Embracing Zero Trust: A strategy for secure and agile business transformation

AWS Prescriptive Guidance

AWS Prescriptive Guidance: Embracing Zero Trust: A strategy for secure and agile business transformation

Table of Contents

Embracing Zero Trust: A strategy for secure and agile business transformation

Greg Gooden, Amazon Web Services (AWS)

December 2023 (document history)

Today, more than ever, organizations are focusing on security as a key priority. This enables a wide range of benefits, from maintaining the trust of their customers, to improving the mobility of their workforce, to unlocking new digital business opportunities. As they do so, they continue to ask an age-old question: What are the optimal patterns to ensure the right levels of security and availability for my systems and data? Increasingly, Zero Trust has become the term used to describe the modern answer to this question.

Zero trust architecture (ZTA) is a conceptual model and an associated set of mechanisms that focus on providing security controls around digital assets that do not solely or fundamentally depend on traditional network controls or network perimeters. Instead, network controls are augmented with identity, device, behavior, and other rich context and signals to make more granular, intelligent, adaptive, and continuous access decisions. By implementing a ZTA model, you can achieve a meaningful next iteration in the continuous maturation of cybersecurity and concepts of defense in depth in particular.

Decision-making processes

Implementing a ZTA strategy requires careful planning and decision-making. It involves evaluating various factors and aligning them with organizational goals. Key decision-making processes for embarking on a ZTA journey include:

1. Stakeholder engagement – It's crucial to engage other CxOs, VPs, and senior managers to understand their priorities, concerns, and vision for your organization's security posture. By involving key stakeholders from the outset, you can align the ZTA implementation with the overall strategic objectives and gains the necessary support and resources.
2. Risk assessment – Conducting a comprehensive risk assessment helps identify issues, excessive surface area, and critical assets, which helps you make informed decisions on security controls and investment. Evaluate your organization's existing security posture, identify potential

weaknesses, and prioritize areas of improvement based on the risk landscape that is specific to your industry and operational environment.

3. Technology evaluation – Assessing the organization's existing technology landscape and identifying gaps helps in selecting appropriate tools and solutions that align with the ZTA principles. This evaluation should include a thorough analysis of the following:
   - Network architecture
   - Identity and access management systems
   - Authentication and authorization mechanisms
   - Unified endpoint management
   - Resource ownership tools and processes
   - Encryption technologies
   - Monitoring and logging capabilities
   - Choosing the right technology stack is crucial for building a robust ZTA model.
4. Change management – Recognizing the cultural and organizational impacts of adopting a ZTA model is essential. Implementing change management practices helps ensure smooth transition and acceptance throughout the organization. It involves educating employees about the principles and benefits of ZTA, providing training on new security practices, and fostering a security-conscious culture that encourages accountability and continuous learning.

This prescriptive guidance aims to provide CxOs, VPs, and senior managers with a comprehensive strategy for implementing ZTA. It will delve into the key aspects of ZTA, including the following:

- Organizational readiness
- Phased adoption approaches
- Stakeholder collaboration
- Best practices to achieve a secure and agile business transformation

By following this guidance, your organization can navigate the ZTA landscape and achieve successful outcomes in your security journey in the Amazon Web Services (AWS) Cloud. AWS offers a variety of services that you can use to implement a ZTA, such as AWS Verified Access, AWS Identity and Access Management (IAM), Amazon Virtual Private Cloud (Amazon VPC), Amazon VPC Lattice, Amazon Verified Permissions, Amazon API Gateway, and Amazon GuardDuty. These services can help to protect AWS resources from unauthorized access.

Targeted business outcomes

This section discusses the expected outcomes associated with defining and implementing a zero trust architecture across your organization.

Improved security posture

By adopting Zero Trust principles, your organization can strengthen its security posture, mitigate security risks, and protect your cloud infrastructure and data. The Zero Trust fundamental principle of granting access on a need-to-know basis, coupled with stringent controls, significantly reduces the surface area, and it limits the potential impact of security events. This proactive approach helps organizations stay ahead of emerging security risks and helps ensure the confidentiality, integrity, and availability of assets.

Seamless cloud adoption

Developing a well-defined zero trust architecture (ZTA) adoption plan can help ensure a smooth and successful transition to the cloud environment. ZTA principles align closely with cloud security best practices by providing a strong foundation for organizations to securely gain the benefits of cloud computing. Incorporating ZTA principles from the beginning helps your organization to design its cloud architecture with security as a core element.

Compliance and regulatory alignment

Implementing ZTA practices can help your organization to meet industry and regulatory requirements and standards. ZTA inherently promotes the principle of least privilege and enforces strict access controls. Access controls are often mandated by regulations such as the following:

- Federal Risk and Authorization Management Program (FedRAMP)
- Health Insurance Portability and Accountability Act (HIPAA)
- Payment Card Industry Data Security Standard (PCI DSS).

By adopting Zero Trust, your organization can help demonstrate its commitment to data protection, privacy, and regulatory compliance while minimizing the potential for penalties or reputational damage.

Enhanced data protection

Organizations can protect sensitive data throughout the cloud adoption process by implementing data encryption, access controls, and regular security assessments. Your organization can take the following specific steps:

- Data encryption – Data encryption – Data encryption is the process of encrypting cleartext data into ciphertext in a way that requires a key to decrypt the data back into the original cleartext form. This makes it much more difficult for unauthorized individuals to access sensitive data, even if they are able to obtain a copy of the data.
- Access controls – Access controls restrict who can access sensitive data and what they can do with it. This can be done by assigning user roles and permissions, and by using multi-factor authentication or other methods to verify user identity.
- Regular security assessments – Regular security assessments can help organizations identify and address security issues and proactively remediate them. These assessments can be conducted by internal security teams or by external security firms.

Zero trust architectures take a comprehensive approach to data protection by implementing a number of security measures. These measure include strong authentication, data encryption, and granular access controls. This approach minimizes the risk of data-related security events, and it safeguards sensitive information from unauthorized access.

Efficient incident response

Organizations can detect and respond to security events more quickly and effectively by establishing monitoring and incident response frameworks in the cloud environment. Zero trust architectures emphasize continuous monitoring, threat intelligence integration, and real-time visibility into user activities, network traffic, and system behavior. Security teams can then proactively identify and mitigate security events. This approach reduces the time to detect and respond to potential issues, and it minimizes the impact on business operations. Key points include the following:

- Testing – Regardless of the incident response framework or methodology your organization aligns with, you should test your incident response plan regularly. Tabletop exercises, simulations, and red teaming provide opportunities to practice incident response in realistic

settings, uncover tooling and capability gaps, and build the experience and confidence of incident responders.

- Monitoring – Continuously monitor your cloud environments for signs of abnormal activity. You can do this by using a variety of tools and techniques, such as log analysis, network monitoring, and vulnerability scanning.
- Threat intelligence integration – Integrate threat intelligence into your monitoring and incident response frameworks. This will help your organization to identify and respond to threats more quickly and effectively.
- Real-time visibility – To identify and respond to security incidents quickly, your organization needs real-time visibility into user activities, network traffic, and system behavior.
- Proactive identification and mitigation – By proactively identifying and mitigating security events, your organization can reduce the time to detect and respond to potential threats, minimizing the impact on business operations.

Improved workforce productivity

The modern workforce requires flexibility to get work done from an increasing array of locations, devices, and times. By implementing a ZTA, you can support these requirements and improve workforce mobility, productivity, and satisfaction, while maintaining or improving the organization's security posture.

Enable digital transformation

Organizations are increasingly pursuing the interconnection of devices, machines, facilities, infrastructure, and processes outside the traditional network perimeter as part of digital transformation. Internet of things (IoT) and operational technology (OT, also known as Industrial Internet of Things, or IIoT) devices often transmit telemetry and predictive maintenance information directly to the cloud. To protect workloads, this requires the application of security controls that extend beyond the traditional perimeter approach.

Section summary

By focusing on these targeted business outcomes, your organization can realize the full potential of ZTA and strengthen your security posture in the cloud. It's important to align these outcomes

with specific organizational goals, tailor them to your unique business requirements, and regularly assess their effectiveness to drive continuous improvement.

Understanding Zero Trust principles

Zero trust architecture (ZTA) is based on a set of core principles that form the foundation of its security model. Understanding these principles is essential for organizations looking to adopt a ZTA strategy effectively. This section covers the core principles of ZTA.

Verify and authenticate

The verify and authenticate principle emphasizes the importance of strong identification and authentication for principals of all types, including users, machines, and devices. ZTA requires continuous verification of identities and authentication status throughout a session, ideally on each request. It doesn't rely solely on traditional network location or controls. This includes implementing modern strong multi-factor authentication (MFA) and evaluating additional environmental and contextual signals during authentication processes. By adopting this principle, organizations can help ensure that resource authorization decisions have the best possible identity inputs.

Least privilege access

The principle of least privilege involves granting principals the minimum level of access required to perform their tasks. By adopting the principle of least privilege access, organizations can enforce granular access controls, so that principals have access only to the resources necessary to fulfill their roles and responsibilities. This includes implementing just-in-time access provisioning, role-based access controls (RBAC), and regular access reviews to minimize the surface area and the risk of unauthorized access.

Micro-segmentation

Micro-segmentation is a network security strategy that divides a network into smaller, isolated segments for authorizing specific traffic flows. You can achieve micro-segmentation by creating workload boundaries and enforcing strict access controls between different segments.

Micro-segmentation can be implemented through network virtualization, software-defined networking (SDN), host-based firewalls, network access control lists (NACLs), and AWS specific features such as Amazon Elastic Compute Cloud (Amazon EC2) security groups or AWS PrivateLink. Segmentation gateways control traffic between segments to explicitly authorize access. Micro-

segmentation and segmentation gateways help organizations restrict unnecessary pathways through the network, particularly those that lead to critical systems and data.

Continuous monitoring and analytics

Continuous monitoring and analytics involve the collection, analysis, and correlation of security-related events and data across your organization's environment. By implementing robust monitoring and analytics tools, your organization can evaluate security data and telemetry in a converged way.

This principle emphasizes the importance of visibility into user behavior, network traffic, and system activities to identify anomalies and potential security events. Advanced technologies such as security information and event management (SIEM), user and entity behavior analytics (UEBA), and threat intelligence platforms play a vital role in achieving continuous monitoring and proactive threat detection.

Automation and orchestration

Automation and orchestration help organizations to streamline security processes, reduce manual intervention, and enhance response times. By automating routine security tasks and using orchestration capabilities, your organization can enforce consistent security policies and rapidly respond to security events. This principle also includes automating access provisioning and deprovisioning processes to help ensure timely and accurate management of user permissions. By embracing automation and orchestration, your organization can improve operational efficiency, reduce human errors, and focus resources on more strategic security initiatives.

Authorization

In a ZTA, each request to access a resource should be explicitly authorized by a gating enforcement point. In addition to the authenticated identity, authorization policies should consider additional context, such as device health and posture, behavior patterns, resource classification, and network factors. The authorization process should evaluate this converged context against the corresponding access policies that are relevant to the resource being accessed. Optimally, machine learning models can provide a dynamic supplement to the declarative policies. When utilized, these models should focus on additional restrictions only, and they should not grant access that wasn't explicitly specified.

Section summary

By adhering to these core principles of ZTA, organizations can establish a robust security model that aligns with the diversity of the modern enterprise environment. Implementing these principles requires a comprehensive approach that combines technology, processes, and people to achieve a zero trust mindset and build a resilient security posture.

Key components of a zero trust architecture

To implement a zero trust architecture (ZTA) strategy effectively, your organization must understand the key components that make up a ZTA. These components work together to continuously improve upon a comprehensive security model that aligns with Zero Trust principles. This section covers those key components of a ZTA.

Identity and access management

Identity and access management forms the foundation of a ZTA by providing robust user authentication and coarse-grain access-control mechanisms. It includes technologies such as single sign-on (SSO), multi-factor authentication (MFA), and identity governance and management solutions. Identity and access management provides a high level of authentication assurance and important context that are integral to making zero trust authorization decisions. At the same time, ZTA is a security model in which access to applications and resources is granted on a per-user, per-device, and per-session basis. This helps to protect organizations from unauthorized access, even if a user's credentials are compromised.

Secure access service edge

Secure access service edge (SASE) is a new approach to network security that virtualizes, combines, and distributes networking and security functions into a single, cloud-based service. SASE can provide secure access to applications and resources, regardless of the user's location.

SASE includes a variety of security features, such as secure web gateways, firewall as a service, and zero trust network access (ZTNA). These features work together to protect organizations from a wide range of threats, including malware, phishing, and ransomware.

Data loss prevention

Data loss prevention (DLP) technologies can help organizations protect sensitive data from unauthorized disclosure. DLP solutions monitor and control data in motion and at rest. This helps organizations to define and enforce policies that prevent data-related security events, helping to ensure that sensitive information remains protected throughout the network.

Security information and event management

Security information and event management (SIEM) solutions collect, aggregate, and analyze security event logs from various sources across an organization's infrastructure. You can use this data to detect security incidents, facilitate incident response, and provide insights into potential threats and vulnerabilities.

For ZTA specifically, a SIEM solution's ability to correlate and understand related telemetry from different security systems is critical to improved detection of and response to abnormal patterns.

Enterprise resource ownership catalog

To properly grant access to enterprise resources, an organization must have a reliable system that catalogs these resources and, importantly, who owns them. This source of truth needs to provide workflows that facilitate access requests, the associated approval decisions, and regular attestations thereof. In time, this source of truth will contain the answers to "who can access what?" within the organization. You can use the answers for both authorization and audit and compliance.

Unified endpoint management

In addition to strongly authenticating the user, a ZTA must also consider the health, posture, and state of the user's device to assess whether corporate data and resource access is secure. A unified endpoint management (UEM) platform provides the following capabilities:

- Device provisioning
- Ongoing configuration and patch management
- Security baselining
- Telemetry reporting
- Device cleansing and retirement

Policy-based enforcement points

In a ZTA, access to each resource should be explicitly authorized by a gating policy-based enforcement point. Initially, these enforcement points can be based on existing enforcement points in existing network and identity systems. The enforcement points can be made incrementally more

capable by considering the wider array of context and signals that ZTA provides. Longer term, your organization should implement ZTA-specific enforcement points that operate on converged context, consistently integrate signal providers, maintain a comprehensive policy set, and are enhanced with intelligence gleaned from combined telemetry.

Section summary

Understanding these key components is essential for organizations planning to adopt a ZTA. By implementing these components and integrating them into a cohesive security model, your organization can establish a strong security posture based on the principles of Zero Trust. The following sections explore organizational readiness, phased adoption approaches, and best practices to help you successfully implement ZTA within your organization.

Assessing organizational readiness for Zero Trust adoption

Adopting a new architecture strategy is a significant undertaking that requires careful planning and consideration of organizational factors. This section focuses on key organizational readiness considerations for Zero Trust adoption across the enterprise. By addressing these considerations, your organization can pave the way for a stronger and more successful security posture.

Leadership alignment and communication

Leadership alignment and communication are essential for the successful implementation of Zero Trust. Leadership must understand the benefits of Zero Trust and the resources required. Leaders must also be willing to make changes to the organization's culture and processes. Communication with employees is necessary for building trust and buy-in. Employees need to understand why the organization is implementing Zero Trust, what it means for them, and how they can help. Communication should be open, transparent, and ongoing.

Leadership support and buy-in

For a successful zero trust architecture (ZTA) implementation, it's crucial that you align key stakeholders and executives on the architecture's goals, benefits, and measures of success. Share the importance of the Zero Trust principles in enhancing security and enabling business agility by moving away from traditional perimeter-based security to a more granular, user-centric approach. By switching to this approach, your organization can adapt to changes and threats more quickly. Executive alignment establishes the tone for the organization and helps overcome potential resistance to change.

Transparent communication

Maintain open and transparent communication with employees throughout the Zero Trust implementation process. Explain the rationale, benefits, and expected outcomes of the adoption, and address concerns promptly. Provide regular updates on the progress of the implementation. This will increase buy-in, reduce resistance, and build trust.

Skill development and training

After leadership is aligned and communication is open, it's important to develop the skills and knowledge of the employees who will implement Zero Trust. This includes understanding the Zero Trust principles, how to implement them in their work, and how to respond to security events. Provide training and development opportunities to help employees acquire these skills.

Cloud knowledge and skills

Assess the organization's skills and knowledge gaps in cloud technologies and Zero Trust principles. Provide training and development programs to upskill employees and equip them with the necessary expertise to work effectively in a cloud-centric and Zero Trust environment. To keep pace with evolving technologies and security practices, foster a culture of continuous learning.

Security culture and awareness

Assess the organization's security culture. Evaluate the level of security awareness among employees, their understanding of security best practices, and their adherence to policies and procedures. Identify any gaps in security knowledge. Consider conducting security-awareness training programs to educate employees about the importance of Zero Trust and their roles in maintaining a secure environment.

Organizational structure and roles

To successfully implement Zero Trust, establish an effective organizational structure and roles. This includes creating a Cloud Center of Excellence (CCoE), reviewing and modifying security operations, and assigning roles and responsibilities for vulnerability management, incident response, and security monitoring.

Cloud Center of Excellence

Establish a CCoE to provide guidance, best practices, and oversight for cloud operations. A CCoE is a team or group of individuals responsible for creating and implementing cloud-related best practices, guidelines, and governance policies. The CCoE should include representatives from different business units and IT teams to help ensure collaboration and alignment. The CCoE plays a crucial role in driving the adoption of Zero Trust principles into cloud-hosted workloads. The CCoE also facilitates knowledge sharing across the organization.

Security operations

To meet the needs of a Zero Trust environment, review and modify the current security operations organization. To improve monitoring, incident response, and threat intelligence capabilities, consider implementing security operations centers (SOCs) or managed security service providers (MSSPs). Establish roles and responsibilities for vulnerability management, incident response, and security monitoring. A well-functioning incident response process is critical to ensuring that minor security events can be detected and remediated quickly to disrupt the sequence of events. This helps to prevent a minor event from evolving into a more impactful one.

IT infrastructure and architecture

Examine the IT architecture and infrastructure of your company to find any constraints or dependencies that might affect the adoption of a Zero Trust approach. Determine whether current applications and systems are compatible with the necessary zero trust architectural components. Analyze whether any infrastructure improvements or adjustments are required to support the successful deployment of Zero Trust principles. For each application or system, consider whether Zero Trust is best implemented in place or through a larger modernization effort.

Risk management, governance, and change control

To successfully implement Zero Trust, Establish effective risk management, governance, and change control processes. This includes aligning risk management with Zero Trust principles, developing an incident response plan, working with legal and compliance departments, and establishing a change control process.

Risk management

Examine the risk management strategy in place at your company and determine how well it adheres to the Zero Trust principles. Analyze the efficiency of the present incident response systems, security measures, and risk assessment procedures. Determine which areas need to be improved to conform to the Zero Trust strategy. Begin developing an automated incident response system or a continuous monitoring and analytics framework to increase speed to resolution.

Change control processes

To help ensure that all cloud-related modifications abide by security and compliance requirements, establish effective change control methods. Establish a systematic change management procedure that includes security configuration analysis, risk evaluations, approvals, and documentation. Review and audit updates frequently to preserve the integrity of the zero trust architecture.

Monitoring and evaluation

To successfully implement Zero Trust, your organization must continuously monitor and evaluate its security posture. This includes establishing key performance indicators (KPIs), monitoring and evaluating the KPIs, and fostering a culture of continuous improvement. By following these steps, organizations can ensure that their Zero Trust implementation is successful and that they are always working to improve their security.

Key performance indicators

Establish pertinent key performance indicators (KPIs) to gauge the success and efficacy of the Zero Trust deployment. These KPIs might measure user satisfaction, equipment and rollout progress, cost reduction, compliance observance, and the number of security occurrences. To track the overall development and find opportunities for improvement, regularly monitor and evaluate these KPIs.

Continuous improvement

Establishing systems to elicit opinions and insights from stakeholders will help to foster a culture of continuous improvement. Encourage staff members to offer thoughts and proposals for improving the cloud environment's security, effectiveness, and user experience. Use this input to streamline procedures, improve security measures, and spur innovation.

Section summary

By addressing these organizational and cultural considerations, your organization can foster a supportive environment for the cloud adoption of a Zero Trust security model. The next section explores phased adoption approaches, providing guidance on how to gradually implement Zero Trust principles in a practical and manageable manner.

Cultivating a Zero Trust mindset

Implementing Zero Trust goes beyond technical implementations. It requires a cultural shift within your organization. Fostering a Zero Trust mindset involves emphasizing the following key aspects.

Zero Trust education and training

Educate employees about the values and advantages of zero trust architecture (ZTA). Provide technical and non-technical explanations of ZTA concepts and approaches through training sessions, workshops, and other resources. Encourage staff members to be aware of their responsibilities in establishing and upholding a Zero Trust security paradigm.

Collaboration and communication

Foster collaboration and transparency across all teams and departments involved in the ZTA implementation. To ensure everyone has a thorough understanding of the plan, promote cross-functional communication, knowledge sharing, and information exchange. Create a culture of shared responsibility where everyone recognizes the importance of their contributions to the overall security of the business.

Continuous learning and improvement

Prioritize continuous learning and improvement in the context of Zero Trust. Encourage employees to stay up to date on the latest security trends, technologies, and best practices. Nurture a culture of innovation and experimentation in which employees are encouraged to explore new solutions and approaches to strengthen the organization's security posture.

Metrics and accountability

Establish clear metrics and accountability mechanisms to measure the effectiveness of the Zero Trust strategy. Define key performance indicators (KPIs) that align with the organization's security goals, and regularly track progress. Hold individuals and teams accountable for their contributions to the implementation and maintenance of Zero Trust principles.

Section summary

By addressing these aspects and cultivating a Zero Trust mindset, organizations can create a solid foundation for successful adoption and implementation of Zero Trust. This cultural shift is essential for helping everyone in the organization to understand the importance of Zero Trust and actively contribute to its success.

The next section explores phased adoption approaches, providing guidance on how to gradually implement Zero Trust principles in a practical and manageable manner.

Phased approach to Zero Trust

Adoption of a zero trust architecture (ZTA) requires careful planning and implementation. We recommend a phased adoption approach to smooth transition and minimize disruption to business operations. This section provides guidance on the key phases involved in adopting a ZTA.

Phase 1: Assessment and planning

The first phase of Zero Trust implementation is assessment and planning. This phase is critical to the success of the overall implementation, because it involves identifying and addressing any gaps in your organization's current security posture. By taking the time to assess your current state and define your security objectives, you can lay the foundation for a successful Zero Trust implementation.

At the same time, a perfectly complete and accurate assessment might not be always realistic. To avoid analysis paralysis that prevents you from moving on to further phases, be prepared to compartmentalize or otherwise accept some level of imperfection.

1. **Assess the current state** – Conduct an assessment of your existing security infrastructure, policies, and controls. Identify potential vulnerabilities, gaps in security, and areas where the implementation of Zero Trust principles can provide improvements.
2. **Define security objectives** – Based on the current state assessment findings, define security objectives that align with the principles of Zero Trust. These security objectives should also align with your organization's overall security strategy and address identified vulnerabilities and gaps.
3. **Design the architecture** – Develop a ZTA that supports your organization's security goals. This architecture should include the necessary components, such as identity and access management solutions, network segmentation mechanisms, and continuous monitoring systems. The architecture should also be scalable, adaptable, and capable of accommodating future growth and technological advancements. Ideally, this architecture should be represented in a format that's easily consumed by the teams responsible for implementing it, such as an AWS CloudFormation template, not just as a document or diagram.
4. **Engage stakeholders** – Involve all stakeholders, including business units, IT teams, and security teams, to gain insights and align their objectives with the ZTA implementation plan. Encourage collaboration and communication to establish a shared understanding of the benefits and requirements of the Zero Trust approach.

Phase 2: Piloting and implementation

The second phase of Zero Trust implementation is piloting and implementation. This phase involves testing the ZTA in a small-scale, controlled environment, and then iteratively deploying it across your organization. It's important to educate employees on the new security measures and their roles in maintaining a Zero Trust environment.

1. **Pilot the deployment** – Test the ZTA in a small-scale, controlled environment. Implement the necessary components and security controls that were defined in the architecture design phase. Monitor the pilot deployment closely, gather feedback, and make any necessary adjustments. Be prepared to be flexible early in the process, when Zero Trust moves from being a hypothetical exercise to one that you're building real experience with.
2. **Deploy iteratively** – Based on the lessons learned from the pilot deployment, begin the iterative deployment of Zero Trust across the organization. Build momentum through a flywheel effect that doesn't require an extensive campaign to achieve critical deployment mass. Reserve leadership mandates or escalations for the longer tail of the rollout where they might be required.
3. **Provide user training and raise awareness** – Educate employees on the new security measures and their roles in maintaining a Zero Trust environment. Emphasize the importance of secure practices, such as strong passwords, multi-factor authentication, and regular security updates.
4. **Manage change** – Create a comprehensive change management plan to address the organizational and cultural changes associated with Zero Trust adoption. Communicate the benefits and rationale behind the adoption to employees, and address any concerns or resistance. Provide ongoing support and guidance to facilitate a smooth transition.

Phase 3: Monitoring and continuous improvement

The third and final phase of Zero Trust implementation is monitoring and continuous improvement. This phase involves establishing a comprehensive monitoring and analytics program, creating a comprehensive incident response plan, and regularly soliciting feedback from stakeholders and users.

1. **Monitor continuously** – Establish a comprehensive monitoring and analytics program to assess the security posture continuously and detect any potential anomalies. Use advanced security tools and technologies to monitor user behavior, network traffic, and system activities.

2. **Plan incident response and remediation** – Create a comprehensive incident response plan that aligns with the Zero Trust principles. Establish clear escalation paths, define roles and responsibilities, and implement automated incident response mechanisms where possible. Regularly test and update the incident response plan.
3. **Obtain feedback and evaluation** – Regularly solicit feedback from stakeholders and users to gather insights into the effectiveness of the zero trust architecture (ZTA). Conduct periodic evaluations and assessments to measure the impact on security posture, operational efficiency, and user experience. Use the feedback and evaluation results to identify areas for improvement. Expect that your ZTAs will change over time, and consider how development teams will implement these updates with minimal effort or disruptions.

Section summary

By following this phased adoption approach, organizations can effectively transition to a ZTA while minimizing risks and disruptions. The next section discusses best practices for achieving success with Zero Trust implementation, covering key considerations and recommendations for CxOs, VPs, and senior managers.

Best practices for achieving success with Zero Trust

Successful adoption of zero trust architecture (ZTA) requires a strategic approach and adherence to best practices. This section presents a set of best practices to guide CxOs, VPs, and senior managers in achieving success with their Zero Trust adoption. By following these recommendations, your organization can establish a strong security foundation and realize the benefits of a Zero Trust approach:

- **Define clear objectives and business outcomes** – Clearly define the objectives and desired business outcomes of the cloud operations. Align these objectives with the principles of Zero Trust to build a strong security foundation while enabling business growth and innovation.
- **Conduct a comprehensive assessment** – Perform a comprehensive evaluation of the current IT infrastructure, applications, and data assets. Identify dependencies, technical debt, and potential compatibility issues. This evaluation will inform the adoption plan and help prioritize workloads based on criticality, complexity, and business impact.
- **Develop an adoption plan** – Incorporate a detailed adoption plan that outlines the step-by-step approach for moving workloads, applications, and data to the cloud. Define adoption phases, timelines, and dependencies. Engage key stakeholders and allocate resources accordingly.
- **Start building early** – Your ability to authentically represent what Zero Trust will look like within your organization will substantially increase after you start building and deploying it (rather than analyzing and talking about it).
- **Obtain executive sponsorship** – Secure executive sponsorship and support for the Zero Trust implementation. Engage other C-level executives to champion the initiative and allocate the necessary resources. Leadership commitment is essential for driving the cultural and organizational changes required for a successful implementation.
- **Implement a governance framework** – Create a governance framework that defines roles, responsibilities, and decision-making processes for the Zero Trust implementation. Clearly define accountability and ownership of security controls, risk management, and compliance. Regularly review and update the governance framework to adapt to evolving security requirements.
- **Support cross-functional collaboration** – Encourage collaboration and communication between different business units, IT teams, and security teams. Create a culture of shared responsibility to foster alignment and coordination throughout the Zero Trust implementation. Encourage frequent interactions, knowledge sharing, and joint problem-solving.

- **Secure your data and applications** – Zero Trust isn't only about end-users accessing resources and applications. Zero Trust principles should also be implemented within and between workloads. Apply the same technical principles—strong identity, micro-segmentation, and authorization—by using all available context within the data center as well.
- **Provide defense in depth** – Implement a defense-in-depth strategy by using multiple layers of security controls. Combine various security technologies, such as multi-factor authentication (MFA), network segmentation, encryption, and anomaly detection, to provide comprehensive protection. Make sure that each layer complements the others to create a strong defense system.
- **Require strong authentication** – Enforce strong authentication mechanisms, such as MFA, for all users accessing all resources. Ideally, consider modern MFA, such as FIDO2 hardware-backed security keys, that provides a high level of authentication assurance for Zero Trust and carries broad security benefits (for example, protection against phishing).
- **Centralize and improve authorization** – Specifically authorize every access attempt. Depending on the protocol specifics, this should be done on a per-connection or per-request basis. Per-request is ideal. Use all available context, including identity, device, behavior, and network information to make more granular, adaptive, and sophisticated authorization decisions.
- **Use the principle of least privilege** – Implement the principle of least privilege to grant users the minimum access rights required to perform their job duties. Regularly review and update access permissions based on job roles, responsibilities, and business needs. Implement just-in-time access provisioning.
- **Use privileged access management** – Implement a privileged access management (PAM) solution to secure privileged accounts and reduce the risk of unauthorized access to critical systems. PAM solutions can provide privileged access controls, session recording, and auditing capabilities to help your organization protect its most sensitive data and systems.
- **Use micro-segmentation** – Divide your network into smaller, more isolated segments. Use micro-segmentation to enforce strict access controls between segments based on user roles, applications, or data sensitivity. Strive to eliminate all unnecessary network pathways, particularly those that lead to data.
- **Monitor and respond to security alerts** – Implement a comprehensive security monitoring and incident response program in the cloud environment. Use cloud-native security tools and services to detect threats in real time, analyze logs, and automate incident response. Establish clear incident response procedures, conduct regular security assessments, and continuously monitor for anomalies or suspicious activities.

- **Use continuous monitoring** – To detect and respond to security incidents quickly and effectively, implement continuous monitoring. Use advanced security analytics tools to monitor user behavior, network traffic, and system activities. Automate alerts and notifications to ensure that incidents are responded to in a timely manner.
- **Promote a culture of security and compliance** – Promote a culture of security and compliance throughout the organization. Educate employees on security best practices, the importance of adhering to Zero Trust principles, and employees' role in maintaining a secure cloud environment. Conduct regular security awareness training to help ensure that employees are vigilant against social engineering and that they understand their responsibilities regarding data protection and privacy.
- **Use social engineering simulations** – Conduct social engineering simulations to assess user susceptibility to social engineering attacks. Use the results of the simulations to tailor training programs for improved user awareness and response to potential threats.
- **Promote continuous education** – Establish a culture of continuous education and learning by providing ongoing security training and resources. Keep users informed about evolving security best practices. Encourage users to stay vigilant and report any suspicious activities promptly.
- **Continuously assess and optimize** – Regularly assess the cloud environment for areas of improvement. Use cloud-native tools to monitor resource usage and performance, and conduct vulnerability assessments and penetration testing to identify and address any weaknesses.
- **Establish a governance and compliance framework** – Develop a governance and compliance framework to help ensure that your organization is aligned with industry standards and regulatory requirements. In the framework, define policies, procedures, and controls to protect data and systems from unauthorized access, use, disclosure, disruption, modification, or destruction. Implement mechanisms for tracking and reporting on compliance metrics, conducting regular audits, and addressing any non-compliance issues promptly.
- **Encourage collaboration and knowledge sharing** – Encourage collaboration and knowledge sharing among teams involved in the ZTA adoption. You can do this by fostering cross-functional communication and collaboration between IT, security, and business units. Your organization can also establish forums, workshops, and knowledge-sharing sessions to promote understanding, address challenges, and share lessons learned throughout the adoption process.

Key takeaways

This guide has explored the essential aspects of developing a successful zero trust architecture (ZTA) strategy. This section summarizes the key takeaways from the prescriptive guidance presented:

- **Understand Zero Trust principles** – Zero Trust is a conceptual model and an associated set of mechanisms that focus on providing security controls around digital assets that do not solely or fundamentally depend on traditional network controls or network perimeters. Instead, network controls are augmented with identity, device, behavior, and other rich context and signals to make more granular, intelligent, adaptive, and continuous access decisions. Familiarize yourself with the core principles of Zero Trust, such as least privilege, micro-segmentation, continuous authentication, and adaptive authorization.
- **Define clear objectives** – Clearly define the objectives and desired business outcomes of the ZTA adoption. Align these objectives with the principles of Zero Trust to help ensure a strong security foundation while enabling business growth and innovation.
- **Conduct comprehensive assessments** – Perform a thorough assessment of your existing IT infrastructure, applications, and data assets. Identify dependencies, technical debt, and compatibility issues to inform your adoption strategy.
- **Develop a ZTA adoption plan** – Create a detailed plan that outlines the step-by-step approach for moving workloads, applications, and data to the cloud. Consider factors such as compliance requirements and application modernization.
- **Implement a robust ZTA** – Design and implement a ZTA that enforces granular access controls, strong authentication mechanisms, and continuous monitoring. For a more efficient ZTA adoption, use cloud-native Zero Trust services, such as AWS Verified Access and Amazon VPC Lattice.
- **Prioritize data and application security** – Apply Zero Trust principles—strong identity, micro-segmentation, and authorization—to provide all available context. Use this context for users accessing systems and resources and for the flow of communications and data within and between backend components.
- **Establish monitoring and incident response frameworks** – Implement robust security monitoring and incident response capabilities in the cloud environment. Use cloud-native security tools for real-time threat detection, log analysis, and incident response automation, such as Amazon Inspector, AWS Security Hub, and Amazon GuardDuty.

- **Foster a culture of security and compliance** – Promote a culture of security awareness and compliance throughout the organization. Educate employees on security best practices and their role in maintaining a secure cloud environment.
- **Continuously assess and optimize** – Regularly assess the cloud environment, security controls, and operational processes. To gather insights and optimize resource utilization, cost management, and performance, use cloud-native analytics and monitoring tools such as Amazon CloudWatch and AWS Security Hub.
- **Establish governance and compliance frameworks** – Develop governance and compliance frameworks that align with industry standards and regulatory requirements. Define policies, procedures, and controls to help ensure adherence to security, privacy, and compliance standards.

Next steps

Adopting a zero trust architecture (ZTA) is one of the most secure ways to improve your organization's posture and reduce risk. This prescriptive guidance has provided you with a comprehensive roadmap for implementing Zero Trust, from understanding the principles to assessing your readiness, to implementing the necessary components.

The next steps in this workstream or domain involve the following:

- Implementing the adoption plan
- Implementing the ZTA
- Conducting regular security assessments
- Continuously optimizing the cloud environment and security controls

ZTA is an ongoing process that requires constant monitoring, evaluation, and adaptation to ensure a strong security foundation. By following the best practices outlined in this guidance, your organization can enhance its security posture, ensure compliance with regulations, and protect sensitive data.

FAQ

This section provides answers to commonly raised questions about designing and implementing a zero trust architecture (ZTA).

What is Zero Trust?

Zero trust is a conceptual model and an associated set of mechanisms that focus on providing security controls around digital assets that do not solely or fundamentally depend on traditional network controls or network perimeters. Instead, network controls are augmented with identity, device, behavior, and other rich context and signals to make more granular, intelligent, adaptive, and continuous access decisions.

What AWS services can help me implement zero trust architecture?

AWS provides several services that can assist in implementing Zero Trust, such as AWS Verified Access, AWS Identity and Access Management (IAM), Amazon Virtual Private Cloud (Amazon VPC), Amazon VPC Lattice, Amazon Verified Permissions, Amazon API Gateway, and Amazon GuardDuty.

How can I ensure data security with AWS?

AWS offers services such as AWS Key Management Service (AWS KMS) for data encryption at rest and in transit, Amazon Virtual Private Cloud (Amazon VPC) for network isolation, and AWS Secrets Manager for secure storage and retrieval of credentials.

Can AWS help with compliance requirements in a Zero Trust environment?

Yes, AWS has compliance programs and services to help meet various regulatory requirements. AWS Artifact provides access to AWS compliance reports, and AWS Config supports continuous monitoring and assessment of compliance.

Are there any AWS tools or services for automating security in a Zero Trust environment?

AWS provides services such as AWS Security Hub, which centralizes and automates security findings, and AWS Config rules for defining and enforcing security policies.

How can I ensure continuous monitoring and incident response in a Zero Trust cloud environment with AWS

AWS offers services such as Amazon CloudWatch for real-time monitoring and AWS CloudTrail for logging and analysis. For incident response best practices, you can use the AWS Security Incident Response Guide.

Resources

References

- What is a cloud center of excellence and why should your organization create one? – This blog post provides an overview of CCoE, best practices for how to create an effective CCoE, and more.
- Zero Trust on AWS – This page provides an overview of Zero Trust security principles and best practices in the AWS environment.
- Zero Trust architecture: An AWS perspective – This blog post shares a definition and guiding principles for the way that Zero Trust is implemented at AWS.
- AWS Identity and Access Management (IAM) User Guide – This guide offers comprehensive documentation on managing user access and permissions in IAM, a crucial component of zero trust architecture.
- AWS Security Hub – Learn about Security Hub, a service that provides a comprehensive view of security alerts and compliance status across your AWS accounts.
- AWS Well-Architected Framework – Explore the Well-Architected Framework, which provides guidance on building secure, high-performing, resilient, and efficient architectures on AWS.
- AWS Security Incident Response Guide – This guide presents an overview of the fundamentals of responding to security incidents within your organization's AWS Cloud environment. It provides an overview of cloud security and incident response concepts and identifies cloud capabilities, services, and mechanisms that are available to customers who respond to security issues.

Tools

- Amazon API Gateway
- AWS Artifact
- AWS CloudTrail
- Amazon CloudWatch
- AWS Config
- Amazon GuardDuty
- AWS Identity and Access Management
- AWS Key Management Service

- AWS Secrets Manager
- AWS Security Hub
- AWS Verified Access

Document history

The following table describes significant changes to this guide. If you want to be notified about future updates, you can subscribe to an RSS feed.

Change	Description	Date
Added updates	Added information to the Key components of a zero trust architecture section, made changes in the Assessing organizational readiness for Zero Trust adoption section, added information to the Best practices section, and made changes to the FAQ.	December 4, 2023
Initial publication	—	June 19, 2023

AWS Prescriptive Guidance glossary

The following are commonly used terms in strategies, guides, and patterns provided by AWS Prescriptive Guidance. To suggest entries, please use the **Provide feedback** link at the end of the glossary.

Numbers

7 Rs

Seven common migration strategies for moving applications to the cloud. These strategies build upon the 5 Rs that Gartner identified in 2011 and consist of the following:

- Refactor/re-architect – Move an application and modify its architecture by taking full advantage of cloud-native features to improve agility, performance, and scalability. This typically involves porting the operating system and database. Example: Migrate your on-premises Oracle database to the Amazon Aurora PostgreSQL-Compatible Edition.
- Replatform (lift and reshape) – Move an application to the cloud, and introduce some level of optimization to take advantage of cloud capabilities. Example: Migrate your on-premises Oracle database to Amazon Relational Database Service (Amazon RDS) for Oracle in the AWS Cloud.
- Repurchase (drop and shop) – Switch to a different product, typically by moving from a traditional license to a SaaS model. Example: Migrate your customer relationship management (CRM) system to Salesforce.com.
- Rehost (lift and shift) – Move an application to the cloud without making any changes to take advantage of cloud capabilities. Example: Migrate your on-premises Oracle database to Oracle on an EC2 instance in the AWS Cloud.
- Relocate (hypervisor-level lift and shift) – Move infrastructure to the cloud without purchasing new hardware, rewriting applications, or modifying your existing operations. You migrate servers from an on-premises platform to a cloud service for the same platform. Example: Migrate a Microsoft Hyper-V application to AWS.
- Retain (revisit) – Keep applications in your source environment. These might include applications that require major refactoring, and you want to postpone that work until a later time, and legacy applications that you want to retain, because there's no business justification for migrating them.

- Retire – Decommission or remove applications that are no longer needed in your source environment.

A

ABAC

See attribute-based access control.

abstracted services

See managed services.

ACID

See atomicity, consistency, isolation, durability.

active-active migration

A database migration method in which the source and target databases are kept in sync (by using a bidirectional replication tool or dual write operations), and both databases handle transactions from connecting applications during migration. This method supports migration in small, controlled batches instead of requiring a one-time cutover. It's more flexible but requires more work than active-passive migration.

active-passive migration

A database migration method in which in which the source and target databases are kept in sync, but only the source database handles transactions from connecting applications while data is replicated to the target database. The target database doesn't accept any transactions during migration.

aggregate function

A SQL function that operates on a group of rows and calculates a single return value for the group. Examples of aggregate functions include `SUM` and `MAX`.

AI

See artificial intelligence.

AIOps

See artificial intelligence operations.

anonymization

The process of permanently deleting personal information in a dataset. Anonymization can help protect personal privacy. Anonymized data is no longer considered to be personal data.

anti-pattern

A frequently used solution for a recurring issue where the solution is counter-productive, ineffective, or less effective than an alternative.

application control

A security approach that allows the use of only approved applications in order to help protect a system from malware.

application portfolio

A collection of detailed information about each application used by an organization, including the cost to build and maintain the application, and its business value. This information is key to the portfolio discovery and analysis process and helps identify and prioritize the applications to be migrated, modernized, and optimized.

artificial intelligence (AI)

The field of computer science that is dedicated to using computing technologies to perform cognitive functions that are typically associated with humans, such as learning, solving problems, and recognizing patterns. For more information, see What is Artificial Intelligence?

artificial intelligence operations (AIOps)

The process of using machine learning techniques to solve operational problems, reduce operational incidents and human intervention, and increase service quality. For more information about how AIOps is used in the AWS migration strategy, see the operations integration guide.

asymmetric encryption

An encryption algorithm that uses a pair of keys, a public key for encryption and a private key for decryption. You can share the public key because it isn't used for decryption, but access to the private key should be highly restricted.

atomicity, consistency, isolation, durability (ACID)

A set of software properties that guarantee the data validity and operational reliability of a database, even in the case of errors, power failures, or other problems.

attribute-based access control (ABAC)

The practice of creating fine-grained permissions based on user attributes, such as department, job role, and team name. For more information, see ABAC for AWS in the AWS Identity and Access Management (IAM) documentation.

authoritative data source

A location where you store the primary version of data, which is considered to be the most reliable source of information. You can copy data from the authoritative data source to other locations for the purposes of processing or modifying the data, such as anonymizing, redacting, or pseudonymizing it.

Availability Zone

A distinct location within an AWS Region that is insulated from failures in other Availability Zones and provides inexpensive, low-latency network connectivity to other Availability Zones in the same Region.

AWS Cloud Adoption Framework (AWS CAF)

A framework of guidelines and best practices from AWS to help organizations develop an efficient and effective plan to move successfully to the cloud. AWS CAF organizes guidance into six focus areas called perspectives: business, people, governance, platform, security, and operations. The business, people, and governance perspectives focus on business skills and processes; the platform, security, and operations perspectives focus on technical skills and processes. For example, the people perspective targets stakeholders who handle human resources (HR), staffing functions, and people management. For this perspective, AWS CAF provides guidance for people development, training, and communications to help ready the organization for successful cloud adoption. For more information, see the AWS CAF website and the AWS CAF whitepaper.

AWS Workload Qualification Framework (AWS WQF)

A tool that evaluates database migration workloads, recommends migration strategies, and provides work estimates. AWS WQF is included with AWS Schema Conversion Tool (AWS SCT). It analyzes database schemas and code objects, application code, dependencies, and performance characteristics, and provides assessment reports.

B

bad bot

A bot that is intended to disrupt or cause harm to individuals or organizations.

BCP

See business continuity planning.

behavior graph

A unified, interactive view of resource behavior and interactions over time. You can use a behavior graph with Amazon Detective to examine failed logon attempts, suspicious API calls, and similar actions. For more information, see Data in a behavior graph in the Detective documentation.

big-endian system

A system that stores the most significant byte first. See also endianness.

binary classification

A process that predicts a binary outcome (one of two possible classes). For example, your ML model might need to predict problems such as “Is this email spam or not spam?" or "Is this product a book or a car?"

bloom filter

A probabilistic, memory-efficient data structure that is used to test whether an element is a member of a set.

blue/green deployment

A deployment strategy where you create two separate but identical environments. You run the current application version in one environment (blue) and the new application version in the other environment (green). This strategy helps you quickly roll back with minimal impact.

bot

A software application that runs automated tasks over the internet and simulates human activity or interaction. Some bots are useful or beneficial, such as web crawlers that index information on the internet. Some other bots, known as *bad bots*, are intended to disrupt or cause harm to individuals or organizations.

botnet

Networks of bots that are infected by malware and are under the control of a single party, known as a *bot herder* or *bot operator*. Botnets are the best-known mechanism to scale bots and their impact.

branch

A contained area of a code repository. The first branch created in a repository is the *main branch*. You can create a new branch from an existing branch, and you can then develop features or fix bugs in the new branch. A branch you create to build a feature is commonly referred to as a *feature branch*. When the feature is ready for release, you merge the feature branch back into the main branch. For more information, see About branches (GitHub documentation).

break-glass access

In exceptional circumstances and through an approved process, a quick means for a user to gain access to an AWS account that they don't typically have permissions to access. For more information, see the Implement break-glass procedures indicator in the AWS Well-Architected guidance.

brownfield strategy

The existing infrastructure in your environment. When adopting a brownfield strategy for a system architecture, you design the architecture around the constraints of the current systems and infrastructure. If you are expanding the existing infrastructure, you might blend brownfield and greenfield strategies.

buffer cache

The memory area where the most frequently accessed data is stored.

business capability

What a business does to generate value (for example, sales, customer service, or marketing). Microservices architectures and development decisions can be driven by business capabilities. For more information, see the Organized around business capabilities section of the Running containerized microservices on AWS whitepaper.

business continuity planning (BCP)

A plan that addresses the potential impact of a disruptive event, such as a large-scale migration, on operations and enables a business to resume operations quickly.

C

CAF

See AWS Cloud Adoption Framework.

canary deployment

The slow and incremental release of a version to end users. When you are confident, you deploy the new version and replace the current version in its entirety.

CCoE

See Cloud Center of Excellence.

CDC

See change data capture.

change data capture (CDC)

The process of tracking changes to a data source, such as a database table, and recording metadata about the change. You can use CDC for various purposes, such as auditing or replicating changes in a target system to maintain synchronization.

chaos engineering

Intentionally introducing failures or disruptive events to test a system's resilience. You can use AWS Fault Injection Service (AWS FIS) to perform experiments that stress your AWS workloads and evaluate their response.

CI/CD

See continuous integration and continuous delivery.

classification

A categorization process that helps generate predictions. ML models for classification problems predict a discrete value. Discrete values are always distinct from one another. For example, a model might need to evaluate whether or not there is a car in an image.

client-side encryption

Encryption of data locally, before the target AWS service receives it.

Cloud Center of Excellence (CCoE)

A multi-disciplinary team that drives cloud adoption efforts across an organization, including developing cloud best practices, mobilizing resources, establishing migration timelines, and leading the organization through large-scale transformations. For more information, see the CCoE posts on the AWS Cloud Enterprise Strategy Blog.

cloud computing

The cloud technology that is typically used for remote data storage and IoT device management. Cloud computing is commonly connected to edge computing technology.

cloud operating model

In an IT organization, the operating model that is used to build, mature, and optimize one or more cloud environments. For more information, see Building your Cloud Operating Model.

cloud stages of adoption

The four phases that organizations typically go through when they migrate to the AWS Cloud:

- Project – Running a few cloud-related projects for proof of concept and learning purposes
- Foundation – Making foundational investments to scale your cloud adoption (e.g., creating a landing zone, defining a CCoE, establishing an operations model)
- Migration – Migrating individual applications
- Re-invention – Optimizing products and services, and innovating in the cloud

These stages were defined by Stephen Orban in the blog post The Journey Toward Cloud-First & the Stages of Adoption on the AWS Cloud Enterprise Strategy blog. For information about how they relate to the AWS migration strategy, see the migration readiness guide.

CMDB

See configuration management database.

code repository

A location where source code and other assets, such as documentation, samples, and scripts, are stored and updated through version control processes. Common cloud repositories include GitHub or Bitbucket Cloud. Each version of the code is called a *branch*. In a microservice structure, each repository is devoted to a single piece of functionality. A single CI/CD pipeline can use multiple repositories.

cold cache

A buffer cache that is empty, not well populated, or contains stale or irrelevant data. This affects performance because the database instance must read from the main memory or disk, which is slower than reading from the buffer cache.

cold data

Data that is rarely accessed and is typically historical. When querying this kind of data, slow queries are typically acceptable. Moving this data to lower-performing and less expensive storage tiers or classes can reduce costs.

computer vision (CV)

A field of AI that uses machine learning to analyze and extract information from visual formats such as digital images and videos. For example, AWS Panorama offers devices that add CV to on-premises camera networks, and Amazon SageMaker provides image processing algorithms for CV.

configuration drift

For a workload, a configuration change from the expected state. It might cause the workload to become noncompliant, and it's typically gradual and unintentional.

configuration management database (CMDB)

A repository that stores and manages information about a database and its IT environment, including both hardware and software components and their configurations. You typically use data from a CMDB in the portfolio discovery and analysis stage of migration.

conformance pack

A collection of AWS Config rules and remediation actions that you can assemble to customize your compliance and security checks. You can deploy a conformance pack as a single entity in an AWS account and Region, or across an organization, by using a YAML template. For more information, see Conformance packs in the AWS Config documentation.

continuous integration and continuous delivery (CI/CD)

The process of automating the source, build, test, staging, and production stages of the software release process. CI/CD is commonly described as a pipeline. CI/CD can help you automate processes, improve productivity, improve code quality, and deliver faster. For more information, see Benefits of continuous delivery. CD can also stand for *continuous deployment*. For more information, see Continuous Delivery vs. Continuous Deployment.

CV

See computer vision.

D

data at rest

Data that is stationary in your network, such as data that is in storage.

data classification

A process for identifying and categorizing the data in your network based on its criticality and sensitivity. It is a critical component of any cybersecurity risk management strategy because it helps you determine the appropriate protection and retention controls for the data. Data classification is a component of the security pillar in the AWS Well-Architected Framework. For more information, see Data classification.

data drift

A meaningful variation between the production data and the data that was used to train an ML model, or a meaningful change in the input data over time. Data drift can reduce the overall quality, accuracy, and fairness in ML model predictions.

data in transit

Data that is actively moving through your network, such as between network resources.

data mesh

An architectural framework that provides distributed, decentralized data ownership with centralized management and governance.

data minimization

The principle of collecting and processing only the data that is strictly necessary. Practicing data minimization in the AWS Cloud can reduce privacy risks, costs, and your analytics carbon footprint.

data perimeter

A set of preventive guardrails in your AWS environment that help make sure that only trusted identities are accessing trusted resources from expected networks. For more information, see Building a data perimeter on AWS.

data preprocessing

To transform raw data into a format that is easily parsed by your ML model. Preprocessing data can mean removing certain columns or rows and addressing missing, inconsistent, or duplicate values.

data provenance

The process of tracking the origin and history of data throughout its lifecycle, such as how the data was generated, transmitted, and stored.

data subject

An individual whose data is being collected and processed.

data warehouse

A data management system that supports business intelligence, such as analytics. Data warehouses commonly contain large amounts of historical data, and they are typically used for queries and analysis.

database definition language (DDL)

Statements or commands for creating or modifying the structure of tables and objects in a database.

database manipulation language (DML)

Statements or commands for modifying (inserting, updating, and deleting) information in a database.

DDL

See database definition language.

deep ensemble

To combine multiple deep learning models for prediction. You can use deep ensembles to obtain a more accurate prediction or for estimating uncertainty in predictions.

deep learning

An ML subfield that uses multiple layers of artificial neural networks to identify mapping between input data and target variables of interest.

defense-in-depth

An information security approach in which a series of security mechanisms and controls are thoughtfully layered throughout a computer network to protect the confidentiality, integrity, and availability of the network and the data within. When you adopt this strategy on AWS, you add multiple controls at different layers of the AWS Organizations structure to help secure resources. For example, a defense-in-depth approach might combine multi-factor authentication, network segmentation, and encryption.

delegated administrator

In AWS Organizations, a compatible service can register an AWS member account to administer the organization's accounts and manage permissions for that service. This account is called the *delegated administrator* for that service. For more information and a list of compatible services, see Services that work with AWS Organizations in the AWS Organizations documentation.

deployment

The process of making an application, new features, or code fixes available in the target environment. Deployment involves implementing changes in a code base and then building and running that code base in the application's environments.

development environment

See environment.

detective control

A security control that is designed to detect, log, and alert after an event has occurred. These controls are a second line of defense, alerting you to security events that bypassed the preventative controls in place. For more information, see Detective controls in *Implementing security controls on AWS*.

development value stream mapping (DVSM)

A process used to identify and prioritize constraints that adversely affect speed and quality in a software development lifecycle. DVSM extends the value stream mapping process originally designed for lean manufacturing practices. It focuses on the steps and teams required to create and move value through the software development process.

digital twin

A virtual representation of a real-world system, such as a building, factory, industrial equipment, or production line. Digital twins support predictive maintenance, remote monitoring, and production optimization.

dimension table

In a star schema, a smaller table that contains data attributes about quantitative data in a fact table. Dimension table attributes are typically text fields or discrete numbers that behave like text. These attributes are commonly used for query constraining, filtering, and result set labeling.

disaster

An event that prevents a workload or system from fulfilling its business objectives in its primary deployed location. These events can be natural disasters, technical failures, or the result of human actions, such as unintentional misconfiguration or a malware attack.

disaster recovery (DR)

The strategy and process you use to minimize downtime and data loss caused by a disaster. For more information, see Disaster Recovery of Workloads on AWS: Recovery in the Cloud in the AWS Well-Architected Framework.

DML

See database manipulation language.

domain-driven design

An approach to developing a complex software system by connecting its components to evolving domains, or core business goals, that each component serves. This concept was introduced by Eric Evans in his book, *Domain-Driven Design: Tackling Complexity in the Heart of Software* (Boston: Addison-Wesley Professional, 2003). For information about how you can use domain-driven design with the strangler fig pattern, see Modernizing legacy Microsoft ASP.NET (ASMX) web services incrementally by using containers and Amazon API Gateway.

DR

See disaster recovery.

drift detection

Tracking deviations from a baselined configuration. For example, you can use AWS CloudFormation to detect drift in system resources, or you can use AWS Control Tower to detect changes in your landing zone that might affect compliance with governance requirements.

DVSM

See development value stream mapping.

E

EDA

See exploratory data analysis.

edge computing

The technology that increases the computing power for smart devices at the edges of an IoT network. When compared with cloud computing, edge computing can reduce communication latency and improve response time.

encryption

A computing process that transforms plaintext data, which is human-readable, into ciphertext.

encryption key

A cryptographic string of randomized bits that is generated by an encryption algorithm. Keys can vary in length, and each key is designed to be unpredictable and unique.

endianness

The order in which bytes are stored in computer memory. Big-endian systems store the most significant byte first. Little-endian systems store the least significant byte first.

endpoint

See service endpoint.

endpoint service

A service that you can host in a virtual private cloud (VPC) to share with other users. You can create an endpoint service with AWS PrivateLink and grant permissions to other AWS accounts

or to AWS Identity and Access Management (IAM) principals. These accounts or principals can connect to your endpoint service privately by creating interface VPC endpoints. For more information, see Create an endpoint service in the Amazon Virtual Private Cloud (Amazon VPC) documentation.

enterprise resource planning (ERP)

A system that automates and manages key business processes (such as accounting, MES, and project management) for an enterprise.

envelope encryption

The process of encrypting an encryption key with another encryption key. For more information, see Envelope encryption in the AWS Key Management Service (AWS KMS) documentation.

environment

An instance of a running application. The following are common types of environments in cloud computing:

- development environment – An instance of a running application that is available only to the core team responsible for maintaining the application. Development environments are used to test changes before promoting them to upper environments. This type of environment is sometimes referred to as a *test environment*.
- lower environments – All development environments for an application, such as those used for initial builds and tests.
- production environment – An instance of a running application that end users can access. In a CI/CD pipeline, the production environment is the last deployment environment.
- upper environments – All environments that can be accessed by users other than the core development team. This can include a production environment, preproduction environments, and environments for user acceptance testing.

epic

In agile methodologies, functional categories that help organize and prioritize your work. Epics provide a high-level description of requirements and implementation tasks. For example, AWS CAF security epics include identity and access management, detective controls, infrastructure security, data protection, and incident response. For more information about epics in the AWS migration strategy, see the program implementation guide.

ERP

See enterprise resource planning.

exploratory data analysis (EDA)

The process of analyzing a dataset to understand its main characteristics. You collect or aggregate data and then perform initial investigations to find patterns, detect anomalies, and check assumptions. EDA is performed by calculating summary statistics and creating data visualizations.

F

fact table

The central table in a star schema. It stores quantitative data about business operations. Typically, a fact table contains two types of columns: those that contain measures and those that contain a foreign key to a dimension table.

fail fast

A philosophy that uses frequent and incremental testing to reduce the development lifecycle. It is a critical part of an agile approach.

fault isolation boundary

In the AWS Cloud, a boundary such as an Availability Zone, AWS Region, control plane, or data plane that limits the effect of a failure and helps improve the resilience of workloads. For more information, see AWS Fault Isolation Boundaries.

feature branch

See branch.

features

The input data that you use to make a prediction. For example, in a manufacturing context, features could be images that are periodically captured from the manufacturing line.

feature importance

How significant a feature is for a model's predictions. This is usually expressed as a numerical score that can be calculated through various techniques, such as Shapley Additive Explanations

(SHAP) and integrated gradients. For more information, see Machine learning model interpretability with :AWS.

feature transformation

To optimize data for the ML process, including enriching data with additional sources, scaling values, or extracting multiple sets of information from a single data field. This enables the ML model to benefit from the data. For example, if you break down the “2021-05-27 00:15:37” date into “2021”, “May”, “Thu”, and “15”, you can help the learning algorithm learn nuanced patterns associated with different data components.

FGAC

See fine-grained access control.

fine-grained access control (FGAC)

The use of multiple conditions to allow or deny an access request.

flash-cut migration

A database migration method that uses continuous data replication through change data capture to migrate data in the shortest time possible, instead of using a phased approach. The objective is to keep downtime to a minimum.

G

geo blocking

See geographic restrictions.

geographic restrictions (geo blocking)

In Amazon CloudFront, an option to prevent users in specific countries from accessing content distributions. You can use an allow list or block list to specify approved and banned countries. For more information, see Restricting the geographic distribution of your content in the CloudFront documentation.

Gitflow workflow

An approach in which lower and upper environments use different branches in a source code repository. The Gitflow workflow is considered legacy, and the trunk-based workflow is the modern, preferred approach.

greenfield strategy

The absence of existing infrastructure in a new environment. When adopting a greenfield strategy for a system architecture, you can select all new technologies without the restriction of compatibility with existing infrastructure, also known as brownfield. If you are expanding the existing infrastructure, you might blend brownfield and greenfield strategies.

guardrail

A high-level rule that helps govern resources, policies, and compliance across organizational units (OUs). *Preventive guardrails* enforce policies to ensure alignment to compliance standards. They are implemented by using service control policies and IAM permissions boundaries. *Detective guardrails* detect policy violations and compliance issues, and generate alerts for remediation. They are implemented by using AWS Config, AWS Security Hub, Amazon GuardDuty, AWS Trusted Advisor, Amazon Inspector, and custom AWS Lambda checks.

H

HA

See high availability.

heterogeneous database migration

Migrating your source database to a target database that uses a different database engine (for example, Oracle to Amazon Aurora). Heterogeneous migration is typically part of a re-architecting effort, and converting the schema can be a complex task. AWS provides AWS SCT that helps with schema conversions.

high availability (HA)

The ability of a workload to operate continuously, without intervention, in the event of challenges or disasters. HA systems are designed to automatically fail over, consistently deliver high-quality performance, and handle different loads and failures with minimal performance impact.

historian modernization

An approach used to modernize and upgrade operational technology (OT) systems to better serve the needs of the manufacturing industry. A *historian* is a type of database that is used to collect and store data from various sources in a factory.

homogeneous database migration

Migrating your source database to a target database that shares the same database engine (for example, Microsoft SQL Server to Amazon RDS for SQL Server). Homogeneous migration is typically part of a rehosting or replatforming effort. You can use native database utilities to migrate the schema.

hot data

Data that is frequently accessed, such as real-time data or recent translational data. This data typically requires a high-performance storage tier or class to provide fast query responses.

hotfix

An urgent fix for a critical issue in a production environment. Due to its urgency, a hotfix is usually made outside of the typical DevOps release workflow.

hypercare period

Immediately following cutover, the period of time when a migration team manages and monitors the migrated applications in the cloud in order to address any issues. Typically, this period is 1–4 days in length. At the end of the hypercare period, the migration team typically transfers responsibility for the applications to the cloud operations team.

I

IaC

See infrastructure as code.

identity-based policy

A policy attached to one or more IAM principals that defines their permissions within the AWS Cloud environment.

idle application

An application that has an average CPU and memory usage between 5 and 20 percent over a period of 90 days. In a migration project, it is common to retire these applications or retain them on premises.

IIoT

See industrial Internet of Things.

immutable infrastructure

A model that deploys new infrastructure for production workloads instead of updating, patching, or modifying the existing infrastructure. Immutable infrastructures are inherently more consistent, reliable, and predictable than mutable infrastructure. For more information, see the Deploy using immutable infrastructure best practice in the AWS Well-Architected Framework.

inbound (ingress) VPC

In an AWS multi-account architecture, a VPC that accepts, inspects, and routes network connections from outside an application. The AWS Security Reference Architecture recommends setting up your Network account with inbound, outbound, and inspection VPCs to protect the two-way interface between your application and the broader internet.

incremental migration

A cutover strategy in which you migrate your application in small parts instead of performing a single, full cutover. For example, you might move only a few microservices or users to the new system initially. After you verify that everything is working properly, you can incrementally move additional microservices or users until you can decommission your legacy system. This strategy reduces the risks associated with large migrations.

Industry 4.0

A term that was introduced by Klaus Schwab in 2016 to refer to the modernization of manufacturing processes through advances in connectivity, real-time data, automation, analytics, and AI/ML.

infrastructure

All of the resources and assets contained within an application's environment.

infrastructure as code (IaC)

The process of provisioning and managing an application's infrastructure through a set of configuration files. IaC is designed to help you centralize infrastructure management, standardize resources, and scale quickly so that new environments are repeatable, reliable, and consistent.

industrial Internet of Things (IIoT)

The use of internet-connected sensors and devices in the industrial sectors, such as manufacturing, energy, automotive, healthcare, life sciences, and agriculture. For more information, see Building an industrial Internet of Things (IIoT) digital transformation strategy.

inspection VPC

In an AWS multi-account architecture, a centralized VPC that manages inspections of network traffic between VPCs (in the same or different AWS Regions), the internet, and on-premises networks. The AWS Security Reference Architecture recommends setting up your Network account with inbound, outbound, and inspection VPCs to protect the two-way interface between your application and the broader internet.

Internet of Things (IoT)

The network of connected physical objects with embedded sensors or processors that communicate with other devices and systems through the internet or over a local communication network. For more information, see What is IoT?

interpretability

A characteristic of a machine learning model that describes the degree to which a human can understand how the model's predictions depend on its inputs. For more information, see Machine learning model interpretability with AWS.

IoT

See Internet of Things.

IT information library (ITIL)

A set of best practices for delivering IT services and aligning these services with business requirements. ITIL provides the foundation for ITSM.

IT service management (ITSM)

Activities associated with designing, implementing, managing, and supporting IT services for an organization. For information about integrating cloud operations with ITSM tools, see the operations integration guide.

ITIL

See IT information library.

ITSM

See IT service management.

L

label-based access control (LBAC)

An implementation of mandatory access control (MAC) where the users and the data itself are each explicitly assigned a security label value. The intersection between the user security label and data security label determines which rows and columns can be seen by the user.

landing zone

A landing zone is a well-architected, multi-account AWS environment that is scalable and secure. This is a starting point from which your organizations can quickly launch and deploy workloads and applications with confidence in their security and infrastructure environment. For more information about landing zones, see Setting up a secure and scalable multi-account AWS environment.

large migration

A migration of 300 or more servers.

LBAC

See label-based access control.

least privilege

The security best practice of granting the minimum permissions required to perform a task. For more information, see Apply least-privilege permissions in the IAM documentation.

lift and shift

See 7 Rs.

little-endian system

A system that stores the least significant byte first. See also endianness.

lower environments

See environment.

M

machine learning (ML)

A type of artificial intelligence that uses algorithms and techniques for pattern recognition and learning. ML analyzes and learns from recorded data, such as Internet of Things (IoT) data, to generate a statistical model based on patterns. For more information, see Machine Learning.

main branch

See branch.

malware

Software that is designed to compromise computer security or privacy. Malware might disrupt computer systems, leak sensitive information, or gain unauthorized access. Examples of malware include viruses, worms, ransomware, Trojan horses, spyware, and keyloggers.

managed services

AWS services for which AWS operates the infrastructure layer, the operating system, and platforms, and you access the endpoints to store and retrieve data. Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB are examples of managed services. These are also known as *abstracted services*.

manufacturing execution system (MES)

A software system for tracking, monitoring, documenting, and controlling production processes that convert raw materials to finished products on the shop floor.

MAP

See Migration Acceleration Program.

mechanism

A complete process in which you create a tool, drive adoption of the tool, and then inspect the results in order to make adjustments. A mechanism is a cycle that reinforces and improves itself as it operates. For more information, see Building mechanisms in the AWS Well-Architected Framework.

member account

All AWS accounts other than the management account that are part of an organization in AWS Organizations. An account can be a member of only one organization at a time.

MES

See manufacturing execution system.

Message Queuing Telemetry Transport (MQTT)

A lightweight, machine-to-machine (M2M) communication protocol, based on the publish/subscribe pattern, for resource-constrained IoT devices.

microservice

A small, independent service that communicates over well-defined APIs and is typically owned by small, self-contained teams. For example, an insurance system might include microservices that map to business capabilities, such as sales or marketing, or subdomains, such as purchasing, claims, or analytics. The benefits of microservices include agility, flexible scaling, easy deployment, reusable code, and resilience. For more information, see Integrating microservices by using AWS serverless services.

microservices architecture

An approach to building an application with independent components that run each application process as a microservice. These microservices communicate through a well-defined interface by using lightweight APIs. Each microservice in this architecture can be updated, deployed, and scaled to meet demand for specific functions of an application. For more information, see Implementing microservices on AWS.

Migration Acceleration Program (MAP)

An AWS program that provides consulting support, training, and services to help organizations build a strong operational foundation for moving to the cloud, and to help offset the initial cost of migrations. MAP includes a migration methodology for executing legacy migrations in a methodical way and a set of tools to automate and accelerate common migration scenarios.

migration at scale

The process of moving the majority of the application portfolio to the cloud in waves, with more applications moved at a faster rate in each wave. This phase uses the best practices and lessons learned from the earlier phases to implement a *migration factory* of teams, tools, and processes to streamline the migration of workloads through automation and agile delivery. This is the third phase of the AWS migration strategy.

migration factory

Cross-functional teams that streamline the migration of workloads through automated, agile approaches. Migration factory teams typically include operations, business analysts and owners, migration engineers, developers, and DevOps professionals working in sprints. Between 20 and 50 percent of an enterprise application portfolio consists of repeated patterns that can be optimized by a factory approach. For more information, see the discussion of migration factories and the Cloud Migration Factory guide in this content set.

migration metadata

The information about the application and server that is needed to complete the migration. Each migration pattern requires a different set of migration metadata. Examples of migration metadata include the target subnet, security group, and AWS account.

migration pattern

A repeatable migration task that details the migration strategy, the migration destination, and the migration application or service used. Example: Rehost migration to Amazon EC2 with AWS Application Migration Service.

Migration Portfolio Assessment (MPA)

An online tool that provides information for validating the business case for migrating to the AWS Cloud. MPA provides detailed portfolio assessment (server right-sizing, pricing, TCO comparisons, migration cost analysis) as well as migration planning (application data analysis and data collection, application grouping, migration prioritization, and wave planning). The MPA tool (requires login) is available free of charge to all AWS consultants and APN Partner consultants.

Migration Readiness Assessment (MRA)

The process of gaining insights about an organization's cloud readiness status, identifying strengths and weaknesses, and building an action plan to close identified gaps, using the AWS CAF. For more information, see the migration readiness guide. MRA is the first phase of the AWS migration strategy.

migration strategy

The approach used to migrate a workload to the AWS Cloud. For more information, see the 7 Rs entry in this glossary and see Mobilize your organization to accelerate large-scale migrations.

ML

See machine learning.

modernization

Transforming an outdated (legacy or monolithic) application and its infrastructure into an agile, elastic, and highly available system in the cloud to reduce costs, gain efficiencies, and take advantage of innovations. For more information, see Strategy for modernizing applications in the AWS Cloud.

modernization readiness assessment

An evaluation that helps determine the modernization readiness of an organization's applications; identifies benefits, risks, and dependencies; and determines how well the organization can support the future state of those applications. The outcome of the assessment is a blueprint of the target architecture, a roadmap that details development phases and milestones for the modernization process, and an action plan for addressing identified gaps. For more information, see Evaluating modernization readiness for applications in the AWS Cloud.

monolithic applications (monoliths)

Applications that run as a single service with tightly coupled processes. Monolithic applications have several drawbacks. If one application feature experiences a spike in demand, the entire architecture must be scaled. Adding or improving a monolithic application's features also becomes more complex when the code base grows. To address these issues, you can use a microservices architecture. For more information, see Decomposing monoliths into microservices.

MPA

See Migration Portfolio Assessment.

MQTT

See Message Queuing Telemetry Transport.

multiclass classification

A process that helps generate predictions for multiple classes (predicting one of more than two outcomes). For example, an ML model might ask "Is this product a book, car, or phone?" or "Which product category is most interesting to this customer?"

mutable infrastructure

A model that updates and modifies the existing infrastructure for production workloads. For improved consistency, reliability, and predictability, the AWS Well-Architected Framework recommends the use of immutable infrastructure as a best practice.

O

OAC

See origin access control.

OAI

See origin access identity.

OCM

See organizational change management.

offline migration

A migration method in which the source workload is taken down during the migration process. This method involves extended downtime and is typically used for small, non-critical workloads.

OI

See operations integration.

OLA

See operational-level agreement.

online migration

A migration method in which the source workload is copied to the target system without being taken offline. Applications that are connected to the workload can continue to function during the migration. This method involves zero to minimal downtime and is typically used for critical production workloads.

OPC-UA

See Open Process Communications - Unified Architecture.

Open Process Communications - Unified Architecture (OPC-UA)

A machine-to-machine (M2M) communication protocol for industrial automation. OPC-UA provides an interoperability standard with data encryption, authentication, and authorization schemes.

operational-level agreement (OLA)

An agreement that clarifies what functional IT groups promise to deliver to each other, to support a service-level agreement (SLA).

operational readiness review (ORR)

A checklist of questions and associated best practices that help you understand, evaluate, prevent, or reduce the scope of incidents and possible failures. For more information, see Operational Readiness Reviews (ORR) in the AWS Well-Architected Framework.

operational technology (OT)

Hardware and software systems that work with the physical environment to control industrial operations, equipment, and infrastructure. In manufacturing, the integration of OT and information technology (IT) systems is a key focus for Industry 4.0 transformations.

operations integration (OI)

The process of modernizing operations in the cloud, which involves readiness planning, automation, and integration. For more information, see the operations integration guide.

organization trail

A trail that's created by AWS CloudTrail that logs all events for all AWS accounts in an organization in AWS Organizations. This trail is created in each AWS account that's part of the organization and tracks the activity in each account. For more information, see Creating a trail for an organization in the CloudTrail documentation.

organizational change management (OCM)

A framework for managing major, disruptive business transformations from a people, culture, and leadership perspective. OCM helps organizations prepare for, and transition to, new systems and strategies by accelerating change adoption, addressing transitional issues, and driving cultural and organizational changes. In the AWS migration strategy, this framework is called *people acceleration*, because of the speed of change required in cloud adoption projects. For more information, see the OCM guide.

origin access control (OAC)

In CloudFront, an enhanced option for restricting access to secure your Amazon Simple Storage Service (Amazon S3) content. OAC supports all S3 buckets in all AWS Regions, server-side encryption with AWS KMS (SSE-KMS), and dynamic PUT and DELETE requests to the S3 bucket.

origin access identity (OAI)

In CloudFront, an option for restricting access to secure your Amazon S3 content. When you use OAI, CloudFront creates a principal that Amazon S3 can authenticate with. Authenticated principals can access content in an S3 bucket only through a specific CloudFront distribution. See also OAC, which provides more granular and enhanced access control.

ORR

See operational readiness review.

OT

See operational technology.

outbound (egress) VPC

In an AWS multi-account architecture, a VPC that handles network connections that are initiated from within an application. The AWS Security Reference Architecture recommends setting up your Network account with inbound, outbound, and inspection VPCs to protect the two-way interface between your application and the broader internet.

P

permissions boundary

An IAM management policy that is attached to IAM principals to set the maximum permissions that the user or role can have. For more information, see Permissions boundaries in the IAM documentation.

personally identifiable information (PII)

Information that, when viewed directly or paired with other related data, can be used to reasonably infer the identity of an individual. Examples of PII include names, addresses, and contact information.

PII

See personally identifiable information.

playbook

A set of predefined steps that capture the work associated with migrations, such as delivering core operations functions in the cloud. A playbook can take the form of scripts, automated runbooks, or a summary of processes or steps required to operate your modernized environment.

PLC

See programmable logic controller.

PLM

See product lifecycle management.

policy

An object that can define permissions (see identity-based policy), specify access conditions (see resource-based policy), or define the maximum permissions for all accounts in an organization in AWS Organizations (see service control policy).

polyglot persistence

Independently choosing a microservice's data storage technology based on data access patterns and other requirements. If your microservices have the same data storage technology, they can encounter implementation challenges or experience poor performance. Microservices are more easily implemented and achieve better performance and scalability if they use the data store best adapted to their requirements. For more information, see Enabling data persistence in microservices.

portfolio assessment

A process of discovering, analyzing, and prioritizing the application portfolio in order to plan the migration. For more information, see Evaluating migration readiness.

predicate

A query condition that returns `true` or `false`, commonly located in a `WHERE` clause.

predicate pushdown

A database query optimization technique that filters the data in the query before transfer. This reduces the amount of data that must be retrieved and processed from the relational database, and it improves query performance.

preventative control

A security control that is designed to prevent an event from occurring. These controls are a first line of defense to help prevent unauthorized access or unwanted changes to your network. For more information, see Preventative controls in *Implementing security controls on AWS*.

principal

An entity in AWS that can perform actions and access resources. This entity is typically a root user for an AWS account, an IAM role, or a user. For more information, see *Principal* in Roles terms and concepts in the IAM documentation.

Privacy by Design

An approach in system engineering that takes privacy into account throughout the whole engineering process.

private hosted zones

A container that holds information about how you want Amazon Route 53 to respond to DNS queries for a domain and its subdomains within one or more VPCs. For more information, see Working with private hosted zones in the Route 53 documentation.

proactive control

A security control designed to prevent the deployment of noncompliant resources. These controls scan resources before they are provisioned. If the resource is not compliant with the control, then it isn't provisioned. For more information, see the Controls reference guide in the AWS Control Tower documentation and see Proactive controls in *Implementing security controls on AWS*.

product lifecycle management (PLM)

The management of data and processes for a product throughout its entire lifecycle, from design, development, and launch, through growth and maturity, to decline and removal.

production environment

See environment.

programmable logic controller (PLC)

In manufacturing, a highly reliable, adaptable computer that monitors machines and automates manufacturing processes.

pseudonymization

The process of replacing personal identifiers in a dataset with placeholder values. Pseudonymization can help protect personal privacy. Pseudonymized data is still considered to be personal data.

publish/subscribe (pub/sub)

A pattern that enables asynchronous communications among microservices to improve scalability and responsiveness. For example, in a microservices-based MES, a microservice can publish event messages to a channel that other microservices can subscribe to. The system can add new microservices without changing the publishing service.

Q

query plan

A series of steps, like instructions, that are used to access the data in a SQL relational database system.

query plan regression

When a database service optimizer chooses a less optimal plan than it did before a given change to the database environment. This can be caused by changes to statistics, constraints, environment settings, query parameter bindings, and updates to the database engine.

R

RACI matrix

See responsible, accountable, consulted, informed (RACI).

ransomware

A malicious software that is designed to block access to a computer system or data until a payment is made.

RASCI matrix

See responsible, accountable, consulted, informed (RACI).

RCAC

See row and column access control.

read replica

A copy of a database that's used for read-only purposes. You can route queries to the read replica to reduce the load on your primary database.

re-architect

See 7 Rs.

recovery point objective (RPO)

The maximum acceptable amount of time since the last data recovery point. This determines what is considered an acceptable loss of data between the last recovery point and the interruption of service.

recovery time objective (RTO)

The maximum acceptable delay between the interruption of service and restoration of service.

refactor

See 7 Rs.

Region

A collection of AWS resources in a geographic area. Each AWS Region is isolated and independent of the others to provide fault tolerance, stability, and resilience. For more information, see Specify which AWS Regions your account can use.

regression

An ML technique that predicts a numeric value. For example, to solve the problem of "What price will this house sell for?" an ML model could use a linear regression model to predict a house's sale price based on known facts about the house (for example, the square footage).

rehost

See 7 Rs.

release

In a deployment process, the act of promoting changes to a production environment.

relocate

See 7 Rs.

replatform

See 7 Rs.

repurchase

See 7 Rs.

resiliency

An application's ability to resist or recover from disruptions. High availability and disaster recovery are common considerations when planning for resiliency in the AWS Cloud. For more information, see AWS Cloud Resilience.

resource-based policy

A policy attached to a resource, such as an Amazon S3 bucket, an endpoint, or an encryption key. This type of policy specifies which principals are allowed access, supported actions, and any other conditions that must be met.

responsible, accountable, consulted, informed (RACI) matrix

A matrix that defines the roles and responsibilities for all parties involved in migration activities and cloud operations. The matrix name is derived from the responsibility types defined in the matrix: responsible (R), accountable (A), consulted (C), and informed (I). The support (S) type is optional. If you include support, the matrix is called a *RASCI matrix*, and if you exclude it, it's called a *RACI matrix*.

responsive control

A security control that is designed to drive remediation of adverse events or deviations from your security baseline. For more information, see Responsive controls in *Implementing security controls on AWS*.

retain

See 7 Rs.

retire

See 7 Rs.

rotation

The process of periodically updating a secret to make it more difficult for an attacker to access the credentials.

row and column access control (RCAC)

The use of basic, flexible SQL expressions that have defined access rules. RCAC consists of row permissions and column masks.

RPO

See recovery point objective.

RTO

See recovery time objective.

runbook

A set of manual or automated procedures required to perform a specific task. These are typically built to streamline repetitive operations or procedures with high error rates.

S

SAML 2.0

An open standard that many identity providers (IdPs) use. This feature enables federated single sign-on (SSO), so users can log into the AWS Management Console or call the AWS API operations without you having to create user in IAM for everyone in your organization. For more information about SAML 2.0-based federation, see About SAML 2.0-based federation in the IAM documentation.

SCADA

See supervisory control and data acquisition.

SCP

See service control policy.

secret

In AWS Secrets Manager, confidential or restricted information, such as a password or user credentials, that you store in encrypted form. It consists of the secret value and its metadata. The secret value can be binary, a single string, or multiple strings. For more information, see What's in a Secrets Manager secret? in the Secrets Manager documentation.

security control

A technical or administrative guardrail that prevents, detects, or reduces the ability of a threat actor to exploit a security vulnerability. There are four primary types of security controls: preventative, detective, responsive, and proactive.

security hardening

The process of reducing the attack surface to make it more resistant to attacks. This can include actions such as removing resources that are no longer needed, implementing the security best practice of granting least privilege, or deactivating unnecessary features in configuration files.

security information and event management (SIEM) system

Tools and services that combine security information management (SIM) and security event management (SEM) systems. A SIEM system collects, monitors, and analyzes data from servers, networks, devices, and other sources to detect threats and security breaches, and to generate alerts.

security response automation

A predefined and programmed action that is designed to automatically respond to or remediate a security event. These automations serve as detective or responsive security controls that help you implement AWS security best practices. Examples of automated response actions include modifying a VPC security group, patching an Amazon EC2 instance, or rotating credentials.

server-side encryption

Encryption of data at its destination, by the AWS service that receives it.

service control policy (SCP)

A policy that provides centralized control over permissions for all accounts in an organization in AWS Organizations. SCPs define guardrails or set limits on actions that an administrator can delegate to users or roles. You can use SCPs as allow lists or deny lists, to specify which services or actions are permitted or prohibited. For more information, see Service control policies in the AWS Organizations documentation.

service endpoint

The URL of the entry point for an AWS service. You can use the endpoint to connect programmatically to the target service. For more information, see AWS service endpoints in *AWS General Reference*.

service-level agreement (SLA)

An agreement that clarifies what an IT team promises to deliver to their customers, such as service uptime and performance.

service-level indicator (SLI)

A measurement of a performance aspect of a service, such as its error rate, availability, or throughput.

service-level objective (SLO)

A target metric that represents the health of a service, as measured by a service-level indicator.

shared responsibility model

A model describing the responsibility you share with AWS for cloud security and compliance. AWS is responsible for security *of* the cloud, whereas you are responsible for security *in* the cloud. For more information, see Shared responsibility model.

SIEM

See security information and event management system.

single point of failure (SPOF)

A failure in a single, critical component of an application that can disrupt the system.

SLA

See service-level agreement.

SLI

See service-level indicator.

SLO

See service-level objective.

split-and-seed model

A pattern for scaling and accelerating modernization projects. As new features and product releases are defined, the core team splits up to create new product teams. This helps scale your organization's capabilities and services, improves developer productivity, and supports rapid innovation. For more information, see Phased approach to modernizing applications in the AWS Cloud.

SPOF

See single point of failure.

star schema

A database organizational structure that uses one large fact table to store transactional or measured data and uses one or more smaller dimensional tables to store data attributes. This structure is designed for use in a data warehouse or for business intelligence purposes.

strangler fig pattern

An approach to modernizing monolithic systems by incrementally rewriting and replacing system functionality until the legacy system can be decommissioned. This pattern uses the analogy of a fig vine that grows into an established tree and eventually overcomes and replaces its host. The pattern was introduced by Martin Fowler as a way to manage risk when rewriting monolithic systems. For an example of how to apply this pattern, see Modernizing legacy Microsoft ASP.NET (ASMX) web services incrementally by using containers and Amazon API Gateway.

subnet

A range of IP addresses in your VPC. A subnet must reside in a single Availability Zone.

supervisory control and data acquisition (SCADA)

In manufacturing, a system that uses hardware and software to monitor physical assets and production operations.

symmetric encryption

An encryption algorithm that uses the same key to encrypt and decrypt the data.

synthetic testing

Testing a system in a way that simulates user interactions to detect potential issues or to monitor performance. You can use Amazon CloudWatch Synthetics to create these tests.

T

tags

Key-value pairs that act as metadata for organizing your AWS resources. Tags can help you manage, identify, organize, search for, and filter resources. For more information, see Tagging your AWS resources.

target variable

The value that you are trying to predict in supervised ML. This is also referred to as an *outcome variable*. For example, in a manufacturing setting the target variable could be a product defect.

task list

A tool that is used to track progress through a runbook. A task list contains an overview of the runbook and a list of general tasks to be completed. For each general task, it includes the estimated amount of time required, the owner, and the progress.

test environment

See environment.

training

To provide data for your ML model to learn from. The training data must contain the correct answer. The learning algorithm finds patterns in the training data that map the input data attributes to the target (the answer that you want to predict). It outputs an ML model that captures these patterns. You can then use the ML model to make predictions on new data for which you don't know the target.

transit gateway

A network transit hub that you can use to interconnect your VPCs and on-premises networks. For more information, see What is a transit gateway in the AWS Transit Gateway documentation.

trunk-based workflow

An approach in which developers build and test features locally in a feature branch and then merge those changes into the main branch. The main branch is then built to the development, preproduction, and production environments, sequentially.

trusted access

Granting permissions to a service that you specify to perform tasks in your organization in AWS Organizations and in its accounts on your behalf. The trusted service creates a service-linked role in each account, when that role is needed, to perform management tasks for you. For more information, see Using AWS Organizations with other AWS services in the AWS Organizations documentation.

tuning

To change aspects of your training process to improve the ML model's accuracy. For example, you can train the ML model by generating a labeling set, adding labels, and then repeating these steps several times under different settings to optimize the model.

two-pizza team

A small DevOps team that you can feed with two pizzas. A two-pizza team size ensures the best possible opportunity for collaboration in software development.

U

uncertainty

A concept that refers to imprecise, incomplete, or unknown information that can undermine the reliability of predictive ML models. There are two types of uncertainty: *Epistemic uncertainty* is caused by limited, incomplete data, whereas *aleatoric uncertainty* is caused by the noise and randomness inherent in the data. For more information, see the Quantifying uncertainty in deep learning systems guide.

undifferentiated tasks

Also known as *heavy lifting*, work that is necessary to create and operate an application but that doesn't provide direct value to the end user or provide competitive advantage. Examples of undifferentiated tasks include procurement, maintenance, and capacity planning.

upper environments

See environment.

V

vacuuming

A database maintenance operation that involves cleaning up after incremental updates to reclaim storage and improve performance.

version control

Processes and tools that track changes, such as changes to source code in a repository.

VPC peering

A connection between two VPCs that allows you to route traffic by using private IP addresses. For more information, see What is VPC peering in the Amazon VPC documentation.

vulnerability

A software or hardware flaw that compromises the security of the system.

W

warm cache

A buffer cache that contains current, relevant data that is frequently accessed. The database instance can read from the buffer cache, which is faster than reading from the main memory or disk.

warm data

Data that is infrequently accessed. When querying this kind of data, moderately slow queries are typically acceptable.

window function

A SQL function that performs a calculation on a group of rows that relate in some way to the current record. Window functions are useful for processing tasks, such as calculating a moving average or accessing the value of rows based on the relative position of the current row.

workload

A collection of resources and code that delivers business value, such as a customer-facing application or backend process.

workstream

Functional groups in a migration project that are responsible for a specific set of tasks. Each workstream is independent but supports the other workstreams in the project. For example, the portfolio workstream is responsible for prioritizing applications, wave planning, and collecting migration metadata. The portfolio workstream delivers these assets to the migration workstream, which then migrates the servers and applications.

WORM

See write once, read many.

WQF

See AWS Workload Qualification Framework.

write once, read many (WORM)

A storage model that writes data a single time and prevents the data from being deleted or modified. Authorized users can read the data as many times as needed, but they cannot change it. This data storage infrastructure is considered immutable.

Z

zero-day exploit

An attack, typically malware, that takes advantage of a zero-day vulnerability.

zero-day vulnerability

An unmitigated flaw or vulnerability in a production system. Threat actors can use this type of vulnerability to attack the system. Developers frequently become aware of the vulnerability as a result of the attack.

zombie application

An application that has an average CPU and memory usage below 5 percent. In a migration project, it is common to retire these applications.