WHITE PAPER

Cyber Threat Predictions for 2022

An Annual Perspective by FortiGuard Labs

2022 Threat Predictions

This past year saw an unprecedented rise in cyber crime. Research by FortiGuard Labs showed an almost 11x increase in ransomware in the 12 months between July 2020 and June 2021. But our challenge going forward is far more than just the rising number of attacks. We are also seeing an increase in attacks on high-profile targets, including the supply chain attack on SolarWinds and the disruption of Colonial Pipeline and JBS Meats, which affect thousands of organizations and millions of people who have nothing to do with IT.

Fresh Blood

While most attacks continue to exploit known vulnerabilities, cyber criminals have also redoubled efforts to target new ones. Last January, the Chinese state-sponsored group Hafnium, for example, began targeting seven new vulnerabilities in Microsoft Exchange Servers—more than two months before patch availability. While three of these vulnerabilities had previously been identified by Microsoft, four were unknown zero-day vulnerabilities. Hafnium's largely automated attacks targeted unpatched Exchange Servers and then used a devised webshell to control the malware remotely, steal data, and gain unauthorized access to critical systems. Tens of thousands of organizations worldwide were impacted, including U.S.-based organizations such as law firms, defense contractors, labs conducting infectious disease research, and non-governmental organizations (NGOs).

In 2022, we can expect the feeding frenzy for fresh vulnerabilities to continue beyond Hafnium. In response to the high-profile attacks of the past year, many organizations are finally dedicating time to basic cyber hygiene. And as patch efforts catch up on the 1- to 3-year-old Common Vulnerabilities and Exposures (CVEs) most cyber criminals seek to exploit, 2022 will most likely be a banner year for the number of CVEs reported (likely breaking 20,000 for the first time), and attackers will begin to leverage those fresh or zero-day vulnerabilities to target unprepared organizations.

Solutions and Countermeasures:

When zero-day attacks are executed locally, they can be detected using advanced extended detection and response (XDR) for networks or endpoint detection and response (EDR) technologies that block critical malware functions, such as contacting a C2 server to download a malicious payload. Sandboxing technology is another essential tool that should be integrated across the network, including next-generation firewalls (NGFWs), secure email gateways, and cloud-based security solutions whether as a standalone solution or as part of a virtual firewall.

Another effective strategy is implementing automation and analytics tools designed to recognize attack patterns and techniques, such as behavioral analytics and deceptive network strategies and solutions combined with honeypots. And given the increasing speed and complexity of today's zero-day threats, any tools designed to detect and interrupt a threat also need to be augmented with advanced artificial intelligence (AI) systems that have been trained to spot anomalous behavior and that can sift through and correlate log files and indicators of compromise (IOCs) to detect complex, multifront attacks.

Convergence—Advanced Persistent Cyber Crime

Cyber crime is often divided into efforts on the "left" or the "right" hand side of the attack kill chain. On the right are the more familiar attacks, such as building and launching malware to corrupt systems, steal data, or hold networks hostage. These attacks are the traditional domain of most cyber criminals. To the left are things like gaining initial access, performing reconnaissance, and the weaponization of vulnerabilities. Advanced persistent threats (APTs) are on the left because much of its work happens before an attack, such as identifying a vulnerable network, gaining unauthorized access, and remaining undetected for an extended period. APTs are usually associated with malicious organizations with lots of resources, such as nation-states or state-sponsored actors.

As cyber crime incidents increase and more gangs compete for a slice of the profits, we expect to see more "left-hand" investment from cyber criminals. Like nation-state-funded APT groups, these efforts will include spending more time and effort on reconnaissance and discovering zero-day capabilities, which will spur the future increase in CVEs.

In addition to more vulnerabilities being discovered, the attacks that exploit them will become more readily available to other attackers and incorporated into other attack kits. The rise in new vulnerabilities will naturally converge with the growth in Malware-as-a-Service. So, not only will cyber criminals discover and weaponize more zero-day vulnerabilities, but those exploits will also be launched at an exponentially higher rate due to the multiplication factor of numerous threat actor affiliates launching attacks simultaneously.

Solutions and Countermeasures:

These services simply amplify the exploitation of vulnerabilities. Organizations should be aware that an increase in new cyber criminals armed with advanced technologies will increase the likelihood and volume of attacks. Standard tools must be able to scale to address potential increases in attack volumes. They also need to be enhanced with AI to detect attack patterns and thwart threats in real time. Critical tools should include EDR, sandbox solutions augmented with MITRE ATT&CK mappings, anti-malware engines using AI detection signatures, advanced intrusion prevention system (IPS) detection, and NGFWs. Ideally, these tools are deployed consistently across the distributed network (data center, campus, branch, multi-cloud, home office, endpoint) using an integrated security platform that can see, share, correlate, and respond to threats as a unified solution.

New Areas of Exploitation

Traditional targets will be high on the menu. Windows 11 is coming out soon, and there will naturally be new vulnerabilities associated with this OS that threat actors will exploit. But that's just the start. In the coming year, we will also see new areas for exploitation being explored.

- **Lesser targeted systems:** Linux runs many of the back-end systems in most networks but has largely been ignored by the hacker community until recently. Vermilion Strike is a malicious implementation of the critical feature "Beacon" within the infamous Cobalt Strike tool. Cobalt Strike is a "threat emulation" software program used by Red Teams (and threat actors) to demonstrate the risk of a network breach, and Beacon is designed to deliver malicious payloads and establish command and control (C2) connections in Windows environments. Vermilion Strike targets Linux systems with remote access capabilities because it runs entirely undetected in a Linux environment.

- **Linux botnets:** New botnet malware is being written for Linux platforms that can both weaponize devices as well as slow down performance. This further expands the threat surface to the edge of the network and increases the threats that need to be defended there. We expect to see even more of this sort of activity targeting edge devices that cyber criminals have traditionally overlooked, such as edge compute devices and servers.

- **New malicious Linux binaries on Windows:** New malicious binaries have been detected targeting Microsoft's Windows Subsystem for Linux (WSL), a compatibility layer for running Linux binary executables natively on Windows 10, Windows 11, and Windows Server 2019. There has already been some movement in this area in 2021, with detected malicious test files acting as loaders, many containing malicious payloads. There will be more to come in 2022 as Microsoft actively integrates WSL 2 with Windows 11, making Linux a potential new attack vector for targeting Windows devices.

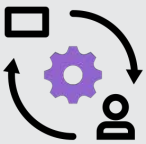
- **Operational technology (OT) networks and other nontraditional targets:** We will see an increase in attacks on nontraditional targets, such as operational technology (OT) systems. According to a recent U.S. Cybersecurity & Infrastructure Security Agency (CISA) report, ransomware attacks are increasingly targeting critical infrastructure and "have demonstrated the rising threat of ransomware to operational technology (OT) assets and control systems." A similar Cybersecurity Advisory from the National Security Agency (NSA) says, "As OT components continue being connected to information technology (IT), IT

exploitation increasingly can serve as a pivot to OT destructive effects." The agency also advised the owners of defense networks to conduct detailed risk analyses of their OT. Traditionally, attacks on OT systems were the domain of highly specialized threat actors, but such capabilities are increasingly being included in attack kits available for purchase on the dark web, making them available to a much broader set of attackers.

- **Quantum targets:** We will begin to see the weaponization of quantum computers, especially quantum encryption. As with any new technology, there will be lots of researchers looking for bugs, vulnerabilities, and bypasses. While this technology is still mainly in the hands of large corporations, universities, and governments, it does not mean it cannot be exploited. Nation-states already gather massive amounts of data, much of it encrypted, from other nations and industries of interest. Very shortly, the power of quantum systems will be used to break the traditional encryption protecting that data, revealing critical information that will be used to execute future/additional attacks.
- **Machine Learning (ML) targets:** We are also likely to see the first attempt to tamper with or evade machine learning (ML)-based systems, especially on external systems facing the internet. Criminals will target internet-facing ML learning nodes located at the network edge. An example of one ML-based exploit would be to surreptitiously train security systems to ignore certain types of attacks. This threat category is severe enough that MITRE has announced a new threat matrix called ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems), designed to help organizations identify and categorize attacks against ML systems.
- **Growth in Crime-as-a-Service (CaaS):** Ransomware as a service has earned billions of dollars for developers and affiliates. New offerings leveraging this business model include a menu of ransomware services, ranging from selling access to pre-compromised organizations, helping with setting and negotiating ransom fees, and laundering money, including cryptocurrency. We should begin to see this model used elsewhere, expanding the CaaS portfolio to include Phishing-as-a-Service and Botnet/SMS-as-a-Service.

We will continue to see this sort of crimeware expansion, especially using destructive ransomware. Cyber criminals have learned that they can make a lot of money reselling their malware online, especially as a service. But rather than directly competing with others offering similar tools, we will also see criminal organizations expand their portfolios to include Linux, OT, and other non-traditional targets, and offerings to target unconventional—and often, less secure—systems. Holding such systems for ransom will be lucrative, but when OT and critical infrastructures are targeted, they can also have dire consequences, including affecting the lives and safety of individuals. These will also become a new attack gateway because as networks are increasingly interconnected, virtually any access point can be exploited to gain entry to the IT network.

Solutions and Countermeasures:



Addressing the expanding attack surface requires developing and deploying a complete security mesh architecture to provide visibility into each attack surface. When endpoint agents, EDR solutions, network-based IPS and anti-malware, firewalls, honeypots, and sandbox solutions work together, they can create internal and external barriers against such attacks. Zero-trust strategies are also incredibly effective against threats targeting new areas of the network, such as areas with no direct IT/OT integration points. They would benefit even more with the addition of AI scanning devices.

Bounty Hunters

One interesting outcome of the rapid growth of cyber crime is that criminal organizations are now finding themselves stepping on each other. Turf wars have already started as malicious organizations compete for a piece of the cyber-crime pie. This isn't new, but current trends indicate that this activity will likely increase, with the end goal being to disrupt their competitors' operations.

Traditionally, threat actors' tactics, techniques, and procedures (TTPs) have been discovered through forensic analysis. But this past summer, a disgruntled "pentester" of the Conti ransomware group leaked insider files to the public, including playbooks, operational how-to documents, and reference files created for Conti affiliates. We should expect more leaks of playbooks and source code.

Hijacking the servers and resources of other cyber criminals has become common enough that some malicious organizations have begun adding digital certificates to their botnet C2 systems. To prevent their infrastructures from being hijacked, they also increasingly require authentication, including multi-factor authorizations (MFA). Along with an increased targeting of digital wallets generally, which we discuss later, we will also see attempts to hijack other gangs' crypto wallets, if that hasn't already happened.

Solutions and Countermeasures:

While the increase in conflicts between cyber criminals may not directly impact business and governmental organizations, leaked threat playbooks can be directly integrated into tools that hunt and protect against attacks. At the same time, there are ongoing red team initiatives to create playbooks based on criminal behaviors and attack fingerprints, so we can expect to see the ingestion and automation of these playbooks into advanced tools like security orchestration, automation, and response (SOAR) that can be integrated into network operations center (NOC) and security operations center (SOC) environments.

Living Off the Edge

The network edge continues to expand, with cloud adoption and work-from-anywhere (WFA) realities seeing significant growth over the past year due to the pandemic. New edges are also being developed, fueled by a growing number of Internet-of-Things (IoT) and endpoint devices, as well as new applications that require real-time computing power. Powered by 5G and AI, smarter edge devices enable the creation of real-time applications such as video processing and analytics, self-driving cars, and automated manufacturing floors. Data collection and processing happen locally at the edge rather than being shuttled back and forth to the cloud or data center in these environments. This allows decisions to be made faster when time to respond is critical.

Last year, we predicted the development of Edge-Access Trojans (EATs) by threat actors designed to target edge environments. This approach has several advantages for a cyber adversary. First, it allows them to collect data and even disrupt critical decisions at the edge of the network, where time sensitivity is often crucial. This would create an entirely new level of urgency to ransomware attacks because disrupting time-sensitive processes often have safety elements, such as critical infrastructure systems. EATs can also be used to corrupt data, which may significantly impact downstream systems that rely on data collected by edge devices. Such edge footholds can also be used to tunnel back to the corporate network.

A new edge-based challenge is also emerging. "Living off the land" allows malware to leverage existing toolsets and capabilities within compromised environments, so attacks and data exfiltration look like normal system activity and go unnoticed. The recent Hafnium Exchange attacks used this technique to live and persist in domain controllers. Living-off-the-land attacks are effective because they use legitimate tools to carry out their nefarious activities.

We expect these two concepts (EATs and living off the land) to converge sometime in 2022. New attacks will be designed to live off-the-edge as edge devices become more powerful, with more native capabilities, and of course, more privileged. Edge malware will monitor edge activities and data and then steal, hijack, or even ransom critical systems, applications, and information while avoiding being detected.

Solutions and Countermeasures:

Defending against these new edge-based threats will require organizations to upgrade end-user devices with advanced EDR technologies along with enhanced access controls, including zero-trust network access (ZTNA). Secure web gateways will also become increasingly vital in protecting the extreme edges of the network.

Localized Wallet Heists

In the last few years, we have seen a steady decline in Trojans targeting bank transactions and wire transfers. Banks are better at detecting and preventing malware and fraud. More security and regulations have been placed around the wiring of funds. Such attacks are also easier to trace, resulting in unwanted attention from law enforcement. But that doesn't mean that cyber criminals aren't interested in stealing someone else's money.

FortiGuard Labs recently documented a new phishing threat that uses a fake Amazon gift card generator to steal cryptocurrency. This malware monitors the victim's clipboard for wallet addresses and replaces them with the attacker's wallet (this is precisely a hijack and digital robbery!). It also uses fake documents to lure victims into potentially giving out confidential information, such as credentials for online shopping sites, credit card numbers, and home addresses. FortiGuard Labs also detected a new phishing campaign last summer that included malware designed to steal crypto wallet information and credentials from a victim's infected device. ElectroRAT is another new tool targeting digital wallets. It combines social engineering with custom cryptocurrency applications and a new Remote Access Trojan (RAT) targeting multiple operating systems, including Windows, Linux, and macOS.

We expect to see more malware designed to target stored crypto credentials and drain digital wallets. Part of the reason for this change is that criminals like to follow the path of least resistance. Capturing wire transfers has become increasingly difficult as organizations encrypt transactions and require multi-factor authentication. Digital wallets, on the other hand, tend to be less secure. And they are a much bigger market. It's sort of the difference between a digital bank robbery and a digital mugging. But while individual wallets may not have as big a payoff, we expect this to change as businesses begin to increasingly use digital wallets and currency for online transactions.

Solutions and Countermeasures:

The best strategy for securing digital wallets is the distribution of EDR technology. Endpoints continue to be the primary point of vulnerability for most organizations, and yet many are still only protected with antivirus or even off-the-shelf endpoint security solutions. EDR solutions should combine AI-enhanced behavioral anomaly detection and enhanced kernel protection to block malicious behaviors, such as contacting a C2 server or triggering a malware payload, effectively disrupting an attack before it can complete its objective.

Ransomware Wipers

Ransomware has always relied on its ability to encrypt, corrupt, and exfiltrate data to force organizations to pay up. Early indications are that they are beginning to up the ante by adding wiper malware to their arsenal to also delete data and cripple critical systems, such as OT or manufacturing equipment and servers, unless a ransom demand is met. Ironically, wiper malware is nearly a decade old. DarkSeoul and Shamoon, for example, were notorious wiper variants back in 2012, so it is interesting to see it make a comeback as part of today's more sophisticated threats.

Cyber criminals have already added extortion to their ransomware activities. They steal critical data and threaten to release it on public servers. Some threat actors have even begun to contact the victim's customers to put pressure on the victim to pay. Extortion works so well that some criminals, like SnapMC, have started to bypass the encryption element entirely. They simply steal data and then demand that companies pay up, or they will make their internal data public.

Ransomware attackers also add to the noise by combining ransomware with distributed denial-of-service (DDoS), another older attack, hoping to overwhelm IT teams so they can't take last-second decisions to mitigate the initial attack's damage. Adding the ticking time bomb of wiper malware, which will not only wreck data but destroy systems and hardware, creates additional urgency for companies to pay up quickly rather than dragging things out, which can sometimes be a tactic victims used to provide law enforcement with a bigger window for discovering the attackers.

Wiper malware has already made a visible comeback this summer, targeting the Olympic Games in Tokyo and disrupting train services in Iran. Given the level of convergence we have seen between attack methods and APTs, we believe that it is just a matter of time before destructive capabilities like wiper malware are added to most ransomware.

Solutions and Countermeasures:

Organizations must select and deploy tools designed to block and interrupt today's increasingly effective—and destructive—ransomware attacks. Many of these predictions start with a robust endpoint security strategy that includes EDR and advanced antivirus. Zero-trust controls, especially when combined with dynamic network segmentation and microsegmentation, can help limit the impact of these attacks. And organizations must have a recovery strategy that includes pristine, off-network backups and devices, red team/blue team activities, and simulated recovery drills to check processes, chains of command, and business continuity strategies.

The Cyber Crime Games

Esports are organized, multiplayer video game competitions, often between professional players and teams. It is a booming industry that is on track to surpass $1 billion in revenue this year. And Newzoo projects it to hit $1.8 billion in 2022. Many of the biggest casinos are also looking to esports to bolster flagging gaming revenue by adding expansive esports competition venues and esports betting to their gaming floors. According to one report, esports betting revenue for the first half of 2021 increased 38.6% to $58.4 million.

Esports seems like a particularly inviting target, whether by ransomware, financial and transactional theft, or social engineering attacks. Given its rate of growth and increasing interest, esports and online gaming are likely to be large attack targets for 2022.

Solutions and Countermeasures:

Service providers and esports providers need to provide secure gaming environments to prevent DDoS and other attacks. They should also include AI-based hunting tools to detect threats lurking in gaming environments. And connected gaming consoles need to have encrypted connections and endpoint protections such as EDR.

Earth to Cyber Crime

We expect to see new proof-of-concept (POC) exploits targeting satellite networks over the next year. Satellite-based internet access continues to grow. New low earth orbit (LEO) satellite systems are becoming faster and increasingly less expensive, making them a viable option not just for remote users but for more mainstream business customers. Viasat, HughesNet, and Starlink (beta) are available now, with new players like OneWeb and Project Kuiper (from Amazon) coming online soon. Starlink is reporting download speeds of over 560 Mbps, with Gigabit speeds promised.

We have already begun to see new attacks. ICARUS is a POC DDoS attack that leverages direct global accessibility to satellites to launch attacks from numerous locations. Every satellite, and its supporting base stations, is a potential entry point to the network. Starlink alone has over 4,000 satellites in place and is projected to eventually include over 30,000 interconnected satellites. And there will be millions of terminals from which to launch an attack. Living-off-the-edge tactics will soon expand to include LEO satellite networks.

The biggest targets will be organizations that rely on satellite-based connectivity to support low-latency activities, like online gaming or delivering critical services to remote locations, and remote field offices, pipelines, or targets in motion, like cruise and cargo ships and airlines. Other attacks, such as ransomware, are sure to follow. This will also expand the potential attack surface as organizations add satellite networks to connect previously off-grid systems, including remote OT devices and systems, to their portfolio of interconnected environments.

Solutions and Countermeasures:

Gateway firewalls, internal segmentation firewalls, and IPS all play a critical role in protecting expanding attack surfaces and new attack vectors. Organizations should also include advanced threat detection solutions to detect and respond to zero-day attacks targeting new environments, including sandboxing and behavioral analytics tools.

Weaponization of AI

We have predicted for some time that AI will begin to be leveraged by cyber criminals to enhance their malicious activities. AI is already used defensively to detect unusual IoT behavior that may indicate an attack, usually by botnets. And now, cyberattackers are leveraging AI to thwart the complicated algorithms used to detect that abnormal activity.

Deepfakes are a growing concern because they leverage AI to mimic human activities and can be used to enhance social engineering attacks. GPT-3 (Generative Pre-trained Transformer) is an AI-based system that uses deep language learning to produce convincing-sounding emails. With it, attackers can leverage hijacked emails by compromising mail servers or running man-in-the-middle attacks to generate emails and email replies that mimic the writing style, word choice, and tone of the person being impersonated, such as a manager or executive, even making references to previous correspondences.

Writing is just the start. There are already software tools online designed to clone someone's voice, with others in development. A vocal fingerprint of someone can be created using just a few seconds of audio and then generate arbitrary speech in real time. And a proof of concept deepfake video was released in August showing the image of the NVIDIA CEO being portrayed by an actor. Although still in initial development, this type of AI-enabled deepfake will become problematic as central processing unit (CPU)/graphics processing unit (GPU) performance becomes more powerful (and cheaper). The bar to creating these deepfakes will also be lowered through the commercialization of advanced applications. These could eventually lead to real-time impersonations over voice and video applications that could pass biometric analysis. The possibilities are endless, including the elimination of voiceprints as a form of authentication.

An open-source tool called Counterfit has just been released to pen test AI systems such as face recognition, image recognition, and fraud detection, etc., to ensure that the algorithms being used are trustworthy. Organizations can also leverage this tool for red/blue wargaming. We can also expect attackers to do the same, using this tool to identify vulnerabilities in AI systems.

Solutions and Countermeasures:

As these POC technologies become mainstream, we will need to change how we detect and mitigate attacks, including using AI to detect minor voice and video anomalies. Currently, our best defenses are segmentation, zero-trust access that restricts users and devices to a predefined set of assets, and integrated security fabrics designed to catch and limit the impact of an attack. We will also need to improve end-user training to detect suspicious or unexpected requests arriving via voice or video in addition to those coming via email. And for those spoofed communications that include embedded malware, organizations need to be monitoring traffic to detect a payload, which means having devices in place that are fast enough to inspect streaming video without impacting user experience.

Crack-a-Mole

Many cybersecurity professionals refer to current efforts to stop cyber crime as "whack-a-mole" because the criminal activity stopped in one location tends to pop back up in another, usually by the same actors. Some of that persistence is automated. Shutting down a C2 server, for example, simply initiates the automated creation of a new server somewhere else, with little to no interruption of criminal activity. As a result, attack cycles are getting shorter as attackers weaponize automation.

One of the big questions cybersecurity professionals have been asking is, How do we slow them down? Disruption is a particularly effective tactic. By "cracking" their armor, meaning we become more effective at coordinating counterstrikes so they stay down longer, we can weaken them by forcing them to regroup. We are already seeing strategic coordination and public/private partnerships increasing our ability to take criminal organizations out of the game for longer. And the better we get at that—using new tools like the new World Economic Forum (WEF) Partnership Against Cybercrime effort and our growing catalog of cyber threat playbooks—the easier it will be to spot them, disrupt their activities, and force them back to the drawing table to reengineer their operations. During this next year, we should expect to see these partnerships grow, especially between law enforcement and the private sector, supported by new cooperation between countries and regions, making the recovery time from disrupted malicious campaigns increase.

Solutions and Countermeasures:

Defending against attacks that target different elements of your attack surface requires a complete, integrated security fabric that provides consistent protections and communications across the distributed network combined with end-user education and awareness.

Defending Against Tomorrow's Threats Today

Many of tomorrow's threats are simply extensions of those we experience today. They tend to be faster, harder to detect, more malicious, or combine existing threats in new combinations. Even new zero-day threats have one thing in common: They all want to get your network or devices to do something you do not want them to do—change, add, or delete a file, add or remove a function, inject something into a normal process, take something or leave something behind. Once we understand that, we can implement security strategies designed to baseline normal operations and detect and intervene when something unexpected happens.

Of course, that's harder than it sounds. It requires solutions designed to interoperate rather than function in isolation. It requires smarter solutions that know how to ingest real-time threat intelligence, detect threat patterns and fingerprints, correlate massive amounts of data to detect anomalies, and automatically initiate a coordinated response. And things like inspecting streaming video or protecting massive environments, like trading floors, gaming environments, or even big data, means that security solutions need to operate at speeds and capacities that most existing tools cannot deliver. But those tools are out there and finding and implementing a fast, adaptive, automated, and fully integrated security strategy is the only way forward if you plan to compete successfully in today's expanding digital economy.

November 15, 2021 4:34 PM

1346406-0-0-EN