AWS WORKSHOP 1

Personal File Backup with AWS S3 Bucket

MARCH 14, 2025

NGUYEN TAI MINH HUY
DINH LE HOANG ANH

Executive Summary

Managing and backing files efficiently is a common challenge in academic and project-based environments. On-premise storage solutions often suffer from high maintenance costs, scalability issues, and data loss. Although traditional cloud storage services such as Google Drive and OneDrive are convenient, they are costly, lack detailed security controls, and offer limited flexibility for custom access management.

The primary objective of this project is to address these challenges by guiding readers on how to create a Secure Personal File Backup System using AWS S3 Bucket. The system employs AWS IAM and MFA to implement rigorous access control, thereby mitigating the likelihood of unauthorized access. Lifecycle management policies transition older files to AWS S3 Glacier for cost optimization, while automated backup mechanism, versioning, and Object Lock features help prevent accidental data loss. In addition to the implementation of AWS CloudTrail and AWS KMS encryption for further enhance security.

Contents

1. Project Background

Efficient file storage and sharing among team members pose considerable challenges when relying on personal on-premise databases or traditional cloud storage solutions. On-premise databases require high maintenance costs, limited scalability, and are prone to data loss due to mismanagement or insufficient backup strategies. This kind of setting makes are not realistic for flexible needs of the team because adding more storage takes a huge amount of time and effort. In addition, having strong controls on user access and security can be complex, which can make working together less efficient and pose potentially damaging data integrity.

Although tradition cloud storage systems such as Google Drive, One Drive, MegaUp, etc… are easier to use, they do have their own problems. Many platforms enforce fixed pricing structures that may not align with actual usage needs. The absence of detailed activity logs makes it harder to track file modifications, and role-based access control systems that fail to perform adequately can let unauthorized data to be exposed. Furthermore, the inconsistent download and upload bandwidth resulting from an uncontrolled database location may hinder the team's working progress.

2. Goals

To mitigate these problems, this project aims to implement a Secure Personal File Backup System on AWS S3 Bucket. AWS S3 provides a highly secure, scalable and cost-effective cloud storage facility that guarantees safe data handling and smooth file-sharing processes. The primary objectives of this project are:

- Implementation of strict access controls and encryption to avoid unauthorized changes in data and data breaches.

- Implementation of secure and efficient file-sharing processes with fine-grained permission control.
- Implementing automated backup policies, versioning, and lifecycle management to reduce data loss risk.
- Designing a storage solution with the scalability needed to accommodate future team growth and shifting project requirements.
- Reducing storage cost through AWS's pricing model based on cost-effectiveness.

3. Scope

This is a design and implementation project for a secure AWS S3-based storage infrastructure that is specifically tailored for internal team use. The project includes the provisioning and configuration of S3 buckets with cost-optimized storage classes and appropriate security controls, deploying IAM roles and policies for managing user permissions, and maintaining data integrity by applying strict access controls. Apart from this, object versioning, lifecycle policies, and automatic backup mechanisms will be established to strengthen the data management and recovery process. The solution will also be designed to ensure compatibility with other AWS services for any eventual integrations in the future, while being modular in design to facilitate simple expansion as team requirements evolve.

4. Services

The project leverages a suite of AWS services:

- **AWS S3 Standard Bucket:** High-availability storage for frequently accessed data.
- **AWS S3 Glacier Bucket:** Cost-effective archival storage for long-term data.
- **AWS CLI:** Command-line interface for efficient management and automation.
- **AWS IAM:** Identity and access management system to enforce secure access control.
- **AWS CloudTrail:** Logging and monitoring service for improved security and compliance tracking.
- **AWS KMS:** Encryption service to protect data from unauthorized access.

5. Diagrams

5.1. Data Flowchart

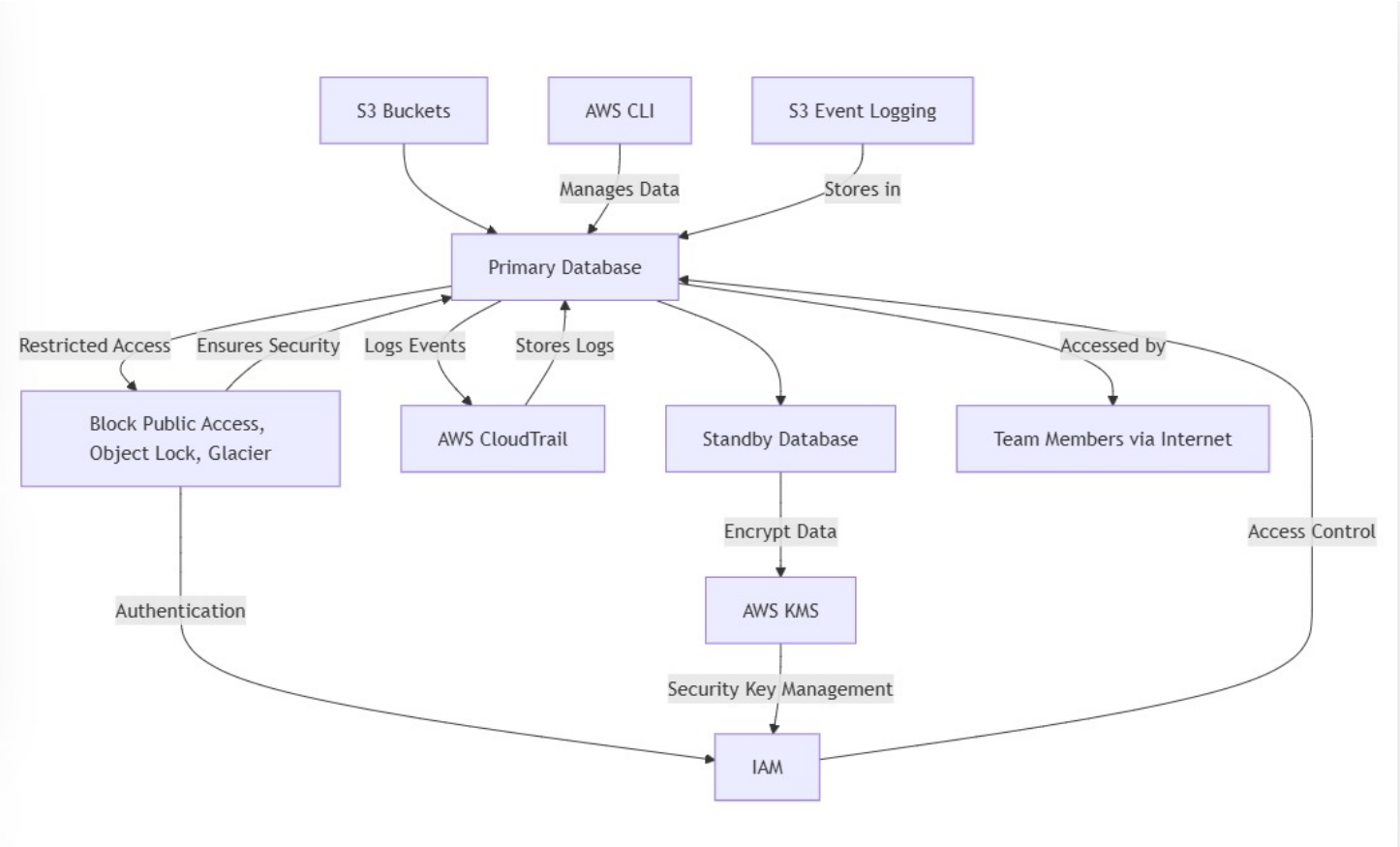


Figure 1. Data Flowchart for the Personal Backup System

This system leverages AWS S3 to provide a secure, scalable, and cost-effective file backup solution. It uses the Primary Database as a storage location and keeps the data intact with the help of AWS CLI, and for security, it stores event logs in AWS CloudTrail. The access to this system is highly limited, with IAM roles and Multi-Factor Authentication (MFA) for authenticated users only to interact with the system.

For additional protection, S3 Bucket policies, Object Lock, and Public Access Restrictions prevent unauthorized modifications or accidental deletions. Inactive files are automatically transferred to S3 Glacier after a certain period using lifecycle management to reduce storage costs. All data at rest is encrypted by AWS KMS, and decryption access is controlled through IAM policies.

The solution provides team members with secure file storage and retrieval over the internet, with detailed access control mechanisms to apply security policies. Full tracking of changes to data and access is enabled through S3 Event Logging and CloudTrail monitoring.

The AWS-based backup solution offers high availability, strong security, compliance monitoring, and cost savings that make it ideal for teams requiring scalable, efficient, and reliable file management.

5.2. AWS Architecture Diagram

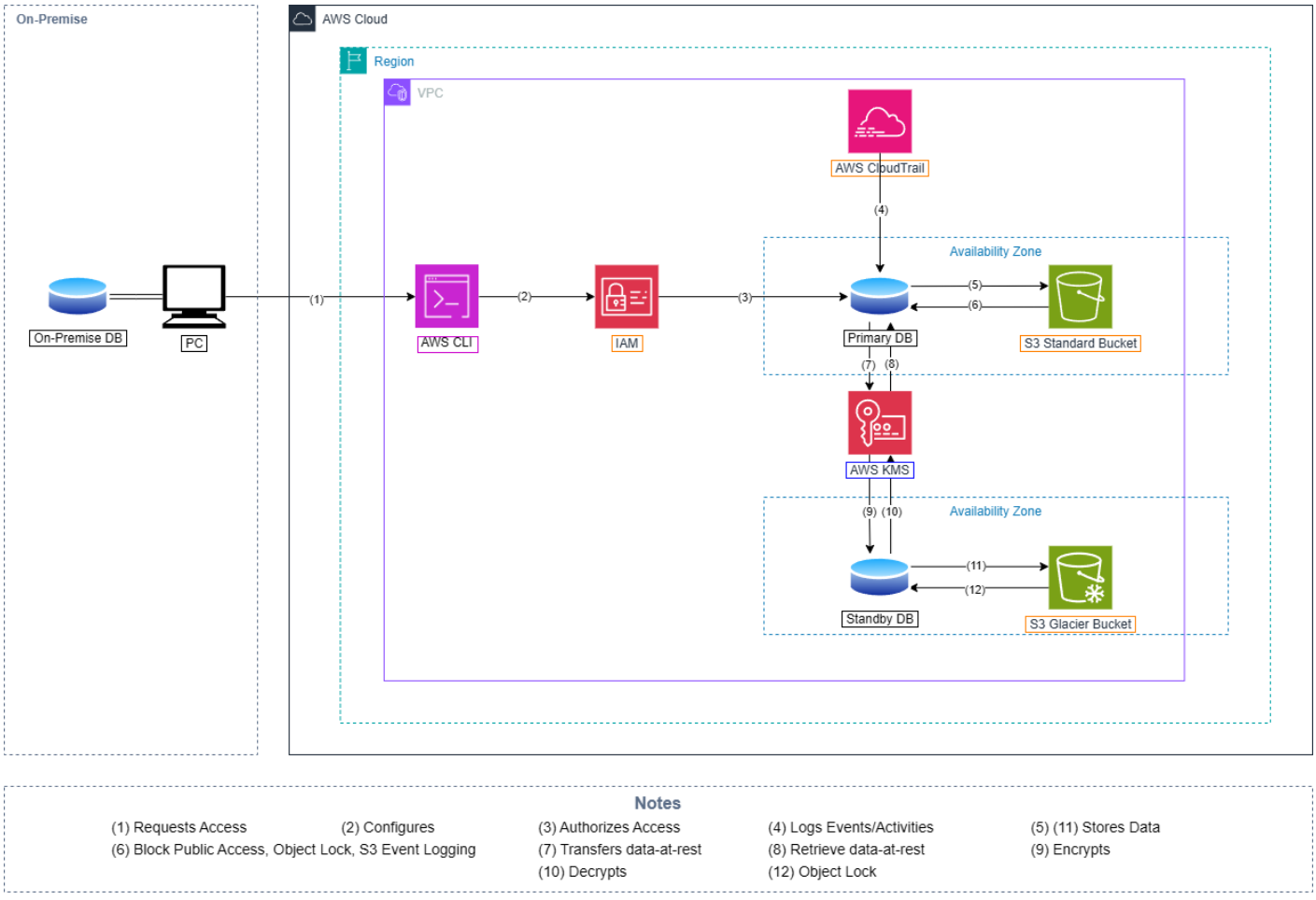


Figure 2. AWS Architecture Diagram for Personal Backup System

The above diagram illustrates the data flow and security mechanisms implemented in this backup system. Team members with an on-premise database and PC initiate access to the backup server using AWS CLI. Before interacting with the primary database, they must

authenticate and be authorized via AWS IAM. Once access is granted, data is securely stored in an AWS S3 Standard Bucket, where public access is blocked, Object Lock is enforced, and S3 Event Logging is enabled to monitor changes. For the purposes of auditing and security compliance, AWS CloudTrail is integrated to record all access events and activities. As per the configured lifecycle policy, any data that remains inactive for 14 days is automatically transferred to a standby database using AWS KMS encryption. In order to retrieve data from the standby database, decryption via AWS KMS is required. The standby database is stored in an AWS S3 Glacier Bucket, located in a different availability zone, with Object Lock enabled to prevent accidental deletions.

6. Creation Process

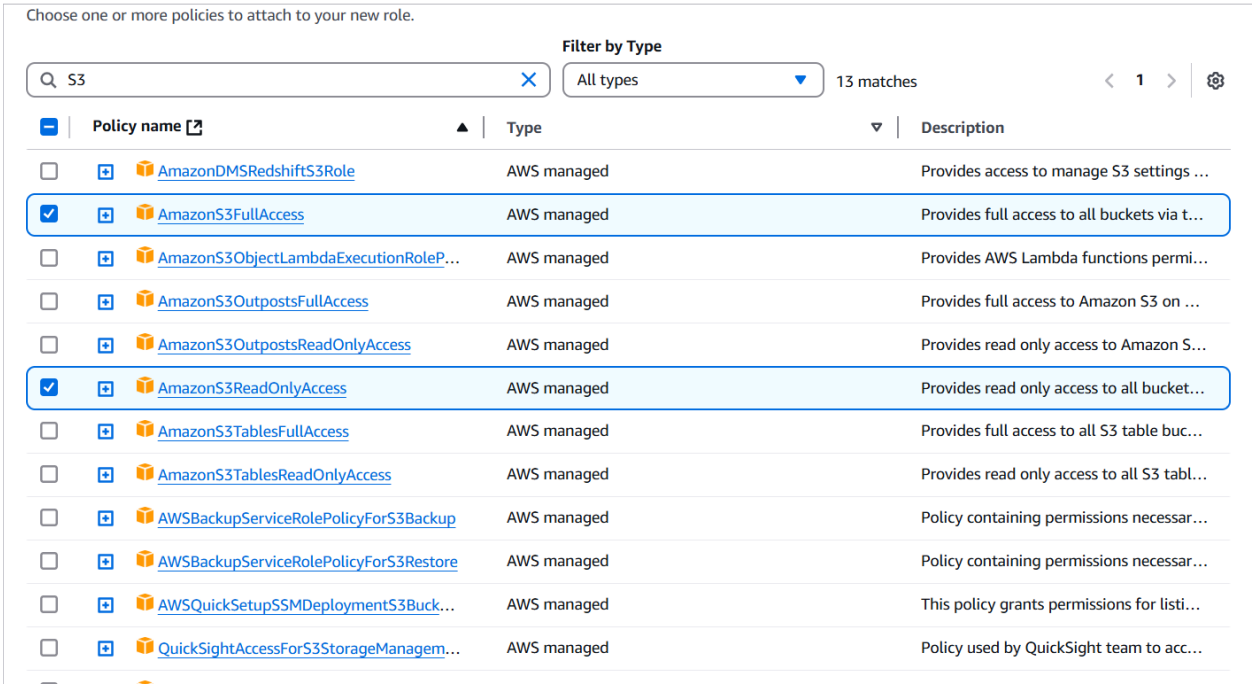
6.1. IAM Role Assignment and MFA for Database Access

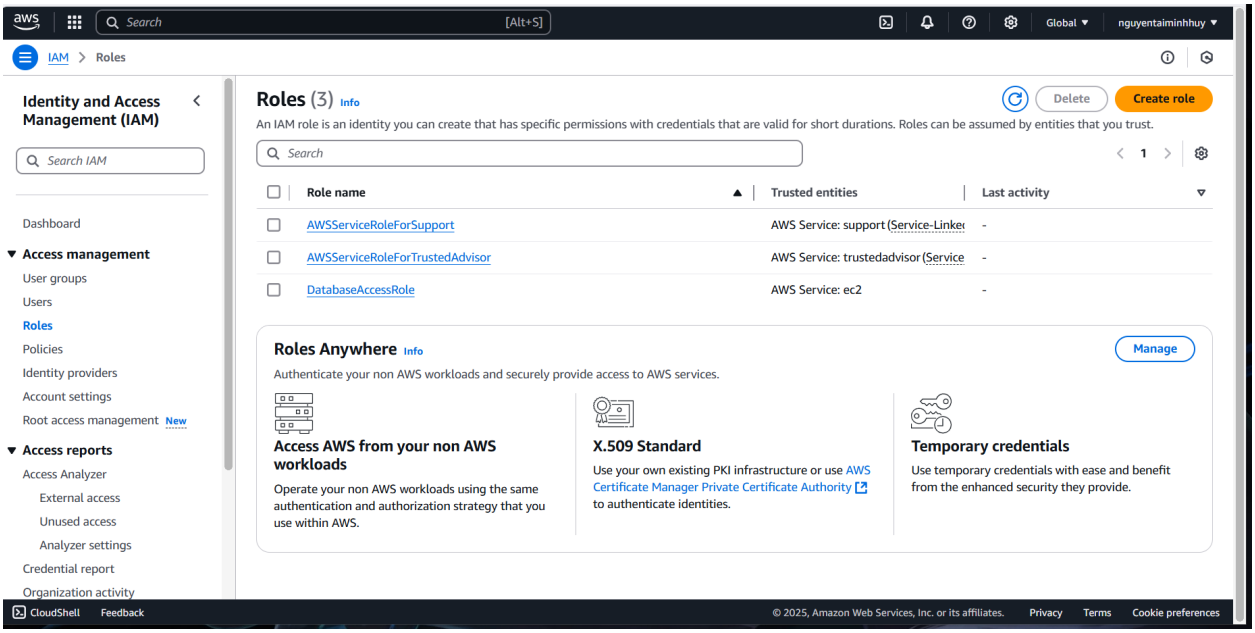
To ensure only authorized team members can access the database, IAM roles and Multi-Factor Authentication (MFA) are enforced.

Implementation Steps:

1. **Create IAM Roles:** Assign IAM roles to users needing database access.
2. **Enable MFA:** Require users to authenticate with an additional security factor before accessing the database.
3. **Restrict Access Policies:** Apply IAM policies restricting database access only to authorized users.
4. **Audit IAM Policies:** Regularly review IAM permissions to prevent excessive privileges.

AWS Services Used: IAM


Choose one or more policies to attach to your new role.
Filter by Type
S3
All types
13 matches
1
Policy name
Type
Description
AmazonDMSRedshiftS3Role
AWS managed
Provides access to manage S3 settings ...
AmazonS3FullAccess
AWS managed
Provides full access to all buckets via t...
AmazonS3ObjectLambdaExecutionRoleP...
AWS managed
Provides AWS Lambda functions permi...
AmazonS3OutpostsFullAccess
AWS managed
Provides full access to Amazon S3 on ...
AmazonS3OutpostsReadOnlyAccess
AWS managed
Provides read only access to Amazon S...
AmazonS3ReadOnlyAccess
AWS managed
Provides read only access to all bucket...
AmazonS3TablesFullAccess
AWS managed
Provides full access to all S3 table buc...
AmazonS3TablesReadOnlyAccess
AWS managed
Provides read only access to all S3 tabl...
AWSBackupServiceRolePolicyForS3Backup
AWS managed
Policy containing permissions necessar...
AWSBackupServiceRolePolicyForS3Restore
AWS managed
Policy containing permissions necessar...
AWSQuickSetupSSMDeploymentS3Buck...
AWS managed
This policy grants permissions for listi...
QuickSightAccessForS3StorageManagem...
AWS managed
Policy used by QuickSight team to acc...


Search
[Alt+S]
Global
nguyentaiminhhuy
IAM > Roles
Identity and Access Management (IAM)
Search IAM
Dashboard
Access management
User groups
Users
Roles
Policies
Identity providers
Account settings
Root access management New
Access reports
Access Analyzer
External access
Unused access
Analyzer settings
Credential report
Organization activity
Roles (3) Info
An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.
Delete
Create role
Search
1
Role name
Trusted entities
Last activity
AWSServiceRoleForSupport
AWS Service: support (Service-Linke
-
AWSServiceRoleForTrustedAdvisor
AWS Service: trustedadvisor (Service
-
DatabaseAccessRole
AWS Service: ec2
-
Roles Anywhere Info
Manage
Authenticate your non AWS workloads and securely provide access to AWS services.
Access AWS from your non AWS workloads
Operate your non AWS workloads using the same authentication and authorization strategy that you use within AWS.
X.509 Standard
Use your own existing PKI infrastructure or use AWS Certificate Manager Private Certificate Authority to authenticate identities.
Temporary credentials
Use temporary credentials with ease and benefit from the enhanced security they provide.
CloudShell
Feedback
© 2025, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

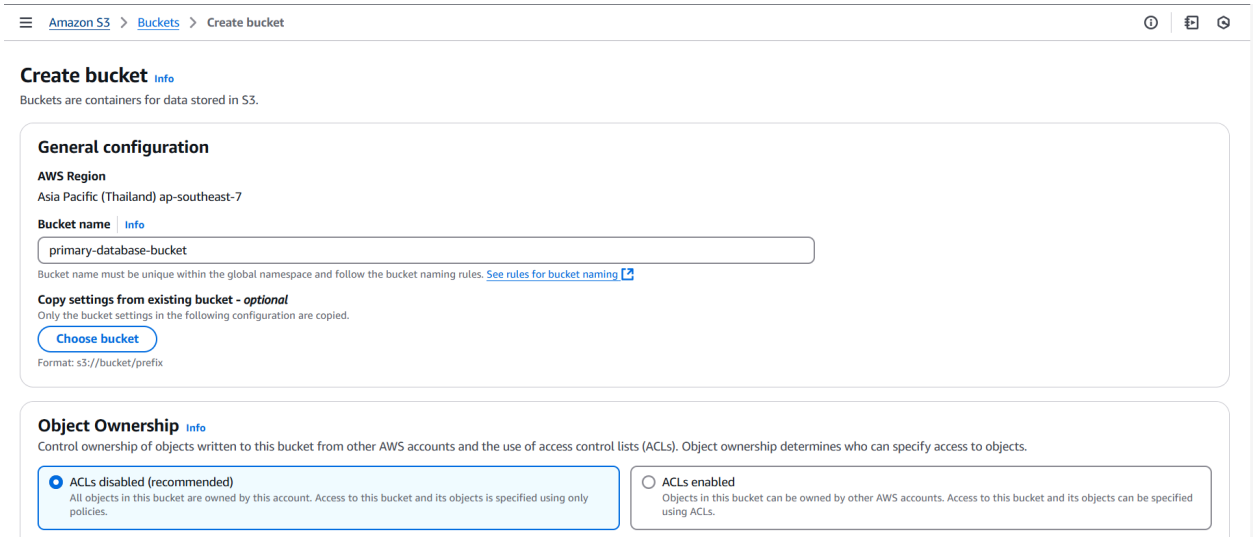
6.2. S3 Bucket Security (Primary Database)

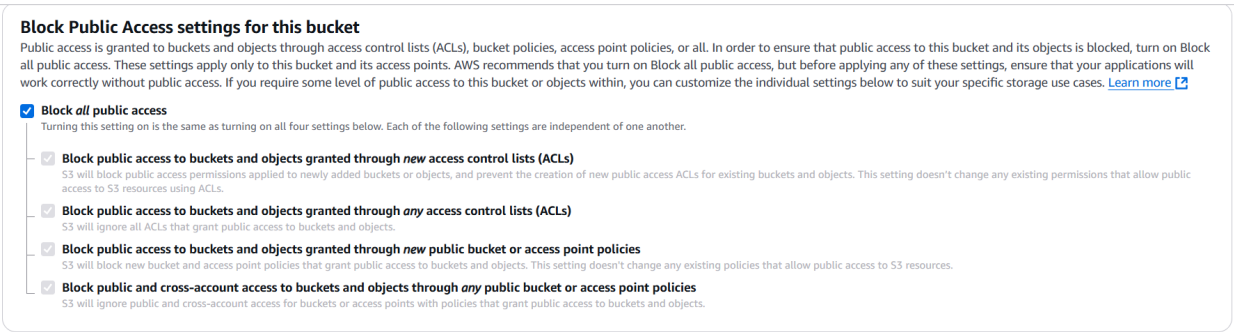
To prevent unauthorized access and accidental deletion, S3 bucket security policies are enforced.

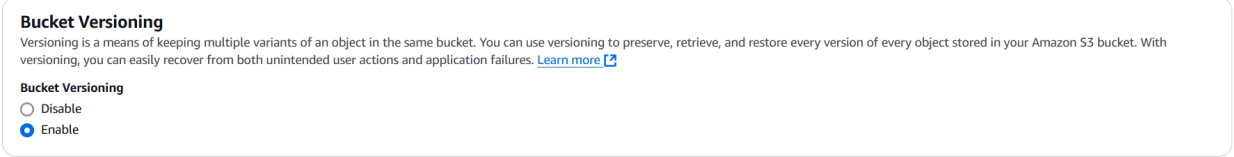
Implementation Steps:

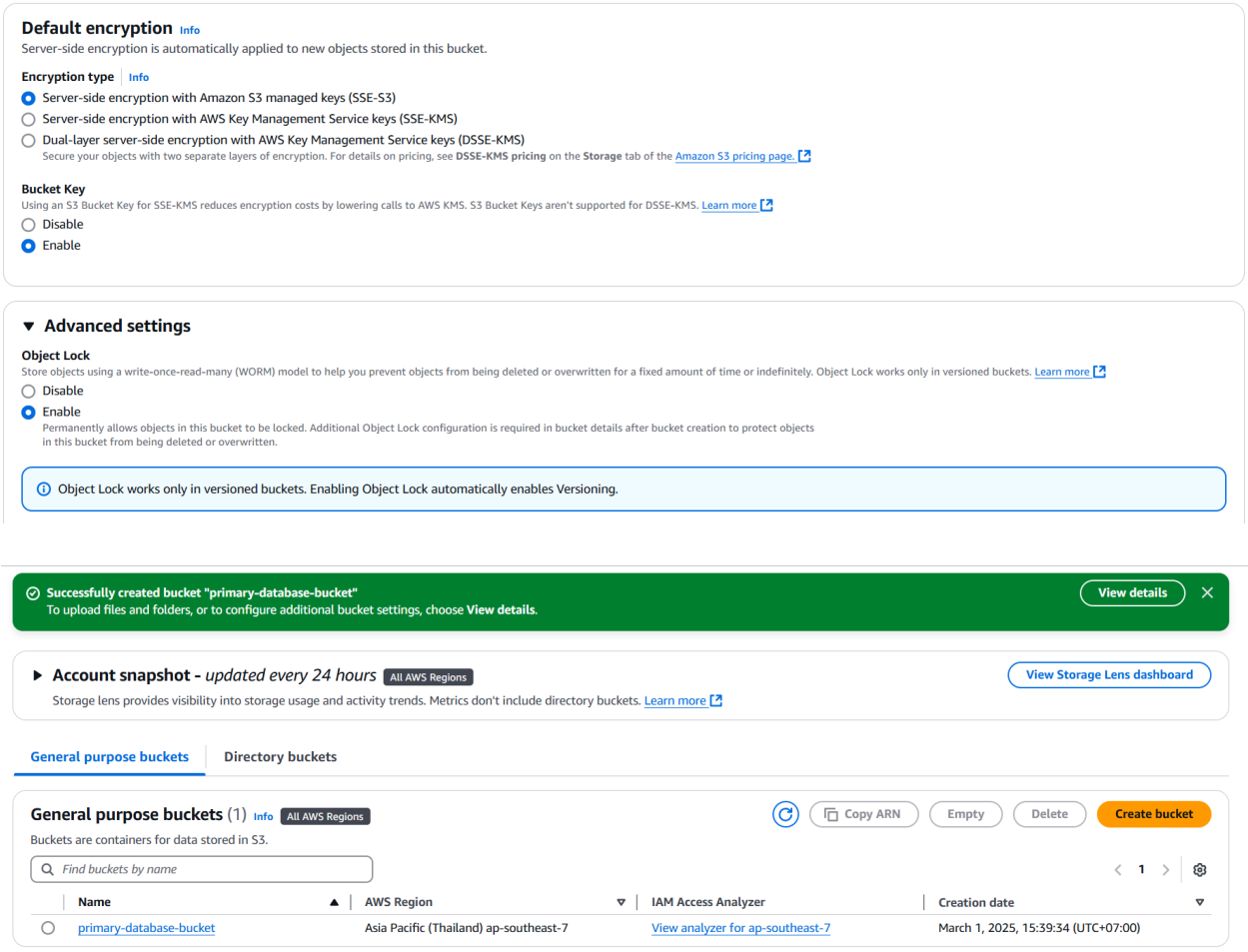
1. **Block Public Access:** Enable the "Block Public Access" setting for the S3 bucket.
2. **Enable Versioning:** Configure versioning to retain previous versions of objects.
3. **Enable Object Lock:** Prevent permanent deletion of objects to protect against accidental removal.

AWS Services Used: Amazon S3, IAM









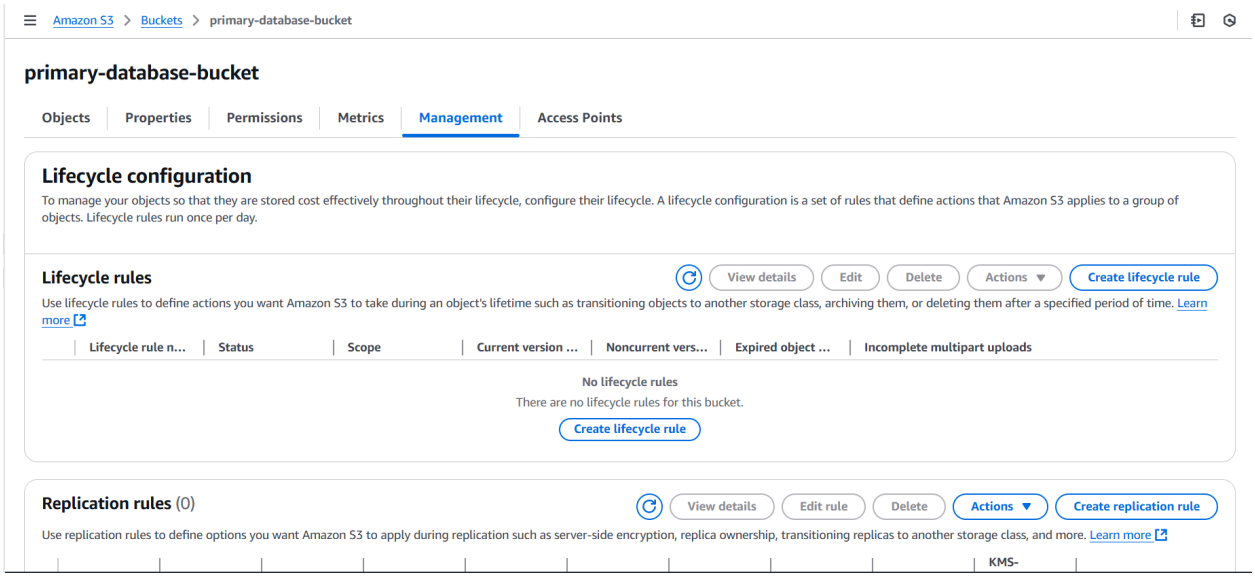
6.3. Automated Data Lifecycle Management

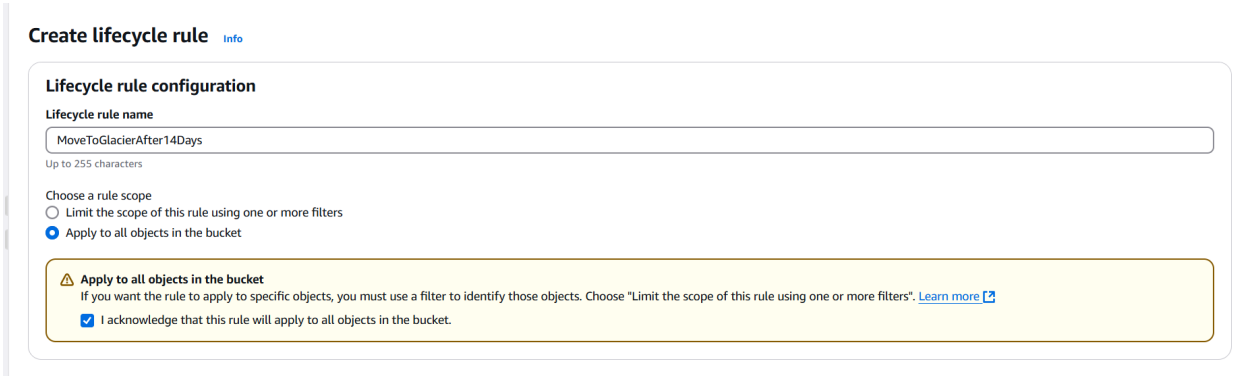
Files in the Standby Database are moved to S3 Glacier after 14 days to optimize storage costs and retention policies.

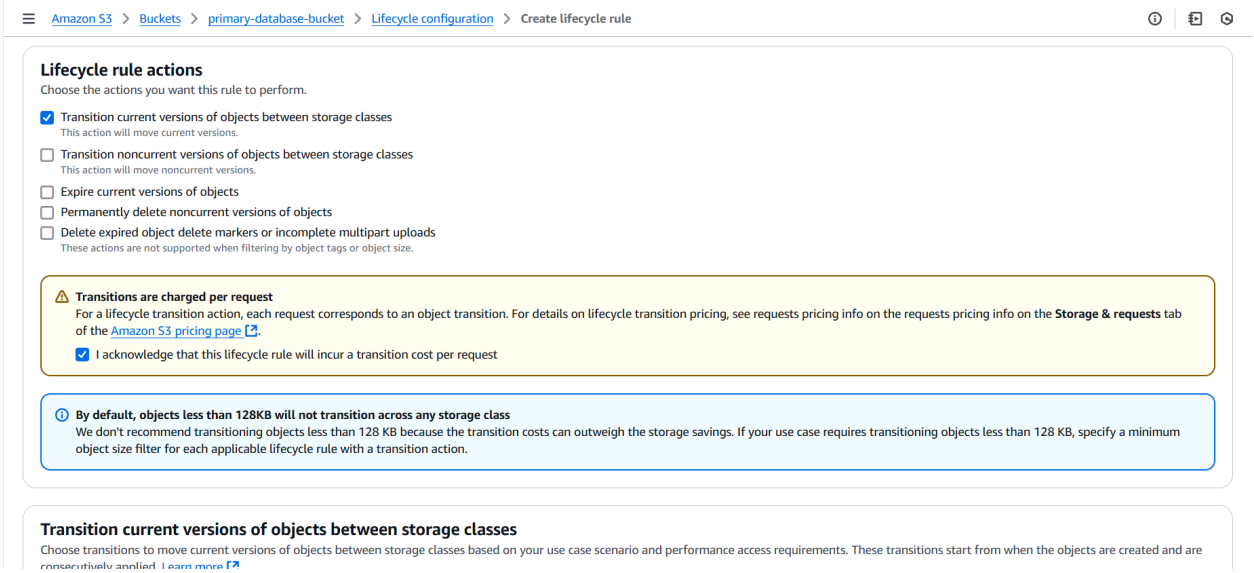
Implementation Steps:

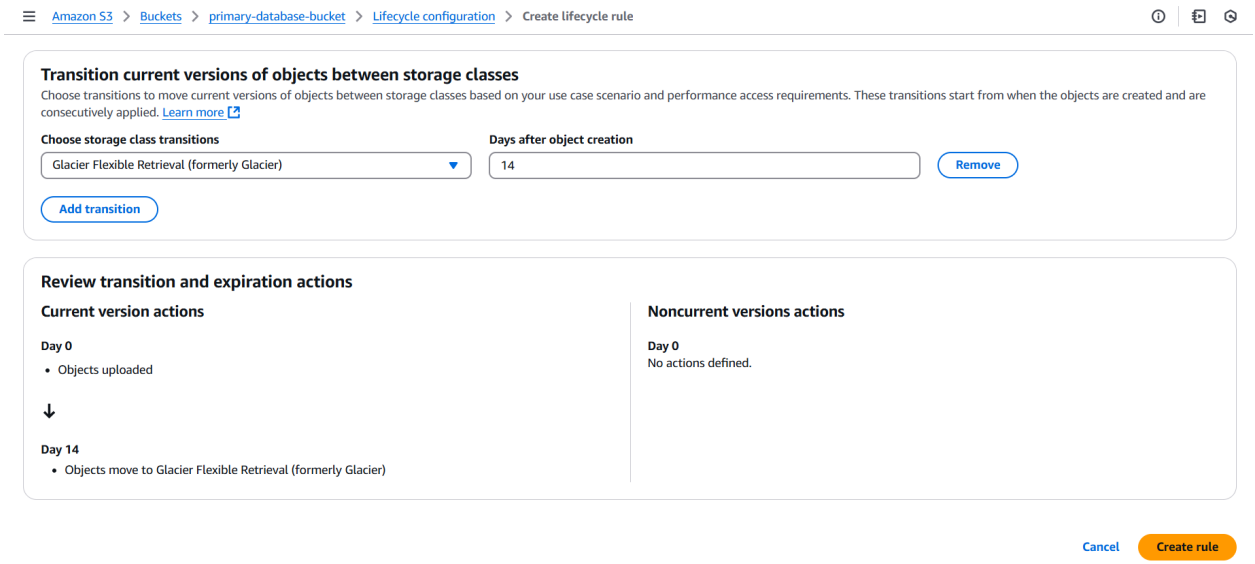
1. **Create Lifecycle Rule:** Define an S3 lifecycle rule to transition objects to Glacier.
2. **Set Transition Period:** Move files to Glacier after 14 days.
3. **Enable Expiration Policy:** Optionally configure deletion policies for obsolete files.

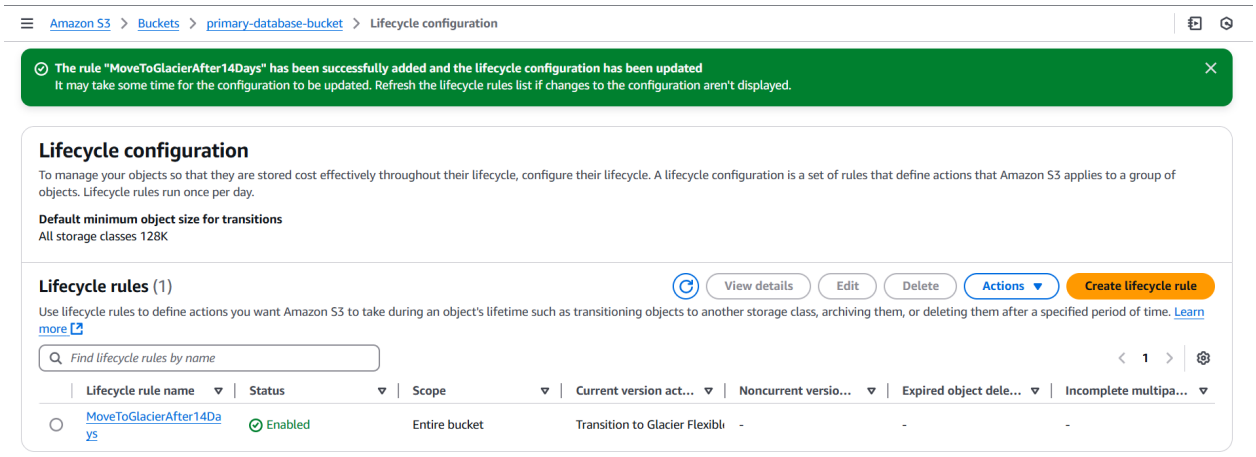
AWS Services Used: Amazon S3, S3 Glacier


Amazon S3 > Buckets > primary-database-bucket
primary-database-bucket
Objects
Properties
Permissions
Metrics
Management
Access Points
Lifecycle configuration
To manage your objects so that they are stored cost effectively throughout their lifecycle, configure their lifecycle. A lifecycle configuration is a set of rules that define actions that Amazon S3 applies to a group of objects. Lifecycle rules run once per day.
Lifecycle rules
View details
Edit
Delete
Actions
Create lifecycle rule
Use lifecycle rules to define actions you want Amazon S3 to take during an object's lifetime such as transitioning objects to another storage class, archiving them, or deleting them after a specified period of time. Learn more
Lifecycle rule n...
Status
Scope
Current version ...
Noncurrent vers...
Expired object ...
Incomplete multipart uploads
No lifecycle rules
There are no lifecycle rules for this bucket.
Create lifecycle rule
Replication rules (0)
View details
Edit rule
Delete
Actions
Create replication rule
Use replication rules to define options you want Amazon S3 to apply during replication such as server-side encryption, replica ownership, transitioning replicas to another storage class, and more. Learn more
KMS-


Create lifecycle rule Info
Lifecycle rule configuration
Lifecycle rule name
MoveToGlacierAfter14Days
Up to 255 characters
Choose a rule scope
Limit the scope of this rule using one or more filters
Apply to all objects in the bucket
Apply to all objects in the bucket
If you want the rule to apply to specific objects, you must use a filter to identify those objects. Choose "Limit the scope of this rule using one or more filters". Learn more
I acknowledge that this rule will apply to all objects in the bucket.


Amazon S3 > Buckets > primary-database-bucket > Lifecycle configuration > Create lifecycle rule
Lifecycle rule actions
Choose the actions you want this rule to perform.
Transition current versions of objects between storage classes
This action will move current versions.
Transition noncurrent versions of objects between storage classes
This action will move noncurrent versions.
Expire current versions of objects
Permanently delete noncurrent versions of objects
Delete expired object delete markers or incomplete multipart uploads
These actions are not supported when filtering by object tags or object size.
Transitions are charged per request
For a lifecycle transition action, each request corresponds to an object transition. For details on lifecycle transition pricing, see requests pricing info on the requests pricing info on the Storage & requests tab of the Amazon S3 pricing page.
I acknowledge that this lifecycle rule will incur a transition cost per request
By default, objects less than 128KB will not transition across any storage class
We don't recommend transitioning objects less than 128 KB because the transition costs can outweigh the storage savings. If your use case requires transitioning objects less than 128 KB, specify a minimum object size filter for each applicable lifecycle rule with a transition action.
Transition current versions of objects between storage classes
Choose transitions to move current versions of objects between storage classes based on your use case scenario and performance access requirements. These transitions start from when the objects are created and are consecutively applied. Learn more





6.4. S3 Event Logging with AWS CloudTrail

Audit and monitor all activities in the database by enabling S3 event logging.

Implementation Steps:

1. **Create an AWS CloudTrail Trail:** Configure a new trail to capture S3 data events.
2. **Enable S3 Event Logging:** Select "Data Events" and specify the S3 bucket.
3. **Store Logs Securely:** Store logs in a dedicated S3 bucket with restricted access.

AWS Services Used: AWS CloudTrail, Amazon S3

6.5. Enable KMS Encryption for Standby Database

To enhance data-at-rest security, AWS KMS encryption is enabled for the Standby Database.

Implementation Steps:

1. **Create a KMS Key:** Generate a new customer-managed KMS key.
2. **Attach Encryption Policy:** Apply a key policy allowing database services access.
3. **Encrypt Database:** Enable encryption at rest for the Standby Database.

AWS Services Used: AWS Key Management Service (KMS), Amazon S3, Amazon RDS

6.6. Secure Database Access with AWS CLI

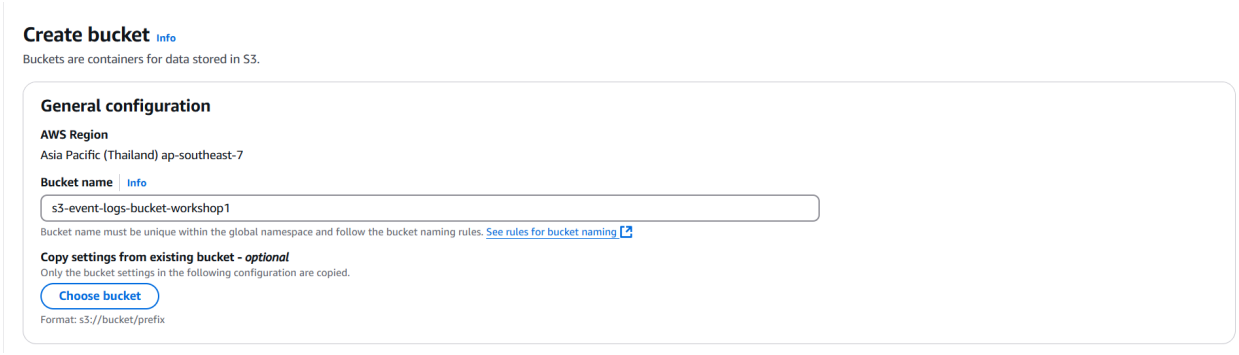
The AWS CLI is used as the primary method for managing the database securely.

Implementation Steps:

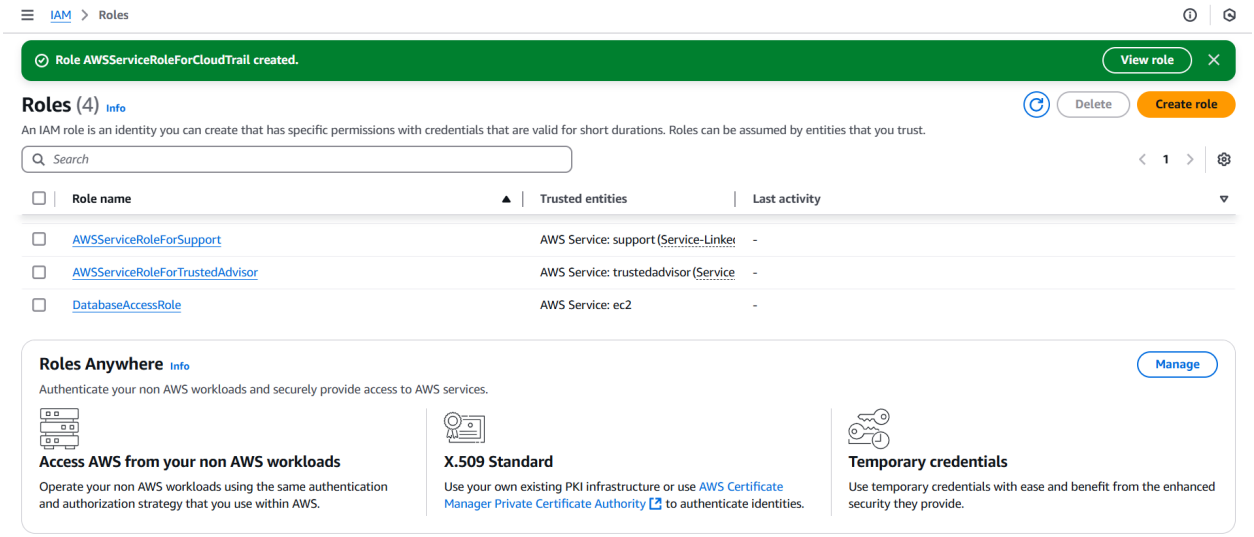
1. **Install and Configure AWS CLI:** Install AWS CLI and configure credentials.
2. **Use IAM Roles for Authentication:** Ensure IAM roles and MFA are used.
3. **Restrict CLI Access to Database:** Apply security policies to limit CLI database access.
4. **Enable CloudTrail for CLI Monitoring:** Monitor CLI actions with AWS CloudTrail logs.

AWS Services Used: AWS CLI, IAM, AWS CloudTrail

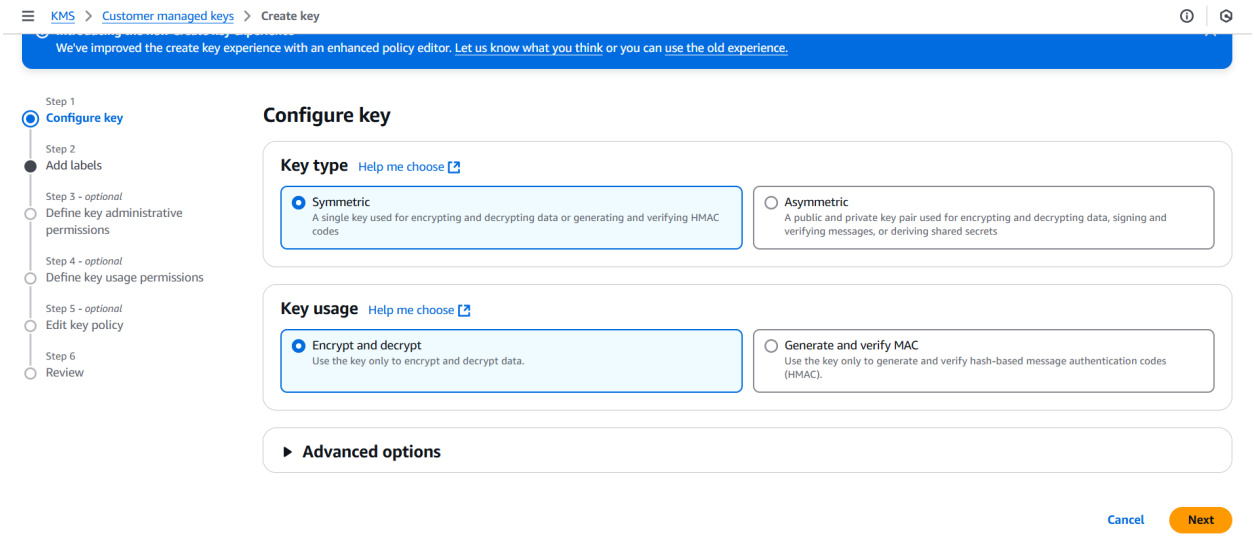
Bucket for CloudTrail:

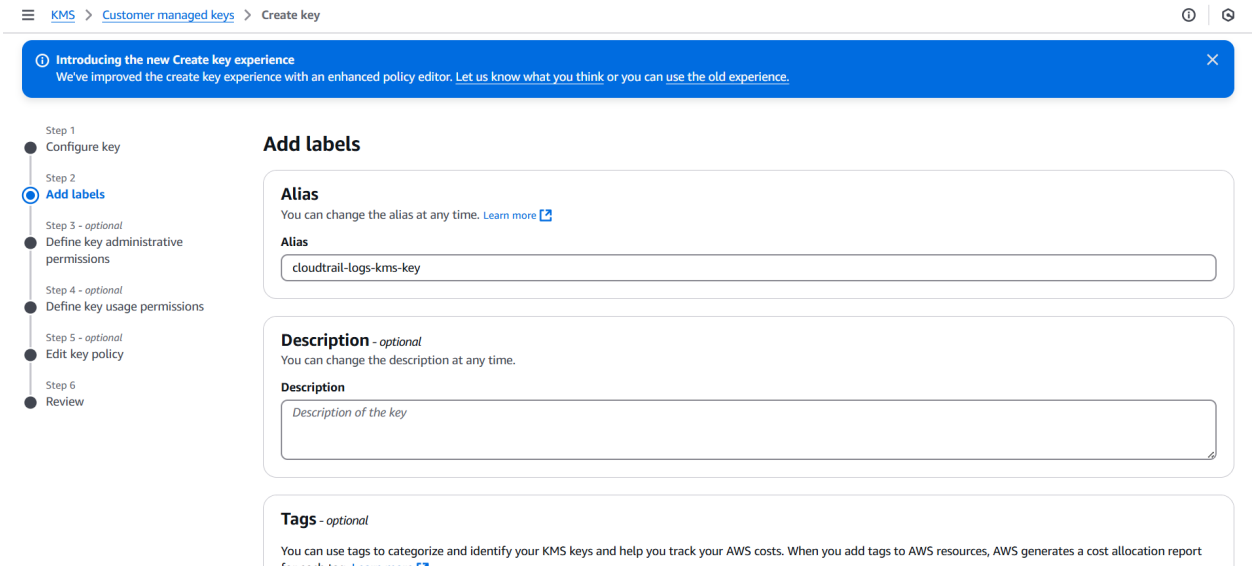


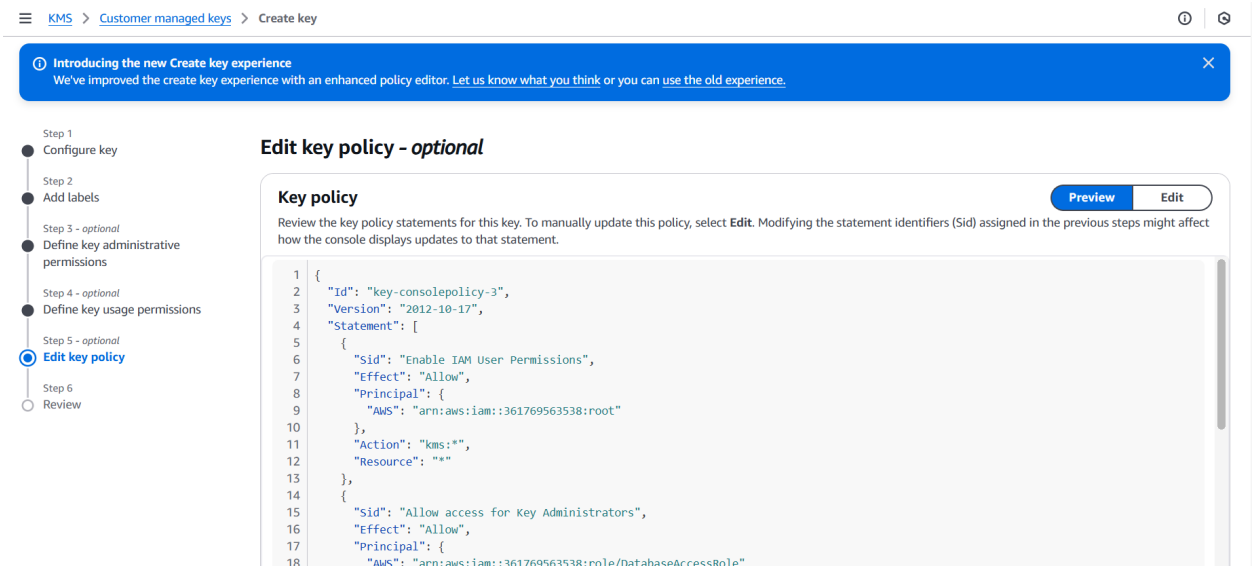
IAM Role for CloudTrail:

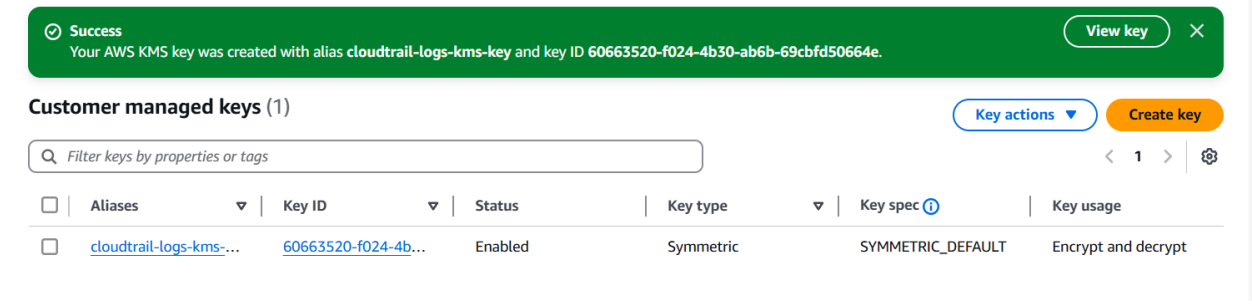


KMS Key:

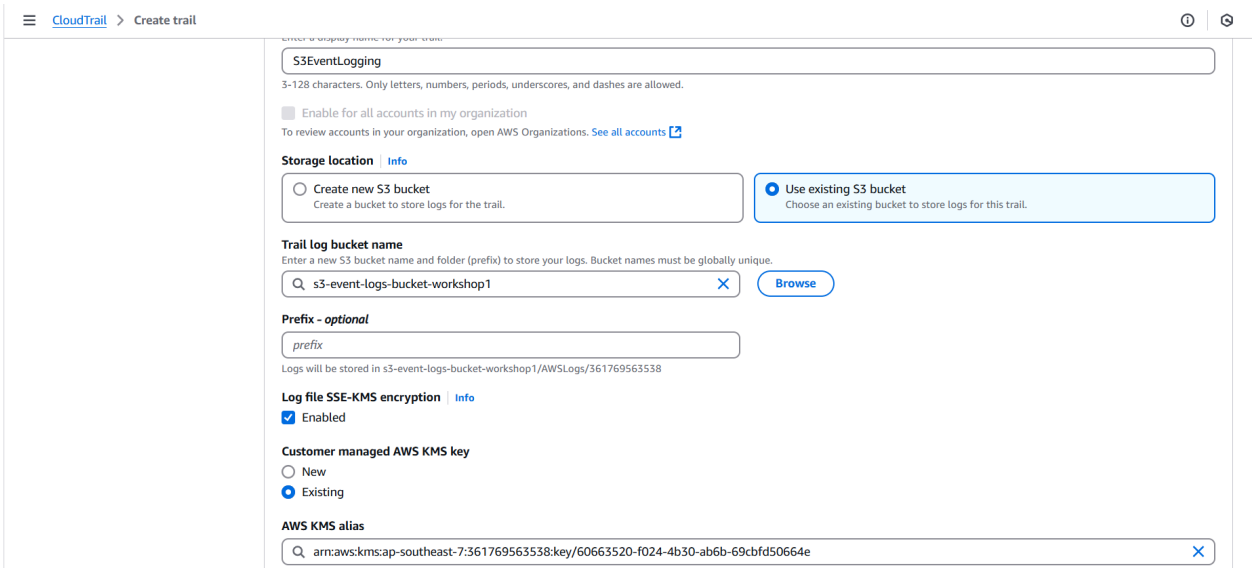


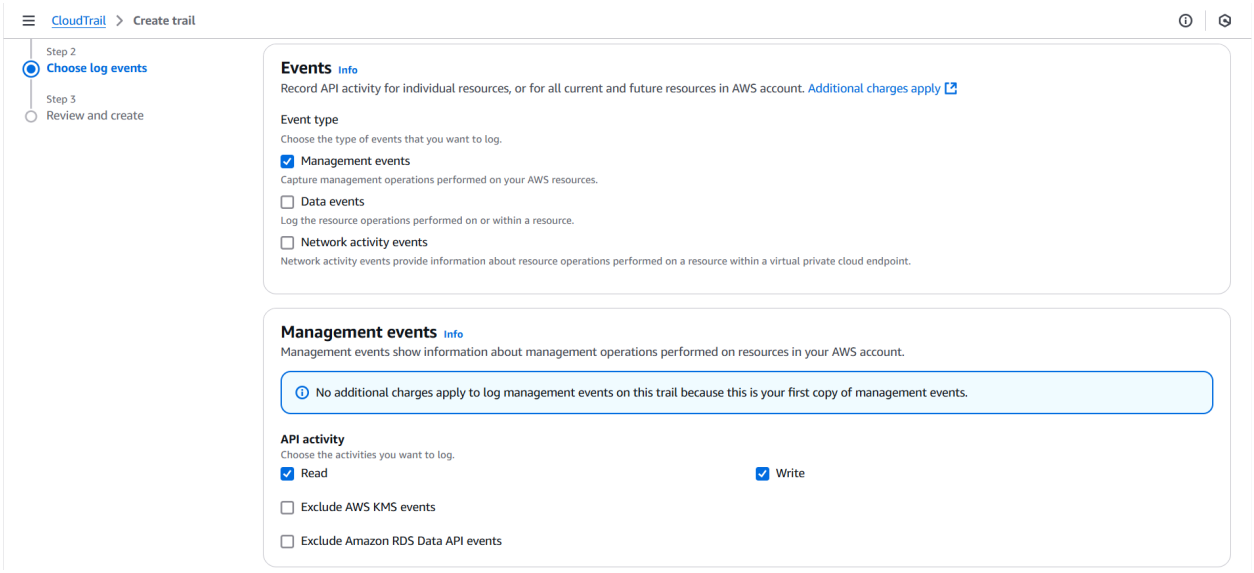

KMS > Customer managed keys > Create key
Introducing the new Create key experience
We've improved the create key experience with an enhanced policy editor. Let us know what you think or you can use the old experience.
Step 1 Configure key
Step 2 Add labels
Step 3 - optional Define key administrative permissions
Step 4 - optional Define key usage permissions
Step 5 - optional Edit key policy
Step 6 Review
Add labels
Alias
You can change the alias at any time. Learn more
Alias
cloudtrail-logs-kms-key
Description - optional
You can change the description at any time.
Description
Description of the key
Tags - optional
You can use tags to categorize and identify your KMS keys and help you track your AWS costs. When you add tags to AWS resources, AWS generates a cost allocation report

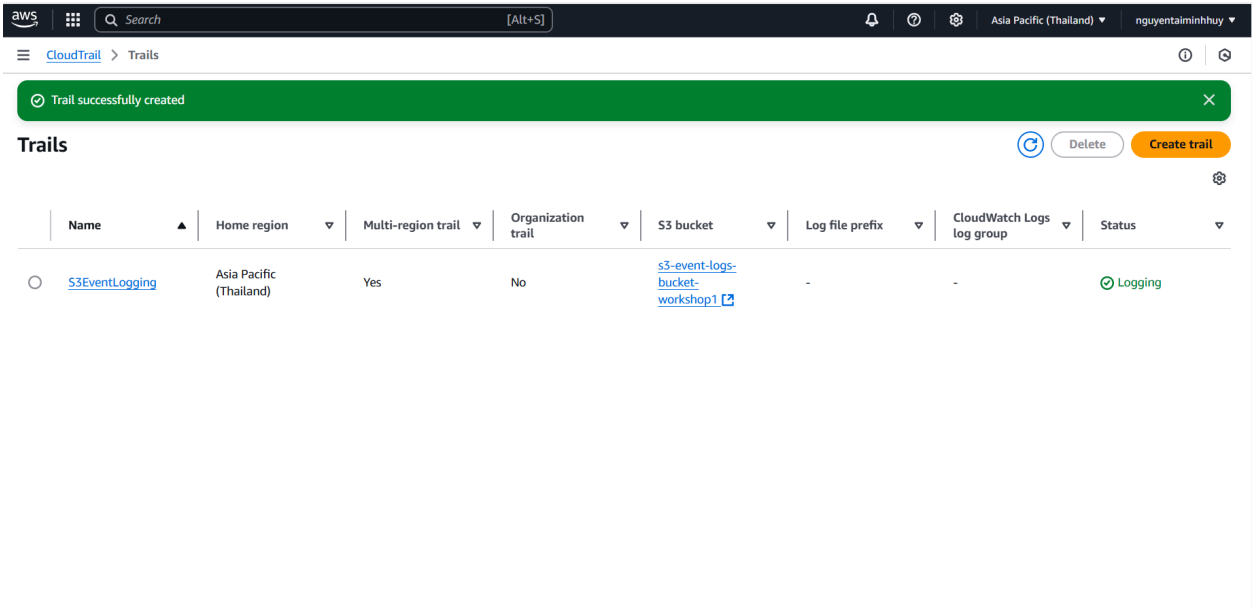

KMS > Customer managed keys > Create key
Introducing the new Create key experience
We've improved the create key experience with an enhanced policy editor. Let us know what you think or you can use the old experience.
Step 1 Configure key
Step 2 Add labels
Step 3 - optional Define key administrative permissions
Step 4 - optional Define key usage permissions
Step 5 - optional Edit key policy
Step 6 Review
Edit key policy - optional
Key policy
Preview
Edit
Review the key policy statements for this key. To manually update this policy, select Edit. Modifying the statement identifiers (Sid) assigned in the previous steps might affect how the console displays updates to that statement.
{
"Id": "key-consolepolicy-3",
"Version": "2012-10-17",
"Statement": [
{
"Sid": "Enable IAM User Permissions",
"Effect": "Allow",
"Principal": {
"AWS": "arn:aws:iam::361769563538:root"
},
"Action": "kms:*",
"Resource": "*"
},
{
"Sid": "Allow access for Key Administrators",
"Effect": "Allow",
"Principal": {


Success
Your AWS KMS key was created with alias cloudtrail-logs-kms-key and key ID 60663520-f024-4b30-ab6b-69cbfd50664e.
View key
Customer managed keys (1)
Key actions
Create key
Filter keys by properties or tags
1
Aliases
Key ID
Status
Key type
Key spec
Key usage
cloudtrail-logs-kms-...
60663520-f024-4b...
Enabled
Symmetric
SYMMETRIC_DEFAULT
Encrypt and decrypt

CloudTrail:


CloudTrail > Create trail
S3EventLogging
3-128 characters. Only letters, numbers, periods, underscores, and dashes are allowed.
Enable for all accounts in my organization
To review accounts in your organization, open AWS Organizations. See all accounts
Storage location | Info
Create new S3 bucket
Create a bucket to store logs for the trail.
Use existing S3 bucket
Choose an existing bucket to store logs for this trail.
Trail log bucket name
Enter a new S3 bucket name and folder (prefix) to store your logs. Bucket names must be globally unique.
s3-event-logs-bucket-workshop1
Browse
Prefix - optional
prefix
Logs will be stored in s3-event-logs-bucket-workshop1/AWSLogs/361769563538
Log file SSE-KMS encryption | Info
Enabled
Customer managed AWS KMS key
New
Existing
AWS KMS alias
arn:aws:kms:ap-southeast-7:361769563538:key/60663520-f024-4b30-ab6b-69cbfd50664e


CloudTrail > Create trail
Step 2
Choose log events
Step 3
Review and create
Events Info
Record API activity for individual resources, or for all current and future resources in AWS account. Additional charges apply
Event type
Choose the type of events that you want to log.
Management events
Capture management operations performed on your AWS resources.
Data events
Log the resource operations performed on or within a resource.
Network activity events
Network activity events provide information about resource operations performed on a resource within a virtual private cloud endpoint.
Management events Info
Management events show information about management operations performed on resources in your AWS account.
No additional charges apply to log management events on this trail because this is your first copy of management events.
API activity
Choose the activities you want to log.
Read
Write
Exclude AWS KMS events
Exclude Amazon RDS Data API events


aws
Search
[Alt+S]
Asia Pacific (Thailand)
nguyentaiminhhuy
CloudTrail > Trails
Trail successfully created
Trails
Delete
Create trail
Name
Home region
Multi-region trail
Organization trail
S3 bucket
Log file prefix
CloudWatch Logs log group
Status
S3EventLogging
Asia Pacific (Thailand)
Yes
No
s3-event-logs-bucket-workshop1
-
-
Logging