DEVICE INTEGRITY AND THE ZERO TRUST FRAMEWORK

The recent shift to a remote work environment has created new challenges for many businesses and government institutions, with profound impacts on organizational security models. Suddenly, many users are no longer protected by the many layers of security found on-premise in the corporate network. Instead, security policies must evolve to support a new reality where users are remote by default and massive amounts of untrusted, inbound connections are the norm. Incorporating security concepts like Zero Trust can be a critical part of securing these remote work environments, which often include a mix of corporate laptops, BYOD devices, and home networking gear.

Zero Trust eschews the old "trusted inside, untrusted outside" model, and instead puts forth that every connection is untrusted until it has been verified. Each session must be assessed from a security perspective before access is granted, and trust is constantly reevaluated. As such, Zero Trust models are constantly adapting to new information, and ideally will incorporate multiple security contexts when rendering access decisions. Is the user's identity verified and are they allowed to access a particular resource? Has the integrity of the user's device been verified? What applications and services are being used to connect?

The security posture and integrity of a device is one of the most fundamentally important of these contexts, yet remains one of the most overlooked aspects of Zero Trust today. With remote workers increasingly outside the protections of network-based security, more and more of the cybersecurity battle is shifted to the users' endpoint devices. The hardware and firmware of a device is arguably the most strategically valuable resource in this battle. If compromised at this fundamental level, attackers can subvert everything on the device including the operating system, endpoint security controls, as well as the user's identity. And while such threats were once the domain of state-sponsored attackers, these techniques are increasingly being observed in more widespread malware and ransomware campaigns. Yet for many organizations, device integrity remains a blind spot where Zero Trust principles are not yet applied, and as a result, security is assumed instead of verified.

However, new security tools and innovations are making it possible for organizations to easily close this gap and incorporate device-level contexts into their overall approach to Zero Trust. From the supply chain to remote work to data centers and cloud computing, we examine how, with minimal effort, Zero Trust principles can be applied to device integrity. With a few simple steps, organizations can ensure that Zero Trust decisions include the most fundamental aspects of device security and begin from the moment a device is powered on.

The Rise of Attacks Against Device Integrity

One of the overarching goals of Zero Trust is for organizations to identify and remove assumptions in their security practice. For example, in the past many organizations assumed that a user inside the network could be trusted and was free from threats. Over the years, attackers learned to take advantage of this assumption by focusing on infecting valid end users with malware to then spread internally and steal data from a trusted position inside the network.

However, this is far from being the only dangerous assumption in cybersecurity. Attackers naturally seek out areas that are assumed to be safe and where security is weakest. In recent years, attackers have increasingly turned their attention to the fundamental layers of the device including its hardware and firmware components. These critical components govern how the machine boots, how the operating system is loaded, and provide some of the most powerful privileges available on the device.

Attackers are increasingly targeting these components both as a way to maintain persistence on a device and to subvert security running in the upper layers. What was once a theoretical class of threats has transitioned to a reality that organizations must address. Key examples include:

- LoJax malware, used by APT28 (a.k.a. Fancy Bear or Sednit) in widespread campaigns compromise the firmware of laptops in order to maintain persistence on infected hosts.
- Ransomware using malicious EFI bootloaders to prevent systems from booting.
- Widespread espionage by APT41 targeting Cisco, Citrix, and Zoho devices.
- ROCKBOOT MBR-based bootkit used to maintain persistence on Windows-based devices.
- Attacks by APT29 attempting to steal Covid-19 research.

These are just a few recent examples of real-world attacks against device integrity. Most notably, these attacks have affected a wide range of industries and a wide range of device types. End user laptops, servers, and networking infrastructure have all proven to be fair game. This means that in the same way security teams can no longer blindly trust internal users, they can no longer simply assume that their devices can be trusted.

Device Context and Integrity in Zero Trust

Device-level contexts and security posture have become an increasingly standard part of Zero Trust-based access decisions. Ultimately the goal is to verify that the connecting device itself can be trusted as part of the access decision. In Gartner's recent Market Guide for Zero Trust Network Access Gartner states,

> *"the new model — zero trust networking — presents an approach that abstracts and centralizes the access mechanisms, so that the security engineers and staff can be responsible for them. ZTNA starts with a default deny posture of zero trust. It grants access based on the identity of the humans and their devices, plus other attributes and context (such as time/date, geolocation and device posture), and adaptively offers the appropriate trust required at the time. The result is a more resilient environment, with improved flexibility and better monitoring. ZTNA will appeal to organizations looking for more-flexible and responsive ways to connect and collaborate with their digital business ecosystems, remote workers and partners"*
>
> (Gartner subscription required).

As BYOD has become a more prevalent and necessary part of the enterprise, it is often not enough to simply classify unmanaged devices as "bad" or untrusted. Instead, organizations increasingly need to assess

the context and posture of the device itself. This can include checking the patch level of the operating system or verifying the presence of an approved antivirus (AV) tool before granting access to a resource. However, these are examples of checking the software that is installed on the device, which is not the same thing as verifying the integrity **of** the device.

Vulnerabilities or threats within the hardware or firmware of a device can subvert the operating system and render all higher-layer device protections moot. State-backed attackers have taken advantage of this ability to hide from security for well over a decade. However these same techniques have been adopted by organized crime and more widespread, opportunistic attackers. As a result, verifying the fundamental integrity of a device must be the first step of establishing trust on the device.

Each organization is ultimately responsible for securing the integrity of its devices. Device manufacturers naturally aim to deliver secure products, and modern devices include a variety of components to defend hardware and firmware. However, the same can be said for operating systems and applications, yet organizations are well aware of the need to continually monitor them for vulnerabilities and threats. Hardware and firmware require similar attention, and this need has been increasingly codified in a variety of security standards and regulations such as PCI, FISMA, and CMMC. The same principles apply to Zero Trust, and it is incumbent on an organization to actively verify the integrity of its devices.

Applying Zero Trust to devices requires a few basic steps. All critical devices need to be addressed. The organization needs to assess the risks associated with each device, including device-level vulnerabilities, misconfigurations, and settings. Lastly, teams need the ability to know if a device has been compromised.

Some of these steps may be new to organizations. However, solutions are available that can automate these functions and allow organizations to efficiently defend their device layer in much the same way they defend their software layers today. For each of the following requirements, we will include key capabilities that are available today to help organizations protect themselves with minimal effort.

BROAD COVERAGE FOR DEVICES

Modern enterprises and their employees rely on a wide array of devices, and virtually all of them should be addressed as part of a Zero Trust approach to security. Any device that can be used to access critical content or systems in the environment should be considered in-scope. In most cases, this will force organizations to consider both managed and unmanaged devices, end-user BYOD technologies, and a wide variety of corporate infrastructure in addition to traditional corporate-issued laptops.

Zero Trust has taken on particular significance recently, specifically as a way to help organizations adapt to workers who are increasingly working from home by default. This working environment naturally lends itself to the use of more employee-owned and unmanaged devices. Likewise, remote users are likely to rely on consumer-grade network routers, which are more prone to attack and have been popular targets for attackers in the wild.

The nature of supporting both managed and unmanaged devices will require device contexts that can be delivered either with or without an agent on the endpoint. This can seem like a challenge for security teams given that device-level contexts are derived from endpoint agents running on the protected device. However, modern solutions can deliver device contexts for unmanaged, BYOD, or home office equipment via network-based scans and analysis.

Organizations will also need to consider their computing and network infrastructure. For instance, the increase in remote work has created an increased reliance on VPN infrastructure for remote connectivity. Vulnerabilities in VPNs also have become a very popular target for attackers. Network and security teams must be able to verify that this critical infrastructure is safe and has not been tampered with in order to ensure the validity of the connection. These same considerations will likely apply to other enterprise infrastructure such as corporate switches, firewalls, and the various servers and management systems that support an organization's applications.

How to Ensure Device Coverage

It can initially seem daunting to extend security to the hardware and firmware layers of so many devices. However, device integrity tools are available that can address these assets. Solutions can use a combination of agent-based and agentless techniques in order to ensure the best coverage for each particular device type. A modern device integrity platform should be able to address all of the following:

- Corporate laptops
- BYOD and personal use devices
- Unmanaged devices
- Home networking gear
- VPN infrastructure
- Corporate networking gear
- Servers and management systems

IDENTIFY DEVICE-LEVEL VULNERABILITIES

NIST's SP 800-207 *Zero Trust Architecture* establishes several tenets of Zero Trust. The fifth tenet states:

> *The enterprise ensures that all owned and associated devices are in the most secure state possible and monitors assets to ensure that they remain in the most secure state possible …. This requires a robust monitoring and reporting system in place to provide actionable data about the current state of enterprise resources.*

Historically speaking, ensuring devices are in the most secure state involved checking that the operating system was up to date and

applications were free from critical vulnerabilities. However, devices can have a wide range of vulnerabilities and problems that lie beneath this traditional waterline. For example, if devices are not properly implementing Secure Boot, attackers could easily subvert the entire operating system during startup. Likewise, vulnerabilities in the system UEFI or BIOS could grant complete control over a device. This same issue can apply to individual components within a device, such as drives, processors, network adapters, and more.

Recognition of these weaknesses is an integral part of understanding the overall risk profile of a connecting device. An organization may not always want to simply block a device due to a vulnerability, but it can provide an essential supporting context. For example, a BYOD laptop with a vulnerability may be limited to accessing basic services, but only systems that are verified to use Secure Boot and are free from critical UEFI vulnerabilities are allowed to access high-value assets.

How to Gain Visibility of Device-Level Vulnerabilities

Scanning for device-level weaknesses should be an automated and ongoing part of an organization's security practice in the same way that software and application vulnerability scanning is implemented today. A device integrity platform can automate this scanning to identify vulnerabilities with established CVEs as well as device-level misconfigurations that can put the device at risk. Teams can use this visibility to identify and prioritize devices that need important updates. Likewise, a platform can provide this context to other systems so that the posture of the device itself can be factored into Zero Trust access decisions in real time.

LOOK FOR SIGNS OF COMPROMISE

Ultimately, Zero Trust access decisions need to answer a very fundamental question—is the device compromised? However, this can lead to a catch-22. Attackers are increasingly using device-level implants, backdoors, and malicious bootloaders as a way to compromise devices without being detected by traditional tools such as antivirus and EDR software.

In order to regain trust at the device level, organizations must be able to verify that the firmware and boot process of the device is secure and hasn't been compromised. This verification can be established by scanning devices with appropriate tools to ensure that all UEFI and component firmware matches valid vendor-approved firmware and that the firmware is free from vulnerabilities or implants.

Additionally, threats in the wild such as the ShadowHammer campaign have demonstrated the potential for valid, vendor-supplied updates to be compromised with malicious code. In these cases, it is important to monitor the hardware and firmware behavior of the device in order to directly detect potential malicious behavior. Once again, the need for behavioral analysis is cited by both Gartner and NIST guidance:

> "In many cases, user and device behavior are continuously monitored for abnormal activity, as described in Gartner's Continuous Adaptive Risk and Trust Assessment (CARTA) framework."
>
> Market Guide for Zero Trust Network Access, Gartner (Gartner subscription required)

> "Behavioral attributes include automated user analytics, device analytics, and measured deviations from observed usage patterns."
>
> SP 800-207 Zero Trust Architecture, NIST

By verifying that only valid, trusted code is run on a device, and then monitoring the actual behavior of that code, organizations can actively verify the integrity of their devices. This provides the necessary solid foundation from which all other device-related contexts can be supported.

How to Detect Signs of Compromise

Once again, security tools are available to extend threat detection and remediation to the device layer. These tools can validate that all firmware and device-level code matches known, valid versions of code from the vendor. Additionally, the security solution should automatically detect the presence of any known backdoors, implants, rootkits, or other malware. Lastly, the tool should monitor the behavior of valid code to reveal any signs of unknown threats or malware. These threat-based contexts should naturally be shared with other security tools and used as part of a Zero Trust access decision.

Zero Trust Examples and Scenarios

Zero Trust principles should be applied consistently and universally across an organization. These concepts should be a standard part of the way organizations approach new and emerging security challenges. It is important to highlight potential problem areas where Zero Trust is often overlooked, allowing presumed trust to creep back into the security model. The following examples highlight some of these areas where organizations may want to consider applying Zero Trust.

SECURE ACCESS FOR REMOTE WORKERS

Virtually overnight, remote work has become the default for many organizations, and, in the process, further accentuated the importance of Zero Trust. Not only can internal users not be implicitly trusted, the majority of users are no longer even in the network. They are connecting from untrusted environments and often relying on potentially insecure personal devices and home routers.

Understanding device-level contexts will be essential to safely enable remote workers while also ensuring that corporate assets remain safe. With the increased dependence on BYOD, organizations need to be able to verify the integrity and audit the security posture of a variety of unmanaged devices. Organizations also may want to consider vulnerabilities and the configuration of users' routers as an extended part of the enterprise attack surface. Vulnerabilities in these types of networking devices have been the target of a variety of large-scale state-based attacks. Organizations may not have the ability to extend direct visibility and protection to these devices, but nevertheless want to provide employees with guidance on updating and maintaining their devices. Organizations will want to verify the safety and integrity of devices prior to granting the device remote access to any resources.

Likewise, organizations should not implicitly trust their networking and VPN infrastructure. A recent alert published by CISA notes that vulnerabilities in corporate VPNs including Citrix and Pulse Secure VPNs have become some of the most popular targets for attackers in 2020. As more corporate traffic runs through VPNs by default, organizations need to be able to ensure their infrastructure has not been compromised.

ZERO TRUST ON DELIVERY

In a Zero Trust model, users and devices are presumed to be compromised until they are actively verified. However, organizations often forget to apply this standard to newly-acquired devices, which are often assumed to be "clean." Weaknesses in the technology supply chain could allow a device to be compromised long before it is ever received by the buying organization. For example, a vulnerability in any of the numerous hardware components within a device could allow attackers to modify firmware and insert a malicious implant to subvert higher-layer controls. The Trusted Computing Group recently spoke on this topic and the need for new industry controls at the RSA 2020 conference. These types of problems have made Supply Chain Risk Management (SCRM) a major focus for many enterprises.

As a result, organizations should ensure that the verification of device integrity and firmware scanning practices described above are applied to all newly acquired devices. This will allow organizations to identify any vulnerabilities or tampering that might have occurred in the supply chain. Likewise, organizations should include firmware scanning as part of the standard pre-purchase evaluation of all prospective technology.

DO NOT BLINDLY TRUST VENDOR UPDATES

Devices naturally need to be regularly updated, and it is easy to assume that a vendor's updates can be trusted. However, this has proven to be a bad assumption. Previous Eclypsium® research found firmware using insecure updating practices, which could allow an attacker to intercept the update traffic and remotely deliver a malicious update to the firmware.

Worse still, a vendor's update infrastructure itself can be compromised. For example, in the aforementioned case of ShadowHammer, attackers were able to infiltrate ASUS update infrastructure and subsequently deliver malware to ASUS customers in the form of updates that were properly signed and delivered through the official ASUS Live Update utility.

It is also important to note that devices often need access to trusted vendor sites as part of the standard update process. This trusted connection to the outside world can present an attacker with an opportunity to quickly cut through an organization's carefully crafted micro-segmentation strategy.

In these cases, organizations will need to have the ability to detect vulnerable update processes such as accepting unsigned firmware updates or insecure connectivity issues. In cases where the valid update from the vendor has been compromised, organizations will need to be able to monitor for abnormal firmware behavior.

DEVICE BEST PRACTICES FOR ZERO TRUST

The points above provide a good starting point for thinking about how the Zero Trust security model applies to device integrity and the underlying firmware. However, it is certainly not an exhaustive list. There are a wide range of details to potentially consider, but we included a series of best practices that are applicable to most environments. These capabilities are available today in modern security tools, and can be integrated into an organization's existing security practices.

Ensure Device Posture: It isn't enough to simply check that a device is corporate-issued or has a recently patched OS. Weaknesses in firmware can subvert even the most trusted devices and up-to-date software. Recommended actions include:

- **Regularly check devices for device-level vulnerabilities.** Vulnerabilities can make it easy for malware or other attackers to add malicious code to a device.
- **Analyze devices for vulnerabilities and insecure updating practices.** Vulnerabilities such as not requiring signed firmware updates can allow an attacker to easily install his own malicious code. Insecure updating over a network could similarly give an attacker the opportunity to deliver malicious firmware remotely.
- **Apply the same vulnerability and integrity checks to device components.** In addition to the system-level BIOS and UEFI firmware, attackers can target firmware in key components such as drives and network adapters.

Ensure Device Integrity: Device-level implants, backdoors, or any unauthorized code, can allow attackers to take full control of a device and evade traditional security controls on the device. Organizations should monitor devices for signs of known and unknown threats. Key steps include:

- **Regularly verify device integrity and check for implants.** Malicious code in firmware allows attackers to persist on a device while gaining the highest possible privileges and control over the device. Teams should scan devices for known and unknown implants or backdoors.
- **Scan all newly acquired hardware.** All new devices should be analyzed at the firmware level to ensure that the firmware is valid, vendor-approved firmware that hasn't been tampered with in the supply chain.
- **Verify the integrity of devices after any security event.** Attackers can implant malicious code in firmware that can survive a complete re-imaging of the system. Staff must verify that the hardware root of trust is intact before returning a device to operation.
- **Monitor device behavior after a firmware update.** If a vendor is compromised, it is possible that even properly signed, valid firmware could contain malicious code. To detect such threats, security should monitor the behavior of firmware to identify any malicious activity.

Conclusions

These are just some of the ways that a Zero Trust approach can be applied to enterprise device security. At its heart, Zero Trust is about rooting out areas where trust is assumed and replacing that assumption with active verification. Unfortunately, for many enterprises, firmware and hardware components have been a persistent blind spot that has been trusted by default.

New tools are allowing organizations to gain the same visibility into firmware vulnerabilities, hardware misconfigurations, and other threats to device integrity, as they have for higher-layer software. This allows organizations to pursue a true Zero Trust approach that begins at the hardware root of trust of every device.

To learn more about device security or Zero Trust, please contact the Eclypsium team at info@eclypsium.com.