AWS PARTNER CERTIFICATION READINESS

Content Review Session

Week 5 – Domain 4

Data Security and Governance

OPTIONAL AWS Skill Builder Subscription

The Skill Builder subscription provides access to official AWS Certification practice exams, self-paced digital training content including open-ended challenges, self-paced labs, and game-based learning. **Please note, the Skill Builder subscription is not required for this Accelerator program.**

Free digital training

LINK HERE

Special features include:

- 600+ digital courses
- Learning plans
- 10 Practice Question Sets
- *AWS Cloud Quest (Foundational)*

Individual subscription

LINK HERE

Everything in free digital training, plus:

- AWS Cloud Quest (Intermediate - Advanced)
- AWS Certification Official Practice Exams
- Enhanced Exam Prep Courses
- Unlimited access to 1000+ hands-on labs
- AWS Jam Journeys (lab-based challenges)
- AWS Digital Classroom (Annual only)

Individual subscriptions are priced at $29 USD per month (*Flexibility to cancel anytime*) or $449 USD per year.

Access **65** Data Engineer - Associate Practice Exam Questions with feedback on your answer choices

Today's Learning Outcomes

During this session, we will cover:

- Apply authentication mechanisms
- Apply authorization mechanisms
- Ensure data encryption and masking
- Prepare logs for audit
- Understand data privacy and governance

AWS PARTNER CERTIFICATION READINESS

Domain 4: Data Security & Governance

Apply authentication mechanisms

Apply authentication mechanisms

Knowledge of:

- VPC security networking concepts
- Differences between managed services and unmanaged services
- Authentication methods (password-based, certificate-based, and role-based)
- Differences between AWS managed policies and customer managed policies

Skills in:

- Updating VPC security groups
- Creating and updating IAM groups, roles, endpoints, and services
- Creating and rotating credentials for password management (for example, AWS Secrets Manager)
- Setting up IAM roles for access (for example, Lambda, Amazon API Gateway, AWS CLI, CloudFormation)
- Applying IAM policies to roles, endpoints, and services (for example, S3 Access Points, AWS PrivateLink)

What is IAM?

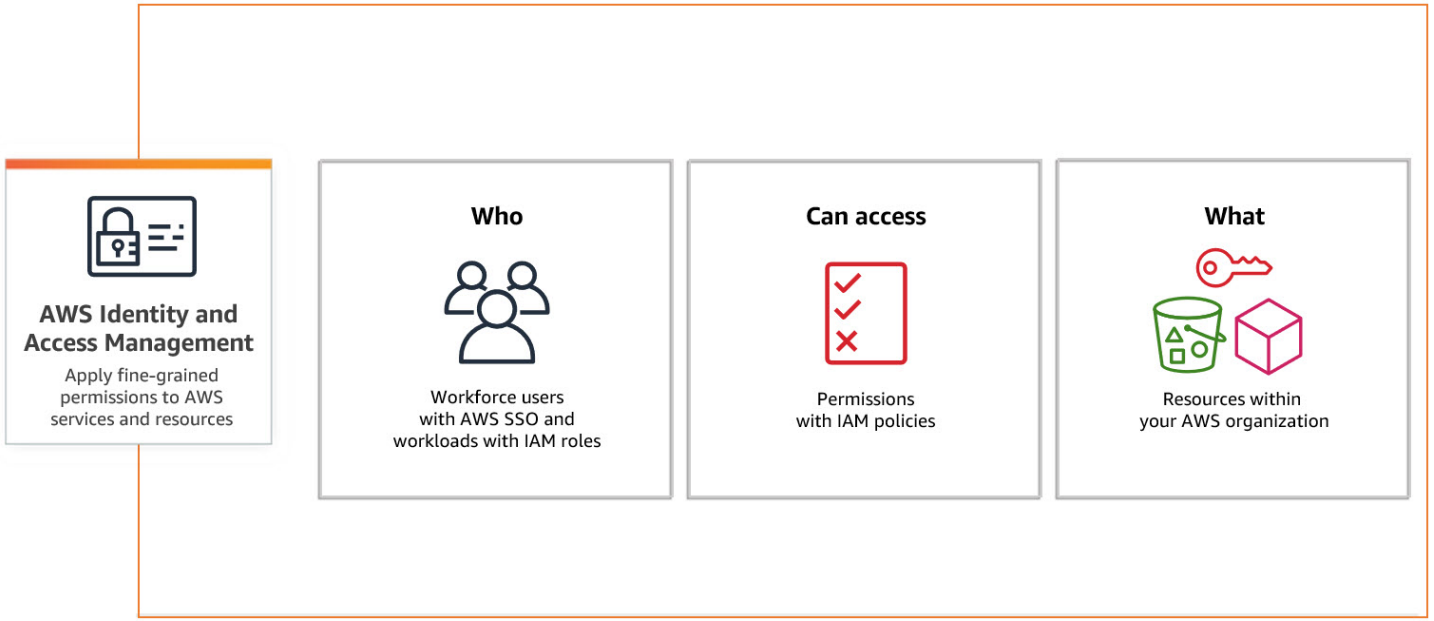
Policies and Technologies used to ensure the appropriate access to technology resources

Overview

AWS IAM provides fine-grained access control across all of AWS. With IAM, you can specify who can access which services and resources, under which conditions.

IAM Policies

IAM Policies allow you to manage permissions for your workforce and systems to ensure least-privileged access.



IAM Users and Groups

The building blocks of AWS Identity and Access Management

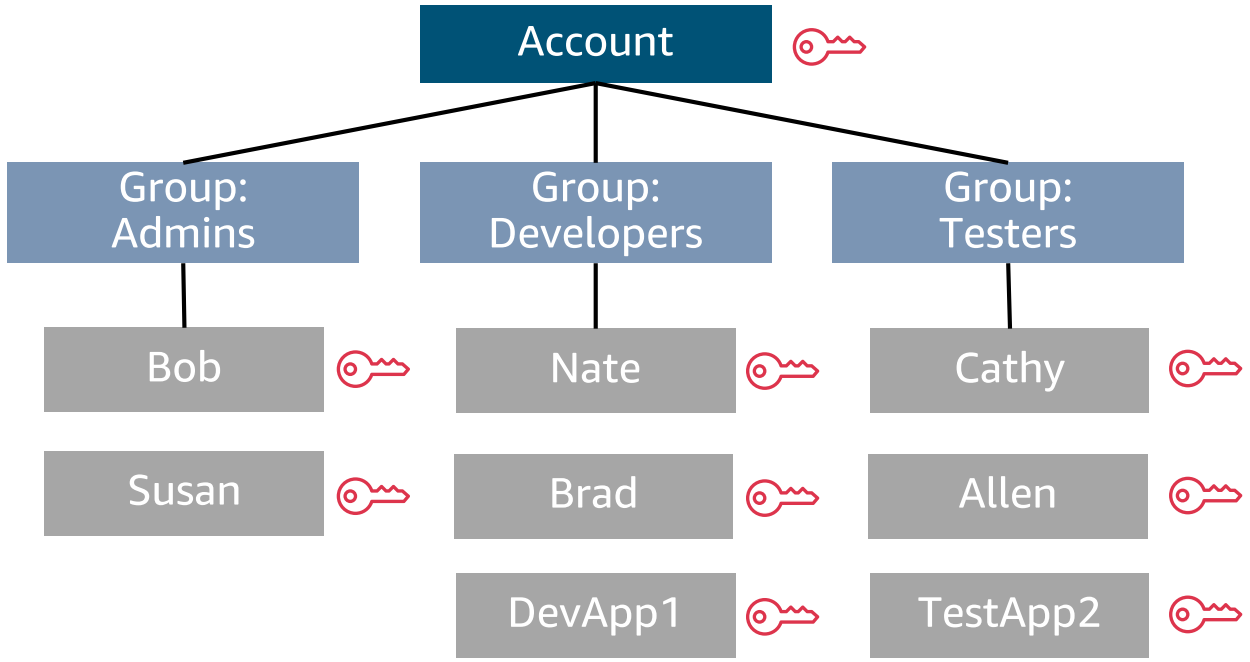
IAM Users

An IAM User is an entity that is created in AWS to represent the person, or application, that uses it to interact with AWS.

IAM Groups

An IAM Group is a collection of IAM users. User groups let you specify permissions for multiple users, which can make it easier to manage the permissions for those users.

Ex: You could have a user group called Admins and give that user group typical administrator permissions.



Policy Interpretation Deep Dive!

IAM Policies are the bedrock of strong IAM security. Understanding how the policies work and being able to interpret them is critical for success as an Architect and on the exam

Identity Policies

Identity Policies are IAM policies that are applied to identities. This can include both users as well as roles that users can assume. These are **different** than resource policies.

Implicit vs. Explicit Allow/Deny

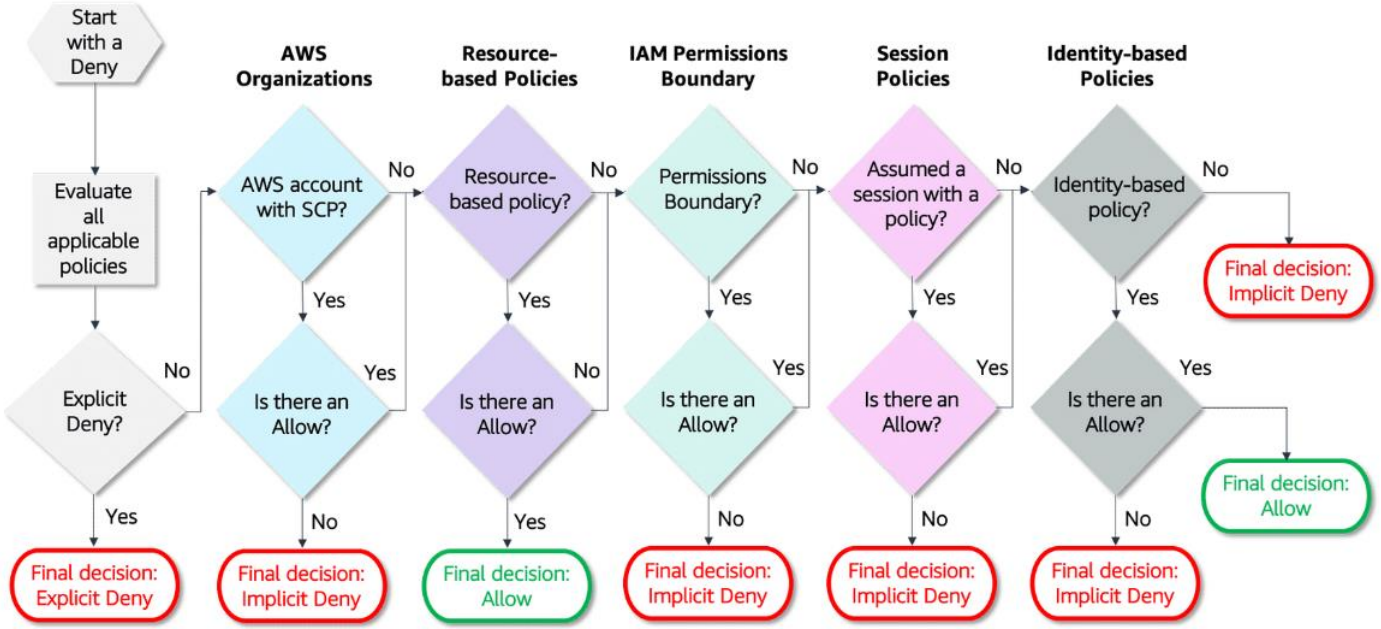
The default response to all requests is an **Implicit Deny**. This 'stance' can be overridden by allowing the user access with a permissions policy – this grants the user access because it has been **Explicitly Allowed.** The same process can be done with an **Explicit Deny** policy. This will deny access regardless of the permissions the user might have.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ExplicitDenyIfNotTheOwner",
            "Effect": "Deny",
            "Action": [
                "datapipeline:ActivatePipeline",
                "datapipeline:AddTags",
                "datapipeline:DeactivatePipeline",
                "datapipeline:DeletePipeline",
                "datapipeline:DescribeObjects",
                "datapipeline:EvaluateExpression",
                "datapipeline:GetPipelineDefinition",
                "datapipeline:PollForTask",
                "datapipeline:PutPipelineDefinition",
                "datapipeline:QueryObjects",
                "datapipeline:RemoveTags",
                "datapipeline:ReportTaskProgress",
                "datapipeline:ReportTaskRunnerHeartbeat",
                "datapipeline:SetStatus",
                "datapipeline:SetTaskStatus",
                "datapipeline:ValidatePipelineDefinition"
            ],
            "Resource": ["*"],
            "Condition": {
                "StringNotEquals": {"datapipeline:PipelineCreator": "${aws:userid}"}
            }
        }
    ]
}
```

Policy Interpretation – Deny vs Allow

Policy Interpretation

This flow chart provides details about how the decision is made as AWS authenticates the principal that makes the request. AWS evaluates the policy types in this order.



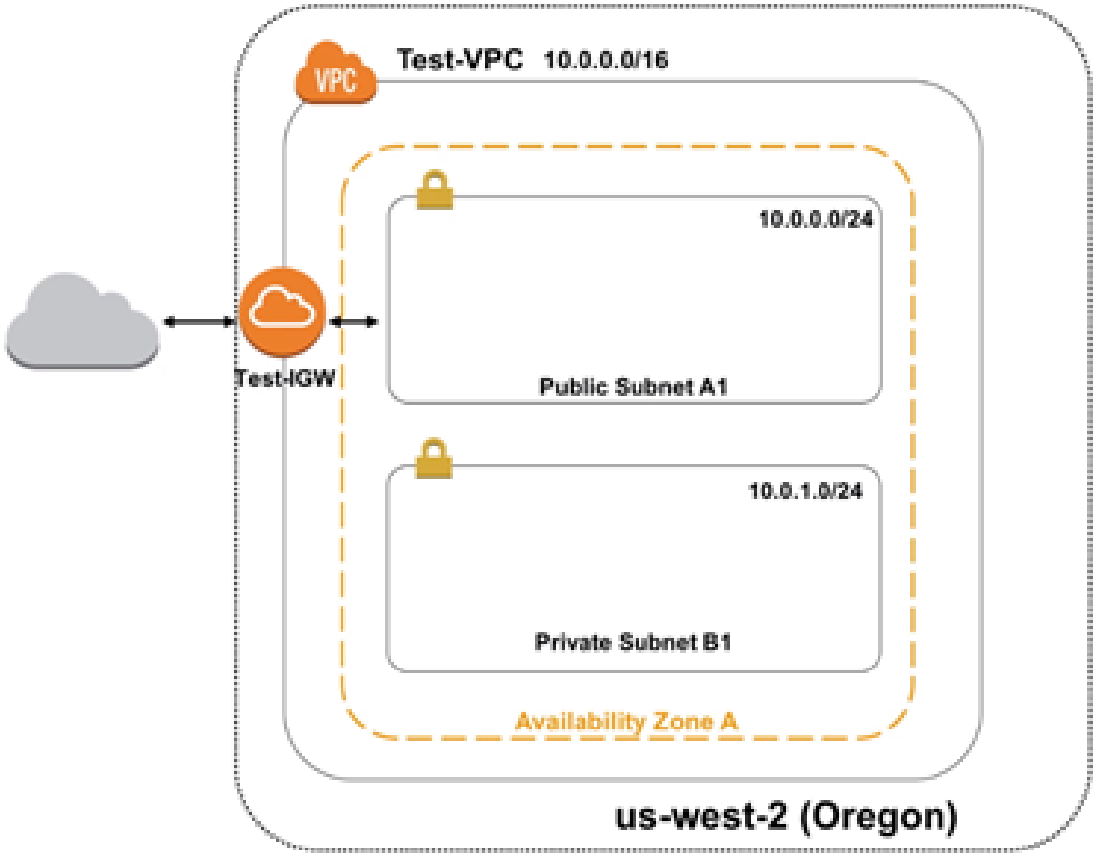
Amazon Virtual Private Cloud (VPC)

Provision a Logically Isolated Section of the AWS Cloud

Utilized with Auto Scaling to define which group scales

Capabilities

- Control your virtual networking environment
  - Subnets
  - Route tables
  - Security Groups
  - Network ACLs
- Connect to your on-premises network via VPN or Direct Connect
- Control if and how your instances access the internet



AWS Security Groups & NACLS

Two AWS features to increase security in your VPC: security groups and network ACLs.

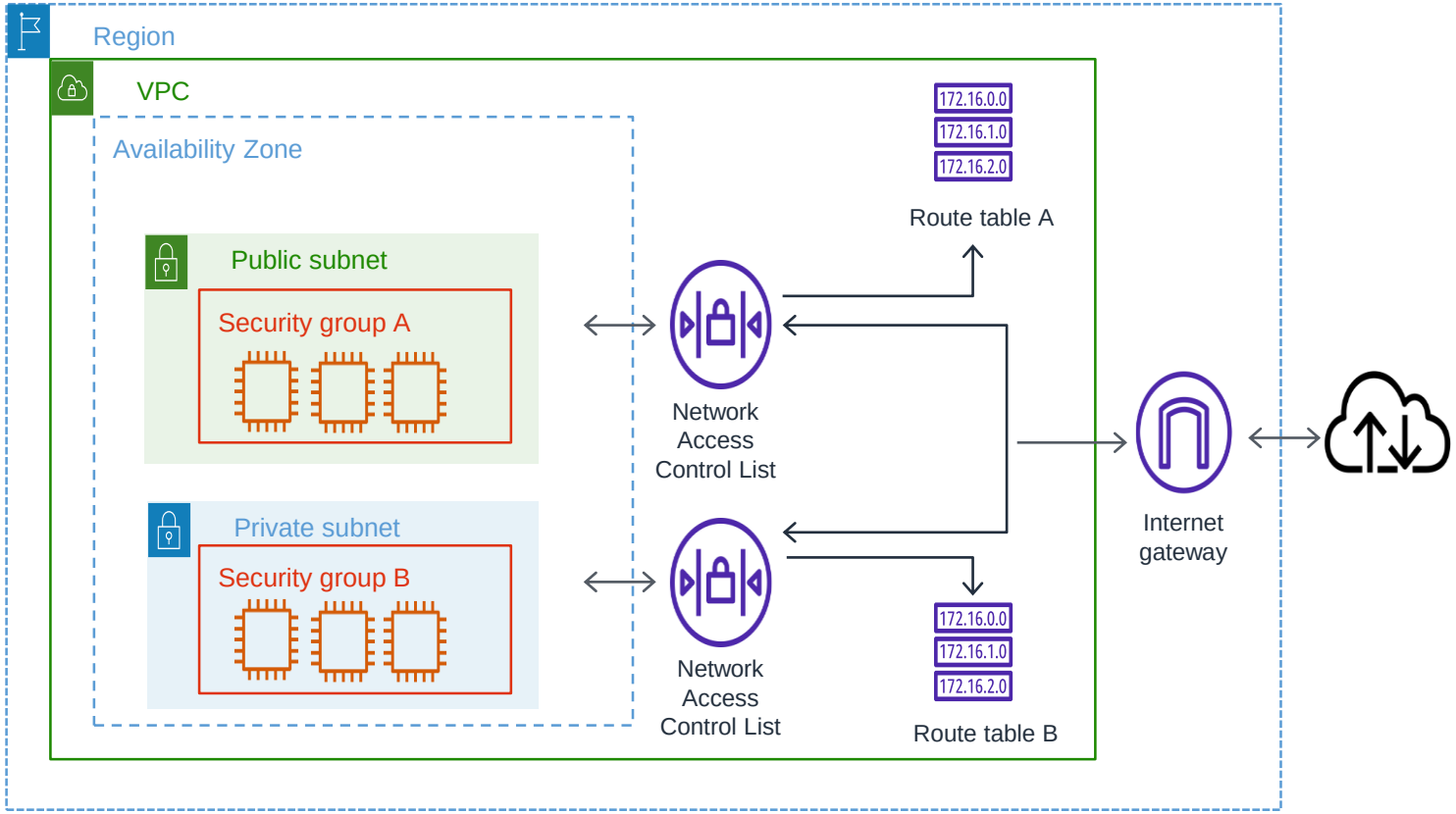
Overview

Security groups control inbound and outbound traffic at the instance level, acting as firewalls for associated EC2 instances.

Network ACLs control inbound and outbound traffic at the subnet level, acting as firewalls for associated subnets.

Key Exam Topics

- Network ACLs are **stateless**
- Security groups are **stateful**



Security Groups and NACLs

Differentiation Cheat Sheet

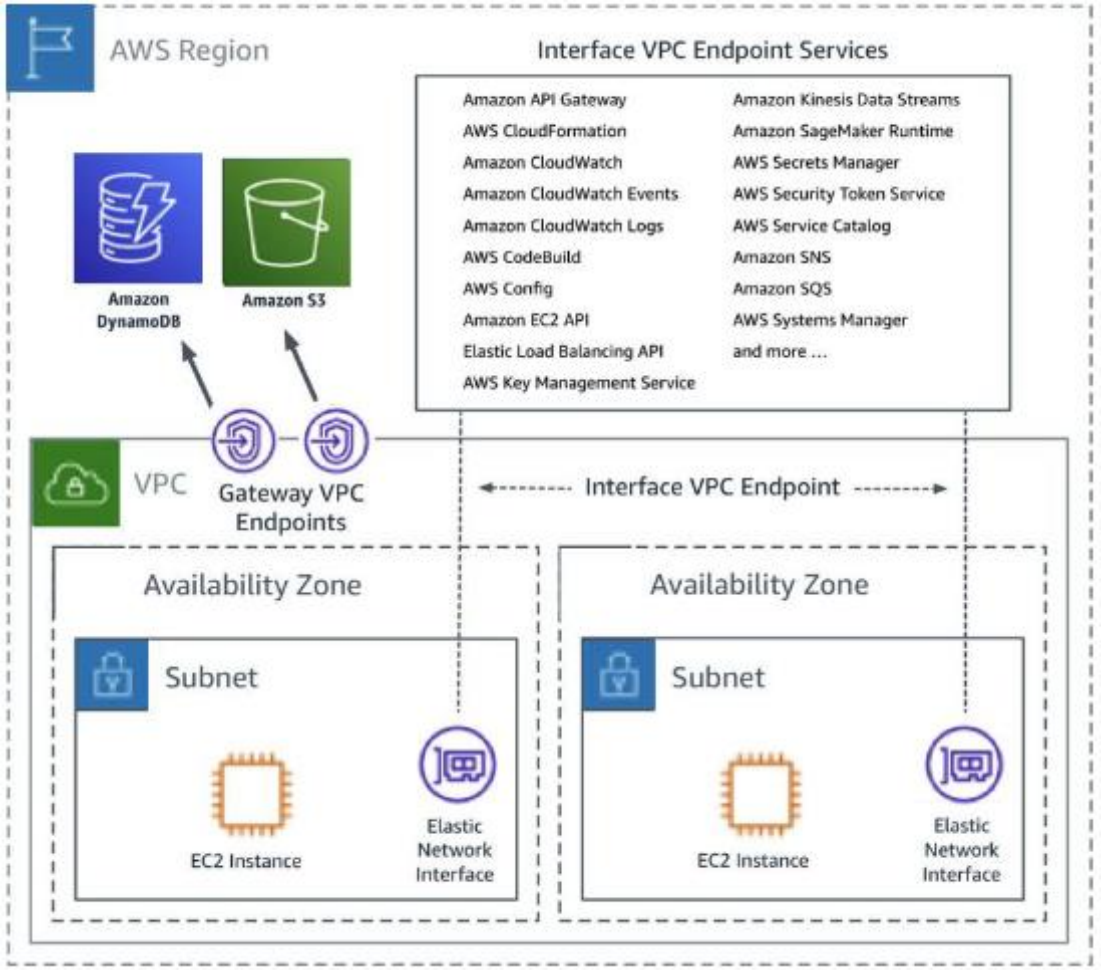
Security Group	Network ACL
Operates at the instance level	Operates at the subnet level
Supports allow rules only	Supports allow rules and deny rules
Is **stateful:** Return traffic is automatically allowed, regardless of any rules	Is **stateless:** Return traffic must be explicitly allowed by rules
We evaluate all rules before deciding whether to allow traffic	We process rules in order, starting with the lowest numbered rule, when deciding whether to allow traffic
Applies to an instance only if someone specifies the security group when launching the instance, or associates the security group with the instance later on	Automatically applies to all instances in the subnets that it's associated with (therefore, it provides an additional layer of defense if the security group rules are too permissive)

Connect Privately to public AWS Services

VPC Endpoints – Interface and Endpoint

Capabilities

- Connect your VPC to:
  - Supported AWS services
  - VPC endpoint services powered by PrivateLink
- Doesn't require public IPs or Internet connectivity
- Traffic does not leave the AWS network.
- Horizontally scaled, redundant, and highly available
- Robust access control



AWS PARTNER CERTIFICATION READINESS

Domain 4: Data Security & Governance

Apply authorization mechanisms

Apply authorization mechanisms

Knowledge of:

- Authorization methods (role-based, policy-based, tag-based, and attribute-based)
- Principle of least privilege as it applies to AWS security
- Role-based access control and expected access patterns
- Methods to protect data from unauthorized access across services

Skills in:

- Creating custom IAM policies when a managed policy does not meet the needs
- Storing application and database credentials (for example, Secrets Manager, AWS Systems Manager Parameter Store)
- Providing database users, groups, and roles access and authority in a database (for example, for Amazon Redshift)
- Managing permissions through Lake Formation

AWS Secrets Manager

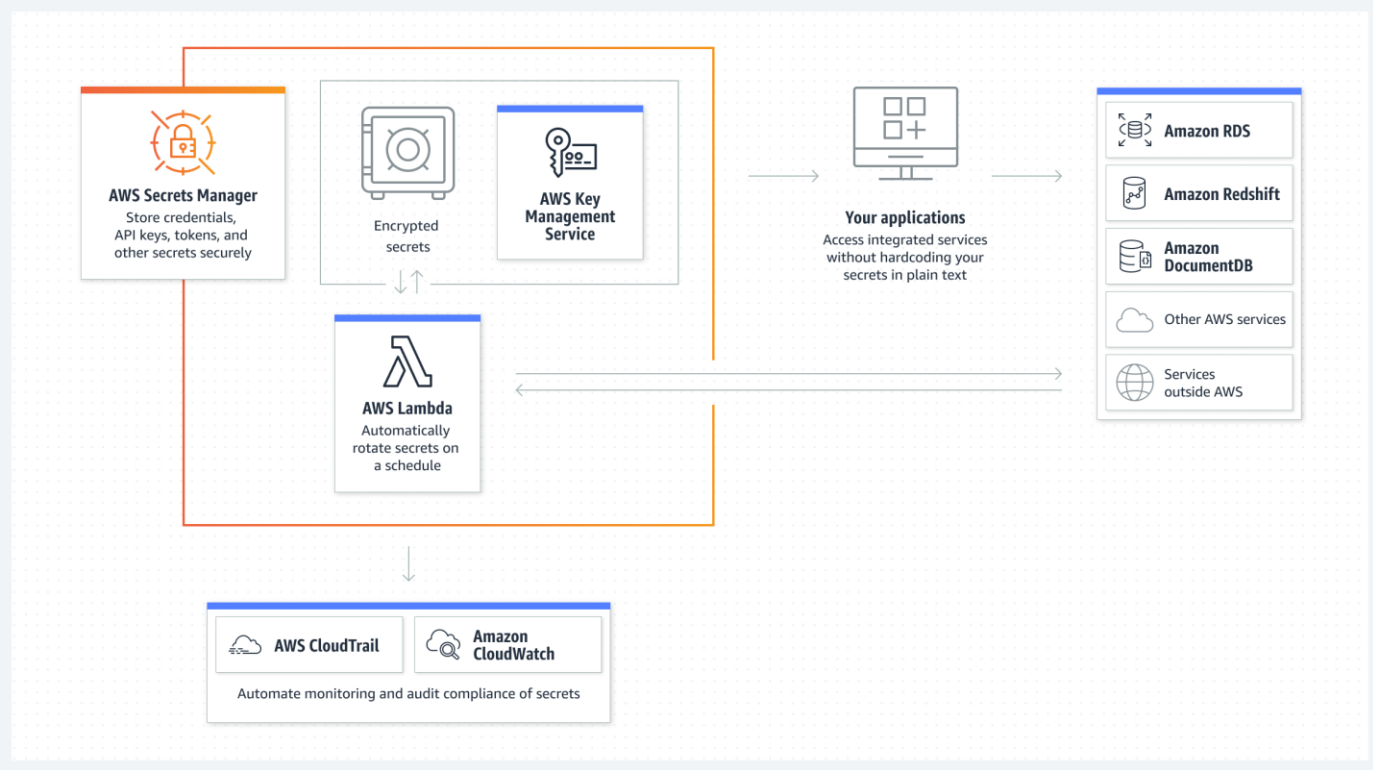
AWS Secrets Manager helps you manage, retrieve, and rotate database credentials, API keys, and other secrets throughout their lifecycles.

Key Features

You can configure Secrets Manager to automatically rotate your secrets without user intervention and on a specified schedule.

Secrets Manager encrypts the protected text of a secret by using AWS Key Management Service (AWS KMS).

Secrets Manager helps you improve your security posture by removing hard-coded credentials from your application source code, and by not storing credentials within the application, in any way.



AWS Systems Manager – Parameter Store

Provides secure, hierarchical storage for configuration data management and secrets management

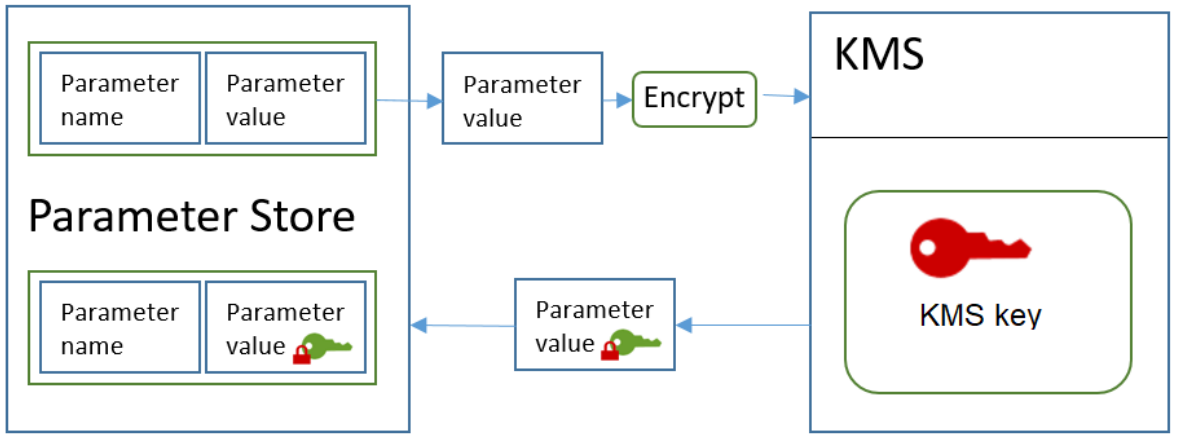
Benefits of Usage

- Separates data from code
- Stores configuration data and encrypted strings in versioned hierarchies
- Granular levels of control and audit access
- Reliable parameter storage with multiple AZ's

Compatibility

Accessible from: Amazon EC2, Amazon ECS, AWS Secrets Manager, AWS Lambda, AWS CloudFormation, AWS CodeBuild, CodePipeline, CodeDeploy

Integrated with: AWS KMS, Amazon SNS, Amazon CloudWatch, Amazon EventBridge, AWS CloudTrail



The Importance of IAM Roles

Roles are a way for users to temporarily gain permissions

What are they?

AWS Roles have the same makeup as an IAM user with the following differences:

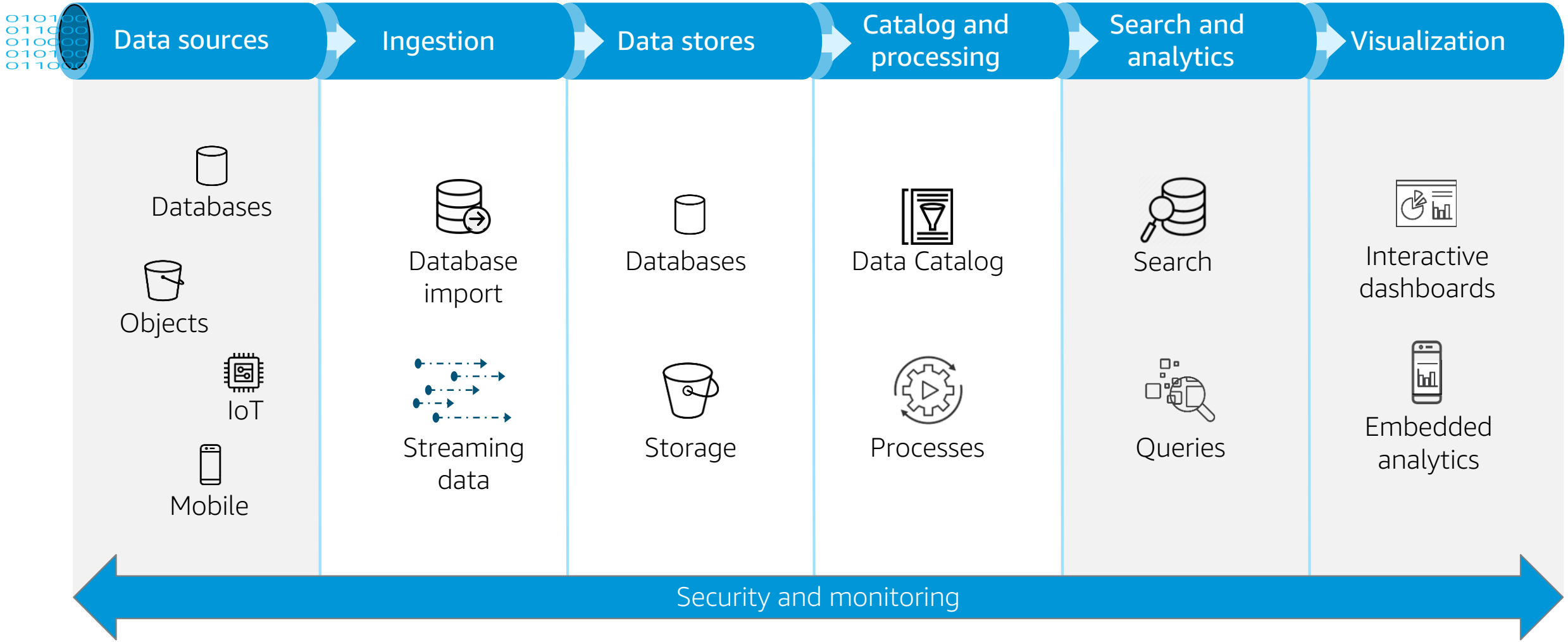
- An IAM role does not have long term credentials associated with it. A principal (user, machine, or authenticated identity) assumes the role and inherits permissions assigned to the user.
- Temporary access is granted using tokens (STS). Token expiration reduces the risks associated with credentials leaking or being reused.
- An IAM role has a trust policy that defines which conditions must be met to allow other principles to assume it.

When should be used?

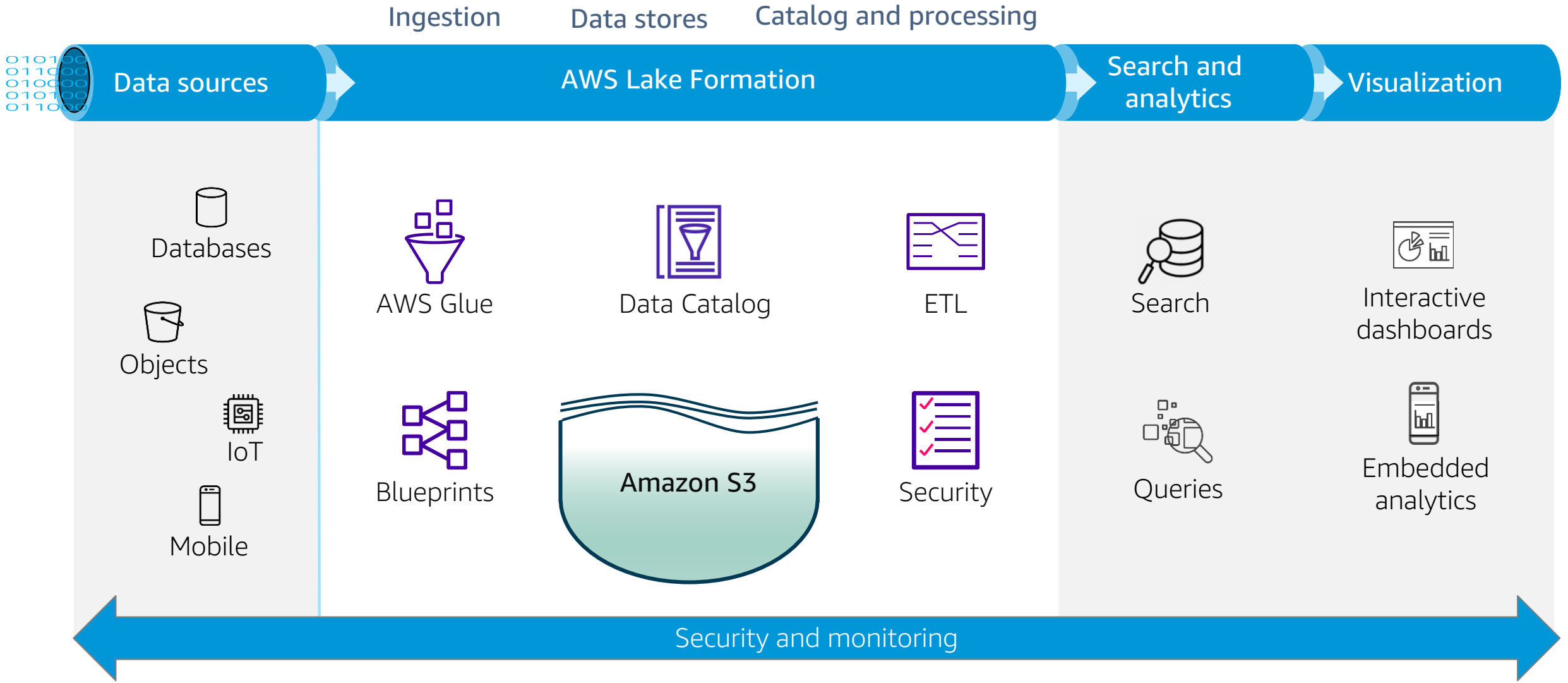
In general, there are four scenarios where IAM roles might be used:

1. One AWS service accesses another AWS Service
2. One AWS account accesses another AWS account
3. A third-party web identity needs access (i.e., Google, Facebook, Cognito)
4. Authentication using SAML2.0 federation (enables SSO)

Building a data lake, manually



AWS Lake Formation



AWS Lake Formation Permissions

Lake Formation uses a combination of Lake Formation permissions and AWS Identity and Access Management (IAM) permissions

Permissions Types

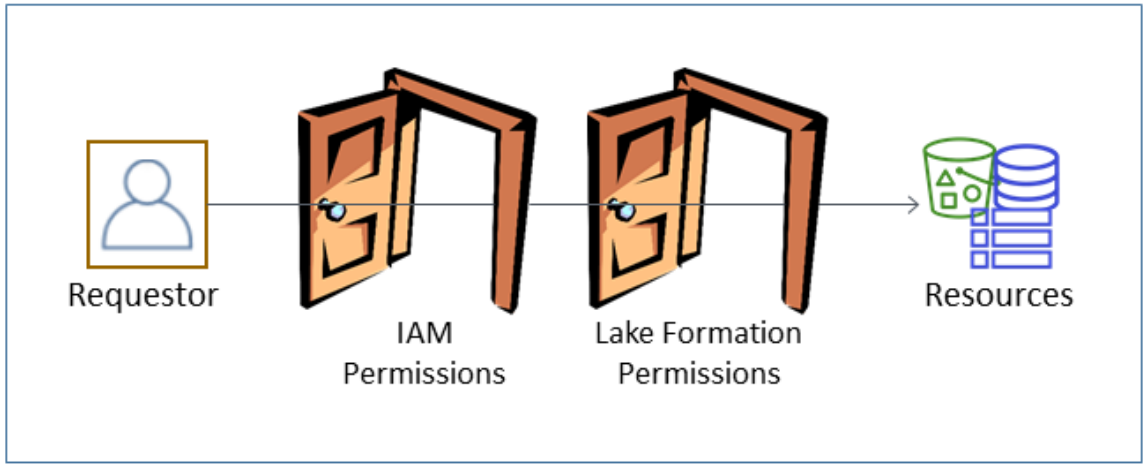
Metadata access – Permissions on Data Catalog resources
Enable principals to create, read, update, and delete metadata databases and tables in the Data Catalog.

Underlying data access – Permissions on locations in Amazon Simple Storage Service (Amazon S3)

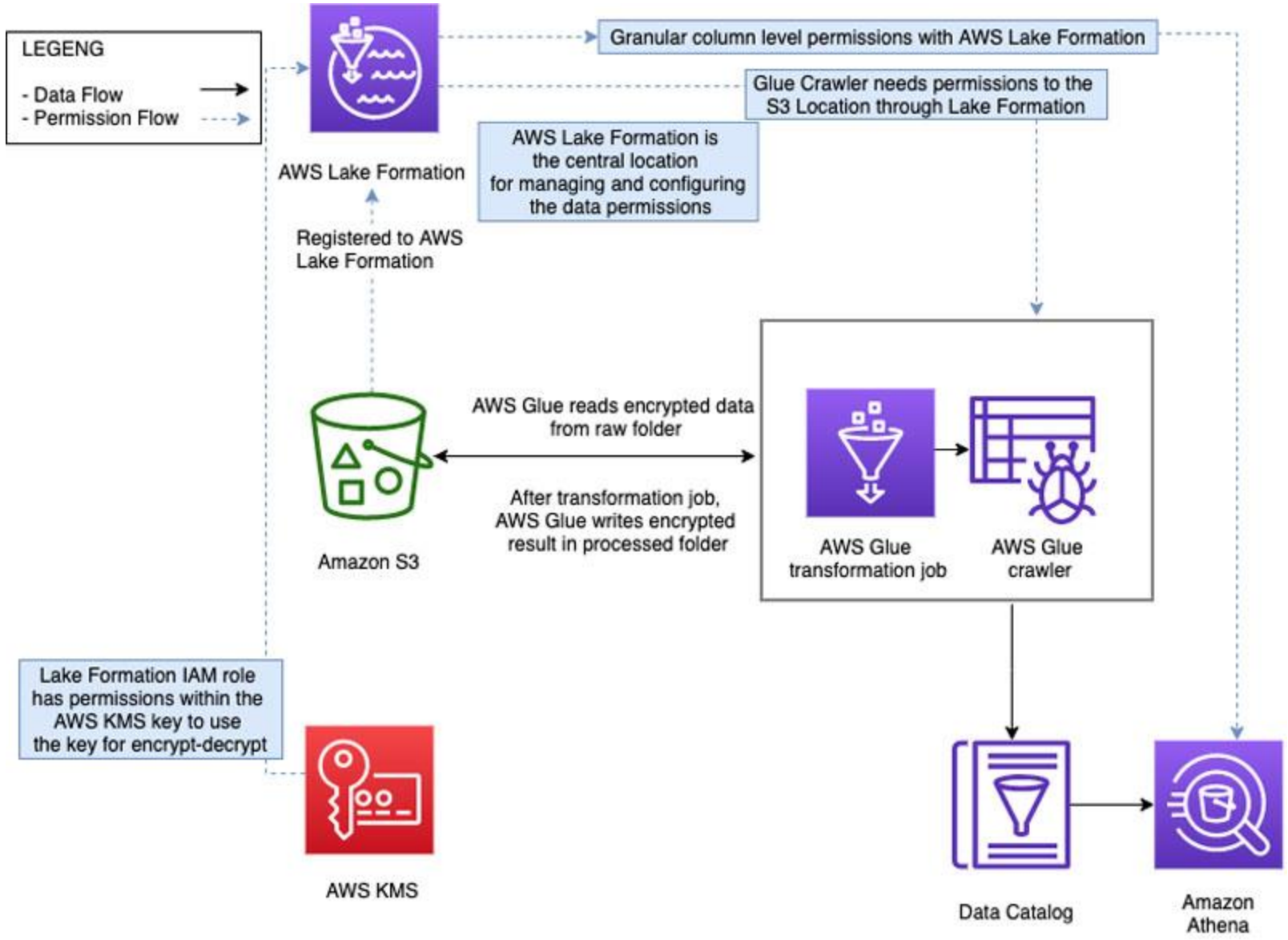
- Data lake permissions enable principals to read and write data to underlying Amazon S3 locations—data pointed to by Data Catalog resources.
- Data location permissions enable principals to create and alter metadata databases and tables that point to specific Amazon S3 locations.

Key Tips

IAM Permissions– Control access to Lake Formation and AWS Glue APIs and resources.
Lake Formation Permissions – Control access to Data Catalog resources, Amazon S3 locations, and underlying data.



Solution overview – Building a secure encrypted data lake with AWS Lake Formation



AWS PARTNER CERTIFICATION READINESS

Domain 4: Data Security & Governance

Ensure data encryption and masking

Ensure data encryption and masking

Knowledge of:

- Data encryption options available in AWS analytics services (for example, Amazon Redshift, Amazon EMR, AWS Glue)
- Differences between client-side encryption and server-side encryption
- Protection of sensitive data
- Data anonymization, masking, and key salting

Skills in:

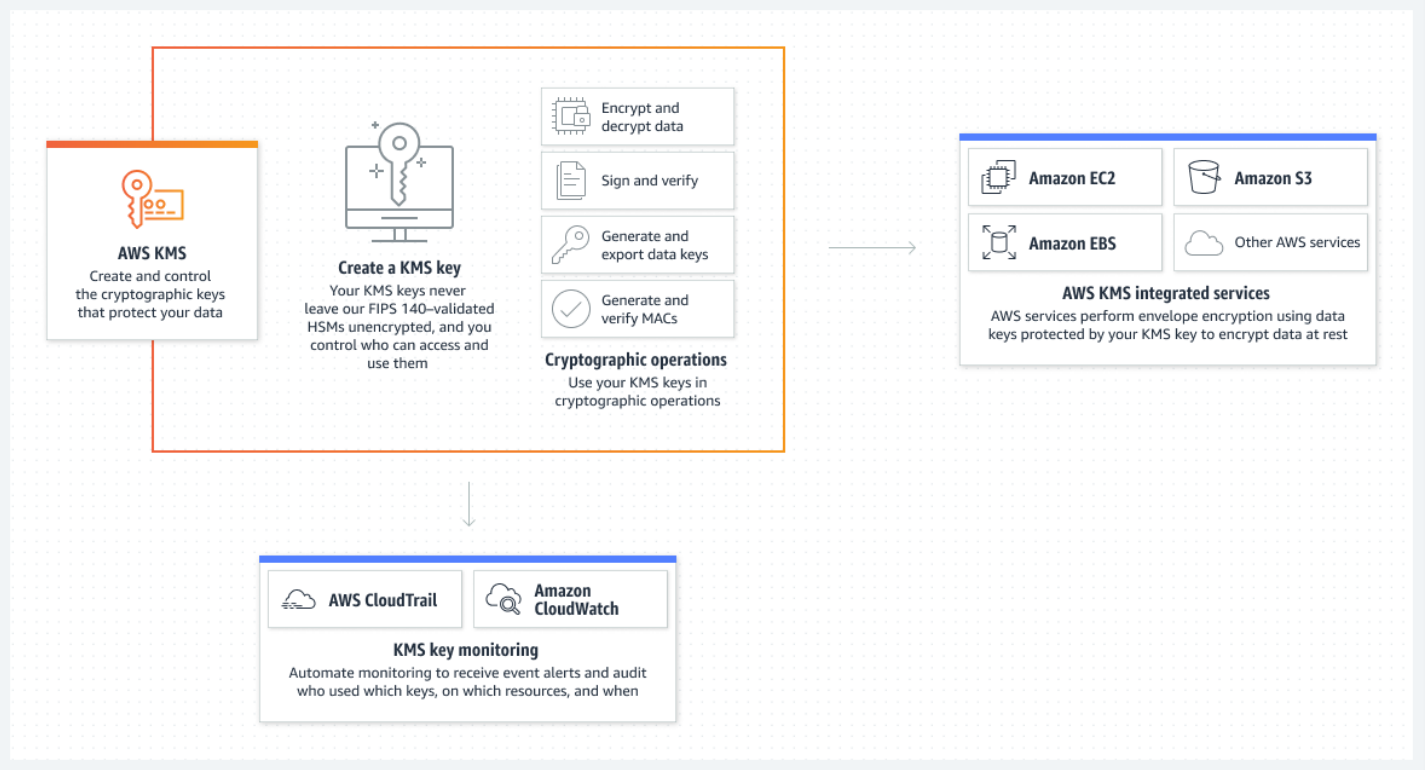
- Applying data masking and anonymization according to compliance laws or company policies
- Using encryption keys to encrypt or decrypt data (for example, AWS Key Management Service [AWS KMS])
- Configuring encryption across AWS account boundaries
- Enabling encryption in transit for data

AWS Key Management Service (KMS)

AWS KMS is a managed service that makes it easy for you to create and control the cryptographic keys that are used to protect your data.

What does it provide?

- You can use a KMS key to encrypt, decrypt, and re-encrypt data.
- AWS services that are integrated with AWS KMS use only symmetric encryption KMS keys to encrypt your data. These services do not support encryption with asymmetric KMS keys.
- All requests to use these keys are logged in AWS CloudTrail so that you can track who used which key, how and when.



AWS Key Management Service (KMS)

AWS KMS is a managed service that makes it easy for you to create and control the cryptographic keys that are used to protect your data

KMS Key Types

CMKs can be broken down into two general types: **AWS-managed** and **customer-managed**.

An AWS-managed CMK is created when you choose to enable server-side encryption of an AWS resource under the AWS-managed CMK for that service for the first time (e.g., SSE-KMS). An AWS-managed CMK can only be used to protect resources within the specific AWS service for which it's created. It does not provide the level of granular control that a customer-managed CMK provides.

For more control, a **best practice is to use a customer-managed CMK** in all supported AWS services and in your applications. A customer-managed CMK is created at your request and should be configured based upon your explicit use case.

	AWS-managed CMK	**Customer-managed CMK**
Creation	AWS generated on customer's behalf	Customer generated
Rotation	Once every three years automatically	Once a year automatically through opt-in or on-demand manually
Deletion	Can't be deleted	Can be deleted
Scope of use	Limited to a specific AWS service	Controlled via KMS/IAM policy
Key Access Policy	AWS managed	Customer managed
User Access Management	IAM policy	IAM policy

Encryption - Redshift

When enabling encryption for a cluster, the data blocks and system metadata are encrypted for the cluster and its snapshots.

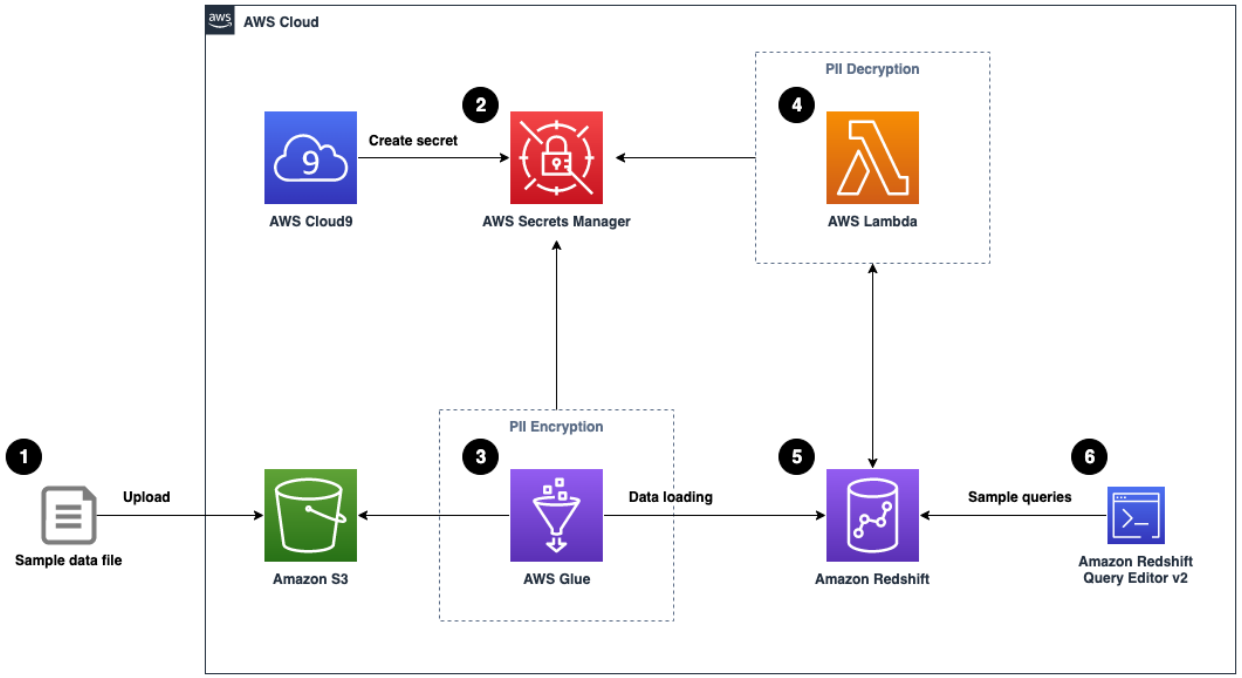
KMS Key Encryption

Enable when the cluster is launched, **or** modify an unencrypted cluster to use AWS KMS encryption.

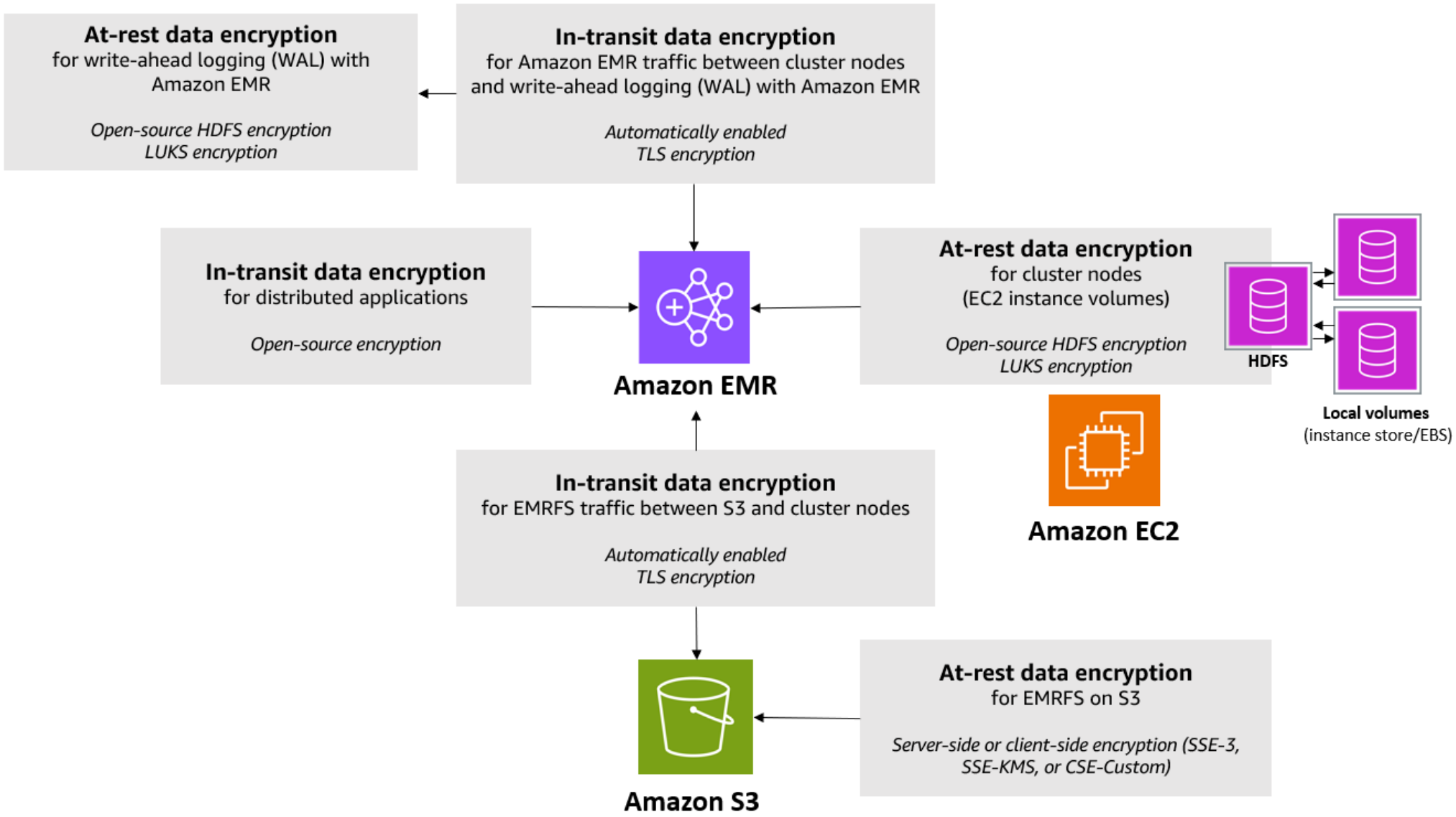
- Modified clusters are automatically migrated to a new encrypted cluster
- Use AWS-managed or customer managed keys

Encrypting RA3 Nodes

Both read and write queries can run during the process with less performance impact from the encryption. The encryption finishes much more quickly. Updated process steps include a restore operation and migration of cluster metadata to a target cluster. When you have petabyte-scale data volumes, the operation has been reduced from weeks to days.



Encryption – Amazon EMR



Encryption – Amazon Glue Data Catalog

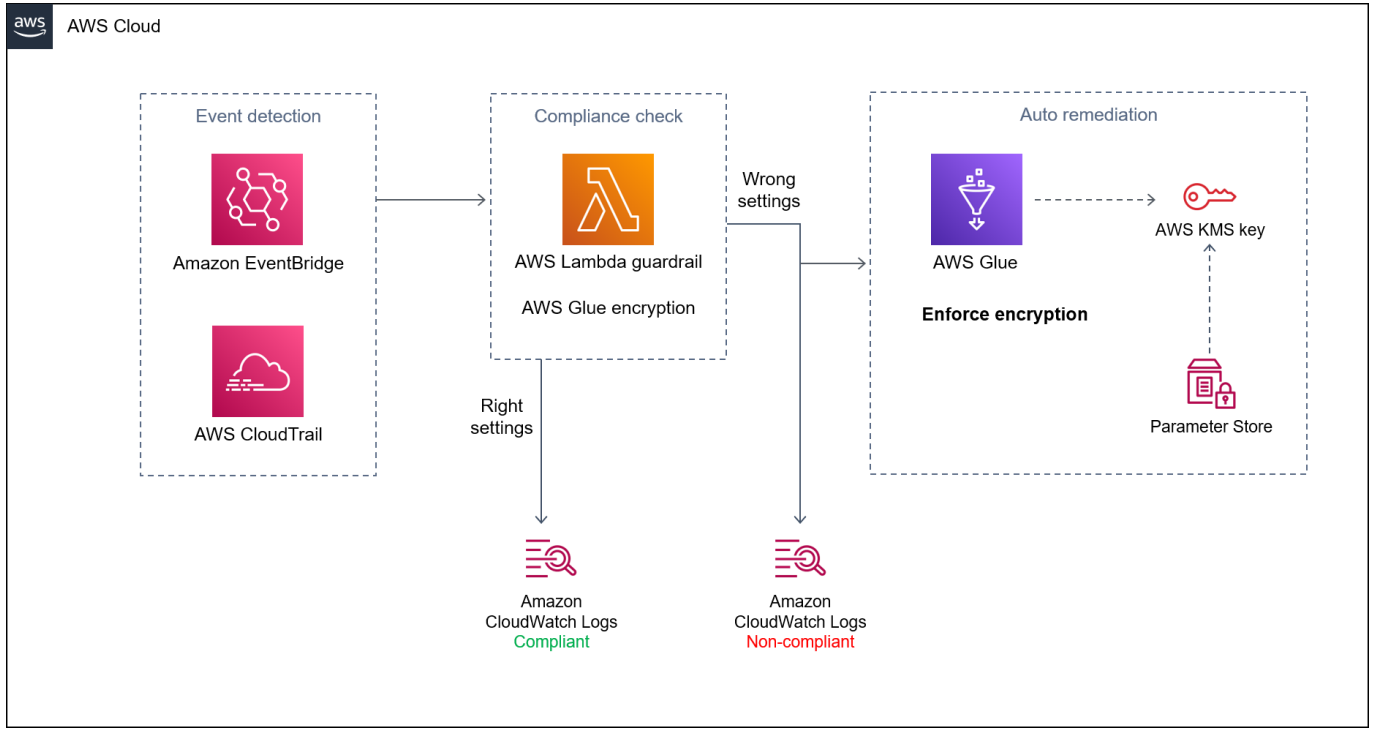
AWS Glue Data Catalog encryption provides enhanced security for your sensitive data.

Behavior

- Integrates with AWS KMS to encrypt metadata stored in the Data Catalog
- Enable or disable using AWS Glue console or CLI
- **Enabled** – all new objects will be encrypted
- **Disabled** – new objects will not be encrypted; existing objects remain encrypted
- Use AWS-managed keys or customer managed keys

Protected Resources

- Databases
- Tables
- Partitions
- Table versions
- Column statistics
- User-defined functions
- Data Catalog views



AWS PARTNER CERTIFICATION READINESS

Domain 4: Data Security & Governance

Prepare logs for audit

Prepare logs for audit

Knowledge of:

- How to log application data
- How to log access to AWS services
- Centralized AWS logs

Skills in:

- Using CloudTrail to track API calls
- Using CloudWatch Logs to store application logs
- Using AWS CloudTrail Lake for centralized logging queries
- Analyzing logs by using AWS services (for example, Athena, CloudWatch Logs Insights, Amazon OpenSearch Service)
- Integrating various AWS services to perform logging (for example, Amazon EMR in cases of large volumes of log data)

Amazon CloudTrail

CloudTrail logs, continuously monitors, and retains account activity related to actions across your AWS infrastructure, giving you control over storage, analysis, and remediation actions.

Some Use Cases

Audit activity

Monitor, store, and validate activity events for authenticity. Easily generate audit reports required by internal policies and external regulations.

Identify security incidents

Detect unauthorized access using the Who, What, and When information in CloudTrail Events. Respond with rules-based EventBridge alerts and automated workflows.

Key Concepts

CloudTrail **records user activity and API usage across AWS services as Events**. CloudTrail Events help you answer the questions of "who did what, where, and when?"

AWS CloudTrail Event history

Always available to view / download last 90 days of events per Region within AWS Account.

CloudTrail Trail outputs

Amazon S3

Saves CloudTrail trail event as series of compressed logs in Amazon S3.

Amazon CloudWatch Logs

Saves each CloudTrail event as a log event into a CloudWatch Logs.

CloudTrail events search

Amazon Athena

Query and analyze CloudTrail logs from Amazon S3 bucket as SQL statements.

Search and analyze CloudTrail events from their CloudWatch Logs group.

Amazon CloudTrail Lake

AWS CloudTrail Lake lets you run SQL-based queries on your events.

Usage

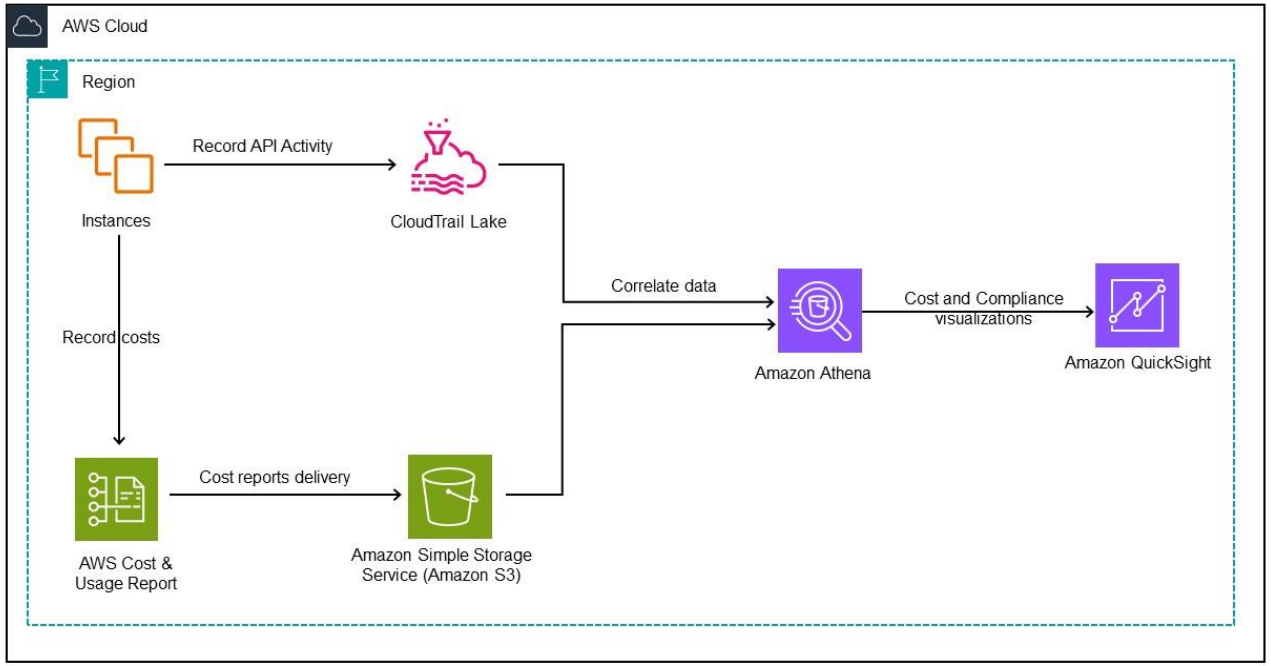
Create event data stores per event type. Supports direct queries, or Metadata can be populated into AWS Glue Data Catalog for querying with Amazon Athena.

Access authorization can be done based on tags. CloudTrail encrypts all events by default.

Visualization - CloudTrail Lake dashboards can be used to visualize the data, **or** Amazon Athena for more customizations and event type compatibility.

Event Types

- CloudTrail Events
- CloudTrail Insights events
- AWS Config configuration items
- AWS Audit Manager evidence
- Non-AWS events



Amazon CloudWatch

Observability of your AWS resources and applications on AWS and on-premises

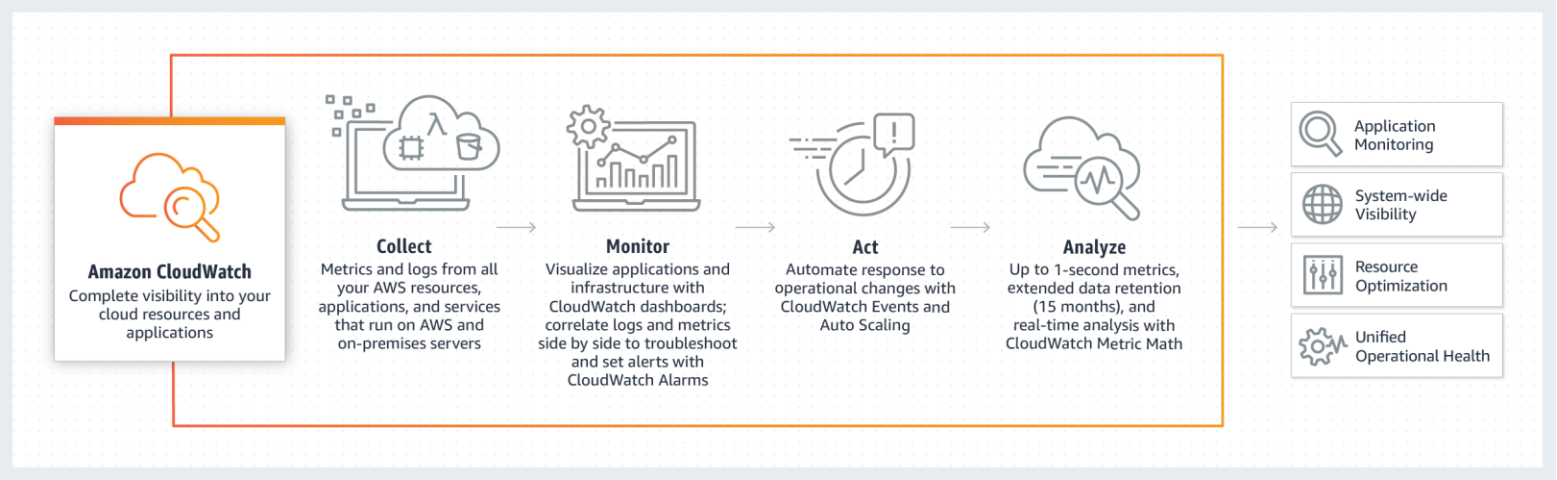
Data and insights to monitor your applications, respond to system-wide performance changes, optimize resource utilization, and get a unified view of operational health.

Infra. Monitoring

Monitor key metrics and logs, visualize your application and infrastructure stack, create alarms, and correlate metrics and logs

Scalability

Take action automatically to enable Amazon EC2 Auto Scaling or stop an instance, for example, so you can automate capacity and resource planning.



Amazon CloudWatch Logs

Monitor, store, and access your log files from EC2, AWS CloudTrail, Route 53 and other sources

Centralize the logs from all of your systems, applications, and AWS services that you use, in a single, highly scalable service.

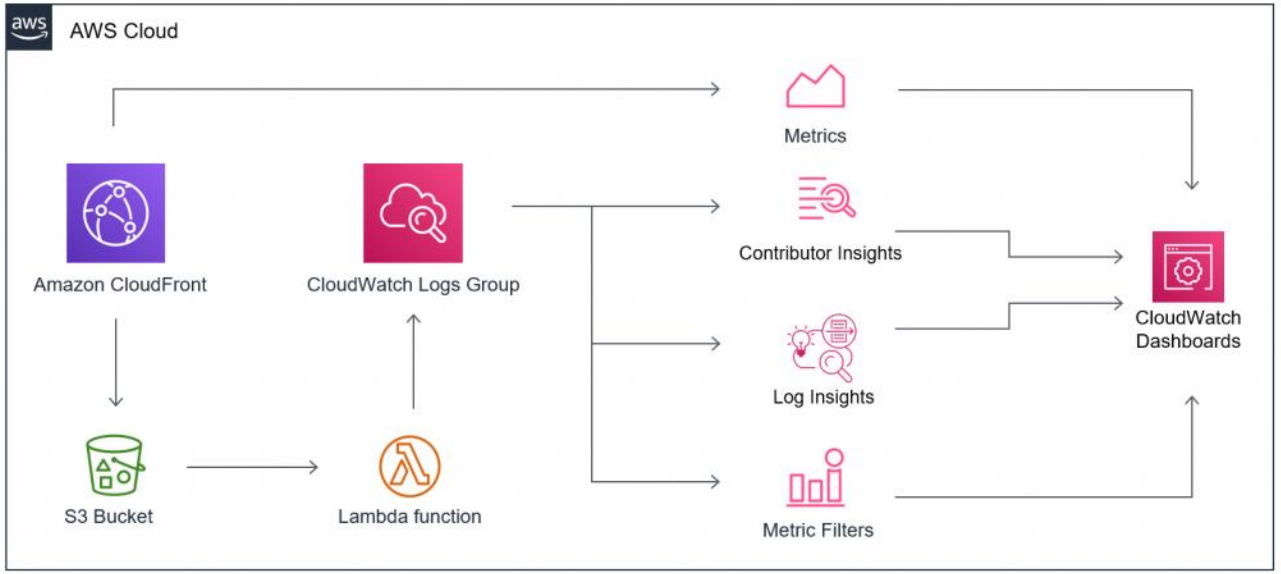
Features

Query – CloudWatch Logs Insights provides a purpose-built query language to search and analyze your log data.

EC2 Instances - Use to monitor applications and systems running on EC2 using log data.

Monitoring – CloudWatch can monitor the logs based on literal terms, which are counted as a CloudWatch metric and can generate Alerts.

Masking – Sensitive data in your logs can be audited and masked using **data protection policies.**



AWS PARTNER CERTIFICATION READINESS

Domain 4: Data Security & Governance

Understand data privacy and governance

Understand data privacy and governance

Knowledge of:

- How to protect personally identifiable information (PII)
- Data sovereignty

Skills in:

- Granting permissions for data sharing (for example, data sharing for Amazon Redshift)
- Implementing PII identification (for example, Macie with Lake Formation)
- Implementing data privacy strategies to prevent backups or replications of data to disallowed AWS Regions
- Managing configuration changes that have occurred in an account (for example, AWS Config)

AWS Config

Track resource inventory & changes

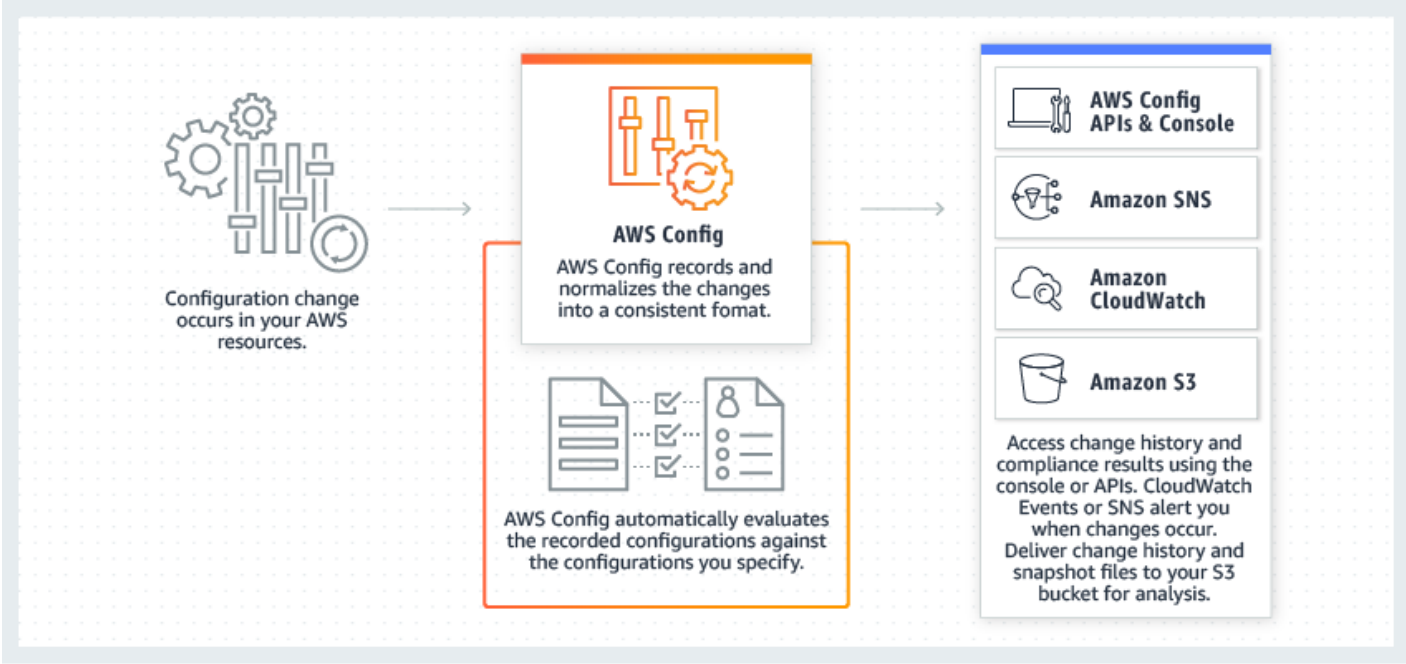
Assess, audit, and evaluate the configurations of AWS resources. Continuously monitors and records AWS resource configurations, allowing automated evaluation against desired configurations

Discovery

AWS Config will discover resources that exist in your account, record their current configuration, and capture any changes to these configurations

Change Management

When your resources are created, updated, or deleted, AWS Config streams these configuration changes to Amazon Simple Notification Service (SNS), so that you are notified of all the configuration changes



Amazon Macie

Amazon Macie continually evaluates your Amazon S3 environment and provides an S3 resource summary across all of your accounts.

Purpose

Amazon Macie uses machine learning and pattern matching to cost efficiently discover sensitive data at scale. Macie automatically detects a large and growing list of sensitive data types, including **personally identifiable information (PII)** such as names, addresses, and credit card numbers. It also gives you constant visibility of the data security and data privacy of **your data stored in Amazon S3.**

Easy to Deploy

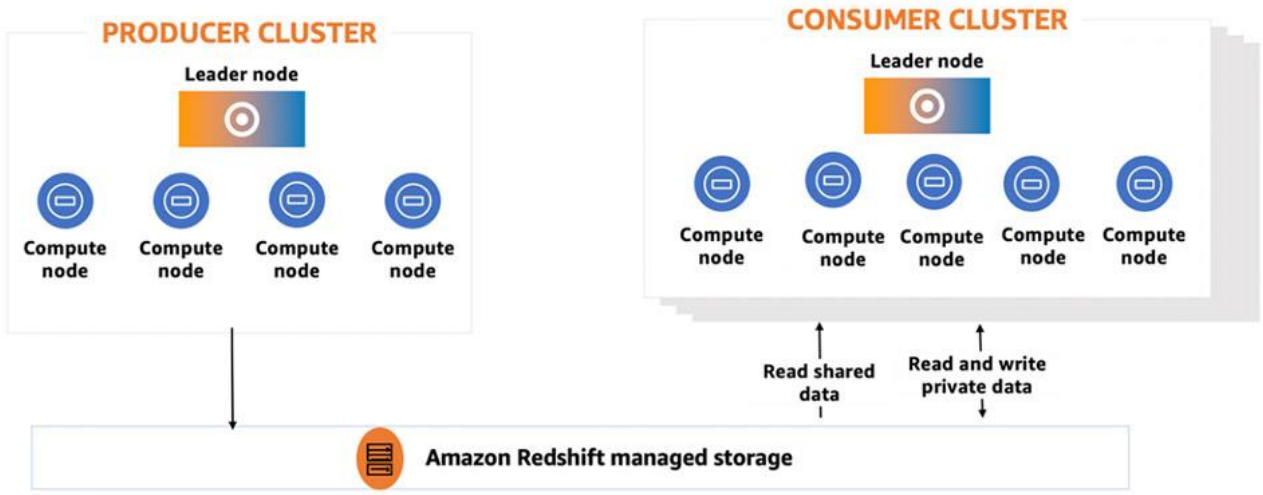
With **one-click** in the AWS Management Console or a single API call, you can enable Amazon Macie in a single account. With a **few more clicks** in the console, you can enable Macie **across multiple accounts**.

Data Sharing – Amazon Redshift

Securely share access to live data across Amazon Redshift clusters, workgroups, AWS accounts, and AWS Regions without manually moving or copying the data

Use Cases

- Supporting different kinds of business-critical workloads
- Enabling cross-group collaboration
- Delivering data as a service
- Sharing data between environments
- Licensing access to data in Amazon Redshift



Thank you for attending this session