AWS Certified Data Engineer - Associate Cheat Sheet

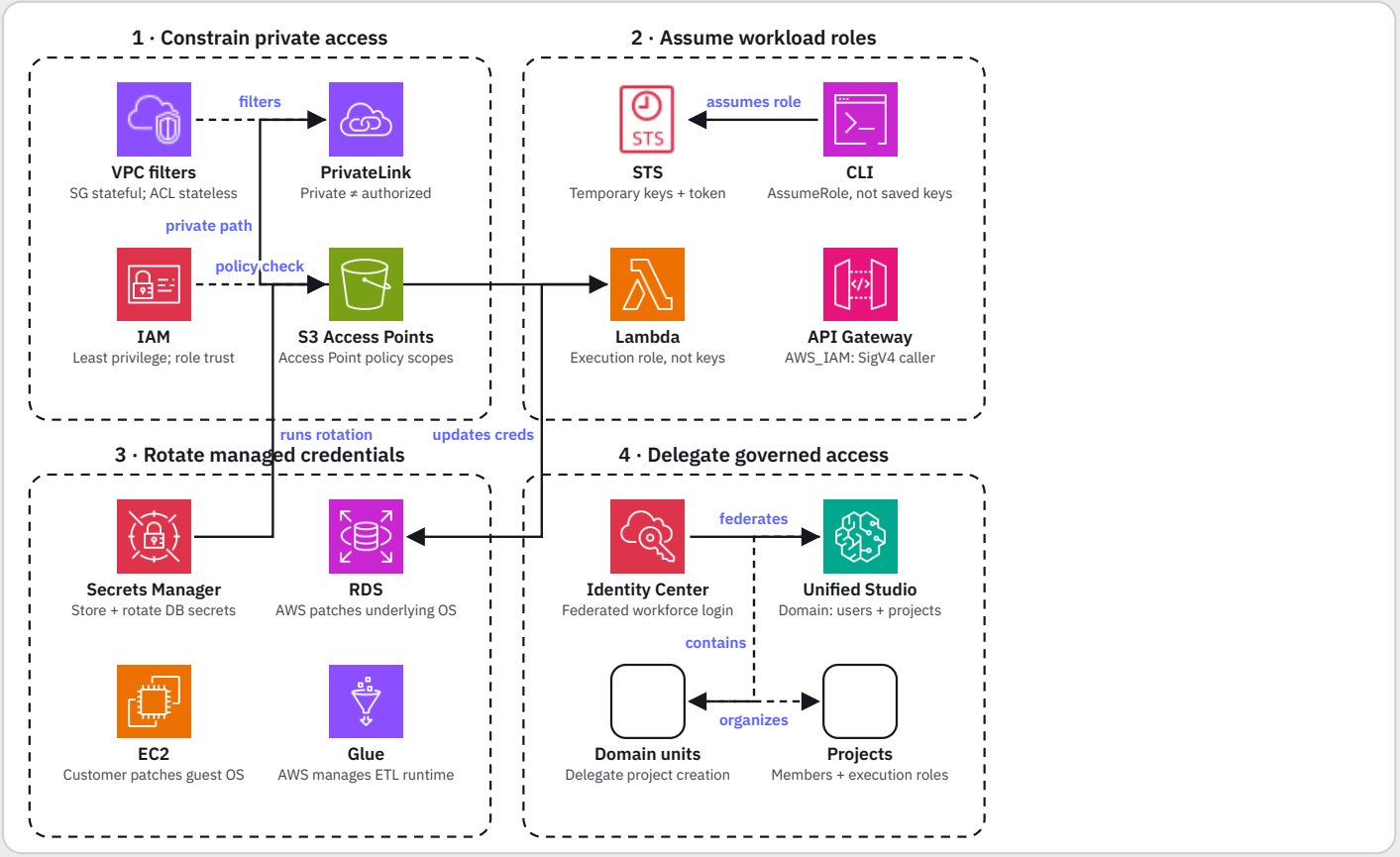
DEA-C01 · Domain 4 of 4

BlowStack

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · 18% OF THE EXAM · TASK STATEMENT 4.1

Apply authentication mechanisms

Filter, assume, rotate, delegate



1 Constrain private access

Use **private endpoints** for private connectivity, not permission grants. Endpoint, resource and identity policies determine effective access; an applicable explicit deny overrides any allow.

2 Assume workload roles

Trust + permissions: trust controls role assumption; permissions control actions. Lambda receives execution-role credentials automatically. CloudFormation can use a service role; IAM groups cannot be assumed.

3 Rotate managed credentials

Rotation must update the target credential and stored secret; consumers must refresh it. Lambda supports custom rotation. RDS and Glue offload infrastructure, not customer responsibility for IAM, data and configuration.

4 Delegate governed access

RBAC separates user authentication from project permissions. Domain units organize and delegate governance; projects assign member roles and execution access. Domain membership alone does not grant data access.

Exam traps

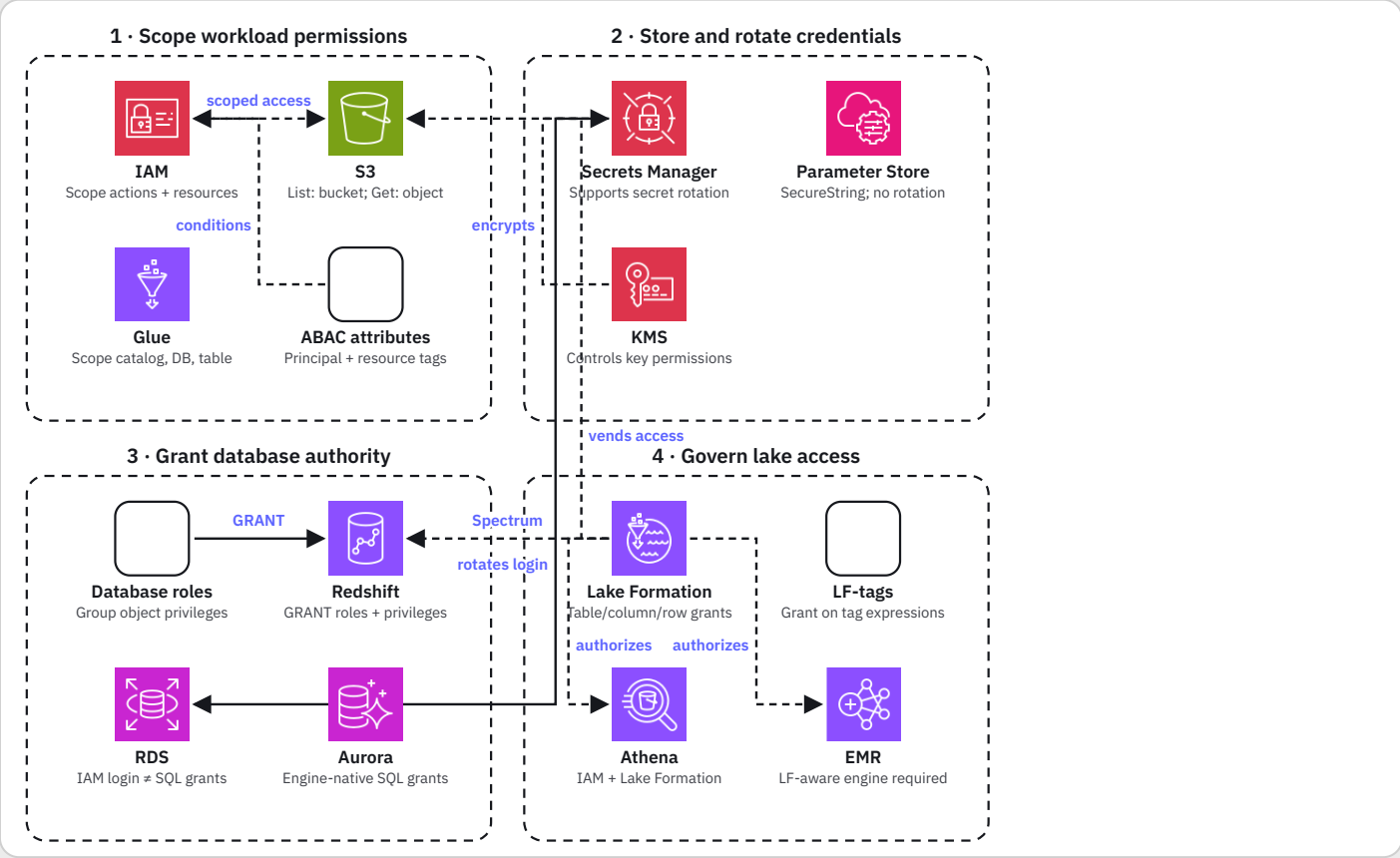
- **Security groups are stateful**, allow-only filters: permitted replies need no reverse rule. NACLs are stateless, support deny rules, and require rules for both directions, including return ports.
- **Rotating a stored password** does not replace workload identity. Roles provide temporary credentials; STS credentials expire rather than being rotated through Secrets Manager.
- **An allow is not sufficient** when an applicable explicit deny exists. Adding another allow cannot override a deny in an endpoint policy, resource policy, identity policy or organizational control.

Get practice exams and lessons for AWS Certified Data Engineer - Associate

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · TASK STATEMENT 4.2

Apply authorization mechanisms

Scope, store, grant, govern



1 Scope workload permissions

Use **least privilege** custom policies when managed policies overgrant: restrict actions, ARNs, and conditions. RBAC assigns roles; ABAC compares trusted principal and resource tags.

2 Store and rotate credentials

Choose **Secrets Manager** for scheduled credential rotation; rotation must be configured. Parameter Store SecureString has no built-in rotation. Reading customer-key-encrypted values also requires KMS decrypt access.

3 Grant database authority

Use **database RBAC** to grant object privileges to roles, then assign users to roles. Redshift supports database roles; RDS and Aurora use engine-specific users, roles, and SQL grants.

4 Govern lake access

Use **Lake Formation** for centralized catalog/data grants, LF-tag policies, and row/column filters. Athena, supported EMR configurations, and Redshift Spectrum honor grants; IAM still controls service API access.

Exam traps

- **Explicit Deny wins** over Allow. Effective access also depends on resource policies, SCPs, boundaries, session policies, and request context; boundaries and SCPs do not grant permissions.
- **IAM database authentication is not SQL authorization**: a login token does not grant table privileges; database GRANTs and role membership still control access.
- **Restrict escalation paths**: broad iam:PassRole or policy-edit rights can defeat least privilege. ABAC also requires controlling who can change authorization tags.

Get practice exams and lessons for AWS Certified Data Engineer - Associate

AWS Certified Data Engineer - Associate Cheat Sheet

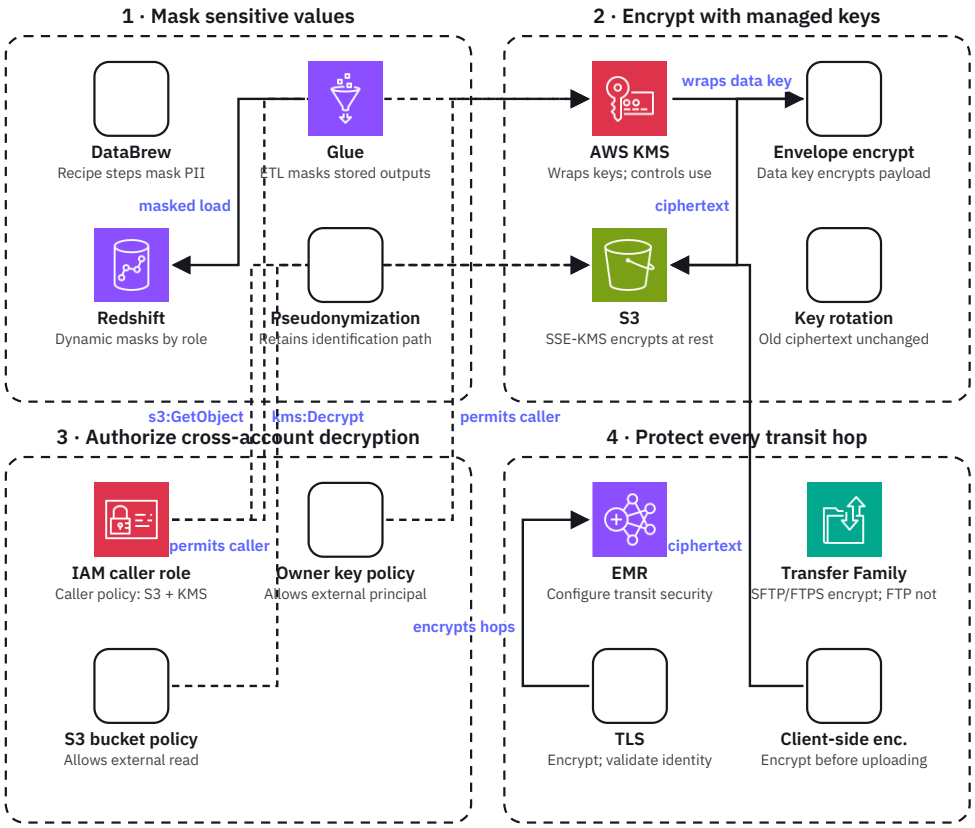
DEA-C01 · Domain 4 of 4

BlowStack

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · TASK STATEMENT 4.3

Ensure data encryption and masking

Mask, encrypt, authorize, protect transit



1 Mask sensitive values

Use recipes or ETL to persist masked outputs; Amazon Redshift **dynamic masking** changes query results, not stored values. Pseudonymization preserves useful links but does not make data anonymous.

2 Encrypt with managed keys

Envelope encryption encrypts data locally with a data key and stores its KMS-encrypted copy with the ciphertext. Automatic KMS rotation changes key material for new operations and retains old material for decryption.

3 Authorize cross-account decryption

Cross-account SSE-KMS reads need S3 authorization and KMS authorization in both accounts. Use a customer managed KMS key; scope principals, actions, objects, and keys to least privilege.

4 Protect every transit hop

Before upload: client-side encryption sends ciphertext to S3. Use an EMR security configuration for in-transit encryption; Transfer Family SFTP uses SSH, FTPS uses TLS, and FTP is unencrypted.

Exam traps

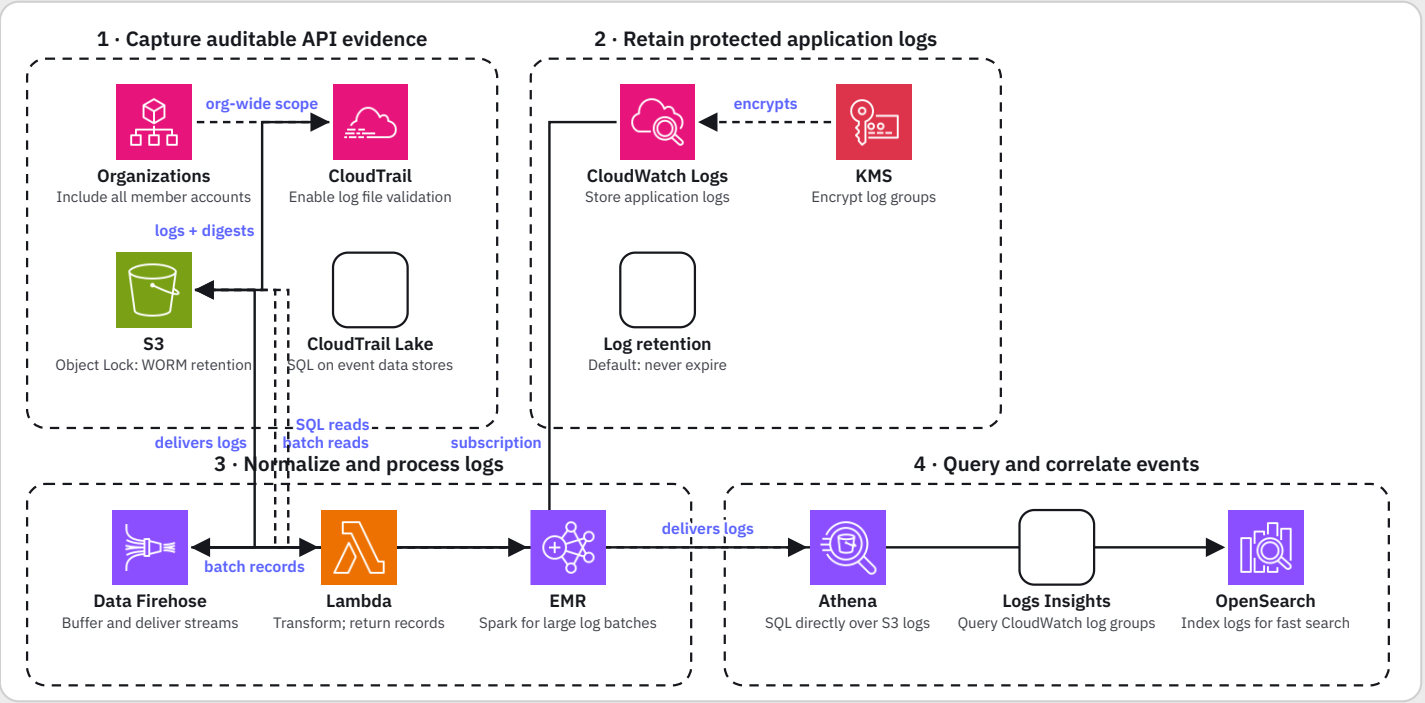
- **Hashing does not guarantee anonymization**: low-entropy identifiers can be guessed by hashing candidate values. Pseudonymized data retains a reidentification path and still needs privacy controls.
- **Bucket permission alone is insufficient** for cross-account SSE-KMS reads: the caller also needs kms:Decrypt and owner-side KMS authorization. The AWS managed aws/s3 key cannot support this sharing.
- **SSE is not transit encryption**: require HTTPS/TLS separately. An S3 bucket policy can deny requests with aws:SecureTransport=false; default encryption does not enforce HTTPS.

Get practice exams and lessons for AWS Certified Data Engineer - Associate

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · TASK STATEMENT 4.4

Prepare logs for audit

Capture, preserve, normalize, query



1 Capture auditable API evidence

An **organization trail** covers member accounts, including new accounts. Organization-enabled Lake event data stores support centralized SQL. Enabling log file validation creates digests; checking them is a separate step.

2 Retain protected application logs

Retention is configured per log group; the default is never expire. KMS controls encryption, not retention. S3 archival requires an explicit export or delivery pipeline.

3 Normalize and process logs

Normalize records with a Firehose Lambda transformation; use EMR for large S3 log batches. Standardize timestamps, identities, and request IDs so events from different sources can be correlated.

4 Query and correlate events

Query location decides: Athena for S3, Logs Insights for log groups, OpenSearch for indexed search. Correlate event times, principals, and request IDs to reconstruct activity across sources.

Exam traps

- CloudTrail Event history is not a durable audit archive: it provides 90 days of regional management events. **Data events**, such as S3 object access, require explicit selection.
- **Log file validation** detects modification or deletion; it does not prevent either. S3 Object Lock provides WORM protection, while access controls restrict who can handle evidence.
- CloudWatch Logs **S3 export is not real time**. Use subscriptions for continuous delivery; setting log-group retention neither archives records to S3 nor configures S3 retention.

Get practice exams and lessons for AWS Certified Data Engineer - Associate

AWS Certified Data Engineer - Associate Cheat Sheet

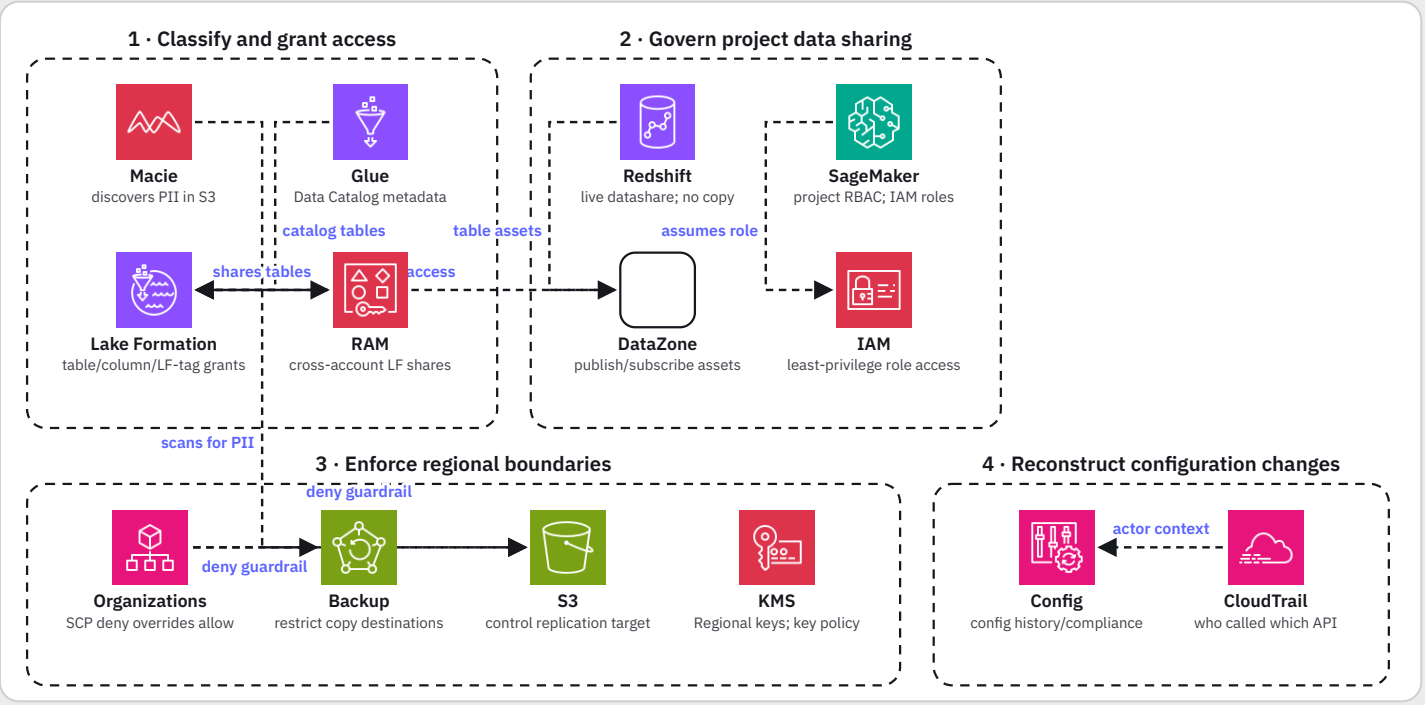
DEA-C01 · Domain 4 of 4

BlowStack

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · TASK STATEMENT 4.5

Understand data privacy and governance

Discover, authorize, constrain, audit



1 Classify and grant access

PII discovery uses Macie on S3; findings do not automatically become Lake Formation grants. Glue supplies catalog metadata; Lake Formation uses AWS RAM for cross-account catalog resource sharing.

2 Govern project data sharing

Project RBAC scopes member roles; approved catalog subscriptions provision supported data access. Redshift datashares read live data without copies. Data mesh combines domain ownership with federated governance.

3 Enforce regional boundaries

Residency covers primary data, backups, and replicas; copy destinations need controls beyond API Regions. SCPs limit permissions, never grant them. Sovereignty also involves jurisdiction, access, and key custody.

4 Reconstruct configuration changes

What changed maps to Config resource history and rule evaluations; API caller identity maps to CloudTrail. Config needs recording enabled; CloudTrail event history covers 90 days of management events.

Exam traps

- **Cross-account sharing** needs producer authorization and consumer permissions; AWS RAM shares resources, not blanket IAM access. Lake Formation grants do not grant unrestricted S3 access.
- **Macie classifies, not masks**: it discovers sensitive data in S3; remediation, ETL masking, and Lake Formation permission changes require separate workflows.
- **Region deny is not destination control**: an allowed-Region request can initiate a copy elsewhere. Restrict backup-copy and S3 replication destinations; encryption alone does not ensure sovereignty.

Get practice exams and lessons for AWS Certified Data Engineer - Associate

DOMAIN 4 · DATA SECURITY AND GOVERNANCE · WHAT TO REMEMBER

Concepts and services of Data Security and Governance

Concepts to know

Least privilege

Grant only required actions on required resources under appropriate conditions. Choose **narrowly scoped workload roles** over shared credentials or broad managed policies; restrict role passing and policy changes that could enable privilege escalation.

Effective Permissions

Usable access depends on applicable identity and resource policies, organizational controls, boundaries, and session restrictions. **An explicit deny overrides an allow**; permissions boundaries and service control policies limit permissions rather than grant them.

Cross-Account Access

Access across account boundaries requires authorization in the relevant accounts. For shared encrypted data, verify **both data and key permissions**; permission to read an object does not by itself authorize decryption.

Core services



AWS Identity and Access Management (IAM)

Use **roles with temporary credentials** for workloads and federation. Scope policies by action, resource, and condition; role trust policies control who can assume a role.



AWS Key Management Service (AWS KMS)

Use **customer managed keys** for cross-account SSE-KMS sharing, with key-policy and caller permissions. Rotation preserves old key material for decryption; it does not re-encrypt existing data.



AWS Lake Formation

Use **centralized fine-grained lake permissions** for supported analytics access. Choose named-resource grants for specific tables or LF-tag-based access control for scalable grants across tagged data.

Get practice exams and lessons for AWS Certified Data Engineer - Associate