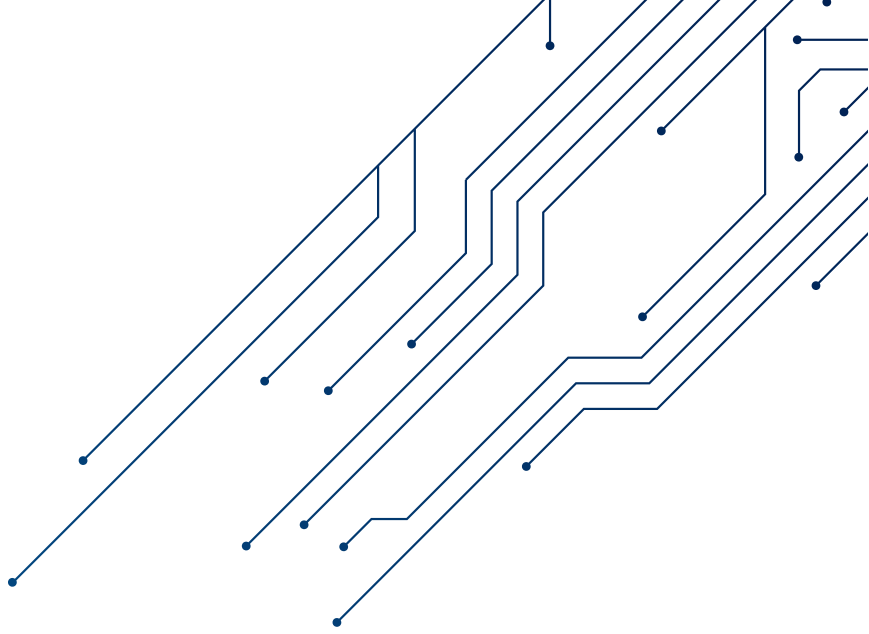




Zero Trust Cybersecurity for the Internet of Things





Abstract

In this paper we review the principles of Zero Trust security, and the aspects of IoT that make proactive application of Zero Trust to IoT different than its application to the workforce. The key capabilities of Zero Trust for IoT are defined for companies with an IoT strategy, and next steps highlight Microsoft solutions enabling your journey of Zero Trust for IoT.








Contents

Executive summary	4
Zero Trust security for the mobile workforce....	5
Zero Trust and IoT	6
Potential impacts of security compromise	6
Technical characteristics of IoT.....	7
Zero Trust capabilities for IoT.....	8
Strong identity	9
Least-privileged access	11
Device health	12
Continual updates.....	14
Security monitoring and response.....	15
Next steps.....	17
Evaluate & deploy a Zero Trust security model	17
Apply Zero Trust to your existing IoT infrastructure.....	17
Use Zero Trust as criteria to select IoT devices and services	18
Microsoft solutions	20
IoT services	20
IoT edge platforms	21
Integrated IoT hardware platforms	22

Executive summary

As organizations increasingly rely on automated systems for core business processes, the importance of improving the security posture of IoT is becoming business-critical. The Zero Trust model based on the principles of “never trust” and “always verify” can be applied to IoT to improve security posture.

There are five key capabilities required to implement a Zero Trust security model to IoT devices and services in your organization:

IoT Capability	Zero Trust principles fulfilled
Strong identity 	Verify explicitly: Mutual authentication of devices and services Least-privileged access: Hardware-backed credentials, and device registry Assume breach: Provision renewable credentials, use a device registry
Least-privileged access 	Verify explicitly: Authorize all IoT device calls Least-privileged access: Just-enough access for IoT devices Assume breach: E2E session encryption, encrypted data, network segmentation
Device health 	Verify explicitly: Evaluate signals like device location, health, and behaviors Least-privileged access: Device health to gate access or flag for remediation Assume breach: Attest regularly
Continual updates 	Verify explicitly: Device health maintained with updates Least-privileged access: Devices must be up to date for access
Security monitoring and response 	Verify explicitly: Proactive security assessments and pen testing Assume breach: Use analytics to get visibility, drive threat detection, and scale defenses through automated response

You can start the Zero Trust journey for IoT in your organization by evaluating and deploying a Zero Trust security model, applying Zero Trust to your existing IoT infrastructure, as well as using Zero Trust capabilities as criteria for selecting new IoT devices and services for your digital transformation.

Zero Trust security for the mobile workforce

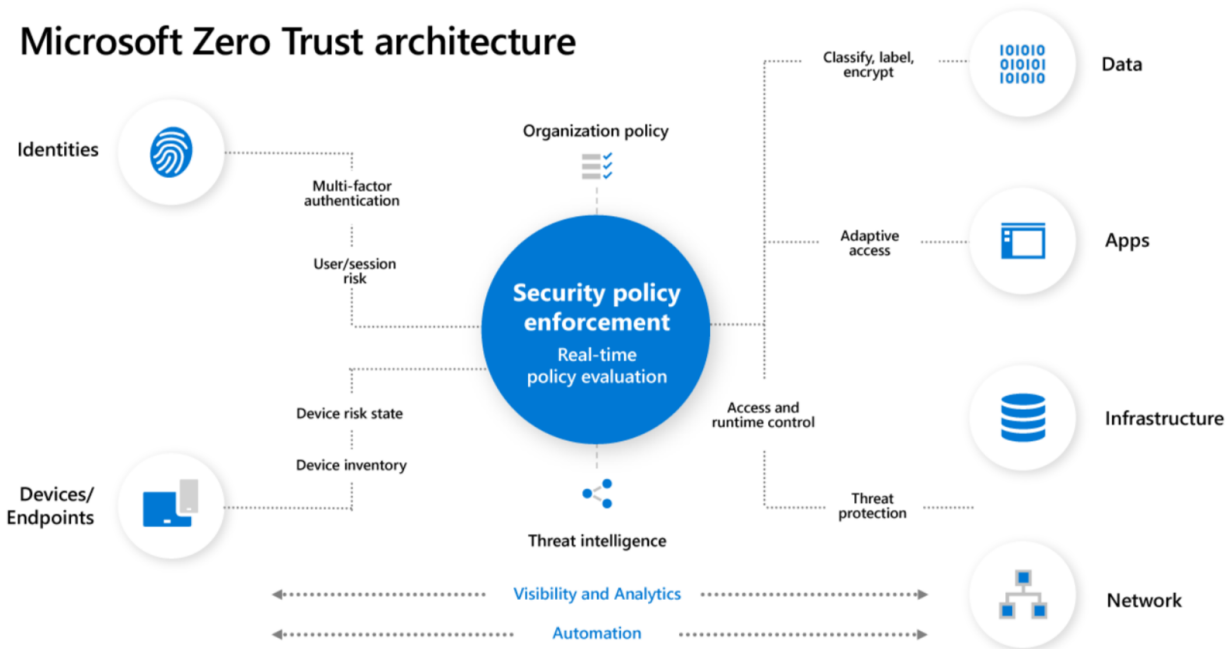
A traditional network security model often doesn't meet the security or user experience needs of modern organizations, including those who have embraced IoT in their digital transformation strategy. User and device interactions with corporate resources and services now often bypass on-premises, perimeter-based defenses. Organizations need a comprehensive security model that more effectively adapts to the complexity of the modern environment, embraces the mobile workforce, and protects their people, devices, applications, and data wherever they are.

Operating under the principles of "never trust" and "verify everything," the Zero Trust security model assumes breach and treats every access request as if it originates from an open network. In a Zero Trust model, every access request is strongly authenticated, authorized within policy constraints, and inspected for anomalies before access is granted. Three key principles of a Zero Trust security model are:

1. **Verify explicitly:** Always authenticate and authorize based on all available data points, including user identity, location, device health, service or workload, data classification, and anomalies.
2. **Use least-privileged access:** Limit user access with just-in-time and just-enough-access (JIT/JEA), risk-based adaptive policies, and data protection to help secure both data and productivity.
3. **Assume breach:** Minimize blast radius for breaches and prevent lateral movement by segmenting access by network, user, devices, and app awareness. Verify all sessions are encrypted end to end. Use analytics to get visibility, drive threat detection, and improve defenses.

The Zero Trust security model enables organizations to mitigate the risks of operating in an increasingly interconnected world. It addresses the challenge of users and applications accessing, analyzing, and acting on data from anywhere, from a variety of traditional and IoT devices, and outside of traditional corporate networks.

Microsoft Zero Trust architecture



Zero Trust and IoT

The Zero Trust security model used to protect today's modern workforce, data, and networks can also be applied to IoT in your organization for a holistic security approach. Application of Zero Trust to IoT is important for your organization given the potential for IoT compromise to cause real business impact. The unique characteristics of IoT make application of Zero Trust distinct when compared to its overall application to your workforce.

Potential impacts of security compromise

Compromise of IoT security posture has a potential for significant impacts to an organization's bottom line:

- **Operational and revenue impact:** IoT devices can be operationally degraded, used for lateral movement, or forced offline by a security incident, all of which can have business-critical impact to production, quality, and core business, leading to revenue loss.
- **Customer impact:** Incidents can degrade customer experience and influence brand reputation.
- **Real-world impact:** Compromise of cyber-physical systems (CPS) can lead to real-world effects beyond traditional IT information disclosure, including potential safety and environmental incidents (see [Compromise of U.S. Water Treatment Facility](#)).
- **Regulation impact:** Non-compliance may impact organizational ability to conform to government and industry regulations, up to and including potential for corporate officer personal liability.
- **Cost impact:** Security solutions to mitigate IoT business risks must be cost-effective in a world of low-margin industry and low-cost IoT devices, many with long lifetime expectancies.

IoT devices are electronic computing devices that collect and exchange data over the Internet. They are programmed to perform specific functions for a variety of applications



Technical characteristics of IoT

IoT solutions have technical challenges to consider beyond today's mobile workforce:

- **IoT devices are "userless" and run automated workloads:** IoT devices are often "userless" devices such as cameras, robots, and controllers. In the Zero Trust model for IoT, the "user" of these devices is actually the device itself acting without human interaction/input. Many workloads running on these devices on the edge are automated, deployed remotely as containers, and running constantly to enable critical business processes.
- **IoT device platforms are varied and integrate into an aging infrastructure:** IoT deployments often leverage existing infrastructure made up of aging devices and equipment designed for a disconnected world. Devices run on a mixture of operating systems from bare metal and [RTOS](#) to rich OS—many with no update capability enabled, often incorporating vulnerable open source components. There is a proliferation of IoT protocols, often proprietary and unencrypted. Devices can be expected to last more than 10 years, especially in cases where they are embedded in critical infrastructure (such as factories or transportation), potentially exposing them to vulnerabilities for much longer than the PCs and smartphones of the workforce.
- **Many IoT devices have limited capability and connectivity:** IoT devices can be small, MCU class, and may not be capable of running a full OS stack, security agents, or encryption. Limited processing power and size is often paired with the power constraints of running on a battery. Networking topologies can impact ability to keep devices and workloads managed, up-to-date, and monitored. Constraints can include remote connectivity using high-latency, low-bandwidth costed networks, completely air-gapped installations, industrial tiered ISA-95

The **edge** is a set of connected systems and devices that gather and analyze data—close to users, the data, or both. Users get real-time insights and experiences, delivered by highly responsive and contextually aware software applications. For additional insight: [Intelligent Edge – Future of Cloud Computing | Microsoft Azure](#)

'Purdue model' deployments, and integrated connectivity across cellular, Wi-Fi, and local (such as Bluetooth) stacks.

- **IoT devices can be high-value targets.** IoT devices used in critical operations and infrastructure can make attractive targets as they provide attackers opportunities for command and control that can have real-world impacts, such as the Triton attack. Even when not used for vital operations, the sheer number of IoT devices make them desirable targets for botnets to compromise at scale. For example, the [Mirai Botnet](#) used IoT devices at scale to cause widespread disruption of internet service. According to [Statista](#), 2020 was estimated as the inflection point when the number of IoT devices surpassed non-IoT devices. By 2025, the estimates are for 3x more IoT devices (31B) than non-IoT devices (10B).
- **IoT devices can be exposed to physical or local attack.** IoT devices are deployed in environments inside and outside of secured organizational spaces. For example, a PLC may be installed in a secured factory, but is exposed to insider threats from employees or contractors connecting to them via laptops or USB. A security camera or wind turbine may be installed outside and exposed to direct physical attack by adversaries. IoT devices deployed in public spaces such as grocery stores may be connected via networks that can also be locally accessed by the public.

Zero Trust capabilities for IoT

The first and foremost responsibility for securing IoT is not IoT specific—it's to implement Zero Trust solutions in the area of highest risk or widest opportunity across your organization's identity management, apps, devices, data, network, and infrastructure.

The [Zero Trust Assessment Tool](#) can help you evaluate where to get started in your Zero Trust journey, and the [Zero Trust Deployment Center](#) provides steps you can take today to begin securing workforce access to IoT solutions.

The potential blast radius of users gaining unauthorized access to IoT services and data in the cloud or on-premises is significant. It can not only lead to both mass information disclosure (for example, leaked production data of a factory), but also to potential elevation of privilege for command and control of cyber-physical systems (for example, stopping a factory production line).

Once your organization has improved the security of its workforce and core systems using the Zero Trust security model, it is essential to secure the IoT devices that are feeding the insights and actions loop of IoT solutions by focusing on these core Zero Trust capabilities for IoT:

- 1 **Strong identity:** To authenticate devices
- 2 **Least-privileged access:** To mitigate blast radius
- 3 **Device health:** To gate access or flag devices for remediation
- 4 **Continual updates:** To keep devices healthy
- 5 **Security monitoring & response:** To detect and respond to emerging threats





Strong identity

The first pillar of Zero Trust for IoT is strong identity for IoT devices. Strong device identity is delivered through tightly integrated capabilities of IoT devices and services, including:

- A hardware root of trust.
- Password-less authentication.
- Renewable credentials.
- Organizational IoT device registry.

Hardware root of trust (RoT): [The Seven Properties of Highly Secured Devices](#) defines a hardware root of trust as its first security property. This means strong identity for IoT devices starts with the device hardware itself, and the process by which devices are manufactured.

- **Lifetime:** Devices with a strong identity are “born” with a cryptographically backed identity that is inseparable from the device hardware. This identity represents the physical device and cannot be changed for the device’s lifetime (immutable). Its identity lifetime is similar to a person’s birth certificate, associated to a person for their entire life and cannot be changed. These immutable, lifetime identities are called a device’s “onboarding” identity.
- **Protection:** The secrets proving the device identity (credentials) must be secured in dedicated, tamper-resistant hardware, where secrets can be generated, stored, and processed. The use of dedicated hardware allows for these identity secrets to be isolated,¹ helping protect from unauthorized access such as malware.
- **Use:** Given their immutability and lifetime, device onboarding identities must only be used to onboard devices into IoT solutions, whereon renewable device credentials are then provisioned and used for regular device access (see [renewable credentials](#)). For a person, their onboarding identity is their name, backed by a birth certificate credential. A person’s “operational” identities can include their driver’s license ID and passport numbers, backed by driver’s license and passport documents issued and renewed by governmental agencies.

Password-less authentication: Password-less authentication, often using standardized X.509 certificates, is a strong mechanism to authenticate (prove) a device’s identity. This strength comes from its use of a private key to cryptographically sign/encrypt, and then using a distinct public key to verify/decrypt. The private/public asymmetric nature of certificate authentication allows for the secret private key to be isolated in hardware and never shared (for example, a RoT such as TPM), while enabling sharing of the known public key with IoT services and devices. This offers greater protection than “secrets”

¹ Supports the ‘small trusted computing base’ property of [The Seven Properties of Highly Secured Devices](#)

such as passwords and symmetric tokens, which rely on a form of secret shared between both parties (such as between the device and the service). To scale certificate usage and management, access is often granted to identities backed by a certificate issued by a designated (trusted) intermediate certificate authority.

After the initial onboarding authentication into an IoT solution, IoT devices should use their regularly provisioned renewable credentials for operational access. See renewable credentials for details.

Supporting IoT devices without strong identity, password-less authentication or renewable credentials: For IoT devices that are not manufactured with or capable of utilizing a strong identity, password-less authentication, or renewable credentials, IoT gateways can be used as “guardians” to locally interface with these less-capable devices, bridging them to access IoT services with strong identity patterns. This enables Zero Trust adoption today, while transitioning to use more capable devices over time.

Renewable credentials: After a device registers with an IoT solution with its onboarding credentials, it must be regularly issued renewable credentials for continuous, secure operations. A device’s renewable credentials back what is known as a device’s operational identity.

Certificates are a strong, standardized mechanism providing renewable, password-less authentication. Operational certificates must be provisioned from a trusted PKI, and have a renewal lifetime appropriate for the security posture of their business use. Their renewal must be automatic (often gated on [device health](#)) to minimize any potential access disruption

due to manual rotation. Any access granted to a device should be granted based on its operational identity. Credential revocation must be supported to enable immediate removal of device access (for example, to respond to compromise or theft).

Organizational IoT device registry: Similar to securing the user identities of an organization’s workforce to achieve Zero Trust security, it is crucial to have a registry for your organization’s IoT devices to manage their lifecycle and audit device access. A cloud-based identity registry is recommended for handling the scale, management, and security of IoT solutions. IoT device registry information is used to onboard devices into an IoT solution by verifying that the device identity and credentials are known and authorized to onboard to the organization. Once onboarded, the device registry contains the core properties of devices, including their operational identity and credentials used to authenticate for everyday use.

In cases where devices do not connect to Azure services for IoT, network sensors can be used to detect and inventory unmanaged IoT devices in an organization’s network for awareness and monitoring.

IoT device registry data can be used to view the inventory of an organization’s IoT devices (including health, patch, and security state), and to query and group devices for scaled operation, management, workload deployment, and access control.



Least-privileged access

Coupled with strong identity provided by integrated devices and services, Zero Trust requires least-privileged access control to limit any potential blast radius from authenticated identities that may have been compromised or running unapproved workloads.

In some cases, the critical asset to protect is the IoT device itself, such as a controller in a power plant. This reinforces the need to implement Zero Trust for your workforce to ensure their access to organizational resources such as critical devices requires privileged access levels.

For IoT scenarios, this means granting access to devices and their workloads using:

- Device and workload access control.
- Conditional access.

Device and workload access control: In cases of smaller, single-purpose IoT devices, granting device access is sufficient to provide access control to the scoped workloads running on the device.

In cases where workloads from multiple parties are running, these workloads should have discrete "app" identities to enable access control and auditing when combined with the device identity. This applies for workload access to:

- **Cloud resources:** For example, a battery level indicator app needs access to device management services, whereas the vehicle detection app needs access to AI services.
- **Local device resources:** For example, a battery level indicator app does not need access to a camera resource on a device, whereas a car detection AI app needs access to the camera feed.

For single-pane-of-glass access management, device and workload access control should be administered in an integrated identity and access management solution. Access management and auditing can be done by device, by application, or by a combination thereof.

Device and workload conditional access: When authorization is granted to devices and their workloads, it is also important to check the operating conditions of the caller's context as "dynamic" access signals. Useful device signals can include:

- **Location:** IP address, GPS, and other device properties can indicate if the connection is from an approved location. For example, if a device was removed from a factory, and is now connecting from another network IP address, access could be denied automatically.

- **Uniqueness:** A device should access services from one location at a time. If a device is accessing from multiple locations concurrently (or within an illogical timeframe), access could be denied. For example, a car accessing IoT services from Paris at 9:00 AM should not be able to access these services from New York at the same time, and not even an hour later at 10:00 AM.
- **Time of day:** Some devices or operations may only be permitted during business hours. For example, a bank vault that can't be opened outside of business hours.

Services can also use device signals to conditionally deploy workloads. For example, sensitive app workloads may be configured to deploy only to devices capable of protecting the workload locally with secure enclaves (such as SGX®/TrustZone®).

The overall "health" of a device itself is also a key signal in the Zero Trust model, as discussed next.

To layer least-privileged access for IoT devices, network segmentation can be used to group IoT devices, mitigating potential blast radius of a potential compromise. A common approach is to connect IoT devices to an "IoT network" (such as for printers, VoIP phones, and Smart TVs) separate from other organizational resources accessed by the corporate workforce.

Network micro-segmentation allows for isolation of less-capable devices at the network layer, either behind a gateway or on a discrete network segment. For example, dedicated OT environments can be logically separated from the corporate IT network using zone (DMZ) network architecture with firewalls. More mature organizations can also implement micro-segmentation policies at multiple layers of the [Purdue Model](#), typically using next-gen firewalls.



Device health

Following the Zero Trust principle of "trust but verify," device health should be used as a key factor in determining the risk profile (such as trust level) of a device. This risk profile can be used as an **access gate** to ensure only healthy devices can access IoT applications and services, or to identify devices in questionable health for **remediation action**.

- **Access gate:** Similar to how device health is used gate enterprise access of mobile workforce devices (such as phones or laptops), device health can be used to prevent IoT devices flagged as unhealthy from accessing IoT solutions. Unhealthy devices can be isolated, investigated, and then mitigated to get to healthy again.

- **Remediation action:** In cases where IoT devices are part of vital business operations and critical infrastructure, taking these IoT devices offline or blocking access may not be a feasible option. These devices can be flagged for mitigation to ensure that action can be taken to get to a healthy state again while remaining online for continuous operation.

There are several facets that can be coalesced to form a singular view on device health. An industry approach and one that many organizations are familiar with for their workforce in the IT space is to use a device vulnerability assessment product to assess device health. When looking at the industry standards that evaluate device health, they tend to offer several core capabilities which are directly applicable to an organization's IoT solutions:

1. **Security configuration assessment:** Answers the question, "Is the device configured securely?" as measured against a defined security configuration posture (such as open ports). A variant of this is the **Compliance configuration assessment**, where devices are measured against defined policy configuration postures (such as PCI).
2. **Vulnerability assessment:** Answers the question, "Is the device running software that is out of date, or software that has known vulnerabilities?" Given that IT and OT staff are overburdened, any system here must not just surface the vulnerable/out-of-date software—it needs to prioritize them. This prioritization can be done through a few different means, such as device count, device value, vulnerability severity, or active threats on the organizational networks.
3. **Insecure credential assessment:** Answers the question, "Is the device using secure credentials (such as certificates) and protocols (such as TLS 1.2+). It also assesses if the device uses weak or default credentials and/or policies. This is called out from the security configuration assessment as credentials are not always configurations and thus need to be managed separately.
4. **Active threats and threat alerts:** Indicates if there are potential threats running on the device. This is a combination of known threats and alerts that might indicate the presence of a threat.
5. **Anomalous behavioral alerts:** For IoT devices that are targeted (purpose-built) in their functionality, there is opportunity to use their behavior as a measure of device health. For example, an MRI device that is communicating with the master patient record database when it normally only writes out images may indicate an anomalous behavior. Or, an industrial gateway device that communicates to local devices outside of normal factory operating hours may indicate anomalous behavior.

All these elements need to be combined into a **device health** posture that is consumed by a risk assessment system for runtime decisions. Some of these elements can be captured through a system such as **attestation** (such a system uses a combination of hardware and cryptography to ensure that the measurements are tamper resistant), however, we must also rely on signals that are not attested. These signals together can be used for **device health**.



Continual updates

The other side of enabling your organization to control device access based on health is the need to proactively maintain production devices in a working, healthy **target state**.

The capabilities to proactively support healthy devices include:

1. **Centralized configuration and compliance management:** Device administrators need to have a straightforward approach to get the devices into a healthy security posture, and to apply policies that allow them to meet their industry-specific compliance needs (such as PCI). For example, enabling firewall rules to only allow inbound connections from known IP ranges.

In addition to managing settings, the centralized configuration also needs to address secure distribution and update of certificates used by IoT devices. These certificates can range from TLS certificates to certificates used by applications. This functionality is the corollary to the "insecure credential assessment" device health measure above, as it enables IT admins to get their devices to a healthy state.

2. **Deployable Updates:** The ability to keep devices up to date is a critical part of keeping devices protected. Customers need to be able to update the full set of software on devices, firmware, drivers, base OS and host applications, and cloud-deployed workloads. The update mechanism should have remote deployment capabilities as well as update rollout verification, and ideally should be integrated with pervasive security monitoring to enable automatic updates for security.

To achieve device health as an access gate and target state for Zero Trust IoT, they must be accessible in an integrated manner that enables organizational teams and their varying job functions to perform their tasked operations. A single pane of glass interface for each of the team roles provides a seamless, productive experience, itself governed through Zero Trust workforce policies enforcing least-privileged access.



Security monitoring and response

Security monitoring: Building on the foundation of strong identity, least privilege, healthy devices, and continual updates, there is also security monitoring necessary to rapidly identify unauthorized or compromised devices.

Monitoring adds an additional layer of protection for managed “greenfield” devices while also providing a compensating control for legacy unmanaged “brownfield” devices that don’t support agents and cannot easily be patched or configured remotely.

As [recommended by CISA](#), this includes:

- Generating an “as-is” asset inventory and network map of all IoT/OT devices.
- Identifying all communication protocols used across IoT/OT networks.
- Cataloging all external connections to and from IoT/OT networks.
- Identifying vulnerabilities in IoT/OT devices and using a risk-based approach to mitigate them.
- Implementing a vigilant monitoring program with anomaly detection to detect malicious cyber tactics like “living off the land” techniques within IoT/OT systems, such as monitoring for unauthorized changes to controllers.

Detection and response: Most IoT attacks follow a familiar “kill chain” pattern in which adversaries establish an initial foothold, elevate their privileges, and move laterally across the network. Often they will leverage privileged credentials to bypass barriers such as next-generation firewalls established to enforce network segmentation across subnets.

Rapidly detecting and responding to these multistage attacks requires a bird’s-eye view across IT, IoT, and OT networks, combined with automation, machine learning, and threat intelligence. By collecting signals from the entire environment—across all users, devices, applications, and infrastructure, both on-premises and in multiple clouds—and analyzing them in centralized Security Information and Event Management (SIEM) and Extended Detection and Response (XDR) platforms, security operations center (SOC) analysts can hunt for and uncover previously unknown threats.

Finally, Security Orchestration and Automated Response (SOAR) platforms are essential for responding to incidents rapidly and mitigating attacks before they have material impact on your organization. This is accomplished by defining playbooks that are automatically executed when specific incidents are detected. For example, you can automatically block or quarantine compromised devices so they are unable to infect other systems.

IoT kill chain example

In April 2019, security researchers in the Microsoft Threat Intelligence Center [discovered a campaign](#) compromising popular IoT devices across multiple customer locations. The attackers exploited simple vulnerabilities to deploy their malware, including default administrative credentials on a VoIP phone and printer and a missing patch on a video decoder.

After establishing initial beachheads on compromised devices, the attackers scanned the network to look for other insecure devices. They were also seen enumerating administrative groups to search for higher-privileged accounts that would grant access to higher-value data such as sensitive intellectual property.

As the actor moved from one device to another, they would drop a shell script to establish persistence on the network for extended access. Analysis of network traffic showed the devices were also communicating with an external command and control (C2) server.

This example demonstrates the need for Zero Trust strategies at all layers of the IoT infrastructure, including the need to protect privileged identities, continuously assess device health, perform centralized configuration and compliance management, monitor for anomalous behaviors, and segment networks to minimize the potential impact of a successful intrusion. It also highlights the need to leverage centralized automation and AI to quickly detect and respond to multistage attacks that touch diverse components of the digital estate.

An important part of security monitoring is security assessments and routine penetration testing done by external parties to detect potential vulnerabilities before adversaries do. IIC's [Security Maturity Model](#) can help assess the security risks for your business. For technical security assessments, these can include automated vulnerability assessments, design reviews, threat modeling, code analysis, as well as [gray box](#) penetration testing.

Next steps

1. Evaluate & deploy a Zero Trust security model

To start your organization's Zero Trust journey, learn more about Microsoft's approach to [Zero Trust](#), and use our Zero Trust Assessment to analyze the gaps in your current protection for identity, endpoints, apps, network, infrastructure and data. Use the recommended solutions from the assessment to prioritize your Zero Trust implementation, and move forward with guidance from the [Microsoft Zero Trust Deployment Center](#).

To help prioritize IoT Zero Trust investments, you can use IIC's [IoT Security Maturity Model](#) to help assess the security risks for your business.

2. Apply Zero Trust to your existing IoT infrastructure

As you deploy a Zero Trust model to your organization's workforce, you can expand its application to your organization's existing IoT devices and services by implementing the following:

- Deploy [Azure Defender for IoT](#) network sensors to **inventory all IoT devices; assess them for vulnerabilities and provide risk-based mitigation recommendations; and continuously monitor devices for anomalous or unauthorized behavior.** Defender for IoT is an agentless solution that continuously monitors network traffic using IoT-aware behavioral analytics to immediately identify unauthorized or compromised IoT devices. Additionally, it is tightly integrated with [Azure Sentinel](#)—Microsoft's cloud-native SIEM/SOAR platform—and supports third-party SOC solutions such as Splunk, IBM QRadar, and ServiceNow.
- Explore adding [Azure Sphere Guardian modules](#) to your critical brownfield devices to enable them to **connect to IoT services with Zero Trust capabilities** including strong identity, end-to-end encryption and regular security updates.
- **Implement granular network segmentation** for IoT devices based on their traffic patterns and risk exposure, in order to minimize the blast radius of compromised devices and ability for adversaries to pivot to higher-value assets. Network segmentation is typically accomplished using next-generation firewalls.

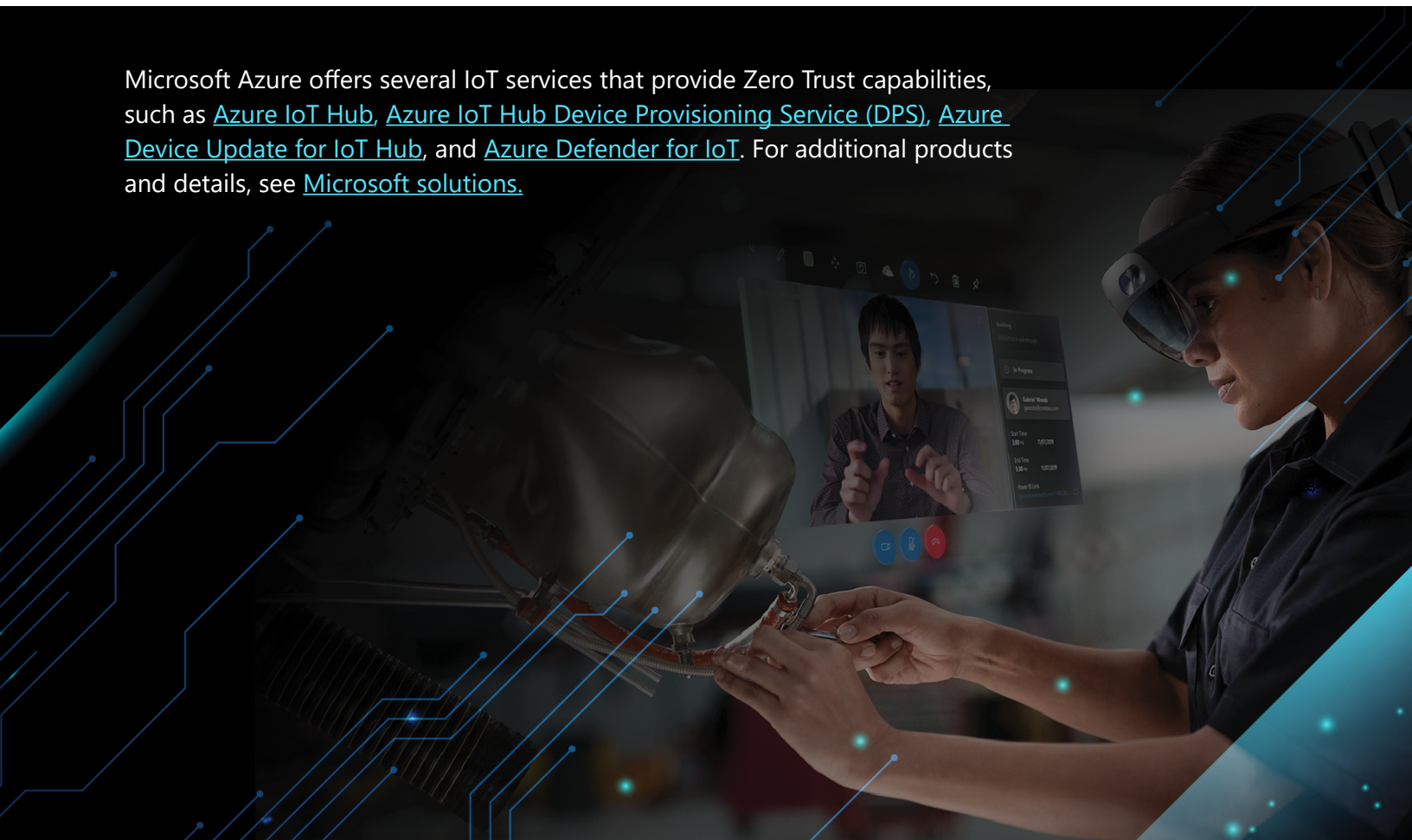
3. Use Zero Trust as criteria to select IoT devices and services

When selecting new IoT devices and services for purchase and integration, ensure they offer key Zero Trust capabilities:

IoT services supporting Zero Trust should:

- ✓ Enable full support for user access control via Zero Trust (for example, require strong user identities, multifactor authentication, conditional user access). Include integration with user access control systems for least-privileged access and conditional controls.
- ✓ Provide a central device registry for full device inventory and device management.
- ✓ Perform mutual authentication, offering renewable device credentials with strong identity verification.
- ✓ Enforce least-privileged device access control, and with conditional access to ensure only devices fulfilling criteria can connect (such as only healthy devices or from known locations.)
- ✓ Support over-the-air (OTA) updates to keep devices healthy.
- ✓ Enable security monitoring of both the IoT services themselves, and of the range of connected IoT devices.

Microsoft Azure offers several IoT services that provide Zero Trust capabilities, such as [Azure IoT Hub](#), [Azure IoT Hub Device Provisioning Service \(DPS\)](#), [Azure Device Update for IoT Hub](#), and [Azure Defender for IoT](#). For additional products and details, see [Microsoft solutions](#).



IoT devices supporting Zero Trust should:

- ✓ Contain a hardware root of trust that is utilized to provide a strong device identity.
- ✓ Leverage renewable credentials for regular operation and access.
- ✓ Enforce least-privileged access control to local device resources (such as cameras, storage, or sensors).
- ✓ Emit proper device health signals to services to enable their enforcement of conditional access.
- ✓ Provide updatability agents and corresponding software updates for the usable lifetime of the device to ensure security updates can be applied, along with device management capabilities to enable cloud-driven device configuration and automated security response.
- ✓ Run security agents that integrate with security monitoring, detection, and response systems.

Ideally, devices should be certified to make your purchase decisions with clarity and confidence.

Microsoft offers edge platforms including runtimes such as [Azure IoT Edge](#) and [Azure IoT platform SDKs](#), and operating systems including [Azure RTOS](#) and [Windows 10 IoT Enterprise](#). Microsoft also offers lightweight endpoint security agents that interoperate with [Azure IoT Hub](#) and [Azure Defender for IoT](#), with support for both Microsoft and Linux IoT platforms.

These Zero Trust-ready software platforms enable device builders and systems integrators to build Zero Trust-capable devices. For the device builder to reflect that their device offers Zero Trust capabilities, they should get the [Edge Secured-core certification](#), one of the certifications in the [Azure Certified Device program](#).

Microsoft also offers Zero Trust devices ready to meet your needs, including [Azure Sphere](#) and [Azure Percept](#).

For additional products and details, see [Microsoft solutions](#).

Microsoft solutions

Microsoft offers a range of service platforms, edge platforms, and integrated hardware platforms which offer Zero Trust capabilities for your organization's IoT deployments:

IoT services

[Azure IoT Hub Device Provisioning Service \(DPS\)](#) provides a central device registry for organizational devices to register for onboarding at scale. DPS accepts device certificates to enable onboarding with strong device identity and renewable credentials, registering devices in IoT Hub for their daily operation.

[Azure IoT Hub](#) supports an operational registry for organizational IoT devices. Accepts device-operational certificates to enable strong identity. Devices can be disabled centrally from Azure IoT Hub to prevent unauthorized connection. Azure IoT Hub supports provisioning of module identities in support of IoT Edge workloads.

[Azure Device Update \(ADU\)](#) enables you to deploy over-the-air updates (OTA) for your IoT devices. It provides a cloud-hosted solution to connect virtually any device. ADU supports a broad range of IoT operating systems—including Linux and [Azure RTOS](#) (real-time operating system)—and is extensible via open source.

[Azure Sentinel](#) is the industry's first cloud-native SIEM/SOAR platform on a hyperscale cloud, providing a bird's-eye view of security across your entire enterprise. By collecting data at cloud scale—across all users, devices, applications, and infrastructure, including firewall, NAC, and network switch devices—Sentinel can quickly spot anomalous behaviors indicating potential compromise of IoT/OT devices.

Azure Sentinel leverages machine learning and large-scale intelligence from decades of Microsoft security experience to accelerate threat detection and response. Additionally, it provides all the benefits of a cloud-based service including simplicity and elastic scaling, reducing costs as much as 48 percent compared to legacy SIEMs².

[Azure Defender for IoT](#) is an agentless, network-layer security platform delivering continuous asset discovery, vulnerability management, and threat detection for IoT and OT devices, including proprietary, embedded OT devices as well as legacy Windows systems commonly found in OT environments.

The solution uses patented IoT/OT-aware behavioral analytics and threat intelligence to quickly detect anomalous or unauthorized activities.

It can be rapidly deployed because it does not require any changes to existing IoT/OT devices. Azure Defender for IoT works with diverse industrial equipment and protocols from all major OT automation vendors (including legacy Windows systems that can't easily be upgraded), and interoperates with Azure Sentinel and other SOC tools (Splunk, IBM QRadar, ServiceNow). Interoperation with Azure Sentinel enables rapid detection of multistage attacks that often cross IT/OT boundaries, as well as automated response using IoT/OT-specific SOAR playbooks.

² [Commissioned study-The Total Economic Impact™ of Microsoft Azure Sentinel](#) conducted by Forrester Consulting, 2020

Azure Defender for IoT offers multiple deployment options including fully on-premises; cloud-connected to benefit from the scalability, simplicity, and continuous threat intelligence updates of a cloud-based service; or hybrid, in which selected alerts are sent to a cloud-based SIEM/SOAR.

For IoT device builders, Azure Defender for IoT also offers lightweight agents to embed strong device-layer security and endpoint detection and response (EDR) capabilities into new IoT/

OT devices. These micro-agents are offered with source code for major IoT operating systems including Linux and Azure RTOS. The agents communicate with the Azure Defender for IoT console in the Azure portal, which also interoperates with Azure Sentinel.

End-user organizations benefit from having multiple layers of protection including network-layer monitoring via the agentless solution and endpoint-layer monitoring via the micro-agents.

IoT edge platforms

Microsoft Azure has several edge platforms that can run on a range of hardware platforms to enable IoT scenarios integrated with Azure services:

[Azure IoT Edge](#) runtime provides an edge runtime connection to IoT Hub and other services.

- Support for certificates to be used as strong device identities.
- Support for PKCS#11 standard enables support for device manufacturing identities (and other secrets for operational identities, for example) to be stored on TPM/HSM.

[Azure IoT SDKs](#): The Azure [IoT platform SDKs](#) are a set of device client libraries, developer guides, samples, and documentation. The device SDKs help you to programmatically configure devices and to securely connect to Azure IoT services.

[Windows 10 IoT Enterprise](#) has significant security features that can be used to help ensure security across three key pillars of the IoT security spectrum:

- **Protect data.** Securing data means protecting it at all times, including at rest, during code execution, and in motion. This is done by using BitLocker Drive Encryption, Secure

Boot, Windows Defender Application Control, Windows Defender Exploit Guard, secure Universal Windows Platform (UWP) applications, Unified Write Filter, a secure communication stack, and security credential management.

- **Monitor and detect.** Device Health Attestation (DHA) lets you start with a trusted device and maintain trust over time. As the device runs, Azure Defender for IoT can help detect and protect against threats.
- **Update and manage.** You can use Device Update Center and Windows Server Update Services (WSUS) to apply the latest security patches. If you determine that a device might be exposed to a threat, you can remediate that threat by using Azure IoT Hub device management features, Microsoft Intune or third-party mobile device management (MDM) solutions, and Microsoft System Center Configuration Manager (Configuration Manager).

[Azure RTOS](#) provides a real-time operating system (RTOS) as C-language libraries that can be deployed on a wide range of embedded/IoT platforms. Included in the offering is a complete TCP/IP stack with TLS (1.2 and 1.3) and basic X.509 capability. Given the wide range of possible configurations and platforms, the responsibility of security for Azure RTOS applications leans heavily on the customer. However, Azure RTOS (along with the Azure IoT Embedded SDK) is designed to integrate with Azure IoT Hub, DPS, and Azure Defender and can utilize some of the same security mechanisms as larger, more expensive IoT devices. With features like X.509 mutual authentication and support for modern TLS cipher suites (ECDHE, AES-GCM), customers can build Azure RTOS applications knowing that the basics of secure network communication are covered.

In addition to the integration with Azure IoT services, Azure RTOS provides support for Zero Trust design on microcontroller platforms that support hardware security features like ARM TrustZone (a memory protection/partitioning architecture), secure element devices like the STSAFE-A110 from ST Microelectronics, and industry standards like the ARM Platform Security Architecture (PSA), which combines hardware and firmware to provide a standardized set of security features including secure boot, cryptography, attestation, and more. Combining a hardware root of trust from secure elements with the memory protection features of TrustZone with PSA-certified firmware, Azure RTOS customers have the ability to implement Zero Trust designs in even the smallest embedded IoT devices.

Along with the edge platforms above, Microsoft provides a program to certify devices, with a device catalog to indicate devices along with their certifications.

Device catalog/certification: The [Azure Certified Device program](#) empowers device partners to easily differentiate and promote devices, and enables solution builders and end customers to find IoT devices built to work well with Azure. From intelligent cameras to connected sensors to edge infrastructure, this enhanced IoT device certification program helps device builders increase their product visibility and saves customers time in building solutions. One of the certifications in this program is the [Edge Secured-core program](#), which validates whether devices meet additional security requirements around device identity, secure boot, operating system hardening, device updates, data protection, and vulnerability disclosures. The Edge Secured-core program requirements have been distilled from various industry requirements and security engineering points of view.

The Edge Secured-core program enables Azure services such as the Azure Attestation service to make conditional decisions based on the posture of the device, thus enabling the Zero Trust model. Some of the highlights consist of requiring the device to include a [hardware root of trust and provide secure boot and firmware protection](#). Attributes such as these can be measured by the attestation service and used by downstream services to conditionally grant access to sensitive resources.

Integrated IoT hardware platforms

Microsoft has several solutions available which provide hardware built specifically for IoT scenarios fully integrated with Azure services:

[Azure Sphere](#) is a fully managed integrated hardware, OS, and cloud platform solution for medium and low-power IoT devices that meets all seven properties of highly secured devices. Azure Sphere has several features that can help an organization implement Zero Trust. Devices are designed for explicit verification and implement certificate-based [Device Attestation and Authentication \(DAA\)](#), which will automatically renew trust. In addition to supporting the Zero Trust principle of explicit verification, Azure Sphere implements least-privileged access, where applications are denied access by default to all peripheral and connectivity options. This even extends to network connectivity, where permitted web domains must be included in the software manifest or the application is not able to connect outside of the device. Azure Sphere is built around the Zero Trust principle of assuming breach, even of its own software applications and OS. Protections are layered with defense in depth throughout the OS design, and a secure-world partition—running in Arm TrustZone on Azure Sphere devices—helps segment breaches to the OS from access to Pluton or hardware resources.

By adopting the principles of Zero Trust, Azure Sphere helps enable devices that can be used by an organization to apply Zero Trust models throughout their IoT deployments. In addition to the devices themselves, Azure Sphere can be applied as a **guardian module** to secure other devices, including existing brownfield systems that were not designed for trusted connectivity. In this scenario, an Azure Sphere guardian module will be deployed with an application designed to interface with an existing product through an interface—like ethernet, serial, or BLE—that doesn't necessarily have direct internet connectivity. The application then translates telemetry and controls from the product's domain to the cloud, utilizing

Azure Sphere's security to ensure continued trust. From the cloud perspective, this device operates no differently than if it was natively built for IoT. This guardian approach can help bring existing organizations into a Zero Trust model more quickly with a high bar for security even with existing products that have limited connectivity.

[Azure Percept](#) is a comprehensive, easy-to-use platform with added security for creating edge AI solutions. The end-to-end edge AI platform includes hardware accelerators integrated with Azure AI and IoT services, pre-built AI models, and solution management to help you start your proof of concept in minutes. Security measures built into your edge AI solution help protect your most sensitive and high-value assets.

Azure Percept uses the Zero Trust security model, so you can help safeguard your sensitive AI models and data in transit and at rest for your edge AI solutions:

- Azure Percept hardware accelerators are designed with a hardware root of trust, which interacts with a device TPM and attestation service to prove device identity and help ensure additional device security.
- Security for Azure Percept DK starts with custom firmware, which supports secure and measured boot as well as an A/B mechanism for resilient over-the-air updates, and leverages the onboard discrete TPM to implement anti-rollback protection features. The proper use of TPM-backed identity, attestation, and storage keys prevents attackers with one-time access to the device TPM from being able to persistently compromise, spoof, or steal device identity.
- The Azure Percept Vision and Audio AI system on module (SOM) also implement secure

- firmware with secure and measured boot and resilient update capabilities. The SOMs attest during every boot, leveraging the Trusted Computing Group's Dice Identifier Composition Engine (DICE) and an attestation service.
- AI inferencing workloads can be configured to protect machine learning models at rest and in transit to Azure Percept devices. Leveraging Azure Attestation services, the release of model encryption keys can be limited to healthy devices (devices that pass attestation).
 - Sensor data can also be protected by leveraging Azure Percept services to ensure that data collected from a vision sensor is encrypted on the device and in transit to the configured cloud storage account.
 - Azure Percept DK ships with [Microsoft CBL-Mariner OS](#). Mariner OS Zero Trust features include:
 - TPM support for hardware root of trust.
 - Platform measurements to validate device health.
 - Available agents for device updates, device management, and security monitoring.
 - Security-focused servicing cadence.



© 2021 Microsoft Corporation. All rights reserved. This whitepaper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT. This document is provided "as is." Information and views expressed in this document, including URL and other Internet website references, may change without notice. You bear the risk of using it. This document does not provide you with any legal rights to any intellectual property in any Microsoft product.