THE THREAT LANDSCAPE FOR HEALTHCARE ORGANIZATIONS_

INTRODUCTION_

The threat landscape for healthcare organizations has shifted tremendously since the outset of COVID-19. The means, opportunities, and motives for threat actors targeting the sector have all changed in nature and increased in intensity. While the theft of private data and ransomware continues to wreak havoc, there has also been a shift towards disruptive or destructive attacks, namely those that are leveraging vulnerabilities in the technology supply chain. One in four of those vulnerabilities known to be exploited by CISA, are device firmware vulnerabilities, and actors like the Conti/Trickbot group have doubled down on developing TTPs (Tools, Tactics and Procedures) that specifically target vulnerabilities in IT supply chain that allow them to evade traditional defenses found in medical environments.

In the US, the nation's leading cybersecurity agency, CISA, has repeatedly sounded the alarm on the rapidly rising risks facing healthcare organizations. Meanwhile the recent US Executive Order on Cyber Security and subsequent Memorandum emphasize the need to secure the software and firmware related to Critical Infrastructure including healthcare.

The risks to healthcare stemming from the vulnerable IT supply chain have cut across all types of organizations and data. The agency has noted an increase in attacks on healthcare bodies, pharmaceutical companies, medical research organizations, local governments, and other organizations involved in the COVID-19 response both nationally and internationally. CISA also recently named healthcare cybersecurity as one of the top priorities for the country, calling out an overall lack of security resources compared to the value of the systems they protect.

Adding to this dynamic, attackers have learned to leverage healthcare's mission of saving lives to exert maximum extortion leverage on hospitals in the form of ransomware. In 2019, 764 healthcare providers suffered attacks, which led to a wide range of clinical impacts including medical records becoming inaccessible/lost, emergency patients being redirected to other hospitals, surgical procedures and diagnostic tests being postponed, and admissions halted. Ransomware attacks in particular had severe impacts on EHR systems, with the average ransomware attack resulting in an average of 15 days of EHR downtime. These and other clinical disruptions from cyberattacks can have serious impacts on patient safety. Research by Vanderbilt shows that the 30-day mortality rate for heart

attacks increases following a data breach, resulting in 36 additional deaths per 10,000 cardiac events per year, potentially thousands of lives lost.

This paper will address a key aspect relevant to recent attacks, which is how quickly, powerfully, and relatively easily, they can be carried out by threat actors. We'll also analyze how the attack surface itself is shifting as devices are increasingly targeted at the hardware/firmware level, along with the reasons why attackers are shifting focus there, including examples.

We'll show a video demonstrating how an attacker attacking from the Internet can remotely exploit both a VPN device and from there, an internal host at the firmware level, all in minutes.

Finally, we'll provide recommendations on how to mitigate the risks associated with these types of attacks.

THREAT LANDSCAPE_

As defenders, cybersecurity strategies and tactics must always remain grounded in and informed by real-world threats. Nowhere is this more true than in healthcare, where organizations are facing marked increases both in the volume and complexity of cyberattacks.

The adversaries behind these attacks are constantly seeking new ways to tip the scales in their favor, and the threat landscape can change quickly when they find a soft spot in an organization's defenses. As such, it is critical for healthcare organizations and their cybersecurity teams to recognize how threats are changing in the wild and to adapt their security efforts appropriately. This is compounded by the observation that healthcare organizations tend to lag behind other industries in the adoption of advanced security controls. Applications and services employed within the organization tend to be based on legacy solutions and therefore impart additional fragility and risk, increasing potential for extensive outages and impacts associated with clinical risk.

TRENDS_

While virtually all industries have suffered through a surge in cyberattacks in recent years, healthcare organizations have been one of the most heavily targeted. Recent studies have found that across all industries, healthcare organizations experienced the "highest increase in the volume of cyber-attacks (69%) as well as the complexity of cyber-attacks (67%)". In particular, ransomware attacks almost doubled in the past year alone.

There have also been important changes in the ways that adversaries target and attack healthcare providers. Attackers have increasingly targeted medical devices both as entry points into the network and as a way to disrupt critical services. Other actors have adopted double extortion techniques by threatening to publish patient data and other provider secrets unless the organization pays a ransom.

Even more recently, political hacktivist actors such as the pro-Russian KILLNET group, have overtly targeted hospitals in the West, even to the point of belligerence. Indeed they capitalize on the potential for loss of life as a primary tool in the context of fear and uncertainty. Outright disruption itself is the primary motive of these groups.

ACTORS_

Given these trends, it is important to know who is behind the rise in attacks. Unfortunately, healthcare environments have been targeted both by state-backed APT groups and hacktivists, as well as financially motivated actors such as ransomware groups.

The critical nature of medical environments has drawn the attention of a wide range of ransomware groups, with a recent FBI IC3 report finding that the Healthcare and Public Health (HPH) Sector had the most reported ransomware attacks of any industry. Hospitals have been specifically targeted by the Daixin Team who have targeted "electronic health records storage, diagnostics, imaging services, and intranet services" within clinical environments. Likewise, the notorious Ryuk ransomware that brought down 400 sites at a large healthcare provider

often attempts to encrypt a device's bootloader files, making the system unbootable when restarted.

Yet another ransomware group, NwGen, has specifically targeted children's hospitals. Recently, members of the defunct Conti ransomware group have re-emerged to target healthcare organizations as part of the Maui and Quantum ransomware families. Per recent research this same Conti/Trickbot group has also been found to target devices at the lowest levels, actively developing malware that targets the BIOS/UEFI and Intel ME/AMT technologies in order to access memory directly and evade every OS-level security control in the process. Additionally, the HHS recently alerted on yet another threat actor group known as Evil Corp that continues to pose a great threat to the healthcare sector, highlighting their extensive history of malware and tactics used against the sector.

Healthcare organizations have also been targeted by a variety of APT groups backed by foreign governments. North Korean APT groups copied the playbook of ransomware operators by using the Maui ransomware to go after hospitals. CISA has also warned that attackers have targeted healthcare and medical research firms by exploiting vulnerabilities known to be used by Russian, Chinese, and Iranian state-based threat actors.

MOTIVES_

Attackers can have very different motivations based on the type of adversary involved.

At a high level, financially-motivated attackers have recognized that the uptime and availability of clinical systems are paramount, and any disruption can put enormous pressure on a hospital to pay the ransom. Disruption to critical systems by ransomware attacks are increasingly contributing to the compromise of patient care. It is likely only a matter of time until such attacks directly cause a patient's death, particularly in scenarios where critical emergency care has been delayed and the patient is unable to receive timely treatment.

Healthcare environments are also home to large amounts of privileged data such as payment card information and protected health information (PHI). Attackers can steal this data both as ransom leverage or simply to sell to other criminal organizations.

APT groups, on the other hand, typically focus on more strategic goals involving intellectual property. In particular, state-sponsored adversaries have specifically targeted pharmaceutical, biotech, and other research firms in order to steal secrets related to new drugs and vaccines.

A final motive not often discussed in research communities is the adversary's potential motivation to not just steal or destroy data and device availability, but instead, manipulate it in ways that either benefit the adversary or cause long-term harm to the organization or research body without the victim easily learning of the attack (unlike ransomware or wiper attacks). The potential for these 'integrity' attacks may have both direct clinical risk impact scenarios, as well as the potential for poisoning / sabotaging medical research data in the context of today's nation-state geopolitical dynamics.

THREATS MOVING "BELOW THE OS" IN THE SUPPLY CHAIN_

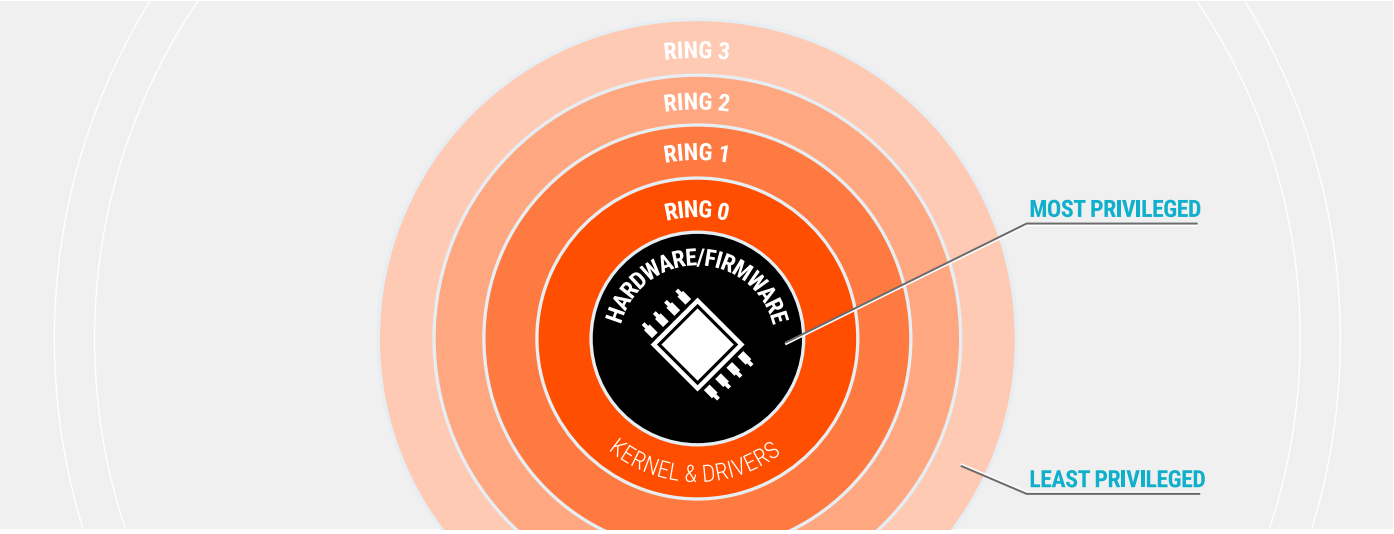
In addition to the steep rise in attacks, healthcare security teams are facing major changes in the ways that attackers operate. One of the most pronounced trends is a shift to target devices "below or beyond the operating system" or BtOS. These techniques target the most fundamental low-level code within a device, and when successful, can give an attacker virtually unmatched control over a device, as well as the OS, applications, passwords and secrets stored on it.

WHAT IS BtOS IN THE CONTEXT OF THE IT SUPPLY CHAIN?

The term BtOS tells us that we will be delving into the areas underlying the operating system of a device, but what does this really entail? In short, we need to consider all the code on a device that runs both below

the operating system as well as the code that runs before the operating system. This can involve a wide range of firmware running at the system level as well as within various physical components within the device. While the operating system governs the user experience and applications, the underlying firmware controls how the hardware actually works and how the device boots. The operating system, by design, is forced to trust the device firmware, and an actor operating below the operating system can completely modify or disable any and all relevant security controls at the operating system level, and enjoy indefinite persistence or the ability to destroy a device outright. This is why attackers are targeting IT supply chain vulnerabilities directly; because traditional defenses are forced to rely upon an assumed trust in their IT supply chain. This mis-guided assumption of trust forms the genesis of these latest attacker trends, opening up a wide berth of means and opportunity to the attacker.

The code running below the user operating system is the most privileged code on a device. In security, we often think of a device and operating system in terms of rings of privileges, with users operating in the least trusted Ring 3 and the operations system kernel in Ring 0. Firmware resides even below Ring 0, and is sometimes referred to as “negative” rings. It forms the root of trust for the entire compute stack “above” it, and has ultimate authority over everything.



Some of the most notable forms of firmware in a device include:

- **System software and firmware** - System firmware is the most important class of firmware in a device and includes the system BIOS and the more modern UEFI firmware. This UEFI firmware controls a wide range of functions, and one of the most important is controlling how a device boots and appropriately brings up the operating system.

- **SMM and runtime firmware** - System management mode (SMM) is another incredibly powerful form of firmware. Unlike UEFI, which focuses heavily on boot-time issues, SMM takes care of runtime coordination across hardware components. SMM includes the ability to interrupt the OS to perform various actions. By interrupting the OS, SMM can take action without the OS having any visibility into what is happening.

- **Remote “lights-out” management remote access management firmware** - The vast majority of modern laptops, desktops, and servers include some capabilities to support the out-of-band management of the device. In servers, this includes firmware in baseboard management controllers (BMCs), and in laptops and desktop workstations there is a physical

microcontroller which is known as the Management Engine (ME) in Intel-based systems. These are entire operating systems that sit below the primary user OS (Windows, Linux, Mac).

- **Component-level firmware** - Firmware is also a key component of virtually every physical component within devices. This can include network controllers, storage drives, GPUs, PCIe controllers, and so on. Once again, while the operating system can control and use these components at an abstract level, it is firmware that controls what the component actually does.

WHY ATTACKERS ARE MOVING BtOS_

Attackers have many reasons to target firmware. First, as the most privileged and the first code to run on a device, firmware gives attackers an opportunity to preempt all the many security controls that run at the level of the operating system. Attackers also have many ways to exploit firmware. Unlike traditional software and operating systems which receive regular updates, firmware is often rarely updated, and this means that vulnerabilities can persist for long periods of time. Even more challenging is the unfortunate reality that patches at the firmware level can often lead to additional vulnerabilities or the re-introduction or prior ones.

Additionally, the firmware on a device is typically quite diverse, including firmware from dozens of suppliers. The boot process relies on many low-level hardware settings, security features, and firmware, often requiring the tight coordination of OEMs, suppliers, OS vendors, chipset vendors, and more. This complexity creates many opportunities for backdoors, vulnerabilities and configuration mistakes that attackers can take advantage of. Even the core functionality of Secure Boot (the user OS's ability to trust the device it is booting from) have vulnerabilities that effectively "forever-days" (an architectural vulnerability that cannot be effectively patched), such as the Baton Drop vulnerability leveraged by the UEFI bootkit "Black Lotus", for sale on criminal forums for only $5000 USD.

There are also many ways that attackers use firmware in the context of an attack. Some of the key examples include:

- **Initial Access** - Firmware can provide a path into the device. Attacks against the firmware in networking devices have become one of the most popular initial access vectors used by attackers in the wild. Remote out-of-band management functionality provided by components such as BMCs and Intel ME can also be abused in order to gain access and control over more traditional devices like servers and laptops. Backdoors, hard-coded passwords, and RCE vulnerabilities are also introduced directly into the ICT supply chain itself, such that devices are already affected before the medical environment even takes delivery of them.

- **Privilege Escalation** - Attackers can also look to firmware as a way of elevating their privileges on a device. Privilege escalation within an operating system (e.g. from User to Admin) has been performed by attackers and malware for decades. Firmware allows attackers to take this a step further and escalate privileges beyond the OS entirely.

- **Security Evasion** - The ability to subvert higher layers most importantly allows attackers to avoid controls and security measures that run at higher layers or only see down to Ring 0. This includes traditional security running at the operating system and virtual machine layers. For example, compromised firmware can easily allow an attacker to control how a system boots, patch the operating system itself, disable OS level security controls such as (Credential Guard, BitLocker, AV/EDR), read privileged data off of hardware, or control assets that the operating system doesn't have visibility into. In the case of servers, compromised firmware can also allow attackers to further compromise the memory, hypervisor, and virtual machine layer in an organization's cloud assets.

- **Persistence** - The ability to hide from and evade the operating system provides attackers with extreme levels of persistence on a compromised device. In addition to evading controls, any malicious code in the firmware is naturally tied to the hardware of the device as opposed to the software. This means the attacker's

code would naturally persist even after a full re-imaging of the system from a backup. Such capability is particularly strategic for attackers, allowing them to remain in the environment undetected with the highest amount of capability possible. Advantages include re-launching 'round two' of a ransomware campaign, employing a 'scorched earth' strategy after the initial exfiltration or extortion efforts, remaining on a Domain Controller, or outright switching to a destructive objective following prior espionage activity.

- **Stealth** - Compromised firmware also enables attackers to perform a variety of critical attack functions without being detected. For example, by controlling the firmware of a hard disk or SSD, an attacker could hide malware in a section of the disk that is not reported to the OS, allowing it to avoid scanning by antivirus tools. Likewise, attacks against the management components of a device such as Intel's Management Engine have allowed attackers to send command-and-control traffic through independent channels that aren't monitored by host-based firewalls running at the OS level, or even logged at the kernel level whatsoever. Perhaps even more profound, is the extreme stealth enjoyed by attackers targeting the firmware of externally facing devices such as firewalls, VPNs, and load balancers; devices that defenders have little visibility or tooling in place to detect such low-level persistence.

- **Disabling Other Security Controls** - One of the key advantages attackers leverage once they have indefinite persistence on devices, is the ability to 'reach up' to the operating system during the boot process, to disable primary security controls such as BitLocker, Microsoft Defender, credential guard, and 3rd party AV/EDR/UEBA/IDM solutions. The Conti/TrickBot group targets Intel ME and the UEFI via TrickBoot primarily for these tactical capabilities, for example. And the Black Lotus UEFI bootkit sold for only $5000 on criminal forums allows any actor to disable these and other controls in order to be able to move up to the OS and laterally across the network without being detected.

- **Damage** - Lastly, access to the firmware layer enables attackers to cause irrevocable damage to a device. By damaging the firmware itself, attackers can "brick" the device permanently. This can potentially cause organizations to rethink their recovery models since it shifts the response from data recovery and reinstallation to complete device replacement. Imagine a Not-Petya campaign that, instead of destroying only the (replaceable) hard drive, bricked the entire motherboard. Additionally, the act of effectively disabling a critical asset can have enormous impacts on an organization by disrupting patient workflows reliant upon key 'choke point' devices.

BtOS EXAMPLES AND TRENDS_

Attacker tactics have also evolved significantly in recent years, and one of the biggest trends is a shift toward device-level attacks. Many of these attacks have focused on the network infrastructure that healthcare organizations and facilities rely on. Beginning in 2020, CISA issued the first of many alerts warning of widespread exploitation of firmware vulnerabilities in popular Pulse Secure and Citrix VPN appliances. The trend continued to gain steam across a broad range of threat actors spreading to additional vendors, including Cisco, F5, Fortinet, Palo Alto Networks, SonicWall, and many others. A lot of these types of attacks leverage the fact that even when there are patches released for these device firmware vulnerabilities, in practice, they don't get updated very frequently, as doing so causes tremendous disruption in a post-covid world. As an example, recent research shows just how aged these devices often are in the field.

These devices have proven to be attractive targets for several reasons. First, they are public-facing devices by nature and rely on firmware and custom device-specific operating systems that are not updated as frequently as traditional user operating systems on IT systems. This makes them the ideal initial access vector into an organization. These devices provide the primary security control function of secure connectivity to a remote workforce and medical application interconnections for an organization's network. They also provide attackers

with an ideal way to spread internally once they have gained access, giving them access to Active Directory, and other devices on the same segment, which is often a 'flat' network in many hospital networks.

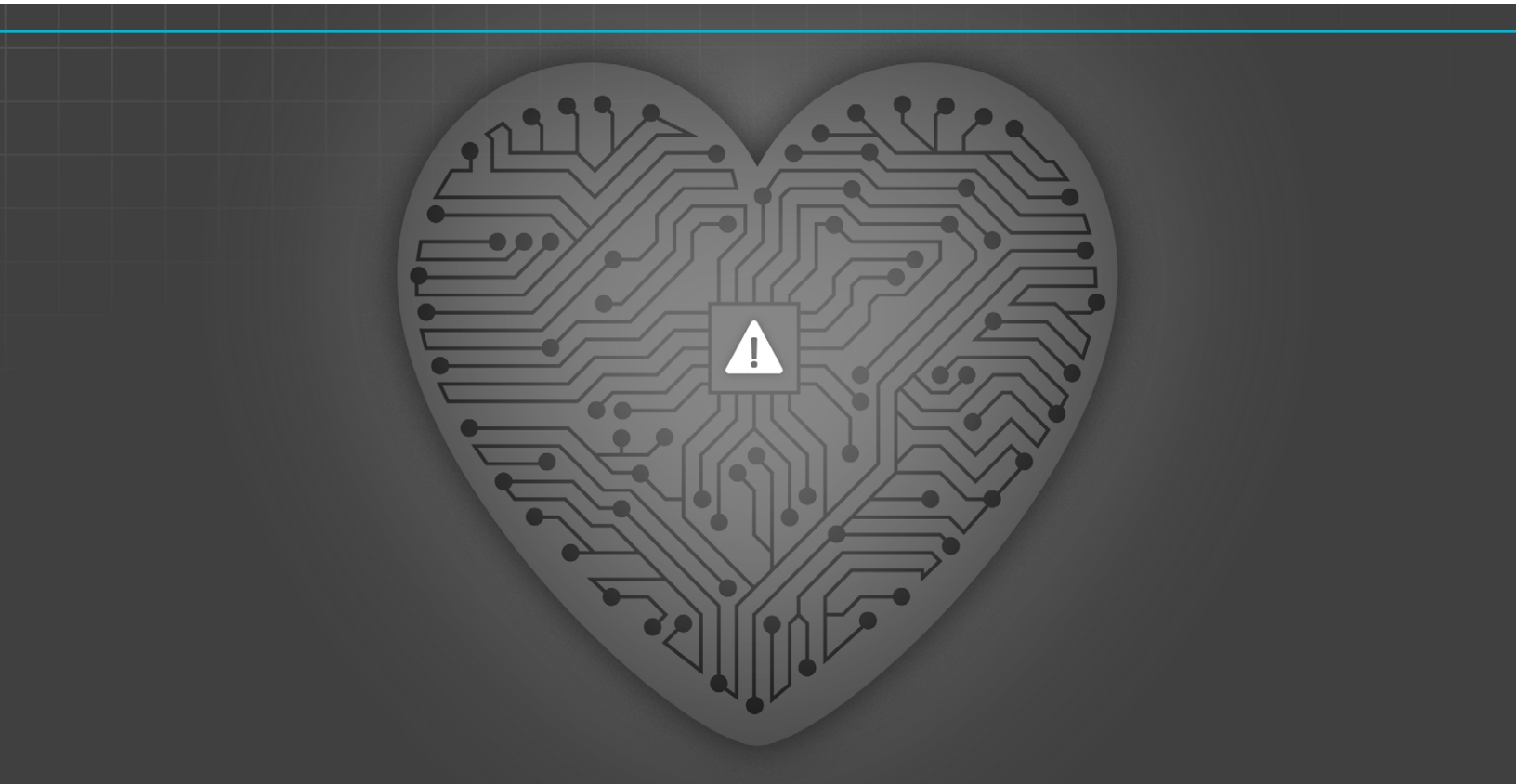
Even network devices within the environment have become a *primary* tactic of those actors most often targeting hospital environments. The TrickBot/Conti group as a whole, often pivots from internal host infections to surrounding devices on the network, leveraging off-the-shelf hacking tools such as RouterScan to find vulnerable devices such as SOHO routers, IoT, IoMT, storage devices, WAPs, and IP cameras, and subsequently, exploit them. These devices serve several purposes once compromised. For one, they buy the attacker time, as the attacker no longer needs to evade AV/EDR and other host-level security controls. Second, the devices themselves provide capabilities perfectly suited for the majority of attacker objectives; evasion, disruption, tunneling, c2, crypto-mining, password stealing, and exfiltration. Attacks against these devices have more than doubled in the past year in healthcare environments specifically.

However, this focus on supply chain firmware-level attacks goes well beyond network devices. Popular malware such as TrickBot incorporated UEFI firmware-targeting capabilities, which was followed by a string of UEFI implants, including FinSpy, MoonBounce, and CosmicStrand, which enable attackers to take complete and persistent control over standard devices such as laptops, servers, and VDI infrastructure.

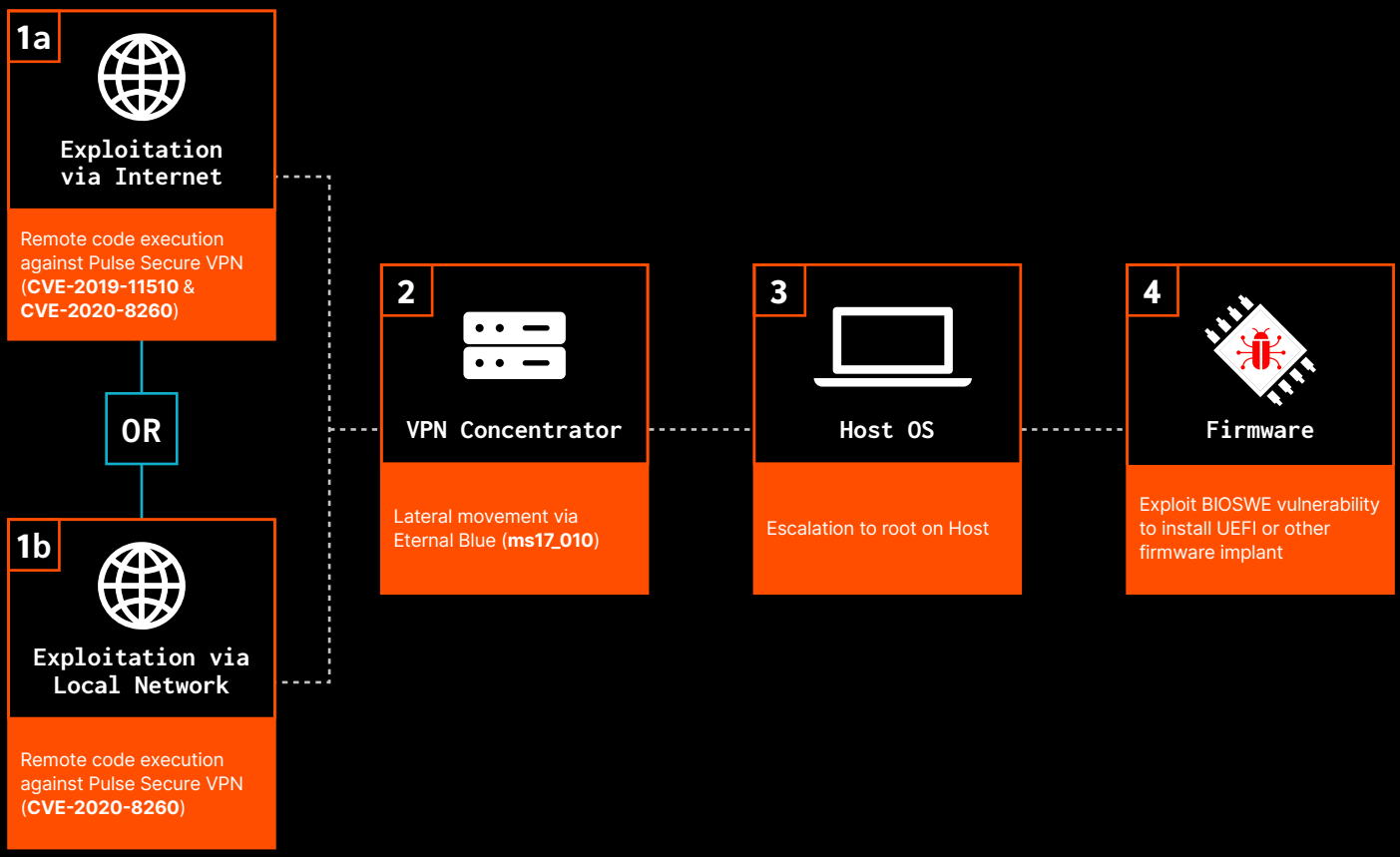
ATTACK SCENARIOS_

Let's look at how a BtOS attack could manifest in a healthcare environment. It is important to note that all of the exploitation, lateral movement, and escalation techniques covered in the following scenarios have been observed in the wild in large-scale attacks. We have chosen these examples to highlight that these are not theoretical risks, but readily available exploits and techniques that attackers can easily use to great effect without the need for particularly specialized skills.

We will look at an attack through two lenses. First, we will cover the basic flow of an attack including how an adversary could gain initial access, spread laterally through the environment, and gain ongoing persistence and privileges on target hosts. And since this is a relatively generic flow that could apply to a wide range of assets, we will next look at how such an attack flow could translate into more tangible impacts based on the types of devices and assets that are compromised.



CORE ATTACK FLOW_



In this example, we will follow an external attacker as they move from gaining initial access to the clinical environment via a vulnerable VPN appliance, through to the installation of a UEFI implant capable of taking complete control of a system such as a laptop, desktop, server, or IoT device. In steps 4 and 5 we are covering specific vulnerabilities and techniques that are relevant to a Windows device, however the same attack flow could apply to other devices such as IoT and Linux-based devices.

1. **Initial access via Pulse Secure VPN** - The attack begins by targeting well-known vulnerabilities in Pulse Secure VPNs that could give an attacker the ability to execute malicious code on the VPN. Such an attack could be performed by a remote, external attacker or by an attacker who is on the local network. A remote attacker would first target CVE-2019-11510 (10.0, Critical) which would allow the remote attacker to capture credentials of the VPN, then exploit CVE-2020-8260 (7.2, High) to use those credentials to execute the attacker's payload and set up a remote shell from the VPN appliance to the attacker. A local attacker could take advantage of the fact that the VPN management interface is likely exposed and directly exploit CVE-2020-8260 without the need for any previous steps.

It is important to note that there are many such

vulnerabilities for attackers to choose from, and this two-step chaining of CVEs against a single device has been used repeatedly in the wild including in recent real-world attacks. Furthermore, while we have used Pulse Secure as an example, the same concept has been frequently used to attack devices made by Cisco, F5, Citrix, Fortinet, and other enterprise infrastructure vendors. In fact, recent research shows both how seldom such devices are patched, as well as how attackers can reside on them indefinitely, persisting even after the devices have been patched to their latest firmware.

These types of RCE (Remote Code Execution) attacks are far more nefarious than run-of-the-mill VPN credential reuse attacks that simply allow access to a device's management UI. Exploits at this level provide attackers with a true Linux shell providing far more power, evasion, and stealth than having access to only the management web interface's features.

2. **Lateral movement via EternalBlue** - Originally disclosed as a part of the Shadow Brokers leak in 2017, the EternalBlue SMB vulnerabilities (MS 17-010) remain in use by a variety of threat actors. These wormable vulnerabilities can allow attackers to easily move to any vulnerable device, and were used extensively during the Wanna-Cry attacks that paralyzed hospitals in 2017. While these vulnerabilities are still very common today and often exploited, there are nearly always Windows vulnerabilities that can be exploited "on any given day" to achieve the same effect.

3. **Windows 10 or Host OS Escalation** - There are many options available to attackers to escalate privileges once on the target host. For Microsoft targets, there is always a way to bypass UAC on any given day (here are 76 methods anyone can use in one github for example). On the Linux side, vulnerabilities also exist that allow for privilege escalation. An example would be CVE-2021-4034, that for the last 12 years, would allow an attacker to escalate to root on every major Linux distribution.

4. **BIOSWE to Overwrite Firmware and Install Implant** - BIOSWE (BIOS Write Enable) is a simple misconfiguration that allows an attacker to write to the device's firmware and implant the UEFI. This has been used in the wild by Trickbot, Lojax, MosaicRegressor, and more. With the implant or bootkit installed, the attacker can subvert the operating system by making changes to the OS before it loads or by simply directing it to an attacker-controlled OS image. At this point, the attacker would have complete, persistent, and nearly invisible control over the device, and their persistence would endure even a full system restore and hard-drive replacement. From here, attackers can destroy the device, modify memory or patient data on the device, or steal and offload patient data on the device.

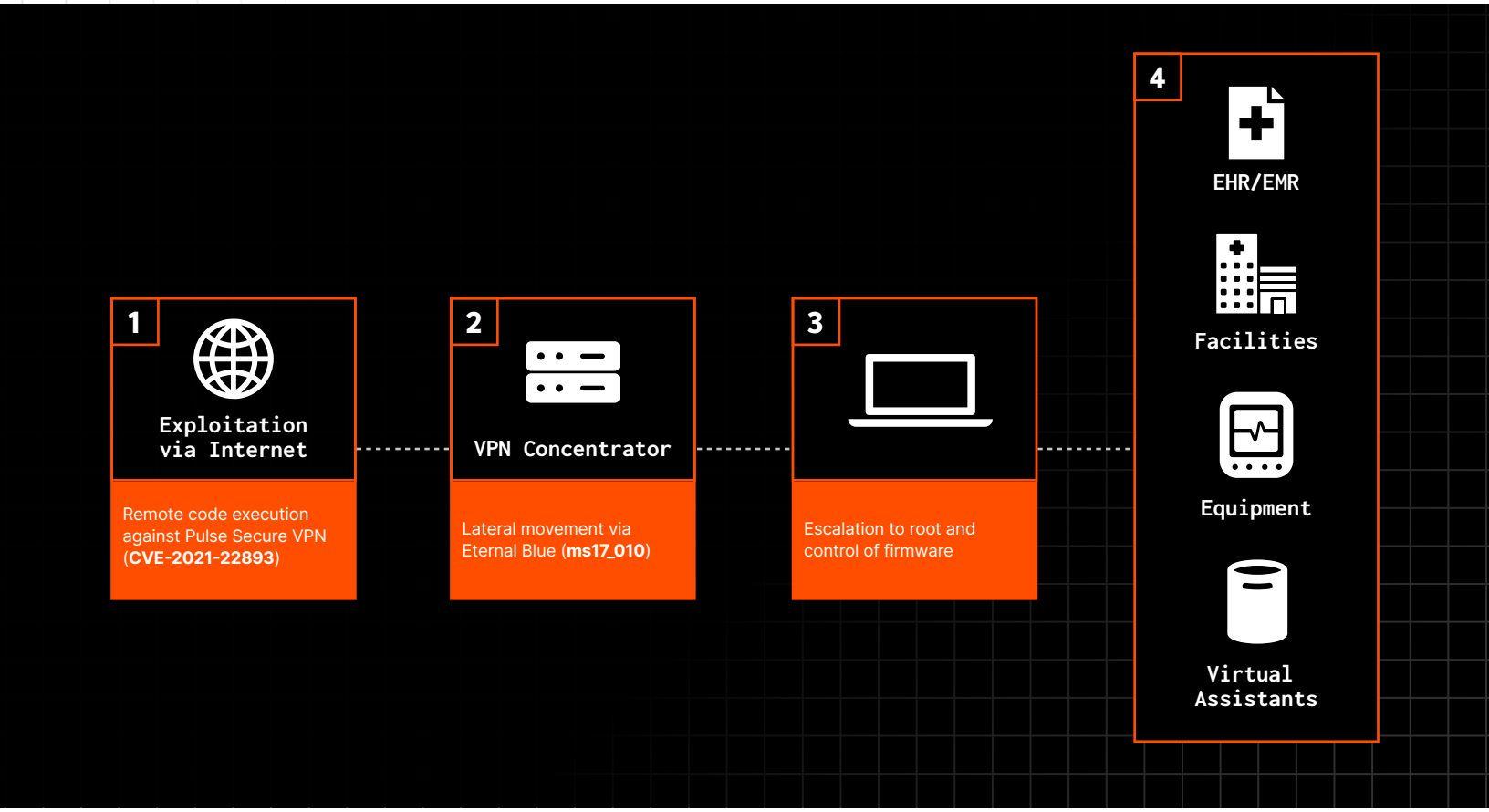
FIRMWARE VIDEO DEMO_

This demo shows how quickly, easily, and effectively an attacker from anywhere on the Internet can compromise an internal Windows host and infect it with a UEFI firmware implant via an all-too-common RCE (Remote Code Execution) vulnerability on a Pulse Secure VPN device.

IMPACT TO HEALTHCARE ASSETS_

It is important to note that attackers could use this basic approach to reach virtually any type of asset in a hospital or medical environment. Even though we showed a Windows laptop in the example, that host could just as easily be a server hosting the facility's EHR system, an infusion pump, a security camera, or virtually any type of device. And likewise, there are many devices in which a loss to the confidentiality, integrity, or availability could have a serious impact on an organization's ability to deliver care.

Let's look at a few examples. We will break things down in terms of systems that are scalable vs non-scalable (i.e. do they affect large numbers of patients or individual patients), and whether those systems support clinical or operational aspects of the facility?



Let's look at a few examples. We will break things down in terms of systems that are scalable vs non-scalable (i.e. do they affect large numbers of patients or individual patients), and whether those systems support clinical or operational aspects of the facility?

- **Scalable, Clinical Assets** - These assets directly support clinical operations and for large numbers of patients, and any compromise to the CIA triad (Confidentiality, Integrity, Availability) of these systems would have devastating impacts on an organization's ability to deliver care. This could include EHR/EMR systems (patient data), LIMS systems (diagnostic/ lab systems), PACS systems (imaging), and others. Likewise, disruption to systems responsible for preserving vaccines, samples, and other sensitive materials could cause widespread disruptions to care.

- **Non-Scalable Clinical Assets** - Devices such as infusion pumps, bedside monitors, and continuous blood glucose monitors are critical to providing care to individual patients. Compromise to these systems can have devastating and even potentially life-threatening impacts to the patient. These devices do not benefit from traditional security prevention or detection controls. However, a compromise to one of these devices would typically not have a facility-wide impact.

- **Scalable Operational Assets** - Hospitals naturally rely on a wide range of non-clinical systems that are ultimately essentials to serving patients and their families. Impacts to these systems can likewise disrupt the facility's ability to function. For example, disruption to systems controlling elevators or parking access can prevent patients and visitors from getting where they need to go. Disruptions to physical security and badging systems can likewise prevent clinicians from getting into controlled areas to do their jobs. Loss of HVAC and other environmental systems can likewise cripple a facility particularly in cases of extreme hot or

cold weather, low-pressure isolations rooms in the ICU, or blood storage facilities.

- **Non-Scalable Operational Assets** - Hospitals are increasingly adopting many IoT devices and systems to support patients and their families. This can include in-room smart devices (e.g. smart light bulbs), tablets, connected TVs and monitors (that may also act as information kiosks) to provide individual visitors with information and enhance the patient experience. Whilst an impact to these systems would rarely impact the quality of direct clinical care, they can have a significant impact on patient confidence and perception of the hospital experience.

- **IT/Admin Assets** - Lastly, we should come first circle and go back and consider the generic laptop. One of the fundamental aspects of BtOS attacks is the ability to gain control over a machine and silently persist out of sight. Compromising an end-user laptop could easily provide an attacker with a path to compromising clinical or operational systems. An attacker could use access on one laptop to move laterally to other machines and ultimately capture administrator credentials, which could then be used to access virtually any asset.

HEALTHCARE IMPACTS_

General Overview

Each of the covered scenarios would involve various degrees of loss of confidentiality, integrity, and availability of an organization's data and systems. Quantification of these risks and associated impacts can be quite challenging to establish, and this often requires the healthcare provider to engage with an external manual consultative exercise (which can span several weeks depending on the footprint of assets involved) to establish the true risk of the organization. Without this type of deeper dive, an overview of healthcare impacts can be made, but often this will be quite a general approach covering the following themes.

Loss of Patient Data

Healthcare data remains by far the most valuable type of data that attackers can steal, with studies indicating that a single patient record can sell for up to $250, almost 50 times more than a payment card. And unlike a stolen credit card number that can be easily replaced, the loss of healthcare data is permanent. This loss can have a variety of financial and reputational impacts on an affected healthcare system. In addition to potential fines due to HIPAA violations, patients are increasingly likely to sue their hospitals in response to a data breach. In fact, the most recent data shows that healthcare organizations had the highest percentage of lawsuits due to data breaches of any industry.

Ransomware and Financial Extortion

Ransomware operators covet healthcare organizations for good reasons - their environments contain the combination of highly valuable data and some of the most critical systems in terms of availability. This gives attackers incredible leverage when demanding a ransom. A ransomware attack can have severe consequences both due to the downtime of any critical systems as well as by shaking the confidence of patients and families during any outages. CISA and other organizations routinely put out alerts focused on specific actor groups targeting healthcare. HHS, for example, recently alerted to a threat actor group (Evil Corp) that has been active for over 15 years, that continues to pose a great threat to the healthcare sector, highlighting their extensive history of malware and tactics used against the sector.

Destructive or Political Attacks

For all the evils of financially-motivated attackers, they can pale in comparison to attackers whose core motivation is to cause damage. Firmware attacks in particular have the potential to permanently disable compromised devices by making them unbootable. This could affect everything from an organization's routers and switches, the mobile computing devices that clinicians depend on, clinical and diagnostic equipment, as well as the servers and cloud assets that handle the organization's data. In these cases, physical devices will need to be replaced or undergo highly technical repairs. Cyber attacks can bring down entire health systems for indefinite periods of time, causing exponential clinical impact. One targeting the NHS has brought systems down for several weeks and threatens to have impacts extending several months before key systems can be brought back online safely.

CONTEXTUAL RISK QUANTIFICATION_

In this paper we have used the AbedGraham Group's Clinical Risk Analytics Service to help us more clearly quantify the risks associated with the compromise of the various healthcare assets outlined above.

The managed service utilizes a proprietary in-house clinical risk analytics tool which contextually analyzes, ranks and visualizes each endpoint based on the risks they present to a health system clinically, organizationally, financially and in terms of regulatory compliance using a standardized 1-12 point scale. This is achieved using algorithmic models that take into account a broad range of behavioral attributes of network endpoints based on their functional behavior across clinical workflows and associated interdependencies. In doing so, a granular asset profile can be determined and different types of attack can be modeled based on the detected vulnerabilities allowing the platform to determine the severity of any potential patient safety risks and their scalability.

The key thematic impact metrics identified are defined as follows:

- **C**linical Risk pertains to the potential severity of patient harm that could occur
- **O**rganizational Risk pertains to the level of clinical workflow disruption or service shut down that could occur
- **F**inancial Risk pertains to the potential level of recovery and regulatory costs, as well as revenue losses that could occur
- **R**egulatory Risk pertains to the severity of intervention from regulators following disruption and degree of reputational damage

ANALYSIS OF ASSETS, VULNERABILITIES AND RISK/IMPACT_

Pulse Secure VPN (Point of Entry to Network)

Device	C	O	F	R	Total
VPN Appliance	**4**	**6**	**6**	**4**	**5**

In this scenario when using the AbedGraham Group's Clinical Risk Analytics Service, the main concern associated with a VPN Appliance containing a vulnerability with a CVSS v3.1 score of 10, is that it will act as a portal of entry into the organization's network, rather than as an asset that has a direct/immediate impact on clinical or organizational workflows, hence a traditionally and comparatively lower clinical risk score (4/12) and organizational risk score (6/12). Organizational and financial costs will be associated more with IT infrastructure management (rather than clinical productivity and efficiencies with patient flow) and with a moderately low financial score (6/12). Depending on the motivations of the attacker, there may not necessarily be any immediate impact apparent to the organization hence a low regulatory risk score (4/12). Reconnaissance to strategize on how best to further compromise the organization to create the most significant impact for an optimal ROI at a later date is likely, but not always the case, as seen in the demonstration provided herein of an attacker jumping from a VPN device's firmware to an internal Windows host and implanting the BIOS/UEFI of the device in a matter of minutes, with the potential to permanently destroy the device at the "motherboard" level, or modify/destroy any of the data or applications (medical services) the device provides (such as an EHR system).

Takeaways:

The key issue associated with compromise of this asset is as a portal of entry into the network, with direct access from the firmware of the device to any devices logically connected to it on the internal network. Although any initial direct immediate impact to the organization may be low (unless the device itself is destroyed), the information gained from reconnaissance often results in a larger attack at a later date resulting in a larger degree of disruption spreading to key mission critical assets. However, as there is some latency until such large-scale disruption may manifest, assets that could result in more immediate disruption to clinical and organizational workflows will score higher (as they would be associated with a higher priority for remediation efforts).

EHR/EMR Server (Scalable Clinical Asset)

Device	C	O	F	R	Total
EHR Server	**10**	**11**	**12**	**12**	**11**

In this scenario when using the AbedGraham Group's Clinical Risk Analytics Service, an EHR/EMR Server supporting a clinical mission critical application containing a vulnerability with a CVSS v3.1 score of 10, would understandably give a relatively high clinical risk score (10/12) and organizational risk score (11/12), as compromised accessibility to the mission critical EHR would delay clinical information gathering, diagnosis, investigation, intervention/treatment and decision making (including formulation of management plans), impacting clinical workflows and affecting both staff and patients at scale across multiple clinical workflows, departments and services throughout the organization. This includes patients being redirected to other hospitals, admissions being halted, surgical procedures and outpatient appointments being canceled, tests being postponed, and discharges delayed. Due to the decreased productivity across the entire estate (applicable to most clinical and administrative settings), this will have knock on effects that would be associated with relatively higher risks in the other financial (12/12) and regulatory (12/12) categories. EHR server downtime can lead to revenue loss, recovery costs and regulatory penalties/litigation costs. This all combines to increase the amount of regulatory scrutiny and reputational damage that an organization would face.

Takeaways:

The key issue associated with compromise of this asset is the resulting decrease in clinical productivity, resulting from a delay rather than cessation of service provision. As the asset is a key mission critical application that supports both clinical and administrative workflow at scale, the severity and extent of disruption is a high concern. This would impact a wide range of activities across inpatient, outpatient and community related settings. Higher financial and regulatory risks would result from decreased productivity at scale across multiple environments, again including re-scheduling/re-organising clinical activities and other associated revenue and remediation costs.

Bedside Monitor (Non-Scalable Clinical Asset)

Device	C	O	F	R	Total
Bedside Monitor	**9**	**7**	**6**	**6**	**7**

In this scenario when using the AbedGraham Group's Clinical Risk Analytics Service, a Bedside Monitor supporting a clinical mission critical application containing a vulnerability with a CVSS v3.1 score of 10, would understandably give a moderately high clinical risk score (9/12) as compromised readings could delay clinical information gathering, diagnosis and decision making (including formulation of management plans). Although this would impact clinical workflows, this would be less than an EHR/EMR as the Bedside Monitor could be replaced relatively quickly with another equivalent device. As it would affect a smaller proportion of staff in the department (and only that one patient being monitored), that lower scale of disruption compared to an EHR results in a moderate organizational risk score (7/12). Although there would be some decreased productivity across a department, this would be limited to the care of one patient and so be associated with relatively moderate risks in the other financial (6/12) and regulatory (6/12) categories.

Takeaways:

The key issue associated with compromise of this asset is the resulting decrease in clinical productivity for a select group of individuals in the department managing this one patient – limiting the scale of disruption. Although clinical care could be somewhat compromised, the device can also be swapped out for another functioning replacement device, although the risk of that patient receiving suboptimal care still exists. Moderate financial and regulatory risks would result from the combination of clinical care disruption and remediation costs.

Physical Access Control System (Scalable Operational Asset)

Device	C	O	F	R	Total
Physical Access Control	**7**	**10**	**8**	**6**	**8**

In this scenario when using the AbedGraham Group's Clinical Risk Analytics Service, a Physical Access Control System containing a vulnerability with a CVSS v3.1 score of 10, may not immediately be associated with a moderate clinical risk score (7/12) as it is associated with hospital security infrastructure rather than clinical care. However, as it impacts the flow of patients and staff to and from clinical settings it would have some impact on clinical workflows (with patient and staff not being present resulting in some mild delay to clinical information gathering, diagnosis, investigation, intervention/ treatment and decision making (including formulation of management plans), and the effects would be observed hospital-wide across multiple clinical workflows, departments and services throughout the organization, resulting in an organizational risk score (10/12). As there could be delays to staff and patients getting to clinical environments and accessing care (resulting in delayed admissions, surgical procedures and outpatient appointments – with some even being canceled), the decreased productivity across the entire estate will have knock on effects that would be associated with relatively higher moderate risks in the other financial (8/12) and regulatory (6/12) categories. Decreased productivity and re-booking appointments can lead to significant revenue loss. Although not directly clinical in nature, the patient experience would be impacted contributing to reputational damage that an organization would face.

Takeaways:

The key issue associated with compromise of this asset is the resulting decrease in clinical and organization productivity, resulting from a delay and even cancellation of service provision (to be re-booked at a later date). As the asset is involved with patient and staff flow at scale across multiple workflows and departments, the severity and extent of disruption is a significant concern. This would impact a wide range of activities across inpatient, outpatient and community related settings. Noticeable financial and some regulatory risks would result from decreased productivity at scale across multiple environments, again including re-scheduling/re-organising clinical activities and other associated revenue and remediation costs.

Smart Light (Non-Scalable Operational Asset)

Device	C	O	F	R	Total
Smart Light	**4**	**6**	**6**	**4**	**5**

In this scenario when using the AbedGraham Group's Clinical Risk Analytics Service, the main concern associated with a Smart Light containing a vulnerability with a CVSS v3.1 score of 7.8, is that (similarly to the VPN appliance above) it will act as a portal of entry into the organization's network, rather than as an asset that has a significant direct/immediate impact on clinical or organizational workflows, hence a comparatively lower clinical risk score (4/12) and organizational risk score (6/12). However, there will be some impact to the staff and patient experience with some mild impact on workflows and it should be noted that although it scores similarly to the VPN appliance, this is with a lower scoring vulnerability. Again organizational and financial costs will be associated more with IT infrastructure management (rather than any major issues with clinical productivity and efficiencies in patient flow) and with a moderately low financial score (6/12). Depending on the motivations of the attacker, there may not necessarily be any immediate impact apparent to the organization hence a low regulatory risk score (4/12). Similarly to the VPN appliance, reconnaissance to strategize on how best to further compromise the organization to create the most significant impact for an optimal ROI at a later date is most likely.

Takeaways:

The key issue associated with compromise of this asset is as a portal of entry into the network. Although any initial direct immediate impact to the organization will be low (though arguably more noticeable than the VPN appliance – such as affecting lighting in clinical areas where procedures occur), the information gained from reconnaissance could result in a larger attack at a later date resulting in a larger degree of disruption spreading to key mission critical assets. However, as there is some latency until such large-scale disruption may manifest, assets that could result in more immediate disruption to clinical and organizational workflows will score higher (as they would be associated with a higher priority for remediation efforts).

RECOMMENDATIONS AND MITIGATIONS_

In its simplest form, cybersecurity risk is often evaluated as the likelihood of a security incident multiplied by the impact. As we have seen, clinical environments face both high likelihood and incredibly high impact from a cybersecurity perspective. As a result, it is critical for organizations to take proactive measures to reduce their risk and protect their assets. This can include efforts in the following key areas.

Utilizing a Clinical Risk Analytics Managed Service

The patient safety and clinical workflow disruption risk metrics produced can be scaled to provide a total health system risk profile and the insights can ultimately guide any remediation strategies and application of security controls. Unlike other managed security service providers and consultancies, the metrics produced by a clinical risk focused managed service factor in the clinical and organizational context, utility and workflow of the asset used in a healthcare environment, providing a more granular and industry specific measure of risk.

Patching

All code will have vulnerabilities, and it is critical for organizations to keep their systems up-to-date with the latest available code. Additionally, security teams will need to continually monitor for any newly discovered vulnerabilities that could pose an immediate risk to the organization. Attackers will often weaponize and begin exploiting critical, high-value vulnerabilities within days or hours of disclosure, making it critical for security teams to periodically apply patches outside of the normal update schedule.

Network Segmentation

If we revisit the attack scenarios covered previously, it is important to note how much of the attack chain occurs inside the network. This is a common trait of modern attacks and is one of the key reasons organizations increasingly implement Zero Trust principles in their networks. While Zero Trust can be applied in many ways, one of the first steps is often to implement granular network segmentation policies. Network segmentation breaks the network into smaller functional units, which can help isolate assets from spreading internal threats and provide more opportunities to detect threats. For example, the ICU could be deployed into its own protected segment, or even classes of devices such as IoT devices can be placed in a tightly controlled segment to ensure that they can only be reached via approved channels.

Firmware Vulnerability and Threat Detection for Threats BtOS

While most organizations are used to looking for traditional threats and vulnerabilities at the OS and application levels, many are blind to these same risks running BtOS. Traditional vulnerability scanning tools typically lack the low-level privileges to see down to the firmware level and simply don't recognize the dozens of low-level configuration issues that can affect the boot process. Likewise, traditional AV and EDR tools have only sparse visibility into firmware and hardware threats, and more importantly, suffer from a fundamental problem in that they rely on the operating system for visibility into the BtOS layers. As we have seen, a threat in these lower layers can easily lie to the operating system, rendering security tools virtually useless.

A dedicated firmware security platform brings specialized capabilities that enable security teams to automate security at the BtOS layers. Teams can now quickly see the full extent of their hardware and firmware inventory, proactively identify vulnerabilities and threats, and take corrective actions. This includes the following key capabilities:

- **Identify** - Automated discovery of network devices and ongoing visibility into the firmware, hardware configuration, and the dozens of components within your network devices and infrastructure. Quickly zero in on important devices, components, attributes, or changes that can impact your security.

- **Verify** - Proactively identify risks from outdated or vulnerable firmware or device misconfigurations. Verify the integrity of all firmware and detect known and unknown firmware threats, including rootkits, implants, and backdoors.

- **Fortify** - Remotely apply patches or updates to proactively mitigate device risks. Receive automated alerts to any firmware integrity changes and drive automated responses via integration with your existing IT and security tools with pre-built integrations with leading SIEMs, vulnerability management, and device management tools.

Supply Chain Integrity Verification

Healthcare organizations must also proactively verify the integrity and security of all newly acquired devices at the firmware level. Modern equipment often contains components and firmware from dozens of suppliers, sub-suppliers, as well as the OEM. Organizations must be able to verify that newly acquired assets are authentic, free from backdoors, implants, and vulnerabilities, and haven't been otherwise tampered with in the supply chain.

Organizations will also need to establish baselines for all critical devices both in terms of the firmware and code in contains and its behaviors. Unlike most applications, firmware remains very predictable. By monitoring the baselines of this code and its actions, security teams can immediately see if there have been any changes and can identify suspicious firmware behavior. This can be invaluable for detecting threats that may be introduced via the vendor's valid update process.

Implement a Device-Level Zero Trust Policy

Clinical security teams should also consider extending Zero Trust policies and capabilities down to the root of trust on each device. In this case, organizations could allow or deny access based on the trustworthiness of the device at the firmware level. For example, connections could be allowed only if the connecting device is found to be free from known vulnerabilities and threats, properly configured, and running a known-approved version of firmware.

Hunt for Threats Below the Surface

Attackers use firmware implants as a way to survive indefinitely, even after system re-imaging or hard drive replacement. With Eclypsium, you can easily check a device for implants before returning it to service after a security incident and include firmware level analysis into host forensic playbooks for better root cause analysis, high-fidelity pivoting, and correlation.

About Eclypsium

Eclypsium's cloud-based platform provides digital supply chain security for critical hardware, firmware and software. Eclypsium defends enterprises and government agencies from the deep implants and exploits that have become the vector of choice for modern adversaries. For more information, visit eclypsium.com.

About The AbedGraham Group

The AbedGraham Group (TAG) is a globally leading healthcare cybersecurity services and technology group based in London, England. Since 2011, our team's unique blend of clinical experience and cybersecurity expertise has been leveraged by government agencies, health IT vendors, cybersecurity companies and healthcare systems to help them analyse threats to patient safety in their environments and shape their clinical risk management plans.

We use a highly powerful proprietary risk analytics tool that has been developed and used internally by physicians at TAG for many years for patient safety based risk quantification. This allows us to augment the existing range of powerful security tools on the market by adding clinical context and prioritising risks based on their patient safety impact.

info@eclypsium.com
eclypsium.com