, is the manageability of devices not owned by the organization but being managed by another organization. A key aspect of this risk is the technical limitation preventing two device management solutions from being deployed on one system. For example, a partner device enrolled in the company's device management system and using the policies defined by the company could not also be enrolled in the Microsoft device management system and use the Microsoft policies. Furthermore, there would be no guarantee or real-time visibility into the compliance state of those policies. As a result, the options are:

1. Assume an implicit trust that the partner device is managed and has the appropriate policies deployed and monitored for health.
2. Put an expectation in place that the partner must unenroll their device from their company management, which would make their device unable to access their own company resources. A security risk associated with remote workforces, as well as third-party vendors and partners, is a personally owned, unmanaged device used from home that accesses organizational resources.
3. Migrate the partner to a virtualized experience, providing them the ability to access resources where the data and controls continue to be managed by the organization, and their company policies continue to be managed on the device.

Learn more:

Security guidance for remote desktop adoption 2020/04/16

Enable a remote workforce by embracing Zero Trust security

Implementing Zero Trust with Microsoft Azure: Identity and Access Management 2020/01/21

Zero Trust Strategy: What Good Looks Like 2019/11/11

Devices and patch management

While many IT departments supported remote work as an option, systems management strategies had been based on the assumption that hardware assets would typically return to the corporate network after only a short leave. This assumption, combined with a desire to reduce the load on VPN concentrators and to avoid unapproved updates, led to policies that prevent laptops from downloading updates when not on the corporate network, and then only from on-premises patch management systems. In other words, policies controlling settings were based on architecture decisions made for a workforce expected to return to the office on a regular basis and weren't even being refreshed in some of these systems. As the COVID-19 outbreak stretched into its third month, many organizations were left struggling with multiple rounds of missing critical patches. Having a remote workforce will necessitate shifts in how IT departments handle patches, as systems management can't rely on assets returning regularly to the corporate network.

IT staff should review patching polices and compliance regularly. If systems are governed by policies that prevent them from receiving new updates when not on the internal network, or are unable to receive new policies, consider alternate methods to deliver policy changes. Cloud-based management solutions and configuration scripts deployed via VPN logons are some of the options available. These alternatives can enable refreshed settings on devices, allowing them to download patches directly from internet sources such as Windows Update instead of waiting to physically return to the organization's internal network.

Attacks on infrastructure: DDoS

DDoS attacks are some of the largest availability and security issues facing customers who are moving their applications to the cloud, making these attacks a significant concern for a remote workforce. A DDoS attack attempts to exhaust an application's resources, making the application unavailable to legitimate users. This type of attack can be costly to the business in terms of time, resources, and even customer attrition.

DDoS attacks can be targeted at any endpoint that's publicly reachable through the internet. While the technology to set up DDoS attacks continues to increase in sophistication, they remain relatively simple and low cost to launch. This low barrier for use is driving up the frequency and ease with which bad actors can wreak havoc on both businesses and users.

Cybercriminals are leveraging the surge in internet traffic owing to the COVID-19 outbreak (with more online meetings, education settings, and other forms of virtual communication) to make DDoS attacks easier, because when trying to disrupt a site (such as banking, hospitals, or governments), the attackers don't have to generate as much traffic for the disruption. In addition to utilizing the increase in traffic, cybercriminals are looking to blend in with the traffic stream to create attack traffic that's harder to distinguish from bona fide traffic.

DDoS attacks are often used as a smoke screen, obfuscating more malicious and harmful infiltrations of an organization's resources. Disruption might be intended solely to bring down a site or to keep the security and IT personnel distracted from the cybercriminal's actual intended target. This kind of misdirection is a useful tool for cybercriminals. They might attack the "front door" of a company while working on a back-end server that they find, keeping the security team busy with what appears to be a higher urgency issue of, say, keeping the company website up. Furthermore, since more personnel are working from home, responses might not be as efficient as they were previously.

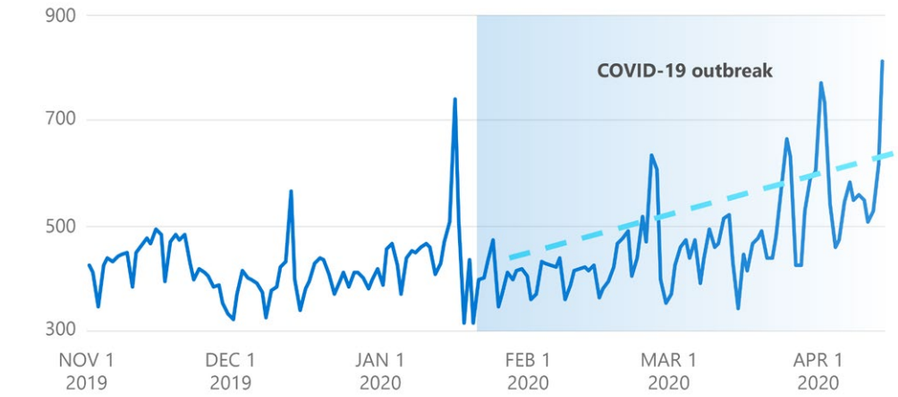
In analyzing cloud-based incoming attacks as seen through Azure DDoS protection, Microsoft threat researchers found that such attacks have been significant in terms of volume and occur on a regular basis. Microsoft observed an increase in DDoS attacks during March 2020, following the COVID-19 outbreak, continuing on an upward trend through April. The company mitigated 600 to 1,000 unique DDoS attacks every day in March, or approximately 50% more than pre-COVID-19 levels.

In developing a strategy for DDoS protection, make sure your cloud and service providers' DDoS protection is enabled.

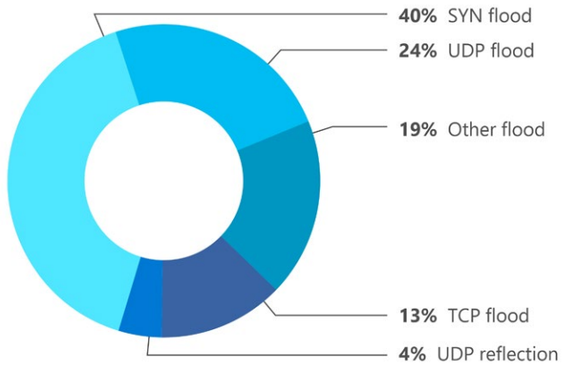
Learn more about best practices in DDoS protection:

Azure DDoS Protection: Designing resilient solutions 2018/10/18

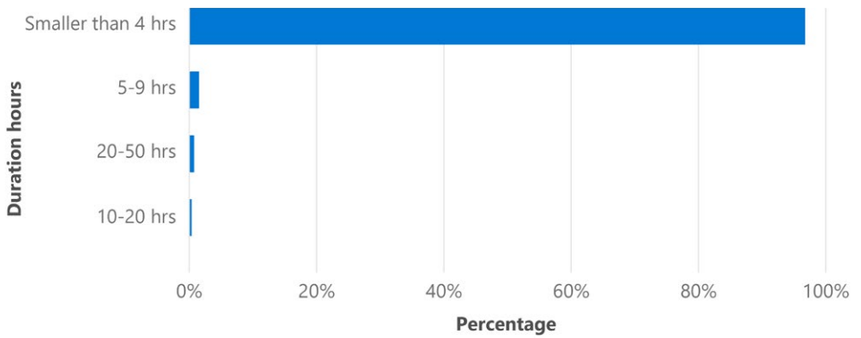
Number of DDoS attacks during COVID-19 outbreak



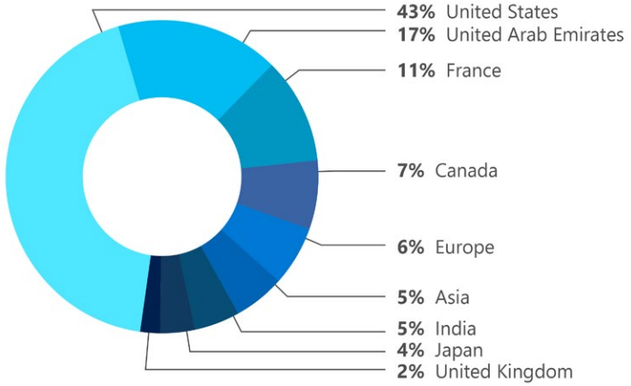
DDoS attack type (January–June 2020)



96.88% of attacks are of small duration (January–June 2020)



Destination region distribution (January–June 2020)



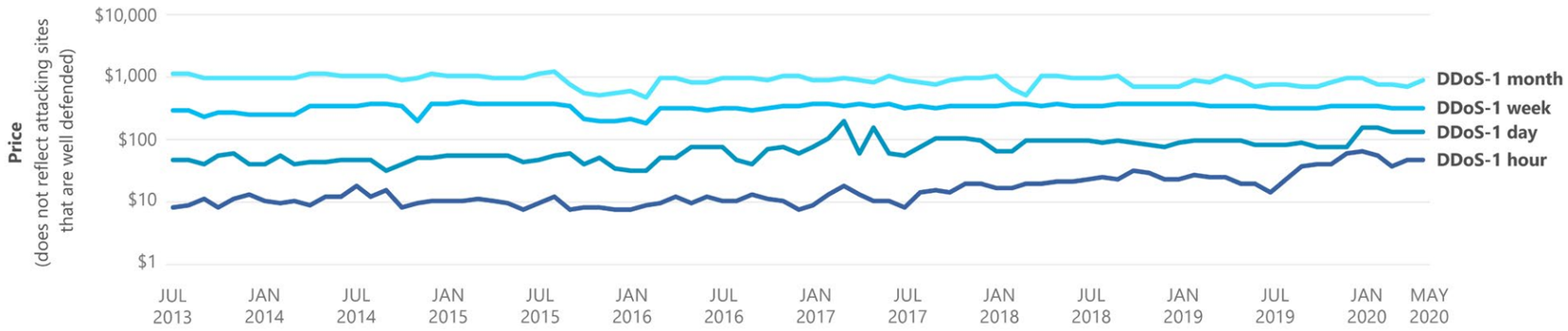
Cybercriminal DDoS services

Professional cybercriminal DDoS services are prevalent on the dark web. Fees paid for these services vary based on factors such as the defenses of the sites being attacked, the type of DDoS attack, the bandwidth needed to conduct the attack, and the regional forum from which the attacker buys the service. In May 2020, the average price of a one-day DDoS attack service was $134.09, with some services offered for as little as $15.00 and the most expensive being $416.67.

As with most mature cybercriminal services, professional DDoS services have reached equilibrium between supply and demand, which has resulted in relatively stable prices over the last seven years, with the exception of shorter duration attacks.

- The average price for DDoS attacks of approximately *one hour* has increased from $14.71 in July 2019 to $48.63 in May 2020.
- The average price for DDoS attacks of approximately *one day* has increased from $74.97 in November 2019 to $134.09 in May 2020.

Trends and approximate average price for cybercriminal DDoS attack services



Data sensitivity, compliance, and protection: Information rights management

Organizations are constantly on the watch to make sure the sensitive information throughout their entire data estate is protected. This concern is driven by a confluence of:

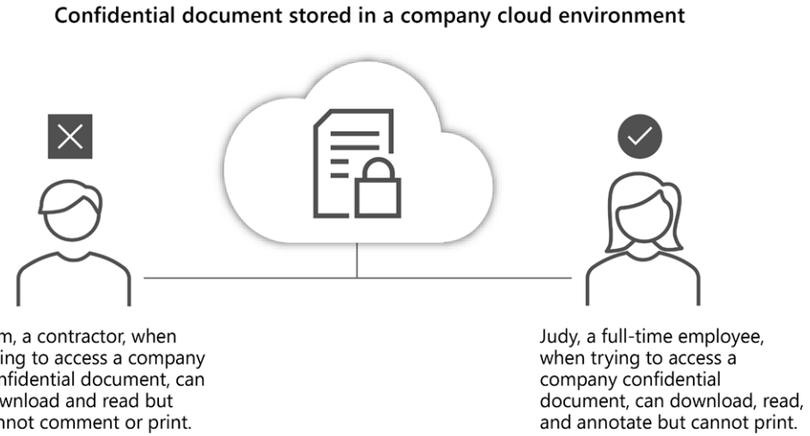
- Growing privacy awareness in customers.
- Increasing government regulation of how sensitive data is handled.
- Needing to ensure that only the right people and system processes have access to sensitive information—and only when needed.

Any new data estate added to an organization must be accounted for as a new variable. Security teams must work continuously to discover sensitive data and to classify, label, and protect it from all the various ingress and egress points in an ever-growing perimeter.

With remote workforces becoming the new paradigm, organizations' information rights management (IRM) practices need to permeate further into their remote working behaviors and solutions, especially when working with sensitive information. This information can range from new marketing campaigns and product design to patents, financial information, or images.

With workforces becoming more remote, more teams need to collaborate on these vital assets without being physically together. Organizations are rising to the challenge with increased use of IRM to enforce policies aimed at protecting confidential information and intellectual property.

With the changes in the workforce environment, we've observed IRM use growing at a steady rate. As of June 2020, there were 22 million monthly active users, with an increase of more than 1 million in just three and a half weeks. An initial dip in use during the COVID-19 outbreak as schools and universities closed was offset by the many enterprises encrypting their content, most notably system integrator customers.

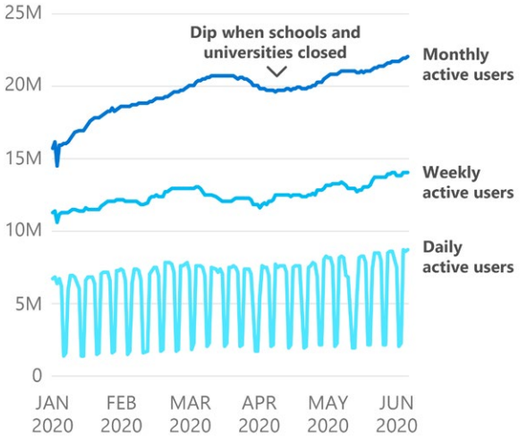


Learn more:

https://docs.microsoft.com/en-us/azure/information-protection/

https://docs.microsoft.com/en-us/Exchange/policy-and-compliance/information-rights-management?view=exchserver-2019

IRM active users



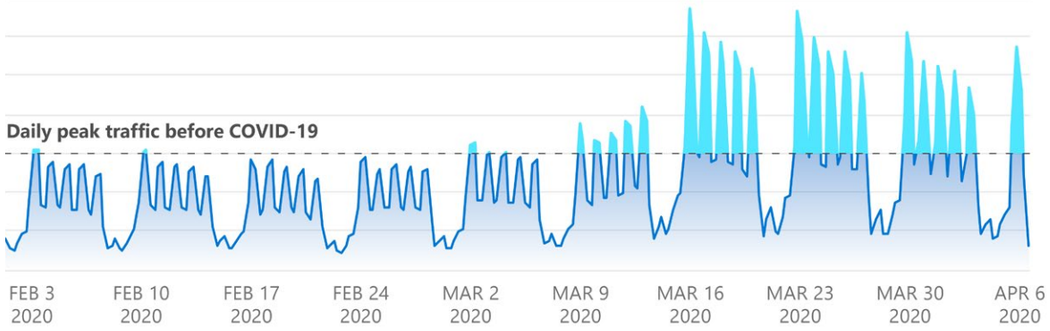
People

Identity and access management

Azure Active Directory (Azure AD), Microsoft's identity platform, is used by organizations for authenticating their users to provide access to their applications and infrastructure. As the first line of defense to better protect users from account compromise, organizations leverage Azure AD's MFA capabilities. MFA is a process in which the system prompts a user for an additional form of identification during sign-in, such as entering a code sent to another device they own or providing a fingerprint scan. Organizations set authentication policies to define the conditions under which users can access resources. For example, a policy might enforce MFA when a user logs in from a new device, new location, or other criteria that enforce their security posture.

As a result of companies closing office access and limiting travel, user authentication patterns have changed. We saw an approximate twofold increase in MFA-enablement requests after the onset of the COVID-19 outbreak, as work-from-home policies were enacted. The higher level of MFA traffic continued as many organizations adopted the use of MFA to support a workforce that was no longer on company networks.

Weekly MFA enablement request volume, February 3–April 6, 2020

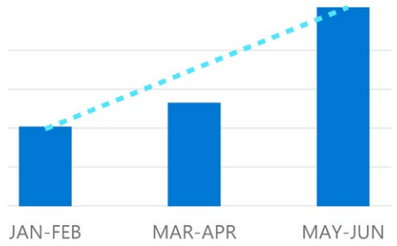


Microsoft data indicates that the vast majority of Microsoft enterprise accounts that were compromised didn't use multi-factor authentication

Identity-based attacks

Azure AD saw an increase in identity-based attacks using brute force on enterprise accounts during the first half of 2020. Cybercriminals appeared to leverage the disruption caused by organizations' COVID-19 response and shift to remote workforces to hide their activities.

Password brute force attempts against Azure AD accounts



Strong authentication, such as passwordless or MFA, are the most effective means of defending against these kinds of attacks.

Most of these attacks happen on legacy authentication protocols, such as POP, SMTP, IMAP, and MAPI, which don't support MFA. Microsoft's research of previous identity-related attacks indicates that more than 99% of password spray attacks use legacy authentication protocols, and more than 97% of credential stuffing attacks use legacy authentication. Given the frequency of passwords being guessed, phished, stolen with malware, or reused, it's critical for people to pair passwords with some form of strong credential. Therefore, disabling legacy authentication and enabling MFA is a critical call to action. Azure AD accounts in organizations that have disabled legacy authentication experience 67% fewer compromises than those in which legacy authentication is enabled.

Insider threats

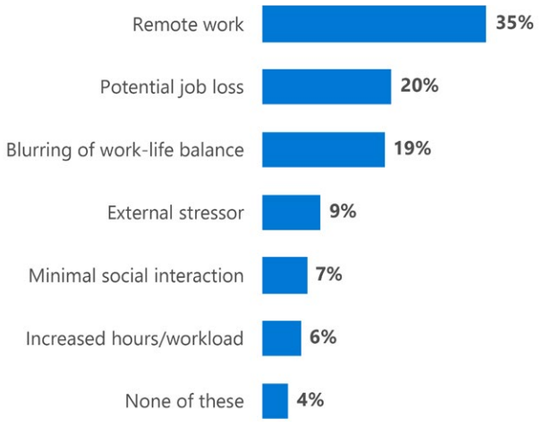
Countless security officers across the world are now asking themselves, "Is my organization effectively prepared to identify and remediate increasing insider risks?" One reason for this question is COVID-19 and the subsequent rapid digital transformation it has forced organizations to undertake. According to a recent survey, the lives of up to 300 million information workers worldwide have been upended, and many are now working remotely with limited resources and increased stress.[46] These employees are not only logging into enterprise environments and line-of-business applications but also accessing, editing, and sharing sensitive data.

Organizations face a broad range of risks from insiders

While some of these activities might happen via a company-managed device, owing to the rapid progression of the pandemic, many information workers are working from their personal PCs or other shared devices. Employees working from home can be subject to increased distractions, such as shared home workspaces and remote schooling for children. According to the SEI CERT institute,[47] user distractions are the cause of many accidental and non-malicious insider risks.

In addition, the current environment has significantly increased stressors such as potential job loss or health and safety concerns, which might lead to some employees participating in malicious activities, such as stealing intellectual property. An additional stressor to consider is the impact of the physical isolation and limited social interaction that has resulted from the abrupt shift to a remote workforce.

Stressors that increase insider risks in a remote work environment[48]



[46] *https://www.bcg.com/publications/2020/covid-remote-work-cyber-security.aspx*

[47] *https://www.sei.cmu.edu/news-events/news/article.cfm?assetId=638958*

[48] *Based on 150 responses from security decision makers in the United States, representing organizations of 300+ employees (June 2020)*

Understanding and managing the complexities of insider risk requires a proactive approach, incorporating organizational, risk management, and cultural considerations. We consider the following elements to be critical to an insider risk management program:

Leading practices for a modern insider threat program

Transparency
Balance employee privacy versus the organization's risk

Cloud first intelligence
Leverage machine learning to identify anomalous, high-risk activity

Dynamic integration
Improve detections with integrated investigative work flows that drive consistency and provide feedback

Insider Risk Management

In talking with customers around the world, we've heard that the move to remote work has amplified the need for an end-to-end insider risk solution.

In a recent survey conducted by Microsoft, 73% of CISOs indicate that their organization encountered leaks of sensitive data and data spillage in the last 12 months, and they plan to spend more on insider risk technology owing to the COVID-19 pandemic.[49]

Customers we spoke to were using a fractured and expensive approach in trying to identify insider risks. They captured signals by using a user activity monitoring solution and fed these signals into a separate user entity behavior analytics solution, with the hopes of finding the proverbial needle in the haystack. We know from our own experience in attempting to deploy these complex types of solutions at Microsoft that not only is the approach not scalable, but often it results in a lot of noise, as they lack enrichment, owing to limited visibility into the sensitivity of the data and lack of broader context.

While having broad visibility into signals from end-user activities, actions, or communications is important, when it comes to effectively identifying the risks, the quality of signals matters most.

On devices, we continue to leverage the agentless capture of signals from Windows 10 endpoints to deliver new signals, including:

- Files transferred to network shares.
- Files uploaded to unallowed domains.
- Files downloaded from enterprise sites.
- Files being renamed.

We're leveraging Microsoft Defender Advanced Threat Protection to capture security signals on the endpoint about:

- Defense evasion (disabling MFA).
- Installation of unwanted software.

In the cloud, native integration with Microsoft 365 allows us to capture billons of additional signals from Teams, SharePoint, and Exchange, including:

- Sharing files, folders, or sites from SharePoint Online to unallowed domains.
- Downloading content from Teams.
- Emailing outside the organization to unallowed domains.

Finally, one of the key early indicators as to whether someone might choose to participate in malicious activities is disgruntlement. We're further enhancing our native HR connector to capture signals related to employee performance which might indicate disgruntlement, including whether the user:

- Is on a performance improvement plan.
- Was recently demoted.
- Had a poor performance review.

Successful insider risk management programs are built on a framework to identify insider risks with integrated processes which enable collaboration between key stakeholders across the organization to take action. It's not just a security issue, but one that requires collaboration between security, HR, and legal teams to ensure insider risks are identified and acted upon in a manner that's compliant with regulatory employment requirements.

Learn more:

Insider risk management in Microsoft 365 2020/07/21

What's new in Microsoft 365 Compliance and Risk Management 2020/06/11

Washington Post: The risks from within 2020/03/11

[49] *Based on 150 responses from security decision makers in the United States, representing organizations of 300+ employees (June 2020)*

Enterprise resilience: The new reality

The circumstances of the extended COVID-19 pandemic, and the corporate response to it, present an opportunity to forge resilience into the enterprise. We must fold the lessons we've learned—and the various strategies that we deployed to survive the pandemic—into our broader enterprise vision.

To some degree, the enterprise response to COVID-19 changed our operational procedures as well as the very vocabulary that we use to describe the reactive measures and lessons learned. Terms such as "extended security boundary," "pandemic resilience," and "human infrastructure" are now the subject of executive analyses and board briefings around the globe. This new vocabulary highlights the emergence of critical areas of resilience that enterprises traditionally didn't scrutinize.

Based on measures that have been put into place to improve our security posture and incorporate the lessons learned since the outbreak of COVID-19, three critical areas have arisen to guide the drive toward enterprise resilience. All three are anchored in planning for the unanticipated.

Critical area 1

Extending the enterprise security boundary beyond the on-site perimeter

The unanticipated shift of such a large percentage of the workforce to a remote work setting hasn't only transitioned work, data, and intellectual property from enterprise-managed facilities to workers' homes, but it has also extended the security boundary beyond the enterprise facility's perimeter. Security teams must now worry about the increased risk of exfiltration or leakage of sensitive data. They must be prepared to respond to breaches in the extended perimeters of the remote workforce's kitchens and home offices. They must ensure the security of assets, services, and infrastructure that extend beyond the corporate perimeter, and possibly even beyond the management and control of corporate policy. They must ensure the privacy and confidentiality of data and any associated data subject rights, and they must protect corporate intellectual property that's distributed across potentially less secure home internet connections and varying grades of Wi-Fi systems and communication channels.

The strategies to support an extended security boundary include expanding VPN capabilities to implementing Zero Trust networking with capabilities such as stronger identity management, device management and device health, alternate access for unmanaged devices, stronger network segmentation and isolation of resources, and checks on application health.

Critical area 2

Prioritizing resilient performance

Just as the pandemic challenged health officials to find ways of protecting public safety, its economic implications stretched corporate leaders to contemplate how to sustain productivity through the massive lift-and-shift exercise where workloads and workforces moved away from their facilities. To better understand what needed to be moved, corporations identified critical services and processes to ensure they weren't abandoned by personnel who needed on-site or break-glass access—in other words, a real-time demonstration of enterprise preparedness. In continuing to learn from this exercise in preparedness, enterprises need to take a closer look at the productivity and performance of critical services and processes.

Critical area 3

Validating the resilience of the human infrastructure

With such a large percentage of the workforce physically distant from management and other coworkers, the potential of new risks around workforce attrition, disengagement, and divestment need to be examined in this new context. Corporate human resource departments are shifting their focus from discussion of work-life balance to work-life integration. A key takeaway in the area of human infrastructure is to embed practices that measure the sustainability of the human capital that the enterprise has invested in, including its workforce, suppliers, partners, and customers.

These three critical areas can serve as the basis for reinforcing enterprise strategy if their core constructs are already included in the vision. If they aren't, they can be readily dovetailed into the vision and strategy as learnings from the COVID-19 pandemic and the response to it, in order to demonstrate and ensure enterprise resilience.

Go to Actionable Learnings

Learn more: *The future of business resilience*

4

Actionable learnings

Steps you can take today

Contributing teams at Microsoft

Steps you can take today

Based on the information learned from the research and intelligence described in this report, we recommend organizations take a proactive approach to shoring up their security and resiliency by using the following controls. While all these learnings are valid and beneficial recommendations for a cybersecurity program, we've selected **our most salient points** for each chapter and mapped them with icons as shown below.

State of Cybercrime
Top recommendations

Nation State Threats
Top recommendations

Security and the Remote Workforce
Top recommendations

Actionable learnings	Chapter mapping
Adopt MFA Multi-factor authentication can stop credential-based attacks dead in their tracks. Without access to the additional factor, the attacker can't access the account or protected resource. MFA should be mandatory for all admin accounts and is strongly recommended for all users. The preferred method is to use an authenticator app rather than SMS or voice where possible.	
Go passwordless For organizations that use modern technologies like Windows 10, we recommend going passwordless by using face authentication, fingerprints, or a PIN code. For organizations with applications or workloads that can't be transitioned to passwordless, we recommend adopting a secure password management solution, such as a password locker or vault and requiring that employees use unique, randomized passwords for access to all sensitive information and on all servers and devices, including IoT and IoT controllers and network infrastructure such as switches, routers, and firewalls.	

Actionable learnings	Chapter mapping

Use good email hygiene

Because 90% of attacks start with an email, preventing phishing (and its voicemail- and text-based variants, vishing and SMiShing) can limit the opportunity for attackers to succeed. Email hygiene platforms that incorporate filtering on the way in and link checking, like Safelinks, when clicked (the way out) provide the most comprehensive protection. Consider limiting or disabling autoforwarding for email unless you have a strong business need for it.

Modernize VPN architectures

Virtual private networks enable secure communications back to a central point. However, most organizations have transitioned significant amounts of workloads to the cloud, making it unnecessary to point 100% of traffic back to the corporate VPN concentrator. A split tunnel allows cloud traffic to connect to cloud resources securely, reducing load and overhead on the corporate VPN.

Patch apps and systems

Vendors are continually making security improvements and fixes, so it's important to ensure that apps or platforms are using the most up-to-date versions, including patches for existing VPN architectures. Ransomware operators and nation state actors have found network devices like gateway and VPN appliances to be a practical target for intrusion. Apply all available security updates for VPN and firewall configurations. Monitor and pay special attention to remote access infrastructure. Any detections from security products or anomalies found in event logs should be investigated immediately. In the event of a compromise, ensure that any account used on these devices has a password reset because the credentials could have been exfiltrated.

Manage configuration changes

Misconfigurations are another prominent attack vector and an example of how small changes can result in big problems. Implementing a robust change management program enables companies to review changes before they're made and confirm that the change won't open a new attack path or otherwise put the organization at risk.

Actionable learnings	Chapter mapping

Implement a secure software development lifecycle

Every organization is engaged in software development, whether they write it themselves or purchase it from a vendor. Even if robust controls are in place lower in the stack, the organization is at risk if the app layer isn't secure. We recommend a robust software development lifecycle that includes threat modeling, design reviews, and static and dynamic application testing, in addition to penetration testing in production.

Take a 3-2-1 approach to backups

Backups are essential for resilience after an organization has been breached. Apply the 3-2-1 rule for maximum protection and availability: 3 copies, original + 2 backups, 2 storage types, and 1 offsite or cold copy.

Monitor cross-cloud security

Most companies today use multiple cloud providers, and without visibility across all the clouds, including the IaaS and SaaS layers, organizations don't have a complete view of their risk profile. Modern SOCs should leverage SIEM tools that can collect and analyze signals from all clouds in use and leverage a cloud access security broker (CASB) as appropriate for audit and control of SaaS.

Limit access with least privilege

To limit the risk of intentional, and unintentional, insider risk, practice the principle of least privilege and maintain good credential hygiene. Avoid the use of domain-wide, admin-level service accounts. Restricting local administrative privileges can help limit installation of remote access trojans and other unwanted applications.

Actionable learnings	Chapter mapping

Leverage machine learning to increase fidelity and reduce alert fatigue

The explosion of data and signals has overwhelmed many SOCs and analysts. Look for modern SIEMs that use technologies like machine learning to reduce the false positives and bubble up the alerts that really matter. ML can also help identify attackers that are using behavioral-based attacks that might not be detected by traditional, rules-based systems.

Closely monitor legacy, certified, and industrial control systems

Keeping operating systems and apps patched and up to date is the best way to protect them. However, some systems can't be upgraded or patched owing to restrictions like certification levels or business limitations. For devices and apps that can't be patched or upgraded, additional monitoring can be tuned to keep a closer eye on access and behavior of legacy devices. Segmentation (covered below) is also an important control here.

Slow attacks with network segmentation

When vulnerabilities are found on a number of systems, an attack can rip through a flat network in a matter of minutes. Additionally, devices and systems with highly sensitive data often need to be protected at higher levels of control. Segmenting networks provides security to limit the spread of attacks and enables organizations to place stricter controls in front of sensitive enclaves. Segmentation can also help to protect unpatched or unpatchable systems from attack.

Actionable learnings	Chapter mapping

Manage the convergence of OT and IT

Traditionally cordoned-off OT networks are converging with IT networks. Many OT networks relied on segmentation and access controls for security, but as they're opened to the internet and the cloud, they're increasingly overlapping with IT. Organizations must have a strategy in place to protect OT post-convergence.

Secure IoT and IIoT

The IoT and the Industrial Internet of Things (IIoT) are pushing the boundary of security management. Many IIoT devices are legacy brownfield and dependent on older OT-type protections. For comprehensive security, organizations must bring IoT and IIoT into their security program, with management and governance for the legacy devices and systems, and by building out modern, trusted solutions for greenfield deployments. For more on this approach, see The Seven Properties of Highly Secure Devices.

Know your perimeter

Endpoints are a common target for attackers. Weak endpoints can become beachheads for network infiltration and reconnaissance activity. With the COVID-19-driven shift to remote workforces, endpoints are farther away from the corporate network and sprawling with new software installs for video conferencing and other collaborative activities. A robust endpoint detection and response solution must be installed on endpoints to keep them protected, and some form of endpoint management is strongly recommended.

Limit perimeter exposure

Understand and control perimeter exposure. Attacks often start with devices that aren't in asset inventories but are still members of the primary Active Directory domain. These devices often have matching local admin passwords or service accounts with highly privileged domain credentials logging on to them. To limit perimeter exposure, an organization needs an up-to-date inventory that includes all members of the Active Directory domain, and these devices should be monitored and managed. Perimeter exposure can also be managed on the network by implementing segmentation (via switches, routers, firewalls) and throughout the cloud with solutions such as CASBs.

Actionable learnings	Chapter mapping

Build a third-party risk program

Partners, suppliers, and contractors interact with data and applications connected to an organization's corporate environments, and attackers are increasingly targeting providers, like MSSPs, in order to gain access to the MSSP's clients' credentials and networks. This means that your partner's risk is your inherited risk, and it should be managed through robust service level agreements, attestations, and shared assessments like SSAE 18 SOC 1 and SOC 2, PCI-DSS, GDPR, and ISO 20001.

Invest in user training (and keep training)

Users can be the weakest link or the first line of defense. When security awareness is institutionalized at an organizational level, end users can be the early responders to activity that might indicate compromise. Train users on what attacks look like and provide them with a way to report unusual activity.

Adopt a Zero Trust mindset

Instead of assuming everything behind the corporate firewall is safe, assume breach and verify each request as though it originates from an open network. Regardless of where the request originates or what resource it accesses, every request is fully authenticated, authorized, and encrypted before granting access.

Contributing teams at Microsoft

The insights in this report, as well as the actionable learnings above, have been provided by a diverse group of security-focused individuals who work across dozens of different teams at Microsoft. Collectively, their goal is to protect Microsoft, Microsoft customers, and the world at large from the threat of cyberattacks. We're proud to share these insights in a spirt of transparency, with a common goal of making the digital world a safer place for everyone.

Cyber Defense Operations Center (CDOC)

A cybersecurity and defense facility that brings together security professionals from across the company to protect Microsoft's corporate infrastructure and the cloud infrastructure to which customers have access. Incident responders sit alongside data scientists and security engineers from across Microsoft's services, products, and devices groups to help protect, detect, and respond to threats 24x7.

Customer Security and Trust (CST)

A cross-disciplinary team driving continuous improvement of customer security in our products and online services. Working with engineering and security teams across the company, the mission of CST is to ensure compliance and enhance security and transparency to protect our customers and promote global trust in Microsoft. They formulate and advocate cybersecurity policy globally, advance digital peace through multi-stakeholder collaboration, focus on digital safety to protect customers from harmful online content and collaborate with public and private organizations to disrupt cyberattacks and support deterrence efforts.

Detection and Response Team (DART)

A Microsoft team whose mission is to respond to security incidents and help Microsoft customers become cyber-resilient. DART leverages Microsoft's strategic partnerships with security organizations around the world and with internal Microsoft product groups to provide the most complete and thorough investigations possible. DART's expertise has been leveraged by government and commercial entities around the world to help secure their most sensitive, critical environments.

Digital Security & Risk Engineering (DSRE)

A Microsoft organization developed with a mission to enable Microsoft to build the most trusted devices and services, while keeping our company safe and our data protected. Across the company, DSRE is continually evolving the security strategy and taking actions to protect Microsoft assets and the data of our customers.

Digital Security Unit (DSU)

A team of cybersecurity attorneys and strategic cyber intelligence analysts who provide legal, operational, geopolitical, and technical subject matter expertise to protect Microsoft and our customers. DSU's analysis and proposed solutions to complex digital security problems help to build trust in Microsoft's enterprise security capabilities and defenses against advanced cyber adversaries worldwide.

GitHub Security Lab

An open-source software-focused security research team. Its mission is to help secure the world's code and build bridges between the security research and software development communities through contributions including security research, tooling, and meetups.

IoT Security Research Team (also known as Research 52)

A team that's composed of domain-expert reverse engineers and data scientists. The team continuously performs reverse-engineering and analysis of large amounts of data related to IoT threats and threat actors to gain better visibility into the IoT landscape and uncover related trends and campaigns.

Microsoft Defender Team

Microsoft's largest global organization of product-focused security researchers, applied scientists, and threat intelligence analysts. The Defender Team delivers innovative detection and response capabilities in Microsoft 365 security solutions and Microsoft Threat Experts.

Microsoft Digital Crimes Unit (DCU)

A team of attorneys, investigators, data scientists, engineers, analysts, and business professionals that fight cybercrime globally through the innovative application of technology, forensics, civil actions, criminal referrals, and public/private partnerships, while protecting the security and privacy of our customers.

Microsoft Security Response Center (MSRC)

Part of the defender community on the front line of security response evolution. For more than 20 years, MSRC has been engaged with security researchers working to protect customers and the broader ecosystem. An integral part of Microsoft's CDOC, MSRC brings together security response experts from across the company to help protect, detect, and respond to threats in real time.

Microsoft Threat Intelligence Center (MSTIC)

Microsoft's centralized team focused on identifying, tracking, and collecting intelligence against the most sophisticated and advanced adversaries impacting Microsoft customers, including nation state threats, malware, phishing, and more. The threat intelligence analysts and engineering teams in MSTIC work closely with Micrsoft security product teams to both develop and refine high-quality detections and defenses across our security product portfolio.

Glossary

Acronyms and terminology

Acronyms and terminology

advanced persistent threat
An adversary that possesses sophisticated levels of expertise and significant resources, which allow it to create opportunities to achieve its objectives by using multiple attack vectors (e.g., cyber, physical, and deception). This term originated as a way to refer to nation state actors but has become a general term to describe organized adversaries.

adversary
An individual, group, organization, or government that conducts or has the intent to conduct cybersecurity attacks.

alert
A notification that a specific attack, anomaly, or suspicious activity has been detected or directed at an organization's information systems. Alerts frequently trigger investigations by security operations/analysts.

analyst
Also known as cybersecurity analyst, a common role within an organization's SOC team that investigates, alerts, or hunts for adversary activities.

asset
An entity of value that could take the form of a person, structure, facility, information and records, IT systems and resources, material, process, relationships, or reputation.

attack
Any attempt to defeat the security assurances of a system or data including confidentiality, integrity, or availability.

attack path
The steps that an adversary takes or might take to plan, prepare for, and execute an attack.

attack pattern
Similar cyber events or behaviors that might indicate an attack has occurred or is occurring, resulting in a security violation or a potential security violation.

attack surface
An information system's characteristics that permit an adversary to probe, attack, or maintain presence in the information system.

attacker
An individual, group, or organization that executes an attack. An attacker might also refer to an adversary or an individual attack operator.

attachment (as in malicious email)
During phishing campaigns, cybercriminals attempt to trick users into selecting an email attachment, which then downloads a malicious executable, infecting the user's computer or mobile device, or, upon opening the attachment the user might be redirected to a fraudulent login site. Attachments can come in various forms, such as a Microsoft Office document, a PDF file, .zip files, etc.

authentication
The process of verifying the identity of an entity (user, process, or device).

authorization
A process of determining whether a subject is allowed to have the specified types of access to a particular resource. This action is typically done by evaluating applicable access control information such as access control lists. In modern cybersecurity approaches, authorization could also incorporate other risk factors such as behavioral analytics and evaluation of threat intelligence.

availability
One of three primary cybersecurity assurances, the property of being accessible and usable upon demand.

banking trojan
A type of malware designed to obtain credentials to banking and other financial services. These trojans use a variety of techniques, including interception of web communications as users access financial services on infected devices. Many known banking trojans are part of botnets that provide cybercrime organizations with persistent access to large numbers of devices.

blast radius
A machine learning method to identify the most impactful users, based on the level of risk to the organization if they become compromised.

botnet
A collection of computers compromised by malicious code and controlled across a network.

breach
Any incident that results in unauthorized access of data, applications, services, networks, and/or devices by bypassing their underlying security mechanisms. A security breach occurs when an individual or an application illegitimately enters a private, confidential, or unauthorized logical IT perimeter.

brownfield
Existing deployed IoT/IIoT devices that might not have modern hardware or functionality. For example, these devices might not have support for over-the-air updates or remote administration.

brute force
An attack technique that uses systematic guessing, static or dynamic lists of passwords, dumped credentials from previous breaches, or other similar methods to forcibly authenticate to a device or service.

business email compromise (BEC)
A technology-facilitated social engineering scam that targets business email accounts and enables cybercriminals to unlawfully redirect and intercept money wires, exfiltrate documents, and launch other cybercrime. BEC is often initiated through some form of credential phishing.

ciphertext
Data or information in its encrypted form used primarily by cryptology experts. Sometimes referred to as encrypted data.

cloud access security broker (CASB)
A CASB is a software tool or service that sits between an organization's on-premises infrastructure and a cloud provider's infrastructure. They might also provide other services such as credential mapping when single sign-on isn't available.

confidentiality
An assurance that information isn't disclosed to unauthorized users, processes, devices, or other entities. One of the three primary cybersecurity assurances (confidentiality, integrity, and availability).

critical infrastructure
The systems and assets, whether physical or virtual, so vital to society that the incapacity or destruction of such might have a debilitating impact on the security, economy, public health or safety, environment, or any combination of these areas.

cybersecurity
The discipline of preserving and rapidly restoring the primary security assurances of confidentiality, integrity, and availability for systems, data, and identities.

data estate
The procedures, services, and infrastructure used to manage corporate data in the digital estate.

data loss prevention (DLP)
A set of procedures and mechanisms to stop sensitive data from leaving a security boundary.

denial of service
An attack that prevents or impairs the authorized use of information system resources or services. A distributed denial of service (DDoS) is a type of attack that uses multiple networked machines to overwhelm a host connected to the internet, temporarily or permanently disrupting service or preventing access.

digital estate
An abstract reference to a collection of tangible owned assets. Those assets include virtual machines (VMs), servers, applications, data, containers, apps, and so on. Essentially, a digital estate is the collection of IT assets that power business processes and supporting operations.

drop account
An email account set up by a criminal to receive credentials provided by an unsuspecting victim.

encryption
The process of transforming plaintext into ciphertext.

event
An observable occurrence in an information system or network.

exfiltration
The unauthorized transfer of information from an information system.

exploit
A technique to breach the security of a network or information system in violation of security policy.

exposure
The condition of being unprotected, thereby allowing access to information or access to capabilities that an attacker can use to enter a system or network.

firewall
A capability to limit network traffic between networks and/or information systems.

fusion
A technology for finding threats that would otherwise fly under the radar. Fusion uses machine learning to combine disparate data from Microsoft and partner datasets, by combining low-fidelity "yellow" anomalous activities into high-fidelity "red" incidents.

greenfield
New or planned IoT/IIoT deployments that support the latest advances in security and technology management.

honeypot
A computer or computer system intended to mimic likely targets of cybercriminals.

human-operated ransomware
A type of ransomware attack that's performed by human operators. During these attacks, human operators use various tools and techniques to compromise and traverse targeted networks, ultimately deploying ransomware on multiple devices on the compromised networks.

hunting
Proactively looking for active adversaries.

identity and access management (IAM)
The methods and processes used to manage subjects and their authentication and authorizations to access specific objects.

indicators of compromise (IOC)
Pieces of forensic data, such as data found in system log entries or files, that identify potentially malicious activity on a system or network.

insider threat
A person or group of persons within an organization who pose a potential risk through violating security policies.

International Organization for Standardization (ISO)
An international standard-setting body composed of representatives from various national standards organizations. Founded in 1947, the organization promotes worldwide proprietary, industrial, and commercial standards.

intrusion
An unauthorized act of bypassing the security mechanisms of a network or information system.

intrusion detection
The process and methods for analyzing information from networks and information systems to determine if a security breach or security violation has occurred.

just in time (JIT)
Provides temporary (measured in hours) elevated access to internal engineers to debug production issues or support customer cases to ensure limited access based on least privilege principles.

kill chain
A cyber kill chain reveals the phases of a cyber attack: from early reconnaissance to the goal of data exfiltration. The kill chain is also a model for identification and prevention of cyber intrusions activity. The model identifies what the adversaries must complete in order to achieve their objective.[50]

machine learning
A type of artificial intelligence focused on enabling computers to use observed data to evolve new behaviors that haven't been explicitly programmed.

macro virus
A type of malicious code that attaches itself to documents and uses the macro programming capabilities of the document's application to execute, replicate, and spread or propagate itself.

maintainer (as in GitHub)
Someone who manages a repository. This person might help triage issues and use labels and other features to manage the work of the repository. This person might also be responsible for keeping the README and contributing files updated.

model inversion
An activity whereby an attacker uses careful queries to recover the secret features used in a machine learning model.

model stealing
An activity whereby an attacker constructs careful queries to recover a machine learning model.

nation state activity group
Cyber threat activity that originates in a particular country with the apparent intent of furthering national interests.

[50] *https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html*

National Institute of Standards and Technology (NIST)
A physical sciences laboratory and a nonregulatory agency of the U.S. Department of Commerce. Its mission is to promote innovation and industrial competitiveness.

npm (as in GitHub)
npm is the package manager for Node.js, and the world's largest software registry.

obfuscation
A method used to hide or obscure an attack payload from inspection by information protection systems.

operator, or attack operator
An individual person who is executing an attack operation. This person might be acting alone, acting on behalf of an organization, or acting in concert with multiple other attack operators in a coordinated campaign.

password spray
High-volume attempts using a large number of common passwords to compromise sourced account information to authenticate and gain access to a network, often leveraging big data algorithms and extensive automation for rapid execution.

phishing
A digital form of social engineering to deceive individuals into providing sensitive information.

phishing kit
A collection of tools assembled to make it easier for individuals with little or no knowledge of phishing practices to launch a phishing exploit.

poisoning attack
Contamination in the training phase of machine learning systems to get an intended result.

Purdue Reference Model, "95"
A reference model for enterprise control, describing a generic view of an integrated manufacturing or production system, expressed as a series of logical levels.

Red Team
A group authorized and organized to emulate a potential adversary's attack or exploitation capabilities against an enterprise's cybersecurity posture.

Red Team exercise or Red Team testing
An exercise, reflecting real-world conditions, that's conducted as a simulated attempt by an adversary to attack or exploit vulnerabilities in an enterprise's information systems.

Remote Desktop Protocol (RDP)
A protocol for remotely connecting to computers running Windows.

resilience
The ability to adapt to changing conditions and prepare for, withstand, and rapidly recover from disruption.

reverse engineering
The reproduction of another manufacturer's product following detailed examination of its construction or composition.

secrets (as in GitHub secrets)
Secrets are tokens, credentials, private keys, or other authentication identifiers that might be used in a service at build or run time. For example, secrets might be used by an application to access an external service.

security development lifecycle (SDL)
The Microsoft SDL introduces security and privacy considerations throughout all phases of the development process, helping developers build highly secure software, address security compliance requirements, and reduce development costs. The guidance, best practices, tools, and processes in the Microsoft SDL are practices we use internally to build more secure products and services. Since first shared in 2008, we've updated the practices as a result of our growing experience with new scenarios, like the cloud, IoT, and artificial intelligence.

security information and event management (SIEM)
An approach to security management that combines SIM (security information management) and SEM (security event management) functions into one security management system. The acronym SIEM is pronounced "sim" with a silent "e."

security operations center (SOC)
A centralized function within an organization employing people, processes, and technology to continuously monitor and improve an organization's security posture while preventing, detecting, analyzing, and responding to cybersecurity incidents.[51]

Security Orchestration, Automation and Response (SOAR)
A solution stack of compatible software programs that allow an organization to collect data about security threats from multiple sources and respond to low-level security events without human assistance.

SMiShing (SMS phishing)
An attack method via a text or SMS message received on a mobile device. An attacker uses SMiShing to trick a user into downloading malware or revealing private information through a fraudulent link.

[51] *https://www.mcafee.com/enterprise/en-us/security-awareness/operations/what-is-soc.html*

spoofing

Faking the sending address of a transmission to gain illegal (unauthorized) entry into a secure system. Website spoofing involves creating a duplicate version of a website that appears to be the original. Hackers use legitimate logos, fonts, colors, and functionality to make the spoofed site look realistic. Even the URL can appear genuine.

supply chain

A system of organizations, people, activities, information, and resources, for creating and moving products including product components and/or services from suppliers through to their customers.

supply chain risk management

The process of identifying, analyzing, and assessing supply chain risk and accepting, avoiding, transferring, or controlling it to an acceptable level considering associated costs and benefits of any actions taken.

threat variant

New or modified strains of an existing virus or malware program; malware family.

vishing (voice phishing)

The telephone equivalent of phishing. It's the act of using the telephone to scam the user into surrendering private information that will be used for identity theft. It can take shape as a phone call or voice message from a live or automated person.

Zero Trust

A security model based on the principle of maintaining strict access controls and not trusting anyone by default, even those already inside the network perimeter.