Introduction to Zero Trust Architecture

CCZT Study Guide

cloud security alliance®

CCZT™
Certificate of
Competence in Zero Trust

The official location for the Zero Trust Working Group is
https://cloudsecurityalliance.org/research/working-groups/zero-trust/

Disclaimer

Cloud Security Alliance designed and created this Zero Trust Training course study guide (the "Work") primarily as an educational resource for security and governance professionals. Cloud Security Alliance makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Version Number: 20240820

About Cloud Security Alliance

The Cloud Security Alliance℠ (CSA) (www.cloudsecurityalliance.org) is the world's leading organization dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment. Cloud Security Alliance harnesses the subject matter expertise of industry practitioners, associations, governments, and its corporate and individual members to offer cloud security-specific research, education, certification, events and products. Cloud Security Alliance activities, knowledge and extensive network benefit the entire community impacted by cloud—from providers and customers, to governments, entrepreneurs and the assurance industry—and provide a forum through which diverse parties can work together to create and maintain a trusted cloud ecosystem.

CSA Address

709 Dupont St.
Bellingham, WA 98225, USA
Phone: +1.360.746.2689
Fax: +1.206.832.3513

Contact us: support@cloudsecurityalliance.org
Website: https://cloudsecurityalliance.org/
Zero Trust Training Page: https://knowledge.cloudsecurityalliance.org/page/zero-trust-training
Zero Trust Advancement Center: https://cloudsecurityalliance.org/zt/
Provide Feedback: support@cloudsecurityalliance.org
CSA Circle Online Community: https://circle.cloudsecurityalliance.org/
Twitter: https://twitter.com/cloudsa
LinkedIn: www.linkedin.com/company/cloud/security/alliance
Facebook: www.facebook.com/csacloudfiles
CSA CloudBytes Channel: http://www.csacloudbytes.com/
CSA Research Channel: https://www.brighttalk.com/channel/16947/
CSA Youtube Channel: https://csaurl.org/youtube
CSA Blog: https://cloudsecurityalliance.org/blog/

Acknowledgments

Dedicated to Juanita Koilpillai, a pioneer in software-defined perimeters whose contributions to the Certificate of Competence in Zero Trust (CCZT), Zero Trust training, and CSA are immeasurable.

The CCZT and Zero Trust training was developed with the support of the Cloud Security Alliance Zero Trust Expert Group, whose members include volunteers from a wide variety of industries across the globe. Made up of subject matter experts with hands-on experience planning and implementing Zero Trust, both as cloud service consumers and providers, the Zero Trust Expert Group includes board members, the technical C-suite, as well as privacy, legal, internal audit, procurement, IT, security and development teams. From cumulative stakeholder input, the Zero Trust Expert Group established the value proposition, scope, learning objectives, and curriculum of the CCZT and Zero Trust training.

To learn more about the CCZT and Zero Trust training and ways to get involved please visit: https://cloudsecurityalliance.org/education/cczt

We would also like to thank our beta testers, who provided valuable feedback on the CCZT and Zero Trust training: https://cloudsecurityalliance.org/contributors/cczt-contributors

Lead Developers:

Abhishek R. Singh
Agnidipta Sarkar
Heinrich Smit
Juanita Koilpilla
Michael Roza
Michael J. Herndon
Michael Shurman
Prasad T.
Richard Lee
Sam Aiello
Vani Murthy

Contributing Editors:

Abbas Kudrati
Anusha Vaidyanathan
Jacob Kline
James Lam
Junaid Islam
Lauren Fishburn
Naresh Kurada
Remo Hardeman
Shruti Kulkarni

Expert Reviewer:

Alex Sharpe
Asad Ali
Matthew Meersman, PhD
Michael J. Herndon
Nishanth Singarapu
Rajesh Ingle, PhD
Robert D. Morris
Ron Martin (Dr.), PhD
Shinesa Cambric
Srinivas Tatipamula

CSA Global Staff:

Anna Schorr
Daniele Catteddu
Hannah Rock
Jenna Morrison
Leon Yen
Ryan Bergsma
Shamun Mahmud
Stephen Smith

Table of Contents

List of Figures

Course Intro

Welcome to *Introduction to Zero Trust* by Cloud Security Alliance. Please note that moving forward we will refer to Zero Trust Architecture as ZTA and to the Cloud Security Alliance as CSA. CSA is dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment across the globe. We hope you are as excited to learn about ZTA as we are about sharing this knowledge with you. This training module is part of a larger series of CSA programs on Zero Trust (ZT) that was created with the support of subject matter experts. If you are interested in volunteering with CSA to help our ongoing research efforts or are just interested in learning more about cloud security, please visit our website at cloudsecurityalliance.org.

In this training, we will provide an introduction to ZTA and ZT. This includes a discussion regarding ZT's relevance, followed by definitions, components, requirements, tenets, pillars, goals, objectives, and benefits of ZT. We'll also cover planning considerations and implementation options for ZTA, as well as use cases demonstrating how different topologies can work together to enhance security in environments assumed to be hostile. Diagrams, explanations, and references are provided to facilitate the learning process.

Course Structure

This introductory course on ZTA consists of seven units, each geared towards helping learners gain competency in a specific area/topic:

- Context of ZTA
- Definitions, Concepts, & Components of ZT
- Objectives of ZT
- Benefits of ZT
- Planning Considerations for ZTA
- ZTA Implementation Examples
- ZT Use Cases

Course Learning Objectives

After completing this course, learners will be able to do the following:

- Understand the foundations of ZT and ZTAs
- Explain ZTA's objectives and benefits
- Discuss possible planning considerations before implementing a ZTA
- Distinguish between the different ZTA implementation options
- Describe ZT use cases and applications

1 Context of ZTA

In this unit, you will learn how the various factors of the evolving technology landscape led to the emergence of ZTA, as well as explore ZT's roots and early approaches in both government and enterprise.

Organizations today are in a cycle of adopting new technologies by leveraging cloud services, either through platforms or by utilizing elastic computing. This means that while transformations are increasingly popular and technology adoption is the strategy for these organizations, their networks and security measures are equally under pressure to keep up with the changing environment and associated new risks.

Changes in the technology landscape, such as cloud computing, edge computing, and IoT, and the evolution of social behavior, such as increased requests for mobility, have led to organizations increasingly adopting distributed environments. Cloud computing, in all its combinations of delivery and deployments models is becoming the leading source of IT services[1]. The result is an increase in complexity for networks and service architectures, due to the need for integrating on-premises IT services with public cloud services, sensors, and actuators. In addition, the need to connect remote offices, remote workers, contractors, smart objects, and others has reinforced the requirement for more flexible, scalable, and secure network capabilities.

Similarly, data often resides in virtual environments outside the organization's premises and its physical control. However, the organization is still responsible and accountable for the data. From a data protection standpoint, traditional security architectures that focus on securing the physical network perimeter are increasingly ineffective in preventing cyber attacks.

This is where ZTA comes into play. ZTA is a model that creates virtual enclaves and grants access to resources inside of that enclave. Every transaction is vetted using the ZT concept of "never trust, always verify". In essence, ZT enables the designing of architectures from the inside out versus outside in.

1.1 History of ZT

ZT was first coined by John Kindervag around 2010 while working as a principal analyst at Forrester[2]. However, this concept was being researched much earlier by the Jericho Forum at the Open Group, and previously by the U.S Defense Information Systems Agency (DISA) and Department of Defense (DOD), with the Black Core project[3]. Kindervag, known as the grandfather of ZT, emphasized that all network traffic is untrusted. His position was that all requests to access data or resources should be verified at each step, with this being termed 'trust but always verify'.

[1] Cloud Security Alliance, "Security Guidance for Critical Areas of Focus in Cloud Computing v4.0," 26th, July 2017, https://cloudsecurityalliance.org/artifacts/security-guidance-v4/

[2] John Kindervag, "Build Security Into Your Network's DNA: The Zero Trust Network Architecture," 5th, November 2010, https://www.virtualstarmedia.com/downloads/Forrester_zero_trust_DNA.pdf

[3] In the CSA's literature on SDP, terms such as "black cloud" or "network darkening" have been discontinued in favor of more neutral terminology.

The earliest concept of ZT was based on a data-centric network design and leveraged micro-segmentation which mandated more granular rules and policies to ultimately limit lateral movement of attackers. As the concept of ZT continued to evolve, it took a more identity-centric approach. This trend accelerated with the adoption of mobility and cloud.

In 2013, Cloud Security Alliance's (CSA) Software-Defined Perimeter (SDP) concept was initiated. SDP was designed to create an invisible perimeter through a security architecture that requires positive identification of network connections from a single packet inspection prior to accessing resources. In 2014, Google implemented ZT for its employees, which motivated it to publish the BeyondCorp model. The approach revolved around the idea that the perimeter had expanded, hence traditional perimeter security and a protected intranet were no longer sufficient to protect against cyber threats. Google's BeyondCorp model shifted the access controls and policies from the perimeter to individual devices and users. It addressed the need to replace the traditional VPN while still allowing users to work securely from any untrusted network with a superior security posture.

Since its inception, the concept of ZT has extended the original security model beyond traditional infrastructure, databases, and network devices to include IoT, cloud environments, big data projects, DevOps environments, containers, and microservices. In 2018, Chase Cunningham and his team at Forrester published the *Zero Trust eXtended (ZTX) Ecosystem* report, which extends the original ZT model beyond its network focus to encompass today's ever-expanding attack surface. In August 2020, NIST announced the final publication of *Special Publication (SP) 800-207, Zero Trust Architecture*, which discusses the core logical components that make up a ZTA[4]. Clearly, ZT is gaining widespread adoption, even as it continues to evolve as a security model.

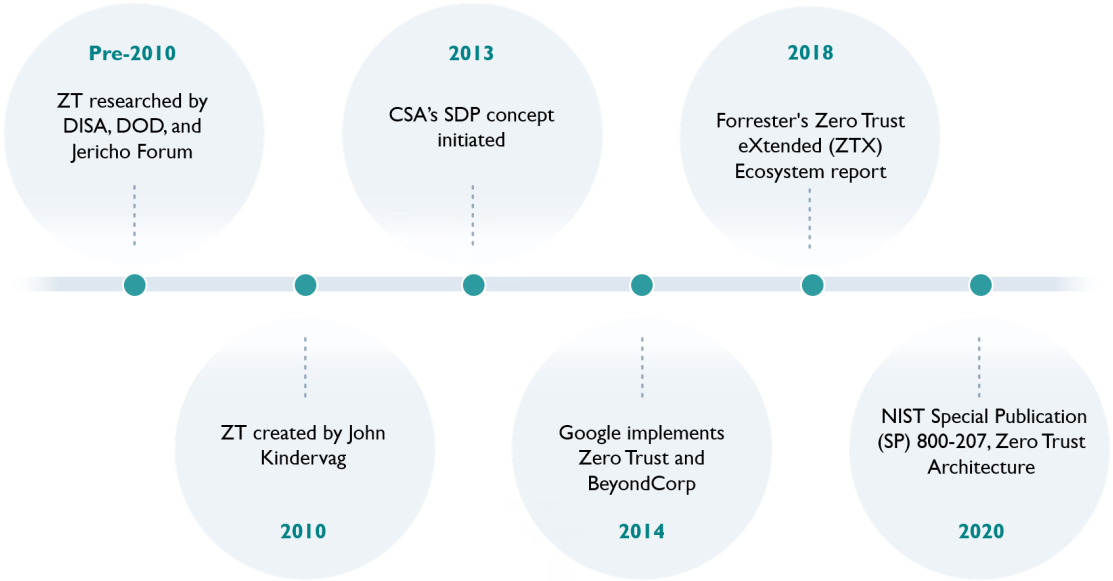


Figure 1: *ZT History and Milestones*

[4] NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://csrc.nist.gov/publications/detail/sp/800-207/final

2 Definitions, Concepts, & Components of ZT

In this unit, you will learn the definitions for key ZT terminology, as well as the concept's main tenets, design principles, pillars, and components and elements.

2.1 Definition[5] of the ZT Concept

ZT is a set of principles and practices designed for reducing cyber risk in today's dynamic IT environments. As a security model, ZT requires strict authentication and verification for each person, device, or service trying to access an IT resource, regardless of whether it is inside or outside the physical network perimeter. Since ZT emphasizes the protection of IT assets rather than network segments, the assessment of a given resource's security posture is not based on its location, but rather on what authentication and authorization controls are in place, and by leveraging risk-based analytics for access verification.

A key aspect of ZT networks is that authentication and explicit authorization must occur prior to network access being granted (e.g., the communication between a requesting entity and the target resource). Encrypting communications between two endpoints will no longer suffice; security practitioners must also ensure that access controls are implemented and each individual flow is confirmed as an authorized connection.

ZT lays out a blueprint for combating both internal and external threat agents trying to access protected assets. Research has shown that 90% of attacks start with a breach via a phishing email[6]. This exploit leads to the creation or compromise of an administrative account, followed by the lateral movement of malware inside the network, finally leading to the exfiltration of enterprise data.

In the context of this training and study guide, CSA defines the ZT concept as a cybersecurity approach that requires the following:

> - Making no assumptions about an entity's trustworthiness when it requests access to a resource
> - Starting with no pre-established entitlements, then relying on a construct that adds entitlements, as needed
> - Verifying all users, devices, workloads, network and data access, regardless of where, who, or to what resource, with the assumption that breaches are impending or have already occurred

Recent trends in enterprise security point to an increasing number of remote users and assets that are based in the cloud versus inside the traditional corporate network[7]. To meet the security challenges brought on by this shift, hardware manufacturers and software vendors are rapidly adopting the ZT model and validating that their products are fit for a ZT implementation.

[5] Note: CSA's working definition of ZT and ZTA is based on existing market definitions of ZT (e.g., as defined by Forrester, NIST, etc.). Throughout this study guide, CSA also incorporates material from normative reference documents developed by the ISO/IEC and IEEE.
[6] CISO, "Cybersecurity Threat Trends," 2021
[7] NIST, "SP 800-207 Zero Trust Architecture," August 2020

2.2 Tenets

A tenet is defined as a principle generally held to be true. According to the USA DOD, ZT has five major tenets[8].

1. **Assume a hostile environment:** Malicious actors reside both inside and outside the network. All users, devices, and networks/environments should be untrusted, by default.
2. **Assume breach:** Most large enterprises experience a barrage of attempted cybersecurity attacks against their networks every day and many have already been compromised. Create, manage, and defend resources with vigilance, assuming that an adversary already has a foothold in your environment. Access and authorization decisions should be scrutinized more closely to improve response outcomes.
3. **Never trust, always verify:** Deny access by default. Every device, user, application/workload, and data flow should be authenticated and explicitly authorized using least privilege, multiple attributes, and dynamic cybersecurity policies.
4. **Scrutinize explicitly:** All resources should be consistently accessed in a secure manner using multiple attributes— both dynamic and static— to derive confidence levels for determining contextual access to resources. Access is conditional and can change based on the action and resulting confidence levels.
5. **Apply unified analytics:** Apply unified analytics and behavioristics to data, applications, assets, and services (DAAS), and log each transaction.

2.3 Design Principles

Several design principles can be used to guide the creation of a ZTA[9]. These design principles include the following:

- Denying access until the requestor has been thoroughly authenticated and authorized withholding access until a user, device, or even an individual packet has been thoroughly inspected, authenticated, and authorized. The access to resources is temporary and reverification is required. The timespan of the access is defined by policies
- Allowing access to the network changes with ZT; requesters (users, machines, processes) aren't allowed access to anything until they authenticate who they are
- Allowing access to resources only after the requesting entity has been authorized
- Enforcing least privilege, specifically, granting the least amount of access required
- Requiring continuous monitoring of existing security controls' implementation and effectiveness (e.g., controls over access or user behavior)

2.4 Pillars

The ZT concept is a work-in-progress with boundaries and definitions that continue to evolve, especially in terms of scope of applicability and use cases. Even so, the industry has reached a certain level of consensus regarding what the fundamental pillars of a ZTA are. CSA emphasizes these seven

[8] DOD, "Department of Defense (DOD) Zero Trust Reference Architecture," February 2021

[9] ISO/IEC/IEEE 42010: 2011 defines "architecture" as: "The fundamental concepts or properties of a system in its environment embodied in its elements, relationships, and in the principles of its design and evolution."

pillars of the DOD ZTA[10].

1. **Users/identities:** Securing, limiting, and enforcing access for person, non-person, and federated entities' to DAAS, encompasses the use of identity, credential, and access management capabilities, such as multi-factor authentication (MFA) and continuous multi-factor authentication (CMFA). Organizations need the ability to continuously authenticate, authorize, and monitor activity patterns to govern users' access and privileges while protecting and securing all interactions. Role-based access control (RBAC) and attribute-based access control (ABAC) will apply to policies within this pillar in order to authorize users to access applications and data.

2. **Device/endpoints:** The ability to identify, authenticate, authorize, inventory, isolate, secure, remediate, and control all devices is essential in a ZT approach. Real-time attestation and patching of devices in an enterprise are critical functions. Some solutions, such as mobile device managers or comply-to-connect (C2C) programs, provide data that can be useful for device confidence assessments. Other assessments (e.g., examinations of compromise state, anomaly detection, software versions, protection status, encryption enablement, etc.) should be conducted for every access request.

3. **Network/environment:** When taking a ZT approach, organizations should logically and physically segment, isolate, and control the on-premise and off-premises network/environment with granular access and policy restrictions. As the perimeter becomes more granular through macro-segmentation, it enables micro-segmentation to provide greater protections and controls over DAAS. It is critical to (a) control privileged access, (b) manage internal and external data flows, and (c) prevent lateral movement.

4. **Applications and workload:** These should include tasks on systems or services on-premises, as well as applications or services running in a cloud environment. ZT workloads should span the complete application stack from application layer to hypervisor. Securing and properly managing the application layer as well as compute containers and virtual machines should be central to the ZT adoption. Application delivery methods like proxy technologies enable additional protections and therefore should also be an important part of ZT decision and enforcement points. Source code developed in-house and common libraries should be vetted through DevSecOps development practices to secure applications from inception.

5. **Data:** ZT protects critical data, assets, applications, and services. A clear understanding of an organization's DAAS is critical for the successful implementation of ZTA. Organizations should categorize their DAAS in terms of mission criticality and use this information to develop a comprehensive data management strategy, as part of their overall ZT approach. This can be achieved through the categorization of data, developing schemas, and encrypting data at rest and in transit. Solutions such as DRM, DLP, software-defined storage and granular data-tagging are crucial for protecting critical data.

6. **Visibility and analytics:** Vital, contextual details should be included to provide a greater understanding of performance, behavior, and activity baselines across the various ZT

[10] DOD, "Department of Defense (DOD) Zero Trust Reference Architecture," February 2021

pillars. This visibility improves the detection of anomalous behavior and provides the ability to make dynamic changes to security policies and real-time contextual access decisions. Additionally, other monitoring data from sensors, in addition to telemetry, are used to provide situational awareness in the environment. This will aid in the triggering of alerts used for response. A ZT enterprise will capture and inspect traffic, looking beyond network telemetry and into the packets themselves to observe threats and bolster defences more appropriately.

7. **Automation and orchestration:** ZT includes automating manual security processes to take policy-based actions across the enterprise with speed and at scale. Security orchestration, automation, and response (SOAR) improves security and decreases incident response times by automating responses to threats. Security orchestration integrates security information and event management (SIEM) with other automated security tools in the management of disparate security systems. In order to provide proactive command and control, automated security responses require defined processes and consistent security policy enforcement across all environments in a ZT enterprise.

8. **Governance:** This is essential to ensure successful implementation and control over goals, requirements, and actions taken. A formal procedure for governance should be established through a review committee that will evaluate the progress made towards meeting objectives, ensuring that plans are funded, and assessing associated risks with future phases.

2.5 Components & Elements

At a high level, ZTA requires three core components before any logic can be applied to allow a decision to be made for access:

1. **Communication:** A request for an entity to access a resource, and the resulting access or session
2. **Identity:** The identity of the entity (e.g., user or device) requesting access to the resources
3. **Resources:** Any assets within the target environment

In addition to these three core components, ZT is also composed of two other fundamental elements:

1. **Policy:** The governance rules that define the *who, what, when, how, why* of access to the target resource access
2. **Data sources:** The contextual information providers can use to keep policies dynamically updated

The applicability of all of these components and elements will depend on your use cases and deployment models.

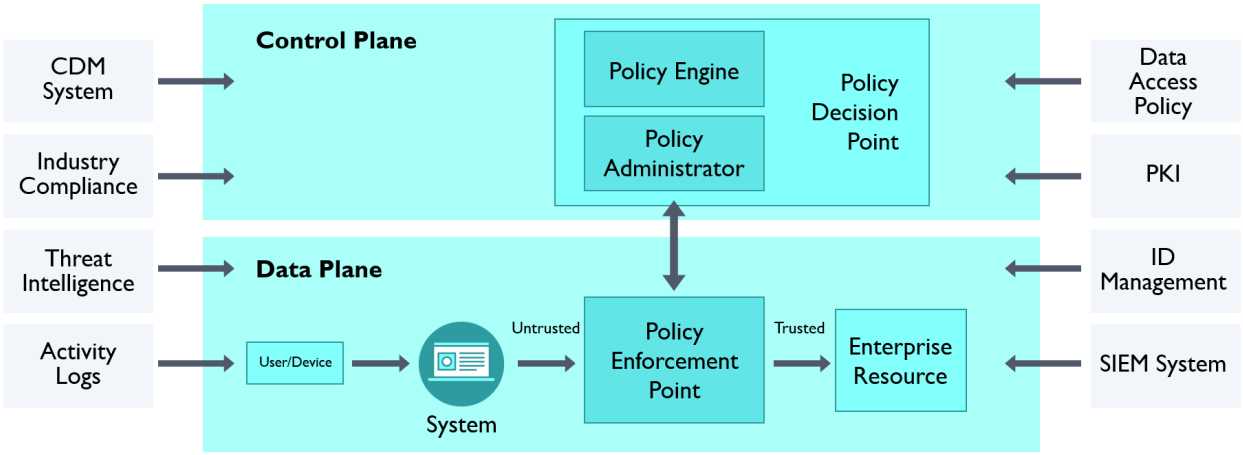


Figure 2: *Key Logical Components of a ZTA*[11]

In the publication, *SP 800-207*, NIST has provided a simple representation of the key logical components of a ZTA (see diagram above). In the NIST ZT workflow the policies are defined, managed, and enforced via the following two mechanisms:

- Policy decision point (PDP)
- Policy enforcement point (PEP)

Together, the PDP and PEP regulate access to resources by being placed in the access workflow of traffic.

The PDP is composed of a policy administrator and policy engine (PE). The PDP determines the *rules* and communicates them to the PEP. The PEP acts as a gateway to ensure that access to an approved resource has been granted to the correct entity, with the correct access levels.

NIST defines the following[12]:

- PDP as the control plane: the component of the logical architecture that has the responsibility to collect, analyze, and transform the data first into intelligence and then into rules to govern the access to resources.
- PEP as the data plane: the component that, based on input passed by the control plane, has the responsibility to enforce the rules and provide access to the resources (i.e., data).

Data sources serve the purpose of feeding data into the PDP, with the goal of maintaining the rules and keeping the overall decision-making process updated. Various sources of intelligence feed into the policy engine and support the policy administrator in defining and/or refining the access rules.

[11] Figure adapted from NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://csrc.nist.gov/publications/detail/sp/800-207/final

[12] NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

The following is a list of the possible information sources for the policy engine:

- Intrusion detection system (IDS)/Intrusion detection and prevention system (IDPS)
- Network devices (e.g., firewalls, proxies, gateways, routers, etc.)
- Threat intelligence feeds (e.g., third party databases of threats, vulnerabilities, weaknesses, and exploits)
- Information sharing systems
- Denylists and blocklists
- Identity providers and access management systems (e.g., Active Directory [AD] or cloud access security brokers [CASBs])
- Legal and regulatory compliance requirements
- Asset/device management and discovery systems
- Public key infrastructure (e.g., certificate revocation lists)

The figure below provides an alternative representation of the data flows and data sources that feed into the PDPs and PEPs.

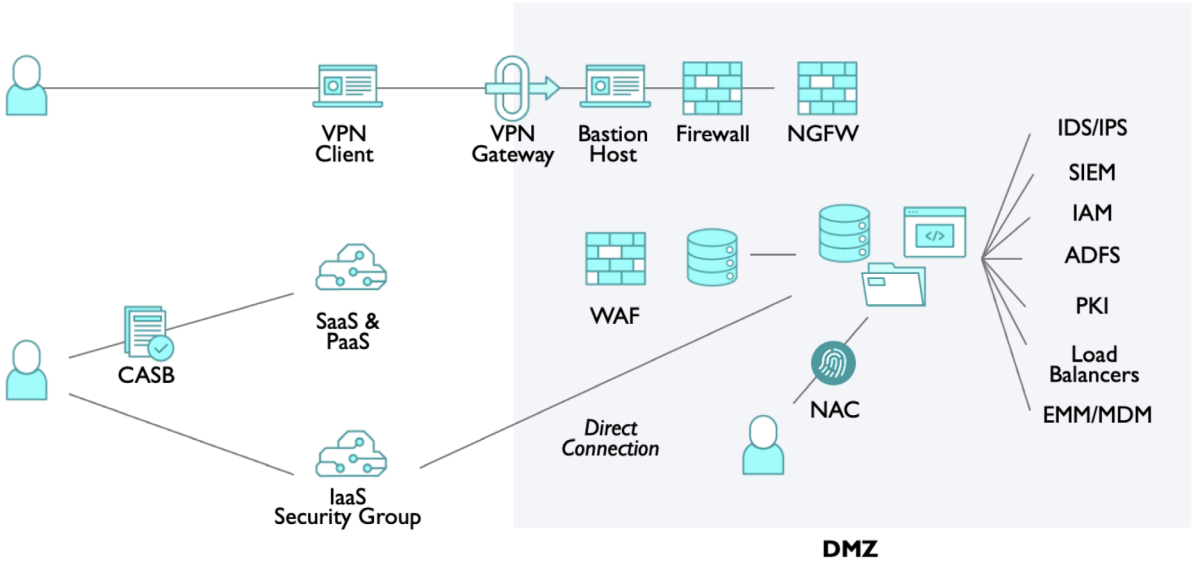


Figure 3: *PDP and PEP Data Flows and Sources*[13]

Security incident and event monitoring databases can be a collection point for any/all of the above sources. Together, these components have telemetry information relating to all the core components of ZTA. This gives enterprises more context to make better informed policy decisions.

Due to the greatly increased number of PEPs, manual management of the access model can be challenging and is not recommended. Instead, automation represents another important characteristic of a ZT environment, as it supports both granular and global control.

[13] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

3 Objectives of ZT

In this unit, you will learn how ZT addresses the main technical and business objectives related to reducing cyber risk in an organization.

As with most security architectures, the primary objective of ZTA is to address security risks inherent in the assumption of trust, and the lack of proper access controls. Typical approaches to addressing these risks include reducing the attack surface and/or improving the effectiveness of security controls.

The motivation behind ZTA is to provide a holistic and consistent security approach for protecting an enterprise against malicious actors both internal and external — threats that exploit inherent or newly-created gaps in conventional protection methods and defense-in-depth controls. The key differentiator in ZTA is the **ephemeral nature of any trust** between data/computing resources and the principals requesting access. This differentiator, combined with capabilities like dynamic policy enforcement and decisioning, bolster an environment's security posture, from the cloud to on premises. This is true for both internal and external attacks that exploit and compromise exposed access mechanisms maliciously.

A ZT approach fulfills both technical and business objectives. Technically, it establishes a framework for protecting resources, simplifies the user experience, reduces the organization's attack surface size and complexity, enforces least privilege, improves control and resilience, and localizes the impact radius of a security failure. From a business perspective, ZT aims to reduce risk, improve governance and regulatory compliance, and align the organization's culture with the risk appetite of its leadership.

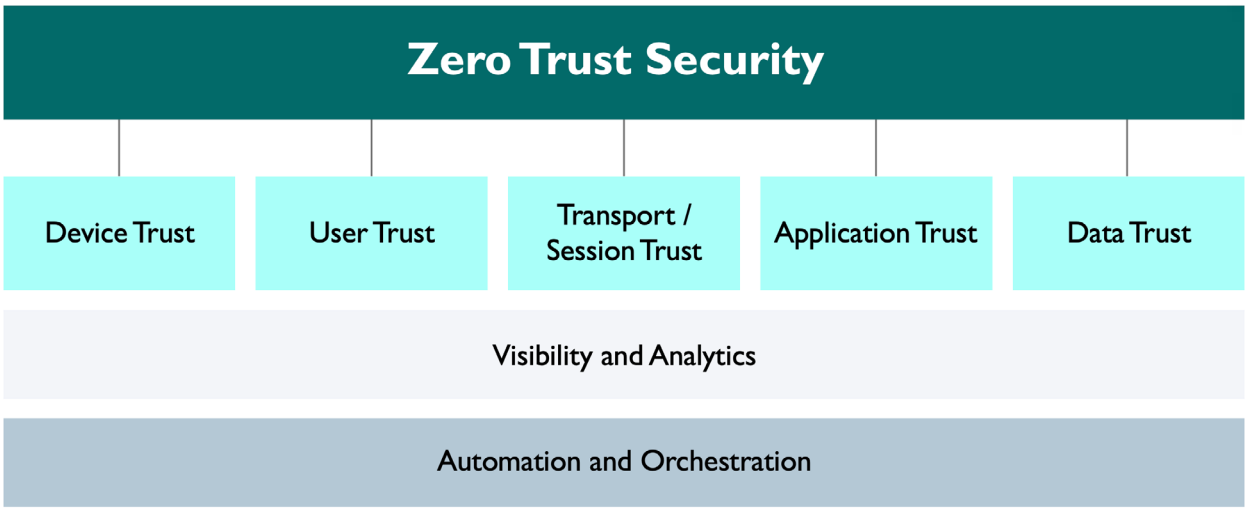
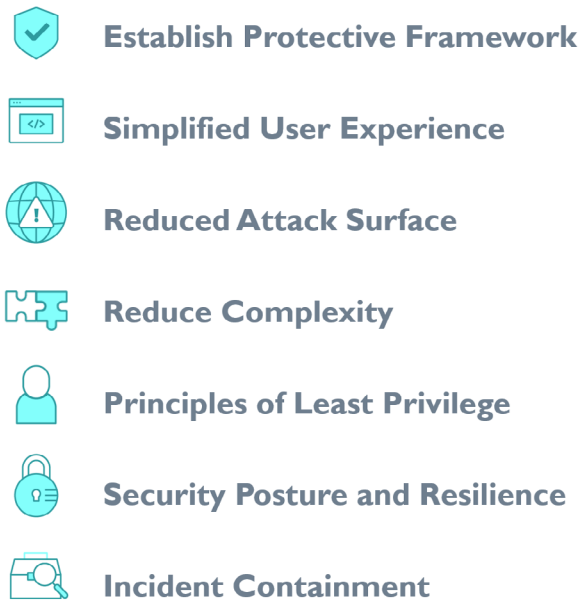


Figure 4: *ZT Concept Framework and Elements*[14]

[14] Figure adapted from ACT-IAC, "Zero Trust Cybersecurity Current Trends,", 18th, April 2019, https://www.actiac.org/system/files/ACT-IAC%20Zero%20Trust%20Project%20Report%2004182019.pdf

3.1 Technical Objectives

The following technical objectives serve as critical milestones for organizations looking to adopt ZTA. These objectives include activities and efforts related to the implementation of specific technologies and supporting security frameworks.



3.1.1 Establishing a Protective Framework

The protective framework established by ZT represents a novel approach to cybersecurity. As mentioned previously, ZT's core premise is that an organization should not inherently trust any entity that comes from within or beyond its boundaries. This new protective framework enables a shift of focus to more business oriented goals, with systems designed around the value of the data and their specific protection needs. Many procedures and strategies that were once considered strong security measures are no longer fully effective; as a result, aged cybersecurity techniques and technology will increasingly yield limited results and inadequate protection.

It is no longer practical to use approaches and frameworks based on physical objects and systems, nor is it effective to rely on signature-based threat detection. The increasing frequency and scale of attacks, combined with today's hyper connected world, virtualized environments, and software-based organizations, requires businesses to reconsider everything from network configurations to detection and prevention approaches.

3.1.2 Reduce Management Overhead

ZTA reduces management overhead by applying a consistent access model throughout the environment for all assets, from network devices to virtual servers and applications. Every request for access, whether explicit or implicit, is met with the same interrogation: Who are you? Do you need this access now? Okay, you get this access to this resource for this period.

To support this uniformity, ZTA models are absent of the following:

- Complicated diagrams of nested groups using legacy access control lists (ACL) with allow and deny parameters producing unexpected results
- Layers of groups managed by potentially irrelevant decision-makers
- Stale and orphaned groups whose owners have long since moved on
- Authorization mechanisms based on antiquated models/labels (e.g., local vs. global)
- Delays in provisioning, deprovisioning, or access revocation, since every request is handled consistently, just-in-time by the PDPs

3.1.3 Reduce Attack Surface

In a traditional security architecture, access decisions made at the network perimeter either allow or deny access. Denied traffic is dropped outside of the perimeter, while allowed traffic enters the perceived secure environment and travels unencrypted, as it is rare for organizations to encrypt internal traffic. Once inside, an attacker may run port scans, find vulnerabilities, launch denial-of-service attacks, steal additional credentials, eavesdrop on privileged network traffic, and move laterally unobstructed with relative ease. In contrast, with the ZT model the same attacker is no better off than if they had not penetrated the system's external defenses, because each internal resource makes a decision as to whether or not to grant access at any given moment. The organization's attack surface effectively contracts from every resource to only improperly secured resources.

3.1.4 Reduce Complexity

An organization's ever-expanding digital footprint makes for an increasingly complex IT environment, especially with some access decisions being made far in advance of being requested/used or even necessary. Access levels often remain, even as the party granting the access has long since moved on, leaving behind orphaned objects with unmanaged permissions. Such complexity represents one of the biggest security challenges for an organization, as it further reduces visibility, complicates configurations, creates weaknesses and vulnerabilities, and generally makes it easier for malicious actors to gain a foothold in the network.

Additionally, the adoption of newer IT paradigms like hybrid cloud implementations, multi-cloud architectures, and edge computing also further complicates the access control policy management. ZT reduces this complexity by assuming that all parties requesting application access are malicious and should therefore be untrusted. Instead of trying to police all the borders and paths across the network, security professionals need only create islands of applications and data to protect in a more focused manner. This is because ZT strategies require far more attributes than standard security mechanisms. As organizations strive for agility by simplifying networks and consolidating data centers, ZT provides a robust security mechanism to reduce security architecture complexity by creating perimeters around applications and identity. This also reduces the number of access points into an enterprise's IT environment, resulting in tighter control over each identity's level of access and privileges, including third parties like vendors and suppliers.

3.1.5 Enforces the Principle of Least Privilege

ZT enforces the principle of least privilege, which dictates that users and programs should only have the necessary privileges to complete their tasks. Per ZT, users get access to exactly what they need to conduct their business, when they need it. ZT also includes the use of micro-segmentation, or the creation of zones in an IT environment to isolate workloads for better security. This enables users to connect to the right application and use only the services they require. This simplified access provisioning makes it easier to manage security operations and governance teams in a continuously evolving security landscape. ZT also includes the use of purpose based dedicated identities also known as identity personas. Identity persona is created for a group of resources that address a common functionality, which helps in limiting the attack surface created by the compromise of an identity.

3.1.6 Improved Security Posture & Resilience

The objective of ZT is to enhance and bolster the resilience and the security posture of an enterprise's IT infrastructure. From outside of the organization, ZTA ensures that malicious actors have reduced visibility into the enterprise's IT infrastructure and individual assets, thereby reducing the potential attack vectors at their disposal. From within the organization, ZTA restricts lateral movement to minimize the risk of cross-site attacks and damage inflicted by insider threats. Because external users are contained and controlled within a small area of the network, any resulting security issues can be quickly contained and addressed. ZT limits the impact radius of security incidents and enables the swift return of systems to their earlier state.

The reduced attack surface ensures that any source scanning and mapping activities initiated by internal or external actors are not successful unless they are authorized within the ZT implementation. The two-layer architecture consisting of a separated control plane and data plan helps ensure that access is granted to the organization's network only after the users and their devices have been properly authenticated and authorized.

3.1.7 Improved Incident Containment & Management

A primary goal of ZTA is to make the incident management process more effective and efficient; to this end, several of ZTA's core design principles like "never trust, always verify" and the presumption of an ongoing breach require continuous behavioral monitoring of all system entities.

Micro-segmentation and the requirement for continuous network access authorization reduces the impact radius of potential breaches, as it restricts a cyber attacker's ability to move laterally. When a breach does occur, damage is limited to a confined area and containment/eradication and remediation efforts can be carried out with respect to the incident's scope.

The continuous monitoring capabilities included in ZTA allow for more effective identification of anomalies and incidents. The incident-related data is also used to update the PDP, allowing for dynamic policy definition/enforcement critical to limiting the impact across the organization's network.

3.2 Business Objectives

The following key business objectives can serve as critical milestones for organizations looking to align ZT adoption efforts with ongoing, high-level operational needs. These include the overall reduction of both compliance and cyber risk, as well as the fostering of a ZT-based organizational culture.

3.2.1 Risk Reduction

A primary business goal of ZTA is the reduction of cyber risk. This is especially critical for organizations dealing with complexity brought on by the proliferation of distributed, open computing infrastructures and the enterprise's migration to the public cloud. The risk reduction objective relates to some of the technical goals and objectives mentioned in the previous section, such as reducing the attack surface and achieving/maintaining an improved and resilient security posture.

Chiefly, ZTA aims to reduce the risk of the following:

- Improper privilege escalation via lateral movement
- Access beyond the need to know requirements
- Access beyond the required time frame
- Access by unsecure devices
- Access via unsecured methods such as unencrypted channels or channels using invalid certificates
- Compromises using methods like brute force, distributed denial-of-service (DDoS), or man in the middle (MITM) attacks
- Unauthorized lateral movement

Additionally, ZT supports the adoption of MFA to protect logins against common brute-force attacks, dictionary attacks, or stolen credentials attacks. In alignment with the ZT model, users and devices are validated before gaining access to protected resources and mutual authentication occurs between the server and client when the connection is being established.

Implemented in all ZTA variants, the principle of least privilege is effective in mitigating the most sophisticated and difficult to detect internal attacks. ZT's level of granularity prevents users from accessing unauthorized resources, as controls and policies are applied separately to every protected resource, for every access request. In addition, all communications between clients and servers flow through mutually authenticated encrypted tunnels, creating extended micro-segmentation systems in lock step.

ZT also includes continuous monitoring as a critical requirement for cyber risk reduction. To maintain a strong security posture, enterprises should continuously monitor all resource access activity and investigate potential signs of compromise. Since ZTA is policy-based, the risk of unauthorized access by compromised accounts can be mitigated, since policies can be conditioned on user and device security posture.

Above all, the ZT model reduces the total risk of running a connected enterprise by using one unified framework, typically provided by a limited number of vendors. This allows an enterprise to mitigate all the major threats that previously required multiple solutions, each with its own drawbacks and security flaws.

3.2.2 Compliance Management

A primary objective of ZT is to help organizations achieve and maintain an optimal compliance posture, reducing both the financial and technical impact of compliance, internal and external. This is mainly achieved through two key ZT features: (1) discovery and (2) mapping out of all networked assets and related access controls. ZT requires that assets are automatically discovered and validated for alignment with the latest compliance requirements since assets and data can only be protected if their presence is known. ZT helps segregate resources based on the relevant legal, regulatory, and contractual compliance requirements.

A proper implementation of ZT verifies authentication and authorization each time traffic moves laterally or inside/outside the network. This approach prevents unauthorized access before data can be accessed, compromised, encrypted for ransom, or exfiltrated. Additionally, it creates an audit trail for satisfying regulatory requirements regarding record keeping and auditing.

The benefits of ZT are instrumental to an organization's efforts in maintaining regulatory compliance. Privacy-related regulations such as General Data Protection Regulation (GDPR)[15] and the California Consumer Privacy Act (CCPA) define stringent requirements for processing and storing personally identifiable information (PII). Organizations must build an accountability framework for maintaining control and visibility over PII: how it is collected, processed, stored, where it resides, for what purpose, how, and by whom; with these components in place, organizations can implement the proper security controls for protecting PII from internal and external threats. ZT enables organizations to better align with standard security practices integrated into existing regulatory requirements' internal controls.

[15] See for instance GDPR Article 30, "Records of processing activities"

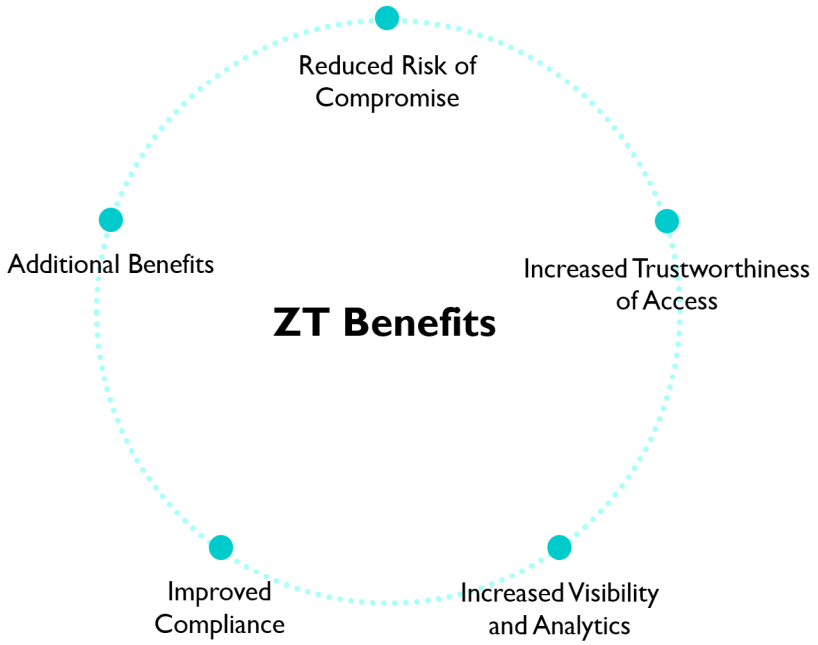
3.2.3 Organizational Improvements

The ZT model's "never trust, always verify" approach results in significant changes to the organization's mindset regarding how resources are accessed, as it requires enterprises to adopt a coordinated, structured approach to cybersecurity. Organizations must shift to a culture based on processes and procedures that support continuous verification — only then can each entity within the company's IT environment be trusted at any given moment in time.

4 Benefits of ZT

In this unit, you will discover the range of benefits that ZT adds to an organization's security efforts, from reducing the risk of compromise to increased visibility and improved compliance.

ZT provides a myriad of benefits for strengthening the cybersecurity posture of an organization, both on-premises and in the cloud. These include, but are not limited to:



Collectively, ZT's benefits enable organizations to bolster their defenses against internal and external threats, reduce cyber risk and improve adherence to compliance frameworks.

4.1 Reduced Risk of Compromise

One of the main benefits of ZT is that it reduces risk of compromise, primarily through the following:

- Reducing the attack surface and limiting the radius of impact
- Reducing an attacker's ability to move laterally
- Reducing the time to detect and contain breaches

4.1.1 Reduced Attack Surface & Impact Radius

The principle of least privilege and "never trust, always verify" are at the very core of ZT. Resources are accessed based on the attributes of the entity or user, security hygiene of the device, context of the request, and relative risk to the environment. This reduces the risk of unauthorized access and escalation of privileges.

In addition, ZTA implementations leverage the concept of resource hiding, where resources are only visible to authenticated, authorized users. This concept is described in various ways depending on the ZTA implementation technique.

As described in NIST *SP 800-207*[16], a user sends a request from the system (e.g., a laptop) to the PEP to access a resource. The PEP forwards the request to the PDP for authorization, which in turn checks if the user has been authenticated and authorized by policy. The PDP then sends its response to the PEP.

This variation of the agent or gateway deployment model implies the use of vetted, compartmentalized applications or processes (e.g., virtual machines, containers, or some other implementation); regardless of what technology is being used, the goal is the same: to protect the application or application instances from potentially compromised hosts or other applications sharing the same server resources. According to this model, the server only runs approved, vetted applications in a sandbox; these applications can communicate with the PEP to request access to resources, but the PEP will refuse requests from other applications running on the server. In this model, the PEP could be an enterprise service running locally or a cloud service.

4.1.2 Reduced Ability to Move Laterally

ZT calls for the implementation of micro-segmentation to restrict lateral movement inside an enterprise IT environment, thereby reducing the attack surface and potential impact radius. Each access attempt to any resource—internal as well as external—is authenticated and authorized before access is granted, regardless of the requester's origin.

4.1.3 Reduced Time to Detect & Contain Breaches

ZTA's centralized authentication and policy enforcement enables improved visibility into all access attempts across multiple cloud providers and on-premises IT infrastructures. This visibility, in conjunction with dynamic access policies, enables organizations to detect malicious access attempts in real-time and mitigate attacks before they cause damage. By adopting ZTA, organizations increase their level of continuous verification and capability in detecting threats like phishing attempts, privilege elevation for accessing applications and services, and/or the use of stolen credentials. Early detection of these threats can often stop attackers from launching a successful intrusion attempt.

[16] NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://csrc.nist.gov/publications/detail/sp/800-207/final

4.2 Increased Trustworthiness of Access

ZTA increases the trustworthiness of data by distrusting anyone inside or outside the organization's perimeter. ZTA considers consolidated identity access management (IAM) and policy solutions capable of managing access across the organization's entire environment, providing a single source of truth for identity, and supporting single sign-on (SSO) as a fundamental capability. User authentication is centralized with authentication being strong, dynamic, and strictly enforced before access is allowed. This is supported by MFA, session timeouts, re-authentication requests, and validation. These steps are equally applied to any layer in the stack.

Granular access and permissions are configured based on roles, context, or attributes as applicable. Access to resources is based on the principles of least privilege and need to know. Access to any data is protected cryptographically based on its sensitivity — whether it is at rest, in motion, or being processed.

In summary, from the perspective of access to the resources, some of the benefits of a ZTA are:

- Granular access and permissions, and ability to grant access based on context
- Authentication of device and user before granting access to network and resources
- Enforcement of the least privilege rule
- Strong authentication, including MFA
- Centralized access control
- Continuous validation of identity, authentication, and authorization to resources
- Improved data protection

Additionally, some ZTA methods incorporate single packet authorization (SPA), which also helps increase the trustworthiness of access. SPA uses a next generation passive authentication technology that features no open ports and service listeners; instead, a specialized encrypted packet is used in the following procedure:

- The first SPA packet sent by the client is rejected
- A second service identifies the SPA packet in the IP stack and attempts to authenticate it
- If successful, the server creates an explicit policy to expose the service to the requesting endpoint

For example, the server may open a port in the firewall (e.g., iptables on Linux systems) for the client to establish a secure, encrypted connection with the service in question. The PEP provides the support to enforce the IAM policy of least privilege for the user identity requesting access by communicating with the PDP, preferably executing MFA—only then is a mutual transport layer security (mTLS) session is created for data transfer. Then an mTLS session is created for data transfer. The device is actively validated in context during this process. Frequent and periodic validation can be part of the IAM policy, which can be enforced either manually or by automation.

Another example of how ZTA increases the trustworthiness of access is described in NIST SP 800-207. The enhanced identity governance approach establishes enterprise resource access policies based on identity and assigned attributes. The main requirement for access is a given entity's access privileges (or lack thereof); in addition, the device used, asset status, and environmental factors also

come into play, as they will affect the ultimate level of access granted to the subject, regardless of its identity privileges.

The user authenticates to the device (e.g., with a username and password), which in turn authenticates to the network. The user authenticates to the network (e.g., using directory services) and their access request to the resource in question is sent to the gateway or portal. The request is forwarded to the policy administrator/policy engine. After authenticating with the identity provider, the result/decision is returned—if approved, access to the resource is granted to the user.

Consider the example of an IEEE 802.1x implementation using network access control (NAC) coupled with Lightweight Directory Access Protocol (LDAP): all corporate laptops have agents installed, and users authenticate to the laptop, which in turn authenticates to the network via IEEE 802.1x. User requests to access resources are vetted by NAC, LDAP, and potentially other access management applications. The request is authorized if the user is verified as part of the appropriate group.

4.3 Increased Visibility & Analytics

ZTA requires logging, monitoring, and alerting capabilities for increased visibility into users' activity: what actions they took, and when they took these actions. Attempts to access privileged resources as well as administrative or root account activity should always be logged, monitored, and reported. Anomaly detection should also be in place for detecting suspicious patterns in both inbound and outbound traffic.

Varying degrees of automation can be developed for these capabilities, as well as automated workflows for faster, more streamlined response and remediation. For example, alert notifications can be created when certain conditions are met, followed by automatic task assignment to the appropriate parties for further action.

To summarize, ZTA's visibility and analytics-related improvements include the following:

- Granular logging and monitoring for greater visibility across the enterprise
- Monitoring analytics over user entities behavior, leading to user entity behavior analytics
- Network isolation and micro-segmentation for improving the ability to quickly detect and resolve errors
- Continuous monitoring across all attack surfaces, making it easier to detect data breaches and enforce appropriate responses
- Minimization of data exfiltration
- Continuous device posture assessment

The specific visibility and analytics benefits will vary depending on the ZTA implementation. In the case of CSA's SDP, IAM policies are enforced when access requests are made to a device or host. Granular records of both successful and failed attempts of all components in the path provides increased visibility and the foundation for analytics. Device posture is evaluated during setup of the mTLS sessions. As logs become more granular and descriptive and user entity behavior analytics evolve, security analytics also become more detailed, making it easier to detect breaches or anomalous behavior. This also enables automation of appropriate responses.

Whereas, NIST SP 800-207 specifies that requirement (3) of ZTA is that it enables "the enterprise to observe all network traffic. The enterprise records packets (i.e., OSI layer 3) seen on the data plane, even if it is not able to perform application layer inspection (i.e., OSI layer 7) on all packets. The enterprise filters out metadata about the connection (e.g., destination, time, device identity) to dynamically update policies and inform the PE as it evaluates access requests."

The DOD ZT Reference Architecture outlines a model for logging, analytics, and automation.

1. Historical user behavioral data and current user actions are sent to the analytics engine to be analyzed.
2. A user's historical and current actions/behaviors are compared against global baselines or unusual activity indicators that house all acceptable trends. These baselines and unusual activity indicators can then be derived from internal analytics metrics or vendor-supplied feeds.
3. The analysis results in a confidence score based on the user's behavior.
4. As users traverse the network, their confidence score and historical behavior patterns dictate the level of access they receive.
5. Monitoring and analysis is continuously occurring in the background.
6. Access to a resource is denied if the users' actions and behavior patterns result in their scores dropping below a certain threshold.
7. If the users' actions and behavior patterns do not appear malicious, they can be informed that their scores do not meet the threshold.
8. If the users' actions and behavior appear malicious, different handling procedures are initiated depending on the specific actions/behaviors and accessed resources.
9. All actions are logged to a SIEM platform, processed by the analytics engine, and handed to a SOAR platform to deploy real time policy access decisions.

4.4 Improved Compliance

ZTA has the potential to improve an organization's compliance posture in several ways. For example, ZTA requires organizations to frequently review access policies to ensure they stay in alignment with requirements as their IT environment evolves. To this end, policies are a key element for security governance as they enable organizations to translate their goals and objectives into the rules that drive their approach to security. Polices also support the organization in remaining accountable to its shareholders and stakeholders. In a ZT approach, policies controlling access to resources are carefully enforced, continuously monitored, and frequently updated based on the current situation. These approaches enable organizations to maintain a strong compliance posture in regards to both external (i.e., legal regulations and oversight measures) and internal (i.e., company policy) requirements.

Continuous monitoring is critical for effective policy management, as it enables the alignment of policy definitions with enforcement measures. This is crucial for organizations looking to implement controls for continuous auditing and compliance.

Finally, micro-segmentation strategies apply access controls to each individual resource via fine-grained authorization mechanisms. The requester's trustworthiness is evaluated prior to access being granted. Policies actively determine access levels and may be based on the user's observable

state/identity, the requesting system, and other behavioral attributes. By implementing micro-segmentation and the principles of need to know and least privilege, organizations effectively reduce their attack surface/risk exposure, which may in turn limit their liability when it comes to laws and regulations. For example, a fewer number of users/devices with access to sensitive data and/or restricted by location reduces the scope of certain compliance measures (e.g., PCI-DSS or GDPR).

4.5 Additional Benefits

A ZT approach can help organizations identify business processes, data flows, users, data, and associated risks. These insights better equip them to reduce risk in their cloud and container deployments while also improving governance and compliance. Organizations can also gain deeper insights into users and devices, identify threats more quickly, and maintain more comprehensive control across a network. A well-architected ZTA also reduces IT complexity while supporting resiliency and defense-in-depth.

Security benefits aside, the advantages of a ZT security framework are numerous and vary depending on the enterprise's organizational landscape, architecture, and operating model. Utilizing cloud technologies to automate ZT functions helps minimize ongoing operational costs and eases the burden on human resources and staffing.

The ZT model provides a unified access control to data, services, applications, and infrastructure. This enables enterprises to counter major threats with one solution, versus a combination of tools (e.g., firewalls, VPNs, CASBs). By unifying the organization's access controls, ZT reduces security costs while improving efficacy, visibility, manageability, and user experience.

The following is a non-exhaustive list of additional ZT benefits:

- Potential cost reduction
- Simplification of IT management design
- Improved data protection (business critical data and customer data)
- Secure remote access
- Improved user experience

5 Planning Considerations for ZTA

In this unit, you will learn about the preliminary activities required to successfully implement ZTA in an organization, as well as some common tools and frameworks for planning.

As mentioned by leading technology vendors as well as public agencies like NIST, the implementation of a ZTA — and more generally the ZT approach and its design principles — is not a one-off task, but rather a process that depends on a number of different factors, including the following:

- The maturity level of the organization's security approach, especially regarding asset
- mapping and classification and identity and access management
- The existing organizational culture, skills, and expertise
- The amount of existing legacy technology and its criticality
- Existing investments

- Available budget
- The complexity of service architecture and data flows
- The end goal and objectives of the organization

Risk management forms the core of any competent cybersecurity approach; subsequently, ZT migration tactics are highly dependent on the risk profile and risk appetite of the organization in question. For some, the ZT design principle will be applied to a limited set of assets; others will apply ZT to all assets across the organization. In either case, the migration to ZT will follow a risk-based staged approach with numerous iterations culminating in the final transformation into a ZT-driven organization.

For example, CISA's *ZT Maturity Model* provides a reference roadmap that organizations can use for charting their transition towards a ZTA.

The CISA *ZT Maturity Model* consists of five pillars and three cross-functional capabilities that together form the crucial foundations for ZT. Each pillar outlines specific examples of traditional, advanced, and optimal ZTA.

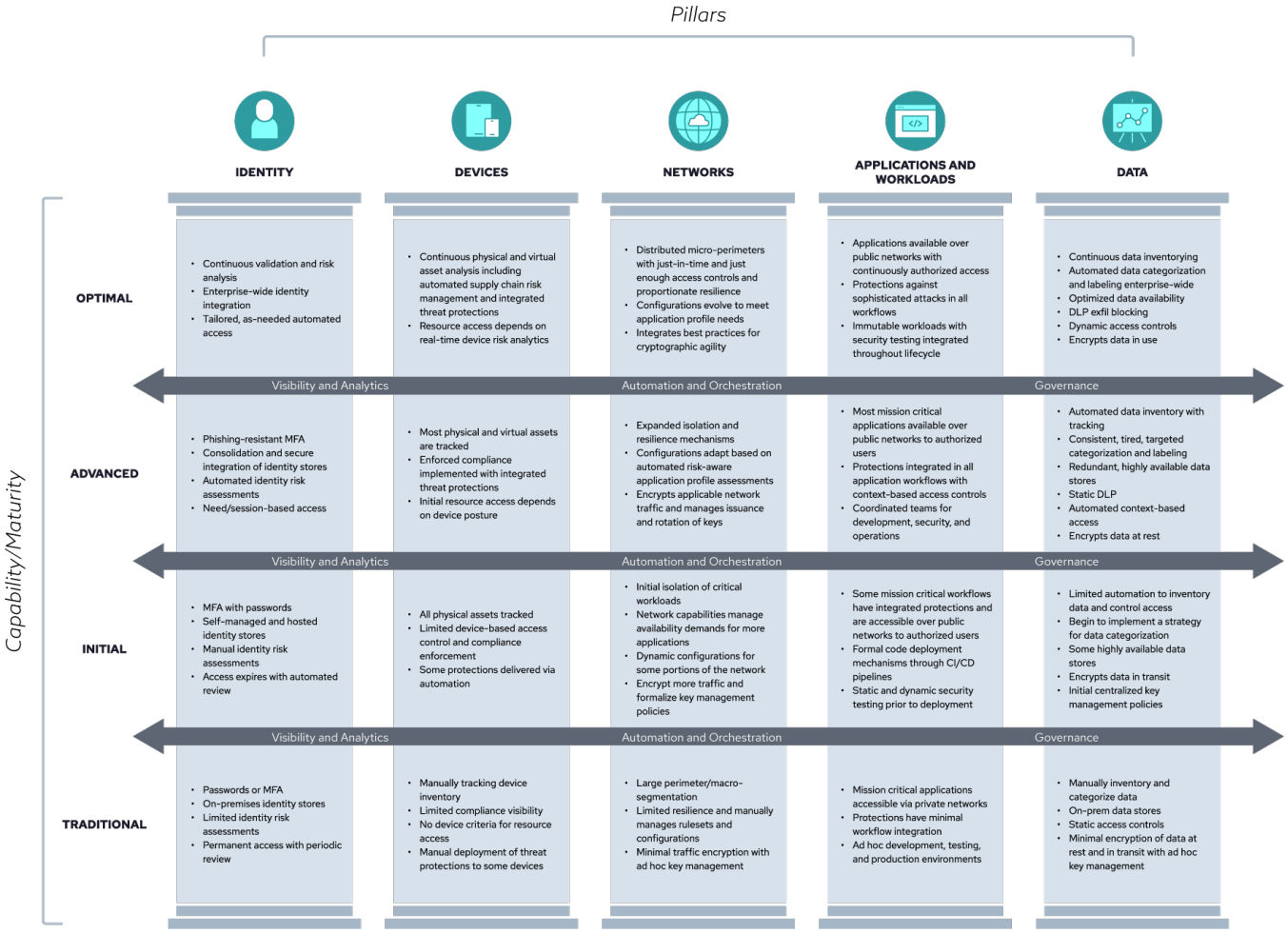


Figure 5 *CISA High-Level Zero Trust Maturity Model*[17]

[17] Figure adpated from CISA, "Zero Trust Maturity Model," June 2021, https://www.cisa.gov/sites/default/files/publications/CISA%20Zero%20Trust%20Maturity%20Model_Draft.pdf

5.1 Organizational & Technical Planning

This section describes the high level set of actions each organization is likely to follow when implementing a ZTA.

5.1.1 Understand Your Needs

The first step in the ZT implementation process is the analysis of the organization's needs at a high level. The ZT champion's role is to guide the organization's decision makers in answering the following questions:

- Why should the organization consider adopting ZT?
- What are the critical assets to be protected?
- What is the mission relevance and criticality of ZT to the organization?
- What are the opportunity costs of adopting versus not adopting ZT?
- Is the organization a cultural fit for ZT? What are the existing gaps, if any?
- How urgent is the ZT adoption and migration?
- What are the success metrics?

5.1.2 Identify Key Stakeholders

The identification of key stakeholders is another foundational step in ZT organization and planning. Like other enterprise-wide risk analysis processes, the organization must ensure that all key stakeholders are engaged and surveyed—this ensures that all the perspectives, requirements, pain points, and possible constraints are collected and considered. Additionally, a critical element in ensuring successful adoption of ZT is support from senior leadership in the organization. Without this, ZT adoption efforts are typically disconnected and uncoordinated; while pockets of success may be realized within the organization, a comprehensive and effective enterprise approach cannot be achieved.

The key stakeholders that should be involved include, but are not limited to:

- Business/service owners
- Application owners
- Infrastructure owners
- Service architecture owners
- CISO/security teams
- Legal officers
- Compliance officers
- Procurement officers
- Any other relevant management

5.1.3 Assemble a Team

Effective team collaboration across multiple groups is critical when assessing the application and server access landscape across the organization. Groups must have cross-team communications channels in place, as well as processes for collating their findings for future planning — this may span multiple phases, based on a formalized roadmap. A detailed explanation of the various technical planning aspects is covered in the following section.

5.1.4 Define Current State

At a high level, the organization needs to determine the level of maturity of its internal approaches and processes, specifically in regards to the following:

- Governance
- Risk management
- Compliance
- Asset management
- Identity and access management
- Cybersecurity

Are these processes and approaches already fully optimized and automated, or are they still ad-hoc and informal? The level of maturity will help create a realistic plan for initial adoption of ZT principles, and a roadmap of future incremental evolutionary steps.

The organization should analyze each one of the seven ZTA pillars identified earlier in this training, in respect to existing processes, procedures and technical solutions related to ZT. These include, but are not limited to the following:

- Asset/data inventory and classification
- Authentication and authorization (e.g., MFA, RBAC/ABAC, federated identity)
- Network segmentation (e.g., micro/nano segmentation)
- Encryption and key management (e.g., for data at rest/in transit, confidential computing)
- Secure software development lifecycle (SDLC) management;

- Continuous integration and continuous delivery (CI/CD)
- Monitoring and analytics
- Transaction flows

Organizations with a greenfield and/or cloud-native IT infrastructures have the opportunity to build ZT into the design of their IT and OT systems from the ground up.

5.1.5 Set Goals

The understanding of the organizational and technological status quo will facilitate the definition of realistic short and medium/long-terms goals.

Is it the final objective of the organization to create a complete transformation to ZTA, or to establish a hybrid of ZTA and legacy perimeter-based controls? What's the percentage of resources that will be affected by the ZT migration?

Once the medium/long term expectations have been set, the organization should answer the following questions:

- What are the priorities (e.g, what needs to be addressed immediately)?
- Are there any quick wins/low hanging fruit?
- What are prerequisites or upstream dependencies?
- Are the existing foundations to start from?

Additionally, the following questions are critical for addressing key factors during the goal setting process:

- What is the level of executive mandate?
- What is the strategy?
- What is the budget?
- What is the roadmap?

5.1.6 Define the Use Cases

This step is a critical process to understanding the organization's needs — specifically, in defining an organization's need for ZTA (i.e., its use cases and applications).

5.1.7 Develop Collaboration Plan

Effective team collaboration is crucial for a successful ZTA deployment. To this end, organizations should establish a unified collaboration plan shared among all team members and stakeholders; this can take the form of a Kanban board or software-based collaboration platform. All project communications regarding the ZTA deployment should be centralized on this platform.

Once a collaboration plan is in place, ZTA planning and deployment teams can move on to addressing the following crucial action items and concerns:

- Determine assets involved (e.g., data or services) and what needs protection– this can be determined through a risk analysis/assessment
- Identify principals in scope (e.g., humans, machines, and processes)
- Define IAM approach and methodology
- Determine processes in scope including both existing processes that need to change and new processes needed
- Select the service architecture
- Design the data and process flow
- Select the ZT implementation model and approach
- Define policies, both new and changes to existing policies
- Test/evaluate/select the technology or solution
- Implement/develop/deploy/deliver the selected approach/solution
- Monitor the ZT implementation for security and performance issues and plan for routine testing of ZTA security control
- Adapt/review/improve based on the results of monitoring and continuous testing, adapt/review/improve the ZTA implementation
- Extend the scope/reiterate the relevant steps of the process

5.2 Risks of Project Implementation

Any project that involves integrating new technologies or adopting novel approaches/methodologies bears some risk of failure; that said, the benefits of ZTA for improving the organization's security posture outweigh any perceived risks.

The following table covers some of the project risks that could arise while implementing a ZTA in an organization, as well as their impact and mitigation tactics.

Description	Implementation Risks	Impact	Mitigation
Failure of the ZTA operational elements such as PDP or PEP	Could hinder users and affected applications from authenticating/ operating properly.	Access to the secured assets could be compromised.	Deploying a high availability system and/or a failover mechanism.
New assessment and review criteria must be applied	Incorrect implementation and compromised operations.	As the new infrastructure solely depends on the architecture, an incorrect assessment of the solution may leave gaps.	A preplanned set of procedures and assessment steps created to validate the ZT implementation.

Security Operations	An interface between two systems in which (a) they are not connected physically and (b) any logical connection is not automated (i.e., data is transferred through the interface only manually, under human control).	Security level is reduced, leaving potential gaps in defenses. Responses to security incidents will use incorrect procedures.	Comprehensive analysis of sensitive data and acceptable routes should be performed early in ZTA's design stages.
Remote API calls	Lack of API protocol support, API request inspection, data leakage monitoring, and API discovery (e.g., for shadow or zombie APIs).	Complexity in parsing API requests and the existence of deprecated versions.	Implement support for all relevant parsers. Provide the right controls to protect sensitive data like PII.
Hybrid implementation complexity resulting in environments that require additional effort/resources to operate, maintain, and support	Unforeseen resource misallocations that could significantly increase implementation costs and deadlines.	ZTA adoption and implementation will likely co-exist with legacy or non-ZTA environments, so operations/technology/ infrastructure must support hybrid architectures.	Alerts for the same network event may be handled differently by an enterprise SIEM per environment.
ZTA integration with existing network and security infrastructure and operations can be challenging	Incompatibility with the legacy systems must be addressed before implementing the ZTA.	Interoperability with the legacy systems is paramount whilst implementing the ZTA.	ZTA integration can be carried out in incremental phases with validation processes and backout contingencies.

Fielding of partial or incomplete ZTA solutions	Fielding without adopting capabilities through the organizational maturity levels may create vulnerabilities that ZTA was intended to mitigate.	Vulnerabilities present within the ZTA will be targeted by adversaries, potentially resulting in technical and/or reputational exposures to the organization.	Validate that the ZTA adoption strategy is properly conceived to ensure that the intent to execute ZTA adoption through the organizational maturity levels is captured. Additionally, confirm that organizational leadership understands that the initial implementation will not be the final end state and will require continuous, iterative development through the maturity model.
Fielding of ZTA solutions without proper operational sustainment/ maintenance planning	Inconsistent enterprise baselines of fielded technologies, solutions/resources that are deteriorated or expended without effective results.	These risks expose the organization to adversarial threats, resulting in elevated technical and reputational risk to the organization.	Ensure that the ZTA adoption strategy properly covers both the initial deployment as well as long term costs and organizational restructuring necessary to support/ maintain ZTA on a long term basis.

Figure 6 *ZTA Project Implementation Risks*

6 Implementation Options of ZTA

In this unit, you will learn about the various ZTA implementation approaches defined by NIST SP 800-207, as well as some real-world ZTA implementation methods and their main characteristics. The options presented in this unit focus on the network architecture domain and align with the NIST approaches "ZTA Using Micro-Segmentation" and "ZTA Using Network Infrastructure and Software-Defined Perimeters". The primary ZTA implementation options covered in this unit are CSA's SDP, Zero Trust Network Access (ZTNA), and Google BeyondCorp.

6.1 NIST Approach to ZT

Organizations looking to adopt NIST's ZT model have several approaches at their disposal for designing their secure workflows. Each approach implements all of the ZT tenets outlined in Section 2.1 of NIST SP 800-207, and a fully-realized ZT solution will incorporate elements from all of the three NIST ZTA approaches:

> - ZTA using Enhanced Identity Governance
> - ZTA using Micro-Segmentation
> - ZTA using Network Infrastructure and Software Defined Perimeters

Depending on factors such as the organization's existing business flows, requirements, and cybersecurity maturity level, a particular approach may be more suitable for a given environment—in turn, the components used and main sources for policy rules will also vary accordingly.

As mentioned previously, this unit focuses on the NIST approaches for "ZTA Using Micro-Segmentation" and "ZTA Using Network Infrastructure and Software-Defined Perimeters''. Subsequent ZT training courses in this series provide a more comprehensive and expanded overview of NIST's approach to ZT.

6.2 Software-Defined Perimeter

CSA's SDP concept is an approach to enabling and enforcing ZT principles. The SDP architecture is designed to provide on demand, dynamically provisioned air-gapped networks: trusted networks that are isolated from all unsecured networks to mitigate network-based attacks.

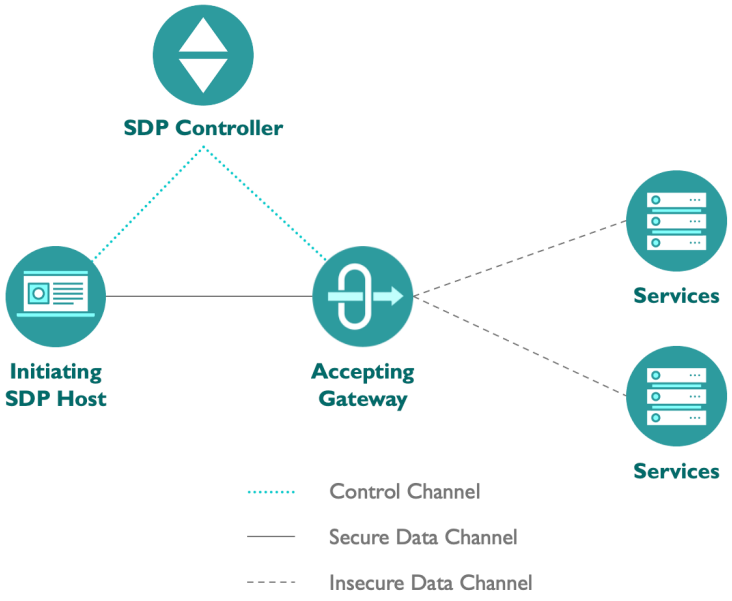


Figure 7: *SDP Pre-Vetting of Connections*[18]

6.2.1 Description

ZT implementations require the verification of anything and everything attempting to access assets, prior to authorization. Additionally, ZT requires continued evaluation of sessions and their risk levels during the entire connection's duration. As described in CSA's *Software-Defined Perimeter (SDP) and Zero Trust*, "a ZT implementation using SDP enables organizations to defend new variations of old attack methods that are constantly surfacing in existing network and infrastructure perimeter-centric networking models. Implementing SDP improves the security posture of businesses that face the challenge of continuously adapting to expanding attack surfaces that are increasingly more complex[19]." The enterprise must monitor the integrity and security posture of the assets. SDP enforces this trust strategy by enabling a default drop-all gateway until users/devices are authenticated and authorized to access the assets hidden by the gateway. By requiring the pre-vetting of connections, SDP enables complete control over who can connect, from which devices to what services, infrastructure, and other conditions and parameters.

As described in the *SDP Architecture Guide v2*, SDP consists of the following major components:

- The client/initiating host (IH)
- The service/accepting host (AH) — also referred to as the PEP per NIST's ZTA model
- An SDP controller to which the AH and IH both connect — also referred to as the PDP per NIST's ZTA model
- An SDP gateway that implements the drop-all firewall

[18] Figure adapted from Cloud Security Allaince, "Software-Defined Perimeter (SDP) Specification v2," 10th, March, 2022, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/

[19] Cloud Security Alliance, "Software-Defined Perimeter (SDP) and Zero Trust," 27th, May 2020, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-and-zero-trust/

According to the *SDP Architecture Guide v2*, SDP works in the following manner:

- The **SDP client** software on the IH opens a connection to the SDP. **IH** devices (e.g., laptops, tablets and smartphones) are user-facing, meaning the SDP client software is run on the devices themselves. The network can be outside the control of the enterprise operating the SDP.
- **AH** devices accept connections from IH and provide a set of SDP-protecting/secured services. AH typically reside on a network under the enterprise's control (and/or under the control of a direct representative).
- An **SDP gateway** provides authorized users and devices with access to protected processes and services. The gateway can also enact monitoring, logging, and reporting on these connections.

IH and AH host devices connect to an **SDP controller**: a device/appliance or process that secures access to isolated services by ensuring the following:

1. Users are authenticated and authorized
2. Devices are validated
3. Secure communications are established
4. User and management traffic remain separate on the network

The controller and AH are protected by SPA, making them invisible and inaccessible to unauthorized users and devices.[20] Six deployment options are available for implementing SDP:

- Client-to-Gateway
- Client-to-Server
- Server-to-Server
- Client-to-Server-to-Client
- Client-to-Gateway-to-Client
- Gateway-to-Gateway

6.2.2 Compliance with ZT Principles

The SDP conforms to the following ZTA principles:

1. The IH and users should first be authenticated and authorized by the controller before connecting to the AH. The AH is cloaked from the IH and its users until authentication is completed.
2. The SDP gateway applies the drop-all policy until the SPA from the IH is verified. The cryptographic mechanism behind the SPA ensures that only authorized devices can communicate with the AH's controller.
3. Every service and AH is protected with its own SDP gateway drop-all policy; communications from the other server should also follow the same access policies. IH and users can therefore only access resources to which they were explicitly granted permissions, ensuring

[20] Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019,: https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2

adherence to the principle of least privilege.
4. The SDP controller and SDP gateway are the chokepoints for all access attempts and communications. Subsequently, they can provide continuous monitoring, logging and reporting of all network communications, to include both legitimate and suspicious access attempts.

6.2.3 Implementation Options

Several options are available for implementing a SDP: controllers may reside on-prem or in the public cloud, the gateway can be deployed on the servers (i.e., the AH) or an external node, and the SDP can be configured to protect a single service or multiple services.

The following are some critical best practices for implementing SDP:

- Because they are single points of failure, controllers should be designed for high availability (HA) in order to withstand DoS/DDoS attacks and other similar malicious activity. HA strategies such as the use of multiple physical server instances with load balancing (e.g., domain name system load balancing) should be considered.
- Gateways can block a service in the event of a case of failure or overload. Different load-balancing schemas can be used (e.g., the controller can act as a load balancer for gateways). Gateways are stateful SDP entities that can maintain mTLS sessions, so switching over to a different gateway may interrupt sessions across the tunnel.
- SDP controllers may use an internal user-to-service mapping or a connection to a third party service (e.g., LDAP, directory service, or other on-premises/cloud-based authorization solution). Authorization is typically based on user roles and more fine-grained information, user or device attributes, or even the specific data element/data flow the user is authorized to access. In effect, the access policies maintained by the SDP controller can be informed by other organizational constructs such as enterprise service directories and identity stores. Per NIST, the dynamic ZT policies enforced by the controller are categorized as a ZT tenet.

6.2.3.1 Service Initiated (Cloud-to-Cloud)

An increasingly common use case for deploying a ZTA entails the use of multiple cloud providers. In this scenario, the enterprise manages a local network but uses two or more cloud service providers to host applications/services and data; occasionally, the application/service is hosted on a cloud service separate from the data source.

As depicted below, the application hosted in Cloud Provider A should directly connect to the data source hosted in Cloud Provider B. This enables better performance and ease of management, as the application isn't forced to tunnel back through the enterprise network.

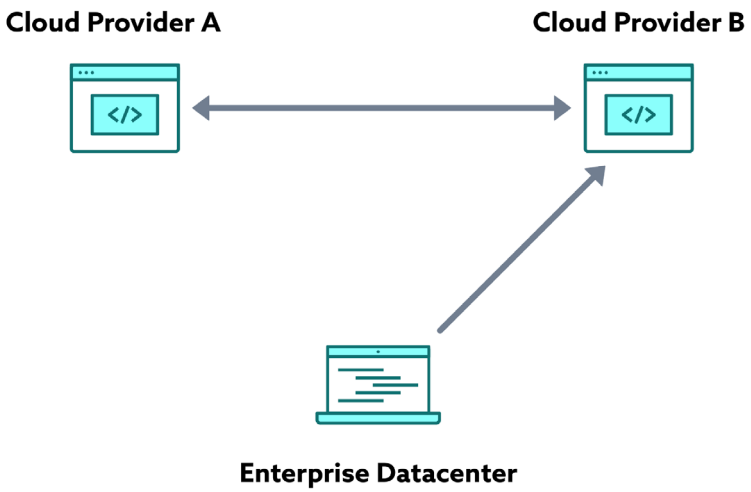


Figure 8: *Cloud-to-Cloud ZTA Service Initiation*[21]

This use case is the server-to-server implementation of the CSA *SDP Specification v2*. A more common example is Cloud Provider A cloud calling Cloud Provider B's LDAP service for authorization/authentication, as part of SSO.

ZTA services are often set up in a mesh configuration. Meshed services lend themselves well to a multi-cloud environment since they facilitate service-to-service communication (to include micro-services communication) via a proxy.

6.2.3.2 Collaboration Across Boundaries

Cross-enterprise collaboration is another prominent ZTA use case. For example, a hypothetical project may involve employees from Enterprise A and Enterprise B. Enterprise A manages the project database but must allow certain members of Enterprise B to access the data.

To meet this requirement, Enterprise A can set up specialized accounts for Enterprise B employees to access the required data, denying access to all other resources; however, this approach can quickly become difficult to manage. Enrolling both organizations in a federated ID management system streamlines the configuration of these permissions, provided both organizations' PEPs can authenticate subjects in a federated ID community.

6.2.4 Characteristics

SDP's main advantages are its maturity and widespread adoption. Early on, prominent enterprises and leading institutions such as the DOD were supporters/adopters; today, organizations across all industries are implementing different flavors of SDP for varying purposes and environments, to include hybrid and multi-cloud deployments, VPN replacement, and securing IoT. Additionally, regular hackathons that test SDP's attack durability continue to add to its popularity.

[21] Figure adapted from NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://csrc.nist.gov/publications/detail/sp/800-207/final

SPA and mTLS are highly effective mechanisms for enforcing ZT principles without sacrificing user experience. SDP is in fact capable of providing robust security while simultaneously improving the user experience—especially when replacing legacy solutions. SDP is also relatively easy to implement and can complement existing solutions in place. Organizations are free to adopt a gradual implementation and/or migration to an SDP.

Because SDP is completely distributed and scalable, it can easily protect highly complex deployments (e.g., hybrid and multi-cloud environments). High availability is also built-in to SDP's architecture.

A major disadvantage of SDP is the requirement for client agent installation on each endpoint that connects to the SDP-protected deployment. Additionally, SDP primarily supports traditional user access methods to enterprise resources; API-based, micro-service, and serverless access methods are not well-supported by SDP.

6.3 Zero Trust Network Access

ZTNA is quite similar in spirit and form to CSA's SDP and Google's BeyondCorp, and all three have influenced each other. Though SDP is distinguished by its use of SPA, ZTNA's premise is nonetheless quite similar to SDP. In fact, some literature have ZTNA deriving its origins from SDP.

6.3.1 Description

ZTNA supports a paradigm where neither users nor the applications they access are sitting behind the perimeter. Often considered a VPN replacement, ZTNA allows users to access services from anywhere, anytime, from any device. ZTNA consists of two distinct architectures: endpoint-initiated ZTNA and service-initiated ZTNA.

Endpoint-initiated ZTNA is fairly similar to the original SDP specification. A lightweight agent is installed on the end-user's device and communicates with a controller, which in turn authenticates the user and provisions the necessary connections to the authorized applications. Because of this agent installation requirement, endpoint-initiated ZTNA is difficult to implement on unmanaged devices.

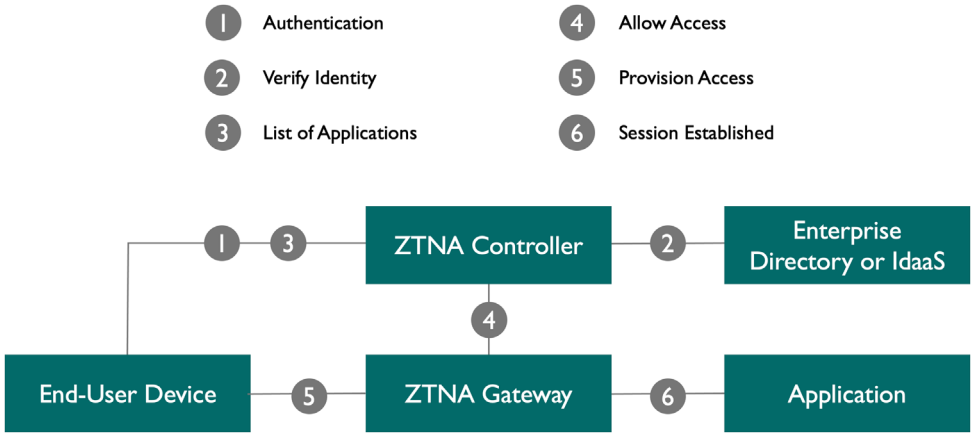


Figure 9: *Endpoint-Initiated ZTNA Communication Flow*[22]

[22] Figure adapted from Gartner, "Market Guide for Zero Trust Network Access (ZTNA)," 8th June, 2020, https://www.gartner.com/en/documents/3986053

On the other hand, service-initiated ZTNA uses a broker between the user and the application. In this case, a lightweight ZTNA connector sits in front of the service, which itself can be located in the data center or in the cloud. The connector establishes an outbound connection from the service to the ZTNA service broker. Upon authentication, traffic passes through the ZTNA broker, isolating services from direct access to unauthenticated users, effectively hiding them and preventing malicious activity like DDoS-type attacks. The service-initiated ZTNA option is suitable for unmanaged devices (e.g., bring your own device [BYOD]) or partner access.

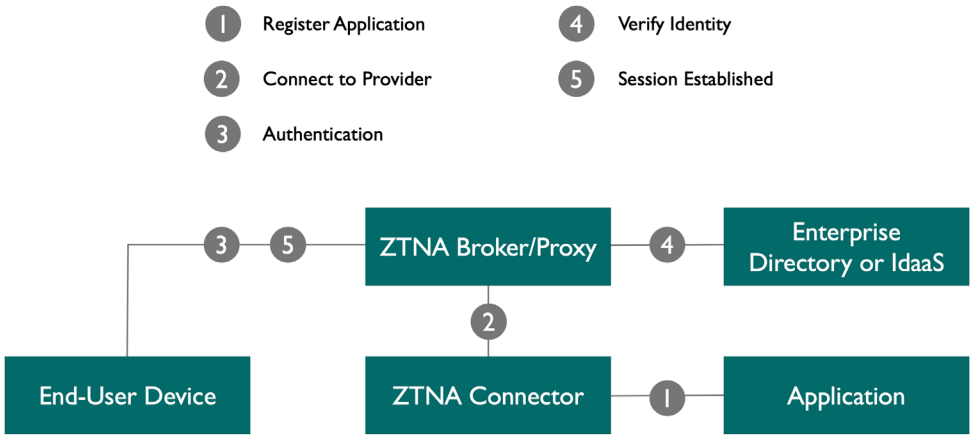


Figure 10: *Service-Initiated ZTNA Communication Flow*[22]

6.3.2 Compliance with ZT Principles

- ZTNA assumes a hostile user access environment. In fact it can operate from unmanaged devices and makes no assumptions about it being pristine.
- ZTNA assumes a breach. The user equipment is unmanaged and can be breached. The authentication and authorization is for a single session between the user and the services.
- Every access to the service is verified in the spirit of "never trust, always verify".
- ZTNA reduces the attack surface by hiding services behind brokers.
- Only authenticated users are allowed access if there is an explicit policy for them to have access.

6.3.3 Implementation Options

ZTNA can be used as a stand-alone product or as a service. In stand-alone mode, the broker runs on the customer's premises, and they are responsible for the deployment and management. Several IaaS cloud providers also offer managed ZTNA services for their customers.

6.3.4 Advantages

ZTNA offers benefits in user experience, agility, adaptability, and simplified policy management. When ZTNA is cloud-based, it has added benefits of scalability and ease of adoption. It is a much favored alternative to traditional VPNs where there is unhindered access once the VPN tunnel is established.

6.3.5 Disadvantages

Endpoint-initiated ZTNA is difficult when the user device is unmanaged (e.g., BYOD). ZTNA cannot guard against malicious actors that are already inside and co-resident with the service. It can only help in cases where the actor is outside of the perimeter where the services are hosted.
Secure access service edge (SASE) is a more recent technology that provides continuous inspection beyond the initial connection authorization and establishment. There is no provision in ZTNA of session revocation based on continuous inspection post establishment.

Policy management (e.g., authorization) is orders of magnitude more complex for programmatic access. The authentication, scale, as well as latency requirements vary significantly. For this reason, ZTNA is mostly applicable for user access and for VPN replacement use cases.

6.4 Google BeyondCorp

As described in the *SDP Architecture Guide v2*: "BeyondCorp is Google's internal network and access security platform, designed to enable their employees access to internal resources." Today, BeyondCorp Enterprise is available to organizations with Google-based IT infrastructures.

6.4.1 Description

The primary component of BeyondCorp is the web proxy: the chokepoint every user/device needs to traverse in order to access the organization's resources.

Some notable features of BeyondCorp include the following:

- Any access to protected resources are done via proxy
- Device and user identities are checked using a device inventory and user/group database
- 802.1x protocol is used to verify the managed devices and provide micro-segmentation
- An access control engine provides authorization for the organization's applications and services
- A data pipeline with additional information such as location, device/user trust levels, and more feeds into the access control engine

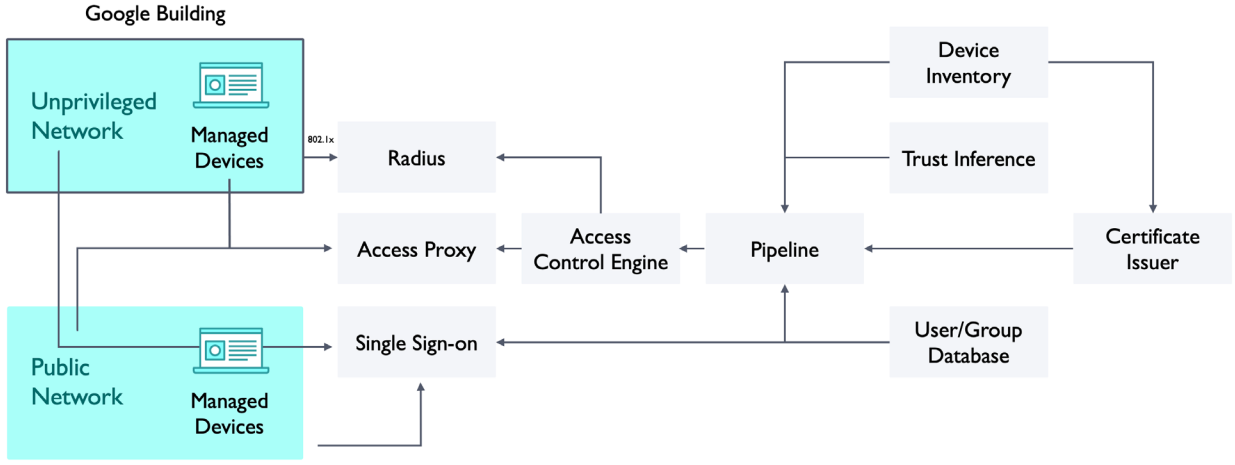


Figure 11: *BeyondCorp Components and Access Flow*[23]

[23] Figure adpated from Google, "BeyondCorp", https://www.beyondcorp.com/

6.4.2 Compliance with ZT Principles

BeyondCorp incorporates ZT principles as follows:

- The device/user should first be authenticated and authorized by the access proxy, prior to establishing a connection to the enterprise application—regardless of whether the device/user is located on the internal or external network.
- The access proxy denies any access request from unauthenticated users or devices.
- Each access request is handled separately by the access proxy, in line with the principle of least privilege.
- The access proxy is the choke point of all access attempts and communication; it should therefore be continuously monitored, with all network communications logged and report—to include both legitimate and illegitimate access attempts.

6.4.3 Implementation Options

As Google's proprietary implementation of ZTA, BeyondCorp offers limited implementation options. Some organizations implement a simplified version of BeyondCorp that only uses an access proxy, leaving out additional components like a device inventory and trust engine.

6.4.3.1 Service Initiated (Remote Application Access)

This implementation approach is in line with BeyondCorp model: a connector is deployed on the same network as the shared applications. Once the connector establishes and maintains a continuous outbound session to the provider's environment, users/devices can authenticate with the provider to access protected applications.

The provider can force the user through an authentication workflow before access is granted. This avoids direct access to the application, as the user/device is allowed to connect to the application server only after the authentication process (e.g., client-initiated ZTNA) is complete. Additionally, this model is agentless (i.e., agent software is not required on the connecting device), with application access over Hypertext Transfer Protocol/Hypertext Transfer Protocol Secure (HTTP/HTTPS)—at layer 7 of the OSI model.

6.4.4 Advantages

BeyondCorp doesn't require client agent installation on connecting devices, though devices should be registered in the device Inventory database and assigned a unique certificate.

6.4.5 Disadvantages

A fully-realized BeyondCorp implementation is less flexible and difficult to integrate with existing security mechanisms such as IAM. Additionally, BeyondCorp's lack of strong cryptographic controls such as SPA and mTLS makes it less secure than SDP, as these controls are required for implementing an *invisible cloud*. Unlike the SDP controller, BeyondCorp's access proxy is an in-line entity that handles both control and data traffic, making for a less scalable/secure model.

7 ZT Use Cases

In this unit, you will learn about various ZT use cases and how they vary — both architecturally as well as in terms of risk mitigation efficacy and limitations/dependencies.

A myriad of ZT use cases can be found across numerous industries. This section provides a non-exhaustive list of example applications. Each use case is broken out by the following:

- Use case description
- Security risks
- ZT mitigation of risks
- Limitations and dependencies

7.1 Remote Access & VPN Replacement

7.1.1 Use Case Description

Enterprises have historically provided employees secure remote access to the corporate network via VPN (i.e., an encrypted tunnel). With the widespread adoption of cloud services, employees now require additional remote access to services residing in one or more clouds and associated environments (e.g., virtual private clouds [VPCs] or virtual networks [VNets]). In the past, secure remote access was limited to applications hosted within the corporate data center. Today, organizations must also provide employees access to applications and services no longer hosted within their corporate data centers.

Traditional VPNs terminate at the organization's perimeter, enabling remote users to access the organization's resources, wherever they are located. The migration of IT resources to the cloud has led to the substantial performance degradation of VPNs. To address this issue and enable optimal access to remote services, organizations are creating encrypted tunnels to external enclaves using new technologies such as cloud proxies and SASE. This allows employees, contractors, and partners to securely access both internal services/applications as well as external IaaS/PaaS/SaaS offerings from other cloud service providers. ZTA bolsters the security posture of remote access processes by including SDP capabilities—namely SPA—in the communications between remote devices/users and external enclaves.

7.1.2 Security Risks

In most VPN solutions, users are allowed into the organizational network via a VPN gateway; once authenticated and granted access, the user has access to enterprise assets. Care must be taken to avoid violating the principle of least privilege. In addition, device authentication should be exercised prior to access, validating that the device has no malicious software or malware. For example, if a remote employee's device is infected with malware, that malware may impact all organizational assets accessible by this user once entering the network.

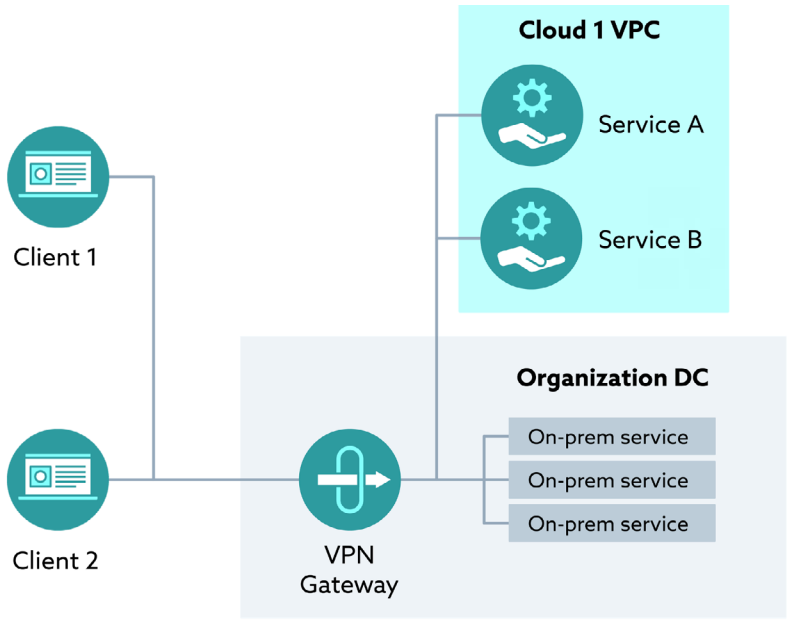


Figure 12: *Traditional VPN Gateway*

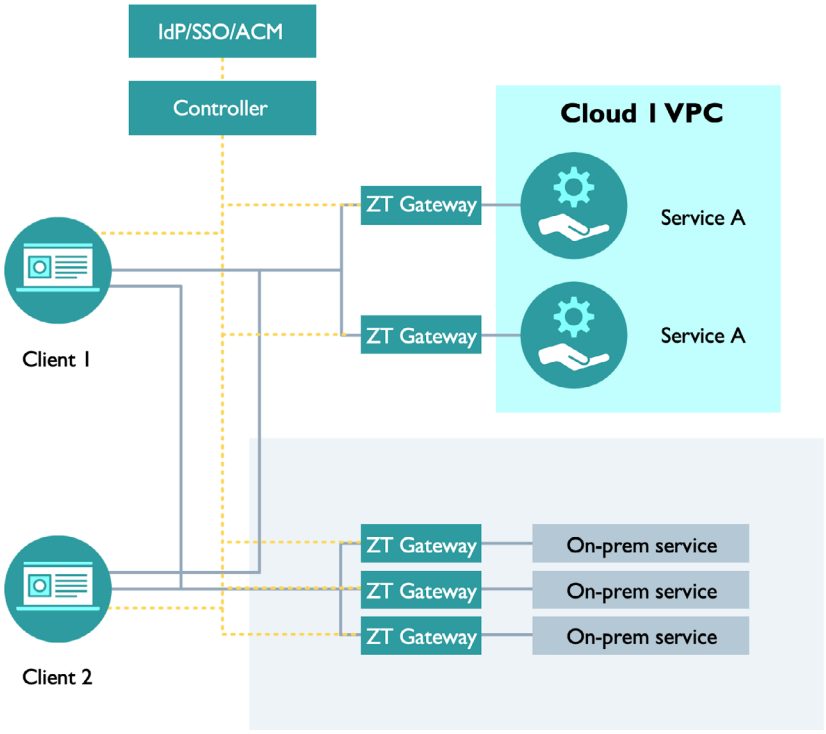


Figure 13: *Protection of Services by ZTA Gateway*

7.1.3 ZT Mitigation of Risks

ZTA effectively avoids or covers many of VPN's inherent security gaps through more granular, contextual security controls. For example, traditional VPN implementations have all user traffic going through a central VPN gateway before reaching a cloud application, creating both high latency as well as a single point of failure/compromise. Additionally, the same policies and security controls are applied to all users regardless of the application and user location.

In contrast, ZTA has each service separately protected by a ZT gateway; each client first connects to the controller, and only after authentication and authorization can they connect to the application over mTLS, via the gateway. Different policies and security controls can also be applied per application.

7.1.3.1 User Experience Impact

With VPN, users — especially mobile users — frequently experience delays, disconnections and connectivity problems. User connectivity to the internet is impacted as well, even if split tunneling is used. Split tunneling is a VPN feature that divides internet traffic and sends some of it through an encrypted VPN tunnel, routing the rest through a separate tunnel on the open network.

7.1.4 Limitations & Dependencies

A ZT environment is flexible and adaptable to change, as the ZT model is based on proven standards including mTLS, SAML, and X.509 certificates, among others. It can be combined with supplemental security systems such as data encryption and remote attestation systems due to its extensible nature. Coupling the evolved encrypted tunnel with the ZTA provides a path for evolution.

7.2 Micro-Segmentation

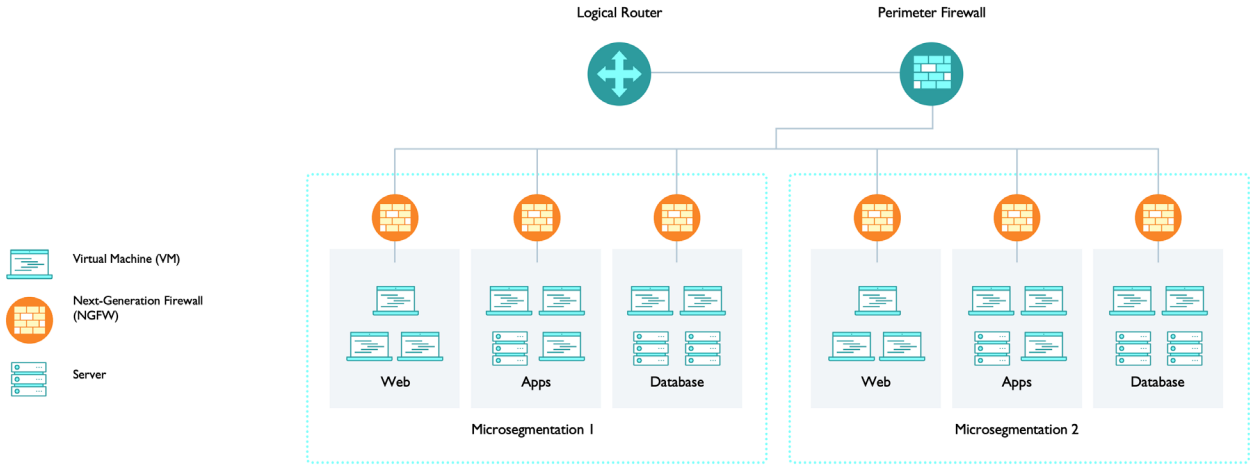


Figure 14: *Micro-Segmentation*[24]

7.2.1 Use Case Description

ZT enforces the separation of connections between the devices on a network. By requiring more granular, policy-based access for device-to-device connections, organizations can prevent the traffic from being visible — even to internal users. This is accomplished by creating dynamic, trusted network zones around applications, effectively hiding them from unauthorized users and devices. On a typical micro-segmented network, each of the connections between servers or devices on a network will be directed through separate layers of authentication and data traffic. Every device must

[24] Figure adapted from TechTarget, "What is Zero Trust? The Ultimate Guide to the Network Security Model," November 2020, https://www.techtarget.com/searchsecurity/definition/zero-trust-model-zero-trust-network

initiate its own encrypted tunnel in order to communicate with servers. Thus, each connection is a separate network impenetrable by other hosts.

Micro-segmentation helps to ensure that device access is limited only to validated, authorized entities, and is highly effective in preventing the spread of a cyber attack across an environment, limiting the impact to the compromised device in question.

Micro-segmentation architectures can be deployed in both cloud environments and on-premise data centers.

7.2.1.1 Types of Micro-Segmentation

Time-based segmentation policies and controls typically become more granular over time. In fact, a direct correlation exists between the granularity of controls and the age of a system's technology stack. Some of the more common micro-segmentation architectural patterns that emerge include the following:

- Traditional network segmentation
- Data center (i.e., east-west) segmentation
- Application micro-segmentation
- Workload micro-segmentation

7.2.2 Security Risks

Once cyber attackers gain a foothold into the network, they typically move laterally in attempts to compromise other machines on the network. Network visibility is usually not restricted to privileged users/devices in VPN and corporate IT environments. The devices themselves are also prone to attacks, since some of these IT assets may be visible from the internet.

7.2.3 ZT Mitigation of Risks

ZT implementations do not implicitly trust any of the devices or applications on the network. Only trusted devices can initiate a connection following a SPA-based request, and later via an encrypted tunnel. The security posture of the IT environment is enhanced by the fact that the devices are completely hidden from unknown users, with users controlled/contained within a tunnel between devices.

7.2.4 Limitations & Dependencies

Because stringent control is maintained over users and devices and their respective access to each application or resource, the architecture and interactions between the devices require careful integration to reduce user/device validation-related latency. Also, the data flowing between devices are not verified/validated, though the connecting device's security posture and identity is verified/validated prior to the connection being granted.

7.3 Software as a Service & ZT

7.3.1 Use Case Description

The rise of the cloud and SaaS deployment models has given organizations access to an unprecedented array of scalable IT resources never before possible. This can fuel innovation and boost productivity, but it also introduces new IT security challenges beyond the traditional corporate firewall. Each SaaS solution in use introduces numerous challenges related to vendor risk management, data protection, access controls, user experience, auditing, monitoring, privileged access management, and more.

7.3.2 Security Risks

In the SaaS shared responsibility model, areas exist where visibility, governance, and control are reduced, leading to varied security risks; SaaS solutions therefore need to be understood, monitored, and reported for risk acceptance. For example, data protection compliance measures apply to SaaS providers, making risk acceptance critical to the implementation process, unless additional controls are added for risk mitigation. Business functions are choosing to procure and use SaaS applications without the knowledge or permission of IT. This phenomenon, also referred to as shadow (or stealth) IT, significantly increases the risk of data breaches and security incidents. Corporate IT should therefore specify in their service level agreements/contracts the requirement for controls with conformance reporting standards.

Due to the rise of the mobile workforce and the proliferation of cloud applications, network-centric security architectures are no longer considered adequate protection. Once a security perimeter is breached using various exploits and attack methods (e.g., phishing, malware, or compromised passwords), threat actors can move freely across other security layers and systems in search of vulnerable data.

Microservices and third-party APIs have also gained widespread adoption in the last decade, enabling SaaS offerings to be integrated with existing systems through publicly supported APIs. Organizations can simply subscribe to these services instead of building them from the ground up; however, this introduces supply chain risk into the ecosystem.

7.3.3 ZT Mitigation of Risks

Adopting the ZT SaaS management model is an effective approach to mitigating cyber risks inherent in SaaS services. This includes the enforcement of policy-based access control in SaaS applications, regardless of the user/device location, as well as the monitoring of all SaaS usage patterns.

In many cases, organizations bolster the security of their SaaS applications with single sign-on security (e.g., SAML) and IP-based access control with a CASB, which may negatively impact the user experience with increased latency and degraded performance. The ZT model adds stronger security to SaaS applications without impacting the user's experience.

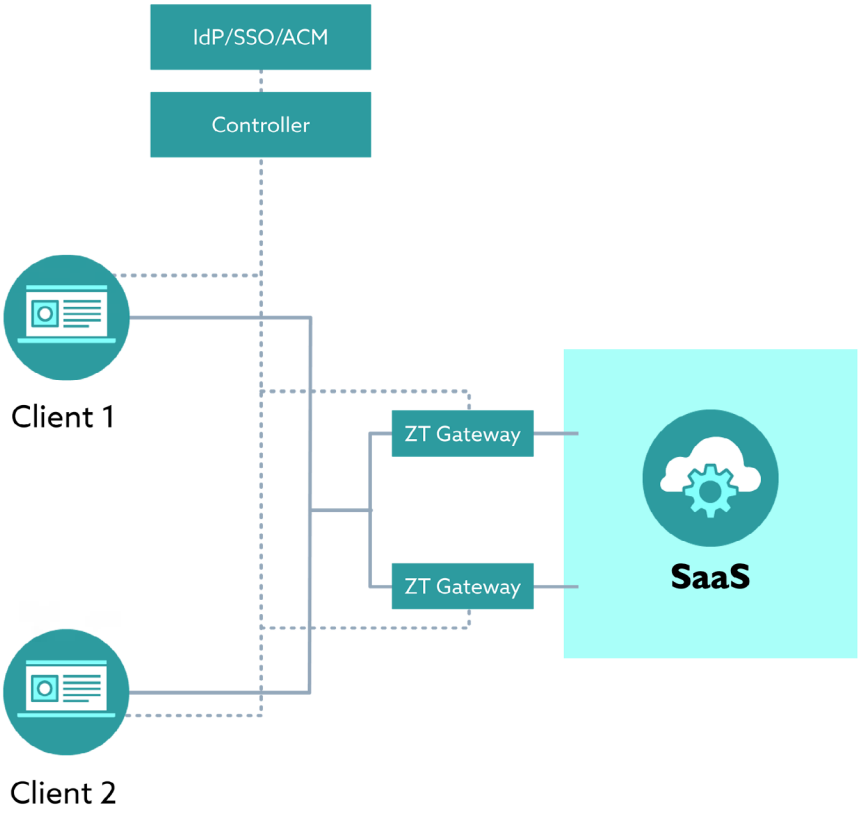


Figure 15: *ZT Model for SaaS Management*

7.3.4 Limitations & Dependencies

ZT SaaS control depends on a SaaS mechanism to control corporate account access. This includes the support of client SSO access lists for a SaaS service — effectively disabling direct access to SaaS services (i.e., bypassing the SSO access mechanism). ZT and SDP are limited in their ability to control the data flow inside a SaaS instance or between different SaaS applications.

7.4 Hybrid, Multi-Cloud, & ZT

7.4.1 Use Case Description

Hybrid clouds combine on-premises solutions or private cloud(s) with one or more public cloud services, with connectivity between each distinct service enabled through technologies like site-to-site VPN and private or dedicated circuits. Many organizations also adopt a multi-cloud strategy in order to leverage several cloud service providers; to this end, organizations can use public, hybrid, or private clouds as part of their overall cloud adoption strategy.

7.4.2 Security Risks

Using multi-cloud, hybrid cloud, or a combination thereof expands the organization's attack surface. Different public cloud providers use varying IAM models, security controls, and connectivity methods between VPCs or between VPCs and private clouds.

The broad level of network access inherent with hybrid and multi-cloud deployments conflicts with ZT's least privilege access model. For example, cloud providers may default to the most open access levels to maintain interoperability—in the case of a site-to-site VPN, the connection between a private and public cloud must be configured for any network access in order for devices on either end to communicate freely.

7.4.3 ZT Mitigation of Risks

If applied across all of an organization's cloud deployments, ZT can mitigate the security risks inherent in publicly exposed cloud services. The following are the guiding principles for accessing an organization's resources across different cloud providers and private clouds:

1. A device/users connection point on a particular network should not determine which cloud services are accessible.
2. Users should be identified, authenticated, and authorized prior to connecting initially, as well as before any subsequent connections to cloud resources.
3. Access to services and resources is granted based on what the organization knows about the user/device, regardless of which cloud service they are connecting to.
4. The same security controls (e.g., tunneling and encryption) are applied to both private and public clouds.

ZTA fulfills these requirements by hiding all the services and resources, regardless of their location. Users in turn have no access to those resources prior to completing authentication and authorization.

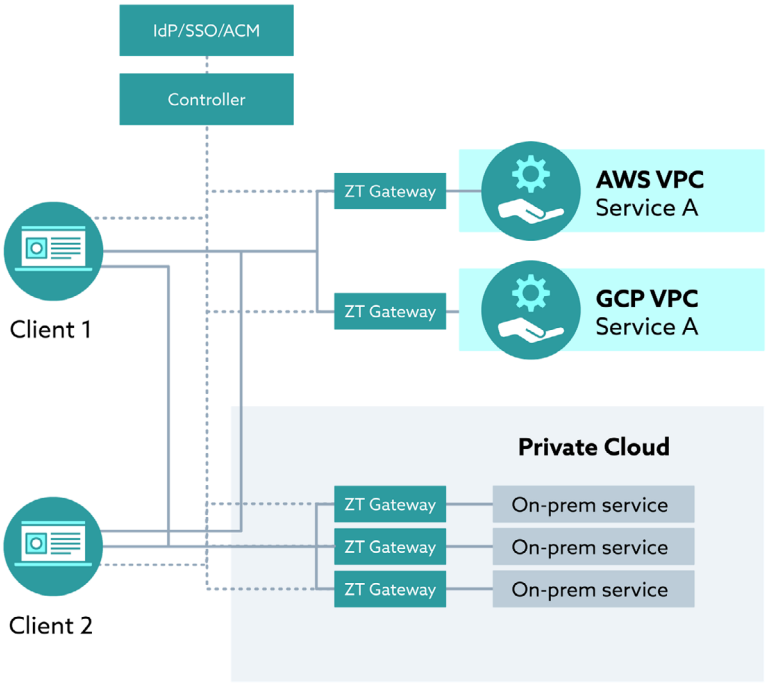


Figure 16: *ZTA Model for VPC and Private Cloud Deployments*

ZTA enforces the use of a mutually encrypted tunnel between the user device and the PEP, per individual service. The least privilege access model is enforced because the access policies are granular and resource/service based, versus network, cloud, or VPC-based.

7.4.4 Limitations & Dependencies

ZT improves the user experience with its distributed architecture, eliminating single choke points that may impose delays and result in single point failures. However, a truly cloud and vendor agnostic implementation of ZT may be difficult to implement due to the varying design patterns of competing cloud providers. For example, the implementation of SSO with Azure AD differs from Azure cloud; similarly, Google Cloud Platform (GCP) differs from an OpenStack-based private cloud.

Lastly, the interconnections between multi-cloud deployments and hybrid-to-public clouds are vendor-dependent. Best practices can be followed, but there isn't one standard protocol or implementation, hence it is not easily designed and implemented.

7.5 Operational Technology

OT primarily exists in industrial environments where processes are carefully regulated and managed to achieve a desired outcome. The systems associated with the OT environment are industrial control systems (ICS) and IIoT devices.

Traditionally, the OT environment was made up of closed, physically air-gapped networks and systems. However, newer OT solutions offer advanced features related to connectivity and automation (e.g., smart OT devices) for an expanding number of industry sectors. Reliance on OT-generated data and features is increasing rapidly, requiring organizations that adopt these new technologies to plan for accessible, secure, and resilient deployments.

Exposing smart OT devices to the internet or public networks can introduce external cyber threats into enterprise networks and environments. For this reason, ZT security best practices mandate that every connected entity has an identity and must be considered an integral part of the ZT Framework—users, devices, virtual infrastructure, and cloud assets[25].

The following section describes several use cases related to ICS and IIoT.

[25] CISA, "Alert (AA20-205A), 23rd, July 2020, https://us-cert.cisa.gov/ncas/alerts/aa20-205a

7.5.1 Use Case Descriptions: CPS, IoT, IIoT, ICS

Per NIST SP 1500-201, cyber-physical systems (CPS) are an integration of physical components, networked systems, embedded computers and software that are linked together for information sharing to make a complete system[26]. CPS serves as the foundation for future smart services, smart cities, smart health care management, and more. As its name implies, CPS are cross-disciplinary in nature and provide seamless integration of cyber and physical systems.

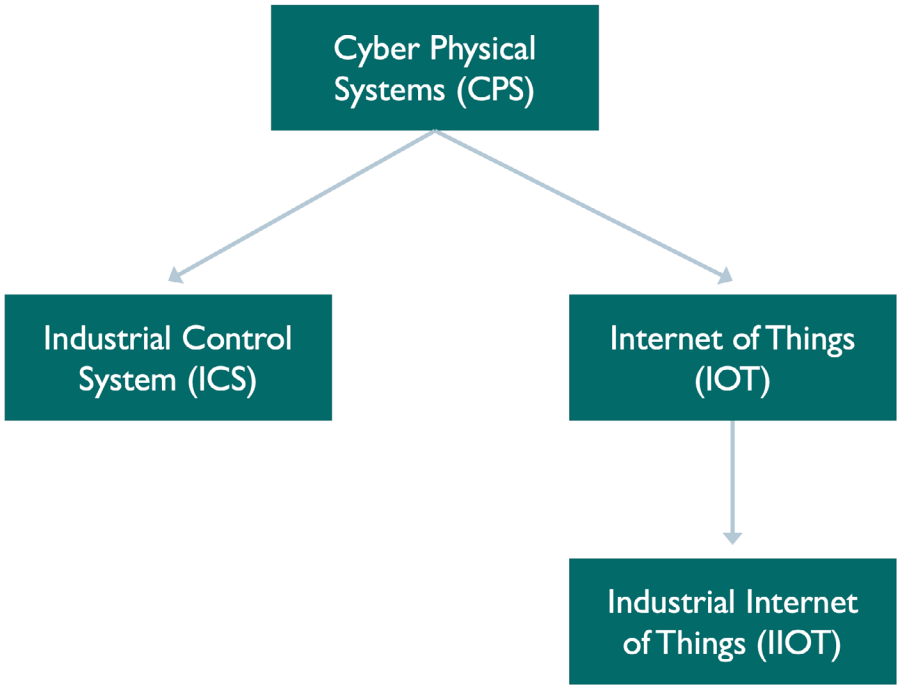


Figure 17: *Cyber-Physical System Types*

7.5.1.1 IoT & IIoT

IoT consists of a network of devices (i.e., things) equipped with software and/or sensors, connected to the internet via wifi or other wireless/wired technology. IoT devices can range from home devices (e.g., home automation solutions, smart doorbells) to industrial equipment (e.g., smart farming devices, assembly line robots). The IIoT is a subset of the IoT that specifically refers to industrial applications. IIoT systems enable industrial enterprises to realize improvements in efficiency and productivity through automation, continuous monitoring, and analysis.

[26] NIST, "SP 1500-201 Framework for Cyber-Physical Systems v1," June 2017, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-201.pdf

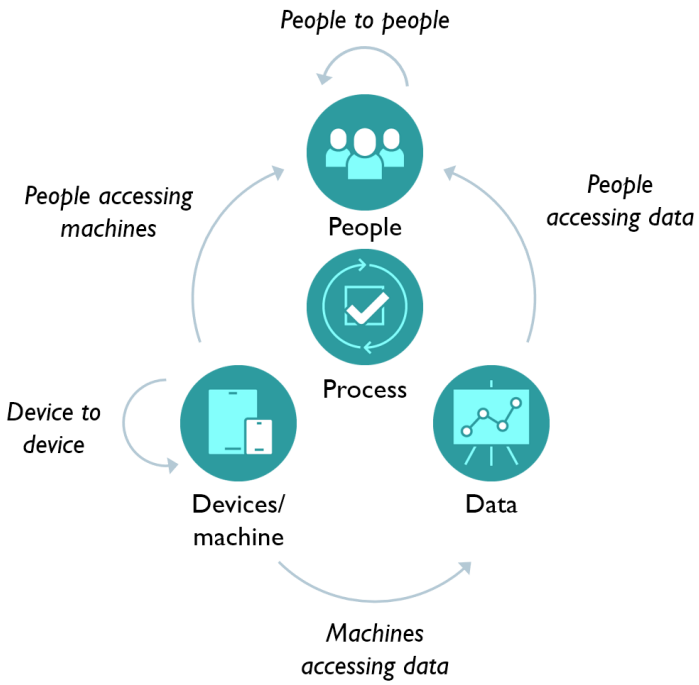


Figure 18: *IoT Entities and Communication Flows*

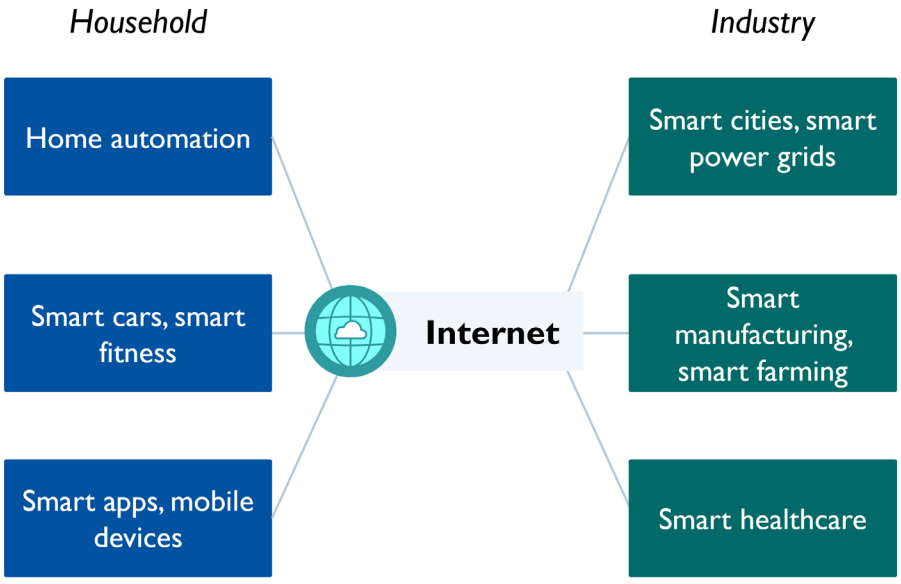


Figure 19: *IoT and IIoT Device Types*

7.5.1.2 Industrial Control Systems

Industrial control systems (ICS) encompass several types of control systems used in industrial production, including the following:

- Supervisory control and data acquisition (SCADA) systems
- Distributed control systems (DCS)
- Programmable logic controllers (PLC), often found in industrial sectors and critical infrastructures[27]

Additionally, commercial-off-the-shelf (COTS) networked devices are increasingly used with industrial automation and control systems (IACS). These COTS devices are typically inexpensive, efficient, and highly automated.

ICS systems typically consist of closed systems with components wired to system controllers in a bus topology. However, organizations increasingly require connectivity between their internal IT network and ICS systems — a requirement that introduces cyber-physical risk into the environment, as ICS systems may enable crucial facility processes for power, lighting, air conditioning, and water management. Organizations should therefore leverage ZT, SDP, and SPA to mitigate the cyber risk created by integrating ICS with an organization's TCP/IP networks.

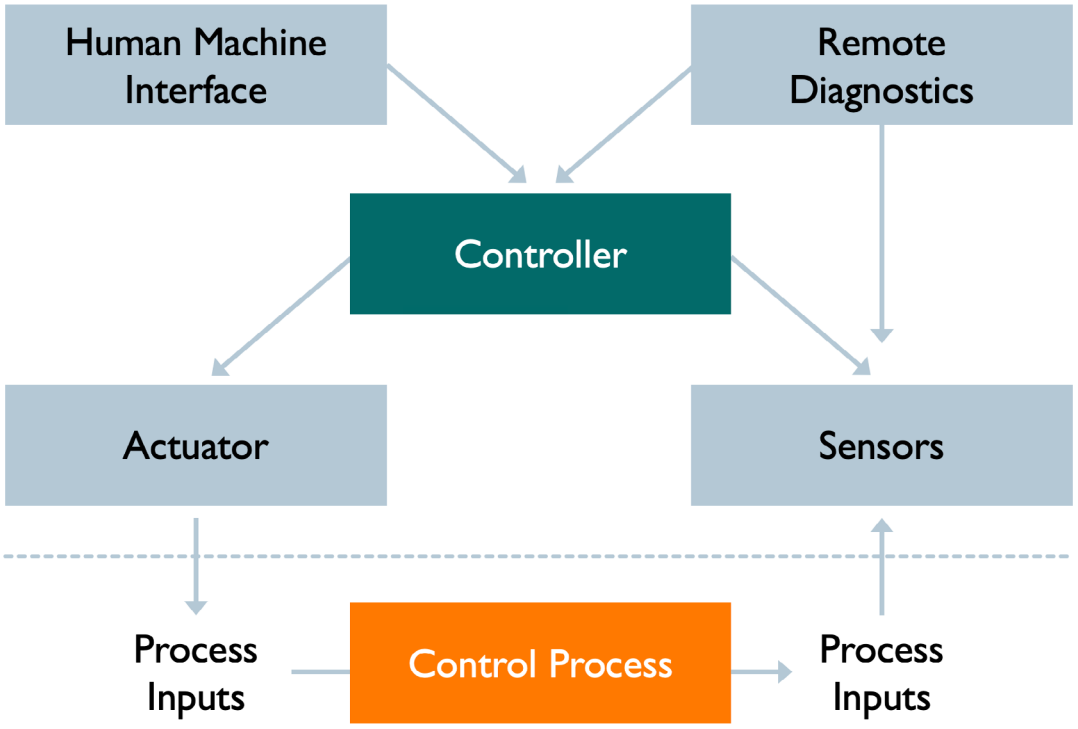


Figure 20: *ICS Communication Flows*[28]

[27] NIST, "SP 800-82 Guide to Industrial Control Systems (ICS) Security," May 2015, https://csrc.nist.gov/publications/detail/sp/800-82/rev-2/final

[28] Figure adapted from NIST, "SP 800-82 Guide to Industrial Control Systems (ICS) Security," May 2015, https://csrc.nist.gov/publications/detail/sp/800-82/rev-2/final

7.5.2 Security Risks

Because IIoT and ICS fall into the domain of industrial or cyber-physical systems, the tenets of confidentiality, integrity and availability (i.e., the CIA Triad) are prioritized differently than traditional IT systems. Instead, availability and integrity take precedence over confidentiality in order to first protect human life and physical assets (e.g., electrical grid). Unfortunately, this lack of confidentiality has led to various high profile incidents in which state-sponsored cyber attackers successfully compromised industrial and cyber-physical systems, causing significant physical damage.

As ICS are widely deployed in critical infrastructure environments such as water, oil/gas, and energy, threats to these systems have a potential for significant harm and loss of life. Subsequently, malicious actors such as terrorists, state-sponsored actors, hacktivists, and criminals have a keen interest in ICS-related vulnerabilities and exploits. To further complicate matters, security hardening and patching is difficult to carry out on these live systems due to their criticality and requirements for high availability.

ICS cyber attacks typically fall into one of the following categories:

1. Attacks that plant malicious software (e.g., Mirai malware) into devices to compromise adjacent resources on the internet/network
2. Attacks that take control of OT devices to steal data or perform unauthorized actions

Over 400 ICS vulnerabilities were disclosed in 2019[29], with over a quarter resulting from unpatched systems. According to United States Computer Emergency Readiness Team (US-Cert) and National Security Agency (NSA), the most common OT threat vectors and exploits include the following:

- Spear phishing to gain a foothold into the organization's IT network, prior to pivoting to the OT network
- Deploying commodity ransomware to encrypt data and adversely impact IT and OT networks
- Connecting to publicly-accessible PLC that require no authentication for initial access
- Exploiting weaknesses in commonly used ports and standard application layer protocols to communicate with controllers and download modified control logic
- Using vendor-supplied engineering software and program downloads compromise systems
- Modifying control logic and parameters on PLCs

7.5.3 ZT Mitigation of Risks

ZT allows organizations to enforce stronger IIoT device integrity and data confidentiality — at both control and data planes — while ensuring the availability of IIoT devices to overall system operations. Additionally, since IIoT devices are considered part of the ICS ecosystem, the ZT model can be leveraged for securely separating IT and OT using micro-segmentation, effectively isolating the business applications on the data plane from those on the control plane.

[29] DRAGOS, "2019 Year in Review ICS Vulnerabilities," 2019, https://www.dragos.com/wp-content/uploads/Year-in-Review-2019_ICS-Vulnerabilities-.pdf

If deployed and/or managed per ZTA specifications, the following OT device types stand to benefit from significant cyber risk reduction:

- **IIoT**: The SDP using SPA reduces the risk exposure of unauthorized user access and rogue IIoT devices (e.g., devices with hardcoded credentials) by enforcing both IIoT device authentication and adaptive risk-based user authentication (e.g., MFA for privileged actions on authorized IIoT devices). Since IIoT endpoints are largely IP-based with one or more network interfaces, standard monitoring solutions used in conjunction with SIEM/SOAR can be used to trigger alerts and defensive measures.
- **ICS**: By bringing ICS components (e.g., SCADA, human machine interface [HMI], DCS) into the fold of SDP with SPA, organizations can limit the highly vulnerable user access to these systems. More specifically, eliminating the bad practices of hardcoded credentials to enforce risk-based user authentication. Furthermore, ZTA with SDP and SPA affords a mechanism to implement micro-segmentation of the control plane of the ICS components with those of the data plane to mitigate the risks due to the interconnectedness of the business applications.

7.5.4 Limitations & Dependencies

In the context of OT, ZT's limitations primarily stem from device resource constraints and ICS systems' use of legacy and/or non-IP based communication protocols at the cyber-physical interface level. For instance, IIoT devices for utility applications may use ZigBee/IEEE 802.15.4 protocols at one interface to communicate with smart meters, while smart meters in turn may use IP protocols to communicate with utility management systems. Furthermore, ICS systems (e.g., SCADA, PLCs) rely on OT protocols such as ModBus or Profinet for control plane functionality. In these scenarios, applying ZT downstream at the edge of the control plane can be challenging.

Because sensors and IIoT controllers are usually limited in their ability to communicate/authenticate with the SDP, architects need to account for these limitations during the ZTA design phase. Ultimately, it may be necessary to incorporate an agentless micro-segmentation or external proxy-based approach, as an agent-based ZTA may not work with OT and IIoT devices.

Additionally, because OT and IIoT devices are harder to patch and/or upgrade, network-based micro-segmentation is critical for protecting adjacent systems against potentially vulnerable devices. Moreover, continuous scanning of traffic and deep packet inspection can be implemented to detect and block known attack types, even if they come from trusted entities.

7.6 5G

7.6.1 Use Case Description

Fifth generation (5G) wireless technology represents a major shift in communications networks, offering new capabilities and connectivity for applications such as smart cities, autonomous vehicles, remote healthcare, and more. With 5G, billions of devices, sensors, and systems will be able to autonomously connect to networks based on time sensitivity, latency, and processing requirements. In addition to faster speeds, greater capacity, and decreased delays, 5G will deliver improved mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communications (uRLLC).

5G uses tiny cells in addition to macro towers to function in the low, mid, and high frequency bands. In highly populated areas, the tiny cells function as signal repeaters, resulting in enhanced speed, network capacity, and reliability. The core network — the backbone of the global communications infrastructure — routes data and connects the different portions of the access network.

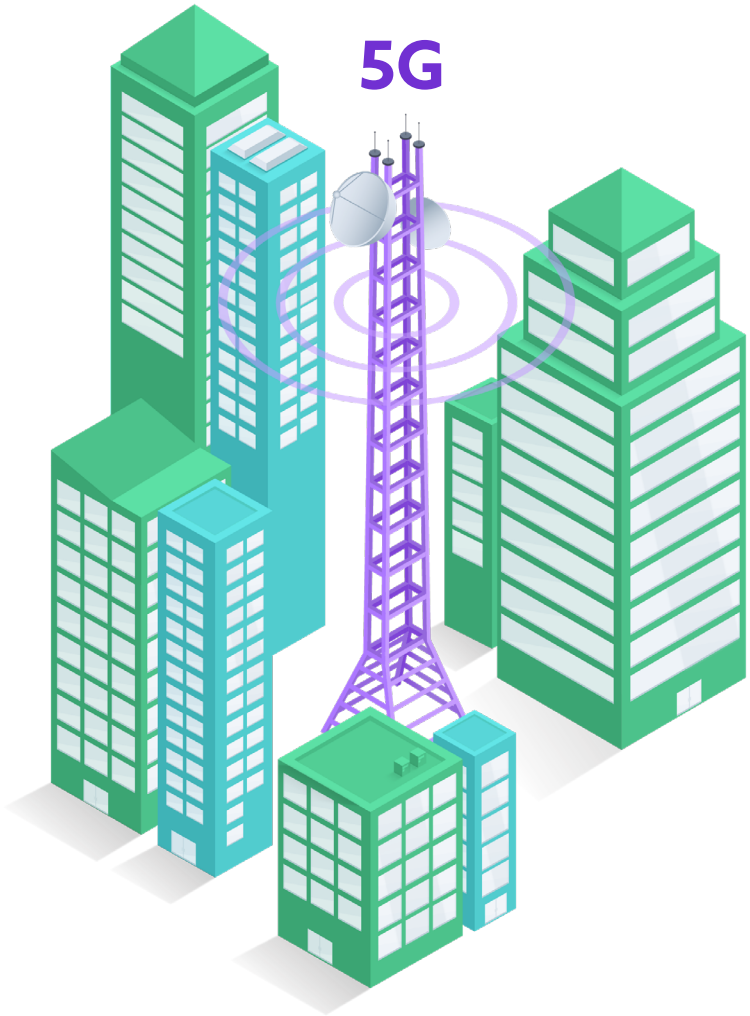


Figure 21: *Fifth Generation*

7.6.2 Security Risks

5G enables several groundbreaking technologies—most notably mobile edge computing (MEC)—that significantly improve application performance and enable unprecedented volumes of real-time data processing. Edge computing places compute and storage resources closer to the customer and/or within the telecommunications network infrastructure. This eliminates performance issues related to backhaul latency (i.e., having to continuously travel back/forth to a central data center).

Unfortunately, these new 5G-enabled configurations also introduce novel security risks into the environment; from the user equipment (UE) to the radio access network (RAN), the mobile edge computing (MEC) to the core nodes, 5G's open architecture makes for an expansive attack surface. Additionally, 5G networks leverage software-defined networking (SDN) technologies; if not properly secured, SDN assets could be compromised by malicious actors, who could in turn reconfigure network devices, monitor all communications, and alter application data.

Also, because 5G brings network devices, storage, computing hardware, and other IT infrastructure closer to the end user, physical security is even more crucial for augmenting ZTA security on the logical layer. It's worth noting this risk factor, along with many others, are common to 4G infrastructures as well, since existing telecommunications devices/equipment operating in remote locations usually lack strong physical security measures for hindering malicious tampering.

7.6.3 ZT Mitigation of Risks

As mentioned previously, 5G networks are software-defined, from the RAN to core and MEC nodes, and are therefore particularly vulnerable to lateral moving malware. ZT device protection can be used to verify the authenticity of software downloads and updates in the system, as well as access to log files.

Because 5G networks support internal compute as well as external cloud resources, they are also vulnerable to MITM attacks. ZT's security model ensures secure connectivity from a 5G UE to a MEC or the cloud. Lastly, ZT data protection can be deployed on IoT-to-5G gateway to ensure only authenticated and authorized systems can access protected data.

7.6.4 Limitations & Dependencies

Integrating ZT requires access to network drivers in 5G infrastructure equipment, which may be difficult to obtain in some vendor implementations. Additionally, ZT's concept of identity and authorization may prove difficult to implement in devices with generic software process names (e.g., store or save). Future ZTA versions will need to support an agentless approach to facilitate the myriad of edge configurations made possible by 5G, as not all edge devices can support agent software installations.

Conclusion

In this introductory ZTA course, we provided learners with an overview of ZT's origins — how it emerged against an increasingly complex technology landscape, why new computing paradigms such as the cloud and virtualization require novel approaches to security, and how early predecessors from both government and enterprise served as models for ZT's foundational concepts. We defined key terminology and principles, followed by an exploration of technical and business benefits that ZT can bring to organizations.

With the historical drivers, early developments, and context of ZTA established, we then outlined planning considerations for ZT adoptions. Learners were given an overview of implementation risks, implementation options, followed by representative use cases to get a sense of how ZTA bolsters security across various industries and application scenarios.

Glossary

For additional terms, please refer to our Cloud Security Glossary, a comprehensive glossary that combines all the glossaries created by CSA Working Groups and research contributors into one place.

Term	Definition	Source
802.1x	An IEEE standard for local and metropolitan area networks–Port-Based Network Access Control. IEEE 802 LANs are deployed in networks that convey or provide access to critical data, that support mission critical applications, or that charge for service. Port-based network access control regulates access to the network, guarding against transmission and reception by unidentified or unauthorized parties, and consequent network disruption, theft of service, or data loss.	https://1.ieee802.org/security/802-1x/
Accepting Host (AH)	The SDP policy enforcement points (PEPs) that control access to any resource (or service) to which an identity might need to connect, and to which the responsible enterprise needs to hide and control access. AHs can be located on-premises, in a private cloud, public cloud, etc.	https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/
Access	To make contact with one or more discrete functions of an online, digital service.	https://csrc.nist.gov/glossary/term/access
Active Directory (AD)	A Microsoft directory service for the management of identities in Windows domain networks.	https://csrc.nist.gov/glossary/term/active_directory
Air-Gapped Networks	An interface between two systems at which (a) they are not connected physically and (b) any logical connection is not automated (i.e., data is transferred through the interface only manually, under human control).	https://csrc.nist.gov/glossary/term/air_gap

Application Programming Interface (API)	A system access point or library function that has a well-defined syntax and is accessible from application programs or user code to provide well-defined functionality.	https://csrc.nist.gov/glossary/term/application_programming_interface
Attribute-Based Access Control (ABAC)	An access control approach in which access is mediated based on attributes associated with subjects (requesters) and the objects to be accessed. Each object and subject has a set of associated attributes, such as location, time of creation, access rights, etc. Access to an object is authorized or denied depending upon whether the required (e.g., policy-defined) correlation can be made between the attributes of that object and of the requesting subject.	https://csrc.nist.gov/glossary/term/abac
Authentication	Verifying the identity of a user, process, or device, often as a prerequisite to allowing access to resources in an information system.	https://csrc.nist.gov/glossary/term/authentication
Authorization	The right or a permission that is granted to a system entity to access a system resource.	https://csrc.nist.gov/glossary/term/authorization
Brute Force Attacks	An attempt to discover a password by systematically trying every possible combination of letters, numbers, and symbols until you discover the one correct combination that works.	https://owasp.org/www-community/controls/Blocking_Brute_Force_Attacks
Cloud Access Security Broker (CASB)	On-premises, or cloud-based security policy enforcement points, placed between cloud service consumers and cloud service providers to combine and interject enterprise security policies as the cloud-based resources are accessed. CASBs consolidate multiple types of security policy enforcement.	https://www.gartner.com/en/information-technology/glossary/cloud-access-security-brokers-casbs
Control Plane	Used by various infrastructure components (both enterprise-owned and from service providers) to maintain and configure assets; judge, grant, or deny access to resources; and perform any necessary operations to set up communication paths between resources.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

Controller (SDP Controller)	Determines which SDP hosts can communicate with each other. The controller may relay information to external authentication services such as attestation, geo-location, and/or identity servers.	https://downloads.cloudsecurityalliance.org/initiatives/sdp/Software_Defined_Perimeter.pdf
Data Plane	Used for communication between software components. This communication channel may not be possible before the path has been established via the control plane.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf
Distributed Denial-of-Service (DDoS)	Involves multiple computing devices in disparate locations sending repeated requests to a server with the intent to overload it and ultimately render it inaccessible.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1800-15.pdf
Firewall	An inter-network connection device that restricts data communication traffic between two connected networks. A firewall may be either an application installed on a general-purpose computer or a dedicated platform (appliance), which forwards or rejects/drops packets on a network. Typically firewalls are used to define zone borders. Firewalls generally have rules restricting which ports are open.	https://csrc.nist.gov/glossary/term/firewall
Gateway (SDP Gateway)	Provides authorized users and devices with access to protected processes and services. The gateway can also enact monitoring, logging, and reporting on these connections.	https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/
Hypertext Transport Protocol Secure (HTTPS)	A secure network communication method, technically not a protocol in itself, HTTPS is the result of layering the Hypertext Transfer Protocol (HTTP) on top of the SSL/TLS protocol, thus adding the security capabilities of SSL/TLS to standard HTTP communications.	https://iapp.org/resources/article/hypertext-transfer-protocol-secure/
Identity (ID)	The set of attribute values (i.e., characteristics) by which an entity is recognizable and that, within the scope of an identity manager's responsibility, is sufficient to distinguish that entity from any other entity.	https://csrc.nist.gov/glossary/term/identity

Identity and Access Management (IAM)	The set of technology, policies, and processes that are used to manage access to resources.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-203.pdf
Identity Provider (IdP)	A trusted entity that issues or registers subscriber authenticators and issues electronic credentials to subscribers. A cloud service provider may be an independent third party or issue credentials for its own use.	https://csrc.nist.gov/glossary/term/identity_provider
Initiating Host (IH)	The host that initiates communication to the controller and to the AHs.	https://downloads.cloudsecurityalliance.org/initiatives/sdp/SDP_Specification_1.0.pdf
Lightweight Directory Access Protocols (LDAP)	A networking protocol for querying and modifying directory services running over TCP/IP.	https://csguide.cs.princeton.edu/email/setup/ldap
Man-in-the-middle (MITM) attacks	An attack where the adversary positions himself in between the user and the system so that he can intercept and alter data traveling between them.	https://csrc.nist.gov/glossary/term/mitm
Micro-segmentation	Is the technique of creating secure zones within a data center and cloud deployments that allow the organization to separate and secure each workload. This makes network security more granular and effective. These secure zones are created based on business services, and rules are defined to secure information workflow.	https://www.techtarget.com/searchnetworking/definition/microsegmentation
Multi-factor Authentication (MFA)	Authentication using two or more factors to achieve authentication. Factors include: (i) something you know (e.g., password/personal identification number (PIN)); (ii) something you have (e.g., cryptographic identification device, token); or (iii) something you are (e.g., biometric).	https://csrc.nist.gov/glossary/term/multi_factor_authentication

Mutual Transport Layer Security (mTLS)	An approach where each microservice can identify who it talks to, in addition to achieving confidentiality and integrity of the transmitted data. Each microservice in the deployment has to carry a public/private key pair and uses that key pair to authenticate to the recipient microservices via mTLS.	https://cheatsheetseries.owasp.org/cheatsheets/Microservices_security.html#mutual-transport-layer-security
Network Access Control (NAC)	A method of bolstering the security of a private or "on-premise" network by restricting the availability of network resources to endpoint devices that comply with a defined security policy.	https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-41r1.pdf
Network Segmentation	Splitting a network into sub-networks, for example, by creating separate areas on the network which are protected by firewalls configured to reject unnecessary traffic. Network segmentation minimizes the harm of malware and other threats by isolating it to a limited part of the network.	https://www.nist.gov/itl/smallbusinesscyber/cybersecurity-basics/glossary
Open Systems Interconnection (OSI)	Qualifies standards for the exchange of information among systems that are "open" to one another for this purpose by virtue of their mutual use of applicable standards.	https://www.ecma-international.org/wp-content/uploads/s020269e.pdf
Phishing	A technique for attempting to acquire sensitive data, such as bank account numbers, through a fraudulent solicitation in email or on a web site, in which the perpetrator masquerades as a legitimate business or reputable person.	https://csrc.nist.gov/glossary/term/phishing
Policy decision point (PDP)	Mechanism that examines requests to access resources, and compares them to the policy that applies to all requests for accessing that resource to determine whether specific access should be granted to the particular requester who issued the request under consideration.	https://csrc.nist.gov/glossary/term/policy_decision_point
Policy enforcement point (PEP)	A system entity that requests and subsequently enforces authorization decisions.	https://csrc.nist.gov/glossary/term/policy_enforcement_point

Port	Another essential asset through which security can be breached. In computer science, ports are of two types - physical ports (which is a physical docking point where other devices connect) and logical ports (which is a well-programmed docking point through which data flows over the internet). Security and its consequences lie in a logical port.	https://www.w3schools.in/cyber-security/ports-and-its-security/
Public Key Infrastructure (PKI)	The framework and services that provide for the generation, production, distribution, control, accounting, and destruction of public key certificates. Components include the personnel, policies, processes, server platforms, software, and workstations used for the purpose of administering certificates and public-private key pairs, including the ability to issue, maintain, recover, and revoke public key certificates.	https://csrc.nist.gov/glossary/term/public_key_infrastructure
Role Based Access Control (RBAC)	Access control based on user roles (i.e., a collection of access authorizations a user receives based on an explicit or implicit assumption of a given role). Role permissions may be inherited through a role hierarchy and typically reflect the permissions needed to perform defined functions within an organization. A given role may apply to a single individual or to several individuals.	https://csrc.nist.gov/glossary/term/role_based_access_control
Security Assertion Markup Language (SAML)	A protocol consisting of XML-based request and response message formats for exchanging security information, expressed in the form of assertions about subjects, between online business partners.	https://csrc.nist.gov/glossary/term/security_assertion_markup_language
Security Orchestration Automation and Response (SOAR)	Refers to technologies that enable organizations to collect inputs monitored by the security operations team. SOAR tools allow an organization to define incident analysis and response procedures in a digital workflow format.	https://www.gartner.com/en/information-technology/glossary/security-orchestration-automation-response-soar

Single Packet Authorization (SPA)	Can authenticate a user to a system for simple remote administration. It is a protocol for allowing a remote user to authenticate securely on a "closed" system (limited or no open services) and make changes to or run applications on the "closed" system.	https://www.blackhat.com/presentations/bh-usa-05/bh-us-05-madhat.pdf
Software-Defined Network (SDN)	An approach to computer networking that allows network administrators to manage network services through abstractions of higher-level functionality. SDNs manage the networking infrastructure. This is done by decoupling the system that makes decisions about where traffic is sent (the control plane) from the underlying systems that forward traffic to the selected destination (the data plane).	https://ieeexplore.ieee.org/abstract/document/6819788
Software-Defined Perimeter (SDP)	A network security architecture that is implemented to provide security at Layers 1-7 of the OSI network stack. An SDP implementation hides assets and uses a single packet to establish trust via a separate control and data plane prior to allowing connections to hidden assets.	https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-and-zero-trust/
Transmission Control Protocol (TCP)	A transport protocol that is used on top of IP to ensure reliable transmission of packets. TCP includes mechanisms to solve many of the problems that arise from packet-based messaging, such as lost packets, out of order packets, duplicate packets, and corrupted packets. Since TCP is the protocol used most commonly on top of IP, the Internet protocol stack is sometimes referred to as TCP/IP.	https://www.khanacademy.org/computing/computers-and-internet/xcae6f4a7ff015e7d:the-internet/xcae6f4a7ff015e7d:transporting-packets/a/transmission-control-protocol--tcp
Transmission Control Protocol/ Internet Protocol (TCP/IP)	A set of protocols covering (approximately) the network and transport layers of the seven-layer Open Systems Interconnection (OSI) network model.	https://www.gartner.com/en/information-technology/glossary/tcpip-transmission-control-protocolinternet-protocol

Introduction to Software-Defined Perimeter

CCZT Study Guide

cloud security alliance®

CCZT™
Certificate of
Competence in Zero Trust

The official location for Software-Defined Perimeter Working Group is
https://cloudsecurityalliance.org/research/working-groups/software-defined-perimeter/

Disclaimer

Cloud Security Alliance designed and created this Zero Trust Training course study guide (the "Work") primarily as an educational resource for security and governance professionals. Cloud Security Alliance makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Version Number: 20240820

About Cloud Security Alliance

The Cloud Security Alliance℠ (CSA) (www.cloudsecurityalliance.org) is the world's leading organization dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment. Cloud Security Alliance harnesses the subject matter expertise of industry practitioners, associations, governments, and its corporate and individual members to offer cloud security-specific research, education, certification, events and products. Cloud Security Alliance activities, knowledge and extensive network benefit the entire community impacted by cloud—from providers and customers, to governments, entrepreneurs and the assurance industry—and provide a forum through which diverse parties can work together to create and maintain a trusted cloud ecosystem.

CSA Address

709 Dupont St.
Bellingham, WA 98225, USA
Phone: +1.360.746.2689
Fax: +1.206.832.3513

Contact us: support@cloudsecurityalliance.org
Website: https://cloudsecurityalliance.org/
Zero Trust Training Page: https://knowledge.cloudsecurityalliance.org/page/zero-trust-training
Zero Trust Advancement Center: https://cloudsecurityalliance.org/zt/
Provide Feedback: support@cloudsecurityalliance.org
CSA Circle Online Community: https://circle.cloudsecurityalliance.org/
Twitter: https://twitter.com/cloudsa
LinkedIn: www.linkedin.com/company/cloud/security/alliance
Facebook: www.facebook.com/csacloudfiles
CSA CloudBytes Channel: http://www.csacloudbytes.com/
CSA Research Channel: https://www.brighttalk.com/channel/16947/
CSA Youtube Channel: https://csaurl.org/youtube
CSA Blog: https://cloudsecurityalliance.org/blog/

Acknowledgments

Dedicated to Juanita Koilpillai, a pioneer in software-defined perimeters whose contributions to the Certificate of Competence in Zero Trust (CCZT), Zero Trust training, and CSA are immeasurable.

The CCZT and Zero Trust training was developed with the support of the Cloud Security Alliance Zero Trust Expert Group, whose members include volunteers from a wide variety of industries across the globe. Made up of subject matter experts with hands-on experience planning and implementing Zero Trust, both as cloud service consumers and providers, the Zero Trust Expert Group includes board members, the technical C-suite, as well as privacy, legal, internal audit, procurement, IT, security and development teams. From cumulative stakeholder input, the Zero Trust Expert Group established the value proposition, scope, learning objectives, and curriculum of the CCZT and Zero Trust training.

To learn more about the CCZT and Zero Trust training and ways to get involved please visit: https://cloudsecurityalliance.org/education/cczt

We would also like to thank our beta testers, who provided valuable feedback on the CCZT and Zero Trust training: https://cloudsecurityalliance.org/contributors/cczt-contributors

Lead Developers:

Juanita Koilpilla
Richard Lee

Contributing Editors:

James Lam
Remo Hardeman

Expert Reviewer:

Anusha Vaidyanathan
Juan Carlos (Charlie) Soto
Matthew Meersman, PhD
Michael J. Herndon
Michael Roza
Nishanth Singarapu
Robert D. Morris
Shinesa Cambric
Vani Murthy

CSA Global Staff:

Anna Schorr
Daniele Catteddu
Hannah Rock
Jenna Morrison
Leon Yen
Ryan Bergsma
Shamun Mahmud
Stephen Smith

Table of Contents

List of Figures

Course Intro

Welcome to your *Introduction to Software-Defined Perimeter* by Cloud Security Alliance. Please note that moving forward we will refer to Software-Defined Perimeter as SDP and to the Cloud Security Alliance as CSA. CSA is dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment across the globe. We hope you are as excited to learn about SDP as we are about sharing this knowledge with you. This training module is part of a larger series of CSA programs on Zero Trust (ZT) that was created with the support of subject matter experts. If you are interested in volunteering with CSA to help our ongoing research efforts or are just interested in learning more about cloud security, please visit our website at cloudsecurityalliance.org.

This course is intended to give a high-level overview of why SDP was created, what it is, what it does, how it can be used, and how it relates to ZT and ZTA. Although it is not within the scope of this course to delve into SDP implementation how-tos, CSA will be releasing additional training courses that will elaborate on ZTA and further explore the details of SDP.[1]

Course Structure

This introductory course on SDP consists of four units, each geared towards helping learners gain competency in a specific area/topic:

- SDP History, Benefits, & Concepts
- Traditional Architecture Issues & SDP Solutions
- Core Tenets, Underlying Technologies, & Architecture
- The Basics of SDP Deployment Models

Course Learning Objectives

After completing this course, learners will be able to do the following:

- Explain what SDP is, how it came about, and what its technology and business benefits are
- Discuss the problems that SDP solves
- Describe some of SDP's underlying technologies
- Distinguish between the basic types of SDP deployments

[1] Cloud Security Alliance, "Zero Trust Architecture Training," https://cloudsecurityalliance.org/education/zero-trust-architecture-training

1 SDP History, Benefits, & Concepts

In this unit, you will be introduced to the concept of SDP, as well as gain a high-level overview of SDP architecture. Part of this introduction includes learning about the basics, such as the history of SDP, its technological and business benefits, as well as other related concepts.

1.1 SDP Definition & Function

CSA defines SDP[2] as a network security architecture implemented to provide security for all layers of the open systems interconnection (OSI) model. An SDP implementation hides assets and uses a single packet to establish trust via a separate control and data plane; only then are assets exposed to the requestor.

Although SDP has different roots than the ZT security model, the evolution of both concepts over time has led to community consensus in categorizing SDP as an implementation option of a ZTA. In order to isolate services from unsecured networks, SDP aims to give infrastructure and application owners the ability to deploy perimeter functionality when and where it's needed. SDP overlays existing physical infrastructure with logical components that should be operated under the control of the application owner. SDP only grants access to the application infrastructure after device attestation and identity verification.

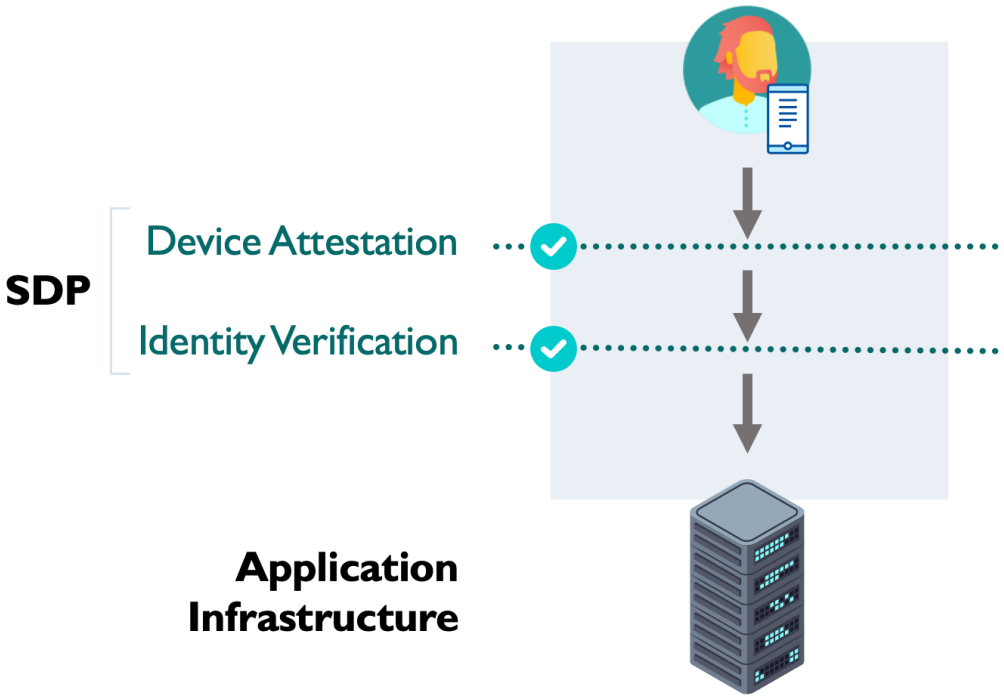


Figure 1: *Access Granted After Device Attestation/Identify Verification*

[2] https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-and-zero-trust/

SDP is based on the premise that organizations should not implicitly trust anything inside or outside the network. It requires users on validated devices to cryptographically sign in to the perimeter created around hidden assets, even as they reside on public infrastructures. An SDP implementation hides assets with a drop-all firewall, uses a single packet to establish trust via a separate control plane, and provides mutual verification of connections in a data plane to hidden assets.

SDP brings together multiple controls that are usually separated by function and therefore hard to integrate: applications, firewalls, and clients, to name a few. These pieces of information need unification in order to establish and ensure secure connections. SDP helps to integrate controls for firewalls, encryption, identity and access management (IAM), session management, and device management into a comprehensive security architecture.

1.2 SDP Principles

The SDP architecture is based on the principles of least privilege and segregation of duties, enforced by implementing the following key controls and processes:

- Dynamic rules on drop-all firewalls
- Hiding servers and services
- Authentication before connections, for example not allowing connections before authorizing users on specific devices
- Using single packet authorization (SPA) and/or bi-directional encrypted communications like mutual transport layer security (mTLS)
- Fine-grained access control and device validation

1.3 Relationship Between SDP & ZT

In this section, you will learn about the relationship between SDP and ZT. ZT is the umbrella category under which SDP falls.

The ZT model is based on the following principles:

- Making no assumptions about the trustworthiness of an entity as it requests access to a resource
- Starting with no pre-established privileges, then relying on a construct which is used to add privileges
- Assuming breach and verifying all workforce, device, workload, network, and data access regardless of where, who, when, or to what resource

In essence, the ZT concept retires the use of trusted entities inside a defined corporate perimeter. Instead, it mandates that enterprises create micro-perimeters around sensitive data assets to maintain control and visibility around data use across the environment. Essentially, ZT aims to defend enterprise assets by distrusting anything inside or outside the perimeter. Implementing ZT requires verifying connection requests to assets before granting access, followed by continuous monitoring and evaluation throughout the entire duration. For additional references on ZT concepts

and architectures, please refer to the existing literature on the topic, and additional CSA training[3].

By comparing the foundational principles of SDP and ZT, it is clear that they are driven by the same high-level principle: "never trust, always verify". In fact, SDP is considered one implementation type of a ZTA; others include Zero Trust Network Access (ZTNA) and Google BeyondCorp, to name a few.

Compared to other ZTA implementations, SDP has some distinctive features and benefits, such as the use of a drop-all rule and the adoption of SPA. While these features are not necessarily unique to SDP, they are foundational to it; however, these features are not necessary requirements of other ZTA implementations.

> **NOTE:** SDP is a ZTA, but not every ZTA conforms with SDP requirements.

1.4 History of SDP

In this section, you will learn about the history of SDP, its origins, and why it was developed.

1.4.1 The Origination of SDP

SDP is a cybersecurity approach that evolved from the U.S. Defense Information Systems Agency's Global Information Grid Black Core Network initiative in 2007[4]. Designed to be extensible and future proof, this approach would later serve as the basis for CSA's SDP framework in 2013. The CSA SDP framework focuses on how to control access to resources based on identity and device attestation. Per SDP, connectivity is provided on a need to know model that verifies device posture and identity before granting access to an application infrastructure. Because the application infrastructure exists without visible domain name system (DNS) information or IP addresses, it is effectively hidden and undetectable unless access is specifically granted.

1.4.2 The Business Case for SDP

As organizations continue to undergo digital transformation, staying ahead of the threat landscape and attack chain curves is becoming increasingly difficult to achieve. Today, rather than managing and securing a single network, most organizations operate a variety of environment types, such as the following:

> - Physical, on-premises networks
> - Private clouds
> - Multiple public clouds
> - Virtual software-defined networking (SDN) environments

[3] Cloud Security Alliance, "Publications," https://cloudsecurityalliance.org/research/artifacts/
[4] DOD, "Vision for a Net-Centric, Service-Oriented DoD Enterprise," June 2007, https://www.acqnotes.com/Attachments/DoD%20GIG%20Architectural%20Vision,%20June%2007.pdf

Within these newer environments, organizations must facilitate the following:

- An expanding wide area network edge
- Information technology and operation technology convergence
- An increasingly mobile workforce

As organizations shift from traditional infrastructures to more virtualized and hybrid architectures, new attack vectors also emerge that require a novel approach to network security. SDP's designers focused on mitigating the most common network-based attacks, including server scanning, denial of service, SQL injection, operating system and application vulnerability exploits, man-in-the-middle, pass-the-hash, pass-the-ticket, to name a few. Despite the evolving cyber threat landscape, SDP continues to hold up against both existing and unknown threats.

1.5 Technology Benefits of SDP

In this section, we will explore the technological benefits of SDP. Some of these include SDP's attack surface reduction and pre-access authentication/authorization. We will also discuss SDP technological benefits such as IAM security and SDP's open specification.

1.5.1 Reduced Attack Surface

Today's network architectures consist of devices with assigned IP addresses used for connectivity. When a device is establishing a connection to another device, a handshake is established and authentication is verified. By reversing this sequence and first verifying the connection, SDP provides key technical benefits, most notably attack surface reduction. Connectivity to an organization's assets is provided only after authentication, validation/authorization, and the determination of which protected assets the user is allowed access to. These steps greatly reduce the attack surface of the application infrastructure.

With SDP, users and devices are no longer granted general access to network segments or subnets. Instead, policies ensure that users and devices only have access to specified hosts, resources, and/or services. Therefore, SDP can be used to protect different types of services or protocols such as Hypertext Transfer Protocol Secure (HTTPS) or remote desktop services (RDS). By controlling the access level that individual users and devices have to specific services, SDP can allow authorized users to access privileged services while hiding them from unauthorized users.

1.5.2 Authenticate & Authorize Before Access

SDP is an inherently comprehensive security architecture implemented using software components overlaid onto physical and virtual infrastructure. SDP uses a drop-all gateway to ensure that authentication and authorization is first performed in the control plane. By performing authentication prior to granting access to the perimeter, SDP ensures only users with appropriate authorization have access to the hidden infrastructure.

This functionality (i.e., providing connectivity to resources after authentication and authorization) is made possible by separating the control and data planes, providing enhanced protection by exposing assets only to verified users and devices. Fine-grained access control is implicit in SDP's design.

Without the SDP gateway's drop-all capability, allowing and enforcing only trusted connections would be prohibitively difficult. SDP's architecture enables pre-access vetting and fine-grained access policies through role and attribute-based permissions, as well as other similar access control mechanisms. Traditional architectures require separate implementations for each of these components, leading to increased complexity and higher maintenance overhead.

In contrast to IP-based alternatives, SDP provides a connection-based security architecture — this means access is granted per each independent connection, versus granting access to a device based on its allowlisted IP address.

SDP is a connection-oriented security architecture: while the physical infrastructure routes packets, SDP secures all connectivity over an infrastructure. This connection-based architecture distinction is important because of the current IP address explosion and the disintegrated perimeter in cloud environments — without SDP, IP-based security protections are ineffective when faced with this increasing complexity. SDP enables validation on the data plane prior to any Transmission Control Protocol/Transport Layer Security (TLS/TCP) handshake and enforces mutually encrypted communications. This practice helps to mitigate threats related to unauthorized access.

1.5.3 Centralized Organizational IAM Security

A prominent technical aspect of SDP is its centralized organizational IAM security. With IAM, a security problem on the front-end only requires an update to the SDP — every subsequent service within the perimeter will adjust to the heightened security measures. Traditional, direct access would require the checking and updating of potentially hundreds of services to address a single flaw. This is another example of how SDP drastically decreases maintenance overhead and complexity.

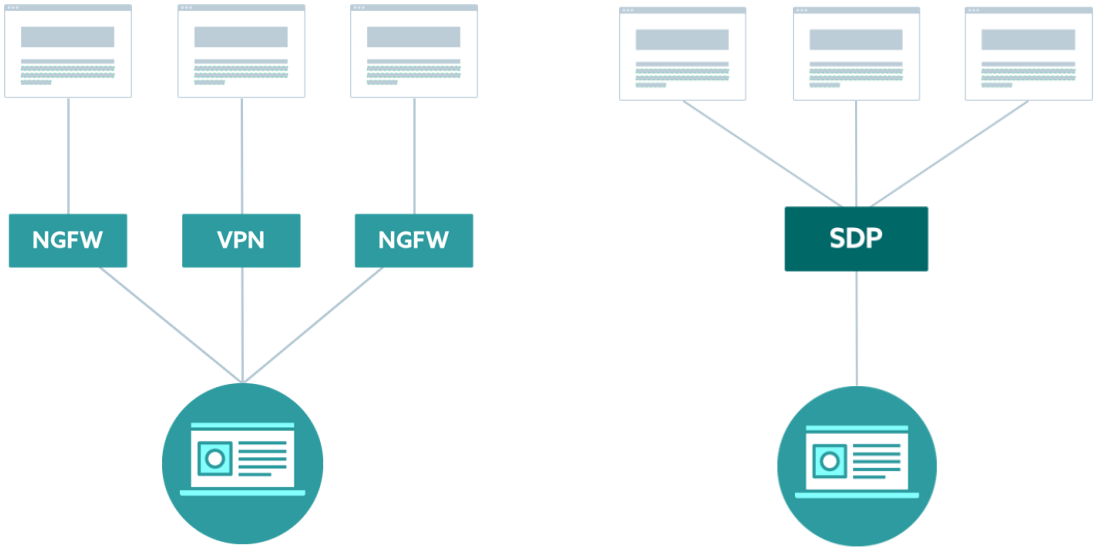


Figure 2: *Traditional IAM Security vs. SDP*

1.5.4 Open Specification

Open specifications are publicly available and therefore directly benefit from greater community contributions. This increases the volume of data flowing in, the validity, and practicality of the specification that is developed, based on a given set of data. With an open specification you can customize the code or implementation to your needs, audit the code as it exists, and receive community feedback on faults and errors.

The SDP specification is open and has been proven on many network implementations, such as SDNs, IoT networks, network functions virtualization, edge computing, 5G, and more. As part of the research efforts, the CSA Software-Defined Perimeter Working Group teamed up with the community at large to research how to create a high availability infrastructure using public clouds with the equivalent robustness of a dedicated data center. The CSA Software-Defined Perimeter Working Group has also created additional reference materials, such as *SDP Architecture Guide v2*[5] and *Software-Defined Perimeter as a DDoS Prevention Mechanism vs. SDP and DDoS*[6] that are publicly available. These documents were created with input from the global cybersecurity community.

1.6 Business Benefits of SDP

In this section, we will discuss the various business benefits that companies gain from implementing SDP. As part of this discussion, we will examine how SDP enhances existing cybersecurity investments, reduces costs and labor, and assists in governance, risk, and compliance (GRC) efforts.

1.6.1 Enhances Existing Cybersecurity Investments

Organizations are under continuous pressure to respond to security events in a timely manner; to this end, they've made substantial investments in cybersecurity. For example, expenditures in vulnerability management, patch management, and configuration management, have allowed organizations to lock down machines that utilize IP addresses for connectivity. Threat intelligence combined with endpoint threat detection and response (EDR) may also be in place, enabling organizations to better understand who the unauthorized users are and what connections they are making. Many organizations also manage their own security operation centers to actively monitor for threats and respond to intrusion alerts and other security events. SDP helps optimize security investments and make them more cost effective as a result of both preventive and reactive security capabilities.

SDP provides a preventive measure against network-based and cross-domain attacks. By hiding resources and applying the drop-all rules, SDP helps companies reduce their attack surface and consequently reduce the amount of security events or alerts that are collected by the security information and event management (SIEM) and sent to the security operation center. In addition, SDP reduces lateral movement in attacks by keeping assets invisible to unauthorized users. SDP helps reduce the

[5] Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

[6] Cloud Security Alliance, "Software-Defined Perimeter as a DDoS Prevention Mechanism," 27th, October 2019, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-as-a-ddos-prevention-mechanism/

complexity of integrating controls like firewalls, IAM, encryption, and device management by maintaining all rules in one place instead of addressing them for each individual implementation. This allows companies to focus internal resources on a smaller set of potentially negative events, therefore increasing the cost-effectiveness of the security investments.

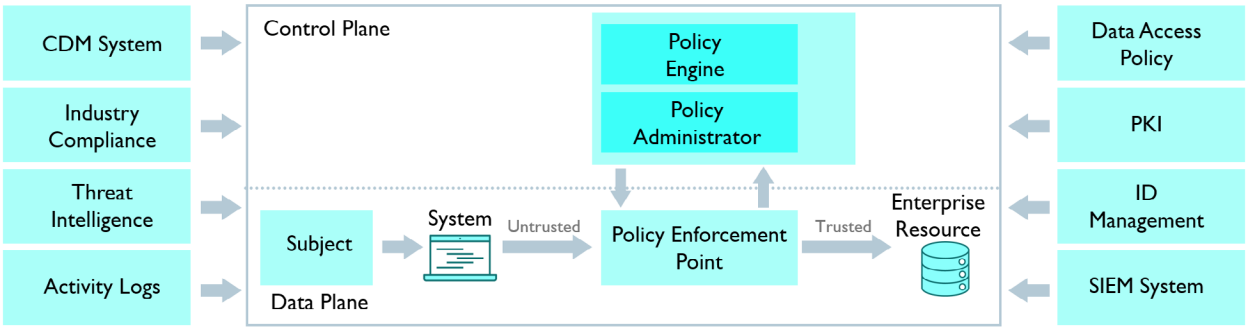


Figure 3: *SDP Ecosystem and Communication Flows*[7]

1.6.2 Cost Reduction & Labor Savings

Replacing traditional network security components with SDP reduces licensing and support costs. Implementing and enforcing security policies using SDP reduces operational complexity and reliance on traditional security tools. SDP also reduces costs of corporate backbone components by reducing or replacing multiprotocol label switching (MPLS) or leased line utilization. As information and communication technology environments change, reliance on corporate backbone is reduced, and more dynamic networks are implemented. SDP allows organizations to achieve dynamic network implementations securely. Ultimately, SDP brings efficiency and simplicity to organizations, which can ultimately help reduce scarce and often expensive labor needs.

1.6.3 Reduces Compliance Scope

As mentioned earlier, two of the main technology benefits of SDP are the reduction of the attack surface and an increased granular control over resource access. These two features, alongside micro-segmentation, are key to helping organizations better face compliance challenges, as they allow the reduction of the scope of compliance. By better controlling where regulated data are processed and stored, and by limiting, both physically and logically, who can have access to that data, organizations can reduce the scope of the compliance requirements. In addition, granular logging and monitoring of who-does-what-when-why support the creation of a much-needed accountability approach, which is foundational to any compliance effort.

[7] Figure adapted from NIST, "SP 800-207 Zero Trust Architecture," August 2020, https://csrc.nist.gov/publications/detail/sp/800-207/final

2 Traditional Architecture Issues & SDP Solutions

This unit reviews various issues that exist within current network security architectures. We will discuss how SDP protects against threats that exist due to those architectural inadequacies. In addition, we will explore how SDP integrates with industry adopted solutions or replaces them.

2.1 Concerns SDP Addresses

In this section you will learn about critical issues that SDP addresses, including the changing perimeter, the IP address challenge, and the integration of security controls.

2.1.1 The Shifting Perimeter

Virtualized networks have superseded the older, fixed network perimeter paradigm that relies on trusted internal network segments protected by network appliances (e.g., load balancers and firewalls). Network protocols of the past are not secure by design and are known to have vulnerabilities. In addition, the plethora of mobile and IoT devices further challenge the validity of a fixed network perimeter.

The introduction of the cloud has drastically changed the composition of organizations' IT environments. With the emergence of bring your own device (BYOD), machine-to-machine connectivity, the rise in remote access, and phishing attacks, legacy security approaches are no longer effective in protecting the shifting physical perimeter. For one, there are more internal devices and varieties of users. For example, contractors working on-site may require temporary access to IT resources, both on-premises and in the cloud. IT environments are also increasingly diversified with the continued enterprise adoption of hybrid architectures. Corporate devices are moving to the cloud, co-located facilities, and in some cases to off-site customer and partner facilities. These migrations further shift the physical perimeter of the organization; SDP addresses the inherent challenges of securing this shifting physical perimeter with a software overlay that creates virtual perimeters dynamically, when and where they are needed.

2.1.2 The IP Address Challenge

Everything on the internet today relies on TCP/IP for trust. This is problematic, because IP addresses have no concept of users' identities. TCP/IP simply addresses connectivity — it doesn't validate the endpoint or the user as being trustworthy.

TCP/IP is a bidirectional protocol, so internal trusted hosts communicating with external untrusted hosts can receive unsafe messages. Any changes to IP addresses may require extensive reconfiguration resulting in potential security group and network access control (NAC) list errors. Unmanaged/forgotten internal hosts can provide an entry point for malicious actors by providing default responses using legacy protocols such as ICMP. This illustrates that common use of network address translation (NAT) tables combined with TCP/IP is inherently open to compromise.

IP addresses should not be used as anchors for network locations because they are location-dependent (i.e., users' devices are assigned new IP addresses when they are relocated). SDP tackles this IP address challenge by securing connections while being IP address agnostic. This means that the SDP is aware of IP addresses but doesn't rely on them for authorizing access to protected resources.

2.1.3 Integrating Security Controls

The integration of multiple security controls like firewalls and identity managers is typically implemented to achieve compliance. However, integrating these controls to work as a whole in protecting the application infrastructure can be challenging. Currently, the integration of controls may be performed by gathering data in an SIEM for analysis; however, correlating disparate streams of security to gain deeper insights (e.g., who is connected, from what device, from where, to what, and more) is resource intensive.

A single point of trust for network connections requires the following:

> - Information about users, provided by the applications
> - Information about the network, provided by firewalls
> - Information about devices, provided by the client

These disparate requirements make it difficult to implement an integrated set of controls for a physical network. Furthermore, integrating identity management prior to allowing access through a firewall requires the routing of packets to a different service — one that is resource-intensive and may or may not be proxied. In addition, most DevOps teams consider application layer firewalls and anti-denial of service/distributed denial of service (DoS/DDoS) protection as an afterthought; moreover, allowing individual applications to control their own security posture may result in catastrophe. Integrating access control, identity management, session management, and firewall management in today's environments is highly difficult; SDP addressed this challenge by providing a unified location for implementing and managing controls for the entire environment, versus using traditional distributed controls.

2.2 Threats SDP Protects Against

In this section, we will analyze the efficacy of SDP for reducing cyber risk and mitigating threats. We will present well-known threats/cyber risks published by the OWASP, Verizon, and CSA that demonstrate the real value of ZTA using the SDP. As illustrated below, the integration of SPA and SDP with enterprise IAM helps raise the bar for security. The tables provide a high-level overview of the relevant risks/threats, results of a successful exploit execution, and how SDP can be leveraged to prevent these security incidents from occurring.

2.2.1 CSA's Egregious 11

Risk/Threat	Result(s) of Successful Exploit Execution	SDP Mitigation
Data breaches	Reputational damage, loss of customer/partner trust, loss of intellectual property to competitors which may impact product releases, regulatory implications that may result in monetary loss, brand damage/ market value loss, legal and contractual liabilities, and financial expenses incurred due to incident response and forensic analysis	SDP has a drop-all firewall that drops packets not explicitly configured, preventing data breaches and/or preventing their scope of damage.
Misconfigurations & inadequate change control	Exposure of data stored in cloud repositories	SDP assists in change control by providing access configured for changes only after approval.
Lack of cloud security architecture & strategy	Financial loss, reputational damage, legal repercussions, and fines	SDP has a ZT policy that outlines a framework with systems designed around the value of the data and its specific protection needs.
Insufficient identity, credential, access, & key management	Unauthorized access, exfiltration, modification, deletion of data, issuing of control plane and management functions, eavesdropping on data in transit, and the release of malicious software that appears to originate from a legitimate source	Authentication, authorization, and mutual factor authorization (MFA) is at the core of SDP; subsequently, using SDP integrated with enterprise and cloud IAM/ identity provider (IdP) reduces the attack surface.

Account hijacking	Complete deletion of organization assets, data and capabilities, data leaks and resulting brand/reputational damage, legal liability due to sensitive personal and business information exposure	Authentication, authorization, and MFA is at the core of SDP; using SDP integrated with enterprise and cloud IAM/ IdP limits the exposure for account hijacking.
Insider threat	Loss of proprietary information and intellectual property, system downtime impacting company productivity, and other customer data losses that reduce their confidence in the organization's services	SDP includes micro-segmentation of the organizational environment to ensure that access to resources are granted on a need to know basis. SDP's continuous logging integrated with user entity behavior analytics can limit the data loss and/or alert on malicious/ abnormal activity and behavior.
Insecure interfaces & APIs	Regulatory and financial impact in the form of fines/penalties, security issues related to confidentiality, integrity, availability and accountability	SDP provides controls defining communication endpoints (as long as the interface and API communications sit behind the SDP controller).
Weak control plane	Data loss, either due to theft or corruption, resulting in a substantial impact on the business — particularly if the incident includes privileged user data, and regulatory punishment for data loss may be incurred	SDP's control plane is protected by both network level controls (e.g., SPA), and strong authentication.
Metastructure/application infrastructure failures	Failures at the cloud service provider level, resulting in customers being severely impacted, and tenant misconfigurations could result in financial losses and operational disruptions	SDP limits the impact of misconfigurations by hiding resources behind the gateway/ controller.

Limited cloud usage visibility	Lack of governance, awareness, and security	SDP logs all inbound activity, providing better visibility and situational awareness.
Abuse of cloud services	Financial losses due to excessive metered cloud use (e.g., attackers using compromised cloud servers as a malware distribution host)	SDP safeguards access to stateful (e.g., security group/ network security group) and stateless (e.g., access control list/network access list) firewall configurations. Coupling security group/ network security group and access control list/ network access control list configurations enables the dropping of unauthorized traffic (e.g., for mining cryptocurrency or distributing malware).

Table 1: Top Threats to Cloud Computing: Egregious Eleven Deep Dive[8]

2.2.2 Verizon's DBIR

Risk/Threat	Result(s) of Successful Exploit Execution	SDP Mitigation
Phishing & social engineering	Acquisition of credentials	SDP's integration with domain-based, message authentication/reporting, as well as its requirements for validating source networks and capabilities (e.g.,MFA and device fingerprinting) reduce the risk of these attacks.

[8] Cloud Security Alliance, "Top Threats to Cloud Computing: Egregious Eleven Deep Dive," 23rd, September 2020, https://cloudsecurityalliance.org/artifacts/top-threats-egregious-11-deep-dive/

Web application attacks	Stolen credentials and successful brute force attempts that enable unauthorized access to web application servers, mail servers, and others IT assets, resulting in compromised privileged data (e.g., medical records, employee data)	SDP's use of SPA for inbound/outbound server traffic coupled with MFA helps prevent these attack types (i.e., requiring authentication prior to authorizing access).
Lost or stolen credentials	Exposure of sensitive data	SDP's MFA requirement minimizes the impact of stolen credentials, since malicious actors are not given explicit access to resources.
Ransomware	Revenue loss and supply chain disruption	SDP prevents the installation of unapproved software and potentially malicious applications (e.g., ransomware) on servers.
Miscellaneous errors compromising security	Eavesdropping, data loss, data exposure, and unauthorized access	SDP requires authentication prior to accessing applications and/or server resources.
DoS	Loss of service and/or service disruption	SDP controls communication endpoints and is therefore stateless; drop-all firewalls block threats such as malware and command and control servers.
System intrusion	Eavesdropping, data loss, data exposure, and unauthorized access	SDP requires the use of SPA to/from the server. Coupled with MFA, these controls help to enforce authentication prior to authorized access.
Privilege abuse	Eavesdropping, data loss, data exposure, and unauthorized access	SDP's MFA requirement prevents unauthorized access and escalation of privileges from occurring.

Table 2: DBIR- 2021 Data Breach Investigations Report[9]

[9] Verizon, "2021 Data Breach Investigations Report," 2021, https://enterprise.verizon.com/resources/reports/2021-data-breach-investigations-report.pdf

2.2.3 OWASP IoT Top 10

Risk/Threat	Result(s) of Successful Exploit Execution	SDP Mitigation
Weak, guessable, or hard coded passwords	Unauthorized access	SDP authenticates users prior to granting them access; additionally, MFA helps mitigate the risk of stolen/lost credentials and devices.
Insecure network services	Unauthorized access	SDP requires encryption for enforcing confidentiality in an insecure network. For example, devices must support encryption in order for SPA over HOTP (HMAC One Time Password) and data communications via mTLS to function (in protecting confidentiality).
Insecure ecosystem interfaces	Unauthorized access	SDP unifies the different ecosystem interfaces into a secure, single source of truth.
Lack of secure update mechanism	Unauthorized access	SDP and SPA provide device authentication and valid endpoints via mTLS, allowing for secure over-the-air authentication and device update mechanisms.
Use of insecure or outdated components	Eavesdropping, data loss/ exposure, and unauthorized access	SDP leverages ZT call flows in the TCP/IP network, thereby protecting legacy, insecure or outdated components.
Insufficient privacy protection	Eavesdropping and data loss/ exposure	SDP requires encryption in order for SPA over HOTP to function and ensure confidentiality.

Insecure data transfer & storage	Eavesdropping and data loss/ exposure	SDP requires encryption in order for SPA over HOTP and data communications via mTLS to function and ensure confidentiality.
Lack of device management	Unauthorized access	SDP provides secure mobile device management by enabling device management and software updates via SPA and mTLS.
Insecure default settings	Unauthorized access	SDP requires micro-segmentation as well as MFA to mitigate the risk of outdated/unpatched and misconfigured devices.
Lack of physical hardening	Unauthorized access	SDP couples automated device auditing with secure device management to validate device security postures.

Table 3: OWASP IoT Top 10[10]

2.2.4 OWASP Top 10

Risk/Threat	Result(s) of Successful Exploit Execution	SDP Mitigation
Broken access control	Unauthorized access	SDP authenticates users and validates that requests are authorized prior to granting access.
Cryptographic failures	Exposure of sensitive data or a compromised system	SDP enforces cryptography requirements (e.g., in mTLS sessions).
Injection	Malicious injection and alteration of responses to compromised application server	SDP mitigates application attacks through its inherent MFA and SPA/drop-all approach.

[10] OWASP, "OWASP IoT Top 10," 2018, https://owasp.org/www-pdf-archive/OWASP-IoT-Top-10-2018-final.pdf

Insecure design	Exploitation of vulnerabilities	SDP helps bolster threat modeling, secure design principles, patterns, and design practices affecting reference architectures and configuration audits.
Security misconfiguration	Exposure of data and exploitation of application vulnerabilities	SDP mandates micro-segmentation and MFA — critical SDP features for mitigating the impact of security misconfigurations. MFA limits the escalation of privileges and reduces the blast radius of attacks.
Vulnerable & outdated components	Exploitation of known vulnerabilities	SDP prevents legacy, insecure, or outdated components from being attacked by hiding the associated services from unauthorized users/devices.
Identification & authentication failures	Privileged access escalation and lateral movement	SDP mandates micro-segmentation and the granting of access to resources based on the requester's need to know/ need for access.
Software & data integrity failures	Insertion of malicious code into critical path continuous integration/continuous delivery (CI/CD) pipelines (e.g., open source software, containing malicious code)	Leveraging the SPA, SDP uses endpoint authentication to assist with verifying the integrity of CI/CD pipelines and software updates.
Security logging & monitoring failures	Lack of visibility into unauthorized or malicious events	SDP enforces comprehensive and continuous monitoring. With logging/monitoring services in place per SDP's requirements, security incidents can be remediated in a timely manner.

Server side request forgery	Unauthorized access and compromise of vulnerable applications and related/ connected back-end systems. Attackers may also use this exploit method to circumvent user input validation	SDP helps mitigate attacks to applications exposed on a network through its inherent MFA and SPA/drop-all approach.

Table 4: 2021 Draft OWASP Top 10[11]

In addition, the following sections address some of the various threats that SDP helps protect against. These include server exploitation and hijacking, among others.

2.2.5 Server Exploitation Threats

SDP features like server isolation, SPA, and dynamic drop-all firewalls bolster application infrastructure security and help protect against server exploitation threats such as the following:

- DoS/DDoS attacks
- Code injection attacks
- Other attacks that exploit server misconfigurations/vulnerabilities

2.2.6 Hijacking Threats

SDP attributes such as encryption, pinned certificates, and non-reliance on DNS protect against connection hijacking threats like the following:

- Man-in-the-middle (MITM) attacks
- Certificate forgery
- DNS poisoning
- Code injections

2.2.7 Other Threats

SDP features like MFA, mTLS, and device fingerprinting protect against the following:

- Phishing
- Keyloggers
- Brute force attacks

For further information, please refer to CSA's *SDP Architecture Guide*[12] and existing research on SDP and ZT[13].

[11] Footnote 12: OWASP, "OWASP Top 10," 2021, https://owasp.org/Top10/
[12] Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/
[13] Cloud Security Alliance, "Software-Defined Perimeter (SDP) and Zero Trust," 27th, May 2020, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-and-zero-trust/

2.3 SDP & Industry Adopted Solutions

In this section, you will learn about various industry adopted solutions and how SDP replaces or works in conjunction with them. This includes NAC, virtual private networks (VPNs), IAM, and next generation firewalls (NGFW).

2.3.1 Network Access Control

NAC typically controls what devices can connect to a given network and which network locations or segments they have access to. These solutions use a combination of standards-based hardware (e.g, 802.1X for port-based NAC) and software to validate devices, prior to granting them network access. NAC typically operates at layer 2 (i.e., the data link layer) of the OSI model.

When a device first appears on the network, the NAC performs device validation followed by assignment to the correct network segment (e.g., virtual local area network). In practice, NACs coarsely assign devices to a small number of networks, as most organizations only have a few networks set up (e.g., guest, employee, and production). Because NACs operate at layer 2 of the OSI model, they more often require specific network equipment, don't operate in cloud environments, and are not used by remote users.

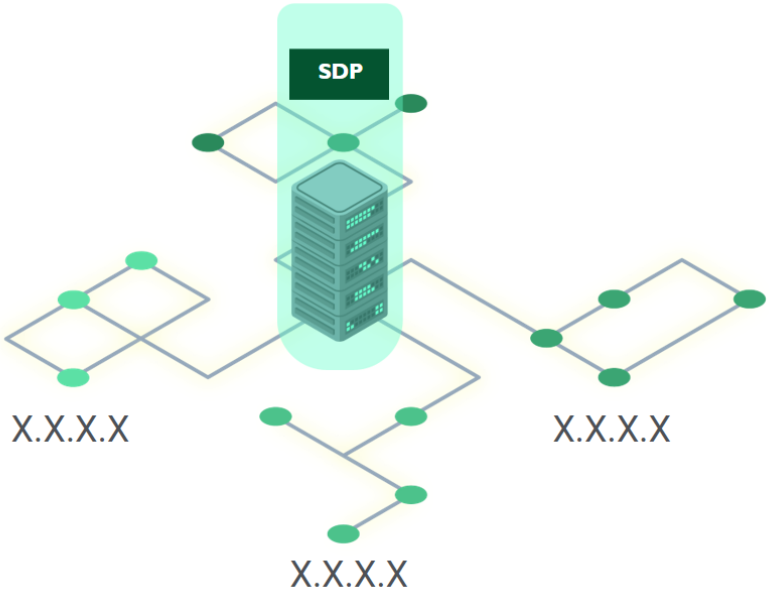


Figure 4: *SDP as NAC Replacement*

In some respects, SDP can be considered a modern replacement for NAC. Though they share similar functionalities, SDP, unlike NAC, does not require specific network hardware to function. This allows for the integration of users and provisioning of device access without a dedicated network appliance. SDP fully supports cloud environments and remote access, overcoming traditional NAC limitations. However, some environments are more suitable for NAC implementations — for example, those with printers, copiers, landline phones, or security cameras. These devices are often 802.1X compliant with built-in support, which means they don't typically support the installation of an SDP client. In this case, the gateway-to-gateway model is a better option for protecting and managing access to these devices.

2.3.2 Virtual Private Network

VPNs establish secure private network connections over untrusted networks. Commonly used for secure remote access (e.g. employee access to a corporate site, secure site-to-site communications, or site-to-site extranets between companies), VPNs use TLS/Secure Sockets Layer (SSL) or Internet Protocol Security (IPSec) to establish an encrypted tunnel.

Although VPNs encapsulate and encrypt network traffic, they also allow unrestricted access to a network segment. This is risky, especially if credentials are compromised. In contrast, SDP will only allow access to specifically assigned applications in the network segments.

On the user experience side, VPNs tend to impose a considerable burden on users, especially in environments undergoing significant cloud-based transformations and migrations. IT may also need to configure VPN for users requiring secure access to multiple sites, as this prevents unintentional network bridging and systems from connecting to multiple locations simultaneously. Ultimately, this shifts the burden and inconvenience of switching back and forth between remote locations on the user.

1. In distributed environments, VPNs may require the unnecessary backhauling of user traffic through a corporate data center, adding latency and bandwidth costs.
2. VPN servers themselves expose the network on the internet. VPN servers contain security vulnerabilities as do most IT components, which an attacker could exploit to gain access and exfiltrate data or perform other malicious activities.
3. VPN licensing costs are not expensive, but anecdotally they can be difficult to implement and maintain. Whenever cloud migration is involved, VPN management balloons in complexity. This is because IT administrators need to configure and sync VPN and firewall policies across multiple locations,making it even more difficult to mitigate unauthorized access.

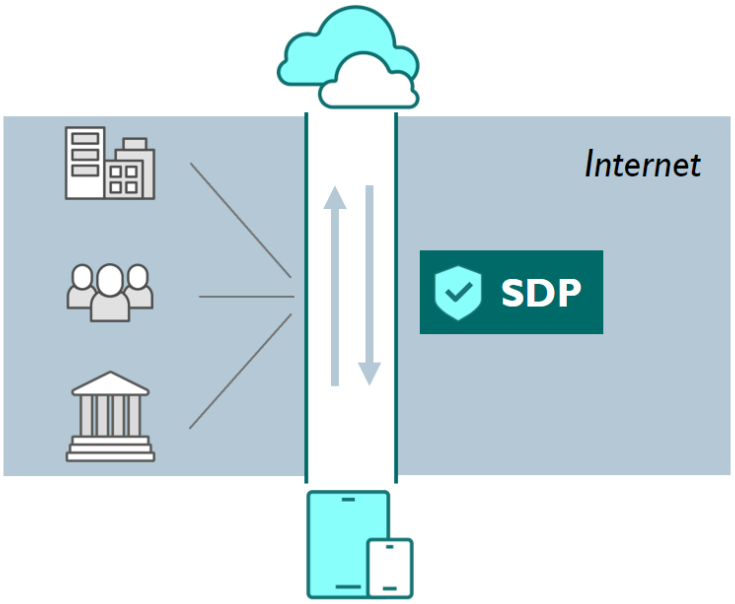


Figure 5: *SDP as VPN Replacement*

VPNs are a prime technology use case for replacement by SDP — however, it's worth noting that SDP can work alongside existing VPNs or replace them entirely, depending on the deployment model. However, both require an installation of a client on the user's device. By using SDP instead of VPNs, organizations can have a single access control platform consistent for secure access to cloud, remote, on-premises, and mobile device users. Since SDPs enable zero visibility via SPA and dynamic firewalls, they are considerably more resilient to cyber attacks than traditional VPN servers.

2.3.3 Identity & Access Management

The SDP architecture is designed to integrate with existing enterprise IAM providers in the cloud, on-premises, or hybrid environments. IAM provides a unified mechanism for users and devices to be validated, authenticated, and authorized. It provides a way to store managed identity attributes and group memberships within a central system using protocols to enable access directly or via federation. SDP supports standard protocols and security mechanisms used by IAM, including Lightweight Directory Access Protocol (LDAP), Active Directory (AD), and Security Assertion Markup Language (SAML).

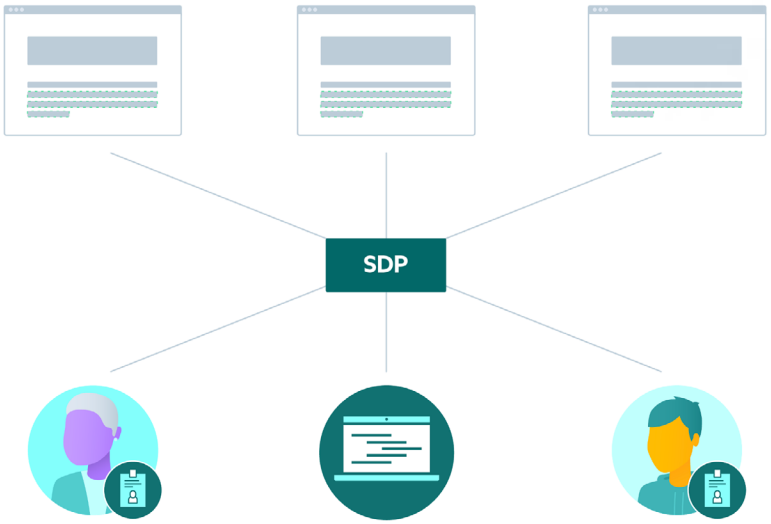


Figure 6: *SDP and IAM*

SDP typically controls access based on business rules. These rules can be built up from IAM attributes and group memberships, as well as from the attributes of devices making the connection and/or the network segments themselves. The telemetry data provided by these sources enables the creation of granular access rules for allowing/restricting access. This ensures only users requesting specific access on registered devices are granted authorization to the resources in question.

Integration of SDP with IAM is not only used for initial user authentication; it's also commonly used in conjunction with step-up authentication (e.g., prompting for a one-time password to access sensitive resources). IAM systems can also communicate with an SDP via API calls, as SDP can respond to identity lifecycle processes in this configuration. Some examples include joiners, mover, and leavers (JML), disabling an account, group membership changes, and dropping user/device connections from certain geographic locations.

In order to authenticate users, SDP must leverage IAM to access identity telemetry information that the SDP controller uses to make authorization decisions. IAM data is not only used to augment the SDP controller's capabilities — it's also used for populating audit logs with additional details regarding user and device access (e.g., access granted/denied details). Compared to traditional network access and IP address information, IAM telemetry correlates application access to users, yielding far more useful data with less overhead. This reduced overhead is leveraged primarily by IT when auditing historical access records in security or compliance use cases.

2.3.3.1 SDP & Identity Lifecycle Management

In identity lifecycle management, IAM tools focus on the business processes for maintaining the identity lifecycle (i.e., the JML process). IAM standardizes how identity information is used to control access to resources, using access methods such as role-based and attribute-based access control.

SDP supports these IAM processes and relies heavily on IAM-managed identity attributes and group memberships. As user attributes or group memberships change, SDP will alter access permissions accordingly without changing IAM telemetry, as SDP is a downstream system. These processes are utilized by SDP via standard protocols like SAML, AD, LDAP, or through the use of APIs.

2.3.3.2 SDP & Open Authentication Protocols

SDP integrates with open authentication protocols such as SAML. Within an SDP deployment, a SAML entity might act as an identity provider for user attributes and/or as an MFA authentication provider.

In addition to SAML, SDP integrates with many other open authentication protocols such as OAuth, OpenID Connect, W3C Web Authentication, and the FIDO Alliance Client-to-Authenticator Protocol. These protocols will be explored in future SDP-related research, but are not in scope for this training.

2.3.4 Next Generation Firewall

NGFWs have all the capabilities of traditional firewalls, along with additional capabilities such as intrusion detection/prevention and deep packet inspection. NGFWs filter data using the information in layers 2 through 4 of the OSI model (i.e., the data-link, network, and transport layers). Additionally, NGFWs use the information in layers 5 through 7 (i.e., the session, presentation, and application layer) to perform additional functions.

Depending on the vendor, NGFWs may provide some or all of the following capabilities:

- Application awareness – recognizes applications to determine what attacks to look for
- Intrusion detection/prevention system (IDPS) – monitors the security status of the network and denies traffic to prevent security problems
- Identity awareness (user and group control) – controls which resources users can access
- VPN – allows for remote user access across an untrusted network

While NGFWs represent a significant improvement over traditional firewalls, they still have their limitations. Some of these include:

- Latency – as is the case with IDPS, NGFWs will cause additional network latency, especially if they're performing file inspection
- Scalability issues – a NGFW requires more robust hardware to scale
- Rule complexity – some NGFW vendors include identity management capabilities such as user/group attribute assignments, but anecdotally these tend to be complex to implement

SDP is a natural complement to existing NGFWs. Enterprises can use SDP for secure user access policies while leveraging their NGFWs for core firewall, IDPS, and traffic inspection capabilities. By integrating SDP with a NGFW, enterprises can at once enforce the zero visibility principle and make them more dynamic.

User access policies can be achieved by integrating NGFWs with IAM or AD. By combining NGFW VPN capabilities with user and application awareness, enterprises can, to some degree, accomplish many of the goals of SDP. However, there are some general architectural differences.

NGFW systems are IP-based and offer limited identity and application-centric capabilities, whereas SDP is connection-based and therefore easier to control authorized connections. Additionally, NGFWs tend to be much less dynamic than SDPs, while the latter often supports the ability to include external systems in access decisions. For example, a prime use case for SDP is to only permit developer access to staging servers during an approved change management window.

Since NGFWs are still firewalls, their network deployment/design patterns still favor traditional perimeter-centric network architectures with site-to-site connections between locations. On the other hand, SDP deployments usually support more distributed and flexible networks, thereby enabling a flexible network segmentation capability.

SDP is fundamentally based on a need to know security principle, which by design hides all unauthorized services from users and leverages SPA and dynamic firewalls to hide connections protected by the SDP. NGFWs are not designed to function this way and typically result in environments that are more visible and therefore higher risk than with SDP. It should be noted that NGFWs have not yet been able to integrate authentication and authorization controls prior to allowing connections.

3 Core Tenets, Underlying Technologies, & Architecture

In this unit you will learn the foundation of how SDP works and how it accomplishes its task of providing network security. We will explain SDP's core tenets, underlying technologies, architectural components, and secure workflow.

3.1 SDP Core Tenets

SDP has three core tenets that govern its implementations: assume nothing, trust no one or thing, and validate everything. These tenets are used as the building blocks of the SDP framework. SDP was designed to secure dynamic workloads, most prominent in cloud mobile environments, by providing the following:

- Software-defined, dynamic, endpoint validation
- A connection-based paradigm
- Integration of firewalls, identity and access, session, encryption, and device management

These design features are covered in CSA's *SDP Specification v2*[14].

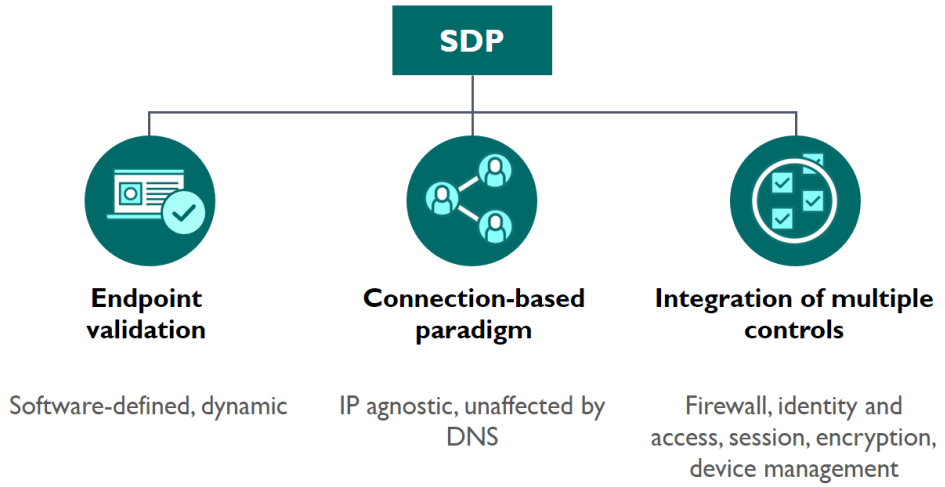


Figure 7: *SDP Core Tenets Tree*

[14] Figure adapted from Cloud Security Allaince, "Software-Defined Perimeter (SDP) Specification v2," 10th, March, 2022, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/

3.2 Underlying Technology

In this section, we will introduce and discuss the underlying technologies that support the SDP architecture, including drop-all firewalls, separate control and data planes, mTLS, and SPA. SDP provides a security architecture designed from the ground up using these foundational technologies.

3.2.1 Drop-All Firewall

The drop-all firewall is the most critical underlying technology of the SDP. Using drop-all rules, these firewalls operate according to the principle of least privilege: all actions not explicitly allowed remain forbidden. This strategy supports the ability to add rules to the firewall dynamically for post-authentication access level changes.

An SDP deployment can identify and deny risky transactions based on the analysis of a single packet. When malicious actors attempt to connect, SDP uses a drop-all rule to drop all unauthorized packets at the perimeter. This approach is highly effective as it only focuses on allowing approved actions, rather than blocking unapproved actions.

3.2.2 Separate Control & Data Planes

The three basic components of an SDP architecture are the data plane, control plane, and management plane. The data plane — also known as the user plane, forwarding plane, carrier plane, or bearer plane — is the part of a network that carries user traffic. In SDP, the control plane and management plane enable the data plane, which bears the traffic that the network carries. The control plane is responsible for establishing connections and dropping unauthorized packets at the perimeter. The control plane takes care of authenticating and authorizing users and devices prior to sending data to the SDP gateway. No devices are allowed to reach the data plane until the user and device in question are validated at the control plane.

In traditional architectures, data and control planes are commonly implemented together. In contrast, SDP architectures place the control plane outside the organization's perimeter; subsequently users and devices do not enter the organization's environment until they are authenticated and authorized. By separating the data plane from the control plane, SDP enables the external control plane to perform authentication and authorization before granting access to resources.

3.2.3 Mutual Transport Layer Security

SDP uses mTLS authentication to ensure that client-server traffic is secure and trusted in both directions. This allows requests that do not log in with an identity provider (e.g., requests from IoT devices) to demonstrate that they are permitted access to a given resource. Client certificate authentication adds an additional security layer for team members who both log in with an identity provider and use a valid client certificate for authentication.

Chiefly, mTLS is ideal for use in the following IT environments:

- A limited number of programmatic and homogeneous clients connect to specific web services
- The operational burden is limited
- Security requirements are more stringent compared to consumer environments

Subsequently, mTLS authentication is more widely used in business-to-business applications.

3.2.4 Single Packet Authorization

SPA is a protocol that allows a user to make a request to a server. This request cannot be replayed and uniquely identifies the user. SDP uses SPA to compensate for the fundamentally open and insecure nature of TCP/IP. In addition, SDP uses SPA to authorize a valid device and authenticate a user identity. SPA then permits access into the perimeter and the relevant system component. The purpose of SPA is to allow assets within the perimeter to be restricted via a default drop-all firewall.

While implementations of SPA may differ slightly, they should share the following common concepts for an SDP implementation:

- An SPA packet must be encrypted and authenticated
- An SPA packet must self-contain all the necessary information
- Packet headers are not considered trustworthy
- A SPA packet must not depend on administrator or root level access in order to generate and send
- There is no raw packet manipulation
- The server must receive and process the SPA packet as silently as possible, no response or verification is sent

3.2.4.1 SPA Benefits

The key advantage of using SPA is service restriction. A default drop-all firewall posture prevents port scanning and other attacker-related reconnaissance techniques. It effectively renders the SPA components invisible to unauthorized users, significantly reducing the attack surface of the SDP system. This compares favorably to systems such as VPNs, with open ports and known vulnerabilities in many implementations.

There are subsequent benefits to restricting services. One is zero-day protection. Any newly discovered vulnerability becomes significantly less critical when only authenticated users can access the affected service. Another benefit is DDoS protection. A relatively small amount of traffic can take an HTTPS service offline if that service is exposed to the public internet for attack. A SPA makes that service visible only to authenticated users. Therefore, a DDoS attack is handled by a default drop-all firewall instead of the protected service itself.

One of the core goals of SDP is to overcome the fundamentally open, or insecure nature of TCP/IP, which follows a connect, then authenticate model. Amid today's threat landscape, it's simply

unacceptable to permit malicious actors to scan and connect to enterprise systems. There are far too many known and unknown vulnerabilities in systems to allow this. SPA and SDP solve this problem in two ways. First, applications using the SDP architecture are hidden behind an SDP gateway so that they're only accessible to authorized users. Second, the SDP components themselves, the controller and gateway, are protected by SPA. This allows them to be securely deployed with internet-facing placement, ensuring that legitimate users have productive and reliable access, while they remain invisible to unauthorized users.

3.2.4.2 SPA Limitations

SPA is only a part of SDP and is not a complete security architecture on its own. While SPA implementations should be designed to be resilient to replay attacks, SPA may be subject to a MITM attack; specifically, if MITM adversaries are able to capture or alter the SPA packet, they can potentially establish the TCP connection to the controller or accepting host (AH) in place of the authorized initiating host (IH). However, these adversaries will be unable to complete the mTLS connection, since it will not have the client's certificate. The controller or AH should therefore reject this connection attempt and close the TCP connection. Even considering this limitation, which only applies to the MITM scenario, SPA is more secure than standard TCP.

3.3 SDP Architecture Components

In this section, we will discuss the foundational SDP architecture components: IH, AH, gateways, SDP clients, and the controller. SDP provides an integrated security architecture that is otherwise hard to achieve with security point products.

SDP integrates the following discrete architectural elements:

- Identity-aware applications
- Client-aware devices
- Network-aware firewalls/gateways

3.3.1 Initiating Hosts

IH initiate connections to the SDP. IH are devices, including laptops, tablets, and smartphones that SDP client software is run on. This host environment may be on a network outside the control of the enterprise operating the SDP.

3.3.2 SDP Client

The SDP client consists of software installed on the IH device. The client initiates connections in order to cryptographically sign in to the SDP. The SDP client typically generates the SPA packet for the SDP gateway after completing the authentication and authorization process with the SDP controller.

3.3.3 Accepting Hosts

AH are devices that accept connections from IH and provide a set of services that are protected by the SDP. They typically reside on a network under the control of the enterprise (and/or direct representative) operating the SDP, and do not acknowledge communications from any other host or respond to non-provisioned requests. To unauthorized users and devices, AH remain cloaked and inaccessible while using SDP's SPA.

3.3.4 Controller

The SDP controller is an appliance or process that secures access to isolated services. It does this by ensuring that users are authenticated and authorized, devices are validated, secure communications are established, and user and management traffic on a network remain separate. Like the AH, the controller is also protected by SPA, making it invisible and inaccessible to unauthorized users and devices. Both IH and AH connect to the SDP controller.

3.3.5 Gateway

The SDP gateway is an appliance or process that provides access through the invisible perimeter for authorized users and devices. Through this gateway, authorized users and devices are able to access protected processes and services. The gateway can also effectively allow monitoring, logging, and reporting on these connections. The functionality of the gateway depends on where it is located.

3.4 SDP Secure Workflow

In this section we will break down SDP's workflow, illustrating how all of the architecture components discussed in the previous section work together.

The following is the most basic SDP workflow for allowing an IH and AH to communicate securely:

1. The AH is cloaked by an SDP gateway on the AH or a similar construct.
2. An SDP controller is added and activated within the SDP and connected to authentication and authorization services (e.g., IAM, public key infrastructure service, device attestation, geolocation, SAML, OpenID, OAuth, LDAP, Kerberos, MFA, and identity federation).
3. An AH is added and activated within the SDP by checking into the SDP controller. It connects to and authenticates with the controller in a secure manner.
4. The IH is added and activated within the SDP, then connects to the SDP controller. The SDP controller authenticates the IH and determines a list of AH the IH is authorized to communicate with.
5. An SPA packet is always sent to establish communications, leaving the application layer cloaked from all but authorized users. In order to establish access after sending an SPA packet, the IH and AH exchange a mutual handshake using TLS for control plane communications.
6. The IH sends a login message request and receives a response from the controller.
7. The controller sends the IH a list of services available (based upon services allowed).

8. The controller also sends a message stating that the IH has been authenticated with the AH.
9. Another SPA packet is sent from the IH to the AH for data plane communications.
10. Finally, a separate mTLS handshake establishes communication for data transfers.

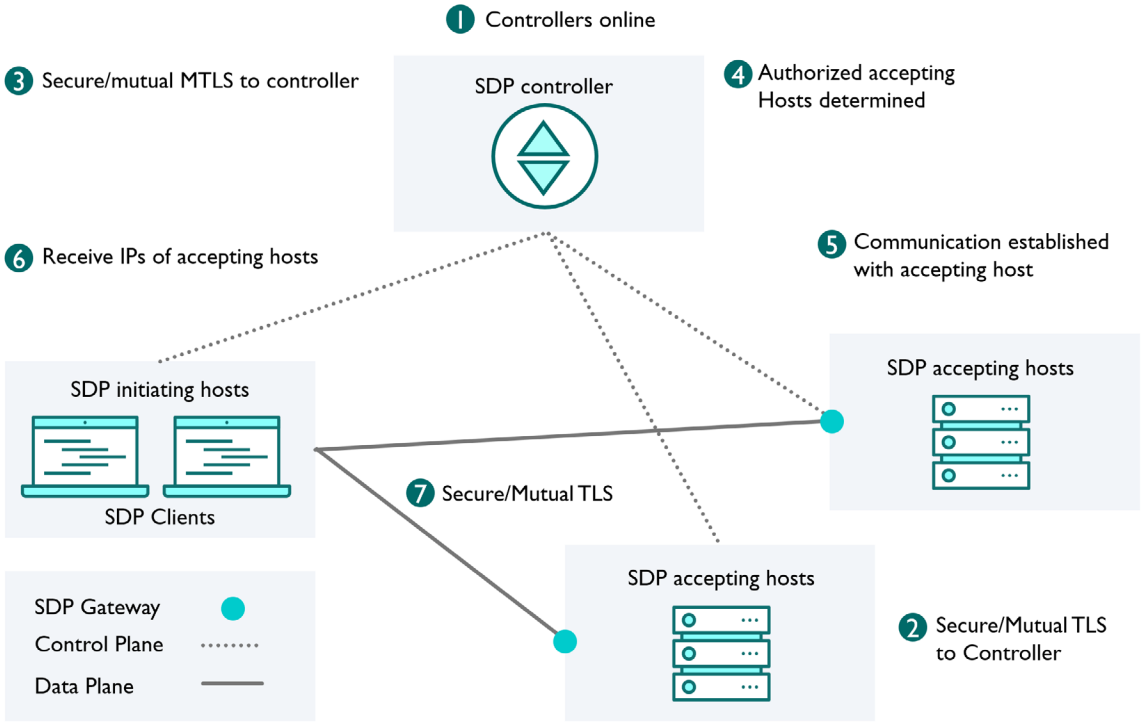


Figure 8: *SDP Secure Workflow*[15]

4 The Basics of SDP Deployment Models

This unit will cover the various SDP architectural considerations to take into account before implementation. Along with this, you will learn the basics of SDP deployment models.

4.1 Architectural Considerations

Several architectural considerations must be taken into account when deploying SDP. For example, organizations should evaluate how an SDP deployment fits into existing network topologies and technologies. Other critical considerations include how SDP impacts users, monitoring, logging, onboarding, application release, and device validation.

[15] Figure adapted from Cloud Security Allaince, "Software-Defined Perimeter (SDP) Specification v2," 10th, March, 2022, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/

4.1.1 Existing Network Topologies & Technologies

Network architects should select the SDP deployment model best suited for their particular use case. However, some models require additional in-line network components like gateways, resulting in network changes like adding firewalls or making routing alterations. This ensures that protected resources are hidden and only accessible through the SDP gateway. To fully leverage the capabilities of SDP, architects should consider proper micro-segmentation, keeping in mind that SDP ensures secure connections irrespective of the underlying network infrastructure.

Enterprise security architectures[16] can be complex, with numerous stakeholders across the organization and business units, as well as governance, risk management, and compliance (GRC) requirements alongside daily IT infrastructure operations and management. Architects should keep these factors in mind when planning their enterprise's SDP deployment.

4.1.2 Monitoring & Logging Systems

SDP affects monitoring and logging architectures. Because it uses mTLS between the IH and AH, SDP also hides network traffic from intermediary services, which may be in place to monitor for security, performance, or reliability. Architects must understand what systems are in operation and how the changes to the network traffic may affect them. However, SDP typically provides richer, identity-centric logging of user access — ideal for augmenting and enhancing existing monitoring systems for a more focused traffic monitoring scope and purpose. In addition, all dropped packets from SDP gateways and controllers can be logged, monitored, and analyzed using security tools like intrusion detection systems/intrusion detection and prevention systems (IDS/IDPS) and SIEMs. With an SDP in place, it is easier to collect the who, what, when, how, why information for every connection versus each individual packet.

4.1.3 Application Release & DevOps

High-velocity application release practices like DevOps[17] and its supporting automation and CI/CD framework require thoughtful integration with SDP. An SDP can be integrated with DevOps to secure authorized users' connections to the various deployment environments (e.g., development, test, staging, and production), as well as used during operations to ensure legitimate users have proper connectivity to protected servers and applications. Ideally, the SDP will be integrated into the application stack to fully leverage its security features. Common DevOps practices such as the use of virtualized environments and containers can further streamline SDP integration; that said, security architects must fully understand the chosen SDP deployment model and how their organization's DevOps mechanisms will interact and integrate with it. When it comes to DevOps toolset integration, security teams should carefully review and evaluate third party APIs supported by their SDP implementation.

[16] Sometimes referred to as enterprise information security architecture

[17] Cloud Security Alliance, "Enterprise Architecture Reference Guide," 18th, May 2021, https://cloudsecurityalliance.org/artifacts/enterprise-architecture-reference-guide-v2/

4.1.4 User Experience

Security teams typically strive to have their solutions work as transparently as possible, with minimal user interruption. SDP is similar to any security control where proper application of least privilege principles balances the user experience with security. Depending on the SDP deployment model, users will need to run the SDP client software on their devices. Security architects should collaborate with IT to model and plan for the user experience, client software distribution, and device onboarding processes.

4.1.5 Onboarding

The onboarding process of SDP controllers, IH, AH, and users will vary depending on the chosen deployment models. SDP systems can be managed via an API or administrative user interface.

A typical SDP onboarding process flow would involve the following steps:

1. One or more SDP controllers are brought online and connected to the appropriate optional authentication and authorization services.
2. One or more AHs are enlisted as SDP gateways. These gateways connect to and authenticate with the controllers.
3. One or more clients on the IHs are onboarded, with each user/entity authenticated by the SDP controller.

Note: Because the onboarding process is distinct from the user authentication process, users are only onboarded once but will require authentication/authorization for each subsequent connection.

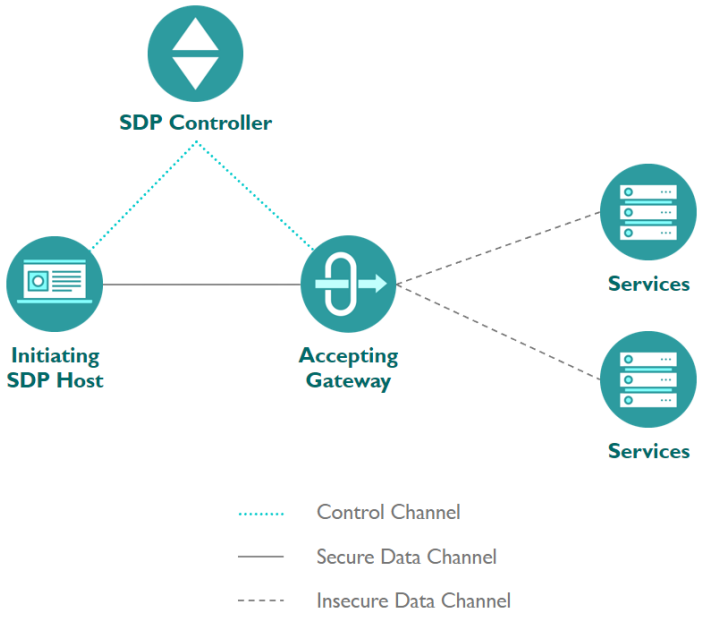


Figure 9: *Onboarding Process Flow*[18]

[18] Figure adapted from Cloud Security Allaince, "Software-Defined Perimeter (SDP) Specification v2," 10th, March, 2022, https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/

4.1.6 Device Validation

mTLS proves that the device requesting access to the SDP possesses a valid, non-expired/non-revoked private key. However, this method can be compromised, as an attacker with a stolen key cannot be distinguished from a legitimate user/key. Device validation can help to further establish a trusted connection based on certificate-based keys. Per SDP, the controller acts as the trusted device because it resides in the most heavily controlled environment. The initiating and AHs must then validate themselves with the controller, thereby preventing unauthorized access via stolen keys.

4.2 Deployment Models

In this section we'll introduce the various SDP deployment models and explore their similarities and differences.

As an architecture, SDP provides the protocol to secure connections at all layers of the network stack. By deploying gateways and controllers at key locations, SDP implementers can focus on securing and protecting the most critical connections from both network-based and cross-domain attacks. All the SDP models support identity-driven network access control/authorization, and most can accommodate existing network security tools like IDS/IDPS and SIEMs by enabling the analysis of dropped packets and unsecured connections. SDP secures the connections between components, as depicted in each of the deployment models described below.

More information on these deployment models can be found in the *SDP Architecture Guide v2*[19].

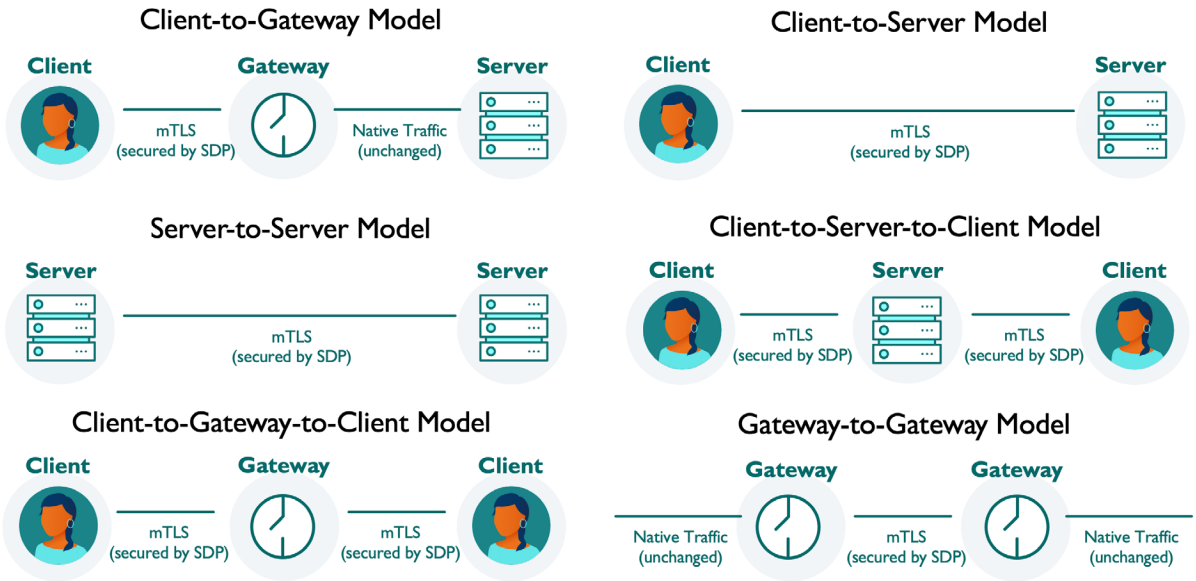


Figure 10: *SDP Deployment Models*[20]

[19] Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

[20] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

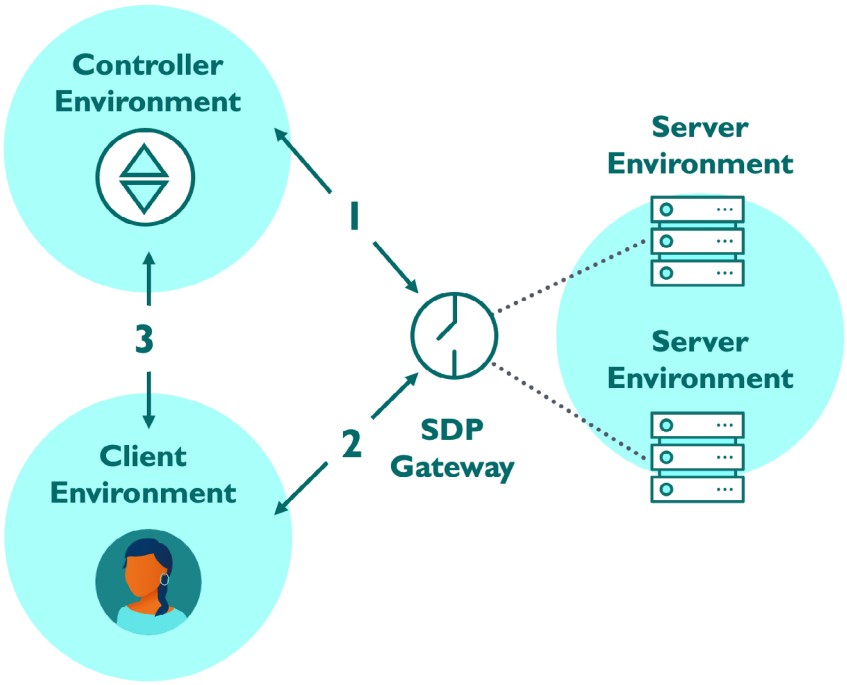


Figure 11: *Client-to-Gateway Model* [21]

4.2.1 Client-to-Gateway Model

The client-to-gateway model is suitable for use cases where one or more servers need to be protected behind a gateway. This approach is preferred when an organization is moving its applications to the cloud or securing on-premises legacy applications. The client (i.e., the IH) and gateway may be in the same location or distributed across the globe. In either case, the connections between the client and the gateway are secured, regardless of the underlying network topology.

In this model, the client is connected to the gateway directly via an mTLS tunnel where the connection terminates. To secure the connection to server environments, additional precautions must be taken. For example, the network on which the server environments reside, will need to be configured to permit inbound connections to protected servers from the gateway only. This prevents unauthorized clients from bypassing the gateway. The gateway should be configured to deny all traffic by default, and explicitly allow approved traffic. The same gateway can be used for the controller and servers by locating the controller in the cloud or near the protected servers.

This model preserves the ability for an organization to use its existing network security components, such as IDSs or IPSs, by deploying them between the SDP gateway and the protected servers.

[21] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

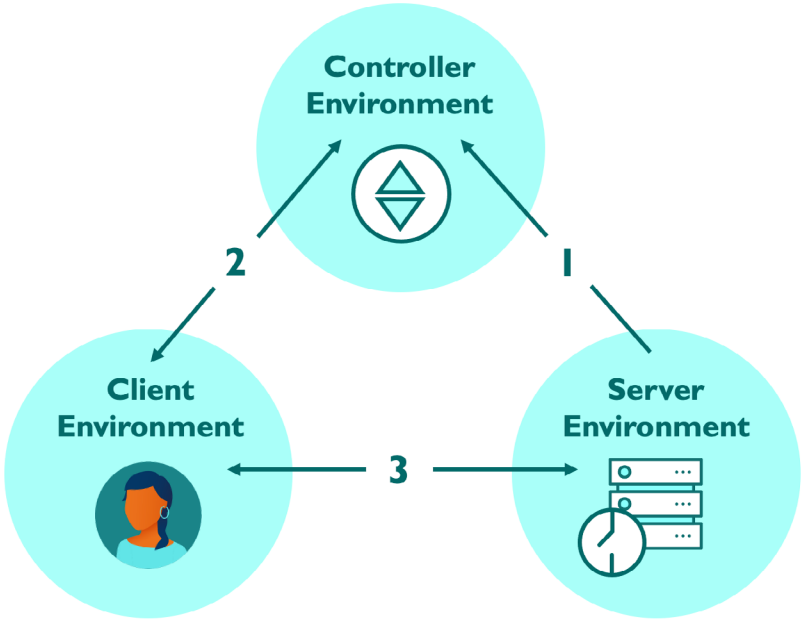


Figure 12: *Client-to-Server Model*[22]

4.2.2 Client-to-Server Model

The client-to-server model is ideal when moving applications to an IaaS provider, as it combines the server and gateway in a single host to ensure connections are secured end-to-end. Organizations are afforded a great deal of flexibility due to the portability of server-gateway combinations between multiple IaaS providers.

Client-to-server is also appropriate for securing on-premises legacy applications that cannot be upgraded. With this model, the protected servers will need to be outfitted with the gateways. The network on which the servers reside do not need configuration to restrict inbound connections to the protected servers, as the gateways or server enforcement points use SPA to prevent unauthorized connections. Secure connections to the servers provided by the gateway may be controlled by the infrastructure owner, as they have full control over the connections. Similar to the client-to-gateway model, the client may be located in the same location or distributed across the globe — in either case, it remains secured. Additionally, this model leaves the data plane completely secure, as there are no breaks in the mTLS tunnel. Traffic can be monitored by analyzing dropped packets from the SDP gateway/protected servers, thereby preserving the mTLS connections between the client and the servers.

[22] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

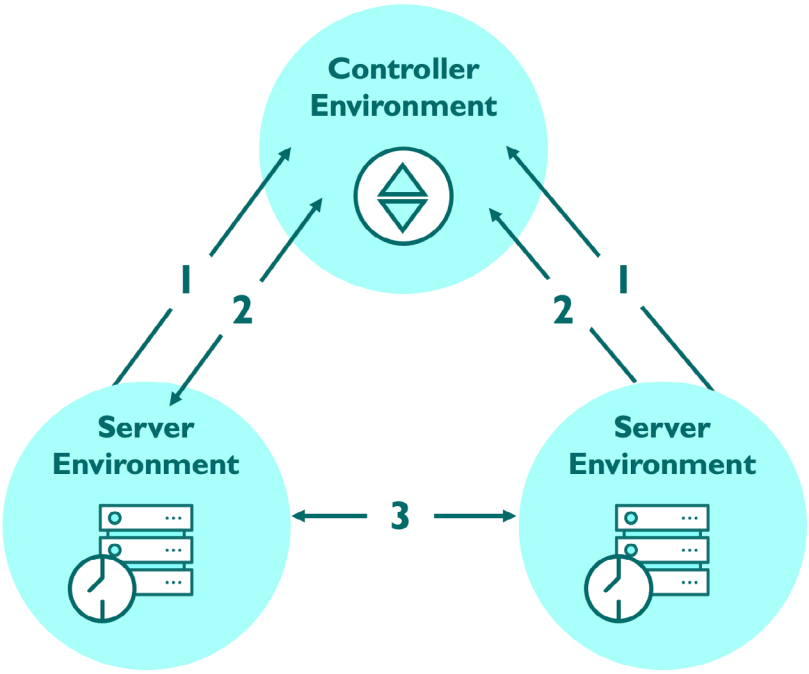


Figure 13: *Server-to-Server Model*[23]

4.2.3 Server-to-Server Model

The server-to-server model is ideal for IoT and virtual machine environments, as it offers full control over connections, regardless of where the server is located, whether in the cloud or on-premises. This model ensures that all connections between servers are encrypted, regardless of the underlying network or IP infrastructure. In addition, it ensures that all communications are explicitly permitted by an SDP allowlist policy. This model enables secure communications between servers across untrusted networks while hiding the servers from all unauthorized connections using the lightweight SPA protocol.

The server-to-server model is similar to the client-to-server model, except that the IH is itself a server and can also act as an AH. Like the client-to-server model, the server-to-server model requires that the SDP gateway, or similar lightweight technology, be installed on each server. This renders all server-to-server traffic hidden to other elements of the security ecosystem. The traffic can also be monitored by analyzing all the dropped packets from the SDP gateway/protected servers. The secure connections to the servers going through the gateway are under the control of the owner of the application/services on the server by default, giving the owner full control of these connections.

[23] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

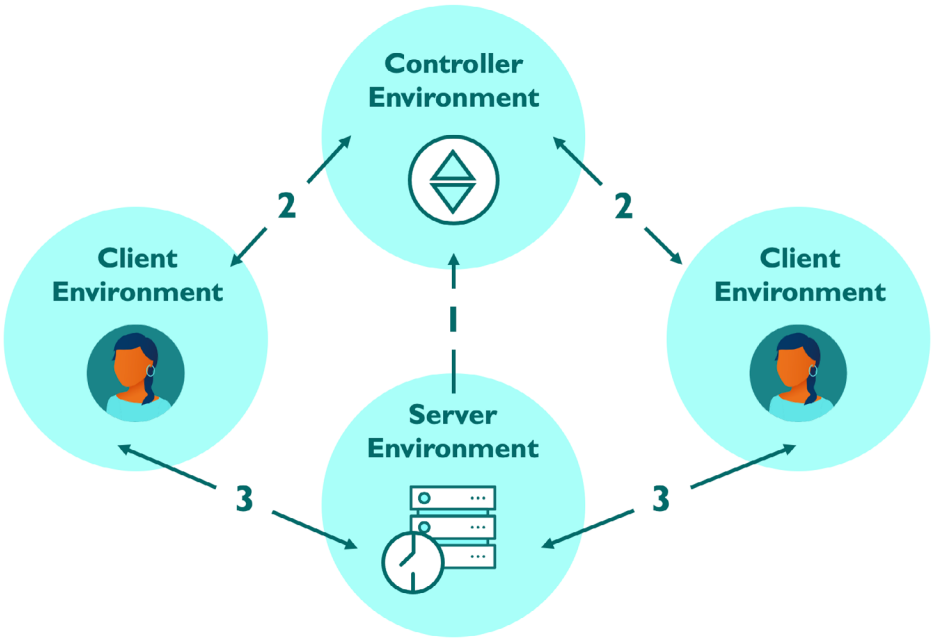


Figure 14: *Client-to-Server-to-Client Model*[24]

4.2.4 Client-to-Server-to-Client Model

The client-to-server-to-client model is well-suited for environments in which organizations are moving their peer-to-peer applications to the cloud, such as IP telephone, chat, or videoconferencing. Regardless of where the server environment is located (cloud or on-premises), organizations can have full control over the connections to the clients. This model results in a logical peer-to-peer relationship between two clients. This can be used for applications in which the traffic must pass through an intermediary server. In these cases, the SDP conceals the IP addresses of the connecting clients, encrypts the network connections between the components, and protects the server/AH from unauthorized network connections by using SPA.

[24] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

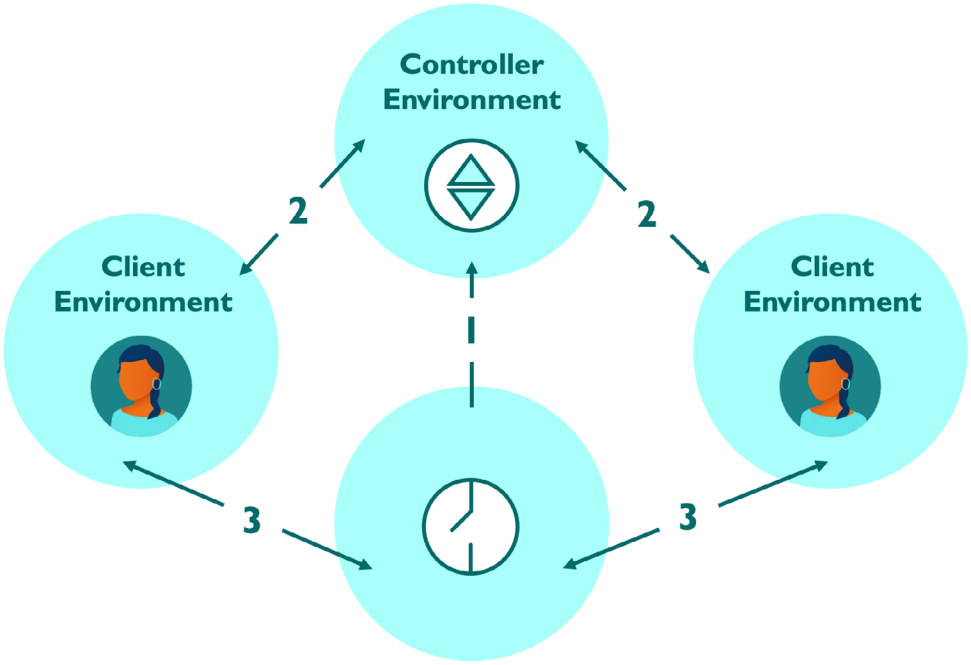


Figure 15: *Client-to-Gateway-to-Client Model*[25]

4.2.5 Client-to-Gateway-to-Client Model

This variation of the client-to-server-to-client model has the advantage of supporting peer-to-peer network protocols that require clients to connect directly to one another, while still enforcing SDP access policies. This results in a logical connection between the clients, each acting as either IH, AH, or both depending on the application protocol. It's worth noting that while the application protocol determines how the clients connect to each other, the SDP gateway continues to perform its standard role as a firewall.

[25] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

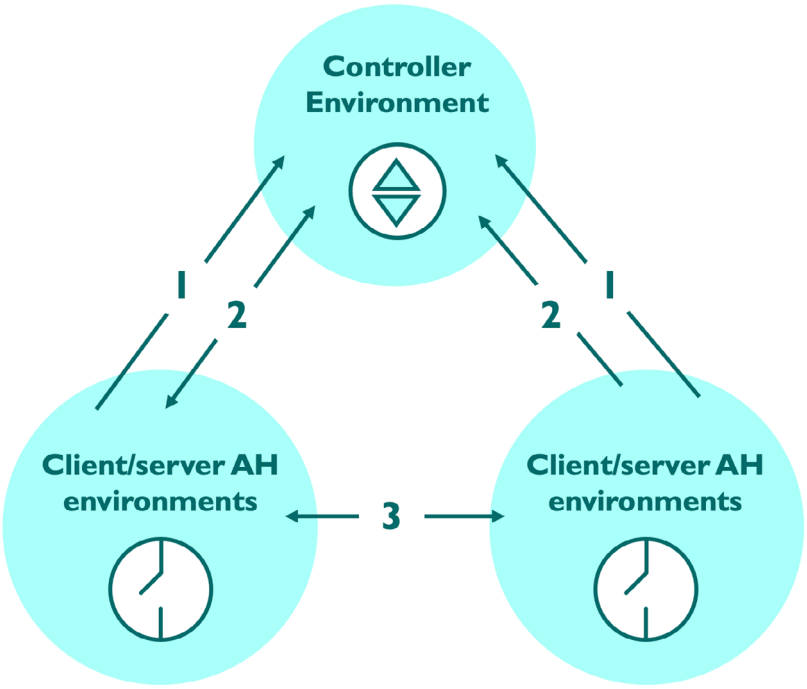


Figure 16: *Gateway-to-Gateway Model*[26]

4.2.6 Gateway-to-Gateway Model

The gateway-to-gateway model is well-suited for certain IoT environments. In this scenario, one or more servers sits behind the AH and acts as a gateway between the clients and the servers. At the same time, one or more clients sits behind an IH that acts as a gateway.

In this SDP model, the IH gateway is running SDP client software, but the client devices are not — they may be incapable of supporting SDP client installation, such as in the case of printers, scanners, sensors, or IoT devices. In this model, the gateway would operate as a firewall or router/proxy, depending on the implementation.

[26] Figure adapted from Cloud Security Alliance, "SDP Architecture Guide v2," 7th, May 2019, https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/

Conclusion

In this introductory SDP course, we provided learners with an overview of SDP's history and how it relates to ZT. We defined key SDP terminology and principles, explored its myriad of technology and business benefits, and walked through current security architecture issues that SDP addresses. Learners were introduced to leading industry cyber risk matrices/lists in order to illustrate how SDP addresses specific, common threats, followed by a deeper dive into its core tenets and underlying technologies.

Lastly, learners were provided with a set of crucial architectural considerations to account for when implementing SDP, followed by the various SDP deployment options and related guidance for selecting the appropriate model.

Glossary

For additional terms, please refer to our Cloud Security Glossary, a comprehensive glossary that combines all the glossaries created by CSA Working Groups and research contributors into one place.

Term	Definition	Source
802.1x	An IEEE standard for local and metropolitan area networks–Port-Based Network Access Control. IEEE 802 LANs are deployed in networks that convey or provide access to critical data, that support mission critical applications, or that charge for service. Port-based network access control regulates access to the network, guarding against transmission and reception by unidentified or unauthorized parties, and consequent network disruption, theft of service, or data loss.	https://1.ieee802.org/security/802-1x/
Accepting Host (AH)	The SDP policy enforcement points (PEPs) that control access to any resource (or service) to which an identity might need to connect, and to which the responsible enterprise needs to hide and control access. AHs can be located on-premises, in a private cloud, public cloud, etc.	https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2/
Access	To make contact with one or more discrete functions of an online, digital service.	https://csrc.nist.gov/glossary/term/access
Active Directory (AD)	A Microsoft directory service for the management of identities in Windows domain networks.	https://csrc.nist.gov/glossary/term/active_directory
Application Programming Interface (API)	A system access point or library function that has a well-defined syntax and is accessible from application programs or user code to provide well-defined functionality.	https://csrc.nist.gov/glossary/term/application_programming_interface

Attribute-Based Access Control (ABAC)	An access control approach in which access is mediated based on attributes associated with subjects (requesters) and the objects to be accessed. Each object and subject has a set of associated attributes, such as location, time of creation, access rights, etc. Access to an object is authorized or denied depending upon whether the required (e.g., policy-defined) correlation can be made between the attributes of that object and of the requesting subject.	https://csrc.nist.gov/glossary/term/abac
Authentication	Verifying the identity of a user, process, or device, often as a prerequisite to allowing access to resources in an information system.	https://csrc.nist.gov/glossary/term/authentication
Authorization	The right or a permission that is granted to a system entity to access a system resource.	https://csrc.nist.gov/glossary/term/authorization
Brute Force Attacks	An attempt to discover a password by systematically trying every possible combination of letters, numbers, and symbols until you discover the one correct combination that works.	https://owasp.org/www-community/controls/Blocking_Brute_Force_Attacks
Certificate Forgery	Data transmitted from an online certificate issuing server to output devices (such as a PC or printer) can be accessed by a hacker and modified into a false certificate.	https://ieeexplore.ieee.org/document/6922060
Client-to-Authenticator Protocol (CTAP)	An application layer protocol for communication between a roaming authenticator and another client/platform, as well as bindings of this application protocol to a variety of transport protocols using different physical media. The application layer protocol defines requirements for such transport protocols.	https://fidoalliance.org/specs/fido-v2.0-id-20180227/fido-client-to-authenticator-protocol-v2.0-id-20180227.html
Control Plane	Used by various infrastructure components (both enterprise-owned and from service providers) to maintain and configure assets; judge, grant, or deny access to resources; and perform any necessary operations to set up communication paths between resources.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

Controller (SDP Controller)	Determines which SDP hosts can communicate with each other. The controller may relay information to external authentication services such as attestation, geo-location, and/or identity servers.	https://downloads.cloudsecurityalliance.org/initiatives/sdp/Software_Defined_Perimeter.pdf
Data Plane	Used for communication between software components. This communication channel may not be possible before the path has been established via the control plane.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf
Device Attestation	The ability to provide proof that elements of the device (e.g., firmware) have not been tampered with.	https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.09082020-draft.pdf
Device Onboarding Process	Involves the installation of the physical device and the setup of credentials so that it can securely communicate with its target cloud or platform. T	https://media.fidoalliance.org/wp-content/uploads/2021/04/Introduction-to-FIDO-Device-Onboard-1.pdf
Distributed Denial-of-Service (DDoS)	Involves multiple computing devices in disparate locations sending repeated requests to a server with the intent to overload it and ultimately render it inaccessible.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1800-15.pdf
Domain Name System (DNS) Poisoning	Results in a DNS resolver storing (i.e., caching) invalid or malicious mappings between symbolic names and IP addresses.	https://www.cs.cornell.edu/~shmat/shmat_securecomm10.pdf
Firewall	An inter-network connection device that restricts data communication traffic between two connected networks. A firewall may be either an application installed on a general-purpose computer or a dedicated platform (appliance), which forwards or rejects/drops packets on a network. Typically firewalls are used to define zone borders. Firewalls generally have rules restricting which ports are open.	https://csrc.nist.gov/glossary/term/firewall
Gateway (SDP Gateway)	Provides authorized users and devices with access to protected processes and services. The gateway can also enact monitoring, logging, and reporting on these connections.	https://cloudsecurityalliance.org/artifacts/sdp-architecture-guide-v2/
Geolocation	Provides access to geographical location information associated with the hosting device.	https://www.w3.org/TR/geolocation/
Hash Message Authentication Code (HMAC)	A message authentication code that uses a cryptographic key in conjunction with a hash function.	https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.198-1.pdf

Hypertext Transport Protocol Secure (HTTPS)	A secure network communication method, technically not a protocol in itself, HTTPS is the result of layering the Hypertext Transfer Protocol (HTTP) on top of the SSL/TLS protocol, thus adding the security capabilities of SSL/TLS to standard HTTP communications.	https://iapp.org/resources/article/hypertext-transfer-protocol-secure/
Identity (ID)	The set of attribute values (i.e., characteristics) by which an entity is recognizable and that, within the scope of an identity manager's responsibility, is sufficient to distinguish that entity from any other entity.	https://csrc.nist.gov/glossary/term/identity
Identity and Access Management (IAM)	The set of technology, policies, and processes that are used to manage access to resources.	https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-203.pdf
Identity Provider (IdP)	A trusted entity that issues or registers subscriber authenticators and issues electronic credentials to subscribers. A cloud service provider may be an independent third party or issue credentials for its own use.	https://csrc.nist.gov/glossary/term/identity_provider
Initiating Host (IH)	The host that initiates communication to the controller and to the AHs.	https://downloads.cloudsecurityalliance.org/initiatives/sdp/SDP_Specification_1.0.pdf
Keyloggers	A reconnaissance tool--with keylogging and screen capture functionality--used for information gathering on compromised systems.	https://attack.mitre.org/software/
Lightweight Directory Access Protocols (LDAP)	A networking protocol for querying and modifying directory services running over TCP/IP.	https://csguide.cs.princeton.edu/email/setup/ldap
Man-in-the-middle (MITM) attacks	An attack where the adversary positions himself in between the user and the system so that he can intercept and alter data traveling between them.	https://csrc.nist.gov/glossary/term/mitm

Micro-segmentation	Is the technique of creating secure zones within a data center and cloud deployments that allow the organization to separate and secure each workload. This makes network security more granular and effective. These secure zones are created based on business services, and rules are defined to secure information workflow.	https://www.techtarget.com/searchnetworking/definition/microsegmentation
Misconfiguration	An incorrect or suboptimal configuration of an information system or system component that may lead to vulnerabilities.	https://csrc.nist.gov/glossary/term/misconfiguration
Multi-factor Authentication (MFA)	Authentication using two or more factors to achieve authentication. Factors include: (i) something you know (e.g., password/personal identification number (PIN)); (ii) something you have (e.g., cryptographic identification device, token); or (iii) something you are (e.g., biometric).	https://csrc.nist.gov/glossary/term/multi_factor_authentication
Multiprotocol Label Switching (MPLS)	An Internet Engineering Task Force (IETF)-specified framework that provides for the efficient designation, routing, forwarding, and switching of traffic flows through the network. MPLS performs the following functions: specifies mechanisms to manage traffic flows of various granularities, remains independent of the Layer-2 and Layber-3 protocols, provides a means to map IP addresses to simple, fixed-length labels used by different packet-forwarding and packet-switching technologies, interfaces to existing routing protocols, and supports the IP, ATM, and frame-relay Layer-2 protocols.	http://tele1.dee.fct.unl.pt/rit1_2020_2021/pages/IEC_MPLS.pdf
Mutual Transport Layer Security (mTLS)	An approach where each microservice can identify who it talks to, in addition to achieving confidentiality and integrity of the transmitted data. Each microservice in the deployment has to carry a public/private key pair and uses that key pair to authenticate to the recipient microservices via mTLS.	https://cheatsheetseries.owasp.org/cheatsheets/Microservices_security.html#mutual-transport-layer-security
Network Access Control (NAC)	A method of bolstering the security of a private or "on-premise" network by restricting the availability of network resources to endpoint devices that comply with a defined security policy.	https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-41r1.pdf

Network Address Translation (NAT)	A function by which internet protocol addresses within a packet are replaced with different IP addresses. This function is most commonly performed by either routers or firewalls. It enables private IP networks that use unregistered IP addresses to connect to the internet. NAT operates on a router, usually connecting two networks together, and translates the private (not globally unique) addresses in the internal network into legal addresses before packets are forwarded to another network.	https://csrc.nist.gov/glossary/term/network_address_translation
Network Segmentation	Splitting a network into sub-networks, for example, by creating separate areas on the network which are protected by firewalls configured to reject unnecessary traffic. Network segmentation minimizes the harm of malware and other threats by isolating it to a limited part of the network.	https://www.nist.gov/itl/smallbusinesscyber/cybersecurity-basics/glossary
Next Generation Firewall (NGFW)	Deep-packet inspection firewalls that move beyond port/protocol inspection and blocking to add application-level inspection, intrusion prevention, and bringing intelligence from outside the firewall. An NGFW should not be confused with a stand-alone network intrusion prevention system (IPS), which includes a commodity or non enterprise firewall, or a firewall and IPS in the same appliance that are not closely integrated.	https://www.gartner.com/en/information-technology/glossary/next-generation-firewalls-ngfws
Open Systems Interconnection (OSI)	Qualifies standards for the exchange of information among systems that are "open" to one another for this purpose by virtue of their mutual use of applicable standards.	https://www.ecma-international.org/wp-content/uploads/s020269e.pdf

Pass-The-Hash	Adversaries may "pass the hash" using stolen password hashes to move laterally within an environment, bypassing normal system access controls. Pass the hash (PtH) is a method of authenticating as a user without having access to the user's cleartext password. This method bypasses standard authentication steps that require a cleartext password, moving directly into the portion of the authentication that uses the password hash.	https://attack.mitre.org/techniques/T1550/002/
Pass-The-Ticket	Adversaries may "pass the ticket" using stolen Kerberos tickets to move laterally within an environment, bypassing normal system access controls. Pass the ticket (PtT) is a method of authenticating to a system using Kerberos tickets without having access to an account's password. Kerberos authentication can be used as the first step to lateral movement to a remote system.	https://attack.mitre.org/techniques/T1550/003/
Phishing	A technique for attempting to acquire sensitive data, such as bank account numbers, through a fraudulent solicitation in email or on a web site, in which the perpetrator masquerades as a legitimate business or reputable person.	https://csrc.nist.gov/glossary/term/phishing
Port	Another essential asset through which security can be breached. In computer science, ports are of two types - physical ports (which is a physical docking point where other devices connect) and logical ports (which is a well-programmed docking point through which data flows over the internet). Security and its consequences lie in a logical port.	https://www.w3schools.in/cyber-security/ports-and-its-security/

Public Key Infrastructure (PKI)	The framework and services that provide for the generation, production, distribution, control, accounting, and destruction of public key certificates. Components include the personnel, policies, processes, server platforms, software, and workstations used for the purpose of administering certificates and public-private key pairs, including the ability to issue, maintain, recover, and revoke public key certificates.	https://csrc.nist.gov/glossary/term/public_key_infrastructure
Role Based Access Control (RBAC)	Access control based on user roles (i.e., a collection of access authorizations a user receives based on an explicit or implicit assumption of a given role). Role permissions may be inherited through a role hierarchy and typically reflect the permissions needed to perform defined functions within an organization. A given role may apply to a single individual or to several individuals.	https://csrc.nist.gov/glossary/term/role_based_access_control
Security Assertion Markup Language (SAML)	A protocol consisting of XML-based request and response message formats for exchanging security information, expressed in the form of assertions about subjects, between online business partners.	https://csrc.nist.gov/glossary/term/security_assertion_markup_language
Security Group	Are sets of IP filter rules that are applied to all project instances, which define networking access to the instance.	https://docs.openstack.org/nova/train/admin/security-groups.html#:~:text=Security%20groups%20are%20sets%20of,networking%20access%20to%20the%20instance.&text=By%20default%2C%20security%20groups%20(and,by%20the%20Neutron%20networking%20service
Single Packet Authorization (SPA)	Can authenticate a user to a system for simple remote administration. It is a protocol for allowing a remote user to authenticate securely on a "closed" system (limited or no open services) and make changes to or run applications on the "closed" system.	https://www.blackhat.com/presentations/bh-usa-05/bh-us-05-madhat.pdf

Software-Defined Network (SDN)	An approach to computer networking that allows network administrators to manage network services through abstractions of higher-level functionality. SDNs manage the networking infrastructure. This is done by decoupling the system that makes decisions about where traffic is sent (the control plane) from the underlying systems that forward traffic to the selected destination (the data plane).	https://ieeexplore.ieee.org/abstract/document/6819788
Software-Defined Perimeter (SDP)	A network security architecture that is implemented to provide security at Layers 1-7 of the OSI network stack. An SDP implementation hides assets and uses a single packet to establish trust via a separate control and data plane prior to allowing connections to hidden assets.	https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-and-zero-trust/
Structured Query Language (SQL) Injection	These attacks, which are still quite common on the Internet, look for web sites that pass insufficiently processed user input to database back-ends and then send carefully-crafted input that will cause exposure of database records, and possibly allow destruction of databases.	https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7682.pdf
Transmission Control Protocol (TCP)	A transport protocol that is used on top of IP to ensure reliable transmission of packets. TCP includes mechanisms to solve many of the problems that arise from packet-based messaging, such as lost packets, out of order packets, duplicate packets, and corrupted packets. Since TCP is the protocol used most commonly on top of IP, the Internet protocol stack is sometimes referred to as TCP/IP.	https://rb.gy/qcorbs
Transmission Control Protocol/ Internet Protocol (TCP/IP)	A set of protocols covering (approximately) the network and transport layers of the seven-layer Open Systems Interconnection (OSI) network model.	https://www.gartner.com/en/information-technology/glossary/tcpip-transmission-control-protocolinternet-protocol
Transport Layer Security (TLS)	A cryptographic protocol, successor to SSL, that provides security for communications over a computer or IP network.	https://csrc.nist.gov/glossary/term/transport_layer_security

Virtual Local Area Network (VLAN)	A broadcast domain that is partitioned and isolated within a network at the data link layer. A single physical local area network (LAN) can be logically partitioned into multiple, independent VLANs; a group of devices on one or more physical LANs can be configured to communicate within the same VLAN, as if they were attached to the same physical LAN.	https://csrc.nist.gov/glossary/term/virtual_local_area_network_vlan
Virtual Private Network (VPN)	A virtual network built on top of existing physical networks that can provide a secure communications mechanism for data and IP information transmitted between networks or between different nodes on the same network.	https://csrc.nist.gov/glossary/term/virtual_private_network
Web Authentication (WebAuth)	Web Authentication (WebAuthn), a core component of FIDO Alliance's FIDO2 set of specifications, is a web-based API that allows websites to update their login pages to add FIDO-based authentication on supported browsers and platforms. FIDO2 enables users to leverage common devices to easily authenticate to online services in both mobile and desktop environments.	https://fidoalliance.org/fido2/fido2-web-authentication-webauthn/

Transport Layer Security (TLS)	A cryptographic protocol, successor to SSL, that provides security for communications over a computer or IP network.	https://csrc.nist.gov/glossary/term/transport_layer_security
Virtual Private Network (VPN)	A virtual network built on top of existing physical networks that can provide a secure communications mechanism for data and IP information transmitted between networks or between different nodes on the same network.	https://csrc.nist.gov/glossary/term/virtual_private_network

Zero Trust Strategy

CCZT Study Guide

cloud security alliance®

CCZT™
Certificate of
Competence in Zero Trust

The official location for SDP and Zero Trust Working Group is
https://cloudsecurityalliance.org/research/working-groups/zero-trust/

Disclaimer

Cloud Security Alliance designed and created this Zero Trust Training course study guide (the "Work") primarily as an educational resource for security and governance professionals. Cloud Security Alliance makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Version Number: 20240820

About Cloud Security Alliance

The Cloud Security Alliance℠ (CSA) (www.cloudsecurityalliance.org) is the world's leading organization dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment. Cloud Security Alliance harnesses the subject matter expertise of industry practitioners, associations, governments, and its corporate and individual members to offer cloud security-specific research, education, certification, events and products. Cloud Security Alliance activities, knowledge and extensive network benefit the entire community impacted by cloud—from providers and customers, to governments, entrepreneurs and the assurance industry—and provide a forum through which diverse parties can work together to create and maintain a trusted cloud ecosystem.

CSA Address

709 Dupont St.
Bellingham, WA 98225, USA
Phone: +1.360.746.2689
Fax: +1.206.832.3513

Contact us: support@cloudsecurityalliance.org
Website: https://cloudsecurityalliance.org/
Zero Trust Training Page: https://knowledge.cloudsecurityalliance.org/page/zero-trust-training
Zero Trust Advancement Center: https://cloudsecurityalliance.org/zt/
Provide Feedback: support@cloudsecurityalliance.org
CSA Circle Online Community: https://circle.cloudsecurityalliance.org/
Twitter: https://twitter.com/cloudsa
LinkedIn: www.linkedin.com/company/cloud/security/alliance
Facebook: www.facebook.com/csacloudfiles
CSA CloudBytes Channel: http://www.csacloudbytes.com/
CSA Research Channel: https://www.brighttalk.com/channel/16947/
CSA Youtube Channel: https://csaurl.org/youtube
CSA Blog: https://cloudsecurityalliance.org/blog/

Acknowledgments

Dedicated to Juanita Koilpillai, a pioneer in software-defined perimeters whose contributions to the Certificate of Competence in Zero Trust (CCZT), Zero Trust training, and CSA are immeasurable.

The CCZT and Zero Trust training was developed with the support of the Cloud Security Alliance Zero Trust Expert Group, whose members include volunteers from a wide variety of industries across the globe. Made up of subject matter experts with hands-on experience planning and implementing Zero Trust, both as cloud service consumers and providers, the Zero Trust Expert Group includes board members, the technical C-suite, as well as privacy, legal, internal audit, procurement, IT, security and development teams. From cumulative stakeholder input, the Zero Trust Expert Group established the value proposition, scope, learning objectives, and curriculum of the CCZT and Zero Trust training.

To learn more about the CCZT and Zero Trust training and ways to get involved please visit: https://cloudsecurityalliance.org/education/cczt

We would also like to thank our beta testers, who provided valuable feedback on the CCZT and Zero Trust training: https://cloudsecurityalliance.org/contributors/cczt-contributors

Lead Developers:

Heinrich Smit
John Kindervag
Michael Roza
Paul Simmonds
Prasad T.
Shruti Kulkarni

Contributing Editors:

Jason Garbis
Mark Schlicting
Richard Lee
Roland Kissoon

Expert Reviewer:

Chase Cunningham
Hannah Day
Jaye Tilson
Jonathan Flack
Matt Lee
Matt Meersman (Dr.), PhD
Ron Martin (Dr.), PhD

CSA Staff:

Adriano Sverko
Andy Ruth
Anna Campbell Schorr
Chandler Curran
Daniele Catteddu
Erik Johnson
Hannah Rock
Judy Bagwell
Stephen Smith

Table of Contents

List of Figures

Course Intro

This training assumes that learners are familiar with the introductory content of the Cloud Security Alliance's (CSA) Zero Trust Training: *Introduction to Zero Trust Architecture*. Additionally, we recommend that students have at least a basic understanding of networks and network security.

This course presents an in-depth exploration of Zero Trust (ZT) from an organizational strategic perspective; and it also delves into the foundational principles of ZT, its benefits, and the critical factors driving organizational buy-in and strategic alignment.

This course comprises several units, each addressing a distinct, strategic ZT aspect:

- We focus on various levels of strategic engagement with ZT;
- We examine ZT's value, drivers and business case;
- We move on to practical tactics for implementing ZT; and
- We cover the broad impact of a ZT strategy on operations, encompassing areas such as cultural and organizational change, training and education, regulatory compliance, legacy systems and infrastructure challenges, user experience, and adapting to the evolving threat landscape.

Course Structure

- **Unit 1:** Levels of Strategy
- **Unit 2:** Zero Trust Drivers & Buy-In
- **Unit 3:** Tactics for Zero Trust
- **Unit 4:** Zero Trust & Operations

Course Learning Objectives

By the end of this course, you will be able to:

- Understand why ZT is a cybersecurity strategy;
- Understand the key principles and components of ZT strategy, and its relationship to the organization;
- Identify how an organization's business goals, and associated IT strategy can be supported with ZT architecture;
- Understand the organization's current state (Business and IT landscape, design architecture, and more);
- Identify tactics and best practices for a ZT implementation; and
- Identify critical cultural, organizational and technical challenges for the implementation of a ZT strategy.

1 Levels of Strategy

Equipping cybersecurity experts with the skills and knowledge they need to successfully implement Zero Trust (ZT) security solutions is a major goal of this course. To effectively implement a ZT strategy, your approach must clearly support existing and new business goals, align with organizational objectives, and secure executive sponsorship and resources. A strong understanding of strategic concepts and an organization's particular set of strategies is essential.

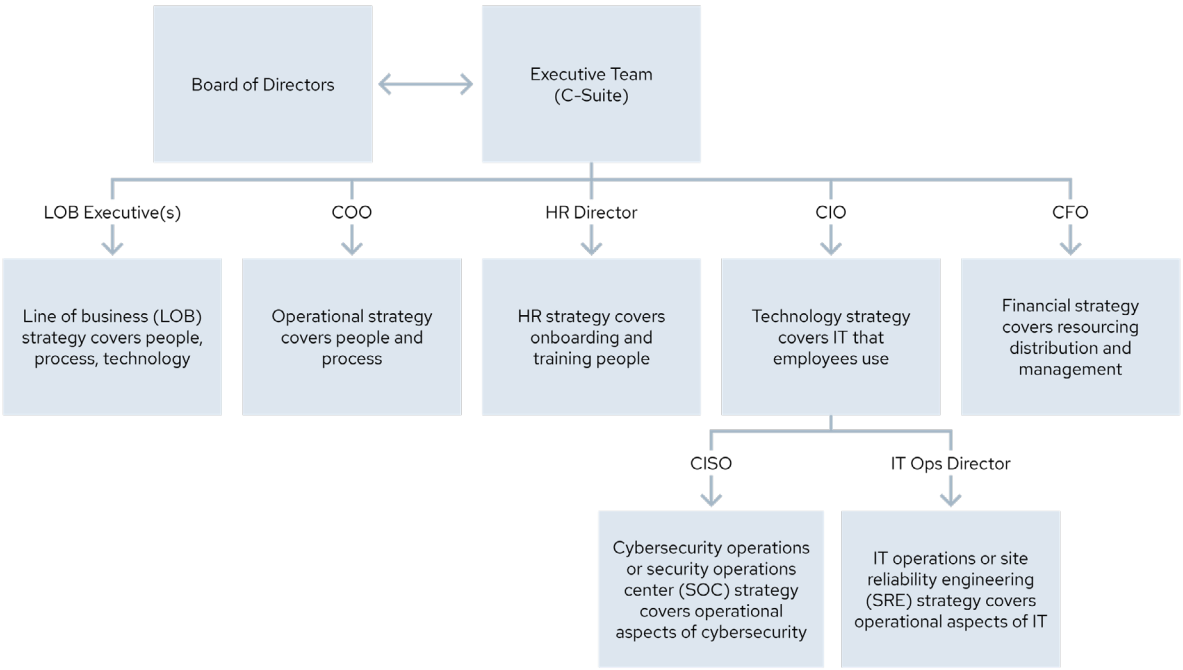


Figure 1: *Strategy Perspective of a Standard Org. Chart*

Though organizational structures vary widely, responsibilities for many roles are more constant, as Figure 1: Strategy Perspective of a Standard Org. Chart, above, illustrates. Hence, since ZT is integral to the cybersecurity footprint, and has a primary focus on technology, it must involve both the IT director and the Chief Information Officer (CIO).

And that's not all. A ZT strategy impacts how product teams develop, deliver and utilize IT products in their line of business (LOB). Collaboration with LOBs is important: If you can foster clarity where there is confusion, especially in the early planning phases, you can effectively convert concepts to intent, and intent to action and results.

Configuration state is important for site reliability, and monitoring for breaches or attacks. Though LOBs must focus on their own strategies and approaches to adding value, their cyber activity must be operated and monitored. In the event of breach, tools and business data must be returned to a known state, and preferably to the expected known state.

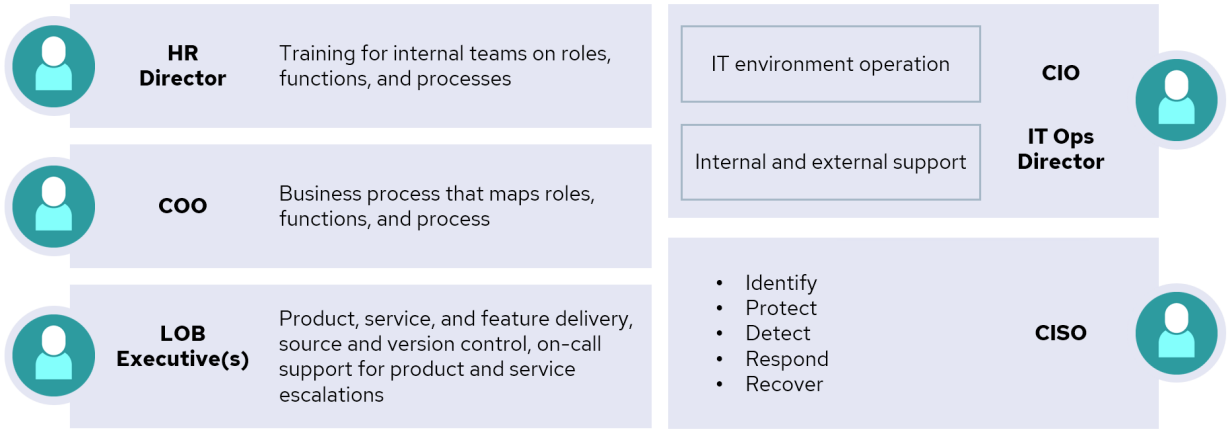


Figure 2: *Example of Roles and Responsibilities*

Regardless of the actual terms used in your organization for each engagement level, if you can clarify how each level influences and informs the others, you are a more effective ZT planner, architect and implementer. We hope that structuring this course around engagement levels helps you better organize and control your ZT-related projects.

One of the main goals in discussing strategic terms is to assist you in thinking and communicating clearly and with authority. Clear, concise communication may drive projects to their milestones and ultimate completion. The table below defines organizing engagement levels used in this course (depending on your organization, actual levels may vary), as follows:

- Organization strategy (this encompasses the business goals and objectives);
- Cybersecurity strategy (usually a part of IT security strategy);
- Technology and IT strategy;
- Tactics; and
- Operations.

Engagement Level	Short Definition	Examples in Practice	ZT Considerations[1]
Organization Strategy	A high-level plan that outlines an organization's goals and objectives.	Corporation: Seamless integration of the organization's third parties, enabling seamless and secure collaboration for outsourcers and joint venture partners. Departments: Gain support from decision-makers across departments.	• Familiarize yourself with the organization's common metrics, such as revenue, net income, margins, cost-related figures & cash flow. • Gain insight from non-financial measurements, such as regulatory compliance, audit results and more. • Embed ZT principles into the organization's mission statement and core values. • Proactively identify and mitigate security risks at the organizational level. • Establish a ZT culture that prioritizes security and privacy.
Cybersecurity Strategy	How an organization protects its information and systems, and responds when there are cyberattacks.	Zero Trust: a security framework that assumes that no user or device can be trusted by default. It implements security controls to verify users and devices before they are granted access to resources. ZT strategy can help organizations protect themselves from cyberattacks, even if the attacker has already gained access to your environment.	• Conduct regular ZT security assessments and penetration tests to identify and remediate security vulnerabilities.

[1] Note: This is *not* an exhaustive list of ZT considerations, instead it is meant to serve as an example to get students to begin thinking about the different strategic levels and how they relate to ZT.

Technology & IT Strategy	How an organization uses technology and IT to achieve its business objectives.	Assets: Take inventory, classify and categorize all assets (e.g., identities, apps, networks, etc.). Risk assessment: Conduct a thorough risk assessment to help prioritize ZT efforts. Compliance and governance: Align with existing compliance requirements for regulatory adherence and to strengthen the organization's security posture.	• Invest in new data centers and cloud computing technologies. • Develop a scalable and reliable cloud computing platform. • Use automation and DevOps to improve efficiency and agility.
Tactics	The things you use. These are the specific tools, methods or actions employed to execute strategy.	Put ZT frameworks into action: ZT Design principles, five steps for ZT implementation, Zero Trust Maturity Model (ZTMM). Integrate with standard business practices: Lean manufacturing practices, JIT inventory management, continuous improvement initiatives.	• Simplify user access and assign clear management responsibilities. • Deploy a micro-segmentation solution to isolate applications and data from each other.
Operations	The way you use them. Details of how these tools and actions are successfully employed in practice to work towards the strategic objectives.	Integrate user experience (UX) and site reliability engineering (SRE) in ZT adoption, focusing on code-driven automation for enhanced operational efficacy.	• Monitor the organization's network and systems for suspicious activity. • Respond to ZT security incidents in a timely and effective manner. • Provide ZT security awareness training to employees.

Table 1: *Org. Engagement Levels with Examples and ZT Considerations*

1.1 Organizational Strategy - The Ultimate Goal

Organizational strategy is the overarching, ultimate goal that guides an organization's actions and decisions. It represents the highest-level objective that an entity aims to achieve. We assume that one of the key approaches that the board of directors and executive team have chosen to improve their cybersecurity strategy is to leverage the principles of ZT.

1.2 Cybersecurity Strategy - Zero Trust

"Zero Trust[2] is a cybersecurity strategy premised on the idea that no entity or asset is implicitly trusted. It assumes that a breach has already occurred or will occur. Therefore, a user should not be granted access to sensitive information by a single verification done at the enterprise perimeter. Instead, each entity (user, device, application, etc.), and transaction must be continually verified."[3] At the strategy level, ZT differs from traditional cybersecurity strategies by not assuming nor providing any implicit or inherited trust in anything.

ZT can impact every person and process inside an organization, as well as the entire technology stack. It should be treated as a holistic cybersecurity strategy that covers all enterprise technology domains. This includes cloud and multi-cloud environments, internal and external endpoints. The strategy also includes organizational and bring your own device (BYOD) scenarios, on-premises and hybrid systems, as well as operational technology (OT) and internet of things (IoT).[4]

1.2.1 Key Tenets of Zero Trust

ZT is a set of principles and practices designed to reduce cyber risk in today's dynamic IT environments. As a cybersecurity strategy, ZT requires strict authentication and verification for all entities (e.g., each person, device or service) trying to access an IT resource. It doesn't matter whether the access is inside or outside the physical network perimeter. ZT emphasizes the protection of individual assets (systems and data) rather than network segments.

Guiding ZT principles, significance and value vary for each organization, depending on factors such as location, industry and individual traits. The following is a list of some of the tenets that have been discussed[5]:

[2] As many organizations familiarize themselves with Zero Trust, they frequently discover a large amount of misinformation that makes navigating difficult. Cloud Security Alliance's Zero Trust Advancement Center (ZTAC) cuts through the noise, focusing on solutions, not vendors, and delivering trusted guidance that helps raise ZT strategy to the next level.
[3] NSTAC. (2022). Report to the President on *Zero Trust and Trusted Identity Management*. Pg.1
[4] Cloud Security Alliance (N.A) *Zero Trust Implementation Primer* - The Five Step Process (Draft). Pg. 6.
[5] See Cloud Security Alliance course *Introduction to Zero Trust Architecture* for an in-depth review of ZT tenets.

- Never trust, always verify: trust no one, either inside or outside the network perimeter (assuming you have one).
- Assume a hostile environment: Malicious actors reside both inside and outside of any environment you manage.
- Presume breach: Operate with the assumption that an adversary already has a presence in your environment. For example, by limiting the blast radius to contain the impact of a breach to a smaller number of impacted devices and services.

1.2.2 Strategic Alignment & Operational Integration

ZT is a holistic endeavor and not just a tactical change. As such, it represents a strategic realignment of the entire security posture. This realignment starts at the highest engagement with the organizational strategic objective[6]. For some organizations, this may be synonymous with *preventing any breach*. For others, it may not be a breach that is most important, but the resiliency in place to limit the impact of it.

Furthermore, as the table that defines the different types of organizing levels mentioned (Table 1: Engagement Levels with Examples and ZT Considerations), Zero Trust Architecture (ZTA) is not just a technical recommendation, but also a cultural shift. The shift demands that security aligns closely with business functions, acknowledging that different departments may have varied security needs. There may be other organizational objectives. Regardless of the specific organizational strategic objective, ZT should be seen as the guiding principle or the *big idea* at the strategy level. ZT should be seen as directly contributing to the organizational strategy.

1.3 IT Strategy & Technology

In the context of ZTA, significant adjustments extend beyond technology and IT strategy. The adjustments encompass a fundamental shift in mindset and organizational culture, embracing the "never trust, always verify" principles. Adopting a *never trust, always verify* approach means that access is continuously validated through rigorous security checks and authentication measures. A necessary transformation in adopting ZT is proliferating and enhancing network segmentation. Segmentation is the sub-dividing of the network environment into smaller, distinct segments to limit access and contain breaches.

At the tactical level, implementing ZT involves specific actions. For example, strict access control on a need-to-know basis and secure access to resources regardless of their location. These topics are discussed in more detail in the tactics and operations sections. Such alignment ensures that security enables business operations, rather than hindering them. It requires an architecture that allows for flexibility, catering to different service level agreements (SLAs), administrative controls, audit requirements, regulations and certifications.

IT strategy also encompasses user and entity behavior analytics (UEBA). Additionally, technology strategy must integrate closely with governance, while rigorously controlling and monitoring access to reduce risk. Cybersecurity goals, including ZT, should align with the organization's overall strategy and board-level roadmap, guiding various projects and technology strategies for the upcoming years.

[6] In some organizations, this is also referred to as the "grand strategic objective."

The operationalization of ZT is where ZT concepts become tangibly interwoven with the day-to-day activities of the organization. This ensures that every aspect of the network is designed from the inside out with a default perspective of verifying everything and trusting nothing. To make such a perspective functional and practical requires consolidating technologies. IT also requires enhancing security measures for critical assets, and applying specialized controls for legacy and critical infrastructure systems.

ZT and technology strategies are closely connected to governance, risk management, and every aspect of security, but it is important to clarify the role of each framework. Governance, focusing on establishing and maintaining policies, standards and guidelines, plays a crucial role in ZT implementations. The governance ensures that ZT practices not only adhere to regulatory requirements but also align with the organization's overall objectives. This relationship positions ZT as a strategic element within the governance landscape. This lesson does not cover risk management or compliance strategies because they are covered in other lessons.

1.4 Tactics

Tactics, within the context of a ZT strategy, are crucial for effectively addressing specific risks and aligning security measures with organizational objectives. These tactics involve prioritizing business goals, adopting an "inside out" security approach, and implementing the principle of least privilege to control resource access. Metrics and reporting improvements are vital for assessing ZT effectiveness. Monitoring and logging network traffic and identifying and protecting critical data, applications, assets and services (DAAS) also plays a pivotal role in these tactics. Additionally, transitioning to ZT requires a phased, risk-based approach that impacts tactics, such as precise policy creation, prioritization, and iterative implementation.

Tactics are fundamental in ensuring a smooth transition to a ZTA and are focused on protecting assets and resources efficiently. ZT policies, detailed access controls, monitoring network traffic, and setting progress metrics contribute to the successful implementation of ZT principles. Organizations can bolster their cybersecurity posture by adopting these tactics and gradually progress along a Zero Trust Maturity Model (ZTMM)[7], ultimately achieving improved security outcomes. Tactics and the maturity model are covered in the tactics units.

NIST *Zero Trust Architecture* (SP 800-207)[8] defines the tenets that are fundamental to a ZT environment. As such, the tenets need to be considered before deploying policy enforcement points (PEPs) and policy decision points (PDPs). To meet these foundational tenets, a dynamic policy must drive the shift away from network access. The policies must motivate organizations to implement measures that reduce the attack surface prone to lateral attacks, such as macro and micro-segmentation.

Because all communications must be secured, regardless of location, a tactical assessment is needed. Assessments can ensure adequate encryption has been used for each application, regardless of destination; and access to resources is on a per-session basis.

[7] CISA. (2023). *Zero Trust Maturity Model* (Version 2.0).
[8] NIST. (2020). *Zero Trust Architecture* (SP 800-207).

The migration to a ZTA might take from a few months to several years, depending on the maturity level of the organization (See Section *Zero Trust Maturity Model*) – the path differs for each organization. As noted above, tactical plans are needed for the platform, the tools, the monitoring and detail metrics must each be assessed for the journey.

1.5 Operations

Operations refers to the activities, processes and procedures involved in managing and maintaining organizational infrastructure and IT infrastructure. This includes a range of tasks, aimed at ensuring the effective and efficient functioning of all IT resources, such as hardware, software, networks and data storage systems.

Embarking on a ZT journey involves cultural and organizational shifts, emphasizing a ZT culture over technology, and securing leadership support for continuous risk management. Training and education geared toward understanding the ZT paradigm is commonly necessary. This includes a strategic appreciation of ZT that targets management, and several other initiatives that emphasize transforming business processes and roles. Regulatory landscapes are also adapting to require robust cybersecurity practices that align with ZT principles.

Achieving ZT success involves several important considerations in managing and executing day-to-day operations. When a ZT strategy is implemented, the identity management process should be automated, as should monitoring and detection. Day-to-day tasks people in the roles listed above perform include:

- Organizing log data so that input logs from different sources can be looked at and analyzed using the same tools and interfaces.
- Adjusting controls and fine tuning the automation regularly to make sure it checks the right parameters according to policy rules.
- Monitoring to ensure that the automatic checks of logs catch any activities that don't follow the policy rule.

Ensuring cybersecurity solutions enhance rather than add friction to a user's productivity and overall experience is an important organizational goal that operational leaders must focus on and defend. Operational processes, such as site reliability engineering (SRE), and a focus on automation and scalable systems, can improve operational efficiency and promote a positive user experience. Furthermore, operational procedures may need updates to align them with a new ZT framework, ensuring that response strategies and daily activities align with ZT principles.

Challenges include integrating ZT with legacy systems, where a tailored approach is necessary. Maintaining vigilance in monitoring the evolving threat landscape ensures ZT remains agile and responsive. These concepts are expanded upon in a unit dedicated to ZT operations.

2 Zero Trust Drivers & Buy-In

For an effective Zero Trust (ZT) implementation strategy, you must align the strategy with organizational values and drivers. The central objective of this approach is to secure buy-in from key stakeholders within the organization. The buy-in ensures that the ZT strategy is not only well-conceived but also well-received and integrated within the organization.

Remember, ZT is a context and risk-based approach: access is granted according to specific situations or conditions. Due to these characteristics, ideas that help deliver and implement ZT are usually clearer and gain wider acceptance if you prepare use-case scenarios and provide contextual applications to support them.

To get started in defining the desired state, you may wish to ask the following questions:

- Identifying Catalysts and Business Drivers:
  - What are the primary catalysts driving the adoption of ZT in our organization?
  - What are the key business drivers aligning with ZT implementation?
- Evaluating Security Posture and Data Access:
  - How does our current security posture align with common attack vectors, especially in the context of our ZT pillars?
  - Is access to confidential and regulated data restricted to registered applications?
- Enhancing Security and Privacy Strategies:
  - How can we develop a more efficient and effective security strategy under ZT principles?
  - What role does privacy play in our overall security and risk management, and how can we define our privacy objectives?
- Access Control and Authentication:
  - Are all privileged accounts secured with FIDO2 or equivalent multi-factor authentication (MFA), or are privileged access workstations (PAWs) necessary?
  - Is MFA mandatory for all (people) identities in our environment?
- Device and Data Access Management:
  - To what extent are personal (non-organization managed) devices allowed access to organizational data?
- Competitive Advantage through Security:
  - Can we gain a market advantage over competitors through a superior security and privacy strategy?
- Compliance with ZT in Development:
  - Are our development teams aligning software testing with ZT standards?

2.1 The Value of Zero Trust

ZT offers numerous potential benefits, including streamlined security and IT infrastructure management, enhanced data protection, regulatory compliance, reduced compliance-related efforts, increased organizational agility, stakeholder confidence, and lower IT and operational costs[9]. In other words, the benefits go beyond the security domain. For these reasons, if positioned correctly, ZT has the potential to be a business enabler rather than a hindrance to the organization or its managers. ZT transforms IT and security, aligns business and security goals, and reduces siloed activities.

[9] Cloud Security Alliance. (2023). *Communicating the Business Value of Zero Trust.*

Having established the benefits of ZT as a business enabler, the next step involves a tailored approach. The tailored approach aims to integrate ZT into the organization's unique culture, and business and cybersecurity practices. Before embarking on your ZT journey, clearly define your organization's goals and challenges, gather relevant information, and align your strategy with business needs. An organization must clearly define its strategy and governance, focusing on what is relevant, what standard (if any) it commits aligning with, who is affected, when and where each applies, and how it's implemented. Be sure to also capture why a particular strategy or governance policy is important. This exercise ensures that you can explain a new architecture, implement according to related policies, and articulate how your effort seamlessly aligns with the organization's business model and rules.

ZT principles make it easier for IT teams and network infrastructure teams to enforce policies consistently and accurately, enabling a more friction-free work environment. This is because implementation of a Zero Trust Architecture (ZTA) requires moving the access enforcement points closer to the protected asset.

2.2 Risk Management as a Driver

In a traditional legacy organization, the risk calculation is predominantly based on a binary trust. If the entity is inside my network environment (or area of control) it is afforded a level of trust. If it's not on my network or in my area of control then it's untrusted, and thus there is a need to provide extra security measures, for example a virtual private network (VPN).

In the realm of ZT, the foundational principle is *never trust, always verify*. An organization shifts its capabilities, such that it can continually assess and contextualize the risk or risks involved in granting an entity access to an asset.

Remember, this assessment that leads to a decision is not just about having a static defense mechanism in place. It is about creating a proactive, dynamic control plane that evolves with the changing risk landscape. In a ZT environment, access is not only based on contextual factors, it is also temporal. The access needs to adapt to new emerging threats, requiring a continuous review of existing controls and emerging threats. The continuous review must ensure the organization can function without friction from security controls while retaining sufficient protection. The emphasis here is on maintaining consistent, measurable effectiveness.

2.2.1 Board-Level Risk Management & Zero Trust Alignment

Aligning a ZTA with an organization's risk appetite is a strategic process, aiming to deliver security solutions that support the board's strategic vision.

The board plays a crucial role in organizational alignment, as they are responsible for setting and defining the risk appetite, setting budgets and determining the appropriate risk oversight structure. Strategic alignment and budget influences technology and control selection, resource allocation, and policy-making, ensuring a viable cybersecurity strategy.

2.2.2 Evolving Threat & Risk Landscape

Monitoring the evolving threat and risk landscape is part of risk management. Cyber threats evolve continuously, and implementing continuous monitoring involves regularly assessing and updating an organization's understanding of potential threats and vulnerabilities. This process includes analyzing cyberattack trends, identifying new methods employed by attackers, and understanding the implications of technological advancements on security. Organizations need to address these challenges to secure their systems. ZT offers a framework that facilitates a shift in mindset that promotes securing and protecting what is important for the organization.

2.3 Create a Case for Zero Trust

Organizations measure their health and progress using key financial metrics like revenue, net income, margins, costs, and cash flow. Organizations also consider non-financial indicators such as stock performance, compliance, audit outcomes, reputation, and employee productivity for a comprehensive view. Different stakeholders prioritize various metrics. Understanding these measures is vital for asking more effective questions, understanding drivers, and constructing a meaningful case for adopting ZT.

The primary goal of the business case, which is defined further during Cloud Security Alliance's *Zero Trust Planning* training, illustrates how an initiative delivers organizational value and return on investment (ROI). The business case also requires alignment with organizational strategy.

Consider these organizational elements during the buy-in phase to lay a solid foundation for a successful and strategic implementation of ZT:

- Alignment with, and assistance in, delivering key business goals and objectives.
- The value to the business in implementing a ZTA, both tangible and intangible.
- Key stakeholder buy-in: ZT is everyone's responsibility, not just the purview of IT or the Chief Information Security Officer (CISO). Gaining support from the key stakeholders across departments is essential.
- Assets inventory, classification, and categorization of business critical assets or asset classes. These must also be defined in terms of risk to the business.
- Compliance and governance: Confirm that any changes made by implementing ZT align with existing compliance requirements. This ensures regulatory adherence and strengthens the organization's security posture.
- Strengths, weaknesses, opportunities, and threats (SWOT) analysis or cost/benefit analysis (CBA): Perform a SWOT analysis to help identify internal strengths, weaknesses, opportunities and threats. CBAs are also helpful. These help guide IT professionals in the initial stages of a ZT implementation.

2.4 Leadership Buy-In

You should look to map the ZT journey so that leadership, especially non-technical leadership roles, can appreciate how it may affect their area of responsibility. The following examples illustrate how

ZT may serve as a foundation for areas such as privacy, security, compliance, third-party risk management (TPRM), and simplicity and efficiency at your organization[10]. These need translating into examples of ZT success using current business problems that your business leaders recognise are improved by ZT.

Privacy

- Data Minimization and Access Control: By adhering to the principle of "never trust, always verify," ZT ensures that access to sensitive data is tightly controlled and monitored, reducing the risk of unauthorized data exposure.
- Enhanced User Privacy: ZTAs can protect user privacy by limiting access to personal data and ensuring that only necessary data is processed and stored.

Security

- Reduced Attack Surface: ZT requires strict access controls without implicit trust and micro-segmentation. The two limit the pathways an attacker can use to move laterally across a network.
- Real-time Monitoring and Response: Continuous monitoring is a key tenet of ZT, allowing for real-time detection and response to threats, thereby enhancing security postures.

Compliance

- Regulatory Alignment: Many regulatory frameworks require strict access controls and data protection measures, which are core components of a ZT model.
- Audit and Reporting: ZT architectures make it easier to log access and changes, thus supporting compliance reporting and auditing requirements.

Third-Party Risk Management (TPRM)

- Vendor Access Limitations: ZT principles can be applied to third-party vendors to ensure they have only the access and visibility that is necessary to perform their functions.
- Continuous Verification of Third-Party Credentials: Regular re-verification of credentials and access rights helps to manage and mitigate the risks associated with third-party partners.

Simplicity and Efficiency

- Often the implementation of ZT-based access simplifies the traditional access mechanism for the user, enhancing productivity. ZTA helps with quick and seamless access to the assets irrespective of location and network boundaries.
- It is essential to designate a person or a limited number of people with the accountability and authority to manage a particular area. A clear owner ensures issues are identified and highlighted at the appropriate level.

[10] Cloud Security Alliance. (2023). *Zero Trust Guiding Principles*.

3 Tactics for Zero Trust

In this unit, we delve into the tactical aspects of implementing Zero Trust (ZT). We discuss nine crucial sub-sections contributing to building a resilient architecture. Alongside these, we'll also introduce CISA's *Zero Trust Maturity Model (ZTMM)*, which, while not a part of the nine steps, is important in understanding the overall progression in ZT implementation. The initial four subsections outline the foundational principles of ZT design, while the subsequent five subsections detail the step-by-step process for ZT implementation.

- ZT Design Principles
  - Focus on Business Outcomes: Understanding how ZT aligns with and supports the organization's primary business goals.
  - Design from the Inside Out: Developing a security strategy that starts within the organization before extending outwards.
  - Determine Who/What Needs Access: Identifying which users and devices require access to specific resources.
  - Inspect and Log Key Traffic: Aim to monitor and record critical activity for potential threats as a targeted approach.
- Foundational Principles of ZT Design
  - **Step 1:** Define Your Protect Surface(s): Identify and secure critical data and resources within the network (environment).
  - **Step 2:** Map the Transaction Flows: Understand the movement of data within and outside the organization and the potential classification of each transaction type.
  - **Step 3:** Build a Zero Trust Architecture (ZTA): Develop the infrastructure and capabilities necessary for ZT.
  - **Step 4:** Create ZT Policy: Establishing guidelines and rules for network, system and data access and security.
  - **Step 5:** Monitor and Maintain the Network (Environment): Continuously oversee the ZT environment to ensure ongoing security and adapt to new threats.

These elements are vital in shaping and executing an effective security approach aligning with an organization's objectives.

3.1 Zero Trust Design Principles[11]

This section explores the foundational design principles that shape an effective ZT security strategy. These principles guide organizations in transitioning from traditional security models to a more robust approach suited for today's dynamic digital landscape. The focus is on setting goals like designing security from the inside out, accurately determining access requirements, and aiming for thorough inspection and logging of network traffic. It is important to note, however, that the realization of these goals may differ based on an organization's specific capabilities and resources.

[11] (2023) *Zero Trust Explained* by John Kindervag

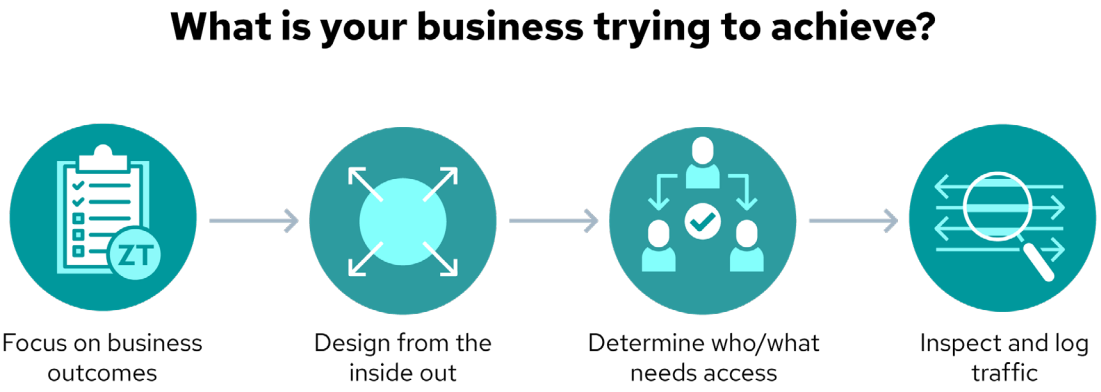


Figure 3: *Zero Trust Design Principles*[12]

3.1.1 Focus on Business Outcomes

Shaping a ZT strategy begins with a clear understanding of the organization's strategic direction and its IT needs. This understanding should incorporate the organization's specific business threats and the broader spectrum of risks posed by factors like organized crime and nation-state actors. The priority is to safeguard business-critical assets—considered the crown jewels—within a ZT framework.

If ZT is the chosen strategy, it's crucial to prioritize business objectives tailored to your organization's specific goals and requirements. From the outset, ZT demands a clear vision, whether it's to manage risks within acceptable limits, reduce compliance costs, or minimize the impact of security incidents. An effective ZT strategy balances security with the cost and value of security and available-resource use to deliver on security initiatives versus other business initiatives like product or feature development, while avoiding excessive measures that could hinder competitiveness.

3.1.2 Design from the Inside Out

ZT marks a shift from traditional perimeter-centric security models, which operate on the obsolete premise that everything inside a network is safe, while external entities pose threats. ZT flips this notion, recognizing that threats can originate from anywhere—both inside and outside the network. This paradigm shift dictates a security architecture designed from the inside out. The design begins with the organization's most critical assets and data at its core and securing access from inside the network, and then extending protection outward. This strategy reorients IT policies to move from a stance of broad threat defense to a focused asset protection approach. The reorientation ensures that the most vital resources are safeguarded at their heart to mitigate the risk of unauthorized access and data breaches.

To connect the design shift to operational strategy, it's important to consider the constraints of limited resources, which all organizations face. This constraint necessitates effective prioritization based on asset value. Conducting a business impact assessment (BIA) or an asset inventory categorized by value helps in identifying critical resources. By ranking assets according to their criticality or value, organizations can efficiently allocate their resources, aligning their security efforts with ZT principles and securing both protect and attack surfaces more effectively.

[12] Figure adapted from: (2023) *Zero Trust Explained* by John Kindervag.

To connect the design shift to operational strategy, it's important to consider the constraints of limited resources, which all organizations face. This constraint necessitates effective prioritization based on asset value. Conducting a business impact assessment (BIA) or an asset inventory categorized by value helps in identifying critical resources. By ranking assets according to their criticality or value, organizations can efficiently allocate their resources, aligning their security efforts with ZT principles and securing both protect and attack surfaces more effectively.

3.1.3 Determine Who & What Needs Access

Today, organizations operate on a global scale, leveraging remote work, joint ventures, outsourced services, and cloud technology. In a ZT security approach, the principle of least privilege (attribute of never trust, always verify) necessitates a precise determination of who or what needs access to certain resources, along with the duration and associated risks of such access. This principle ensures that each entity – be it a user or a system – has access strictly as per their need, thus narrowing the attack surface and enhancing security. An asset's visibility should strictly conform to the need-to-know basis, remaining invisible to those without a legitimate requirement for access.

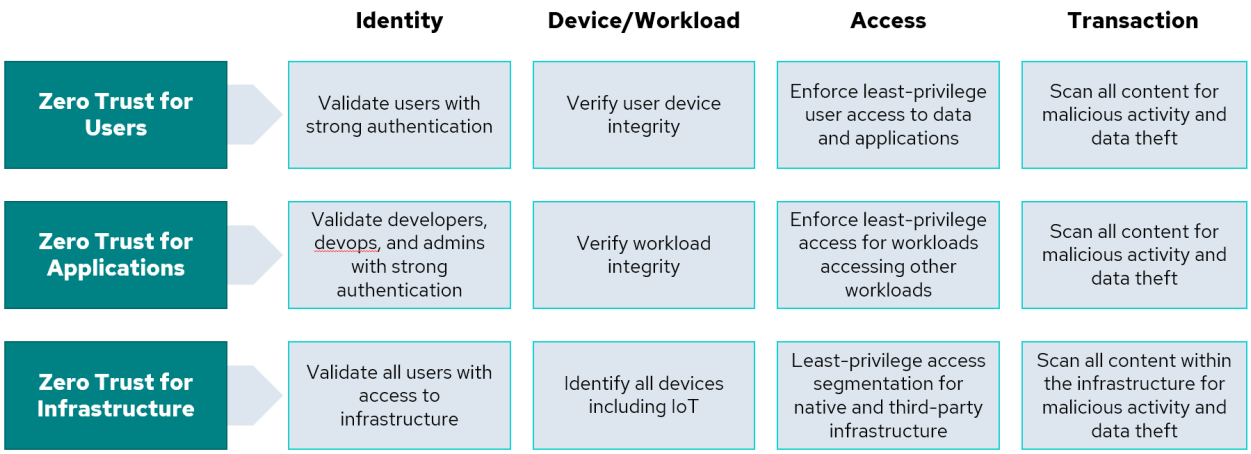


Figure 4: *Zero Trust From a People Perspective*[13]

The concept derived from the Identity Security Alliance, as depicted in Figure 4 Zero Trust From a People Perspective, encompasses seven elements: users, applications, infrastructure, identity, device/workload, access, and transaction. Training individuals outside of security roles, like network teams and developers, to identify and manage trust relationships across these elements is a key challenge. Critical tasks include mapping out where trust is established. Examples include between users and identities or infrastructure and identities, and adopting secure practices like proper firewall configurations and secure coding. A thorough understanding of these trust points allows for the effective identification and mitigation of vulnerabilities that cybercriminals target. Benefit is derived from the insights gained through extensive penetration testing experience.

[13] Cloud Secuirty Alliance. (2023) *The Most Important Part of Zero Trust: People* by George Finney

3.1.4 Inspect & Log Traffic

Two key principles of ZT, as outlined in NIST *Zero Trust Architecture* (SP 800-207)[14], are: Continuously monitoring and assessing the security and integrity of all assets and resources. Gathering extensive information on the current state of assets, network infrastructure, and communications to enhance security measures.

In the journey towards ZT adoption, organizations require some sort of logging and monitoring capabilities. The level of sophistication will vary greatly, depending on the level of organizational maturity, and the resources available.

This process typically begins with the establishment of foundational log management practices. This means starting with the basic yet important step of implementing systems to gather user and entity activity logs, particularly focusing on privileged credentials, coupled with routine manual analysis. This initial phase should cover all essential ZT pillars, laying the groundwork for more advanced security measures.

As the organization's maturity in the ZT framework advances, supplemented by adequate resources and expertise, it can evolve these practices into more sophisticated systems. A key development in this evolution is the integration of a security information and event management (SIEM) system. SIEM serves as a pivotal tool for automated log aggregation and analysis, setting the stage for the adoption of security, orchestration, automation, and response (SOAR) capabilities.

In scenarios where the organization has control over network-level infrastructure or can log traffic at the access gateway, it's strategically important to incorporate relevant and contextual ZT logs into a SIEM system or log management tool. This integration not only enhances the organization's security posture but also aligns with the fundamental principles of ZT. The integration ensures continuous monitoring and adaptation to the ever-evolving security landscape. The ability to assess and log relevant and contextual traffic from both internal and external sources can significantly enhance operational intelligence.

Additionally, at high levels of maturity, the carefully selected log data from various layers or applications can be unified into a common data structure. Data captured includes device, time, user, and the resource or asset access (e.g., server, service, application, etc.) requested. By coupling monitoring and logging, engineers can continuously improve security by rapidly countering any suspicious activity. Continuously scrutinizing traffic patterns in such a manner is a powerful, strategic asset.

Capturing data and monitoring it in real time requires the development of reactive controls, including system and organization controls (SOC) assessments, analysis, response staff and automation in the response pipeline. Logging is only any good if you do something with it, but for the many organizations without a SOC in place, there is no reason for a major consolidated log database. It is acceptable for the ZTA configuration to simply monitor and log:

- At the policy decision points (PDPs);
- All admin operations; and
- All user access event logs.

[14] NIST. (2020). *Zero Trust Architecture* (SP 800-207)

3.2 Zero Trust Maturity Model

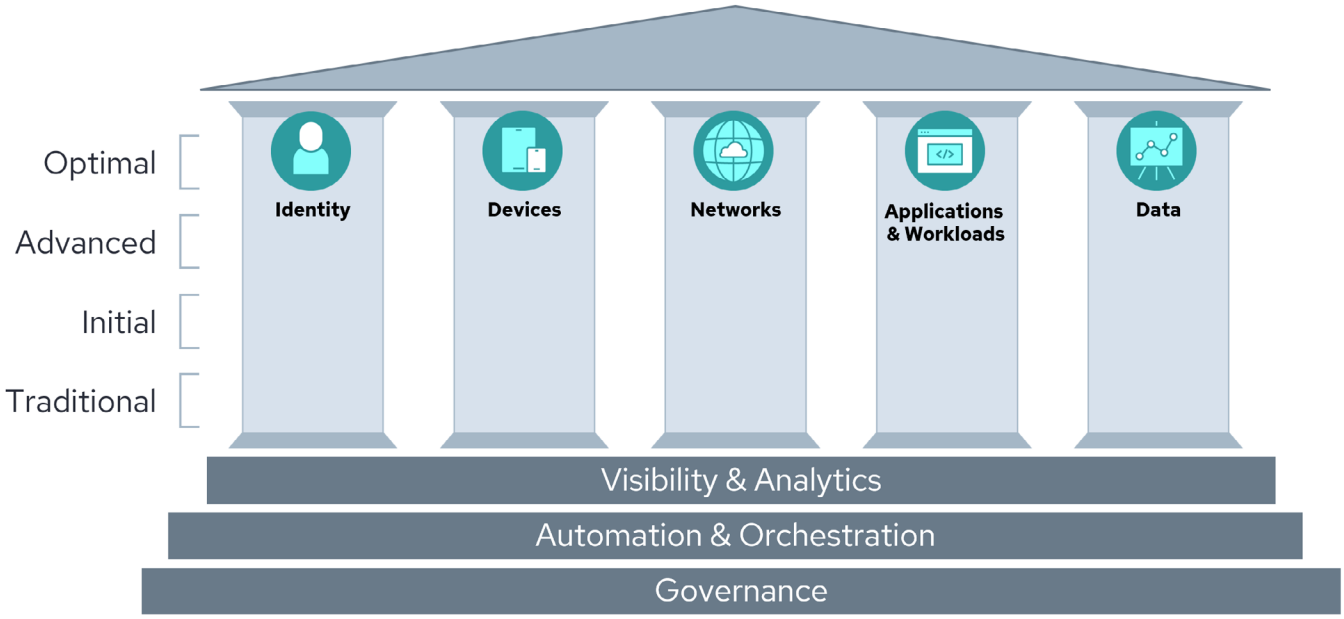


Figure 5: *CISA Zero Trust Maturity Model (ZTMM)*[15]

This section covers the CISA *Zero Trust Maturity Model (ZTMM)*[16], which helps organizations enhance their ZT strategies. The CISA ZTMM outlines maturity stages – Traditional, Initial, Advanced, Optimal – across ZT pillars (Identity, Devices, Networks, Applications and Workloads, and Data) and capabilities (visibility, automation, governance). These maturity stages help organizations assess, plan and implement the necessary measures to progress toward a more secure ZTA. The CISA ZTMM journey, depicted in the accompanying figure, represents a path towards achieving optimal ZT maturity. This journey, a practical visual representation, shows how companies advance through ZT's various maturity levels.

[15] Figure adapted from: CISA. (2023). *Zero Trust Maturity Model* (Version 2.0).
[16] CISA. (2023). *Zero Trust Maturity Model* (Version 2.0).

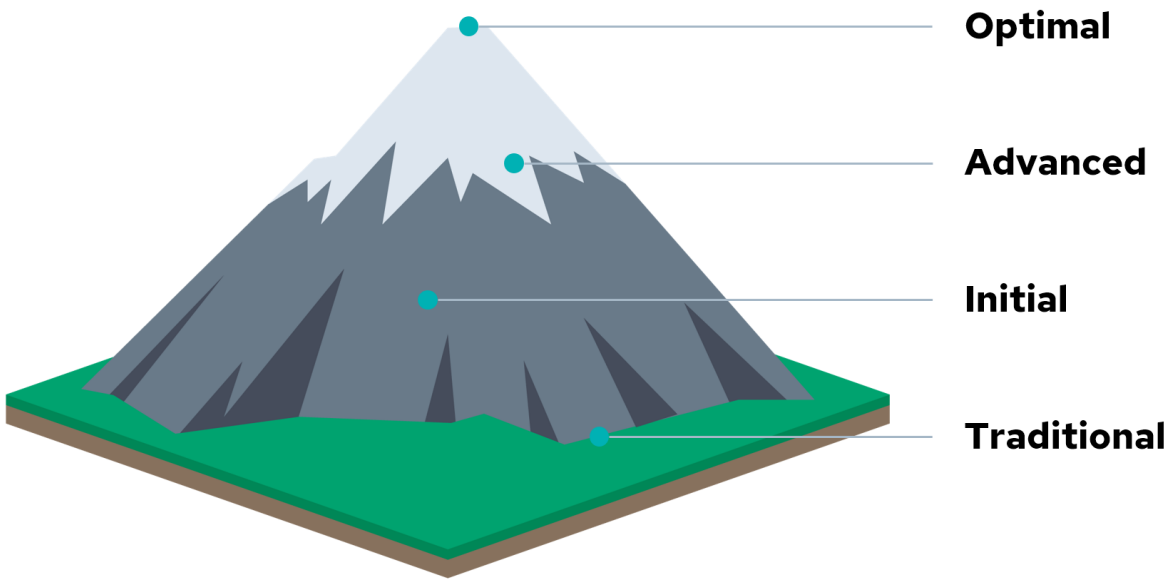


Figure 6: *Zero Trust Maturity Journey*[17]

To utilize the CISA ZTMM effectively, grasp the framework, refine your functions and assess your current ZT maturity. Finally, plan steps for maturity advancement and align them with organizational projects and priorities, using a prioritization model to guide you.

3.2.1 Zero Trust Maturity Model in Practice

Tailoring the CISA ZTMM to fit your needs may seem overwhelming. It is not advisable to strive to achieve optimal maturity across all pillars simultaneously. Nor is it advisable to focus on a single pillar and expect to perfect it across the entire organization. Attempting to perfect one pillar (like identity) across all systems before moving to the next is not only impractical but can lead to stagnation in overall security posture improvement. It is important, instead, to evaluate each protect surface, using worksheets such as the one illustrated below, which is based on the NSTAC report. Each worksheet identifies the protect surface and critical data, assets, application, and services (DAAS) element being evaluated, with 5 (optimized) representing the best possible score for each attribute. The total perfect score on a worksheet would be 25. This is a rare occurrence. Such worksheets help teams prioritize projects, based on safeguarding business-critical assets. This targeted approach allows for a more accurate assessment of maturity gaps and enables the development of specific projects to enhance the security and maturity of each protect surface. Furthermore, by evaluating each protect surface individually, organizations can create a more nuanced and actionable cybersecurity roadmap. Finally, all the protect surfaces can aggregate to define an overall score for the organization as well as an average score per protect surface.

[17] Figure adapted from: (2023) Zero Trust Explained by John Kindervag.

Protect Surface: _____ DAAS Element: _____	Initial	Repeatable	Defined	Managed	Optimized
1. Define your protect surface	1	2	3	4	5
2. Map the transaction flows	1	2	3	4	5
3. Architect a zero trust environment	1	2	3	4	5
4. Create zero trust policy	1	2	3	4	5
5. Monitor and maintain the network	1	2	3	4	5

Total Score: _____

Figure 7: *Zero Trust Maturity Model Worksheet*[18]

This methodology simplifies the complexity inherent in managing multiple and discrete identity solutions across an organization. For example, if an organization focuses on improving the security maturity of its directory services (as a protect surface), it can methodically elevate the maturity level in this specific area, thereby making tangible progress and ensuring continuous improvement in cybersecurity defense. Finally, it helps you monitor progress across various ZT projects to stay aligned with your organization's IT strategy and cybersecurity strategy.

3.2.2 CISA-Based Maturity Model

You may also wish to explore this interactive CISA ZTMM Spreadsheet model[19], a comprehensive tool with status bars for monitoring progress. After a ZT assessment, approach the journey systematically, with the same considerations that we suggested if you choose to use the National Security Telecommunications Advisory Committee (NSTAC) based assessment model:

- Analyze all functions, adjusting the depth as needed;
- Avoid tackling all functions simultaneously so as to not be overwhelmed;
- Focus on enhancing specific areas within individual projects, addressing a single protect surface at a time; and
- Ensure projects align with business drivers and deliver tangible business value, not just security benefits.

You are encouraged to tailor the ZTMM approach to your organization's needs. This pragmatic approach ensures that the journey towards a mature ZT environment is both achievable and manage-

[18] Figure adapted from: NSTAC. (2022). NSTAC Report to the President on *Zero Trust and Trusted Identity Management*. Pg. A-1

[19] Jason Garbis and Numberline Security have created The Zero Trust Maturity Model Resource Center and associated worksheets (GCP Sheets and Excel), aligned with the CISA ZTMM. Learn more about these tools here.

able, because it keeps you flexible. Remember that the goal is delivering tangible value to your business, with ZT maturity serving as a measure of progress and a guide for prioritizing enhancements in your implementation.

3.3 The Five Steps for Zero Trust Implementation

In the journey towards an ideal ZT architecture, there are five essential steps to follow to operationalize each protect surface project. These steps provide a structured approach to enhance cybersecurity and ensure a successful transition to a ZT paradigm. Organizations can gain a deeper understanding of their data interactions by:

- Beginning with the definition of protect surface(s) and a risk-based strategy in Step 1;
- Mapping transaction flows in Step 2;
- Building and implementing protect surface projects (tailoring the ZTA), that emphasize flexibility and customization to work alongside existing network environments in Step 3;
- Focusing on creating precise ZT policies, addressing the who, what, where, when, why, how, and for how long of access controls in Step 4; and
- Continuous monitoring and maintaining the network (environment) as it enters production (fundamental to the sustained success of a ZTA) in Step 5.

These five steps collectively form the foundation for implementing a comprehensive ZT strategy.

3.3.1 Step 1: Define Your Protect Surface(s)

As you embark on your ZT journey, shift your perspective to focus on what you're protecting rather than what you're defending against. Visualize your end goal and prioritize safeguarding critical and vulnerable components within your protect surface, known as DAAS. Organizations should prioritize identifying protect surfaces, and then document attack surfaces to complement them, steering clear of a traditional, attack-surface-centric approach. Examples include:

- Data: Sensitive information. Examples include:
  - Payment card industry (PCI);
  - Protected health information (PHI);
  - Personally identifiable information (PII); and
  - Intellectual property (IP) that can cause significant harm if compromised.
- Applications: Software interacting with sensitive data or controlling essential assets and processes related to the business.
- Assets: IT, OT, or IoT devices such as point-of-sale (PoS) terminals, supervisory control and data acquisition (SCADA) controls, and networked medical devices.
- Services: Examples include:
  - Domain Name System (DNS);
  - Dynamic Host Configuration Protocol (DHCP);
  - Active Directory; and
  - Network Time Protocol (NTP).

To deploy ZT environments, organizations should focus on two factors: the criticality of the protect surface and the duration of the ZT journey, ideally ongoing. Data classification – as identified above (data sensitivity) – is a critical starting point. Start with low-sensitivity learning protect surfaces, like lab environments or non-critical web pages, allowing for safe experimentation and failure. Progress to practice protect surfaces, which are more sensitive but not the organization's most critical assets. This step-by-step approach builds confidence in ZT principles before moving to the most sensitive areas. After securing high-value assets, the focus shifts to less critical protect surfaces, gradually covering all significant areas in the ZT environment.

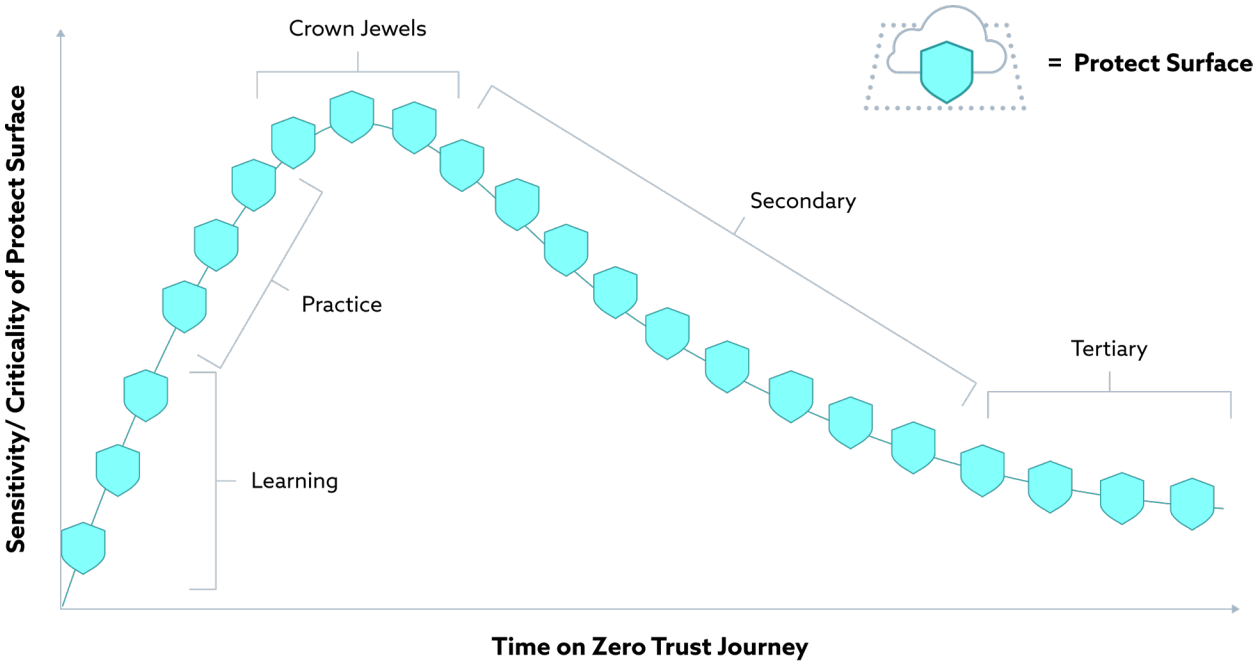


Figure 8: *Zero Trust Learning Curve*[20]

3.3.2 Step 2: Map & Prioritize the Transaction Flows

The fundamental objective of this step is to prove to your audience that you have an understanding of how the whole cybersecurity system works. Mapping transaction flows for each protect surface is critical for understanding how DAAS components interact (how the system works). The mapping is also critical for determining the optimal placement of controls for data protection. These network traffic patterns, specifically tailored to the protect surface data, are essential for shaping the overall design.

Once the transaction flows have been mapped, the next task involves prioritizing, which may also be shaped by the reality of the readily available resources. This process involves determining how resources, such as personnel, time and budget, should be allocated to these prioritized flows to implement the ZTA efficiently.

From a strategic viewpoint, defining protect surfaces and prioritizing transaction flows are necessary inputs to request and allocate necessary resources (e.g., budget and personnel). For example,

[20] Figure adapted from: (2023) *Zero Trust Explained by John Kindervag.*

business processes involving sensitive data should be prioritized, as these processes are crucial for establishing access rights and conditions within the ZTA[21]. To ensure a smooth transition to ZTA, starting with a low-risk business process is advisable, minimizing disruption and gaining valuable experience before moving on to more critical processes.

3.3.3 Step 3: Build a Zero Trust Architecture

Implementing a ZTA through protect surface projects is a journey that modifies the existing infrastructure and processes rather than replaces what has already been implemented. Designing protect surface projects involves mapping transaction flows, identifying controls and secondary protect surfaces, and ultimately designing a system or solution. Even in a completely new environment, transitioning to ZTA within a single technology refresh cycle is improbable. Adapting existing workflows to ZTA likely necessitates, at the very least, a partial overhaul. How an enterprise migrates to a strategy depends on its current cybersecurity posture and operations. Migrating to ZTA requires an organization to have detailed knowledge of its assets (physical and virtual), subjects (including user privileges), and business processes.

Let us put these principles into plain and practical language. The protect surface with the most sensitive assets is in most need of ZT. Temptation: address this surface first. However, the services, assets or business data contained therein might need approvals from more than one department. As we mentioned earlier, as a strategic thinker, you may benefit from delivering a faster or easier win. To establish confidence and trust within the organization, you can opt to improve a protect surface that needs less approvals and less time to complete – the low-hanging fruit.

Another strategy might be to look at protect surfaces where you can build some shared services, or consolidate some technologies. Your benefit here is in showing value and then repeating what you have done on other protect surfaces. With each consecutive instance producing better results in less time. Something complex, such as centralization of Identity Providers (IdP centralization), may be challenging in situations where you are implementing a centralized IdP in a large or complex organization, running legacy systems and diverse application environments. Visible benefits may include simplified management, a better user experience or improved compliance with regulatory requirements[22].

ZT frameworks are not tied to any specific technology, allowing organizations to fully customize their security measures based on their unique protection needs. This flexibility allows for a security approach that is focused on critical protect surfaces within the organization. Dividing the network into smaller, distinct segments to limit access and contain potential breaches, ensuring that even if one segment is compromised, others remain secure heightens security and control over data flow within the organization. Enterprises can adopt various approaches to implement ZTA, emphasizing different components and policy rules. These approaches, namely governance-driven enhanced identity, logical micro-segmentation, network-based segmentation, cloud usage and outsourcing, and even removal of the corporate network altogether, all can adhere to ZT principles.

[21] NIST. (2020). *Zero Trust Architecture* (SP 800-207)

[22] See NSTAC. (2022). NSTAC Report to the President on *Zero Trust and Trusted Identity Management*. Appendix A and B for more ideas.

Typically, a complete ZT solution incorporates many different elements. The suitability of each approach varies, depending on the strategic business direction and risk defined by the board, and should flow down into the architectural solution chosen. While one approach may align seamlessly with a chosen use case and policies, it doesn't imply that other approaches wouldn't work. Indeed, alternative approaches, while more challenging to implement, may provide better long term alignment with the overall strategic business direction.

Depending on the enterprise, multiple ZTA deployment models may be employed within a particular organization for various business processes.[23]

3.3.4 Step 4: Create Zero Trust Policy

ZT policies form the cornerstone of a secure ZTA. While these policies are initially static, they should be designed to evolve dynamically in tandem with the organization's progression in implementing and maturing its ZTA.

To effectively implement ZT, organizations should use the 5 W's plus How for policy creation. This method helps the effort focus on defining granular access controls and considerations for resource access. It also helps you to write specific policy statements and procedures, tailored to the protect surface access perspective. The list below outlines the key aspects that should be factored into any risk evaluation when creating ZT policies:

- **Who:** Determine which entities (people, devices, organizations, code, agents, etc.) should be allowed to access a particular resource.
- **What:** Understand the context in which the entity tries to access systems and/or data
- **When:** Define the time frames or conditions under which the entity may access the resource.
- **Where:** Identify the location, network, or geo-fence that allows the entity access.
- **Why:** Establish why the entity (the "Who") needs access to the resource, emphasizing the justification.
- **How:** Define the technological controls necessary to deliver appropriate risk-based controls to satisfy the 5 W's.

3.3.5 Step 5: Monitor & Maintain the Network

In the CISA ZTMM, visibility and analytics provide the insights that improve ZT operations. Knowing the current and dynamic state of each protect surface's security posture within the network (environment) is critical to any potential response. This involves a focus on logging, monitoring and prompt alerting. These components enable continuous improvement and an effective incident response framework. Regular feedback loops, efficient incident detection, a robust response plan, and the ongoing monitoring of activities are key to maintaining and updating policy rules.

It's also important to regularly review and modify the protect surface and automated policies, which can be achieved through quarterly reviews of ZT identity, devices, access, policies, and protect surfaces.

[23] To learn more about model variations, review these Cloud Security Alliance courses: *Introduction to Software-Defined Perimeter* and *Architectures and Components of Software-Defined Perimeter*.

By continually monitoring and updating each subsequent protect surface, organizations can progressively strengthen their security posture. Such continuous oversight not only enhances security but also improves operational efficiency, speed of access, and flexibility, contributing to overall productivity. Communicating these returns on investment to leadership is essential to acknowledge the long-term benefits of the ZT strategy.

4 Zero Trust & Operations

When organizations conduct a detailed technology landscape assessment, they should identify specific areas where Zero Trust (ZT) principles can and should be applied to optimize or extend existing controls. These enhancements encompass a range of security technologies, including continuous authentication and authorization, user and entity behavior analytics (UEBA), and dynamic policy enforcement points (PEPs). Automation and orchestration, based on the designed Zero Trust Architecture (ZTA), are ZT enablement items.

Here is a list of common operational areas impacted by ZT strategy:

- System administration;
- Network management;
- Data management;
- Performance monitoring;
- Helpdesk and support; and
- DevOps and engineering (access workflow).

This section delves into the multifaceted approach necessary to effectively adopt and integrate ZTA. It also emphasizes the need for a shift in corporate culture, tailored to each organization's unique business type and directorial objectives. Education initiatives are vital for both staff and senior management to understand and communicate the business value of ZT. This educational aspect is pivotal for gaining board buy-in and aligning ZT with the organization's strategic goals.

In response to the evolving cybersecurity regulatory landscape and the inadequacy of traditional security models, ZT offers a proactive and comprehensive framework to protect sensitive data and infrastructure. Organizations need to be aware of regulatory requirements in different regions and adapt their ZT strategy accordingly, especially those with legacy systems. The organization may need to adopt a vendor-based readymade solution to construct the automated workflow to integrate multiple ZTA elements. More orchestration at each step, such as during access and monitoring, can make the operation easier and more adoptable.

Finally, the integration of user experience (UX) and site reliability engineering (SRE) plays a critical role in the successful adoption of ZT. By focusing on UX and automated, code-driven solutions, organizations can foster greater team support, reduce human error, and ensure that security measures are both effective and user-friendly, ultimately enhancing their security posture and operational efficiency.

4.1 Cultural & Organizational Shift

The following list highlights areas for corporate culture shifts, tailored to each organization's specific business type and directorial objectives.

- Cultivate a ZT culture:
  - Emphasize people, processes and organizational aspects over technology acquisition.
  - Implement continuous monitoring, logging and responsive actions.
- Change the tone from the top:
  - Secure executive endorsement and support for ZT initiatives, ensuring leadership commitment.
  - Develop a communications plan for consistent stakeholder alignment and guidance on the ZT journey.
- Instill a culture of continuous risk management:
  - Continuously assess and measure risk to guide access decisions and align with risk appetite.

4.2 Training & Education

Educational initiatives help ensure IT staff, senior management and line-of-business (LOB) managers understand the new ZT paradigm. ZT-informed executives are key to communicating ZT's business value, especially in getting board buy-in. This involves demonstrating how ZT aligns with the organization's strategic objectives. In parallel, it is important to educate the broader workforce. This education should focus on differentiating ZT principles from mere technology tools, helping employees understand the fundamental concepts of ZT. Training reaching the broader workforce should also provide an understanding of revised roles within the ZT framework.

Where applicable, the organization's audit functions (both internal and external) need to participate in the educational process. Auditors need to be informed about how ZT architecture enhances organizational security and resilience.

Lastly, ZT training should be integrated into the existing training program for all staff. This integration ensures that future updates, scheduling and necessary refreshes are consistently applied and not overlooked by the organization's training and education functions.

4.3 Regulatory & Compliance Shift

The cybersecurity regulatory landscape is undergoing a dynamic transformation, spurred by the escalating complexity and frequency of cyber threats. Traditional security models are increasingly inadequate in this environment, prompting governments and industry regulators to endorse proactive and comprehensive frameworks like ZT for safeguarding sensitive data and critical infrastructure.

In this evolving scenario, specific regulations and compliance standards, such as General Data

Protection Regulation (GDPR)[24] and Health Insurance Portability and Accountability Act (HIPAA)[25], are being updated to necessitate the adoption of ZT-aligned security controls. This trend is especially pronounced in the finance, healthcare and government sectors, where the sensitivity of stored personal data heightens the urgency. While not all regulations mandate ZT principles yet, the shift is undeniable in these highly regulated industries, where compliance is not just a legal formality but a critical defense against modern cyber threat.

4.3.1 Regional Regulations

Organizations must stay informed about the regulatory requirements in the countries and regions where they store data and operate. The advent of new regulations often brings the need for specific assessments or attestations, particularly during transitions to ZTAs.

In the United States, for instance, compliance with the Federal Information Security Management Act (FISMA) becomes crucial for US federal government entities, and their suppliers and service providers. This often necessitates optimization and automation of compliance tasks. The reasoning behind this is linked to the requirements of FISMA, which mandates that agencies undergo a rigorous cycle of assessment and reauthorization of systems, especially when making significant changes like adopting ZT. The challenge lies in legacy environments, where agencies frequently find it difficult to keep pace with these demanding tasks, resulting in potential delays or constraints in fully transitioning to a ZT framework.

4.4 Legacy Systems & Infrastructure

Specialized technologies – sometimes legacy-based – such as OT, IoT or industrial control systems (ICS) devices, are often deployed within critical infrastructure services and often have significant technical constraints in key areas, such as patching and access control. This and similar technologies may require implementing specialized micro-perimeter access control technologies and strategies to achieve ZT objectives for such infrastructure.

Organizations with legacy systems and traditional trust models often encounter challenges in adopting ZT, particularly due to limited network and asset visibility. As we have mentioned in other sections, the transition to ZT varies with each organization's unique attributes, including its maturity level, mission and specific challenges. Not all legacy systems require immediate ZT upgrades, but any updates should be strategically planned to address emerging threats and system modernization.

Legacy infrastructure influences the adoption of ZT models. For example, the Information Security Continuous Monitoring (ISCM) model requires adaptable systems for its data movement workflows. Legacy systems' rigidity can hinder the implementation of such models. Additionally, an organization's experience with measurement programs affects its ability to adopt ZT, with more mature organizations adapting more easily than those with less developed measurement capabilities.

[24] General Data Protection Regulation is designed to protect data and privacy of European Union citizens.

[25] Health Insurance Portability and Accountability Act is United States legislation designed to, in part, protect a patient's health information.

4.5 Usability & Friction

This section explores the integral role of user experience (UX) and SRE[26] in promoting the adoption of ZT architecture in organizations. The focus is on refining UX to boost the acceptance of ZT principles while shifting towards a more automated, code-driven approach. This shift not only enhances team support but also minimizes human error, thus strengthening SRE practices. The synergy between UX and SRE ensures that security measures are not only effective but also user-friendly, enhancing the organization's security posture and ensuring smooth operational processes. The key to fostering ZT acceptance among employees is to prioritize UX and implement solutions through code and automation, leading to greater team buy-in and improved SRE outcomes.

4.5.1 User Experience

Incorporating UX helps encourage ZT acceptance and adoption within an organization. A key aspect of this is transitioning from manual processes to code-based automation. By leveraging automation and code, team acceptance is increased, and the likelihood of human error is significantly reduced. This shift improves SRE practices. A well-designed UX ensures that security measures are robust and user-friendly, fostering a more secure and efficient work environment.

4.5.2 Site Reliability Engineering

SRE combines software engineering and IT operations to build scalable and reliable systems. Focused on proactive management through continuous monitoring, automation, orchestration and scalability, SRE planning is a key part of ZT security, helping to maintain system integrity and resilience, including early vulnerability detection and efficient resource management.

Applicable to both cloud-based and on-premises environments, SRE's principles, such as automation, performance monitoring and incident management, universally enhance system reliability, regardless of the hosting setup.

Automation and orchestration (AO) are usually coupled terms, enabling ZT improvement in two important ways. First, AO provide automated feedback that improves access controls, policies, and enforcement, based on feedback loops.

Second, with infrastructure as code (IaC) and automated compliance checks, automated scripts and tools can continuously check compliance with ZT policies, ensuring that any deviations are quickly detected and rectified. AO also enables rapid response to detected threats by automatically adjusting access controls and network configurations in real-time. IaC helps prevent infrastructure drift – the phenomenon where the live state of the network diverges from the state defined in code. This alignment is vital for maintaining the integrity of ZT policies.

4.5.2.1 Monitoring & Understanding System Compromises

In ZT security, monitoring the technology stack is crucial for vulnerability detection, with SRE enhancing this through continuous system monitoring and logging. This approach enables quick iden-

[26] Google. (2016) Site Reliability Engineering.

tification of potential breaches and supports proactive security measures. Additionally, SRE aids in understanding system compromises through postmortem analysis and learning from failures, which is essential for secure recovery and resilience enhancement. Practices like thorough incident documentation and blameless postmortems help teams understand root causes and reinforce system defenses.

4.5.2.2 Resource & Component Management

In the context of ZT security, deploying immutable resources may play a crucial role, and this is where SRE becomes significant. Immutable resources refer to infrastructure components that, once deployed, are not modified. Instead, if changes are needed, new instances of the resources are deployed. SRE facilitates this by automating the deployment process, ensuring that new instances are consistent, reliable, and verifiable. This approach reduces the risk of configuration drift and unauthorized changes, aligning well with the ZT principle of "never trust, always verify." SRE's focus on automation and reliability ensures that deploying immutable resources is efficient and secure.

A decisive and swift response may be necessary when a system component is compromised. This approach is akin to rapidly decommissioning and replacing – effectively and quickly removing and substituting the compromised component with a new, secure instance. SRE supports this rapid response strategy with practices like infrastructure as code and automated deployment pipelines. These practices allow for the quick rollout of new, unaffected instances, minimizing downtime and exposure to threats. By automating the replacement process, SRE ensures that the response to security incidents is fast and reliable.

Conclusion

In Zero Trust (ZT), the levels of strategic engagement include several components. At the top is the organization's strategy, guiding overall actions and decisions. Below this, at the strategy level, ZT redefines traditional trust concepts in computing, emphasizing continuous verification due to the inevitability of breaches.

Aligning ZT with organizational values involves understanding its adoption drivers, like compliance and security enhancement, and how it offers competitive advantages such as streamlined security and cost reduction. Risk management is key, focusing on protecting digital assets and requiring clear ownership for risk handling.

Building a business case for ZT involves assessing financial and performance impacts, gaining cross-departmental stakeholder buy-in, and aligning it with organizational strategy. Tactics for ZT implementation include focusing on specific business outcomes, internal security design, and managing access permissions.

Successful ZT adoption necessitates a cultural shift, integrating continuous risk management, executive support, and comprehensive education across all organizational levels. It also involves adapting to regulatory changes. Overall, ZT is a cybersecurity approach that requires strategic alignment, planning, and execution for full effectiveness.

Glossary

For additional terms, please refer to our Cloud Security Glossary, a comprehensive glossary that combines all the glossaries created by CSA Working Groups and research contributors into one place.

Acronym List

Acronym	Term
AD	Active Directory
AO	Automation and Orchestration
BYOD	Bring Your Own Device
C-Suite	Chief-Suite
CBA	Cost/Benefit Analysis
CFO	Chief Financial Officer
CIO	Chief Information Officer
CISO	Chief Information Security Officer
COO	Chief Operating Officer
DAAS	Data, Applications, Assets and Services
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
FISMA	Federal Information Security Management Act
HR	Human Resources
IaC	Infrastructure as Code
ICS	Industrial Control Systems
IdP	Identity Providers
IoT	Internet of Things
IP	Intellectual property
ISCM	Information Security Continuous Monitoring
IT Ops	Information Technology Operations
LOB	Line of Business
MFA	Multi-Factor Authentication
NSTAC	National Security Telecommunications Advisory Committee
NTP	Network Time Protocol
OT	Operational Technology

PAW	Privileged Access Workstations
PCI	Payment card industry
PDPs	Policy Decision Points
PEPs	Policy Enforcement Points
PHI	Protected health information
PII	Personally identifiable information
PoS	Point-of-Sale
ROI	Return on Investment
SCADA	Supervisory Control and Data Acquisition
SIEM	Security Information and Event Management
SLA	Service Level Agreement
SOAR	Security, Orchestration, Automation, and Response
SOC	Security Operation Center
SRE	Site Reliability Engineering
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TPRM	Third-Party Risk Management
UEBA	User and Entity Behavior Analytics
UX	User Experience
VPN	Virtual Private Network
ZT	Zero Trust
ZTA	Zero Trust Architecture
ZTMM	Zero Trust Maturity Model

Zero Trust Planning

CCZT Study Guide

cloud security alliance®

CCZT™
Certificate of
Competence in Zero Trust

The official location for SDP and Zero Trust Working Group is
https://cloudsecurityalliance.org/research/working-groups/zero-trust/

Disclaimer

Cloud Security Alliance designed and created this Zero Trust Training course study guide (the "Work") primarily as an educational resource for security and governance professionals. Cloud Security Alliance makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Version Number: 20240820

About Cloud Security Alliance

The Cloud Security Alliance℠ (CSA) (www.cloudsecurityalliance.org) is the world's leading organization dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment. Cloud Security Alliance harnesses the subject matter expertise of industry practitioners, associations, governments, and its corporate and individual members to offer cloud security-specific research, education, certification, events and products. Cloud Security Alliance activities, knowledge and extensive network benefit the entire community impacted by cloud—from providers and customers, to governments, entrepreneurs and the assurance industry—and provide a forum through which diverse parties can work together to create and maintain a trusted cloud ecosystem.

CSA Address

709 Dupont St.
Bellingham, WA 98225, USA
Phone: +1.360.746.2689
Fax: +1.206.832.3513

Contact us: support@cloudsecurityalliance.org
Website: https://cloudsecurityalliance.org/
Zero Trust Training Page: https://knowledge.cloudsecurityalliance.org/page/zero-trust-training
Zero Trust Advancement Center: https://cloudsecurityalliance.org/zt/
Provide Feedback: support@cloudsecurityalliance.org
CSA Circle Online Community: https://circle.cloudsecurityalliance.org/
Twitter: https://twitter.com/cloudsa
LinkedIn: www.linkedin.com/company/cloud/security/alliance
Facebook: www.facebook.com/csacloudfiles
CSA CloudBytes Channel: http://www.csacloudbytes.com/
CSA Research Channel: https://www.brighttalk.com/channel/16947/
CSA Youtube Channel: https://csaurl.org/youtube
CSA Blog: https://cloudsecurityalliance.org/blog/

Acknowledgments

Dedicated to Juanita Koilpillai, a pioneer in software-defined perimeters whose contributions to the Certificate of Competence in Zero Trust (CCZT), Zero Trust training, and CSA are immeasurable.

The CCZT and Zero Trust training was developed with the support of the Cloud Security Alliance Zero Trust Expert Group, whose members include volunteers from a wide variety of industries across the globe. Made up of subject matter experts with hands-on experience planning and implementing Zero Trust, both as cloud service consumers and providers, the Zero Trust Expert Group includes board members, the technical C-suite, as well as privacy, legal, internal audit, procurement, IT, security and development teams. From cumulative stakeholder input, the Zero Trust Expert Group established the value proposition, scope, learning objectives, and curriculum of the CCZT and Zero Trust training.

To learn more about the CCZT and Zero Trust training and ways to get involved please visit: https://cloudsecurityalliance.org/education/cczt

We would also like to thank our beta testers, who provided valuable feedback on the CCZT and Zero Trust training: https://cloudsecurityalliance.org/contributors/cczt-contributors

Lead Developers:

Alex Sharpe
Clement Betacorne
Heinrich Smit
Mark Schlicting
Michael Herndon
Michael Roza
Prasad T.
Richard Lee
Shruti Kulkarni
Sky Hackett

Contributing Editors:

Aunudrei Oliver
Emilio Mazzon
Ledy Eng
Matt Lee
Ron Kearns

Expert Reviewer:

Agnidipta Sarkar
James Lam
Jaye Tillson
Robert Morris
Roland Kissoon
Ron Martin (Dr.) , PhD
Vani Murthy
Farid Gurbnov

CSA Global Staff:

Anna Schorr-Campbell
Chandler Curran
Adriano Sverko
Daniele Cattedue
Noelle Scheck
Hannah Rock
Leon Yen
Stephen Smith

Table of Contents

List of Figures

Course Intro

Welcome to Zero Trust Planning by the Cloud Security Alliance (CSA). This training module is part of a larger series titled the Certificate of Competence in Zero Trust (CCZT). In this course, learners will get an in-depth look at the crucial facets of ZT planning, from initial considerations such as stakeholder identification and supply chain risk, to organizational security policies, to compliance. Use cases for prioritization, scoping, and gap analysis are also covered.

Course Structure

This course consists of 8 units, each geared towards helping learners gain competency in the following topics:

1. Starting the Zero Trust journey
2. Planning considerations
3. Scope and priority
4. Gap analysis
5. Defining the protect surface and attack surface
6. Documenting transaction flows
7. Defining policies for Zero Trust
8. Developing a target architecture

Course Learning Objectives

After completing this course, learners will be able to:

- Demonstrate understanding of the ZT maturity model, and how it supports an organization's ZT planning process
- Identify the crucial ZT planning steps and key considerations
- Understand ZT pre-requisites and common ZT use cases
- Possess a working knowledge of how industry-recognized methods (e.g., gap analysis, risk register, RACI diagrams) fit into a ZT planning process
- Demonstrate an understanding of the concepts of protect and attack surface
- Demonstrate understanding of how to map organizational data flows within the scope of the ZT approach
- Demonstrate an understanding of how to plan ZT policies
- Demonstrate an understanding of variables to consider when planning for a ZT target architecture

1 Starting the Zero Trust Journey

Congratulations, your board of directors and senior management are committed to starting the organization's ZT effort. Now your journey begins!

The following roadmap identifies the primary phases of your organization's journey to ZT and maps them to the respective units and sections covered in this module.

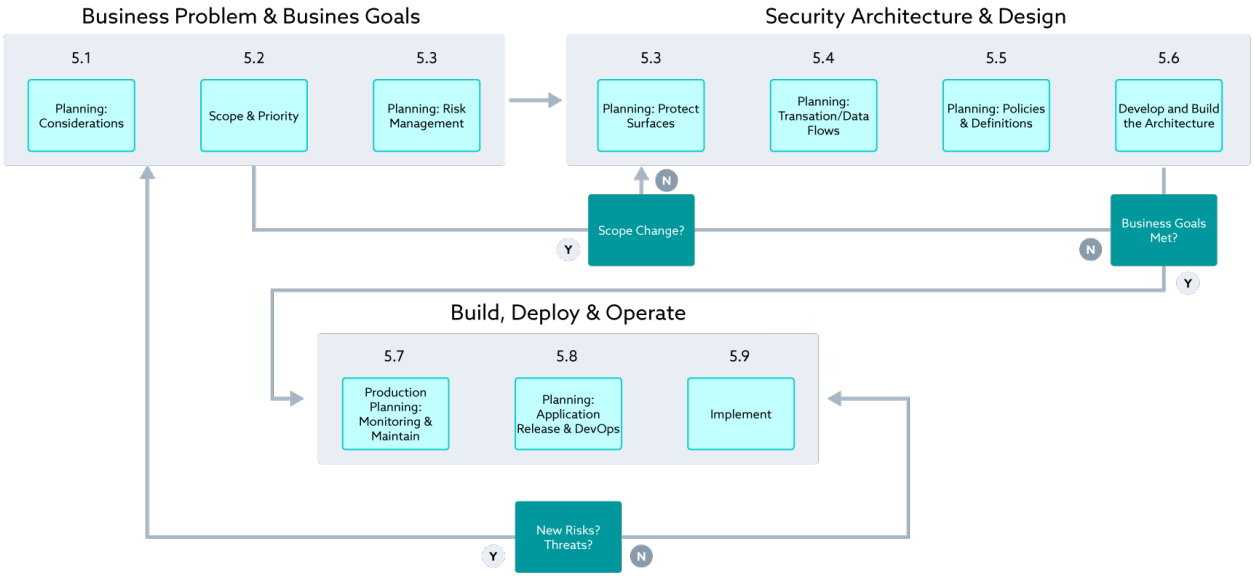


Figure 1: *The ZT Journey Roadmap*

During the course of this module, we explore the various considerations and steps to plan your ZT journey. During ZT planning, the primary focus should be aligning activities and resources to achieve business outcomes, with acceptable risk levels defined by the board of directors and senior leadership.

In this unit we will cover:

- Module assumptions
- Initial considerations

1.1 Module Assumptions

It is most likely that ZT will be implemented in an existing environment with existing controls (in either on-premises, hybrid, or cloud-only scenarios). However, this module applies to completely new implementations as well. While these considerations are similar to implementation in existing environments, they have little or no dependencies on existing business systems and may be implemented more deeply and quickly. Additionally, this module focuses on items specific to ZT and presumes that learners possess foundational knowledge in areas like project planning, enterprise risk

management (ERM), information security (InfoSec), systems engineering, and enterprise architecture design.

The module also assumes an understanding of core ZT principles, including the following:

- **Never trust; always verify:** Claimed identities and authorized entitlements should be verified before access is granted to assets.
- **Inside-out security:** Starting with the assets, the mission-critical elements that are most valuable or vulnerable should be protected.
- **Risk-based security approach:** ZT planning efforts should stem from risk-driven business decisions; that is, assuming budget scarcity, the organization should allocate resources based on risk and opportunity. For example, the decision to protect a specific asset should depend on how it contributes to the company's financial value and its criticality to the organization's mission.

To foster learning and clarity, this course treats the ZT initiative as an atomic unit; in reality, a single organization may pursue a portfolio of initiatives with different motivations and success criteria. In general, larger organizations will likely pursue a portfolio of ZT initiatives based on geography, line of business (LOB), function, regulatory concerns, and more, while smaller organizations may only have a single ZT effort. For example, a multinational, publicly-traded enterprise may pursue three ZT initiatives—one in Europe to address General Data Protection Regulation (GDPR) compliance requirements, another for the firm's manufacturing business in South America, and a third for its U.S.-based operations. In contrast, a privately held small and medium-sized business may pursue a single ZT project to fulfill requirements when bidding for government projects with ZT-related requirements.

1.2 Initial Considerations

A plan for implementing ZT philosophy, approach, and design principles should consider the following five steps, as outlined in the 2022 U.S. National Security Telecommunications Advisory Committee (NSTAC) Report to the President[1]:

1. Define the protect surface: Identify the data, applications, assets, and services (DAAS) elements to protect.
2. Map the transaction flows: Understand how the networks work by mapping the transaction flows to and from the protect surface, including how various DAAS components interact with other resources on the network. These transaction flows provide insight to help determine where to place proper controls.
3. Build a Zero Trust Architecture (ZTA): Design your ZTA, tailored to the protect surface, determined in steps 1 and 2. The way traffic moves across the network specific to the data in the protect surface determines design. The architectural elements cannot be predetermined, though a good rule of thumb is to place the controls as close as possible to the protect surface.

[1] NSTAC. (2022). NSTAC Report to the President on Zero Trust and Trusted Identity Management.

4. Create a ZT policy: Instantiate ZT as an application layer policy statement. Use the Kipling Method[2] of ZT policy writing to determine who or what can access your protect surface. Consider person and non-person (services, applications, and bots) entities.
5. Monitor and maintain the network: Inspect and log all traffic, all the way through the application layer. The telemetry gathered and processed from this process helps prevent significant cybersecurity events and provides valuable security improvement insights over the long term. As a result, each subsequent protect surface can become more robust and better protected over time.

Figure 2: *Five-Step Process for ZT Implementation*[3]

Proper risk management should form the basis of any competent cybersecurity approach, as establishing a framework for identifying and mitigating risks is crucial for minimizing project failure and, in the case of already existing system deployments, disrupting existing systems and business processes. ZT migration tactics depend on the organization's risk profile and risk appetite. For some, ZT design principles will be applied to a limited set of assets; others will apply ZT to all assets across the organization. In either case, the migration to ZT will follow a risk-based, staged approach with numerous iterations culminating in the final transformation into a ZT-driven organization.

Frameworks and models such as the Cybersecurity and Infrastructure Security Agency (CISA) *Zero Trust Maturity Model*[4] can provide organizations at the start of their ZT journey with a reference roadmap for charting their transition towards a ZTA.

[2] NSTAC. (2022). NSAC Report to the President on Zero Trust and Trusted Identity Management. Table 3: Key Zero Trust Foundational Concepts and Definitions.
[3] Figure adapted from: NSTAC. (2022). NSTAC Report to the President on Zero Trust and Trusted Identity Management.
[4] CISA. (2023). Zero Trust Maturity Model (Version 2.0).

1.2.1 CISA High-Level Zero Trust Maturity Model

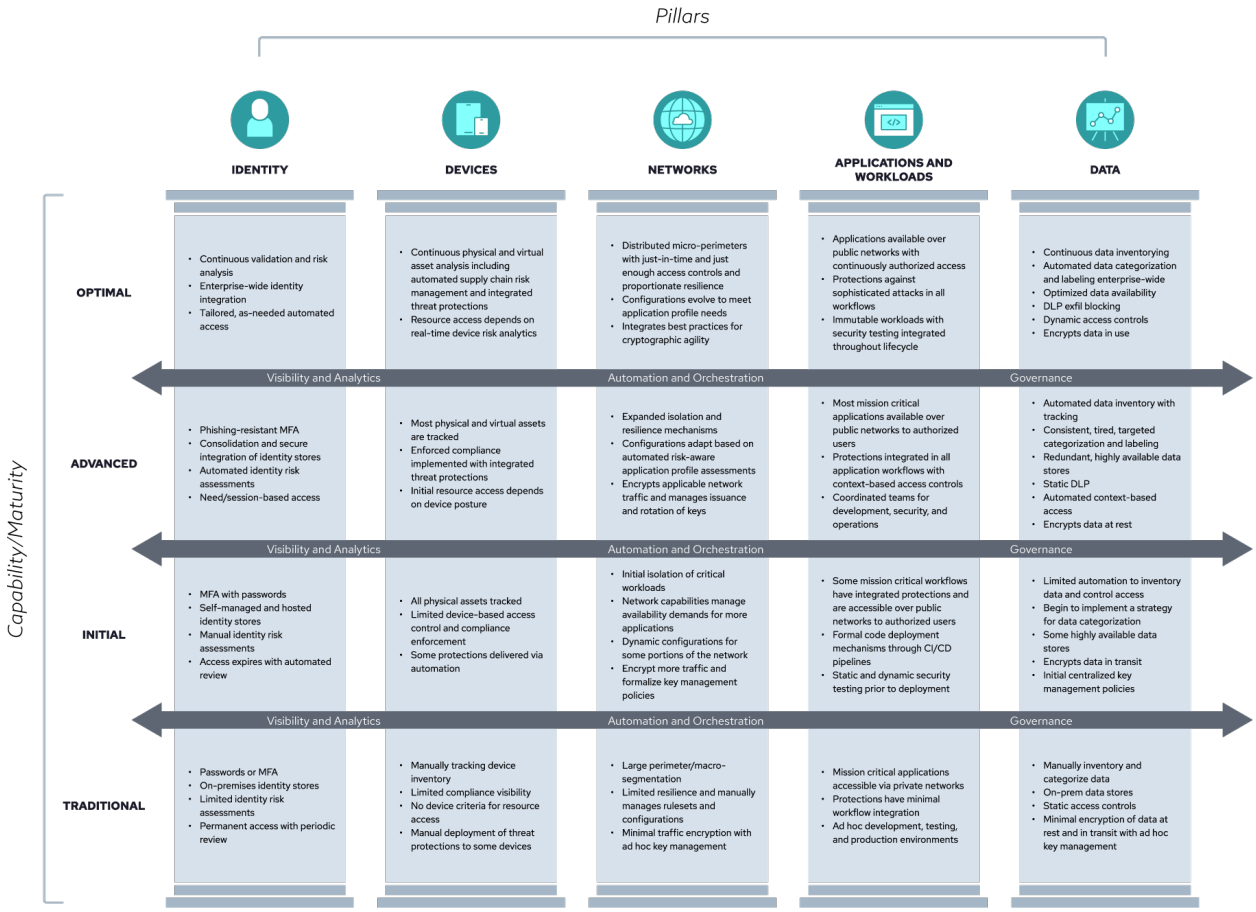


Figure 3: *CISA High-Level ZT Maturity Model*[5]

CISA's *Zero Trust Maturity Model* consists of five pillars, three cross-cutting capabilities, and four cross-functional maturity stages that together form the crucial foundations for ZT. In some diagrams (i.e., from the US Department of Defense[6]), the cross-cutting capabilities (Visibility and Analytics, Automation and Orchestration, and Governance), are depicted as foundational pillars. This representation emphasizes the significance of incorporating these capabilities into the planning process as they assist in defining objectives for the five pillars.

The five pillars are:

- Identity
- Devices
- Networks
- Applications & Workloads
- Data

[5] Figure adapted from: CISA. (2023). Zero Trust Maturity Model (Version 2.0).
[6] U.S. Department of Defense. (2022). DoD Zero Trust Strategy.

In this course, as well as in our other courses, the three capabilities (represented by arrows) are discussed individually due to their frequent need for revision and modification. Additionally, the arrow that lists these capabilities is intentionally repeated several times to highlight the importance of implementing tasks or projects to adapt them as your organization progresses from a traditional model to more advanced security stages.

CISA's *Zero Trust Maturity Model* also consists of several stages. Traditional, the first stage, is where most companies will likely find themselves before embarking on their ZT implementation journey. Attempting to reach the highest level of maturity in a single implementation is impractical and virtually impossible. By incorporating the maturity model into planning discussions, teams will be able to focus on setting clear expectations regarding the desired outcomes for each iterative ZT implementation they commit their resources to.

The four maturity stages, along with a brief example of their criteria, are as follows:

> - **Traditional** - Utilizes multi-factor authentication (MFA), employs manual deployment of threat protection, and maintains an on-premises network
> - **Initial** - Implements MFA with passwords, tracks all physical assets, initiates isolation of critical workloads, and employs formal deployment mechanisms through the CI/CD pipeline
> - **Advanced** - Implements phishing-resistant MFA, tracks most physical and virtual assets, makes most mission-critical apps available over public networks, automates data inventory with tracking, and encrypts data at rest
> - **Optimal** - Engages in continuous validation and risk analysis, grants resource access based on real-time device risk analysis, establishes distributed micro perimeters with just-in-time (JIT) and just-enough access controls, conducts continuous data inventorying, and encrypts data in use

The CISA *Zero Trust Maturity Model* outlines specific examples of Traditional, Initial, Advanced, and Optimal ZTA elements within each pillar. For example, an organization beginning its ZT journey (e.g., it has not yet implemented ZT) may find itself at the traditional tier. According to ZT principles, the organization still would not have enough protection or security even after moving to the initial tier. The two lower maturity stages fail the organization because they lack essential features for a secure and effective ZTA. In this example, the organization would want to move to a future state of advanced and optimal tiers to manage access control effectively. Which would require implementing authentication and identity management, authorization, encryption, monitoring and logging, data protection, and segregation of duties. In later sections (Gap Analysis, Developing a Target Architecture), the CISA *Zero Trust Maturity Model* will be covered more in-depth as a tool for moving across tiers.

2 Planning Considerations

Planners should consider several key factors and variables prior to undertaking the organization's journey to ZT. These include, but are not limited to:

- Stakeholders to engage
- Technology strategy
- Business impact analysis (BIA) results
- The risk register
- Supply chain risk management
- Organizational security policies
- Architecture options
- Compliance requirements
- Workforce training

These key considerations have far-reaching implications on the organization's ZT planning efforts. For example, results from stakeholder identification, BIA, and risk register development activities should dictate how subsequent policies are created.

2.1 Stakeholders

Though seemingly straightforward, stakeholder identification is a critical step that, in practice, requires a significant, concerted time and energy investment. More than any other, this stage can make or break the organization's ZT effort.

Stakeholders include, but are not limited to:

- Business/service owners
- Application owners
- Infrastructure owners
- Service architecture owners
- CISO/security teams
- Legal officers
- Compliance officers
- Procurement officers

Once stakeholders are identified, planning efforts should proceed to mapping out their respective responsibilities, and a communications plan should be developed.

2.1.1 Stakeholder Responsibilities

Stakeholder identification efforts should result in a Responsible, Accountable, Consulted, and Informed (RACI) chart and communications plan. Also referred to as a responsibility assignment matrix, a RACI chart maps out task roles and responsibilities to streamline project management efforts. The RACI chart should reflect the cloud's shared responsibility model, as well as the ability to delegate responsibility, but not accountability—the risk register also shares both attributes.

IT will likely run the organization's ZT initiative daily, with sponsorship by business units, risk management, compliance, or the CISO. Both sponsors and stakeholders should be relevant to the ZT initiative's expected business outcomes. As a starting point, governing documents approved by senior management and the board of directors should designate the executive sponsor and provide insights into reporting expectations.

The most critical ZT-specific role is the asset owner, who will more than likely reside in the business units. As part of their data governance role, the asset owners determine valid users, valid roles, privileges, data usage, and more. Because ZT is an inside-out strategy based on asset value, identifying both assets and asset owners is crucial. However, asset owners should not be confused with asset custodians (e.g., database administrators), who are responsible for implementing directives set by asset owners. Asset owners typically exist in the business while asset custodians are almost always part of IT.

Organizations pursuing ZT should not lose focus on other internal users and groups in human resources (HR), legal, risk management, audit teams, end users, and senior management. An effective, well-informed ZT initiative must consist of stakeholders spread across the organization and at all levels, including functional areas. Functional areas should be consulted or informed, at the bare minimum. For example, HR should serve as the primary source of truth for the organization's identity, while procurement should serve as the source of truth for contractors and vendors. Internal audit, compliance, and the CISO office will likely play crucial roles in the go-live approval.

Bringing in stakeholders across the organization early and keeping them engaged helps the ZT initiative remain well-balanced and focused. To this end, stakeholders should be well-informed of the organization's collective mission and ongoing priorities in order to avoid operational conflicts, aid in prioritization, and ensure the most efficient assignment of resources.

2.1.2 Stakeholder Communications

A communications plan is an essential enterprise tool and is especially critical to a ZT initiative. Because of ZT's prescribed, fundamental philosophical changes and enterprise nature, organizations should develop and adhere to a well-designed communications plan; chiefly, the document should serve as a roadmap for team communications with stakeholders, staff, customers, business partners, and regulators. At a minimum, the communications plan should:

- Define a communication strategy, including tools and any required guidance
- Establish cadence (e.g., forums, format, etc.)
- Incorporate mechanisms for setting proper expectations with interested parties
- Include a means to communicate and document key decisions

2.2 Technology Strategy

Most organizations have a technology strategy consisting of the principles, objectives, methods, plans, processes, and budget for using technology to achieve business objectives. At the beginning of their ZT journey, organizations need to ensure that ZT planning activities are happening in the context of the broader technology strategy. In other words, the ZT strategy and planning need to take into account the existing technology strategy and then update that technology strategy.

During the planning process, organizations should be asking themselves the following essential questions:

- How does the ZT strategy fit into the organization's technology strategy?
- How does the ZT strategy need to be updated to incorporate the technology strategy?
- How does the ZT strategy impact existing plans, processes, and procedures?
- How does the ZT strategy affect existing budgets and investments?
- How does the ZT strategy affect existing internal standards and best practices?

2.3 Business Impact Assessment

Larger organizations operating in highly regulated environments and firms with mature ERM programs are likely to have already carried out a recent BIA. A BIA provides organizations with a list of assets followed by their relative values and owners, valuable information like recovery point objective (RPO) and recovery time objective (RTO), interdependencies and priorities, and an assessment of resources required to restore and maintain each asset. Based on this information, organizations can establish more comprehensive and accurate service level agreements (SLAs), business continuity/disaster recovery (BC/DR) plans, third-party risk management (TPRM) programs, as well as streamline prioritization and stakeholder identification efforts for ZT planning.

2.4 Risk Register

In a similar vein, organizations with a mature ERM or InfoSec program are likely to have developed a risk register containing an inventory of potential risk events, recorded and tracked by likelihood, impact, and description. The risk register should also contain controls for reducing risk levels within the organization-defined risk appetite thresholds, along with the risk owner and the control owner.

With a well-developed risk register, organizations are better equipped to understand what cyber risks their ZT implementation will mitigate. However, the risk register will require continuous updating as the organization adopts new technologies and its infrastructure evolves. For example, the ongoing shift to expanded connectivity and the cloud mandates shared responsibility, and while responsibilities can be outsourced or delegated, accountability cannot. In this case, the risk register must be updated to reflect the cloud's shared responsibility model.

2.5 Supply Chain Risk Management

Modern organizations exist as ecosystem players on a myriad of fronts, from retailer order fulfillment logistics to outsourced human resources. When it comes to technology acquisitions and implementation, the same applies—whether software, hardware, or cloud-based services. Solutions on the market are, to a greater or lesser degree, an assemblage of components developed by third parties. As a result, an organization's visibility into its supply chain is limited by nature, since many components are outside the organization's control. Attestations regarding the validity, security, and quality of third-party components are typically the primary driver behind the organization's technology acquisition decisions.

Crucially, ZT planning considerations should address supply chain risk, since lack of visibility into potential third-party exposures and security glitches (e.g., coding errors, intentional or unintentional hardware or software back doors, unpatched libraries) could result in a data breach or compromise. In the absence of a ZT approach, supply chain participation requires organizations to inherently trust that the initial processes, degree of scrutiny, and approvals to use third-party components in downstream offerings (e.g., hardware, software, or systems) were sufficient; as a result, the required assumption is that the third-party risk of that technology acquisition was and remains acceptable.

Several tools and frameworks can help organizations better understand and mitigate supply chain risk in their ZT implementations. The National Institute of Standards and Technology (NIST) Special Publication (SP) 800-207 presents ZT tenets that apply to a supply chain and its supplier organizations across all ZT pillars, namely Identity, Devices, Networks, Applications and Workloads, and Data. Since 2018, the National Telecommunications and Information Administration (NTIA) has pursued the concept of a software bill of materials (SBOM) as a tool for advancing supply chain risk management, with CISA having announced its intent to support this work. This effort aims to create a mechanism for exposing software components, enabling organizations to make better risk decisions when deciding what software products to incorporate into their products or offerings.

Additionally, the following non-exhaustive list of tools and resources can help organizations in determining supply chain risk:

- CSA STAR Program[7] (STAR Level 1 and STAR Level 2)
- ISO 27001 assessments
- SOC 1 and 2 assessments
- Systems audits
- Bridge letters & attestations
- Supplier organization and service offering reputation research

With these frameworks, tools, and resources at their disposal, organizations can apply ZT principles in evaluating potential supply chain risk exposures more comprehensively and effectively.

2.6 Organizational Security Policies

ZT planners should keep in mind that various policies will change across all domains (e.g., HR, identity and access management [IAM], technical, and privacy). Also, pre-existing policies affect how ZT will be implemented. From a ZT perspective, organizational policies affecting identity, devices, networks, applications and workloads, and data should be considered for updates, at the very least.

These policies are designed to provide direction across the enterprise. The policies updated (or created) for ZT will be provided to the team(s) for implementation.

The most relevant policies will fit into roughly three categories:

1. Policies that dictate or constrain the ZT initiative
2. Policies that require updating due to ZT
3. Policies that need to be created to support ZT

[7] https://cloudsecurityalliance.org/star/

While the list of relevant policies and how they fit into each category vary widely between organizations and potentially groups within organizations, the following are common policy types for a ZT initiative:

- General IT and security
- ZT
- Data governance
- Cloud
- Key management policy
- Incident response
- User and IAM
- Monitoring
- Disaster recovery (DR)
- Business continuity (BC)

2.7 Architecture

During the ZT planning process, especially in the early stages, planners should identify the relevant architecture capabilities and components that could impact ZT or require updating due to ZT. These capabilities may include architectural frameworks such as The Open Group Architecture Framework (TOGAF)[8], Sherwood Applied Business Security Architecture (SABSA)[9], CSA's *Enterprise Architecture Reference Guide*[10], and other less formal frameworks and standard organizational configurations. It is also necessary to identify key components such as architecture requirements repositories, architecture landscapes, solution landscapes, and standards information bases. Architecture will be discussed in greater depth in later sections.

2.8 Compliance

At this time the United States is at the global forefront in the pursuit of ZT. For instance, U.S. government agencies have produced artifacts that provide critical ZT guidance like the NSTAC *Report to the President on Zero Trust and Trusted Identity Management*, and the NIST Cyber Security White Paper (CSWP) 20[11], to name a few. Other jurisdictions like Europe and Asia are also preparing ZT guidance or regulations.

However, even without fully realized ZT-based regulations and laws, the ZT approach can be invaluable in achieving compliance with existing cybersecurity and data privacy laws and regulations. A ZT approach will be helpful in two ways:

- First, it will increase control over regulated data by enforcing controls that foster accountability and by segregating data within dedicated micro-segments.
- Second, it will drive better overall cybersecurity, which in many cases exceeds most existing legal and regulatory requirements.

[8] The Open Group Architectural Framework. (2022). The TOGAF Standard, 10th Edition.
[9] Sherwood Applied Business Security Architecture. (2009). SABSA White Paper (W100).
[10] Cloud Security Alliance. (2021). Enterprise Architecture Reference Guide.
[11] NIST. (2022). Planning for a Zero Trust Architecture: A Planning Guide for Federal Administrators (CSWP 20).

Implementing ZT across an organization or system impacts all in-scope architectures. This implies that potentially every system, control, and process will change. These potential impacts across the organization should be kept in mind during the planning phase. Potential impacts, considerations, and updating needs may surface in unforeseen areas (e.g., infrastructure support, incident management, BC/DR, and end-user support).

2.9 Workforce Training

As one the most critical aspects for the success of a ZT journey, training is undoubtedly a key component of every cyber security program and represents a foundational component of the ZT approach, despite being often left to the last minute. Because ZT is essentially a paradigm shift, it will benefit from a shift in existing training programs along with the typical user training that accompanies any technology rollout. Your organization most likely has a training and awareness program and a security awareness program. ZT principles should be integrated into the security awareness program in all phases — onboarding, role changes, yearly reviews, drip feed, and termination.

Special attention needs to be paid to the training of the:

- Staff who determines access controls
- Staff who configures the access control rules
- Support Team, including the Help Desk, who need to be ready to handle the paradigm shift is paramount to a smooth transition
- Staff who audit what has been done, including IT audit and security audit
- Upper management who need to fully embrace the cultural shift that ZT might impose

Finally, it is important that the board of directors and CEO have the necessary level of awareness to be able to fully understand the progress and challenges of the ZT project.

3 Scope, Priority, & Business Case

As mentioned at the start of this module, the ZT approach should be regarded as a journey of several stages, eventually taking the organization to a state where the business operates on a ZT model. Each stage of the journey should be seen as an individual project.

The organization may start its ZT journey with a project focused on a critical protect surface, then expand onto the rest of the organization's protect surfaces. In the event that the organization has identified several critical protect surfaces, the key questions are prioritization-related: Where does the ZT journey start? How does that define the scope of the initial ZT project? The answers to these questions are discussed in the next section, while protect surfaces are covered more in-depth in a later section.

ZT concerns securing the protect surface to reduce the risk of data and process-compromise. To do this, the protect surface's data, assets, processes, and the identities that access it must be comprehensively understood and mapped out. The organization may choose to include one or several protect surfaces in a project.

Regardless of the number of protect surfaces in a project, organizations at this stage should start with the prerequisites for understanding the protect surface, followed by a definition of the ZT project's scope and priorities, and lastly, a development of a business case. The last section in this unit provides several examples for assessing scope and priorities per use case.

3.1 Prerequisite to Understanding the Protect Surface

The first step in defining the priorities within a ZT journey is understanding what the organization wants to protect using a ZT approach. In other words, the starting point for prioritizing ZT efforts is identifying the data and assets that the organization seeks to secure, the location of the data, the asset where data is hosted, and the services, processes, and classifications.

To this effect, several prerequisite actions need to take place in order to have a clear understanding of the organization:

- Data and asset discovery and inventory
- Data and asset classification
- Entities/user discovery and inventory

3.1.1 Data & Asset Discovery & Inventory

To effectively protect its data, the organization needs to know where that data resides. This can be achieved with data discovery activities. At a minimum, more mature organizations will have a current asset inventory that contains a list of itemized data, devices, applications, services, and more, followed by an assessment of each asset's value.

Ideally, the asset inventory will exist in the form of an up-to-date, automatically updated configuration management database (CMDB) that contains all the relevant information about the assets (e.g., hardware, software, devices, etc.) and the inter-component relationships. As ZT is driven by the asset's value, the CMDB should be viewable based on these models and parameters.

Alternatively, in the absence of an asset inventory, the organization may decide to run a data discovery activity using automated tools, followed by the population of a CMDB with the metadata obtained during the data discovery activity.

3.1.2 Data & Asset Classification

With a new or existing asset inventory on hand, the organization must classify data and assets based on the sensitivity of data handled by the business transaction. Data and asset classification activities are meant to be a prerequisite for any ZT project. This helps in the identification of the protect surface and enables organizations to plan for the proper security controls in their ZT implementations. It also plays a crucial role in identifying relevant regional laws and regulations that may apply to the organization.

3.1.3 Entities/User Discovery

The discovery of entities, both person and non-person users, is another essential prerequisite before the scope of a ZT project can be defined. In order to be able to access the organization's IT assets and data and carry out business transactions, those entities need to have an identity assigned and eventually a set of different personas.

These entities may be person or non-person users (e.g., machines, service accounts, APIs). Organizations should understand whether the entities run transactions in the background and whether they are authorized to run the transactions after being authenticated. The discovered entities should be mapped to all relevant protect surfaces, (creating an inventory of entities/users) and their identities used to define the ZT policies discussed in later sections.

3.2 Scope

Once the prerequisites are met, the organization needs to define the scope of the ZT project. Scope would typically include:

- Success criteria identified for the ZT projects
- Business units that are identified for the ZT journey
- Protect surfaces that are part of the business units, including identification of:
  - The data and the assets that are part of the protect surface
  - The identities that access the protect surface
  - The entities mapped to the identities/personas

3.3 Priority

Once the scope is identified, the organization needs to determine how and in what priority to implement ZT. Some approaches to this include the following:

- **Prioritization based on complexity:** Building from simple to complex, the organization may choose to select a smaller, simpler protect surface as a pilot project and progress to more complex protect surfaces. This approach allows a better understanding of the ZT project life cycle, document learnings, and apply them to the next set of protect surfaces. Starting small and simple makes it easier to apply the relevant planning considerations.
- **Prioritization based on risks:** Selecting a protect surface high on the risk register may help in scenarios where the organization has experienced security compromises or incidents involving protect surfaces. This approach may help reduce any cyber risks brought about by access control. After completing the high-risk protect surface projects, the organization may then move on to lower-risk projects.
- **Prioritization based on use case:** This approach is suitable for organizations with a definite use case in mind. Use case examples are provided later in this unit.

3.4 Development of a Business Case for ZT Planning

After the identification of data, assets and identities, classification of data, and the critical processes are identified, the organization can move forward and define a business case that would justify why a certain asset should move under the protection of a ZT approach.

The business case is supposed to be briefed and approved by senior leadership and most likely, the board of directors. This will outline expectations, motivations, funding, and any other requirements that the team may choose to share.

Most mature organizations have existing business case templates which should be utilized for this purpose. Factors to be considered in the business case would include:

- The BIA
- The risks that the ZT program is designed to address
- The cost of the project (e.g., capital costs, operational costs, resourcing and administration costs)
- The cost of not doing the project (i.e., the impact of not implementing ZT), to include costs incurred due to any data breaches or security incidents involving access controls
- What the organization stands to gain through ZT (e.g., ease of access administration, reduction of the visible attack surface, and more)
- Additional benefits that come about through improving the organization's security culture

ZT adoption may help the organization position itself favorably among competitors. For example, a software as a service (SaaS) provider may include ZT in its marketing collateral and sales materials to demonstrate the optimal security posture of its platform as well as its forward-thinking commitment to protecting customer privacy.

3.5 Use Case Examples

The following use case examples can help organizations anticipate priority and scope-related concerns regarding specific access types and environments.

3.5.1 Role Based Access Control for Internal Staff

Assuming that the organization has implemented network segmentation, network zones, and micro-segmentation with different security requirements, administrators can define policies accordingly. For example, a soap manufacturing company may place all the trade secrets related to soap recipes and formulas in a network segment that only server administrators and recipe/formula engineers can access. Implementing ZT means that any malicious movement using compromised credentials is preempted with device verification, thus securing trade secrets.

3.5.2 Remote Access

Remote access is the new normal way of working. Remote access users include (but are not limited to) employees, contractors, temporary staff, suppliers, etc. Remote access also opens the possibilities for lateral movement via compromised access controls. Administrators mitigate this risk with technology like virtual desktop infrastructure (VDI) and/or corporate cloud workstation resources and by publishing applications and resources. However, application jailbreaking may be a residual risk in these scenarios.

Using ZT, administrators can define policies such that remote workers access only those applications and resources for which they are authorized. This reduces the attack surface that is available to remote workers. The attack surface can also be reduced with device authentication before granting access to users. As you may recall, device authentication relates to ZT's "verification before granting access." Administrators may also integrate opportunistic MFA with their ZT controls for behavior analysis and geofencing.

3.5.3 Services Accessed Using Mobile Devices

Organizations subscribe to services that can be accessed from mobile devices (e.g., smartphones, tablets). Services like HR portals, portals for salary/wage slips, and office directories can be accessed via mobile applications and web applications run on browsers. To ensure that compromise of users' credentials do not lead to compromise of data, ZT policies can aid in authenticating the users and their devices before granting access to the services. Additionally, MFA can be configured with ZT. As ZT policies can be made as granular as possible, separation of duties between the users and administrators help prevent any privilege escalation attacks. A caveat is that it should not be assumed that ZT can prevent access to these services via stolen devices.

3.5.4. Third-Party Service Providers with Remote Access

Administrators can leverage ZT policies to authenticate third-party users and their devices to determine the required access privileges for resources while hiding all other assets to prevent any lateral movement. This helps reduce the attack surface for any supply chain risk materialization.

3.5.5 Staff Access to Assets in Hybrid Environments

Staff access to root accounts for cloud services such as AWS and Azure should be tightly controlled. Lack of awareness or speed to market may make staff miss out on controls like configuring MFA for such resources. Administrators can configure ZT policies for such accounts and subscriptions, thus ensuring that the same policies are applied to all accounts. In addition, these accounts and subscriptions remain hidden behind the policies leading to reduced visibility in the public domain resulting in reduced attack surface.

3.5.6 SaaS & PaaS

SaaS and platform as a service (PaaS) require access at two levels. One is access to the service and the second is access to the features within the service. Implementing ZT will help define attribute-based access control (ABAC) for the features within the service. For example, granting database administrators (DBAs) access to the master database in a SQL database-as-a-service but not to data persisted in user databases.

The applications are often consumed for managing the organization's private or sensitive data. It is important to ensure that only legitimate users can access the application, though it is a cloud-hosted one. ZTA can be designed such that access to the application or platform is allowed only for the traffic coming from the ZT gateway. Thus the user and the entity can be subjected to the validations and policies before sharing access to the assets. The design can be achieved via SAML authorization, where the SAML requests from the gateways alone are accepted at the SAML service provider residing at the SaaS/PaaS application.

3.5.7 Application Release & DevOps

High-velocity application release practices like DevOps and its supporting automation and continuous integration/continuous delivery (CI/CD) framework require thoughtful integration with ZTA. ZTA can be integrated with DevOps to secure authorized connections to the various deployment environments (e.g., development, test, staging, and production) to ensure proper connectivity to protected servers and applications. ZTA can provide a better developer experience by streamlining access provisioning. Ideally, ZTA should be integrated into the application stack to fully leverage its security features.

During planning for ZT implementation, the following usage areas need to be considered:

- Secure remote access during the application release cycle
- Access to individual protected servers and applications
- Integration of ZTA into the application stack

Common DevOps practices such as the use of virtualized environments and containers can streamline ZTA integration; that said, security architects must fully understand the chosen ZTA deployment model and how their organization's DevOps mechanisms will interact and integrate with it. When it comes to DevOps toolset integration, security teams should carefully review and evaluate third-party APIs and repositories supported by their ZTA implementation.

3.5.8 Industrial Control Systems, Operational Technology, & Internet of Things

Industrial control systems (ICS), operational technologies, and the Internet of Things (IoT) rely on generic non-user identities (service accounts, resource accounts, roles, etc.) to access resources. However, these identities can be enabled with interactive logon rights for users—a feature that can be potentially compromised or abused. Furthermore, investigating security events involving interactive

logon rights is challenging, as logging only records generic identity names, not the name of the user behind the generic identity. Implementing ZT in these environments ensures that identities have only the required access to assets for the task at hand, thereby limiting the attack surface in the event the identities are compromised.

4 Gap Analysis

A gap analysis is an industry-accepted tool that allows organizations to determine how to best realize their objectives. At its core, a gap analysis is a three-step process that compares where the organization is with where it wants to be and then defines a road map to close the gap. Most organizations have a preferred gap analysis framework like Strengths, Weaknesses, Opportunities and Threats (SWOT)[12], the McKinsey 7-S Framework[13], or the Nadler-Tushman Congruence Model[14], to name a few. Depending on the size of your enterprise, you may undertake several gap analyses for different business units, geographies, and functions.

A gap analysis consists of the following steps:

- Determine current state
- Determine target state
- Create a roadmap to close the gap
- Requirements

4.1 Determine Current State

The first step in the gap analysis is to make an objective, comprehensive assessment of the organization. Ideally, prior third-party assessments, maturity models, frameworks, and other existing resources can help inform this effort. For example, the *CISA Zero Trust Maturity Model* provides organizations with a framework to assess their current state regarding ZT adoption.

The following are crucial steps for determining the organization's current state:

- Define the current protect surface(s) and the implications for each ZT pillar: Identity, Devices, Networks, Applications & Workloads, and Data
- List current controls for each pillar, focusing on the protect surface for each respective pillar
- Determine and declare the current CISA maturity stage for each pillar

For example, an organization defining its current protect surface regarding data has determined that most of its data-at-rest is being stored unencrypted. Additionally, the organization is still using traditional password-based authentication for its systems and continues to rely on local authorization for security access to application workloads.

[12] Humphrey, A. (1960). SWOT Analysis.

[13] McKinsey & Company. (2008, March 1). Enduring Ideas: The 7-S Framework. McKinsey Quarterly. Retrieved 2023, January 20.

[14] Nadler, D., and Tushman, M. (1980). A Model for Diagnosing Organizational Behavior. Organizational Dynamics 9, no. 2.

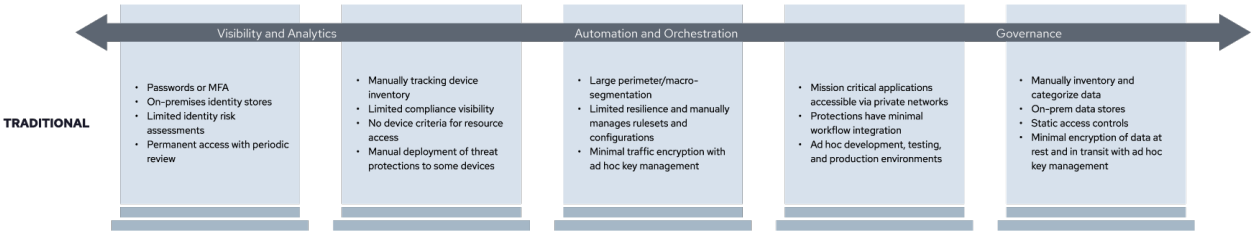


Figure 4: *CISA Zero Trust Maturity Model: Traditional*[15]

Per the CISA *Zero Trust Maturity Model*, a firm storing its data unencrypted falls into the Traditional tier; the same is true of this organization's application workload and identity protect surfaces. Part of this process also involves determining risk appetite, which should feed into scoping activities and decisions when the future state is decided/selected.

4.2 Determine the Target State

Once the planner, user, or organization has a solid understanding of the current state, the next step in the gap analysis is to determine the target state. During this second phase, the goal is to: Define the protect surface and the impact for each in-scope pillar across the organization (i.e., what each should look like when ZT has been implemented).

Determine and declare the desired target CISA maturity stage for each pillar. The CISA *Zero Trust Maturity Model* represents a gradient of implementation attributes across five distinct pillars, where minor advancements can be made over time toward optimization.

Regarding the previous example, the organization may determine that achieving an Optimal ZT maturity stage, while ideal, may be prohibitive due to several factors. The organization may require more long-term vetting of AI/ML technologies and may be unable to encrypt all of its stored data across each environment. The organization may elect to adopt MFA as its future state to bolster the identity protect surface, and to start with encryption at rest for cloud and remote environments to bolster the data protect surface.

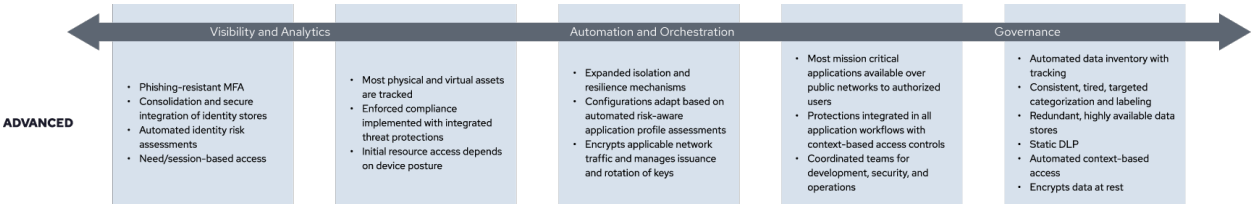


Figure 5: *CISA Zero Trust Maturity Model: Advanced*[16]

The organization's target ZT maturity stage and future state should ultimately fall under the Advanced or Optimal tier. However, achieving this requires a gradual evolution through incremental steps, first exhibiting characteristics of the Initial tier, then proceeding to the Advanced tier, and finally reaching the Optimal tier in all areas. As risk appetite was defined when determining the

[15] Figure adapted from: CISA. (2023). Zero Trust Maturity Model (Version 2.0).
[16] Figure adapted from: CISA. (2023). Zero Trust Maturity Model (Version 2.0).

current state, the scope is defined when determining the target state based on those previous assessments.

4.3 Create a Roadmap to Close the Gaps

Once an organization knows where it is and where it is going, it should create a roadmap of the required future state across all pillars. During the roadmap phase, the organization should compare the current state and maturity stage to the desired state and maturity stage, listing the future controls required to raise the current maturity stage to the future desired state.

For example, the organization mentioned previously must now plan for bridging the identified gaps between its current Traditional maturity stage and target state. By establishing the foundational ZTA, the ZTA can evolve to effectively manage access control, implement more encryption, increase data protection, and add segregation of duties and principles. Once this is integrated into operations, detailed risk assessments can be made, along with appropriate response plans for all risks related to compromised data or access control points.

The roadmap should include all the controls and procedures necessary to bring the organization from a Traditional to an Advanced maturity stage. As a simple example, in the Traditional approach, it is perhaps enough to have a login screen and get to enterprise-sensitive data or business functionality with a single password. For your initial ZT implementation (Initial stage), you couple the login screen with MFA requirements. In a subsequent ZT implementation project (an Advanced stage), you add safeguarding technologies so that the MFA process cannot be leveraged by a bad actor for phishing attacks. In a ZT implementation that further advances your login and MFA processes, you add enterprise-wide agents that can track network activity (Optimal stage). Now, your organization can set up monitoring consoles and security experts can be flagged to spot seemingly dangerous activity, enabling them to take corrective measures before a security breach can occur.

4.4 Requirements

One of the key outputs of the gap analysis will be requirements for ZTA implementation. There is a large body of work for requirements analysis. The following section focuses on key items unique to ZT.

How to define and document your requirements will largely depend on whether your ZT effort is stand-alone or part of a larger effort. If ZT is part of a larger effort, it is recommended to collect your requirements in a ZT-specific section to maintain focus. This may not be practical in all situations but should be the objective. If ZT is a stand-alone effort, you have the luxury of a dedicated requirements document that can be used by the project team.

Either way, a primary focus in the early phase(s) of planning will be to solidify your identification, entitlement, and access control infrastructure. At a minimum, you want to be sure you have requirements defined for:

- Source of truth for unique identities
- Management of those identities through the full life cycle of employees, contractors, and vendors

- Definition, provisioning, and management of entitlements
- Definition, provisioning, and management of access controls
- Segmentation/micro-segmentation
- Incident detection and response
- Reporting and analytics
- Special considerations (e.g., devices)
- Concept of least privilege
- Segregation of duties

5 Define the Protect Surface & Attack Surface

The following unit covers the crucial activities for identifying the protect surface and attack surface, outlines how the protect surface and attack surface are interrelated, and provides key considerations for designing the two surfaces to complement each other.

5.1 Identify the ZTA Protect Surface

Malicious actors compromise data confidentiality, integrity, and availability via improper access. Subsequently, ZT aims to reduce cyber attacks and data breaches through more stringent access requirements, that is, by requiring authentication and authorization prior to granting access to resources. Hence, to reduce cyber risk in this manner, organizations must understand and identify data and their locations. As data cannot exist in a vacuum and needs a house (i.e., the asset) to live in, the data and asset both need to be identified, as well as their respective criticality levels.

Extending this premise to the organization at large, ZT planners should define what needs to be protected in an organization, also known as its protect surface. NSTAC defines the protect surface as the area the ZT policies protect. Each protect surface contains a single DAAS element, and in turn, each ZT environment will have multiple protect surfaces.

5.2 Identify the Attack Surface

Along with defining and defending the protect surface, organizations should also define the attack surface—the surface through which data and assets can be attacked. NIST defines an attack surface as "The set of points on the boundary of a system, a system element, or an environment where an attacker can try to enter, cause an effect on, or extract data from, that system, system element, or environment." [17]

Organizations more often experience difficulties in defining the attack surface for defense purposes, since defining an attack surface at a given point in time for most organizations can be a moving target due to the evolving usage patterns of devices and assets (e.g., BYOD, SaaS). In contrast, a protect surface has a defined boundary.

[17] NIST. (2018, October). Glossary: attack surface. NIST Computer Security Resource Center. Retrieved 2023, January 20.

The diagram below illustrates the Zero Trust Architecture as defined by NIST, originally in the SP 800-207[18], and then further elaborated in SP 1800-35B[19] draft.

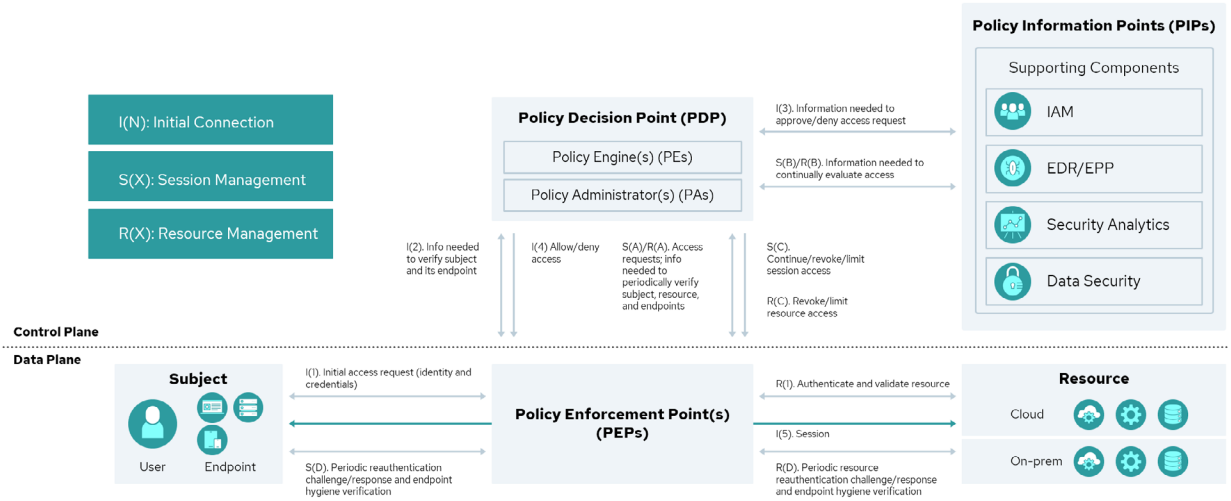


Figure 6: *General ZTA Reference Architecture*[20]

Applying the NIST definition to the above diagram, the attack surface is identified as follows:

- Endpoint
- Information flow between the endpoint and policy enforcement points (PEPs)
- Information flow between PEP and resource
- Information flow between PEP and policy decision point (PDP)
- Information flow between PDP and policy information points (PIPs)
- Policies
- Identity and access management used by ZT
    - End users' identities
    - API identities
- The application stack for PEP, PDP, and PIP
- The solution in the supply chain

The identified attack surface can be analyzed for threats, abuse cases, and mitigations as illustrated in the table below, an example of attack surface-focused threat modeling using STRIDE, a common threat modeling methodology championed by Microsoft.[21]

[18] NIST. (2020). Zero Trust Architecture (SP 800-207).

[19] NIST. (2022). Implementing a Zero Trust Architecture (SP 1800-35B). Second preliminary draft.

[20] Figure adapted from: NIST. (2022). Implementing a Zero Trust Architecture (SP 1800-35B). Second preliminary draft.

[21] Microsoft. (2009, November 12). The Stride Threat Model. Microsoft Learn. Retrieved 2023, January 20.

Attack Surface	Threat	Abuse Case	Mitigation
• Information flow between PEP and Resource • Information flow between PEP and PDP • Information flow between PDP and PIPs	• Spoofing • Tampering • Information disclosure	• A malicious actor can perform a man-in-the-middle attack to spoof the user of the resource • A malicious actor can intercept and manipulate data on the information flow channel between the resource and the PEP, between the PEP and the PDP, and between the PDP and the PIP	• Use TLS certs certifications as described in the Encryption section • Use mTLS for 2-way authentication
Endpoint and its environment	• Spoofing	• A malicious actor (malware/phishing attack), may try to harvest credentials used by the endpoint to log into the ZT endpoint agent	• Onboard the ZT endpoint agent as described in the earlier modules • Employ user- based / machine- based certificate for authentication (as recommended in Introduction to Zero Trust Architecture)
Policies configured on PIP/PEP/PDP	• Tampering • Information disclosure	• A malicious insider can try to add/ modify policies	• Use the supplier due diligence process to check if this is a possibility in the vendor's environment- (background checks, RBAC on the backend, etc.) • Logging and possible sharing of logs with customers

PDP and PIP administration console	• Spoofing • Elevation of privileges • Repudiation	• A malicious actor may try to spoof administrators to access the administration console for policies	• Use MFA to address spoofing threats via credential harvesting • Check for assurance from the vendor that an administrator cannot access the data belonging to a different organization • Logging of all actions carried out by an administrator with a possibility of sharing the logs with the customers
IAM for ZT users	• Spoofing	• Identity providers may become compromised leading to the harvesting of users' credentials by malicious actors	• Identity provider is assessed for security to make sure it is fit for purpose

Supply chain of ZT	• Spoofing • Tampering • Repudiation • Information Disclosure • Denial of Service (DoS) • Elevation of privileges	• Lack of governance, risk, and compliance in the ZT organization leading to lack of line-of-sight visibility to security posture in the organization. • Insider threat leads to the exfiltration of customer data • Vulnerabilities in the management plane / software components leads to the compromise of policies • Lack of vendor controls results in DoS for customers • DoS at the application layer • Lack of OS hardening, secure configuration, host-level intrusion detection, and network layer intrusion detection • Vulnerabilities on the management console lead to SQLi, XSS, and lateral privilege escalation for customers	• Information security management system implemented and practiced in the organization • Conduct background checks for the staff that works with customer data • Vulnerability management program to upgrade and update technologies that compromise the ZTA • Secure SDLC to ensure the management console is developed securely • Web application firewall (WAF) drops any packets that can result in DoS at the management console • Underlying infrastructure components (server endpoints, web servers, application servers, containers, etc.) are hardened, secure configuration is applied, host intrusion detection is enabled, file integrity monitoring is enabled, etc.

Table 1: *Example STRIDE Threat Model*

5.3 Illustration of Protect Surface & Attack Surface

The credit card example below illustrates the protect surface and attack surface:

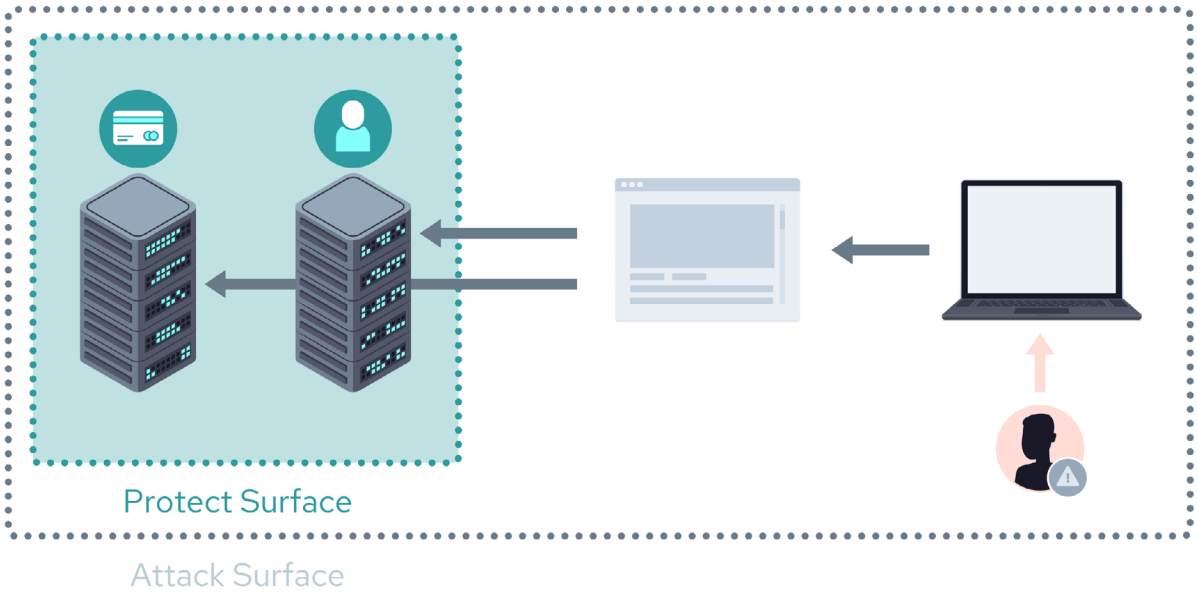


Figure 7: *Attack Surface and Protect Surface: Credit Card Example*

The cardholder data, personally identifiable information (PII), and underlying assets (i.e., server) comprise the protect surface, since any unauthorized access to the assets may subsequently lead to unauthorized data access, resulting in a data breach. Consequently, these assets should be covered with ZT policy that requires entity verification for asset and data access to ensure that such breaches do not occur.

Then the organization permits end-users and administrators to access cardholder data and PII housed on the server via an application accessible through their laptop's browsers. These laptops, browsers, and servers are potential entry points for malicious actors; any vulnerabilities or lack of hardening on the asset may enable malicious actors to compromise the server and data. Hence, the attack surface encompasses the end-user and administrators' devices and applications, as well as the protect surface.

The attack surface can increase with the addition of another laptop and a cloud service. This is because the entry points to the data increase with added assets and devices.

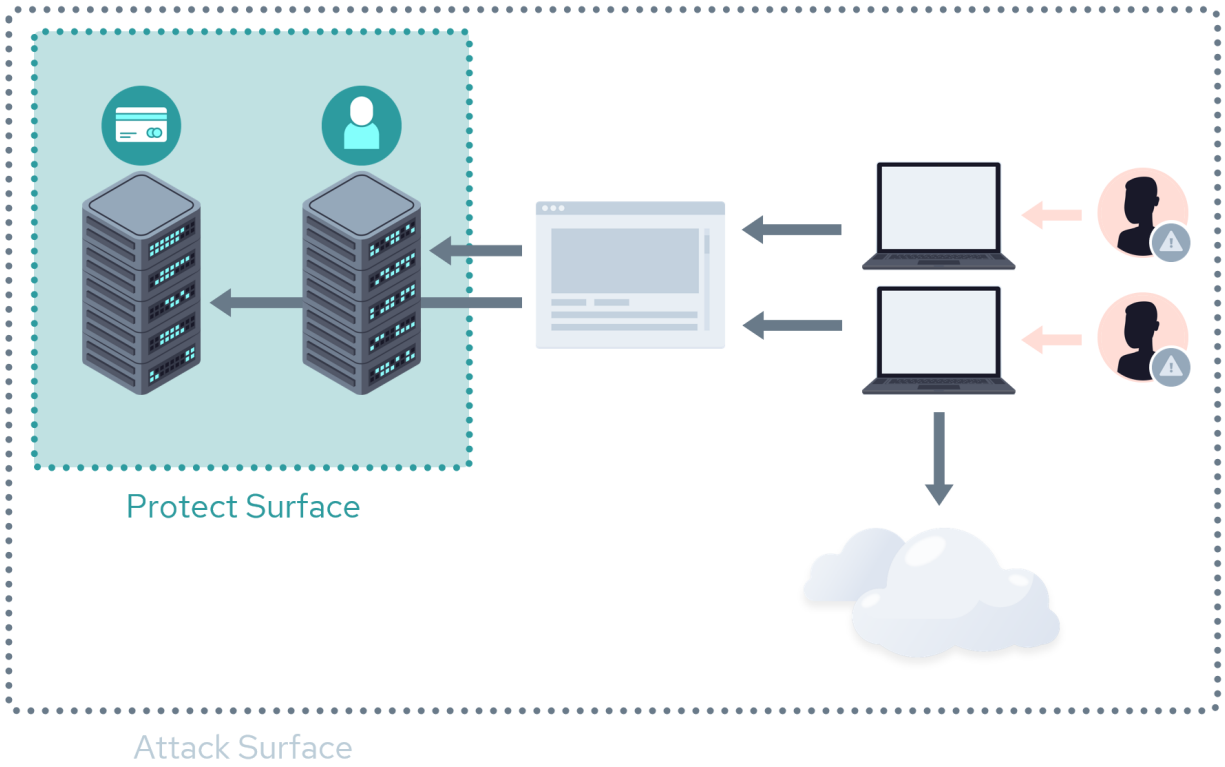


Figure 8: *Laptop and Cloud Services Expand Attack Surface*

However, the protect surface remains the same and is, therefore, more stable and constant than the attack surface. That said, the stability of the protect surface means:

- At the onset, it essentially identifies all the data and assets an organization should protect
- Along with data, it allows the organization to identify the location of the data, the assets that hold the data, and the critical services that the data requires to provide business functions

Once data, the assets, and the critical services are identified, the protect surface allows the organization to move controls closer to the assets at hand and essentially minimize the risk of compromise for critical assets via attack vectors like lateral privilege escalation and visibility to a public network.

In reference to the previous diagram, if a malicious actor successfully compromises the entry points via the application running on a browser, it is only a matter of time before cardholder data or PII is compromised via a vulnerability exploit of the asset. Identifying the assets hosting cardholder data and PII (i.e., the protect surface) enables the organization to move controls like role-based access control (RBAC), system hardening, and secure configurations closer to these assets. For example, the base image of a server build can be hardened before deployment, and the web server hosted on the server asset may be separated from the database host; that is, the database may be moved to another physical server.

Additionally, the organization may create two protect surfaces with micro-segmentation while aligning itself to NSTAC's protect surface definition.

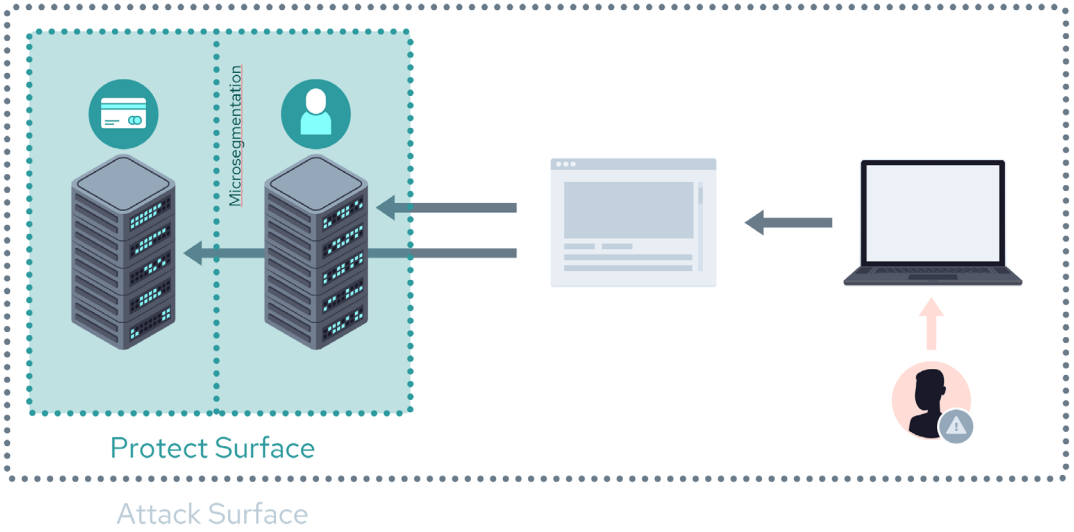


Figure 9: *Two Protect Surfaces Created with Micro-Segmentation*

5.4 Protect & Attack Surface Considerations

While the protect surface provides an inside-out view of the organization, the attack surface provides an outside-in view, or a view from the vantage point of the attacker trying to break in. The protect surface and attack surface complement each other in helping organizations identify what needs protection and how to optimally secure the most critical assets.

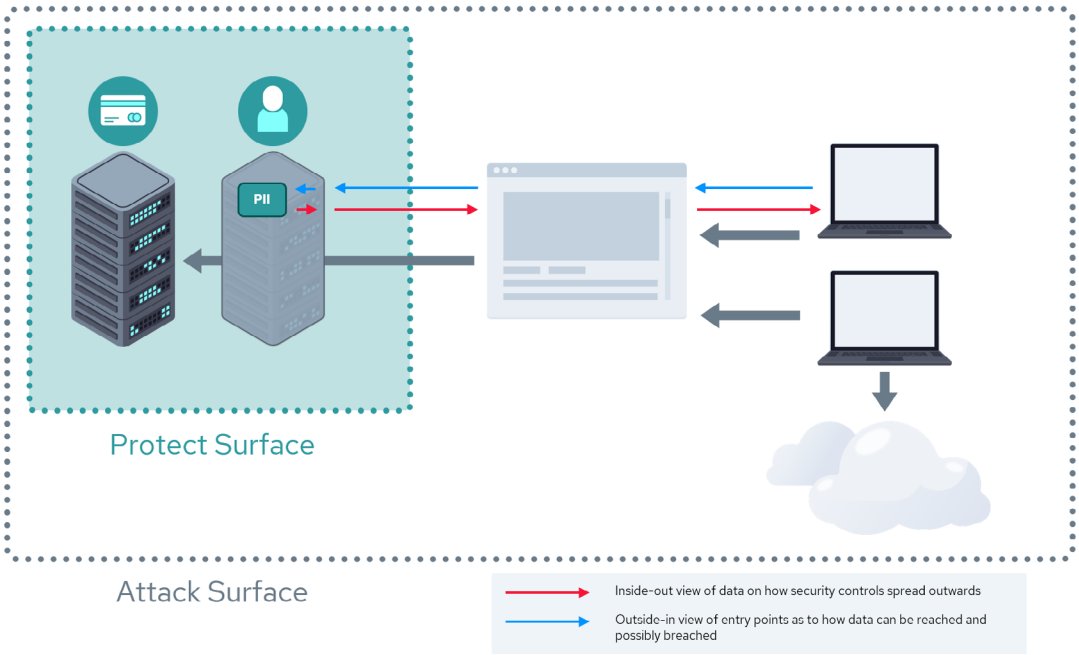


Figure 10: *Different Views of the Organization*

ZT calls for protecting access to any applications or servers shared privately with internal or external users. It is important to assess and plan for each of the assets. Each user should be granted the minimum required access to the assets following the principle of least privilege.

The following are crucial considerations for defining the protect surface:

- Data types involved (e.g., cardholder data, personally identifiable data, health data, trade secrets, by-products of business processes)
- Data and asset classification (e.g., public, internal-only, confidential, and restricted)
- Applications and/or services that handle the identified data
- Critical business functions (e.g., turbine health in a nuclear power plant)

Organizations should design ZTA protect surfaces that incorporate the following:

- The policies defined for ZT access
- The data defining the users (e.g., username, password, identification of the asset held by a user)
- The data defining the assets (e.g., servers, required services, and connections)
- The transport layer
- Business execution algorithms
- The logical and physical relationships between the asset at the core of the protect surface and other business and IT functions

6 Document Transaction Flows

As ZT planners acquire an understanding of the requirements for the system being built and define their protect surfaces, they should also identify what transactions occur with those protect surfaces and how they interrelate. Understanding and tracing of the data flows, application transactions, and business processes allows the planners to understand if the controls in place are sufficient to safeguard the protect surface. In other words, documenting the data and transaction flows ensures that access of entities to the protect surface happens within the defined risk appetite of the organization. Transactions in a system are often derived from the underlying business requirements as part of solution development. The transaction within a system exists because it addresses a business need and is often tied to maintaining business continuity. Business requirements can change over time as will the security considerations.

In the context of ZT, a transaction is any action within a system that needs verification. This could be any component in the architecture from person and non-person entities, an internal or external device to the system, or the process itself that owns the transaction in question. A number of these components are identified in the initial sections of this module.

If you are new to transaction flows and the tools available to create them, one way to look at this task is to envision the data's life cycle. What business transactions occur and what services are invoked to complete a transaction? This could be as simple as the data flow of an online retail purchase, or it could be more in-depth like a customer relationship management (CRM) transaction generating a sales prospect or lead record for a sales organization.

6.1 Example Transaction Flow: eCommerce

The following is a simple illustration of an eCommerce transaction, tracing the steps that occur when a purchase is made from the perspective of the credit card transaction.

In this example the identified protect surface is the payment process components.

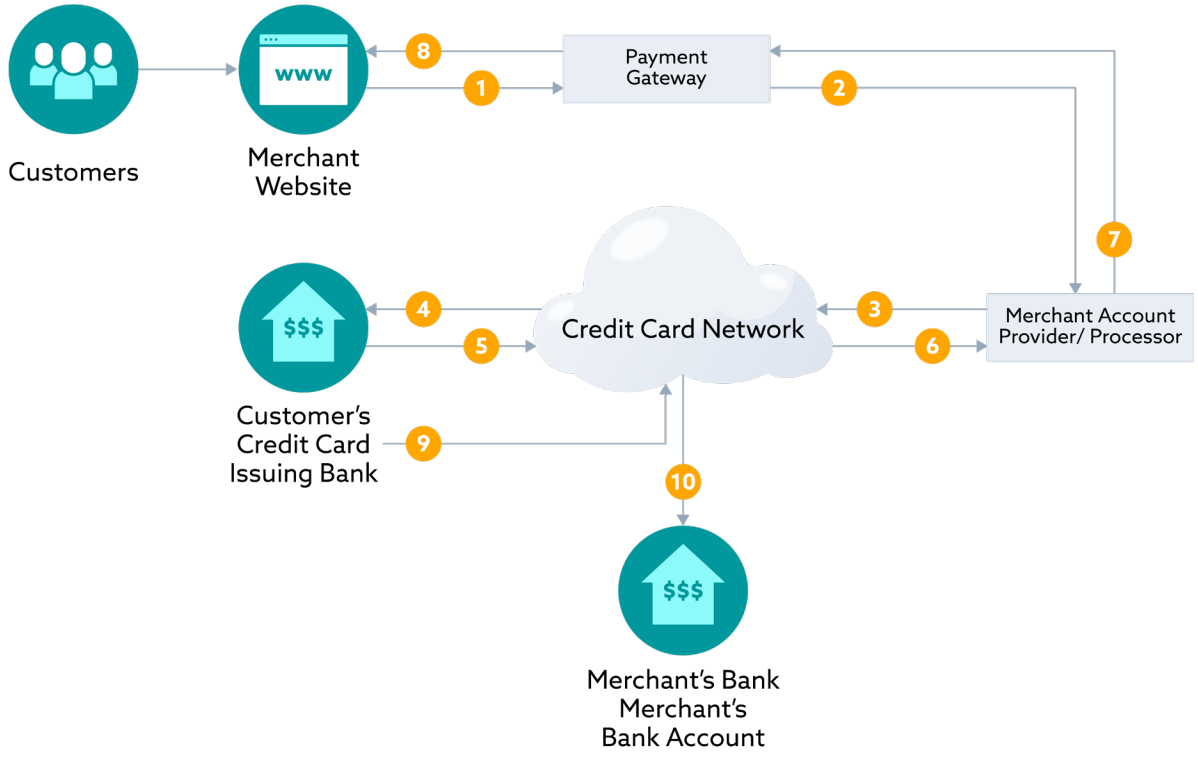


Figure 11: *Example Transaction Flow: eCommerce Payment Process*

1. The merchant's website sends a credit card transaction to the payment gateway via a secure connection.
2. The payment gateway receives the credit card transaction request and submits it using a secure connection to the merchant bank. For accepting payments, merchants need an account with a payment gateway.
3. The merchant bank's processor sends the transaction request to the credit network to process.
4. The credit network forwards the complete transaction to the institution which issued the card.
5. The credit card issuing bank accepts or refuses the transaction based on the card's valid card number, cardholder's name, expiration date, and card verification code and sends back the results to the credit network.
6. The credit network transmits the results to the processor for the merchant's bank account.
7. The merchant's bank processor sends the results back to the payment gateway.
8. The payment gateway saves the transaction results and transmits those results to the merchant's website, which in turn delivers them to the end customer. This is the end of the

approval process.

9. The customer's card issuing institution or bank checks the card number and name, and approves the transaction, sending the funds to the credit network.
10. The credit network sends the funds to the merchant's bank. The merchant's bank deposits the funds into the merchant's bank account and the payment is complete.

Several steps occur in an end-to-end transaction, all in a matter of seconds or milliseconds.

ZT planners should consider approaching the questions of **who, what, where, when, how** and **why** for the steps in this end-to-end transaction, and why each step needs to happen. This will help to ensure that the controls to guard the protect surface fall within the architecture being defined and adhere to organization policies and standards without introducing unmitigated risk. For instance, if the **who** element is a customer, the transaction may be handling PII. The protect surface must therefore be designed to keep the customer's PII data secure during the transaction. This analysis should occur at each step in the transaction process. Refer to the previous section discussing how the protect surface and attack surface are defined. When you measure the risk of transactions, the protect surface will need to be already defined. The attack surface could be modified as new transactions are added or existing ones are modified to keep the architecture in line with business requirements.

An example is provided below for illustration purposes. The transaction focus and protect surface in the last two right-hand columns are where you compare your risk and validate it against the protect surface discussed earlier in this module.

Who	What	Where	When	How	Why	Transaction Focus	Protect Surface Focus
Customer	Person or entity initiating a transaction	Anywhere	24x7 365 days a year	Application Interface	In need of the result the transaction will provide	Disclosing too much unnecessary data	Customer PII including PCI related Information
Merchant/ eCommerce Presence	Selling things	Online and possibly at a physical location	24x7 365 days a year	Through an online commerce portal	To make money	Web front end, customer data, inventory, business records	Customer data in motion and at rest. Entry point device(s), PCI compliance
Bank	Holder of all things monetary	In a giant vault and in the ether	24x7 365 days a year	Online transactions, transfers, in person	Money needs to be deposited or withdrawn	Customer data, PII information, account information, availability, fraud alerts	PII data, customer assets and corporate assets

Credit Card Issuer	Provide credit cards and credit limits to consumers	Everywhere	24x7 365 days a year	Online transactions, physical card, tied to an account	Ease of use, make money through interest	Similar to the banks	PII data, customer assets and/or corporate assets, PCI compliance
Payment Gateway	Provides a connector from the merchant	Between the merchant and credit card transaction process	24x7 365 days a year	API connections from merchant accounts, e.g., PayPal	Provide a mechanism for merchants to send credit card data to a credit card issuer	Inbound connections from ecommerce site, outbound connections to credit card industry	PII in motion and data at rest. PCI compliance

Table 2: *Transaction Flow: eCommerce Payment Process*

Following this eCommerce example, the question to ask is if the protect surface is meeting all requirements to ensure that customer data and financial information is not disclosed. Are you storing any credit card numbers as part of your interaction between the payment gateway and the credit card issuer? Is that transaction of data stored properly with the proper controls around it according to organizational and regulatory requirements (e.g., PCI-DSS)? Putting a matrix together such as this will help validate whether the protect surface is defined, monitored, and enforced properly. More details on data, monitoring, and transactions are discussed next. Additionally, the matrix that follows identifies transactions in your solution which may not be your responsibility, such as the banking institution itself, but that still need to be considered when planning your protect surface.

6.2 Transaction Discovery: Functional Analysis & Tooling

A critical, initial step is determining what data flows are involved. Establishing the proper visibility is crucial for discovering the transactions, their data, and data types (i.e., classification).

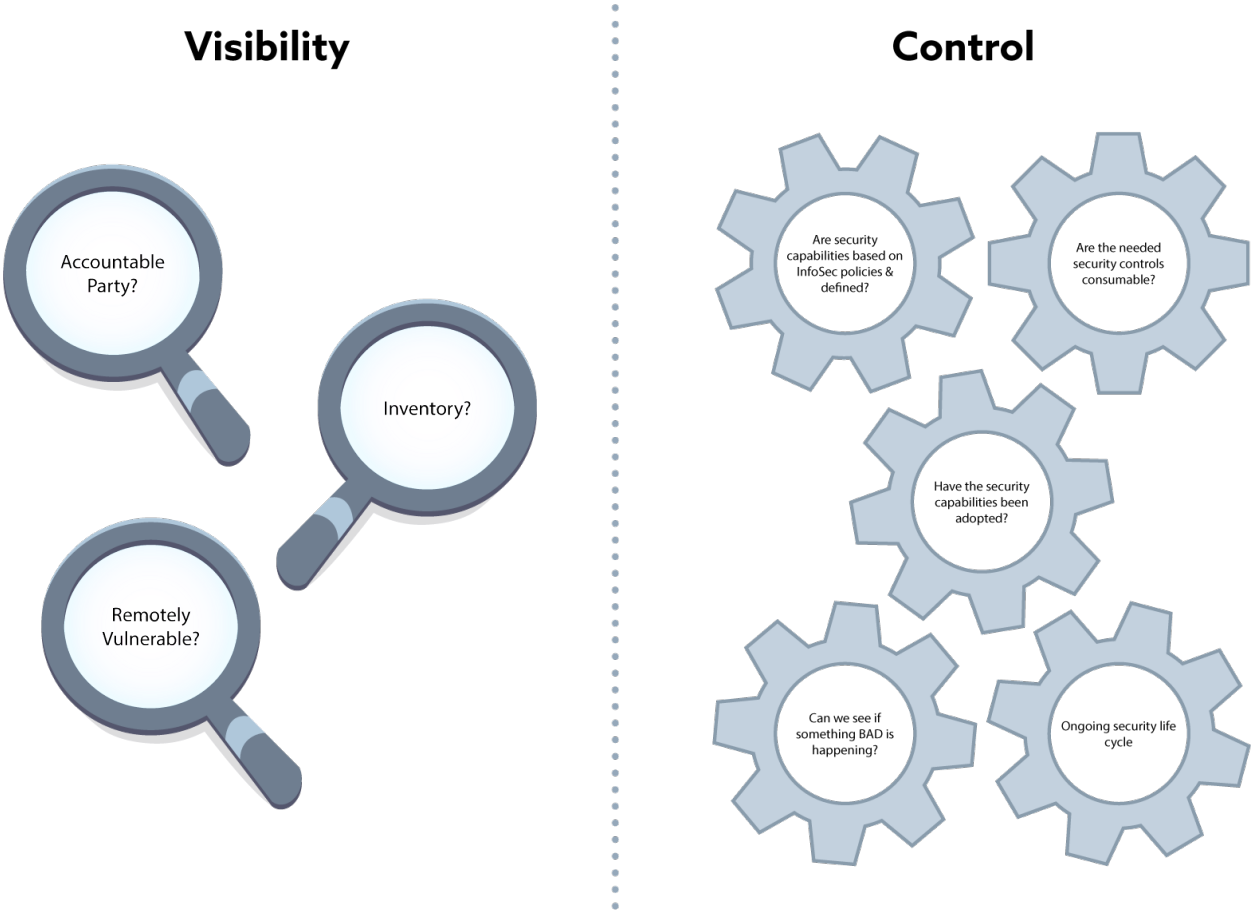


Figure 12: *Transaction Visibility & Control*

Whether the organization is operating in a private, public, or hybrid cloud, the analysis, tooling, and automation to identify what transactions are critical to the business and business processes are fundamentally the same. Below are some key functional areas to consider.

6.2.1 Collecting Data

To define a transaction, begin with an initial understanding of the data—what it is, where it is, and where it goes. Start with existing knowledge about the organization's business process and underlying architecture. It may help to leverage numerous sources such as packet captures, logs, or more sophisticated methods of traffic analysis between the service entities comprising the systems.

6.2.2 Discovery of Known & Unknown Transactions

From the data collection methodologies used, organizations will discover what transactions look like within the service or system being analyzed. They may discover unknown transactions they were previously unaware of. For instance, in the eCommerce example, there may be a transaction between the payment gateway and the eCommerce application to determine what tax to add to the full value of the transaction. Perhaps a tax rate changed based on doing business in a new market, or a shipping method changed or was added due to a new delivery requirement. While not specific to the method, this represents an ancillary process uncovered in the data collection phase.

6.2.2.1 Transaction Inventory

In completely new deployments, transaction inventories will be defined and developed during the architecture and design phase; nonetheless, organizations should create an inventory as part of the planning exercises. If organizations are pursuing an already existing deployment to migrate to a ZTA, they should collect and inventory the known transactions to maintain, highlight the ones that will change or become deprecated, and create entries for new transactions expected to be part of the solution as it is being developed.

6.2.2.2 Transaction Records

Transaction records are the historical paper trail of the behavior of transactions and the types of data involved. Recall that at every transaction layer, ZT controls will continuously analyze the processes and compare them to existing policy enforcement rules. Keeping and analyzing transaction records are part of the planning process and remnant samples of monitoring and analytics.

6.2.3 Monitoring & Analytics

Once you know what transactions are in your system, you will want to monitor them to collect statistics and profile the behavior. Again, as shown in the eCommerce example, monitoring the payment gateway and collecting analytics from dollar values being processed through your system will start to build a view of the behavior and trends of the process you ultimately want to protect. Monitoring and analytics are covered more in-depth later in this module. However, it is important to understand that ZT prescribes a continuous assessment of all transactions.

6.2.4 Identifying Anomalies & Edge Cases

During transaction discovery, ZT planners should identify what behaves unexpectedly or abnormally from the baseline. Edge cases may be greater in number than expected. Does this change the protect surface area as a result? Are new transactions required to mitigate any risks? Do any of them change?

For instance, in the eCommerce example, imagine the entire platform was designed at the baseline to allow for free shipping within the United States. The shipping provider charges the merchant for this behind the scenes. Suddenly, due to a marketing campaign in Hawaii and Alaska, there is great demand for overnight shipping to both locations which can't be met by the current provider integrated into your shipping workflow. Even though it is a small corner case sales and sales transactions, you need to bring in an additional third-party shipping partner to meet the shipping SLA. In doing so, do you need to supply them with any additional customer data? Do you have to set up a separate shipping workflow to integrate with them? Can they maintain the same level of privacy? Do you have to do any additional monitoring to ensure this new shipping partner doesn't create any new risk to you or the customer?

While this seems like a simple edge case on the surface, the seemingly benign act of offering a specific delivery mechanism to a specific subset of customers requires going through this new transaction again and checking all the boxes to ensure that your protect surface doesn't require additional controls.

7 Define Policies for Zero Trust

In a ZT approach, the visibility of and access to resources by any user or device is regulated and controlled by policies. Policy planning should be carried out with utmost care, with a granular understanding of who should get access to what resource, which actions are allowed, under which conditions, and for how long or at what time of day.

The policies need to be planned prior to implementation as doing so can help the stringent control for each of the user or user groups. The controls and policies allowed for each asset need to be documented for a better organized implementation and maintenance of the policies. Each of the newly introduced changes needs to be tracked down with a separate tracker that is peer-reviewed.

The user and the access are trusted after the authorization at the zero-trust gateway as the policies enable the authorization of the access.

Figure 13: *PDP/PEP & Zone Interactions*[22]

ZT systems enforce validation of the user and the device before permitting any access, hence the ZT policies allow organizations to plan and create access policies based on user or device attributes and contextual risks. By leveraging aspects such as directory group membership, IAM-assigned attributes and roles, location, and device posture, organizations can define and control access to cloud or data center resources in a way that is meaningful to business, security, and compliance teams.

7.1 The Policy

ISO 9001 defines policies as documents that include information about a set of standards. Within organizations, policies might appear in a hierarchical structure, and each policy defines the set of rules used to govern a different area of the business. To avoid any confusion, we want to clarify that in this section we are referring to rule-based security[23] policies that set the rules which control the access and the entitlements to a planned set of IT assets. The ZT policies are a set of rules which control the access and entitlements to a planned set of IT assets. The PDP will have two components/functions, an engine to maintain the set of rules, and another component to administer the rules at the user interface. The ZT policy can be applied to a combination of applications, application groups, user, or user groups based on the implementation planning. The policies can also vary the access, entitlements, and enforcements based on the dynamic contextual risks. The session management procedures take actions that enable the PDP to continually evaluate the session once it

[22] Figure adapted from: NIST. (2020). Zero Trust Architecture (SP 800-207).
[23] ISO. (1989). Rule-based security policy. In ISO 7498-2:1989(en)

has been established. The session will be revoked once the technical conditions are not met. The policy may define different levels of access and entitlements based on the attributes of the identity. For example, users coming from different locations or times of day may end up getting different levels of access.

ZT implementations typically use device, network, environment, user, and entity behavior analytics (UEBA), and IAM attributes for users (e.g., directory group memberships, directory attributes, roles) as elements for policies. For example, a policy may state that all users in the directory group **HR**, using an endpoint device with a level of security hygiene **medium**, may access server HR-Portal on port 443 from a specific country from 10:00 AM PST to 8:00 PM PST and perform actions A, B, and C, under standard **low** risk conditions. This example illustrates how a system can add value to, and extend the power of, an existing IAM deployment.

7.2 The Policy Workflow

ZTA policy planning should have the gateway enforcing access policies on a per-user/group basis, achieving the principle of least privilege by denying access by default. Additionally, the PEP/gateways should be situated at the entry point of each private cloud network controlling all inbound traffic based on the policies defined at the policy engine.

The planning phase should involve the planning for the required policies. The policy administrator (i.e., the previously mentioned logical component of PDP) allows for the planning and definition of policies at the policy engine. The PEP allows the ZTA to apply the policies based on which access can be managed for various assets such as, for instance, web application or secure shell (SSH) access.

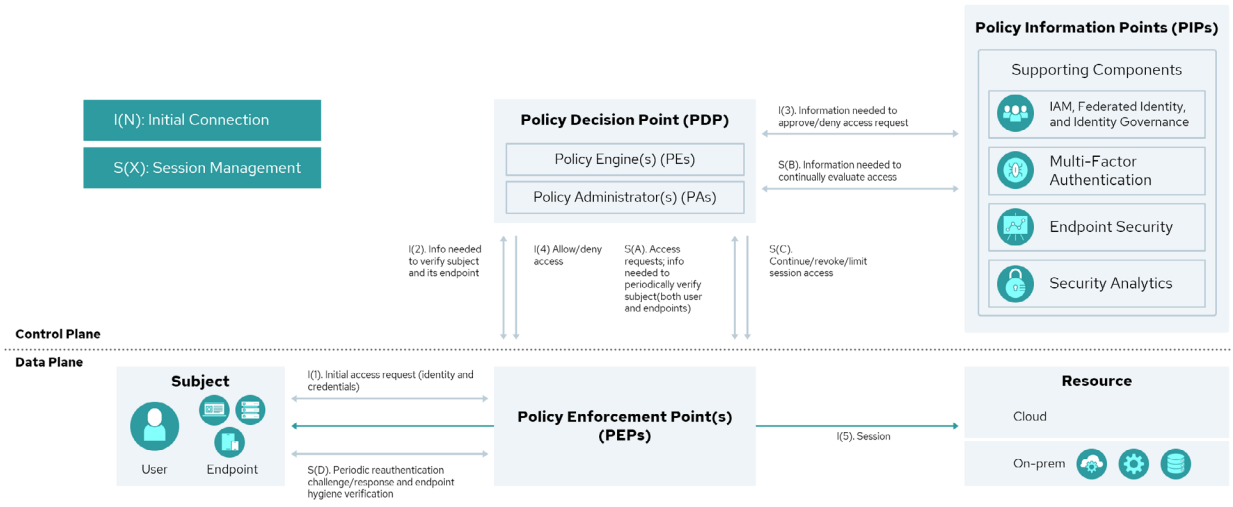


Figure 14: *Zero Trust Entities & Policy Workflow*[24]

The policies kept at the policy engine have to be planned with utmost care and acceptance from the respective decision makers. The policy administrator helps to add/modify and maintain the policies in a continual improvement model.

[24] Figure adapted from: NIST. (2022). Implementing a Zero Trust Architecture (SP 1800-35B). Second preliminary draft.

The PEP will authorize access based on the policies defined at the policy engine. The policy data will be percolated down to the PEP by the PDP. In turn, the subject is either granted or denied access to the protected resource in question.

Throughout the lifetime of the session, the PEP may periodically challenge the subject to re-authenticate itself, depending on the level of risk associated with the transaction. After doing so, the PEP will provide the PDP with the identity and credentials that the subject provided. Similarly, throughout the lifetime of the session, the PEP will request hygiene information from the subject's endpoint. After obtaining this hygiene information, the PEP will provide it to the PDP. The frequency with which the subject should be issued authentication challenges is determined by enterprise policy, as is the frequency with which the hygiene of its endpoint should be validated. In both cases the policy is a reflection of the risk-based decision.

The connection between the PEP and the subject may be terminated or reconfigured based on changes to the endpoint, resource, or operating environment that indicate the subject no longer conforms to enterprise policy. The policy may enforce various levels of access and entitlements based on the risk and current context of the user, device, endpoint, network.

It is important to highlight that the policy workflow is a logical representation that doesn't correspond to the actual physical architecture. The components defined are meant to represent logical functions, not physical devices.

7.3 Policy Considerations & Planning

As mentioned in the previous sections, ZT policies are the logic behind enforcement of ZT principles like least privilege and need to know. These policies are the gatekeepers that greet each incoming access request with a set of predefined questions such as:

- The identity of the requesting identity
- The target assets a particular entity/user is entitled to access
- The timeframe the entity/users is allowed access to the asset
- The geographical or logical origin of the request and requestor
- The device attributes
- The entitlements of the entity/user

Access should be planned at the group level as much as possible versus at the individual user and application level. The plan should define which group of users can access which group of applications (e.g., HR for HR Applications) while policies are planned for various levels of access and entitlements.

The following is a list of factors to consider in the definition and eventual enforcement of the ZT policies:

- **User attributes:** Is the requesting user authenticated? What is the level of certainty about the requesting user's assertion of identity? Is the user internal or external? Where, geographically, is the user located? Where do we expect the user to be located at the time of the request? How often should the user re-authenticate to the asset, and under which

circumstances? Which network/IP is consumed by the user?

- **Target resource attributes:** What is the classification of the data? Which services, devices, data, and other resources are available and allowed for a particular request? What level of access should a particular requestor be granted at the time of the request?
- **Time:** During which time windows do we expect requests from a particular requestor to a particular resource? When should access end?
- **Device attributes:** Which devices are registered with the enterprise, and do we allow unregistered devices to originate requests for access? Which attributes do we expect from an originating device (e.g., MAC address, profiles created by an agent)? Do we require that the originating device be registered and/or authenticated with the enterprise? Which patch levels and/or software suites do we require from the originating device? What is the overall security hygiene level expected from the device?
- **Entitlements:** What level of access should be granted to a particular user over a particular resource? How do the entitlements change based on a person's/user's attributes? How entitlements change depending on the level of hygiene of the device? How do entitlements change depending on risk factors?

ZT is ultimately about dynamic risk management, so the policies need to reflect the changes in the risk levels and allow access and actions depending on that risk context. The risk in a specific context is a reflection of all the previously mentioned variables (e.g., type of user/entity, environmental conditions, device posture, user behavior, etc.) Therefore, the possibility to access resources and perform certain actions changes depending on risk levels.

The policy should take into consideration the behavior of the user/entities. In case of anomalies, for instance, there should be a rule to ensure that the access is revoked or the entitlements are limited based on the level of deviation from the baseline measured behavior of the user/entity dynamically during the session.

It is important to note that establishing and enforcing policies based on behavior and risk assumes that the organization can collect and analyze telemetry data from the PIP.

In the planning phase, it is important to realistically define what, within the context of the protect and attack surface, can be subject to continuous monitoring and what data source the organization is able to collect and analyze.

7.4 Continual Improvement

The process of continual improvement should be well planned and documented with a track of reviews and approvals. The access rules and entitlements should always be subjected to recurring reviews and improvements based on need, risk, and context. The reviews can be planned to repeat after a fixed duration and scheduled accordingly. The changes introduced over the collection of attributes such as user, endpoints, applications, or the attack surface of the organization can often change the access and entitlements needs. Hence the planning should involve the possible processes and definition of controls in place.

New access points may be created according to needs that may arise over time; these changes should be reviewed and planned for ahead of time during the course of ongoing discussions

7.5 Automation & Orchestration

The effective implementation of ZT needs to ensure that the automation and orchestration are planned at each stage of the operation such as authentication, MFA, access provisioning, policy enforcement, and dynamic evaluation of the posture.

The strategy for automation and orchestration should be planned well in advance. The automated enforcement of access policies reduces the need to manually update and test firewall rules in response to user or server changes. In larger organizations, this is typically part of the daily workload for IT, and therefore presents an opportunity to reduce both workload and labor costs (especially in an outsourced model). It also accelerates business and technical user productivity, which while worthwhile in its own right, can also reduce hard costs (particularly for hourly or outsourced workers). The need for automation and orchestration varies for each organization and kind of access requirements. Hence, we need planning and preparation for the same.

8 Developing a Target Architecture

Identifying and developing a target architecture is the last step of the ZT planning phase. This step is about defining how your service and network architecture looks. The ZT target architecture will likely be an evolution of the existing architecture; alternatively, the change could be revolutionary.

The target architecture will be business-driven, as expected, by the nature of the business, technological environment, and consequently, the challenges that the organization is trying to address through the ZT approach. Is the organization addressing the challenges of multi or hybrid cloud? Is it about the security of a highly distributed supply chain and/or production chain? Is it about solving the issue of a highly distributed workforce with a need for remote access? Is the challenge related to securing an ICS, or more general operational technologies?

The definition of the target architecture should consider a number of technical variables which are described in the ZT five pillars (i.e., Identity, Devices, Networks, Applications and Workloads, and Data) and the cross-cutting functions (i.e., Visibility and Analytics, Automation and Orchestration, and Governance). This conceptual framework might be used for determining their **current** state, a desired **future** state, and finally a path towards developing a target architecture (**roadmap**). This topic of assessing the current state and determining the roadmap for the target architecture has been analyzed in an earlier unit.

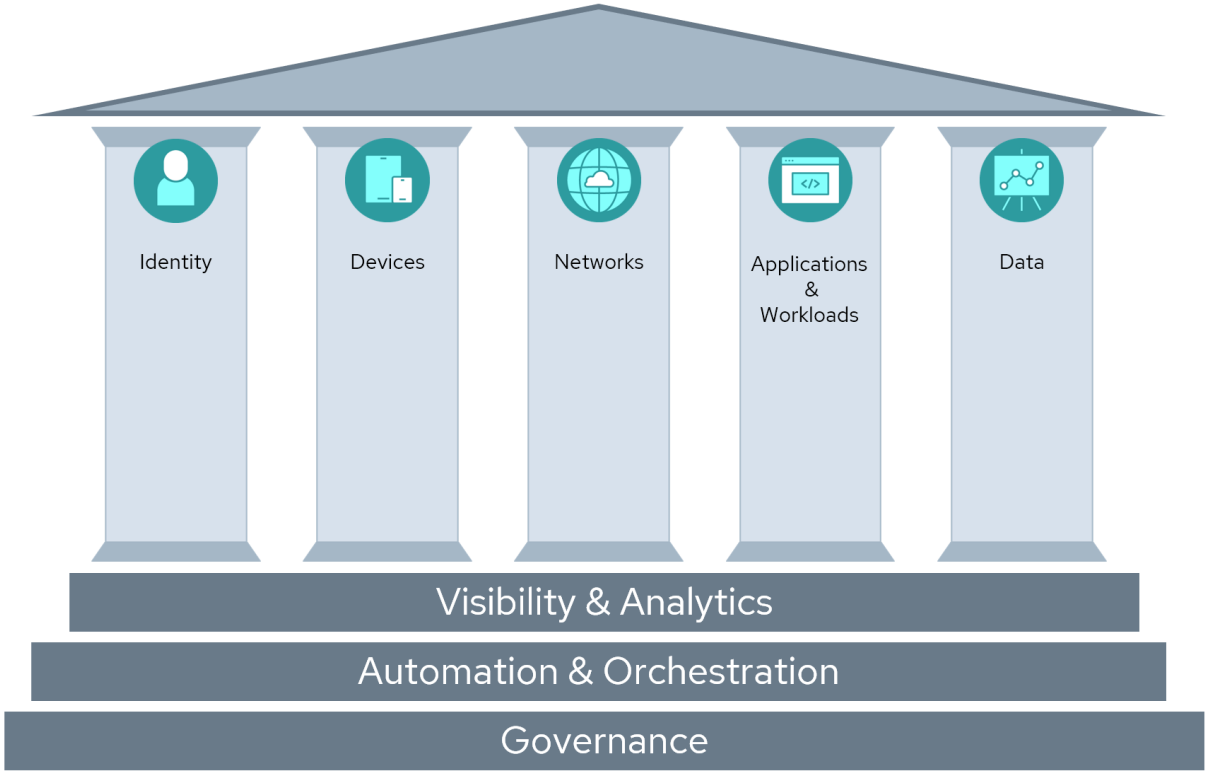


Figure 15: *Zero Trust Pillars & Foundations*[25]

In this unit, we discuss important considerations when planning for a target architecture in the context of these main pillars and functions.

8.1 Identity Considerations

Proper identity validation of the entity requesting access to a resource is paramount in a ZT environment. The principle of least privilege, which is a key component of ZT, can only be assured with validated entities (e.g., human user, computer services, etc.). This validation should be performed when the entity is requesting access to a resource and also periodically throughout this access, with the frequency and technology determined by the sensitivity of the information being accessed (data classification). As a rule, MFA should be leveraged to validate the identity of the entity. In some cases, step-up/adaptive/conditional authentication (additional rigorous authentication steps) should be required for access to more sensitive information. As ZT maturity improves, real-time machine learning analysis to highlight any user or device behavior that is unusual should be performed for further analysis and follow-up to ensure the security of the information protected by ZT.

Identity stores contain the entities and associated information. These stores are queried during the authentication process by the ZT process. Mature ZT implementations include a global identity store that can be leveraged for both on-premise and across cloud environments.

[25] Figure adapted from: CISA (2023). Zero Trust Maturity Model (Version 2.0).

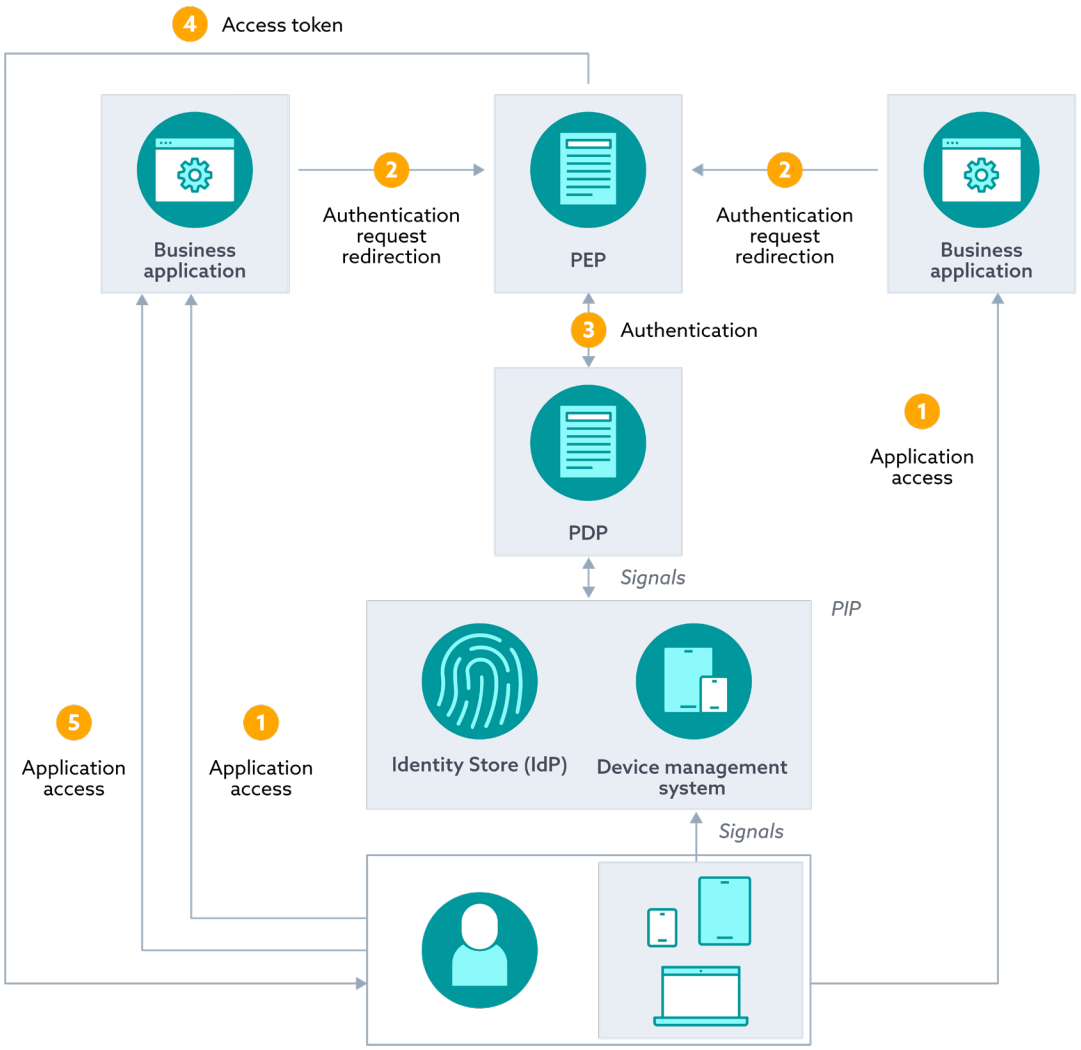


Figure 16: *Validating SaaS Application Access*[26]

A process should also exist in order to ensure that the user identity is mapped to a real user when possible (identity proofing). The accuracy of claims should also be controlled during the lifecycle of the user. When possible, an integration with a public key infrastructure (PKI) should be considered because it can be integrated with the identity system.

8.2 Device & Endpoint Considerations

In a ZT environment, devices and endpoints (e.g., laptops, mobile phones, IoT devices, servers, etc.) also require authentication validation before they can access the resources protected by the ZTA. In addition, the device's security posture (e.g., device hardware and software patch level, the status of installed security software, etc.) should also be validated against an organization's security policies before the device is allowed access to the resources it requests. In more mature implementations, these validation steps are performed continuously, and the device's behavior is also analyzed to identify any unusual activity.

[26] Figure inspired by: NIST. (2020). Zero Trust Architecture (SP 800-207).

Performing a complete and accurate inventory of all devices and endpoints is a highly sought-after goal for most organizations. While some achieve this goal, many large organizations fail primarily due to the vast number and relatively short life cycle of devices deployed in their environments. A mature ZTA deployment will ensure that only properly registered and secured devices can access the organization's resources. This ZTA requirement will help even the largest organizations achieve their device and endpoint inventory data quality goals. Unmanaged devices (e.g., contractor devices, etc.) should also be incorporated in the ZTA design, and solutions leveraging a gateway or VDI should be explored.

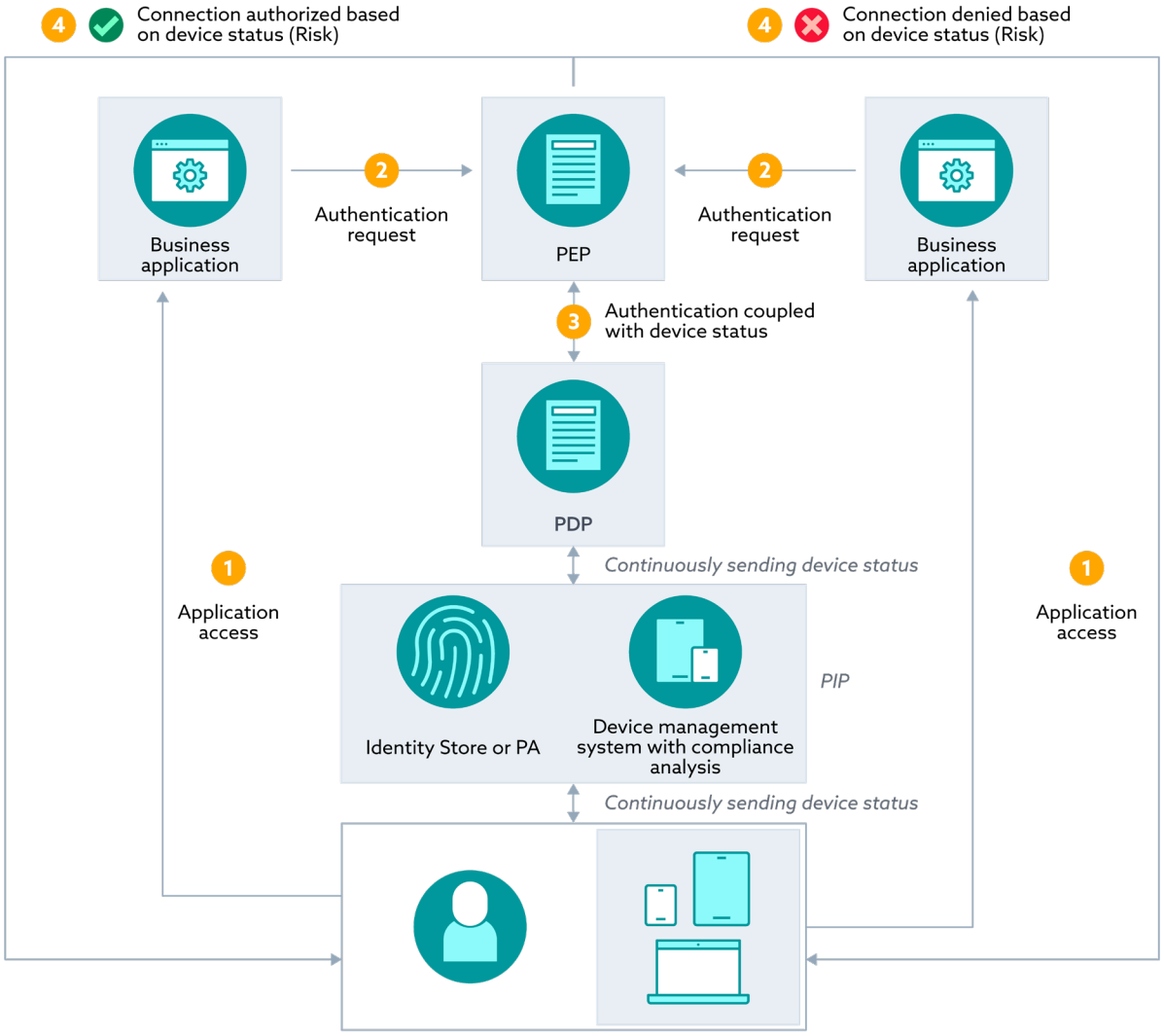


Figure 17: *Access Decisions with Endpoint Risk Analysis*[27]

8.3 Network & Environment Considerations

A typical ZTA implementation will use micro-segmentation coupled with encryption in order to improve the security posture of the network. This means that the data plane used for application/service communication and the control plane used for network communication control should be separated.

[27] Figure inspired by: NIST. (2020). Zero Trust Architecture (SP 800-207).

The decision to allow access to the application is made over the control plane, and the actual application interaction and data exchange with the requesting device occurs via the data plane. To achieve micro-segmentation technology like host-based firewalls, software agents, intelligent routers, or next-generation firewalls can be used. Segmenting traffic based on the data flow (internal data flows vs. external data flows) should also be considered.

8.4 Workload & Application Considerations

All applications should use a centralized authentication, authorization, monitoring, and attributes system. This configuration provides better visibility and enables real-time risk analytics in ZTA. Access authorization should be continuously evaluated and consider real-time risk analytics, which means that the application should adapt to environmental changes. For internally developed applications, security testing should be implemented in all stages of the CI/CD cycle.

Whenever possible, applications should also be integrated into the monitoring system in order to send internal insights (e.g., from which country a user is accessing the application, the type of device or browser if possible) and be accessible without VPN.

8.5 Data Considerations

An organization's data classification policy should codify the required data security controls and processes used for each of the data classes defined by the organization. These security controls can include secure encryption (at rest and in transit) and network segmentation for highly sensitive data to simply read-only access for publicly available data. The processes outlined by the policy should include how entities gain access to the data (leveraging least privilege principles) and what steps are required to properly dispose of the data when its end-of-life has been reached.

8.6 Visibility & Analytics Capability Considerations

As mentioned in the introduction of this unit, visibility and analytics is one of the important functions needed in a ZT architecture and helps support the pillars noted above. When optimally deployed, this function increases security by the following three means:

- Leveraging UEBA to continually evaluate the user's behavior against a baseline of previous activity to identify any unusual action
- Running regular device posture assessments to ensure the device being used to access the application or data is properly configured and secured
- Monitoring application health and security by leveraging systems and sensors external to the application

The feedback from these visibility and analytics capabilities should be directed towards the PEP so it can make real-time decisions about granting and revoking access to the requested application and data.

8.7 Automation & Orchestration Capability Considerations

Automation and orchestration should be used to support every pillar. When optimally deployed, this function increases security by:

- Taking advantage of automation by using infrastructure-as-code to deploy network and environment configurations and consolidating with the CI/CD pipeline
- Orchestrating and automating the identity lifecycle, including dynamic user identity and group membership, JIT access to applications and data, and revoking access when required

8.8 Governance Capability Considerations

In a ZTA deployment, governance is the most important function because it ensures that business, risk, and IT perspectives are aligned. Governance helps to define ZTA policies; for example, to access and process data, a device must be encrypted. From a non-technical perspective, governance should also manage and reduce complexity. In order to do that, the focus should be on the protect surface, with governance policies enforced by the PEP.

8.9 Examples of Zero Trust Architecture

From a practical perspective, several reliable sources can serve as a model for defining the target architecture to meet the organization's specific business objectives. A good reference is the NIST SP 800-207.

There are two important items to note. First, every situation is unique with its own business objectives and constraints. These sources are templates designed to inspire and guide you in developing an architecture that fits your needs and meets your objectives. Second, more sophisticated enterprises may have a suite of target architectures. For example, in the case of equipment and devices on a manufacturing floor, an organization may design a different architecture for them than the one it uses for core IT functions.

In essence, ZTA approach variations typically fall into one of the following three categories:

1. **ZTA using enhanced identity governance:** As the name implies, at its core, ZTA is driven by identity and rooted in the proper governance of the access privileges and entitlements for specific assets.
2. **ZTA using micro-segmentation:** This approach is based on the logical segmentation of the network. The organization uses devices such as next generation firewalls (NGFWs) or gateways to act as a PEP and enforce the logical boundaries of the protect surface. This approach assumes that a fully functioning enhanced identity governance program is enforced. It also assumes the organization updates access rules to accommodate changes in business objectives, threats, context, user behavior, and other factors. Micro-segmentation has the added advantage of limiting the impact radius in the event of an incident.
3. **ZTA using network infrastructure and software-defined perimeters (SDPs):** This approach focuses on the network architecture to achieve ZTA. The SDP approach uses the PDP as a

network controller (SDP controller) to restrict visibility of the asset and which entities can interact with the resources part of the protect surface. This approach was developed by CSA. The specific details can be found in the training modules entitled *Introduction to SDP*[28], *Key Features & Technologies of SDP*[29], and *Architectures & Components of SDP*[30].

As NIST highlights, each approach varies based on the specific situation. NIST specifically addresses components and the source of truth for the organization's policy rules.

A full ZT solution will eventually include elements of all three approaches. The selection of the starting point depends on the considerations mentioned previously, including the maturity of the organization, business objectives, technology strategy, use cases, etc. We described/discussed/outlined how these conceptual architectures are geared/crafted to inspire the design of the architecture for your specific situation. NIST lists the following variations:

- Device agent / gateway-based deployment
- Enclave-based deployment
- Resource portal-based deployment
- Device application sandboxing.

Additional discussion can be found in *Architectures and Components of SDP*. These variations are described in more detail and are also compared with the SDP architecture deployment variations in that module.

Conclusion

This module covered the planning activities and considerations for an organization moving to ZT. Learners were instructed on how to identify ZT stakeholders, prioritize and scope a ZT implementation, carry out a gap analysis, map out the protect and attack surfaces, and define the ZT technology policies. Lastly, critical considerations for planning for a target architecture were described in the context of ZT's main pillars and functions.

Glossary

Please refer to our Cloud Security Glossary, a comprehensive glossary that combines all the glossaries created by CSA Working Groups and research contributors into one place.

[28] Cloud Security Alliance. (2022.) Introduction to Software-Defined Perimeter.
[29] Cloud Security Alliance. (2022.) Key Features & Technologies of Software-Defined Perimeter.
[30] Cloud Security Alliance. (2022.) Architectures & Components of Software-Defined Perimeter.

Zero Trust Implementation

CCZT Study Guide

cloud security alliance®

CCZT™
Certificate of
Competence in Zero Trust

The official location for SDP and Zero Trust Working Group is
https://cloudsecurityalliance.org/research/working-groups/zero-trust/

Disclaimer

Cloud Security Alliance designed and created this Zero Trust Training course study guide (the "Work") primarily as an educational resource for security and governance professionals. Cloud Security Alliance makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

Version Number: 20240820

About Cloud Security Alliance

The Cloud Security Alliance℠ (CSA) (www.cloudsecurityalliance.org) is the world's leading organization dedicated to defining and raising awareness of best practices to help ensure a secure cloud computing environment. Cloud Security Alliance harnesses the subject matter expertise of industry practitioners, associations, governments, and its corporate and individual members to offer cloud security-specific research, education, certification, events and products. Cloud Security Alliance activities, knowledge and extensive network benefit the entire community impacted by cloud—from providers and customers, to governments, entrepreneurs and the assurance industry—and provide a forum through which diverse parties can work together to create and maintain a trusted cloud ecosystem.

CSA Address

709 Dupont St.
Bellingham, WA 98225, USA
Phone: +1.360.746.2689
Fax: +1.206.832.3513

Contact us: support@cloudsecurityalliance.org
Website: https://cloudsecurityalliance.org/
Zero Trust Training Page: https://knowledge.cloudsecurityalliance.org/page/zero-trust-training
Zero Trust Advancement Center: https://cloudsecurityalliance.org/zt/
Provide Feedback: support@cloudsecurityalliance.org
CSA Circle Online Community: https://circle.cloudsecurityalliance.org/
Twitter: https://twitter.com/cloudsa
LinkedIn: www.linkedin.com/company/cloud/security/alliance
Facebook: www.facebook.com/csacloudfiles
CSA CloudBytes Channel: http://www.csacloudbytes.com/
CSA Research Channel: https://www.brighttalk.com/channel/16947/
CSA Youtube Channel: https://csaurl.org/youtube
CSA Blog: https://cloudsecurityalliance.org/blog/

Acknowledgments

Dedicated to Juanita Koilpillai, a pioneer in software-defined perimeters whose contributions to the Certificate of Competence in Zero Trust (CCZT), Zero Trust training, and CSA are immeasurable.

The CCZT and Zero Trust training was developed with the support of the Cloud Security Alliance Zero Trust Expert Group, whose members include volunteers from a wide variety of industries across the globe. Made up of subject matter experts with hands-on experience planning and implementing Zero Trust, both as cloud service consumers and providers, the Zero Trust Expert Group includes board members, the technical C-suite, as well as privacy, legal, internal audit, procurement, IT, security and development teams. From cumulative stakeholder input, the Zero Trust Expert Group established the value proposition, scope, learning objectives, and curriculum of the CCZT and Zero Trust training.

To learn more about the CCZT and Zero Trust training and ways to get involved please visit: https://cloudsecurityalliance.org/education/cczt

We would also like to thank our beta testers, who provided valuable feedback on the CCZT and Zero Trust training: https://cloudsecurityalliance.org/contributors/cczt-contributors

Lead Developers:

Clement Betacorne
Heinrich Smit
Mark Schlicting
Michael Roza
Prasad T.
Shruti Kulkarni

Contributing Editors:

AJ (Alexander) Stein
Ledy Eng
Richard Lee
Robert Morris
Roland Kissoon
Ross Kovelman

Expert Reviewer:

Abbas Kudrati
Adish Jain
Amit Butail
Andy Radle
Ashwini Siddhi
Aunudrei Oliver
Charlie Soto
Chris Willman
David Skrdla
Donald ByersA
Gustavo Vallejo
Jaye Tillson
John Kindervag
Karthik Ramamurthy
Madhav Chablani
Matt Lee
Matt Meersman (Dr.), PhD
Michael Herndon
Naresh Kurada
Philip TM Pearson
Reto Kaeserm
Roeland van Zeijst
Ron Kearns
Ron Martin (Dr.), PhD
Sakthiswaran Rangaraju
Shain Singh
Shinesa Cambric
Sky Hackett
Vani Murthy

CSA Global Staff:

Adriano Sverko
Anna Campbell Schorr
Chandler Curran
Daniele Catteddu
Hannah Rock
Noelle Sheck
Stephen Smith

Table of Contents

List of Figures

Course Introduction

Welcome to *Zero Trust Implementation* by Cloud Security Alliance (CSA). This training module is part of a larger series titled Zero Trust Training (ZTT). It builds upon and extends the concepts discussed in the CSA *Zero Trust Planning*[1] and *Introduction to Zero Trust Architecture*[2] courses. In this course, learners get an in-depth look at the crucial facets of Zero Trust (ZT) implementation, from creating project kick-off documents and disaster planning, to setting up the network environment, deploying agents to devices, and adding automation.

Course Structure

This course consists of four units, each geared towards gaining increased competency in the following topics:

1. Continuing the ZT journey
2. ZT project implementation considerations
3. Implementation preparation activities
4. ZT target architecture implementation

Course Learning Objectives

After completing this course, learners will be able to:

- Identify the assumptions and considerations for continuing the ZT journey
- Explain the main ZT project implementation preparatory activities
- Outline Zero Trust Architecture (ZTA) implementation steps
- Leverage ZT pillars and cross-cutting capabilities to define and prioritize implementation tasks
- Visualize and document security workflow architecture using transaction flow diagrams and tables
- Design testing procedures that can be repeated and generate audit trails
- Define success criteria and review the success of ZT implementation

1 Continuing the ZT Journey

Before we jump into the content of the *Zero Trust Implementation* course, let's recap a few key points that we already addressed in the *Zero Trust Planning*[3] module. The ZT project plan is the roadmap that serves as the team's checklist to implement this plan.

[1] https://knowledge.cloudsecurityalliance.org/zero-trust-planning
[2] https://knowledge.cloudsecurityalliance.org/introduction-to-zero-trust-architecture
[3] https://knowledge.cloudsecurityalliance.org/zero-trust-planning

The *Zero Trust Planning* module covered:

- Starting the ZT journey
- Planning considerations
- Scope, priority, and business case
- Gap analysis
- Defining the protect and attack surfaces
- Documenting transaction flows
- Defining ZT policies
- Developing a target architecture

Before diving into implementation of ZT, it's important to set the stage and make some general assumptions, which we cover in the next section. This is due to the varying types of industries and companies that exist; there isn't enough room in this course to delve into all the specifics for each industry, such as healthcare, finance, or energy.

1.1 Training Assumptions

This training assumes that the learner reviewed and understood the preceding Zero Trust Trainings (ZTT): *Introduction to Zero Trust Architecture* and *Zero Trust Planning*. Together, they form a body of work, including defining ZTA, ZT pillars (Identity, Devices, Networks, Applications and Workloads, and Data), and ZT cross-cutting capabilities (Visibility and Analytics, Automation and Orchestration, and Governance). For this training, our ZT implementation is designed around the ZT pillars and ZT cross-cutting capabilities, defined in the Cybersecurity and Infrastructure Security Agency (CISA) *Zero Trust Maturity Model*[4].

Additionally, this training assumes the student knows and understands basic software project management and can create a project plan for implementation. Let's quickly review the main ZT project management implementation steps we have already covered:

1. Project organization (covered in *Zero Trust Planning*): A company defines the protect surfaces and priorities, determines what objectives must be met and by whom, and identifies the steering committee.
2. Project design (covered in *Zero Trust Planning*): The project team maps the transaction flow, defines the ZT policies, and designs the ZT environment.
3. Implementation (covered in *Zero Trust Implementation*): During ZT implementation, the solution is set up and documented. Frequent status updates are needed for the project manager and the steering committee to ensure the project is on time and within budget. During this phase, the security team also creates a plan to monitor and maintain the ZT policies and network.
4. Testing (covered in *Zero Trust Implementation*): After implementation, various types of tests are run to prove acceptance criteria are met. These can be classified as systems readiness testing (SRT) and operational readiness testing (ORT).

Lastly, given the comprehensive, enterprise-wide scope of ZT, implementations are usually

[4] CISA. (2023). Zero Trust Maturity Model (Version 2.0).

incremental and iterative (as opposed to entirely new or cut-over implementations). In alignment with project management practices, this course refers to a single iteration of implementation as a **project**, and a collection of implementation projects pertaining to the ZTA goal and scope as a **program**.

The following unit will discuss the ZT project implementation tasks that should be considered before beginning your implementation preparation activities.

2 ZT Project Implementation Considerations

Before implementing ZT, a list of tasks and requirements should be considered:

- The gap analysis report should be consolidated and approved by stakeholders.
- Organizational security policies must include ZT-related security objectives.
- Migration from existing architecture to ZTA requirements needs to be addressed.
- Stakeholders must determine what part of ZT implementation will be done in-house.

2.1 Gap Analysis Report[5]

When you are ready to implement ZT, let's revisit the gap analysis your organization put together during the planning phase (see *ZT Planning*). Simply put, a gap analysis looks at what you have and compares it to what you need, reminding you to focus on each individual protect surface. By looking at each protect surface and what needs to be done, you can prioritize accordingly and start with your ZTA implementation–one protect surface at a time. This makes implementing ZT much more manageable than trying to do it all at once.

The gap analysis report identifies the steps needed to build a target ZTA, which should be prioritized and agreed upon by stakeholders who must coordinate their alignment with the ZT plan and its associated pillars and cross-cutting capabilities. For instance, those concerned with the Identity pillar might identify a security control missing from authentication and opt to introduce multi-factor authentication (MFA) as an appropriate measure to bridge the gap in authentication requirements. In a later section, we delve further into the details of ZT pillars and cross-cutting capabilities.

2.2 Aligning Information Security Policies with ZT

It is important to understand how information security policies connect with ZT principles. ZT is a concept that helps organizations strengthen their security measures. When you align ZT with information security policies, you need to consider the risks that ZT addresses and the security requirements of its core pillars. By considering these factors, you can develop stronger policies that enhance your organization's security posture.

[5] Learn more about "Gap Analysis" here: https://knowledge.cloudsecurityalliance.org/zero-trust-planning

Let's take the example of the ZT Identity pillar. To ensure alignment, you should examine your organization's identity and access management policies, procedures, and processes. Check if all the necessary elements, like the access control review procedure, are in place. If any of these elements are missing, or if ZT requires additional procedures, you may need to create new ones or make changes to existing ones.

2.3 Migration From Existing Architectures to ZTA

To effectively implement a ZTA, it's crucial to clearly understand the project scope and communicate it effectively with your team. This includes determining whether we are implementing an entirely new architecture or migrating from an existing one.

In this training, we assume that ZTA will be implemented in an already-existing environment with controls, whether on-premises, hybrid, or in a cloud-only scenario. To successfully achieve the desired ZTA, assessing the impacts and requirements of the technology involved is essential. When evaluating the architecture for any technology gaps, the team should consider the following:

> - Weighing the benefits and costs of introducing new technologies versus collaborating with existing vendors to enhance their product or service capabilities
> - Finding ways to simplify the environment if it becomes overly complex, aiming for a more streamlined approach
> - Moving beyond simply replacing technology, but instead advancing capabilities in a manner that aligns with the organization's goals and growth path

By considering these factors, you'll be better equipped to make informed decisions, ensure a successful ZTA implementation, and avoid unnecessary complexity.

Similar considerations apply in IT environments that have not yet implemented ZT; however, these systems have fewer dependencies on existing business systems and can be implemented more quickly. When the ZT system or modifications are operational, existing (and now redundant) systems should be decommissioned. This training applies to both situations.

2.4 Managed Service & In-House Implementation

ZTA is a combination of in-house and managed services. In-house implementation can be achieved internally by leveraging teams, such as InfoSec, identity and access management (IAM), and infrastructure, with possible contributions from specialized implementation consultants. A managed services approach includes various vendors that can provide solutions designed to support ZT strategies. While your team may strive to keep the implementation of ZTA in-house, there will always be a level of shared responsibility that needs to be orchestrated and considered.

To choose the best method, the following list identifies some considerations for ZTA implementation planners:

- Cost/benefit
- Capability
- Resource availability and your organization's skill-set availability
- Support
- Shared responsibility
- Policy
- Proof of concept and business proposal

3 Implementation Preparation Activities

Although we previously discussed implementation preparation in Zero Trust Planning, we summarize some essential kick-off activities here:

- Define ZT project deliverables
- Communicate ZT changes to users
- Create an implementation checklist

3.1 Defining ZT Project Deliverables

Prior to starting the ZT implementation, the overarching ZT project team must build some common deliverables that apply across the entire organization (or in-scope target architectures, if the scope is narrower than organization-wide) for each ZT pillar. By doing this with an agile approach, you can work quickly and efficiently, especially in areas that incorporate new automation or involve the development team. Due to the high likelihood that third-party stakeholders will be involved, use waterfall-based project planning models to help you develop milestones, where partial payments can be made for the percentage of work completed. Ideally, all pillars and cross-cutting capabilities should be worked on simultaneously while following the five-step process[6] outlined in the *Zero Trust Planning* training:

- Define the protect surface
- Map the transaction flows
- Build a ZTA
- Create a ZT policy
- Monitor and maintain the network

[6] NSTAC. (2022). NSTAC Report to the President on *Zero Trust and Trusted Identity Management*.

3.2 Communicate ZT Change to Users

All-hands meetings should be held on a regular cadence to ensure everyone is aware of the collective progress when it comes to implementing ZT pillars. Moving forward, ensure users are informed about upcoming operational changes, such as bringing in a new external consultant or introducing a cloud-based service. Communicating the exposed assets and their priority level will help with the successful adaptation of the ZT solution.

Changes to workflows should be communicated early, which has these and other benefits:

- Teams can plan for any adoption impacts
- Implementation teams can reach out to the group to ensure that testing covers all use cases
- Impact on infrastructure access during vital deliveries or engineering cycles can be minimized

3.3 Create an Implementation Checklist

Before kicking off the implementation, create a checklist of milestone activities that maps, where appropriate, the changes required. The list should include changes that will need to be made to:

- Organization's governance
- Compliance
- Risk management
- Operations and maintenance
- Visibility and analytics
- Incident management
- Change management
- Vulnerability and patch management
- Problem management
- Business continuity planning (BCP) and disaster recovery (DR)
- Training and awareness

3.3.1 Organization's Governance

ZT will likely require an update to the organization's governance approach and organizational policies, procedures, guidelines, and security controls. As the organization implements a ZTA, governance practices will guide the implementation functions, activities, and outcomes. The organization's senior leadership will rely on the technical expertise of functional technicians to develop and implement security controls that involve input validation, session management, and password storage. Integration examples of these controls are authentication, authorization, cryptography, input validation, output encoding, auditing and logging, and monitoring and alerting. Organizational governance assures sound stewardship of resources throughout the implementation process.

3.3.2 Compliance

When implementing ZT, it's crucial to remember that existing change control processes, compliance, and auditing requirements are followed, even if changes are needed. Whenever an information system within the scope of ZT undergoes changes, it's essential to follow established change control processes to ensure compliance.

As you keep working on implementing and refining your ZT approach, brace yourself for future laws, standards, rules, and regulations likely to impose more stringent requirements on ZT. Keep a keen eye out for changes in the compliance landscape as governments and organizations increase their insistence on security.

3.3.3 Risk Management

The implementation of a ZT strategy can change an organization's risk posture and risk management approach. To respond quickly to changes in the risk landscape, organizations must have a culture of continuous risk evaluation and policy adjustment. This means that the existing risk analysis and assessment process should include the following elements:

- An assessment frequency that allows for the rapid identification of new risks and new threats
- Metrics (tailored to the organization's networking trends), which are easy to spot in reports and monitoring tools, and that should also be able to send out alerts in extreme cases
- Data from various sources should be pooled together for more useful analysis

3.3.4 Operational Requirements

Operational or business-as-usual requirements promulgated by one or several departments will impose conditions or restrictions on ZT. To ensure that the ZT implementation stays within budget and on schedule, it is important to be prepared by identifying these processes and requirements before the start of an implementation. For example, integration between ZT automation and a configuration database must be agreed-upon and then installed and tested to ensure the operational readiness of agents or API calls.

3.3.5 Visibility & Analytics Integration

Log management and monitoring of cyber events drive visibility, which in turn supplies analytics that 'inform policy decisions, facilitate response activities and build a risk profile to develop proactive security measures[7]. For logs to be useful and to provide value when monitored, the following needs to be identified:

- Log scoping: for example, domain name system (DNS) logs, network address translation (NAT) logs, and intrusion detection systems/intrusion prevention systems (IDS/IPS)
- Log sources: for example, the identity provider (IdP), policy enforcement point/policy decision point (PEP/PDP)

[7] CISA. (2023). Zero Trust Maturity Model (Version 2.0). April 2023, page 11.

- Security events that need to be monitored: for example, failed authentication requests
- Anomalies that need to be detected: for example, five failed authentications within five minutes
- Correlation rules to identify threats or potential threats: for example, an identity that drops a production database is deleted
- Dashboards for visualization of logs: for example, a dashboard that shows privileged activities carried out by all administrators
- Monitoring: for example, continuous monitoring of any network connections made to a known command-and-control
- Alerting: for example, using emails, SMS, security information and event management (SIEM), or security orchestration, automation, and response (SOAR)

The above planning and design ensures that your implementation team leverages all log sources, collects and sends security events to SIEM, and identifies SIEM interfaces (e.g., add-on applications, API calls, and event collectors). These activities prevent mishaps caused by implementers scrambling to get these interfaces together during the implementation phase.

3.3.6 Vulnerability Scanning & Patch Management

While vulnerability scanning and patch management are standard IT practices, ZT acts as a control gate to ensure all systems are patched and business operations can only be performed on a patched system. Identifying the vulnerability scanning and patching requirements for components, such as IdP or PEP, of ZT implementation ensures that implementation services become part of the existing patching process that the organization has. By identifying vulnerability checks and tests, wherever possible, your organization can continue to work quickly and harvest all the benefits of technology without leaving the organization vulnerable to a hack. By identifying patching requirements, you can build patching checkpoints that monitor for unpatched or vulnerable software components within your environment and ensure all components are kept up to date. It may also assess the current security status of an operating system or application, identifying any available patches and upgrades, testing these to ensure compatibility with existing applications, implementing necessary updates, and monitoring for successful patch installation.

The absence of vulnerability scanning and patch management may defeat the very purpose of implementing ZT because a vulnerability may diminish the objective of authentication before authorization.

3.3.7 Change Management Process

ZT needs to be included in the change management process. During this activity, it is important to design and track separate workflows for:

- A service request
- A change request

For example, the onboarding of a device can be treated as a service request. However, adding a policy to the PEP, also known as a gateway, may be treated as a change request. Suppose a change

request is needed because an unforeseen gap has been found during the ZT implementation. This change may require an emergency budget, developer- or engineering-based redesign, or consultation from a third party. This type of change may need modifications to the code, the network, or the environment and will need to be queued into the problem management process.

3.3.8 Problem Management Process

ZT should be integrated into the organization's problem management process to address any recurring incidents, future incidents, and methods for reducing the impact of incidents that cannot be prevented. Problem management processes should cover problem detection, problem logging, error control, and root cause analysis, to name a few.

3.3.9 Incident Management

Incident management is focused on addressing incidents in real-time and may need to be revised based on the changes the ZT approach will bring to the organization. As part of this, the ZT implementation team needs to define what constitutes a security incident in a ZT solution and what constitutes a significant security incident. The team also needs to identify the list of people that must be contacted in the event of a significant incident, along with their contact info.

3.3.10 Business Continuity Planning & Disaster Recovery

BCP and DR are important processes that should be aligned with ZT implementations. As ZT deals with access to resources to carry out IT activities, existing access processes may not work as expected due to ZT implementation activities. Business continuity activities should be planned during implementation such that disaster recovery can be addressed during this phase.

3.3.11 Training & Awareness Programs

New technologies, architectures, and solutions introduce new workflows and ways of working. While preparing for ZTA implementation, it is important to keep training requirements in focus. An implementation solution will be productive and effective if assigned teams are comfortable using it, which comes from regular training and awareness-building.

To do this, list all ZT implementation training and awareness programs that need to be scheduled. Also, identify the frequency at which the programs need to be scheduled. This will help organize and budget training. While the details will vary, depending on what the implementation entails and the nature of your organization, training requirements may be categorized into two groups:

- Business and operational goals
- Technical goals

Business and operational goals include items, such as:

- Training and awareness that publicizes the ZT implementation business goals
- Building awareness about responsibilities, such as the budget owner, asset and process owner, legal team, architects, security team, and IT team
- Training architects
- Training support personnel
- Training for risk managers and compliance officers
- Training to understand new forthcoming operational processes, such as change management, incident management, BCP drills, and so on

Technical goals include items, such as:

- Enrolling endpoints to ZT architecture
- Creating and maintaining policies
- PEP and PDP maintenance
- Creating and maintaining micro-segmentation
- Maintaining identity provider services and databases by working with identity creation, access control reviews, and suspension of identities belonging to offboarded employees

Feedback should be obtained during these training sessions so that updates can be made to the training and awareness program.

4 ZT Target Architecture Implementation

We have just explored a series of milestone activities that must be reviewed and considered to support a ZT target architecture implementation. With the above objectives in mind, this section will explore the following implementation themes:

- ZT helps in reducing access related compromises by authenticating and validating the access request.
- Organizations must have the right personnel and systems to monitor suspicious activity and policy violations.
- Integrate ZT testing efforts into the cybersecurity program and promote collaboration between departments to ensure the organization maintains secure operations.
- Prioritize risk management, review transaction flows, continual improvement, and project closure.

Imagine a real-life scenario: We have planned and prepared the following ZT reference architecture, as shown in the diagram below.[8]

[8] Note: Please refer to CSA trainings: Introduction to Zero Trust Architecture and Zero Trust Planning for a review of the components depicted in this diagram and their respective functions. These components are also defined in CSA's Cloud Security Glossary.

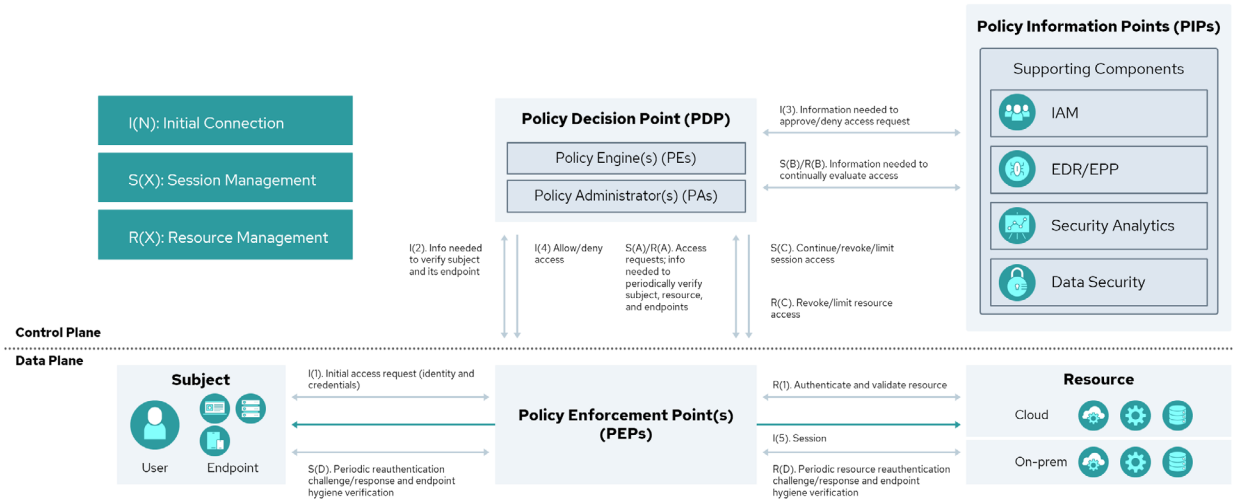


Figure 1: *General ZTA Reference Architecture*[9]

This diagram captures the flow of operations when the subject accesses a resource. When a valid subject wants access to a resource, the request is initiated from the endpoint to the PEP in the following steps:

1. An initial client authentication request is sent from an endpoint to the PEP.
   NOTE: The operation may repeat based on the identified ZT requirements.
2. Information required to verify the subject and endpoint is collected by the PEP and shared with the PDP.
3. The PDP validates the device and subject authentication. At the final, advanced stages of ZTA implementation, this point can be integrated with various policy information point (PIP) solutions and technologies. If the validation succeeds, the PDP decides on the type or level of authorization needed.
4. The PDP informs the PEP about the authentication status for the connection and authorization details if obtained.
5. The session established to check a user's credentials or endpoint is now terminated. The established sessions will undergo periodic validations and terminate per the predefined rules.

It is important to note that this flow of operations is separated by the control and data plane. The control plane decides the path for sending packets or frames and directs how these packets should be forwarded. The data plane is where the action takes place; it's all about the functions and processes that move those packets from one interface to another.

Having discussed the ZTA reference architecture above, let us dive into the ZT pillars and cross-cutting capabilities with the intent of studying their attributes in more detail. This will help us better design the ZT implementation across all pillars and cross-cutting capabilities.

[9] Figure adapted from: NIST. (2022). Implementing a Zero Trust Architecture (SP 1800-35B). Second preliminary draft.

4.1 Zero Trust Pillars & Cross-Cutting Capabilities

One way ZTA implementation can be coordinated is within and across the ZT pillars and cross-cutting capabilities. As ZT functions are added, each pillar should mature and evolve from the traditional level at the start to the optimal level as you near the final implementation phases.

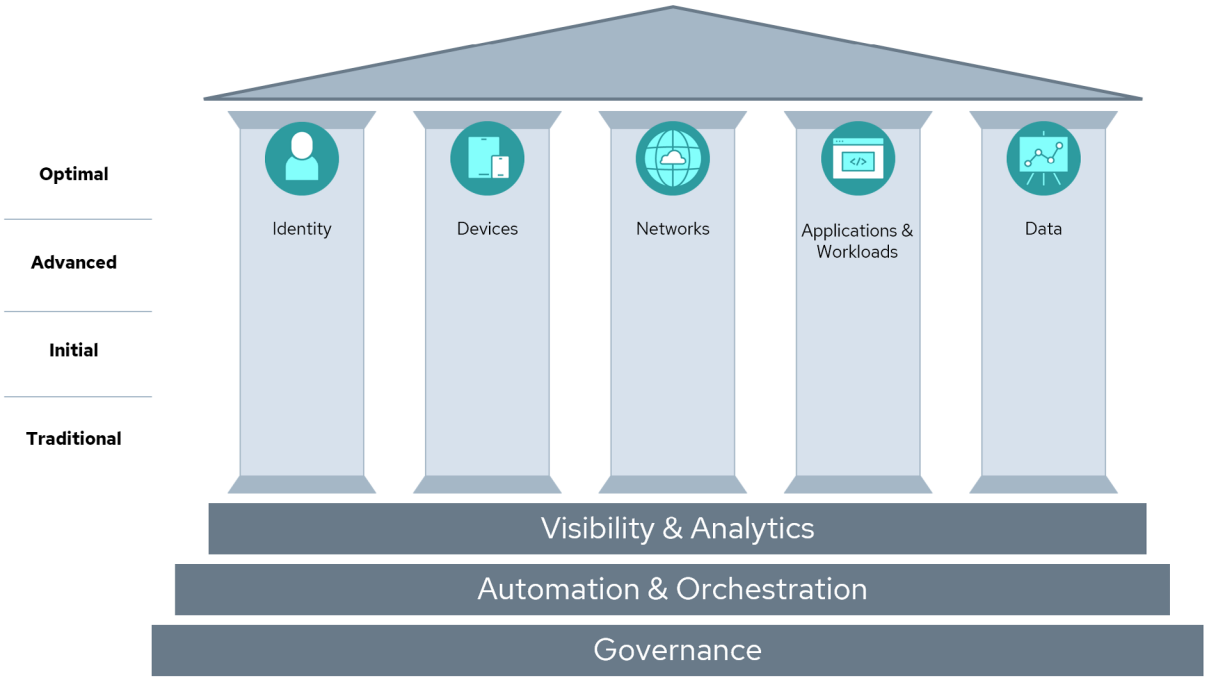


Figure 2: *Zero Trust Maturity Evolution*[10]

As displayed above, the five ZT pillars are:

- Identity
- Devices
- Networks
- Applications and Workloads
- Data

Cutting across these pillars are capabilities, referred to as cross-cutting capabilities, that improve at every maturity level and interact with each pillar:

- Visibility and Analytics, primarily by aggregating output
- Automation and Orchestration
- Governance, by introducing governance and compliance software

In some diagrams (i.e., from the US Department of Defense[11]), the cross-cutting capabilities (Visibility and Analytics, Automation and Orchestration, and Governance), are depicted as foundational pillars. This representation emphasizes the significance of incorporating these capabilities into the implementation process as they assist in defining objectives for the five pillars.

[10] Figure adapted from: CISA (2023). Zero Trust Maturity Model (Version 2.0).
[11] U.S. Department of Defense. (2022). DoD Zero Trust Strategy.

Before covering each pillar and cross-cutting capability in depth, the relevant attributes for each pillar and cross-cutting capability should be identified and defined, as noted in the following table.

Pillar	Pillar	Attribute	Notes
P	Identity	• Identify all the identities that are required to be part of the ZTA • Identify the identities that requireaccess to the protect surface • Define permissions for identities according to least privilege using custom roles (a group of individual permissions) and assigned to a role (group), except in situations where access may be restricted to a single service account or other non-human identity • Monitor modification/drift for the membership in the role/group as well as the individual permissions associated in the custom role	• This should be treated as infrastructure components
P	Devices	• Identify all devices that need to be enrolled • Identify business & security applications running on devices	
P	Networks	• Identify the necessary macro-segmentation in data centers& micro-segementation	Micro-segmentation between hosts within the: • VLAN • VPC • CNet
P	Application & Workloads	• Identify your applications and workloads that exist in your on-prem or cloud infrastructure	
P	Data	• Identify data sources that are part of your protect surface • Identify transaction flows	
C	Visibility & Analytics	• Identify the registry (logs) repository in all entities (users/ identities, devices/endpoints, network and environment, and applications and workload) • Identify the external data that you need to enrich the visibility • Identify the parameters you need to monitor the performance, behavior, and activity of the ZT deployment	

C	Automation & Orchestration	• Identify the security operations to perform when a policy allows or denies actions • Identify the standards that support the security and non-security technologies to communicate with others • Identify the conditions and technologies of the security playbooks	
C	Governance	• Identify the governance structure that ZTrequires	
		• Policies and procedures	• Organizational security policies: • Identify the information security policies that ZT deployment needs to adhere to
		• Controls	• Identify the security controls that need to be applied to the devices, network, applications and workloads, and data • This should not be confused with the transaction flow controls • Identify the controls to be implemented across the user agent, such as authentication, authorization and various other aspects • Identify the rule-based access policies that are part of the PDP, also referenced as the controller, and have been identified in the planning session
		• Risk management	• Identify the risk managment requirements
		• Compliance	• Identify the risk compliance requirements

Table 1: *Implementing ZT Across Pillars & Cross Capabilities*

Now that we have seen overviews and attribute summaries of the pillars and cross-cutting capabilities, the remainder of this unit will discuss implementation-specific details and considerations that tie these to executing the implementation.

4.1.1 Identity

Identity is the pillar involving authentication and authorization, including privileged access management (PAM). Effective use of identity as a data source will include centralized directories (as

well as related onboarding strategies) and federation between enterprises. The rules and validations related to identity are implemented at the PDP, where a predefined IdP provides user management and other validations, such as risk assessments and device posture validations. The IdP can be within the enterprise or external to the enterprise (e.g., as a SaaS application). When the IdP is within the organization's network, the PDP can reach the IdP through the PEP.

4.1.1.1 PDP Identity

PDP users that monitor and manage the PDP–we can call them PDP admins–must create policies and perform maintenance on the ZTA at the control plane. To reduce the impact radius of potential cyber-attacks, each of these PDP administrators, and any other user with elevated permissions, needs at least two identity profiles:

One for their day-to-day activities, such as reading emails, surfing the web, or using the ticketing system. This primary identity should hold no elevated entitlements or roles.
Another is for elevated access, such as ZTA management permissions (for creating ZT and software-defined perimeter [SDP] policies, applying patches–often automated with a PAM system, and reading logs).

4.1.2 Applications & Workloads

Imagine that the ZT implementation team tasked with determining the organization's current state has identified the IPs, applications, and workloads during the planning phase (see *Zero Trust Planning*). To satisfy access needs, these elements should be configured and organized separately at a PDP (also referred to as a controller) for onboarding, and it should be organized based on access needs. This will require centralized authentication, authorization, and monitoring, as well as segmentation of application groups.

Policies are defined at the policy administrator level to provide the high-level access needed for these assets. At the traditional stage, access needs are often organized by job function or role. This should be revamped as ZTA implementations mature through the stages; traditional, initial, advanced, and optimal, and permissions and work should instead be organized according to access needs, enabling efficient organization based on security policies.

4.1.3 Networks & Environments

To ensure the security of networks and environments from malicious actors, unauthorized visibility, and unauthorized access, organizations should embrace ZTA security best practices. These include identity-based policies, session establishment and termination, micro-segmentation, installation of PDPs (to validate authentication) and PEPs (access configuration and enforcement decisions), as well as redundant PEPs for failover and load balancing so that services can be continued even if one component fails. The ZTA diagram below summarizes one way to structure the various security components, dividing them between what is handled in the control plane versus what is handled in the data plane.

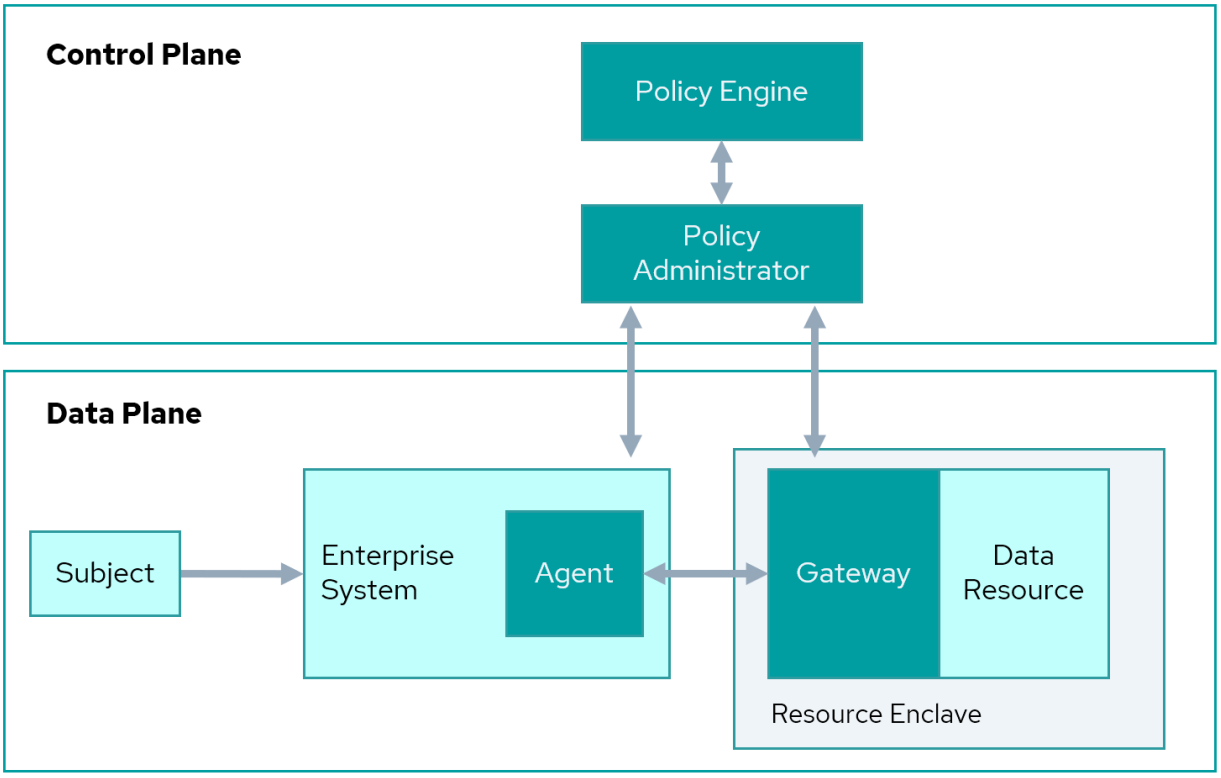


Figure 3: *Enclave Gateway Model*[12]

To assist in fulfilling the goals of this pillar, break down the network and environment considerations into these six types of network signals, which we discuss in more detail in the remainder of this section:

- Initial client authentication request reaching from agent to PEP
- Authentication request validation request (AR/VR) from PEP to PDP
- Decision transmission from PDP to PEP
- Session establishment and termination from client to resources
- Micro-segmentation
- PEP installation and access configuration

4.1.3.1 Initial Client Authentication Request

Establishing a secure connection between the application and server is essential to ensure mutual authentication. This ensures that both parties are verified before communicating with applications with sensitive data access. To initiate client authentication safely for PEP, you should:

- Position the PEP at the network perimeter while keeping other components, such as PDP and resources, on separate network segments
- Ensure that the agent can send the initial authentication request to PEP which will, in turn, forward the request to PDP
- Configure an encrypted channel for authentication request transmission

Building ZTA with user-agent-initiated access at PDP for user authentication is also possible.

[12] Figure adapted from: NIST. (2020). Zero Trust Architecture (SP 800-207).

4.1.3.2 Authentication Request/Validation Request

AR/VR is an important part of ZTA. AR/VRs help ensure that only authorized requests can be processed and approved by authorized entities, helping to prevent identity spoofing and other malicious activities.

The user agent can securely share their credentials with the PEP, which will be forwarded to the PDP for validation. The PDP then verifies the user or subject's credentials and initiates an additional MFA process. Once verification is complete, the authorization data is shared with the PEP. To ensure secure communication between the PEP and PDP, network access should be configured to allow for incoming and outgoing transmissions only between these two parties. Additionally, authentication and a secure channel of communication should be established.

4.1.3.3 Decision Transmission

Decision transmission is an essential component of ZTA, as it enables the PDP to make an informed decision about access based on user- and context-based information. This ensures that users are granted only the least amount of access required to perform their job duties, protecting data from unauthorized access.

To ensure secure transmission of data between the PDP and the PEPs, you must:

- Configure your network access to accept incoming and outgoing transmissions from only the PDP and PEPs
- Set up authentication between the two
- Perform periodic re-authentication challenges

Doing so will help to guarantee that data is kept safe and secure.

4.1.3.4 Session Establishment or Termination

To ensure secure access to their network, organizations must establish and terminate client sessions in a way that verifies the identity of clients, validates session data, and prevents person-in-the-middle attacks. Session termination is particularly important for businesses that allow privileged professionals such as company directors or medical doctors to log on from any machine within the work environment. To properly establish and terminate client sessions requesting resources, organizations should:

- Configure their PEPs to respond only to the initial authentication request
- Manage client sessions according to the authorization decided by the PDP

4.1.3.5 Micro-Segmentation

Micro-segmentation is a key component of ZTA solutions, which helps improve network security and simplify its management. Instead of creating multiple rules based on addresses, identity-based policies can be used to secure segments effectively. It is achieved by dividing resources into several

distinct network segments using either network devices, such as switches and routers, or by using host-based micro-segmentation with software agents and endpoint firewalls. The security gateway then grants access based on authorization obtained from identity attributes and must be managed to act as a PEP for protecting resources from unauthorized access.

The ultimate aim of micro-segmentation is to establish boundaries between resources within the same network zone and ensure that only authorized entities have access to secured assets.

4.1.3.6 PEP Installation & Access Configuration

Once the PEP is installed, the following checks should be conducted to improve the security stack which already supports ZT, such as port knocking and single packet authorization (SPA) for obfuscation: assess the accessibility of the device to both the PDP and endpoints at the edge of the network. The following checks can be done once the PEP is installed:

- PEPs should be able to enforce the identified authorization for access based on the policies defined at the PDP
- Ensure the PEP receives policy updates from the PDP/policy engine (PE)
- User endpoints at each in-scope location are able to reach the PEP through the network
- Ensure all in-scope PEPs and data feeds are integrated
- Establish redundant PEPs of the same kind for failover and load balancing based on scale and requirements. Similar to running a next-generation firewall (NGFW) in high availability mode, this ensures that if one PEP component fails, a second one will take over

It is important to continuously monitor the network for suspicious activity and regularly review access controls to ensure that your system's security remains intact.

4.1.4 Data

An important part of the protect surface is data (along with other resources). All data (not just information) can be better protected with a ZT implementation because ZT mandates that access decisions be made as close to the resource as possible. To apply ZT to data, it is necessary to discover, inventory, categorize (or label), and control data.

The implementations that are part of the Data pillar are done at the PDP. Using the data inventory to identify where data is located, the identity store to identify those who need access to it, and the PDP to define transaction flows for authentication and authorization of access, organizations can ensure that the necessary security measures are in place.

Logs from the PDP and PEP should be shipped to a SIEM to maintain visibility for granted accesses. This allows organizations to understand better who has access to which data and when. Ultimately, this helps them maintain the security of their data and protect their data from any potential misuse.

4.1.5 Devices

ZT can be implemented or enabled for any device in an organization. These include but are not limited to PCs, servers, mobile devices, and any OT or IoT device, to name a few. The scope for supported devices is predefined in the planning. The architecture implementation can take one of two forms:

1. Agent-based access: when a software client is installed onto the device. This sometimes encompasses other features besides the ZT agent, bundling traditional endpoint security with productivity tools or business apps; or
2. Agentless access: agentless options are deployed to devices lacking the ability to have an agent installed but can also be deployed to devices that can accept an agent. Agentless can further be subdivided into two subsets:
    - On devices that support browsers: In such scenarios, a connection using a secure tunnel to cloud/SaaS services is established through a plugin or manual configuration, which then handles all inspections.
    - On devices that cannot support browsers and agents: The entire site or micro-segment is connected using a tunnel to the cloud or SaaS service. This option may be more applicable to OT and IoT devices.

Agent-based access can also be configured on OT systems based on the implementation strategy and OT architecture.

Sometimes, you will also encounter a bring your own device (BYOD) environment, which refers to employees being allowed, and sometimes encouraged, to use personal devices to complete their work for the employer. Whether your organization's ZT policy is to use agent-based approaches or agentless access methods, BYOD scenarios will impact aspects of your ZT implementation. For example, deploying agents will require you to add a privacy notice if this agent is being installed on a personal device. Your team will need to coordinate with governance, compliance, and legal teams to confirm your messages are correct and in accordance with local law.

When an employee-organization agreement terminates, you need to ensure that the decommissioning of ZT-related agents and apps involves the decoupling of security solutions from a BYOD device. To complete these tasks, employees need to know the policy and procedure. They may also need to be reminded that some of the protection that they previously relied upon is being removed.

4.1.5.1 Deploying Agent-Based Access

In the case of agent-based approaches, a software agent must be installed and run on all endpoints. It is then the agent's job to collect the user identity and share the security posture data of the device and connection. Agents must be regularly updated – either automatically or as part of the company's patch process. To ensure compatibility with different device types, such as Mac OS, Linux, and Windows laptops, they should be tested appropriately before being deployed on end-user devices. The setup process and usability factor should also be evaluated prior to mass deployment.
The available options will depend heavily on the target environment (e.g., Windows, Linux, Android,

iOS), and the selected solution may vary in how the agent is finally deployed.

Based on current IT trends, agent deployment will likely be based on unified endpoint management (UEM) and mobile device management (MDM). However, most vendors provide a download console for agent installs.

4.1.5.2 Deploying Agentless Access

The agentless access method can be used to deploy ZT onto an endpoint device, such as with a browser or browser-based application. The browser is responsible for assessing the security posture and connection of the device prior to allowing access. This deployment method is also useful when dealing with light devices that do not have a browser available, like OT devices. In this case, a verified network with a proxy handling the agent load and building the connection can be established. After authentication of the user through an IdP (which may incorporate single sign-on [SSO] or MFA), they will then be redirected to their requested resources if authorization is granted.

4.1.6 Visibility & Analytics

Implementing ZT requires a platform-based approach to security, empowered with analytic and visibility dashboards that authorized personnel can use to make policy changes. To achieve this, agents or APIs should be implemented to gather logs from various log sources.

Depending on the requirements of the log aggregation tool, a log collector may be necessary. Aggregated logs enable event correlation that is more powerful in threat analysis and discovery than if the logs are kept separate and analyzed separately. The visualization component of the central log aggregation tool can provide dashboards, enabling visibility of the inner workings of each pillar.

To achieve the above-mentioned set of goals:

- Implement agents or APIs to acquire logs from each of the logs' sources.
- If required, implement a log collector.
- Ship the logs to the log aggregation tool using either a push or a pull mechanism, depending on the tool's requirement.
- Implement a search tool that can provide a query language that can be used to search logs for events and event correlation.
- Design dashboards to view query and event correlation output.

Once a visibility dashboard has been fully implemented, potential threats can be identified and monitored through careful observation of the dashboards.

4.1.7 Automation & Orchestration

The ability to orchestrate and automate the deployment of the target architecture's logical components is key to a successful ZTA implementation. This includes tasks such as updating the ZTA components' security posture, dynamic access and authorization policy updates, patch

management, and change management. These deployments can span on-premises, cloud, or hybrid implementations.

Depending on the deployment model identified in the target architecture, two types of orchestration methods can be used for ZTA deployments: application pipeline and infrastructure as code (IaC) pipeline. In most cases, a combination of both is used. The IaC pipeline is typically used when the PEP is a single component acting as a gateway for the subject requests. In contrast, the application and pipeline are commonly used when an AuthZ module and other modules must be deployed directly on the component. This could include deployment within the application code.

Depending on the orchestration methods and target architecture, different functional and non-functional orchestration requirements will exist. However, it is important to note automation and orchestration can be achieved in some cases but not all. For example, achieving it in OT or industrial control systems (ICS) is very challenging. It is important to factor in this while implementing ZT for such technologies.

4.1.8 Governance

Governance of the ZT program and individual ZT projects is essential to ensure successful implementation and control over goals, requirements, and actions taken. A formal procedure for governance should be established through a review committee that will evaluate the progress made towards meeting objectives, ensuring that plans are funded, and assessing associated risks with future phases.

As part of a successful ZTA implementation, it is essential to establish a formal review process headed by senior management. This committee will ensure that appropriate ZT requirements are observed and that the organization has the necessary resources to complete the ZTA plan. Their main objectives include:

- Verifying that each phase is completed with success
- Ensuring sufficient funding for the next phase
- Assessing the risks associated with continuing to the following phase[13]

At the start of implementation, metrics need to be defined and collected to measure set parameters. These metrics can range from high-level indicators, such as the number of goals achieved, budget consumption, and impact on organizational policies, to lower-level metrics, which include the number of support tickets raised, complaints by end-users, policy changes, failed policies, and downtime incurred. All these should be identified.

4.1.8.1 ZT Policies

ZT policies are used to bridge the gap between a business's mission and its risk management requirements. These policies are documented and set up within the ZT planning process. The PDP and PEP communicate in near real-time to ensure that authorization decisions made by the PDP,

[13] United States Department of Health and Human Services. (2012). Enterprise Performance Life Cycle Framework.

based on the policies defined at the PE, are enforced. Policies can be applied to users, applications, and workloads during the PDP onboarding stage. These policies will define access conditions for each user or device based on parameters such as location and time.

To effectively manage a ZTA implementation, it is important to save some of your authorization work for the implementation stage. This is because a macro-level approach will focus too much on the overall architecture and will create extra work each time the overall architecture changes. Instead, your detailed policy rules should focus on each PDP and PEP technology separately. For example, a privileged access workstation (PAW) can only be accessed by a specific user or administrator from their approved device. In that case, there are various places where the policy could be updated. One could be the local permissions on the PAW host. Another could be the micro-segmentation firewall rules for the user and device hostname. Lastly, a network access control (NAC) or UEM solution could check the user's local device posture.

> - Attack surfaces change quickly, altering risk. Policies must be updated regularly by:
> - Re-evaluating the updated transaction flows for changes in risk
> - Re-implementing a macro- and micro-level approach to access and authorization policies
> - Ensuring that updated policies are applied to the policy engine at the PDP

4.2 Transaction Flow Architecture Review

Maintaining a transaction inventory allows you to reevaluate the behavior of the data within each transaction at regular intervals and, more importantly, detect any changes or abnormalities.

In the planning phase (see *Zero Trust Planning*), we discussed the need for a detailed analysis and mapping of existing transaction flows. When the time comes to implement any identified changes to these transaction flows as part of your ZTA implementation, there will be some considerations to address, such as the addition of ZT nodes, services, and components. You will also need to address legacy controls that need to be replaced to protect existing transaction flows.

Mappings may need to be reviewed against your planning notes. This will happen often, both during and after your implementation. As we discussed in the previous section (ZT Policies), you don't want to focus too much on the overall architecture; not only will it be too complex to manage easily, but it will create extra work each time any part of the architecture changes. Instead, your detailed transaction flows should focus on each individual protect surface.

By orienting your transaction flow diagrams or transaction inventory to each protect surface, you can easily manage change down the road.

4.2.1 Transaction Flow Mapping

In the context of ZT, we could record the transactions that are part of the subject-initiated access to the applications and network or between the access-controlled applications, as long as we are centered on the surface we need to protect.

The way that you map your flow is always unique to your implementation. More often than not, this process is manual. While some automation tools to help you with mapping are emerging, such technology is still catching up to the need.

The best way to maintain a transaction inventory of the transactions you are managing, along with their dependencies, will depend on the unique characteristics of your team and the organization's IT architecture. Nevertheless, in the next section, we provide you with a brief example.

4.2.2 Converting Flow Maps to Transaction Lists

In the example below, the management team chose to invest the time to map all transactions involving their protect surfaces with a professional flowchart tool that their network engineers are proficient in. This will help them communicate with the vendors they use to receive security-oriented services. However, they also want to have each transaction listed because they plan on using this as a checklist during implementation and, down the road, as a potential set of requirements that can guide future development projects.

The ZT implementation team determined that this particular protect surface involves six transactions, each with a unique identifier (see Figure below). Dependencies are shown.

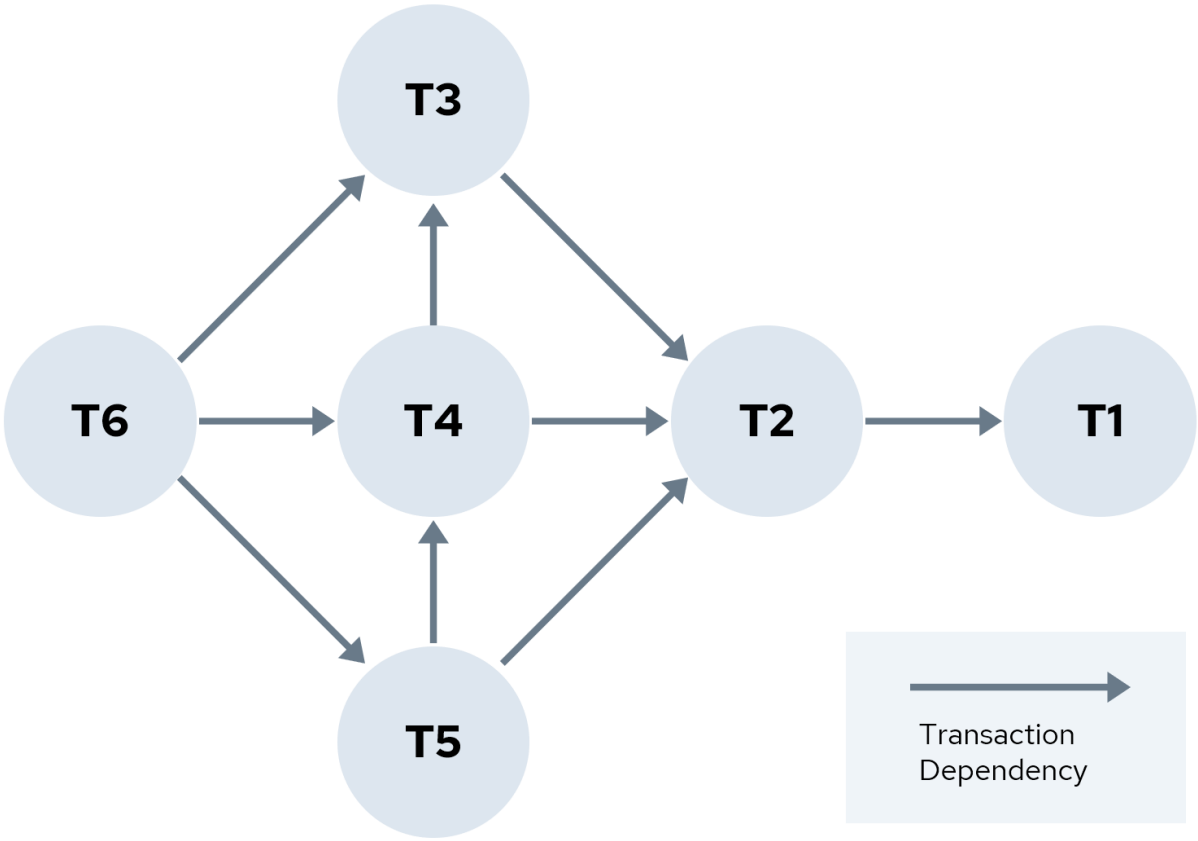


Figure 5: *Transaction Inventory*

Next, each transaction path is assigned a **Transaction ID**. Any transaction it depends upon is entered in **Input**, and any transactions dependent upon that identified transaction are labeled **Output** (see Table below). The table can also include optional values, such as a **Service ID**, a **Title** (not shown), or **Description** (not shown).

Transaction ID	Input	Output	Service ID
Transaction 1 (T1)	T2		Service Baseline A
Transaction 2 (T2)	T3, T4, T5	T1	Service Baseline B
Transaction 3 (T3)	T6, T4	T2	Service Baseline C
Transaction 4 (T4)	T6, T5	T3, T2	Service Baseline D
Transaction 5 (T5)	T6	T4, T2	Service Baseline E
Transaction 6 (T6)		T3, T4, T5	Service Baseline F

Table 2: *Transaction Configuration Management Inventory*

This table can be entered into a configuration management system, unified modeling language (UML) tool, or programmable logic controller in reverse. Such tools can help provide indexable and searchable data that can help troubleshoot problems and document fixes. You can even use it to create more complex ladder or signaling diagrams.

4.3 Testing

After an organization has completed implementation, it must develop and maintain relevant policies, procedures, and agile testing scripts that define how the ZT testing process works and should be conducted so that the testing methodology remains consistent from one implementation stage to the next. To ensure that the planned ZTA delivers the intended service levels, and before the legacy architecture can be decommissioned, a test cycle must be completed for ZT implementation, including testing on non-production and production environments. It is important to isolate whether a problem originates from the implementation itself or is the result of new technology merely catching a problem that existed in a pre-existing solution or data but was not caught due to weaknesses in the previous setup.

Security testing must ensure that access controls are working correctly and the network is protected from threats. This testing includes vulnerability scans, penetration tests, application security assessments, system readiness testing, operational readiness testing, and other forms of security testing. Testing must:

- Confirm that the ZT objectives were achieved.
- Provide evidence that continuous authentication and authorization over all communications, users, systems, and networks is taking place.

- Provide evidence that employees are able to complete their work with a minimum of disruption.
- Create a secure baseline from which future changes can be monitored for potential threats or vulnerabilities.
- Confirm that there is a robust audit trail to monitor suspicious activity and policy violations.
- Ensure that the organization is on the path to maintaining effective communication between monitoring systems and personnel to ensure the efficiency and thoroughness of the ZT process.

Regardless of the test type or environment you are testing, each testing activity must align or refer back to the ZT plan and strategy. Integrate ZT testing efforts into the overall cybersecurity program; doing so will provide enhanced protection. Additionally, each planning objective in each pillar needs to be sufficiently tested to confirm that your team has met its objectives.

Finally, promote collaboration between departments to ensure ZT testing results in secure and efficient operations across the organization. When all test phases are completed, production can begin cut-over, also in phases.

4.4 Continual Improvement

During implementation, the ZT project should proceed in a continuous feedback loop, with each pillar, sub-project, and related effort recorded in a task repository that can be analyzed by a project management expert and analysis tool.

However, don't stop with task monitoring. Additionally, organizations should perform regular audits to ensure their policies and practices are followed. These audits should test current security measures and identify any potential gaps in the system that malicious actors could exploit. This way, organizations can ensure their networks are secure and compliant with best practices.

This feedback loop drives future ZT adjustments, as needed, in response to any encountered challenges, timeline modifications, or other reasons. ZT projects can take considerable time to implement. Any technology changes in the environment require review and consideration at the ZT project level to achieve the original goals and possibly make adjustments to those goals, with approval from all relevant stakeholders.

A key component of any ZT project should be proper risk management. As the ZT project is implemented and adjustments are made upon encountering a changing environment, the feedback loop should trigger a re-evaluation of risks to the ZT project, which, in turn, triggers ZT-related change control.

4.5 Project Closure

Successful ZT implementation would include a complete inventory of all transactions, dependencies, and services with associated IDs. Policies and procedures should be developed to ensure consistent testing methodology across different stages of the implementation. Security tests such as

vulnerability scans, penetration tests, application security assessments, and system readiness testing should have been conducted to protect the network from potential threats. Tests should be carried out on both non-production and production environments to ensure the quality of the implementation. All legacy architecture should be decommissioned once all tests have been completed successfully. Finally, sufficient documentation must be created to ensure future troubleshooting is easier and fixes can be documented accurately.

With all the project-related tasks completed, the project needs to be formally closed. For successful operations and maintenance, important policies, procedures, and processes must be reviewed and maintained. Final sign-offs must be obtained from key stakeholders, and a "go-live" date must be communicated to end users. From here, operations and maintenance cycles will then follow.

Conclusion

ZT is an important security strategy that must be planned, tested, and monitored for optimal effectiveness. ZT implementations must be iterative so that their efficacy can be evidenced while assessing results; this better prepares future implementations that lead to higher ZT performance levels with the ultimate goal of having proactive monitoring and logging feedback before a malicious actor can achieve a breach.

Furthermore, organizations should integrate ZT testing into their overall cybersecurity program to provide enhanced protection. Additionally, regular audits of policies and practices should be conducted to identify potential security gaps in the system. A key component of any ZT project should also involve proper risk management and knowledge management to ensure lessons learned from incidents are not repeated.

With the help of these strategies, organizations can increase their ZT efficiency and proactively protect their operations or security against malicious actors.

Glossary

Please refer to our Cloud Security Glossary, a comprehensive glossary that combines all the glossaries created by CSA Working Groups and research contributors into one place.

Acronym List

Acronym	Term
API	Application programming interface
AR/VR	Authentication request/validation request
BYOD	Bring your own device
BCP	Business continuity planning
CSA	Cloud Security Alliance
CISA	Cybersecurity and Infrastructure Security Agency
DR	Disaster recovery
DNS	Domain name system
IdP	Identity provider
ICS	Industrial control systems
IaC	Infrastructure as code
IoT	Internet of Things
IP	Internet Protocol
IDS	Intrusion detection systems
IPS	Intrusion prevention systems
MDM	Mobile device management
MFA	Multi-factor authentication
NAC	Network access control
NAT	Network address translation
NGFW	Next-generation firewall
OSI	Open Systems Interconnection
ORT	Operational readiness testing
OT	Operational technology
PC	Personal computer
PDP	Policy decision point
PEP	Policy enforcement point

PE	Policy engine
PIP	Policy information point
PAM	Privileged access management
PAW	Privileged access workstation
SIEM	Security information and event management
SOAR	Security orchestration, automation, and response
SMS	Short message service
SPA	Single packet authorization
SSO	Single sign-on
SaaS	Software as a service
SDP	Software-defined perimeter
SRT	Systems readiness testing
UEM	Unified endpoint management
UML	Unified modeling language
ZT	Zero Trust
ZTA	Zero Trust Architecture
ZTT	Zero Trust Training