National Security Agency | Cybersecurity Information Sheet

Advancing Zero Trust Maturity Throughout the Data Pillar

Executive summary

This cybersecurity information sheet (CSI) provides recommendations for maturing data security and enforcing access to data at rest and in transit, ensuring that only those with authorization can access the data. It further discusses how these capabilities integrate into a comprehensive Zero Trust (ZT) framework, as described in Embracing a Zero Trust Security Model. [1] Traditional security approaches have often relied on perimeter defenses alone to secure networks. Recent events highlight that adversaries who are successful at gaining a foothold in information systems often readily gain unfettered access to all data in those systems. By applying the recommendations in the data pillar, including identifying risks to data, integrating granular data attributes into access control mechanisms, and monitoring data access and use, organizations will reduce the impact and consequences of breaches and identify suspect activity earlier in the cyber intrusion lifecycle.

To protect data, an organization needs to know what data it has and track how it moves and is accessed inside and outside the enterprise. Tracking data can be a significant task, so having an automated method for identifying data of value on the network or performing a data inventory operation is recommended. Data protection ensures that data is only accessed by authorized entities. Granular control of data not only keeps it safe within the enterprise, but also ensures that it can be safely shared with other organizations and partners to achieve interoperability. Implementing these activities will limit the ability of adversaries to reach targeted data assets. It will also provide visibility to system managers of compromised assets that require mitigation should adversaries be successful in their efforts.

Introduction

In September 2017, a major credit reporting agency (CRA) reported it had been the victim of a data breach resulting in the theft of records from 148 million American customers. The stolen data included highly sensitive personally identifiable information (PII), such as social security numbers, credit card numbers, dates of birth, residential records, and driver's license numbers. [2] The incident began with access to a vulnerable server, whereupon PII from dispute resolution documents was stolen and additional login credentials obtained. The cyber threat actors then used those credentials to penetrate deeper into the network and pilfer a staggering amount of data over a 76-day period in which they accessed 51 different databases. [3]

As one of the nation's largest CRAs, this company's data was highly valuable and the loss of it extremely costly to itself and its customers. The CRA agreed in 2022 to a global settlement with the Federal Trade Commission of $425 million paid to those affected by the breach. [4] If the data had resided within a ZT enabled environment, the breach could have been prevented, or at least lessened due to controls on data access and use. The ZT security model assumes that a breach is inevitable or has likely occurred already, so it constantly limits access to only what is needed and looks for anomalous or malicious activity. [1]

"Zero Trust is a security model, a set of system design principles, and a coordinated cybersecurity and system management strategy based on an acknowledgement that threats exist both inside and outside traditional network boundaries. The Zero Trust security model eliminates implicit trust in any one element, node, or service and instead requires continuous verification of the operational picture via real-time information fed from multiple sources to determine access and other system responses." [1] This guidance focuses on the data pillar, which specifically addresses data cataloging, governance, attributes and tags, monitoring, encryption, loss prevention, and access control.

The information presented in this report is not a definitive guide with a standardized solution that fits all organizational needs, but rather provides suggestions and considerations for adopting ZT. Discovering and identifying the assets that need to be secured to support the organization's mission will help build a picture of the current architecture for applying the recommendations in these seven ZT pillar reports. This picture of the current architecture will help all stakeholders identify organizational risks and gaps and ultimately inform building a mature ZT architecture for the organization. The ultimate goal is to integrate these principles into a comprehensive ZT strategy aligned with the organization’s security objectives.

Adopting ZT principles is not accomplished overnight. Implementing them is achieved through careful and deliberate planning and continuous incremental improvements that bring cybersecurity protections, responses, and operations to maturity over time. Building capabilities aligned to a mature ZT framework requires integrating every system in the enterprise with the appropriate security controls, best practices, configuration management, and vulnerability management for each of the seven pillars: User, Device, Network & Environment, Data, Application & Workload, Visibility & Analytics, and Automation & Orchestration. Each pillar constitutes a key focus area of ZT implementation, with the data pillar effectively secured by the other six. [5]

Audience

This report provides guidance primarily intended for National Security System (NSS), Department of Defense (DoD), and Defense Industrial Base (DIB) networks, but may be useful for owners and operators of other systems that might be targeted by sophisticated malicious actors. Guidance for other system owners and operators is also available via the National Institute of Standards and Technology (NIST) [6] and the Cybersecurity and Infrastructure Security Agency (CISA). [7] This guidance is compatible with the DoD ZT guidance referenced at the end of this document. [5]

Background

The President’s Executive Order on Improving the Nation’s Cybersecurity (EO 14028) [8] and National Security Memorandum 8 (NSM-8) [9] direct the Federal Civilian Executive Branch (FCEB) agencies and NSS owners and operators to develop plans to adopt a ZT cybersecurity framework.

In the NSA report, Embracing a Zero Trust Security Model, the concept of ZT is defined and contextualized along with the undergirding principles of the seven pillars [1] as illustrated in the following figure. The pillars are made up of several capabilities that earmark the progressive maturity of a comprehensive ZT framework. The capabilities described in this report are intended to continually mature cybersecurity protections, responses, and operations over time. Progression of capabilities in each pillar should be seen as a cycle of continuous improvement based on evaluation and monitoring of threats.

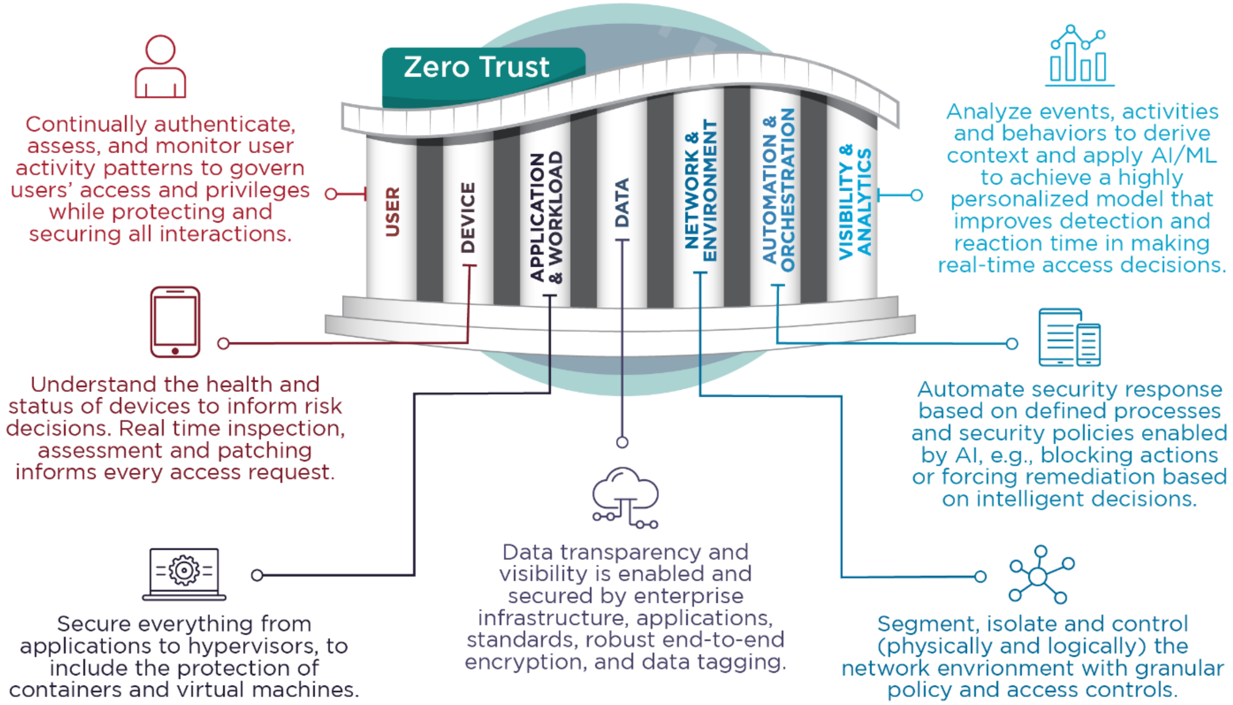
Zero Trust
USER
DEVICE
APPLICATION & WORKLOAD
DATA
NETWORK & ENVIRONMENT
AUTOMATION & ORCHESTRATION
VISIBILITY & ANALYTICS
Continually authenticate, assess, and monitor user activity patterns to govern users' access and privileges while protecting and securing all interactions.
Analyze events, activities and behaviors to derive context and apply AI/ML to achieve a highly personalized model that improves detection and reaction time in making real-time access decisions.
Understand the health and status of devices to inform risk decisions. Real time inspection, assessment and patching informs every access request.
Automate security response based on defined processes and security policies enabled by AI, e.g., blocking actions or forcing remediation based on intelligent decisions.
Data transparency and visibility is enabled and secured by enterprise infrastructure, applications, standards, robust end-to-end encryption, and data tagging.
Secure everything from applications to hypervisors, to include the protection of containers and virtual machines.
Segment, isolate and control (physically and logically) the network envrionment with granular policy and access controls.

Figure 1: Description of the seven pillars of ZT

Figure 1 depicts the ZT pillars, including the data pillar. The capabilities and milestones for the data pillar component of the ZT maturity model are described in detail throughout this report. The pillars are not independent; many capabilities in the data pillar depend on or align with capabilities in other pillars as indicated.

Data pillar

An organization’s data is extremely important and valuable. It is data, in its many forms, that is targeted by malicious entities. Customer records, user credentials, proprietary information, employee personally identifiable information (PII), intellectual property, personal emails, etc. are all fundamental to an organization. The ZT architecture is designed as a data-centric security model that draws on each connected pillar to ensure the confidentiality, integrity, and availability of an organization’s data, whether it exists within or outside of the network.

The data pillar focuses on securing and enforcing access to data at rest and in transit through various methods, including encryption, tagging and labeling, data loss prevention (DLP) strategies, and application of data rights management (DRM) tools. Additionally, securing data so it is accessed exclusively by authorized users is a primary responsibility of the data pillar and should not be taken for granted. The data pillar derives security benefits from capabilities performed by the other six pillars. Those capabilities are mapped to the DoD Chief Information Office (CIO) ZT Strategy, and NIST SP 800-207: Zero Trust Architecture. [10], [6]

This report identifies the following capabilities and aligns them to ZT maturity levels:

- Data catalog risk alignment
- Enterprise data governance
- Data labeling and tagging
- Data monitoring and sensing
- Data encryption and rights management
- Data loss prevention
- Data access control

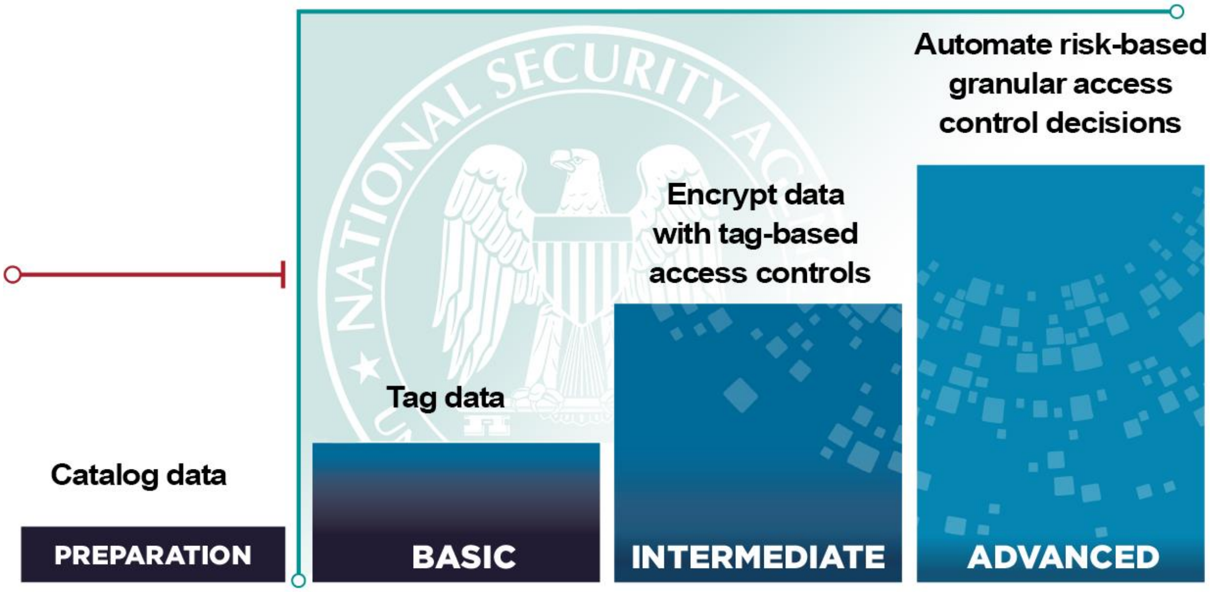
Automate risk-based granular access control decisions
Encrypt data with tag-based access controls
Tag data
Catalog data
PREPARATION
BASIC
INTERMEDIATE
ADVANCED

Figure 2: ZT data pillar maturity

Data catalog risk alignment

The first step in controlling data against threats is to identify all types of data in the environment and assess their risks of exposure, loss of availability, and loss of integrity. An enterprise data catalog should be a comprehensive inventory of data within the enterprise available for reference. This catalog, while not containing the data itself, includes metadata about the data, governance policies, and data usage. [5]

Data owners within an organization are aware of the details and purpose of their data. They must ensure their data is identified, inventoried, and categorized in the data catalog. This enterprise view of the data helps to facilitate data governance activities. When data owners review the catalog, they can identify potential risks or risk levels related to data loss, breach, or any other unauthorized alteration and/or access to data.

Table 1: Data catalog risk alignment maturity

Preparation	Basic	Intermediate	Advanced
Data landscape is reviewed to identify potential risks related to data loss, breach, or any other unauthorized alteration and/or access. Data ownership is identified, and data is catalogued based on resource criticality.	Critical organization data is manually identified and inventoried. Current state is recorded and data baseline set.	Automated processes are established to identify and monitor the data landscape within the catalog. Processes are enabled to ensure data is automatically detected and included within the catalog. Data usage patterns are established.	Data is known and can be collected, tagged, and protected according to risk levels in alignment with a prioritization framework, and encrypted for protection. Data is continuously analyzed to evaluate risk. Tooling is employed to discover improperly tagged sensitive data and alert/quarantine the data.

Enterprise data governance

Enterprise data governance ensures that data is controlled, accessed, and shared across organizations according to defined policies based on inputs from their cybersecurity infrastructure. Enterprise data labeling and tagging, access control and sharing policies, along with Data as a Service (DaaS) capabilities where applicable, ensure enforceability at the data object level. [11]

Table 2: Enterprise data governance maturity

Preparation	Basic	Intermediate	Advanced
Organization develops enterprise data labeling/tagging and access control/sharing policies that are enforceable. Data tagging and interoperability standards are defined.	Data is tagged and labeled in compliance with applicable enterprise policies. Data is encrypted with published enterprise frameworks according to enterprise policies.	Data protection policies are assessed and refined for interoperability across networks and partner organizations. Organization establishes just-in-time and just-enough data access control policies.	Rules and access controls are automated through central policy management. Policies are reviewed on a periodic basis and solutions regularly updated to remain in compliance.

Data labeling and tagging

Establishing granular data attributes integrated into access control systems (e.g. data tagging) consistently and correctly is required for machine enforceable data access controls, risk assessment, and situational awareness. As data attribute tagging and labeling practices mature, labeling should become automated to meet scaling demands and provide better labeling accuracy. Organizations should apply granular attributes to security and mission critical data on high value assets first.

Organizations should tag data in accordance with enterprise policies. Phases of implementation should advance toward full automation to enable accurate tagging at scale. Once data is properly labeled and tagged, the organization should establish automated data access controls, risk assessments, and monitoring for situational awareness based on enterprise governance policies. [12]

Table 3: Data labeling and tagging maturity

Preparation	Basic	Intermediate	Advanced
Data tagging standards are defined and tools configured to support enterprise policies.	Data tagging and classification tools are implemented. Data owners manually label and tag data in compliance with enterprise governance on labeling/tagging policy.	Machine enforceable data access controls are implemented. Automated tooling is created and implemented to meet scaling demands and provide better accuracy.	Data tagging and support is fully automated. Continuous analysis is employed to ensure data is properly tagged and labeled, and automation procedures remedied as needed.

Data monitoring and sensing

Data should always be detectable and observable by those who should have access to it and those who are required to manage it. Data metadata should be observable for tracking and alerting, although sometimes only partially since metadata can have sensitivities and access controls. Data owners and automated management solutions should ensure all data has associated metadata that includes current information about the access, sharing, transformation, and use of the data assets. This ensures basic integration with monitoring systems, and data owners with authorized access will make decisions about potential corruption or compromise. Organizations must have enforcement points in place to enable logging and policy enforcement.

Security Information and Event Management (SIEM) tools, which will be discussed more in the Visibility & Analytics pillar, play a role in this capability, providing data owners with the ability to gather and analyze security data from information systems using a single interface.

Table 4: Data monitoring and sensing maturity

Preparation	Basic	Intermediate	Advanced
Data owners identify and capture active	Database monitoring solutions are procured and	File monitoring tools are used to monitor all regulatory protected	Logs and analytics from all the data monitoring

Preparation	Basic	Intermediate	Advanced
metadata that provides insight into access, sharing, transformation, and use of data assets. Analysis is conducted to determine where tooling should be deployed for logging and enforcement points.	implemented across all databases containing regulated data types (CUI, PII, PHI, etc.) Data file monitoring tools are utilized to monitor critical data in applications, services, and repositories. Analytics from monitoring is fed into the SIEM with basic data attributes.	data in applications, services, and repositories. Extended integration is used to send data to appropriate inter/intra-pillar solutions, such as DLP, DRM, and User & Entity Behavior Analytics. Data outside of DLP and DRM scope, such as file shares and databases, are actively monitored for anomalous and malicious activity using alternative tooling.	solutions are fed into the SIEM for monitoring and response. Analytics are fed into cross pillar activities to better inform decision making. Additional data attributes to meet ZT advanced functionalities are integrated into analytics.

Data encryption and rights management

Data encryption and rights management combines technology with policy to protect data against unwanted access, modification, or redistribution. Data should be automatically encrypted based on data attributes assigned through tagging and labeling. By encrypting the data, organizations can be more assured their data is protected even if it is exfiltrated or lost as long as a malicious actor does not have the associated decryption keys. For additional security or if encryption is impossible, other data controls can be applied to protect data; this includes using DRM tools that prevent a user from forwarding, editing, saving, or printing data.

Table 5: Data encryption and rights management maturity

Preparation	Basic	Intermediate	Advanced
Organizations establish a strategy for encrypting data at rest and in transit following enterprise standards and requirements.	Encryption gaps are identified; enterprise-managed devices and centralized key management are employed. Organizations procure encryption tools as needed to implement the data at rest and in transit encryption strategy. Initial DRM implementations are used to reduce data exposure outside of enterprise-managed systems, focusing on protecting critical data in high-risk data repositories.	Encryption keys are automatically managed. All data is encrypted across the entire enterprise environment. DRM is expanded to all scoped data repositories.	Data tags are integrated with DRM; data is automatically encrypted at rest based on data tags. Additional tags are created to protect extended data repositories with DRM solutions designed to track and protect data. Machine learning models are used to detect and alert on anomalous usage of data. These models are integrated with encryption and DRM tools.

Data loss prevention

Data loss prevention (DLP) is a security strategy focused on detecting and preventing data leakage or loss through unauthorized use, exfiltration, or destruction. DLP tools deployed only at a system boundary are inadequate to address corruption of data throughout the system. Therefore, DLP tools are placed at identified enforcement points throughout the architecture to detect and mitigate data breaches and exfiltration. Organizations must establish a baseline for data usage before enabling the prevention capabilities of DLP tools. When implemented correctly along with the other capabilities of the data pillar, DLP tools, established throughout an organization’s network and not just at the perimeter, can more reliably secure an organization’s data.

Insider threats can pose a great risk to an organization. Entities with access to sensitive data from within can leak, destroy, or steal that data, intentionally or unintentionally. As examples, a vexed former employee could steal data to sell to a competitor, or one might accidentally leak sensitive data by using it in an AI tool, such as a large language model (LLM). DLP can help stop the unauthorized forwarding, copying, or destroying of sensitive data by tracking sensitive information within the network.

Among other scenarios, external threats can target data for exfiltration (theft), or use ransomware to manipulate and destroy data to make it inaccessible to authorized users. DLP can help prevent malicious cyber actors from successfully obtaining or encrypting internal data. DLP is a proactive solution for protecting data, but there should still be a plan in place for data recovery should data loss occur in spite of these efforts due to hardware failure, ransomware, or other causes.

Table 6: Data loss prevention maturity

Preparation	Basic	Intermediate	Advanced
Organizations scope enforcement points to deploy DLP solutions. Techniques for identifying sensitive data are established, such as key terms, fingerprints, pattern matching, and file matching.	A DLP solution is deployed to the in-scope enforcement points. DLP solution is set to “monitor-only” and/or “learning” mode to limit impact.	DLP solution results are analyzed and policy is fine-tuned to manage risk to an acceptable level. The DLP solution is updated from monitor mode to prevention mode. Basic manual data tags are utilized for the DLP solution and a logging schema is integrated with manual tags.	The DLP solution is updated to integrate data tags based on parallel automation activities for data tagging. DLP data scope is extended, utilizing the automated data tags to identify sensitive data. Automated data monitoring identifies missing enforcement points for additional DLP deployment.

Data access control

Data access control seeks to limit access to and use of data-based properties and attributes associated with the data and a user/device tuple along with any other relevant information. This capability is dependent on the others and brings into focus the ultimate job of the data pillar to enforce granular access controls and utilize all available data attributes for access decisions. This ensures unauthorized entities or entities on unauthorized devices cannot access the data. It also ensures those users and devices with access to data will continue to have their attributes inspected through various policy decision and enforcements points within the architecture.

The data protection needs of organizations will differ, and organizations must decide how they will use Role Based Access Control (RBAC), Policy Based Access Control (PBAC), Attribute Based Access Control (ABAC), and other options to control access. Organizations should mature through the phases as follows:

Table 7: Data access control maturity

Preparation	Basic	Intermediate	Advanced
Organizational policy is developed with enterprise-wide central management solutions in mind. Ensure appropriate access to, and use of, data based on the data and user/NPE/device properties. A software defined storage (SDS) policy and an enterprise Identity Provider (IdP)	Central management solutions, such as SDS and automation tools, are integrated with established policy and DRM tooling in a phased approach to measure results, improve protections, and adjust accordingly. Policy Based Access Controls (PBAC) are established. PBACs inform data access decisions using attributes determined by policy rules.	Attribute Based Access Controls (ABAC) are defined and established, ensuring identity attributes correspond to appropriate data objects. Roles are defined and implemented ensuring access to data dependent on proper user roles within the organization.	Individual and policy based access controls are established and automated central management solutions are fully integrated to manage changes from the central controller. ABAC, RBAC, and PBAC controls are further refined to provide more

Preparation	Basic	Intermediate	Advanced
integration plan are developed.			granular access regulations.

Data pillar guidance at a glance

- Develop enterprise data classification and labeling/tagging standards.
- Ensure all data is properly tagged and encrypted.
- Ensure that data tags are integrated with encryption policies.
- Ensure that all sensitive data is protected using proper encryption tools, such as DRM for data that moves beyond enterprise systems.
- Develop a DLP framework that counters internal and external threats to data security.
- Enforce data access controls based on enterprise policies and all information available about the access request.
- Monitor data for unauthorized movement, access, or alteration of data.

Conclusion

The need to protect data, a critical asset of any organization, is the driving force behind ZT. Data is protected through effective cataloging, labeling, and encryption while at rest and in transit. ZT strategy is ultimately centered on protecting an organization’s data through constant verification, so it is important that data owners take the steps necessary to survey their data to design and implement effective controls. Once in place, those controls should be tested and the maturity evaluated. Implementing an effective data management plan within the ZT framework will limit data breaches, and if a breach does occur, will provide the necessary information on the assets that were compromised to minimize the damage.

Further guidance

NSA is assisting DoD customers that are implementing ZT capabilities, coordinating ZT activities with NIST, CISA, NSS, and DoD, and developing additional ZT guidance to support system developers through the challenges of integrating ZT within NSS, DoD, and Defense Industrial Base (DIB) environments. Upcoming additional guidance will help organize, contextualize, and guide incorporation of ZT principles and designs into enterprise networks.

Supplementary NSA guidance on implementing a ZT architecture and ensuring a secure and defensible network environment are available at https://www.nsa.gov/cybersecurity-guidance:

- Embracing a Zero Trust Security Model
- NSA's Top Ten Cybersecurity Mitigation Strategies
- Defend Privileges and Accounts
- Continuously Hunt for Network Intrusions
- Segment Networks and Deploy Application-aware Defenses
- Transition to Multi-factor Authentication
- Actively Manage Systems and Configurations
- Performing Out-of-Band Network Management
- Hardening SIEM Solutions
- Mitigating Cloud Vulnerabilities
- Selecting Secure Multi-Factor Authentication Solutions

Partners at NIST, CISA, DoD, and others have produced guidance related to ZT architecture and capabilities, including:

- NIST SP 800-53 rev 5: Assessing Security and Privacy Controls in Information Systems and Organizations
- NIST SP 800-63-3: Digital Identity Guidelines (overview and parts a, b, c)
- NIST IR 8149: Developing Trust Frameworks to Support Identity Federations
- Federal ICAM Architecture
- NIST SP 800-207: Zero Trust Architecture
- CISA Zero Trust Maturity Model
- DoD Zero Trust Reference Architecture (Version 2.0)

Works cited

[1] National Security Agency. Embracing a Zero Trust Security Model. 2021. https://media.defense.gov/2021/Feb/25/2002588479/-1/-1/0/CSI_EMBRACING_ZT_SECURITY_MODEL_UOO115131-21.PDF

[2] Electronic Privacy Information Center. Equifax Data Breach. 2024. https://archive.epic.org/privacy/data-breach/equifax/

[3] Government Accountability Office. GAO Report to Congressional Requesters: Actions Taken by Equifax in Response to the 2017 Breach. 2018. https://www.warren.senate.gov/imo/media/doc/2018.09.06%20GAO%20Equifax%20report.pdf

[4] Federal Trade Commission. Equifax Data Breach Settlement. 2022. https://www.ftc.gov/enforcement/refunds/equifax-data-breach-settlement

[5] Department of Defense. Department of Defense (DoD) Zero Trust Reference Architecture Version 2.0. 2022. https://dodcio.defense.gov/Portals/0/Documents/Library/(U)ZT_RA_v2.0(U)_Sep22.pdf

[6] National Institute of Standards and Technology. Special Publication 800-207: Zero Trust Architecture. 2020. https://csrc.nist.gov/publications/detail/sp/800-207/final

[7] Cybersecurity and Infrastructure Security Agency. CISA Zero Trust Maturity Model. 2021. https://cisa.gov/zero-trust-maturity-model

[8] The White House. Executive Order 14028: Improving the Nation's Cybersecurity. 2021. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/

[9] The White House. National Security Memorandum 8: Improving the Cybersecurity of National Security, Department of Defense, and Intelligence Community Systems. 2022. https://www.whitehouse.gov/briefing-room/presidential-actions/2022/01/19/memorandum-on-improving-the-cybersecurity-of-national-security-department-of-defense-and-intelligence-community-systems/

[10] Department of Defense. DoD Zero Trust Strategy. 2022. https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-ZTStrategy.pdf

[11] M. Rouse. Data as a Service. 2017. https://www.techopedia.com/definition/28560/data-as-a-service-daas

[12] Department of Defense. DoD Zero Trust Capability Execution Roadmap (COA 1). 2022. https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-ZTExecutionRoadmap.pdf

Disclaimer of endorsement

The information and opinions contained in this document are provided "as is" and without any warranties or guarantees. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the United States Government, and this guidance shall not be used for advertising or product endorsement purposes.

Purpose

This document was developed in furtherance of NSA's cybersecurity missions, including its responsibilities to identify and disseminate threats to National Security Systems, Department of Defense, and Defense Industrial Base information systems, and to develop and issue cybersecurity specifications and mitigations. This information may be shared broadly to reach all appropriate stakeholders.

Contact

Cybersecurity Report Feedback: CybersecurityReports@nsa.gov
General Cybersecurity Inquiries or Customer Requests: Cybersecurity_Requests@nsa.gov
Defense Industrial Base Inquiries and Cybersecurity Services: DIB_Defense@cyber.nsa.gov
Media Inquiries / Press Desk: NSA Media Relations: 443-634-0721, MediaRelations@nsa.gov