aws

AWS Privacy Reference Architecture (AWS PRA)

AWS Prescriptive Guidance

AWS Prescriptive Guidance: AWS Privacy Reference Architecture (AWS PRA)

Table of Contents

AWS Privacy Reference Architecture (AWS PRA)

Amazon Web Services ([contributors](#))

March 2024 ([document history](#))

> We would love to hear from you. Please provide feedback on the AWS PRA by taking a [short survey](#).

Notices

This guide is provided for the purposes of information only. It isn't legal advice and shouldn't be relied on as legal advice. AWS encourages its customers to obtain appropriate advice on their implementation of privacy and data protection environments, and more generally, applicable laws relevant to their business.

Customers are responsible for making their own independent assessment of the information in this document. This document: (a) is for informational purposes only, (b) represents current AWS product offerings and practices, which are subject to change without notice, and (c) doesn't create any commitments or assurances from AWS and its affiliates, suppliers, or licensors. AWS products or services are provided "as is" without warranties, representations, or conditions of any kind, whether express or implied.

The responsibilities and liabilities of AWS to its customers are controlled by AWS agreements, and this document isn't part of, nor does it modify, any agreement between AWS and its customers.

Introduction

The AWS Privacy Reference Architecture (PRA) provides a set of guidelines specific to the design and configuration of privacy-supporting controls in AWS services. This guide can help you make decisions about people, process, and technology that help support privacy in the AWS Cloud.

The AWS shared responsibility model and privacy

In the AWS Cloud, you share responsibility for security and compliance with AWS. AWS is responsible for security *of* the cloud, which means that AWS is responsible protecting the infrastructure that runs all of the services offered in the AWS Cloud. You are responsible for

security *in* the cloud, which means that you are responsible for configuring and managing AWS services in accordance with security and privacy requirements. For more information, see the [AWS shared responsibility model](#).

AWS services provide capabilities that allow you to implement your own privacy controls in the cloud in order to support your privacy requirements. Your privacy responsibility varies based on many factors, including the AWS services and AWS Regions you choose, the integration of those services into your IT environment, and the laws and regulations applicable to your organization and workload.

When using AWS services, you maintain control over your content. Specifically, *content* is defined as software (including machine images), data, text, audio, video, or images that you or any end user transfer to us for processing, storage, or hosting by AWS services in connection with your account. It also includes any computational results that you or an end user derive by using AWS services. You are responsible for managing the following decisions, which are under your control:

- The data you choose to collect, store, or process on AWS
- The AWS services you use with the data
- The AWS Region where you collect, store, or process data
- The format and structure of your data and whether it's masked, anonymized, or encrypted
- How you define, store, rotate, and operate your cryptographic keys for encryption
- Who has access and when they have access to your data, and how those access rights are granted, managed, and revoked

Once you understand the AWS shared responsibility model and how it generally applies to operating in the cloud, you must determine how it applies to your use case. The AWS services that you choose to use determine the amount of configuration you must perform as part of your organization's privacy responsibilities. For example, a service such as Amazon Elastic Compute Cloud (Amazon EC2) is categorized as infrastructure as a service (IaaS). As such, if you use Amazon EC2, you must perform all of the necessary privacy configurations for guest operating systems and for the application software or utilities you install on your EC2 instances. When you use an abstracted service, such as Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB, AWS is responsible for the infrastructure layer, the operating system, and platforms. Your responsibility is to manage and classify the data and to configure the policies used to access the endpoints in order to store and retrieve data. For more information about how AWS helps you protect data and privacy, see [Data protection and privacy at AWS](#).

Understanding the AWS PRA

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The section describes the relationship between the AWS Privacy Reference Architecture (AWS PRA) and other AWS guidance. This section also reviews the general layout and structure of the example AWS multi-account environment in the AWS PRA.

This section contains the following topics:

- Using the AWS PRA and the AWS SRA
- AWS Organizations and the dedicated account structure
- Operationalizing AWS privacy services

Using the AWS PRA and the AWS SRA

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The AWS PRA provides patterns that customers have found helpful in planning foundational and application-level privacy controls for their infrastructure and workloads in AWS. The AWS Security Reference Architecture (AWS SRA) provides a set of guidelines for building an architecture that implements and supports the right set of security controls across your AWS landing zone and applications. In order to establish the privacy controls detailed in this guide, the AWS PRA assumes many of the same foundational guidelines and account structure that are described in the AWS SRA. The AWS PRA and AWS SRA detail many of the same key AWS services. This guide includes only brief descriptions of these services. You can learn more about these services and how they're used in a security context in the AWS SRA.

The AWS SRA can help you design, implement, and manage AWS security services so that they align with AWS recommended practices. You can use the AWS SRA as a standalone guide, or you can use the AWS SRA and AWS PRA as companion guides. Many of the security guidelines detailed

in the AWS SRA can be followed in tandem with the privacy controls that are detailed in the AWS PRA. Similar to security, there are foundational privacy considerations that can be helpful to make early in your AWS Cloud journey because these decisions can affect the design of the organization's account structure. For example, some questions you might consider include:

- How does my organization define personal data?

- Does my organization support applications that process personal data?

- What about applications that process other types of regulated data?

- What organization-level controls can I implement to keep my developers and cloud engineers as far away from personal data as possible?

- How do I segregate personal data from other types of data?

- What are my organization's cross-border data transfer requirements?

The answers to many of these questions can have implications for the design of your cloud environment, such as your AWS account structure, service control policies, and AWS Identity and Access Management (IAM) roles.

AWS Organizations and the dedicated account structure

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

AWS Organizations is an account management service that helps you centrally manage and govern multiple AWS accounts. Use of AWS Organizations is the basis of a well-architected, multi-account AWS environment. For more information, see Establishing your best practice AWS environment.

The following diagram shows the high-level account and organizational unit (OU) structure of the AWS PRA. For the most part, the organizational structure of the AWS PRA matches the organizational structure of the AWS SRA.
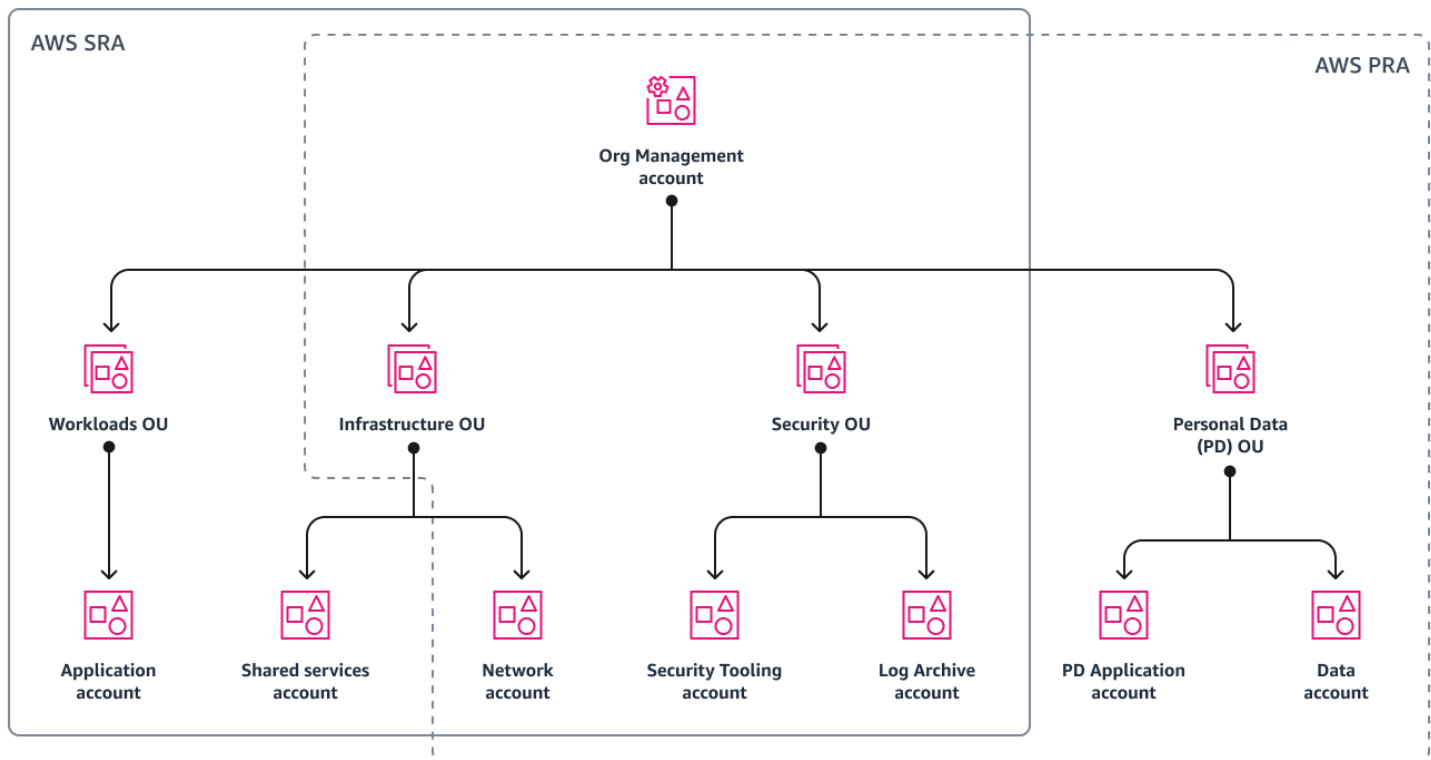
The deviations from the AWS SRA organization include:

- The AWS PRA adds the Personal Data (PD) OU, which is dedicated for collecting, storing, and processing personal data. This structural separation provides flexibility so that you can define specific, fine-grained controls to help protect personal data from unintended disclosure.

- In the Infrastructure OU, the AWS PRA doesn't currently include additional guidance for the Shared Services account that is described in the AWS SRA.

- The AWS PRA doesn't currently include additional guidance for the Workloads OU that is described in the AWS SRA. Applications that collect or process personal data are located in dedicated accounts in the PD OU.

You can use AWS Control Tower for overall foundational governance and automated deployment of security and privacy controls across your organization. If AWS Control Tower isn't in use today in your organization, you can still deploy many of the security and privacy controls in AWS Control Tower, such as service control policies and AWS Config rules, in their respective services.

You might find it helpful to consider the processing of personal data when you plan your account and OU structure, including an account segmentation strategy. You might need to consider the types of data you are processing for their unique use cases and applicable laws and regulations. For

example, cardholder data is protected under the Payment Card Industry Data Security Standard (PCI DSS), and protected health information might be subject to the Health Insurance Portability and Accountability Act (HIPAA). You might want to review which environments contain personal data and plan your segmentation strategy heavily on that. A typical account segmentation strategy can include dedicated AWS accounts that align to the software development lifecycle (SDLC), such as dedicated accounts for development, staging or quality assurance (QA), and production. A segmentation strategy such as this can be a critical component in the overall design discussion, and your OUs might need to align to your specific regulatory requirements.

Operationalizing AWS privacy services

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

For many, privacy is cross-cutting. Many different teams have a part to play, including regulatory, compliance, and engineering teams. When your organization has started to define the key people and policy components of your privacy program, you can map controls against a privacy compliance framework for consistent operations. A framework can serve as a rubric for implementing foundational and application-specific privacy controls for personal data in your AWS environment.

Regardless of the framework that customers use to categorize their privacy requirements, privacy compliance, privacy engineering, and application teams often need to work together to achieve implementation goals. For example, regulatory and compliance teams might provide the high-level requirements, and engineering and application teams configure AWS services and features to align to these requirements. Starting with a control framework can help you define more prescriptive organizational and technical controls.

When defining the technical controls of AWS services and features, another key decision is whether a control should apply to the entire organization, an OU, an account, or a specific resource. Some services and features are a great fit for implementing controls across your full AWS organization. For example, blocking public access to Amazon S3 buckets is a specific control that is preferably configured at the organization root rather than individually for each account. However, your retention policies might vary from application to application, which means that you might apply the control at the resource level.

To help you accelerate operationalizing privacy in your organization, AWS offers audit and compliance advisory services for your AWS workloads. For more information, contact AWS SAS.

The AWS Privacy Reference Architecture

We would love to hear from you. Please provide feedback on the AWS PRA by taking a [short survey](#).

The following diagram illustrates the AWS Privacy Reference Architecture (AWS PRA). This is an example of an architecture that connects many privacy-related AWS services and features. This architecture is built on a landing zone that is governed by AWS Control Tower.
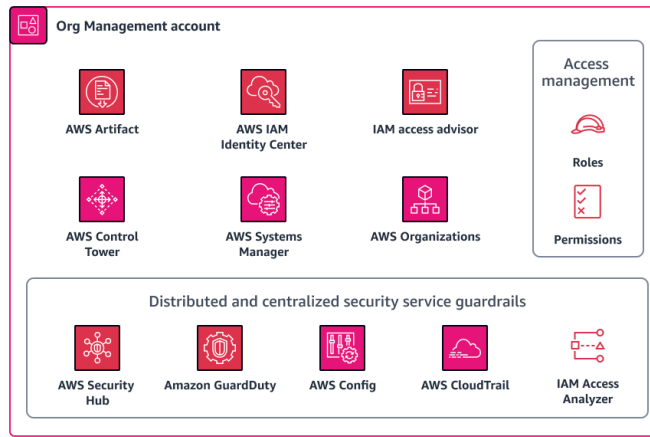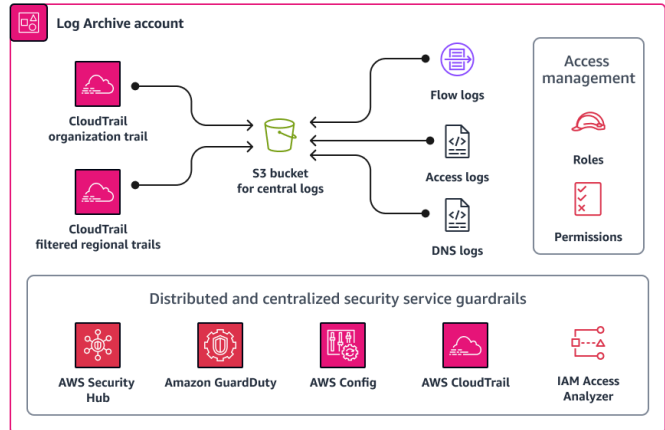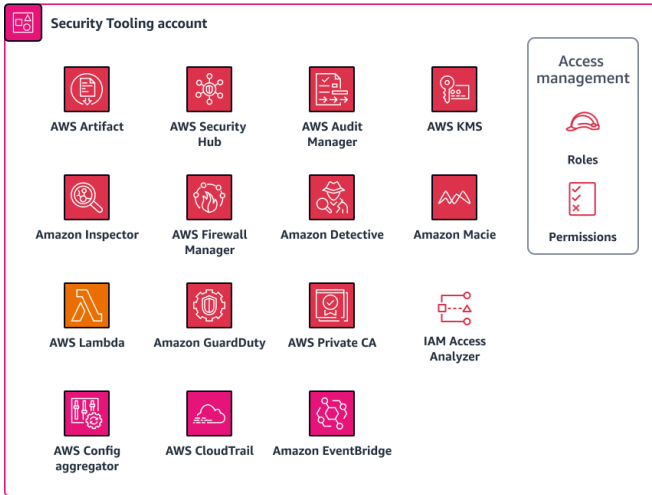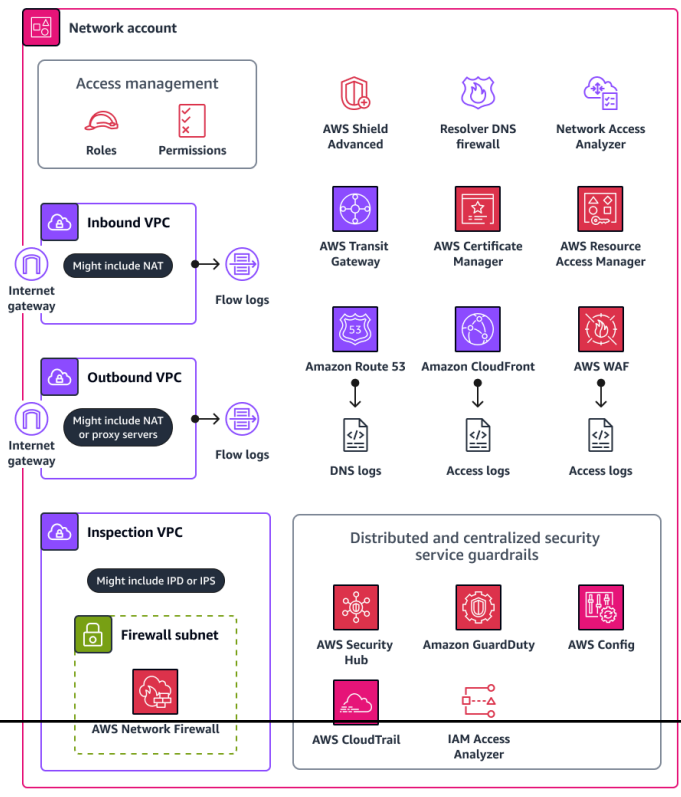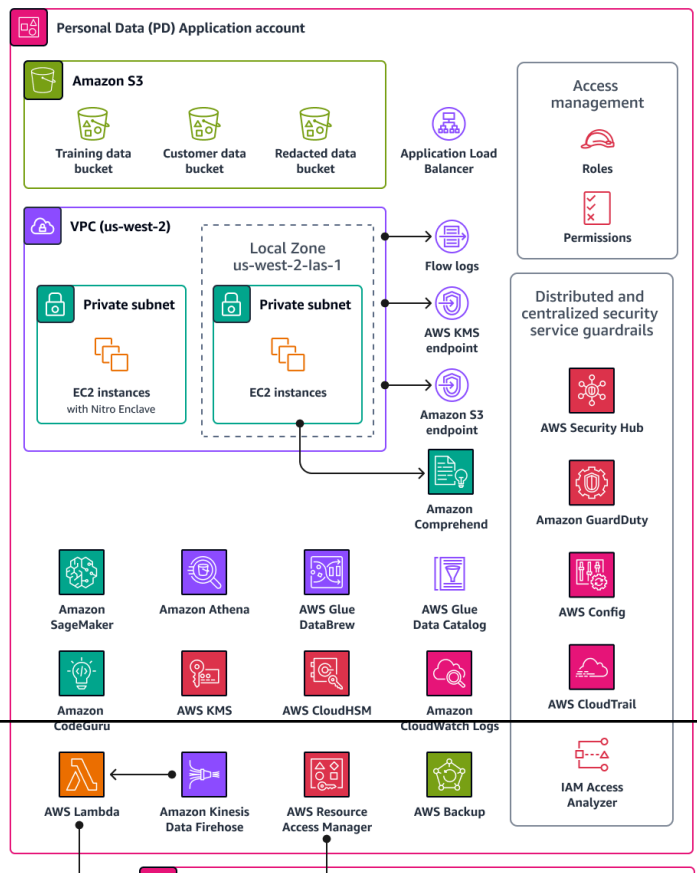
The AWS PRA includes a serverless web architecture that is hosted in the Personal Data (PD) Application account. The architecture in this account is an example workload that collects personal data directly from consumers. In this workload, users connect through a web tier. The web tier interacts with the application tier. This tier receives inputs from the web tier, processes and stores the data, allows authorized internal teams and third parties to access the data, and eventually archives and deletes the data when it's no longer required. The architecture is purposefully modular and event-driven in order to demonstrate many of the foundational privacy engineering techniques without delving into specific use cases, such as data lakes, containers, compute, or Internet of Things (IoT).

Next, this guide describes each account in the organization in detail. It discusses the privacy-related services and features, considerations and recommendations, and diagrams for each of the following accounts:

- Org Management account
- Security OU – Security Tooling account
- Security OU – Log Archive account
- Infrastructure OU – Network account
- Personal Data OU – PD Application account

Org Management account

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The Org Management account is primarily used to manage resource configuration drift for the foundational privacy controls across all of the accounts in your organization, which is managed by AWS Organizations. This account is also where you can deploy new member accounts consistently, with many of the same security and privacy controls. For more information about this account, see the AWS Security Reference Architecture (AWS SRA). The following diagram illustrates the AWS security and privacy services that are configured in the Org Management account.
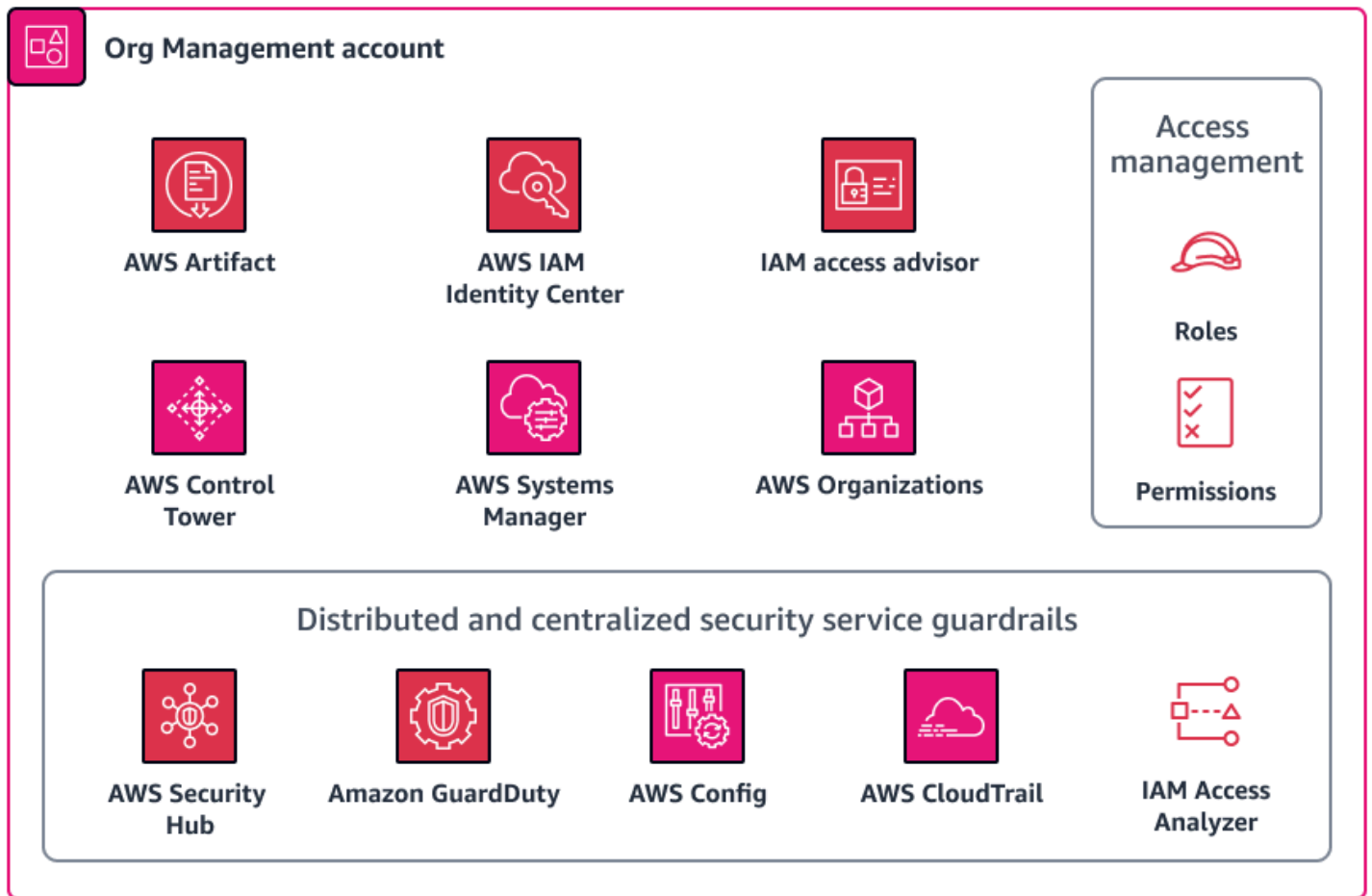
This section provides more detailed information about the following AWS services that are used in this account:

- AWS Artifact
- AWS Control Tower
- AWS Organizations

AWS Artifact

AWS Artifact can help you with audits by providing on-demand downloads of AWS security and compliance documents. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

This AWS service helps you understand the controls that you inherit from AWS and determine what controls may be remaining for you to implement in your environment. AWS Artifact provides access to AWS security and compliance reports, such as System and Organization Controls (SOC)

reports and payment card industry (PCI) reports. It also provides access to certifications from accreditation bodies across geographies and compliance verticals that validate the implementation and operating effectiveness of AWS controls. Using AWS Artifact, you can provide the AWS audit artifacts to your auditors or regulators as evidence of AWS security controls. The following reports might be useful to demonstrate the effectiveness of AWS privacy controls:

- **SOC 2 Type 2 Privacy report** – This report demonstrates the effectiveness of AWS controls for how personal data is collected, used, retained, disclosed, and disposed of. For more information, see the SOC FAQ.

- **SOC 3 Privacy report** – The SOC 3 Privacy report is a less detailed description of the SOC privacy controls, for general circulation.

- **ISO/IEC 27701:2019 certification report** ––ISO/IEC 27701:2019 describes requirements and guidelines to establish and continuously improve a privacy information management system (PIMS). This report details the scope of this certification and can serve as proof of AWS certification. For more information about this standard, see ISO/IEC 27701:2019 (ISO website).

AWS Control Tower

AWS Control Tower helps you set up and govern an AWS multi-account environment that follows prescriptive security best practices. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

In AWS Control Tower, you can also automate the deployment of a number of proactive, preventative, and detective controls, also known as *guardrails*, that align to your data residency and data protection requirements. For example, you can specify guardrails that limit the transfer of data to only approved AWS Regions. For even more granular control, you can choose from more than 17 guardrails that are designed to control data residency, such as *Disallow Amazon Virtual Private Network (VPN) connections*, *Disallow internet access for an Amazon VPC instance*, and *Deny access to AWS based on the requested AWS Region*. These guardrails consist of a number of AWS CloudFormation hooks, service control policies, and AWS Config rules that can be uniformly deployed across your organization. For more information, see Controls that enhance data residency protection in the AWS Control Tower documentation.

If you need to deploy privacy guardrails beyond data residency controls, AWS Control Tower includes a number of mandatory controls. These controls are deployed by default across every OU when you set up your landing zone. Many of these are preventative controls that are designed to

protect logs, such as *Disallow Deletion of Log Archive* and *Enable Integrity Validation for CloudTrail Log File*.

AWS Control Tower is also integrated with AWS Security Hub to provide detective controls. These controls are known as [Service-Managed Standard: AWS Control Tower](). You can use these controls to monitor for configuration drift of privacy-supporting controls, such as encryption at rest for Amazon Relational Database Service (Amazon RDS) database instances.

AWS Organizations

The AWS PRA uses AWS Organizations to centrally manage all accounts within the architecture. For more information, see [AWS Organizations and the dedicated account structure]() in this guide. In AWS Organizations, you can use service control policies (SCPs) and [management policies]() to help protect personal data and privacy.

Service control policies (SCPs)

[Service control policies (SCPs)]() are a type of organization policy that you can use to manage permissions in your organization. They provide centralized control over the maximum available permissions for AWS Identity and Access Management (IAM) roles and users in the target account, organization unit (OU), or entire organization. You can create and apply SCPs from the Org Management account.

You can use AWS Control Tower to deploy SCPs uniformly across your accounts. For more information about the data residency controls you can apply through AWS Control Tower, see [AWS Control Tower]() in this guide. AWS Control Tower includes a full complement of preventative SCPs. If AWS Control Tower isn't currently used in your organization, you can also deploy these controls manually.

Using SCPs to address data residency requirements

It's common to manage personal data residency requirements by storing and processing data within a specific geographic region. In order to verify that a jurisdiction's unique data residency requirements are met, we recommend that you work closely with your regulatory team to confirm your requirements. When these requirements have been determined, there are a number of AWS foundational privacy controls that can help support. For example, you can use SCPs to limit which AWS Regions can be used to process and store data. For a sample policy, see [Restrict data transfers across AWS Regions]() in this guide.

Using SCPs to restrict high-risk API calls

It is important to understand which security and privacy controls AWS is responsible for and which ones you are responsible for. For example, you are responsible for the results of API calls that could be made against the AWS services that you use. You are also responsible for understanding which of those calls could result in changes to your security or privacy posture. If you are concerned about maintaining a certain security and privacy posture, you can enable SCPs that deny certain API calls. These API calls may have implications, such as unintended disclosure of personal data or violations of specific cross-border data transfer. For example, you might want to prohibit the following API calls:

- Enabling public access to Amazon Simple Storage Service (Amazon S3) buckets
- Disabling Amazon GuardDuty or creating suppression rules for data exfiltration findings, such as the Trojan:EC2/DNSDataExfiltration finding
- Deleting AWS WAF data exfiltration rules
- Sharing Amazon Elastic Block Store (Amazon EBS) snapshots publicly
- Removing a member account from the organization
- Disassociating Amazon CodeGuru Reviewer from a repository

Management policies

Management policies in AWS Organizations can help you centrally configure and manage AWS services and their features. The types of management policy you choose determine how policies affect the OUs and accounts that inherit them. Tag policies are an example of a management policy in AWS Organizations that directly relates to privacy.

Using tag policies

Tags are key value pairs that help you manage, identify, organize, search for, and filter AWS resources. It can be useful to apply tags that distinguish the resources in your organization that handle personal data. The use of tags supports many of the privacy solutions in this guide. For example, you might want to apply a tag that indicates the general data classification of the data being processed or stored within the resource. You can write attribute-based access control (ABAC) policies that limit access to resources that have a particular tag or set of tags. For example, your policy might specify that the SysAdmin role can't access resources that have the dataclassification:4 tag. For more information and a tutorial, see Define permissions to access AWS resources based on tags in the IAM documentation. In addition, if your organization uses AWS Backup to apply data retention policies broadly across your backups in many accounts, you can apply a tag that puts that resource within scope for that backup policy.

Tag policies help you maintain consistent tags throughout your organization. In a tag policy, you specify rules that apply to resources when they are tagged. For example, you can require resources to be tagged with specific keys, such as `DataClassification` or `DataSteward`, and you can specify valid case treatments or values for keys. You can also use enforcement to prevent noncompliant tagging requests from completing.

When using tags as a core component of your privacy control strategy, consider the following:

- Consider the implications of placing personal data or other types of sensitive data within tag keys or values. When you contact AWS for technical assistance, AWS might analyze tags and other resource identifiers to help resolve the issue. In this case, you might want to deidentify tag values and then reidentify them by using a customer-controlled system, such as an IT service management (ITSM) system. AWS recommends not including personally identifiable information in tags.

- Consider that some tag values need to be made immutable (unmodifiable) to prevent circumvention of technical controls, such as ABAC conditions that rely on tags.

Security OU – Security Tooling account

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The Security Tooling account is dedicated to operating security and privacy foundational services, monitoring AWS accounts, and automating security and privacy alerting and response. For more information about this account, see the AWS Security Reference Architecture (AWS SRA). The following diagram illustrates the AWS security and privacy services that are configured in the Security Tooling account.
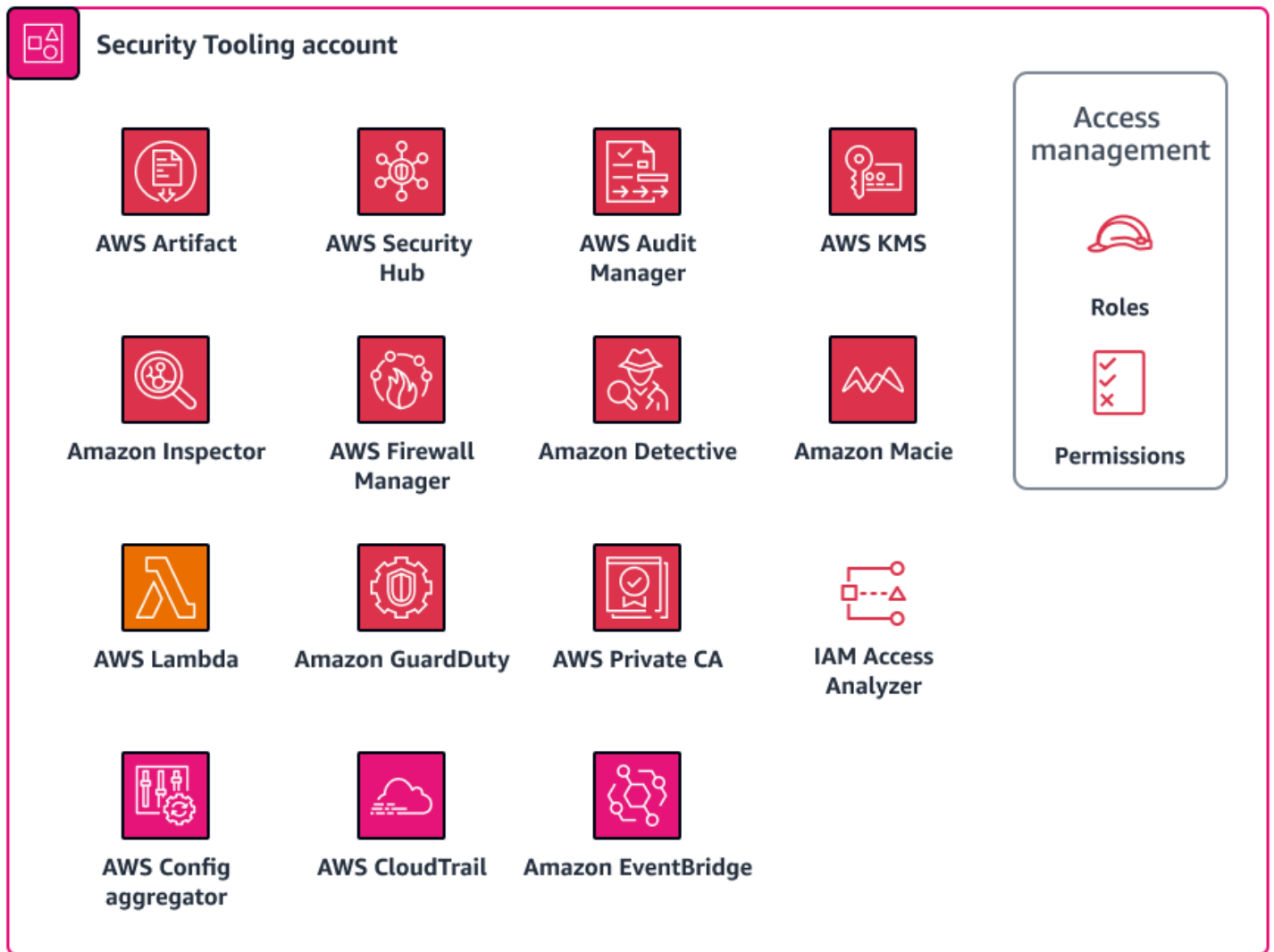
This section provides more detailed information about the following in this account:

- AWS CloudTrail
- AWS Config
- Amazon GuardDuty
- IAM Access Analyzer
- Amazon Macie

AWS CloudTrail

AWS CloudTrail helps you audit the overall API activity in your AWS account. Enabling CloudTrail in all AWS accounts and AWS Regions that store, process, or transmit personal data can help you

track the use and disclosure of this data. The AWS Security Reference Architecture recommends enabling an organization trail, which is a single trail that logs all events for all accounts in the organization. However, enabling this organization trail aggregates the multi-Region log data into a single Amazon Simple Storage Service (Amazon S3) bucket in the Log Archive account. For accounts that handle personal data, this can bring some additional design considerations. Log records might contain some references to personal data. To meet your data residency and data transfer requirements, you might need to reconsider aggregating cross-Region log data into a single Region where the S3 bucket is located. Your organization might consider which regional workloads should be included or excluded from the organization trail. For workloads that you decide to exclude from the organization trail, you could consider configuring a Region-specific trail that masks personal data. For more information about masking personal data, see the Amazon Data Firehose section of this guide. Ultimately, your organization might have a combination of organization trail and regional trails that aggregate into the centralized Log Archive account.

For more information about configuring a single-Region trail, see the instructions for using the AWS Command Line Interface (AWS CLI) or the console. When you create the organization trail, you can use an opt-in setting in AWS Control Tower, or you can create the trail directly in the CloudTrail console.

For more information on the overall approach and how to manage centralization of logs and data transfer requirements, see the Centralized log storage section in this guide. Whatever configuration you choose, you might want to separate trail management in the Security Tooling account from log storage in the Log Archive account, according to the AWS SRA. This design helps you create least-privilege access policies for those who need to manage logs and those who need to use the log data.

AWS Config

AWS Config provides a detailed view of the resources in your AWS account and how they're configured. It helps you identify how resources relate to one another and how their configurations have changed over time. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

In AWS Config, you can deploy conformance packs, which are sets of AWS Config rules and remediation actions. Conformance packs provide a general-purpose framework that is designed to enable privacy, security, operational, and cost-optimization governance checks by using managed or custom AWS Config rules. You can use this tool as a part of a larger set of automation tools

to track whether your AWS resource configurations comply with your own control framework requirements.

The [Operational Best Practices for NIST Privacy Framework v1.0](#) conformance pack is aligned to a number of privacy-related controls in the NIST Privacy Framework. Each AWS Config rule applies to a specific AWS resource type, and it relates to one or more NIST Privacy Framework controls. You can use this conformance pack to track privacy-related continuous compliance across resources in your accounts. The following are some of the rules included in this conformance pack:

- `no-unrestricted-route-to-igw` – This rule helps prevent data exfiltration on the data plane by continually monitoring VPC route tables for default `0.0.0.0/0` or `::/0` egress routes to an internet gateway. This helps you restrict where internet-bound traffic can be sent, especially if there are CIDR ranges that are known to be malicious.

- `encrypted-volumes` – This rule checks whether Amazon Elastic Block Store (Amazon EBS) volumes that are attached to Amazon Elastic Compute Cloud (Amazon EC2) instances are encrypted. If your organization has specific control requirements that pertain to the usage of AWS Key Management Service (AWS KMS) keys for protection of personal data, you can specify specific key IDs as a part of the rule to check that the volumes are encrypted with a specific AWS KMS key.

- `restricted-common-ports` – This rule checks whether Amazon EC2 security groups allow unrestricted TCP traffic to specified ports. Security groups can help you manage network access by providing stateful filtering of ingress and egress network traffic to AWS resources. Blocking ingress traffic from `0.0.0.0/0` to common ports, such as TCP 3389 and TCP 21, on your resources helps you restrict remote access.

AWS Config can be used for both proactive and reactive compliance checks of your AWS resources. In addition to considering the rules found in the conformance packs, you can incorporate these rules in both detective and proactive evaluation modes. This helps implement privacy checks earlier in your software development lifecycle because the application developers can start incorporating predeployment checks. For example, they can include hooks in their AWS CloudFormation templates that check the declared resource in the template against all privacy-related AWS Config rules that have proactive mode enabled. For more information, see [AWS Config Rules Now Support Proactive Compliance](#) (AWS blog post).

Amazon GuardDuty

AWS offers multiple services that might be used to store or process personal data, such as Amazon S3, Amazon Relational Database Service (Amazon RDS), or Amazon EC2 with Kubernetes. Amazon GuardDuty combines intelligent visibility with continuous monitoring to detect indicators that might be related to unintended disclosure of personal data. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

With GuardDuty, you can identify potentially malicious, privacy-related activity throughout an attack lifecycle. For example, GuardDuty can alert you about connections to blacklisted sites, unusual network port traffic or traffic volumes, DNS exfiltration, unexpected EC2 instance launches, and unusual ISP callers. You can also configure GuardDuty to stop alerts for trusted IP addresses from your own *trusted IP lists* and alert on known malicious IP addresses from your own *threat lists*.

As recommended in the AWS SRA, you can enable GuardDuty for all AWS accounts in your organization and configure the Security Tooling account as the GuardDuty delegated administrator. GuardDuty aggregates findings from across the organization into this single account. For more information, see Managing GuardDuty accounts with AWS Organizations. You can also consider identifying all privacy-related stakeholders in the incident response process, from detection and analysis to containment and eradication, and involving them in any incidents that might involve data exfiltration.

IAM Access Analyzer

Many customers want continual assurance that personal data is being shared appropriately with preapproved and intended third-party processors and no other entities. A data perimeter is a set of preventive guardrails designed to allow only trusted identities from expected networks to access trusted resources in your AWS environment. As you define controls for unintended and intended disclosure of personal data, you can define trusted identities, trusted resources, and expected networks.

With AWS Identity and Access Management Access Analyzer (IAM Access Analyzer), organizations can define an AWS account zone of trust and configure alerting for violations to that zone of trust. IAM Access Analyzer analyzes IAM policies to help identify and resolve unintended public or cross-account access to potentially sensitive resources. IAM Access Analyzer uses mathematical logic and inference to generate comprehensive findings for resources that can be accessed from outside an AWS account. Finally, for responding to and remediating overly-permissive IAM policies, you

can use IAM Access Analyzer to validate existing policies against IAM best practices and provide suggestions. IAM Access Analyzer can generate a least-privilege IAM policy that is based on an IAM principal's prior access activity. It analyzes CloudTrail logs and generates a policy that grants only the permissions required to continue performing those tasks.

For more information about how IAM Access Analyzer is used in a security context, see the AWS Security Reference Architecture.

Amazon Macie

Amazon Macie is a service that uses machine learning and pattern matching to discover sensitive data, provides visibility into data security risks, and helps you automate protections against those risks. Macie generates findings when it detects potential policy violations or issues with the security or privacy of your Amazon S3 buckets. Macie is another tool that organizations can use to implement automation in order to support compliance efforts. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

Macie can detect a large and growing list of sensitive data types, including personally identifiable information (PII), such as names, addresses, and other identifiable attributes. You can even create custom data identifiers in order to define detection criteria that reflects your organization's definition of personal data.

As your organization defines preventative controls for your Amazon S3 buckets that contain personal data, you can use Macie as a validation mechanism to provide continual reassurance of where your personal data lives and how it's protected. To start, enable Macie and configure automated sensitive data discovery. Macie continually analyzes objects in all of your S3 buckets, across accounts and AWS Regions. Macie generates and maintains an interactive heat map that depicts where personal data resides. The automated sensitive data discovery feature is designed to reduce costs and minimize the need to manually configure discovery jobs. You can build on top of the automated sensitive data discovery feature and use Macie to automatically detect new buckets or new data in existing buckets and then validate the data against the assigned data classification tags. Configure this architecture to notify the appropriate development and privacy teams about misclassified or unclassified buckets in a timely manner.

You can enable Macie for every account in your organization by using AWS Organizations. For more information, see Integrating and configuring an organization in Amazon Macie.

Security OU – Log Archive account

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The Log Archive account is where you centralize infrastructure, service, and application log types. For more information about this account, see the AWS Security Reference Architecture (AWS SRA). With a dedicated account for logs, you can apply consistent alerting across all log types and to confirm that incident responders can access an aggregate of these logs from one place. You can set up security controls and data retention policies all from one place as well, which can simplify the privacy operational overhead. The following diagram illustrates the AWS security and privacy services that are configured in the Log Archive account.
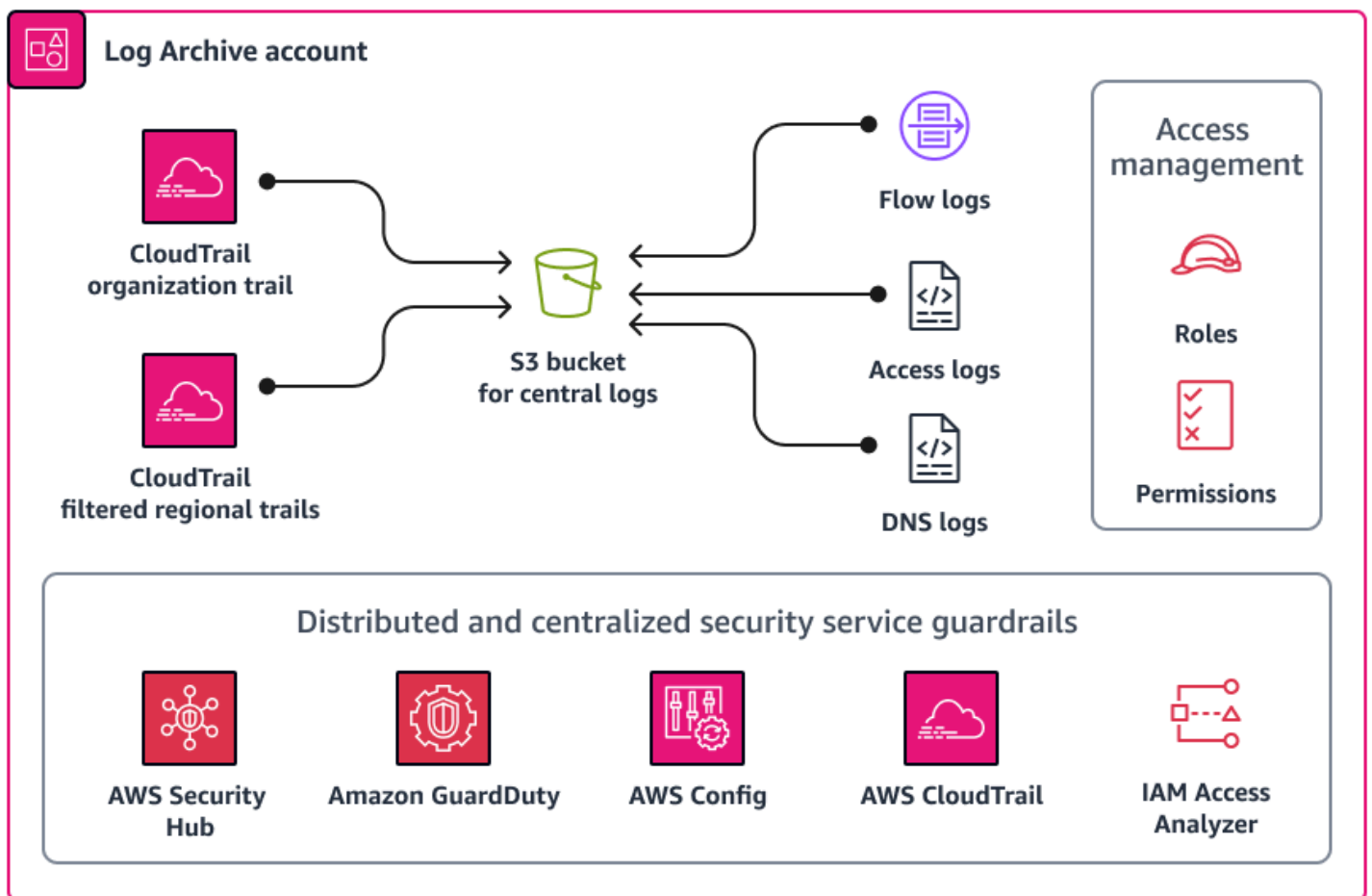
Centralized log storage

Log files (such as AWS CloudTrail logs) might contain information that could be considered personal data. Some organizations choose to use an organization trail in order to aggregate CloudTrail logs across AWS Regions and across accounts into one central location, for visibility purposes. For more information, see AWS CloudTrail in this guide. When implementing centralization of CloudTrail logs, the logs are typically stored in an Amazon Simple Storage Service (Amazon S3) bucket in a single Region.

Depending on your organization's definition of personal data and applicable regional privacy regulations, you might need to consider cross-border data transfers. If your organization needs to meet data transfer requirements for regional privacy regulations, the following options can help support:

1. If your organization is providing services in the AWS Cloud to data subjects in multiple countries, you might choose to aggregate all logs in the country that has the most stringent data residency requirements. For example, if you're operating in Germany and it has the most stringent requirements, you might aggregate data in an S3 bucket in the `eu-central-1` AWS Region so that data collected in Germany doesn't leave the borders of Germany. For this option, you can configure a single organization trail in CloudTrail that aggregates logs from across all accounts and AWS Regions into the target Region.

2. Redact the personal data that needs to stay in the AWS Region before the data is copied and aggregated to another region. For example, you can mask the personal data in the application's host Region before you transfer the logs to a different Region. For more information about masking personal data, see the Amazon Data Firehose section of this guide.

Work with your legal counsel to determine which personal data is in scope and what AWS Region-to-Region transfers are allowed.

Infrastructure OU – Network account

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

In the Network account, you manage the networking between your virtual private clouds (VPCs) and the broader internet. In this account, you can implement broad disclosure control mechanisms

by using AWS WAF, use AWS Resource Access Manager (AWS RAM) to share VPC subnets and AWS Transit Gateway attachments, and use Amazon CloudFront to support targeted service usage. For more information about this account, see the AWS Security Reference Architecture (AWS SRA). The following diagram illustrates the AWS security and privacy services that are configured in the Network account.
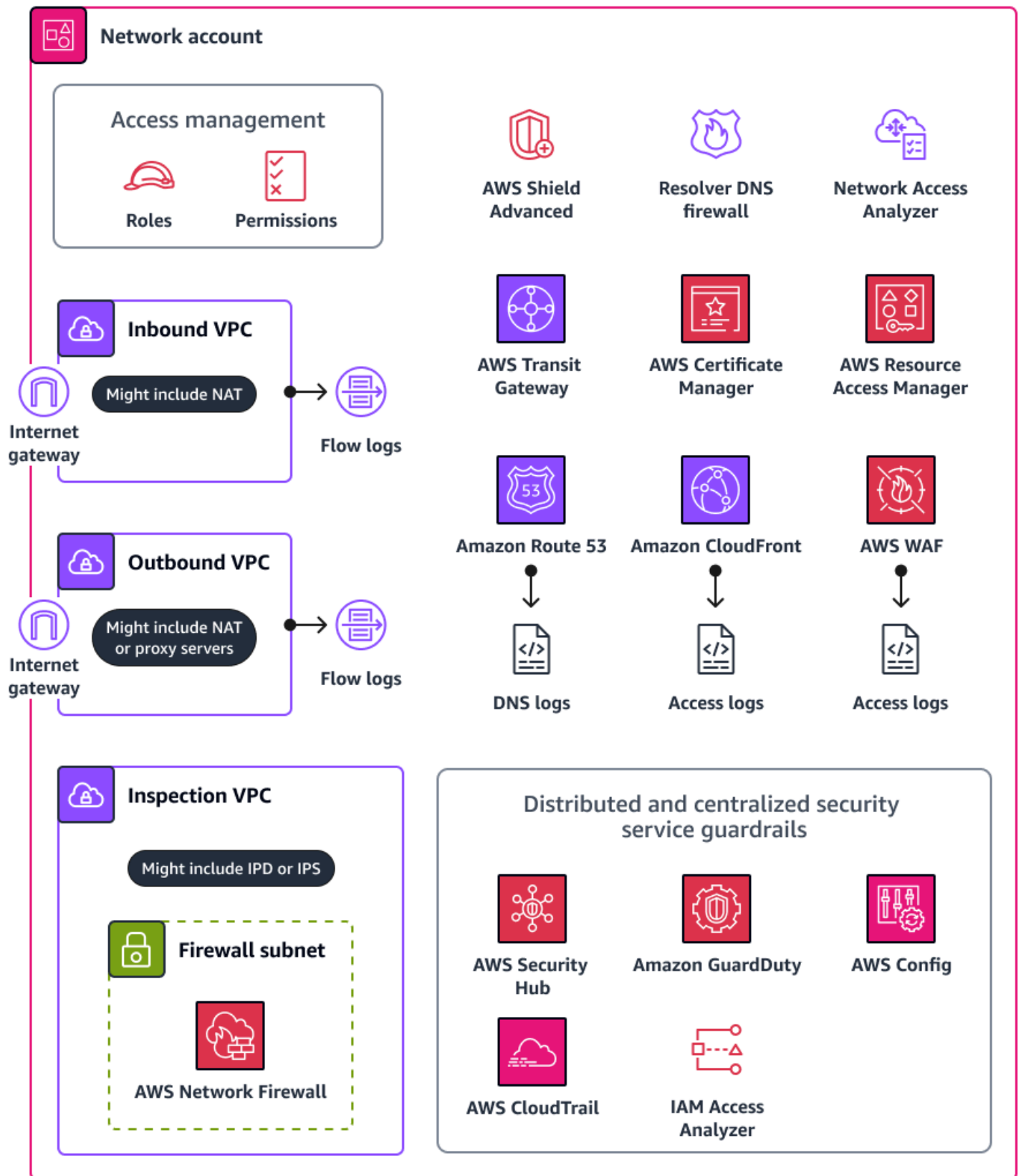
This section provides more detailed information about the following AWS services that are used in this account:

- Amazon CloudFront

- AWS Resource Access Manager

- AWS Transit Gateway

- AWS WAF

Amazon CloudFront

Amazon CloudFront supports geographic restrictions for frontend applications and file hosting. CloudFront can deliver content through a worldwide network of data centers that are called *edge locations*. When a user requests content that you're serving with CloudFront, the request is routed to the edge location that provides the lowest latency. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

You can use CloudFront geographic restrictions to prevent users in specific geographic locations from accessing content that you are distributing through a CloudFront distribution. For more information and configuration options for geographic restrictions, see Restricting the geographic distribution of your content in the CloudFront documentation.

You can also configure CloudFront to generate access logs that contain detailed information about every user request that CloudFront receives. For more information, see Configuring and using standard logs (access logs) in the CloudFront documentation. Finally, if CloudFront is configured to cache content at a series of edge locations, you might consider where caching occurs. For some organizations, cross-Regional caching might be subject to cross-border data transfer requirements.

AWS Resource Access Manager

AWS Resource Access Manager (AWS RAM) helps you securely share your resources across AWS accounts to reduce operational overhead and provide visibility and auditability. With AWS RAM, organizations can restrict which AWS resources can be shared with other AWS accounts in their organization or with third-party accounts. For more information, see Shareable AWS resources. In the Network account, you can use AWS RAM to share VPC subnets and transit gateway connections. If you use AWS RAM to share a data plane connection with another AWS account, consider establishing processes to check that the connections are made to preapproved AWS Regions.

In addition to sharing VPCs and transit gateway connections, AWS RAM can be used to share resources that don't support IAM resource-based policies. For a workload hosted in the Personal Data OU, you can use AWS RAM to access personal data that is located in a separate AWS account. For more information, see AWS Resource Access Manager in the *Personal Data OU – PD Application account* section.

AWS Transit Gateway

If you want to deploy AWS resources that collect, store, or process personal data in AWS Regions that align with your organizational data residency requirements and you have the appropriate technical safeguards, consider implementing guardrails to prevent unapproved cross-border data flows on the control and data planes. On the control plane, you can limit Region usage and, as a result, cross-Region data flows by using IAM and service control policies.

There are multiple options for controlling cross-Region data flows on the data plane. For example, you can use route tables, VPC peering, and AWS Transit Gateway attachments. AWS Transit Gateway is a central hub that connects virtual private clouds (VPCs) and on-premises networks. As a part of your larger AWS landing zone, you can consider the various ways data can traverse AWS Regions, including out through internet gateways, through direct VPC-to-VPC peering, and through inter-Region peering with AWS Transit Gateway. For example, you can do the following in AWS Transit Gateway:

- Confirm that the east-west and north-south connections between your VPCs and on-premises environments are aligned with your privacy requirements.
- Configure VPC settings according to your privacy requirements.
- Use a service control policy in AWS Organizations and IAM policies to help prevent modifications to your AWS Transit Gateway and Amazon Virtual Private Cloud (Amazon VPC) configurations. For a sample service control policy, see Restrict changes to VPC configurations in this guide.

AWS WAF

To help prevent unintended disclosure of personal data, you can deploy a defense-in-depth approach for your web applications. You can build input validation and rate limiting into your application, but AWS WAF can serve as another line of defense. AWS WAF is a web application firewall that helps you monitor HTTP and HTTPS requests that are forwarded to your protected web application resources. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

With AWS WAF, you can define and deploy rules that inspect for specific criteria. The following activities might be associated with unintended disclosure of personal data:

- Traffic from unknown or malicious IP addresses or geographical locations

- Open Worldwide Application Security Project (OWASP) Top 10 attacks, including exfiltration-related attacks such as SQL injection

- High rates of requests

- General bot traffic

- Content scrapers

You can deploy AWS WAF rule groups that are managed by AWS. Some managed rule groups for AWS WAF can be used to detect threats to privacy and personal data, for example:

- SQL database – This rule group contains rules designed to block request patterns associated with exploitation of SQL databases, such as SQL injection attacks. Consider this rule group if your application interfaces with a SQL database.

- Known bad inputs – This rule group contains rules designed to block request patterns that are known to be invalid and are associated with exploitation or discovery of vulnerabilities.

- Bot Control – This rule group contains rules designed to manage requests from bots, which can consume excess resources, skew business metrics, cause downtime, and perform malicious activities.

- Account takeover prevention (ATP) – This rule group contains rules designed to prevent malicious account takeover attempts. This rule group inspects the login attempts sent to your application's login endpoint.

Personal Data OU – PD Application account

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

The Personal Data (PD) Application account is where your organization hosts services that collect and process personal data. Specifically, you might store what you define as personal data in this account. The AWS PRA demonstrates a number of example privacy configurations through a multi-

tier serverless web architecture. When it comes to operating workloads across an AWS landing zone, privacy configurations should not be considered one-size-fits-all solutions. For example, your goal might be to understand the underlying concepts, how they can enhance privacy, and how your organization can apply solutions to your particular use cases and architectures.

For AWS accounts in your organization that collect, store, or process personal data, you can use AWS Organizations and AWS Control Tower to deploy foundational and repeatable guardrails. Establishing a dedicated organizational unit (OU) for these accounts is critical. For example, you might want to apply data residency guardrails to only a subset of accounts where data residency is a core design consideration. For many organizations, these are the accounts that store and process personal data.

Your organization might support a dedicated Data account, which is where you store the authoritative source of your personal datasets. An authoritative data source is a location where you store the primary version of data, which might be considered the most reliable and accurate version of the data. For example, you might copy the data from the authoritative data source to other locations, such as Amazon Simple Storage Service (Amazon S3) buckets in the PD Application account that are used to store training data, a subset of customer data, and redacted data. By taking this multi-account approach to separate complete and definitive personal datasets in the Data account from the downstream consumer workloads in the PD Application account, you can reduce the scope of impact in the event of unauthorized access to your accounts.

The following diagram illustrates the AWS security and privacy services that are configured in the PD Application and Data accounts.

Personal Data (PD) Application account

Amazon S3

Training data bucket

Customer data bucket

Redacted data bucket

Application Load Balancer

VPC (us-west-2)

Local Zone us-west-2-las-1

Private subnet

EC2 instances
with Nitro Enclave

Private subnet

EC2 instances

Flow logs

AWS KMS endpoint

Amazon S3 endpoint

Amazon Comprehend

Access management

Roles

Permissions

Distributed and centralized security service guardrails

AWS Security Hub

Amazon GuardDuty

AWS Config

AWS CloudTrail

IAM Access Analyzer

Amazon SageMaker

Amazon Athena

AWS Glue DataBrew

AWS Glue Data Catalog

Amazon CodeGuru

AWS KMS

AWS CloudHSM

Amazon CloudWatch Logs

AWS Lambda

Amazon Kinesis Data Firehose

AWS Resource Access Manager

AWS Backup

Data account

Access management

This section provides more detailed information about the following AWS services that are used in these accounts:

- Amazon Athena
- Amazon CloudWatch Logs
- Amazon CodeGuru Reviewer
- Amazon Comprehend
- Amazon Data Firehose
- AWS Glue
- AWS Key Management Service
- AWS Local Zones
- AWS Nitro Enclaves
- AWS PrivateLink
- AWS Resource Access Manager
- Amazon SageMaker AI
- AWS features that help manage the data lifecycle
- AWS services and features that help segment data

Amazon Athena

You can also consider data query limitation controls to meet your privacy goals. Amazon Athena is an interactive query service that helps you analyze data directly in Amazon S3 by using standard SQL. You don't have to load the data into Athena; it works directly with the data stored in S3 buckets.

A common use case for Athena is providing data analytics teams with tailored and sanitized datasets. If the datasets contain personal data, you can sanitize the dataset by masking entire columns of personal data that provide little value to the data analytics teams. For more information, see Anonymize and manage data in your data lake with Amazon Athena and AWS Lake Formation (AWS blog post).

If your data transformation approach requires additional flexibility outside of the supported functions in Athena, you can define custom functions, called user-defined functions (UDF). You can invoke UDFs in a SQL query submitted to Athena, and they run on AWS Lambda. You can use UDFs in SELECT and FILTER SQL queries, and you can invoke multiple UDFs in the same query.

For privacy, you can create UDFs that perform specific types of data masking, such as showing only the last four characters of every value in a column.

Amazon CloudWatch Logs

Amazon CloudWatch Logs helps you centralize the logs from all your systems, applications, and AWS services so you can monitor them and archive them securely. In CloudWatch Logs, you can use a data protection policy for new or existing log groups to help minimize the risk of disclosure of personal data. Data protection policies can detect sensitive data, such as personal data, in your logs. The data protection policy can mask that data when users access the logs through the AWS Management Console. When users require direct access to the personal data, according to the overall purpose specification for your workload, you can assign `logs:Unmask` permissions for those users. You can also create an account-wide data protection policy and apply this policy consistently across all accounts in your organization. This configures masking by default for all current and future log groups in CloudWatch Logs. We also recommend that you enable audit reports and send them to another log group, an Amazon S3 bucket, or Amazon Data Firehose. These reports contain a detailed record of data protection findings across each log group.

Amazon CodeGuru Reviewer

For both privacy and security, it's vital to many organizations that they support continuous compliance during both deployment and post-deployment phases. The AWS PRA includes proactive controls in deployment pipelines for applications that process personal data. Amazon CodeGuru Reviewer can detect potential defects that might expose personal data in Java, JavaScript, and Python code. It offers suggestions to developers for improving the code. CodeGuru Reviewer can identify defects across a wide range of security, privacy, and general best practices. For more information, see the Amazon CodeGuru Detector Library. It's designed to work with multiple source providers, including AWS CodeCommit, Bitbucket, GitHub, and Amazon S3. Some of the privacy-related defects that CodeGuru Reviewer can detect include:

- SQL injection
- Unsecured cookies
- Missing authorization
- Client-side AWS KMS re-encryption

Amazon Comprehend

Amazon Comprehend is a natural-language processing (NLP) service that uses machine learning to uncover valuable insights and connections in English text documents. Amazon Comprehend can detect and redact personal data in structured, semi-structured, or unstructured text documents. For more information, see Personally identifiable information (PII) in the Amazon Comprehend documentation.

You can use the AWS SDKs and Amazon Comprehend API to integrate Amazon Comprehend with many applications. An example is using Amazon Comprehend to detect and redact personal data with Amazon S3 Object Lambda. Organizations can use S3 Object Lambda to add custom code to Amazon S3 GET requests to modify and process data as it's returned to an application. S3 Object Lambda can filter rows, dynamically resize images, redact personal data, and more. Powered by AWS Lambda functions, the code runs on infrastructure that is fully managed by AWS, which eliminates the need to create and store derivative copies of your data or to run proxies. You don't need to change your applications to transform objects with S3 Object Lambda. You can use the `ComprehendPiiRedactionS3Object` Lambda function in AWS Serverless Application Repository to redact personal data. This function uses Amazon Comprehend to detect personal data entities and redacts those entities by replacing them with asterisks. For more information, see Detecting and redacting PII data with S3 Object Lambda and Amazon Comprehend in the Amazon S3 documentation.

Because Amazon Comprehend has many options for application integration through AWS SDKs, you can use Amazon Comprehend to identify personal data in many different places where you collect, store, and process data. You can use Amazon Comprehend ML capabilities to detect and redact personal data in application logs (AWS blog post), customer emails, support tickets, and more. The architecture diagram for the PD Application account shows how you can perform this function for application logs on Amazon EC2. Amazon Comprehend offers two redaction modes:

- `REPLACE_WITH_PII_ENTITY_TYPE` replaces each PII entity with its types. For example, **Jane Doe** would be replaced with **NAME**.

- MASK replaces the characters in PII entities with a character of your choice (!, #, $, %, &, , or @). *For example, **Jane Doe** could be replaced with ****** *****.*

Amazon Data Firehose

[Amazon Data Firehose](#) can be used to capture, transform, and load streaming data into downstream services, such as Amazon Managed Service for Apache Flink or Amazon S3. Firehose is often used to transport large quantities of streaming data, such as application logs, without having to build processing pipelines from the ground up.

You can use Lambda functions to perform customized or built-in processing before the data is sent downstream. For privacy, this capability supports data minimization and cross-border data transfer requirements. For example, you can use Lambda and Firehose to transform multi-Region log data before it's centralized in the Log Archive account. For more information, see [Biogen: Centralized Logging Solution for Multi Accounts](#) (YouTube video). In the PD Application account, you configure Amazon CloudWatch and AWS CloudTrail to push logs to a Firehose delivery stream. A Lambda function transforms the logs and sends them to a central S3 bucket in the Log Archive account. You can configure the Lambda function to mask specific fields that contain personal data. This helps prevent the transfer of personal data across AWS Regions. By using this approach, the personal data is masked before the transfer and centralization, rather than after. For applications in jurisdictions that aren't subject to cross-border transfer requirements, it is typically more operationally efficient and cost effective to aggregate logs through the organizational trail in CloudTrail. For more information, see [AWS CloudTrail](#) in the *Security OU – Security Tooling account* section of this guide.

AWS Glue

Maintaining datasets that contain personal data is a key component of [Privacy by Design](#). An organization's data might exist in structured, semi-structured, or unstructured forms. Personal datasets without structure can make it difficult to perform a number of privacy-enhancing operations, including data minimization, tracking down data attributed to a single data subject as a part of a data subject request, ensuring consistent data quality, and overall segmentation of datasets. [AWS Glue](#) is a fully managed extract, transform, and load (ETL) service. It can help you categorize, clean, enrich, and move data between data stores and data streams. AWS Glue features are designed to help you discover, prepare, structure, and combine datasets for analytics, machine learning, and application development. You can use AWS Glue to create a predictable and common structure on top of your existing datasets. AWS Glue Data Catalog, AWS Glue DataBrew, and AWS Glue Data Quality are AWS Glue features that can help support your organization's privacy requirements.

AWS Glue Data Catalog

AWS Glue Data Catalog helps you establish maintainable datasets. The Data Catalog contains references to data that is used as sources and targets for extract, transform, and load (ETL) jobs in AWS Glue. Information in the Data Catalog is stored as metadata tables, and each table specifies a single data store. You run an AWS Glue crawler to take inventory of the data in a variety of data store types. You add built-in and custom classifiers to the crawler, and these classifiers infer the data format and schema of the personal data. The crawler then writes the metadata to the Data Catalog. A centralized metadata table can make it easier to respond to data subject requests (such as right to erasure) because it adds structure and predictability across disparate sources of personal data in your AWS environment. For a comprehensive example of how to use Data Catalog to automatically respond to these requests, see Handling data erasure requests in your data lake with Amazon S3 Find and Forget (AWS blog post). Finally, if your organization is using AWS Lake Formation to administer and provide fine-grained access across databases, tables, rows, and cells, Data Catalog is a key component. Data Catalog provides cross-account data sharing and helps you use tag-based access control to manage your data lake at scale (AWS blog post).

AWS Glue DataBrew

AWS Glue DataBrew helps you clean and normalize data, and it can perform transformations on the data, such as removing or masking personally identifiable information and encrypting sensitive data fields in data pipelines. You can also visually map the lineage of your data to understand the various data sources and transformation steps that the data has been through. This feature becomes increasingly important as your organization works to better understand and track personal data provenance. DataBrew helps you mask personal data during data preparation. You can detect personal data as part of a data profiling job and gather statistics, such as the number of columns that might contain personal data and potential categories. You can then use built-in reversible or irreversible data transformation techniques, including substitution, hashing, encryption, and decryption, all without writing any code. You can then use the cleaned and masked datasets downstream for analytics, reporting, and machine learning tasks. Some of the data masking techniques available in DataBrew include:

- **Hashing** – Apply hash functions to the column values.

- **Substitution** – Replace personal data with other, authentic-looking values.

- **Nulling out or deletion** – Replace a particular field with a null value, or delete the column.

- **Masking out** – Use character scrambling, or mask certain portions in the columns.

The following are the available encryption techniques:

- **Deterministic encryption** – Apply deterministic encryption algorithms to the column values. Deterministic encryption always produces the same ciphertext for a value.

- **Probabilistic encryption** – Apply probabilistic encryption algorithms to the column values. Probabilistic encryption produces different ciphertext each time that it's applied.

For a complete list of provided personal data transformation recipes in DataBrew, see Personally identifiable information (PII) recipe steps.

AWS Glue Data Quality

AWS Glue Data Quality helps you automate and operationalize the delivery of high-quality data across data pipelines, proactively, before they are delivered to your data consumers. AWS Glue Data Quality provides statistical analysis of data quality issues across your data pipelines, can trigger alerts in Amazon EventBridge, and can make quality rule recommendations for remediation. AWS Glue Data Quality also supports rules creation with a domain-specific language so that you can create custom data quality rules.

AWS Key Management Service

AWS Key Management Service (AWS KMS) helps you create and control cryptographic keys to help protect your data. AWS KMS uses hardware security modules to protect and validate AWS KMS keys under the FIPS 140-2 Cryptographic Module Validation Program. For more information about how this service is used in a security context, see the AWS Security Reference Architecture.

AWS KMS integrates with most AWS services that offer encryption, and you can use KMS keys in your applications that process and store personal data. You can use AWS KMS to help support a variety of your privacy requirements and safeguard personal data, including:

- Using customer managed keys for greater control over strength, rotation, expiration, and other options.

- Using dedicated customer managed keys to protect personal data and secrets that allow access to personal data.

- Defining data classification levels and designating at least one dedicated customer managed key per level. For example, you might have one key to encrypt operational data and another to encrypt personal data.

- Preventing unintended cross-account access to KMS keys.

- Storing KMS keys within the same AWS account as the resource to be encrypted.

- Implementing separation of duties for KMS key administration and usage. For more information, see How to use KMS and IAM to enable independent security controls for encrypted data in S3 (AWS blog post).

- Enforcing automatic key rotation through preventative and reactive guardrails.

By default, KMS keys are stored and can be used only in the Region where they were created. If your organization has specific requirements for data residency and sovereignty, consider whether multi-Region KMS keys are appropriate for your use case. Multi-Region keys are special-purpose KMS keys in different AWS Regions that can be used interchangeably. The process of creating a multi-Region key moves your key material across AWS Region boundaries within AWS KMS, so this lack of regional isolation might not be compatible with your organization's compliance goals. One way to solve for this is to use a different type of KMS key, such as a Region-specific customer managed key.

AWS Local Zones

If you need to comply with data residency requirements, you can deploy resources that store and process personal data in specific AWS Regions to support these requirements. You can also use AWS Local Zones, which helps you place compute, storage, database, and other select AWS resources close to large population and industry centers. A Local Zone is an extension of an AWS Region that is in geographic proximity to a large metropolitan area. You can place specific types of resources within a Local Zone, near the Region to which the Local Zone corresponds. Local Zones can help you meet data residency requirements when a Region is unavailable within the same legal jurisdiction. When you use Local Zones, consider the data residency controls that are deployed within your organization. For example, you might need a control to prevent data transfers from a specific Local Zone to another Region. For more information about how to use SCPs to maintain cross-border data transfer guardrails, see Best Practices for managing data residency in AWS Local Zones using landing zone controls (AWS blog post).

AWS Nitro Enclaves

Consider your data segmentation strategy from a processing perspective, such as processing personal data with a compute service such as Amazon Elastic Compute Cloud (Amazon EC2). Confidential computing as a part of a larger architecture strategy can help you isolate personal data processing in an isolated, protected, and trusted CPU enclave. Enclaves are separate, hardened, and highly-constrained virtual machines. AWS Nitro Enclaves is an Amazon EC2 feature

that can help you create these isolated compute environments. For more information, see [The Security Design of the AWS Nitro System](#) (AWS whitepaper).

Nitro Enclaves deploy a kernel that is separated from the parent instance's kernel. The parent instance's kernel doesn't have access to the enclave. Users can't SSH or remotely access the data and applications in the enclave. Applications that process personal data can be embedded in the enclave and configured to use the enclave's [Vsock](#), the socket that facilitates communication between the enclave and the parent instance.

One use case where Nitro Enclaves can be useful is joint processing between two data processors that are in separate AWS Regions and that might not trust each other. The following image shows how you can use an enclave for central processing, a KMS key for encrypting the personal data before it's sent to the enclave, and an AWS KMS key policy that verifies that the enclave requesting decryption has the unique measurements in its attestation document. For more information and instructions, see [Using cryptographic attestation with AWS KMS](#). For a sample key policy, see [Require attestation to use an AWS KMS key](#) in this guide.
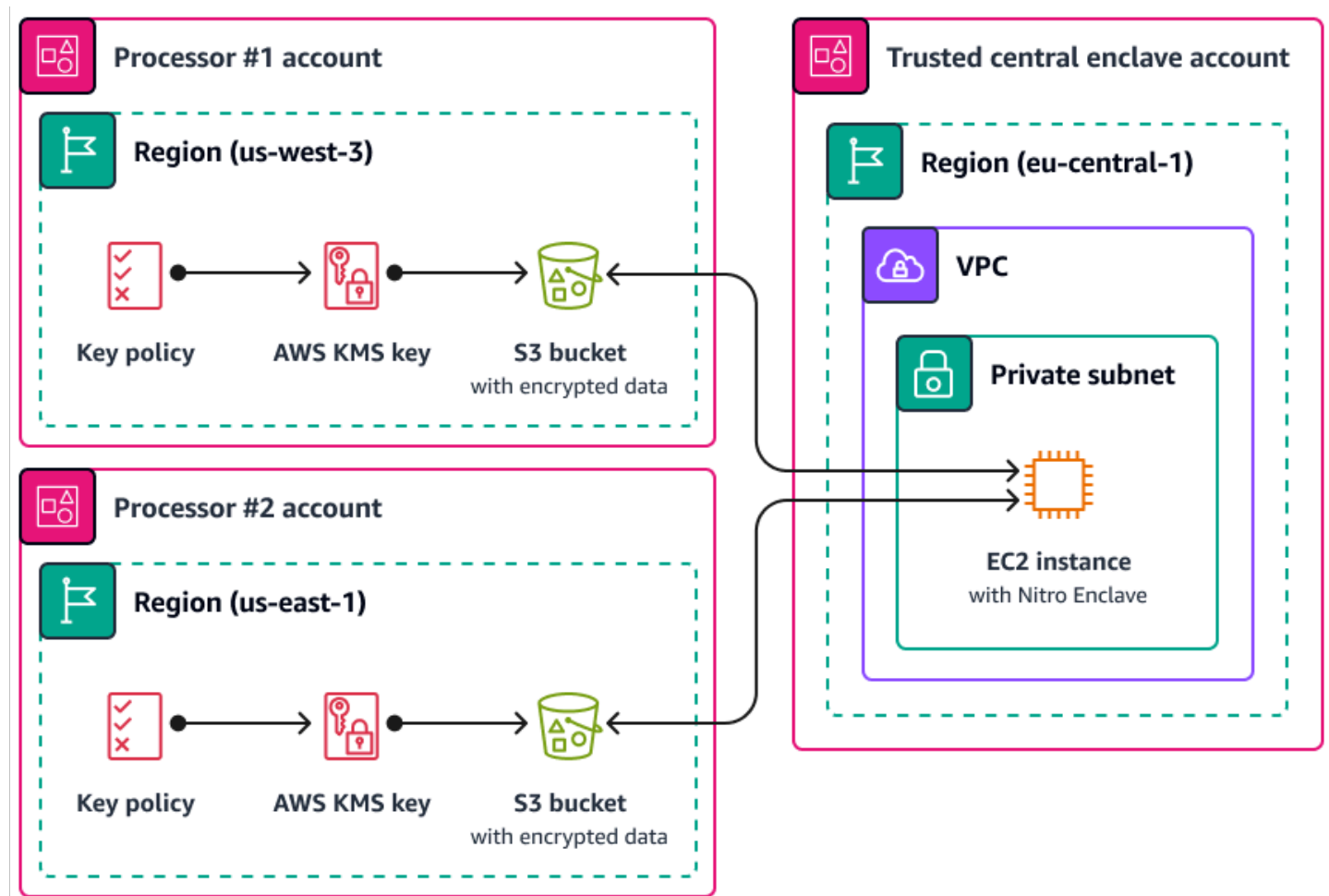
With this implementation, only the respective data processors and the underlying enclave have access to the plaintext personal data. The only place the data is exposed, outside of the respective data processors' environments, is in the enclave itself, which is designed to prevent access and tampering.

AWS PrivateLink

Many organizations want to limit the exposure of personal data to untrusted networks. For example, if you want to enhance the privacy of your overall application architecture design, you can segment networks based on data sensitivity (similar to the logical and physical separation of datasets that is discussed in the AWS services and features that help segment data section). AWS PrivateLink helps you create unidirectional, private connections from your virtual private clouds (VPCs) to services outside of the VPC. Using AWS PrivateLink, you can set up dedicated private connections to the services that store or process personal data in your environment; there is no need to connect to public endpoints and transfer this data over untrusted public networks. When you enable AWS PrivateLink service endpoints for the in-scope services, there is no need for an internet gateway, NAT device, public IP address, AWS Direct Connect connection, or AWS Site-to-Site VPN connection in order to communicate. When you use AWS PrivateLink to connect to a service that provides access to personal data, you can use VPC endpoint policies and security groups to control access, according to your organization's data perimeter definition. For a sample VPC endpoint policy that allows only IAM principles and AWS resources in a trusted organization to access a service endpoint, see Require organization membership to access VPC resources in this guide.

AWS Resource Access Manager

AWS Resource Access Manager (AWS RAM) helps you securely share your resources across AWS accounts to reduce operational overhead and provide visibility and auditability. As you plan your multi-account segmentation strategy, consider using AWS RAM to share the personal data stores that you store in a separate, isolated account. You can share that personal data with other, trusted accounts for the purposes of processing. In AWS RAM, you can manage permissions that define what actions can be performed on shared resources. All API calls to AWS RAM are logged in CloudTrail. Also, you can configure Amazon CloudWatch Events to automatically notify you for specific events in AWS RAM, such as when changes are made to a resource share.

Though you can share many types of AWS resources with other AWS accounts by using resource-based policies in IAM or bucket policies in Amazon S3, AWS RAM provides several additional

benefits for privacy. AWS provides data owners with additional visibility over how and with whom the data is shared across your AWS accounts, including:

- Being able to share a resource with an entire OU instead of manually updating lists of account IDs

- Enforcement of the invitation process for share initiation if the consumer account isn't part of your organization

- Visibility into which specific IAM principals have access to each individual resource

If you've previously used a resource-based policy to manage a resource share and want to use AWS RAM instead, use the PromoteResourceShareCreatedFromPolicy API operation.

Amazon SageMaker AI

Amazon SageMaker AI is a managed machine learning (ML) service that helps you build and train ML models and then deploy them into a production-ready hosted environment. SageMaker AI is designed to make it easier to prepare training data and create model features.

Amazon SageMaker AI Model Monitor

Many organizations consider data drift when training ML models. Data drift is 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. If the statistical nature of the data that a ML model receives in production drifts away from the nature of the baseline data it was trained on, the accuracy of the predictions might decline. Amazon SageMaker AI Model Monitor can continuously monitor the quality of Amazon SageMaker AI machine learning models in production and monitor data quality. Early and proactive detection of data drift can help you implement corrective actions, such as retraining models, auditing upstream systems, or fixing data quality issues. Model Monitor can alleviate the need to manually monitor models or build additional tooling.

Amazon SageMaker AI Clarify

Amazon SageMaker AI Clarify provides insight into model bias and explainability. SageMaker AI Clarify is commonly used during ML model data preparation and the overall development phase. Developers can specify attributes of interest, such as gender or age, and SageMaker AI Clarify runs a set of algorithms to detect any presence of bias in those attributes. After the algorithm runs, SageMaker AI Clarify provides a visual report with a description of the sources and measurements

of possible bias so that you can identify steps to remediate the bias. For example, in a financial dataset that contains only a few examples of business loans to one age group as compared to others, SageMaker could flag imbalances so that you can avoid a model that disfavors that age group. You can also check already trained models for bias by reviewing its predictions and by continuously monitoring those ML models for bias. Finally, SageMaker AI Clarify is integrated with Amazon SageMaker AI Experiments to provide a graph that explains which features contributed most to a model's overall prediction-making process. This information could be useful to meet explainability outcomes, and it could help you determine if a particular model input has more influence than it should on the overall model behavior.

Amazon SageMaker Model Card

Amazon SageMaker Model Card can help you document critical details about your ML models for governance and reporting purposes. These details can include the model owner, general purpose, intended use cases, assumptions made, risk rating of a model, training details and metrics, and evaluation results. For more information, see Model Explainability with AWS Artificial Intelligence and Machine Learning Solutions (AWS whitepaper).

AWS features that help manage the data lifecycle

When personal data is no longer required, you can use lifecycle and time-to-live polices for data in many different data stores. When configuring data retention policies, consider the following locations that might contain personal data:

- Databases, such as Amazon DynamoDB and Amazon Relational Database Service (Amazon RDS)
- Amazon S3 buckets
- Logs from CloudWatch and CloudTrail
- Cached data from migrations in AWS Database Migration Service (AWS DMS) and AWS Glue DataBrew projects
- Backups and snapshots

The following AWS services and features can help you configure data retention policies in your AWS environments:

- Amazon S3 Lifecycle – A set of rules that define actions that Amazon S3 applies to a group of objects. In the Amazon S3 Lifecycle configuration, you can create expiration actions, which define when Amazon S3 deletes expired objects on your behalf. For more information, see Managing your storage lifecycle.

- [Amazon Data Lifecycle Manager](#) – In Amazon EC2, create a policy that automates the creation, retention, and deletion of Amazon Elastic Block Store (Amazon EBS) snapshots and EBS-backed Amazon Machine Images (AMIs).

- [DynamoDB Time to Live (TTL)](#) – Define a per-item timestamp that determines when an item is no longer needed. Shortly after the date and time of the specified timestamp, DynamoDB deletes the item from your table.

- [Log retention settings in CloudWatch Logs](#) – You can adjust the retention policy for each log group to a value between 1 day and 10 years.

- [AWS Backup](#) – Centrally deploy data protection policies to configure, manage, and govern your backup activity across a variety of AWS resources, including S3 buckets, RDS database instances, DynamoDB tables, EBS volumes, and many more. Apply backup policies to your AWS resources by either specifying resource types or provide additional granularity by applying based on existing resource tags. Audit and report on backup activity from a centralized console to help meet backup compliance requirements.

AWS services and features that help segment data

Data segmentation is the process by which you store data in separate containers. This can help you to provide differentiated security and authentication measures to each dataset and to reduce the scope of impact of exposure for your overall dataset. For example, instead of storing all customer data in one large database, you may segment this data into smaller, more manageable groups.

You can use physical and logical separation to segment personal data:

- **Physical separation** – The act of storing data in separate data stores or distributing your data into separate AWS resources. Though the data is physically separated, both resources might be accessible to the same principals. This is why we recommend combining physical separation with logical separation.

- **Logical separation** – The act of isolating data by using access controls. Different job functions require different levels of access to subsets of personal data. For a sample policy that implements logical separation, see [Grant access to specific Amazon DynamoDB attributes](#) in this guide.

The combination of a logical and physical separation provides flexibility, simplicity, and granularity when writing identity-based and resource-based policies to support differentiated access across job functions. For example, it can be operationally complex to create the policies that logically

separate different data classifications in a single S3 bucket. Using dedicated S3 buckets for each data classification simplifies policy configuration and management.

Sample privacy-related policies

We would love to hear from you. Please provide feedback on the AWS PRA by taking a [short survey](#).

Many organizations that handle sensitive data take a preventative-forward approach, with layers of detective and reactive controls implemented throughout. This section provides examples of privacy-related policies for AWS Identity and Access Management (IAM), AWS Organizations, and AWS Key Management Service (AWS KMS). These policies can help your organization meet various use, disclosure limitation, and cross-border data transfer privacy goals by using a preventative approach. Many of these policies are referenced in previous sections in this guide.

This section contains the following sample policies:

- [Require access from specific IP addresses](#)
- [Require organization membership to access VPC resources](#)
- [Restrict data transfers across AWS Regions](#)
- [Grant access to specific Amazon DynamoDB attributes](#)
- [Restrict changes to VPC configurations](#)
- [Require attestation to use an AWS KMS key](#)

Require access from specific IP addresses

We would love to hear from you. Please provide feedback on the AWS PRA by taking a [short survey](#).

This policy allows the `john_stiles` user to assume IAM roles only if the call is coming from an IP address in the ranges `192.0.2.0/24` or `203.0.113.0/24`. This policy can help prevent unintended disclosure of personal data and unwanted cross-border data transfers. For example, if your organization has customer support staff that require access to personal data, you might want that support staff to access that data only from offices that are located in a subset of specific AWS Regions.

Require organization membership to access VPC resources

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

This VPC endpoint policy allows only AWS Identity and Access Management (IAM) principals and resources from the o-1abcde123 organization to access Amazon Personalize (Amazon S3) endpoints. This preventative control helps establish a zone of trust and define the personal data perimeter. For more information about how this policy can help protect privacy and personal data in your organization, see AWS PrivateLink in this guide.

Restrict data transfers across AWS Regions

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

With the exception of two AWS Identity and Access Management (IAM) roles, this service control policy denies API calls to regional AWS services in AWS Regions other than eu-west-1 and eu-central-1. This SCP can help prevent the creation of AWS storage and processing services in unapproved Regions. This can help prevent personal data from being handled by AWS services in those Regions altogether. This policy uses a NotAction parameter because it accounts for global AWS services, such as IAM, and services that integrate with global services, such as AWS Key Management Service (AWS KMS) and Amazon CloudFront. In the parameter values, you can specify those global and other non-applicable services as exceptions. For more information about how this policy can help protect privacy and personal data in your organization, see AWS Organizations in this guide.

Grant access to specific Amazon DynamoDB attributes

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

As your organization discusses strategies to physically and logically separate personal data, consider which AWS storage services support fine-grained access control policies in AWS Identity and Access Management (IAM). The following identity-based policy allows retrieval of only the `UserID`, `SignUpTime`, and `LastLoggedIn` attributes from an Amazon DynamoDB table named `Users`. For example, you might attach this policy to a customer support role instead of giving this role access to the full personal dataset. For more information about how this policy can help protect privacy and personal data in your organization, see AWS services and features that help segment data in this guide.

Restrict changes to VPC configurations

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

After you have designed and deployed the AWS infrastructure that supports your cross-border data transfer requirements, which includes network data flows, you might want to prevent modifications. The following service control policy helps prevent VPC configuration drift or unintentional modification. It denies new internet gateway attachments, VPC peering connections, transit gateway attachments, and new VPN connections. For more information about how this policy can help protect privacy and personal data in your organization, see AWS Transit Gateway in this guide.

Require attestation to use an AWS KMS key

We would love to hear from you. Please provide feedback on the AWS PRA by taking a [short survey](#).

The following AWS Key Management Service (AWS KMS) key policy allows AWS Nitro Enclave instances to use a KMS key only if the enclave's attestation document in the request matches the measurements in the condition statement. This policy allows only trusted enclaves to decrypt the data. For more information about how this policy can help protect privacy and personal data in your organization, see [AWS Nitro Enclaves](#) in this guide. For a complete list of AWS KMS condition keys that can be used in key policies and in AWS Identity and Access Management (IAM) policies, see [Condition keys for AWS KMS](#).

Resources

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

AWS Prescriptive Guidance

- AWS Security Reference Architecture (AWS SRA)

AWS documentation

- Data protection (AWS Well-Architected Framework)
- Data classification (AWS whitepaper)
- Amazon Web Services: Risk and Compliance (AWS whitepaper)
- Hybrid architectures to address personal data processing requirements (AWS whitepaper)
- Navigating GDPR Compliance on AWS (AWS whitepaper)
- Building a data perimeter on AWS (AWS whitepaper)
- AWS Security Documentation

Other AWS resources

- AWS Compliance Programs
- AWS Shared Responsibility Model
- Data Privacy FAQ
- AWS Security Assurance Services
- AWS Digital Sovereignty Pledge: Control without compromise (AWS blog post)
- AWS Security Learning

Contributors

We would love to hear from you. Please provide feedback on the AWS PRA by taking a short survey.

This guide was authored by the AWS Security Assurance Services team. For support implementing the recommendations in this guide and operationalizing your workloads, contact the AWS Security Assurance Services team.

Primary authors

- Daniel Nieters, AWS Principal Privacy Consultant
- Amber Welch, AWS Senior Privacy Consultant
- Robert Carter, AWS Technical Program Manager

Contributors

- Avik Mukherjee, AWS Senior Security Consultant
- David Bounds, AWS Senior Solutions Architect
- Jeff Lombardo, AWS Senior Security Solutions Architect
- Ram Ramani, AWS Principal Security Solutions Architect
- Vanessa Jacobs, AWS Senior Security Consultant

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
Signficant updates	We made significant updates throughout.	March 26, 2024
Initial publication	—	October 2, 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, Amazon SageMaker AI 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.

EDI

See electronic data interchange.

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.

electronic data interchange (EDI)

The automated exchange of business documents between organizations. For more information, see What is Electronic Data Interchange.

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.

few-shot prompting

Providing an LLM with a small number of examples that demonstrate the task and desired output before asking it to perform a similar task. This technique is an application of in-context learning, where models learn from examples (*shots*) that are embedded in prompts. Few-shot prompting can be effective for tasks that require specific formatting, reasoning, or domain knowledge. See also zero-shot prompting.

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.

FM

See foundation model.

foundation model (FM)

A large deep-learning neural network that has been training on massive datasets of generalized and unlabeled data. FMs are capable of performing a wide variety of general tasks, such as understanding language, generating text and images, and conversing in natural language. For more information, see What are Foundation Models.

G

generative AI

A subset of AI models that have been trained on large amounts of data and that can use a simple text prompt to create new content and artifacts, such as images, videos, text, and audio. For more information, see What is Generative AI.

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.

golden image

A snapshot of a system or software that is used as a template to deploy new instances of that system or software. For example, in manufacturing, a golden image can be used to provision software on multiple devices and helps improve speed, scalability, and productivity in device manufacturing operations.

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.

holdout data

A portion of historical, labeled data that is withheld from a dataset that is used to train a [machine learning](#) model. You can use holdout data to evaluate the model performance by comparing the model predictions against the holdout data.

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 language model (LLM)

A deep learning AI model that is pretrained on a vast amount of data. An LLM can perform multiple tasks, such as answering questions, summarizing documents, translating text into other languages, and completing sentences. For more information, see What are LLMs.

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.

LLM

See large language model.

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

A system engineering approach that takes privacy into account through the whole development 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.

prompt chaining

Using the output of one LLM prompt as the input for the next prompt to generate better responses. This technique is used to break down a complex task into subtasks, or to iteratively refine or expand a preliminary response. It helps improve the accuracy and relevance of a model's responses and allows for more granular, personalized results.

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).

RAG

See Retrieval Augmented Generation.

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.

Retrieval Augmented Generation (RAG)

A generative AI technology in which an LLM references an authoritative data source that is outside of its training data sources before generating a response. For example, a RAG model might perform a semantic search of an organization's knowledge base or custom data. For more information, see What is RAG.

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 by design

A system engineering approach that takes security into account through the whole
development process.

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.

system prompt

A technique for providing context, instructions, or guidelines to an LLM to direct its behavior. System prompts help set context and establish rules for interactions with users.

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.

zero-shot prompting

Providing an LLM with instructions for performing a task but no examples (*shots*) that can help guide it. The LLM must use its pre-trained knowledge to handle the task. The effectiveness of zero-shot prompting depends on the complexity of the task and the quality of the prompt. See also few-shot prompting.

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.